



**HAL**  
open science

# Typage des souches de *Clostridium botulinum* du groupe III : développement d'outils et application à l'épidémiologie du botulisme animal

Thibault Le Gratiet

► **To cite this version:**

Thibault Le Gratiet. Typage des souches de *Clostridium botulinum* du groupe III : développement d'outils et application à l'épidémiologie du botulisme animal. Microbiologie et Parasitologie. Université de Rennes, 2022. Français. NNT : 2022REN1B023 . tel-03860659

**HAL Id: tel-03860659**

**<https://theses.hal.science/tel-03860659v1>**

Submitted on 18 Nov 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THESE DE DOCTORAT DE

L'UNIVERSITE DE RENNES 1

ECOLE DOCTORALE N° 600

*Ecole doctorale Ecologie, Géosciences, Agronomie et Alimentation*

*Spécialité : Microbiologie, Virologie et Parasitologie*

Par

**Thibault Le Gratiet**

**Typage des souches de *Clostridium botulinum* du groupe III :  
Développement d'outils et application à l'épidémiologie du  
botulisme animal.**

**Thèse présentée et soutenue à Ploufragan, le 24 Mai 2022**

**Unité de recherche :**

**Anses Laboratoire de Ploufragan-Plouzané-Niort**

**Unité Hygiène et Qualité des Produits Avicoles et Porcins (HQPAP).**

**Rapporteurs avant soutenance :**

Dr. Frédéric Carlin

Directeur de recherche, INRAE, Avignon, France

Pr. Henri-Jean Boulouis

Professeur en bactériologie, ENVA, Maisons-Alfort, France

**Composition du Jury :**

**Président :**

Dr. Frédéric Carlin

Directeur de recherche, INRAE, Avignon, France

**Examineurs :**

Dr. Frédéric Carlin

Directeur de recherche, INRAE, Avignon, France

Pr. Henri-Jean Boulouis

Professeur en bactériologie, ENVA, Maisons-Alfort, France

Dr. Christelle Mazuet

Directrice du CNR des Bactéries anaérobies et Botulisme, Institut Pasteur, Paris, France

Dr. Patrick Fach

Directeur de recherche, Anses, Maisons-Alfort, France

**Directeur de thèse :**

Dr. Marianne Chemaly

Directeur de projets de recherche, Anses, Ploufragan, France

**Co-directeur de thèse :**

Dr. Caroline Le Maréchal

Chargée de projet, Anses, Ploufragan, France

**Membres invités :**

Dr. Didier Hilaire

Adjoint au chef de la division Biologie, DGA - Centre d'études du Bouchet, Vert-le-Petit, France

Dr. Laurence Delbrassinne

Chef de travaux, Sciensano, Bruxelles, Belgique



## Remerciements

Je souhaiterais remercier dans un premier temps l'ensemble des membres de mon jury, Frédéric Carlin, Henri-Jean Boulouis, Christelle Mazuet, Patrick Fach, Laurence Delbrassinne et Didier Hilaire pour l'intérêt qu'ils ont porté à ce projet en acceptant de l'évaluer.

Je remercie également les membres du comité de pilotage de thèse, Sergine Even, Patrick Fach, Cédric Woudstra et Christelle Mazuet pour leurs conseils et les échanges scientifiques très enrichissants que nous avons pu avoir.

Ensuite je tiens à remercier ma directrice de thèse Marianne Chemaly, pour la confiance qu'elle m'a accordée, son soutien et ses précieux conseils.

Je tiens à remercier plus particulièrement Caroline Le Maréchal qui m'a guidé tout au long de ce projet. Son expérience et son professionnalisme ont permis de rendre cette expérience très enrichissante.

Merci à Cédric Woudstra qui m'a permis, au cours du premier confinement, de collaborer avec lui pour effectuer des travaux sur les systèmes CRISPR-Cas. Cette expérience fut très intéressante et fut la bienvenue dans un contexte très particulier.

Merci à l'ensemble de l'équipe du LNR botulisme aviaire, Caroline, Sandra, Typhaine, Marie, Emmanuelle, Amandine et Léa pour leur aide précieuse lors des différentes manipulations que j'ai pu effectuer au cours de ces 3 ans. Merci à Sandra, Typhaine et Caroline pour m'avoir formé à la manipulation de *Clostridium botulinum* du groupe III au début du projet.

Merci à Fabrizio Anniballi et son équipe pour leur accueil lors de mon séjour de recherche à Rome et pour m'avoir formé à la méthode MLVA et à l'assemblage des génomes.

Merci à l'équipe de la plateforme Identypath, Patrick, Sabine et Mai-Lan pour leur accueil à Maisons-Alfort et pour m'avoir formé à la préparation des bibliothèques pour le séquençage Illumina.

Merci à l'unité GVB, en particulier Yannick Blanchard et Arnaud Felten, pour leur aide et disponibilité. Merci à l'unité VIPAC, en particulier Francois-Xavier Briand, pour m'avoir permis d'utiliser le séquenceur pour le séquençage des amplicons MLVA.

Je remercie tous mes collègues de l'unité HQPAP avec qui j'ai passé de très bons moments, ça a été un plaisir de vous côtoyer pendant un peu plus de 3 ans. Enfin, merci à ma famille et mes amis pour leur soutien tout au long de la thèse.

## Résumé

**Mots clés :** *Clostridium botulinum*, botulisme animal, outils de typage, WGS, CRISPR-Cas, MLVA

Le botulisme animal, qui touche principalement les volailles, l'avifaune sauvage et les bovins, se manifeste par une paralysie flasque causée par l'action des neurotoxines botuliques. Les toxines les plus fréquemment associées aux épisodes de botulisme animal sont celles de type C, D, C/D et D/C, produites par *Clostridium botulinum* du groupe III, une bactérie anaérobie stricte sporulante, regroupée avec les espèces *Clostridium novyi* et *Clostridium haemolyticum*, sous le terme *Clostridium novyi sensu lato*. L'épidémiologie du botulisme animal est encore méconnue et les données d'épidémiologie moléculaire relatives aux souches impliquées dans le botulisme animal sont quasi inexistantes. Ceci s'explique principalement par l'absence d'outils de typage pour caractériser les souches du groupe *C. novyi sensu lato*, liée aux difficultés pour isoler les souches de ce groupe. L'objectif de la thèse est de développer des outils de typage adaptés à ce contexte particulier pour permettre de mener des investigations épidémiologiques lors d'épisodes de botulisme animal et pouvoir à terme tracer les souches.

