

République Algérienne Démocratique et Populaire  
وزارة التعليم العالي والبحث العلمي  
Ministère de l'Enseignement Supérieure et de la Recherche Scientifique  
جامعة محمد الصديق بن يحيى - جيجل  
Université Mohammed-Seddik Benyahia – Jijel

Faculté des Sciences de la Nature et  
de la Vie  
Département de Biologie Moléculaire  
et Cellulaire



كلية: علوم الطبيعة والحياة  
قسم: البيولوجيات الجزيئية والخلوية

Mémoire de Fin d'Études  
En vue de l'obtention du diplôme : de Master Académique en Sciences de La Nature  
Et de la Vie  
Filière: Sciences Biologiques  
Option: Sciences Pharmacologiques

## Thème

---

Analyse *in silico* des groupes de gènes biosynthétiques PKS  
et NRPS pour la découverte de nouveaux composés  
bioactifs dérivés d'actinobactéries

---

Membres de Jury :

Présidente: Dr. LAHOUEL Asma

Examinatrice: Dr. RIANE Karima

Encadrante: Mme AZZOUZ Ouassila

Présenté par :

GUEDJALI Chaima

TAIEB Manel

Année Universitaire 2022-2023

Numéro d'ordre (bibliothèque) : .....

---

---

# DEDECAS

---

---

Je Dédie ce mémoire :

À mes chers parents que dieu me les garde, qui ont toujours été là pour moi et m'ont donné un magnifique modèle de labeur et de persévérance.

À mes adorables sœurs, frères et en particulier ma chère sœur "Amina".

À toute ma famille

À mes amis

Enfin et surtout, je tiens à me dédier. Je souhaite exprimer ma gratitude envers moi-même pour avoir cru en mes capacités, pour ne jamais avoir abandonné et pour être resté fidèle à qui je suis.

Chaima

---

---

# DEDECAS

---

---

Je souhaite dédier ce travail à tous ceux qui ont généreusement partagé leurs conseils incessants, leur assistance et leur soutien moral. C'est un témoignage de ma gratitude et de mon affection, avec l'espoir que vous en serez fiers.

À ma mère, qui est la personne la plus chère à mes yeux et qui m'a toujours encouragée tout au long de mes études. Je prie pour que Dieu te protège et te réserve une longue vie.

À mon père bien-aimé, mon exemple dans cette vie, qui m'a toujours soutenue et encouragée.

À mes sœurs bien-aimées, RAYEL et ASMAA.

À toute ma famille.

À mes chères amies d'enfance : Hiba, Karima, Narimane, Manal et Romaissa.

À mes collègues de la promotion spécialisée en sciences pharmacologiques et en particulier : Chaa,,Djihane, Ibtissam et Yousra.Z .

**TAIEB MANEL**



---

---

## REMERCIEMENTS

---

---

Tout d'abord, nous rendons grâce à Dieu, le tout-puissant, qui nous a donné la force, la volonté, la patience et le courage pour accomplir ce travail.

Nous tenons à exprimer notre gratitude et notre profonde reconnaissance envers nos familles qui nous ont soutenus inlassablement tout au long de ces années d'études.

Nous souhaitons également remercier notre promotrice, Madame AZZOUZ Ouassila, pour ses orientations et ses conseils précieux.

Nous adressons nos remerciements aux membres du jury, le Dr. LAHOUEL Asma et le Dr. RIANE Karima, qui ont accepté d'évaluer notre travail.

Enfin, nous voulons exprimer notre profonde reconnaissance envers tous les enseignants qui ont contribué à notre formation. À tous, nous adressons un grand merci.

---

---

# LISTE DES ABREVIATIONS

---

---

**ACP** : Acyl Carrier Protein

**ANI**: Average Nucleotide Identity

**antiSMASH**: Antibiotics and Secondary Metabolite Analysis Shell

**AT**: AcylTransferase

**BGC**: *Biosynthetic gene clusters*

**BLAST**: *Basic Local Alignment Search Tool*

**BRIG**: *BLAST Ring Image Generator*

**CDS**: CoDing Sequence

**CoA**: *Coenzyme A*

**DH**: Dehydratase

**ER**: Enoyl reductase

**FAS**: Fat Acid Synthases

**GOC**: *Gene Order Conservation*

**HR**: Highly Reducing

**KS**: KetoSynthase

**KR**: Ketoreductase

**MEGA**: Molecular Evolutionary Genetics Analysis

**MIBiG**: Minimum Information about a Biosynthetic Gene cluster

**MT**: Methyltransferase

**MUSCLE**: MUltiple Sequence Comparison by Log-Expectation

**NCBI**: National Center for Biotechnology Information

**NGS**: New generation sequencing

**NR**: Non-Reducing

**NRPS**: Non-Ribosomal Peptides Synthetase

**ORF**: *Open Reading Frames*

**PCP:** Peptidyl Carrier Protein

**PR:** Partially Reducing

**PKS:** Polyketide synthases

**RAST:** Rapid Annotation using Subsystem Technology

**RAxML:** Randomized Axelerated Maximum Likelihood

**WGS:** Whole-genome sequencing

---

---

# LA LISTE DES FIGURES

---

---

<b>FIGURE 1.</b> CYCLE DE DÉVELOPPEMENT DES ACTINOBACTÉRIES .....	6
<b>FIGURE 2.</b> APPARENCE DES ISOLATS D'ACTINOBACTÉRIES SUR UNE PLAQUE DE GÉLOSE AMIDON-CASÉINE. A, C VUE EN PLAQUE DES ISOLATS D'ACTINOBACTÉRIES. B, D MORPHOLOGIE DES COLONIES ISOLÉES.....	8
<b>FIGURE 3.</b> RÉPARTITION DES 6 CLASSES D'ACTIVITÉS BIOLOGIQUES DE 1272 PEPTIDES NON-RIBOSOMIQUES RECENSÉS DANS LA BASE DE DONNÉES NORINE .....	18
<b>FIGURE 4.</b> PRODUITS PEPTIDIQUES NATURELS. UNE SÉLECTION DE PEPTIDES NON SYNTHÉTISÉS PAR LES RIBOSOMES. LES CARACTÉRISTIQUES STRUCTURELLES SONT MIS EN ÉVIDENCE.....	20
<b>FIGURE 5.</b> CHAÎNE D'ASSEMBLAGE CLASSIQUE D'UNE NRPS .....	21
<b>FIGURE 6.</b> RÉACTION CATALYSÉE PAR LE DOMAINE D'ADÉNYLATION DES NRPS. ....	22
<b>FIGURE 7.</b> RÉACTION CATALYSÉE PAR LE DOMAINE DE CONDENSATION DES NRPS. ....	22
<b>FIGURE 8.</b> RÉACTION CATALYSÉE PAR LE DOMAINE DE THIOLATION DES NRPS. ....	23
<b>FIGURE 9.</b> RÉACTION CATALYSÉE PAR LE DOMAINE DE THIOESTÉRISE DES NRPS. ....	23
<b>FIGURE 10.</b> EXEMPLES DE LA COMPLEXITÉ ET DE LA DIVERSITÉ DES POLYKÉTIDES ET DE LEUR FONCTION ASSOCIÉE .....	25
<b>FIGURE 11.</b> COMPARAISON DES MOLÉCULES ISSUES DE SYNTHÈSE ET DE BIOSYNTHÈSE UTILISÉES EN MÉDECINE .....	26
<b>FIGURE 12.</b> REPRÉSENTATION D'UNE PKS DE TYPE I MODULAIRE .....	29
<b>FIGURE 13:</b> SCHÉMA D'ASSEMBLAGE DES PKS DE TYPE II .....	30
<b>FIGURE 14.</b> SCHÉMA D'ASSEMBLAGE DES PKS DE TYPE III .....	30
<b>FIGURE 15.</b> DIAGRAMME DE VENN ILLUSTRANT UN PAN-GÉNOME THÉORIQUE .....	34
<b>FIGURE 16.</b> UN DIAGRAMME AFFICHANT LES PROCESSUS QUI PEUVENT CONDUIRE À, ET DÉFINIR, DES ORTHOLOGUES ET DES PARALOGUES.....	35
<b>FIGURE 17.</b> VUE SCHÉMATIQUE D'UN OPÉRON : PROMOTEUR (P) ; OPÉRATEUR (O); GÈNE OPÉRON ET TERMINATEUR (T).....	37
<b>FIGURE 18.</b> ANNOTATION DU GÉNOME - FLUX DE TRAVAIL .....	40
<b>FIGURE 19.</b> (A) RÉPARTITION GÉOGRAPHIQUE DES SOURCES THERMALES ALGÉRIENNES ÉTUDIÉES DANS CE TRAVAIL .....	47

<b>FIGURE 20.</b> LA NORME DE DONNÉES MIBiG ET LE SYSTÈME DE SOUMISSION CONDUIRONT À UN ENSEMBLE DE DONNÉES EN CROISSANCE CONTINUE .....	53
<b>FIGURE 21.</b> ANALYSE PHYLOGÉNÉTIQUE MOLÉCULAIRE EN UTILISANT LES SÉQUENCES DE L'ARNr 16S, DE STREPTOMYCES AMBOFACIENS ATCC 23877, STREPTOMYCES LEEUWENHOEKII STRAIN C34 ET STREPTOMYCES SP. KY75 .....	55
<b>FIGURE 22.</b> ARBRES PHYLOGÉNÉTIQUES BASÉ SUR LES SÉQUENCES DU GÉNOME ENTIER DES SOUCHES STREPTOMYCES AMBOFACIENS ATCC 23877, STREPTOMYCES LEEUWENHOEKII ET STREPTOMYCES SP. KY75.....	56
<b>FIGURE 23.</b> COMPARAISON DU GÉNOME DES TROIS SOUCHES DE STREPTOMYCES .....	58
<b>FIGURE 24.</b> DIAGRAMME DE VENN DU NOMBRE DE PROTÉINES PARTAGÉES ET UNIQUES ENTRE <i>STREPTOMYCES AMBOFACIENS</i> ATCC 23877 CP012382, <i>STREPTOMYCES LEEUWENHOEKII</i> STRAIN C34, <i>STREPTOMYCES SP. KY75</i> (NZ_LR881953.1) ET LES SOUCHES LES PLUS PROCHEES <i>STREPTOMYCES AMBOFACIENS</i> DSM 406, <i>STREPTOMYCES CYANOGRISEUS SUBSP NONCYANOGENUS</i> NMWTET <i>STREPTOMYCES SP. KY70</i> , TEL QUE CRÉÉ PAR ORTHOVENN3. ....	59
<b>FIGURE 25.</b> INFORMATIONS SUR LES SOUS-SYSTÈMES DES GÈNES POUR LA SOUCHE <i>STREPTOMYCES AMBOFACIENS</i> ATCC 23877, OBTENU PAR LE SERVEUR PATRIC .....	63
<b>FIGURE 26.</b> APERÇU SUR LES SOUS-SYSTÈMES DES GÈNES POUR LA SOUCHE <i>STREPTOMYCES LEEUWENHOEKII</i> C34 (= DSM 42122 = NRRL B-24963) (1437453.21), OBTENU PAR LE SERVEUR PATRIC. ....	64
<b>FIGURE 27.</b> INFORMATIONS SUR LES SOUS-SYSTÈMES DES GÈNES POUR LA SOUCHE <i>STREPTOMYCES SP. KY75</i> , OBTENU PAR LE SERVEUR PATRIC. ....	64
<b>FIGURE 28.</b> IDENTIFICATION DES RÉGIONS DE MÉTABOLITES SECONDAIRES À L'AIDE DU SERVEUR ANTISMASH. ....	65
<b>FIGURE 29.</b> GROUPES DE GÈNES PUTATIFS CODANT POUR DES MÉTABOLITES SECONDAIRES DANS <i>STREPTOMYCES AMBOFACIENS</i> ATCC 23877. ....	65
<b>FIGURE 30.</b> GROUPES DE GÈNES PUTATIFS CODANT POUR DES MÉTABOLITES SECONDAIRES DANS <i>STREPTOMYCES SP. KY75</i> . ....	66
<b>FIGURE 31.</b> GROUPES DE GÈNES PUTATIFS CODANT POUR DES MÉTABOLITES SECONDAIRES DANS <i>STREPTOMYCES LEEUWENHOEKII</i> C34. ....	66
<b>FIGURE 32.</b> INDICATEURS DE COULEURS (ALAM ET AL., 2022). ....	66
<b>FIGURE 33.</b> DIAGRAMME DE CORDE CIRCULAIRE POUR VISUALISER LES CLASSES DE BGC PRÉDITES DANS LES SOUCHES ÉTUDIÉES.....	70



<b>FIGURE 34.</b> LES RÉSULTATS DES GÈNES BIOSYNTHÉTIQUES OBTENUS PAR ANTISMASH. ....	71
<b>FIGURE 35.</b> DISTRIBUTION DES DOMAINES PKS ET NRPS DANS LES SOUCHES <i>STREPTOMYCES KY75</i> , <i>STREPTOMYCES LEEWENHOEKI C34</i> ET <i>STREPTOMYCES AMBOFACIEN ATCC278992</i> .....	72
<b>FIGURE 36.</b> REPRÉSENTATIONS GRAPHIQUES DES CLUSTERS CONNUS LES PLUS SIMILAIRES EN TANT QUE RÉSULTAT DE SPÉCIFIÉ PAR KNOWNBLASTER DANS LE RÉFÉRENTIEL MIBIG.....	74
<b>FIGURE 37.</b> PRÉDICTION DES STRUCTURES MOLÉCULAIRES À ACTIVITÉS THÉRAPEUTIQUES .....	79

---

---

# LA LISTE DES TABLEAUX

---

---

<b>TABLEAU 1.</b> MARQUEURS CHIMIOTAXONOMIQUES APPLIQUÉS DANS L'APPROCHE POLYPHASIQUE POUR L'IDENTIFICATION DES ACTINOBACTÉRIES.....	9
<b>TABLEAU 2.</b> EXEMPLES DES ANTIBIOTIQUES PRODUITS PAR LES ACTINOBACTÉRIES. ....	12
<b>TABLEAU 3.</b> EXEMPLES DE SUBSTANCES ANTI-OXYDANTES EXTRAITES DES ACTINOBACTÉRIES. ....	14
<b>TABLEAU 4.</b> EXEMPLES DE SUBSTANCES ANTI- TUMORALES PRODUITES PAR LES ACTINOBACTÉRIES. ....	15
<b>TABLEAU 5.</b> QUELQUES ENZYMES PRODUITS PAR ACTINOBACTÉRIES .....	16
<b>TABLEAU 6 :</b> CARACTÉRISTIQUES DES PKS DE TYPE I,II,III.....	27
<b>TABLEAU 8.</b> CARACTÉRISTIQUES GÉNOMIQUES DES SOUCHES ÉTUDIÉES STREPTOMYCES KY75, STREPTOMYCES LEEWENHOEKI C34 ET STREPTOMYCES AMBOFACIEN ATCC278992.....	61
<b>TABLEAU 9.</b> GROUPES DE GÈNES BIOSYNTHÉTIQUES DE MÉTABOLITES SPÉCIALISÉS IDENTIFIÉS DANS LES SOUCHES ANALYSÉES. ....	67
<b>TABLEAU 10.</b> COMBINAISON DES RÉSULTATS D'ANTISMASH ET MIBIG, VARIATIONS DES RÉGIONS NRPS ET PKS ET LES TYPES DE MOLÉCULES PRODUITES CHEZ DIFFÉRENTES SOUCHES BACTÉRIENNES.....	74

---

---

# SOMMAIRE

---

---

DEDECAS .....	I
REMERCIEMENTS .....	III
LISTE DES ABREVIATIONS .....	IV
LA LISTE DES FIGURES .....	VI
LA LISTE DES TABLEAUX .....	IX
SOMMAIRE.....	X
INTRODUCTION .....	1

## APERÇU BIBLIOGRAPHIQUE

CHAPITRE I .....	4
I. <i>Généralités sur les actinobactéries</i> .....	4
I.1.    Historique .....	4
I.2.    Généralité et définition .....	5
I.3.    Cycle de développement .....	5
I.4.    Habitats.....	6
I.5.    Critères d'identification des actinobactéries.....	8
I.5.1.    Les critères morphologiques .....	8
I.5.2.    Les critères chimiotaxonomiques.....	8
I.5.3.    Les critères physiologiques.....	9
I.5.4.    Les critères génétiques.....	10
I.6.    Métabolites secondaires des actinobactéries .....	10
I.6.1.    Principales substances bioactives issues des actinobactéries.....	11
I.6.1.1.    Les antibiotiques.....	12
I.6.1.2.    Les antifongiques.....	13
I.6.1.3.    Les substances antioxydantes .....	13
I.6.1.4.    Les substances anti- cancéreuses.....	14
I.6.1.5.    Les enzymes.....	15
CHAPITRE II.....	17
II. <i>Généralités sur les groupes de gènes biosynthétiques PKS et NRPS (BGC)</i> .....	17
II.1.    Les peptides bactériens non ribosomiques (NRP).....	17
II.1.1.    Voie de synthèse des protéines.....	17
II.1.2.    Voie de synthèses des peptides bactériens non ribosomiques (NRP) .....	18
II.1.3.    Organisation des NRPS .....	21
II.1.4.    La détection des NRPS .....	23
II.2.    Les gènes biosynthétiques de type PKS.....	24
II.2.1.    Découvert et répartition génétique .....	26
II.2.2.    Les types PKS et leur capacité biosynthétiques.....	27
CHAPITRE III.....	31
III. <i>Annotation et prédiction génomique des PKS et NRPS chez les actinobactéries</i> .....	31
III.1.    Introduction.....	31

III.2.	Notion des multi-omiques .....	32
III.3.	Caractéristiques du génome des actinobactéries .....	33
III.3.1.	Diversité dans la taille du génome .....	33
III.3.2.	Pourcentage GC et stabilité du génome.....	33
III.3.3.	Core-génome et pan-génome .....	34
III.3.4.	Genès orthologues et Genès paralogues.....	35
III.3.5.	Détection des ORF .....	36
III.3.6.	Importance des séquences codantes CDS.....	37
III.3.7.	Conservation de l'ordre des gènes biosynthétiques GOC .....	37
III.4.	Détection et prédiction des groupes de gènes biosynthétique .....	38
III.4.1.	Méthodes de prédiction .....	38
III.4.2.	Séquençage.....	39
III.4.3.	Annotation automatique .....	40
III.4.3.1.	Annotation structurale .....	41
III.4.3.2.	Annotation fonctionnelle.....	41

## **PARTIE PRATIQUE**

MATERIELS ET METHODES .....	44
I. <i>Choix des souches actinomycétales pour l'étude in silico</i> .....	45
II. <i>Analyses phylogénétiques et phylogénomiques</i> .....	47
II.1.  Analyse phylogénétique .....	47
II.1.1.  Analyse phylogénétique basée sur le gène de l'ARNr 16S .....	48
II.1.2.  Analyse phylogénétique basée sur le génome entier .....	48
II.2.  Comparaisons des génomes entiers.....	49
II.3.  Détermination de l'identité nucléotidique moyenne (ANI).....	49
II.4.  Analyse de similitude des protéines prédites.....	50
III. <i>Annotation génomique</i> .....	50
III.1.  Annotation structurale et fonctionnelle du génome .....	50
III.2.  Analyses des gènes des métabolites secondaires bioactifs et des composés.....	51
III.2.1.  Détection des régions de métabolites secondaires bioactifs .....	51
III.2.2.  Profilage des métabolites secondaires .....	51
III.2.3.  Prédiction des composés secondaires bioactifs .....	52
RESULTATS ET DISCUSSIONS.....	54
I. <i>Analyses phylogénétiques et phylogénomiques</i> .....	54
I.1.  Analyse phylogénétique .....	54
I.1.1.  Analyse phylogénétique basée sur le gène de l'ARNr 16S .....	54
I.1.2.  Analyse phylogénétique basée sur le génome entier .....	55
I.2.  Comparaisons des génomes entiers.....	57
I.3.  Détermination de l'identité nucléotidique moyenne (ANI).....	58
I.4.  Analyse de similitude des protéines prédites.....	59
II. <i>Annotation génomique</i> .....	60
II.1.  Annotation structurale et fonctionnelle du génome .....	60
II.2.  Analyses des gènes des métabolites secondaires bioactifs et des composés.....	65
II.2.1.  Détection des régions de métabolites secondaires bioactifs .....	65

II.2.2. Profilage des métabolites secondaires.....	70
II.2.3. Prédiction des composés secondaires bioactifs.....	72
II.2.3.1. Groupes de gènes connus similaires de MIBiG 3.1.....	73
CONCLUSION .....	83
LES REFERENCES .....	85
ANNEXE	
RESUME	

---

---

# INTRODUCTION

---

---

Les polykétides (PKS) synthases et les peptides synthétases non ribosomiques (NRPS) sont des enzymes impliqués dans les voies biosynthétiques des métabolites secondaires naturels importants sur les plans pharmaceutique et biotechnologique. Ces enzymes sont codées par des groupes de gènes qui se regroupent en clusters appelés groupes de gènes biosynthétiques ou BGC (Biosynthetic Gene Clusters), qui sont souvent regroupés dans le génome bactérien (Komaki & Tamura, 2020).

Les actinobactéries, des bactéries Gram positives filamenteuses, largement distribuées dans la nature. Elles sont considérées comme une importante source incontournable de métabolites secondaires bioactifs, utilisés à des fins pharmaceutiques. Parmi les genres d'actinomycètes les mieux étudiés, le genre *Streptomyces* est particulièrement remarquable en raison de sa capacité exceptionnelle à produire des substances bioactives. La plupart des composés d'origine microbienne découverts à ce jour avec des activités antibiotiques, antitumorales ou immunosuppressives, antibiotiques, antitumoraux ou immunosuppresseurs sont dérivés de *Streptomyces*. Il a été suggéré que ces bactéries pourraient produire beaucoup plus de métabolites que ceux qui ont été identifiés à ce jour. Cette caractéristique est due à la présence de nombreux clusters de gènes biosynthétiques dans son génome, notamment les PKS et NRPS (Djinni et al., 2019).

Le séquençage du génome microbien a considérablement augmenté au cours de la dernière décennie, donnant accès à une énorme quantité d'informations sur la capacité génétique des micro-organismes à produire des produits naturels. Des analyses bio-informatiques ont révélé que le nombre de groupes de gènes biosynthétiques putatifs de produits naturels dépasse souvent largement le nombre de métabolites rapportés, suggérant qu'une grande partie du potentiel chimique de la nature reste inexplorée (Weber & Kim, 2016).

L'analyse des groupes de gènes biosynthétiques PKS et NRPS offre une approche efficace pour la découverte de nouveaux composés bioactifs. Elle permet d'exploiter le potentiel métabolique des actinobactéries et de guider la recherche de nouveaux antibiotiques et autres agents thérapeutiques.

La présente étude est divisée en deux grandes parties :

- Une première partie bibliographique, composée de trois chapitres dont :

Le premier chapitre de cette étude se concentre sur les généralités concernant les actinobactéries. Il fournit des informations sur leur importance en tant que productrices principales des métabolites secondaires précieux et leurs modes d'action.

Le deuxième chapitre est consacré aux généralités sur les groupes de gènes biosynthétiques polyketide synthase (PKS) et nonribosomal peptide synthetase (NRPS). Ce chapitre examine les différentes classes de PKS et NRPS et leur diversité fonctionnelle, en expliquant leur rôle dans la biosynthèse des métabolites secondaires chez les actinobactéries et souligne leur importance dans la production de composés bioactifs.

Le troisième chapitre se concentre sur l'annotation et la prédiction génomique des PKS et NRPS chez les actinobactéries. Il décrit les méthodes utilisées pour identifier et caractériser ces groupes de gènes biosynthétiques dans les génomes bactériens.

Une deuxième partie pratique, réalisée *in silico*, qui porte sur l'étude de trois souches de *Streptomyces* provenant de deux écosystèmes extrêmes différents en Algérie. Cette étude comprend :

- Une analyses phylogénétiques et phylogénomiques.
- Une annotation structurelle et fonctionnelle du génome.
- Une annotation et prédiction génomiques des groupes de gènes biosynthétiques (BGC) PKS et NRPS.

# APERÇU BIBLIOGRAPHIQUE



---

---

# CHAPITRE I

---

---

## I. Généralités sur les actinobactéries

### I.1. Historique

La découverte, la caractérisation, la nomenclature et l'utilisation des actinobactéries se répartissent en plusieurs périodes :

- Tout d'abord, c'est la découverte de la première actinobactérie *Streptothrix foeresteri* afin d'étudier les organismes pathogènes et leur rôle en pathologie par Cohn en 1875 (Anibou et al., 2008).
- Ensuite, la découverte des conditions d'habitat saprophytes des actinobactéries et les premières tentatives de distinction entre les groupes pathogènes et saprophytes (Baldacci, 1962).
- De plus, la découverte des antibiotiques produits par les actinobactéries en 1943. Cette période a vu une augmentation du nombre d'espèces, des brevets de spécialités antibiotiques et des réglementations orientant les nouvelles espèces (Baldacci, 1962). En 1943, Waksman et Henrici créent le genre *Streptomyces* (combinant les noms des genres *Streptothrix* et *Actinomyces*) qui regroupe les actinobactéries dont le mycélium aérien produit des chaînes de spores portées par des sporophores.
- D'autre part, la prise de conscience qu'un même antibiotique peut être produit par différentes actinobactéries. Cette période est caractérisée par la révision des travaux effectués et des règles de nomenclature. C'est donc la période d'élaboration des critères morphologiques et biochimiques de la classification des actinobactéries.
- En 1944 : la découverte de la streptomycine produite par *Streptomyces griseus* (Le Minor Léon, 1989).
- A partir des années 1960, le début de l'utilisation des méthodes génétiques par Hopwood (Hopwood & Wright, 1973; Chater, 1999). La génomique (Hopwood, 2003) a bouleversé la classification des espèces (Ventura et al., 2007) et l'étude de l'utilisation biotechnologique de ces microorganismes.

## I.2. Généralité et définition

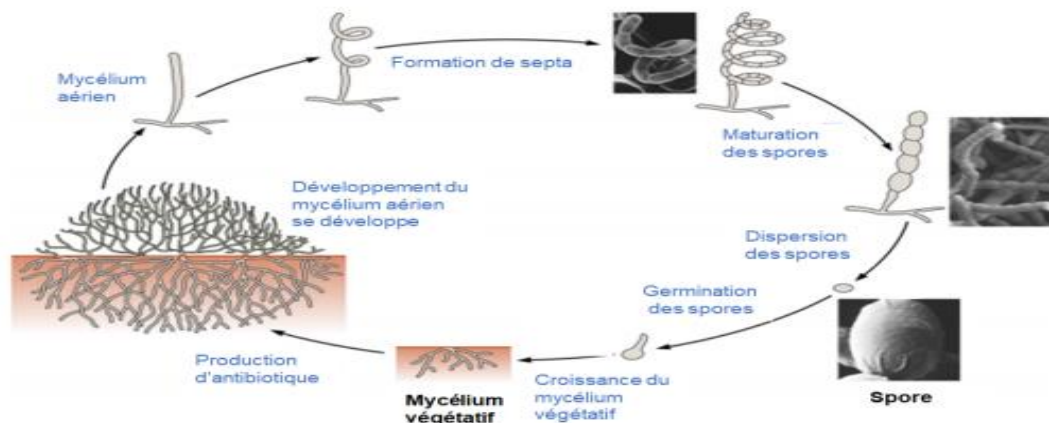
Les actinobactéries représentent un grand groupe de procaryotes constituant l'ordre des Actinomycétales (Mariat & Sebald, 1990). Ces bactéries sont connus comme des bactéries Gram positif et se caractérisent par la formation de mycélium aérien et de substrat sur des milieux solides, ainsi que par la présence de spores avec différentes surfaces de spores, et une teneur élevée en Guanine + Cytosine de l'ADN (Goodfellow & Williams, 1983).

Les actinobactéries ont souvent été confondues avec les mycètes en raison de leur morphologie fongicoïde (présence de filaments ramifiés, sporulation) et également en raison de l'aspect mycosique des maladies de certaines d'entre elles, ce qui explique leur dénomination. En effet, le mot Actinomycète a été dérivé des mots grecs «Aktis» qui signifie rayon et «mykes» qui signifie champignon «Champignons à rayons» ou «Champignons rayonnants» (Reponen et al., 1998).

## I.3. Cycle de développement

Comme les eucaryotes multicellulaires, les actinobactéries ont un cycle de vie qui est le produit de trois processus physiologiques majeurs : la croissance végétative, la différenciation et la sénescence cellulaire, puis la mort. Les actinobactéries ont un cycle de vie similaire à celui de certains champignons, mais leur structure procaryote sans noyau distinct les a classées parmi les bactéries (Barka et al., 2015).

La figure 1 montre le cycle de développement des actinobactéries. Sur un milieu solide, le cycle commence par la germination d'une spore qui donne naissance à un mycélium végétatif formé d'hyphes multi-nucléaires, ramifiés et ancrés dans le milieu solide. Un mycélium aérien se développe sur ce mycélium végétatif, utilisant ce dernier comme substrat. En fait, le mycélium végétatif s'autolyse et les produits de la lyse sont cannibalisés par le mycélium aérien (Miguélez et al., 1999). La croissance hyphale est apicale (se produit par le haut) et s'accompagne de la formation de septa, conduisant à des unités unigéniques. Les cellules se différencient ensuite pour former des spores. Si les spores sont situées dans des sporanges, elles sont appelées sporangiospores. Généralement ces spores ne sont pas résistantes à la chaleur, mais résistent bien à la dessiccation et ont donc une valeur adaptative importante, les actinobactéries sont immobiles, sauf les spores de certains genres (*Actinoplan*, *Spirillospora*...etc.) (Flärdh, 2010).



**Figure 1.** Cycle de développement des actinobactéries (Barka et al., 2015).

#### I.4. Habitats

Les actinomycètes sont capables de survivre dans différents habitats et sont largement répandus dans les écosystèmes naturels. En se basant sur des critères morphologiques et chimiques, les actinomycètes ont été regroupés en différents genres ; *Streptomyces* est le genre le plus communément isolé de l'ordre Actinomycetales en raison de sa grande importance en médecine, en écologie, en biotechnologie médicale et en industrie biotechnologique (Karupiah & Mustafa, 2013).

Le sol est l'habitat le plus important pour les actinobactéries. On peut les trouver à  $10^6$  -  $10^9$  cellules par gramme de sol (Goodfellow & Williams, 1983). D'autres facteurs environnementaux, à savoir la température, le pH, l'humidité du sol ont également un effet important sur la densité de population, ainsi que sur les types d'actinobactéries. Certaines actinobactéries mésophiles sont trouvées avec des températures de croissance optimales dans la gamme de 25°C à 30°C. D'autre part, les actinobactéries dont la température de croissance se situe entre 50°C et 60°C sont dites thermophiles (Edwards, 1993). Une faible humidité favorise la croissance végétative des actinobactéries, tandis que la croissance des actinobactéries peut être affectée dans les sols secs avec une moindre teneur en humidité (Barka et al., 2016) .

Les plantes actinomycorrhiziennes sont caractérisées par la formation dans leurs racines des nodules de fixation de l'azote en combinaison avec des actinomycètes comme *Frankia sp.* Il a été mentionné qu'un nouvel antibiotique, la calcimycine, a été produit par l'espèce *Frankia sp.* (Udwary et al., 2011).

Il a été connu ces dernières années que l'eau douce constitue une source importante d'actinomycètes, où le genre *Streptomyces* est dominant dans l'eau de la rivière tandis que, le genre *Micromonospora* est dominant dans les sédiments de la rivière. Ces actinomycètes ont une activité antifongique dans cet habitat. Cet habitat pourrait être de nouvelle source de substances à activité antimycosique contre les levures et les moisissures qui affectent les patients atteints d'une maladie en phase terminale (Rifaat, 2003).

La distribution des actinomycètes dans l'environnement aquatique a été dans une large mesure non découverte et l'existence d'actinomycètes marins primitifs dans les océans reste un mystère. Cela est dû à l'absence de tentative de découverte des taxons d'actinomycètes marins, alors que les actinomycètes des habitats terrestres ont été, jusqu'à récemment, une source utile de nouveaux composés bioactifs (Murthi & Agrawal, 2010).

La distribution des actinobactéries dans l'environnement marin a été rapportée par de nombreux chercheurs. Les bactéries marines peuvent également constituer un domaine de recherche important, car elles peuvent produire des produits naturels importants (Bull & Stach, 2007). La plupart des actinobactéries marines contiennent des polykétides synthétases (PKS) et des polykétides synthétases non ribosomiques (NRPS), indiquant leur capacité à produire des métabolites secondaires (Salomon et al., 2004; Cho et al., 2006; Fenical & Jensen, 2006; Jensen et al., 2007).

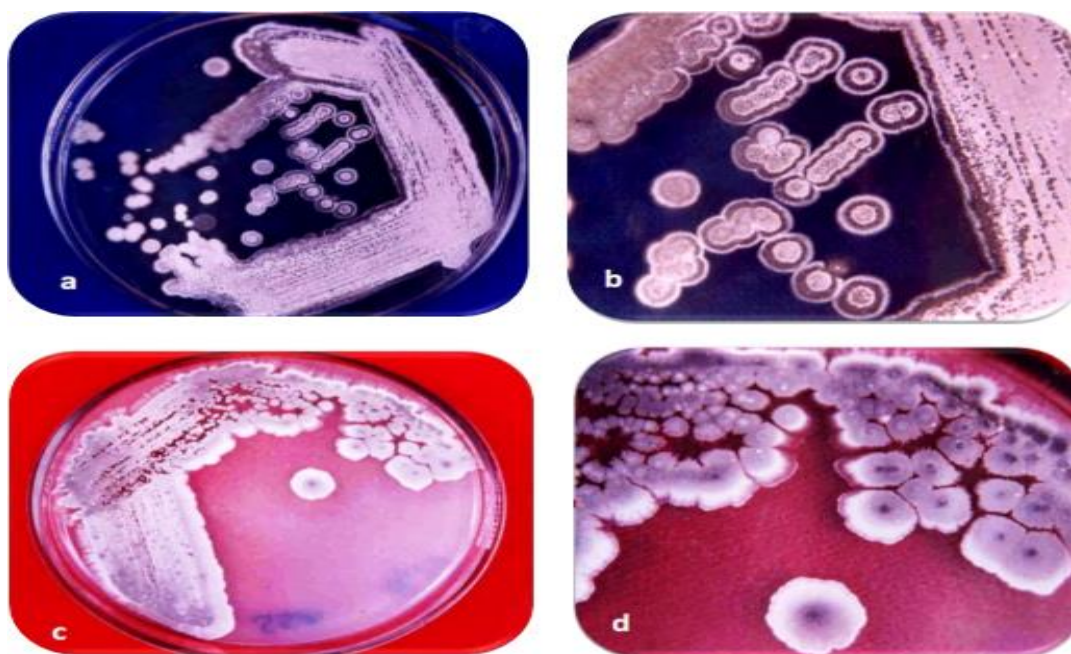
Les actinomycètes sont nés dans des conditions marines inhabituelles, par exemple dans les réservoirs d'hydrates de gaz et les agrégats organiques des grands fonds marins, où ils sont les principaux composants des communautés microbiennes. Dans la mer des Wadden, les actinomycètes isolés des agrégats organiques marins qui s'y trouvent présentaient une forte activité antagoniste au sein de cette communauté. Ces actinobactéries marines produisent des composés précieux qui sont utilisés dans l'industrie pharmaceutique (Lam, 2006).

La capacité des spores d'actinomycètes à se trouver dans l'air a été détectée, l'air contient différents types de leurs spores. Il a été déclaré que les spores d'actinomycètes aéroportés comme *Nocardia sp.* sont responsables de différentes productions antimicrobiennes (McNeil & Brown, 1994 ; Asan et al., 2004).

## I.5. Critères d'identification des actinobactéries

### I.5.1. Les critères morphologiques

Les actinobactéries comprennent un groupe de micro-organismes unicellulaires ramifiés, dont la plupart forment un mycélium aérobie appelé substrat et aérien. Ils se reproduisent par fission binaire ou en produisant des spores ou des conidies. La sporulation des actinobactéries se fait par fragmentation et segmentation ou par formation de conidies. L'aspect morphologique des actinobactéries est compact, souvent coriace, donnant un aspect conique avec une surface sèche sur les milieux de culture et sont fréquemment couverts de mycélium aérien (Figure 2) (Ranjani et al., 2016).



**Figure 2.** Apparence des isolats d'actinobactéries sur une plaque de gélose amidon-caséine. a, c Vue en plaque des isolats d'actinobactéries. b, d Morphologie des colonies isolées (Ranjani et al., 2016).

### I.5.2. Les critères chimiotauxonomiques

La chimiotauxonomie des actinobactéries s'intéresse à la distribution de substances chimiques spécifiques de l'enveloppe cellulaire des actinobactéries, telles que les acides aminés, les sucres, les lipides polaires, les ménaquinones, l'acide mycolique et les acides gras (tableau 1), en utilisant des techniques chimiques, notamment l'extraction, le fractionnement, la purification, la détermination de la composition chimique, et la résolution des composés cibles (Goodfellow & o'donnell, 1994).

**Tableau 1.** Marqueurs chimiotaxonomiques appliqués dans l'approche polyphasique pour l'identification des actinobactéries (Ranjani et al., 2016).

Catégories	Site dans la cellule	Composition	Constituants	Especce
Chimiotaxonomie	Paroi cellulaire	Acide aminé	-L- DAP+Glycine (type I)	Nocardioïdes et Streptomyces
			-Lysine+ Ornithine (type V)	- Actinomyces
			- Acide Aspartique et de galactose (Type VI)	-Actinomyces et Arcanobacterium
	Cellule	Sucres	-Arabinose et galactose  -Madurose (3-O- méthyl-D- galactose)  -Ne synthétisent aucun glucide  -Arabinose et des xyloses	-Nocardia et Saccharopolyspora  -Actinomodura et Streptosporangium  -Streptomyces  -Actinoplanes et Micromonospora
Membranes	Lipides polaires	- Phospholipides	- Actinomadura et Streptomyces	
Membranes	Acides gras	-12 à 20 atomes de Carbones	- Streptomyces	
		-20 à 90 atomes de Carbones	-Nocardia et Mycrobacterium	

### I.5.3. Les critères physiologiques

Un large éventail de caractéristiques physiologiques a été évalué, notamment le profil d'utilisation des glucides, le profil d'utilisation des sources d'azote, la dégradation ou l'hydrolyse de

nombreux substrats, et la sensibilité à divers inhibiteurs (Shirling & Gottlieb, 1966 ; Messaoudi et al., 2015).

#### I.5.4. Les critères génétiques

L'identification des actinobactéries *via* les méthodes traditionnelles telles que les caractéristiques phénotypiques n'est pas aussi précise que les méthodes génotypiques. L'analyse de la séquence du gène de l'ARNr 16S a été reconnue comme un outil puissant pour l'identification des souches mal décrites, rarement isolées ou phénotypiquement aberrantes et peut conduire à une analyse phylogénétique unique des souches nouvellement isolées (Heuer et al., 1997; Monciardini et al., 2002; Busti et al., 2006). Cependant, il n'est pas encore possible d'achever une classification supragénérique complète des actinobactéries à partir des résultats du séquençage partiel de l'ARNr 16S et il n'a pas été possible de distinguer des espèces étroitement apparentées, voire des genres (Colquhoun et al., 1998). Une taxonomie actualisée de l'embranchement des actinobactéries basée sur les arbres d'ARNr 16S a été récemment décrite (Korn-Wendisch & Schneider, 1992). Cette classification actualisée a supprimé les rangs taxonomiques des sous-classes et des sous-ordres, faisant passer les sous-classes et sous-ordres précédents au rang de classes et d'ordres, respectivement (Barka et al., 2016). L'embranchement est donc réparti en six classes distinctes : Actinobacteria, Acidimicrobiia, Coriobacteria, Nitrospirae, Rubrobacteria et Thermoleophilia. La classe des Actinobacteria contient 16 ordres, comprenant les ordres précédemment suggérés, Actinomycetales et Bifidobacteriales (Hasegawa et al., 1983). L'ordre Actinomycetales est maintenant délimité uniquement aux membres de la famille Actinomycetaceae (Barka et al., 2016). Plusieurs genres ont été identifiés par l'analyse des séquences du gène de l'ARNr 16S, tels que *Streptomyces*, *Micromonospora*, *Kribbella*, *Actinomadura* et *Saccharopolyspora* (Patel et al., 2004).

#### I.6. Métabolites secondaires des actinobactéries

Les métabolites secondaires sont des molécules organiques qui ne participent pas à la croissance et au développement normal d'un organisme. Ces composés constituent un groupe extrêmement diversifié de produits naturels synthétisés par les bactéries. La plupart des métabolites secondaires, tels que les terpènes, les composés phénoliques et les alcaloïdes, sont classés en fonction de leur origine biosynthétique. Les différentes classes de ces composés sont souvent associées à un ensemble étroit d'espèces au sein d'un groupe phylogénétique (Agostini-Costa et al., 2012).

Le développement des techniques de séquençage de nouvelle génération (NGS) a permis l'exploration de multiples génomes microbiens ainsi que de métagénomes, révélant un grand nombre

de nouveaux groupes de gènes impliqués dans la biosynthèse de métabolites spécialisés, d'enzymes et de protéines, élargissant ainsi le potentiel de découverte de produits naturels, des gènes aux molécules (Um et al., 2013; Elsayed et al., 2015). Exploitation de ces génomes pour la présence de groupes de gènes biosynthétiques (BGCs) qui permettent aux organismes de produire des métabolites secondaires représente une meilleure possibilité de découverte de nouveaux composés (Lee et al., 2020; Z. Yang et al., 2020).

Ce nombre représente près de 17 % de tous les métabolites connus et isolés jusqu'ici des microbes. Parmi ceux-ci, le groupe des actinomycétales filamenteux produit plus de 10000 des composés bioactifs, hors dont près de 7500 composés dérivés de *Streptomyces* et 2500 des rares espèces d'actinomycètes (Nair & Abraham, 2020). Le groupe rare d'actinomycètes représente environ 46% des producteurs de métabolites bioactifs. En regardant plus en détail la littérature des 5 dernières années, depuis 2017, de nombreux métabolites secondaires bioactifs ont été découverts à partir de source marine, La plupart de ces nouveaux composés biologiquement actifs intégrant la classe antibiotique polykétide (Rai & Bai, 2022).

### **I.6.1. Principales substances bioactives issues des actinobactéries**

Au total, près de 80 années de recherche sur les actinobactéries ont abouti à la découverte de nombreux médicaments cruciaux qui sauvent encore aujourd'hui de nombreuses vies humaines dans le monde entier (Jose et al., 2021). Le groupe des actinobactéries, et en particulier le genre *Streptomyces*, qui a été historiquement le plus important producteur d'antibiotiques connus, constitue toujours une source prometteuse pour la découverte de nouveaux composés (Rd et al., 2011; Bj et al., 2015).

L'alnumycine, les munumbicines A à D et les coronamycines ont été produit par diverses espèces de *Streptomyces* et qui présentent différentes activités biologiques, notamment des agents antimicrobiens, anticancéreux, immunosuppresseurs, larvicides et antiviraux (Benhadj et al., 2020). De plus, les deux anticancéreux importants, soit les anthraquinones et les lupinacidines A et B, ont été synthétisés par le ce genre (Supong et al., 2016).



