

République algérienne Démocratique et populaire
Ministère de l'enseignement Supérieur et de la recherche Scientifique
Université de Blida -1-
Faculté de science de la nature et de la vie
Département de Biotechnologie et Agroécologie



Mémoire de fin d'étude
En vue de l'obtention du Diplôme de Master 2 en
Biotechnologie
Option : Biotechnologie Microbienne

La violacéine chez *Chromobacterium*
***violaceum* : Quorum-Sensing, étude *in silico*.**

Réalisé par :

SELLAM Imen Chahinez

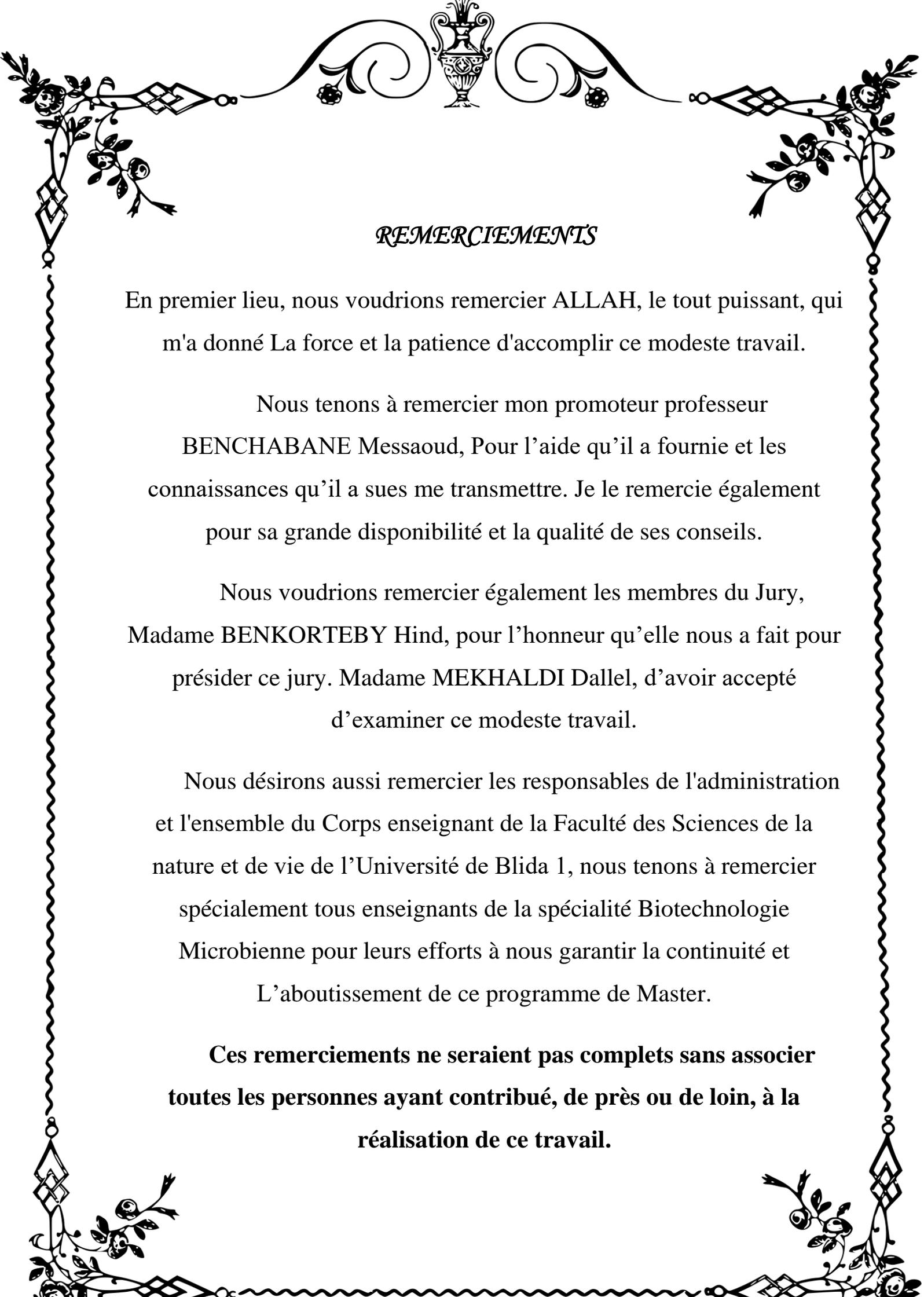
SLIMANI Hadjer

Soutenu le : 12/07/2022

Devant le jury composé de :

Présidente :	Mme BENKORTEBY H.	MAA	USDB1
Examinatrice :	Mme MEKHALDI D.	MAB	USDB1
Promoteur :	Mr BENCHABANE M.	Professeur	USDB1

Année universitaire 2021/2022



REMERCIEMENTS

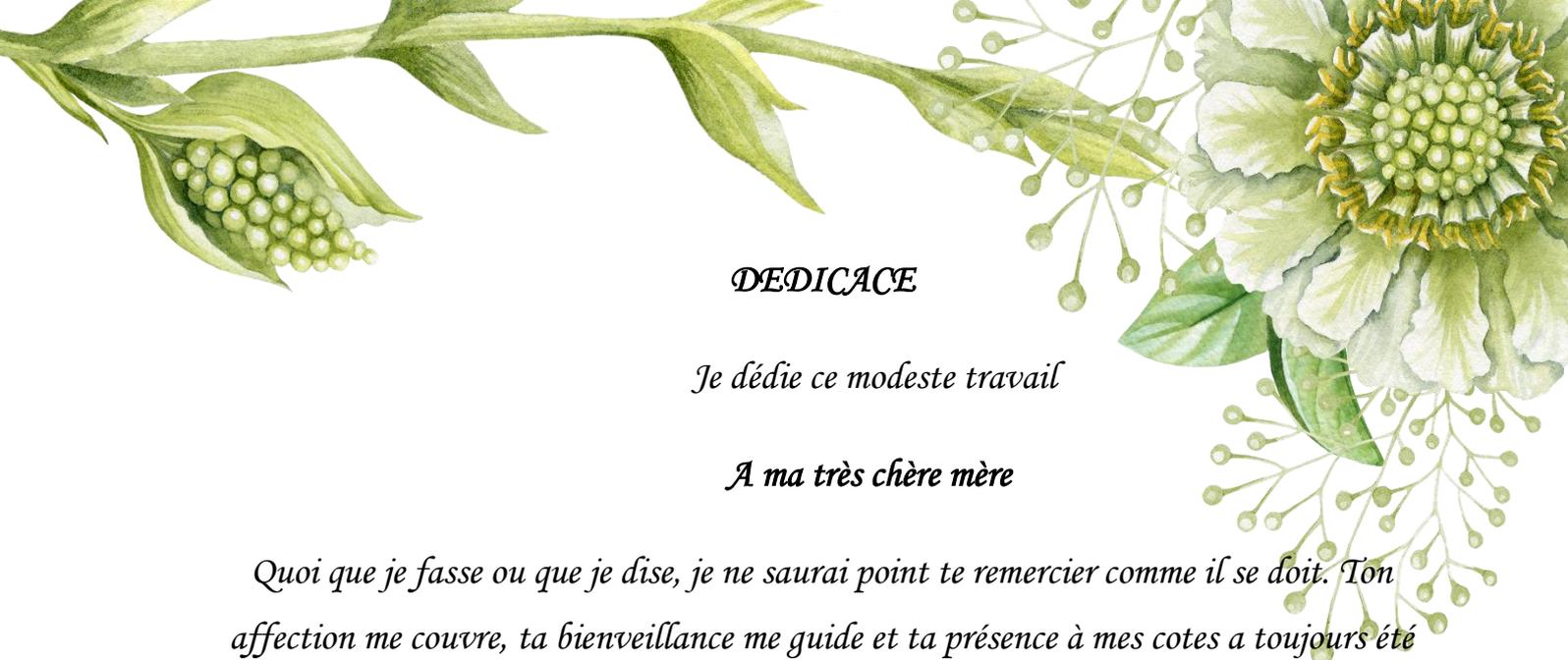
En premier lieu, nous voudrions remercier ALLAH, le tout puissant, qui m'a donné La force et la patience d'accomplir ce modeste travail.

Nous tenons à remercier mon promoteur professeur BENCHABANE Messaoud, Pour l'aide qu'il a fournie et les connaissances qu'il a sues me transmettre. Je le remercie également pour sa grande disponibilité et la qualité de ses conseils.

Nous voudrions remercier également les membres du Jury, Madame BENKORTEBY Hind, pour l'honneur qu'elle nous a fait pour présider ce jury. Madame MEKHALDI Dallel, d'avoir accepté d'examiner ce modeste travail.

Nous désirons aussi remercier les responsables de l'administration et l'ensemble du Corps enseignant de la Faculté des Sciences de la nature et de vie de l'Université de Blida 1, nous tenons à remercier spécialement tous enseignants de la spécialité Biotechnologie Microbienne pour leurs efforts à nous garantir la continuité et L'aboutissement de ce programme de Master.

Ces remerciements ne seraient pas complets sans associer toutes les personnes ayant contribué, de près ou de loin, à la réalisation de ce travail.



DEDICACE

Je dédie ce modeste travail

A ma très chère mère

Quoi que je fasse ou que je dise, je ne saurai point te remercier comme il se doit. Ton affection me couvre, ta bienveillance me guide et ta présence à mes cotes a toujours été ma source de force pour affronter les différents obstacles de ma vie, Merci maman

A mon très cher père

Tu as toujours été à mes cotes pour me soutenir et m'encourager que ce travail traduit ma gratitude et mon affection.

A mes chères sœurs HIBA, DALAL et MALAK

A mon cher grand frère CHAOUKI

A l'âme de ma grand mère DOUADI YAMINA, que dieu vous garde dans son vaste paradis.

A mes grands parents, que dieu vous protège

A mes oncles YUCEF, MUSTAPHA, MOHAMED et ABED LKRIM

A ma binôme et ma meilleur amie SLIMANI HADJER

A mes amies proche RANIA, MANEL, RAYANE et NESRINE

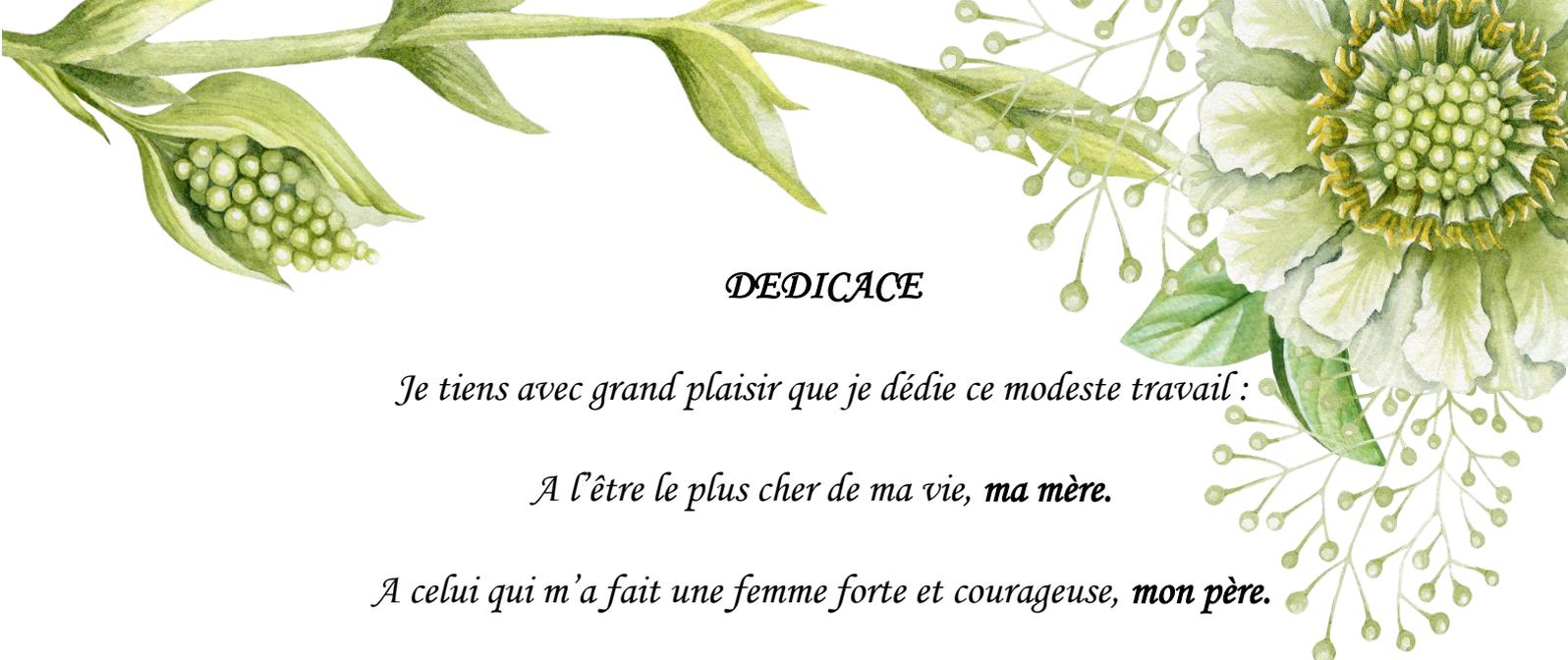
A ma famille de grands jusqu' aux petits.

Un grand merci à vous...!



Chahinez





DEDICACE

Je tiens avec grand plaisir que je dédie ce modeste travail :

*A l'être le plus cher de ma vie, **ma mère.***

*A celui qui m'a fait une femme forte et courageuse, **mon père.***

*A mon cher frère **ABDELNACER** et mes chères sœurs **NOUR EL IMENE,**
MEROUA et **ASMAA.***

*A l'âme de papa-sidou **EL HANACHI**, que dieu vous garde dans son vaste
paradis.*

A mes grands-parents, que dieu vous accorde une longue vie.

*A mes chers cousins **FATIMA, NABILA** et **SHAIMAA.***

*Au professeur **KRIMI ZOULIKHA** et **HAMOUCHE BACHIR**, sans eux, je ne
serais pas où je suis aujourd'hui.*

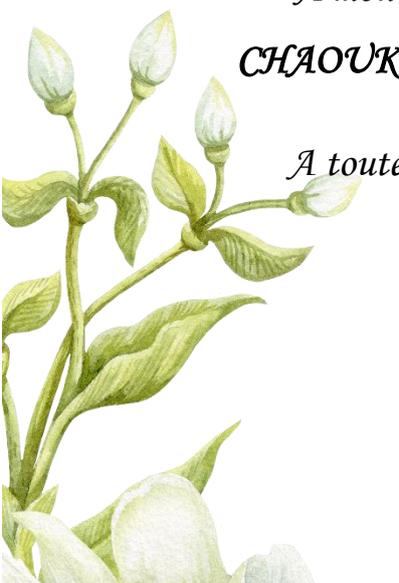
*A mes sœurs d'âme **SARAH** et **RANIA.***

*Aux membres du club **BIBLIOPHILES**, surtout **KHAOULA** et **KAWTHER,***

*A mon binôme **CHAHINEZ**, mon ami **MOHAMED** et mon grand frère
CHAOUKI qui ont été à mes côtés pendant toutes mes années dans l'université.*

A toutes mes connaissances et à tous ceux dont j'ai un jour croisé la route

Je vous remercie ... !



Hadjer

Résumé

Chromobacterium violaceum est une Betaproteobacteria Gram-négatif, de la famille de Neisseriaceae, coccobacille anaérobie facultative et non sporulant. Cette bactérie est largement utilisée dans les laboratoires de microbiologie impliqués dans la recherche sur le quorum sensing (QS). Parmi les caractéristiques régulées par le quorum sensing, la production de violacéine a reçu la plus grande attention. Ce pigment a une valeur d'application médicale potentielle, en qualité d'antibactériens, d'antiulcéreux et d'anticancéreux. Du fait de son intérêt pour des applications biotechnologiques, la compréhension et l'étude des mécanismes de sa régulation génétique est évidente et plus qu'indispensable. Après avoir consulté les nouveautés des publications scientifiques, nous avons élaboré une revue bibliographique pour réaliser une synthèse sur les caractéristiques microbiologiques de base de cette bactérie et la biosynthèse de son pigment, tout en abordant l'aspect global du concept du quorum sensing, notamment sa régulation chez *C. violaceum*.

A l'aide de méthodes de recherche bioinformatiques (*in silico*), nous chercherons de comparer les phylogénies moléculaires (régulation génétique et biosynthèse) pour souligner les propriétés génétiques et protéiques de ce pigment. Les résultats de l'analyse bioinformatique ont montré que la protéine régulatrice CviR chez *C. violaceum* ATCC12472 est une homologue avec presque toutes les protéines régulatrices LuxR, elle présente une structure bien précise qui assure une régulation et un contrôle pour l'expression des gènes de l'opéron *Vio* qui sont en nombre de 5 gènes (*VioA*, *VioB*, *VioC*, *VioD* et *VioE*). Ces gènes sont disponibles chez toutes les souches de *C. violaceum* et existent chez d'autres espèces du genre *Chromobacterium*, mais ils ne sont pas complètement identiques, ce qui explique pourquoi seulement les espèces de *C. violaceum* ont la capacité de synthèse du pigment de violacéine.

Mots clés:

Chromobacterium violaceum, Violacéine, Régulations des Gènes, Quorum sensing, *in silico*

Abstract**Violacein in *Chromobacterium violaceum*: Quorum-Sensing, in silico study**

Chromobacterium violaceum is a Gram-negative Betaproteobacteria, family Neisseriaceae, facultative anaerobic and non-spore forming coccobacillus. This bacterium is widely used in microbiology laboratories involved in quorum sensing (QS) research. Among the characteristics regulated by quorum sensing, the production of violacein has received the most attention. This pigment has potential medical application value as antibacterial, antiulcer and anticancer agents. Because of its interest for biotechnological applications, the understanding and study of the mechanisms of its genetic regulation is obvious and more than necessary. After having consulted the latest scientific publications, we have elaborated a bibliographical review to realize a synthesis on the basic microbiological characteristics of this bacterium and the biosynthesis of its pigment, while approaching the global aspect of the concept of quorum sensing, especially its regulation in *C. violaceum*.

Using bioinformatics research methods (*in silico*), we will seek to compare molecular phylogenies (genetic regulation and biosynthesis) to highlight the genetic and protein properties of this pigment. The results of the bioinformatics analysis showed that the regulatory protein CviR in *C. violaceum* ATCC12472 is homologous with almost all LuxR regulatory proteins, it has a very precise structure that ensures regulation and control for the expression of the genes of the *Vio* operon which are in number of 5 genes (*VioA*, *VioB*, *VioC*, *VioD* and *VioE*) These genes are available in all strains of *C. violaceum* and exist in other species of the genus *Chromobacterium*, but they are not completely identical, which explains why only species of *C. violaceum* have the ability to synthesize the violacein pigment.

Key words:

Chromobacterium violaceum, Violacein, Regulation of Genes, Quorum sensing, *in silico*

ملخص

الفيولاسين في *Chromobacterium violaceum*: نظام الإشارات، دراسة *in silico*.

Chromobacterium violaceum وهو Betaproteobacteria سلبي الغرام، ينتمي إلى عائلة Neisseriaceae عصيات لاهوائية اختيارية وغير بوجية. تم استخدام هذه البكتيريا على نطاق واسع في مختبرات الأحياء الدقيقة المشاركة في أبحاث نظام الإشارات، من بين الخصائص المنظمة لنظام الإشارات لهذه البكتيريا، حظي إنتاج الفيولاسين بأكبر قدر من الاهتمام. هذه الصبغة لها قيمة تطبيقية محتملة كأدوية مضادة للبكتيريا والقرحة ومضادة للسرطان. نظرًا لتطبيقاتها التكنولوجية الحيوية الهامة، فإن فهم ودراسة آليات تنظيمها الجيني واضح وأكثر من الضروري. بعد الاطلاع على أحدث المنشورات العلمية، قمنا بإعداد مراجعة ببيوغرافية لتحقيق توليفة حول الخصائص الميكروبيولوجية الأساسية لهذه البكتيريا والتركيب الحيوي لصبغتها، مع الاقتراب من الجانب العالمي لمفهوم نظام الإشارات، وخاصة تنظيمه في

Chromobacterium violaceum.

باستخدام طرق البحث في المعلوماتية الحيوية (*in silico*)، سنسعى إلى مقارنة السلالات الجزئية (التنظيم الجيني الحيوي) لتسليط الضوء على الخصائص الجينية والبروتينية لهذه الصبغة. أظهرت نتائج تحليل المعلوماتية الحيوية أن البروتين التنظيمي CviR في *Chromobacterium violaceum* ATCC12472 متماثل مع جميع بروتينات تنظيم LuxR تقريبًا، فهو يقدم بنية دقيقة تضمن التنظيم والتحكم في التعبير عن جينات المشغل *Vio* والتي عددها 5 جينات (*VioA*, *VioB*, *VioC*, *VioD* و *VioE*) هذه الجينات متوفرة في جميع أنواع *Chromobacterium violaceum* ويمكن أن تكون متاحة في جنس آخر من *Chromobacterium sp* لكنها ليست متطابقة تمامًا، وهو ما يفسر سبب وجود أنواع فقط من بكتيريا *Chromobacterium violaceum* لديها القدرة على إنتاج صبغة الفيولاسين.

الكلمات المفتاحية:

Chromobacterium violaceum, فيولاسين، تنظيم الجينات، نظام الإشارات، *in silico*

Liste des abréviations

AINS : Anti-Inflammatoires Non Stéroïdiens

AIP : Peptides Auto-Inducteurs

BLAST : Basic Local Aligement Search Tool

EBI : L'Institut Européen de Bioinformatique

EMBL : European Molecular Biology Laboratory

FASTA : FAST All

HCD : Hight Cellular Density (haute densité cellulaire)

IgE : Immunoglobuline E

LCD : Low Cellular Density (faile densité cellulaire)

NCBI : National Center for Biotechnology Information

NIH : National Institutes of Health

NJ : Neighbor Joining

NLM : National Library of Medicine

PCR : Réaction de Polymérisation en Chaîne

PDB : Protein Data Bank

PIR : Protein Information Resource

QS : Quorum sensing

Uniprot : Universal Protein Resource

YAC : Yeast Artificial Chromosome (chromosome artificiel de levure)

Liste des figures

Figure 1: Croissance de <i>Chromobacterium violaceum</i> sur gélose MacConkey (à gauche) et sur gélose au sang de mouton à 5% (à droite).	5
Figure 2 : <i>Chromobacterium violaceum</i> sur microscope optique Gx40	5
Figure 3 : La structure chimique des métabolites de <i>Chromobacterium violaceum</i>	7
Figure 4 : Biosynthèse de la violacéine.	8
Figure 5 : Schéma résumé de la biosynthèse de la violacéine.	9
Figure 6 : Représentation graphique circulaire du génome de <i>Chromobacterium violaceum</i>	11
Figure 7 : Test de sensibilité aux antimicrobiens de <i>Chromobacterium violaceum</i>	12
Figure 8 : Structure chimique de violacéine qui dérive de deux molécules de tryptophane modifiées. Il est constitué de trois noyaux principaux : du 5-hydroxyindole (à gauche), de l'oxindole (en haut, à droite) et de la 2-pyrrolidone (au milieu).	14
Figure 9 : Schéma de biosynthèse de la violacéine. La liaison du complexe CviR-AHL au site promoteur du VioA déclenche l'expression du <i>Vio</i> opéron, conduisant à la synthèse du pigment violet violacéine.	15
Figure 10 : Voie biosynthétique de la violacéine (noire), de la staurosporine (bleu) et de la rebeccamycine (rouge).	18
Figure 11 : Les réactions chimiques entrant dans la biosynthèse du violacéine par <i>Chromobacterium violaceum</i>	19
Figure 12 : Mécanisme de quorum sensing.	25
Figure 13 : structure d'autoinducteur-2, AI-2.	26
Figure 14 : A. <i>Vibrio fischeri</i> bioluminescent sous un microscope et B. un calmar à queue de cheval, avec un manteau d'une longueur d'environ 1,2 pouce.	27
Figure 15 : Détection du quorum chez <i>Vibrio fischeri</i> ; une signalisation LuxIR circuit. Les triangles rouges indiquent l'autoinducteur qui est produit par LuxI. OM , membrane externe ; IM , membrane interne.	27

Figure 16 : Structure d’autoinducteur AHL.	28
Figure 17 : Divers peptides de détection du quorum, illustrant la diversité structurelle observée parmi les peptides de détection du quorum.	29
Figure 18 : (Panneau de gauche) Mécanisme typique de détection du quorum des bactéries Gram-négatives (AHL). (Panneau de droite) Mécanisme typique de détection du quorum des bactéries Gram-positives (Peptide).....	30
Figure 19 : Multi discipline de la bio-informatique	32
Figure 20 : Professeur Margaret Oakley Dayhof	33
Figure 21 : Exemple de clustering de données.....	37
Figure 22 : représente les ressources en bioinformatique. Les ressources web et les ressources lignes de commandes.....	39
Figure 23 : Format FASTA de la séquence qui code pour la protéine CviR chez <i>Chromobacterium violaceum</i> ATCC12472.....	48
Figure 24 : Format FASTA de la séquence qui code pour le gène <i>VioA</i> chez <i>Chromobacterium violaceum</i>	48
Figure 25 : Format FASTA de la séquence qui code pour le gène <i>VioB</i> chez <i>Chromobacterium violaceum</i> ATCC12472.....	49
Figure 26 : Format FASTA de la séquence qui code pour le gène <i>VioC</i> chez <i>Chromobacterium violaceum</i> ATCC12472.....	49
Figure 27 : Format FASTA de la séquence qui code pour le gène <i>VioD</i> chez <i>Chromobacterium violaceum</i> ATCC12472.....	50
Figure 28 : Format FASTA de la séquence qui code pour le gène <i>VioE</i> chez <i>Chromobacterium violaceum</i> ATCC12472.....	50
Figure 29 : Comment rechercher la structure de la protéine CviR dans PDB.....	51
Figure 30 : Comment rechercher la structure de la protéine CviR dans UNIPROT	51
Figure 31 : Cartographie du gène <i>VioA</i>	54

Liste des figures

Figure 32 : Cartographie du gène <i>VioB</i>	54
Figure 33 : Cartographie du gène <i>VioC</i>	55
Figure 34 : Cartographie du gène <i>VioD</i>	55
Figure 35 : Cartographie du gène <i>VioE</i>	55
Figure 36 : A : Structure 3D de la protéine régulatrice CviR lié à son auto-inducteur chez <i>Chromobacterium violaceum</i> 12472 prise à partir du PDB. B : Structure 3D de la protéine régulatrice CviR lié à son auto-inducteur chez <i>Chromobacterium violaceum</i> ATCC12472 prise à partir de l'UNIPROT.	56
Figure 37 : Séquences produisant des alignements CviR significatifs.....	57
Figure 38 : Séquences produisant des alignements <i>VioA</i> significatifs	58
Figure 39 : Séquences produisant des alignements <i>VioB</i> significatifs	59
Figure 40 : Séquences produisant des alignements <i>VioC</i> significatifs	60
Figure 41 : Séquences produisant des alignements <i>VioD</i> significatifs.....	61
Figure 42 : Séquences produisant des alignements <i>VioE</i> significatifs	62
Figure 43 : Arbre phylogénétique de la protéine CviR	63
Figure 44 : Arbre phylogénétique du gène <i>VioA</i>	63
Figure 45 : Arbre phylogénétique du gène <i>VioB</i>	64
Figure 46 : Arbre phylogénétique du gène <i>VioC</i>	64
Figure 47 : Arbre phylogénétique du gène <i>VioD</i>	65
Figure 48 : Arbre phylogénétique du gène <i>VioE</i>	65

Liste des tableaux

Tableau 1 : Taxonomie de <i>Chromobacterium violaceum</i>	4
Tableau 2 : Résultats biochimiques de la bactérie <i>Chromobacterium violaceum</i>	6
Tableau 3 : Caractéristiques générales du génome de <i>Chromobacterium violaceum</i>	10
Tableau 4 : Sensibilités aux antibiotiques chez <i>Chromobacterium violaceum</i>	12
Tableau 5 : Gènes d’opéron <i>Vio</i>	15
Tableau 6 : Divers traits régulés par QS de <i>Chromobacterium violaceum</i>	31
Tableau 7 : Les dates clés de l’essor de la bio-informatique.	34
Tableau 8 : Les disciplines de la bioinformatique et leurs utilités.	43
Tableau 9 : Informations générales sur NCBI.....	46
Tableau 10 : Type et fonction du BLAST utilise dans notre travail.	52

Table Des Matières

Introduction	1
Partie I : Synthèse bibliographique	3
Chapitre I : Généralité sur <i>Chromobacterium violaceum</i>	4
I. Généralités sur <i>Chromobacterium violaceum</i>	3
1. Historique et classification	3
2. Écologie et mode de transmission	4
3. Caractéristiques morphologiques et culturales	4
4. Caractéristiques biochimiques	6
5. Métabolisme secondaire chez <i>C. violaceum</i>	7
6. Caractéristiques génomiques	10
7. Pathogénicité	11
8. Sensibilité aux antibiotiques	12
Chapitre II : <i>Chromobacterium violaceum</i> : Quorum sensing, biosynthèse de violacéine et ces applications.	14
I. La violacéine : Caractéristiques, biosynthèse et intérêts pratique :	14
1. La violacéine	Error! Bookmark not defined.
2. Biosynthèse de la violacéine	14
3. Voie biosynthétique de la violacéine	16
4. Activités biologiques	20
4.1. Potentiel antimicrobien	20
4.2. Potentiel antitumoral	21
4.3. Potentiel immunomodulateur	21
5. Activités industrielles	22
6. Activités agricoles	23
II. Généralités sur le Quorum sensing (QS)	23

1.	Découverte et définitions.....	23
2.	Notion et mécanismes	24
3.	Le quorum sensing chez les bactéries	25
3.1.	Bactéries Gram négatives	26
3.2.	Bactéries Gram positives	28
4.	Quorum sensing chez <i>Chromobacterium violaceum</i>	30
	<i>Chapitre III : Bio-informatique (in silico)</i>	32
I.	Généralités sur la bio-informatique.....	32
1.	C'est quoi la bioinformatique ?	32
2.	Historique.....	32
3.	Objectifs.....	35
4.	Méthodes de la bio-informatique	36
4.1.	Classification.....	36
4.2.	Clustering.....	36
4.3.	Prédication	37
4.4.	Règle d'association	38
5.	Les outils de la bio-informatique.....	38
5.1.	Fichiers et formats.....	40
5.2.	Les algorithmes en bio-informatique	40
6.	Les bases de données biologiques	42
7.	Les disciplines de la bioinformatique	43
8.	Les applications de la bioinformatique.....	43
8.1.	La médecine personnelle.....	43
8.2.	Gestion des déchets	43
8.3.	Industries laitières	44

8.4.	Agriculture.....	44
<i>Partie II : Partie expérimentale (Analyse in silico)</i>		45
<i>Chapitre I : Matériels et méthodes</i>		46
1.	Matériels.....	45
1.1.	National Center for Biotechnology Information NCBI	45
1.2.	Protein Data Bank (PDB)	46
1.3.	UniProt	47
1.4.	Séquences génétiques et protéiques	47
1.5.	Structure	50
1.6.	Basic Local Alignment Search Tool (BLAST).....	52
2.	Méthodes	53
2.1.	Etude phylogénétique.....	53
<i>Chapitre II : Résultats et Discussion</i>		54
1.	Résultats	54
1.1.	Cartographie génétique (localisation)	54
1.2.	Structure	56
1.3.	Bilan BLAST.....	56
1.4.	Arbres phylogénétiques	62
2.	Discussion	66
<i>Conclusion</i>		71

Introduction

Chromobacterium violaceum est un bacille à Gram-négatif appartenant aux β protéobactéries, mobile et saprophyte. Cette bactérie vit dans le sol et l'eau des régions tropicales et subtropicales et a été reconnue comme un pathogène opportuniste des humains et des animaux (Richard, 1993). Elle est à l'origine d'une mortalité grave associée à des infections telles que les bactériémies et les abcès. Elle est également résistante à plusieurs antibiotiques, provoquant ainsi des infections récurrentes (Bali et *al.*, 2019).

Le génome de cette bactérie fût le premier séquencé par le réseau brésilien de séquençage. La séquence complète de son génome a révélé des voies alternatives très variées pour la production d'énergie, des systèmes complexes et étendus pour l'adaptation au stress, la motilité et l'utilisation généralisée du quorum sensing pour le contrôle de ses systèmes inductibles. Toutes ces découvertes démontrent la versatilité et l'adaptabilité de l'organisme (Konzen et *al.*, 2006).

Une des principales caractéristiques de ce microorganisme est la capacité de produire un pigment violet insoluble dans l'eau, connu sous le nom de violacéine (Konzen et *al.*, 2006 ; Orohoshi et *al.*, 2010), qui est responsable de la coloration violette intense, caractéristique des colonies de *C. violaceum*. La violacéine est un composé indolocarbazole naturel formé par la condensation de deux molécules de tryptophane, considérée un nouveau réactif pharmaceutique potentiel en raison de ses propriétés biologiques étendues telles que ses activités antitumorales, antibactériennes, antivirales, antioxydantes, trypanocides, et génotoxiques (Sun et *al.*, 2016).

La production de ce pigment chez *C. violaceum* est contrôlée par le quorum sensing (QS), un processus de communication bactérienne cellule-cellule dans lequel les cellules régulent la transcription des gènes spécifiques responsables de la production d'antibiotiques, de la différenciation du biofilm, de la division cellulaire, de la bioluminescence et d'autres processus (Bali et *al.*, 2019). Les gènes de biosynthèse de la violacéine chez *C. violaceum* sont au nombre de quatre *VioA*, *VioB*, *VioC*, *VioD* ; regroupées dans l'opéron *Vio*. Ce dernier possède un système de détection du quorum composé de CviI/CviR, un homologue de LuxI/LuxR (Dessaux et *al.*, 2004 ; Stauff et *al.*, 2011).