Dans un premier temps, une méthode d'isolement a été développée et validée sur des souches impliquées dans des épisodes de botulisme animal en France. Cette méthode a permis d'isoler 35 souches associées à des épisodes de botulisme aviaire. Puis un outil de typage MLVA (Multi Loci Variable number of tandem repeats analysis) a été évalué sur la collection de souches isolées puis appliqué en conditions réelles dans le cadre d'investigations menées dans deux épisodes de botulisme animal. Une autre piste de typage a ensuite été explorée via l'étude des systèmes CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats – CRISPR associated protein). Les résultats ont révélé une prévalence de ces systèmes chez 95 % des souches de *C. novyi sensu lato*, avec une grande diversité (6 types de système CRISPR-Cas). Cette grande diversité n'a pas permis de proposer un outil de typage basé sur les CRISPR-Cas. Enfin, après une première étape d'optimisation des différentes étapes en amont du séquençage, les génomes de *C. novyi sensu lato* ont été comparés afin d'identifier de nouvelles cibles pertinentes.

Ces travaux ont permis de développer une approche permettant de séquencer les souches impliquées dans le botulisme aviaire (depuis l'isolement jusqu'au séquençage de l'isolat) et de mettre en place une méthode MLVA qui se révèle particulièrement adaptée au contexte du botulisme aviaire. Ils ont également permis d'identifier de nouvelles cibles très prometteuses qui permettront de proposer d'autres outils de typage dans le futur, applicables au botulisme animal de manière général.

## Abstract

**Keywords:** *Clostridium botulinum*, animal botulism, typing tools, WGS, CRISPR-Cas, MLVA.

Animal botulism, which affects mainly poultry, wildbirds and cattle, is a neuroparalytic disease caused by the action of botulinum neurotoxins type C, D, C/D and D/C. They are produced by *Clostridium botulinum* group III, an anaerobic, spore-forming bacterium belonging to the genospecies *Clostridium novyi sensu lato* alongside *Clostridium novyi* and *Clostridium haemolyticum*. The epidemiology of animal botulism is still poorly understood and molecular typing data are missing. This is mainly due to the unavailability of typing tools for *C. novyi sensu lato* characterization, mostly because strain isolation is very difficult. The objective of the PhD project is to develop typing tools adapted to this specific context to be able to perform epidemiological investigations during animal botulism outbreaks and strain tracking in the future.

First, an isolation method was developed and validated using strains involved in animal botulism outbreaks in France. This method allowed the isolation of 35 strains from avian botulism outbreaks. Then, a MLVA (Multi Loci Variable number of tandem repeats analysis) typing tool was evaluated on the isolated strains and used in real conditions during the investigation of two animal botulism outbreaks. Another potential typing approach was then explored through the study of CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats – CRISPR associated protein) systems. The results revealed the presence of these systems in 95 % of *C. novyi sensu lato* genomes, with a great diversity (6 different types of CRISPR-Cas system). This high diversity prevents the development of a typing tool based on these CRISPR-Cas systems. Finally, following the optimization of the different steps preceding sequencing, *C. novyi sensu lato* genomes were compared to identify new relevant targets.

This work allowed the development of an approach to sequence strains involved in avian botulism (from strain isolation to whole genome sequencing) and to set up an MLVA method that is particularly adapted to the study of avian botulism. It has also allowed the identification of new targets that are very promising for the development of new tools in the future, usable to investigate all animal botulism outbreaks.

## Table des matières

Remerciements .....	3
Liste des Abréviations .....	11
Liste des figures .....	13
Liste des équations .....	15
Liste des Tableaux.....	15
Introduction .....	16
I. Généralités sur <i>Clostridium botulinum</i> et le botulisme.....	16
1. Histoire du botulisme, de la découverte de la bactérie et de la toxine au bioterrorisme	16
1.1. Découverte de la bactérie et premières descriptions des symptômes.....	16
1.2. Découverte de plusieurs types de neurotoxine botulique et mode d'action.....	16
1.3. Utilisation détournée de la toxine botulique .....	17
1.3.1. Utilisation comme traitements thérapeutiques et comme produits cosmétiques .....	17
1.3.2. Utilisation comme arme biologique.....	18
2. <i>Clostridium botulinum</i> , la bactérie.....	20
2.1. Généralités sur la bactérie .....	20
2.2. Classification des bactéries productrices de BoNT.....	21
3. La toxine botulique .....	25
3.1. Structures et fonctions de la toxine et des protéines associées. ....	25
3.2. Mécanisme d'action de la toxine.....	27
3.3. Types et sous-types de BoNT.....	28
4. Le botulisme chez l'humain.....	33
II. Le Botulisme animal.....	34
1. Botulisme aviaire .....	34
1.1. Types toxiques associés et sensibilité .....	35
1.2. Données de surveillance du botulisme aviaire .....	37
1.3. Signes cliniques.....	38
1.4. Botulisme aviaire en élevage.....	38
1.4.1. Pathogénèse.....	38
1.4.2. Véhicules et sources de contamination .....	39
1.4.3. Mesures préventives et traitement .....	39
1.6. Botulisme aviaire en faune sauvage .....	40

1.6.1.	Pathogénèse.....	40
1.6.2.	Réservoirs de <i>C. botulinum</i> .....	40
1.6.3.	Mesures préventives et traitement .....	43
1.7.	Portage sain .....	44
2.	Botulisme chez les bovins.....	45
2.1.	Types toxiques associés et sensibilité .....	45
2.2.	Données de surveillance du botulisme chez le bovin.....	45
2.3.	Pathogénèse.....	45
2.4.	Sources de contamination.....	46
2.5.	Signes cliniques.....	46
2.6.	Portage sain .....	47
2.7.	Mesures préventives et traitement.....	47
3.	Réglementation concernant le botulisme animal .....	48
III.	<i>Clostridium novyi sensu lato</i> .....	50
1.	Physiologie des souches de <i>C. novyi sensu lato</i> .....	51
1.1.	Caractéristiques microbiologiques .....	51
1.2.	Caractéristiques macroscopiques .....	51
1.3.	Métabolisme, sporulation et toxinogénèse .....	52
1.4.	Isolement des souches .....	53
2.	Caractéristiques des génomes de <i>C. novyi sensu lato</i> .....	54
2.1.	Chromosome .....	54
2.2.	Plasmidome .....	55
2.3.	Transferts horizontaux de gènes.....	55
2.4.	Répartition des souches en 4 lignées.....	56
2.5.	Système CRISPR-Cas .....	57
3.	Présentation des outils et méthodes de typage existants .....	60
3.1.	Méthodes de caractérisation utilisant des enzymes de restrictions .....	62
3.1.1.	Pulse Field Gel Electrophoresis (PFGE).....	62
3.1.1.1.	Principe de la méthode .....	62
3.1.1.2.	Historique .....	62
3.1.1.3.	Utilisation de la PFGE pour le typage de <i>C. botulinum</i> .....	62
3.1.1.4.	Limites .....	64
3.1.2.	Amplified Fragment-Length Polymorphism (AFLP).....	64
3.1.2.1.	Principe de la méthode .....	64
3.1.2.2.	Utilisation de la méthode pour <i>C. novyi sensu lato</i> .....	64