### I.6.1.1. Les antibiotiques

Les actinobactéries produisent un grand nombre d'antibiotiques de structure chimique très variées (aminoglycosides, anthracyclines, glycopeptides, beta-lactamines, tetracyclines, macrolides, nucléosides ...) qui ont de nombreuses applications thérapeutiques (Imada & Hotta, 1992) .

Les substances antibiotiques sont importantes dans la lutte contre les maladies infectieuses qui peuvent causer la mort due à la résistance aux médicaments par des agents pathogènes. Les pathogènes ont développé une résistance contre divers composés antibactériens (Karthik, 2022).

Parmi les genres connus pour la production des composés antimicrobiens : *Streptomyces*, *Micromonospora*, *Kitasatospora*, *Amycolatopsis*, *Actinoplanes*, *Saccharopolyspora*, *Nocardioopsis*, *Pseudonocardia*, et *Nocardia*. En 1940, le premier antibiotique, actinomycine, a été isolé de *Streptomyces antibioticus*. Après ce temps, on a découvert de nombreux antibiotiques de différentes espèces d'actinobactéries (Barka et al., 2015; C et al., 2019). Les actinobactéries marines sont utilisés pour développer de nouveaux médicaments efficaces sans aucun effets secondaires (Karthik, 2022).

Actuellement, plus de 60% des actinomycètes synthétisent des antibiotiques qui peuvent être utilisés contre diverses infections, notamment le groupe des *Streptomyces* (Abdullah Al-Dhabi et al., 2020). Le tableau 2 montre quelques exemples d'antibiotiques produits par les actinobactéries.

**Tableau 2.** Exemples des antibiotiques produits par les actinobactéries.

Actinomycètes	Substances	Références
<i>Amycolatopsis sp. IRD-009</i>	Pradimicin-IRD	(Bauermeister et al., 2019)
<i>Kocuria marina CMG S2</i>	Kocumarin	(Uzair et al., 2018)
<i>Micromonospora sp. CA-214,671</i>	Phocoenamicin B	(Pérez-Bonilla et al., 2018)
<i>Micromonospora sp. TP-A0468</i>	16-demethylrifamycin S	(Zhou et al., 2019)
<i>S. Globisporus subsp. Globisporus</i>	Globimycin	(« Editorial Board », 2018)
<i>S. Lusitanus OUCT16-27*</i>	Grincamycin L	(Yang et al., 2019)
<i>Streptomyces sp. 120,454</i>	Mayamycin B	(Bo et al., 2018)
<i>Streptomyces sp. MMI68-141F8</i>	Quadoctomycin	(Jose et al., 2021)

<i>Verrucosipora sp. FIM06-0036</i>	2-ethylhexyl1H-imidazole-4-carboxylate	(Chen et al., 2021)
<i>Micromonospora harpali</i>	Microsporانات A_F	(Gui et al., 2017)
<i>Streptomyces sp. 8P21H-1</i>	Griseoviridin	(Wang et al., 2021)

### I.6.1.2. Les antifongiques

De vastes travaux de recherche sont en cours pour identifier de nouveaux composés antifongiques d'actinobactéries inhabituelles qui sont efficaces contre les champignons pathogènes (Wanner, 2009). Par exemple, les composés bioactifs tels que la scopafungine, la candiplanecine, rapamycine, phtalates, etc., ont des propriétés antifongiques extraites des actinobactéries marines. Certains enzymes et leurs dérivés (Ex. La chitinase) sont connus pour leur pouvoir antifongique hautement biocompatible. Ces métabolites secondaires ont été exploités pour leurs applications dans le domaine de la biomédecine, par exemple, le cartilage, l'ingénierie tissulaire, le transport de médicaments et la régénération nerveuse (Patra et al., 2020).

Une étude a montré qu'un pourcentage élevé de souche d'actinobactéries avaient une activité antagoniste contre les champignons phytopathogènes. 56 des 107 souches testées (52,3 %) ont montré une activité antifongique contre les champignons phytopathogènes. Ces résultats suggèrent que les actinobactéries représentent une ressource prometteuse pour explorer de nouveaux antifongiques (Wang et al., 2020).

Plus récemment, (Sangkanu et al., 2021) ont extrait et identifié le n-hexadécanoïque acide, acide tétradécanoïque et acide pentadécanoïque de *Streptomyces sp.* Tous ces composés étaient capables d'inhiber *Talaromyces marneffeii*, un champignon pathogène dimorphique (Karthik, 2022).

### I.6.1.3. Les substances antioxydantes

Un antioxydant est une substance ou un composé qui évite la détérioration, les dommages ou la destruction par le processus d'oxydation. Ces composés bioactifs confèrent une protection contre le stress oxydatif dans un organisme en bloquant ou en retardant les dommages oxydatifs causés par les espèces réactives d'oxygène (ROS). Les antioxydants utilisent de

multiples mécanismes pour éliminer les ROS, comme l'inhibition de la formation de radicaux libres, la récupération de molécules d'oxygène et la chélation des prooxydants métalliques. Une liste des composés antioxydants extraits de rares actinobactéries (écosystème marin) est présentée dans le tableau 3, ci-dessous (Nair & Abraham, 2020).

Les métabolites pyrrolo[1-a]pyrazine-1, 4-dione et hexahydro-3(2-méthylpropyl) sont identifiés dans *Streptomyces sp.* MUM 256 et *Streptomyces malaysiense sp.* MUSC 136 (Tan et al., 2015). Ce composé a montré une activité antioxydante puissante et une activité cytotoxique élevée contre les cellules HCT-116 (Ser et al., 2016), et potentiellement actif contre les cellules cancéreuses du poumon (A549) et du col de l'utérus (hela) en fonction de la dose (Lalitha et al., 2016).

**Tableau 3.** Exemples de substances anti-oxydantes extraites des actinobactéries.

Actinomycètes	Substances	Références
<i>Pseudonocardia autotrophica</i> KCTC9441	Polyene B1	(Park et al., 2020)
<i>Saccharothrix yanglingensis</i> Hhs.015	10-deoxyfungichromin (WH02)	(Wang et al., 2019)
<i>S. Morookaense</i> AM25	Gloeosporiocide	(Slany et al., 2019)
<i>Umezawaea sp.</i> RD066910 <i>Tsukamurella pulmonis</i> TP-B0596	Umezawamides A	(Hoshino et al., 2018)

#### I.6.1.4. Les substances anti- cancéreuses

Les métabolites secondaires produits par les actinobactéries ont une activité antitumorale (Ex., bléomycine et doxorubicine) (Tableau 4) (Grasso et al., 2016).

*Nocardiosis sp.* G057 a été identifié pour produire 12 nouveaux composés. Ces composés ont montré une forte activité cytotoxique contre les lignées cellulaires des tumeurs à kératine (KB), des lignées cellulaires du cancer du poumon (LU-1), des lignées cellulaires du cancer du foie humain (hepg-2) et des cellules cancéreuses du sein (MCF-7) (Thi et al., 2016).

La cytotoxicité des niphimycines C, E et F étaient évidentes contre les lignées cellulaires hela humaines cancéreuses (CI50 : de 3,0 à 9,0 µm). N-acétylborrelidine B (3 un nouveau microlide naturel antibiotique a été obtenu de *Streptomyces mutabilis sp.* MII qui a démontré un

effet cytotoxique puissant même dans l'extrait brut contre les lignées cellulaires cancéreuses du col de l'utérus humain (KB-3-1) dans des conditions *in vitro* (Hamed et al., 2018).

La culture fermentative de *Streptomyces sp.* Le SCSIO 41 a permis la production de l'aranciamycine et à l'isotirandamycine qui ont montré des activités cytotoxiques *in vitro* contre les lignées cellulaires K560 (Cong et al., 2019).

La désertomycine G (50) a été obtenue de *Streptomyces althioticus* MSM3, la première fois que l'on signale une activité anti-tumorale contre les cellules d'adénocarcinome colorectal (DLD-1) et adénocarcinome du cancer du sein humain (MCK-7) (Braña et al., 2019).

**Tableau 4.** Exemples de substances anti-tumorales produites par les actinobactéries.

Actinomycètes	Substances	Références
<i>Actinoalloteichus hymeniacidonis</i> 179DD-027	Dokdolipid B	(Choi et al., 2019)
<i>Micromonospora sp.</i> HS-HM-036	Naphthalenepropanoic acid analog	(Gao et al., 2017)
<i>Nonomuraea sp.</i> AKA32	Akazamicin	(Yang et al., 2019)
<i>Saccharomonospora sp.</i> UR22 <i>Dietzia sp.</i> UR66#	Saccharomonosporine A	(El-Hawary et al., 2018)
<i>S. Curacoï</i> NBRC 12,761	Curacozole	(Lee et al., 2020)
<i>Streptomyces sp.</i> 166	Sekgranaticin	(Lv et al., 2019)

### I.6.1.5. Les enzymes

Les actinobactéries sont signalés pour produire un grand nombre d'enzymes extracellulaires tels que la cellulase, amylase, chitinase, lipase, protéase, xylanase et pectinases (Hazarika & Thakur, 2020)(voir tableau 5). Ces enzymes sont d'une importance vitale dans les industries du textile ou du papier, l'industrie alimentaire et la fermentation, en plus des applications biotechnologiques. Par conséquent, les actinobactéries se sont révélées être une bonne source de L-asparaginase. La kératinase est une enzyme produite partir de *Nocradiopsis sp.* (Verma et al., 2017).

Diverses actinobactéries, souvent isolées des sols, comme *Nocardia sp.*, *Streptomyces albidoflavus*, *S. Griseus* et *Streptomyces karnatakensis* produisent plusieurs enzymes (Lewin et al., 2016).

**Tableau 5.** quelques enzymes produits par actinobactéries(Hazarika & Thakur, 2020).

Enzymes	Les genres
Amylase	<i>Streptomyces erumpens</i> and <i>Termobifida fusca</i>
Cellulase	<i>Streptomyces</i> , <i>Cellulomonas</i> , <i>Micromonospora</i> , <i>Actinopolyspora</i> , <i>Actinoplanes</i> , <i>Microbispora</i> , <i>Thermomonospora</i> , <i>Rhodococcus</i> , <i>Nocardia</i> and <i>Thermoactinomyce</i>
Chitinase	<i>Streptomyces thermoviolaceus</i> <i>Nocardiopsis prasina</i> , <i>Microbispora spp.</i>
Lipase	<i>Streptomyces exfoliates</i> <i>Nocardiopsis alba</i> , <i>Microbacterium spp.</i>
Protease	<i>Streptomyces</i> , <i>Nocardia</i> and <i>Nocardiopsis</i>
Xylanase	<i>Actinomadura spp.</i> <i>Streptomyces spp.</i>

---

---

## CHAPITRE II

---

---

### **II. Généralités sur les groupes de gènes biosynthétiques PKS et NRPS (BGC)**

Les gènes impliqués dans les voies biosynthétiques des métabolites secondaires se regroupent en clusters appelés Biosynthetic Gene Clusters (BGC). Ces BGC sont composés d'un gène central codant une enzyme structurale définissant le composé principal du métabolite, et de gènes adjacents codant des enzymes de modification pour produire des variants (Keller, 2015).

Les BGC permettent la régulation coordonnée de l'expression des gènes et la production efficace des métabolites secondaires. Ils sont cruciaux pour la découverte de nouveaux composés bioactifs et le développement de médicaments, produits agricoles et industriels. L'étude des BGC implique des approches bioinformatiques et expérimentales, elle consiste en l'analyse des séquences génomiques, la comparaison avec des bases de données existantes, et l'étude de l'expression des gènes et la production des métabolites (Marcia Soares Almeida, 2022).

#### **II.1. Les peptides bactériens non ribosomiques (NRP)**

Certains micro-organismes, tels que les bactéries et les champignons, utilisent une voie de synthèse peptidique différente de la synthèse ribosomiale pour produire des molécules nécessaires à leur développement et à leur survie. Cette voie de synthèse, appelée synthèse peptidique non-ribosomiale, ne dépend pas de l'intervention des ribosomes. Cette dernière est remarquable en raison de la diversité des activités biologiques des produits qui en résultent. Elle est caractéristique des micro-organismes et n'a pas encore été observée chez les organismes supérieurs (Vanvlassenbroeck, 2012).

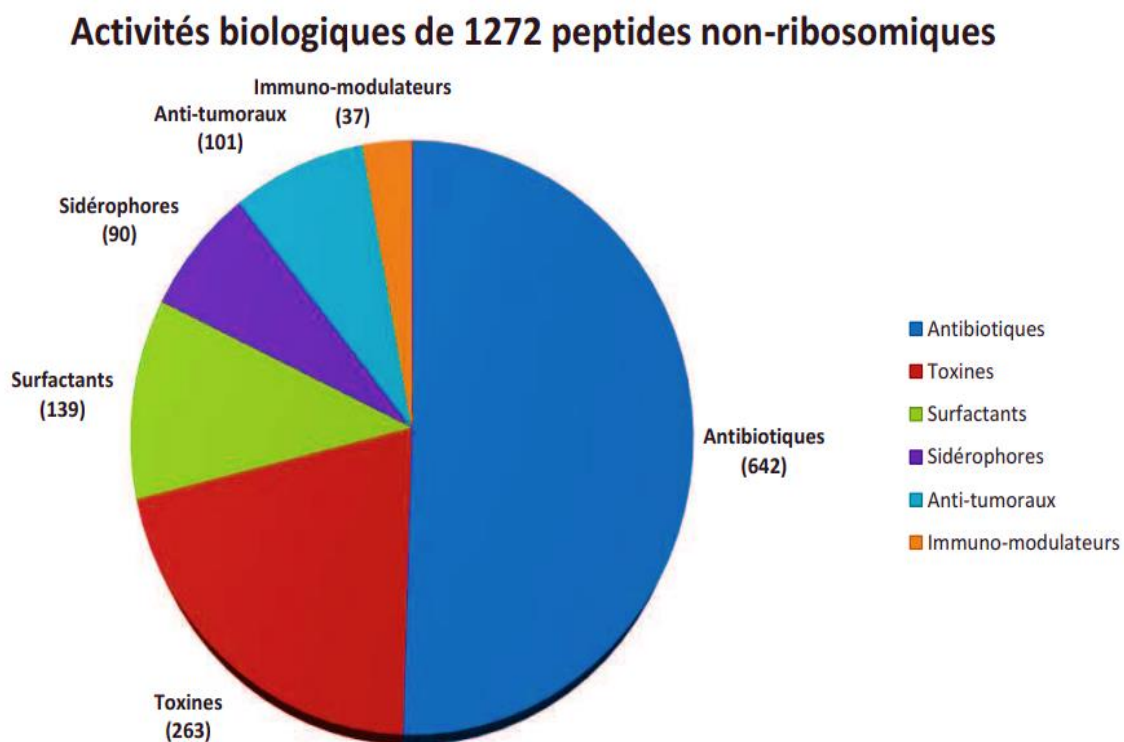
##### **II.1.1. Voie de synthèse des protéines**

La synthèse ribosomique est la voie de synthèse des protéines. L'ADN est d'abord transcrit en ARN à l'aide d'une polymérase, puis, comme son nom l'indique, cette synthèse utilise des organites ribonucléoprotéiques appelés ribosomes. Cela aligne les acides aminés un par un à partir de l'ARN messager pour générer la séquence protéique primaire (Chevalier, 2018).

Les acides aminés impliqués dans la synthèse des protéines sont appelés acides aminés protéinogènes (Ambrogelly et al., 2007). Il en existe typiquement 20, et seuls deux autres (sélénocystéine et pyrrolysine) ont été récemment décrits chez les microbes (Chevalier, 2018).

### II.1.2. Voie de synthèses des peptides bactériens non ribosomiques (NRP)

Les peptides non ribosomiques sont un groupe de métabolites secondaires et de produits non associés à la croissance qui n'utilisent pas la machinerie traductionnelle médiée par les ribosomes de la cellule pour la synthèse. Au lieu de cela, ils sont synthétisés par des enzymes multi-domaines modulaires appelées synthétases peptidiques non ribosomales (NRPS) (Martínez-Núñez & López, 2016). L'utilisation d'un tel système permet la synthèse d'un peptide court avec de nombreuses caractéristiques uniques, telles que des structures macrocycliques et ramifiées, la présence à la fois d'acides aminés protéinogènes et non protéinogènes. Ces caractéristiques structurelles leur confèrent une diversité fonctionnelle, allant des propriétés tensioactives aux activités antimicrobiennes et cytotoxiques (Figure 3) (Sieber & Marahiel, 2003; Singh et al., 2012).



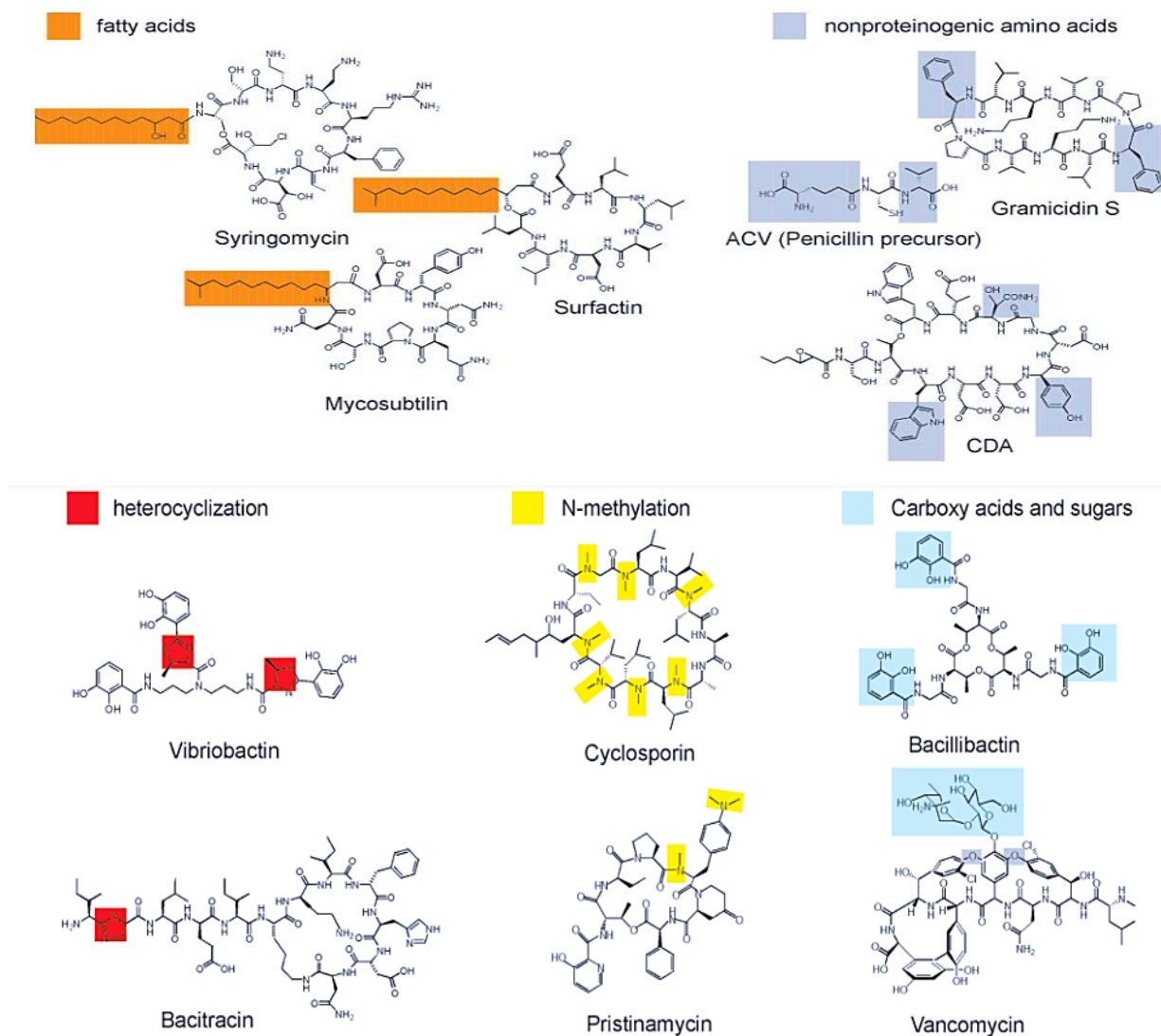
**Figure 3.** Répartition des 6 classes d'activités biologiques de 1272 peptides non-ribosomiques recensés dans la base de données Norine (Tambadou, 2014).

Contrairement à la synthèse peptidique classique qui utilise les 20 acides aminés protéogéniques, la biosynthèse non-ribosomique utilise en plus de ces 20 acides aminés d'autres molécules non

protéogéniques. C'est pour cette raison que nous utiliserons le terme de monomères pour désigner les molécules précurseurs incorporées dans les peptides non-ribosomiques. Actuellement le nombre de monomères identifiés par la base de données Norine sur l'ensemble des peptides non-ribosomiques est supérieur à 500 (Caboche et al., 2008, 2010). Les synthétases peuvent aussi contenir des domaines facultatifs qui sont capables de modifier la structure des monomères incorporés, ce qui augmente encore la diversité de ces peptides. Des produits tels que des lipides ou des hydrates de carbone issus d'autres voies de biosynthèse peuvent également être incorporés (Tambadou, 2014).

En plus de la diversité de leur composition, les peptides non-ribosomiques présentent des structures particulières qui sont souvent complexes. On retrouve des structures linéaires (comme le tripeptide ACV, précurseur de la pénicilline), branchées (la vibriobactine), partiellement cycliques (les polymyxines), totalement cycliques (la cyclosporine A), voire complexes (la vancomycine) (Caboche et al., 2010). Ces structures peuvent contenir des N-méthylations, des N-formylations, des glycosylations, des insertions d'acétate ou de propionate ainsi que des acides gras (Schwarzer et al., 2003) (Figure 4). Ainsi, la diversité des monomères et des structures permet d'obtenir des molécules complexes ayant des propriétés pharmacologiques intéressantes.





**Figure 4.** Produits peptidiques naturels. Une sélection de peptides non synthétisés par les ribosomes. Les caractéristiques structurales sont mis en évidence (Sieber & Marahiel, 2005).

En conséquence, les NRP bactériens possèdent une haute valeur en tant que molécules bioactives cliniquement et industriellement pertinentes, avec des applications telles que les antibiotiques, les médicaments anticancéreux, la dégradation des déversements d'hydrocarbures, l'assainissement de l'environnement, la récupération microbienne améliorée du pétrole, la thérapie par chélation des métaux, les biocapteurs, les biopesticides et bien d'autres (McIntyre et al., 1999; Mukherjee et al., 2009; Sen, 2010a; Roongsawang et al., 2011; Khan et al., 2018).

Selon la base de données BiG-FAM (à partir de 2022/8), le NRPS est la classe la plus répandue (29,30 %) sur les 1 060 938 groupes de gènes biosynthétiques (BGC) détectés à partir de 170 585 génomes bactériens, et 46,81 % sur les 123 939 BGC de 5 588 génomes fongiques (Kautsar et al., 2021)

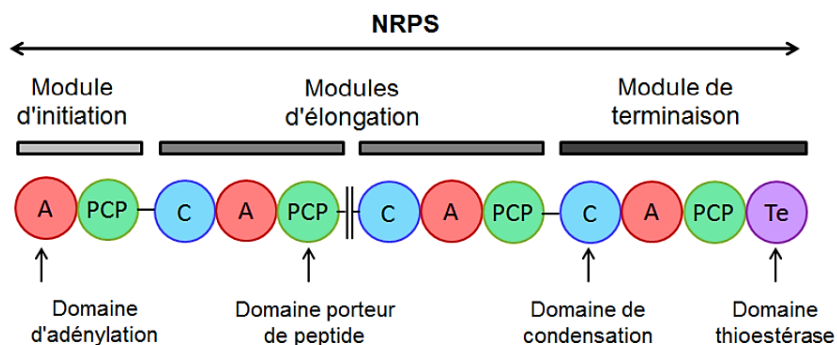
Les membres de la famille des NRP bactériens comprennent les lipopeptides (surfactine, iturine, engycine), les glycopeptides (vancomycine), les peptides cycliques (bacitracine, tyrocidine) et les sidérophores (entérobactine, ornibactine) et leurs structures correspondantes ont été données dans la figure 4 (Mitra et al., 2022).

Malgré l'immense importance commerciale de ces composés, les NRP bactériens ont acquis la réputation d'être des produits « de grande valeur et à faible volume » (Sen, 2010), en raison de leur faible rendement et des pertes importantes lors de la purification. De nombreuses tentatives ont été faites pour améliorer le rendement et l'efficacité de la production en isolant des souches productrices très efficaces et en optimisant ultérieurement les composants moyens et le processus (Mitra et al., 2022).

### II.1.3. Organisation des NRPS

Les peptides non ribosomaux constituent l'une des principales familles de métabolites secondaires. Ces molécules sont assemblées par des méga complexes enzymatiques appelés synthétases de peptides non ribosomaux (non ribosomal peptide synthetase : NRPS).

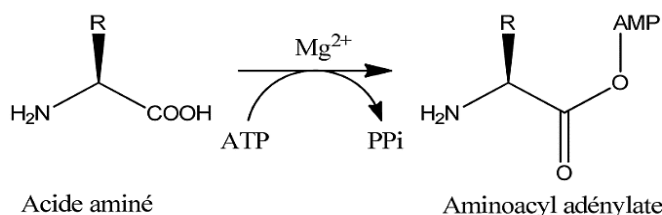
Les NRPS sont constituées d'un ou plusieurs modules qui ont pour rôle l'incorporation d'un acide aminé dans la chaîne d'assemblage du peptide. Chacun de ces modules est lui-même subdivisé en domaines. Tous les modules contiennent obligatoirement le domaine d'adénylation (A), le domaine porteur de peptide (peptidyl carrier protein : PCP) et le domaine de condensation (C), à l'exception du module d'initiation qui ne contient pas de domaine C. Il existe d'autres domaines, accessoires, qui ne sont pas nécessairement présents mais qui permettent d'apporter des modifications au peptide en cours de synthèse. Ce sont, par exemple, les domaines méthyltransferase (MT) ou épimérase (E). Enfin, le dernier module contient un domaine thioestérase (Te) capable de libérer le peptide (Figure 5)(Haas, 2015).



**Figure 5.** Chaîne d'assemblage classique d'une NRPS (Haas, 2015).

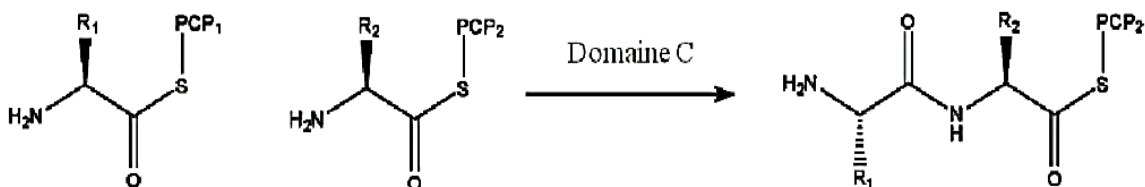
NRPS synthétise les peptides à la chaîne en utilisant ses unités de module répétitives. Chaque module est généralement composé de trois domaines clés :

- **Le domaine d'adénylation (A)**, qui reconnaît le substrat et l'active en tant qu'adénylate d'aminocyclo. Ce domaine est responsable non seulement de l'incorporation plus ou moins spécifique du monomère mais aussi de son activation en aminoacyl-adénylate. Cette forme « activée », de l'acide aminé, lui apporte l'énergie nécessaire pour être utilisé par le domaine suivant (Figure 6)(Chevalier, 2018).



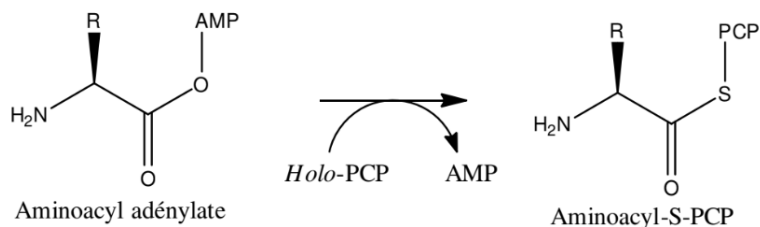
**Figure 6.** Réaction catalysée par le domaine d'adénylation des NRPS.

- **Le domaine de condensation (C)** catalyse la formation de liaison amide entre deux substrats, il forme la liaison peptidique entre le monomère porté par le domaine PCP d'un module et le monomère porté par le domaine du module PCP suivant (Figure 7) (Chevalier, 2018).



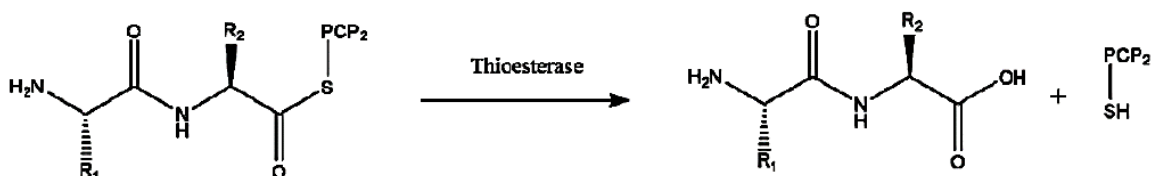
**Figure 7.** Réaction catalysée par le domaine de condensation des NRPS.

- **Le domaine de thiolation (T)**, fait la navette entre les substrats et les intermédiaires peptidiques entre les domaines catalytiques (He et al., 2023). Le domaine de thiolation (domaine T, ou peptidyl carrier protein, PCP) porte un groupement prosthétique dérivé de la pantéthéine (phosphopantéthéine) qui réalise la liaison entre la synthétase et le monomère afin d'assurer le transport au cours de l'élongation (Figure 8) (Chevalier, 2018).



**Figure 8.** Réaction catalysée par le domaine de thiolation des NRPS.

- D'autres domaines facultatifs sont également importants, tels que le domaine thioestérase (TE) ou le domaine terminal de type condensation ( $C_T$ ) (Gao et al., 2012), qui cyclisent ou libèrent des peptides du NRPS et se produisent généralement à la fin d'un NRPS (Figure 9); domaine d'épimérisation (E), qui catalyse la conversion du substrat d'acides aminés L en acides aminés D et suit généralement le domaine C au cas où cela se produirait (He et al., 2023).



**Figure 9.** Réaction catalysée par le domaine de thioestérase des NRPS.

#### II.1.4. La détection des NRPS

Afin de créer de nouveaux produits, de nombreux efforts de synthèse ont été faits pour réorganiser ces éléments constitutifs des NRPS. Par exemple, en raison du rôle central des domaines A dans la reconnaissance et l'activation du substrat, les premiers travaux se sont concentrés sur ce domaine, y compris sa substitution, la reprogrammation des résidus codant pour le substrat et la substitution du sous-domaine A. Plus tard, compte tenu des rôles possibles des domaines C dans la sélectivité des substrats, les chercheurs ont tenté de réorganiser les domaines catalytiques avec le domaine A, par exemple en substituant un module complet, ou le domaine C+A et le domaine T+C+A (Brown et al., 2018). Une compréhension systématique de la conception rationnelle du NRPS reste à développer (He et al., 2023).

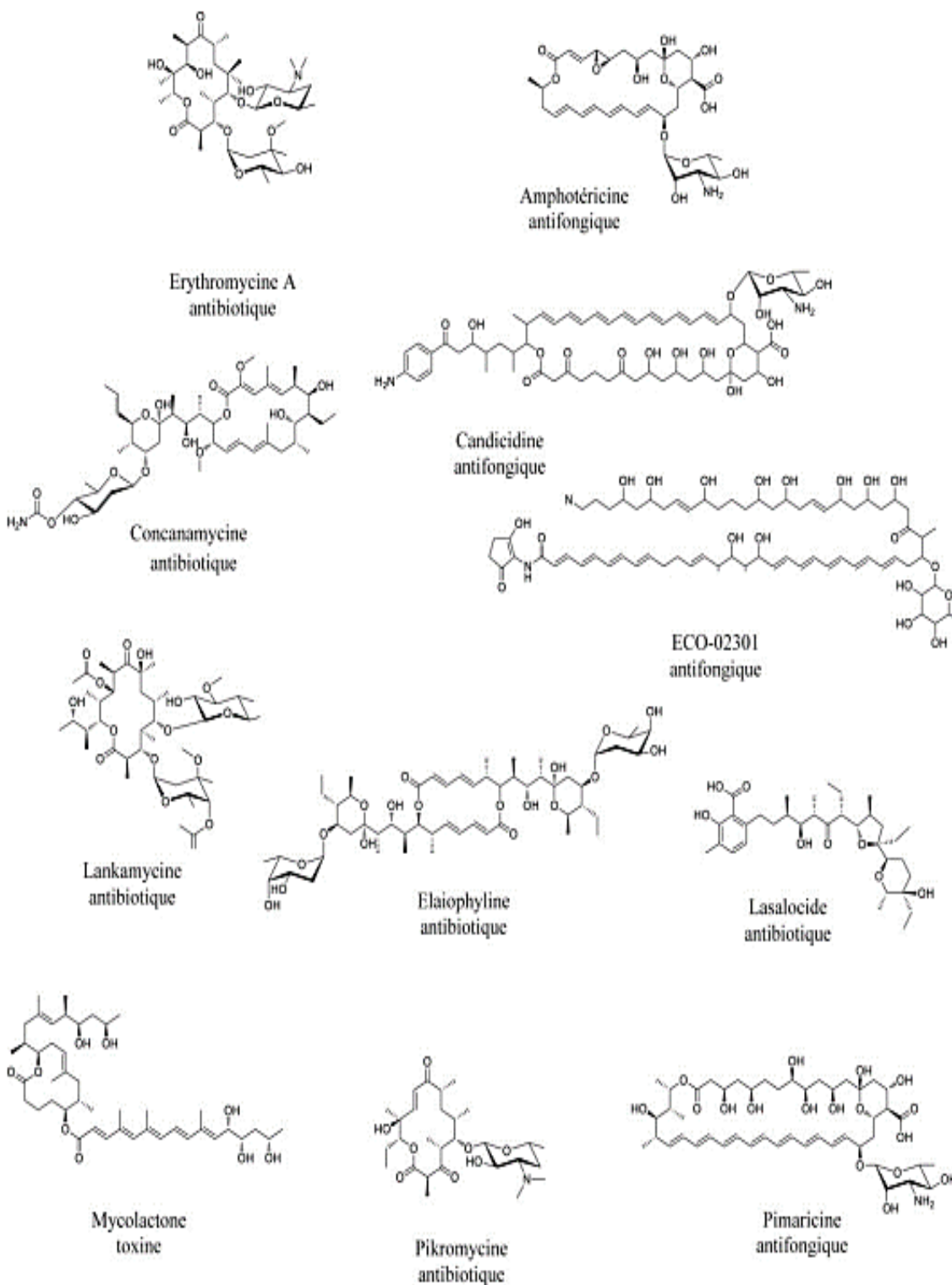
L'expansion massive des données de séquençage microbien a créé des opportunités pour une compréhension globale des NRPS (Medema et al., 2014). En octobre 2021, la base de données Integrated Microbial Genomes Atlas of Biosynthetic gene Clusters (IMG-ABC) contient un total de 411 027 BGC basés sur l'annotation de séquences (Palaniappan et al., 2020), et la base de données Minimum

Information about a Biosynthetic Gene Cluster (MiBiG) contient 1926 BGC confirmés biologiquement actifs (Kautsar et al., 2021). Parallèlement à l'expansion accélérée des bases de données, les plateformes bio-informatiques qui prédisent et annotent spécifiquement les métabolites secondaires sont également en développement rapide (Medema & Fischbach, 2015). Ils ont été utilisés pour aider à la découverte de nouveaux antibiotiques NRPS (Park et al., 2017). Cependant, différents outils et plates-formes bio-informatiques utilisent différentes normes et algorithmes d'annotation, et leurs définitions de domaines et inter-domaines dans NRPS sont clarifiées (He et al., 2023).

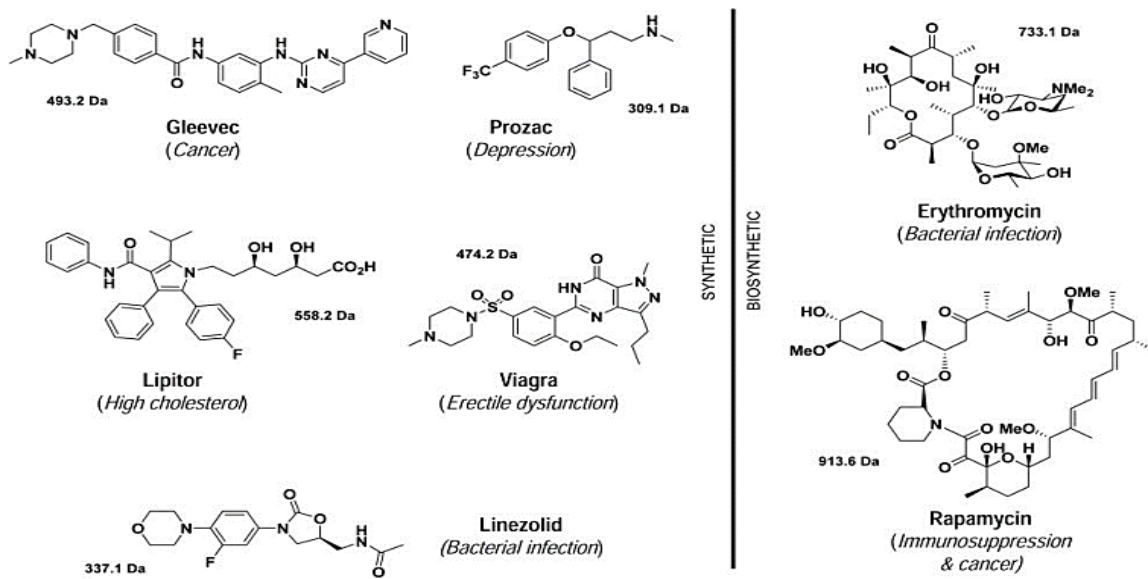
## **II.2. Les gènes biosynthétiques de type PKS**

Les polykétides constituent une grande famille de produits naturels que l'on trouve dans les bactéries, les champignons et les plantes. Ils comprennent de nombreux médicaments d'importance clinique tels que la tétracycline (antibiotiques), la daunorubicine (anticancéreux), l'érythromycine (antibiotique), la rapamycine (immunosuppresseur) et la lovastatine (anticholestérolémique) (B, 2003). Les polykétides jouent donc un rôle important dans le domaine médical et présentent un grand intérêt économique.

En plus de leur diversité fonctionnelle, les polykétides présentent une riche diversité structurale (Figure 10). Leur complexité structurale est supérieure à celle des molécules synthétiques utilisées en médecine (Figure 11). En général, les polykétides sont des molécules rigides caractérisées par une haute densité de groupes fonctionnels avec une stéréochimie bien définie (Clardy & Walsh, 2004).



**Figure 10.** Exemples de la complexité et de la diversité des polykétides et de leur fonction associée (Annaval, 2015).



**Figure 11.** Comparaison des molécules issues de synthèse et de biosynthèse utilisées en médecine (Clardy & Walsh, 2004).

Les polykétides synthases (PKS) sont des enzymes composées d'un ou plusieurs domaines qui jouent un rôle essentiel dans la synthèse des polykétides, puis de les incorporer dans des chaînes en croissance et de les modifier. La voie chimique de synthèse des polycétides est très proche de la synthèse des acides gras, la lipogenèse, par les synthases d'acides gras (FAS : fat acid synthases). Cette homologie mécaniste est en corrélation avec l'homologie au niveau de la séquence primaire de ces deux types de synthases, suggérant leur origine évolutive commune (Annavaal, 2015).

### II.2.1. Découvert et répartition génétique

La compréhension des mécanismes de biosynthèse des polykétides n'a considérablement progressé qu'à partir de 1953, lorsque Birch et Donovan ont émis l'hypothèse d'une voie de biosynthèse potentielle similaire à celle des acides gras (Birch & Donovan, 1953).

Depuis les premiers rapports sur les PKS bactériennes de type I en 1990 (Cortes et al., 1990; Donadio et al., 1991), les PKS de type II en 1984 (Malpartida & Hopwood, 1984; Motamedi & Hutchinson, 1987) et les PKS de type III en 1999 (Funa et al., 1999). Les paradigmes des PKS ont servi la communauté scientifique au-delà de l'appel du devoir, en fournissant la base moléculaire pour expliquer la vaste diversité structurale observée dans les produits naturels polykétidiques, et la plate-forme

biotechnologique pour produire des produits naturels "non naturels" par des méthodes de biosynthèse combinatoire avec des PKS modifiées.

D'autres avancées dans le domaine des polykétides ont conduit à un consensus selon lequel les polykétides sont généralement biosynthétisés par condensations décarboxylatives successives d'unités dérivées du coenzyme A (CoA), en un composé polycyclique complexe à plusieurs carbones contenant des groupes céto ou hydroxyle (Dunn & Khosla, 2013). Des enzymes modificatrices en aval sont ensuite utilisées pour obtenir des flavonoïdes et d'autres substances bioactives.

Des études récentes sur l'ingénierie et la caractérisation structurale des polykétides synthases (PKS) ont facilité l'utilisation d'enzymes cibles comme biocatalyseurs pour produire de nouveaux polykétides optimisés sur le plan fonctionnel, qui peuvent servir de pistes médicamenteuses potentielles [(Abe, 2012), (Till & Race, 2014)].

### II.2.2. Les types PKS et leur capacité biosynthétiques

Les PKS ont été classées en trois grandes catégories : les types I, II, et le type III. Cependant, ces distinctions ne sont ni rigoureuses ni rapides, car de nombreuses PKS présentent un comportement hybride (Moss et al., 2004; Wenzel & Müller, 2005). Les PKS de type II et III génèrent toutes deux des molécules aromatiques, mais utilisent une approche fondamentalement différente pour construire une chaîne polycarbonyle réactive et la canaliser vers un destin cyclique particulier (Tableau 6).

**Tableau 6 :** Caractéristiques des PKS de type I,II,III (Weissman, 2009).

Type de PKS	Mode de fonctionnement	Mode d'activation du substrat	Produit typique
I	Modulaire	ACP	Réduit
I	Iterative	ACP	Aromatique et réduit
II	Iterative	ACP	Aromatique
III	Iterative	CoA	Aromatique

#### II.2. 2. 1. Les PKS de type I

Les PKS de type I se sont révélés être à la fois les plus compliqués et les plus polyvalents des trois classes de catalyseurs. La caractéristique diagnostique de ces systèmes est la présence de sites actifs multiples dans chaque polypeptide. Son mode opératoire est de type soit itératif soit modulaire (Cox, 2007).



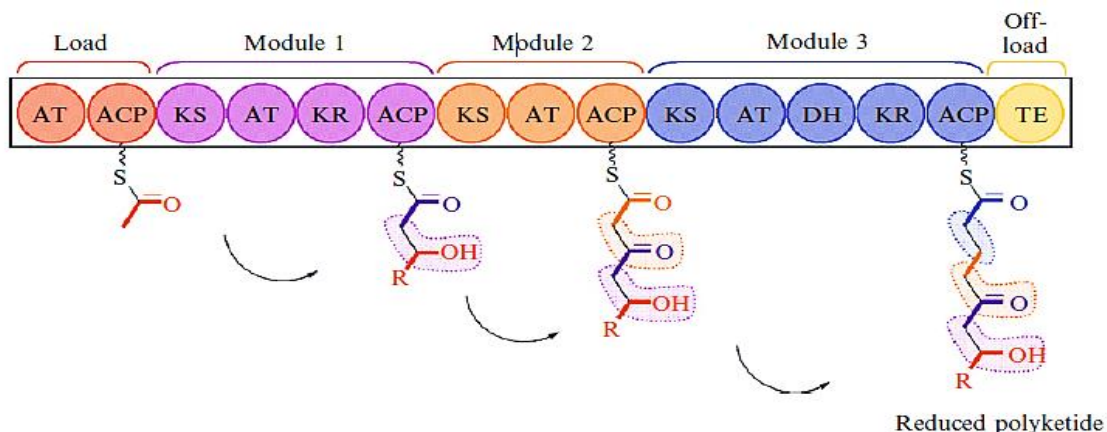
Ce type de PKS se compose de modules qui eux-mêmes consistent en un ensemble de domaines minimaux requis pour incorporer de nouveaux blocs de construction : AT, ACP et KS. Des domaines optionnels peuvent être ajoutés à cet ensemble minimal de domaines pour permettre la modification des blocs de construction : KR, DH, ER et MT (Annavaal, 2015).