En consultant les publications scientifiques, nous avons constaté une évolution, sans précédent, avec les progrès technologiques récents, en fournissant une grande quantité de données, sur l'espèce *C. violaceum* et les voie de biosynthèse utilisé pour la production du pigment violacéine (Diniz et *al.*, 2017). Même avec toutes ces connaissances, les mécanismes ne sont pas connus de façon très détaillés vu qu'il y a une masse de donnée très importante, d'où la nécessité

Introduction

d'accompagner toutes les expérimentations *in vitro* et *in vivo* par des analyses *in silico*. Cette dernière est devenue une méthode d'investigation et un outil indispensable aux biologistes, car en bio-informatique, le développement d'outils (statistiques et informatiques) est nécessaire pour aider à comprendre les mécanismes sous-jacents aux questions biologiques (Diniz et *al.*, 2017). En biologie, le principal défi est de donner un sens à l'énorme quantité de données, c'est donc L'ère d'une « nouvelle biologie » émergente accompagnée par la naissance/le développement d'autres sciences, telles que la bio-informatique (Diniz et *al.*, 2017)

En plus de notre consultation et analyse des informations bibliographiques sur *C. violaceum*, issues des publications scientifiques, le principal objectif de notre travail consiste à réaliser en premier lieu une recherche bibliographique ciblée surtout sur les processus biologiques du quorum sensing et de la biosynthèse de la violacéine. L'intérêt est focalisé sur une analyse bioinformatique (*in silico*) sur la protéine de la transcription CviR et les gènes de l'opéron *VioABCDE* chez *C. violaceum* ATCC12472, suivie d'une prédiction de structure de protéine en 3D. Enfin, nous tentons d'analyser les arbres phylogénétiques de la protéine CviR et les gènes *VioA*, *VioC*, *VioC*, *VioD* et *VioE*, pour des fins de comparaisons moléculaires avec d'autres Procaryotes et/ou Eucaryotes dans un but d'étude des potentialités biotechnologiques. Cette étude *in silico* a été réalisée principalement en ligne, en sollicitant en ligne la base de données numérique et informatisée NCBI (National Center for Biotechnology Information) : (www.ncbi.nlm.nih.gov) et les banques de données PDB (www.rcsb.org/pdb) et UNIPROT (UniProt).

Partie I :

Synthèse bibliographique

Chapitre I :

Généralité sur

Chromobacterium violaceum

I. Généralités sur *Chromobacterium violaceum*

1. Historique et classification

L'espèce bactérienne *Chromobacterium violaceum* est composée du mot grec *Chromo* (couleur), *bacterium* (baktêrion « petit bâton ») et *violaceum* signifie la couleur violette en latin.

- En 1976, deux colonies de bactéries différentes : une blanche et une violette ont été isolées dans une station de traitement de l'eau de la ville de Manaus (Amazonie-Brésil) (Caldas et al., 1978 in Durán et al., 2001). Le professeur Wilson Chagas de Araujo, de l'Institut de microbiologie de l'Université fédérale de Rio de Janeiro a identifié la violette comme *Chromobacterium violaceum*. C'était la première fois que ce microorganisme est isolé au Brésil (Durán et al., 1980 in Durán et al., 2001).

- En 1880, Dans une étude indépendante, Bergonzini fit une découverte accidentelle alors qu'il travaillait à l'Université de Modène en Italie. Vers la fin de mars 1880, il préparait des solutions d'ovalbumine pour étudier les « mécanismes d'action qui causent le retard de la putréfaction ». Après les expériences, il a oublié d'éliminer les solutions de contrôle et en fin d'avril il a observé un aspect singulier des échantillons. Le chercheur italien a réduit le volume par évaporation et a obtenu une solution recouverte d'un fin film très dense de couleur violette. Au début, il soupçonnait *Cromococcus violaceus*, qui était la seule bactérie connue à présenter cette coloration ; toutefois, cette possibilité a été éliminée en raison de l'insolubilité de tous les composants. Bergonzini après quelques tests a conclu qu'il s'agissait d'une nouvelle bactérie qu'il a nommé *Cromobacterium violaceum* (sans ch) par la suite il a publié en 1881 la découverte d'une "nouvelle bactérie colorée" (Bergonzini, 1881). Sur la base des spectres d'absorption visible enregistrés par Boisbaudran, De Moss a indiqué qu'il s'agit bien de la *Chromobacterium violaceum* et de la violacéine, respectivement (Moss, 1967 in Durán et al., 2001).

- En 1881, ZIMMERMAN a corrigé l'orthographe de *Cromobacterium*, donnée par BERGONZINI, en *Chromobacterium* (Zimmerman, 1967 in Durán et al., 2001). À cette époque, *Bacillus violaceus* n'était plus utilisé. Aujourd'hui, le genre *Chromobacterium*, décrit par SNEATH, est celui qui se trouve dans le Manuel de bactériologie systématique de Bergey (Holt et Krieg, 1984 in Durán et al., 2001).

- En 1882, *Comptes Rendus de l'Académie des sciences* a publié un article de Boisbaudran, qui décrivait une coloration violette sur des préparations de fleurs de riz que l'auteur attribuait à un « petit organisme » faisant référence à un microorganisme (Boisbaudran, 1882 in Durán et al., 2001).

La classification de cette bactérie est résumée selon le tableau ci-dessous (Tableau 1).

Tableau 1 : Taxonomie de *C. violaceum* (Bergonzini, 1880 in Durán et al., 2001).

Domaine	Bacteria
Phylum	Proteobacteria
Classe	Betaproteobacteria
Ordre	Neisseriales
Famille	Neisseriaceae
Genre	Chromobacterium
Espèce	<i>Chromobacterium violaceum</i>

2. Écologie et mode de transmission

C. violaceum est un composant abondant du microbiote du sol et de l'eau dans les régions tropicales et subtropicales du monde entier (Batista et al., 2017).

Les modes de transmission sont l'ingestion d'eau contaminée, l'exposition de plaies et de lésions traumatiques au sol et à l'eau contaminés, l'insertion d'une sonde urinaire ou l'utilisation de différents équipements médicaux en milieu hospitalier (Alim et al., 2022).

3. Caractéristiques morphologiques et culturelles

C. violaceum est une espèce de bactéries qui sont des petits bâtonnets ou des coccobacilles Gram-négatifs, non sporulants, non acidogènes. Ils ont une taille de 0,6-0,9µm x 1,5-3,0µm et présentent une coloration bipolaire (figure 2). La motilité de *Chromobacterium violaceum* est

obtenue au moyen d'un seul flagelle polaire et de jusqu'à quatre flagelles latéraux antigéniquement et structurellement distincts (Tucker et *al.*, 1979).

Les *C. violaceum* sont des anaérobies facultatifs dont la plage de croissance s'étend de 15 à 40°C. La croissance optimale est obtenue à 30-35°C. Ils produisent de manière caractéristique des colonies violettes sur la gélose nutritive et se développent généralement sur la gélose MacConkey (figure 1). Les colonies se développent facilement sur les milieux de culture courants, qui conviennent à l'isolement des souches à partir de spécimens de patients. Cependant, ce ne sont pas les meilleurs milieux pour isoler ces organismes du sol ou de l'eau car ils ne sont présents qu'en petit nombre. Les grains de riz stériles et la gélose nutritive à concentration réduite ont été utilisés avec succès pour les isoler du sol ou de l'eau. Les colonies sont faiblement convexes, violettes, lisses et non gélatineuses. Le pigment produit par la plupart des souches est la violacéine, qui donne aux colonies un éclat métallique distinctif, violet foncé (Tucker et *al.*, 1979), (Dyer et *al.*, 1998).



Figure 1: Croissance de *C. violaceum* sur gélose MacConkey (à gauche) et sur gélose au sang de mouton à 5% (à droite) (Cristina et *al.*, 2005).



Figure 2 : *C. violaceum* sous microscope optique Gx40 (James et *al.*, 2011).

4. Caractéristiques biochimiques

La caractérisation biochimique de *chromobacterium violaceum* a été effectuée par des tests biochimiques conventionnels tels que le test de la catalase, le test de l'oxydase, le test de l'agar-fer à triple sucre, l'utilisation du citrate, l'hydrolyse de l'urée, la motilité, le test d'oxydation-fermentation, le test du rouge de méthyle de Voges-Proskauer, les tests de décarboxylase (lysine, arginine et ornithine) (Tableau 2).

Tableau 2 : Résultats biochimiques de la bactérie *Chromobacterium violaceum* (Parajuli et al., 2016).

Tests biochimiques effectués	Résultats
Coloration de Gram	Bacilles à Gram négatif
Test de catalase	Positive
Test à l'oxydase	Positive
Production de gaz	Négative
Production de sulfure d'hydrogène	Négative
Fermentation du glucose	Positive
Fermentation du lactose et/ou du saccharose	Négative
Motilité	Positive
Production de sulfure d'hydrogène	Négative
Indole	Négative
Test d'hydrolyse de l'urée	Négative
Test d'utilisation du citrate	Positive
Test du rouge de méthyle	Négative
Test de Voges-Proskauer	Négative
Lysine décarboxylase	Négative
Arginine-dihydrolase	Positive
Ornithine décarboxylase	Négative
Test ADNse	Positive
Pigment	Violet (violacéine)

5. Métabolisme secondaire chez *Violaceum*

Comme mentionné depuis 1867, la couleur du pigment produit par l'une des souches de *C. violaceum*, la violacéine, est une grande attraction. Reilly et Pyne ont été les premiers à étudier en détail la structure chimique du pigment violet à partir du début de la décennie 1930. Depuis cette époque, plusieurs propositions concernant les structures chimiques de la violacéine sont apparues, qui se sont toutes avérées incorrectes par la suite. C'est à l'Université de Liverpool en 1958, par des études de dégradation et par synthèse, que la structure de ce composé a été corrigée (Figure 3), (Ballantine et al., 1958 et 1960).

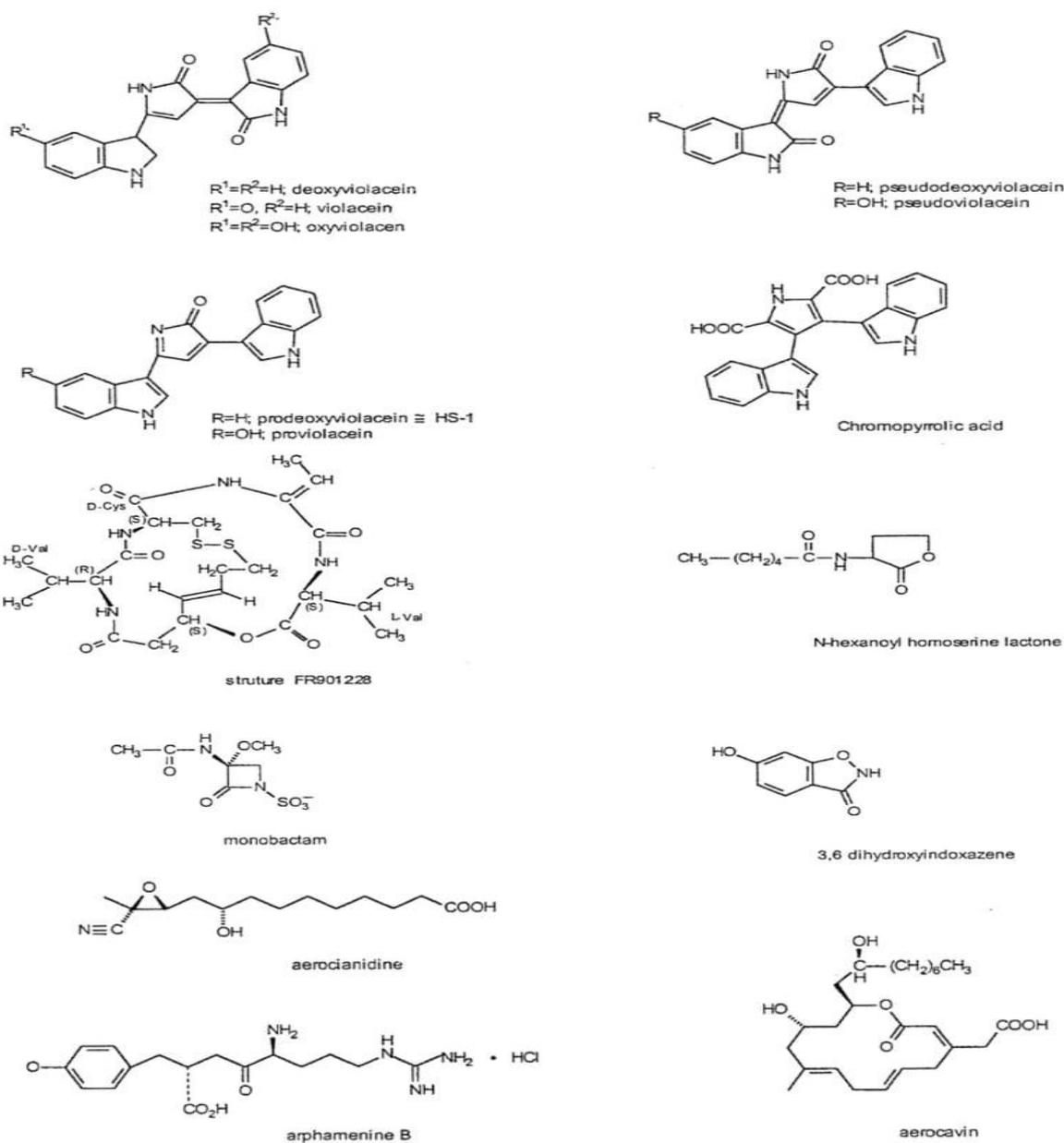


Figure 3 : La structure chimique des métabolites de *Chromobacterium violaceum* (Ballantine et al., 1958 ; Ballantine et al., 1960).

squelette de la violacéine, et ils ont démontrés que la liaison C-C de C2 de l'anneau naturel à C2 de la chaîne 16 latérale a été complètement conservé pendant la formation de la partie droite (côté oxindole) pendant l'ensemble du processus de biosynthèse (Figure 5) (Hoshino et *al.*,1986).

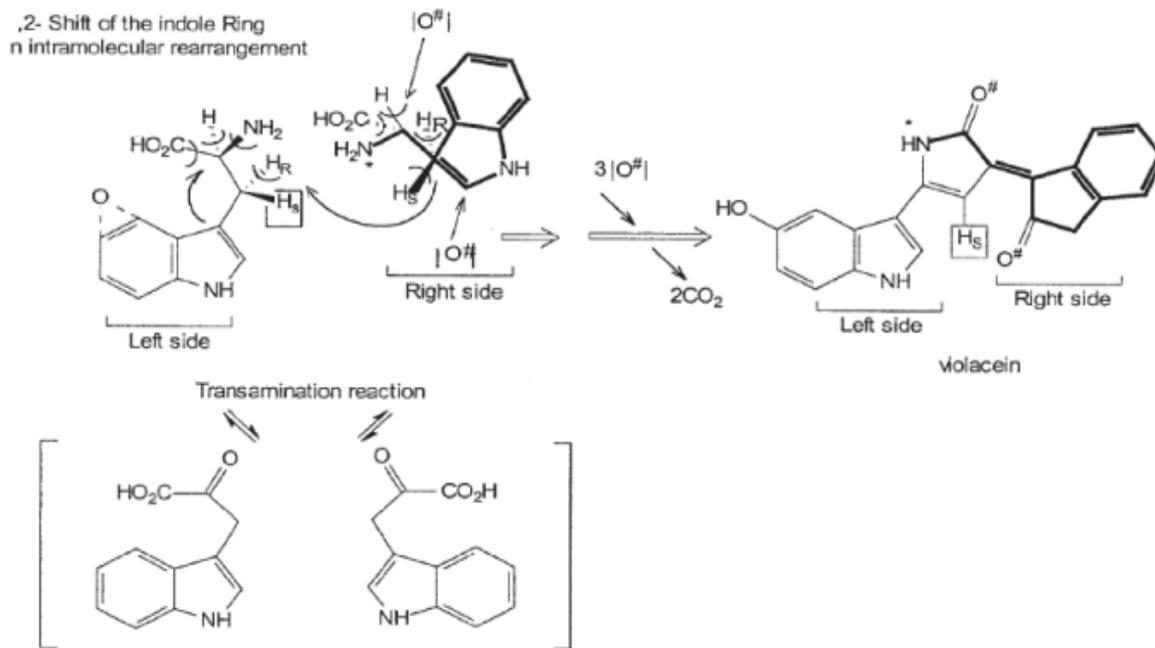


Figure 5 : Schéma résumé de la biosynthèse de la violacéine (Momen et *al.*,2000).

En 1989, des études ont commencé sur d'autres paramètres importants de la biosynthèse de la violacéine. Durán et *al.*, dans le cadre d'une étude portant sur radio-isotope, a suggéré que, outre le L-tryptophane, *Chromobacterium violaceum* est capable de synthétiser la violacéine à partir de l'acide indol-3-acétique en tant que métabolite précurseur du L-tryptophane, et a proposé un intermédiaire amidique. Dans cette étude, la méthode de production et la purification de désoxyviolacéine libre de violacéine ont été définies (Durán et *al.*, 2001).

D'autres études ont révélé que *Chromobacterium violaceum* présente un groupe d'enzymes capables de synthétiser des composés structurellement similaires à la violacéine, tels que l'oxyviolacéine, le HS-1 (prodéoxyviolacéine) et la pseudodésoxyviolacéine (Figure 3), à partir de substrats autres que le L-tryptophane. Les lactones acétylées de N- (homosérine) induisant la production de violacéine ont été décrites comme régulateurs génétiques dans des populations à haute densité dans une variété de systèmes bactériens à Gram négatif et est appelé détection de quorum (Durán et *al.*, 2001).

6. Caractéristiques génomiques

Le génome complet de *Chromobacterium violaceum* (figure 6) consiste en un seul chromosome circulaire de 4 751 080 pb avec une teneur moyenne en G+C de 64,83 % (tableau 3). Il y a 4 431 ORF (Open Reading Frame) codant pour des protéines, répartis uniformément, qui couvrent 89 % du génome et ont une longueur moyenne de 954 pb. Parmi ceux-ci, 2 717 (61,3 %) ont pu se voir attribuer des fonctions putatives, tandis que 958 (21,6 %) ont été identifiés comme des protéines hypothétiques conservées. Les 756 restants (17,1 %) ont été désignés comme des protéines hypothétiques. Parmi les ORF hypothétiques conservés, 499 ont des motifs protéiques contenus à la fois dans interpro et cog, tandis que 242 ont des motifs contenus dans l'un ou l'autre. Parmi les ORF hypothétiques, 68 ont des motifs contenus dans les deux et 135 dans une seule des deux bases de données. Sur les 131 familles paralogues, 111 (84,7 %) contiennent deux membres, mais certaines contiennent jusqu'à six ORF. Les fonctions d'environ un tiers des familles sont liées au transport, et environ un quart ont des fonctions inconnues. Il y a 98 gènes d'ARNt représentant tous les 20 acides aminés et 8 opérons d'ARNr qui sont identiques dans leur région codante, bien que 6 contiennent un insert de 100 pb dans la région intercalaire. L'origine probable de la réplication est identifiable sur la base du skew (l'inclinaison) G+C et des positions de dnaA, dnaN, et gyrA (Carraro et *al.*, 2003).

Tableau 3 : Caractéristiques générales du génome de *C. violaceum* (Carraro et *al.*, 2003).

Longueur pb	4, 751,080
Contenu G+C	64.83%
Nombre total des ORF	4,431
Pourcentage du génome constituant des régions codantes	89%
Longueur moyenne des ORF, pb	954
Nombre des protéines connues	2,717
Nombre des protéines hypothétiques conservées	958
Nombre des protéines hypothétiques	756
ARNr	8 × (16S-23S-5S)
ARNt	98

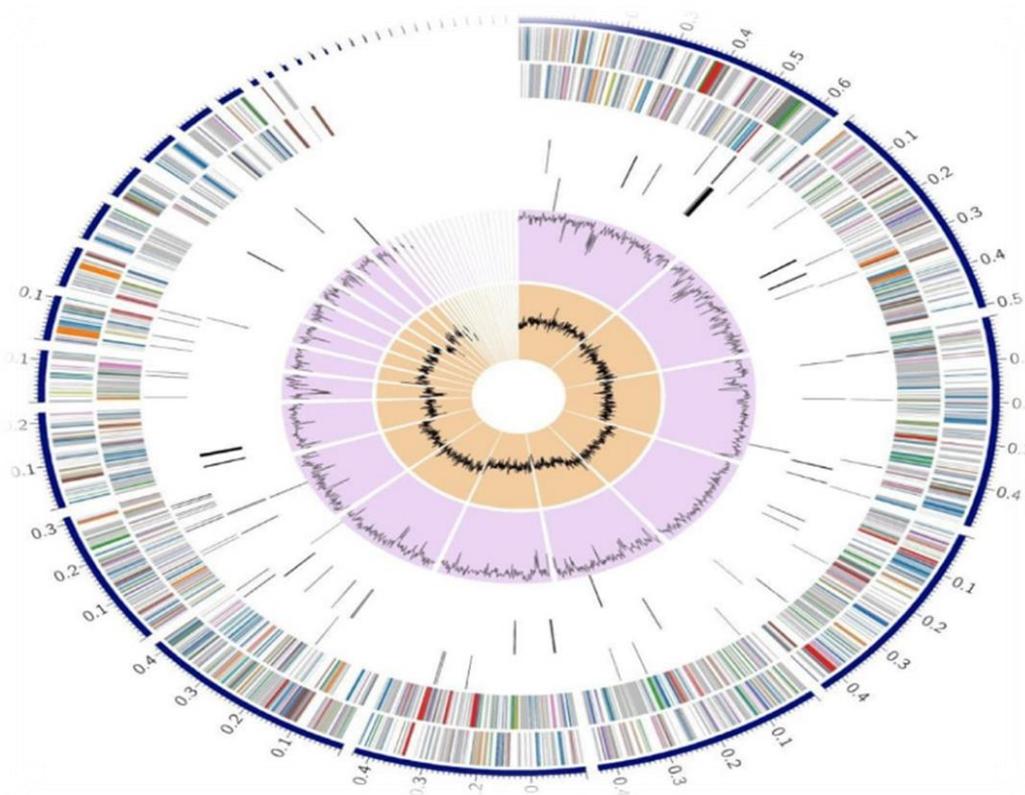


Figure 6 : Représentation graphique circulaire du génome de *Chromobacterium violaceum* (Subbiah et al., 2021).

7. Pathogénicité

C. violaceum se définit comme une bactérie saprophyte, normalement considérée comme non pathogène pour l'homme, mais à l'occasion peut devenir un pathogène opportuniste responsable de certaines infections graves et parfois mortelles chez l'homme qui ont été signalées en Argentine, en Australie, au Brésil, à Cuba, au Nigeria, à Singapour, à Taiwan, aux États-Unis et au Vietnam. *C. violaceum* peut également causer des infections chez les animaux (porcs, singes, moutons, et des chiens) (Durán et al., 2001).

Bien que *C. violaceum* soit fréquemment trouvé dans la nature et soit à l'origine d'infections graves chez l'homme, l'incidence est très faible. Après une analyse et un diagnostic des individus infectés, *Chromobacterium violaceum* est donc déclaré comme un agent pathogène de faible degré susceptible de provoquer 7 des infections graves principalement chez les patients immunodéprimés. Cependant, en accord avec des études récentes, cette généralisation est discutable (Durán et al., 2001).

À l'occasion, elle peut agir comme un agent pathogène opportuniste chez les animaux et les humains et provoquer une fatale septicémie liée à des lésions cutanées présentant de nombreuses lésions hépatiques et abcès du poumon (Durán et *al.*, 2001).

En outre des cas de granulomatose chronique associés à *C. violaceum* d'adénite en tant que complication de la granulomatose chronique et de problèmes associés à *C. violaceum* dans le cadre d'une ostéomyélite, d'une cellulite péri orbitale et d'une infection oculaire ont été reportés. Tous les cas mentionnés jusqu'à présent impliquent une bactérie pigmentée, mais il y a aussi des cas d'infections causées par une souche non pigmentée. Cela implique que le pouvoir pathogène de cette bactérie est peut être indépendant de la violacéine (Durán et *al.*, 2001).

8. Sensibilité aux antibiotiques

Le test de susceptibilité aux antibiotiques de l'isolat a été effectué par des méthodes de diffusion en disque d'agar (Figure 7).

Les résultats ont révélé que *C. violaceum* est résistant à l'ampicilline, polymixine B, Sulfate de colistine, l'ampicilline sulbactam, Tobramycine, Céfépime et Ceftriaxone Est sensible à Pipéracilline/tazobactam, Céfotaxime, Amikacine, Gentamicine, Ciprofloxacine, Levofloxacine et Triméthoprime/sulfaméthoxazole (tableau 4). (Kim et *al.*, 2005).

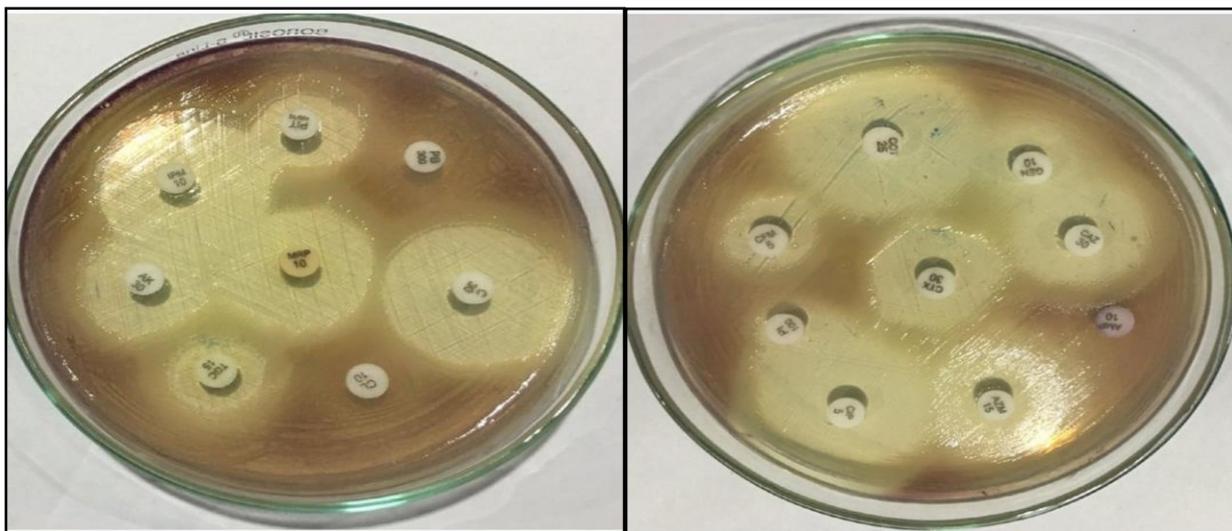


Figure 7 : Test de sensibilité aux antimicrobiens de *C. violaceum* (Parajuli et *al.*, 2016).

Tableau 4 : Sensibilités aux antibiotiques chez *C. violaceum* (Kim et al., 2005).

Antibiotiques	Résultats
Ampicilline	Résistant
Ampicilline/Sulbactam	Résistant
Pipéracilline/tazobactam	Sensible
Céfotaxime	Sensible
Ceftriaxone	Résistant
Céfépime	Résistant
Imipénème	Intermédiaire
Amikacine	Sensible
Tobramycine	Résistant
Ciprofloxacine	Sensible
Levofloxacine	Sensible
Triméthoprim/sulfaméthoxazole	Sensible
Sulfate de colistine	Résistant
Gentamicine	Sensible

Chapitre II :

Chromobacterium violaceum :
Quorum Sensing, biosynthèse de
violacéine et ses applications.

I. La violacéine : Caractéristiques, biosynthèse et intérêts pratique :

1. Caractéristiques

La violacéine est un pigment pourpre naturel de *Chromobacterium violaceum* (Inchagova et al., 2019) (Füller et al., 2016), Celle-ci est polyvalent bleu-violet, insoluble dans l'eau (Kothari et al., 2017) et responsable de la coloration violette intense, caractéristique des colonies de *Chromobacterium violaceum* (Dessaux et al., 2004). Elle est composée de trois unités structurales (figure 8), c'est-à-dire un 5-hydroxyindole, un oxindole et une 2-pyrrolidone : deux anneaux indoles sont impliqués, donc ce pigment est également appelé un composé bisindole (Hoshino, 2011).

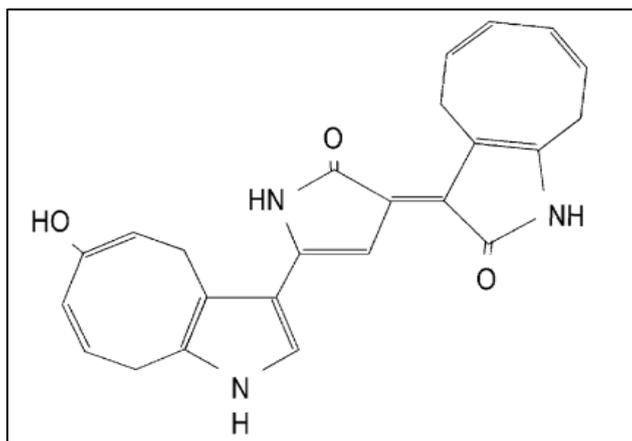


Figure 8 : Structure chimique de violacéine qui dérive de deux molécules de tryptophane modifiées. Il est constitué de trois noyaux principaux : du 5-hydroxyindole (à gauche), de l'oxindole (en haut, à droite) et de la 2-pyrrolidone (au milieu) (Dessaux et al., 2004).

2. Biosynthèse de la violacéine

La production de violacéine implique l'expression de l'opéron *Vio* (tableau 5) consistant en cinq gènes contigus codant des enzymes (*VioA*, *VioB*, *VioC*, *VioD* et *VioE*) qui sont transcrits dans une direction. Ces enzymes sont codées par un fragment d'ADN de 7,3 ko dans le format numérique (format FASTA), Il a été démontré que les enzymes *VioA*, *VioC* et *VioD* sont des monooxygénases dépendantes du FAD (Flavine Adénine Dinucléotide) (Balibar et al., 2006 ; Shinoda et al., 2007 in Kothari et al., 2017). *VioB* est une protéine hémique contenant Fe^{2+} , considérée comme une polykétide synthase (August et al., 2000 in Kothari et al., 2017).

Tableau 5 : Gènes d’opéron *Vio* (Kothari et al., 2017).

Gène	Action Enzymatique
<i>VioA</i>	Tryptophane 2-monooxygénase dépendant du flavin
<i>VioB</i>	Considéré comme une synthase de polykératide contenant de la protéine hémique
<i>VioC</i>	Monooxygénase dépendante du FAD
<i>VioD</i>	Monooxygénase dépendant de la flavine
<i>VioE</i>	Responsable de la conversion du flavanone en isoflavone

Différents rapports attachent une importance variable à différentes enzymes de la voie de biosynthèse de la violacéine (figure 9). Bien que *VioE* ait été montré comme l’enzyme clé par Hirano et al. Car il peut convertir rapidement le dimère de courte durée (IPA) en PVA qui permet la biosynthèse continue de la violacéine ; d’autres rapports ont montré *VioB* comme enzyme limitant la vitesse. Dans ce cas, l’augmentation de la production de violacéine semble provenir d’une expression plus élevée des gènes participant au métabolisme du glucose par voie phosphatée de pentose, ce qui donne lieu à une plus grande disponibilité de 4-phosphate érythrocytaire, à utiliser pour la biosynthèse du tryptophane, qui est le précurseur de la synthèse de la violacéine (Kothari et al., 2017).

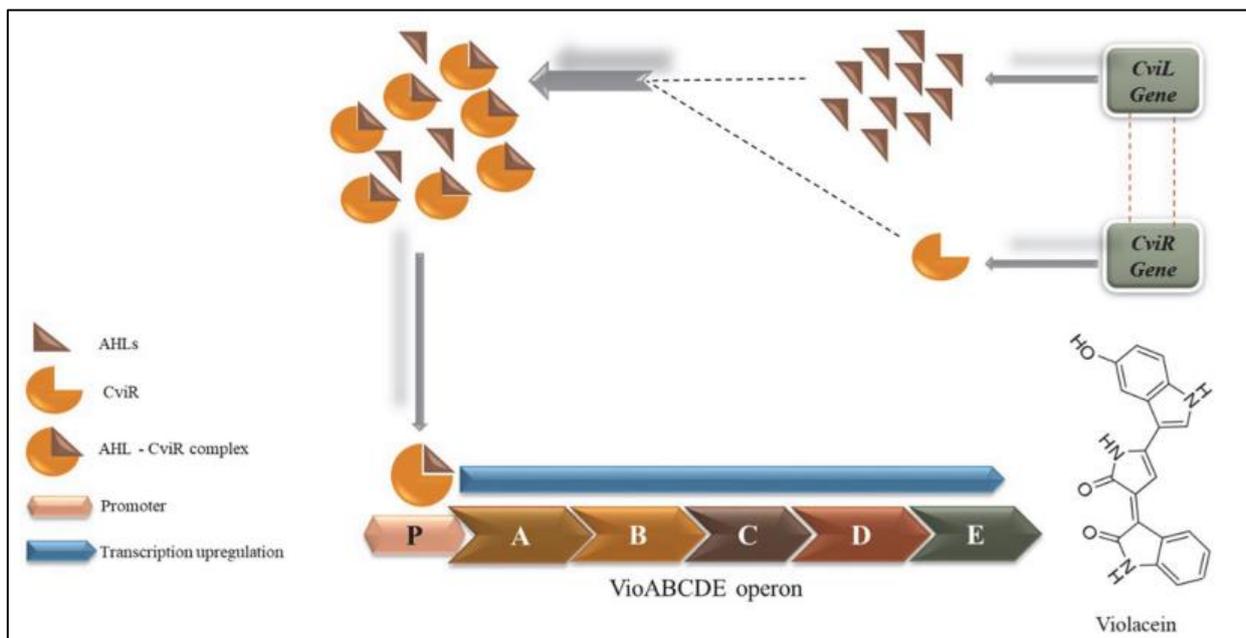


Figure 9 : Schéma de biosynthèse de la violacéine. La liaison du complexe CviR-AHL au site promoteur du *VioA* déclenche l’expression du *Vio* opéron, conduisant à la synthèse du pigment violet violacéine (Ahmed et al., 2021).