3.1.2.3. Limites .....	65
3.2. Méthodes de caractérisation basées sur l’outil PCR .....	66
3.2.1. Random Amplified Polymorphic DNA (RAPD) .....	66
3.2.1.1. Principe de la méthode .....	66
3.2.1.2. Utilisation chez <i>C. botulinum</i> .....	66
3.2.1.3. Limites .....	66
3.2.2. Multi Loci VNTR Analysis (MLVA) .....	67
3.2.2.1. Généralités .....	67
3.2.2.2. Historique .....	67
3.2.2.3. Principe de la méthode .....	67
3.2.2.4. Utilisation de la méthode MLVA chez <i>C. botulinum</i> .....	68
3.2.3. Multi Loci Sequence Typing (MLST) .....	69
3.2.3.1. Généralités .....	69
3.2.3.2. Principe de la méthode .....	69
3.2.3.3. Utilisation chez <i>C. novyi sensu lato</i> .....	69
3.2.3.4. Limites .....	69
3.2.4. Comparaison des souches de <i>C. botulinum</i> sur la base du gène <i>fliC</i> .....	70
3.2.4.1. Généralités .....	70
3.2.4.2. Utilisation chez <i>C. novyi sensu lato</i> .....	70
3.2.5. Développement de PCR ciblant des EGM .....	72
3.2.5.1. Généralités .....	72
3.2.5.2. Utilisation chez <i>C. novyi sensu lato</i> .....	72
3.2.5.3. Conclusion .....	74
3.2.6. PCR ciblant les systèmes CRISPR-Cas .....	74
3.3. Autres méthodes de caractérisation .....	75
3.3.1. Analyse comparative du génome complet .....	75
3.3.1.1. Comparaison par Alignement fragmenté des génomes .....	75
3.3.1.2. Analyse des Single Nucleotide Polymorphism (SNP) .....	76
3.3.2. Matrix Assisted Laser Desorption Ionization - Time of Flight Mass Spectrometry (MALDI-TOF MS) .....	78
3.3.2.1. Généralités sur la spectrométrie de masse .....	78
3.3.2.2. Principe .....	78
3.3.2.3. Utilisation du MALDI-TOF pour l’identification et le typage de <i>C. novyi sensu lato</i> .....	79
3.4. Conclusion .....	81
Objectifs de la thèse .....	82

Résultats obtenus au cours de la thèse .....	84
Isolement des souches de <i>Clostridium botulinum</i> du groupe III .....	84
1. Contexte de l'étude et présentation de l'article.....	85
2. Article n°1 .....	87
Évaluation d'une méthode MLVA comme outil de typage et de traçabilité des souches de <i>C. novyi sensu lato</i> .....	100
1. Développement de la méthode MLVA pour la caractérisation des souches de <i>C. botulinum</i> du groupe III .....	101
1.1. Identification des régions VNTR d'intérêt et design des amorces.....	101
1.2. Évaluation de la méthode MLVA .....	101
1.3. Transfert de la méthode au laboratoire de Ploufragan-Plouzané-Niort.....	105
2. Évaluation de la méthode MLVA sur des échantillons du LNR botulisme aviaire .	105
2.1. Changement de kit d'électrophorèse capillaire .....	105
2.2. Évaluation de la méthode sur différents types de souches de <i>C. novyi sensu lato</i>	106
2.3. Application de la méthode à l'étude d'épisodes de botulisme animal .....	107
2.3.1. Exemple 1 d'application : Épisode de botulisme de type D/C .....	108
2.3.1.1. Présentation de l'épisode et du suivi dans le temps.....	108
2.3.1.2. Utilisation de la MLVA.....	109
2.3.1.3. Article n°2 .....	110
2.3.2. Exemple 2 d'application : Épisode de botulisme de type C/D .....	123
2.3.2.1. Présentation de l'épisode et du suivi dans le temps.....	123
2.3.2.2. Utilisation de la MLVA pour tracer la contamination.....	123
3. Calcul du pouvoir discriminant de la méthode MLVA.....	125
3. Stabilité de la méthode .....	126
4. Conclusion et perspectives sur la méthode MLVA .....	126
5. Publications en lien avec la méthode MLVA .....	127
Optimisation des étapes du séquençage et analyses comparatives des génomes de <i>C. novyi sensu lato</i> .....	132
1. Optimisation des étapes menant au séquençage du génome complet .....	133
1.1. Comparaison des kits d'extraction d'ADN .....	133
1.1.1. Contexte de l'étude .....	133
1.1.2. Article N°3 .....	135
1.1.3. Conclusions générales de l'étude et perspectives .....	146
1.2. Préparation des librairies (Séquençage short-reads) .....	146
2. Analyses comparatives des génomes de <i>C. novyi sensu lato</i> .....	149

2.1.	Analyses comparatives de l'ensemble du génome .....	149
2.1.1.	Comparaison de la taille des génomes par lignée .....	149
2.1.2.	Séquençage de souches isolées pendant le projet de thèse .....	150
2.1.3.	Analyses kmer.....	151
2.1.3.1.	Résultats de la comparaison des génomes .....	153
2.1.3.2.	Séquençage d'isolats impliqués dans un épisode de botulisme aviaire... 154	
2.1.3.3.	Recherche de cibles d'intérêt pour déterminer la lignée d'une souche ... 157	
2.2.	Analyses des systèmes CRISPR-Cas .....	160
2.2.1.	Présentation de l'étude .....	160
2.2.2.	Article N°4.....	162
2.2.3.	Utilisation des systèmes CRISPR-Cas pour le développement d'un outil de typage 180	
	Discussion et Perspectives .....	181
1.	Développement d'une méthode d'isolement des souches de <i>C. botulinum</i> du groupe III 182	
2.	Évaluation de l'outil de typage MLVA.....	183
3.	Sélection des protocoles d'extraction d'ADN et séquençage du génome .....	184
4.	Recherche de cibles d'intérêt pour le typage de <i>C. novyi sensu lato</i> .....	187
5.	Analyse des systèmes CRISPR-Cas.....	188
	Bibliographie.....	191
	ANNEXES .....	207