### **A. Les PKS de type I itératives**

Les PKS itératives de type I peuvent être subdivisées en trois types, en fonction du degré de réduction de la PK résultante : les PKS non réductrices (NR), partiellement réductrices (PR) et fortement réductrices (HR) (Robbins et al., 2016). Ils contiennent au moins un module central minimal de trois domaines - kétosynthase (KS), acyltransférase (AT) et un domaine de protéine porteuse d'acyle (ACP) - pour catalyser la condensation de Claisen tête-bêche de l'acétyl-CoA (Hertweck, 2009; Herbst et al., 2018). Pour les PKS itératives de type I, l'extension de chaîne polykétide se produit dans un seul module. Par conséquent, une seule multi-enzyme est utilisée à plusieurs reprises. Pour ses produits, Itérative PKS permet la synthèse d'acides gras polyinsaturés à longue chaîne (PUFA), notamment dans les membranes cellulaires des bactéries marines (Metz et al., 2001). La PKS itérative permet également la synthèse de polykétides de type ènediène dans les bactéries. Ces enzymes partagent une forte homologie de séquence avec celles qui synthétisent les acides gras polyinsaturés (Van Lanen & Shen, 2008).

### **B. Les PKS de type I modulaires**

Contrairement aux PKS itératives (mono-modulaires), les PKS de type I modulaires ajoutent des unités d'extension spécifiques au niveau de différents modules successifs. Ainsi un module, à l'intérieur d'une grande chaîne polypeptidique, ne catalyse la condensation et la modification que d'un acyl-CoA à la chaîne en cours de croissance (Cox, 2007).



**Figure 12.** Représentation d'une PKS de type I modulaire (Weissman, 2009).

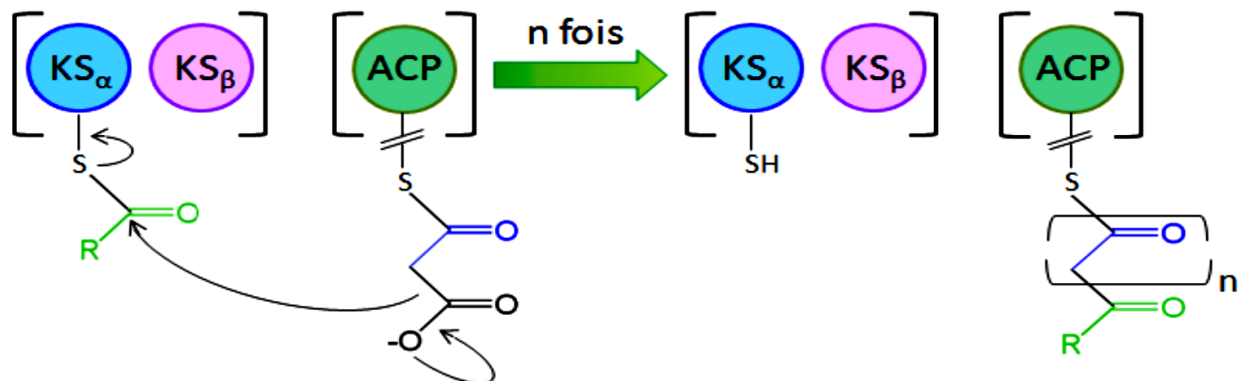
Les modules individuels des PKS de la chaîne d'assemblage sont classés en modules cis-AT et trans-AT. Les modules cis-AT contiennent les trois domaines essentiels (KS, AT et ACP) comprenant une PKS, tandis que dans les modules trans-AT, l'unité d'extension est transacylée sur le domaine ACP par un AT autonome qui est souvent partagé par plusieurs modules (Helfrich & Piel, 2016) (Figure 12). Les modules sont reliés soit par des connecteurs intermodulaires (Gokhale et al., 1999), soit, si une PKS s'étend sur plusieurs polypeptides, par des domaines d'amarrage qui établissent des interactions non covalentes spécifiques entre les modules successifs (Broadhurst et al., 2003). Les architectures des PKS cis-AT sont souvent colinéaires à leur codage génétique (c'est-à-dire que l'ordre dans lequel les modules sont codés au niveau de l'ADN correspond à l'ordre dans lequel ils opèrent), alors que les modules des lignes d'assemblage trans-AT ne sont souvent pas colinéaires (Helfrich & Piel, 2016).

### II.2. 2. 2. Les PKS de type II

Les PKS de type II sont des complexes de protéines discrètes agissant de manière itérative, chacune d'entre elles ayant une fonction spécifique dans la voie. Elles possèdent une forte homologie avec les synthèses d'acides gras bactériennes, tant au niveau des mécanismes d'extension de chaîne que dans la composition de certains domaines (Hertweck et al., 2007).

L'initiation de la chaîne (le plus souvent avec l'acétate) et l'élongation avec le malonate sont accomplies par la "PKS minimale" qui consiste en deux enzymes de condensation de type kétosynthase ( $KS_{\alpha}$  et  $KS_{\beta}$ , dont seule la première fournit le site actif pour la condensation), et une acyle carrier protein (ACP), à laquelle la chaîne en croissance est attachée de manière covalente sous la forme d'un thioester. Une activité malonyl-CoA:ACP transférase (MCAT) contribue probablement à recruter des unités d'extension à partir du métabolisme primaire. D'autres sous-unités enzymatiques (kétoréductases [KR],

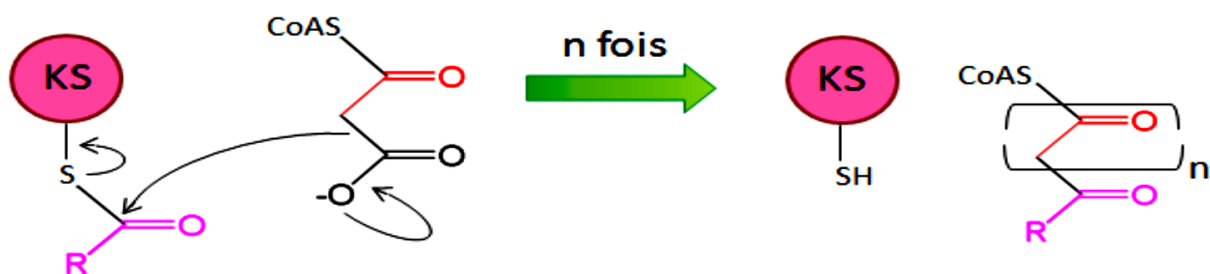
cyclases [CYC] et aromatasés [ARO] coopèrent avec la PKS minimale pour diriger le schéma de pliage de la chaîne polykétide naissante. Les produits polyphénoliques de la PKS sont souvent remodelés en profondeur par des oxygénases et décorés par des glycosyl- et N-, O- et C-méthyltransférases (Rix et al., 2002) (Figure 13à).



**Figure 13:** Schéma d'assemblage des PKS de type II (Weissman, 2009).

### II.2. 2. 3. Les PKS de type III

Ce type de PKS, également connu sous le nom de chalcone et de stilbène synthases, est très répandu chez les plantes. Elle joue un rôle clé dans la première étape de la biosynthèse des flavonoïdes (Austin & Noel, 2003). Cependant, il est également présent dans les bactéries et les champignons. Ce type d'enzyme montre une architecture simple, c'est-à-dire composée d'une kétosynthase homodimérique multifonctionnelle, contrairement aux PKS de type I et II qui présentent une organisation multienzymatique (Annaval, 2015). Les PKS de type III sont considérés comme itératives car elles utilisent ce site actif unique de KS pour subir des condensations répétées de plusieurs acyl-CoA pour former des polykétides (Figure 14). A l'issue de ces condensations répétées, une étape de cyclisation se produit, suivie d'une étape de modification du produit formé par des enzymes post-PKS. Contrairement aux PKS de type I et II, dans lesquelles l'unité d'extension est ancrée au niveau de l'ACP, l'unité d'extension des PKS de type III est toujours chargée sur la molécule de coenzyme A (Austin & Noel, 2003).



**Figure 14.** Schéma d'assemblage des PKS de type III (Weissman, 2009)

---

---

## CHAPITRE III

---

---

### **III. Annotation et prédiction génomique des PKS et NRPS chez les actinobactéries**

#### **III.1. Introduction**

Les bactéries, en particulier les actinobactéries, parmi les micro-organismes qui produisent un éventail varié de petites molécules bioactives ayant un potentiel d'utilisation important en médecine (O'Brien & Wright, 2011). Ces métabolites secondaires sont principalement formés par l'activation de groupes de gènes cryptiques qui sont inactifs dans certaines conditions; par conséquent, l'augmentation de l'expression de ces groupes pourrait aider à exploiter la variabilité chimique des micro-organismes (Pettit, 2011; Guzmán-Trampe et al., 2017; Xu et al., 2019).

Les actinobactéries sont des bactéries Gram positifs, ont un contenu élevé de G + C dans leur ADN. Ils habitent des environnements divers comme les sols salins, les sédiments d'eau douce, les plantes médicinales, les forêts profondes, les sources chaudes, les milieux marins...etc (Demain & Sanchez, 2009). En raison de leurs diverses fonctions écologiques ont de multiples applications dans plusieurs domaines (Demain et al., 2019). L'application d'outils d'omique (génomique, protéomique, transcriptomique, métabolomique et NGS) aux actinobactéries a eu une incidence sur la détection et la production de nouveaux composés bioactifs (Orsi et al., 2016).

Au cours de la dernière décennie, l'incapacité à dévoiler le plein potentiel des actinobactéries productrices de produits naturels (NP) était due au manque de compréhension de leur génome et de l'expression de groupes de gènes biosynthétiques (BGC). L'avènement de technologies de séquençage rapides et économiques a révélé que les BGC dans les actinobactéries sont en nombre plus important que les molécules déjà découvertes qu'elles synthétisent (Schorn et al., 2016; Tokovenko et al., 2016). L'évolution des outils génomiques et bioinformatiques a augmenté de façon exponentielle la compréhension des informations génétiques utiles qui peuvent être exploitées pour découvrir des composés bioactifs (Machado et al., 2015; Weber & Kim, 2016; Albarano et al., 2020).

Le séquençage du génome de *Streptomyces* a révélé qu'ils contiennent des BGC en grand nombre. *Streptomyces clavuligerus* a 58 BGC, *Streptomyces bottropensis* a 21 BGC, tandis que *S. avermitilis* contient

30 BGC (M. M. Zhang et al., 2017, p. 3). Les génomes des actinobactéries sont séquencés rapidement depuis 2013 pour la chasse à la découverte de nouvelles NP bioactives. Jusqu'à récemment, un total de 1 749 génomes de *Streptomyces* ont été déposés au 6 février 2020 et sont disponibles dans la base de données RefSeq. Les 1 749 génomes de *Streptomyces* contiennent 867 niveaux de contig, 646 niveaux d'échafaudage, 36 niveaux de chromosome, tandis que 200 génomes complets sont inclus (Lee et al., 2020).

### **III.2. Notion des multi-omiques**

Les mécanismes biologiques ne se limitent pas à l'état des molécules cellulaires elles-mêmes, mais sont étendus par la façon dont ces molécules interagissent et sont impliquées dans le processus cellulaire par des relations interdépendantes, la régulation transcriptionnelle, les interactions protéines-protéines et la régulation métabolique. Ainsi, les analyses centrées sur une seule couche omique ne peuvent pas élucider l'ensemble des réponses biologiques (Yugi et al., 2016).

Les multi-omiques, peuvent fournir une approche encore plus complète pour aborder la complexité biologique (Nguyen et al., 2022). Encouragent une approche intégrative et systémique des mécanismes d'actions liés aux substances bioactives. Les chercheurs ont incorporé des modèles d'expression génétique et de méthylation pour mieux comprendre les effets secondaires à long terme des médicaments (Selevsek et al., 2020; Hsu et al., 2021).

Les approches multi-omiques sont aussi appelées « omiques verticales » ou « trans-omiques » (Yugi et al., 2016; Hasin et al., 2017); Cependant, ils font tous référence à l'intégration de plusieurs types de données omiques pour constituer un flux d'informations et aider à révéler multiples mécanismes d'action (Nguyen et al., 2022).

Une enquête de Droste et al. (2021) a souligné que les données du transcriptome et du protéome sont précieuses pour améliorer l'annotation de la souche *Streptomyces lividans* TK24 génome. La souche TK24 pourrait être utilisée pour examiner l'analyse des groupes de gènes des métabolites secondaires.

Martín et al., (2021) ont révélé le rôle des modifications post translationnelles dans la modulation du métabolisme bactérien, en particulier sur leurs effets sur la biosynthèse des métabolites secondaires.

Singh et al., (2021) ont décrit que l'actinobactérie est bien connue pour produire divers métabolites secondaires et ont soutenu les développements de l'approche génomique comme passerelle pour l'examen et la manipulation de nouveaux groupes de gènes antibiotiques.

### **III.3. Caractéristiques du génome des actinobactéries**

#### **III.3.1. Diversité dans la taille du génome**

Ces dernières décennies, la biologie moléculaire s'est imposée comme un outil puissant et incontournable en taxonomie. À l'heure actuelle, on ne peut plus proposer une nouvelle espèce sans procéder à des analyses génétiques (Ibrahimi, 2020).

Les génomes de plusieurs espèces de *Streptomyces* ont été séquencés et assemblés. C'est le cas des génomes *Streptomyces* JA03T. La taille du génome de la souche JA03T est de 9 Mb (Wai et al., 2022) Le génome de *Streptomyces* sp. KD18 est de taille de 7,2 Mb (Khushboo et al., 2023). Les génomes de *Streptomyces* sont de grande taille (Haas, 2015), La taille du génome des actinobactéries variait de 1 à 12 Mb (Větrovský & Baldrian, 2013), soit deux fois la taille des génomes de *E. coli* (4,6 Mb) . Ils ont également la particularité d'être linéaires et de posséder un haut taux de G+C (aux alentours de 70%). La densité du gène est élevée avec approximativement 1 gène toutes les 1200 paires de base. Il s'agit en particulier du premier exemple d'une bactérie contenant davantage de gènes dans son génome que l'eucaryote *Saccharomyces cerevisiae*. (7825 chez *S. coelicolor* contre 6607 chez *S. cerevisiae*) (Haas, 2015), avec de puissants groupes de gènes biosynthétiques (BGC) qui codent pour des composés biologiquement actifs (Větrovský & Baldrian, 2013). Le terme « actinobactéries douées » a été inventé pour la première fois par Richard H. Baltz pour les actinobactéries avec une grande taille de génome, codant pour plus de production des métabolites secondaires que les autres taxons (Baltz, 2017) .

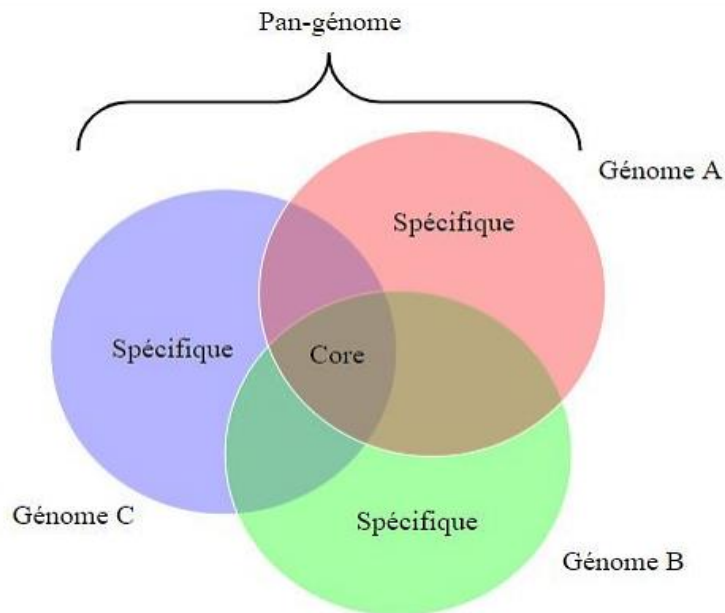
#### **III.3.2. Pourcentage GC et stabilité du génome**

En 1949, Chargaff et al., signalent que le contenu en bases puriques et pyrimidiques de l'ADN pouvait varier d'un individu à l'autre mais était constant pour les individus d'une même espèce. Chez les bactéries, le taux de G+C varie entre 25 et 75%. Pour les actinobactéries, ce coefficient est supérieur à 55%, il se situe en général entre 60 et 78% (Euzéby, 2003). Le pourcentage GC % qui est obligatoirement demandé lors d'une proposition de nouveaux genres. Ces critères ont permis de retracer l'ensemble de la phylogénèse des actinobactéries. (Ibrahimi, 2020).

L'ADN des *Streptomyces* est hautement enrichi en bases G/C, qui représentent entre 70 et 74% des nucléotides selon les espèces. Ce taux reste globalement le même dans les régions non codantes (Haas, 2015).

### III.3.3. Core-génomique et pan-génomique

Au sein d'un même groupe taxonomique, comme par exemple l'espèce, la variabilité génomique peut être très importante (Touchon et al., 2009). Le terme pan-génomique fut proposé pour représenter l'ensemble de la diversité génétique d'un ensemble d'individus (Tettelin et al., 2005), c'est-à-dire l'ensemble des gènes de tous les génomes analysés. Par opposition, les gènes partagés par tous les individus du groupe, définissent le core-génomique (Figure 15) (Lorenzi, 2020).



**Figur 15.** Diagramme de Venn illustrant un pan-génomique théorique. Chaque cercle représente le génome d'un organisme et les intersections les gènes partagés entre ces organismes. Le core-génomique (intersection des 3 cercles) correspond aux gènes partagés par tous les génomes, le génome accessoire à tous les gènes qui n'appartiennent pas au core-génomique (Lorenzi, 2020).

Les comparaisons de génomes complets ont permis de prendre conscience non seulement de la variabilité existante au sein de certaines espèces mais également de l'absence de variabilité au sein d'autres espèces. Deux situations extrêmes sont bien décrites dans la littérature (Choulet, 2006).

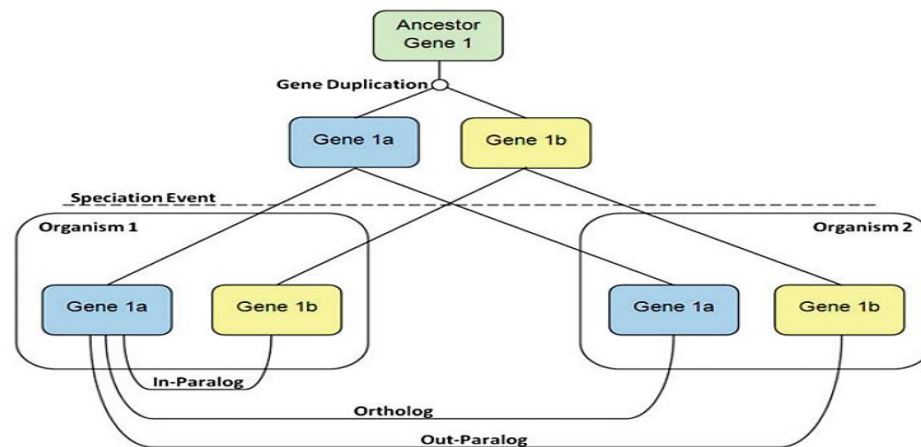
De nombreux travaux portent sur les répertoires de gènes et se distinguent principalement par le niveau phylogénétique étudié. Une analyse des répertoires de gènes à l'échelle du règne des bactéries (utilisation de 573 génomes bactériens) indique que seuls environ 250 gènes forment le core-génomique (Lapierre & Gogarten, 2009). Cette valeur est inférieure au génome minimal de 473 gènes, permettant de supposer que ces 250 gènes constituent le "squelette" de base d'un génome sur lequel se greffent d'autres

gènes essentiels, mais dépendant de l'environnement des organismes et/ou ayant trop divergés pour être identifiés comme orthologues à l'échelle du règne des bactéries (Hutchison et al., 2016).

À des niveaux phylogénétiques plus restreints (intra genre, espèce, population), la comparaison génomique a permis de mettre en évidence certains cas de l'absence de variabilité. Chez *E. coli* où deux études différentes utilisant un ensemble de 17 souches (Rasko et al., 2008) et 21 souches (Touchon et al., 2009) ont mis en évidence une grande diversité génétique chez cette espèce, où chaque nouvelle souche considérée augmente la diversité en gènes définissant un pan-génoome ouvert. Le core génome représente quant à lui environ la moitié du génome d'*E. coli*. La diversité entre les génomes ne peut se résumer à la composition en gènes, il faut également considérer l'organisation et la structure des génomes pour comprendre cette diversité (Lorenzi, 2020).

### III.3.4. Genès orthologues et Genès paralogues

La définition des gènes orthologues et paralogues est d'une grande importance lors de l'annotation de nouveaux génomes. «Homologie» fait référence à des gènes qui partagent une origine commune, «Horthologie» fait référence à des gènes qui surviennent par spéciation et «Paralogie» fait référence à des gènes qui surviennent par duplication. La figure 16 montre certains processus qui peuvent conduire à, et définir, des orthologues et des paralogues.



**Figure 16.** Un diagramme affichant les processus qui peuvent conduire à, et définir, des orthologues et des paralogues. Les événements de duplication et de spéciation des gènes créent des relations évolutives complexes entre les gènes. Le gène 1a dans l'organisme 1 a trois homologues après l'événement de spéciation. Le gène 1a dans l'organisme 1 et le gène 1a dans l'organisme 2 sont des orthologues car ils n'ont été séparés que par l'événement de spéciation. Le gène 1a dans l'organisme 1 et le gène 1b dans l'organisme 1 sont des paralogues, car ils n'ont été séparés que par l'événement de duplication de gène. Enfin, le gène 1a dans l'organisme 1 et le gène 1b dans l'organisme 2 sont des paralogues sortants, car ils ont été séparés par la duplication du gène et l'événement de spéciation (Richardson & Watson, 2013).



Ces processus ne sont pas seulement cruciaux pour définir les relations évolutives, mais aussi les relations fonctionnels, car les orthologues ont tendance à conserver des fonctions similaires, tandis que les paralogues ont tendance à diverger avec le temps pour remplir des fonctions différentes (Koonin, 2005). Il est essentiel que les orthologues soient définis avec précision. Il existe plusieurs approches de calcul qui peuvent être utilisées pour définir avec précision les orthologues (Kristensen et al., 2011). Les approches basées sur les arbres phylogénétiques tentent de reconstruire la relation évolutive entre les séquences de gènes et définissent ainsi les orthologues et les paralogues (Overbeek et al., 1999).

### **III.3.5. Détection des ORF**

Classiquement, les cadres de lecture ouverts ou ORF (Open Reading Frames) sont définis comme une étendue de triplets de nucléotides (codons) consécutifs et qui peuvent être traduits en protéines, où il doit également être initié avec un codon de départ dans le cadre (AUG) et se terminer avec l'un des trois codons d'arrêt (UAA, UAG, UGA) (Leong et al., 2022). De manière théorique, Olexiouk et al. (2018) ont estimé que la probabilité de générer aléatoirement un codon de départ dans l'espace nucléotidique est de 1 sur 64, et que les chances de trouver un codon d'arrêt dans les 99 prochains codons sont d'environ 99 %. Par conséquent, cela signifie que ~ 1,5 % du génome peut coder des ORF < 100 codons (T. A. Brown, 2002; Olexiouk et al., 2018).

La définition la plus courante d'un ORFs petits cadres de lecture ouverte (Small open reading frames) est simplement un ORF de moins de 100 acides aminés. Ils peuvent être situés dans des transcrits codants (5'UTR, CDS ou 3'UTR), les ARN circulaires et les ARN mitochondriaux (Orr et al., 2020). Alors que certains d'eux s'initient avec le codon d'initiation canonique (AUG) (Kearse & Wilusz, 2017).

Naturellement, cela se traduit par un nombre déraisonnablement élevé des ORFs putatifs, dont les chances d'être transcrits et traduits en polypeptides fonctionnels semblent exagérées. Par conséquent, un seuil de 300 nucléotides a été introduit, car la plupart de ces ORFs ont été jugés sans signification et aléatoires (Ladoukakis et al., 2011; Peeters & Menschaert, 2020). La région codante est généralement le cadre de lecture ouvert le plus long de l'ARNm, mais de nombreux autres plus courts sont également souvent présents dans le transcrit (Kute et al., 2022).

### **III.3.6. Importance des séquences codantes CDS**

Les régions codantes ou CDS (Coding DNA Sequence) sont les gènes impliqués dans la biosynthèse des antibiotiques, et autres métabolites appartenant au métabolisme secondaire, sont localisés dans le génome du microorganisme, et rarement dans les plasmides. Ils sont regroupés en opérons très longs, dont l'assemblage général est schématisé en Figure 17. Généralement, il est possible d'identifier des séquences codifiantes (gènes) et des séquences régulatrices (P et O) capables de promouvoir ou d'inhiber l'expression des gènes en réponse à des changements environnementaux spécifiques. Ces gènes codent pour des enzymes impliquées dans l'assemblage et l'édition du composé bioactif (Craney et al., 2013; Nodwell, 2017).



**Figure 17.** Vue schématique d'un opéron : promoteur (P) ; opérateur (O); gène opéron et terminateur (T) (De Simeis & Serra, 2021).

On distingue le « cœur », partie centrale contenant tous les gènes essentiels : métabolisme primaire, réplication de l'ADN, division cellulaire, ribosomes et ARNt. Les extrémités du gène, contiennent des gènes impliqués dans le métabolisme secondaire. Mais aussi certains gènes du métabolisme secondaire sont présents dans la partie centrale. Enfin, les TIR, régions inversées répétées (Terminal Inverted Repeats), sont comme leur nom l'indique des séquences identiques inversées situées aux extrémités du chromosome. Leur taille est très variable, pouvant aller de 174 pb chez *S. avermitilis* (Ikeda et al., 2003) (Haas, 2015).

### **III.3.7. Conservation de l'ordre des gènes biosynthétiques GOC**

Un indice, le GOC (Gènes Ordre conservation) a été défini comme la proportion du nombre total de paires d'orthologues entre deux génomes par rapport au nombre total d'orthologues entre ces deux génomes (Rocha, 2006).

Le GOC correspond à la fréquence relative des réarrangements du matériel génétique commun (transféré verticalement à partir de l'ancêtre commun) entre deux espèces en éliminant le transfert horizontal. Le GOC varie donc entre 1 (conservation parfaite de l'ordre des gènes) et 0 (aucune conservation). Une valeur élevée de GOC peut correspondre à deux situations, soit l'organisation des génomes comparés est stable soit ces deux génomes ont divergés récemment. Pour distinguer ces situations, le GOC a été modélisé en fonction de la distance phylogénétique séparant les espèces considérées pour les gènes isolés ou en opérons. Ainsi, la présence d'opérons impact sur l'évolution de la valeur de GOC avec la distance phylogénétique; l'organisation en opéron tend à biaiser à la hausse la valeur de GOC, l'organisation en

opéron étant plus conservée. De plus, il est apparu que le nombre de réarrangements endogènes, perturbant l'ordre des gènes ne serait pas aussi important qu'il est apparu à travers les premières études de génomique comparée. Il a également pu être montré qu'il n'existe pas de corrélation directe entre instabilité génomique et mode de vie de l'organisme. Cependant, certains organismes apparaissent plus instables que d'autres et en particulier les *Streptomyces* (Lorenzi, 2020).

### **III.4. Détection et prédiction des groupes de gènes biosynthétique**

#### **III.4.1. Méthodes de prédiction**

Les gènes essentiels, en tant que sous-ensemble minimal de gènes dans les organismes, sont nécessaires à la survie, au développement ou à la fertilité (Itaya, 1995 ; Kobayashi et al., 2003). Par conséquent, la prédiction et l'identification de ces gènes sont non seulement intéressantes, mais elles revêtent également une importance théorique et pratique. Une meilleure connaissance des gènes essentiels permet de mieux comprendre la structure primaire du réseau complexe de régulation des gènes dans une cellule (Yu et al., 2004 ; Papp et al., 2004 ; Yu et al., 2004 ; Gerdes et al., 2006). Elle aide à élucider la relation entre le génotype et le phénotype (D'Elia et al., 2009 ; Dowell et al., 2010), à identifier les maladies humaines (Dickerson et al., 2011), à découvrir des cibles médicamenteuses potentielles dans de nouveaux pathogènes (Chalker & Lunsford, 2002 ; Cole, 2002) et à réorganiser les micro-organismes (Gibson et al., 2010).

Deux types d'approches sont principalement utilisés pour prédire les gènes essentiels : les techniques expérimentales de laboratoire et les techniques informatiques. Les premières sont de façon aléatoire ou systématique pour inactiver les gènes essentiels potentiels, et l'essentialité des gènes peut être déterminée en fonction de la situation de vie de l'organisme. Les stratégies générales de perturbation génique comprennent l'inactivation d'un seul gène (Giaever et al., 2002), les knockouts conditionnels (Roemer et al., 2003) , l'interférence ARN (Cullen & Arndt, 2005), et la mutagenèse par transposon (Salama et al., 2004). Malheureusement, les techniques expérimentales présentent des durées longues et des coûts élevés. En outre, le spectre de l'essentialité des gènes varie en fonction des conditions de croissance (D'Elia et al., 2009 ; Fang et al., 2005).

Les techniques informatiques sont devenues populaires ces dernières années pour plusieurs raisons. Tout d'abord, les gènes essentiels connus de dizaines de micro-organismes fournissent du matériel d'instruction et de formation. Deuxièmement, les séquences génomiques et génomes disponibles obtenues par séquençage à haut débit offrent des possibilités sans précédent d'étudier le sous-ensemble minimal de

gènes dans divers organismes. Enfin, et surtout, le développement d'outils bio-informatiques améliore notre capacité à explorer les gènes essentiels (Cheng et al., 2013).

Plusieurs modèles de prédiction ont été développés *in silico* pour identifier les gènes essentiels. Parmi ces modèles, le plus simple est la prédiction des gènes essentiels basée sur l'essentialité connue des gènes homologues (Wang et al., 2004 ; Holman et al., 2009). Bien que ces modèles de prédiction présentent des niveaux de confiance élevés, ils présentent encore deux limites : premièrement, les orthologues conservés entre espèces ne représentent qu'une petite partie du génome (Brucoleri et al., 1998), d'autre part, les orthologues, en particulier dans des espèces très proches, présentent souvent des variations dans la régulations et fonctions des gènes (D'Elia et al., 2009; D.-U. Kim et al., 2010), ce qui conduit à une diversité potentielle dans l'essentialité des gènes. Pour contourner ces limitations, des modèles basés sur les caractéristiques ont été élaborés pour distinguer les gènes essentiels des gènes non essentiels sur la base de caractéristiques communes ou similaires parmi les gènes essentiels (Seringhaus et al., 2006; Deng et al., 2011).

### **III.4.2. Séquençage**

La génomique des procaryotes a connu une explosion du nombre de projets génomiques, grâce à l'avènement du séquençage de nouvelle génération (NGS), ce qui a permis de réduire considérablement le temps et l'argent investis dans les projets de génomique (MacLean et al., 2009).

Cependant, au début des années 2000, la seule technique automatisée de séquençage de l'ADN était la méthode des didésoxynucléotides mise au point par Sanger et ses collaborateurs (Sanger et al., 1977). Les progrès réalisés dans les années 2000, notamment en matière d'informatique et d'algorithmes d'assemblage (Myers et al., 2000), le séquençage Sanger était trop coûteux et exigeait trop de travail pour fournir une couverture suffisante pour une approche de routine du génome entier, nécessitant la création, le tri et le séquençage de bibliothèques génomiques avant l'assemblage complet du génome (Myers et al., 2000 ; Bentley et al., 2002).

Entre 2004 et 2006, les technologies de "séquençage de nouvelle génération (NGS)" ont été introduites, ce qui a transformé la recherche biomédicale et entraîné une augmentation spectaculaire de la production de données de séquençage. L'augmentation significative de la production de données est due aux principes et aux innovations de la nanotechnologie qui ont permis le séquençage massivement parallèle de molécules d'ADN uniques. Les caractéristiques combinées du séquençage à haut débit et du séquençage d'une seule molécule d'ADN sont les caractéristiques de la NGS, quelle que soit la plateforme de séquençage. Les

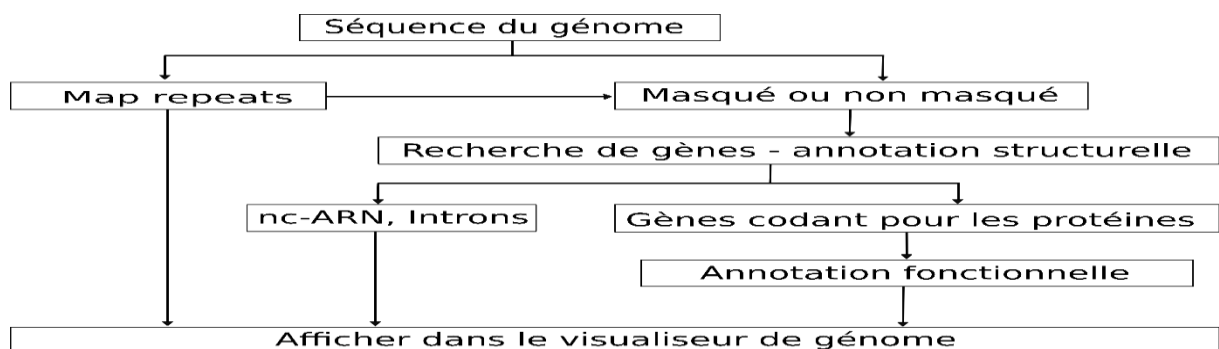
procédures évoluées de la technologie ont été mieux fusionnées avec l'acquisition et l'analyse des données, libérant la communauté des approches historiques de séquençage Sanger, plus intensives en main-d'œuvre et peu efficaces, et facilitant une augmentation extraordinaire de la production de données (Mardis, 2013).

Les technologies NGS rendent abordable le séquençage à haut débit des génomes bactériens, ce qui, associé aux progrès constants des algorithmes informatiques et de bases de données pour le balayage (Gomez-Escribano et al., 2016).

### III.4.3. Annotation automatique

L'annotation est un moyen d'extraire des informations encodées dans la multitude de séquences différentes des quatre nucléotides (A, T, C et G). Le terme "annotation du génome" a évolué, passant de l'annotation des gènes codant pour des protéines à l'annotation de nucléotides uniques sur des milliers de génomes individuels. Une annotation réussie dépend de la qualité de l'assemblage du génome. Plusieurs méthodes statistiques sont utilisées pour décrire l'exhaustivité et la contiguïté d'un assemblage (Yandell & Ence, 2012). Les améliorations apportées au séquençage, notamment les technologies long-read (English et al., 2012) et linked-read (Weisenfeld et al., 2017), ont permis de mettre à disposition des assemblages de génomes de haute qualité à des prix plus bas. La disponibilité d'assemblages de génomes de haute qualité a constitué une source solide d'informations phylogénétiques, ce qui a permis d'améliorer les alignements et les annotations de génomes entiers qui reposent en grande partie sur des modèles de souris et d'êtres humains (J et al., 2019).

La nécessité d'une annotation automatisée est devenue urgente avec le développement des technologies de séquençage de nouvelle génération, mais l'annotation des génomes reste un défi et repose toujours sur un processus comprenant de nombreuses étapes manuelles (Figure 18) (Bryson et al., 2006; Génolevures Consortium et al., 2009).



**Figure 18.** Annotation du génome - flux de travail

### **III.4.3.1. Annotation structurelle**

La recherche des caractéristiques de l'ADN (exons, introns, promoteurs, transposons, etc.) est connue sous le nom d'annotation structurelle. Alors que l'annotation structurelle tente de trouver des gènes dans une séquence génomique, la définition des gènes a évolué avec les progrès de la génomique moderne. Un gène peut être défini comme "une région de séquence nécessaire pour générer des produits fonctionnels" (Spieth & Lawson, 2006). Les produits fonctionnels des gènes sont les protéines et les ARN. Les gènes qui conduisent à la production de protéines sont appelés gènes codant pour des protéines. D'autres gènes qui ne codent pas de protéines, mais des molécules d'ARN fonctionnelles, sont appelés gènes non codants. Les gènes ARN comprennent les gènes de l'ARN ribosomique (ARNr), de l'ARN de transfert (ARNt), du microARN (ARNm), du petit ARN nucléaire et de l'ARN nucléolaire (snRNA et snoRNA, respectivement) (B. Zhang et al., 2016) et long ARN non codant (lncRNA). Les annotations structurelles identifient également les pseudogènes. Ils étaient initialement considérés comme sans fonction et comme des culs-de-sac évolutifs (evolutionary dead-ends). Nous savons aujourd'hui qu'ils participent parfois à la régulation des gènes. Leur prédiction améliore donc la compréhension des génomes (Xiao et al., 2016).

La première étape de l'annotation structurelle consiste à masquer les répétitions. Les répétitions d'ADN sont présentes à la fois dans les organismes procaryotes et eucaryotes. Les répétitions représentent de 0 % à plus de 42 % du génome procaryote (Ejigu & Jung, 2020).

L'alignement des transcriptions et des protéines après masquage est la deuxième étape de l'annotation structurelle avant l'identification des gènes, bien qu'elle ne soit pas obligatoire. BLAST (Altschul et al., 1990) ou BLAT (Kent, 2002) peuvent être utilisés pour aligner les transcriptions et les protéines. En outre, les données RNA-seq peuvent être alignées à l'aide de TopHat (D. Kim et al., 2013) ou de HISAT (D. Kim et al., 2015).

L'identification des gènes codant pour des protéines et d'autres éléments de régulation occupe une place centrale dans l'annotation des gènes. La prédiction des gènes est un processus complexe, en particulier pour l'ADN eucaryote. La taille variable des introns (séquences non codantes) entre les exons et les variantes d'épissage rendent difficile la prédiction de la structure des gènes (Yandell & Ence, 2012).

### **III.4.3.2. Annotation fonctionnelle**

Bien que l'annotation manuelle soit toujours considérée comme l'étalon-or, cette approche est difficilement extensible. Il est donc nécessaire d'utiliser des méthodes d'annotation automatisées, afin de

s'adapter à la pléthore de données génomiques actuellement générées par la technologie NGS. La prédiction automatique de la fonction peut être directement, en utilisant des outils d'alignement local, tels que BLAST, qui recherchent dans une base de données de protéines les alignements les plus performants. La fonction est alors attribuée à la séquence inconnue de la requête sur la base d'une séquence de résultat connue, à condition qu'il s'agisse de l'alignement le mieux noté de toutes les séquences, au-dessus d'une certaine valeur seuil spécifiée (Ejigu & Jung, 2020). L'hypothèse de transfert de fonction sur laquelle s'appuient les outils de type BLAST est que la fonction est conservée dans les protéines qui ont des séquences similaires et qui ont évolué à partir d'un seul ancêtre (Ejigu & Jung, 2020). En d'autres termes, les outils identifient les relations évolutives en découvrant les relations orthologues et paralogues entre les séquences. Les orthologues sont des gènes qui proviennent d'un seul gène ancestral dans le dernier ancêtre commun des génomes comparés, tandis que les paralogues sont des gènes au sein du même génome. sont des gènes du même génome issus de duplications (Koonin, 2005). Les annotations fonctionnelles sont simples à utiliser et donnent de bons résultats dans de nombreux cas. Cependant, elles présentent certains inconvénients. Les exemples incluent l'erreur de source de la base de données, la relativité du seuil d'alignement, faible sensibilité/spécificité, et le transfert excessif d'annotations à partir d'une région locale de de similarité sans rapport (Sasson et al., 2006).

# **PARTIE PRATIQUE**



---

---

# MATERIELS ET METHODES

---

---

La présente étude *in silico* s'inscrit dans le cadre de la bio- prospection et l'exploration du génome bactérien. Sur la base du développement rapide des technologies de séquençage et des progrès de l'exploration du génome « genome mining », il est devenu impératif d'explorer les espèces nouvellement isolées pour trouver d'éventuels nouveaux métabolites secondaires afin de découvrir de nouveaux produits naturels bioactifs.

En s'engageant dans cette voie pour contribuer *via* cette étude *in silico* à l'avancement scientifique et technologique dans la détection des gènes biosynthétiques. Ainsi, notre étude s'est déroulée dans le cadre d'une analyse bio-informatique, en explorant toutes les étapes allant de l'analyse du génome bactérien jusqu'à l'identification des protéines, tout en déduisant leurs activités pharmacologiques. Nous avons focalisé notre attention sur les gènes de type PKS et NRPS, ainsi que sur les métabolites issus des souches bactériennes de la famille des Actinobactéries plus précisément le genre *Streptomyces*.

Il est convié de souligner que les membres du genre *Streptomyces* sont réputés pour leur capacité exceptionnelle à survivre dans des milieux extrêmes, ce qui se traduit par des caractéristiques génétiques uniques par rapport aux autres espèces présentes dans des environnements naturels plus communs. De plus, nous avons également pris en considération la disponibilité d'informations complètes sur le séquençage entier de leurs génomes, publiées dans la prestigieuse base de données NCBI.

Cette démarche nous a permis d'élargir le champ de notre recherche et d'obtenir des données plus approfondies fournissant, ainsi une base solide pour nos analyses bio-informatiques. En combinant les informations génétiques spécifiques à ces souches bactériennes avec nos connaissances en pharmacologie, nous souhaitons faire progresser la compréhension de l'activité des clusters gènes biosynthétiques et de leurs implications pharmacologiques, ouvrant ainsi la voie à de la prédiction de nouvelles substances ayant applications potentielles dans le domaine pharmaceutique et médical.

## **I. Choix des souches actinomycétales pour l'étude *in silico***

Les environnements extrêmes sont caractérisés par une salinité élevée, un pH élevé ou faible, des conditions arides, une faible teneur en nutriments et en oxygène, des températures élevées ou basses et une forte exposition aux rayons UV, et qui seraient préjudiciables aux conditions « normales » nécessaires à la survie (Zenova et al., 2011). Ces dernières années, les chercheurs se sont davantage concentrés sur l'environnement extrême en tant que source puissante de nouvelles espèces de streptomycètes à fort activité biologique (Sivalingam et al., 2019).

Ce groupe de bactéries a la capacité de survivre dans de multiples conditions de ce type car elles possèdent des caractéristiques adaptatives distinctives telles que la production d'enzymes spécifiques, la commutation entre différents modes métaboliques et l'antibiose (L & T, 2015).