3. Voie biosynthétique de la violacéine

Le squelette de carbone, d'hydrogène et d'azote de la violacéine est entièrement dérivé de deux molécules de L-tryptophane et les atomes d'oxygène sont dérivés de dioxygène. La première enzyme dans la voie, VioA, fonctionne comme un acide L-amino oxydase dépendant du flavin qui catalyse la production de deux molécules de 2-imino-3-(indol-3-yl) propanoate (IPA imine) à partir de deux molécules de L-tryptophane. Cette réaction se retrouve également dans les voies de biosynthèse antibiotique staurosporine biosynthèse et K-252 biosynthèse, chez Actinobacteria (Fulcher, 2020).

Les réactions clés dans la voie convertissent le 2-imino-3-(indol-3-yl) propanoate en protodésoxyviolacéinate et sont catalysées par le deuxième et troisième enzymes VioB et VioE. Le produit de VioB, qui est le substrat de VioE, reste spéculatif, car il est trop labile pour être isolé (Fulcher, 2020).

VioB est une enzyme contenant de l'hème Fe^{2+} . Il catalyse une réaction qui convertit le 2-imino-3-(indol-3-yl) propanoate en un intermédiaire dimérique de courte durée de vie qui résulte de la formation d'une liaison C-C entre les atomes de β -carbone de deux molécules de 2-imino-3-(indol-3-yl) propanoate. Des structures et des transitions structurales possibles pour cet intermédiaire dimérique ont été proposées. VioB a une homologie de séquence d'acides aminés élevée avec RebD et StaD dans les voies de biosynthèse de rebeccamycine et de biosynthèse de staurosporine, respectivement (Fulcher, 2020).

L'intermédiaire dimérique de courte durée de vie subit une désamination et une décarboxylation par formation d'anneaux cyclopropane. VioE a été proposé pour catalyser ces transitions ainsi que le 1,2-décalage unique de l'anneau indole vu dans son protodésoxyviolacéinate produit. VioE est une enzyme intéressante qui a peu d'homologies de séquence, et aucune exigence de cofacteur n'a été trouvée. La structure cristalline de VioE a été déterminée (Fulcher, 2020).

En l'absence de VioE, l'intermédiaire dimérique de courte durée formé à partir du 2-imino-3-(indol-3-yl) propanoate par VioB peut subir une désamination et une aromatisation spontanée pour produire du chromopyrrolate, sans le 1,2-décalage de l'anneau indole. Dans la biosynthèse de staurosporine de voie, la synthèse de chromopyrrolate est catalysée par StaD, et dans la biosynthèse de voie K-252 par NokB (Fulcher, 2020).

VioD catalyse un hydroxylation de protodésoxyviolacéinate pour produire le protoviolacéinate intermédiaire. VioC hydroxyle ensuite la position 2 de l'anneau indole opposé

du protoviolacéinate pour produire du violacéinate. En l'absence de VioD, VioC peut oxyder le protodésoxyviolaceinate en désoxyviolaceinate. Il a été démontré que VioC et VioD étaient des oxygénases dépendantes du NADPH et de la flavine. La réaction finale dans la voie principale qui produit la violacéine se produit par l'autooxydation non enzymatique (spontanée) et la décarboxylation. Le métabolite désoxyviolacein peut être produit de façon similaire (Fulcher, 2020).

Contrairement aux composés protodésoxyviolacéinate et protoviolacéinate qui sont vrais, mais les intermédiaires labiles de la voie, les composés prodeoxyviolacein et proviolacein sont considérés comme des artefacts. Ils sont générés à partir de protodésoxyviolacéinate et de protoviolacéinate en l'absence de VioC par autooxydation et décarboxylation non enzymatique (spontanée) (Fulcher, 2020).

Tel que noté ci-dessus, une caractéristique unique de la biosynthèse de la violacéine est le décalage indole médié par VioE qui génère du protodésoxyviolacéinate (par rapport à aucun décalage indole dans chromopyrrolate). Ce changement ne se produit pas lors de la biosynthèse de la rebeccamycine ou de la biosynthèse de la staurosporine, et un homologue VioE fait défaut dans ces groupes de gènes. Une autre caractéristique de la voie est la combinaison de réactions catalysées enzymatiques et non enzymatiques (spontanée) autooxydation et des réactions de décarboxylation. Une troisième caractéristique est que l'oxygène est incorporé dans la violacéine via VioD et VioC, ainsi qu'une étape non enzymatique finale (Figures 10 et 11) (Fulcher, 2020)

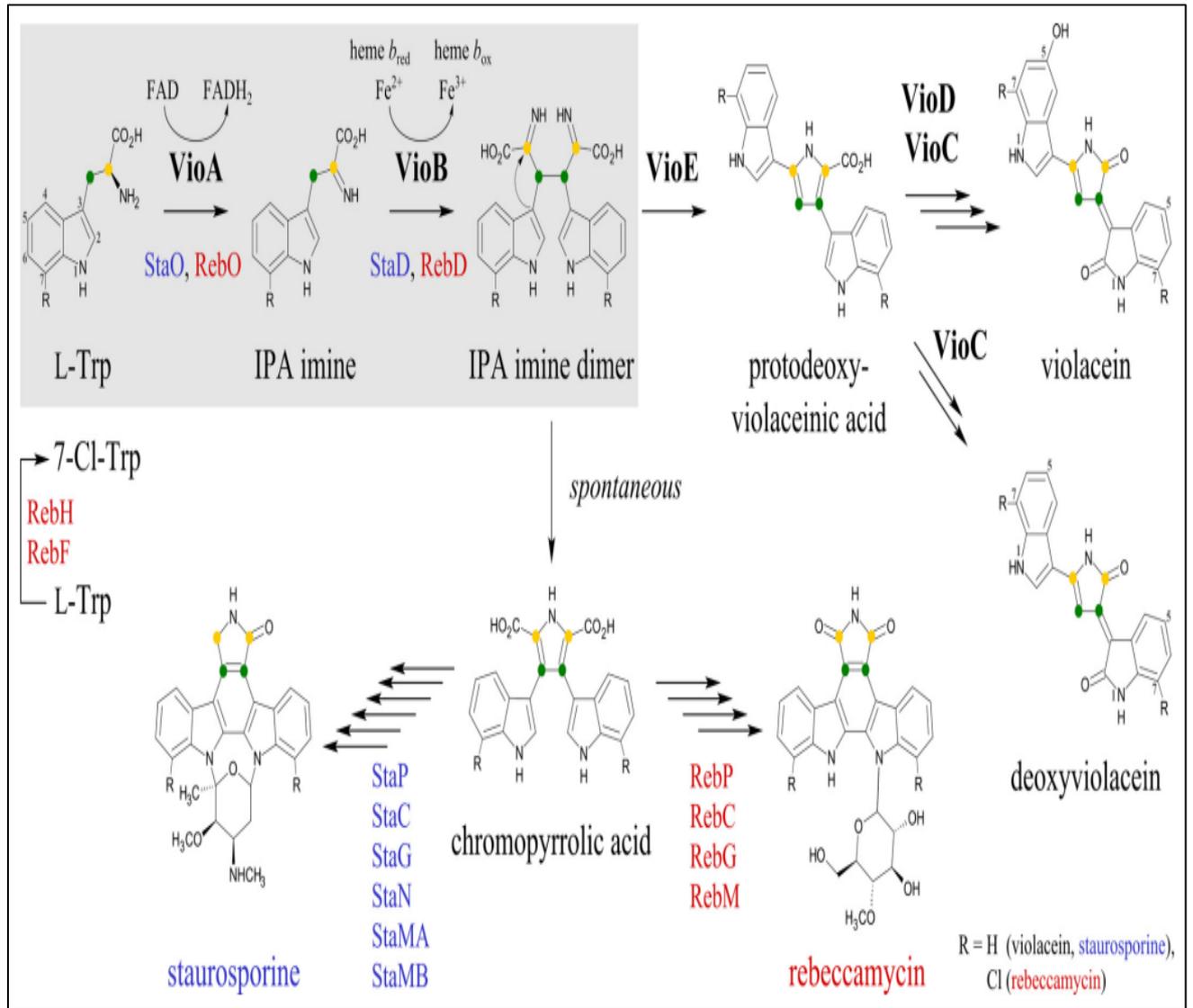


Figure 10 : Voie biosynthétique de la violacéine (noire), de la staurosporine (bleu) et de la rebeccamycine (rouge) (Füller et al., 2016).

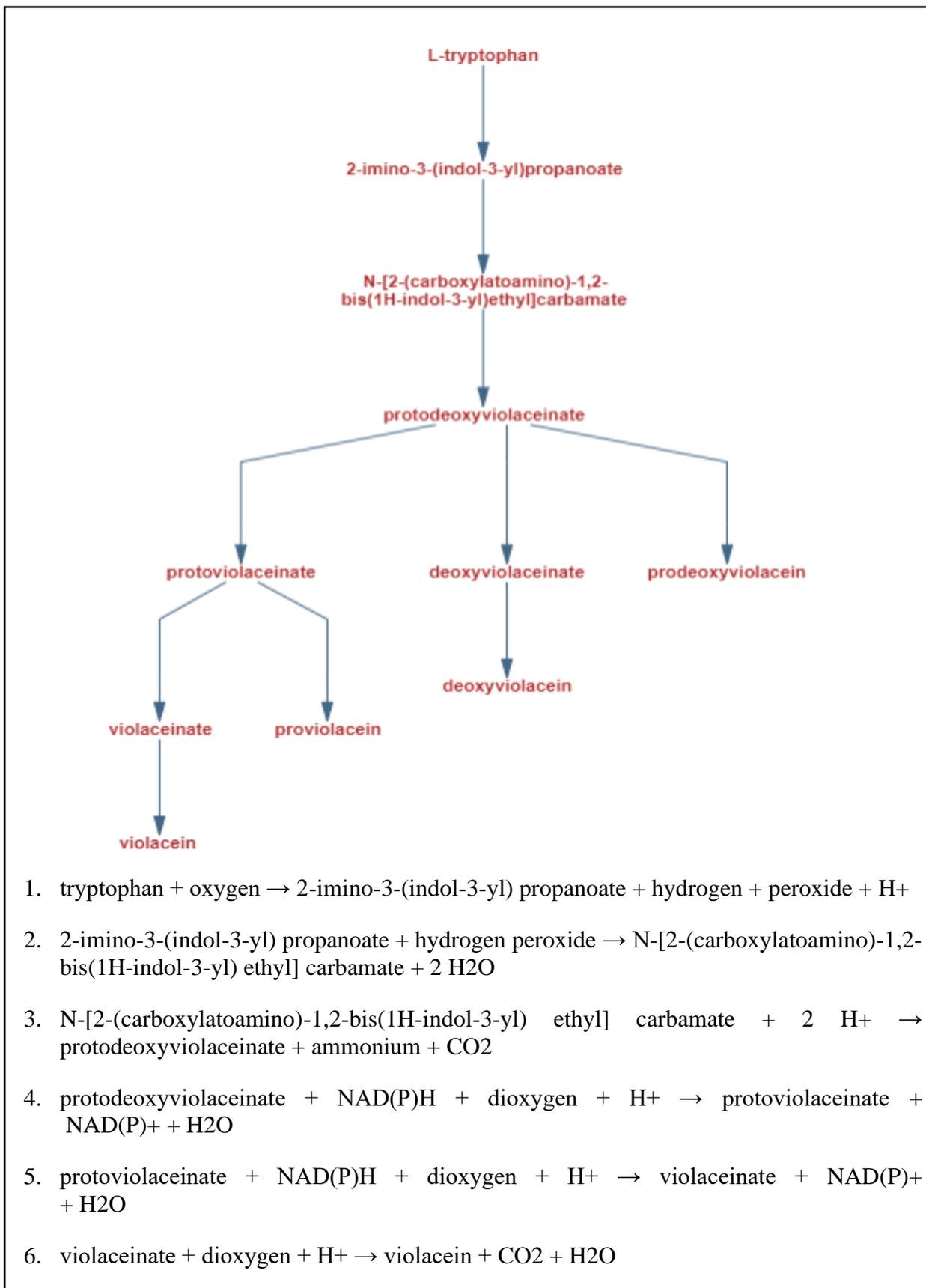


Figure 11 : Les réactions chimiques entrant dans la biosynthèse du violacéine par *Chromobacterium violaceum* (Fulcher, 2020).

4. Activités biologiques

Depuis la caractérisation de *Chromobacterium violaceum* et la purification de sa violacéine, ce dérivé indole de couleur pourpre a été étudié pour son potentiel biologique et a suscité l'intérêt de la communauté scientifique, ce qui a considérablement accru le nombre d'études dans la littérature. D'un côté, ceux qui travaillent dans le domaine des anti-infectieux recherchent une inhibition de la production de violacéine régulée par la détection du quorum (QS) chez *Chromobacterium violaceum*, ceux qui se concentrent sur la violacéine en tant que molécule bioactive cherchent à améliorer le rendement de fermentation de la violacéine (Durán et al., 2016).

La violacéine présente plusieurs activités biologiques intéressantes ayant une pertinence clinique, dont certaines ont déjà été examinées. Les premières enquêtes ont révélé son rôle en tant qu'agent antioxydant chez la bactérie, suivi de l'explication de son importance chez deux autres espèces de bactéries productrices de violacéine, *Pseudoalteromonas tunicata* et *Janthinobacterium lividum*. Ces études attirent l'attention sur un domaine d'activités de la violacéine encore inexploré, qui permettra de découvrir de nouvelles connaissances sur son rôle dans les environnements écologiques (Durán et al., 2016).

4.1. Potentiel antimicrobien

L'action antibactérienne du mélange de violacéine et de désoxyviolacéine a été démontrée contre les bactéries Gram-positives *Bacillus lichenniformis*, *Bacillus subtilis*, *Bacillus megaterium*, *Staphylococcus aureus*, et Gram-négatives *Pseudomonas aeruginosa*, à une concentration élevée du pigment (Durán et al., 2012). Il a été signalé aussi des activités antivirales contre le virus herpétique simplex et les poliovirus (Durán et al., 2016). La violacéine a été capable de provoquer la mort cellulaire (Durán et al., 2012).

L'activité anti-protozoaire de la violacéine a été étudiée contre diverses espèces de flagellés, de ciliés et d'amibes protozoaires, ainsi que contre le ver nématode bactérien *Caenorhabditis elegans*, un organisme modèle génétique de premier plan. La violacéine est toxique pour le prédateur bactérien *Caenorhabditis elegans* (Durán et al., 2016).

La violacéine a également été utilisée comme marqueur bactérien robuste pour l'imagerie photo-acoustique non invasive présentant un potentiel élevé pour la recherche fondamentale et les futures applications théranostiques dans le ciblage de tumeurs bactériennes (Durán et al., 2016).

L'utilisation potentielle de violacéine en tant qu'agent antifongique, en particulier contre les champignons dermatophytiques, *Trichophyton rubrum*, un agent pathogène humain majeur, a également été indiquée. Une équipe dirigée par Stefanie Lopes a montré que des concentrations micro-molaires de violacéine éliminaient efficacement les souches de *Plasmodium falciparum* sensibles et résistantes à la chloroquine in vitro. (Durán et al., 2016).

4.2. Potentiel antitumoral

Il est bien connu que la mort cellulaire programmée dans le cancer n'est pas bien établie pour la plupart des agents chimio-thérapeutiques et des types de cellules tumorales. À cet égard, la violacéine est un composé modèle présentant des caractéristiques de mort cellulaire spécifiques à la lignée et des mécanismes d'action (Durán et al., 2016).

La violacéine montre également des (Durán et al., 2016):

- Effets cytotoxiques contre plusieurs lignées de cellules tumorales, activité antiproliférative sur les lignées de cellules cancéreuses.
- Un effet d'apoptose (le processus d'autodestruction cellulaire) sur les cellules leucémiques.
- Inhibition de la transduction de signal dans les cellules cancéreuses du côlon humain.
- Inhibition de la croissance contre les tumeurs.

Ces activités peuvent offrir des applications cliniques prometteuses pour le traitement du cancer.

4.3. Potentiel immunomodulateur

Malgré le nombre d'études démontrant les activités biologiques de la violacéine, ce n'est qu'après 2010 que les premières preuves du potentiel immuno- modulateur de la violacéine sont apparues (Antonisamy et al., 2010 ; Antonisamy et al., 2014 ; Verinaud et al., 2015 in Durán et al., 2016). Antonisamy et Ignacimuthu (2010) ont signalé diverses activités de violacéine chez le rat, y compris des activités analgésiques, immunomodulatrices et antipyrétiques.

La violacéine administrée par voie orale a inhibé l'hypersensibilité immunitaire provoquée contre les globules rouges du mouton, probablement en modulant les médiateurs inflammatoires. De plus, ce composé a pu contrôler la réaction anaphylactique à IgE de type I. La violacéine a également induit une réponse antinociceptive centrale et périphérique. Les études impliquant l'activité analgésique de la violacéine incluent le test de formol, qui évalue son activité dans une

phase antérieure de douleur et une phase inflammatoire tardive. L'activité antipyrétique de la violacéine souligne davantage son action inhibitrice possible sur la biosynthèse de la prostaglandine (PG) (Antonisamy et al., 2010 in Durán et al., 2016).

Dans une autre étude, Antonisamy et al. (2014) ont signalé que la violacéine a des effets gastro-protecteurs importants contre les lésions gastriques induites par les AINS (Antonisamy et al., 2014).

La violacéine inhibe également la migration des neutrophiles vers la cavité péritonéale. La violacéine a entraîné une amélioration de l'évolution clinique dans l'inflammation chronique. Ces effets étaient liés à la modulation des cellules dendritiques et des médiateurs inflammatoires (Verinaud et al., 2022). Il est important de mentionner qu'aucune toxicité n'a été signalée en accord avec les rapports précédents (Bromberg et al., 2005, 2010 in Durán et al., 2016).

5. Activités industrielles

Un autre domaine important pour l'application commerciale de la violacéine est dans les additifs commerciaux pour la protection solaire de la peau humaine, en raison de ses activités antioxydantes et antimicrobiennes. La violacéine ajoutée aux écrans solaires a été signalée pour améliorer l'absorption des UV dans la gamme 290 à 320 nm. Ces caractéristiques étaient évidentes lorsque la violacéine était utilisée pour compléter les extraits de feuilles d'aloé-vera et *Cucumis sativus* (concombre), qui présentent une activité photo-protectrice.

L'industrie textile s'est également intéressée à l'application de la violacéine, car il s'agit d'un pigment pourpre. Par exemple, nous pouvons pointer les fibres teintées à la lumière qui ont été préparées en traitant les fibres teintées avec de la violacéine, de la désoxyviolacéines ou des mélanges de leurs solutions aqueuses en présence de thio-urée, produisant un tissu de soie tissé de couleur pourpre bleuâtre (Durán et al., 2012).

Les pigments bactériens offrent d'excellentes possibilités en raison de leurs effets positifs accrus sur l'environnement et de leurs caractéristiques de performance supérieures. Une étude a démontré que les mots imprimés pouvaient être mis en évidence à l'aide d'encre avec violacéine en présence d'acide citrique et de glycérol. Il s'agit d'une bonne indication que les études sur les pigments des bactéries doivent être élargies pour répondre à certains besoins industriels (Venil et al., 2013, 2015 in Durán et al., 2016).

La violacéine peut être utilisée dans une large gamme de produits de l'industrie alimentaire comme colorant dans les systèmes de modèles alimentaires, où le yogourt et la gelée ont été sélectionnés pour tester le colorant violet produit (Venil et *al.*, 2015 in Durán et *al.*, 2016).

6. Activités agricoles

Les applications médicales et épidémiologiques ont été largement discutées, mais l'un des buts de cette application de pigment est dans le contrôle des phytopathogènes, un domaine important pour le développement agricole (Durán et *al.*, 2016).

La violacéine dérivée de *Chromobacterium violaceum*, *Alteromonas luteoviolacea* ou *Janthinobacterium lividum* peut être appliquée pour la lutte contre les parasites. Ce pigment inhibe la croissance des champignons phytopathogènes et des nématodes phyto-parasites, et peut donc être utilisé pour lutter efficacement contre l'antracnose, la pourriture sclérotique des tiges, la brûlure pythienne des graminées et la fonte des semis des germes d'haricots, ainsi que des maladies des nématodes parasites des plantes, telles que les maladies de la pastèque provoqué par *Meloidogyne sp.* (Durán et *al.*, 2016).

II. Généralités sur le quorum sensing (QS)

1. Découverte et définitions

Avant 1994, la détection du quorum était communément appelée "auto-induction" (Nealson et *al.*, 1970 ; Fuqua et *al.*, 1994). L'auto-induction a été décrite à l'origine pour l'organisme marin Gram-négatif *Vibrio fischeri* au début des années 1970. La série d'expériences menées par Kempner et Hanson (1968) a révélé l'induction de la bioluminescence chez *Vibrio fischeri* fraîchement inoculé. La culture s'est illuminée en réponse à un milieu qui avait été préalablement conditionné par la même souche bactérienne. Nealson et *al.* (1970) ont été les premiers à proposer que l'auto-induction de la luminescence chez *Vibrio fischeri* se produit au niveau transcriptionnel et que le processus est régulé par des composants sécrétés extracellulairement (Eberhard, 1972).

Le terme "quorum sensing" a été introduit par le Dr Steven Winans en 1994, alors qu'il rédigeait l'un des premiers articles de synthèse sur l'auto-induction chez les bactéries. D'une certaine manière, le mot "autoinducteur", un terme utilisé pour décrire les petites molécules diffusibles impliquées dans le processus, ne semblait pas convenir au jeune professeur. Cette aversion était en partie due à la confusion courante entre le terme "auto-induction" et

"autorégulation" (Fuqua et *al.*, 1994). En outre, l'induction de la bioluminescence entre espèces avait été signalée par Greenberg et *al.* (1979), de sorte qu'en 1994, le terme "auto-induction" était devenu quelque peu inexact. Winans était déterminé à trouver un nouveau nom innovant, descriptif et, surtout, accrocheur. Assisté de son beau-frère à l'esprit littéraire, le Dr Winans a généré des dizaines de termes possibles, dont "gridlockins", "communiolins" et "quoromones". Aucun de ces termes n'est devenue populaire, mais la notion de quorum a été acceptée par les collègues de Winan et a fini par devenir le titre du chapitre. Le terme "quorum sensing" s'est répandu comme une traînée de poudre, se retrouvant dans pratiquement tous les articles sur l'auto-induction écrits par la suite (Winans, 1994).

2. Notion et mécanismes

Le quorum sensing (QS) est un processus de communication entre cellules bactériennes qui implique la production, la détection et la réponse à des molécules de signalisation extracellulaires appelées auto-inducteurs (AI). Les AI s'accumulent dans l'environnement à mesure que la densité de la population bactérienne augmente (Figure 12), et les bactéries surveillent ces informations pour suivre les changements dans leur nombre de cellules et modifier collectivement l'expression des gènes. QS contrôle des gènes qui dirigent des activités bénéfiques lorsqu'elles sont réalisées par des groupes de bactéries agissant de manière synchrone. Les processus contrôlés par QS comprennent la bioluminescence, la sporulation, la compétence, la production d'antibiotiques, la production des pigments, la formation de biofilms et la sécrétion de facteurs de virulence (Heilmann et *al.*, 2015).

Malgré des différences dans les composants régulateurs et les mécanismes moléculaires, tous les systèmes QS connus dépendent de trois principes de base :

- Premièrement, les membres de la communauté produisent des AI, qui sont les molécules de signalisation. À faible densité cellulaire (LCD), les AI se diffusent et sont donc présents à des concentrations inférieures au seuil requis pour la détection. À haute densité cellulaire (HCD) (Figure 12), la production cumulative d'AI conduit à une concentration locale élevée, permettant la détection et la réponse (Kaplan et *al.*, 1985).
- Deuxièmement, les AI sont détectées par des récepteurs qui existent dans le cytoplasme ou dans la membrane.

- Troisièmement, en plus d'activer l'expression des gènes nécessaires aux comportements coopératifs, la détection des AI entraîne l'activation de la production d'AI. Cette boucle d'auto-induction en amont favorise vraisemblablement la synchronisation de la population (Rutherford et al., 2012).

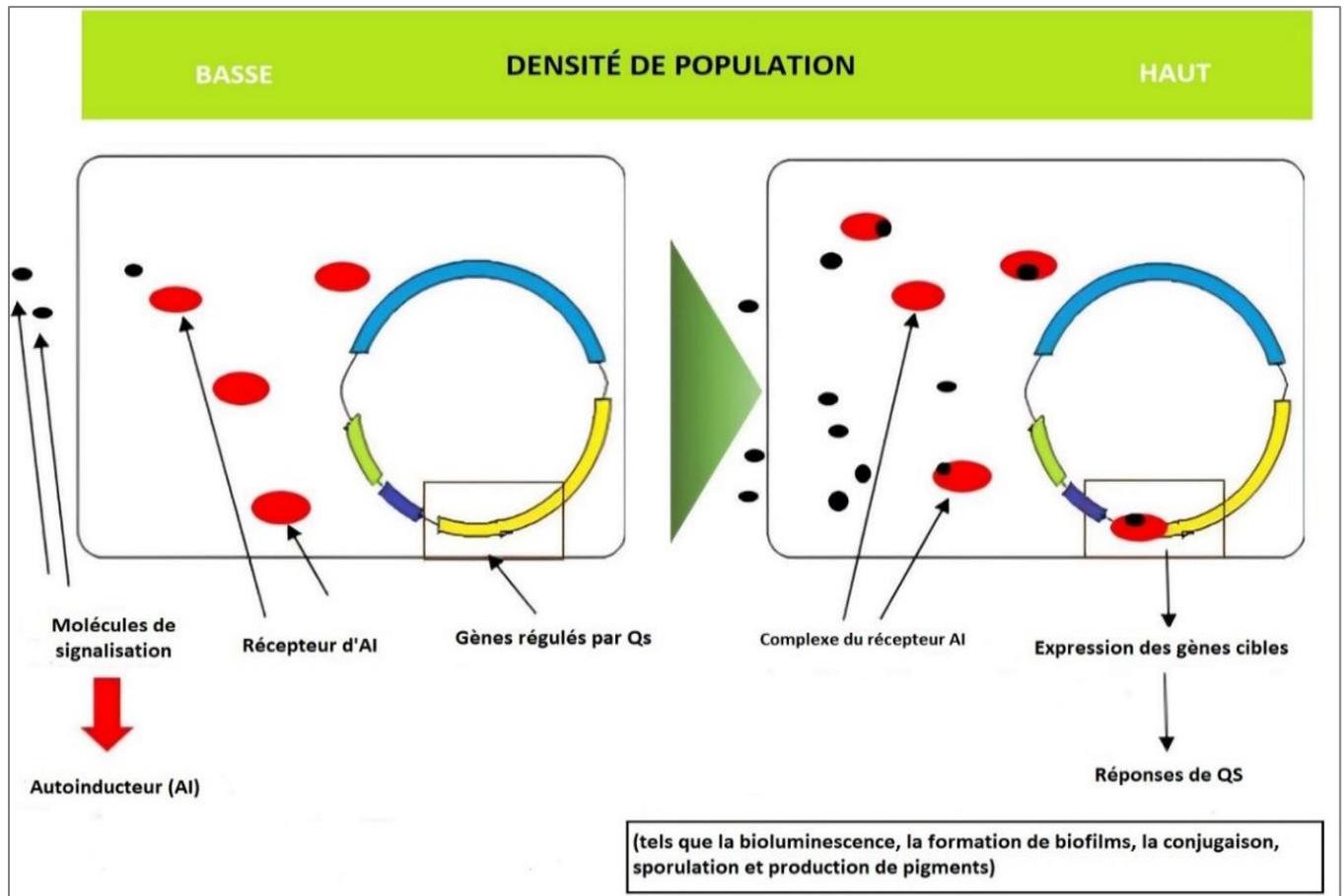


Figure 12 : Mécanisme de quorum sensing (Sarwar et al., 2021).

3. Le quorum sensing chez les bactéries

Les bactéries Gram-négatives et Gram-positives utilisent la détection du quorum pour communiquer, mais elles produisent des auto-inducteurs différents. Les bactéries Gram-négatives dépendent principalement des molécules de N-acyl homosérine lacton (AHL) (autoinducteur-1, AI-1) tandis que les bactéries Gram-positives utilisent des oligopeptides modifiés (peptides auto-inducteurs, AIP). Un troisième type d'est constitué de molécules de signalisation dérivées du bore furane (autoinducteur-2, AI-2) (figure 13) et sont produits et détectés par les bactéries Gram-négatives et Gram-positives.

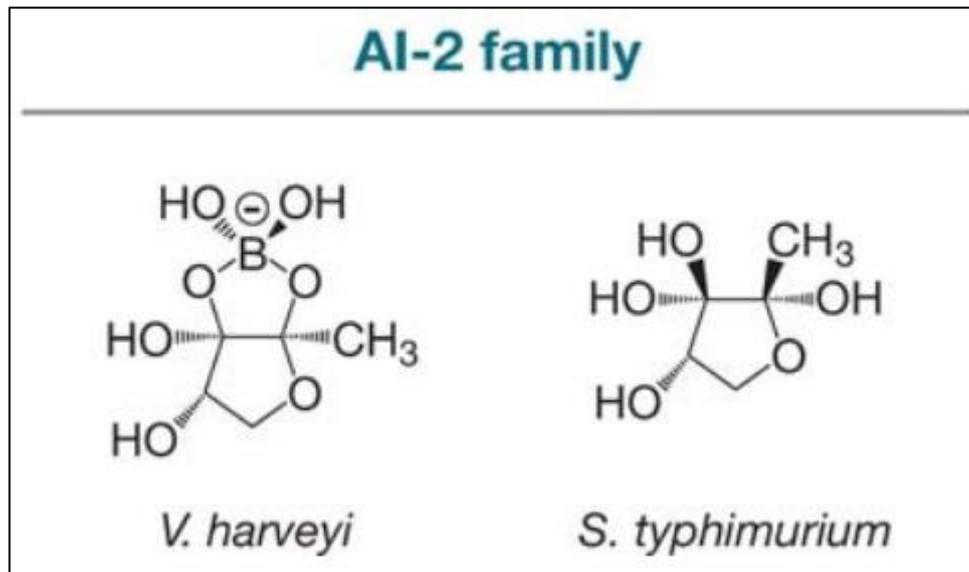


Figure 13 : structure d'autoinducteur-2, AI-2 (Waters et *al.*, 2005).

3.1. Bactéries Gram négatives

Le mécanisme de détection du quorum des bactéries Gram-négatives peut être décrit à l'aide de l'exemple de *Vibrio fischeri* (figure 13 et 14) : un autoinducteur synthase intracellulaire (LuxI) synthétise les molécules de signal AHL (figure 12) en catalysant une réaction entre la S- adénosyl méthionine et une protéine porteuse d'acyle. En raison de leur petite taille et de leur caractère lipophile, les auto-inducteurs AHL passe facilement la membrane cellulaire par diffusion passive. Si la concentration d'AHL est suffisamment élevée, l'autoinducteur AHL se lie à la protéine LuxR intracellulaire et provoque la révélation du domaine de liaison à l'ADN de LuxR.

Par la suite, la protéine LuxR se lie à l'ADN, provoquant l'activation de la transcription du gène cible (Figure 15) (Verbeke et *al.*, 2017).

Plus de 100 espèces de bactéries Gram-négatives appliquent un système de type LuxI/LuxR avec une synthase auto-inductrice (par exemple, LuxI) et un régulateur transcriptionnel (par exemple, LuxR) (Verbeke et *al.*, 2017).

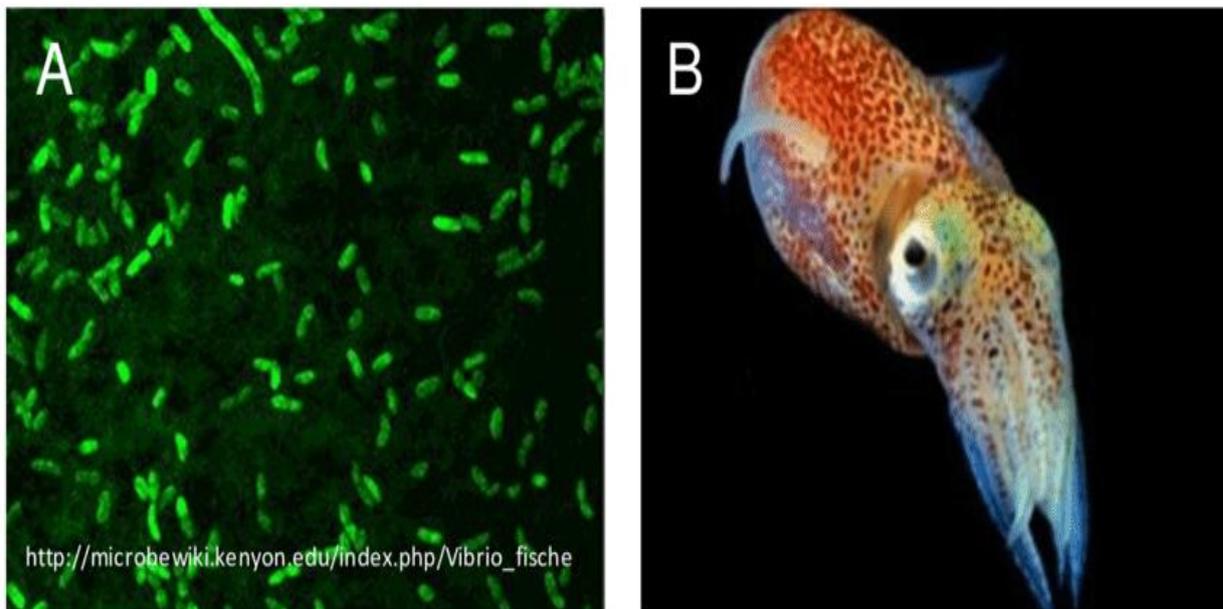


Figure 14 : A. *Vibrio fischeri* bioluminescent sous un microscope et B. un calmar à queue de cheval, avec un manteau d'une longueur d'environ 1,2 pouce (Mann et al., 2015).