## Liste des Abréviations

ADN	Acide DésoxyriboNucléique
AFLP	Amplified Fragment Length Polymorphism
AMM	Autorisation de Mise sur le Marché
ARN	Acide RiboNucléique
ATU	Autorisation Temporaire d'Utilisation
BLAST	Basic Local Alignment Search Tool
<i>Bont</i>	Botulinum neurotoxin gene
BoNT	Botulinum NeuroToxin
Cas	CRISPR Associated protein
<i>C. botulinum</i>	<i>Clostridium botulinum</i>
<i>C. novyi sensu lato</i>	<i>Clostridium novyi sensu lato</i>
<i>C. haemolyticum</i>	<i>Clostridium haemolyticum</i>
<i>C. novyi</i>	<i>Clostridium novyi</i>
CDC	Centers for Disease Control and Prevention
CNR	Centre National de Référence
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DN	DNeasy Ultraclean Microbial kit
EGM	Elément Génétique Mobile
EYA	Egg Yolk Agar
FCMM	Forced Cooked Meat Medium
FliC	Flagellin protein C
GVB	Génétique Virale et Biosécurité
HA	Hémagglutinine
HQPAP	Hygiène et Qualité des Produits Avicoles et Porcins
ISS	Istituto Superiore di Sanità
kDa	KiloDalton
LNR	Laboratoire National de Référence
NAP	Neurotoxin Associated Protein
NTNH	Non-Toxin Non-Haemagglutin Protein
MALDI-TOF	Matrix Assisted Laser Desorption Ionisation-Time Of Flight
Mb	Mégabase
MDO	Maladie à Déclaration Obligatoire
MLST	Multi Loci Sequence Typing
MLVA	Multi Loci VNTR Analysis
MOT	Micro-Organismes et Toxines hautement pathogènes
MPC	MasterPure Complete DNA and RNA purification kit
MPG	MasterPure Gram positive DNA purification kit
MRC	Maladie Réputée Contagieuse
PCR	Polymerase Chain Reaction
PFGE	Pulsed-Field Gel Electrophoresis
PG	Plasmid Group
PTC	Progenitor Toxin Complex
QA	Qiagen QIAamp DNA Mini Kit

RAPD	Random Amplified Polymorphic DNA
RKI	Robert Koch Institute
RNOEA	Réseau National d'Observations Épidémiologiques en Aviculture Pathologie aviaire
SNARE	Soluble N-éthylmaleimide-sensitive-factor Attachment protein REceptor
SNAP	Synaptosomal-Associated Protein
SNP	Single Nucleotid Polymorphism
TeNT	Tetanus NeuroToxin
TPGY	Tryptone Peptone Glucose Yeast Extract
VAMP	Vesicle-Associated Membrane Protein
VNTR	Variable Number of Tandem Repeats
WGS	Whole Genome Sequencing

## Liste des figures

Figure 1 : Micrographies électroniques des différentes étapes de la sporulation chez <i>Clostridium botulinum</i> du groupe III.....	21
Figure 2 : Structure cristallographique de la neurotoxine botulique (BoNT) .....	25
Figure 3 : Structure du precursor toxin complex (PTC).....	26
Figure 4 : Fonctionnement normal d'une cellule nerveuse lors de la transmission du signal nerveux aux cellules musculaires. ....	27
Figure 5 : Mécanisme d'action de la neurotoxine botulique. Action de la toxine sur la libération de l'acétylcholine après son internalisation dans la cellule nerveuse. ....	28
Figure 6 : Dendrogramme représentant l'ensemble des sous-types de BoNT basé sur leur différence au niveau de la composition en acides aminés de la toxine. ....	30
Figure 7 : Représentation schématique de la structure des neurotoxines botuliques produites par les souches de <i>Clostridium botulinum</i> du groupe III (C, D, C/D et D/C) .....	32
Figure 8 : Pathogénèse du botulisme chez les volailles. ....	39
Figure 9 : Cycle de contamination « carcass-maggot » des oiseaux sauvages aquatiques. ....	41
Figure 10 : Représentation schématique du cycle épidémiologique de <i>C. botulinum</i> dans l'écosystème d'eau douce.....	42
Figure 11 : Influence des conditions environnementales sur les pics de mortalité lors d'un évènement de botulisme aviaire. ....	43
Figure 12 : Heat-map présentant la similitude moyenne des génomes complets de <i>C. novyi sensu lato</i> et leur regroupement par lignée en utilisant le logiciel Gegenees. ....	50
Figure 13 : Géloses EYA ensemencées avec des souches de <i>C. botulinum</i> du groupe III .....	52
Figure 14 : Liste des différentes classes et types de systèmes CRISPR-Cas et rôles des différents gènes <i>cas</i> . ....	58
Figure 15 : Fonctionnement d'un système CRISPR-Cas.....	59
Figure 16 : Gel d'électrophorèse en champ pulsé testé sur 4 souches de <i>C. botulinum</i> type C responsables de botulisme aviaire .....	63
Figure 17 : Dendrogramme des pulsotypes de 14 souches isolées de <i>C. botulinum</i> du groupe III. ....	63
Figure 18 : Dendrogramme de 174 souches de <i>C. botulinum</i> basé sur les résultats obtenus avec la méthode AFLP. ....	65

Figure 19 : Résultats des différents profils de plasmidome obtenus sur 110 extraits d'ADN positifs à différentes toxines de <i>C. botulinum</i> groupe III. ....	73
Figure 20 : Matrice de similarités des 61 éléments génétiques mobiles caractérisés chez <i>C. novyi sensu lato</i> . ....	76
Figure 21 : Arbre phylogénétique obtenu après analyse des SNP de plusieurs génomes de <i>C. botulinum</i> du groupe III. ....	77
Figure 22 : Empreintes spectrales obtenues suite à l'analyse au MALDI-TOF de trois échantillons de <i>C. novyi sensu lato</i> ( <i>C. novyi</i> ATCC 25758, <i>C. haemolyticum</i> ATCC 9650 et <i>C. botulinum</i> type C CCUG 7970). ....	80
Figure 23 : Dendrogramme calculé à partir des codes MLVA des échantillons testés lors du développement de la méthode à l'ISS. ....	104
Figure 24 : Dindes présentant des signes cliniques dans un des trois bâtiments affectés. ....	123
Figure 25 : Arbre phylogénétique présentant la distribution des souches de <i>C. novyi sensu lato</i> sur la base d'une analyse kmer des génomes. ....	152
Figure 26 : Arbre phylogénétique présentant la distribution des souches de <i>C. novyi sensu lato</i> sur la base d'une analyse kmer des génomes. ....	156
Figure 27 : Arbre phylogénétique présentant la distribution des souches de <i>C. novyi sensu lato</i> sur la base d'une analyse kmer des génomes. ....	158