Cette première étude réalisée en Algérie ouvre de nouvelles perspectives passionnantes dans le domaine de la biotechnologie et de la pharmacologie. En exploitant le potentiel des actinobactéries pour la production de métabolites secondaires, l'Algérie peut non seulement renforcer ses capacités de recherche, mais aussi contribuer de manière significative au développement de solutions innovantes dans les domaines médicaux,

Pour assurer la pertinence et la solidité de notre étude, nous avons mené une recherche bibliographique exhaustive. Cette étape cruciale nous a permis de sélectionner des souches spécifiques de *Streptomyces* à étudier. Au total, nous avons examiné environ 37 souches appartenant à différentes espèces du genre *Streptomyces*, toutes isolées à partir de divers écosystèmes extrêmes en Algérie. Ces souches ont déjà fait l'objet d'études et de publications par des chercheurs algériens renommés dans le domaine de la génomique bactérienne, et nous les avons identifiées grâce à des recherches dans des bases de données tels que PubMed, ScienceDirect, ResearchGate et Frontiers.

Le choix des souches à étudier a été déterminé avec soin, en tenant compte de leur intérêt scientifique et de leur pertinence par rapport à nos objectifs de recherche. Finalement, trois souches ont été retenues pour leur caractère particulièrement intéressant :

- *Streptomyces ambofaciens* ATCC 23877(Medjemadj et al., 2020).
- *Streptomyces leeuwenhoekii* souche type souche (C34 = DSM 42122 = NRRL B-24963)(Medjemadj et al., 2020).
- *Streptomyces sp.* KY75(Tighidet et al., 2022).

Ces souches ont suscité notre intérêt en raison de leurs caractéristiques uniques, de leur potentiel biochimique ou de leurs propriétés adaptatives dans des environnements extrêmes. Leur étude pourrait fournir des informations essentielles pour mieux comprendre le genre *Streptomyces* dans des contextes environnementaux spécifiques et explorer leurs applications potentielles en biotechnologie, pharmacologie ou agriculture.

Il est important de souligner que l'utilisation de souches provenant d'écosystèmes extrêmes en Algérie ajoute une dimension particulière à notre recherche, car ces environnements présentent souvent des conditions physico-chimiques inhabituelles et des défis uniques pour la survie et l'adaptation des bactéries. En nous appuyant sur les travaux préalables de chercheurs algériens compétents, nous bénéficions également d'une expertise locale pour mieux contextualiser nos résultats.

Le 05 avril 2023, les séquences génomiques complètes (complete WGS) des trois souches de *Streptomyces* ont été récupérées à partir de la base de données NCBI, avec les numéros d'accèsion suivants :

- *Streptomyces ambofaciens* ATCC 23877 avec un numéro d'accèsion sur NCBI : (CP012382). Elle a été isolée à partir de la source thermale de Béniharoun, wilaya de Mila.
- *Streptomyces leeuwenhoekii* strain C34 avec un numéro d'accèsion sur NCBI : (LN831790). Elle a été isolée à partir de la source thermale d'Essalihine, wilaya de Khenchela.

Les sources chaudes sont pauvres en matière organique, mais riches en cations (tels que le sodium et le calcium) et en anions (tels que les chlorures, les bicarbonates et les sulfates). Cette composition chimique, combinée à la température élevée de ces eaux, confère aux micro-organismes un écosystème des plus extrêmes et des plus uniques. Par conséquent, les bactéries oligotrophes, et en particulier les actinobactéries, qui existent dans ces endroits, peuvent faire preuve d'une capacité d'adaptation exceptionnelle et présenter des caractéristiques métaboliques intéressantes.

- *Streptomyces sp.* KY75 avec un numéro d'accèsion sur NCBI : (NZ\_LR881953). C'est une nouvelle souche actinomycétale qui a été isolée à partir des sols argilo-sableux de la mine fermée de " Kherzet Youcef " à Ain Azel, wilaya de Sétif. C'est une ancienne mine de plomb et de zinc située dans le nord-est de l'Algérie.

Pour survivre dans des écosystèmes pollués, cette souche a une capacité de résister aux métaux lourds. Le zinc est un métal présent dans la croûte terrestre, mais pas à l'état libre. Il se trouve sous

forme d'oxyde, de silicates et de carbonates. On le trouve souvent associé au plomb et au cadmium dans les minéraux, et son extraction se fait par des procédés hydro-métallurgiques ou pyrométallurgiques (Belabed, 2019).

La localisation des sites d'étude (origine des souches) est représentée sur la carte géographique ci-dessous (Figure 19).



**Figure 19.** (A) Répartition géographique des sources thermales algériennes étudiées dans ce travail. En couleur, les provinces où sont situés les sites d'échantillonnage : Setif (bleu), Khenchela (rose) et Mila (vert). (B) Source thermale de Béniharoun (Mila, 36° 33' 193 N, 6° 16' 213 E). (C) Source chaude d'Essalihine (Khenchela, 35° 26' 253 N, 7° 05' 2043). (D) sols argilo-sableux de " Kherzet Youcef " à Ain Azel (setif, 35°50'36"N, 5°31'19"E).

## **II. Analyses phylogénétiques et phylogénomiques**

### **II.1. Analyse phylogénétique**

La « phylogénie » est l'étude systématique de la reconstruction de l'histoire évolutive passée d'espèces ou de taxons existants, sur la base de données actuelles, telles que des caractéristiques morphologiques ou des informations moléculaires (données de séquence). Cette histoire évolutive ou phylogénie est idéalement représentée sous la forme d'un arbre binaire (Jarvis et al., 2017).

Les méthodes phylogénétiques jouent un rôle déterminant dans l'identification et la classification des bactéries. Ici, nous examinons l'utilité et l'efficacité des phylogénies traditionnelles basées sur l'alignement pour reconstruire les relations évolutives entre les espèces bactériennes. Nous considérons les approches du génome entier (ainsi que sans alignement) développées plus récemment. Enfin, nous fournissons des exemples où des techniques phylogénétiques ont été utilisées en conjonction avec des informations protéomiques et biochimiques pour identifier des sous-groupes fonctionnels au sein de grandes superfamilles de gènes (Armstrong et al., 2019).

### **II.1.1. Analyse phylogénétique basée sur le gène de l'ARNr 16S**

Les séquences du gène de l'ARNr 16S des souches étudiées ont été alignées pour la recherche d'homologie en comparant la séquence avec celles présentes dans la base de données publique (NCBI) à l'aide de l'outil standard Basic Local Alignment Search Tool (BLAST). L'analyse phylogénétique a été réalisée à l'aide du logiciel Molecular Evolutionary Genetics Analysis (MEGA) version 11 (Tamura et al., 2021, p. 11) et les alignements multiples des données ont été effectués par MUSCLE (Edgar, 2004). L'arbre phylogénétique a été reconstruit à l'aide de l'algorithme de jonction des voisins. La topologie de l'arbre résultant a été évaluée par une analyse bootstrap de l'ensemble des données de l'association de voisins, sur la base de 1000 répliques (Felsenstein, 1985).

### **II.1.2. Analyse phylogénétique basée sur le génome entier**

Le personnel du National Center for Biotechnology Information (NCBI) sélectionne et classe manuellement les génomes de référence et les génomes représentatifs qu'il considère comme étant de grande qualité et importants pour la communauté des chercheurs. PATRIC fournit les génomes de référence et les génomes représentatifs et les inclut dans l'analyse phylogénétique qui fait partie du rapport Comprehensive Genome Analysis. Les génomes de référence et représentatifs les plus proches ont été identifiés par Mash/MinHash (Ondov et al., 2016). Les familles de protéines globales PATRIC (PGFams) (Davis et al., 2016) ont été sélectionnées à partir de ces génomes pour déterminer la position phylogénétique de ce génome. Les séquences protéiques de ces familles ont été alignées avec MUSCLE (Davis et al., 2016), et les nucléotides de chacune de ces séquences ont été mis en correspondance avec l'alignement des protéines. L'ensemble des alignements d'acides aminés et de nucléotides a été concaténé dans une matrice de données, et RaxML (Stamatakis, 2014) a été utilisé pour analyser cette matrice, avec un bootstrapping rapide (Stamatakis et al., 2008) pour générer les valeurs de soutien dans l'arbre.

## **II.2. Comparaisons des génomes entiers**

Les comparaisons des génomes ont été effectuées à l'aide du BLAST Ring Image Generator (BRIG). Il existe un besoin évident d'un programme convivial capable de produire des comparaisons de génomes pour un grand nombre de génomes procaryotes en mettant l'accent sur l'utilisation rapide de données génomiques inachevées ou non assemblées. BRIG est une application multiplateforme qui permet la génération interactive d'images génomiques comparatives *via* une simple interface utilisateur graphique (Alikhan et al., 2011).

Il permet de générer des images montrant de multiples comparaisons de génomes de procaryotes nécessite l'installation de Java 1.6 ou plus et NCBI BLAST+ ou BLAST legacy. Tout d'abord, les séquences du génome entier ont été obtenus auprès du NCBI, puis les séquences ont été téléchargées sous forme FASTA. Ensuite, les séquences du génome entier ont été sélectionné. Après ces étapes les anneaux ont été dénommés et personnalisés par des couleurs aussi, le pourcentage d'identification est resté par défaut.

Les lectures des 3 souches ont été soumises au service complet d'analyse du génome dans le Pathosystems Resource Integration Center (PATRIC), ce site est le centre de ressources en bio-informatique bactérienne (<https://www.patricbrc.org>), il fournit aux chercheurs une ressource en ligne qui stocke et intègre une variété de types de données, par exemple la génomique, la transcriptomique, les interactions protéine-protéine, les structures protéiques 3D et les données de typage de séquences, et métadonnées associées. Ce site Web est principalement centré sur l'organisme, avec différents niveaux de données génomiques et des informations associées liées à chaque organisme inclus. L'interface repensée permet aux chercheurs d'accéder directement aux outils et aux données, et l'accent a été mis sur les groupes de génomes créés par l'utilisateur, avec des résumés détaillés et des vues des données que les chercheurs ont sélectionnées. Le changement le plus important a peut-être été la capacité accrue des chercheurs à analyser leurs données privées et à les comparer aux données publiques disponibles (Wattam et al., 2017).

## **II.3. Détermination de l'identité nucléotidique moyenne (ANI)**

La valeur ANI représente la proportion des nucléotides identiques dans les zones alignées. Ces chiffres fournissent une évaluation quantitative de la similarité entre deux génomes ou bien plus. Nous avons effectué une analyse du génome entier sur les trois génomes de référence en utilisant le serveur JSpeciesWS avec l'algorithme BLAST (<http://jspecies.ribohost.com/jspeciesws/>) (Richter et al., 2016), afin de déterminer la variabilité génétique entre les génomes ou d'établir des lignées non

ambiguës. La comparaison entre ces génomes a été évaluée au moyen de mesures de l'ANI sur la base des résultats de l'alignement du génome entier.

#### **II.4. Analyse de similitude des protéines prédites**

Pour faciliter cette recherche, OrthoVenn3 a été développé comme un puissant outil basé sur le Web qui permet aux utilisateurs d'identifier et d'annoter efficacement les clusters orthologues et de déduire les relations phylogénétiques à travers une gamme d'espèces et illustré les similitudes des protéines. OrthoVenn3 comprend plusieurs nouvelles fonctionnalités importantes, notamment une meilleure précision d'identification des grappes orthologues, des capacités de visualisation améliorées pour les grands ensembles de données et l'ajout d'outils d'analyse phylogénétique (<https://orthovenn3.bioinfotoolkits.net>).

OrthoVenn3 a été utilisé pour comparer les protéines des trois souches étudiées et les protéines de leurs plus proches parents, nous avons sélectionné trois espèces génétiquement proches à chacune des trois *Streptomyces* étudiées à savoir :

- *Streptomyces ambofaciens* strain DSM 40697
- *Streptomyces cyaneogriseus subsp. noncyanogenus* strain NMWT
- *Streptomyces sp.*KY70

Nous avons utilisé les séquences protéiques du génome complet des six *Streptomyces*, disponibles sur NCBI au format FASTA. Enfin, nous avons utilisé la fonction de similarité des protéines sur OrthoVenn3 pour générer les résultats.

### **III. Annotation génomique**

L'annotation génomique comprend à la fois l'annotation structurelle pour délimiter les éléments et les caractéristiques générale du génome, et l'annotation fonctionnelle pour associer des informations fonctionnelles à ces éléments. Elle permet une meilleure compréhension des données génomiques et facilite les études sur les processus biologiques et l'évolution des espèces (Bright et al., 2009).

#### **III.1. Annotation structurelle et fonctionnelle du génome**

L'annotation structurelle ou syntaxique d'une séquence d'ADN est essentielle pour valoriser ses caractéristiques telles que les promoteurs, la longueur de ADN, les ARN, le pourcentage G/C, les CDS, le nombre de protéines, etc. Celle-ci est généralement accompagnée d'une annotation fonctionnelle visant à déterminer les fonctions de ces différentes régions. Ce dernier type d'annotation

a pour but de donner une vision intégrée des relations entre les produits des gènes. Différentes analyses bio-informatiques peuvent être utilisées pour prédire la structure du gène, ce qui contribue à une meilleure compréhension et utilisation des séquences d'ADN (Abril & Castellano, 2019).

Les génomes complets des souches bactériennes étudiées ont été annotés à l'aide du logiciel RAST (Rapid Annotation using Subsystem Technology) (<https://rast.nmpdr.org/>) qui est un service entièrement automatisé pour annoter des génomes bactériens et archéens complets ou presque complets. Il fournit des résultats de haute qualité pour ces génomes dans l'ensemble de l'arbre phylogénétique. RAST peut fournir un appel de gènes automatisé de haute qualité et une annotation fonctionnelle, avec la capacité de prendre en charge à la fois les séquences génomiques complètes et les brouillons de génomes. Les génomes soumis sont annotés et disponibles dans les 12 à 24 heures suivantes à la soumission du service (Aziz et al., 2008).

### **III.2. Analyses des gènes des métabolites secondaires bioactifs et des composés**

#### **III.2.1. Détection des régions de métabolites secondaires bioactifs**

L'outil de serveur web antiSMASH (antibiotics and Secondary Metabolite Analysis Shell) (version bactérienne) (<https://antismash.secondarymetabolites.org/>) utilise une approche basée sur des règles pour identifier de nombreux types différents de voies de biosynthèse impliquées dans la production des métabolites secondaires. Des analyses plus approfondies sont effectuées pour la détection des BGC codant pour les synthétases de peptides non ribosomiques (NRPS), les synthases de polykétides de type I et de type II (PKS) et les classes de peptides synthétisés par le ribosome et modifiés après la traduction (RiPP) et les thiopeptides...ect. Pour ceux-ci, des analyses spécifiques au cluster peuvent fournir plus d'informations sur les étapes de biosynthèse réalisées et ainsi fournir également des prédictions plus détaillées sur le ou les composés produits (Blin et al., 2023).

Tout d'abord, les séquences du génome entier et les numéros d'accèsion des souches bactériennes ont été obtenus auprès du NCBI, puis les séquences ont été téléchargées comme données d'entrée sous forme gb. Known Cluster Blast, Sub Cluster Blast et Active Site Finder ont été marqués comme options vedettes. Les régions de métabolites secondaires et leurs types ont été identifiés à partir de la sortie graphique.

#### **III.2.2. Profilage des métabolites secondaires**

Les diagrammes de corde circulaires sont très utiles pour visualiser des données complexes. Ils représentent admirablement les données comportant un grand nombre de catégories, affichent des



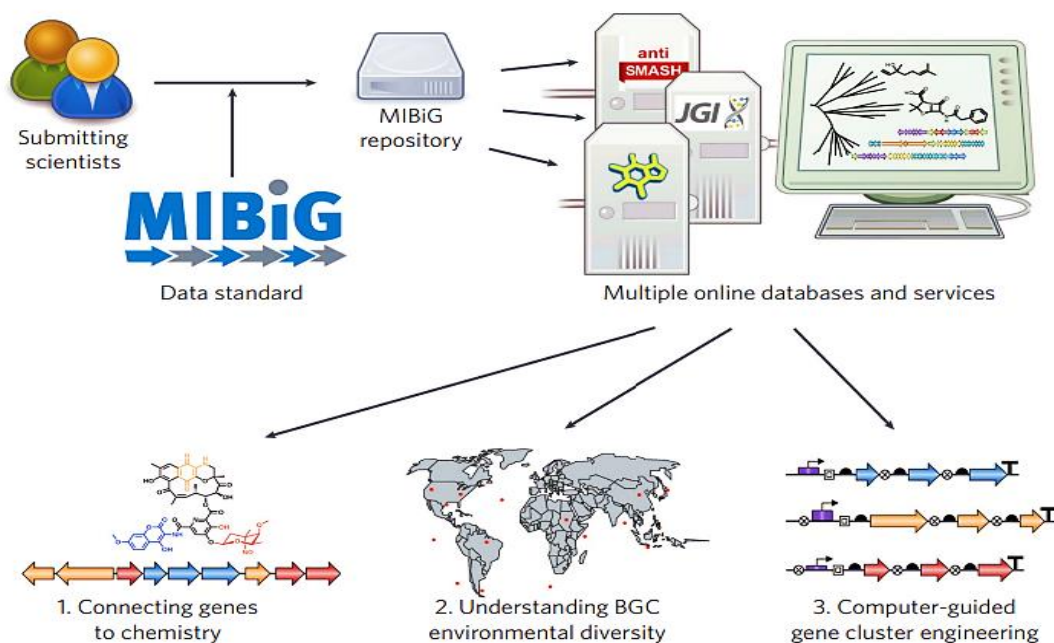
données avec plusieurs pistes se concentrant sur le même objet et montrent facilement les relations entre les éléments.

Nous avons utilisé le diagramme de corde circulaire pour représenter la diversité des groupes de gènes biosynthétiques (BGC), en se basant sur les résultats de la détection et l'identification des BGC obtenus par le serveur antiSMASH pour chaque génome des trois souches étudiées. Après une curation manuelle de ces résultats en générant un tableau représentant les résultats d'antiSMASH des trois souches. Ensuite, nous avons converti ce tableau en matrice de données dans l'environnement RStudio (Ariel de Lima et al., 2022), en utilisant le paquet Circlize afin de tracer un diagramme de corde circulaire.

### **III.2.3. Prédiction des composés secondaires bioactifs**

Pour faciliter la détection et l'analyse comparative des BGC prédits avec les BGC connus, et pour caractériser l'interaction entre la séquence, la structure et la fonction, à cette fin, les chercheuses ont développé la norme Minimum Information about a Biosynthetic Gene cluster (MIBiG) et construit une base de données qui contient des entrées standardisées pour des BGC validés expérimentalement et dont la fonction est connue (Figure 20) (Kautsar et al., 2020; M. Medema et al., 2015).

Chaque entrée contient au minimum des informations sur l'entrée nucléotidique et les coordonnées du locus génomique concerné, la taxonomie de l'organisme producteur, la classe biosynthétique, le nom du (des) composé(s) produit(s) et la (les) référence(s) bibliographique(s) (Terlouw et al., 2023).



**Figure 20.** La norme de données MIBiG et le système de soumission conduiront à un ensemble de données en croissance continue (stocké dans le dépôt en ligne MIBiG) qui sera chargé dans plusieurs bases de données et services web. La partie inférieure de la figure montre le triple potentiel de MIBiG pour l'étude des BGC, qui permettra (1) de relier systématiquement les gènes et la chimie en identifiant quels gènes sont responsables de la biosynthèse de quelles entités chimiques ; (2) de comprendre la diversité génétique naturelle des BGC dans leur contexte environnemental et écologique, en combinant les ensembles de métadonnées dérivées de MIBiG et de MiXS ; et (3) de développer un registre de pièces basé sur des preuves pour l'ingénierie des voies de biosynthèse et des groupes de gènes par le biais de la biologie synthétique (M. Medema et al., 2015).

Les bases de données antiSMASH et Minimal Information about a Biosynthetic Gene cluster-MIBiG (<https://mibig.secondarymetabolites.org/>) ont été utilisées pour identifier les régions de métabolites secondaires proches. En fonction du groupe, différents composés ont été trouvés par Knowledge Blast.

---

---

# RESULTATS ET DISCUSSIONS

---

---

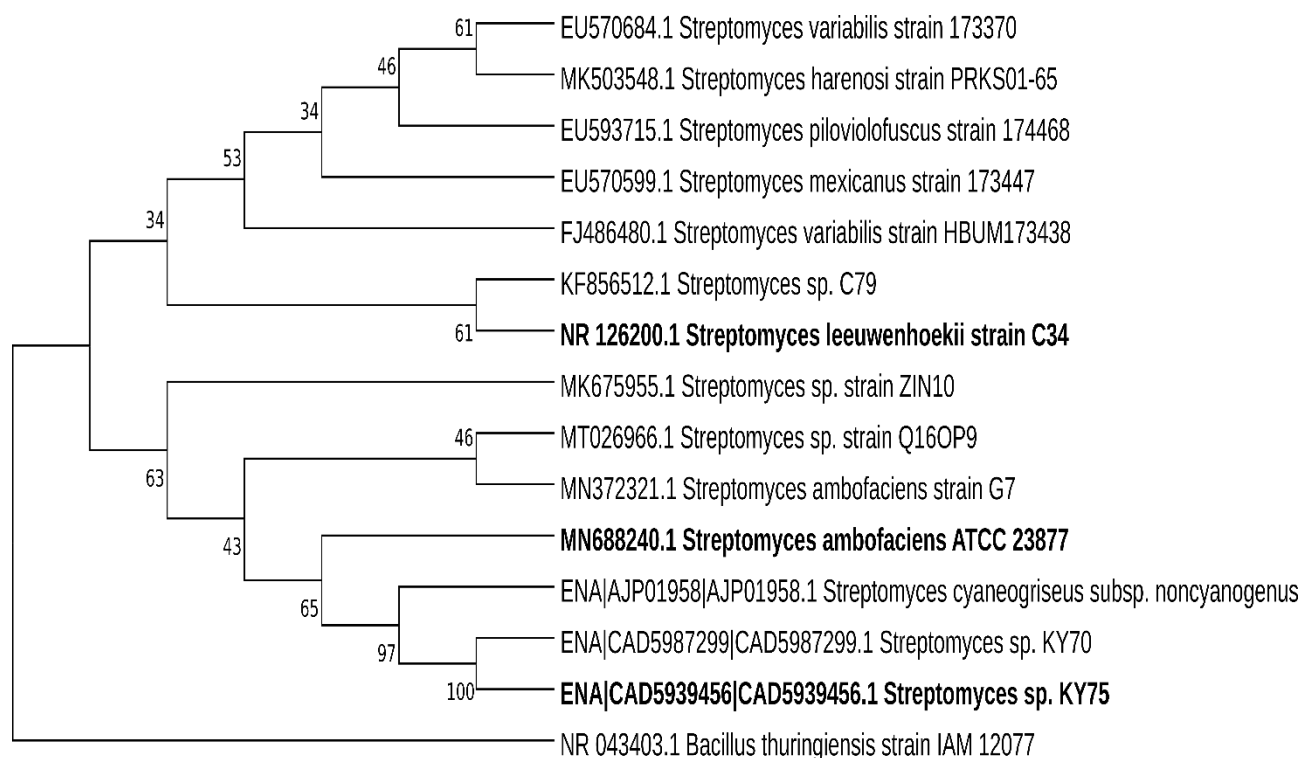
## I. Analyses phylogénétiques et phylogénomiques

### I.1. Analyse phylogénétique

Les gènes orthologues proviennent d'un ancêtre commun lors d'événements de spéciation et sont généralement synténiques entre les espèces proches. Les paralogues partagent également un ancêtre commun, mais résultent d'événements de duplication de séquence au sein d'une espèce, et présentent souvent une synténie limitée et une divergence plus liée à la spéciation. Si des gènes orthologues de plusieurs espèces présentent une grande similarité de séquence, il est probable qu'ils continuent à remplir des fonctions biologiques similaires (Xu et al., 2019). D'autre part, les orthologues avec des séquences qui montrent une plus grande divergence par rapport aux autres espèces sont plus susceptibles d'accomplir des fonctions distinctes (Fang et al., 2010).

#### I.1.1. Analyse phylogénétique basée sur le gène de l'ARNr 16S

L'analyse phylogénétique basée sur les séquences de l'ARNr 16S a été réalisée pour les souches de *Streptomyces ambofaciens* ATCC 23877, *Streptomyces eleuwenhoekii* et *Streptomyces sp.* KY75. Cette analyse a été effectuée à l'aide du logiciel MEGA 11 pour déterminer leur positionnement phylogénétique ce résultat généré a permis de visualiser les relations évolutives entre ces souches et d'autres espèces de *Streptomyces* étudiées (Figure 21). Cette approche a révélé des informations sur la diversité génétique et l'évolution de ces bactéries du sol.



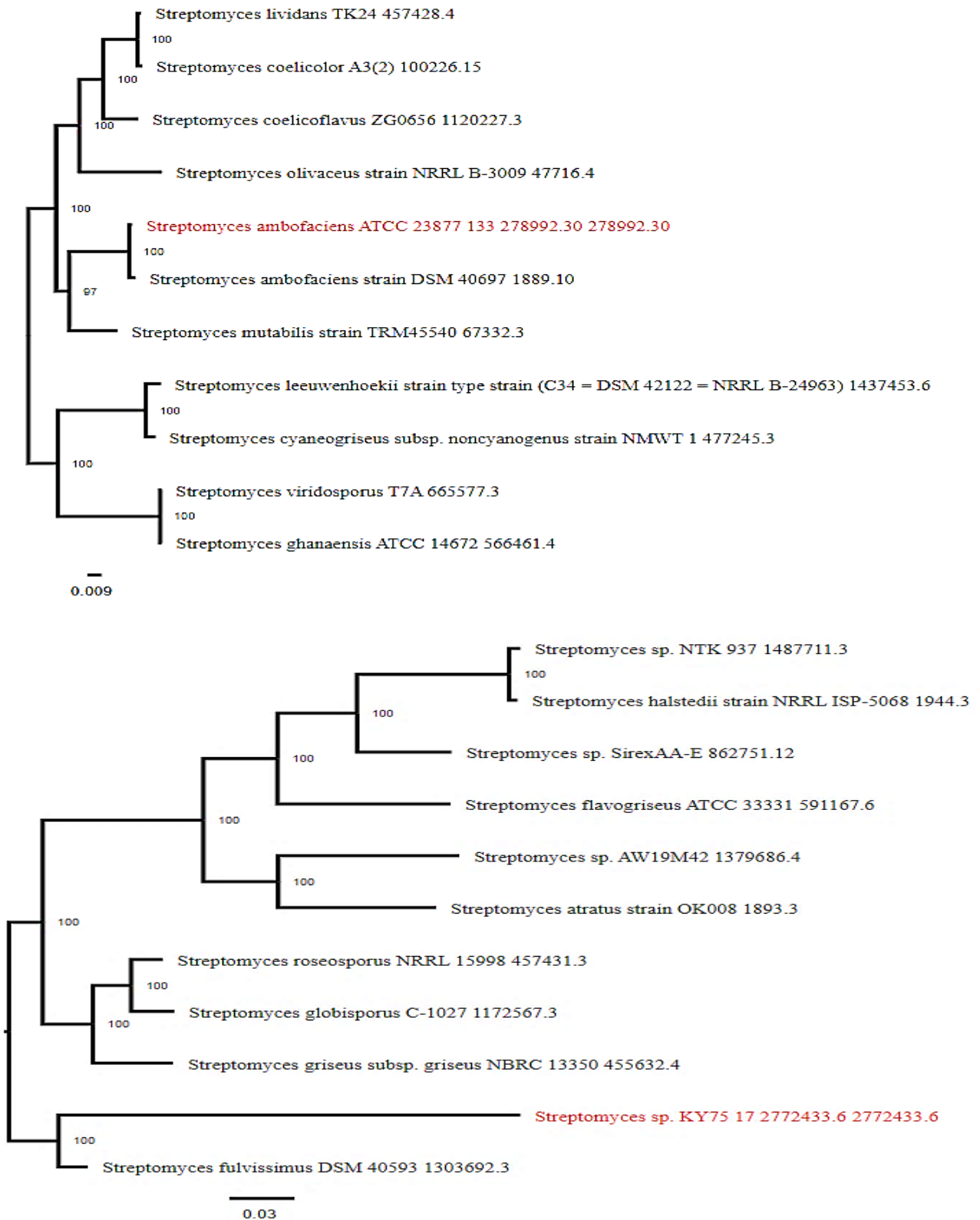
**Figure 21.** Analyse phylogénétique moléculaire en utilisant les séquences de l'ARNr 16S, de *Streptomyces ambofaciens* ATCC 23877, *Streptomyces leeuwenhoekii* strain C34 et *Streptomyces* sp. KY75. L'arbre a été construit par MEGA11 en utilisant la méthode du maximum de vraisemblance basée sur le modèle général réversible dans le temps.

Cette analyse a contribué à notre compréhension de la classification et de la phylogénie des *Streptomyces* étudiés. Les résultats de l'analyse phylogénétique suggèrent que les espèces étudiées sont des orthologues, ce qui signifie qu'elles ont évolué à partir d'un ancêtre commun et partagent une origine génétique similaire. Malgré les mutations et les divergences au fil du temps, les orthologues ont tendance à maintenir des fonctions fondamentales, ce qui implique que les gènes d'une espèce peuvent avoir des fonctions équivalentes dans une autre espèce. Les différences entre ces génomes sont souvent influencées par des adaptations spécifiques à l'écosystème dans lequel les bactéries évoluent, ainsi que par les interactions qu'elles entretiennent avec d'autres microorganismes présents dans leur environnement.

### I.1.2. Analyse phylogénétique basée sur le génome entier

Sur la base de l'analyse phylogénétique des séquences du génome entier, pour les souches de *Streptomyces ambofaciens* ATCC 23877, *Streptomyces leeuwenhoekii* et *Streptomyces* sp. KY75, les arbres phylogénétiques représentés dans la figure 22 ont été construits, en utilisant le serveur PARTIC, pour déterminer les positions phylogénétiques de ces génomes.

Les résultats des deux arbres phylogénétiques mettent en évidence que les trois souches étudiées ont le même ancêtre malgré les différences génétiques.



**Figure 22.** Arbres phylogénétiques basé sur les séquences du génome entier des souches *Streptomyces ambofaciens* ATCC 23877, *Streptomyces leeuwenhoekii* et *Streptomyces sp.* KY75. PATRIC intègre ces

génomés et sélectionne les familles de protéines globales pour déterminer leur position phylogénétique. Les alignements d'acides aminés et de nucléotides sont ensuite utilisés pour construire l'arbre phylogénétique à l'aide de RaxML, en obtenant des valeurs de soutien grâce au bootstrapping rapide.

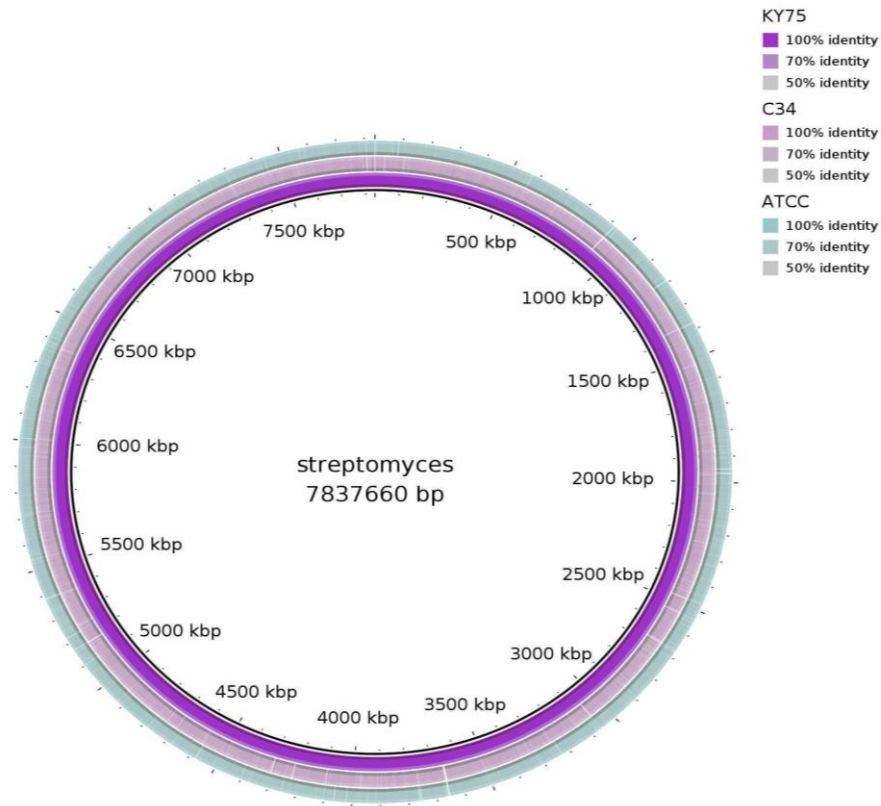
## **I.2. Comparaisons des génomes entiers**

L'analyse réalisée avec BRIG a permis d'obtenir des informations détaillées sur la comparaison du génome de *Streptomyces sp. KY75* avec les deux autres génomes de *Streptomyces ambofaciens* ATCC 23877 CP012382 et *Streptomyces leeuwenhoekii* strain C34 (LN831790). Les anneaux colorés dans la figure 23 représentent les correspondances BLAST et permettent de visualiser les variations d'identité entre les séquences génomiques.

L'observation des couleurs des anneaux montre une correspondance élevée entre les séquences, indiquant un pourcentage élevé d'identité. En effet, la synténie, qui représente la conservation de l'ordre des gènes sur le chromosome, est forte et atteint 99% d'identité au niveau de l'ADN entre les souches étudiées. Cela suggère une conservation significative de la structure chromosomique et une proximité génomique entre ces espèces.

Cependant, il est important de noter l'existence de plusieurs régions de délétion sur les deux chromosomes des séquences de *Streptomyces* en comparaison avec la séquence de référence *Streptomyces sp. KY75*. Ces régions de délétion peuvent entraîner des différences fonctionnelles et des variations dans les capacités métaboliques entre les souches.

Ces résultats soulignent l'importance de l'analyse comparative des génomes pour mieux comprendre la diversité génétique entre différents génomes. Ils mettent également en évidence la nécessité de prendre en compte ces variations lors de l'interprétation des caractéristiques biologiques et des capacités métaboliques de chaque souche.



**Figure 23.** Comparaison du génome des trois souches de *Streptomyces*. De l'extérieur vers l'intérieur, la légende indique le gradient de couleur pour le % de similarité. Anneau 1 *Streptomyces ambofaciens* ATCC 23877 CP012382, anneau 2 *Streptomyces leeuwenhoekii* strain C34 (LN831790) et Anneau3 *Streptomyces sp.* KY75 (NZ\_LR881953.1).

### I.3. Détermination de l'identité nucléotidique moyenne (ANI)

L'analyse de l'identité nucléotidique moyenne (ANI) est une méthode utilisée pour évaluer de manière quantitative les différences et les similitudes génétiques entre les souches étudiées. Dans notre cas, l'outil jspeciesWS a été utilisé pour calculer l'ANI et confirmer de manière plus précise la diversité génétique distincte entre les souches étudiées (Tableau 7).

**Tableau 7.** Valeurs moyennes d'identité nucléotidique des souches comparées.

	<i>Streptomyces ambofaciens</i> ATCC 23877	<i>Streptomyces leeuwenhoekii</i> C34	<i>Streptomyces sp.</i> KY75
<i>Streptomyces ambofaciens</i> ATCC 23877	*	87.42	85.09
<i>Streptomyces leeuwenhoekii</i> C34	87.42	*	85.11
<i>Streptomyces sp.</i> KY75	85.10	85.11	*

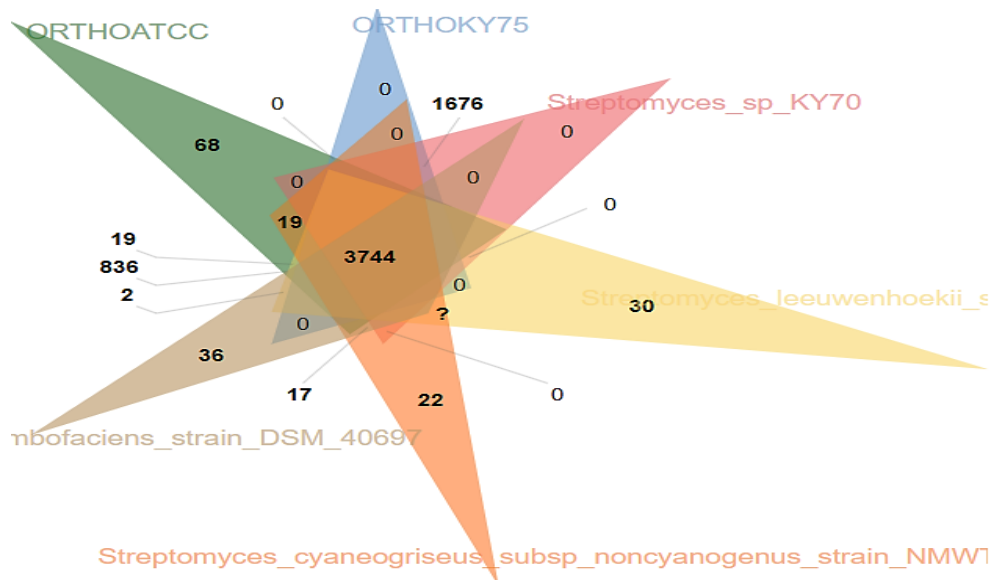
Légende : Au-dessus du seuil (> 95 %) En dessous du seuil (< 95 %)

Les résultats ont révélé des pourcentages d'ANI de 85,09%, 87,42% et 85,11%, ce qui indique des variations significatives par rapport au seuil de 95% généralement utilisé pour déterminer l'appartenance à la même espèce. La comparaison entre *S.sp.* KY75 et *S. ambofaciens* ATCC a montré un pourcentage d'ANI de 85,10%, renforçant ainsi la conclusion selon laquelle ces souches sont génétiquement différentes.

Ces résultats démontrent clairement la diversité substantielle entre ces souches et confirment qu'elles représentent des entités distinctes sur le plan génomique. Il est pertinent de noter que ces souches ont été sélectionnées en prenant en compte des paramètres liés à des écosystèmes extrêmes, ce qui ajoute une dimension intéressante à leur diversité génétique et souligne leur adaptation à des conditions environnementales spécifiques.

#### I.4. Analyse de similitude des protéines prédites

Afin de soutenir les observations précédentes, une analyse de similarité des clusters de protéines a été effectuée pour les trois souches et leurs voisins les plus proches génétiquement, pour déduire si ces souches sont orthologues ou paralogues (Figure 24).



**Figure 24.** Diagramme de Venn du nombre de protéines partagées et uniques entre *Streptomyces ambofaciens* ATCC 23877 CP012382, *Streptomyces leeuwenhoekii* strain C34, *Streptomyces sp.* KY75 (NZ\_LR881953.1) et les souches les plus proches *Streptomyces ambofaciens* DSM 406, *Streptomyces cyaneogriseus subsp noncyanogenus* NMWT et *Streptomyces sp.* KY70, tel que créé par OrthoVenn3.

Les résultats obtenus grâce à l'outil d'orthovenn3 ont permis de générer un diagramme illustrant les similitudes entre les gènes codant pour les protéines. Il est clairement visible que ces six souches partagent 3744 clusters de gènes en commun.



La souche *Streptomyces ambofaciens* DSM 40697 montre une proximité marquée avec *Streptomyces ambofaciens* ATCC 23877, présentant un nombre significatif de 836 clusters de gènes en commun.

De manière similaire, *Streptomyces cyaneogriseus subsp noncyanogenus* NMWT est étroitement liée à *Streptomyces leeuwenhoekii* C34 (= DSM 42122 = NRRL B-24963), bien que le nombre exact de clusters partagés ne soit pas défini.

Il est intéressant de noter que *Streptomyces sp.* KY75 possède un gène similaire à celui de *Streptomyces sp.* KY70, en soulignons que ces deux souches présentent dans le même environnement pollué ont un nombre relativement élevé de 1676 clusters de gènes en commun.

En revanche, *Streptomyces sp.* KY75 ne présente aucun cluster spécifique partagé avec les autres espèces, à l'exception de *S. sp.* KY70. Concernant *Streptomyces leeuwenhoekii* C34, il partage 19 clusters de gènes spécifiques avec *Streptomyces ambofaciens* ATCC 23877.

Cela suggère que ces espèces sont des orthologues vu qu'elles ont un nombre énorme de clusters en commun.

## **II. Annotation génomique**

### **II.1. Annotation structurelle et fonctionnelle du génome**

Les génomes complets de *Streptomyces KY75*, *Streptomyces Leewenhoeki C34* et *Streptomyces Ambofacien* ATCC278992 ont été obtenus à partir de GenBank et annotés par les serveurs RAST et PATRIC (voir Annexe). Après la comparaison des génomes des souches *Streptomyces KY75*, *Streptomyces Leewenhoeki C34* et *Streptomyces Ambofacien* ATCC278992, nous avons observé des similarités et des différences significatives dans leurs caractéristiques génétiques. Les caractéristiques génomiques de ces souches sont analysées et présentées dans le tableau 8.

**Tableau 8.** Caractéristiques génomiques des souches étudiées *Streptomyces* KY75, *Streptomyces leeuwenhoekii* C34 et *Streptomyces ambofaciens* ATCC278992.

Caractéristiques		Caractéristiques du chromosome	
Génome	<i>Streptomyces</i> sp. KY75 (Taxonomy ID : 2772433)	<i>Streptomyces leeuwenhoekii</i> C34 (= DSM 42122 = NRRL B-24963) (Taxonomy ID : 1437453)	<i>Streptomyces ambofaciens</i> ATCC 23877 (Taxonomy ID : 278992)
Taxonomie	<i>Actinomycètes</i> ; <i>Actinomycètes</i> ; <i>Kitasatosporales</i> ; <i>Streptomycètes</i> ; <i>Streptomyces</i> ; <i>Streptomyces</i> sp. KY75	<i>Actinomycètes</i> ; <i>Actinomycètes</i> ; <i>Kitasatosporales</i> ; <i>Streptomycètes</i> ; <i>Streptomyces</i> ; <i>Streptomyces leeuwenhoekii</i> ; <i>Streptomyces leeuwenhoekii</i> C34 (= DSM 42122 = NRRL B-24963)	<i>Actinomycetota</i> ; <i>Actinomycetes</i> ; <i>Kitasatosporales</i> ; <i>Streptomycetaceae</i> ; <i>Streptomyces</i> ; <i>Streptomyces ambofaciens</i> ATCC 23877 ;
Taille	7 837 660	7 903 895	<b>8 303 940</b>
Contenu GC	71,6	72,8	<b>72,2</b>
Nombre de sous-systèmes	312	306	<b>326</b>
Nombre d'ARN	84	85	<b>87</b>
CDS	12034	7097	<b>7608</b>
Régions répétées	87	57	<b>316</b>
ARNt	65	66	<b>70</b>
ARNr	17	19	<b>17</b>
Caractéristiques des protéines			
Protéines hypothétiques	4340	2180	<b>2090</b>
Protéines avec assignation fonctionnelle	7694	4917	<b>5518</b>
Protéines avec affectation de Groupe opérationnel GO	1889	1116	<b>1124</b>

Protéines avec des attributions de voies métaboliques	1709	1011	1025
---	------	------	------

Tout d'abord, en termes de taille, *S. sp.* KY75 possède un génome de taille de 7 837 660 paires de bases, *S. Leewenhoeki* C34 est légèrement plus grand avec 7 903 895 paires de bases, tandis que *S. Ambofacien* possède le génome le plus long avec une taille de 8 303 940 paires de bases. La variation de la taille des génomes est largement influencée par la niche écologique des organismes. Les organismes vivant dans des niches complexes et changeantes tendent à avoir des génomes de grande taille. Par exemple, des bactéries telles que *B. japonicum* (9 105 828 paires de bases) et *Streptomyces scabies* (10 148 695 paires de bases), qui vivent dans le sol, possèdent des génomes parmi les plus grands génomes séquencés (Gallois, 2007).