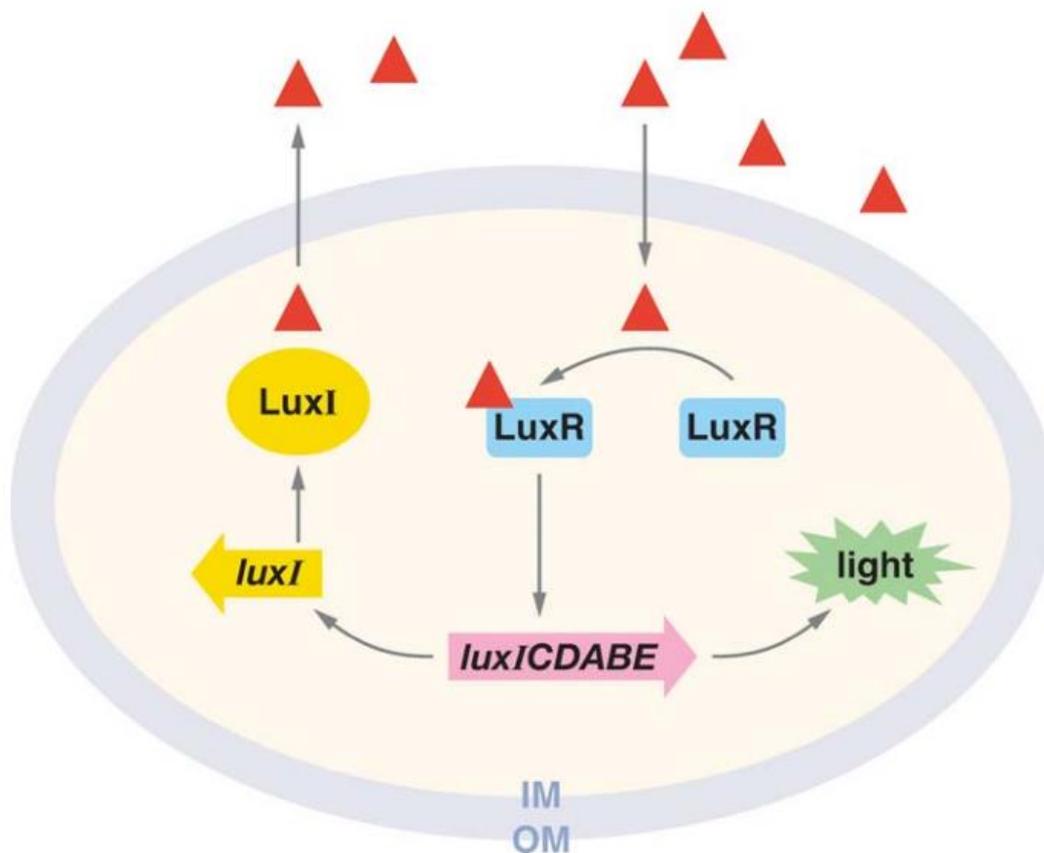


Figure 15 : Détection du quorum chez *Vibrio fischeri* ; une signalisation LuxIR circuit. Les triangles rouges indiquent l'autoinducteur qui est produit par LuxI. OM, membrane externe ; IM, membrane interne (Waters et al., 2005).

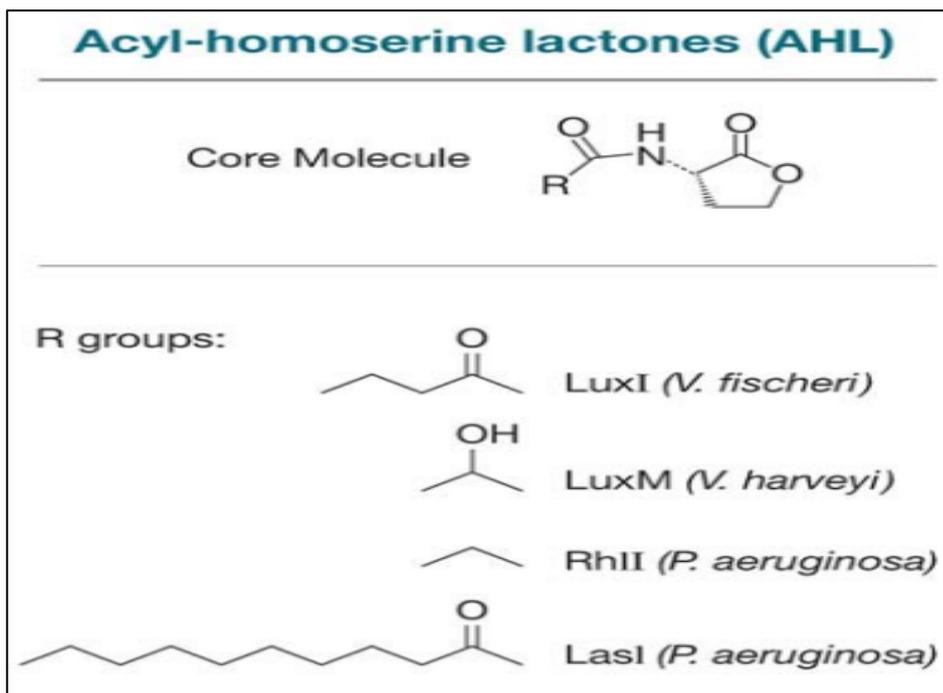


Figure 16 : Structure d'autoinducteur AHL (Waters et *al.*, 2005).

3.2. Bactéries Gram positives

Les bactéries Gram-positives utilisent des peptides comme auto-inducteurs pour la détection du quorum (figure 18, panneau de droite). Des exemples de ce groupe hétérogène de peptides sont montrés dans la figure 17. Ces peptides sont synthétisés par les ribosomes en tant que peptides précurseurs et subissent des modifications post-traductionnelles pendant l'excrétion pour être activés et stabilisés. En général, la sécrétion du PIA est facilitée par un transporteur ABC (ATP-binding cassette) associé à la membrane. Lorsque la densité de la population augmente, les PIA s'accumulent dans l'environnement. Lorsqu'un certain seuil est atteint, la liaison d'un AIP à un récepteur déclenche l'activation de la kinase du récepteur par phosphorylation sur un résidu histidine conservé. Par la suite, le récepteur kinase activée transfère le groupe phosphorylé à un résidu aspartate conservé du régulateur de réponse intracellulaire, qui sera à son tour activé. Le régulateur de réponse activé influence la transcription de gènes cibles, notamment les gènes AIP, les gènes du récepteur kinase et du régulateur de réponse et les gènes du transporteur ABC. Selon l'espèce, la nomenclature des mécanismes de détection du quorum peut être différente, en raison des gènes et des récepteurs impliqués. Par exemple, les espèces de staphylocoques utilisent le système de détection du quorum agr, les espèces de streptocoques utilisent le

système de détection du quorum ComX et les espèces de bacilles utilisent le système de détection du quorum Rap (Verbeke et *al.*, 2017).

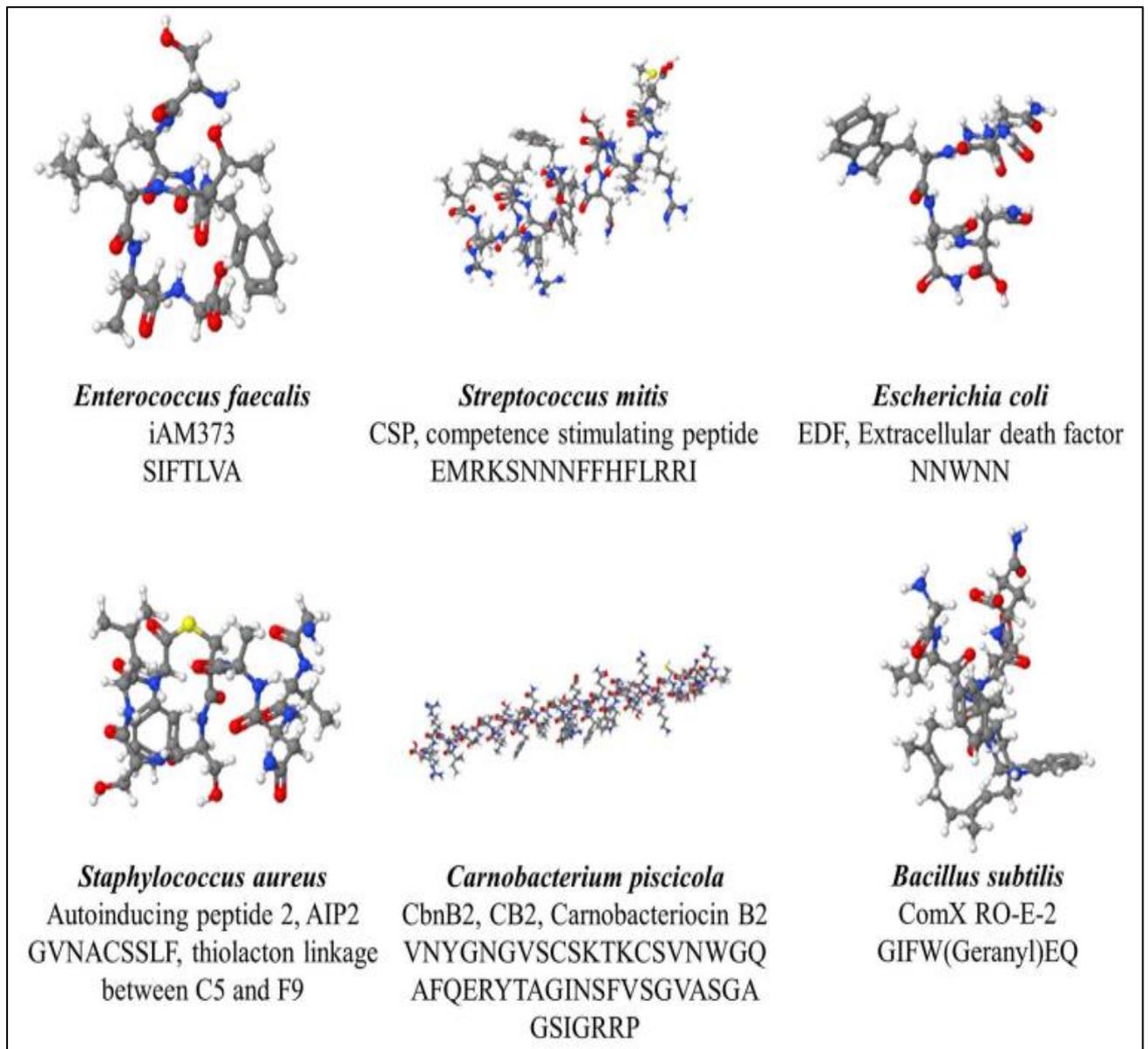


Figure 17 : Divers peptides de détection du quorum, illustrant la diversité structurelle observée parmi les peptides de détection du quorum (Verbeke et *al.*, 2017).

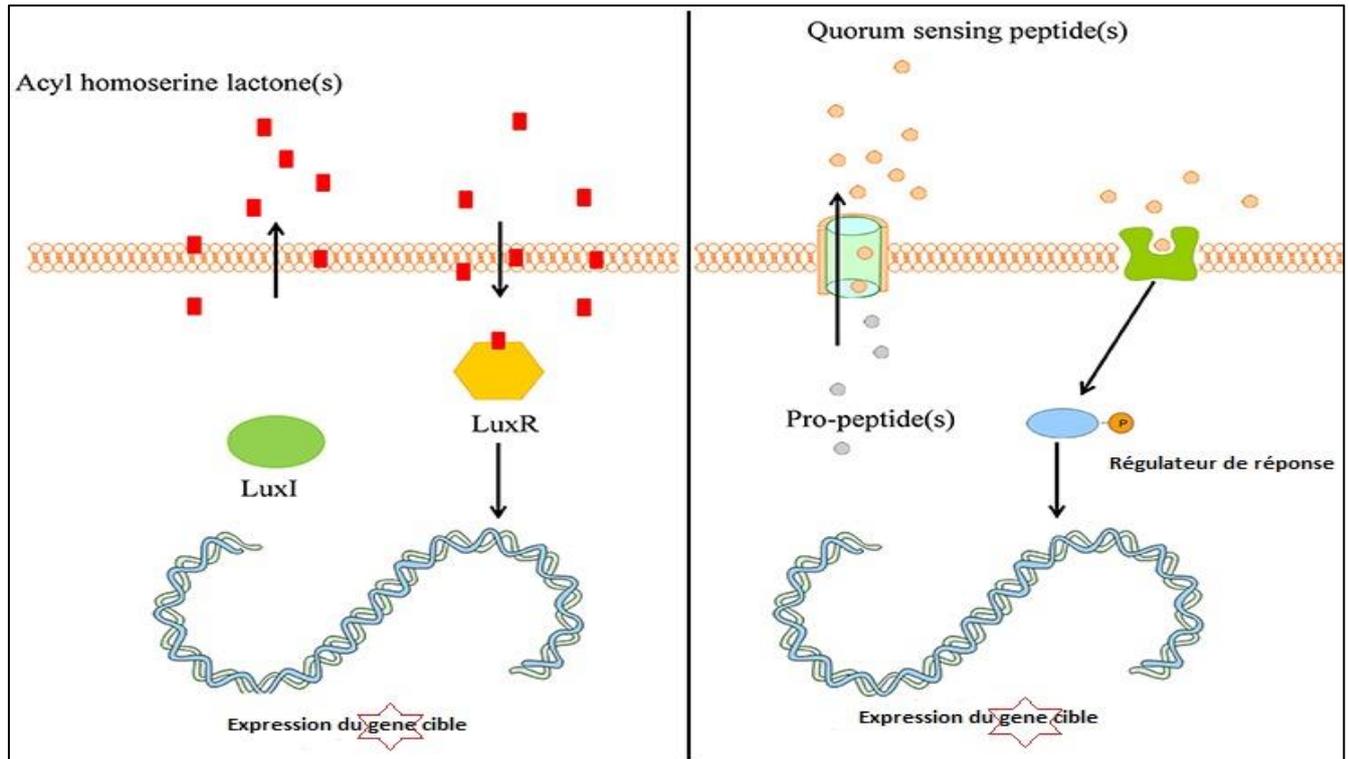


Figure 18 : (Panneau de gauche) Mécanisme typique de détection du quorum des bactéries Gram-négatives (AHL). (Panneau de droite) Mécanisme typique de détection du quorum des bactéries Gram-positives (Peptide) (Verbeke et *al.*, 2017).

4. Quorum sensing chez *Chromobacterium violaceum*

Le système QS de *C. violaceum* comprend quatre composants principaux :

- La CviI synthase (la molécule diffusible des AHL) (Stauff et *al.*, 2011).
- Un récepteur de signal de type CviR et certains gènes cibles qui impliquent l'activation de l'activité lytique comme les chitinases, et les exoprotéases (Stauff et *al.*, 2011).
- Un facteur de virulence comme le gène lié à la sécrétion de type VI (Stauff et *al.*, 2011).
- Des gènes liés à un régulateur transcriptionnel, une guanine désaminase et la *cviI* responsable de la synthèse des AHL et l'opéron violacéine (*VioA, B, C, D, E*) entre autres (Stauff et *al.*, 2011).

La production de violacéine n'est pas le seul trait qui est régulé par QS du *C. violaceum*, ils Ya d'autre caractéristiques phénotypiques comme la production d'élastase, et la production de cyanure, sont également contrôlées par un circuit QS médié par la molécule signal, l'acyl homosérine lactone (AHL) (tableau 6) (Kothari et *al.*, 2017).

Tableau 6 : Divers traits régulés par QS de *C. violaceum* (Kothari et al., 2017).

Trait régulé par QS	Gène(s)/Opéron codant pour le trait
Production de violacéine	Opéron <i>Vio</i> ABCDE (<i>VioA</i> , <i>VioB</i> , <i>VioC</i> , <i>VioD</i> , <i>VioE</i>)
Production de cyanure d'hydrogène	Opéron <i>hcnABC</i>
Dégradation des cyanures	<i>cynT</i> (cyanate permease : CV1881) et <i>cynS</i> (cyanase : CV1880)
Production d'élastase Pilus	<i>pilE2</i> (codes for type IV pilus protéine)
Formation du Biofilm	HmsHNFR

Chapitre III :

Bio-informatique (*in silico*)

I. Généralités sur la bio-informatique

1. C'est quoi la bioinformatique ?

La bioinformatique est une branche émergente de la science biologique qui est née de la combinaison de la biologie et des technologies de l'information. Il s'agit d'un domaine d'étude interdisciplinaire qui fait appel à la biologie, à la chimie, aux mathématiques, aux statistiques, aux ingénieries et à l'informatique (figure 18), qui sont fusionnées pour former une seule discipline. Ce secteur est principalement impliqué dans l'analyse des données biologiques, le développement de nouveaux logiciels utilisant des outils biologiques (Sheokand *et al.*, 2019).

Selon le NCBI (National Center for Biotechnology Information), branche de la NLM (National Library of Medicine) et du NIH (National Institutes of Health), la bioinformatique est l'analyse, la collecte, la classification, la manipulation, la récupération, le stockage et la visualisation de toutes les informations biologiques à l'aide de la technologie informatique (Eric *et al.*, 2014).

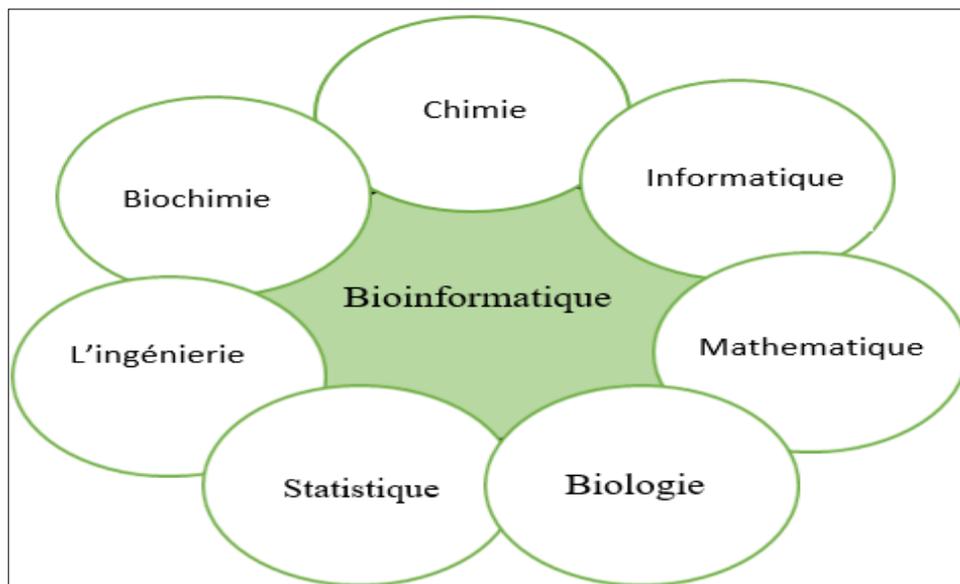


Figure 19 : Multi discipline de la bio-informatique

2. Historique

La bio-informatique est une discipline relativement récente, le terme ayant été créé dans les années 80 (Gérard, 2012). Le terme bioinformatique a été inventé par Paulien Hogeweg de

l'Université d'Utrecht (Pays-Bas) en 1979 pour l'étude des processus informatiques dans les systèmes biotiques, mais le domaine de la bio-informatique n'a été reconnu que dans les années 1990 (Eric et *al.*, 2014). Cette notion englobe l'ensemble des applications de l'informatique aux sciences de la vie, domaine très vaste qui recouvre tous les axes de recherche, allant des applications en robotique aux techniques les plus avancées en intelligence artificielle. Pour la plupart des membres de la communauté scientifique, cette notion semble dans la pratique s'adapter plus particulièrement aux outils informatiques qui permettent de stocker, d'analyser et de visualiser les informations contenues dans les séquences des gènes et des protéines des êtres vivants. L'histoire de la bio-informatique est donc étroitement liée à celle de la biologie moléculaire, l'étude des molécules du vivant. (Gérard, 2012).

- **Margaret Oakley Dayhoff** (1925-1983), (figure 21), physicochimiste américaine était la pionnière de l'application des méthodes informatiques en biochimie en 1960. La contribution de Dayhoff dans ce domaine est aussi considérée avec la même importance que les fonctions exercées par David J Lipman, directeur du National Center for Biotechnology Information (NCBI). (Gauthier et *al.*, 2019).



Figure 20 : Professeur Margaret Oakley Dayhof (Gauthier et *al.*, 2019).

Tableau 7 : Les dates clés de l'essor de la bio-informatique (Gauthier et *al.*, 2019).

1965	Première compilation de protéines ("Atlas of Protein Sequences")
1970	Algorithme pour l'alignement global de séquences
1971	Premier microprocesseur Intel 4004
1972	Clonage de fragments d'ADN dans un virus, l'ADN recombiné
1973	Découverte des enzymes de restriction qui coupe spécifiquement l'ADN. Méthode de transfection (introduction d'un ADN étranger) des cellules eucaryotes grâce à un virus (vecteur).
1974	Programme de prédiction de structures secondaires des protéines : "Prediction of Protein Conformation"
1977	Micro-ordinateurs. Techniques de séquençage d'ADN
1978-1980	Mutagenèse dirigée : Michael Smith, Séquençage du 1er génome à ADN, le bactériophage phiX174 Premières bases de données : EMBL, GenBank, PIR
1981	Micro-ordinateur IBM-PC 8088. 370.000 nucléotides séquencés. Programme d'alignement local de séquences
1983	IBM-XT disque dur (10 Mb)
1984	Amplification de l'ADN : réaction de polymérisation en chaîne (PCR) MacIntosh : interface graphique & souris
1985	"FASTA" : Programme d'alignement local de séquences
1987	Nouveau vecteur permettant de cloner des fragments d'ADN 20 fois plus grands : le YAC (Yeast Artificial Chromosome). Rend possible l'étude de grands génomes.
1988	Taq polymérase, enzyme thermostable pour la PCR.
1989	INTERNET succède à ARPANET (En 1990 le CERN - Centre Européen de Recherche Nucléaire- propose au monde le protocole HTTP et le Web)
1990	Clonage positionnel et premier essai de thérapie génique. "BLAST" : Programme d'alignement local de séquences
1992	Séquençage complet du chromosome III de levure
1996	Séquençage complet de la levure (consortium Européen)
1997	11 génomes bactériens séquencés
1998	Séquençage de 2 millions de nucléotides par jour
2000	Séquençage du 1er génome de plante : <i>Arabidopsis thaliana</i>

2001	Séquence "premier jet" complète du génome humain
2010	Création de la première bactérie synthétique, contrôlée par un génome synthétisé chimiquement. (Science 2 July 2010 : 40-41)
Septembre 2012	3708 génomes eucaryotes et procaryotes séquencés et plus de 14000 en projet (Genomes OnLine) ! Plus de 430.000.000.000 nucléotides !
Septembre 2013	Nombre de génomes séquencés : 6880. Nombre de génomes en cours : 21058
Avril 2014	Plus de 776.000.000.000 nucléotides séquencés
Décembre 2015	Plus de 18.900 génomes eucaryotes et procaryotes séquencés et des milliers en projet. Plus de 1.363.000.000.000 nucléotides séquencés.
Mars 2019	303 milliards de nucléotides 49 millions séquencées d'acide aminé

3. Objectifs

La bio-informatique organise les données de manière à permettre aux chercheurs d'accéder aux informations existantes et également de soumettre de nouvelles entrées à des bases de données telles que Protein Data Bank, et aussi cherche à développer des outils d'analyse des données dans le but d'utiliser ces outils pour analyser et interpréter les résultats des données (Eric et *al.*, 2014), donc les objectifs sont (Zucker, 2010) :

- ✓ Collecter et filtrer, stocker et accéder des informations (données biologiques) dans des bases de données).
- ✓ Analyser, structurer, comparer, traiter les informations biologiques (par exemple fournir des outils de comparaison de séquences (protéiques ou nucléotidiques).
- ✓ Formaliser des problèmes de biologie moléculaire.
- ✓ Modéliser des processus biologiques
- ✓ Prédire des résultats biologiques.
- ✓ Faire avancer les connaissances en biologie, en génétique humaine, en théorie de l'évolution, etc.,
- ✓ Aider à la conception des médicaments.
- ✓ Comprendre les maladies complexes.
- ✓ Développement de logiciels pour la biologie.

4. Méthodes de la bio-informatique

Pour répondre aux objectifs suscités, les pratiques en bio-informatique s'organisent autour de méthodes de regroupement, de partitionnement, d'association et de prédiction...

4.1. Classification

Est la technique la plus couramment appliquée, qui utilise un ensemble d'exemples pré-classifiés pour développer un modèle capable de classifier la population d'enregistrements au sens large. Les applications de détection des fraudes et des risques de crédit sont particulièrement bien adaptées à ce type d'analyse. Cette approche utilise fréquemment des algorithmes de classification basés sur des arbres de décision ou des réseaux neuronaux. Le processus de classification des données implique l'apprentissage et la classification. Dans l'apprentissage, les données de formation sont analysées par l'algorithme de classification. Dans la classification, des données de test sont utilisées pour estimer la précision des règles de classification. Si la précision est acceptable, les règles peuvent être appliquées aux nouveaux tuples de données. Pour une application de détection des fraudes, il s'agirait d'enregistrements complets des activités frauduleuses et valides, déterminés sur une base d'enregistrement par enregistrement. L'algorithme d'apprentissage du classificateur utilise ces exemples pré-classifiés pour déterminer l'ensemble des paramètres nécessaires à une discrimination correcte. L'algorithme code ensuite ces paramètres dans un modèle appelé classificateur (Ramageri et *al.*,2017).

- Types de modèles de classification :
 - Classification par induction d'arbre de décision
 - Classification bayésienne
 - Réseaux neuronaux
 - Machines à vecteurs de support (SVM)
 - Classification basée sur les associations

4.2. Clustering

Le clustering peut être considéré comme l'identification de classes d'objets similaires. En utilisant les techniques de clustering, nous pouvons identifier les régions denses et éparses dans l'espace objet et découvrir le modèle de distribution global et les corrélations entre les attributs des données. L'approche de la classification peut également être utilisée comme moyen efficace de distinguer des groupes ou des classes d'objets, mais elle devient coûteuse. Le clustering peut

donc être utilisé comme approche de prétraitement pour la sélection de sous-ensembles d'attributs et la classification. Par exemple, pour former des groupes de clients en fonction de leurs habitudes d'achat, pour classer des gènes ayant une fonctionnalité similaire (Muhammad, 2015 ; Ramageri et *al.*,2017).

- Types de méthodes de regroupement
 - Méthodes de partitionnement
 - Méthodes d'agglomération hiérarchique (division)
 - Méthodes basées sur la densité
 - Méthodes basées sur la grille
 - Méthodes basées sur des modèles

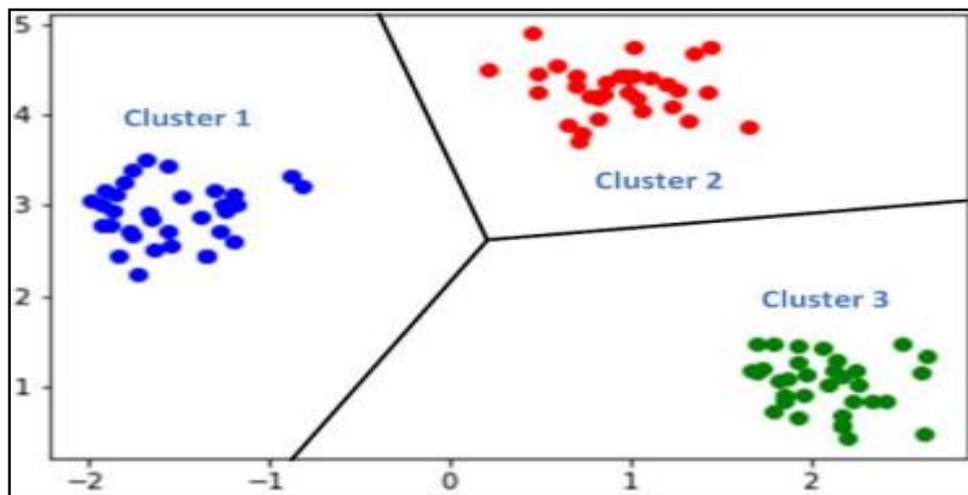


Figure 21 : Exemple de clustering de données (Muhammad, 2015).

4.3. Prédiction

La technique de régression peut être adaptée à la prédiction. L'analyse de régression peut être utilisée pour modéliser la relation entre une ou plusieurs variables indépendantes et des variables dépendantes. Dans le data mining, les variables indépendantes sont des attributs déjà connus et les variables de réponse sont ce que nous voulons prédire. Malheureusement, de nombreux problèmes du monde réel ne sont pas simplement de la prédiction. Malheureusement, de nombreux problèmes du monde réel ne sont pas de simples prédictions. Par exemple, les volumes de ventes, les prix des stocks et les taux de défaillance des produits sont tous très difficiles à prédire car ils peuvent dépendre d'interactions complexes entre plusieurs variables prédictives. Par conséquent, des techniques plus complexes (par exemple, la régression logistique, régression

logistique, arbres de décision ou réseaux neuronaux) peuvent être nécessaires pour prévoir les valeurs futures. Les mêmes types de modèles peuvent souvent être utilisés pour la régression et la classification. Par exemple, l'algorithme d'arbre de décision CART peut être utilisé pour construire des arbres de classification et des arbres de régression (pour prévoir des variables de réponse continues).

Les réseaux neuronaux peuvent également créer des modèles de classification et de régression (Ramageri et *al.*, 2017).

- Types de méthodes de régression
 - Régression linéaire
 - Régression linéaire multivariée
 - Régression non linéaire
 - Régression non linéaire multivariée

4.4. Règle d'association

L'association et la corrélation permettent généralement de trouver des résultats d'ensembles d'éléments fréquents parmi de grands ensembles de données. Ce type de recherche aide les entreprises à prendre certaines décisions, comme la conception de catalogues, le marketing croisé et l'analyse du comportement d'achat des clients. Les algorithmes de règles d'association doivent être capables de générer des règles avec des valeurs de confiance inférieures à un. Cependant, le nombre de règles d'association possibles pour un ensemble de données donné est généralement très élevé et une grande partie des règles n'ont généralement que peu (ou pas) de valeur (Ramageri et *al.*, 2017).

- Types de règles d'association
 - Règle d'association multi-niveaux
 - Règle d'association multidimensionnelle
 - Règle d'association quantitative

5. Les outils de la bio-informatique

- ❖ **Les outils lignes de commandes** : Ces outils peuvent être difficile à utiliser pour la plupart des biologistes, mais offrent presque toujours plus d'options pour l'exécution des programmes. Ils sont plus appropriés pour analyser des ensembles de données grande échelle

qui sont rencontrés actuellement en bio-informatiques (figure 23) ,(Rechenman et Quinkal, 2004 ; Siddharthan, 2005 ; Jamet, 2006 ; Pevnsner, 2015).

- ❖ **Les outils Web (web based software) :** parfois appelés (point-and-click) ne nécessitent pas de connaissances en programmation et sont immédiatement accessibles la communauté scientifique (figure 23) (Pvenser, 2015).

Le domaine de la bioinformatique s'appuie fortement sur Internet pour accéder aux données de séquence, aux logiciels utiles pour analyser les données moléculaires et pour intégrer différents types de ressources et d'informations relatives la biologie. Nous allons décrire une variété de sites Web. Dans un premier temps, nous nous concentrerons sur les principales bases de données accessibles au public qui servent de référentiels pour les données sur l'ADN et les protéines (Biegert et *al.*, 2006), (Pevnsner, 2015). Ceux-ci comprennent :

- Le Centre national d'information sur la biotechnologie (NCBI), qui héberge la GenBank et d'autres ressources
- L'Institut européen de bioinformatique (EMBI)
- Ensemble, qui comprend un navigateur génomique et des ressources pour étudier des dizaines de génomes
- Le site de bioinformatique du génome de L'Université de Californie Santa Cruz (UCSC), comprenant un navigateur Web et un navigateur de tableaux pour diverses espèces.