## Liste des équations

Équation 1 : Indice Hunter-Gaston.....	125
--	-----

## Liste des Tableaux

Tableau 1 : Caractéristiques phénotypiques des 6 différents groupes de clostridies capables de produire la neurotoxine botulique. ....	24
Tableau 2 : Liste des différentes espèces bactériennes capables de produire la toxine botulique et leurs types et sous-types toxiques associés .....	29
Tableau 3 : Liste des principaux animaux affectés par le botulisme et les types de BoNT associés pour chaque hôte .....	34
Tableau 4 : Toxicité des différentes neurotoxines botuliques selon l'espèce aviaire testée.....	36
Tableau 5 : Liste des méthodes de typage existantes pour la caractérisation de <i>C. novyi sensu lato</i> .....	61
Tableau 6 : Matrice de similarités des 7 types de gène <i>fliC</i> caractérisés chez <i>C. botulinum</i> du groupe III.....	71
Tableau 7 : Liste des amorces pour l'amplification des 9 régions VNTR de la méthode MLVA développée pour le typage des souches de <i>Clostridium botulinum</i> du groupe III. Tableau de correspondance des tailles des amplicons selon le nombre de répétitions pour chaque VNTR étudié. ....	103
Tableau 8 : Liste des échantillons testés pour la deuxième étude terrain et leurs codes MLVA correspondants.....	124
Tableau 9 : Indices de discrimination Hunter et Gaston (HGDI) calculés pour chaque VNTR et pour le schéma MLVA avec 7 VNTR.....	125
Tableau 10 : Quantités d'ADN amplifiés selon le kit de préparation des bibliothèques utilisé....	147
Tableau 11 : Exemples de kits utilisés pour la préparation des bibliothèques pour le séquençage Illumina de <i>C. novyi</i> et différents groupes de <i>C. botulinum</i> .....	149
Tableau 12 : Tailles moyennes des génomes (chromosome et EGM) de <i>C. novyi sensu lato</i> en fonction des lignées. ....	150



## Introduction

### **I. Généralités sur *Clostridium botulinum* et le botulisme**

#### **1. Histoire du botulisme, de la découverte de la bactérie et de la toxine au bioterrorisme**

##### **1.1. Découverte de la bactérie et premières descriptions des symptômes**

Les plus anciens documents historiques mentionnant des cas de botulisme humain datent de la fin du 18<sup>ème</sup> siècle où le manque de mesures sanitaires dans la production de nourriture en Europe a conduit à de nombreux décès liés à des maladies d'origine alimentaire (Ting & Freiman, 2004). La première source de botulisme identifiée est la consommation de saucisses fumées qui a été décrite entre 1817 et 1822 par le médecin allemand Justinus Andreas Christian Kerner. Dans ses publications (Kerner, 1820, 1822), il décrit entre autres les symptômes cliniques du botulisme causés par « le poison de saucisse ». Le poison se développe en conditions anaérobies dans les saucisses putrides et va entraîner une paralysie flasque causée par l'arrêt de la transmission du signal moteur dans le système nerveux périphérique. Ce poison est létal à petite dose. Les principaux symptômes décrits sont une mydriase, une diminution de la sécrétion des glandes salivaires, une paralysie gastro-intestinale et de la vessie, une dysphagie et des insuffisances respiratoires (Erbguth, 2008; Ting & Freiman, 2004).

La première personne à avoir isolé la bactérie est Emile Pierre Marie van Ermengem, un microbiologiste de l'université de Ghent, qui, à la suite d'un épisode de botulisme en 1895, parvient à mettre en corrélation de la saucisse de porc contaminée et la mort de 3 personnes qui avaient consommé l'aliment. Il a isolé une bactérie en forme de bacille capable de produire des spores à partir des restes de viande crue et des foies des victimes de l'intoxication alimentaire. Il a nommé cette bactérie *Bacillus botulinus* (*botulus* signifie saucisse en latin) qui deviendra *Clostridium botulinum* quelques années plus tard. En 1904, un épisode de botulisme dû à la consommation d'une salade de haricots blancs contaminés en conserve a permis de démontrer que la bactérie et la toxine peuvent aussi être présentes dans des aliments qui ne sont pas de la viande ou du poisson (Erbguth, 2008; Popoff, 1995).

##### **1.2. Découverte de plusieurs types de neurotoxine botulique et mode d'action.**

Dans la première partie du 20<sup>ème</sup> siècle, il a été démontré que les toxines retrouvées en 1895 et en 1904 (épisodes décrits dans le paragraphe précédent) étaient distinctes sérologiquement. Elles ont été désignées par la suite comme les toxines de type A et B par Georgina Burke (Burke, 1919). La

toxine de type C a été découverte en 1922 aux Etats-Unis par Bengtson (Bengtson, 1922) et en Australie par Seddon (Seddon, 1922). Les types D et E ont été caractérisés respectivement en 1928 par Meyer et Gunnison aux Etats-Unis (Meyer & Gunnison, 1929) et en 1936 par Bier en Ukraine (Gunnison et al., 1936). Enfin, les toxines de type F et G ont été découvertes dans la deuxième partie du 20<sup>ème</sup> siècle respectivement en 1960 par Moller et Scheibel en Scandinavie (Moller & Scheibel, 1960) et en 1970 par Gimenez et Ciccarelli en Argentine (Erbguth, 2008; Geiges, 2002; Gimenez & Ciccarelli, 1970).

Les mécanismes d'action de la toxine botulique ont par la suite été largement étudiés. Dès 1949, des chercheurs de Londres ont découvert que la toxine botulique agissait en bloquant la libération de l'acétylcholine au niveau de la jonction neuromusculaire (Sellin, 1981). Il a ensuite fallu attendre les années 90 pour que Giampetro Schiavo et son équipe établissent que la toxine botulique est composée d'une métalloprotéase spécifique des protéines SNARE (Soluble N-éthylmaleimide-sensitive-factor Attachment protein REceptor) au même titre que la toxine tétanique. La synaptobrevine est le premier substrat cible identifié (Schiavo et al., 1992). Ce sont les sérotypes B, D, F et G qui sont capables de cibler cette protéine SNARE. La même équipe a ensuite déterminé que la cible des neurotoxines des sérotypes A, C et E est la protéine SNARE SNAP-25 (Blasi et al., 1993a; Schiavo et al., 1993). Enfin, une autre équipe de recherche a mis en évidence que la neurotoxine du sérotype C est capable de cliver une autre protéine de la membrane présynaptique appelée syntaxine (Blasi et al., 1993b). Les mécanismes sont présentés plus en détail dans le chapitre I.3.2.

### **1.3. Utilisation détournée de la toxine botulique**

#### **1.3.1. Utilisation comme traitements thérapeutiques et comme produits cosmétiques**

Dans les années 60, il a été démontré avec succès sur des singes puis sur des hommes atteints de strabisme et de blépharospasme que l'injection de quelques picogrammes de toxine entraîne une paralysie des muscles sans effets secondaires. L'utilisation de la toxine pour traiter le strabisme et le blépharospasme a été approuvée aux États-Unis par la Food and Drug Administration (FDA) en 1989 sous le nom de « Oculinum » (Scott, 2004), plus tard renommée Botox® lors de l'acquisition des droits de distribution de la toxine thérapeutique par l'entreprise Allergan (Ting & Freiman, 2004). Dans les années 90, le Dr Carruthers observa la disparition des rides au-dessus des sourcils suite à l'utilisation du Botox® sur des patients atteints de blépharospasme. Avec son mari dermatologiste, ils publièrent un article sur l'utilisation du Botox® à des fins cosmétiques (Carruthers et al., 1996). Encore à ce jour, de nombreux traitements utilisent le Botox® dans les domaines de la dermatologie et de la cosmétique (Naik, 2021; Ting & Freiman, 2004).