Une autre différence notable réside dans la teneur en guanine-cytosine (GC) de ces génomes. *S. sp.* KY75 présente un pourcentage GC de 71,6%, *S. leewenhoeki* C34 de 72,8% de GC, et *S. ambofacien* de 72,2% de GC. Ces pourcentages sont comparables avec ceux qui ont été obtenus pour des isolats à partir d'échantillons d'eau du lac Fetzara (Annaba, Algérie). Les séquences préliminaires du génome ont 7,0 à 8,5 Mb avec une teneur moyenne en G + C de 71,8%, ce qui est en bon accord avec le genre (Vicente et al., 2018).

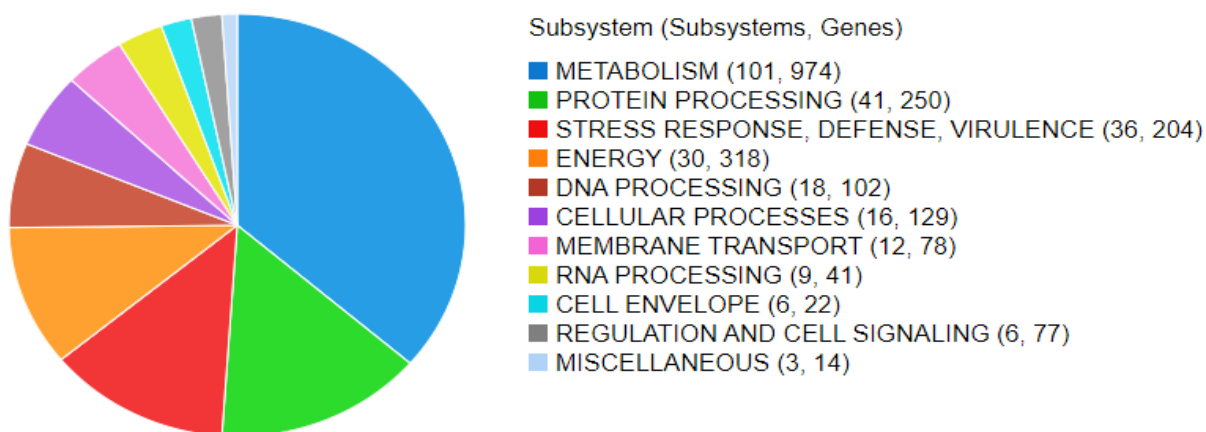
La teneur génomique bactérienne en guanine + cytosine (G+C) peut varier énormément selon les espèces. Il a été démontré que cette plage s'étendait 75 % en moles (Luciano, 2013). Les facteurs influençant cette variation sont le biais mutationnels agit sur les génomes. On pensait que ce biais, associé à des facteurs environnementaux, exerçait une pression de sélection vers la composition du génome la plus adaptée pour un habitat donné. Des recherches ultérieures ont suggéré que ce biais mutationnel agit généralement sur toutes les espèces bactériennes (Almpanis et al., 2018).

Les séquences codantes (CDS) présentent également des différences entre les souches. La souche *S. sp.* KY75 compte un total de 12 034 CDS, parmi lesquelles 7 694 sont des clusters gènes qui code pour des protéines fonctionnelles dont 1 709 code pour des protéines aux fonctions métaboliques. De plus, cette souche, prélevée dans un environnement pollué par le zinc et le plomb, possède 84 séquences codant pour des ARN, représentant le plus grand nombre de CDS et BGC qui code pour des protéines par rapport aux souches provenant d'environnements thermaux. *S. ambofacien*, quant à lui, possède 7 608 CDS, dont 5 518 correspondent à des BGC de protéines fonctionnelles et 87 clusters correspondent à des ARN. Le génome de *S. leewenhoeki* C34 présente le

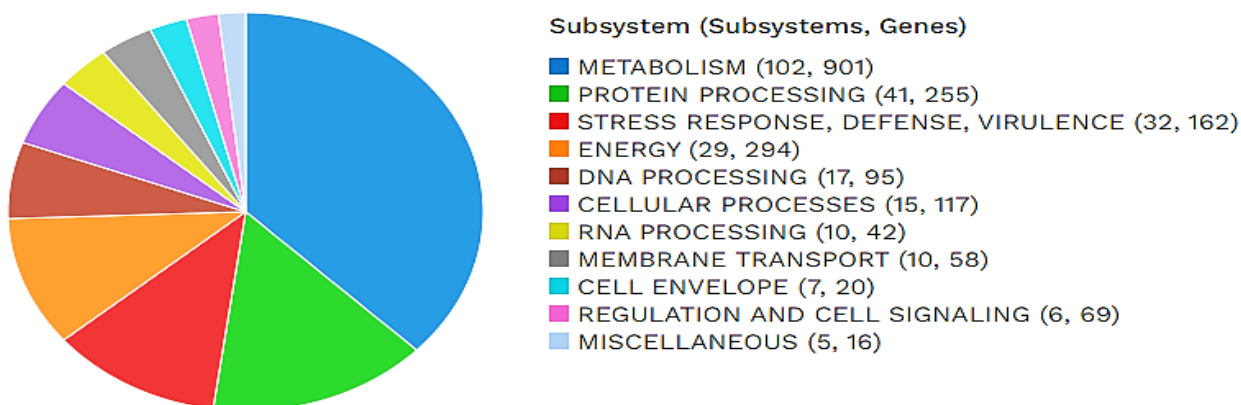
nombre le plus faible de 7 097 séquences CDS, comprenant 4 917 gènes codant pour des protéines fonctionnelles et 85 séquences codant pour des ARN.

Les comparaisons des génomes de ces souches révèlent quelques différences significatives. Les variations de taille et de composition en guanine-cytosine indiquent une diversité génomique entre ces souches. De plus, les variations dans le nombre et les types de séquences codantes reflètent des caractéristiques spécifiques de chaque génome, notamment en ce qui concerne les fonctionnalités des protéines et les activités métaboliques.

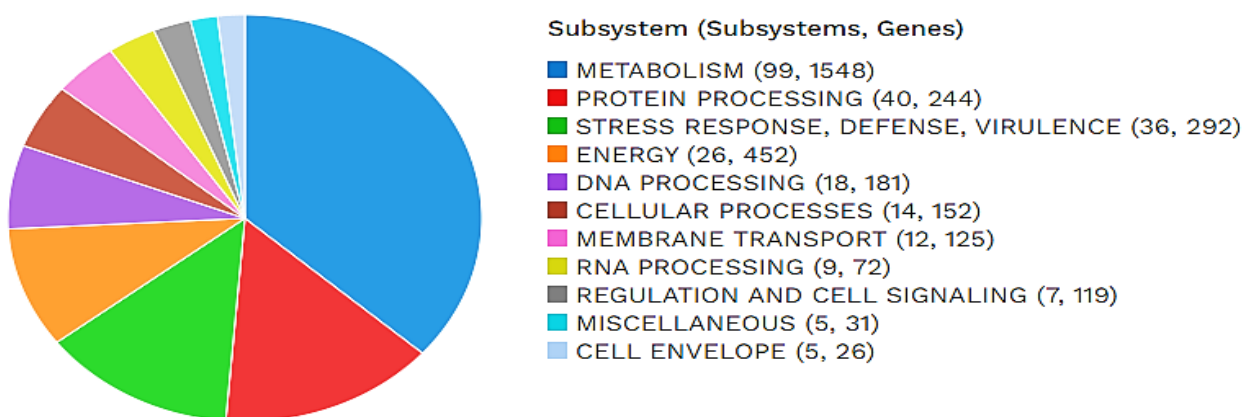
Les gènes et les protéines prédites des *Streptomyces* présents dans des écosystèmes pollués et thermaux sont également très similaires (98 à 99%) à ceux des autres espèces de *Streptomyces*. Cependant, en raison des conditions environnementales spécifiques de ces écosystèmes, il peut y avoir des différences dans les mécanismes moléculaires associés à la résistance à la dessiccation et à la tolérance aux radiations.



**Figure 25.** Informations sur les sous-systèmes des gènes pour la souche *Streptomyces ambofaciens* ATCC 23877, obtenu par le serveur PATRIC.



**Figure 26.** Aperçu sur les sous-systèmes des gènes pour la souche *Streptomyces leeuwenhoekii* C34 (= DSM 42122 = NRRL B-24963) (1437453.21), obtenu par le serveur PATRIC.



**Figure 27.** Informations sur les sous-systèmes des gènes pour la souche *Streptomyces* sp. KY75, obtenu par le serveur PATRIC.

Dans le cas des *Streptomyces* présents dans des écosystèmes pollués, des études ont montré des adaptations spécifiques pour faire face aux stress environnementaux causés par la pollution. Par exemple, certaines souches ont développé des mécanismes de dégradation enzymatique pour résister et métaboliser les contaminants présents dans leur environnement.

D'autre part, les *Streptomyces* trouvés dans des écosystèmes thermaux sont exposés à des températures élevées, ce qui nécessite des mécanismes de résistance spécifiques. Des recherches ont révélé que ces bactéries thermophiles possèdent des adaptations moléculaires pour stabiliser leurs protéines et maintenir leur fonctionnalité à des températures élevées.

Dans l'ensemble, bien que les gènes et les protéines prédites des *Streptomyces* des écosystèmes pollués et thermaux soient similaires à ceux des autres espèces de *Streptomyces*, ils peuvent présenter des variations adaptatives spécifiques à leur environnement. Ces adaptations peuvent inclure des

mécanismes de dégradation des contaminants pour les souches des écosystèmes pollués et des mécanismes de stabilisation des protéines pour les souches des écosystèmes thermaux.

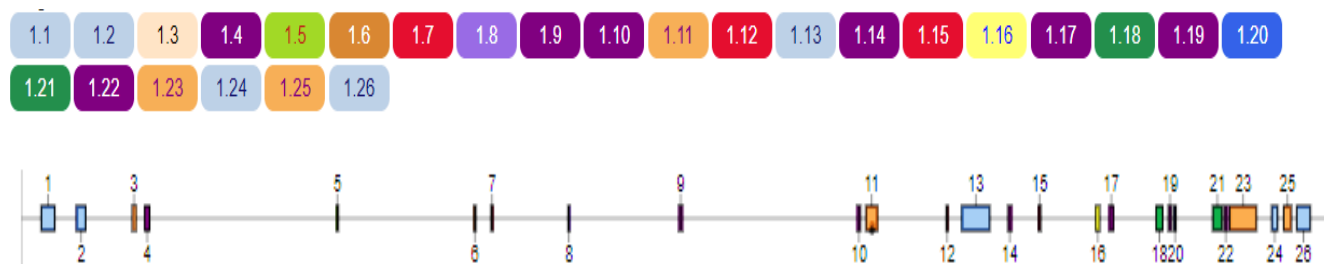
## II.2. Analyses des gènes des métabolites secondaires bioactifs et des composés

### II.2.1. Détection des régions de métabolites secondaires bioactifs

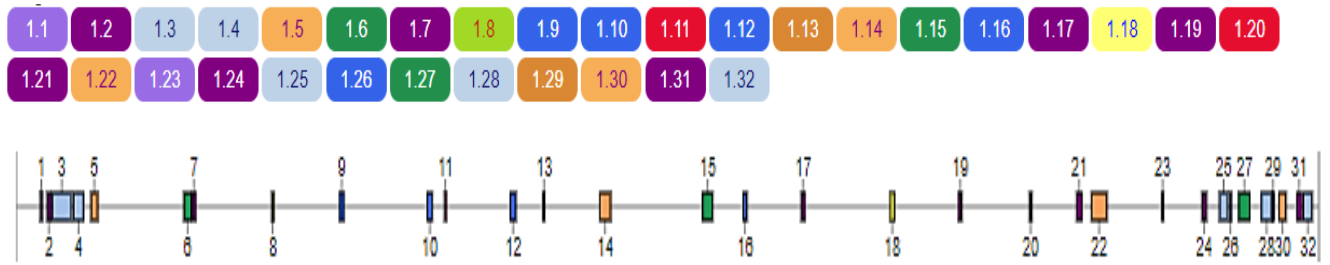
Grâce à l'utilisation du serveur antiSMASH, nous avons pu mettre en évidence les différentes régions de métabolites secondaires présentes dans les génomes des souches bactériennes étudiées (Figures 25, 26, 27 et 28). Cette information est cruciale pour comprendre la capacité de ces bactéries à produire des métabolites d'intérêt biologique et leur potentiel dans des applications médicales, pharmaceutiques ou environnementales.



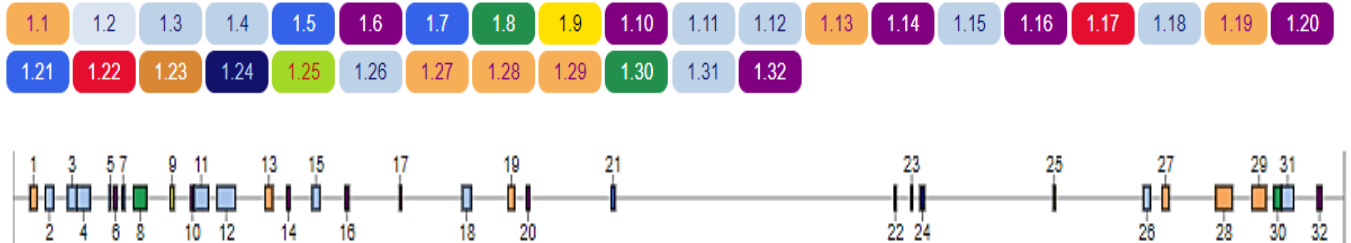
**Figure 28.** Identification des régions de métabolites secondaires à l'aide du serveur antiSMASH.



**Figure 29.** Groupes de gènes putatifs codant pour des métabolites secondaires dans *Streptomyces ambofaciens* ATCC 23877.

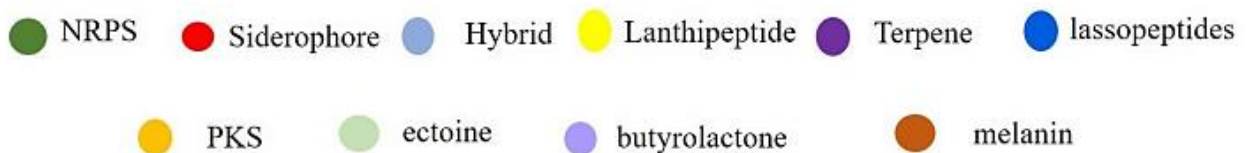


**Figure 30.** Groupes de gènes putatifs codant pour des métabolites secondaires dans *Streptomyces sp.* KY75.



**Figure 31.** Groupes de gènes putatifs codant pour des métabolites secondaires dans *Streptomyces leeuwenhoekii* C34.

*Streptomyces leeuwenhoekii* et *Streptomyces sp.* KY75 ont montré un total de 32 BGC, tandis que *Streptomyces ambofaciens* ATCC 23877 en avait 26. Ces BGC sont des régions spécifiques du génome impliquées dans la biosynthèse des métabolites secondaires.



**Figure 32.** Indicateurs de couleurs (Alam et al., 2022).

La coloration associée à ces régions (Figure 29) permet de les distinguer visuellement et d'évaluer la diversité des métabolites produits par chaque souche. Cette caractérisation des BGC revêt une importance capitale dans la compréhension du potentiel métabolique des souches bactériennes, notamment dans des domaines tels que la recherche pharmaceutique et la biotechnologie. Elle ouvre la voie à de nouvelles découvertes et à l'exploitation de ces métabolites dans diverses applications scientifiques.

Les séquences du génome ont été analysées avec antiSMASH 7.0 pour l'identification des grappes de gènes. La catégorie des grappes de gènes hybrides PKS-NRPS. D'autres clusters hybrides qui ont été détectés ont été séparés en leurs éléments constitutifs pour mieux visualiser la diversité des

clusters. Les clusters codant pour les betalacton, les lantipeptides et les lassopeptides sont regroupés dans la catégorie "RiPP", et la catégorie "Autre" comprend d'autres voies biosynthétiques et des clusters moins courants tels que les indoles et les butyrolactones, entre autres (Tableau 9).

**Tableau 9.** Groupes de gènes biosynthétiques de métabolites spécialisés identifiés dans les souches analysées.

BGC	<i>Streptomyces</i> KY75	<i>Streptomyces</i> <i>ambofaciens</i> ATCC 23877	<i>Streptomyces</i> <i>leeuwenhoekii</i> C34
<b>T2PKS</b>	1	3	1
<b>Butyrolactone</b>	2	3	1
<b>Terpene</b>	6	8	7
<b>T1PKS</b>	2	5	9
<b>NRPS</b>	6	4	3
<b>Indole</b>	-	1	-
<b>Ectoine</b>	-	1	1
<b>Melanin</b>	3	1	1
<b>NI-siderophore</b>	1	3	3
<b>Furan</b>	-	1	-
<b>Oligosaccharide</b>	-	1	-
<b>Lanthipeptide class iii</b>	1	1	-
<b>RiPP-like</b>	2	1	
<b>NRP-Metallophore</b>	2	1	1
<b>Lanthipeptide class-i</b>	-	1	-
<b>NRPS-like</b>	-	-	6
<b>Autre</b>	-	-	1
<b>RRE-containg</b>	-	-	1
<b>LAP</b>	1	-	1
<b>PKS-like</b>	-	-	1
<b>TransAT-Pks</b>	1	-	1
<b>Betalactone</b>	1	-	1
<b>Lassopectide</b>	1	-	1
<b>Amglyccycl</b>	-	-	1
<b>CDPS</b>	-	-	1
<b>T3PKS</b>	1	-	-
<b>Thiopeptide</b>	1	-	-



À l'aide de l'outil antiSMASH 7.0, nous avons détecté les BGC dans les génomes des souches bactériennes analysées. Un total de 27 clusters ont été identifiés, dont 8 sont communs entre les trois souches.

La souche *Streptomyces sp.* KY75 abrite-le moins de BGC, avec 32 identifiés, tandis que la souche *Streptomyces ambofaciens* ATCC 23877 en possède le plus, avec 35 clusters. La corrélation entre le nombre de BGC et la taille du génome n'est conforme aux attentes, où la souche ayant le génome le plus petit et le plus grand nombre de cuistres.

Il est intéressant de noter que la souche *Streptomyces leeuwenhoekii* C34 présente le plus grand nombre de BGC, avec 41 identifiés. Les BGC découverts dans les trois souches appartiennent aux classes principales de métabolites, incluant les polykétide synthases (PKS), les synthétases non ribosomales (NRPS) et les hybrides PKS-NRPS. Ces résultats mettent en évidence la diversité et le potentiel métabolique des souches bactériennes étudiées.

Lors de l'analyse des résultats de prédiction des BGC par AntiSMASH, 26 clusters ont été identifiés dans *Streptomyces ambofaciens* ATCC 23877 et 32 dans le génome de *S. leewenhoekii*, présentant des similarités variables avec des clusters connus est d'autres sont non identifier. Parmi ces clusters, 3 clusters ne sont pas identifier dans les deux bactéries, 13 ont montré une similarité de plus de 80% avec des clusters de gènes connus dans *S. ambofaciens* ATCC 23877 et 10 dans le génome de *S. leehenhokii*, ce qui suggère une grande capacité biosynthétique pour la production de métabolites tels que l'antimycine, l'ectoïne, la mélanine, l'albaflavenone, la géosmine, la spiramycine, le coelichelin, et d'autres.

Ces métabolites ont des rôles importants dans la survie et l'adaptation des bactéries de genre *Streptomyces* dans les écosystèmes extrêmes, notamment dans des environnements thermiques. Par exemple, la présence de clusters de gènes impliqués dans la biosynthèse de la mélanine suggère une capacité de protection contre les radiations UV et d'autres stress environnementaux (Hullo et al., 2001; Moeller et al., 2005). De même, la production d'ectoïne peut offrir une protection contre le stress osmotique (Reshetnikov et al., 2011), tandis que la géosmine peut jouer un rôle dans la communication chimique et l'attraction des organismes environnants, Il est important de noter que ces composés sont produits principalement pendant la sporulation dans le cycle de vie classique de *Streptomyces* (Zambri et al., 2022).

En outre, la présence de clusters de gènes associés à des métabolites tels que l'antimycine, l'albaflavenone et la spiramycine indique la capacité des *Streptomyces* à produire des composés antimicrobiens et potentiellement thérapeutiques.

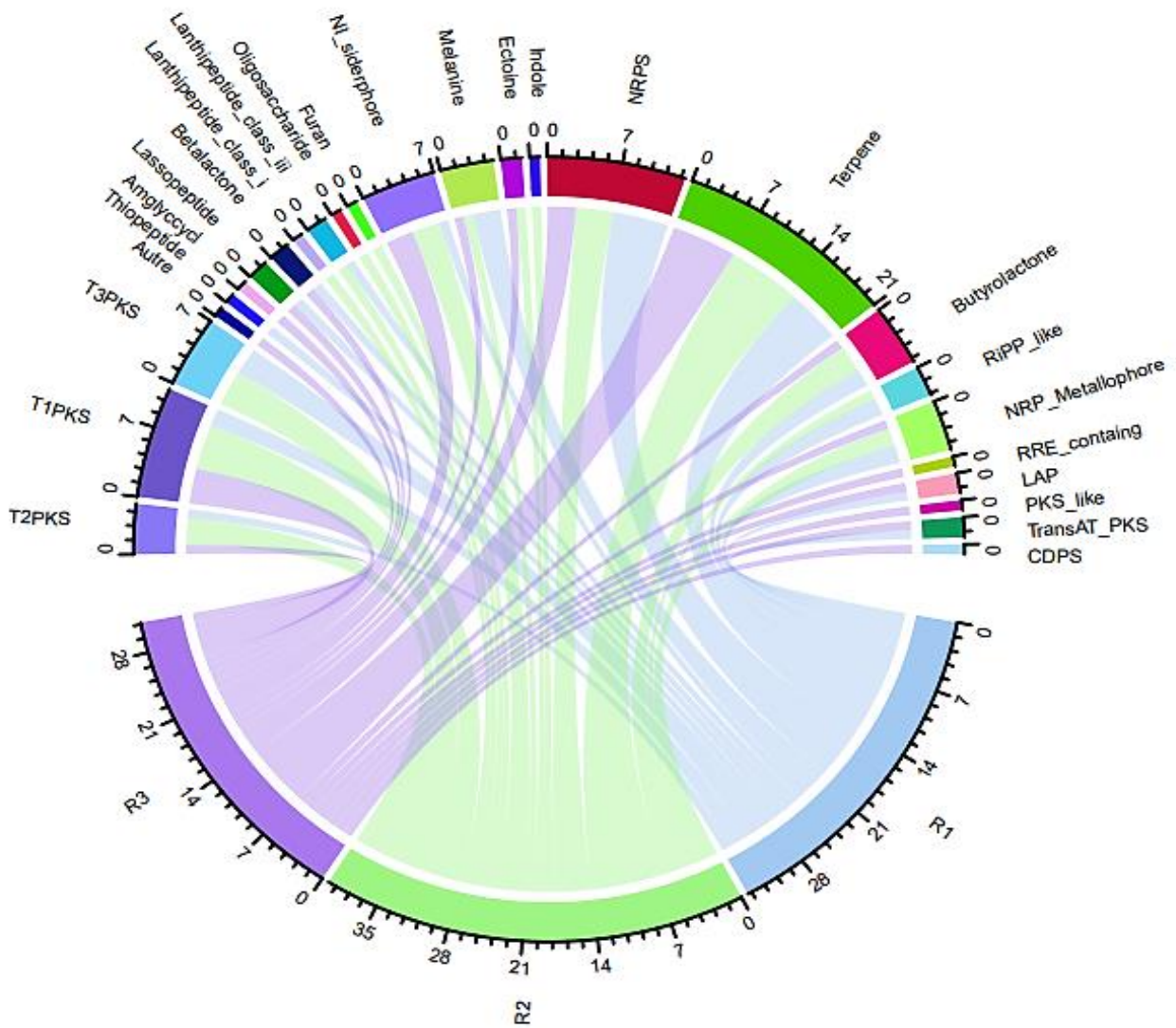
Une étude sur de deux bactéries du désert algériens a été faite par Zerouki et al. (2021) pour prédire la présence de différents types de grappes de gènes de biosynthèse (BGC) dans les souches Babs14 et Osf17 de *Streptomyces*. Les résultats ont révélé un total de 29 groupes de gènes biosynthétiques dans chaque souche, dont 26 groupes présentaient une similitude avec des BGC connus. Huit groupes de gènes ont été identifiés comme contenant tous les gènes d'un groupe connu à 100%. Certains de ces groupes étaient associés à des BGC impliqués dans la production de métabolites spécifiques tels que l'ectoïne, géosmine, qui sont souvent trouvés dans les souches de *Streptomyces*. Par exemple, l'ectoïne offre une protection contre le stress osmotique et sert de nutriment polyvalent (Schulz et al., 2017). En comparant les quatre souches provenant d'écosystèmes thermiques différents, nous avons constaté que les bactéries étudiées du Sahara algérien présentent une capacité biosynthétique plus élevée que les bactéries *Streptomyces ambofaciens* ATCC 23877 et *S. leewenhoekii*. Cette observation confirme l'impact du degré de température sur l'adaptation des *Streptomyces* pour assurer leur survie. Lors de la comparaison entre les souches provenant des milieux thermiques, des différences significatives ont été observées au niveau de leurs capacités biosynthétiques.

La souche *S. sp.* KY75 présente 14 clusters de gènes, dont la majorité montre une forte similarité avec des clusters déjà identifiés. Parmi ceux-ci, 12 clusters présentent une similitude de 100% avec des gènes connus. Les clusters de gènes de cette souche sont principalement associés à la synthèse de terpènes et de polykétides, en particulier de type 3, qui ne sont pas détectés dans la souche *S. Ambofaciens*. Ces différences dans les capacités de synthèse reflètent l'adaptation génomique de *S. sp.* KY75 à son environnement. Il est important de noter que cette a été isolée d'un environnement contaminé par le zinc et le plomb, mettant en évidence sa capacité d'adaptation à des conditions environnementales difficiles.

De plus, l'analyse a révélé que *S. KY75* partage des capacités de synthèse communes avec d'autres *Streptomyces* présents dans des milieux thermiques, tout en présentant également des capacités spécifiques, telle que la production d'alkylrésorcinol, qui lui est propres.

### II.2.2. Profilage des métabolites secondaires

Les résultats de l'annotation des BGC par antiSMASH ont été visualisés par d'un diagramme de corde circulaire (Figure 30), permettant de mettre en évidence les classes de BGC prédites dans chaque espèce. Cette approche a permis de mettre en évidence les points communs des voies biosynthétiques présentes dans les souches étudiées.

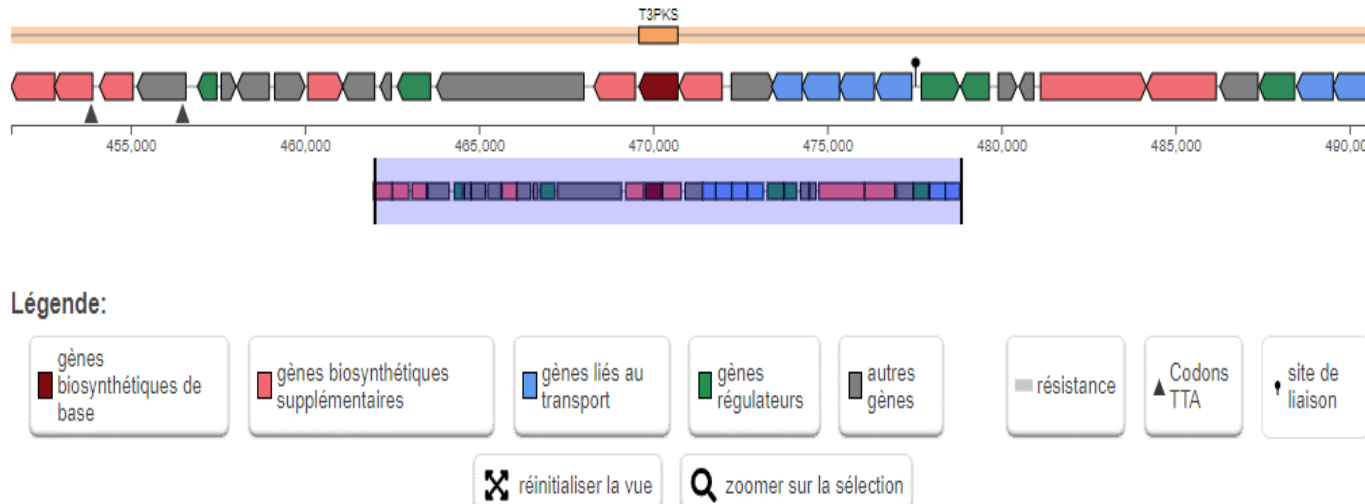


**Figure 33.** Diagramme de corde circulaire pour visualiser les classes de BGC prédites dans les souches étudiées. L'arc inférieur représente les génomes/plasmides R1 (*Streptomyces* sp. KY75), R2 (*Streptomyces* ambofaciens ATCC 23877) et R3 (*Streptomyces* leeuwenhoekii C34), tandis que l'arc supérieur représente les différentes classes de BGC. Les rubans de couleur indiquent la présence spécifique d'un BGC dans chaque espèce.

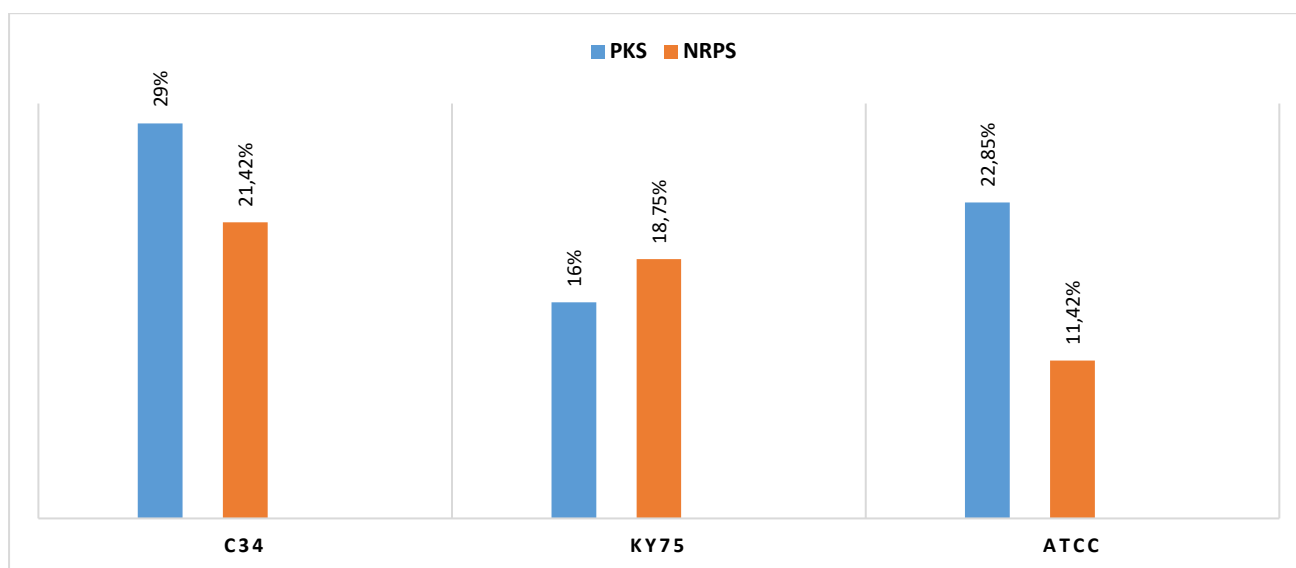
En comparant les trois souches, on observe que la souche *Streptomyces ambofaciens* présente la plus grande proportion de clusters biosynthétiques de terpènes (22,85%), suivie par *Streptomyces sp.* KY75 avec 18,75% et *Streptomyces leeuwenhoekii* avec 17%.

En revanche, les clusters biosynthétiques de PKS de type 1 sont plus fréquents dans *S. Leeuwenhoekii*, tandis que *S. sp.* KY75 présente une plus grande proportion de clusters NRPS. De plus, des clusters biosynthétiques tels que les PKS de type 2, les NRP\_métallophores, les buttirolactones et les butyrolactones Ni\_sidrophores sont également présents, mais à des taux plus faibles et de manière similaire entre les souches. Ces résultats mettent en évidence à la fois les similitudes et les différences dans les profils de clusters biosynthétiques des souches étudiées.

Les Clusters de Gènes Biosynthétiques (BGC) sont distribués dans le génome bactérien et sont prédits par l'outil AntiSMASH. Dans la figure 31, on peut observer la présence de polykétides de type 3 dans la région 19 du génome de *S. Lee*. Cette région contient plus de 5 gènes associés aux processus fondamentaux de biosynthèse, de régulation et de transport, qui travaillent ensemble pour produire une molécule possédant des propriétés thérapeutiques. Pour identifier cette molécule, on utilise la fonction KnownBlaster en combinaison avec la base de données MIBiG.



**Figure 34.** Les résultats des gènes biosynthétiques obtenus par antiSMASH.



**Figure 35.** Distribution des domaines PKS et NRPS dans les souches *Streptomyces* KY75, *Streptomyces leewenhoeki* C34 et *Streptomyces ambofacien* ATCC278992.

L'analyse comparative des pourcentages de PKS et NRPS dans les trois souches (Figure 32) révèle des variations significatives. On constate que le pourcentage de PKS est plus élevé dans la souche *S. leewenhoeki* C34, atteignant 29%. Ensuite, la souche *S. ambofacien* TCC présente un taux de 22,85%, tandis que la souche *S.sp.* KY75 affiche le taux le plus bas, avec seulement 16%. Ces résultats suggèrent que les bactéries présentes dans les écosystèmes thermiques ont une plus grande proportion de PKS, ce qui indique une capacité accrue à produire des métabolites secondaires de ce type.

En ce qui concerne les NRPS, la souche *S. leewenhoeki* C34 présente un pourcentage élevé de 21,42%. La souche *S.sp.* KY75 affiche un taux modéré de 18,75%, tandis que la souche *S. ambofacien* ATCC présente le taux le plus faible de NRPS, avec seulement 11,42%. Ces variations dans les taux de gènes NRPS suggèrent une variation dans la production de métabolites secondaires et dans leurs interactions et fonctions biologiques.

En résumé, l'analyse comparative des pourcentages de PKS et NRPS entre les souches indique des différences significatives dans la capacité de production de métabolites secondaires et suggère des variations dans les interactions et les fonctions biologiques associées à ces gènes.

### II.2.3. Prédiction des composés secondaires bioactifs

Les protéines sont des molécules essentielles dans de nombreux processus biologiques, permettant la régulation enzymatique, la transmission de signaux cellulaires, et d'autres fonctions

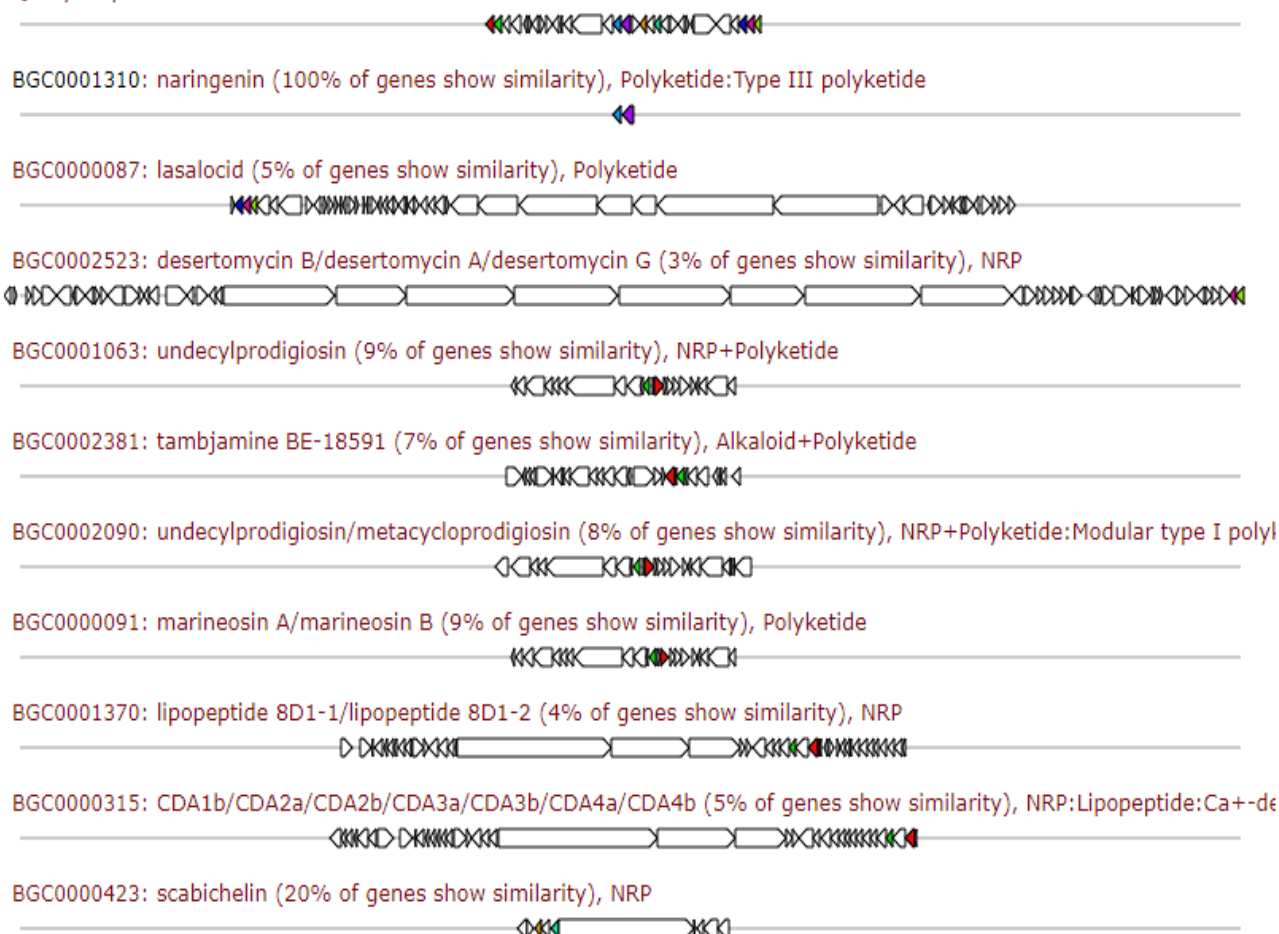
biologiques. Les polykétide synthases (PKS) et les nonribosomal peptide synthétases (NRPS) sont des enzymes responsables de la biosynthèse de métabolites spécialisés. Ces métabolites ont des activités biologiques importantes et sont utilisés par les organismes pour interagir avec leur environnement, notamment dans la défense contre les pathogènes ou la communication intercellulaire (Haas, 2015; Mitra et al., 2022)

### **II.2.3.1. Groupes de gènes connus similaires de MIBiG 3.1**

Ici, les analyses bio-informatiques de la séquence du génome entier des trois souches étudiées ont mis en évidence de nombreuses voies de biosynthèse. Cependant, l'analyse bio-informatique a permis de découvrir de nombreux groupes de gènes ne sont pas liés aux groupes reconnus (Figure 36 et Tableau 10). Des recherches supplémentaires en utilisant la base de données MIBIG seront donc nécessaires à l'avenir pour améliorer, isoler et identifier les nouvelles molécules bioactives de ces souches.

L'affichage des clusters de la base de données MIBIG présente les similitudes avec la région actuelle. Les gènes marqués de la même couleur sont interdépendants, tandis que les gènes blancs n'ont aucune relation. Les gènes de référence permettent d'accéder aux détails des similitudes avec les gènes de la région actuelle. Vous pouvez utiliser le numéro d'accès pour ouvrir cette entrée dans la base de données MiBIG.

Query sequence



**Figure 36.** Représentations graphiques des clusters connus les plus similaires en tant que résultat de spécifié par KnownBlaster dans le référentiel MIBiG. Les codes de couleur représentent les régions génétiques similaires appartenant au groupe de PKS3.