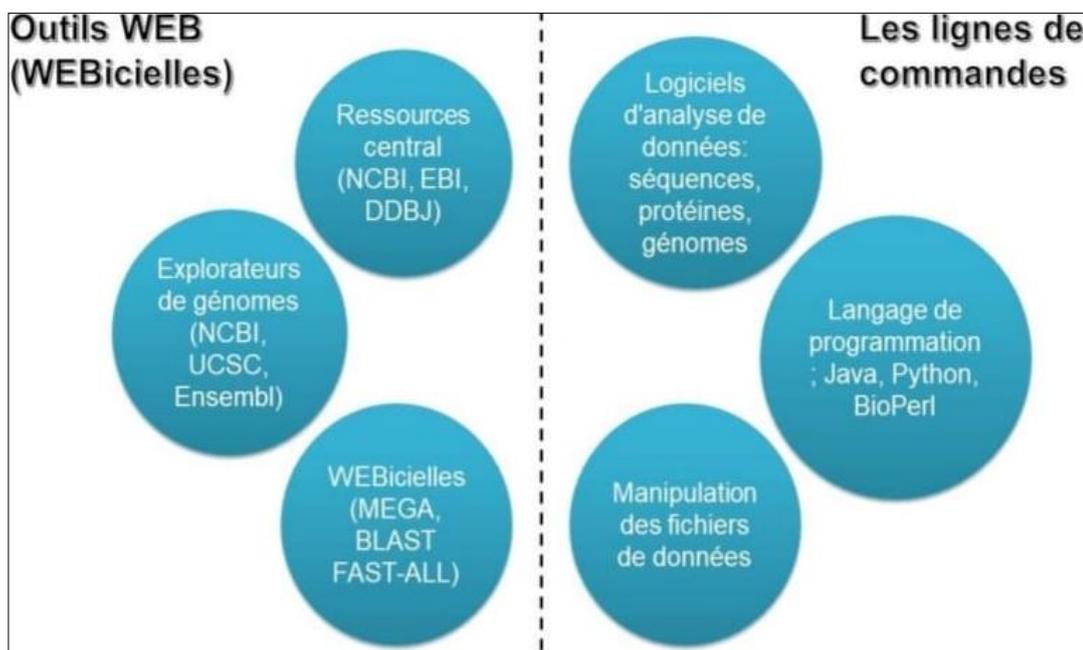


Figure 22 : représente les ressources en bioinformatique. Les ressources web et les ressources lignes de commandes (Pvenser, 2015).

5.1. Fichiers et formats

Les séquences, que l'on peut appeler lectures brutes, générées par le séquençage sur différentes plateformes sont stockées dans de multiples fichiers de lectures courtes. Après le séquençage, les données brutes sont générées et peuvent être stockées dans différents formats de fichiers tels que FASTQ, FASTA, SAM/BAM, etc. (Simons, 2004).

Le format permet :

- Une mise en forme automatisée
 - Le stockage homogène de l'information
 - Le traitement informatique ultérieur de l'information
- Exemples de quelques formats en bio-informatique :
- a) **Format FASTA** : Sans doute le format de fichier le plus répandu car très simple et l'un des plus pratique. Ce format est généralement utilisé lors de l'alignement ou de la cartographie du génome de référence par différents outils et logiciels (Simons, 2004).
 - b) **Format FASTQ** : FASTQ est le format de fichier le plus couramment utilisé. Il permet de stocker des données avec des valeurs de qualité correspondantes connues sous le nom de Phred scores (Simons, 2004).
 - c) **Format Stockholm** : est un format d'alignement de séquences multiples utilisé par Pfam et Rfam pour diffuser les alignements de séquences de protéines et d'ARN (Finn et *al.*, 2008; Gardner et *al.*, 2009).
 - d) **BAM/SAM** : Les données de séquence brutes générées par le séquenceur ne contiennent aucune information génomique et doivent être alignées sur un génome de référence. Après la cartographie ou l'alignement sur un génome de référence, la sortie est générée au format SAM/BAM. SAM qui stocke les séquences dans un format aligné sur le génome de référence (Simons, 2004).

5.2. Les algorithmes en bio-informatique

On distingue deux types d'algorithmes se basant sur l'alignement des séquences :

❖ **Les algorithmes d'alignement globaux : (Needleman&Wunsch, 1970) :**

L'un des premiers et des plus importants algorithmes d'alignement de deux séquences de protéines a été décrit par Needleman et Wunsch (1970). Cet algorithme est important car il produit un alignement optimal de séquences de protéines ou d'ADN, même en permettant l'introduction de lacunes. Le résultat est optimal, mais il n'est pas nécessaire d'évaluer tous

les alignements possibles. Une comparaison exhaustive par paires serait trop coûteuse à réaliser en termes de calcul (Hatzimaikatis, 2000).

Nous pouvons décrire l'approche Needleman-Wunsch de l'alignement global de séquences en trois étapes : (1) établissement d'une matrice ; (2) évaluation de la matrice ; et (3) identification de l'alignement optimal (Hatzimaikatis, 2000).

❖ **Les algorithmes d'alignement locaux : (Smith & Waterman, 1981) :**

L'algorithme d'alignement local de Smith et Waterman (1981) est la méthode la plus rigoureuse par laquelle des sous-ensembles de deux séquences de protéines ou d'ADN peuvent être alignés. L'alignement local est extrêmement utile dans une variété d'applications telles que la recherche de bases de données dans lesquelles nous pouvons souhaiter aligner des domaines de protéines (mais pas les séquences entières). Un algorithme d'alignement local de séquences ressemble à celui de l'alignement global en ce sens que deux protéines sont disposées dans une matrice et qu'un chemin optimal le long d'une diagonale est recherché. Cependant, il n'y a pas de pénalité pour commencer l'alignement à une position interne, et l'alignement ne s'étend pas nécessairement jusqu'aux extrémités des deux séquences (Hatzimaikatis, 2000) (ex : BLAST est un programme qui adopte cet algorithme).

❖ **Versions rapides et heuristiques de Smith-Waterman : FASTA - Pearson & Lipman (1988) et BLAST Altschul et al. (1990) :**

Deux algorithmes d'alignement local populaires ont été développés et offrent des alternatives rapides à Smith-Waterman : FASTA (Pearson et Lipman, 1988) et BLAST (Basic Local Alignment Search Tool ; Altschul et al., 1990). Chacun de ces algorithmes nécessite moins de temps pour effectuer un alignement. Le gain de temps est dû au fait que FASTA et BLAST restreignent la recherche en balayant une base de données pour trouver des correspondances probables avant d'effectuer des alignements plus rigoureux. Il s'agit d'algorithmes heuristiques qui sacrifient une certaine sensibilité en échange de la rapidité ; contrairement à Smith-Waterman, ils ne sont pas garantis de trouver des alignements optimaux (Hatzimaikatis, 2000).

- **Programme FASTA - Pearson et Lipman (1988) :** Le programme FASTA ne considère que les séquences présentant une région de forte similitude avec la séquence recherchée. La codification numérique des séquences, c'est-à-dire la décomposition de la séquence en courts motifs transcodés en entiers, confère à l'algorithme l'essentiel de sa rapidité.

- **Le programme BLAST** (et ses dérivés) - Altschul et *al.* (1990) : Le programme BLAST (Basic Local Alignment Search Tool) s'appuie sur une méthode heuristique qui utilise la méthode de Smith & Waterman. La rapidité de BLAST permet la recherche des similarités entre une séquence requête et toutes les séquences d'une base de données.

6. Les bases de données biologiques

Les fichiers contenant l'information biologique sous la forme de séquences est l'élément central autour duquel les banques de données se sont constituées. Il existe un grand nombre de bases de données d'intérêt biologique (Bahia, 2011). On peut distinguer :

- les banques de données ou bases de données généralistes : elles correspondent à une collecte des données la plus exhaustive possible et qui offrent un ensemble plutôt hétérogène d'informations (Kanehisa & Bork, 2003).
 - les bases de données ou bases de données spécialisées : elles correspondent à des données plus homogènes établies autour d'une thématique et qui offrent une valeur ajoutée (Kanehisa et *al.*, 2003).
- Exemple de grandes banques généralistes :
- EMBL** : Banque européenne créée en 1980 et financée par l'EMBO (European Molecular Biology Organisation). Elle est aujourd'hui diffusée par l'EBI (European Bioinformatics Institute, Cambridge) (Bahia, 2011).
 - Genbank** : Créée en 1982 par la société Intelli Genetics et diffusée maintenant par le NCBI (National Center for Biotechnology Information, Los Alamos) (Bahia, 2011).
 - DDBJ** (DNA Data Bank of Japan) : Créée en 1986 et diffusée par le NIG (National Institute of Genetics, Japon) (Bahia, 2011).
- Exemple de banques spécialisées :
- Disulfide Bridge DataBase** (J.-M. Richer, G. Hunault & E. Jaspard) : Cette base de données contient un grand nombre d'informations structurales sur les cystéines (Bahia, 2011).
 - LEAPdb** (G. Hunault, L. Châtelain & E. Jaspard) : Cette base de données contient un grand nombre d'informations sur les "Late Embryogenesis Abundant proteins". Elle a pour but d'étudier la relation structure - fonction de ces protéines impliquées entre autres dans la protection contre les stress biotiques et abiotiques (Bahia, 2011).

7. Les disciplines de la bioinformatique

La bioinformatique comprend plusieurs disciplines (tableau 8) (Amara, 2020).

Tableau 8 : Les disciplines de la bioinformatique et leurs utilités (Amara, 2020).

Discipline	Utilité
Stockage et gestion des données	Banques de données généralistes et spécialisées
Structures moléculaires	Visualisation, analyse, classification, prédiction
Analyse de séquences	Alignements, recherches de similarités ...
Génomique structurale	Annotation des génomes, génomique comparative
Génomique fonctionnelle	Transcriptome, protéome, interactome
Phylogénie	Relations évolutives entre gènes, entre génomes, entre organismes, inférence de scénarios évolutifs
Analyses des réseaux bimoléculaires	Réseaux métaboliques, d'interactions protéiques, de régulation génétique, ...

8. Les applications de la bioinformatique

Le domaine de la bioinformatique, qui comprend divers outils et bases de données, est utilisé dans un grand nombre de domaines. Il s'agit notamment de l'analyse de la génétique liée aux micro-organismes, de la biotechnologie, de la gestion des déchets, de la thérapie génique, de l'agriculture et de la médecine. La combinaison de l'informatique et de la biologie a rendu les découvertes plus pratiques et les résultats plus précis (Thapar, 2010).

8.1. La médecine personnelle

La médecine personnelle consiste à personnaliser les soins de santé en fonction du patient en utilisant des informations génétiques. Les informations génétiques peuvent être largement déterminées en utilisant des outils et des bases de données bio-informatique (Thapar, 2010).

8.2. Gestion des déchets

La gestion des déchets par l'utilisation de micro-organismes est appelée bioremédiation. Le processus de bioremédiation utilise des couches de complexité à plusieurs étapes pour éliminer les déchets toxiques. Les bases de données sont impliquées dans la prévision des informations sur

les produits chimiques, la toxicité, l'évaluation des risques, le degré du processus de dégradation (Thapar, 2010).

8.3. Industries laitières

Grâce aux bases de données relatives à la composition génétique des organismes, les micro-organismes peuvent être explorés pour améliorer la valeur nutritionnelle et la qualité des produits. Ils peuvent également être utilisés comme vecteurs de médicaments dans la science pharmaceutique (Thapar, 2010).

8.4. Agriculture

- a) Cultures - Il est possible d'approfondir les connaissances sur les séquences génomiques de diverses cultures comme *Arabidopsis*, le riz, le pois, la papaye, etc. à l'aide de la bioinformatique (Thapar, 2010).
- b) Résistance aux insectes - Cette technologie est devenue un succès grâce à la base de données des séquences du génome de micro-organismes, dont les gènes de résistance aux insectes ont été incorporés dans les plantes pour les rendre résistantes aux insectes. Cette technologie a permis de diminuer la quantité d'insecticides dans les plantes, d'augmenter la productivité et la valeur nutritionnelle des plantes (Thapar, 2010).

Partie II :

Partie expérimentale

(Analyse in silico)

Chapitre I :

Matériels et méthodes

1. Matériels

Notre travail consiste à réaliser une analyse bio-informatique (*in silico*) sur la protéine régulatrice de la transcription CviR et les gènes *VioA*, *VioB*, *VioC*, *VioD* et *VioE* de la biosynthèse du violacéine en situation de quorum sensing chez *Chromobacterium violaceum*.

L'étude s'est réalisée totalement sur la plateforme NCBI (Centre national de biotechnologie et information bio-informatique), en utilisant les outils d'analyse disponible en ligne, par application directe. Pour l'analyse relatif à la protéine régulatrice CviR, son analyse structurale s'est effectuée, aussi en ligne, en sollicitant les données protéomiques des bases de données PDB (Protein Data Bank) et Uniprot (Universal Protein Resource).

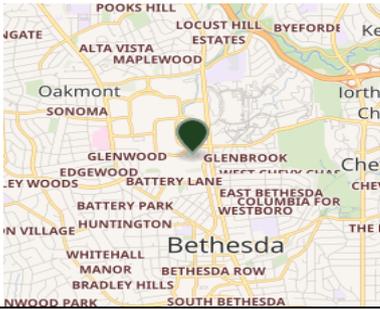
1.1. National Center for Biotechnology Information NCBI

Le Centre américain pour les informations biotechnologiques (NCBI), en anglais National Center for Biotechnology Information, est un institut national américain pour l'information biologique moléculaire. Fondé en 1988 à Bethesda dans le Maryland (tableau 9), cet organisme fait partie de la Bibliothèque américaine de médecine, un des Instituts américains de la santé (Stephen, 2007).

Le NCBI abrite une série de bases de données pertinentes pour la biotechnologie et la biomédecine et renferme une ressource importante pour les outils et services bioinformatiques. Les principales bases de données comprennent GenBank pour les séquences d'ADN et PubMed, une base de données bibliographiques pour la littérature biomédicale. D'autres bases de données comprennent la base de données Epigenomics du NCBI. Toutes ces bases de données sont accessibles en ligne par le biais du moteur de recherche Entrez. Le NCBI a été dirigé par David Lipman, l'un des auteurs originaux du programme d'alignement de séquences BLAST (Basic Local Alignment Search Tool) et une figure largement respectée de la bioinformatique (Stephen, 2007).



Tableau 9 : Informations générales sur NCBI (stephen, 2007).

National Center for Biotechnology Information	
Abréviation	NCBI
Fondée	4 novembre 1988
Siège	Bethesda
Emplacement	 États-Unis
Langue officielle	Anglais
Directeur	Stephen Sherry (since Mar 31, 2020)
Organisation mère	United States National Library of Medicine
Affiliations	National Institutes of Health
Website	www.ncbi.nlm.nih.gov
Coordonnées	38° 59' 45" N, 77° 05' 56" O 

1.2. Protein Data Bank (PDB)

Protein Data Bank (PDB) a été créée au Brookhaven National Laboratoires (BNL) en 1971, considérée comme la principale réserve d'archives, à l'échelle mondiale de données structurales de macromolécules biologiques. Elle comprend des données obtenues par cristallographie aux rayons X et par spectrométrie de résonance magnétique nucléaire (RMN), soumises par des biologistes et des biochimistes du monde entier. Actuellement, la PDB est sous la responsabilité de la Banque mondiale de données sur les protéines (wwPDB), est organisée en réseau de quatre organisations :

- Research Collaboratory for Structural Bioinformatics (RCSB) PDB (États-Unis)
- PDB in Europe (PDBe) (Europe)
- PDB Japan (PDBj) (Japon)
- Biological Magnetic Resonance Data Bank (BMRB) (États-Unis)

Dont la mission est de "maintenir des archives uniques de la Banque de données sur les protéines pour les données structurales macromoléculaires qui sont librement et publiquement disponibles pour la communauté mondiale" (Joo et Shoba ,2013).

Actuellement en 2022, plus de 191565 structures macromoléculaires biologiques ont été déposées dans la PDB selon le site www.rcsb.org/stats/growth/growth-released-structures. La base de données est librement accessible à l'adresse : www.rcsb.org/pdb



1.3. UniProt

UniProt est une base de données librement accessible contenant des informations sur la séquence et les fonctions protéiques, de nombreuses entrées étant dérivées de projets de séquençage génomique. Elle contient une grande quantité d'informations sur la fonction biologique des protéines provenant de la littérature scientifique. Elle est maintenue par le consortium UniProt, qui se compose de plusieurs organisations européennes de bioinformatique et d'une fondation de Washington, DC, États-Unis (Apweiler, 2008). Toutes les données UniProt sont fournies gratuitement et sont disponibles sur le Web à l'adresse : UniProt .



1.4. Séquences génétiques et protéiques

- La Protéine régulatrice CviR et les gènes d'opéron *Vio* (*VioA*, *VioB*, *VioC*, *VioD* et *VioE*) du quorum sensing régulateur de la synthèse du pigment violacéine chez *Chromobacterium violaceum*.

Après notre analyse bibliographique, réalisé précédemment, nous avons localisé la protéine régulatrice de la biosynthèse de la violacéine CviR et les gènes de biosynthèse (*VioA*, *VioB*, *VioC*, *VioD* et *VioE*). En premier lieu nous avons effectué le téléchargement la séquence protéiques de CviR ainsi que les séquences génétiques du gènes *VioA*, *VioB*, *VioC*, *VioD* et *VioE* en date de 17

juin 2022 sur le site : [Home - Gene - NCBI \(nih.gov\)](https://www.ncbi.nlm.nih.gov/). Le travail consiste à identifier les numéros d'enregistrement (NC : numéro de concession). Sur la plate-forme NCBI nous avons pu obtenir les formats FASTA (Fast-All) de chaque séquence. Il faut savoir que le format FASTA (ou format Pearson) est un format de fichier texte utilisé pour stocker des séquences biologiques de nature nucléique ou protéique. Ces séquences sont représentées par une suite de lettres codant pour des acides nucléiques ou des acides aminés selon la nomenclature de l'Union internationale de chimie pure et appliquée (Cock et *al.*, 2009).

- CviR (AAP32919.1) (Figure 23)
- *VioA* (NC_005085.1 :c3566288-3565032) (Figure 24)
- *VioB* (NZ_JYKA01000001.1 :c193857-190861) (Figure 25)
- *VioC* (NC_005085.1 :c3561959-3560670) (Figure 26)
- *VioD* (NC_005085.1 :c3560670-3559549) (Figure 27)
- *VioE* (NC_005085.1 :c3559539-3558964) (Figure 28)

```
>AAP32919.1 CviR [Chromobacterium violaceum]
MVISKPINARPLPAGLTASQQWTLLEWIHMAGHIETENELKAFDQVLSQAPSERLLLALGRLNNQNQIQ
RLERVLNVSYPDWLDQYMKENYAQHDPILRIHLGQGPVMEERFNRAKGAEKRFIAEATQNGMGSGIT
FSAASERNNIGSILSIAGREPGNMPLSQREYDIFHWMSRGKTNWEIATILDISERTVKFHVANVIRKLNA
NNRTHAIVLGMHLAMPSTVANE
```

Figure 24 : Format FASTA de la séquence qui code pour la protéine CviR chez *C. violaceum*

```
>NC_005085.1:c3566288-3565032 Chromobacterium violaceum ATCC 12472, complete sequence
ATGAAGCATTCTCCGATATCTGCATTGTGCGCGCCGGCATCAGCGGCCTGACCTGCGCCAGCCATCTGC
TCGACTCGCCCGCTTGGCCGCGCCTGTGCTGCGCATCTTCGACATGCAGCAGGAGGCGGGCGGCCGCAT
CCGCTCGAAGATGCTGGATGGCAAGGCGTCGATAGAGCTGGGCGCGGGGCGATACTCCCCGAGCTGCAC
CCGCATTTCCAGAGCGCGATGCAGCATTACAGCCAGAAGAGCGAGGTGTATCCGTTACCCAGCTGAAAT
TCAAGAGCCATGTCCAGCAGAAGCTGAAGCGGGCGATGAACGAGTTGTGCCCAGGCTGAAAGAGCATGG
CAAGGAATCCTTTCTCCAGTTCGTCAGCCGCTACCAGGGCCATGACAGCGCGGTGGGCATGATCCGCTCC
ATGGGCTACGACGCGCTGTTCTGCCCCGACATCTCGGCCGAGATGGCCTACGACATCGTCGGCAAGCACC
CGGAAATCCAGAGCGTGACCGATAACGACGCCAACCAGTGGTTCGCGGCGGAAACGGGCTTTGCGGGCCT
GATCCAGGGCATCAAGGCCAAGGTCAAGGCTGCCGGCGCGCCTTCCAGCCTGGGTTACCGGCTGCTGTCG
GTGAGGACGGACGGCGACGGCTACCTGCTGCAACTGGCCGGCGACGACGGCTGGAAGCTGGAACACCGGA
CCCGCCATCTGATCCTGGCCATTCTCCGTCGGCGATGGCCGGGCTCAATGTCGACTTCCCCGAGGCGTG
GAGCGGCGCGCTACGGCTCGTGGCGTGTTCAGGGTTTCTCACCTACGGCGAGCCATGGTGGCTG
GACTACAAGCTGGACGACCAGGTGCTGATCGTGCACAACCCGCTGCGCAAGATCTACTTCAAGGGCGACA
AGTACCTGTTCTTACACCGACAGCGAGATGGCCAATTACTGGCGCGGCTGCGTGGCCGAAGGAGAGGA
CGGCTACCTGGAGCAGATCCGCACCCATCTGGCCAGCGCGCTGGGCATCGTTCCGAGCGCATTCCCCAG
CCCCTCGCCCATGTGCACAAGTATTGGGCGCATGGCGTGGAGTTCTGCCGCGACAGCGATATCGACCATC
CGTCCGCGCTCAGCCACCGCGACAGCGGCATCATCGCTGTTCCGACGCCTACACCGAGCACTGCGGCTG
GATGGAGGGCGGCGCTGCTCAGCGCCCGGAAGCCAGCCGCTGCTGCTGCAGCGCATCGCCGCGTGA
```

Figure 23 : Format FASTA de la séquence qui code pour le gène *VioA* chez *C. violaceum* ATCC12472

```
>NC_005085.1:c3564957-3561961 Chromobacterium violaceum ATCC 12472, complete sequence
ATGAGCATTCTGGATTTTCCACGCATCCATTTCCGCGGCTGGGCGGGTCAACGCGCCACCGCCAACC
GCGATCCGCACGGCCACATCGACATGGCCAGCAATACGGTGGCCATGGCAGGCGAACC GTTCGACCTCGC
GCGCCATCCGACCGAGTTCCACCGCCACCTGCGGTCGCTGGGGCCGCGTTTCGGCCTGGACGGCCGGGCT|
GACCCGGAAGGGCCGTTACGCTGGCCGAGGGCTACAACGCGGCGGCAACAACCATTTCTCTGGGAGA
GCGCCACCGTCAGCCACGTGCAGTGGGATGGCGGCGAAGCGGACCGCGGCGACGGCCTGGTCGGCGCCAG
GCTGGCGCTGTGGGGCATTACAACGATTACCTGCGCACCACCTTCAACCGCGCGCGCTGGGTGGACAGC
GACCCACCCGCGCGACGCGGGCGAGATCTACGCGGGCAGTTACGATCAGCCGCGCGCGCGCGGAC
CGGGACGCCCTGGCTGTTACCGCGGACATCGACGACAGCCACGCGCGCGCTGGACGCGCGGGCGGCA
CATCGCCGAGCGGGCGGCCATTTCTGGACGAGGAGTTGGCCCTGGCGCGGCTGTTCCAGTTCTCGGTG
CCCAAAGACCATCCGCACTTCTGTTCCACCGGGGCCATTGATTCCGAAGCCTGGCGCAGGCTGCAGC
TGGCGCTGGAGGACGACGACGTGCTCGGCTGACGGTGCAGTACGCGCTGTTCAATATGTCGACGCCGCC
GCAACCCAACCTCGCGGTTCACAGCATGGTCGGCGTGGTGGCCGTGGCGGCGCGGCGAAGTGGCC
AGCTACCCGCGCGGGCGGCTGCTGCGTCCGCGCCAGCCCGGGTGGGGGATCTGACGCTGCGCGTAAGCG
GCGGCGCGTGGCGTGAATCTGGCCTGCGCCATTCCGTTCTCCACCGGGCGGCGCAGCCGTCGCGCC
GGACAGGCTGACGCCGATCTCGGGGCCAAGTGCCTGGGCGACCTGCTGCTGCGCGACGAGGACGGC
GCGTTGCTGGCGGGGTGCCGAGGCGCTTTACCAGGATTACTGGACGAACCACGGCATCGTCGACCTGC
CGCTGCTGCGCGAGCCAGGGGCTCGTGACGCTGTCCAGCGAGCTGGCCGAATGGCGCGAGCAGGACTG
GGTACGCGAGTCCGACGCCCAATCTTTATTTGGAAGCGCCGGACCGCGCCACGGCCGTTTCTTTCCG
GAAAGCATCGCGTGCAGCTATTTCCGCGCGAGGCCCGCGCGCGCCGGACATTTCCACCGGATCG
AGGGGATGGGTCTGGTGGCGTGGAGTCCGCGCCAGGACGGCGATGCCCGCAATGGCGGCTGACCGGCT
GCGGCCCGGGCCGCGCGCATCGTCTGACGACGGCGCGGAGCGATCCCGCTGCGGGTGC TGCCGGAC
GACTGGGCGTTGGACGACGCGACGGTGGAGGAGTTCGATTACGCTTCTGTACCGGCGACGTGATGGCCT
ATTACGAGCTGGTCTACCCGTTATGTCCGACAAGGTGTTACGCTGGCCGACCGCTGCAAGTGCAGAC
CTACGCCAGGCTGATGTGGCAGATGTGCGATCCGAGAACCGGAACAAGAGCTACTACATGCCAGCACCC
CGCGAGCTGTGCGGCCCAAGGCCAGGCTGTTCTCAAATACCTGGCCCATGTGAGGGCCAGGCCAGGC
TGCAGGCGCCGCGCGGGCCGGGCGGCGCATCGAGAGCAAGGCCAGCTGGCGGCGGAGCTGCGCAA
GGCGGTGGATCTGGAGTTGTCGGTATGCTGCAGTACCTGTACGCCGCTATTCCATTCCCAATTACGCC
```

Figure 26 : Format FASTA de la séquence qui code pour le gène *VioB* chez *C. violaceum* ATCC12472

```
>NC_005085.1:c3561959-3560670 Chromobacterium violaceum ATCC 12472, complete sequence
ATGAAAAGAGCAATCATAGTCGGAGGCGGGCTCGCCGGCGGGCTGACCGCCATCTACCTGGCGAAGCGCG
GCTACGAGGTCCACGTGGTGGAAAAGCGCGCGACCCGCTGCGGGACCTGTCTTCTTACGTGGATGTGGT
CAGCTCGCGGGCGATAGGCGTCAGCATGACCGTGC GCGGCATCAAGTCGGTGTGGCGGCCGGCATTCCG
CGCGGGAGCTGGACGCTGCGGCGAACCATCGTGGCGATGGCGTTTTTCCGTGGCGGCCAGTACCGGA
TGCGGGAGCTCAAGCCGCTGGAGGATTTCCGCCGCTGTCGCTGAACCGCGCGGCGTTTCAGAAGCTGCT
GAACAAGTACGCCAACCTGGCCGGCGTCCGCTACTACTTCGAGCACAAAGTGCCTGGACGTGGATCTGGAC
GGCAAGTCGGTGTGATCCAGGGCAAGGACGGCCAGCCGACGCGCTTGCAGGGCGATATGATCATCGGCG
CCGACGGCGCGCACTCGGCGGTGCGGCGAGGCGATGCAGAGCGGGTTGCGCCGCTTCGAATTCAGCAGAC
TTTCTTCCGCCACGGCTACAAGACGCTGGTGTGCTGCCGACGCGCAGGCGCTGGGCTACCGCAAGGACAG
CTGTATTTCTTGGCATGGACTCCGGCGGCTGTTGCGCGGCGCGCGCCACCATCCCGGACGGCAGCG
TCAGCATCGCGGTCTGCCGTACAGCGGCGAGCCCGAGCTGACCACCACCGACGAGCCGACGATGCG
CGCCTTTTTTCGACCGTTACTTCCGGCGGCTGCGCGGGACGCGCGGACGAGATGCTGCGCCAGTTCCTG
GCCAAGCCAGCAACGACCTGATCAACGTCCGTTCCAGCACCTTCCACTACAAGGGCAATGTGCTGCTGC
TGGGCGACGCCGCCACGCCACCGCGCCTTTCTCGGCCAGGGCATGAACATGGCGCTGGAGGACGCGCG
CACCTTCGTCGAGCTGCTGGACCGCCACAGGGCGACCAGGACAAGGCCTTTCCCGAGTTCACCGAGCTG
CGCAAGGTGCAGGCCGACGCGATGCAGGACATGGCGCGGCCAACACGACGTGCTGAGTGTCCCAATC
CCATCTTTCATGCGGGCCCGCTACACCCGCTACATGCATAGCAAGTTTCCCGGCTTTTACCCGCGGGA
CATGGCGGAGAAGCTGTACTTACGTCGAGCCGTACGACAGACTGCAGCAGATCCAGAGAAAACAGAAC
GTTTGGTACAAGATAGGGAGGGTCAACTGA
```

Figure 25 : Format FASTA de la séquence qui code pour le gène *VioC* chez *C. violaceum* ATCC12472

```
>NC_005085.1:c3560670-3559549 Chromobacterium violaceum ATCC 12472, complete sequence
ATGAAGATTCTGGTCATCGGCGCGGGCCGGCCGGCCTGGTGTTCGCCAGCCAACTGAAACAGGCGCGTC
CGCTGTGGGCGATAGACATCGTCGAAAAGAACGACGAGCAGGAAGTGTGGGGTGGGGCGTGGTGTGCC
CGGCCGGCCCGCCAGCATCCGGCCAATCCGCTGTCTACCTGGACGCGCCGGAGAGGCTGAATCCGCAG
TTCCTGGAAGACTTCAAGCTGGTCCACCACAACGAGCCCAGCTGATGAGCACCGGCGTGTGTGTGCG
GCGTGGAGCGCCGCGCCTGGTGCACGCCTTGC GCGACAAGTGCCGCTCGCAGGGCATCGCCATCCGCTT
CGAATCGCCGCTGTGGAGCATGGCGAGCTGCCGCTGGCCGACTACGACCTGGTGGTGTGGCCAACGGC
GTCAATCACAAAGACCGCCACTTCACCGAGGCGCTGGTGCAGGCTGGACTACGGCCGCAACAAGTACA
TCTGGTACGGCACCAGCCAGCTGTTTCGACCAGATGAACCTGGTGTTCGCGACCCACGGCAAGGACATTTT
CATCGCCACGCCTACAAGTACTCGGACACGATGAGCACCTTCATCGTCGAGTGCAGCGAGGAGACCTAT
GCCCCGCGCCCGCTGGGCGAGATGTCCGGAAGAGCGTCGGCCGAATACGTCGCCAAGGTGTTCAGGCCG
AGCTGGGCGGCCACGGCCTGGTGTAGCCAGCCCGCCTCGGCTGGCGCAACTTCATGACCCCTGAGCCACGA
CCGCTGCCACGACGGCAAGCTGGTGTGTGTGGGCGACGCGCTGCAGTCCGGCCACTTCTCCATCGGCCAC
GGCACCACGATGGCGGTGGTGGTGGCGCAGCTGTGGTGAAGGCGCTGTGCACCGAGGACGGCGTGC CGG
CCGCGCTGAAGCGCTTCGAGGAGCGCGCCTGCCGCTGGTCCAGCTGTTCCGCGGCCATGCCGACAACAG
CCGGGTC TGGTTCGAGACGGTGGAGGAGCGCATGCACCTGTCCAGCGCCGAGTTCGTGCAGAGCTTCGAC
GCGCGCCGCAAGTCGCTGCCGCCGATGCCGGAAGCGCTGGCGCAGAACCTGCGCTACGCGCTGCAACGCT
GA
```

Figure 27 : Format FASTA de la séquence qui code pour le gène *VioD* chez *C. violaceum* ATCC12472

```
>NC_005085.1:c3559539-3558964 Chromobacterium violaceum ATCC 12472, complete sequence
ATGGAAAACCGGAACCGCCGCTGCTGCCGGCGCGCTGGAGCAGCGCCTATGTGTCTACTGGAGTCCGA
TGCTGCCGGATGACCAGCTGACGTCCGGCTACTGCTGGTTCGACTACGAGCGCGACATCTGTCCGATAGA
CGGCCTGTTCAATCCCTGGTCCGAGCGCGACACCGGCTACCGGCTGTGGATGTCCGAGGTCGGCAACGCC
GCCAGCGGCCACCTGGAAGCAGAAGGTGGCCTATGGCCGCGAGCGGACCGCCCTGGGCGAGCAGCTGT
GCGAGCGGCCGCTGGACGACGAGACCGGCCCGTTCCGCCGAGCTGTTCCGCGCGCGACGTGTGCGCCG
GCTGGGCGCCCGCCATATCGGCCCGCCGCTGGTGTGTGGCAGGGAAGCCGACGGCTGGCGCTACCAGCGT
CCGGGCAAGGGGCGCTCCACGTTGTACCTGGACGCCCGCAGCGGTACGCCGCTGAGGATGGTACCAGGGG
ACGAGGCGTGC GCGCGCTGCTGCGCGATTCCCAACGTCAGCGAGGCGGAGATTCCCGACGCCGCTCTT
CGCCGCCAAGCGCTAG
```

Figure 28 : Format FASTA de la séquence qui code pour le gène *VioE* chez *C. violaceum* ATCC12472

1.5. Structure

Les Plateformes des bases de données des protéines telles que Protein Data base (PDB) et Uniprot constitue une méthode importante pour prédire les structures des protéines car Il est important d'étudier la structure d'une protéine pour comprendre et déterminer sa fonction.

Il nécessaire juste de mettre le nom de la protéine et l'organisme producteur de cette protéine dans la case de recherche dans les plateformes citées pour avoir la structure.

Dans notre travail, nous avons utilisé ces deux banques de protéines PDB (figure 30), et Uniprot (figure 31), pour avoir la structure de la protéine régulatrice de transcription CviR chez *Chromobacterium violaceum*.