Au Royaume-Uni, la recherche et production de la toxine fut effectuée par le « Centre for Applied Microbiology and Research » (CAMR) et des expérimentations réalisées pour traiter des patients atteints de contractions du muscle perioriculaire (Rowe & Noonan, 2017). L'utilisation de la toxine botulique a été approuvée pour traiter des personnes atteintes de dystonie sous le nom de Dysport®. Aujourd'hui, cette toxine est également utilisée pour traiter des blépharospasmes et des dystonies cervicales (Naik, 2021).

D'autres toxines botuliques sont actuellement commercialisées : Myobloc™ aux USA, Xeomin® et Neurobloc® en Europe (Benecke et al., 2005; Lew et al., 2000) sont utilisées pour traiter les dystonies cervicales (Naik, 2021). Une toxine de type A est aussi produite en Chine (Erbguth, 2008).

En oncologie, la neurotoxine botulique est utilisée comme analgésique durant et après le traitement de certains types de cancer pour calmer des douleurs locales liées à la chirurgie ou à la chimiothérapie. Des études sur l'utilisation de la toxine pour réduire la taille des tumeurs et induire l'apoptose chez les cellules cancéreuses ont montré des résultats prometteurs *in vitro* et *in vivo* (sur des animaux et humains). Ces types de traitements pourraient être utilisés plus fréquemment dans un futur proche, en tenant compte des effets secondaires possibles du fait de l'utilisation de la toxine botulique (Grenda et al., 2021).

En dermatologie et cosmétologie, la neurotoxine botulique peut aider à traiter des conditions telles que l'alopecie, la sécrétion excessive de sébum rendant la peau grasse, la prévention des cicatrices post opératoires, le psoriasis, le syndrome de Raynaud, l'hyperhydrose, la notalgie paresthésique et d'autres problèmes de peau (Naik, 2021). Elle peut aussi traiter temporairement le sourire gingival (Zengiski et al., 2021).

### **1.3.2. Utilisation comme arme biologique**

L'utilisation de la toxine botulique comme arme biologique a été envisagée dès les années 30 par les armées japonaise, américaine, canadienne et soviétique lors de la seconde guerre mondiale. Malgré de nombreuses avancées dans les connaissances relatives à la toxine au cours de cette période, le projet fut un échec et la toxine n'a pas été utilisée à des fins militaires lors de cette guerre (Tatu & Feugeas, 2021). La forme cristallographique de la toxine A a été obtenue pour la première fois à l'aide des techniques de concentration et de cristallisation mises en place à Fort Detrick (USA) par les équipes de Carl Lamanna et James Duff en 1946, soit un an après la fin de la guerre (Lamanna et al., 1946). Les recherches ont continué aux Etats-Unis jusqu'au début des années 70, période durant laquelle le président Nixon a pris la décision d'arrêter les programmes de recherches liés au

développement d'armes biologiques. L'Irak et la Russie ont produit des armes chargées avec de la toxine botulique malgré la signature en 1972 d'une convention interdisant la recherche, production et utilisation d'armes biologiques.

Récemment, les inquiétudes se sont plus portées sur l'utilisation de la toxine à des fins de bioterrorisme. Depuis 2001, *C. botulinum* fait partie des agents biologiques représentant une des menaces les plus importantes avec *Bacillus anthracis*, *Yersina pestis* ou encore *Francisella tularensis* (Broussard, 2001). Trois utilisations potentielles à des fins malveillantes ont été identifiées : son utilisation sous forme d'injection, la contamination d'aliments ou boissons avec la toxine et sa libération sous forme aéroportée (Cenciarelli et al., 2019). Néanmoins, plusieurs facteurs rendent l'utilisation de la toxine botulique comme arme biologique très difficile : la quantité nécessaire et le taux de mortalité variable lors de l'ingestion de la toxine, sa dégradation rapide à l'air libre, le fait que la maladie ne soit pas transmissible d'une personne à l'autre et que le traitement soit relativement facile, sous réserve d'une prise en charge rapide (antitoxine, antibiotiques) (Cenciarelli et al., 2019; Ting & Freiman, 2004).

Du fait de ce contexte, le botulisme (les toxines de *C. botulinum*) fait partie des maladies et agents biologiques classés dans la catégorie A par le CDC (qui comprend les maladies et agents hautement prioritaires présentant un risque majeur pour la sécurité nationale).

En France, *C. botulinum* et l'ensemble des toxines botuliques sont classés dans la liste des microorganismes et toxines hautement pathogènes (MOT). Cette liste comprend des agents dont l'utilisation pourrait présenter un risque pour la santé publique (Annexe 2 de l'article L. 5139-1 du code de la santé publique dont l'arrêté date du 30 Avril 2012). Les opérations de production, fabrication, transport, importation, exportation, détention, offre, cession et emploi portant sur ces agents MOT sont soumises à des autorisations qui sont délivrées par l'Agence Nationale de Sécurité du Médicament (ANSM). Pour manipuler la bactérie ou la toxine, une habilitation est nécessaire pour le personnel des laboratoires autorisés à manipuler ces agents MOT et le laboratoire doit détenir les autorisations pour pouvoir mener des travaux de recherche sur les agents pathogènes listés dans ce texte réglementaire.

Le botulisme est rapporté depuis l'antiquité et étudié depuis le XIX<sup>ème</sup> siècle. Les travaux portant sur *Clostridium botulinum* se placent essentiellement dans un contexte de sécurité alimentaire. La toxine botulique a quant à elle été largement étudiée et ses propriétés détournées à des fins cosmétiques, thérapeutiques et militaires, aboutissant à la mise en place d'une réglementation contraignante en France (Règlementation MOT).

## **2. *Clostridium botulinum*, la bactérie**

### **2.1. Généralités sur la bactérie**

*Clostridium botulinum* est une espèce bactérienne appartenant au genre *Clostridium*, dont la classification a été remise en question (Collins et al., 1994) suite aux analyses génomiques menées du fait de la disponibilité croissante de génomes de ce genre bactérien grâce aux techniques de séquençage de seconde génération. Une étude portant sur la comparaison de génomes d'espèces bactériennes du genre *Clostridium* a en effet révélé de nombreuses incohérences taxonomiques suggérant que plusieurs genres distincts seraient en fait présents dans ce groupe (Cruz-Morales et al., 2019). Cela a par exemple abouti à la modification de l'espèce *Clostridium difficile* en *Clostridioides difficile* (Lawson et al., 2016).