**Tableau 10.** Combinaison des résultats d'antiSMASH et MIBiG, variations des régions NRPS et PKS et les types de molécules produites chez différentes souches bactériennes

Souche	Type	Region	Composé	Référence Mibig	Activité Pharmacologique
<b>Streptomyces Ambofaciens ATCC 23877</b>	NRPS	1.18	Netropsin	<a href="#">BGC0000327</a>	Antibiotique Et Antivirale
			Anthelvencin A,B,C	<a href="#">BGC0002042</a>	Anthelminthique Et Antibactérien Modéré
			Indigoidine	<a href="#">BGC0000375</a>	Antioxydant Et Antimicrobien
		1.21	Coelichelin	<a href="#">BGC0000325</a>	
			Potensibactin	<a href="#">BGC0002017</a>	
			Enterobactin	BGC0002476	Antibiotiques Améliore Les Propriétés Antibactériennes
			Chaetocin	<a href="#">BGC0000321</a>	Un Antibiotique Antitumoral

			Cyclo-(D-Phe-L-Phe-D-Val-L-Val)	<a href="#">BGC0000357</a>	Antibactériennes
			Xefoampeptides A-G	<a href="#">BGC0001824</a>	
			Oxepinamide F	<a href="#">BGC0002208</a>	
			AM-Toxin	<a href="#">BGC0001261</a>	Antibiotiques
	PKS	1.11	Spore Pigment	<a href="#">BGC0000272</a>	
			Curamycin	<a href="#">BGC0000215</a>	
			Anthrabenoxocinone	<a href="#">BGC0001513</a>	
			Griseusin A, B	<a href="#">BGC0000231</a>	
			Resistomycin, Resistoflavine	<a href="#">BGC0000264</a>	Antibiotiques
			Jadomycin	<a href="#">BGC0000234</a>	Anticancéreux
			Enterocin A,B,C	<a href="#">BGC0002366</a>	Antibactérien
		1.23	Stambomycin A, B,C,D	<a href="#">BGC0000151</a>	Antibiotiques Et Anticancéreux
			Oronofacic Acid	<a href="#">BGC0000041</a>	
			Amycomicin	<a href="#">BGC0001504</a>	Antimicrobien
			Asperlactone	<a href="#">BGC0001273</a>	Antibactérien Et Antifongique Et Anti- Inflamatoire
			1-Heptadecene	<a href="#">BGC0001163</a>	
			6-Methylsalicylic Acid	<a href="#">BGC0001275</a>	Antibiotique Et Anticancéreux
			Arsono-Polyketide	<a href="#">BGC0001283</a>	
		1.25	Cepacin A	<a href="#">BGC0001897</a>	Antimicrobien
			Nostovalerolactone, 9- Dehydronostovalerolactone	<a href="#">BGC0002711</a>	Antimicrobien
			Nonactin	<a href="#">BGC0000252</a>	Antibiotique
			Aurachin A, D	<a href="#">BGC0001343</a>	Antioxydants
			Dinactin, Trinactin, Tetranactin	<a href="#">BGC0000243</a>	Antibiotique, Antibactérienne
			Youssofene A1, B1,B2,B3,B4	<a href="#">BGC0002547</a>	Antimicrobien
<b>Streptomyces Sp. KY75</b>	NRPS	1.6	Minimycin	<a href="#">BGC0002295</a>	Antibactérienne
			Chromobactin	<a href="#">BGC0002679</a>	
			Indigoidine	<a href="#">BGC0000375</a>	Antioxydant And Antimicrobien
			Mycotrienin I	<a href="#">BGC0000957</a>	Antibiotiques
		1.15	Amphi-Enterobactin 1,2,3,4	<a href="#">BGC0001890</a>	
			Curacomycin	<a href="#">BGC0001561</a>	Antibiotiques Et Antimicrobien



			Pyreudione A,B,C,D,E	<a href="#">BGC0002075</a>	Amibicide, Antibactérienne, Antiproliférative Et Cytotoxique
			Enterobactin	<a href="#">BGC0000343</a>	
			Benarthin, Dibenarthin	<a href="#">BGC0002688</a>	
			Bacillibactin	<a href="#">BGC0001185</a>	Antibiotiques
			Dimethylcoprogen	<a href="#">BGC0001249</a>	Antioxydant
			Aminochelin, Azotochelin, Protochelin	<a href="#">BGC0002528</a>	
			Metachelin C,A, A-CE, B,	<a href="#">BGC0002710</a>	Antioxydant
		1.27	Potensibactin	<a href="#">BGC0002017</a>	
			Ambactin	<a href="#">BGC0001131</a>	Antibactériens
			Holrhizin	<a href="#">BGC0001844</a>	
			Icosalide A, B	<a href="#">BGC0001833</a>	Antibactériens
			Benarthin, Dibenarthin	<a href="#">BGC0002688</a>	Inhibiteur De La Pyroglutamyl Peptidase
			Bovienimide A	<a href="#">BGC0002135</a>	
			Rhodochelin	<a href="#">BGC0000417</a>	Antibiotique
			Vicibactin	<a href="#">BGC0000457</a>	
			Syringafactin A,C	<a href="#">BGC0002518</a>	
			Xenortide A,B,C,D	<a href="#">BGC0000465</a>	Cytotoxique
	PKS	1.5	Naringenin	<a href="#">BGC0001310</a>	Antioxydants, Antitumoraux, Antiviraux, Antibactériens, Anti- Inflammatoires, Antiadipogènes Et Cardioprotecteurs
			Flaviolin, 1,3,6,8- Tetrahydroxynaphthalene	<a href="#">BGC0002127</a>	Anticancéreuses, Antioxydantes, Anti- Inflammatoires Et Antivirales.
			2,4-Diacetylphloroglucinol	<a href="#">BGC0000280</a>	Antibactérien Et Antifongique
			Flaviolin	<a href="#">BGC0000902</a>	Anticancéreuses, Antioxydantes, Anti- Inflammatoires Et Antivirales.
			Flaviolin, Flaviolin Rhamnoside, 3,3'- Diflaviolin	<a href="#">BGC0000285</a>	Anticancéreuses, Antioxydantes, Anti- Inflammatoires Et Antivirales.
		1.14	Oronofacic Acid	<a href="#">BGC0000041</a>	
			Geldanamycin	<a href="#">BGC0000068</a>	En Se Liant Au Site ATP De La HSP90 Et En Interférant Avec Ses

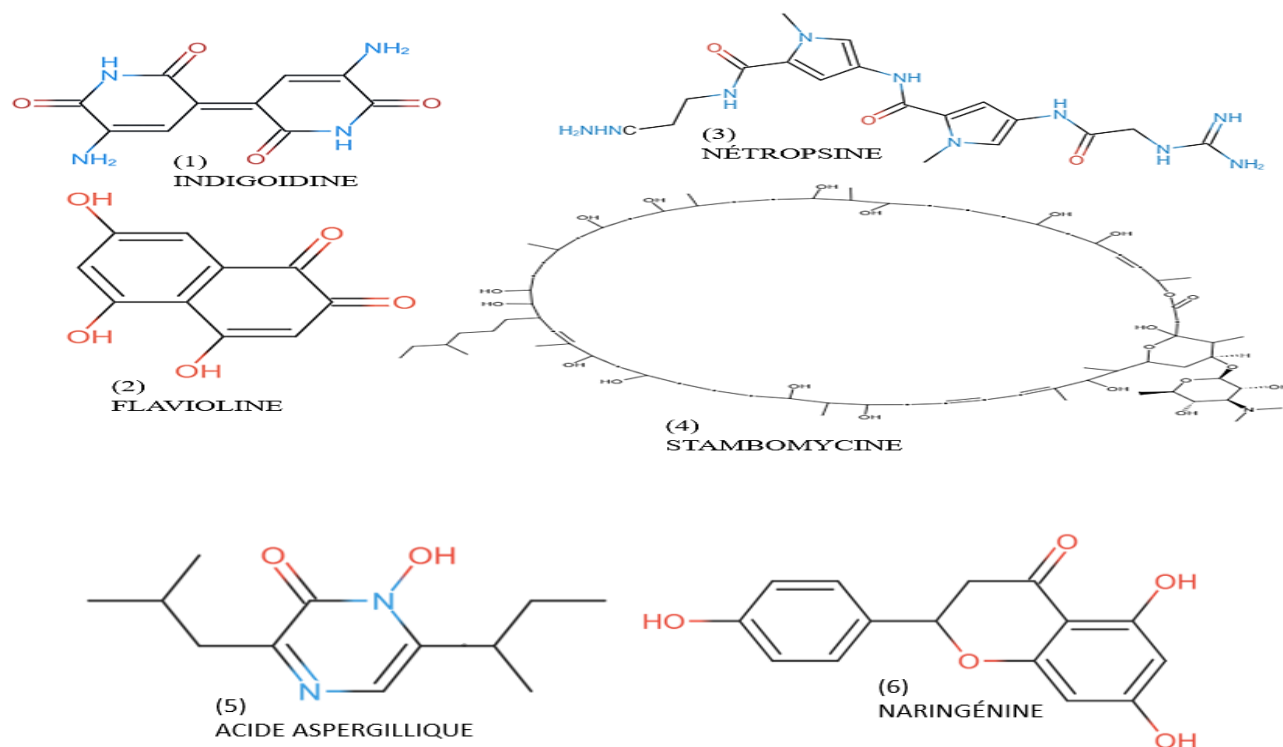
					Fonctions De Chaperonnage
			6-Methylsalicylic Acid	<a href="#">BGC0001275</a>	
			Betaenone A,B,C	<a href="#">BGC0001264</a>	Phytotoxique
			Asperlactone	<a href="#">BGC0001273</a>	Antibactérien, Antifongique, Cytotoxique, Nématocide, Piégeur De Radicaux, Ovicide Et Régulateur De La Croissance Des Insectes
			(-)-Mellein	<a href="#">BGC0001244</a>	Anti-Inflammatoire, Antioxydant, Anticancéreux, Antimicrobien, Antiviral, Antihépatotoxique
			6-MSA, Aculin A	<a href="#">BGC0002229</a>	Antibactérien, Antifongique Et Antiprotozoaire
		1.22	Griseusin A,B	<a href="#">BGC0000231</a>	Antibiotiques
			Jadomycin	<a href="#">BGC0000234</a>	Antimicrobien
			Oviedomycin	<a href="#">BGC0001723</a>	
			Urdamycin	<a href="#">BGC0000278</a>	Antibiotique
			Angolamycin	<a href="#">BGC0000019</a>	Anti-Bactérien
		1.30	BE-24566B, Zunyimycin A	<a href="#">BGC0001366</a>	Antibiotique
			Alkylpyrone-407, 393	<a href="#">BGC0001831</a>	
			Loseolamycin A1,A2	<a href="#">BGC0002362</a>	
			Alkyl-O-Dihydrogeranyl- Methoxyhydroquinones	<a href="#">BGC0001077</a>	Antibiotique
			Alkylresorcinol	<a href="#">BGC0000282</a>	Antimicrobien
<b>Streptomyces leeuwenhoekii strain C34</b>	NPRS	1.8	Taxllaid A	<a href="#">BGC0001133</a>	Antiprotozoaire
			Potensibactin	<a href="#">BGC0002017</a>	
			Thermochelin	<a href="#">BGC0002467</a>	
			Bovienimide Asalinichelins	<a href="#">BGC0002135</a>	Antimicrobien
			Salinichelins	<a href="#">BGC0001767</a>	
			Thermoactinoamide A	<a href="#">BGC0002437</a>	Cyclopeptide Lipophile Antibiotique
			Saccharochelin A,B,C,	<a href="#">BGC0002424</a>	
			Bicornutin A1, A2	<a href="#">BGC0001135</a>	Antifongique Et Antibactérien
			Syringafactin A, C	<a href="#">BGC0002518</a>	Stress Oxidative
			Icosalide A, B	<a href="#">BGC0001833</a>	Antimicrobien
		1.30	Livipeptin	<a href="#">BGC0001168</a>	
			Aspergillilic Acid	<a href="#">BGC0001516</a>	Antibiotique And Antifongique

			FR901483	<a href="#">BGC0002146</a>	Inhibition De La Biosynthèse Des Nucléotides Puriques
	PKS	1.1	Deoxyhangtaimycin	<a href="#">BGC0002657</a>	
			6-Methylsalicylic Acid	<a href="#">BGC0001276</a>	Antibiotique Et Anticancéreux
			Asperlactone	<a href="#">BGC0001273</a>	Antioxydant
			Scytalone, T3HN	<a href="#">BGC0002161</a>	
		1.13	Strobilurin A	<a href="#">BGC0001909</a>	
			Fusarubin, 1233A, 1233B, NG-391, Lucilactaene	<a href="#">BGC0002525</a>	Antimicrobien Et Antitumoral
			UNII-YC2Q1O94PT	<a href="#">BGC0001252</a>	
			Solanapyrone A	<a href="#">BGC0002515</a>	
			Monacolin K	<a href="#">BGC0000088</a>	
			Alternapyrone	<a href="#">BGC0000012</a>	Antivirale
			Solanapyrone D	<a href="#">BGC0000146</a>	
			Prolipyrone B, Gibepyrone D	<a href="#">BGC0002191</a>	
			Betaenone A,B,C	<a href="#">BGC0001264</a>	
		1.19	Naringenin	<a href="#">BGC0001310</a>	Antioxydant, Antitumoral, Antiviral, Antibactérien, Anti-Inflammatoire, Antiadipogénique Et Cardioprotecteur
			2,4-Diacetylphloroglucinol	<a href="#">BGC0000281</a>	Antibactérien Et Antifongique
		1.27	Flaviolin, 1,3,6,8-Tetrahydroxynaphthalene	<a href="#">BGC0002127</a>	Anticancéreuses, Antioxydantes, Anti-Inflammatoires Et Antivirales.
			Alkylpyrone-407, 393	<a href="#">BGC0001831</a>	
		1.28	Oronofacic Acid	<a href="#">BGC0000041</a>	
			Angolamycin	<a href="#">BGC0000018</a>	
			Phenylannolone A,B,C	<a href="#">BGC0000122</a>	Anticancéreux
			6-Methylsalicylic Acid	<a href="#">BGC0001276</a>	Antibiotique Et Anticancéreux
				<a href="#">BGC0000068</a>	
			Geldanamycin		
			Asperlactone	<a href="#">BGC0001273</a>	Anti-Inflammatoire, Antibactérien Et Antifongique
			(-)-Mellein	<a href="#">BGC0001244</a>	Anti-Inflammatoire, Antioxydant, Anticancéreux,

					Antimicrobien, Antiviral, Antihépatotoxique
			Altemicidin, SB-203207, SB-203208	<a href="#">BGC0001925</a>	
	1.29		Phenylannolone A,B,C	<a href="#">BGC0000122</a>	Anticancéreux
			Pyranonigrin E	<a href="#">BGC0001124</a>	Antioxydant
			Oronofacic Acid	<a href="#">BGC0000041</a>	
			Nocuolactylates	<a href="#">BGC0002597</a>	Molécules Cytotoxiques
			Asperlactone	<a href="#">BGC0001273</a>	Anti-Inflammatoire
			Amycomicin	<a href="#">BGC0001504</a>	Antibactériens
			Trichoxide	<a href="#">BGC0002233</a>	
			6-Methylsalicyclic Acid	<a href="#">BGC0001275</a>	

Le tableau ci-dessus montre la diversité des capacités biosynthétiques des différentes souches actinomycétales, avec des variations dans le nombre de régions NRPS et PKS, ainsi que dans les types de molécules produites. Cela souligne l'importance de ces souches en tant que sources potentielles de molécules thérapeutiques et ouvre des perspectives intéressantes dans la recherche et le développement de nouveaux médicaments et molécules bioactives.

Sur la base des BGC identifiés, nous avons effectué une prédiction structurale des composés proposés par ces BGC, comme le montre la figure 37.



**Figure 37.** Prédiction des structures moléculaires à activités thérapeutiques

**1/L'Indigoidin** : est un composé organique faisant partie du groupe des azachinones. C'est un pigment bleu produit par certaines espèces de bactéries et sécrété dans leur milieu environnant (Abramovitch, 1974). Ces Propriétés chimiques. La couleur de l'indigoidine est attribuée à un chromophore indigoïde. Les études spectroscopiques RMN sur les dérivés ont confirmé la structure symétrique indiquée. En raison des atomes d'azote dans les cycles quinoïdes , l'indigoidine est classée dans la classe des azaquinones (Takahashi et al., 2007) .

**2/Le Flavioline** : Ce composé appartient à la classe des composés organiques, qui consiste en un cycle benzénique fusionné linéairement à une benzène-1,4-dione (quinone). Activité oxydoréductase, agissant sur donneurs appariés, avec incorporation ou réduction d'oxygène moléculaire. Il catalyse la réaction de couplage CC oxydative pour polymériser la flavioline et former des pigments hautement conjugués qui protègent la bactérie du sol des effets délétères de l'irradiation UV (Zeng et al., 2012).

**3/La nétropsine** : est un polyamide ayant des propriétés antibiotiques et antivirales. Elle a été initialement isolée à partir de *Streptomyces netropsis*, un actinomycète (Finlay et al., 1951). La nétropsine appartient à la classe des pyrrole-amidines et possède une activité contre les bactéries à Gram négatif et à Gram positif ,La nétropsine se lie spécifiquement au petit sillon de l'ADN bicaténaire, en particulier aux séquences riches en paires de bases A-T. (Zimmer & Wähnert, 1986).

**4/Les stambomycines**: sont une famille de macrolides bioactifs isolés de *Streptomyces ambofaciens*. Mis à part deux stéréocentres installés par oxydation du cytochrome P450, leur stéréochimie a été prédite par analyse de séquence de la polykétide synthase. Il existe 3 types de différentes structures stambomycines A,B,C,D (Lim et al., 2021).

**5/L'acide aspergillique** :  $C_{12}H_{20}O_2N_2$  , un antibiotique isolé des filtrats de culture d'*Aspergillus flavus* (Woodbury, 1940) , donne avec un rendement élevé un produit de réaction neutre, l'acide désoxyaspergillique (Dutcher, 1947), il s'agit d'un acide hydroxamique cyclique apparenté à la pyridine (Masaki et al., 1966).

**6/La naringénine** : une flavanone , se trouve essentiellement dans la peau de la tomate et d'autres agrumes comme le raisin. Il possède de nombreuses propriétés pharmacologiques telles que antidiabétique, antidyslipidémique ,antinécroscrotique, antifibrotique et antiobésité . Il possède également des propriétés antioxydantes . Il inhibe l'absorption de glucose basale et stimulée par l'insuline en Colombie-Britannique. Fait important, il a été démontré que la naringénine inhibe la stimulation de la PIP qui sont induites par l'insuline (Singhal et al., 2021). Il régule également à la

baisse le TNF- $\alpha$  et la COX-2. Ainsi, il est déterminé que la naringénine empêche la prolifération des cellules du cancer du sein en altérant l'absorption du glucose (Keating & Martel, 2018).

En comparant les différentes souches bactériennes mentionnées, on observe des similitudes et des différences dans leurs capacités biosynthétiques. La souche *S. ambofaciens* ATCC et la souche *Streptomyces sp.* KY75 présentent toutes deux des régions NRPS et PKS, mais avec des nombres différents. La souche *S. ambofaciens* ATCC se distingue par la présence de deux régions BGC spéciales de type NRPS et trois régions codants à des PKS. Tandis que *Streptomyces sp.* KY75 possède trois régions NRPS et quatre régions PKS. Cela suggère une plus grande diversité de molécules produites par cette souche.

En ce qui concerne les molécules spécifiques, la souche *S. ambofaciens* ATCC produit des molécules avec des structures et des actions diverses grâce à ses régions NRPS spéciales. La Nétropsine, par exemple, est une molécule antibiotique et antivirale produite par les gènes NRPS dans la région 18.

La souche *Streptomyces sp.* KY75, quant à elle, synthétise l'indigoidine à partir de l'une de ses régions NRPS, ainsi que la molécule flavioline à partir de ses régions PKS, qui possède des propriétés antioxydantes.

La souche *Streptomyces leeuwenhoekii* C34 se distingue des autres souches mentionnées en ayant un plus grand nombre de régions de PKS (6) et de NRPS (2). Elle est capable de synthétiser plusieurs désignés, tels que l'acide Aspergillique par les régions NRPS et la Naringénine par les régions PKS.

On observe la présence de certains composés en commun. Deux de ces composés partagés sont l'indigoidine et Flaviolin.

La présence de la flavioline, un composé trouvé dans deux souches bactériennes (*Streptomyces sp.* KY75 et *Streptomyces leeuwenhoekii* C34), pourrait offrir un avantage dans des environnements à stress oxydatif élevé.

La flavioline possède des propriétés antioxydantes, ce qui aide à protéger les cellules bactériennes contre les dommages causés par les espèces réactives de l'oxygène (ROS) associées au stress oxydatif, aux variations de température, aux rayons UV élevés et à la pollution. En plus de son rôle antioxydant, la flavioline a des propriétés antivirales, anticancéreuses et anti-inflammatoires, ce

qui peut renforcer la compétitivité des souches bactériennes dans des environnements où la lutte contre les virus.

L'indigoidine est un composé trouvé dans *Streptomyces ambofaciens* ATCC 23877 et *Streptomyces sp.* KY75, ce qui suggère une certaine similarité biochimique entre elles. L'indigoidine est connue pour ses propriétés antioxydantes et antimicrobiennes, ce qui pourrait indiquer une activité protectrice contre les radicaux libres et une action contre certains agents pathogènes.

La présence de ces composés communs indique une certaine conservation des voies de biosynthèse ou des mécanismes d'acquisition de ces composés dans les trois souches. Cela peut être le résultat d'une évolution convergente de gènes ou de clusters génétiques impliqués dans la synthèse de ces composés vue que les trois souches appartiennent à des environnements avec des conditions extrêmes.

Cependant, il convient de noter que ces deux composés ne représentent qu'une petite fraction des produits naturels par ces souches bactériennes. Chaque souche peut avoir un profil métabolique unique, avec de nombreux autres produits composés qui présentent des caractéristiques spécifiques.

---

---

# CONCLUSION

---

---

En conclusion, cette étude *in silico*, qui porte sur la bio-prospection et l'exploration du potentiel génomique de trois souches actinomycétales habitant des écosystèmes extrêmes différents en Algérie, souligne le rôle clé des *Actinobactéries* du genre *Streptomyces* en tant que sources riches en métabolites secondaires bioactifs. Les résultats de l'analyse phylogénétique et les caractéristiques génomiques observées soutiennent l'idée que les espèces étudiées sont des orthologues ayant évolué à partir d'un ancêtre commun et l'analyse comparative des génomes a révélé une conservation élevée de l'ordre des gènes et une proximité génomique entre les souches étudiées. Cependant, des régions non similaires ont été repérés, ce qui peut entraîner des variations fonctionnelles et métaboliques entre les souches.

Structuralement, la taille du génome varie pour les trois souches *Streptomyces sp.* .KY75, *Streptomyces leewenhoeki* C34 et *Streptomyces ambofacien* ayant des génomes de tailles respectives de 7 837 660, 7 903 895 et 8 303 940 paires de bases. La composition en guanine-cytosine (GC) diffère entre ces souches, avec des pourcentages de 71,6%, 72,8% et 72,2% respectivement. En ce qui concerne les séquences codantes (CDS), *S. sp.* KY75 présente le nombre le plus élevé avec un total de 12 034 CDS, dont 7 694, sont des clusters gènes codant pour des protéines fonctionnelles. Le génome de *S. ambofacien* comporte 7 608 CDS, parmi lesquels 5 518 sont des clusters gènes biosynthétiques codant pour des protéines fonctionnelles. *S. leewenhoeki* C34 présente le nombre le plus bas avec 7 097 CDS, dont 4 917 codent pour des protéines fonctionnelles.

Ces différences dans la taille du génome, la composition en GC et le nombre de CDS indiquent une diversité génomique entre les souches étudiées. Ces variations peuvent refléter des caractéristiques spécifiques de chaque souche.

Il est important de noter que les *Streptomyces* présents dans des écosystèmes pollués et thermaux ont développé des mécanismes d'adaptations spécifiques pour faire face aux contraintes environnementales. Ces adaptations incluent des mécanismes de dégradation des contaminants pour les souches des écosystèmes pollués et des mécanismes de résistance à la chaleur pour les souches des écosystèmes thermaux, cela est prouvée grâce aux résultats d'annotation par le serveur antiSMASH.



L'analyse des génomes des souches bactériennes a révélé la présence de 27 clusters de gènes biosynthétiques (BGC), dont 8 étaient communs entre les trois souches. *Streptomyces sp.* KY75 avait le moins de BGC, soit 32, tandis que *Streptomyces ambofaciens* ATCC 23877 en avait plus, soit 35. *Streptomyces leeuwenhoekii* C34 avait 41 BGC identifiés, c'est remarquablement le nombre le plus élevé. Les BGC identifiés appartenait aux classes principales de métabolites, y compris les PKS, les NRPS.

Ces résultats suggèrent une capacité biosynthétique importante pour la production de métabolites divers tels que l'antimycine, l'ectoïne, la mélanine, l'albaflavenone, la géosmine, la spiramycine et le coelichelin.

L'étude de ces souches a révélé la présence de composés communs, tels que l'indigoidine et la flavioline, qui présentent des propriétés antioxydantes, antivirales, anticancéreuses et anti-inflammatoires. Ces résultats soulignent leurs capacités à produire une large gamme de métabolites bioactifs, offrant ainsi des opportunités prometteuses pour le développement de nouveaux médicaments et la recherche de nouvelles molécules thérapeutiques.

Cette étude démontre l'importance de la détection des BGC dans les *Actinobacteries* en tant que sources pour la découverte de nouveaux composés bioactifs. La recherche sur ces micro-organismes offre des perspectives fascinantes.

Il est également essentiel d'exploiter ces données et de passer à des études *in vitro* et *in vivo* pour tester les capacités et les activités de ces molécules bioactives.

L'évaluation du potentiel métabolique des actinomycètes offre de grandes perspectives de recherche et ouvre des opportunités de production et de commercialisation de nouvelles substances naturelle bioactives à intérêts thérapeutiques, pharmaceutique, biotechnologique et industriel.

---

---

## LES REFERENCES

---

---

- A, P.-M., Fl, S., Jj, F., C, T., S, U., L, F., L, G., R, D., N, K., Ad, K., L, A., Sa, N., Kc, F., J, C., M, T., G, C., G, T., Jg, M., & Aw, T. (2021). Pan-genome analysis identifies intersecting roles for *Pseudomonas* specialized metabolites in potato pathogen inhibition. *eLife*, *10*. <https://doi.org/10.7554/eLife.71900>
- Abdullah Al-Dhabi, N., Ali Esmail, G., Mohammed Ghilan, A.-K., Valan Arasu, M., & Duraipandiyar, V. (2020). Metabolite profiling of *Streptomyces* sp. Al-Dhabi-100 isolated from the marine environment in Saudi Arabia with anti-bacterial, anti-tubercular and anti-oxidant potentials. *Journal of King Saud University - Science*, *32*(2), 1628-1633. <https://doi.org/10.1016/j.jksus.2019.12.021>
- Abe, I. (2012). Novel applications of plant polyketide synthases. *Current Opinion in Chemical Biology*, *16*(1-2), 179-185. <https://doi.org/10.1016/j.cbpa.2011.12.016>
- Abramovitch, R. A. (Éd.). (1974). *Pyridine and Its Derivatives, Volume 14, Part 1 Supplement* (99e édition). Wiley-Interscience.
- Abril, J. F., & Castellano, S. (2019). Genome Annotation. In S. Ranganathan, M. Gribskov, K. Nakai, & C. Schönbach (Éds.), *Encyclopedia of Bioinformatics and Computational Biology* (p. 195-209). Academic Press. <https://doi.org/10.1016/B978-0-12-809633-8.20226-4>
- Agostini-Costa, T., Vieira, R., Bizzo, H., Silveira, D., & Gimenes, M. (2012). *Secondary Metabolites*. <https://doi.org/10.5772/35705>
- Alam, K., Hao, J., Zhong, L., Fan, G., Ouyang, Q., Islam, Md. M., Islam, S., Sun, H., Zhang, Y., Li, R., & Li, A. (2022). Complete genome sequencing and in silico genome mining reveal the promising metabolic potential in *Streptomyces* strain CS-7. *Frontiers in Microbiology*, *13*. <https://www.frontiersin.org/articles/10.3389/fmicb.2022.939919>
- Albarano, L., Esposito, R., Ruocco, N., & Costantini, M. (2020). Genome Mining as New Challenge in Natural Products Discovery. *Marine Drugs*, *18*(4), 199. <https://doi.org/10.3390/md18040199>

- Alikhan, N.-F., Petty, N. K., Ben Zakour, N. L., & Beatson, S. A. (2011). BLAST Ring Image Generator (BRIG) : Simple prokaryote genome comparisons. *BMC Genomics*, 12, 402. <https://doi.org/10.1186/1471-2164-12-402>
- Almpanis, A., Swain, M., Gatherer, D., & McEwan, N. (2018). Correlation between bacterial G+C content, genome size and the G+C content of associated plasmids and bacteriophages. *Microbial Genomics*, 4(4), e000168. <https://doi.org/10.1099/mgen.0.000168>
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403-410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Ambrogelly, A., Palioura, S., & Söll, D. (2007). Natural expansion of the genetic code. *Nature Chemical Biology*, 3(1), Article 1. <https://doi.org/10.1038/nchembio847>
- Anibou, M., Chait, A., Ziad, A., Taourirt, M., Ouhdouch, Y., & Benherref, A. (2008). Actinomycetes from Moroccan habitats : Isolation and screening for cytotoxic activities. *World Journal of Microbiology and Biotechnology*, 24(10), 2019-2025. <https://doi.org/10.1007/s11274-008-9705-7>
- Annaval, T. (2015). *Etude multidisciplinaire des aspects clés de la biosynthèse des polykétides par des polykétide synthases modulaires.*
- Ariel de Lima, D., Helito, C. P., de Lima, L. L., Clazzer, R., Gonçalves, R. K., & de Camargo, O. P. (2022). HOW TO PERFORM A META-ANALYSIS : A PRACTICAL STEP-BY-STEP GUIDE USING R SOFTWARE AND RSTUDIO. *Acta Ortopedica Brasileira*, 30(3), e248775. <https://doi.org/10.1590/1413-785220223003e248775>
- Armstrong, J., Fiddes, I. T., Diekhans, M., & Paten, B. (2019). Whole-Genome Alignment and Comparative Annotation. *Annual Review of Animal Biosciences*, 7(1), 41-64. <https://doi.org/10.1146/annurev-animal-020518-115005>
- Asan, A., İlhan, S., Sen, B., Erkara, I. P., Filik, C., Cabuk, A., Demirel, R., Türe, M., Okten, S. S., & Tokur, S. (2004). Airborne Fungi and Actinomycetes Concentrations in the Air of Eskisehir City (Turkey). *Indoor and Built Environment*, 13(1), 63.

- Austin, M. B., & Noel, J. P. (2003). The chalcone synthase superfamily of type III polyketide synthases. *Natural Product Reports*, 20(1), 79-110. <https://doi.org/10.1039/B100917F>
- Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek, R. A., McNeil, L. K., Paarmann, D., Paczian, T., Parrello, B., ... Zagnitko, O. (2008). The RAST Server : Rapid annotations using subsystems technology. *BMC Genomics*, 9, 75. <https://doi.org/10.1186/1471-2164-9-75>
- B, S. (2003). Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Current Opinion in Chemical Biology*, 7(2). [https://doi.org/10.1016/s1367-5931\(03\)00020-6](https://doi.org/10.1016/s1367-5931(03)00020-6)
- Baldacci, E. (1962). [Current trends in the classification of actinomycetes]. *Annales De La Societe Belge De Medecine Tropicale (1920)*, 42, 633-649.
- Baltz, R. H. (2017). Gifted microbes for genome mining and natural product discovery. *Journal of Industrial Microbiology & Biotechnology*, 44(4-5), 573-588. <https://doi.org/10.1007/s10295-016-1815-x>
- Barka, E. A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Klenk, H.-P., Clément, C., Ouhdouch, Y., & van Wezel, G. P. (2015). Taxonomy, Physiology, and Natural Products of Actinobacteria. *Microbiology and Molecular Biology Reviews : MMBR*, 80(1), 1-43. <https://doi.org/10.1128/MMBR.00019-15>
- Barka, E. A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Meier-Kolthoff, J. P., Klenk, H.-P., Clément, C., Ouhdouch, Y., & van Wezel, G. P. (2016). Correction for Barka et al., Taxonomy, Physiology, and Natural Products of Actinobacteria. *Microbiology and Molecular Biology Reviews: MMBR*, 80(4), iii. <https://doi.org/10.1128/MMBR.00044-16>
- Bauermeister, A., Calil, F. A., das C. L. Pinto, F., Medeiros, T. C. T., Almeida, L. C., Silva, L. J., de Melo, I. S., Zucchi, T. D., Costa-Lotufo, L. V., & Moraes, L. A. B. (2019). Pradimicin-IRD from *Amycolatopsis* sp. IRD-009 and its antimicrobial and cytotoxic activities. *Natural Product Research*, 33(12), 1713-1720. <https://doi.org/10.1080/14786419.2018.1434639>

- Belabed, B. E. (2019). *La pollution par les métaux lourds dans la région d'Annaba : « Sources de contamination des écosystèmes aquatiques »*. Éditions universitaires européennes.
- Benhadj, M., Metrouh, R., Menasria, T., Gacemi-Kirane, D., Slim, F. Z., & Ranque, S. (2020). Broad-spectrum antimicrobial activity of wetland-derived *Streptomyces* sp. ActiF450. *EXCLI Journal*, *19*, 360.  
<https://doi.org/10.17179/excli2020-1124>
- Bentley, S. D., Chater, K. F., Cerdeño-Tárraga, A. M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C. W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., ... Hopwood, D. A. (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature*, *417*(6885), 141-147.  
<https://doi.org/10.1038/417141a>
- Birch, A. J., & Donovan, F. W. (1953). Studies in relation to Biosynthesis. I. Some possible routes to derivatives of Orcinol and Phloroglucinol. *Australian Journal of Chemistry*, *6*(4), 360-368.  
<https://doi.org/10.1071/ch9530360>
- Bj, B., Ga, H., Kl, D., & Da, M. (2015). A prevalent peptide-binding domain guides ribosomal natural product biosynthesis. *Nature Chemical Biology*, *11*(8). <https://doi.org/10.1038/nchembio.1856>
- Blin, K., Shaw, S., Augustijn, H. E., Reitz, Z. L., Biermann, F., Alanjary, M., Fetter, A., Terlouw, B. R., Metcalf, W. W., Helfrich, E. J. N., van Wezel, G. P., Medema, M. H., & Weber, T. (2023). antiSMASH 7.0 : New and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Research*, gkad344. <https://doi.org/10.1093/nar/gkad344>
- Bo, S. T., Xu, Z. F., Yang, L., Cheng, P., Tan, R. X., Jiao, R. H., & Ge, H. M. (2018). Structure and biosynthesis of mayamycin B, a new polyketide with antibacterial activity from *Streptomyces* sp. 120454. *The Journal of Antibiotics*, *71*(6), 601-605. <https://doi.org/10.1038/s41429-018-0039-x>
- Braña, A. F., Sarmiento-Vizcaíno, A., Pérez-Victoria, I., Martín, J., Otero, L., Palacios-Gutiérrez, J. J., Fernández, J., Mohamedi, Y., Fontanil, T., Salmón, M., Cal, S., Reyes, F., García, L. A., & Blanco, G. (2019). Desertomycin G, a New Antibiotic with Activity against *Mycobacterium tuberculosis* and

- Human Breast Tumor Cell Lines Produced by *Streptomyces althoticus* MSM3, Isolated from the Cantabrian Sea Intertidal Macroalgae *Ulva* sp. *Marine Drugs*, *17*(2), 114.  
<https://doi.org/10.3390/md17020114>
- Bright, L. A., Burgess, S. C., Chowdhary, B., Swiderski, C. E., & McCarthy, F. M. (2009). Structural and functional-annotation of an equine whole genome oligoarray. *BMC Bioinformatics*, *10*(S11), S8.  
<https://doi.org/10.1186/1471-2105-10-S11-S8>
- Broadhurst, R. W., Nietlispach, D., Wheatcroft, M. P., Leadlay, P. F., & Weissman, K. J. (2003). The structure of docking domains in modular polyketide synthases. *Chemistry & Biology*, *10*(8), 723-731.  
[https://doi.org/10.1016/s1074-5521\(03\)00156-x](https://doi.org/10.1016/s1074-5521(03)00156-x)
- Brown, A. S., Calcott, M. J., Owen, J. G., & Ackerley, D. F. (2018). Structural, functional and evolutionary perspectives on effective re-engineering of non-ribosomal peptide synthetase assembly lines. *Natural Product Reports*, *35*(11), 1210-1228. <https://doi.org/10.1039/c8np00036k>
- Brown, T. A. (2002). Understanding a Genome Sequence. In *Genomes. 2nd edition*. Wiley-Liss.  
<https://www.ncbi.nlm.nih.gov/books/NBK21136/>
- Bruccoleri, R. E., Dougherty, T. J., & Davison, D. B. (1998). Concordance analysis of microbial genomes. *Nucleic Acids Research*, *26*(19), 4482-4486. <https://doi.org/10.1093/nar/26.19.4482>
- Bryson, K., Loux, V., Bossy, R., Nicolas, P., Chaillou, S., van de Guchte, M., Penaud, S., Maguin, E., Hoebeke, M., Bessières, P., & Gibrat, J.-F. (2006). AGMIAL : Implementing an annotation strategy for prokaryote genomes as a distributed system. *Nucleic Acids Research*, *34*(12), 3533-3545.  
<https://doi.org/10.1093/nar/gkl471>
- Bull, A. T., & Stach, J. E. M. (2007). Marine actinobacteria : New opportunities for natural product search and discovery. *Trends in Microbiology*, *15*(11), 491-499. <https://doi.org/10.1016/j.tim.2007.10.004>
- Busti, E., Monciardini, P., Cavaletti, L., Bamonte, R., Lazzarini, A., Sosio, M., & Donadio, S. (2006). Antibiotic-producing ability by representatives of a newly discovered lineage of actinomycetes. *Microbiology*, *152*(3), 675.

- C, R., T, M., & J, W. (2019). Biosynthesis of Polyketides in Streptomyces. *Microorganisms*, 7(5).  
<https://doi.org/10.3390/microorganisms7050124>
- Caboche, S., Leclère, V., Pupin, M., Kucherov, G., & Jacques, P. (2010). Diversity of Monomers in Nonribosomal Peptides : Towards the Prediction of Origin and Biological Activity. *Journal of Bacteriology*, 192(19), 5143-5150. <https://doi.org/10.1128/JB.00315-10>
- Caboche, S., Pupin, M., Leclère, V., Fontaine, A., Jacques, P., & Kucherov, G. (2008). NORINE : A database of nonribosomal peptides. *Nucleic Acids Research*, 36(Database issue), D326-331.  
<https://doi.org/10.1093/nar/gkm792>
- Chalker, A. F., & Lunsford, R. D. (2002). Rational identification of new antibacterial drug targets that are essential for viability using a genomics-based approach. *Pharmacology & Therapeutics*, 95(1), 1-20.  
[https://doi.org/10.1016/s0163-7258\(02\)00222-x](https://doi.org/10.1016/s0163-7258(02)00222-x)
- Chater, K. (1999). David Hopwood and the emergence of Streptomyces genetics. *International Microbiology: The Official Journal of the Spanish Society for Microbiology*, 2(2), 61-68.
- Chen, M.-H., Lian, Y.-Y., Fang, D.-S., Chen, L., Jia, J., Zhang, W.-L., Lin, R., Xie, Y., Bi, H.-K., & Jiang, H. (2021). Identification and antimicrobial properties of a new alkaloid produced by marine-derived *Verrucosipora* sp. FIM06-0036. *Natural Product Research*, 35(22), 4211-4217.  
<https://doi.org/10.1080/14786419.2019.1689498>
- Cheng, J., Wu, W., Zhang, Y., Li, X., Jiang, X., Wei, G., & Tao, S. (2013). A new computational strategy for predicting essential genes. *BMC Genomics*, 14(1), 910. <https://doi.org/10.1186/1471-2164-14-910>
- Chevalier, M. (2018). *Mise en place d'un workflow d'identification de microorganismes et de leurs métabolites secondaires d'origine non ribosomique par spectrométrie de masse* [These de doctorat, Université de Lille (2018-2021)]. <https://www.theses.fr/2018LILUR049>
- Cho, J. Y., Kwon, H. C., Williams, P. G., Kauffman, C. A., Jensen, P. R., & Fenical, W. (2006). Actinofuranones A and B, Polyketides from a Marine-Derived Bacterium Related to the Genus Streptomyces (Actinomycetales). *Journal of Natural Products*, 69(3), 425-428. <https://doi.org/10.1021/np050402q>

- Choi, B.-K., Lee, H.-S., Kang, J. S., & Shin, H. J. (2019). Dokdolipids A–C, Hydroxylated Rhamnolipids from the Marine-Derived Actinomycete Actinoalloteichus hymeniacidonis. *Marine Drugs*, 17(4), Article 4. <https://doi.org/10.3390/md17040237>
- Choulet, F. (2006). *Evolution du génome des Streptomyces : Transfert horizontal et variabilité des extrémités chromosomiques* (p. 210 p.) [Phdthesis, Université Henri Poincaré - Nancy 1]. <https://hal.univ-lorraine.fr/tel-01754311>
- Clardy, J., & Walsh, C. (2004). Lessons from natural molecules. *Nature*, 432(7019), 829-837. <https://doi.org/10.1038/nature03194>
- Cole, S. T. (2002). Comparative mycobacterial genomics as a tool for drug target and antigen discovery. *European Respiratory Journal*, 20(36 suppl), 78s-86s. <https://doi.org/10.1183/09031936.02.00400202>
- Colquhoun, J. A., Mexson, J., Goodfellow, M., Ward, A. C., Horikoshi, K., & Bull, A. T. (1998). Novel rhodococci and other mycolate actinomycetes from the deep sea. *Antonie Van Leeuwenhoek*, 74(1-3), 27-40. <https://doi.org/10.1023/a:1001743625912>
- Cong, Z., Huang, X., Liu, Y., Liu, Y., Wang, P., Liao, S., Yang, B., Zhou, X., Huang, D., & Wang, J. (2019). Cytotoxic anthracycline and antibacterial tirandamycin analogues from a marine-derived Streptomyces sp. SCSIO 41399. *The Journal of Antibiotics*, 72(1), Article 1. <https://doi.org/10.1038/s41429-018-0103-6>
- Cortes, J., Haydock, S. F., Roberts, G. A., Bevitt, D. J., & Leadlay, P. F. (1990). An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of Saccharopolyspora erythraea. *Nature*, 348(6297), 176-178. <https://doi.org/10.1038/348176a0>
- Cox, R. J. (2007). Polyketides, proteins and genes in fungi : Programmed nano-machines begin to reveal their secrets. *Organic & Biomolecular Chemistry*, 5(13), 2010-2026. <https://doi.org/10.1039/B704420H>
- Craney, A., Ahmed, S., & Nodwell, J. (2013). Towards a new science of secondary metabolism. *The Journal of Antibiotics*, 66(7), 387-400. <https://doi.org/10.1038/ja.2013.25>



- Cullen, L. M., & Arndt, G. M. (2005). Genome-wide screening for gene function using RNAi in mammalian cells. *Immunology and Cell Biology*, 83(3), 217-223. <https://doi.org/10.1111/j.1440-1711.2005.01332.x>
- Davis, J. J., Gerdes, S., Olsen, G. J., Olson, R., Pusch, G. D., Shukla, M., Vonstein, V., Wattam, A. R., & Yoo, H. (2016). PATtyFams : Protein Families for the Microbial Genomes in the PATRIC Database. *Frontiers in Microbiology*, 7. <https://www.frontiersin.org/articles/10.3389/fmicb.2016.00118>
- De Simeis, D., & Serra, S. (2021). Actinomycetes : A Never-Ending Source of Bioactive Compounds—An Overview on Antibiotics Production. *Antibiotics*, 10(5), 483. <https://doi.org/10.3390/antibiotics10050483>
- D'Elia, M. A., Pereira, M. P., & Brown, E. D. (2009). Are essential genes really essential? *Trends in Microbiology*, 17(10), 433-438. <https://doi.org/10.1016/j.tim.2009.08.005>
- Demain, A. L., Gómez-Ortiz, B., Ruiz-Villafán, B., Rodríguez-Sanoja, R., & Sánchez, S. (2019). Recent findings of molecules with anti-infective activity : Screening of non-conventional sources. *Current Opinion in Pharmacology*, 48, 40-47. <https://doi.org/10.1016/j.coph.2019.04.003>
- Demain, A. L., & Sanchez, S. (2009). Microbial drug discovery : 80 years of progress. *The Journal of Antibiotics*, 62(1), 5-16. <https://doi.org/10.1038/ja.2008.16>
- Deng, J., Deng, L., Su, S., Zhang, M., Lin, X., Wei, L., Minai, A. A., Hassett, D. J., & Lu, L. J. (2011). Investigating the predictability of essential genes across distantly related organisms using an integrative approach. *Nucleic Acids Research*, 39(3), 795-807. <https://doi.org/10.1093/nar/gkq784>
- Dickerson, J. E., Zhu, A., Robertson, D. L., & Hentges, K. E. (2011). Defining the role of essential genes in human disease. *PLoS One*, 6(11), e27368. <https://doi.org/10.1371/journal.pone.0027368>
- Djinni, I., Defant, A., Kecha, M., & Mancini, I. (2019). Actinobacteria Derived from Algerian Ecosystems as a Prominent Source of Antimicrobial Molecules. *Antibiotics*, 8(4), 172. <https://doi.org/10.3390/antibiotics8040172>