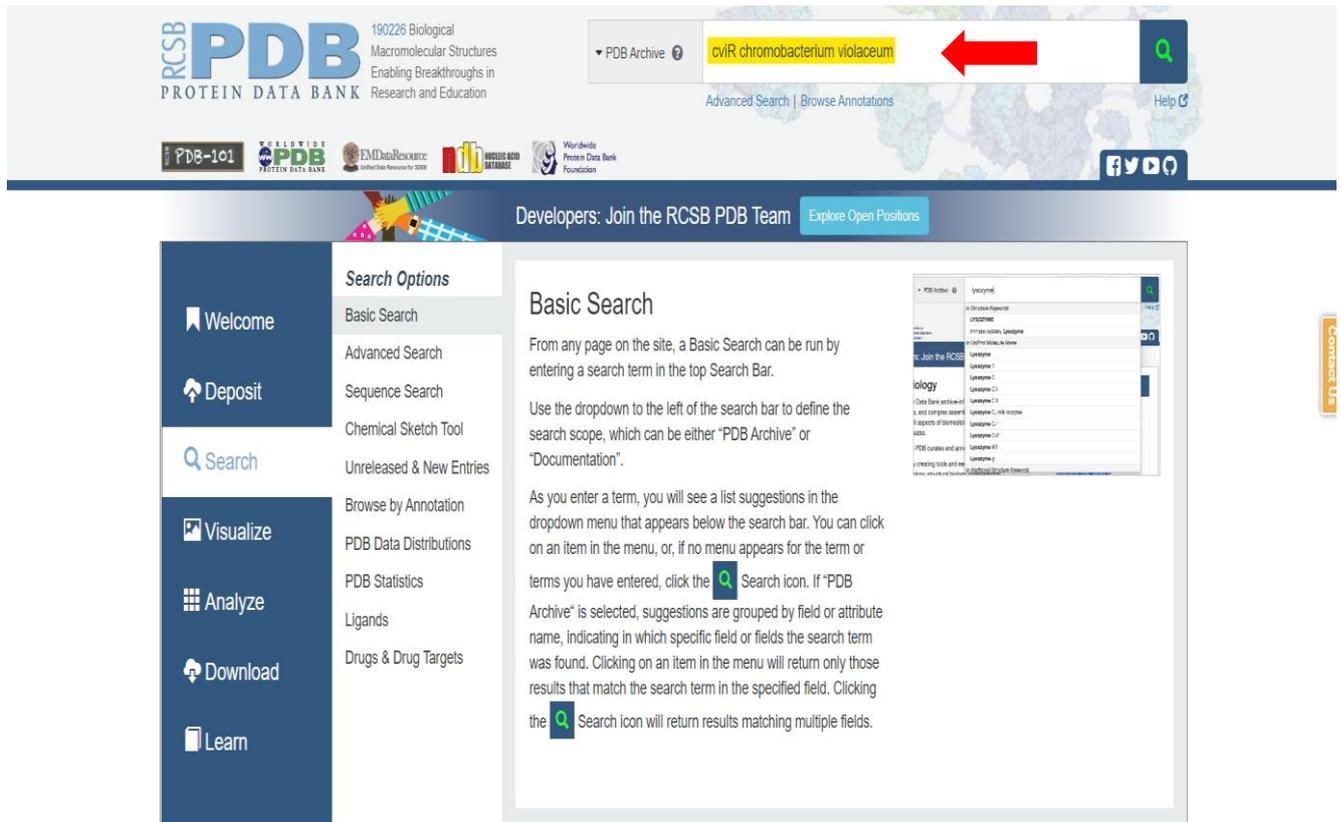
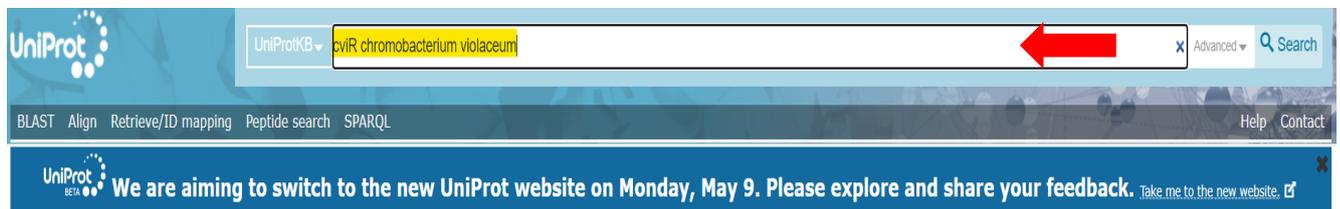


Figure 29 : Comment rechercher la structure de la protéine CviR dans PDB



The mission of UniProt is to provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information.

UniProtKB

UniProt Knowledgebase

Swiss-Prot (566,996)

Manually annotated and reviewed.

Records with information extracted from literature and curator-evaluated computational analysis.

TrEMBL (230,328,648)

Automatically annotated and not reviewed.

Records that await full manual annotation.

UniRef

The UniProt Reference Clusters (UniRef) provide clustered sets of sequences from the UniProt Knowledgebase (including isoforms) and selected UniParc records.

UniParc

UniParc is a comprehensive and non-redundant database that contains most of the publicly available protein sequences in the world.

Proteomes

A proteome is the set of proteins thought to be expressed by an organism. UniProt provides proteomes for species with completely sequenced genomes.

New UniProt portal for the latest SARS-CoV-2 coronavirus protein entries and receptors, updated independent of the general UniProt release cycle.

[View SARS-CoV-2 Proteins and Receptors](#)

News

Forthcoming changes

Planned changes for UniProt

UniProt release 2022_01

A phospholipase for clear vision | Cross-references to MANE-Select

UniProt release 2021_04

3TC: Inhibitor of the RNA polymerase

Supporting data

Literature citations Taxonomy Subcellular locations

Figure 30 : Comment rechercher la structure de la protéine CviR dans UNIPROT

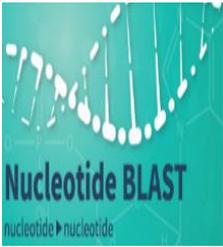
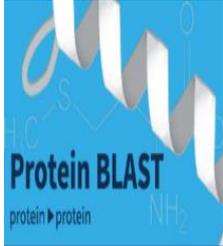
1.6. Basic Local Alignment Search Tool (BLAST)

Sert à calculer le taux de similarité entre nos séquences, c'est-à-dire le degré de parenté et Similitudes entre les séquences et les séquences stockées dans la base de données (NCBI), Nous avons utilisé le logiciel BLAST comme outil principal pour cette étude in silico.

Il y a plusieurs façons de faire une Recherche de similarité de séquence, mais la méthode la plus populaire est fiable c'est " Basic Local Alignment Search Tool " (BLAST) ou "Outil de recherche d'alignement" base local" (Madden, 2013). Le NCBI a produit la première version de BLAST vers 1990. BLAST utilise des heuristiques pour produire des résultats rapidement. Il calcule également La "valeur attendue", qui estime le nombre de correspondances qui se seraient produites à un score donné par hasard, ce qui peut aider un utilisateur à juger le degré de confiance à avoir dans un alignement (Madden, 2013).

Il existe de nombreux types de recherches BLAST, dans notre travail nous avons utilisé le BLASTp (tableau 10) pour la protéine régulatrices CviR et pour les gènes *VioA*, *VioB*, *VioC*, *VioD* et *VioE* nous avons utilisé le BLASTn (tableau 10) (Madden, 2013).

Tableau 10 : Type et fonction du BLAST utilise dans notre travail (Madden, 2013).

Programme	Type de séquence à analyser	Type de banque	Fonction
BlastN 	Nucléotidique	Nucléotidique	Comparaison d'une séquence d'interrogation nucléotidique avec une base de données de séquences nucléotidiques.
BlastP 	Protéique	Protéique	Compare une séquence de requête d'acides aminés à une base de données de séquences de protéine

2. Méthodes

2.1. Etude phylogénétique

Arbre phylogénétique, également appelé dendrogramme, est un diagramme montrant les interrelations évolutives d'un groupe d'organismes dérivés d'une forme ancestrale commune. L'ancêtre se trouve dans le "tronc" de l'arbre ; les organismes qui en sont issus sont placés aux extrémités des "branches" de l'arbre. La distance entre un groupe et les autres groupes indique le degré de parenté ; c'est-à-dire que les groupes étroitement apparentés sont situés sur des branches proches les unes des autres. Les arbres phylogénétiques, bien que spéculatifs, constituent une méthode pratique pour étudier les relations phylogénétiques (Stuessy, 1983).

Dans notre travail, on a utilisé deux méthodes basées sur la distance que le logiciel Blast utilise pour la construction des arbres phylogénétiques la méthode Neighbor-joining et la méthode d'évolution minimale. Ces méthodes-là sont développées à partir d'une méthode ancienne appelée UPGMA (Unweighted Paired Group Mean Arithmetic), Cette méthode permet de construire des arbres ultra métriques. Dans ces arbres, toutes les feuilles sont à équidistance de leur ancêtre commun. Cette propriété répond parfaitement au principe de l'horloge moléculaire qui suppose que les substitutions s'accumulent à une vitesse constante dans le temps et que les distances évolutives sont le reflet (strict) des temps géologiques qui sépare deux OTU de leur ancêtre commun (Sneath et *al.*, 1973).

- **Méthode Neighbor-joining** : méthode NJ a été initialement proposée par Saitou et Nei (1987) puis modifiée par Studier et Kepler (1988) (Pearson et *al.*, 1999). La méthode NJ est un pilier de la reconstruction de la phylogénie chez les biologistes (St. John et *al.*, 2003). Elle commence avec un arbre en étoile non résolu. Chaque paire est évaluée pour être et la somme de la longueur de toutes les branches de l'arbre résultant est calculée. L'arbre résultant. La paire qui donne la plus petite somme est considérée comme la plus proche voisine et est donc jointe. Considérée comme la plus proche voisine et est donc jointe. Une nouvelle branche est insérée entre eux et le reste de l'arbre et la longueur des branches est la longueur de la branche est recalculée. Ce processus est répété jusqu'à ce qu'il n'y ait plus qu'un seul terminal (Bhanbri et *al.*, 2017).
- **Méthode d'évolution minimale** : L'évolution minimale est une méthode de distance employée dans la modélisation phylogénétique. Elle partage avec la parcimonie maximale l'aspect de la recherche de la phylogénie qui a la somme totale des longueurs de branche la plus courte (Rzhetsky et *al.*, 1993).

Chapitre II :

Résultats et Discussion

1. Résultats

1.1. Cartographie génétique (localisation)

L'analyse in silico exprime la cartographie génétique des 5 gènes et leurs positions (localisation) par rapport aux autres gènes (figures : 31,32, 33, 34, 35).

- D'après les résultats obtenus à partir de la base de données NCBI, le gène *VioA* est situé entre 3565032k et 3566288k (figures 31).

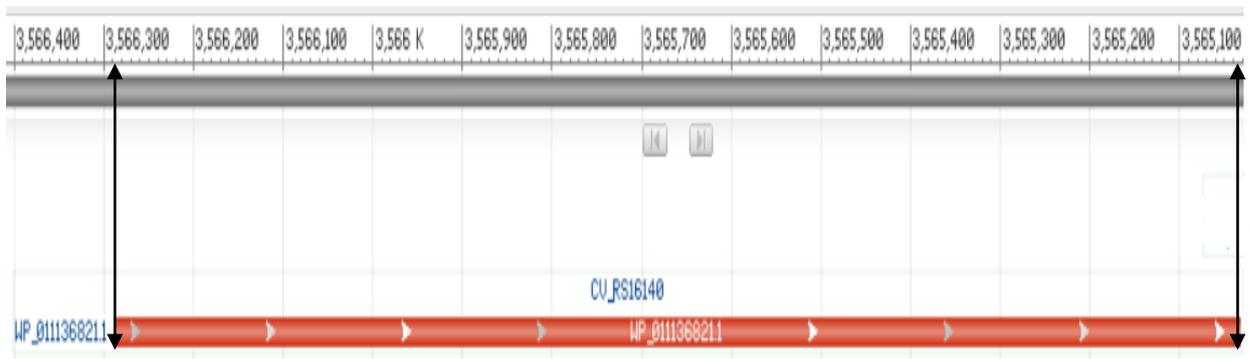


Figure 31 : Cartographie du gène *VioA*

- D'après la cartographie du gène *VioB* obtenu, ce gène est situé entre 3564958k et 3561960k (figure 32).

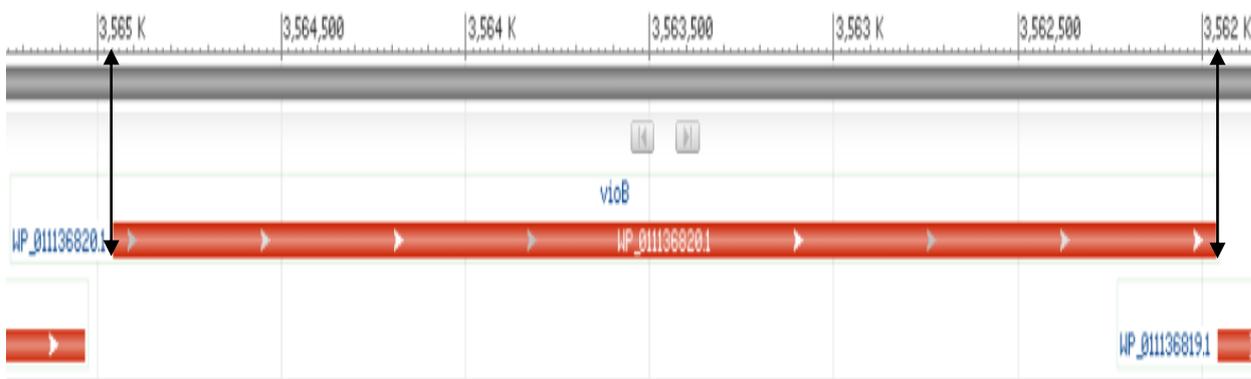


Figure 32 : Cartographie du gène *VioB*

- D'après la cartographie du gène *VioC* obtenu, ce gène est situé entre 3561959k et 3560669k (figure 34).

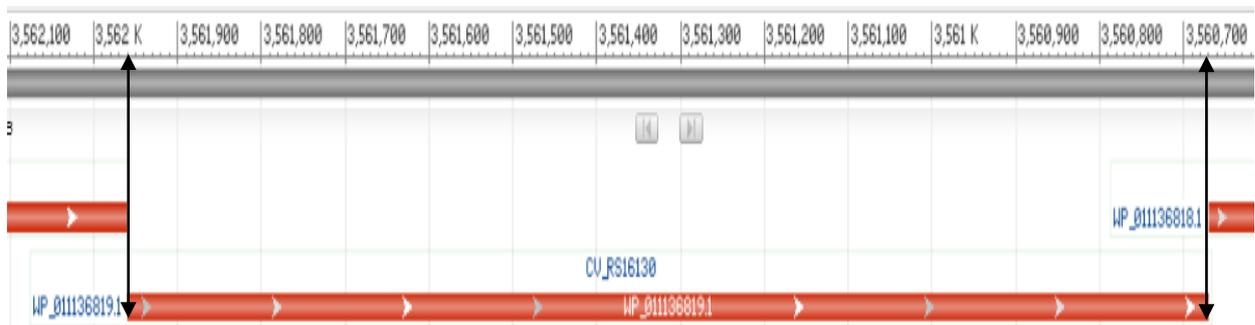


Figure 33 : Cartographie du gène *VioC*

- D'après la cartographie du gène *VioD* obtenu, ce gène est situé entre 3560673k et 3559549k (figure 35).

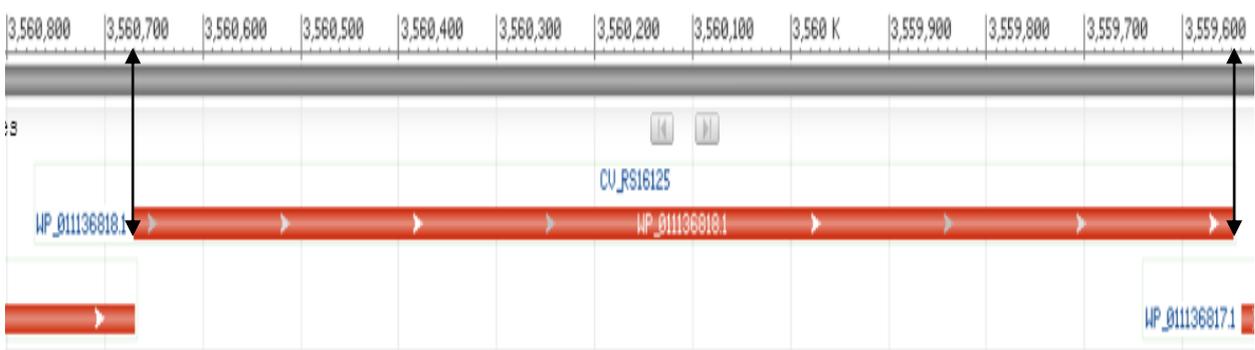


Figure 34 : Cartographie du gène *VioD*

- D'après la cartographie du gène *VioE* obtenu, ce gène est situé entre 3559540k et 3558964k (figure 36).



Figure 35 : Cartographie du gène *VioE*

1.2. Structure

Avec l'aide de la plateforme PDB et Uniprot, nous sommes capables de prédire la structure 3D de la protéine régulatrice CviR lié à son autoinducteur C10-HSL chez *C. violaceum* 12472 (figure 37).

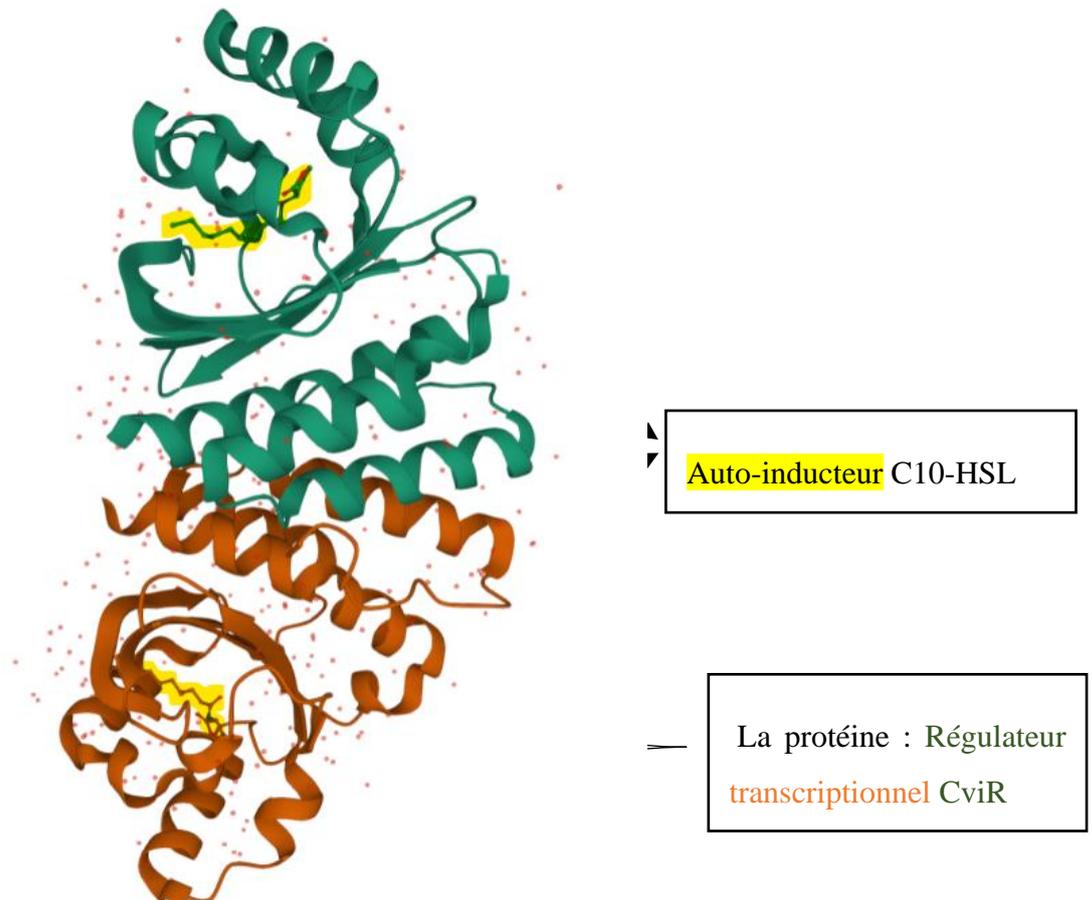


Figure 36 : **A :** Structure 3D de la protéine régulatrice CviR lié à son auto-inducteur chez *C. violaceum* 12472 prises à partir du PDB. **B :** Structure 3D de la protéine régulatrice CviR lié à son auto-inducteur chez *C. violaceum* ATCC12472 prise à partir de l'UNIPROT.

1.3. Bilan BLAST

Après avoir exécuté BLASTp pour la protéine régulatrice CviR et BLASTn pour les cinq gènes *Vio* (A, B, C, D et E) nous avons obtenu des listes d'un ensemble de séquences similaires à notre requête, avec leurs pourcentages d'identité et numéros d'accèsion.

✚ D'après les résultats obtenus à partir du BLASTp :

- Pour la protéine régulatrice CviR les séquences obtenues présentent des pourcentages d'identité très varié : pourcentage le plus élevé est **100%** et le pourcentage le plus bas est **28,49%**.

Description	Scientific Name	Per. Ident	Accession
CviR [Chromobacterium violaceum]	Chromoba...	100.00%	AAP32919.1
LuxR family transcriptional regulator [Chromobacterium]	Chromoba...	87.92%	WP_052257945.1
LuxR family transcriptional regulator [Chromobacterium subtsugae]	Chromoba...	87.55%	WP_220154015.1
LuxR family transcriptional regulator [Chromobacterium subtsugae]	Chromoba...	87.55%	WP_065202624.1
LuxR family transcriptional regulator [Chromobacterium sphagnoi]	Chromoba...	82.64%	WP_071114352.1
LuxR family transcriptional regulator [Chromobacterium sp. ATCC 53434]	Chromoba...	80.75%	WP_101709717.1
LuxR family transcriptional regulator [Chromobacterium]	Chromoba...	79.25%	WP_166440056.1
hypothetical protein DK842_11120 [Chromobacterium phragmitis]	Chromoba...	76.23%	AXE30399.1
LuxR family transcriptional regulator [Chromobacterium phragmitis]	Chromoba...	76.23%	WP_168191878.1
LuxR family transcriptional regulator [Chromobacterium violaceum]	Chromoba...	80.78%	WP_210543942.1
LuxR family transcriptional regulator [Chromobacterium violaceum]	Chromoba...	76.52%	WP_080509080.1
LuxR family transcriptional regulator [Chromobacterium violaceum]	Chromoba...	76.52%	WP_080762180.1
LuxR family transcriptional regulator [Chromobacterium violaceum]	Chromoba...	76.52%	WP_011137637.1
LuxR family transcriptional regulator [Chromobacterium violaceum]	Chromoba...	76.14%	WP_081571415.1
LuxR family transcriptional regulator [Chromobacterium paludis]	Chromoba...	75.85%	WP_168209814.1
LuxR family transcriptional regulator [Chromobacterium paludis]	Chromoba...	75.85%	QEL57650.1
LuxR family transcriptional regulator [Chromobacterium vaccini]	Chromoba...	75.76%	WP_081665865.1
Transcriptional activator protein luxR [Chromobacterium violaceum]	Chromoba...	76.14%	STB69171.1
LuxR family transcriptional regulator [Chromobacterium violaceum]	Chromoba...	76.14%	WP_084194192.1
LuxR family transcriptional regulator [Chromobacterium vaccini]	Chromoba...	75.76%	WP_186338552.1
LuxR family transcriptional regulator [Chromobacterium violaceum]	Chromoba...	76.14%	WP_081527328.1
LuxR family transcriptional regulator [Chromobacterium violaceum]	Chromoba...	76.14%	WP_204086452.1
LuxR family transcriptional regulator [Chromobacterium sp. ASV23]	Chromoba...	75.47%	WP_199154190.1
LuxR family transcriptional regulator [Chromobacterium sp. MWU14-2602]	Chromoba...	76.23%	WP_158253730.1
hypothetical protein C2119_15350 [Chromobacterium sp. MWU14-2602]	Chromoba...	76.23%	POZ61066.1
LuxR family transcriptional regulator [Chromobacterium]	Chromoba...	72.83%	WP_071110073.1
hypothetical protein C2134_09615 [Chromobacterium sp. MWU13-2610]	Chromoba...	72.83%	POA98831.1
LuxR family transcriptional regulator [Chromobacterium phragmitis]	Chromoba...	72.45%	MBM2886301.1
LuxR family transcriptional regulator [Chromobacterium sp. ASV5]	Chromoba...	73.15%	WP_088739358.1
LuxR family transcriptional regulator [Chromobacterium]	Chromoba...	68.34%	WP_081862609.1
LuxR family transcriptional regulator [Chromobacterium haemolyticum]	Chromoba...	68.34%	WP_081546380.1
LuxR family transcriptional regulator [Chromobacterium]	Chromoba...	67.95%	WP_107797968.1
LuxR family transcriptional regulator [Chromobacterium]	Chromoba...	67.95%	WP_118266473.1
LuxR family transcriptional regulator [Chromobacterium]	Chromoba...	67.05%	WP_114153105.1
LuxR family transcriptional regulator [Chromobacterium]	Chromoba...	66.67%	WP_043594052.1
LuxR family transcriptional regulator [Chromobacterium]	Chromoba...	67.05%	WP_082150892.1
hypothetical protein B0T40_20665 [Chromobacterium haemolyticum]	Chromoba...	67.95%	OQS32060.1
LuxR family transcriptional regulator [Chromobacterium alkanivorans]	Chromoba...	66.29%	WP_205465812.1
hypothetical protein UF16_14965 [Chromobacterium violaceum]	Chromoba...	74.04%	KJH66595.1
hypothetical protein B0T37_21120 [Chromobacterium violaceum]	Chromoba...	73.62%	OQS20670.1
LuxR family transcriptional regulator [Chromobacterium vaccini]	Chromoba...	73.19%	WP_021475642.1
hypothetical protein BS642_16600 [Chromobacterium violaceum]	Chromoba...	73.62%	OLZ76252.1
hypothetical protein B0T49_14100 [Chromobacterium violaceum]	Chromoba...	73.62%	OQS49455.1
LuxR family transcriptional regulator [Chromobacterium piscinae]	Chromoba...	70.00%	WP_052247053.1
hypothetical protein VK98_09180 [Chromobacterium sp. LK11]	Chromoba...	67.37%	KMN82264.1
hypothetical protein VI26_04930 [Chromobacterium sp. LK1]	Chromoba...	65.98%	KMN37033.1
LuxR family transcriptional regulator [Chromobacterium haemolyticum]	Chromoba...	66.95%	WP_176218509.1
Crystal structure of CviR ligand-binding domain bound to the native ligand C6...	Chromoba...	100.00%	3QP1_A
Transcriptional activator protein lasR [Chromobacterium violaceum]	Chromoba...	74.26%	YEB45420.1
Crystal structure of CviR (Chromobacterium violaceum 12472) ligand-binding...	Chromoba...	84.62%	3QP8_A
LuxR C-terminal-related transcriptional regulator [Aquitalea denitrificans]	Aquitalea...	53.85%	WP_159874843.1
helix-turn-helix transcriptional regulator [Vogesella urethralis]	Vogesella u...	78.79%	WP_206210867.1
helix-turn-helix transcriptional regulator [Vogesella oryzae]	Vogesella o...	34.55%	WP_174874860.1
Response regulator containing a CheY-like receiver domain and an HTH DNA...	Vogesella s...	36.19%	SKK15977.1
Uncharacterised protein [Chromobacterium violaceum]	Chromobac...	92.16%	YEB45419.1
helix-turn-helix transcriptional regulator [Trinickia dinghuensis]	Trinickia d...	60.71%	WP_115536331.1
hypothetical protein C0Z18_02350 [Trinickia dabaoshanensis]	Trinickia da...	58.18%	PMS23080.1
helix-turn-helix transcriptional regulator [Methylophilaceae bacterium]	Methylophil...	46.48%	QYL45268.1
LuxR family transcriptional regulator [Halomonas sp. CSM-2]	Halomonas...	28.49%	WP_085920098.1

Figure 37 : Séquences produisant des alignements CviR significatifs

✚ D'après les résultats obtenus à partir du BLASTn :

- Pour le gène *VioA* les séquences obtenues présentent des pourcentages d'identité très varié : le pourcentage le plus élevé est **100%** et le pourcentage le plus bas est **81,99%**.

	Description	Scientific Name	Per. Ident	Accession
<input checked="" type="checkbox"/>	Chromobacterium violaceum ATCC 12472, complete genome	Chromoba...	100.00%	AE016825.1
<input checked="" type="checkbox"/>	Chromobacterium violaceum violacein biosynthetic gene cluster (vio A, vio B, y...	Chromoba...	100.00%	AB032799.1
<input checked="" type="checkbox"/>	Chromobacterium violaceum strain CV1197 chromosome	Chromoba...	99.76%	CP024029.1
<input checked="" type="checkbox"/>	Chromobacterium violaceum strain CV1192 chromosome	Chromoba...	99.76%	CP024028.1
<input checked="" type="checkbox"/>	Chromobacterium violaceum strain FDAARGOS_635 chromosome, complete g...	Chromoba...	99.52%	CP050992.1
<input checked="" type="checkbox"/>	Chromobacterium violaceum strain NCTC9695 genome assembly, chromosom...	Chromoba...	99.36%	LR134182.1
<input checked="" type="checkbox"/>	Chromobacterium violaceum violacein biosynthetic gene cluster, complete seq...	Chromoba...	99.28%	AF172851.1
<input checked="" type="checkbox"/>	Chromobacterium violaceum strain FDAARGOS_1273 chromosome, complete...	Chromoba...	99.05%	CP069587.1
<input checked="" type="checkbox"/>	Chromobacterium violaceum strain FDAARGOS_1274 chromosome, complete...	Chromoba...	99.05%	CP069442.1
<input checked="" type="checkbox"/>	Chromobacterium vaccinii strain 21-1, complete genome	Chromoba...	89.35%	CP017707.1
<input checked="" type="checkbox"/>	Chromobacterium vaccinii strain GIMC1602:ChrSima_v chromosome	Chromoba...	89.35%	CP060046.1
<input checked="" type="checkbox"/>	Chromobacterium vaccinii strain GIMC1601:ChrSima_w chromosome	Chromoba...	89.35%	CP060044.1
<input checked="" type="checkbox"/>	Chromobacterium vaccinii strain XC0014 chromosome, complete genome	Chromoba...	89.19%	CP022344.1
<input checked="" type="checkbox"/>	Chromobacterium phragmitis strain IIBBL_112-1 chromosome, complete geno...	Chromoba...	88.22%	CP029495.1
<input checked="" type="checkbox"/>	Chromobacterium phragmitis strain IIBBL_274-1 chromosome, complete geno...	Chromoba...	88.06%	CP029554.1
<input checked="" type="checkbox"/>	Chromobacterium paludis strain IIBBL_257-1 chromosome, complete genome	Chromoba...	87.64%	CP043473.1
<input checked="" type="checkbox"/>	Chromobacterium violaceum DNA, violacein biosynthesis related genes, compl...	Chromoba...	81.99%	LC000628.1

Figure 38 : Séquences produisant des alignements *VioA* significatifs

- Pour le gène *VioB* les séquences obtenues présentent des pourcentages d'identité très varié : le pourcentage le plus élevé est **100%** et le pourcentage le plus bas est **71,30%**.