*C. botulinum* produit la neurotoxine botulique (BoNT), responsable du botulisme chez l'homme et l'animal. C'est une bactérie anaérobie stricte se présentant sous la forme d'un bacille avec une coloration de Gram positif (Figure 1) (Portinha et al., 2022). Elle possède une mobilité grâce à la présence de flagelles autour de la bactérie. Elle est capable de sporuler, ce qui lui permet de persister dans des environnements peu propices à sa croissance. C'est une espèce ubiquitaire, le réservoir est l'environnement (sol, sédiments marins et d'eau douce) mais des spores sont occasionnellement retrouvées en portage chez l'homme et certains animaux (Peck, 2009).

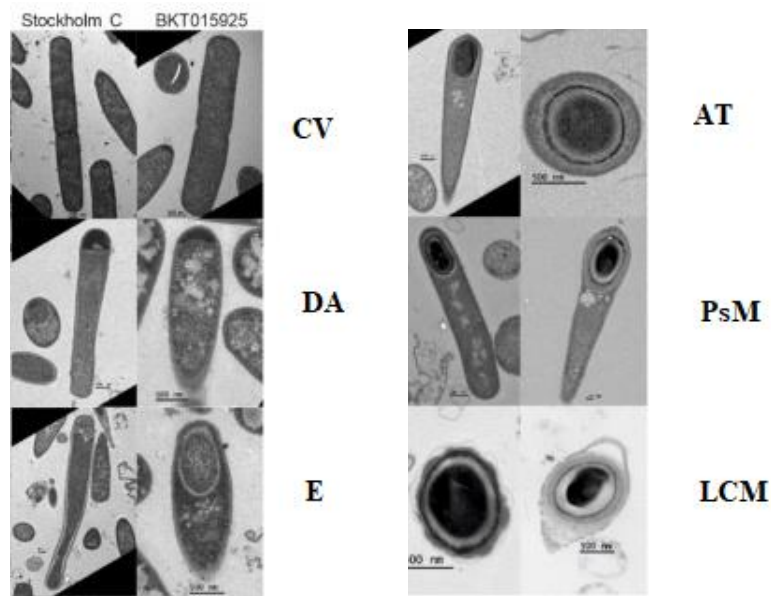


Figure 1 : Micrographies électroniques des différentes étapes de la sporulation chez *Clostridium botulinum* du groupe III (souches C/D BKT015925 et C Stockholm). CV = Cellule végétative – DA = Division asymétrique – E = Englobement de la préspore – AT = Assemblage de la tunique – PsM = Préspore mature – LCM = Lyse de la cellule végétative mère.

Source : Portinha et al, 2022.

Les souches de *C. botulinum* possèdent le gène codant la neurotoxine botulique. En effet, une souche ne possédant pas ce gène ne peut être considérée comme *C. botulinum* mais comme une clostridie équivalente non toxigène (Peck, 2009; Smith et al., 2018). Cependant, d'autres espèces de clostridies sont capables de produire des toxines botuliques comme *Clostridium butyricum* et *Clostridium baratii* (Aureli et al., 1986; Hall et al., 1985).

Enfin, *C. botulinum* est classé dans le groupe 2 des agents pathogènes (Arrêté du 18 Juillet 1994, modifié le 16 février 2018). La bactérie doit être manipulée dans un confinement de classe L2 avec des équipements de protection adéquats et adaptés. Comme évoqué précédemment, ce pathogène fait partie de la liste des MOT, ce qui implique des contraintes et restrictions en termes de sécurité et sûreté biologique pour répondre aux exigences de cette réglementation.

## 2.2. Classification des bactéries productrices de BoNT

Les clostridies capables de produire la neurotoxine botulique sont classées en 6 groupes distincts selon le type de neurotoxine produite, les conditions de croissance, la résistance des spores

à la chaleur et plusieurs caractéristiques phénotypiques et biochimiques (Tableau 1). Sept types de BoNT ont été décrits historiquement, classés de A à G, et plus de 40 sous types (Peck et al., 2017).

Plus récemment, deux nouveaux types de BoNT ont été découverts :

- BoNT H qui est une mosaïque des BoNTs F et A dont le gène a été retrouvé dans une souche de *C. botulinum* bivalente portant aussi le gène codant BoNT B (Gonzalez-Escalona et al., 2014).
- BoNT X qui a été identifiée et caractérisée en 2017 (Zhang et al., 2017) à partir d'une souche de *C. botulinum* isolée lors d'un cas de botulisme infantile en Mars 1995 dans un échantillon où la toxine BoNT B a été détectée (Kakinuma et al., 1996).

Les souches désignées comme *C. botulinum* sont seulement présentes dans les trois premiers groupes (sur les 6 groupes actuellement décrits). Les souches du groupe I vont produire les BoNT A, B, F, H et X. Elles ont une activité protéolytique : elles sécrètent une protéase capable d'activer les neurotoxines botulique (en coupant le pont disulfure) à l'extérieur de la bactérie. Elles sont lipases positives mais ne possèdent cependant pas de lécithinase. Ces spores sont thermorésistantes avec une valeur  $D_{121^{\circ}\text{C}}$  de 0,19 minute représentant le temps nécessaire pour diviser par 10 la population bactérienne à  $121^{\circ}\text{C}$  et une valeur de  $z$  d'environ  $11,3^{\circ}\text{C}$  représentant la variation de température nécessaire pour faire varier d'un facteur 10 la valeur  $D$  (Diao et al., 2014). Les souches du groupe I impliquées dans les foyers de botulisme humain sont majoritairement liées à du botulisme alimentaire (conserves de légumes ou produits de charcuterie) (Lindström et al., 2006). L'équivalent non toxique de *C. botulinum* du groupe I possédant les mêmes caractéristiques biochimiques est l'espèce *C. sporogenes*.

Les souches du groupe II, produisant les BoNTs B, E et F, ne sont pas protéolytiques (l'activation de la toxine se fait par une protéase exogène non produite par la bactérie elle-même, elle peut avoir lieu par des protéases d'origine digestive telle que la trypsine par exemple) et la thermorésistance de leurs spores est modérée avec une valeur  $D_{80^{\circ}\text{C}}$  se situant entre 1 et 1,5 minute et une valeur de  $z$  se situant entre 6 et  $6,5^{\circ}\text{C}$ . Elles peuvent cependant croître et produire de la toxine à des températures relativement basses ( $3,5^{\circ}\text{C}$ ). Les spores de *C. botulinum* du groupe II sont donc capables de croître et produire la toxine botulique au réfrigérateur, constituant un challenge pour garantir la sécurité sanitaire des denrées sensibles conservées à ces températures. Les souches de *C. botulinum* du groupe II responsables de botulisme vont principalement être retrouvées dans des aliments transformés industriellement conservés à basse température et dans des aliments à base de poisson (Lindström et al., 2006). Son équivalent non toxique est *C. taeniosporum*.