- Donadio, S., Staver, M. J., McAlpine, J. B., Swanson, S. J., & Katz, L. (1991). Modular organization of genes required for complex polyketide biosynthesis. *Science (New York, N.Y.)*, 252(5006), 675-679.  
<https://doi.org/10.1126/science.2024119>
- Dowell, R. D., Ryan, O., Jansen, A., Cheung, D., Agarwala, S., Danford, T., Bernstein, D. A., Rolfe, P. A., Heisler, L. E., Chin, B., Nislow, C., Giaever, G., Phillips, P. C., Fink, G. R., Gifford, D. K., & Boone, C. (2010). Genotype to phenotype : A complex problem. *Science (New York, N.Y.)*, 328(5977), 469.  
<https://doi.org/10.1126/science.1189015>
- Droste, J., Rückert, C., Kalinowski, J., Hamed, M. B., Anné, J., Simoens, K., Bernaerts, K., Economou, A., & Busche, T. (2021). Extensive Reannotation of the Genome of the Model Streptomyete *Streptomyces lividans* TK24 Based on Transcriptome and Proteome Information. *Frontiers in Microbiology*, 12. <https://www.frontiersin.org/articles/10.3389/fmicb.2021.604034>
- Dunn, B. J., & Khosla, C. (2013). Engineering the acyltransferase substrate specificity of assembly line polyketide synthases. *Journal of the Royal Society, Interface*, 10(85), 20130297.  
<https://doi.org/10.1098/rsif.2013.0297>
- Dutcher, J. D. (1947). ASPERGILLIC ACID : AN ANTIBIOTIC SUBSTANCE PRODUCED BY ASPERGILLUS FLAVUS: I. GENERAL PROPERTIES; FORMATION OF DESOXYASPERGILLIC ACID; STRUCTURAL CONCLUSIONS. *Journal of Biological Chemistry*, 171(1), 321-339. [https://doi.org/10.1016/S0021-9258\(17\)41131-8](https://doi.org/10.1016/S0021-9258(17)41131-8)
- Edgar, R. C. (2004). MUSCLE : Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), 1792-1797. <https://doi.org/10.1093/nar/gkh340>
- Editorial Board. (2018). *Tetrahedron Letters*, 59(4), ii. [https://doi.org/10.1016/S0040-4039\(17\)31613-1](https://doi.org/10.1016/S0040-4039(17)31613-1)
- Edwards, C. (1993). Isolation properties and potential applications of thermophilic actinomycetes. *Applied Biochemistry and Biotechnology*, 42(2-3), 161-179. <https://doi.org/10.1007/BF02788050>
- Ejigu, G. F., & Jung, J. (2020). Review on the Computational Genome Annotation of Sequences Obtained by Next-Generation Sequencing. *Biology*, 9(9), 295. <https://doi.org/10.3390/biology9090295>

- El-Hawary, S. S., Sayed, A. M., Mohammed, R., Khanfar, M. A., Rateb, M. E., Mohammed, T. A., Hajjar, D., Hassan, H. M., Gulder, T. A. M., & Abdelmohsen, U. R. (2018). New Pim-1 Kinase Inhibitor From the Co-culture of Two Sponge-Associated Actinomycetes. *Frontiers in Chemistry*, *6*, 538.  
<https://doi.org/10.3389/fchem.2018.00538>
- Elsayed, S. S., Trusch, F., Deng, H., Raab, A., Prokes, I., Busarakam, K., Asenjo, J. A., Andrews, B. A., van West, P., Bull, A. T., Goodfellow, M., Yi, Y., Ebel, R., Jaspars, M., & Rateb, M. E. (2015). Chaxapeptin, a Lasso Peptide from Extremotolerant *Streptomyces leeuwenhoekii* Strain C58 from the Hyperarid Atacama Desert. *The Journal of Organic Chemistry*, *80*(20), 10252-10260.  
<https://doi.org/10.1021/acs.joc.5b01878>
- English, A. C., Richards, S., Han, Y., Wang, M., Vee, V., Qu, J., Qin, X., Muzny, D. M., Reid, J. G., Worley, K. C., & Gibbs, R. A. (2012). Mind the gap : Upgrading genomes with Pacific Biosciences RS long-read sequencing technology. *PLoS One*, *7*(11), e47768. <https://doi.org/10.1371/journal.pone.0047768>
- Euzeby, J. P. (2003). New taxons of veterinary interest during 2002. *Revue de Médecine Vétérinaire (France)*.  
[https://scholar.google.com/scholar\\_lookup?title=New+taxons+of+veterinary+interest+during+2002&author=Euzeby%2C+J.P.+%28%28Soci%C3%A9t%C3%A9+de+Bact%C3%A9riologie+Syst%C3%A9matisation+et+V%C3%A9t%C3%A9rinaire+%28France%29%29%29&publication\\_year=2003](https://scholar.google.com/scholar_lookup?title=New+taxons+of+veterinary+interest+during+2002&author=Euzeby%2C+J.P.+%28%28Soci%C3%A9t%C3%A9+de+Bact%C3%A9riologie+Syst%C3%A9matisation+et+V%C3%A9t%C3%A9rinaire+%28France%29%29%29&publication_year=2003)
- Fang, G., Bhardwaj, N., Robilotto, R., & Gerstein, M. B. (2010). Getting started in gene orthology and functional analysis. *PLoS Computational Biology*, *6*(3), e1000703.  
<https://doi.org/10.1371/journal.pcbi.1000703>
- Fang, G., Rocha, E., & Danchin, A. (2005). How essential are nonessential genes? *Molecular Biology and Evolution*, *22*(11), 2147-2156. <https://doi.org/10.1093/molbev/msi211>
- Felsenstein, J. (1985). CONFIDENCE LIMITS ON PHYLOGENIES : AN APPROACH USING THE BOOTSTRAP. *Evolution; International Journal of Organic Evolution*, *39*(4), 783-791.  
<https://doi.org/10.1111/j.1558-5646.1985.tb00420.x>

- Fenical, W., & Jensen, P. R. (2006). Developing a new resource for drug discovery : Marine actinomycete bacteria. *Nature Chemical Biology*, 2(12), Article 12. <https://doi.org/10.1038/nchembio841>
- Finlay, A. C., Hochstein, F. A., Sobin, B. A., & Murphy, F. X. (1951). Netropsin, a New Antibiotic Produced by a Streptomyces. *Journal of the American Chemical Society*, 73(1), 341-343. <https://doi.org/10.1021/ja01145a113>
- Flårdh, K. (2010). Cell polarity and the control of apical growth in Streptomyces. *Current Opinion in Microbiology*, 13(6), 758-765. <https://doi.org/10.1016/j.mib.2010.10.002>
- Funa, N., Ohnishi, Y., Fujii, I., Shibuya, M., Ebizuka, Y., & Horinouchi, S. (1999). A new pathway for polyketide synthesis in microorganisms. *Nature*, 400(6747), 897-899. <https://doi.org/10.1038/23748>
- Gallois, A. (2007). *Dynamique des extrémités du chromosome linéaire de Streptomyces ambofaciens* [These de doctorat, Nancy 1]. <https://www.theses.fr/2007NAN10108>
- Gao, M.-Y., Qi, H., Li, J.-S., Zhang, H., Zhang, J., Wang, J.-D., & Xiang, W.-S. (2017). A new naphthalenepropanoic acid analog from the marine-derived actinomycetes *Micromonospora* sp. HS-HM-036. *Journal of Asian Natural Products Research*, 19(9), 930-934. <https://doi.org/10.1080/10286020.2016.1264393>
- Gao, X., Haynes, S. W., Ames, B. D., Wang, P., Vien, L. P., Walsh, C. T., & Tang, Y. (2012). Cyclization of fungal nonribosomal peptides by a terminal condensation-like domain. *Nature Chemical Biology*, 8(10), 823-830. <https://doi.org/10.1038/nchembio.1047>
- Génolevures Consortium, Souciet, J.-L., Dujon, B., Gaillardin, C., Johnston, M., Baret, P. V., Cliften, P., Sherman, D. J., Weissenbach, J., Westhof, E., Wincker, P., Jubin, C., Poulain, J., Barbe, V., Ségurens, B., Artiguenave, F., Anthouard, V., Vacherie, B., Val, M.-E., ... Talla, E. (2009). Comparative genomics of protoploid Saccharomycetaceae. *Genome Research*, 19(10), 1696-1709. <https://doi.org/10.1101/gr.091546.109>

- Gerdes, S., Edwards, R., Kubal, M., Fonstein, M., Stevens, R., & Osterman, A. (2006). Essential genes on metabolic maps. *Current Opinion in Biotechnology*, 17(5), 448-456.  
<https://doi.org/10.1016/j.copbio.2006.08.006>
- Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B., Arkin, A. P., Astromoff, A., El-Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., ... Johnston, M. (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature*, 418(6896), 387-391.  
<https://doi.org/10.1038/nature00935>
- Gibson, D. G., Glass, J. I., Lartigue, C., Noskov, V. N., Chuang, R.-Y., Algire, M. A., Benders, G. A., Montague, M. G., Ma, L., Moodie, M. M., Merryman, C., Vashee, S., Krishnakumar, R., Assad-Garcia, N., Andrews-Pfannkoch, C., Denisova, E. A., Young, L., Qi, Z.-Q., Segall-Shapiro, T. H., ... Venter, J. C. (2010). Creation of a bacterial cell controlled by a chemically synthesized genome. *Science (New York, N.Y.)*, 329(5987), 52-56. <https://doi.org/10.1126/science.1190719>
- Gokhale, R. S., Tsuji, S. Y., Cane, D. E., & Khosla, C. (1999). Dissecting and exploiting intermodular communication in polyketide synthases. *Science (New York, N.Y.)*, 284(5413), 482-485.  
<https://doi.org/10.1126/science.284.5413.482>
- Gomez-Escribano, J. P., Alt, S., & Bibb, M. J. (2016). Next Generation Sequencing of Actinobacteria for the Discovery of Novel Natural Products. *Marine Drugs*, 14(4), 78. <https://doi.org/10.3390/md14040078>
- Goodfellow, M., & O'donnell, A. (1994). *Chemical methods in prokaryotic systematics*.  
<https://www.semanticscholar.org/paper/Chemical-methods-in-prokaryotic-systematics-Goodfellow-O%27donnell/41bfeea81f93e955bbad7eadf8420b4dd84bffb6>
- Goodfellow, M., & Williams, S. T. (1983). ECOLOGY OF ACTINOMYCETES. *Annual Review of Microbiology*, 37(1), 189-216. <https://doi.org/10.1146/annurev.mi.37.100183.001201>

- Grasso, L. L., Martino, D. C., Alduina, R., Grasso, L. L., Martino, D. C., & Alduina, R. (2016). Production of Antibacterial Compounds from Actinomycetes. In *Actinobacteria—Basics and Biotechnological Applications*. IntechOpen. <https://doi.org/10.5772/61525>
- Gui, C., Zhang, S., Zhu, X., Ding, W., Huang, H., Gu, Y.-C., Duan, Y., & Ju, J. (2017). Antimicrobial Spirotetronate Metabolites from Marine-Derived Micromonospora harpali SCSIO GJ089. *Journal of Natural Products*, 80(5), 1594-1603. <https://doi.org/10.1021/acs.jnatprod.7b00176>
- Guzmán-Trampe, S., Ceapa, C. D., Manzo-Ruiz, M., & Sánchez, S. (2017). Synthetic biology era : Improving antibiotic's world. *Biochemical Pharmacology*, 134, 99-113. <https://doi.org/10.1016/j.bcp.2017.01.015>
- Haas, D. (2015). *Métabolisme secondaire de Streptomyces ambofaciens : Exploration génomique et étude du groupe de gènes dirigeant la synthèse du sphydrofurane* [Phdthesis, Université Paris Sud - Paris XI]. <https://theses.hal.science/tel-01505073>
- Hamed, A., Abdel-Razek, A. S., Frese, M., Wibberg, D., El-Haddad, A. F., Ibrahim, T. M. A., Kalinowski, J., Sewald, N., & Shaaban, M. (2018). N-Acetylborrelidin B : A new bioactive metabolite from Streptomyces mutabilis sp. MII. *Zeitschrift Für Naturforschung C*, 73(1-2), 49-57. <https://doi.org/10.1515/znc-2017-0140>
- Hasegawa, T., Takizawa, M., & Tanida, S. (1983). A Rapid Analysis for Chemical Grouping of Aerobic Actinomycetes. *The Journal of General and Applied Microbiology*, 29(4), 319-322. <https://doi.org/10.2323/jgam.29.319>
- Hasin, Y., Seldin, M., & Lusi, A. (2017). Multi-omics approaches to disease. *Genome Biology*, 18(1), 83. <https://doi.org/10.1186/s13059-017-1215-1>
- Hazarika, S. N., & Thakur, D. (2020). Chapter 21—Actinobacteria. In N. Amaresan, M. Senthil Kumar, K. Annapurna, K. Kumar, & A. Sankaranarayanan (Éds.), *Beneficial Microbes in Agro-Ecology* (p. 443-476). Academic Press. <https://doi.org/10.1016/B978-0-12-823414-3.00021-6>

- He, R., Zhang, J., Shao, Y., Gu, S., Song, C., Qian, L., Yin, W.-B., & Li, Z. (2023). Knowledge-guided data mining on the standardized architecture of NRPS : Subtypes, novel motifs, and sequence entanglements. *PLoS Computational Biology*, *19*(5), e1011100. <https://doi.org/10.1371/journal.pcbi.1011100>
- Helfrich, E. J. N., & Piel, J. (2016). Biosynthesis of polyketides by trans-AT polyketide synthases. *Natural Product Reports*, *33*(2), 231-316. <https://doi.org/10.1039/c5np00125k>
- Herbst, D. A., Townsend, C. A., & Maier, T. (2018). The architectures of iterative type I PKS and FAS. *Natural Product Reports*, *35*(10), 1046-1069. <https://doi.org/10.1039/c8np00039e>
- Hertweck, C. (2009). The biosynthetic logic of polyketide diversity. *Angewandte Chemie (International Ed. in English)*, *48*(26), 4688-4716. <https://doi.org/10.1002/anie.200806121>
- Hertweck, C., Luzhetskyy, A., Rebets, Y., & Bechthold, A. (2007). Type II polyketide synthases : Gaining a deeper insight into enzymatic teamwork. *Natural Product Reports*, *24*(1), 162-190. <https://doi.org/10.1039/B507395M>
- Heuer, H., Krsek, M., Baker, P., Smalla, K., & Wellington, E. M. (1997). Analysis of actinomycete communities by specific amplification of genes encoding 16S rRNA and gel-electrophoretic separation in denaturing gradients. *Applied and Environmental Microbiology*, *63*(8), 3233-3241.
- Holman, A. G., Davis, P. J., Foster, J. M., Carlow, C. K. S., & Kumar, S. (2009). Computational prediction of essential genes in an unculturable endosymbiotic bacterium, Wolbachia of *Brugia malayi*. *BMC Microbiology*, *9*, 243. <https://doi.org/10.1186/1471-2180-9-243>
- Hopwood, D. A. (2003). Streptomyces genes : From Waksman to Sanger. *Journal of Industrial Microbiology and Biotechnology*, *30*(8), 468-471. <https://doi.org/10.1007/s10295-003-0031-7>
- Hopwood, D. A., & Wright, H. M. (1973). A Plasmid of *Streptomyces coelicolor* Carrying a Chromosomal Locus and its Inter-specific Transfer. *Microbiology*, *79*(2), 331-342. <https://doi.org/10.1099/00221287-79-2-331>
- Hoshino, S., Wong, C. P., Ozeki, M., Zhang, H., Hayashi, F., Awakawa, T., Asamizu, S., Onaka, H., & Abe, I. (2018). Umezawamides, new bioactive polycyclic tetramate macrolactams isolated from a

- combined-culture of *Umezawaea* sp. And mycolic acid-containing bacterium. *The Journal of Antibiotics*, 71(7), 653-657. <https://doi.org/10.1038/s41429-018-0040-4>
- Hsu, C.-H., Tomiyasu, H., Liao, C.-H., & Lin, C.-S. (2021). Genome-wide DNA methylation and RNA-seq analyses identify genes and pathways associated with doxorubicin resistance in a canine diffuse large B-cell lymphoma cell line. *PLOS ONE*, 16(5), e0250013. <https://doi.org/10.1371/journal.pone.0250013>
- Hullo, M.-F., Moszer, I., Danchin, A., & Martin-Verstraete, I. (2001). CotA of *Bacillus subtilis* Is a Copper-Dependent Laccase. *Journal of Bacteriology*, 183(18), 5426-5430. <https://doi.org/10.1128/jb.183.18.5426-5430.2001>
- Hutchison, C. A., Chuang, R.-Y., Noskov, V. N., Assad-Garcia, N., Deerinck, T. J., Ellisman, M. H., Gill, J., Kannan, K., Karas, B. J., Ma, L., Pelletier, J. F., Qi, Z.-Q., Richter, R. A., Strychalski, E. A., Sun, L., Suzuki, Y., Tsvetanova, B., Wise, K. S., Smith, H. O., ... Venter, J. C. (2016). Design and synthesis of a minimal bacterial genome. *Science (New York, N.Y.)*, 351(6280), aad6253. <https://doi.org/10.1126/science.aad6253>
- Ibrahimi, M. (2020). *Extraction et caractérisation de nouveaux antibactériens produits par les actinobactéries prédatrices d'origine marine* [Phdthesis, Université de Poitiers ; Université Cadi Ayyad, Marrakech, Maroc]. <https://theses.hal.science/tel-03340984>
- Imada, A., & Hotta, K. (1992). Historical Perspectives of Approaches to Antibiotics Discovery. In J. A. Sutcliffe & N. H. Georgopapadakou (Éds.), *Emerging Targets in Antibacterial and Antifungal Chemotherapy* (p. 1-23). Springer US. [https://doi.org/10.1007/978-1-4615-3274-3\\_1](https://doi.org/10.1007/978-1-4615-3274-3_1)
- Itaya, M. (1995). An estimation of minimal genome size required for life. *FEBS Letters*, 362(3), 257-260. [https://doi.org/10.1016/0014-5793\(95\)00233-y](https://doi.org/10.1016/0014-5793(95)00233-y)
- J, A., It, F., M, D., & B, P. (2019). Whole-Genome Alignment and Comparative Annotation. *Annual Review of Animal Biosciences*, 7. <https://doi.org/10.1146/annurev-animal-020518-115005>



- Jarvis, P. D., Holland, B. R., & Sumner, J. G. (2017). Phylogenetic Invariants and Markov Invariants☆. In *Reference Module in Life Sciences*. Elsevier. <https://doi.org/10.1016/B978-0-12-809633-8.06903-X>
- Jensen, P. R., Williams, P. G., Oh, D.-C., Zeigler, L., & Fenical, W. (2007). Species-specific secondary metabolite production in marine actinomycetes of the genus *Salinispora*. *Applied and Environmental Microbiology*, 73(4), 1146-1152. <https://doi.org/10.1128/AEM.01891-06>
- Jose, P., Maharshi, A., & Jha, B. (2021). Actinobacteria in natural products research : Progress and prospects. *Microbiological Research*, 246, 126708. <https://doi.org/10.1016/j.micres.2021.126708>
- Karthik, L. (2022). *Actinobacteria : Microbiology to Synthetic Biology*. Springer Nature.
- Karupiah, P., & Mustaffa, M. (2013). Antibacterial and antioxidant activities of *Musa* sp. Leaf extracts against multidrug resistant clinical pathogens causing nosocomial infection. *Asian Pacific Journal of Tropical Biomedicine*, 3(9), 737-742. [https://doi.org/10.1016/S2221-1691\(13\)60148-3](https://doi.org/10.1016/S2221-1691(13)60148-3)
- Kautsar, S. A., Blin, K., Shaw, S., Navarro-Muñoz, J. C., Terlouw, B. R., van der Hooft, J. J. J., van Santen, J. A., Tracanna, V., Suarez Duran, H. G., Pascal Andreu, V., Selem-Mojica, N., Alanjary, M., Robinson, S. L., Lund, G., Epstein, S. C., Sisto, A. C., Charkoudian, L. K., Collemare, J., Linington, R. G., ... Medema, M. H. (2020). MIBiG 2.0 : A repository for biosynthetic gene clusters of known function. *Nucleic Acids Research*, 48(D1), D454-D458. <https://doi.org/10.1093/nar/gkz882>
- Kautsar, S. A., Blin, K., Shaw, S., Weber, T., & Medema, M. H. (2021). BiG-FAM : The biosynthetic gene cluster families database. *Nucleic Acids Research*, 49(D1), D490-D497. <https://doi.org/10.1093/nar/gkaa812>
- Kearse, M. G., & Wilusz, J. E. (2017). Non-AUG translation : A new start for protein synthesis in eukaryotes. *Genes & Development*, 31(17), 1717-1731. <https://doi.org/10.1101/gad.305250.117>
- Keating, E., & Martel, F. (2018). Antimetabolic Effects of Polyphenols in Breast Cancer Cells : Focus on Glucose Uptake and Metabolism. *Frontiers in Nutrition*, 5, 25. <https://doi.org/10.3389/fnut.2018.00025>

- Keller, N. P. (2015). Translating biosynthetic gene clusters into fungal armor and weaponry. *Nature Chemical Biology*, 11(9), 671-677. <https://doi.org/10.1038/nchembio.1897>
- Kent, W. J. (2002). BLAT--the BLAST-like alignment tool. *Genome Research*, 12(4), 656-664. <https://doi.org/10.1101/gr.229202>
- Khan, A., Singh, P., & Srivastava, A. (2018). Synthesis, nature and utility of universal iron chelator – Siderophore : A review. *Microbiological Research*, 212-213, 103-111. <https://doi.org/10.1016/j.micres.2017.10.012>
- Khushboo, null, Singhvi, N., Gupta, V., Dhaka, N., & Dubey, K. K. (2023). Draft genome sequence of *Streptomyces* sp. KD18, isolated from industrial soil. *3 Biotech*, 13(1), 34. <https://doi.org/10.1007/s13205-022-03453-3>
- Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT : A fast spliced aligner with low memory requirements. *Nature Methods*, 12(4), 357-360. <https://doi.org/10.1038/nmeth.3317>
- Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., & Salzberg, S. L. (2013). TopHat2 : Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology*, 14(4), Article 4. <https://doi.org/10.1186/gb-2013-14-4-r36>
- Kim, D.-U., Hayles, J., Kim, D., Wood, V., Park, H.-O., Won, M., Yoo, H.-S., Duhig, T., Nam, M., Palmer, G., Han, S., Jeffery, L., Baek, S.-T., Lee, H., Shim, Y. S., Lee, M., Kim, L., Heo, K.-S., Noh, E. J., ... Hoe, K.-L. (2010). Analysis of a genome-wide set of gene deletions in the fission yeast *Schizosaccharomyces pombe*. *Nature Biotechnology*, 28(6), 617-623. <https://doi.org/10.1038/nbt.1628>
- Kobayashi, K., Ehrlich, S. D., Albertini, A., Amati, G., Andersen, K. K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., Bessieres, P., Boland, F., Brignell, S. C., Bron, S., Bunai, K., Chapuis, J., Christiansen, L. C., Danchin, A., Débarbouille, M., Dervyn, E., ... Ogasawara, N. (2003). Essential *Bacillus subtilis* genes. *Proceedings of the National Academy of Sciences of the United States of America*, 100(8), 4678-4683. <https://doi.org/10.1073/pnas.0730515100>

- Komaki, H., & Tamura, T. (2020). Polyketide Synthase and Nonribosomal Peptide Synthetase Gene Clusters in Type Strains of the Genus *Phytohabitans*. *Life*, *10*(11), 257. <https://doi.org/10.3390/life10110257>
- Koonin, E. V. (2005). Orthologs, paralogs, and evolutionary genomics. *Annual Review of Genetics*, *39*, 309-338. <https://doi.org/10.1146/annurev.genet.39.073003.114725>
- Korn-Wendisch, F., & Schneider, J. (1992). Phage typing—A useful tool in actinomycete systematics. *Gene*, *115*(1-2), 243-247. [https://doi.org/10.1016/0378-1119\(92\)90565-7](https://doi.org/10.1016/0378-1119(92)90565-7)
- Kristensen, D. M., Wolf, Y. I., Mushegian, A. R., & Koonin, E. V. (2011). Computational methods for Gene Orthology inference. *Briefings in Bioinformatics*, *12*(5), 379-391. <https://doi.org/10.1093/bib/bbr030>
- Kute, P. M., Soukarieh, O., Tjeldnes, H., Trégouët, D.-A., & Valen, E. (2022). Small Open Reading Frames, How to Find Them and Determine Their Function. *Frontiers in Genetics*, *12*. <https://www.frontiersin.org/articles/10.3389/fgene.2021.796060>
- L, S., & T, S. (2015). Thermophilic and alkaliphilic Actinobacteria : Biology and potential applications. *Frontiers in Microbiology*, *6*. <https://doi.org/10.3389/fmicb.2015.01014>
- Ladoukakis, E., Pereira, V., Magny, E. G., Eyre-Walker, A., & Couso, J. P. (2011). Hundreds of putatively functional small open reading frames in *Drosophila*. *Genome Biology*, *12*(11), R118. <https://doi.org/10.1186/gb-2011-12-11-r118>
- Lalitha, P., Veena, V., Vidhyapriya, P., Lakshmi, P., Krishna, R., & Sakthivel, N. (2016). Anticancer potential of pyrrole (1, 2, a) pyrazine 1, 4, dione, hexahydro 3-(2-methyl propyl) (PPDHMP) extracted from a new marine bacterium, *Staphylococcus* sp. Strain MB30. *Apoptosis: An International Journal on Programmed Cell Death*, *21*(5), 566-577. <https://doi.org/10.1007/s10495-016-1221-x>
- Lam, K. S. (2006). Discovery of novel metabolites from marine actinomycetes. *Current Opinion in Microbiology*, *9*(3), 245-251. <https://doi.org/10.1016/j.mib.2006.03.004>
- Lapierre, P., & Gogarten, J. P. (2009). Estimating the size of the bacterial pan-genome. *Trends in Genetics: TIG*, *25*(3), 107-110. <https://doi.org/10.1016/j.tig.2008.12.004>

- Le Minor Léon. (1989). *Bactériologie médicale* / [publié par] Léon Le Minor, Michel Véron (2e édition).  
Médecine-Sciences Flammarion.
- Lee, N., Hwang, S., Kim, J., Cho, S., Palsson, B., & Cho, B.-K. (2020). Mini review : Genome mining approaches for the identification of secondary metabolite biosynthetic gene clusters in Streptomyces. *Computational and Structural Biotechnology Journal*, 18, 1548-1556.  
<https://doi.org/10.1016/j.csbj.2020.06.024>
- Leong, A. Z.-X., Lee, P. Y., Mohtar, M. A., Syafruddin, S. E., Pung, Y.-F., & Low, T. Y. (2022). Short open reading frames (sORFs) and microproteins : An update on their identification and validation measures. *Journal of Biomedical Science*, 29(1), Article 1. <https://doi.org/10.1186/s12929-022-00802-5>
- Lewin, G. R., Carlos, C., Chevrette, M. G., Horn, H. A., McDonald, B. R., Stankey, R. J., Fox, B. G., & Currie, C. R. (2016). Evolution and Ecology of Actinobacteria and Their Bioenergy Applications. *Annual Review of Microbiology*, 70, 235-254. <https://doi.org/10.1146/annurev-micro-102215-095748>
- Lim, J., Chintalapudi, V., Gudmundsson, H. G., Tran, M., Bernasconi, A., Blanco, A., Song, L., Challis, G. L., & Anderson, E. A. (2021). Synthesis of the C1–C27 Fragment of Stambomycin D Validates Modular Polyketide Synthase-Based Stereochemical Assignments. *Organic Letters*, 23(19), 7439-7444.  
<https://doi.org/10.1021/acs.orglett.1c02650>
- Lorenzi, J.-N. (2020). *Dynamique des génomes du genre Streptomyces* [Phdthesis, Université Paris-Saclay].  
<https://theses.hal.science/tel-02936363>
- Luciano, B. (2013). The GC Content of Bacterial Genomes. *Journal of Phylogenetics & Evolutionary Biology*.  
<https://doi.org/10.4172/2329-9002.1000e108>
- Lv, Q., Fan, Y., Tao, G., Fu, P., Zhai, J., Ye, B., & Zhu, W. (2019). Sekgranaticin, a SEK34b-Granaticin Hybrid Polyketide from Streptomyces sp. 166#. *The Journal of Organic Chemistry*, 84(14), 9087-9092.  
<https://doi.org/10.1021/acs.joc.9b01022>

- Machado, H., Sonnenschein, E. C., Melchiorson, J., & Gram, L. (2015). Genome mining reveals unlocked bioactive potential of marine Gram-negative bacteria. *BMC Genomics*, *16*(1), 158. <https://doi.org/10.1186/s12864-015-1365-z>
- MacLean, D., Jones, J., & Studholme, D. (2009). MacLean D, Jones JDG, Studholme DJ.. Application of 'next-generation' sequencing technologies to microbial genetics. *Nat Rev Microbiol* 7 : 287-296. *Nature reviews. Microbiology*, *7*, 287-296. <https://doi.org/10.1038/nrmicro2122>
- Malpartida, F., & Hopwood, D. A. (1984). Molecular cloning of the whole biosynthetic pathway of a Streptomyces antibiotic and its expression in a heterologous host. *Nature*, *309*(5967), Article 5967. <https://doi.org/10.1038/309462a0>
- Marcia Soares Almeida, H. (2022). *Learning to discover biosynthetic gene clusters in fungi* [Thèse ou essai doctoral accepté, Université du Québec à Montréal]. <https://archipel.uqam.ca/15707/>
- Mardis, E. R. (2013). Next-generation sequencing platforms. *Annual Review of Analytical Chemistry (Palo Alto, Calif.)*, *6*, 287-303. <https://doi.org/10.1146/annurev-anchem-062012-092628>
- Mariat, F., & Sebald, M. (1990). Actinomycètes In : Bactériologie Médicale. *Flammarion (2nd Edn)*, editor, 935-949.
- Martín, J. F., Liras, P., & Sánchez, S. (2021). Modulation of Gene Expression in Actinobacteria by Translational Modification of Transcriptional Factors and Secondary Metabolite Biosynthetic Enzymes. *Frontiers in Microbiology*, *12*. <https://www.frontiersin.org/articles/10.3389/fmicb.2021.630694>
- Martínez-Núñez, M. A., & López, V. E. L. y. (2016). Nonribosomal peptides synthetases and their applications in industry. *Sustainable Chemical Processes*, *4*(1), 13. <https://doi.org/10.1186/s40508-016-0057-6>
- Masaki, M., Chigira, Y., & Ohta, M. (1966). Total Syntheses of Racemic Aspergillic Acid and Neoaspergillic Acid. *The Journal of Organic Chemistry*, *31*(12), 4143-4146. <https://doi.org/10.1021/jo01350a062>

- McIntyre, J. J., Bunch, A. W., & Bull, A. T. (1999). Vancomycin production is enhanced in chemostat culture with biomass-recycle. *Biotechnology and Bioengineering*, 62(5), 576-582.  
[https://doi.org/10.1002/\(sici\)1097-0290\(19990305\)62:5<576::aid-bit9>3.0.co;2-p](https://doi.org/10.1002/(sici)1097-0290(19990305)62:5<576::aid-bit9>3.0.co;2-p)
- McNeil, M. M., & Brown, J. M. (1994). The medically important aerobic actinomycetes : Epidemiology and microbiology. *Clinical Microbiology Reviews*, 7(3), 357-417.
- Medema, M. H., Cimermancic, P., Sali, A., Takano, E., & Fischbach, M. A. (2014). A systematic computational analysis of biosynthetic gene cluster evolution : Lessons for engineering biosynthesis. *PLoS Computational Biology*, 10(12), e1004016. <https://doi.org/10.1371/journal.pcbi.1004016>
- Medema, M. H., & Fischbach, M. A. (2015). Computational approaches to natural product discovery. *Nature chemical biology*, 11(9), 639-648. <https://doi.org/10.1038/nchembio.1884>
- Medema, M., Kottmann, R., Yilmaz, P., Cummings, M., Biggins, J., Blin, K., I, de, Chooi, Y.-H., Claesen, J., Coates, R. C., Cruz-Morales, P., Duddela, S., Düsterhus, S., Edwards, D., Fewer, D., Garg, N., Geiger, C., JP, G.-E., Greule, A., & Glöckner, F. (2015). Minimum Information about a Biosynthetic Gene cluster. *Nature Chemical Biology*, 11, 625-631. <https://doi.org/10.1038/nchembio.1890>
- Medjemadj, M., Escuder-Rodríguez, J.-J., Boudemagh, A., & González-Siso, M.-I. (2020). *Actinobacteria isolated from Algerian hot spring waters : A potential source of important enzymes*. 26, 1145-1157.
- Messaoudi, O., Bendahou, M., Benamar, I., & Abdelwouhid, D.-E. (2015). Identification and preliminary characterization of non-polyene antibiotics secreted by new strain of actinomycete isolated from sebkha of Kenadsa, Algeria. *Asian Pacific Journal of Tropical Biomedicine*, 5(6), 438-445.  
<https://doi.org/10.1016/j.apjtb.2015.04.002>
- Metz, J. G., Roessler, P., Facciotti, D., Levering, C., Dittrich, F., Lassner, M., Valentine, R., Lardizabal, K., Domergue, F., Yamada, A., Yazawa, K., Knauf, V., & Browse, J. (2001). Production of polyunsaturated fatty acids by polyketide synthases in both prokaryotes and eukaryotes. *Science (New York, N.Y.)*, 293(5528), 290-293. <https://doi.org/10.1126/science.1059593>

- Miguélez, E. M., Hardisson, C., & Manzanal, M. B. (1999). Hyphal Death during Colony Development in *Streptomyces antibioticus* : Morphological Evidence for the Existence of a Process of Cell Deletion in a Multicellular Prokaryote. *Journal of Cell Biology*, *145*(3), 515-525.  
<https://doi.org/10.1083/jcb.145.3.515>
- Mitra, S., Dhar, R., & Sen, R. (2022). Designer bacterial cell factories for improved production of commercially valuable non-ribosomal peptides. *Biotechnology Advances*, *60*, 108023.  
<https://doi.org/10.1016/j.biotechadv.2022.108023>
- Moeller, R., Horneck, G., Facius, R., & Stackebrandt, E. (2005). Role of pigmentation in protecting *Bacillus* sp. Endospores against environmental UV radiation. *FEMS Microbiology Ecology*, *51*(2), 231-236.  
<https://doi.org/10.1016/j.femsec.2004.08.008>
- Monciardini, P., Sosio, M., Cavaletti, L., Chiocchini, C., & Donadio, S. (2002). New PCR primers for the selective amplification of 16S rDNA from different groups of actinomycetes1. *FEMS Microbiology Ecology*, *42*(3), 419-429. <https://doi.org/10.1111/j.1574-6941.2002.tb01031.x>
- Motamedi, H., & Hutchinson, C. R. (1987). Cloning and heterologous expression of a gene cluster for the biosynthesis of tetracenomycin C, the anthracycline antitumor antibiotic of *Streptomyces glaucescens*. *Proceedings of the National Academy of Sciences of the United States of America*, *84*(13), 4445-4449. <https://doi.org/10.1073/pnas.84.13.4445>
- Mukherjee, S., Das, P., Sivapathasekaran, C., & Sen, R. (2009). Antimicrobial biosurfactants from marine *Bacillus circulans* : Extracellular synthesis and purification. *Letters in Applied Microbiology*, *48*(3), 281-288. <https://doi.org/10.1111/j.1472-765X.2008.02485.x>
- Murti, Y., & Agrawal, T. (2010). *Marine derived pharmaceuticals- Development of natural health products from marine biodiversity*.
- Myers, E. W., Sutton, G. G., Delcher, A. L., Dew, I. M., Fasulo, D. P., Flanigan, M. J., Kravitz, S. A., Mobarry, C. M., Reinert, K. H., Remington, K. A., Anson, E. L., Bolanos, R. A., Chou, H. H., Jordan, C. M., Halpern, A. L., Lonardi, S., Beasley, E. M., Brandon, R. C., Chen, L., ... Venter, J. C. (2000). A whole-genome

- assembly of *Drosophila*. *Science (New York, N.Y.)*, 287(5461), 2196-2204.  
<https://doi.org/10.1126/science.287.5461.2196>
- Nair, S., & Abraham, J. (2020). *Natural Products from Actinobacteria for Drug Discovery* (p. 333-363).  
[https://doi.org/10.1007/978-981-15-2195-9\\_23](https://doi.org/10.1007/978-981-15-2195-9_23)
- Nguyen, N., Jennen, D., & Kleinjans, J. (2022). Omics technologies to understand drug toxicity mechanisms. *Drug Discovery Today*, 27(11), 103348. <https://doi.org/10.1016/j.drudis.2022.103348>
- Nodwell, J. R. (2017). Antimicrobials : Expressing antibiotic gene clusters. *Nature Microbiology*, 2, 17061.  
<https://doi.org/10.1038/nmicrobiol.2017.61>
- O'Brien, J., & Wright, G. D. (2011). An ecological perspective of microbial secondary metabolism. *Current Opinion in Biotechnology*, 22(4), 552-558. <https://doi.org/10.1016/j.copbio.2011.03.010>
- Olexiouk, V., Van Criekinge, W., & Menschaert, G. (2018). An update on sORFs.org : A repository of small ORFs identified by ribosome profiling. *Nucleic Acids Research*, 46(D1), D497-D502.  
<https://doi.org/10.1093/nar/gkx1130>
- Ondov, B. D., Treangen, T. J., Melsted, P., Mallonee, A. B., Bergman, N. H., Koren, S., & Phillippy, A. M. (2016). Mash : Fast genome and metagenome distance estimation using MinHash. *Genome Biology*, 17(1), 132. <https://doi.org/10.1186/s13059-016-0997-x>
- Orr, M. W., Mao, Y., Storz, G., & Qian, S.-B. (2020). Alternative ORFs and small ORFs : Shedding light on the dark proteome. *Nucleic Acids Research*, 48(3), 1029-1042. <https://doi.org/10.1093/nar/gkz734>
- Orsi, W. D., Smith, J. M., Liu, S., Liu, Z., Sakamoto, C. M., Wilken, S., Poirier, C., Richards, T. A., Keeling, P. J., Worden, A. Z., & Santoro, A. E. (2016). Diverse, uncultivated bacteria and archaea underlying the cycling of dissolved protein in the ocean. *The ISME Journal*, 10(9), 2158-2173.  
<https://doi.org/10.1038/ismej.2016.20>
- Overbeek, R., Fonstein, M., D'Souza, M., Pusch, G. D., & Maltsev, N. (1999). The use of gene clusters to infer functional coupling. *Proceedings of the National Academy of Sciences of the United States of America*, 96(6), 2896-2901. <https://doi.org/10.1073/pnas.96.6.2896>



- Palaniappan, K., Chen, I.-M. A., Chu, K., Ratner, A., Seshadri, R., Kyrpides, N. C., Ivanova, N. N., & Mouncey, N. J. (2020). IMG-ABC v.5.0 : An update to the IMG/Atlas of Biosynthetic Gene Clusters Knowledgebase. *Nucleic Acids Research*, *48*(D1), D422-D430. <https://doi.org/10.1093/nar/gkz932>
- Papp, B., Pál, C., & Hurst, L. D. (2004). Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast. *Nature*, *429*(6992), 661-664. <https://doi.org/10.1038/nature02636>
- Park, H. B., Perez, C. E., Barber, K. W., Rinehart, J., & Crawford, J. M. (2017). Genome mining unearths a hybrid nonribosomal peptide synthetase-like-pteridine synthase biosynthetic gene cluster. *ELife*, *6*, e25229. <https://doi.org/10.7554/eLife.25229>
- Park, H.-S., Kim, H.-J., Han, C.-Y., Nah, H.-J., Choi, S.-S., & Kim, E.-S. (2020). Stimulated Biosynthesis of an C10-Deoxy Heptaene NPP B2 via Regulatory Genes Overexpression in *Pseudonocardia autotrophica*. *Frontiers in Microbiology*, *11*, 19. <https://doi.org/10.3389/fmicb.2020.00019>
- Patel, J. B., Wallace, R. J., Brown-Elliott, B. A., Taylor, T., Imperatrice, C., Leonard, D. G. B., Wilson, R. W., Mann, L., Jost, K. C., & Nachamkin, I. (2004). Sequence-Based Identification of Aerobic Actinomycetes. *Journal of Clinical Microbiology*, *42*(6), 2530-2540. <https://doi.org/10.1128/JCM.42.6.2530-2540.2004>
- Patra, J. K., Shukla, A. C., & Das, G. (2020). *Advances in Pharmaceutical Biotechnology : Recent Progress and Future Applications*. Springer Nature.
- Peeters, M. K. R., & Menschaert, G. (2020). The hunt for sORFs : A multidisciplinary strategy. *Experimental Cell Research*, *391*(1), 111923. <https://doi.org/10.1016/j.yexcr.2020.111923>
- Pérez-Bonilla, M., Oves-Costales, D., de la Cruz, M., Kokkini, M., Martín, J., Vicente, F., Genilloud, O., & Reyes, F. (2018). Phocoenamicins B and C, New Antibacterial Spirotetronates Isolated from a Marine *Micromonospora* sp. *Marine Drugs*, *16*(3), 95. <https://doi.org/10.3390/md16030095>
- Pettit, R. K. (2011). Small-molecule elicitation of microbial secondary metabolites. *Microbial biotechnology*, *4*(4), 471-478. <https://doi.org/10.1111/j.1751-7915.2010.00196.x>

- Rai, R. V., & Bai, J. A. (2022). *Natural Products from Actinomycetes : Diversity, Ecology and Drug Discovery*. Springer Nature.
- Ranjani, A., Dharumadurai, D., & P M, G. (2016). *An Introduction to Actinobacteria* (p. 3-37).  
<https://doi.org/10.5772/62329>
- Rasko, D. A., Rosovitz, M. J., Myers, G. S. A., Mongodin, E. F., Fricke, W. F., Gajer, P., Crabtree, J., Sebaihia, M., Thomson, N. R., Chaudhuri, R., Henderson, I. R., Sperandio, V., & Ravel, J. (2008). The pangenome structure of *Escherichia coli* : Comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *Journal of Bacteriology*, *190*(20), 6881-6893. <https://doi.org/10.1128/JB.00619-08>
- Rd, K., Yi, Y., Y, X., P, C., Sj, N., W, F., Ma, F., Bs, M., & Pc, D. (2011). A mass spectrometry-guided genome mining approach for natural product peptidogenomics. *Nature Chemical Biology*, *7*(11).  
<https://doi.org/10.1038/nchembio.684>
- Reponen, T. A., Gizenko, S. V., Grinshpun, S. A., Willeke, K., & Cole, E. C. (1998). Characteristics of Airborne Actinomycete Spores. *Applied and Environmental Microbiology*, *64*(10), 3807-3812.
- Reshetnikov, A. S., Khmelenina, V. N., Mustakhimov, I. I., & Trotsenko, Y. A. (2011). Chapter Two—Genes and Enzymes of Ectoine Biosynthesis in Halotolerant Methanotrophs. In A. C. Rosenzweig & S. W. Ragsdale (Éds.), *Methods in Enzymology* (Vol. 495, p. 15-30). Academic Press.  
<https://doi.org/10.1016/B978-0-12-386905-0.00002-4>
- Richardson, E. J., & Watson, M. (2013). The automatic annotation of bacterial genomes. *Briefings in Bioinformatics*, *14*(1), 1-12. <https://doi.org/10.1093/bib/bbs007>
- Richter, M., Rosselló-Móra, R., Oliver Glöckner, F., & Peplies, J. (2016). JSpeciesWS : A web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics (Oxford, England)*, *32*(6), 929-931. <https://doi.org/10.1093/bioinformatics/btv681>