	Description	Scientific Name	Per. Ident	Accession
✓	Chromobacterium violaceum ATCC 12472, complete genome	Chromobac...	100.00%	AE016825.1
✓	Chromobacterium violaceum violacein biosynthetic gene cluster (vio A, vio B, vio...	Chromobac...	99.97%	AB032799.1
✓	Chromobacterium violaceum violacein biosynthetic gene cluster, complete seque...	Chromobac...	99.00%	AF172851.1
✓	Chromobacterium violaceum strain FDAARGOS_635 chromosome, complete gen...	Chromobac...	98.63%	CP050992.1
✓	Chromobacterium violaceum strain FDAARGOS_1273 chromosome, complete ge...	Chromobac...	98.53%	CP069587.1
✓	Chromobacterium violaceum strain FDAARGOS_1274 chromosome, complete ge...	Chromobac...	98.53%	CP069442.1
✓	Chromobacterium violaceum strain NCTC9695 genome assembly, chromosome: 1	Chromobac...	98.53%	LR134182.1
✓	Chromobacterium violaceum strain CV1197 chromosome	Chromobac...	98.06%	CP024029.1
✓	Chromobacterium violaceum strain CV1192 chromosome	Chromobac...	98.06%	CP024028.1
✓	Chromobacterium vaccinii strain XC0014 chromosome, complete genome	Chromobac...	92.74%	CP022344.1
✓	Chromobacterium vaccinii strain GIMC1602:ChrSima_v chromosome	Chromobac...	92.24%	CP060046.1
✓	Chromobacterium vaccinii strain GIMC1601:ChrSima_w chromosome	Chromobac...	92.24%	CP060044.1
✓	Chromobacterium vaccinii strain 21-1, complete genome	Chromobac...	92.07%	CP017707.1
✓	Chromobacterium phragmitis strain IIBL 112-1 chromosome, complete genome	Chromobac...	89.93%	CP029495.1
✓	Chromobacterium phragmitis strain IIBL 274-1 chromosome, complete genome	Chromobac...	89.89%	CP029554.1
✓	Chromobacterium paludis strain IIBL 257-1 chromosome, complete genome	Chromobac...	88.42%	CP043473.1
✓	Cloning vector pTHSSe_71, complete sequence	Cloning vect...	81.27%	MH029788.1
✓	Cloning vector pTHSSe_70, complete sequence	Cloning vect...	81.27%	MH029787.1
✓	Cloning vector pET15b-vioB, complete sequence	cloning vect...	81.27%	KX461960.1
✓	Chromobacterium violaceum DNA, violacein biosynthesis related genes, complet...	Chromobac...	79.42%	LC000628.1
✓	Massilia sp. NR 4-1, complete genome	Massilia sp...	74.54%	CP012201.1
✓	Janthinobacterium sp. HH102 chromosome, complete genome	Janthinoba...	77.89%	CP062062.1
✓	Massilia sp. MB5 chromosome, complete genome	Massilia sp...	78.00%	CP092780.1
✓	Janthinobacterium lividum strain EIF2 chromosome, complete genome	Janthinoba...	77.50%	CP049828.1
✓	Janthinobacterium lividum DNA, violacein biosynthesis related genes, complete c...	Janthinoba...	77.43%	LC000629.1
✓	Janthinobacterium lividum contig 186, complete sequence	Janthinoba...	77.43%	DQ074977.1
✓	Janthinobacterium lividum strain EIF1 chromosome, complete genome	Janthinoba...	77.31%	CP048832.1
✓	Massilia violaceinigra strain B2 chromosome	Massilia vio...	77.29%	CP024608.1
✓	Duganella sp. B2 violacein biosynthetic gene cluster, complete sequence	Duganella s...	77.63%	GQ266676.1
✓	Janthinobacterium sp. 1_2014MBL_MicDiv chromosome, complete genome	Janthinoba...	77.74%	CP011319.1
✓	Janthinobacterium sp. PLB04 chromosome	Janthinoba...	77.07%	CP088968.1
✓	Janthinobacterium sp. PLB02 chromosome	Janthinoba...	77.07%	CP071520.1
✓	Janthinobacterium lividum strain BP01 clone pBPpur genomic sequence	Janthinoba...	76.90%	EF063591.1
✓	Collimonas sp. MPS11E8 putative periplasmic phosphate/phosphonate transpor...	Collimonas ...	76.51%	FJ965838.1
✓	Saccharothrix algeriensis strain NRRL B-24137 chromosome	Saccharothr...	76.09%	CP072788.1
✓	Uncultured bacterium clone NM343 methylarcyriarubin tryptophan dimer gene cl...	uncultured...	74.34%	KF551863.1
✓	Uncultured bacterium clone AB234 erdasporine-like tryptophan dimer gene cluste...	uncultured...	72.66%	KF551864.1
✓	Uncultured bacterium clone AR1973 erdasporine-like tryptophan dimer gene clus...	uncultured...	78.32%	KF551873.1
✓	Streptomyces venezuelae strain ATCC 15068 chromosome, complete genome	Streptomyc...	80.11%	CP029194.1
✓	Uncultured bacterium clone AZ25P204 genomic sequence	uncultured...	71.30%	KP274855.1
✓	Sphingopyxis macrogoltabida strain 203N, complete genome	Sphingopyxi...	100.00%	CP013344.1
✓	Sphingopyxis macrogoltabida strain 203 chromosome, complete genome	Sphingopyxi...	100.00%	CP009429.1

Figure 39 : Séquences produisant des alignements *VioB* significatifs

- Pour le gène *VioC* les séquences obtenues présentent des pourcentages d'identité très varié : le pourcentage le plus élevé est 100% et le pourcentage le plus bas est 73,35%.

	Description	Scientific Name	Per. Ident	Accession
✓	Chromobacterium violaceum ATCC 12472, complete genome	Chromobac...	100.00%	AF016825.1
✓	Chromobacterium violaceum violacein biosynthetic gene cluster (vio A, vio B...	Chromobac...	100.00%	AB032799.1
✓	Chromobacterium violaceum strain FDAARGOS_1273 chromosome, complet...	Chromobac...	99.53%	CP069587.1
✓	Chromobacterium violaceum strain FDAARGOS_1274 chromosome, complet...	Chromobac...	99.53%	CP069442.1
✓	Chromobacterium violaceum strain CV1197 chromosome	Chromobac...	99.38%	CP024029.1
✓	Chromobacterium violaceum strain CV1192 chromosome	Chromobac...	99.38%	CP024028.1
✓	Chromobacterium violaceum violacein biosynthetic gene cluster, complete s...	Chromobac...	99.22%	AF172851.1
✓	Chromobacterium violaceum strain FDAARGOS_635 chromosome, complete...	Chromobac...	99.15%	CP050992.1
✓	Chromobacterium violaceum strain NCTC9695 genome assembly, chromoso...	Chromobac...	98.84%	LR134182.1
✓	Chromobacterium vaccinii strain XC0014 chromosome, complete genome	Chromobac...	94.42%	CP022344.1
✓	Chromobacterium paludis strain IIBBL_257-1 chromosome, complete genome	Chromobac...	94.12%	CP043473.1
✓	Chromobacterium vaccinii strain GIMC1602:ChrSima_v chromosome	Chromobac...	93.18%	CP060046.1
✓	Chromobacterium vaccinii strain GIMC1601:ChrSima_w chromosome	Chromobac...	93.18%	CP060044.1
✓	Chromobacterium vaccinii strain 21-1, complete genome	Chromobac...	93.10%	CP017707.1
✓	Chromobacterium phragmitis strain IIBBL_112-1 chromosome, complete gen...	Chromobac...	91.96%	CP029495.1
✓	Chromobacterium phragmitis strain IIBBL_274-1 chromosome, complete gen...	Chromobac...	91.65%	CP029554.1
✓	Chromobacterium violaceum DNA, violacein biosynthesis related genes, com...	Chromobac...	86.04%	LC000628.1
✓	Cloning vector pET15b-vioC, complete sequence	cloning_vec...	82.43%	KX461961.1
✓	Cloning vector pTU3-A-RebFHOD_VioCDE_KanR, complete sequence	Cloning_vec...	82.43%	MW388777.1
✓	Massilia sp. NR 4-1, complete genome	Massilia sp...	78.29%	CP012201.1
✓	Massilia sp. MB5 chromosome, complete genome	Massilia sp...	78.14%	CP092780.1
✓	Duganella sp. ZLP-XI violacein biosynthesis gene cluster, complete sequence	Duganella s...	78.19%	KJ131413.1
✓	Massilia sp. CCM 8941 chromosome, complete genome	Massilia an...	77.79%	CP065053.1
✓	Collimonas sp. MPS11E8 putative periplasmic phosphate/phosphonate tran...	Collimonas...	77.36%	FJ965838.1
✓	Uncultured bacterium violacein gene cluster, complete sequence	uncultured_...	77.25%	AF367409.1
✓	Janthinobacterium agaricidamnosum NBRC_102515 = DSM_9628, complete ...	Janthinoba...	76.85%	HG322949.1
✓	Janthinobacterium lividum strain EIF2 chromosome, complete genome	Janthinoba...	76.88%	CP049828.1
✓	Janthinobacterium lividum strain EIF1 chromosome, complete genome	Janthinoba...	76.80%	CP048832.1
✓	Janthinobacterium sp. PLB04 chromosome	Janthinoba...	76.78%	CP088968.1
✓	Janthinobacterium sp. PLB02 chromosome	Janthinoba...	76.78%	CP071520.1
✓	Janthinobacterium lividum DNA, violacein biosynthesis related genes, compl...	Janthinoba...	76.72%	LC000629.1
✓	Janthinobacterium lividum contig 186, complete sequence	Janthinoba...	76.72%	DQ074977.1
✓	Janthinobacterium lividum DNA, violacein biosynthesis related genes, compl...	Janthinoba...	76.57%	LC000630.1
✓	Janthinobacterium sp. LM6, complete genome	Janthinoba...	75.68%	CP019510.1
✓	Myxococcus stipitatus DSM_14675, complete genome	Myxococcu...	75.59%	CP004025.1
✓	Myxococcus xanthus strain GH5.1.9c20 chromosome, complete genome	Myxococcu...	73.60%	CP017170.1
✓	Myxococcus xanthus strain KF4.3.9c1 chromosome, complete genome	Myxococcu...	73.60%	CP017172.1
✓	Myxococcus xanthus strain GH3.5.6c2 chromosome, complete genome	Myxococcu...	73.47%	CP017169.1
✓	Myxococcus xanthus strain MC3.3.5c16 chromosome, complete genome	Myxococcu...	73.35%	CP017173.1
✓	Myxococcus xanthus strain MC3.5.9c15 chromosome, complete genome	Myxococcu...	73.35%	CP017174.1
✓	Arenimonas daejeonensis strain KACC_14631 chromosome	Arenimona...	85.90%	CP042453.1
✓	Beauveria bassiana ARSEF_2860 kynurenine 3-monooxygenase partial mRNA	Beauveria b...	100.00%	XM_008596699.1
✓	Beauveria bassiana strain HN6 chromosome 3	Beauveria b...	100.00%	CP045884.1

Figure 40 : Séquences produisant des alignements *VioC* significatifs

- Pour le gène *VioD* les séquences obtenues présentent des pourcentages d'identité très varié : le pourcentage le plus élevé est **100%** et le pourcentage le plus bas est **76,15 %**.

	Description	Scientific Name	Per. Ident	Accession
✓	Chromobacterium violaceum ATCC 12472, complete genome	Chromobac...	100.00%	AE016825.1
✓	Chromobacterium violaceum violacein biosynthetic gene cluster (vio A, vio B, vio...	Chromobac...	100.00%	AB032799.1
✓	Chromobacterium violaceum strain CV1197 chromosome	Chromobac...	99.73%	CP024029.1
✓	Chromobacterium violaceum strain CV1192 chromosome	Chromobac...	99.73%	CP024028.1
✓	Chromobacterium violaceum strain FDAARGOS_635 chromosome, complete gen...	Chromobac...	99.64%	CP050992.1
✓	Chromobacterium violaceum strain NCTC9695 genome assembly, chromosome: 1	Chromobac...	99.55%	LR134182.1
✓	Chromobacterium violaceum strain FDAARGOS_1273 chromosome, complete ge...	Chromobac...	99.29%	CP069587.1
✓	Chromobacterium violaceum strain FDAARGOS_1274 chromosome, complete ge...	Chromobac...	99.29%	CP069442.1
✓	Chromobacterium violaceum violacein biosynthetic gene cluster, complete seque...	Chromobac...	99.29%	AF172851.1
✓	Chromobacterium paludis strain IIBBL_257-1 chromosome, complete genome	Chromobac...	96.52%	CP043473.1
✓	Chromobacterium vaccinii strain XC0014 chromosome, complete genome	Chromobac...	95.72%	CP022344.1
✓	Chromobacterium vaccinii strain GIMC1602:ChrSima_v chromosome	Chromobac...	94.30%	CP060046.1
✓	Chromobacterium vaccinii strain GIMC1601:ChrSima_w chromosome	Chromobac...	94.30%	CP060044.1
✓	Chromobacterium vaccinii strain 21-1, complete genome	Chromobac...	94.21%	CP017707.1
✓	Chromobacterium phragmitis strain IIBBL_274-1 chromosome, complete genome	Chromobac...	93.67%	CP029554.1
✓	Chromobacterium phragmitis strain IIBBL_112-1 chromosome, complete genome	Chromobac...	93.58%	CP029495.1
✓	Chromobacterium violaceum DNA, violacein biosynthesis related genes, complet...	Chromobac...	85.40%	LC000628.1
✓	Massilia sp. NR 4-1, complete genome	Massilia sp...	78.73%	CP012201.1
✓	Massilia sp. MB5 chromosome, complete genome	Massilia sp...	78.73%	CP092780.1
✓	Collimonas sp. MPS11E8 putative periplasmic phosphate/phosphonate transport...	Collimonas...	77.03%	FJ965838.1
✓	Janthinobacterium sp. LM6, complete genome	Janthinoba...	79.36%	CP019510.1
✓	Janthinobacterium lividum strain EIF1 chromosome, complete genome	Janthinoba...	79.08%	CP048832.1
✓	Janthinobacterium lividum DNA, violacein biosynthesis related genes, complete c...	Janthinoba...	79.11%	LC000630.1
✓	Janthinobacterium agaricidamosum strain BHSEK chromosome, complete geno...	Janthinoba...	77.48%	CP033019.1
✓	Janthinobacterium lividum strain EIF2 chromosome, complete genome	Janthinoba...	78.86%	CP049828.1
✓	Janthinobacterium lividum DNA, violacein biosynthesis related genes, complete c...	Janthinoba...	78.86%	LC000629.1
✓	Janthinobacterium sp. PLB04 chromosome	Janthinoba...	78.91%	CP088968.1
✓	Janthinobacterium sp. PLB02 chromosome	Janthinoba...	78.91%	CP071520.1
✓	Massilia sp. CCM 8941 chromosome, complete genome	Massilia ant...	76.67%	CP065053.1
✓	Janthinobacterium lividum contig 186, complete sequence	Janthinoba...	78.86%	DQ074977.1
✓	Duganella sp. ZLP-XI violacein biosynthesis gene cluster, complete sequence	Duganella s...	76.49%	KJ131413.1
✓	Janthinobacterium sp. HH102 chromosome, complete genome	Janthinoba...	76.15%	CP062062.1
✓	Janthinobacterium agaricidamosum NBRC 102515 = DSM 9628, complete geno...	Janthinoba...	76.26%	HG322949.1
✓	Janthinobacterium lividum strain BP01 clone pBPpur genomic sequence	Janthinoba...	76.19%	EF063591.1
✓	Massilia violaceinigra strain B2 chromosome	Massilia vio...	76.83%	CP024608.1
✓	Duganella sp. B2 violacein biosynthetic gene cluster, complete sequence	Duganella s...	76.83%	GQ266676.1
✓	Janthinobacterium sp. 1_2014MBL_MicDiv chromosome, complete genome	Janthinoba...	77.88%	CP011319.1
✓	Uncultured bacterium contig00106 genomic sequence	uncultured ...	100.00%	JX827796.1
✓	Rhizobium tropici CIAT 899 plasmid pRtrCIAT899c, complete sequence	Rhizobium t...	100.00%	CP004018.1

Figure 41 : Séquences produisant des alignements *VioD* significatifs

- Pour le gène *VioE* les séquences obtenues présentent des pourcentages d'identité très varié : le pourcentage le plus élevé est **100%** et le pourcentage le plus bas est **91,33 %**.

	Description	Scientific Name	Per. Ident	Accession
<input checked="" type="checkbox"/>	Chromobacterium violaceum ATCC 12472, complete genome	Chromobac...	100.00%	AE016825.1
<input checked="" type="checkbox"/>	Chromobacterium violaceum violacein biosynthetic gene cluster (vio A, vio B, vio ...	Chromobac...	100.00%	AB032799.1
<input checked="" type="checkbox"/>	Cloning vector pET21-vioE, complete sequence	cloning vect...	98.95%	KX461963.1
<input checked="" type="checkbox"/>	Chromobacterium violaceum strain NCTC9695 genome assembly, chromosome: 1	Chromobac...	98.09%	LR134182.1
<input checked="" type="checkbox"/>	Chromobacterium violaceum strain CV1197 chromosome	Chromobac...	97.92%	CP024029.1
<input checked="" type="checkbox"/>	Chromobacterium violaceum strain CV1192 chromosome	Chromobac...	97.92%	CP024028.1
<input checked="" type="checkbox"/>	Chromobacterium violaceum strain FDAARGOS_635 chromosome, complete gen...	Chromobac...	97.57%	CP050992.1
<input checked="" type="checkbox"/>	Chromobacterium violaceum strain FDAARGOS_1273 chromosome, complete ge...	Chromobac...	97.22%	CP069587.1
<input checked="" type="checkbox"/>	Chromobacterium violaceum strain FDAARGOS_1274 chromosome, complete ge...	Chromobac...	97.22%	CP069442.1
<input checked="" type="checkbox"/>	Chromobacterium violaceum violacein biosynthetic gene cluster, complete seque...	Chromobac...	95.00%	AF172851.1
<input checked="" type="checkbox"/>	Chromobacterium vaccinii strain GIMC1602:ChrSima_v chromosome	Chromobac...	94.27%	CP060046.1
<input checked="" type="checkbox"/>	Chromobacterium vaccinii strain GIMC1601:ChrSima_w chromosome	Chromobac...	94.27%	CP060044.1
<input checked="" type="checkbox"/>	Chromobacterium vaccinii strain XC0014 chromosome, complete genome	Chromobac...	93.92%	CP022344.1
<input checked="" type="checkbox"/>	Chromobacterium vaccinii strain 21-1, complete genome	Chromobac...	93.92%	CP017707.1
<input checked="" type="checkbox"/>	Chromobacterium phragmitis strain IIBBL 274-1 chromosome, complete genome	Chromobac...	93.15%	CP029554.1
<input checked="" type="checkbox"/>	Chromobacterium phragmitis strain IIBBL 112-1 chromosome, complete genome	Chromobac...	93.15%	CP029495.1
<input checked="" type="checkbox"/>	Chromobacterium paludis strain IIBBL 257-1 chromosome, complete genome	Chromobac...	91.33%	CP043473.1

Figure 42 : Séquences produisant des alignements *VioE* significatifs

1.4. Arbres phylogénétiques

Après l'obtention des résultats à l'aide des logiciels BLASTp et BLASTn pour la protéine régulatrice CviR et les gènes *Vio* (*A*, *B*, *C*, *D*, *E*), des dendrogrammes ont été générés (Figures 43, 44, 45, 46, 47 et 48).

- L'arbre phylogénétique de la protéine régulatrice CviR du quorum sensing chez *C. violaceum* a été construit à l'aide de la plate-forme NCBI et logiciel BLASTp en utilisant la méthode neighbor joining (NJ) (figure 43).

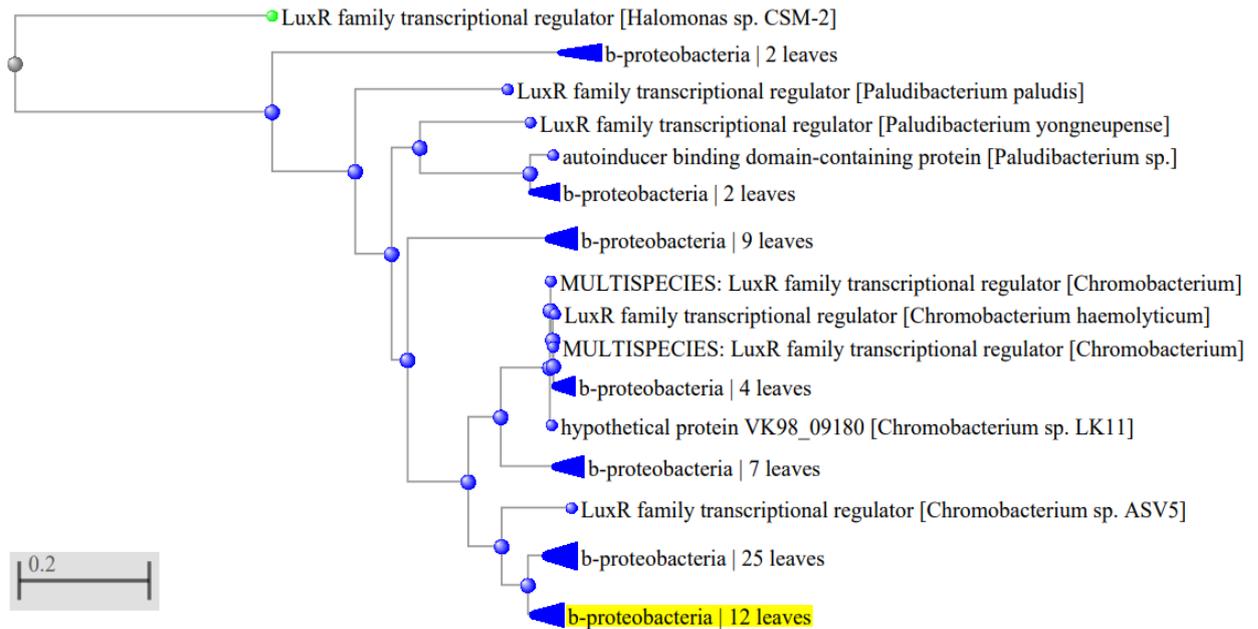


Figure 43 : Arbre phylogénétique de la protéine CviR

- L'arbre phylogénétique du gène *VioA* chez *C. violaceum* 12472 a été construit à l'aide de la plate-forme NCBI et logiciel BLASTn en utilisant la méthode neighbor joining (NJ) (figure 44).

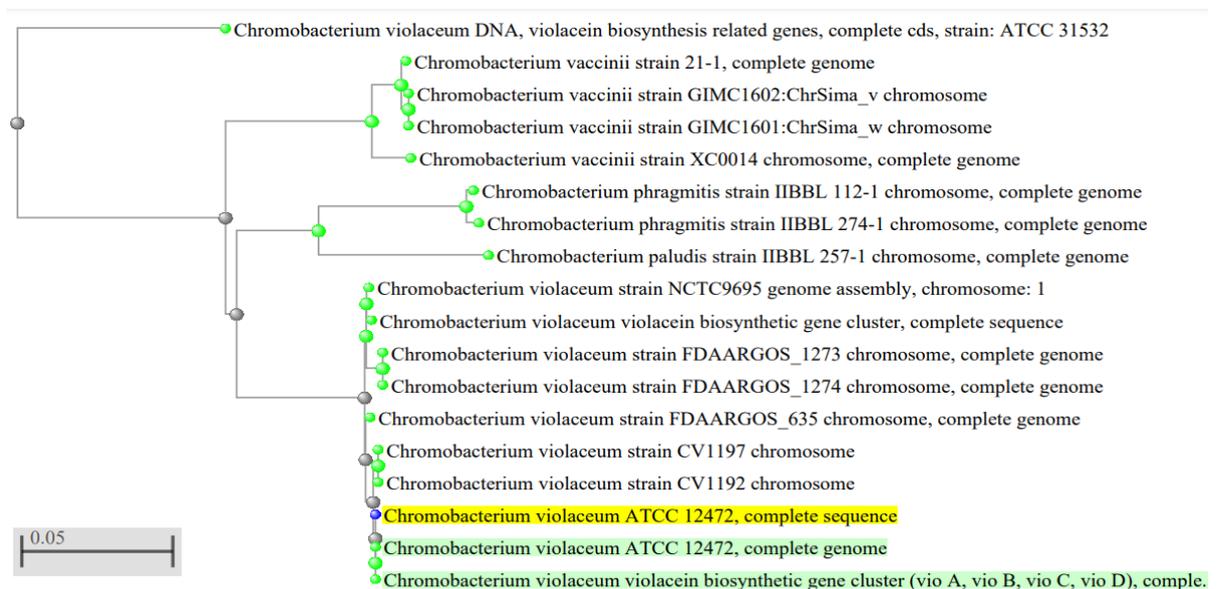


Figure 44 : Arbre phylogénétique du gène *VioA*

- L'arbre phylogénétique du gène *VioB* chez *C. violaceum* 12472 a été construit à l'aide de la plate-forme NCBI et logiciel BLASTn en utilisant la méthode neighbor joining (NJ) (figure 45).

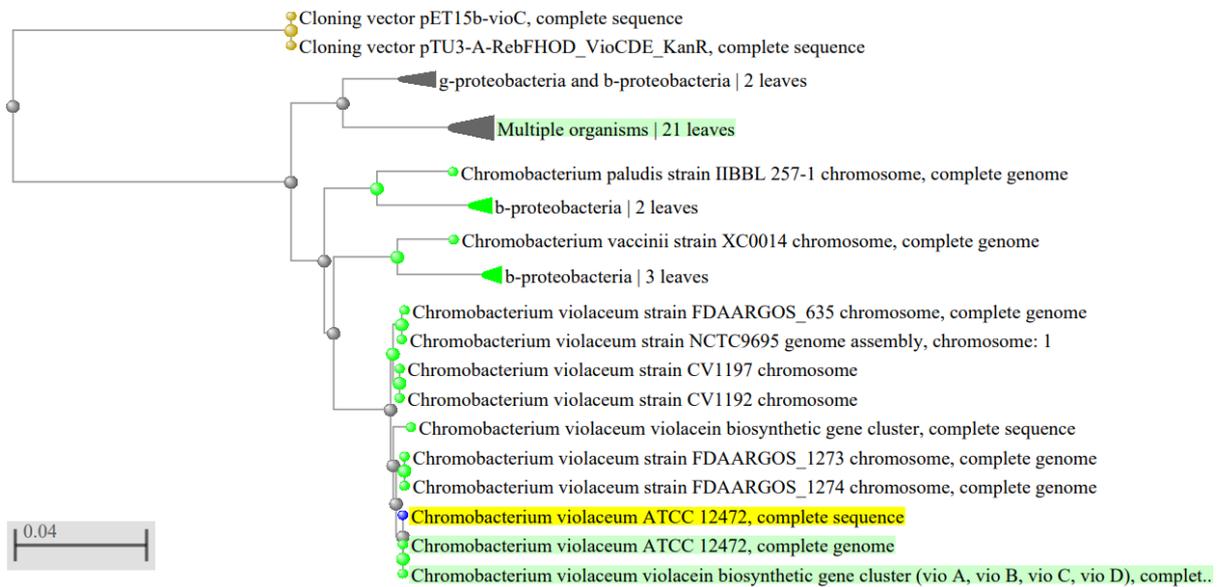


Figure 45 : Arbre phylogénétique du gène *VioB*

- L'arbre phylogénétique du gène *VioC* chez *C. violaceum* 12472 a été construit à l'aide de la plate-forme NCBI et logiciel BLASTn en utilisant la méthode neighbor joining (NJ) (figure 46)

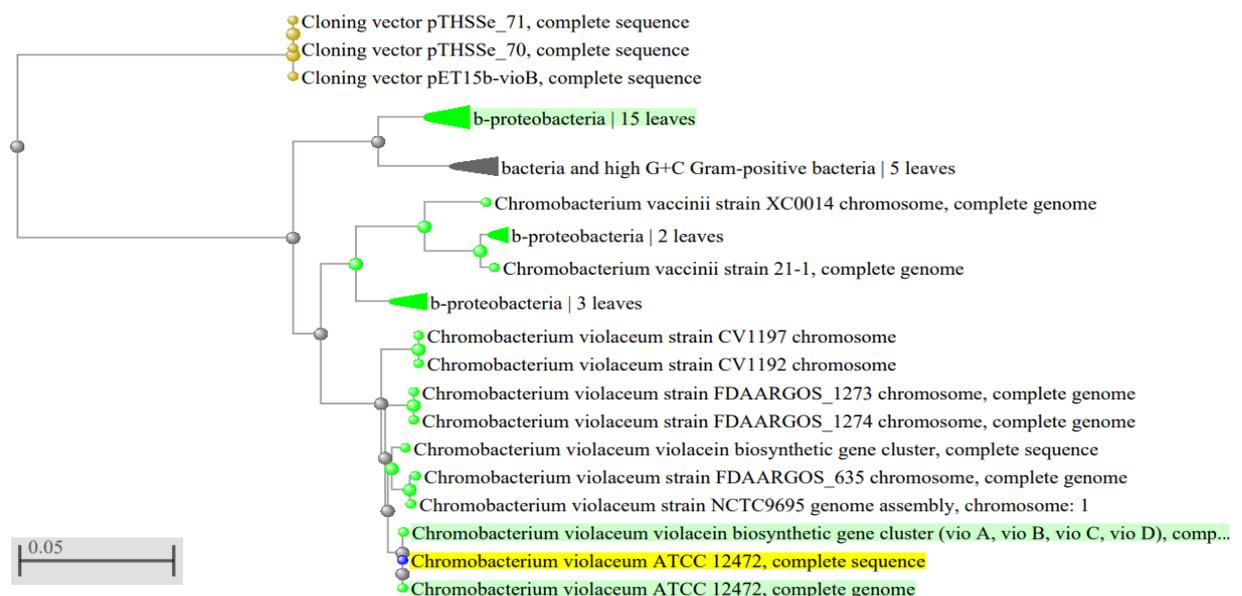


Figure 46 : Arbre phylogénétique du gène *VioC*

- L'arbre phylogénétique du gène *VioD* chez *C. violaceum* 12472 a été construit à l'aide de la plate-forme NCBI et logiciel BLASTn en utilisant la méthode neighbor joining (NJ) (figure 47).

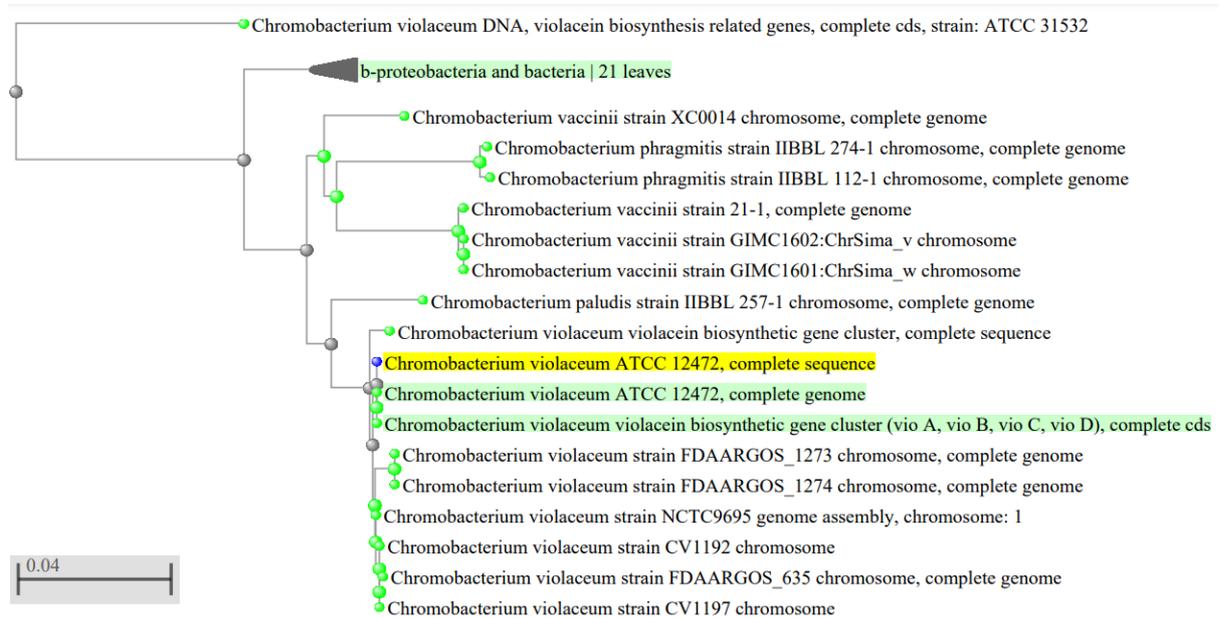


Figure 47 : Arbre phylogénétique du gène *VioD*

- L'arbre phylogénétique du gène *VioE* chez *C. violaceum* 12472 a été construit à l'aide de la plate-forme NCBI et logiciel BLASTn en utilisant la méthode neighbor joining (NJ) (figure 48).

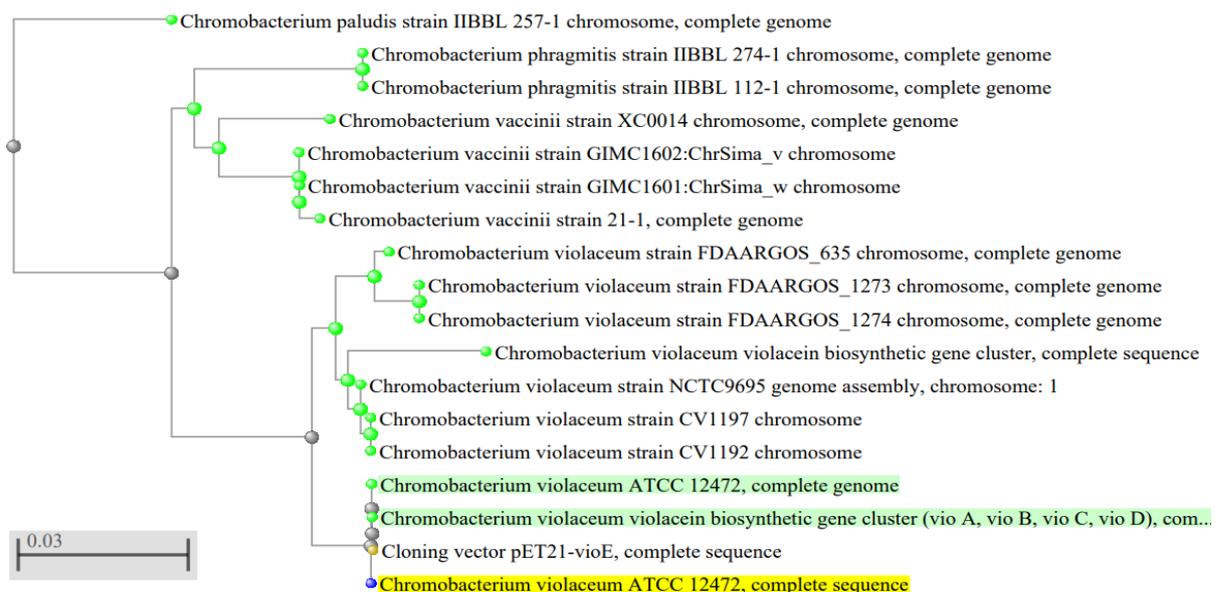


Figure 48 : Arbre phylogénétique du gène *VioE*

2. Discussion

À partir de nos résultats obtenus, nous constatons que la protéine régulatrice CviR du quorum sensing chez *C. violaceum* présente 8 ancêtres communs (figure 43). Le premier (l'ancêtre le plus ancien.) présente une homologie, c'est à dire une similarité provenant d'un ancêtre commun qui est généré d'un évènement de spéciation due probablement à une pression évolutive. Donc il est suggéré comme une évolution divergente, ainsi cette protéine peut être homologue avec la majorité des protéines des familles LuxR Régulateur de la transcription, comme l'espèce *Halomonas sp* et le genre *Paludibacterium paludis*, mais plus homologue avec les espèces *Chromobacterium sp*, car ils partagent un ancêtre commun le plus récent, c'est pour cela la protéine CviR possède presque les mêmes propriétés et la même structure de cette famille des LuxR.