- Rifaat, H. M. (2003). *The biodiversity of Actinomycetes in the River Nile exhibiting antifungal activity*.  
<https://www.semanticscholar.org/paper/The-biodiversity-of-Actinomycetes-in-the-River-Nile-Rifaat/ef018bbb9bb5ebeb7866fe0fd3561f4075ba4cbd>
- Rix, U., Fischer, C., Remsing, L. L., & Rohr, J. (2002). Modification of post-PKS tailoring steps through combinatorial biosynthesis. *Natural Product Reports*, 19(5), 542-580.  
<https://doi.org/10.1039/B103920M>
- Robbins, T., Liu, Y.-C., Cane, D. E., & Khosla, C. (2016). Structure and mechanism of assembly line polyketide synthases. *Current Opinion in Structural Biology*, 41, 10-18.  
<https://doi.org/10.1016/j.sbi.2016.05.009>
- Rocha, E. P. C. (2006). Inference and Analysis of the Relative Stability of Bacterial Chromosomes. *Molecular Biology and Evolution*, 23(3), 513-522. <https://doi.org/10.1093/molbev/msj052>
- Roemer, T., Jiang, B., Davison, J., Ketela, T., Veillette, K., Breton, A., Tandia, F., Linteau, A., Sillaots, S., Marta, C., Martel, N., Veronneau, S., Lemieux, S., Kauffman, S., Becker, J., Storms, R., Boone, C., & Bussey, H. (2003). Large-scale essential gene identification in *Candida albicans* and applications to antifungal drug discovery. *Molecular Microbiology*, 50(1), 167-181. <https://doi.org/10.1046/j.1365-2958.2003.03697.x>
- Roongsawang, N., Washio, K., & Morikawa, M. (2011). Diversity of Nonribosomal Peptide Synthetases Involved in the Biosynthesis of Lipopeptide Biosurfactants. *International Journal of Molecular Sciences*, 12(1), Article 1. <https://doi.org/10.3390/ijms12010141>
- Salama, N. R., Shepherd, B., & Falkow, S. (2004). Global transposon mutagenesis and essential gene analysis of *Helicobacter pylori*. *Journal of Bacteriology*, 186(23), 7926-7935.  
<https://doi.org/10.1128/JB.186.23.7926-7935.2004>
- Salomon, C. E., Magarvey, N. A., & Sherman, D. H. (2004). Merging the potential of microbial genetics with biological and chemical diversity : An even brighter future for marine natural product drug discovery. *Natural Product Reports*, 21(1), 105-121. <https://doi.org/10.1039/b301384g>

- Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, 74(12), 5463-5467.  
<https://doi.org/10.1073/pnas.74.12.5463>
- Sangkanu, S., Rukachaisirikul, V., Suriyachadkun, C., & Phongpaichit, S. (2021). Antifungal activity of marine-derived actinomycetes against *Talaromyces marneffeii*. *Journal of Applied Microbiology*, 130(5), 1508-1522. <https://doi.org/10.1111/jam.14877>
- Sasson, O., Kaplan, N., & Linial, M. (2006). Functional annotation prediction : All for one and one for all. *Protein Science : A Publication of the Protein Society*, 15(6), 1557-1562.  
<https://doi.org/10.1110/ps.062185706>
- Schorn, M. A., Alanjary, M. M., Aguinaldo, K., Korobeynikov, A., Podell, S., Patin, N., Lincecum, T., Jensen, P. R., Ziemert, N., & Moore, B. S. (2016). Sequencing rare marine actinomycete genomes reveals high density of unique natural product biosynthetic gene clusters. *Microbiology (Reading, England)*, 162(12), 2075-2086. <https://doi.org/10.1099/mic.0.000386>
- Schulz, A., Hermann, L., Freibert, S.-A., Bönig, T., Hoffmann, T., Riclea, R., Dickschat, J. S., Heider, J., & Bremer, E. (2017). Transcriptional regulation of ectoine catabolism in response to multiple metabolic and environmental cues. *Environmental Microbiology*, 19(11), 4599-4619.  
<https://doi.org/10.1111/1462-2920.13924>
- Schwarzer, D., Finking, R., & Marahiel, M. A. (2003). Nonribosomal peptides : From genes to products. *Natural Product Reports*, 20(3), 275-287. <https://doi.org/10.1039/B111145K>
- Selevsek, N., Caiment, F., Nudischer, R., Gmuender, H., Agarkova, I., Atkinson, F. L., Bachmann, I., Baier, V., Barel, G., Bauer, C., Boerno, S., Bosc, N., Clayton, O., Cordes, H., Deeb, S., Gotta, S., Guye, P., Hersey, A., Hunter, F. M. I., ... Kleinjans, J. (2020). Network integration and modelling of dynamic drug responses at multi-omics levels. *Communications Biology*, 3(1), Article 1.  
<https://doi.org/10.1038/s42003-020-01302-8>

- Sen, R. (2010a). Surfactin : Biosynthesis, genetics and potential applications. *Advances in Experimental Medicine and Biology*, 672, 316-323. [https://doi.org/10.1007/978-1-4419-5979-9\\_24](https://doi.org/10.1007/978-1-4419-5979-9_24)
- Sen, R. (2010b). Surfactin : Biosynthesis, Genetics and Potential Applications. In R. Sen (Éd.), *Biosurfactants* (p. 316-323). Springer. [https://doi.org/10.1007/978-1-4419-5979-9\\_24](https://doi.org/10.1007/978-1-4419-5979-9_24)
- Ser, H.-L., Palanisamy, U. D., Yin, W.-F., Chan, K.-G., Goh, B.-H., & Lee, L.-H. (2016). *Streptomyces malaysiense* sp. nov. : A novel Malaysian mangrove soil actinobacterium with antioxidative activity and cytotoxic potential against human cancer cell lines. *Scientific Reports*, 6, 24247. <https://doi.org/10.1038/srep24247>
- Seringhaus, M., Paccanaro, A., Borneman, A., Snyder, M., & Gerstein, M. (2006). Predicting essential genes in fungal genomes. *Genome Research*, 16(9), 1126-1135. <https://doi.org/10.1101/gr.5144106>
- Shirling, E. B., & Gottlieb, D. (1966). Methods for characterization of *Streptomyces* species. *International Journal of Systematic Bacteriology*, 16(3), 313-340. <https://doi.org/10.1099/00207713-16-3-313>
- Sieber, S. A., & Marahiel, M. A. (2003). Learning from Nature's Drug Factories : Nonribosomal Synthesis of Macrocylic Peptides. *Journal of Bacteriology*, 185(24), 7036-7043. <https://doi.org/10.1128/JB.185.24.7036-7043.2003>
- Sieber, S. A., & Marahiel, M. A. (2005). Molecular Mechanisms Underlying Nonribosomal Peptide Synthesis : Approaches to New Antibiotics. *Chemical Reviews*, 105(2), 715-738. <https://doi.org/10.1021/cr0301191>
- Singh, R., Singh, P., Sharma, R., & Mohapatra, T. (2012). *NONRIBOSOMAL PEPTIDE SYNTHESIS IN MICROBES* (p. 183).
- Singh, T. A., Passari, A. K., Jajoo, A., Bhasin, S., Gupta, V. K., Hashem, A., Alqarawi, A. A., & Abd\_Allah, E. F. (2021). Tapping Into Actinobacterial Genomes for Natural Product Discovery. *Frontiers in Microbiology*, 12. <https://www.frontiersin.org/articles/10.3389/fmicb.2021.655620>
- Singhal, S., Singh, M., Singh, R. K., Tiwari, V. K., & Bajpai, S. (2021). Chapter 11—Molecular Mechanisms Underlying Breast Cancer and Role of Plant Products in Targeted Therapy. In G. Brahmachari (Éd.),

- Discovery and Development of Anti-Breast Cancer Agents from Natural Products* (p. 295-351). Elsevier. <https://doi.org/10.1016/B978-0-12-821277-6.00011-8>
- Sivalingam, P., Hong, K., Pote, J., & Prabakar, K. (2019). Extreme Environment *Streptomyces* : Potential Sources for New Antibacterial and Anticancer Drug Leads? *International Journal of Microbiology*, 2019, e5283948. <https://doi.org/10.1155/2019/5283948>
- Slany, M., Dziedzinska, R., Babak, V., Kralik, P., Moravkova, M., & Slana, I. (2019). Toxoplasma gondii in vegetables from fields and farm storage facilities in the Czech Republic. *FEMS Microbiology Letters*, 366(14), fnz170. <https://doi.org/10.1093/femsle/fnz170>
- Spieth, J., & Lawson, D. (2006). Overview of gene structure. *WormBook: The Online Review of C. Elegans Biology*, 1-10. <https://doi.org/10.1895/wormbook.1.65.1>
- Stamatakis, A. (2014). RAxML version 8 : A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), 1312-1313. <https://doi.org/10.1093/bioinformatics/btu033>
- Stamatakis, A., Hoover, P., & Rougemont, J. (2008). A Rapid Bootstrap Algorithm for the RAxML Web Servers. *Systematic Biology*, 57(5), 758-771. <https://doi.org/10.1080/10635150802429642>
- Supong, K., Thawai, C., Choowong, W., Kittiwongwattana, C., Thanaboripat, D., Laosinwattana, C., Koohakan, P., Parinthawong, N., & Pittayakhajonwut, P. (2016). Antimicrobial compounds from endophytic *Streptomyces* sp. BCC72023 isolated from rice (*Oryza sativa* L.). *Research in Microbiology*, 167(4), 290-298. <https://doi.org/10.1016/j.resmic.2016.01.004>
- Takahashi, H., Kumagai, T., Kitani, K., Mori, M., Matoba, Y., & Sugiyama, M. (2007). Cloning and Characterization of a *Streptomyces* Single Module Type Non-ribosomal Peptide Synthetase Catalyzing a Blue Pigment Synthesis\*. *Journal of Biological Chemistry*, 282(12), 9073-9081. <https://doi.org/10.1074/jbc.M611319200>
- Tambadou, F. (2014). *Étude de la production de peptides non-ribosomiques chez des souches de Paenibacillus* [These de doctorat, La Rochelle]. <https://www.theses.fr/2014LAROS019>

- Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11 : Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*, 38(7), 3022-3027. <https://doi.org/10.1093/molbev/msab120>
- Tan, L. T.-H., Ser, H.-L., Yin, W.-F., Chan, K.-G., Lee, L.-H., & Goh, B.-H. (2015). Investigation of Antioxidative and Anticancer Potentials of Streptomyces sp. MUM256 Isolated from Malaysia Mangrove Soil. *Frontiers in Microbiology*, 6. <https://www.frontiersin.org/articles/10.3389/fmicb.2015.01316>
- Terlouw, B. R., Blin, K., Navarro-Muñoz, J. C., Avalon, N. E., Chevrette, M. G., Egbert, S., Lee, S., Meijer, D., Recchia, M. J. J., Reitz, Z. L., van Santen, J. A., Selem-Mojica, N., Tørring, T., Zaroubi, L., Alanjary, M., Aleti, G., Aguilar, C., Al-Salihi, S. A. A., Augustijn, H. E., ... Medema, M. H. (2023). MIBiG 3.0 : A community-driven effort to annotate experimentally validated biosynthetic gene clusters. *Nucleic Acids Research*, 51(D1), D603-D610. <https://doi.org/10.1093/nar/gkac1049>
- Tettelin, H., Massignani, V., Cieslewicz, M. J., Donati, C., Medini, D., Ward, N. L., Angiuoli, S. V., Crabtree, J., Jones, A. L., Durkin, A. S., Deboy, R. T., Davidsen, T. M., Mora, M., Scarselli, M., Margarit y Ros, I., Peterson, J. D., Hauser, C. R., Sundaram, J. P., Nelson, W. C., ... Fraser, C. M. (2005). Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae : Implications for the microbial « pan-genome ». *Proceedings of the National Academy of Sciences of the United States of America*, 102(39), 13950-13955. <https://doi.org/10.1073/pnas.0506758102>
- Thi, Q. V., Tran, V. H., Mai, H. D. T., Le, C. V., Hong, M. L. T., Murphy, B. T., Chau, V. M., & Pham, V. C. (2016). Secondary Metabolites from an Actinomycete from Vietnam's East Sea. *Natural Product Communications*, 11(3), 1934578X1601100320. <https://doi.org/10.1177/1934578X1601100320>
- Tighidet, S., Plewniak, F., Sayeh, A., Gil, L., Vandecasteele, C., Lopez-Roques, C., Kecha, M., & Bertin, P. (2022). Dyes Biodegradation in the Presence of Hexavalent Chromium by Streptomyces sp. KY75 : A Novel Heavy Metal Resistant Strain Isolated in Algeria. *Water Air and Soil Pollution*, 233. <https://doi.org/10.1007/s11270-022-05756-w>

- Till, M., & Race, P. R. (2014). Progress challenges and opportunities for the re-engineering of trans-AT polyketide synthases. *Biotechnology Letters*, 36(5), 877-888. <https://doi.org/10.1007/s10529-013-1449-2>
- Tokovenko, B., Rebets, Y., & Luzhetskyy, A. (2016). *Automating Assessment of the Undiscovered Biosynthetic Potential of Actinobacteria* (p. 036087). bioRxiv. <https://doi.org/10.1101/036087>
- Touchon, M., Hoede, C., Tenaillon, O., Barbe, V., Baeriswyl, S., Bidet, P., Bingen, E., Bonacorsi, S., Bouchier, C., Bouvet, O., Calteau, A., Chiapello, H., Clermont, O., Cruveiller, S., Danchin, A., Diard, M., Dossat, C., Karoui, M. E., Frapy, E., ... Denamur, E. (2009). Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genetics*, 5(1), e1000344. <https://doi.org/10.1371/journal.pgen.1000344>
- Udwary, D. W., Gontang, E. A., Jones, A. C., Jones, C. S., Schultz, A. W., Winter, J. M., Yang, J. Y., Beauchemin, N., Capson, T. L., Clark, B. R., Esquenazi, E., Eustáquio, A. S., Freel, K., Gerwick, L., Gerwick, W. H., Gonzalez, D., Liu, W.-T., Malloy, K. L., Maloney, K. N., ... Moore, B. S. (2011). Significant Natural Product Biosynthetic Potential of Actinorhizal Symbionts of the Genus *Frankia*, as Revealed by Comparative Genomic and Proteomic Analyses ▽. *Applied and Environmental Microbiology*, 77(11), 3617-3625. <https://doi.org/10.1128/AEM.00038-11>
- Um, S., Kim, Y.-J., Kwon, H., Wen, H., Kim, S.-H., Kwon, H. C., Park, S., Shin, J., & Oh, D.-C. (2013). Sungsanpin, a lasso peptide from a deep-sea streptomycete. *Journal of Natural Products*, 76(5), 873-879. <https://doi.org/10.1021/np300902g>
- Uzair, B., Menea, F., Khan, B. A., Mohammad, F. V., Ahmad, V. U., Djeribi, R., & Menea, B. (2018). Isolation, purification, structural elucidation and antimicrobial activities of kocumarin, a novel antibiotic isolated from actinobacterium *Kocuria marina* CMG S2 associated with the brown seaweed *Pelvetia canaliculata*. *Microbiological Research*, 206, 186-197. <https://doi.org/10.1016/j.micres.2017.10.007>
- Van Lanen, S. G., & Shen, B. (2008). Biosynthesis of Eneidiyne Antitumor Antibiotics. *Current topics in medicinal chemistry*, 8(6), 448-459.



- Vanvlassenbroeck, A. (2012, juillet 17). *Etude expérimentale et in silico du potentiel de synthèse NRPS chez les Pseudomonas fluorescents*. <https://www.semanticscholar.org/paper/Etude-exp%C3%A9rimentale-et-in-silico-du-potentiel-de-Vanvlassenbroeck/10a19e01be1810257f8c5272c93d7339ef689920>
- Ventura, M., Canchaya, C., Tauch, A., Chandra, G., Fitzgerald, G. F., Chater, K. F., & van Sinderen, D. (2007). Genomics of Actinobacteria : Tracing the Evolutionary History of an Ancient Phylum. *Microbiology and Molecular Biology Reviews : MMBR*, 71(3), 495-548. <https://doi.org/10.1128/MMBR.00005-07>
- Verma, A., Singh, H., Anwar, S., Chattopadhyay, A., Tiwari, K. K., Kaur, S., & Dhilon, G. S. (2017). Microbial keratinases : Industrial enzymes with waste management potential. *Critical Reviews in Biotechnology*, 37(4), 476-491. <https://doi.org/10.1080/07388551.2016.1185388>
- Větrovský, T., & Baldrian, P. (2013). The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. *PLoS One*, 8(2), e57923. <https://doi.org/10.1371/journal.pone.0057923>
- Vicente, C. M., Thibessard, A., Lorenzi, J.-N., Benhadj, M., Hôtel, L., Gacemi-Kirane, D., Lespinet, O., Leblond, P., & Aigle, B. (2018). Comparative Genomics among Closely Related Streptomyces Strains Revealed Specialized Metabolite Biosynthetic Gene Cluster Diversity. *Antibiotics*, 7(4), Article 4. <https://doi.org/10.3390/antibiotics7040086>
- Wai, K. Z., Luechapudiporn, R., Tedsree, N., Phongsopitanun, W., Malisorn, K., & Tanasupawat, S. (2022). Streptomyces barringtoniae sp. Nov., isolated from rhizosphere of plant with antioxidative potential. *International Journal of Systematic and Evolutionary Microbiology*, 72(5), 005364. <https://doi.org/10.1099/ijsem.0.005364>
- Wang, H., Tian, R., Tian, Q., Yan, X., Huang, L., & Ji, Z. (2019). Investigation on the Antifungal Ingredients of Saccharothrix Yanglingensis Hhs.015, an Antagonistic Endophytic Actinomycete Isolated from Cucumber Plant. *Molecules*, 24(20), Article 20. <https://doi.org/10.3390/molecules24203686>

- Wang, S., Sim, T. B., Kim, Y.-S., & Chang, Y.-T. (2004). Tools for target identification and validation. *Current Opinion in Chemical Biology*, 8(4), 371-377. <https://doi.org/10.1016/j.cbpa.2004.06.001>
- Wang, T., Li, F., Lu, Q., Wu, G., Jiang, Z., Liu, S., Habden, X., Razumova, E. A., Osterman, I. A., Sergiev, P. V., Dontsova, O. A., Hu, X., You, X., & Sun, C. (2021). Diversity, novelty, antimicrobial activity, and new antibiotics of cultivable endophytic actinobacteria isolated from psammophytes collected from Taklamakan Desert. *Journal of Pharmaceutical Analysis*, 11(2), 241-250. <https://doi.org/10.1016/j.jpha.2020.06.004>
- Wang, Z., Yu, Z., Zhao, J., Zhuang, X., Cao, P., Guo, X., Liu, C., & Xiang, W. (2020). Community Composition, Antifungal Activity and Chemical Analyses of Ant-Derived Actinobacteria. *Frontiers in Microbiology*, 11, 201. <https://doi.org/10.3389/fmicb.2020.00201>
- Wanner, L. A. (2009). A Patchwork of Streptomyces Species Isolated from Potato Common Scab Lesions in North America. *American Journal of Potato Research*, 86(4), 247-264. <https://doi.org/10.1007/s12230-009-9078-y>
- Wattam, A. R., Davis, J. J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., Conrad, N., Dietrich, E. M., Disz, T., Gabbard, J. L., Gerdes, S., Henry, C. S., Kenyon, R. W., Machi, D., Mao, C., Nordberg, E. K., Olsen, G. J., Murphy-Olson, D. E., Olson, R., ... Stevens, R. L. (2017). Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. *Nucleic Acids Research*, 45(D1), D535-D542. <https://doi.org/10.1093/nar/gkw1017>
- Weber, T., & Kim, H. U. (2016). The secondary metabolite bioinformatics portal : Computational tools to facilitate synthetic biology of secondary metabolite production. *Synthetic and Systems Biotechnology*, 1(2), 69-79. <https://doi.org/10.1016/j.synbio.2015.12.002>
- Weisenfeld, N. I., Kumar, V., Shah, P., Church, D. M., & Jaffe, D. B. (2017). Direct determination of diploid genome sequences. *Genome Research*, 27(5), 757-767. <https://doi.org/10.1101/gr.214874.116>
- Weissman, K. J. (2009). Introduction to polyketide biosynthesis. *Methods in Enzymology*, 459, 3-16. [https://doi.org/10.1016/S0076-6879\(09\)04601-1](https://doi.org/10.1016/S0076-6879(09)04601-1)

- Wenzel, S. C., & Müller, R. (2005). Formation of novel secondary metabolites by bacterial multimodular assembly lines : Deviations from textbook biosynthetic logic. *Current Opinion in Chemical Biology*, 9(5), 447-458. <https://doi.org/10.1016/j.cbpa.2005.08.001>
- Woodbury, A. M. (1940). Antler-Eating by Rodents. *Science*, 92(2380), 127-128. <https://doi.org/10.1126/science.92.2380.127.b>
- Xiao, J., Sekhwal, M. K., Li, P., Ragupathy, R., Cloutier, S., Wang, X., & You, F. M. (2016). Pseudogenes and Their Genome-Wide Prediction in Plants. *International Journal of Molecular Sciences*, 17(12), 1991. <https://doi.org/10.3390/ijms17121991>
- Xu, F., Wu, Y., Zhang, C., Davis, K. M., Moon, K., Bushin, L. B., & Seyedsayamdost, M. R. (2019). A Genetics-Free Method for High-Throughput Discovery of Cryptic Microbial Metabolites. *Nature chemical biology*, 15(2), 161-168. <https://doi.org/10.1038/s41589-018-0193-2>
- Xu, L., Dong, Z., Fang, L., Luo, Y., Wei, Z., Guo, H., Zhang, G., Gu, Y. Q., Coleman-Derr, D., Xia, Q., & Wang, Y. (2019). OrthoVenn2 : A web server for whole-genome comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Research*, 47(W1), W52-W58. <https://doi.org/10.1093/nar/gkz333>
- Yandell, M., & Ence, D. (2012). A beginner's guide to eukaryotic genome annotation. *Nature Reviews Genetics*, 13(5), Article 5. <https://doi.org/10.1038/nrg3174>
- Yang, L., Hou, L., Li, H., & Li, W. (2019). Antibiotic angucycline derivatives from the deepsea-derived *Streptomyces lusitanus*. *Natural Product Research*, 34, 1-7. <https://doi.org/10.1080/14786419.2019.1577835>
- Yang, T., Yamada, K., Zhou, T., Harunari, E., Igarashi, Y., Terahara, T., Kobayashi, T., & Imada, C. (2019). Akazamicin, a cytotoxic aromatic polyketide from marine-derived *Nonomuraea* sp. *The Journal of Antibiotics*, 72(4), 202-209. <https://doi.org/10.1038/s41429-018-0139-7>

- Yang, Z., He, J., Wei, X., Ju, J., & Ma, J. (2020). Exploration and genome mining of natural products from marine Streptomyces. *Applied Microbiology and Biotechnology*, 104(1), 67-76.  
<https://doi.org/10.1007/s00253-019-10227-0>
- Yu, H., Greenbaum, D., Xin Lu, H., Zhu, X., & Gerstein, M. (2004). Genomic analysis of essentiality within protein networks. *Trends in Genetics*, 20(6), 227-231. <https://doi.org/10.1016/j.tig.2004.04.008>
- Yugi, K., Kubota, H., Hatano, A., & Kuroda, S. (2016). Trans-Omics : How To Reconstruct Biochemical Networks Across Multiple 'Omic' Layers. *Trends in Biotechnology*, 34(4), 276-290.  
<https://doi.org/10.1016/j.tibtech.2015.12.013>
- Zambri, M. P., Williams, M. A., & Elliot, M. A. (2022). Chapter Five - How Streptomyces thrive : Advancing our understanding of classical development and uncovering new behaviors. In R. K. Poole & D. J. Kelly (Éds.), *Advances in Microbial Physiology* (Vol. 80, p. 203-236). Academic Press.  
<https://doi.org/10.1016/bs.ampbs.2022.01.004>
- Zeng, J., Decker, R., & Zhan, J. (2012). Biochemical Characterization of a Type III Polyketide Biosynthetic Gene Cluster from Streptomyces toxytricini. *Applied Biochemistry and Biotechnology*, 166(4), 1020-1033. <https://doi.org/10.1007/s12010-011-9490-x>
- Zenova, G. M., Manucharova, N. A., & Zvyagintsev, D. G. (2011). Extremophilic and extremotolerant actinomycetes in different soil types. *Eurasian Soil Science*, 44(4), 417-436.  
<https://doi.org/10.1134/S1064229311040132>
- Zerouki, C., Bensalah, F., Kuitinen, S., Pappinen, A., & Turunen, O. (2021). Whole-genome sequencing of two Streptomyces strains isolated from the sand dunes of Sahara. *BMC Genomics*, 22(1), 578.  
<https://doi.org/10.1186/s12864-021-07866-x>
- Zhang, B., Han, D., Korostelev, Y., Yan, Z., Shao, N., Khrameeva, E., Velichkovsky, B. M., Chen, Y.-P. P., Gelfand, M. S., & Khaitovich, P. (2016). Changes in snoRNA and snRNA Abundance in the Human, Chimpanzee, Macaque, and Mouse Brain. *Genome Biology and Evolution*, 8(3), 840-850.  
<https://doi.org/10.1093/gbe/evw038>

- Zhang, M. M., Wong, F. T., Wang, Y., Luo, S., Lim, Y. H., Heng, E., Yeo, W. L., Cobb, R. E., Enghiad, B., Ang, E. L., & Zhao, H. (2017). CRISPR-Cas9 strategy for activation of silent *Streptomyces* biosynthetic gene clusters. *Nature Chemical Biology*. <https://doi.org/10.1038/nchembio.2341>
- Zhou, Q., Luo, G.-C., Zhang, H., & Tang, G.-L. (2019). Discovery of 16-Demethylrifamycins by Removing the Predominant Polyketide Biosynthesis Pathway in *Micromonospora* sp. Strain TP-A0468. *Applied and Environmental Microbiology*, *85*(4), e02597-18. <https://doi.org/10.1128/AEM.02597-18>
- Zimmer, C., & Wähnert, U. (1986). Nonintercalating DNA-binding ligands : Specificity of the interaction and their use as tools in biophysical, biochemical and biological investigations of the genetic material. *Progress in Biophysics and Molecular Biology*, *47*(1), 31-112. [https://doi.org/10.1016/0079-6107\(86\)90005-2](https://doi.org/10.1016/0079-6107(86)90005-2)

---



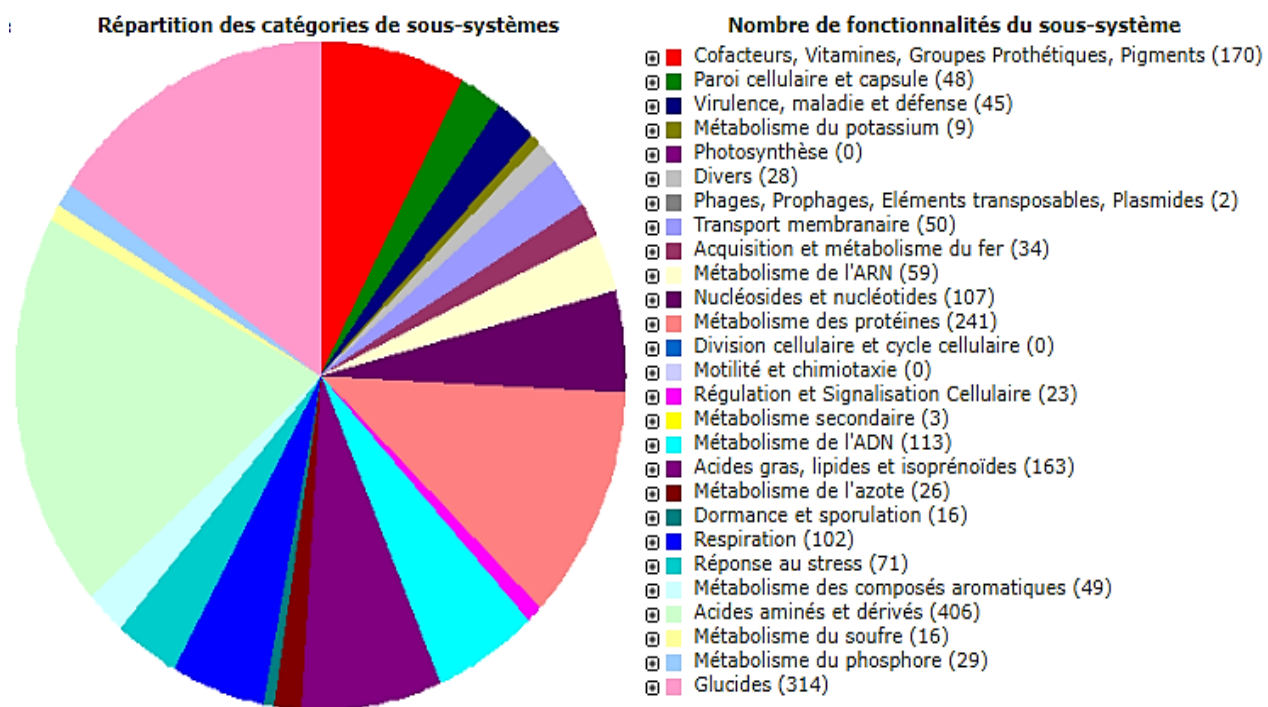
---

# ANNEXE

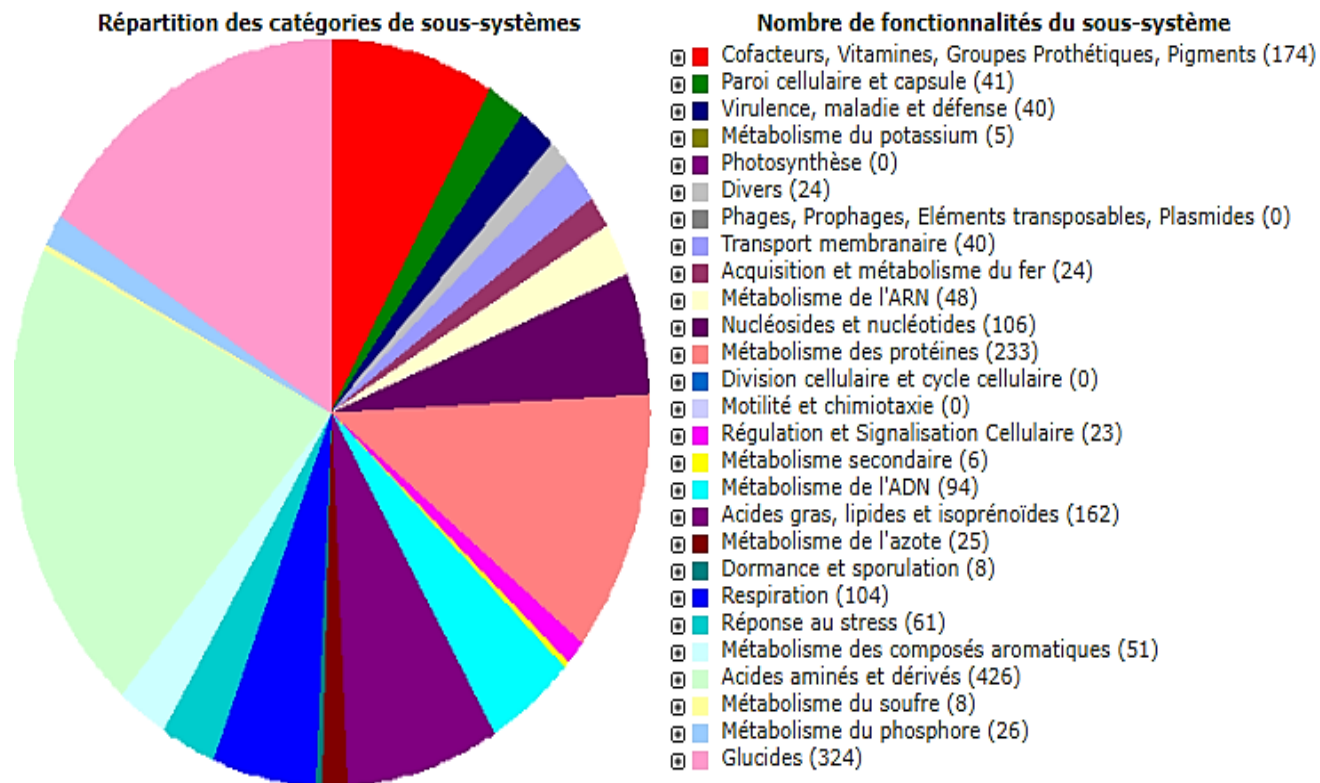
---



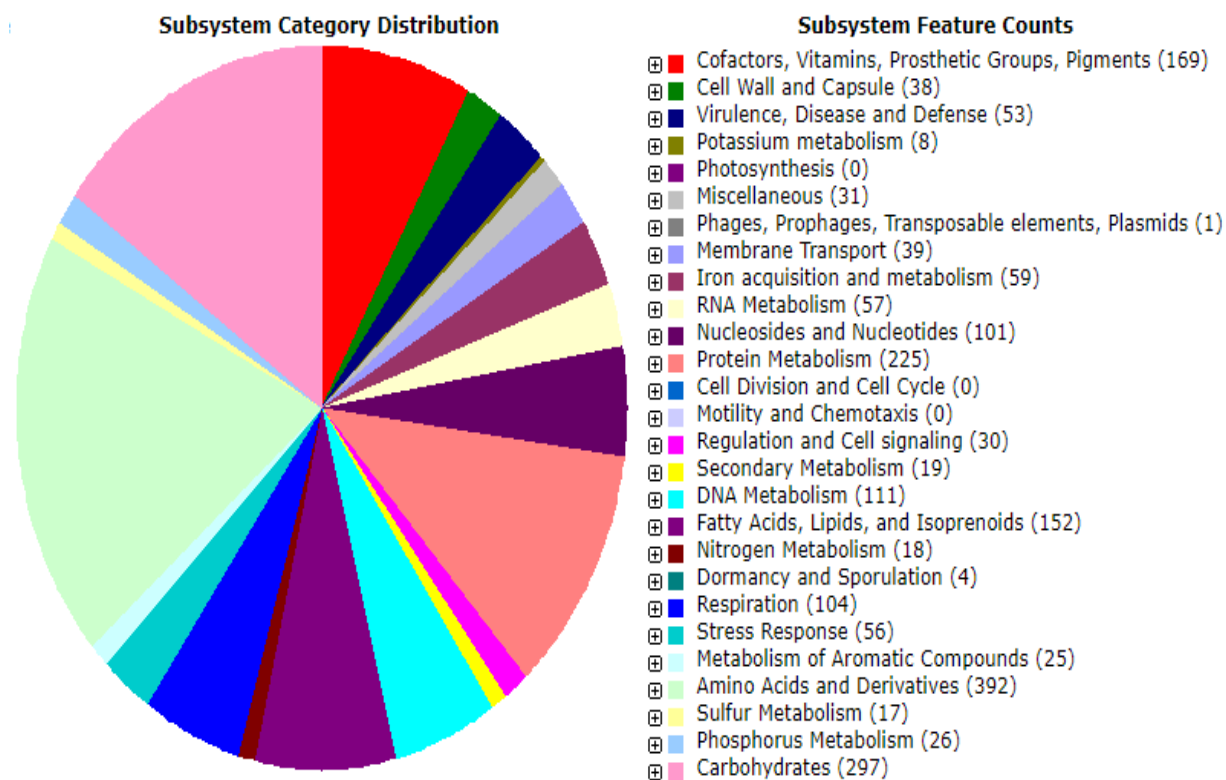
---



**Annexe.1** les sous-système de *Streptomyces ambofaciens* ATCC 23877 et leurs fonctions, résultat de Rast.



**Annexe.2** les sous-système de *Streptomyces leeuwenhoekii* C34 (= DSM 42122 = NRRL B-24963) et leurs fonctions.



**Annexe.3** les sous-système de *Streptomyces* sp. KY75 et leurs fonctions.



---

---

# RESUME

---

---

Les polykétides synthases (PKS) et les peptides synthétases non ribosomiques (NRPS) sont des enzymes responsables de la biosynthèse de nombreux métabolites naturels thérapeutiquement et biotechnologiquement importants. Les actinobactéries, notamment le genre *Streptomyces*, sont connues pour leur capacité exceptionnelle à produire des substances bioactives grâce à la présence de clusters de gènes biosynthétiques dans leur génome. Les avancées bio-informatiques nous ont permis d'identifier et de prédire ces clusters présents dans trois souches étudiées *Streptomyces ambofaciens* ATCC 23877, *Streptomyces leeuwenhoekii* C34 et *Streptomyces sp.* KY75 qui ont été isolées à partir de différents écosystèmes extrêmes en Algérie, tels que les sources thermales et les zones polluées. Cela a permis d'analyser et de comparer *in silico* les groupes de gènes biosynthétiques PKS et NRPS afin de repérer de nouveaux composés bioactifs qui permettent aux actinomycètes de survivre dans ces conditions. Les séquences génomiques complètes ont été récupérées à partir de la base de données NCBI (*National Center for Biotechnology Information*). L'analyse phylogénétique et comparative des génomes des souches étudiées a révélé une diversité génomique entre elles. Les différences comprennent la taille du génome (7 837 660 à 8 303 940 paires de bases) et la composition en guanine-cytosine (71,6% à 72,8%). Le nombre de séquences codantes (CDS) varie également, avec des totalités de 7 097 à 12 034 CDS, dont certains codent pour des protéines fonctionnelles liées à des fonctions métaboliques. Le seueur antiSMASH 7.0 a permis de détecter un total de 27 clusters de gènes biosynthétiques (BGC), avec 8 clusters communs entre les trois souches. La souche *Streptomyces leeuwenhoekii* C34 présente le plus grand nombre de BGC (41), suivie de *Streptomyces ambofaciens* ATCC 23877 (35), tandis que *Streptomyces sp.* KY75 en a le moins (32). Les BGC identifiés appartiennent aux classes principales de métabolites, notamment les PKS et les NRPS. Ces résultats soulignent le potentiel métabolique important et la diversité des souches étudiées, offrant des perspectives pour la découverte de nouveaux composés bioactifs. La recherche sur ces actinobactéries constitue une source prometteuse pour le développement de nouveaux médicaments et peut ouvrir des opportunités de commercialisation. Notre étude représente une avancée dans la recherche scientifique en Algérie, car elle est considérée comme la première de ce genre à être menée dans le pays.

**Mots clés :** PKS, NRPS, actinobactéries, BGC, *Streptomyces*.

## Abstract

Polyketide synthases (PKS) and non-ribosomal peptide synthetases (NRPS) are enzymes responsible for the biosynthesis of many therapeutically and bio-technologically important natural metabolites. Actinobacteria, notably the genus *Streptomyces*, are known for their exceptional ability to produce bioactive substances thanks to the presence of biosynthetic gene clusters in their genome. Bioinformatics advances enabled us to identify and predict these clusters present in three strains studied *Streptomyces ambofaciens* ATCC 23877, *Streptomyces leeuwenhoekii* C34 and *Streptomyces* sp. KY75, which were isolated from different extreme ecosystems in Algeria, such as thermal springs and polluted areas. This enabled in silico analysis and comparison of the PKS and NRPS biosynthetic gene clusters to identify novel bioactive compounds that enable actinomycetes to survive in these conditions. Complete genomic sequences were retrieved from the NCBI (National Center for Biotechnology Information) database. Phylogenetic and comparative analysis of the genomes of the strains studied revealed genomic diversity between them. Differences include genome size (7,837,660 to 8,303,940 base pairs) and guanine-cytosine composition (71.6% to 72.8%). The number of coding sequences (CDS) also varies, with totalities ranging from 7,097 to 12,034 CDS, some of which code for functional proteins linked to metabolic functions. The antiSMASH 7.0 server detected a total of 27 biosynthetic gene clusters (BGCs), with 8 clusters common to all three strains. The *Streptomyces leeuwenhoekii* C34 strain had the highest number of BGCs (41), followed by *Streptomyces ambofaciens* ATCC 23877 (35), while *Streptomyces* sp. KY75 had the fewest (32). The BGCs identified belong to the main classes of metabolites, notably PKSs and NRPSs. These results underline the significant metabolic potential and diversity of the strains studied, offering prospects for the discovery of new bioactive compounds. Research into these actinobacteria is a promising source for the development of new drugs, and may open up opportunities for commercialization. Our study represents a breakthrough in scientific research in Algeria, as it is considered the first of its kind to be carried out in the country.

**Key words:** PKS, NRPS, actinobacteria, BGC, *Streptomyces*.

## ملخص:

إن مُركبات البوليكيتيد وأنزيمات تركيب الببتيد غير الريبوزومي هي الإنزيمات المسؤولة عن إنتاج العديد من المركبات الطبية والحيوية التي تتميز بأهمية كبيرة في مجال التكنولوجيا الحيوية. تعتبر الاكتينوبكتيريا، وعلى وجه الخصوص جنس الستربتوميس (*Streptomyces*)، معروفة بقدرتها الاستثنائية على إنتاج مواد حيوية نشطة بفضل وجود مجموعات الجينات التركيبية الحيوية في جينومها. تمكّنا التطورات في مجال البيوانفورماتيك من تحديد وتنبؤ هذه المجموعات الموجودة في ثلاثة سلالات مدروسة، وهي *Streptomyces leeuwenhoekii* C34 و *Streptomyces ambofaciens* ATCC 23877 و *Streptomyces sp.* KY75، والتي تم عزلها من بيئات متطرفة مختلفة في الجزائر، مثل الينابيع الحاررية والمناطق الملوثة. تمكّنا ذلك من إجراء تحليل حوسبي ومقارنة لمجموعات الجينات التركيبية للبوليكيتيد والببتيدات غير الريبوزومية لتحديد مركبات حيوية جديدة تمكن الأكتينومييس من البقاء في هذه الظروف. تم استرداد تسلسلات الجينوم الكاملة من قاعدة بيانات NCBI (National Center for Biotechnology Information) أظهر التحليل الفيلوجيني والمقارن لجينومات السلالات المدروسة تنوعًا جينوميًا بينها. تتفاوت الاختلافات بمقدار الجينوم (من 7,837,660 إلى 8,303,940 زوج قاعدي) وتركيب الغوانين-السيوسين (71.6% إلى 72.8%). كما تختلف عدد تسلسلات الترميز (CDS)، حيث يتراوح الإجمالي بين 7097 و 12034 CDS، بعضها يُرمز لبروتينات وظيفية مرتبطة بالوظائف الأيضية. كشف خادم antiSMASH 7.0 عن إجمالي 27 مجموعة جينات تركيبية حيوية (BGCs)، حيث كانت 8 من هذه المجموعات مشتركة بين السلالات الثلاثة. تعتبر سلالة *Streptomyces leeuwenhoekii* C34 لديها أعلى عدد من مجموعات الجينات التركيبية (41)، تليها سلالة (35) *Streptomyces ambofaciens* ATCC 23877، في حين تحتوي *Streptomyces sp.* KY75 على الأقل (32). تنتمي المجموعات الجينية التي تم التعرف عليها إلى الفئات الرئيسية للمركبات الأيضية، بما في ذلك البوليكيتيد والببتيدات غير الريبوزومية. تؤكد هذه النتائج الإمكانيات الأيضية والتنوع الكبير للسلالات المدروسة، مما يفتح آفاقًا لاكتشاف مركبات حيوية جديدة. يمثل البحث في هذه الأكتينوبكتيريا مصدرًا واعدًا لتطوير أدوية جديدة وقد يفتح آفاقًا للتسويق التجاري. يمثل بحثنا اختراقًا في البحث العلمي في الجزائر، حيث تعتبر أول دراسة من نوعها تُجرى في الجزائر.

**الكلمات المفتاحية:** الاكتينوبكتيريا، مجموعة جينات تركيبية حيوية (BGCs)، مُركبات البوليكيتيد، تركيب الببتيد غير الريبوزومي، الستربتوميس.