La protéine régulatrice de la transcription CviR va former un complexe avec l'AHL pour activer la transcription des gènes d'opéron *VioABCDE* responsable de la synthèse de la violacéine. A l'exception de la souche *ATCC12472* que nous avons analysé, elle possède un système AHL QS et, étonnamment, elle produit et répond à différents AHLs. Le système AHL QS de *C. violaceum ATCC12472* est codé par les gènes génétiquement liés CviI et CviR qui produisent et répondent avec la plus grande affinité à la N-décanoyl-L-homosérine lactone (C10-HSL) (Devescovi et al., 2017).

Nous avons remarqué selon les résultats de BLASTn (figure 37), un taux de similitude élevé entre les séquences produisant des alignements CviR chez la famille *LuxR* spéciale à l'espèce *Chromobacterium sp*, et une similitude moins élevée chez toutes les séquences des familles LuxR appartenant à d'autres genres bactériens. Cela est dû à une forte homologie entre la protéine régulatrice CviR du *Chromobacterium violaceum* et la famille LuxR des *Chromobacterium sp*, et une moindre homologie avec les familles LuxR des autres bactéries.

Grâce à la plate-forme PDB, on a pu avoir la structure de la protéine CviR (PDB ID : 3QP8) liée à son auto-inducteur C10-HSL (N-[(3S)-2-oxotétrahydrofuran-3-yl]décanamide) (figure 36) à une résolution de 1,60 Å. Cette protéine possède quatre chaînes identiques appartenant à la famille des Autoind_bind (A, B, C et D), et à la classe des alpha-beta, chacune de ces chaînes est constituée de deux couches sandwich (two layer sandwich), et présente une topologie similaire aux bêta-lactamase.

Les protéines de type LuxR, comme les protéines CviR sont des homodimères, chaque monomère est constitué de deux domaines, un domaine de liaison au ligand (LBD) et un domaine de liaison à l'ADN (DBD). À la réception du signal correspondant via le LBD, elles subissent certains changements de conformation permettant ainsi l'expression des gènes.

Après le déclenchement de l'expression de l'opéron *Vio* (*Vio ABCDE*) grâce aux complexe CviR-HSL. Le premier produit enzymatique, VioA (Flavin-dépendant tryptophan 2-monooxygénase) catalyse l'oxydation du tryptophane en indole-3-pyruvic acid imine (IPA) avec la réduction du cofacteur de la FAD. L'IPA est ensuite convertie en dimère imine de courte durée par le processus de dimérisation, par l'activité de VioB. Le dimère de courte durée de vie est soit transformé spontanément en acide chromo pyrrolique qui est impliqué dans la biosynthèse indolocarbazole, ou VioE modifie le dimère rapidement en acide protodésoxyviolacéinique (PVA) qui mène à la biosynthèse de la violacéine. On peut donc dire que VioE joue un rôle clé dans la construction du squelette moléculaire de la violacéine. De plus, VioD produit de la proviolacéine en catalysant l'hydroxylation à la cinquième position d'un anneau indole de PVA avec l'oxydation du NADPH en NADP. La proviolacéine est ensuite convertie en violacéine avec l'oxydation du NADPH en NADP⁺ accompagnée de la formation d'eau et de CO₂ sous le contrôle de *VioC* (Kothari et al., 2017). L'activation des gènes *vio* (opéron *Vio ABCDE*) de la biosynthèse de la violacéine, est selon l'enchaînement suivant ; *VioA* suivie par *VioB*, *VioE*, *VioD* et *VioC*.

Les gènes *VioA* appartiennent à la famille 2 des glutathion réductases des oxydoréductases dépendantes du FAD selon le domaine de liaison du cofacteur structurellement conservé. Le domaine de liaison au substrat de *VioA* est principalement responsable de la reconnaissance spécifique du L-tryptophane.

Selon, les résultats du BLASTn obtenus pour le gène *VioA* (figure 38) chez la souche ATCC12472 de *Chromobacterium violaceum*, nous avons eu 17 séquences qui produisant des alignements *VioA* significatifs. Toutes les souches de *C. violaceum* peuvent produire des gènes *VioA* très identique (99 % jusqu'à 100 %), avec les gènes de la souche ATCC12472 de *C. violacerum* étudiée. On remarque que d'autres genres de *Chromobacterium* ont ces gènes *VioA*, mais moins similaires et identiques à celles de l'espèce de *C. violaceum*, en général et à la souche ATCC12472 précisément. Le pourcentage de similitude le plus faible a été noté avec la souche ATCC31532 de l'espèce *C. violaceum*.

Les gènes *VioB*, nommée aminophenyl-pyruvate dimère synthase synthétise des enzymes. Après l'action d'une L-aminoacide oxydase flavino-dépendante qui convertit le L-tryptophane en

indole-3-pyruvate imine, les enzymes VioB, peuvent ligaturer deux molécules, formant un dimère couplé d'iminophényl-pyruvate. Dans la voie de biosynthèse de la violacéine, ce composé est actionné par VioE avant de se cycliser spontanément en acide chromopyrrolique.

Nos résultats BLASTn obtenus sur le gène *VioB* (figure 39) de la souche *ATCC12472* de *C. violaceum*, e que la majorité des souches de cette espèce ont la capacité de produire ce gène, qui est presque complètement identiques à celle des *C. violaceum ATCC12472*. Nous avons constaté qu'il y a d'autres genres de *Chromobacterium* qui possèdent ces gènes (*VioB*) mais avec un pourcentage d'identité moins faible avec le gène de référence étudié. À l'exception de la souche 203 de *Sphingopyxis macrogoltabida* qui possède des gènes *VioB* 100 % identiques. L'origine de cette similarité entre les gènes *VioB* de ces deux espèces bactériennes n'est pas assez expliquée.

Le VioC est un monooxygénase flavino-dépendant et NADPH-dépendant qui catalyse la dernière étape enzymatique de la voie de biosynthèse de la violacéine. Il présente une homologie élevée avec la p-hydroxybenzoate hydroxylase et les enzymes apparentées. La VioC est principalement responsable d'un hydroxylation en position 2 du cycle indole opposé du protoviolacéinate pour produire du violacéinate et également un hydroxylation de la position 16 du protoviolacéinate et du protodésoxyviolacéinate pour former respectivement la violacéine et la désoxyviolacéine.

Par rapport au bilan BLASTn issu du gène *VioC* (figure 40) chez la souche *ATCC12472* de *C. violaceum*. Nous avons remarqué que la plupart des souches de *C. violaceum* ont la capacité de produire les gènes *VioC*, qui sont presque complètement identiques à *C. violaceum ATCC12472*. Il y a d'autres espèces bactériennes qui possèdent les gènes *VioC* mais leurs pourcentages d'identité varient entre 82% et 73%. Il y a aussi une similitude de 100% entre les souches *ARSEF2860* et *HN6* de *Beauveria bassiana* et la souche *ATCC12742* de *C. violaceum* avec une origine de similitude indéterminée.

Le VioD ou tryptophane hydroxylase VioD est un gène qui appartient à la famille des oxydoréductases flavino-dépendant et NADPH-dépendant, ce gène catalyse la dernière étape enzymatique de la voie de biosynthèse de la violacéine. Le VioD catalyse l'oxygénation de la position 6 du protodésoxyviolacéinate pour former du protoviolacéinate dans la voie de biosynthèse de la violacéine.

A partir du bilan BLASTn résulte du gène *VioD* (figure 41) chez la souche *ATCC12742* de *C. violaceum*. Nous avons observé que la plupart des souches de *C. violaceum* ont la capacité de

produire les gènes *VioC*, qui sont presque complètement identique à celle des *C. violaceum* ATCC12742, en outre il y a d'autres espèce bactériennes qui possèdent ces gènes (*VioD*) mais le pourcentage d'identité est moins faible avec celle du gène de référence étudié sauf qu'il y a deux souches exceptionnelles *contig00106* de *Uncultred bacterium* et *CIAT 899* de *Rhizobium tropici* qui possèdent un pourcentage de similarité 100% identique avec le gène de référence étudié celle de la souche ATCC12742 de *C. violaceum*, mais nous ne connaissons pas encore son origine.

Le *VioE* un gène qui code pour une enzyme inhabituelle sans homologues caractérisés qui joue un rôle d'un acteur clé dans la construction du squelette moléculaire de la violacéine. Il partage un pli avec les protéines transporteuses des lipoprotéines, et le substrat de la *VioE* est un intermédiaire dimérique à courte durée de vie.

Sans cofacteurs ou métaux liés, *VioE*, de la bactérie *Chromobacterium violaceum*, est le médiateur d'un déplacement 1,2 d'un cycle indole et de la chimie oxydative pour générer la prodéoxyviolacéine, un précurseur de la violacéine.

Grâce au bilan BLASTn obtenu du gène *VioE* (figure 42), chez la souche ATCC12742 de *C. violaceum*, Nous avons pu distinguer que la majorité des souches de *Chromobacterium violaceum* peuvent produire des gènes *VioE*, qui sont presque identique à celle des *Chromobacterium violaceum* ATCC12742. Nous avons fait l'attention qu'il y a trois autres espèces de *Chromobacterium* qui possèdent les gènes *VioE*, l'espèce *Chromobacterium vaccinii*, l'espèce *Chromobacterium phragmitis* et l'espèce *Chromobacterium paludismatis*, le pourcentage d'identité de ces souches est remarquable parce qu'il dépasse 90%, donc ils sont presque complètement identiques avec la souche étudiée.

D'après nos résultats, l'arbre phylogénétique global du gène *VioA* de *C. violaceum* ATCC12742 (figure 44) représente 7 ancêtres communs. Le premier (un ancêtre le plus ancien.) présente une homologie donc c'est une évolution divergente, qui a généré un événement de spéciation due probablement à une pression évolutive. Ce gène *VioA* est homologue avec tous les genres de *Chromobacterium* mais plus homologue avec *C. violaceum* à cause d'un ancêtre commun récent. A l'exception de la souche ATCC31532, considérée comme moins homologue à cause de l'ancêtre commun le plus ancien, ce qui explique pourquoi la souche ATCC31532 de *C. violaceum* au pourcentage d'identité le plus faible selon le bilan BLASTn (figure 38).

A Partir de de nos résultats, l'arbre phylogénétiques global des quatre gènes restant *VioB*, *VioC*, *VioD* et *VioE*, les deux gènes *VioB* et *VioC* représente 7 ancêtres communs (figure 45 et

46), tandis que le gène *VioD* représente 8 ancêtres communs (Figure 47) et le gène *VioE* possède 5 ancêtre communs (figure 48). Le premier (l'ancêtre le plus ancien.) de chaque de ces gènes présente une homologie donc c'est une évolution divergente, qui a généré un événement de spéciation due probablement à une pression évolutive. Ces gènes sont homologues avec tous les genres de *Chromobacterium*, mais plus homologues avec le genre *C. violaceum* à cause d'un ancêtre commun récente, ce qui explique les pourcentages d'identités très élevées des *VioB*, *VioC*, *VioD* et *VioE* chez les *C. violaceum* obtenu à partir des résultats de BLASTn (figure 39,40,41,42).

Nous avons fait des tentatives de reformulation de ces arbres phylogénétiques selon la méthode de Fast Minimum, mais nous avons trouvées presque les mêmes profils dont les distances entre les espèces présentés par les deux méthodes sont presque identiques, donc nous avons choisi la méthode la plus exploitée, la méthode de Neighbor Joining (NJ).

Conclusion

Conclusion

Les travaux et les études scientifiques, à travers notre synthèse bibliographique, montrent clairement que *Chromobacterium violaceum* est une bactérie importante comme une source de connaissances sur la génétique et la pathogénèse, et aussi comme un producteur de divers métabolites secondaires. Ces dernières années, l'intérêt est devenu et orienté, notamment, sur les propriétés biologiques et les applications de son pigment, la violacéine, qui revêt une importance économique pour des utilisations industrielles et des applications pharmacologiques. La production de ce pigment peut être considérée comme une partie de la stratégie compétitive pour prolonger la durée de vie de la colonie microbienne (Durán *et al.*, 2016; LAATTAR, 2019).

La biosynthèse de violacéine chez *C. violaceum* est sous le contrôle d'un mécanisme de régulation dépendant de la densité bactérienne, appelé quorum sensing. Chez cette bactérie il comprend quatre composants principaux : CviI synthase, la molécule diffusible des AHLs (Homéoméries Lactones), à un récepteur de signal de type CviR et des gènes liés à un régulateur transcriptionnel, les gènes d'opéron *Vio* (*VioA*, *VioB*, *VioC*, *VioD* et *VioE*). Un système de détection de quorum médié par HHL est également présent chez *C. violaceum*, et son organisation semble être très similaire à celle de *Pseudomonas aeruginosa* (Federal *et al.*, 2004; Durán *et al.*, 2016). Le QS régulateur de la production du pigment violacéine chez *Chromobacterium violaceum*, est considéré comme un quorum quenching pour la formation du biofilm, de plusieurs bactéries pathogènes tell que : *Staphylococcus epidermidis* et *Staphylococcus aureus*.

Les résultats, de notre analyse *in silico*, montrent que la protéine régulatrice CviR chez *C. violaceum* ATCC12472 présente une structure bien définie qui assure la liaison avec l'auto-inducteur pour assurer la régulation et le contrôle de l'expression des gènes de l'opéron *Vio*. Cette protéine régulatrice CviR présente une ressemblance avec presque toutes les protéines régulatrices appartenant à la famille des LuxR, responsable de la régulation du quorum sensing chez plusieurs bactéries.

Nos résultats indiquent que les cinq gènes de l'opéron *Vio*, sont indispensables pour la biosynthèse de la violacéine et leur activation se fait en cascades du *VioA*, *VioB*, *VioE*, *VioD* jusqu'à *VioC*. Ces gènes existent chez toutes les souches de *C. violaceum* et peuvent exister chez d'autre espèces du genre *Chromobacterium*, néanmoins les gènes codant l'opéron *Vio* chez *C. violaceum* ne sont pas complètement identique avec les gènes *Vio* d'autre espèces *Chromobacterium*, ce qui explique pourquoi seulement les souches de *C. violaceum* ont la capacité de produire le pigment violacéine.

Conclusion

Les données de séquence, consultées en ligne *in silico*, représentent une excellente ressource pour identifier l'expression différentielle des gènes en question. En effet, ces ressources sont des points de départ populaires dans de nombreux programmes de recherche sur la découverte de gènes de maladies. Ces ensembles de données sont collectés et annotés dans des bases de données en ligne hautement organisées (NCBI). Les scientifiques ont donc la possibilité de dresser le profil génomique, pour certaines maladies, de distinctions d'intérêt et d'identifier ainsi les gènes exprimés de manière différentielle.

L'étude *in silico* avec ses outils d'analyse et de prédiction peut constituer une approche pour offrir de meilleures stratégies de diagnostic pour des alternatives thérapeutiques avec de meilleures performances. Également de telles connaissances scientifiques peuvent être sollicitées pour mieux comprendre et étudier les potentialités biotechnologiques de plusieurs microorganismes. Ainsi, la connaissance des mécanismes moléculaires du QS offre l'opportunité de développer de nouvelles cibles thérapeutiques.

Enfin, l'utilisation de stratégies d'exploration génétique *in silico* fournit un excellent cadre pour l'identification initiale de gènes clés et de groupes de gènes dont l'expression est altérée dans les tissus malades. Les données générées par ces études constituent un point de départ pour les recherches visant à délimiter la base moléculaire de certaines pathologies humaines.

*Les références
bibliographiques*

Les références bibliographiques

- Ahmed A, Ahmad A, Li R, Al-Ansi W, Fatima M, Mushtaq BS, Basharat S, Li Y, Bai Z. Recent Advances in Synthetic, Industrial and Biological Applications of Violacein and Its Heterologous Production. *J Microbiol Biotechnol.* 2021 Nov 28;31(11):1465-1480. doi: 10.4014/jmb.2107.07045. PMID: 34584039.
- Alim R. , Safiullah, S. , Munwar, S. , Mazhar, I. , Zaman, S. and Bari, M. (2022) Wound Infection Caused by *Chromobacterium violaceum*: A Case Report from a Tertiary Care Hospital in Bangladesh. *Advances in Microbiology*, 12, 83-89. doi: [10.4236/aim.2022.122007](https://doi.org/10.4236/aim.2022.122007).
- AMARA M. et KORBA R. Université Mohamed El Bachir El Ibrahimi - Bordj Bou Arréridj - Faculté SNV-STU Département des Sciences Biologiques'. (2020) 'Ver . MARS / 2020 IFB . <http://www.france-bioinformatique.fr/fr/evenements/EBA2017> .
- Antonisamy P, Kannan P, Aravinthan A, Durairandiyan V, Arasu MV, Ignacimuthu S, Al-Dhabi NA, Kim JH. Gastroprotective activity of violacein isolated from *Chromobacterium violaceum* on indomethacin-induced gastric lesions in rats: investigation of potential mechanisms of action. *ScientificWorldJournal.* 2014;2014:616432. doi: 10.1155/2014/616432. Epub 2014 Aug 5. PMID: 25162059; PMCID: PMC4138890.
- Bahia B. (2011). Etude Bioinformatique Et Dosage Chimique Des Tanins Produit Par Les Plantes Medicinales Exposees Au Stress De La Pollution. Universite Abdelhamid Ibn Badis Mostaganem , Faculte Des Sciences Et De La Technologie , Department De Genie Des Procedes, 104.
- Bali EB, Türkmen KE, Erdönmez D, Sağlam N. Comparative Study of Inhibitory Potential of Dietary Phytochemicals Against Quorum Sensing Activity of and Biofilm Formation by *Chromobacterium violaceum* 12472, and Swimming and Swarming Behaviour of *Pseudomonas aeruginosa* PAO1. *Food Technol Biotechnol.* 2019 Jun;57(2):212-221. doi: 10.17113/ftb.57.02.19.5823. PMID: 31537970; PMCID: PMC6718966.
- Balibar C.J. and Walsh, C.T. (2006) 'In Vitro Biosynthesis of Violacein from L -Tryptophan by the Enzymes VioA - E from *Chromobacterium violaceum* †', pp. 15444–15457.
- Ballantine J A, Beer R J S, Crutchley D J, Dodd G M, Palmer D R: The chemistry of bacteria. Part VIII. The synthesis of violacein and related compound, *J.Chem. Soc* 1960; 2292–

Les références bibliographiques

- Batista J.H. and Neto, J.F. d. S. (2017) ‘Chromobacterium violaceum pathogenicity: Updates and insights from genome sequencing of novel Chromobacterium species’, *Frontiers in Microbiology*, 8(NOV), pp. 1–7. doi:10.3389/fmicb.2017.02213.
- Bergonzini C: Um nuevo bacterio colorato, *Ann. Soc. Natural. Modena. Ser.2* 1881; 14: 149– 58.
- Bhanbri P. and Gupta, O.P. (2017) ‘Phylogenetic Tree Construction for Distance based Methods’, *International Journal of Scientific Research in Research Paper. Computer Science and Engineering*, 5(3), pp. 142–149.
- Boisbaudran L : Matière colorante se formant dans la colle de farine, *Comp. Rend. Acad. Sci* 1882 ; 94 : 562-3.
- Caldas L R, Leitão A A C, Santos S M, Tyrrell R M: Preliminary experiments on the photobiological properties of violacein, *Intern. Symp. Curr. Topics Radiol. Photobiol. Acad. Brasil. Cienc. Rio de Janeiro* 1978; 121–32.
- Carraro D. M., Camargo, A. A., Salim, A. C., Grivet, M., Vasconcelos, A. T., & Simpson, A. J. (2003). The complete genome sequence of *Chromobacterium violaceum* reveals remarkable and exploitable bacterial adaptability\rPCR-assisted contig extension: stepwise strategy for bacterial genome closure. *100(20)*, 11660-5.
- Castillo-Juárez, Israel & García-Contreras, Rodolfo & Guadarrama, Norma & soto-hernandez, Marcos & Martinez-Vazquez, Mariano. (2013). Amphypterygium adstringens Anacardic Acid Mixture Inhibits Quorum Sensing-controlled Virulence Factors of Chromobacterium violaceum and Pseudomonas aeruginosa. *Archives of medical research*. 44. 10.1016/j.arcmed.2013.10.004.
- Cristina de Siqueira, I., Dias, J., Ruf, H., Ramos, E. A. G., Maciel, E. A. P., Rolim, A., Jabur, L., Vasconcelos, L., & Silvany, C. (2005). *Chromobacterium violaceum* in Siblings, Brazil. *Emerging Infectious Diseases*, 12(9), 1443–1445. <https://doi.org/10.3201/eid1209.050278>.
- Dessaux Y., Elmerich, C. and Faure, D. (2004) ‘Violacein: A molecule of biological interest originating from the soil-borne bacterium Chromobacterium violaceum’, *Revue de Medecine Interne*, 25(9), pp. 659–662. doi:10.1016/j.revmed.2004.03.012.
- Diniz WJ, Canduri F. REVIEW-ARTICLE Bioinformatics: an overview and its applications. *Genet Mol Res*. 2017 Mar 15;16(1). doi: 10.4238/gmr16019645. PMID: 28301675.

Les références bibliographiques

- Durán M., Ponezi, A.N., Faljoni-Alario, A., Teixeira, M.F.S., Justo, G.Z. (2012) ‘Potential applications of violacein: A microbial pigment’, *Medicinal Chemistry Research*, 21(7), pp. 1524–1532. doi:10.1007/s00044-011-9654-9.
- Durán N, Faljoni-Alario A: Bacterial chemistry-I: studies of a potential phototherapeutic substance from *Chromobacterium violaceum*, *An. Acad. Brasil. Ciênc* 1980; 52: 297–302.
- Durán N, Justo GZ, Durán M, Brocchi M, Cordi L, Tasic L, Castro GR, Nakazato G. Advances in *Chromobacterium violaceum* and properties of violacein-Its main secondary metabolite: A review. *Biotechnol Adv.* 2016 Sep-Oct;34(5):1030-1045. doi: 10.1016/j.biotechadv.2016.06.003. Epub 2016 Jun 8. PMID: 27288924.
- Durán N. and Menck, C.F.M. (2001) ‘*Chromobacterium violaceum*: A review of pharmacological and industrial perspectives’, *Critical Reviews in Microbiology*, 27(3), pp. 201–222. doi:10.1080/20014091096747.
- Eberhard A. Inhibition and activation of bacterial luciferase synthesis. *J. Bacteriol.* 1972; 109 :1101–1105.
- Federal U. and Catarina, D.S. (2004) ‘Genetic analysis of violacein biosynthesis by’, *Genetics and Molecular Research*, 3(1), pp. 85–91.
- Finn RD, Tate J, Mistry J, Coghill PC, Sammut SJ, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer EL, Bateman A. The Pfam protein families database. *Nucleic Acids Res.* 2008 Jan;36(Database issue):D281-8. doi: 10.1093/nar/gkm960. Epub 2007 Nov 26. PMID: 18039703; PMCID: PMC2238907.
- Fuqua WC, Winans SC, Greenberg EP. Quorum Sensing in bacteria: The LuxR-LuxI family of cell density-responsive transcriptional regulators. *J. Bacteriol.* 1994; 176 :269–275.
- Gauthier J, Vincent AT, Charette SJ, Derome N. A brief history of bioinformatics. *Brief Bioinform.* 2019 Nov 27;20(6):1981-1996. doi: 10.1093/bib/bby063. PMID: 30084940.
- Greenberg EP, Hastings JW, Ulitzer S. Induction of luciferase synthesis in *Beneckea harveyi* by other marine bacteria. *Arch. Microbiol.* 1979; 120 :87–91.
- Hatzimaikatis V. (2000) Bioinformatics and functional genomics: Challenges and opportunities, *AICHE Journal*. doi:10.1002/aic.690461202.

Les références bibliographiques

- Hirano S, Asamizu S, Onaka H, Shiro Y, Nagano S. Crystal structure of VioE, a key player in the construction of the molecular skeleton of violacein. *J Biol Chem.* 2008 Mar 7;283(10):6459-66. doi: 10.1074/jbc.M708109200. Epub 2008 Jan 1. PMID: 18171677.
- Holt J G, Krieg, N R (Eds): *Bergey's Manual of Systematic Bacteriology*, Vol 1, Williams & Wilkins, Baltimore 1984; 580–2.
- Hoshino T. (2011) 'Violacein and related tryptophan metabolites produced by *Chromobacterium violaceum*: Biosynthetic mechanism and pathway for construction of violacein core', *Applied Microbiology and Biotechnology*, 91(6), pp. 1463–1475. doi:10.1007/s00253-011-3468-z.
- James Versalovic; Karen C Carroll; Guido Funke; James H Jorgensen; Marie L Landry. (2011) *Manual of Clinical Microbiology 10th Edition*.
- Kanehisa, M. and Bork, P. (2003) 'Bioinformatics in the post-sequence era', *Nature Genetics*, 33(3S), pp. 305–310. doi:10.1038/ng1109.
- Katherine S. Ryan, Carl J. Balibar, Kaitlyn E. Turo, Christopher T. Walsh, A. and Catherine L. Drennan (2017) 'HHS Public Access', *Physiology & behavior*, 176(3), pp. 139–148. doi:10.1074/jbc.M708573200.The.
- Kempner ES, Hanson FE. Aspects of light production by *Photobacterium fischeri*. *J. Bacteriol.* 1968; 95 :975–979.
- Kim M. H., Lee, H. J., Suh, J. T., Chang, B. S., & Cho, K. S. (2005). A case of *Chromobacterium* infection after car accident in Korea. *Yonsei Medical Journal*, 46(5), 700–702. <https://doi.org/10.3349/ymj.2005.46.5.700>.
- Konzen, Marlon & Marco, Daniela & Cordova, Clarissa & Vieira, Tiago & Antonio, Regina & Creczynski-Pasa, tânia. (2007). Antioxidant properties of violacein: Possible relation on its biological function. *Bioorganic & medicinal chemistry*. 14. 8307-13. 10.1016/j.bmc.2006.09.013.
- Kothari V., Sharma, S. and Padia, D. (2017) 'Recent research advances on *Chromobacterium violaceum*', *Asian Pacific Journal of Tropical Medicine*, 10(8), pp. 744–752. doi:10.1016/j.apjtm.2017.07.022.
- Mann Elizabeth & Kroutil Emily. (2015). *Determining Bacterial Growth Rates Using Flow Cytometry: An Educational Exercise*.

Les références bibliographiques

- McClean KH, Winson MK, Fish L, Taylor A, Chhabra SR, Camara M, Daykin M, Lamb JH, Swift S, Bycroft BW, Stewart GSAB, Williams P. Quorum sensing and *Chromobacterium violaceum*: exploitation of violacein production and inhibition for the detection of N-acylhomoserine lactones. *Microbiology (Reading)*. 1997 Dec;143 (Pt 12):3703-3711. doi: 10.1099/00221287-143-12-3703. PMID: 9421896.
- Momen A Z M R, Hoshino T: Biosynthesis of violacein: Intact incorporation of the tryptophan molecule on the oxindole side, with intramolecular rearrangement of the indole ring on the 5-hydroxyindole side, *Biosci. Biotechnol. Biochem* 2000; 64: 539–49.
- Morohoshi T, Fukamachi K, Kato M, Kato N, Ikeda T. Regulation of the violacein biosynthetic gene cluster by acylhomoserine lactone-mediated quorum sensing in *Chromobacterium violaceum* ATCC 12472. *Biosci Biotechnol Biochem*. 2010;74(10):2116-9. doi: 10.1271/bbb.100385. Epub 2010 Oct 7. PMID: 20944413.
- N.W. Dyer, DVM, MS. "*Chromobacterium violaceum* septicemia in a Chinese Red Panda." [online] November 1998.<http://www.ndsu.nodak.edu/ndsu/veterinary_science/vetdiag/case.ht>.
- Nadia Sarwar, Sarah Sarwar, Sharoon Ejaz, Abdulqader Al-Adeeb, Waleed AL-Ansi, Ye Li, Zhonghu Bai, P.(2021).Metabolic Engineering of Microorganisms to Increase Production of Violacein. *International Journal of Environment Agriculture and Biotechnology*(ISSN: 2456-1878).6(1), 295-306.10.22161/ijeab.61.37.
- Neilson KH, Platt T, Hastings JW. Cellular control of the synthesis and activity of the bacterial luminescent system. *J. Bacteriol*. 1970; 104 :313–322.
- Parajuli N. P., Bhetwal, A., Ghimire, S., Maharjan, A., Shakya, S., Satyal, D., Pandit, R., & Khanal, P. R. (2016). Bacteremia caused by a rare pathogen – *Chromobacterium violaceum*: A case report from Nepal. *International Journal of General Medicine*, 9, 441–446. <https://doi.org/10.2147/IJGM.S125183>.
- Paul P. Gardner, Jennifer Daub, John G. Tate, Eric P. Nawrocki, Diana L. Kolbe, Stinus Lindgreen, Adam C. Wilkinson, Robert D. Finn, Sam Griffiths-Jones, Sean R. Eddy, Alex Bateman, Rfam: updates to the RNA families database, *Nucleic Acids Research*, Volume 37, Issue suppl_1, 1 January 2009, Pages D136–D140, <https://doi.org/10.1093/nar/gkn766>.

Les références bibliographiques

- Ramageri B.M. et Bharati M. Ramageri / Indian Journal of Computer Science and Engineering Vol. 1 No. 4 301-305 DATA MINING TECHNIQUES AND APPLICATIONS', 1(4), pp. 301–305.
- Ravichandran V, Zhong L, Wang H, Yu G, Zhang Y, Li A. Virtual Screening and Biomolecular Interactions of CviR-Based Quorum Sensing Inhibitors Against *Chromobacterium violaceum*. *Front Cell Infect Microbiol.* 2018 Sep 4; 8:292. doi: 10.3389/fcimb.2018.00292. PMID: 30234025; PMCID: PMC6132053.
- Reilly J, Pyne G: CXL. On the pigment produced by *Chromobacterium violaceum*, *Biochem. J* 1927; 21: 1059–64.
- Richard C. *Chromobacterium violaceum*, bactérie pathogène opportuniste des régions tropicales et subtropicales [*Chromobacterium violaceum*, opportunist pathogenic bacteria in tropical and subtropical regions]. *Bull Soc Pathol Exot.* 1993;86(3):169-73. French. Erratum in: *Bull Soc Pathol Exot* 1993;86(4):307. PMID: 8219788.
- Rutherford S. T., & Bassler, B. L. (2012). Bacterial Quorum Sensing: Its role in virulence and possibilities for its control. *Cold Spring Harbor Perspectives in Medicine*, 2(11), 1–26. <https://doi.org/10.1101/cshperspect.a01242>
- Shinoda K, Hasegawa T, Sato H, Shinozaki M, Kuramoto H, Takamiya Y, Sato T, Nikaidou N, Watanabe T, Hoshino T. Biosynthesis of violacein: a genuine intermediate, protoviolaceinic acid, produced by VioABDE, and insight into VioC function. *Chem Commun (Camb).* 2007 Oct 28;(40):4140-2. doi: 10.1039/b705358d. Epub 2007 Aug 8. PMID: 17925955.
- Siddharthan R. (2005) 'Bioinformatics : Tasks , techniques , tools Introduction : Molecular Biology 101', pp. 1–10.
- Simons S. (2004) *Advances in Bioinformatics Technologies*, Genetic Engineering News.
- Stauff Devin & Bassler, Bonnie. (2011). Quorum Sensing in *Chromobacterium violaceum*: DNA Recognition and Gene Regulation by the CviR Receptor. *Journal of bacteriology.* 193. 3871-8. 10.1128/JB.05125-11.

Les références bibliographiques

- Subbiah V. K., Mazlan, Z., Mastor, N. N., & Hoque, M. Z. (2021). Whole genome sequence data of *Chromobacterium violaceum* WCH4, a human pathogenic strain from Sabah, Malaysia. *Data in Brief*, 39, 0–6. <https://doi.org/10.1016/j.dib.2021.107533>.
- Sun H., Zhao, D., Xiong, B. *et al.* Engineering *Corynebacterium glutamicum* for violacein hyper production. *Microb Cell Fact* **15**, 148 (2016). <https://doi.org/10.1186/s12934-016-0545-0>.
- Thapar P. (2010) ‘Bioinformatics: tools and applications’, *Choice Reviews Online*, 47(10), pp. 47-5624-47–5624. doi:10.5860/choice.47-5624.
- Tucker RE; Winter WG Jr; Wilson HD. Osteomyelitis associated with *Chromobacterium violaceum* sepsis: A case report. *J Bone Joint Surg. [Am]* 1979 Sep;61(6A):949-51 (ISSN: 0021-9355)
- Verbeke F., De Craemer, S., Debunne, N., Janssens, Y., Wynendaele, E., Van de Wiele, C., & De Spiegeleer, B. (2017). Peptides as Quorum Sensing molecules: Measurement techniques and obtained levels in vitro and in vivo. *Frontiers in Neuroscience*, 11(APR), 1–18. <https://doi.org/10.3389/fnins.2017.00183>
- Verinaud L, Lopes SC, Prado IC, Zanucoli F, Alves da Costa T, Di Gangi R, Issayama LK, Carvalho AC, Bonfanti AP, Niederauer GF, Duran N, Costa FT, Oliveira AL, Höfling MA, Machado DR, Thomé R. Violacein Treatment Modulates Acute and Chronic Inflammation through the Suppression of Cytokine Production and Induction of Regulatory T Cells. *PLoS One*. 2015 May 4;10(5):e0125409. doi: 10.1371/journal.pone.0125409. PMID: 25938431; PMCID: PMC4418714.
- Waters C.M. and Bassler, B.L. (2005) ‘Quorum Sensing: Cell-to-cell communication in bacteria’, *Annual Review of Cell and Developmental Biology*, 21, pp. 319–346. doi:10.1146/annurev.cellbio.21.012704.131001.
- Zimmerman B: Review of Bergonzini on *Cromobacterium*, *Bot. Centralbl* 1882; 4: 1528-30 Moss R D: Violacein, *Antibiotics*, 2nd 1967; 77– 81.