Enfin, les souches du groupe III produisent les BoNTs C, D ainsi que les mosaïques C/D et D/C. Elles peuvent croître à des températures allant jusqu'à 40°C, ne possèdent pas d'activité protéolytique. Elles sont lipases positives et la plupart des souches ont une activité lécithinase. La thermorésistance des spores est intermédiaire par rapport aux deux premiers groupes avec une valeur  $D_{104^{\circ}\text{C}}$  se situant entre 0,1 et 0,9 minutes et une valeur  $D_{90^{\circ}\text{C}}$  se situant entre 11,5 et 11,9 minutes pour une valeur de  $z$  de 5,7 °C (Portinha et al., 2022; Segner & Schmidt, 1971). Le botulisme animal est principalement associé à *C. botulinum* du groupe III. Leur équivalent non toxique est *Clostridium novyi* (Peck, 2009; Smith et al., 2018). Les particularités des souches de *C. botulinum* du groupe III sont décrites plus précisément dans le chapitre dédié au groupe d'espèces *C. novyi sensu lato* dont fait partie *C. botulinum* du groupe III (cf III).

Les autres groupes ne sont pas désignés comme *C. botulinum* malgré leur capacité à produire la toxine botulique. Cela est lié aux différences phénotypiques et génétiques entre ces espèces et les trois groupes de *C. botulinum* (Peck, 2009). Les souches du groupe IV, initialement désignées comme *C. botulinum*, sont maintenant classées comme étant l'espèce *C. argentinense*. Elles sont capables de produire la BoNT G, ont une activité protéolytique mais ne possèdent pas d'activité lipase et lécithinase. Elles sont génétiquement proches de l'espèce *C. subterminale* (Suen et al., 1988). Cette toxine n'a jamais été associée à un cas clinique. Enfin, certaines espèces de *C. butyricum* (groupe V) et *C. baratii* (groupe VI) ont la capacité de produire respectivement des BoNTs de type E et F et ont déjà été associées à des cas cliniques humains (Aureli et al., 1986; Dykes et al., 2015; Hall et al., 1985; Tréhard et al., 2016).

Suite au séquençage d'un grand nombre de souches et à l'augmentation du nombre de données permettant de mieux discriminer les souches des 6 groupes, un changement dans la classification des clostridies capables de produire la neurotoxine botulique a été proposé par de nombreux scientifiques (Peck, 2009; Peck et al., 2017; Skarin et al., 2011; Smith et al., 2018).

Sur la base d'analyses bio-informatiques, il a aussi été démontré que des espèces bactériennes n'appartenant pas au genre *Clostridium* possèdent des gènes présentant des homologies et structurations similaires aux toxines botuliques. Une étude en 2015 a ainsi identifié un gène apparenté au gène *bont* (BoNT I ou BoNT Wo) dans le génome de l'espèce *Weisella oryzae*, une bactérie Gram positif non sporulante présente dans le riz fermenté (Mansfield et al., 2015). Un autre gène a été identifié dans le génome d'une bactérie Gram négatif, *Chryseobacterium piperi*, bien qu'aucune production de BoNT n'ait été démontrée (Wentz et al., 2017). Une souche d'*Enterococcus faecium* isolée à partir d'un prélèvement de fèces de vache possède un gène apparenté au gène *bont* sur un



plasmide et il a été démontré que la toxine produite *in vitro* (nommée BoNT J ou BoNT/En) est fonctionnelle (Zhang et al., 2018).

Tableau 1 : Caractéristiques phénotypiques des 6 différents groupes de clostridies capables de produire la neurotoxine botulique.

Sources : Peck et al., 2009 et Smith et al., 2018.

	Groupe I	Groupe II	Groupe III	Groupe IV	Groupe V	Groupe VI
Clostridie toxigène	<i>C. botulinum</i>			<i>C. argentinense</i>	<i>C. butyricum</i>	<i>C. baratii</i>
BoNTs produites	A, B, F, H, X	B,E,F	C, D	G	E	F
Équivalent non toxigène	<i>C. sporogenes</i>	<i>C. taeniosporum</i>	<i>C. novyi</i>	<i>C. subterminale</i>	/	/
T°C optimale de croissance	35-40°C	18-25°C	37-40°C	37°C	30-37°C	30-45°C
Activité protéolytique	+	-	-	+	-	-
Lipase	+	+	+	-	-	-
Lécithinase	-	-	+/-	-	-	+
Liquéfaction de la gélatine	+	+	+	+	-	-
Digestion de la caséine	+	-	-	+	-	-
Fermentation du glucose	+	+	+	-	+	+
Résistance à la chaleur des spores	+++	--	+	+	-	/

Les clostridies capables de produire la toxine botulique sont réparties en 6 groupes, présentant des caractéristiques phénotypiques et biochimiques très différentes. Elles sont toutes anaérobies et sporulantes. Les toxines botuliques produites varient en fonction des groupes et sont détaillées dans le paragraphe suivant.

### 3. La toxine botulique

#### 3.1. Structures et fonctions de la toxine et des protéines associées.

La neurotoxine botulique est une protéine polypeptidique qui fait partie de la famille des métalloendopeptidases, d'un poids moléculaire d'environ 150 kDa et qui contient un atome de zinc au niveau du domaine catalytique de la protéine. La protéine est composée d'une chaîne légère (LC) d'environ 50 kDa et d'une chaîne lourde (HC) d'environ 100 kDa (Figure 2A) (Lim et al., 2019). Ces chaînes sont reliées par un pont disulfure qui se forme entre 2 cystéines, une présente à l'extrémité C-terminale de la chaîne légère et une présente à l'extrémité N-terminale de la chaîne lourde (Figure 2B) (Lim et al., 2019). La chaîne lourde se divise en deux sous-unités aux fonctions bien distinctes : la partie proche de l'extrémité C-terminale va permettre la liaison entre la toxine et le récepteur cible présent au niveau de la membrane cellulaire des neurones ainsi que l'internalisation de la toxine par endocytose. L'autre partie proche de l'extrémité N-terminale va quant à elle permettre la translocation de la chaîne légère de la vésicule vers le cytoplasme, qui va ainsi pouvoir rejoindre sa cible. La chaîne légère est une métalloprotéase où se trouve le domaine catalytique permettant le clivage de sites spécifiques présents sur les 3 protéines SNARE (Soluble N-ethylmaleimide sensitive factor Attachment Receptor) cibles (synaptobrévine, SNAP-25 Synaptosomal-Associated Protein-25, une protéine de 206 acides aminés associée aux membranes lipidiques via des cystéines, et syntaxine).

Selon le type de neurotoxine, la protéine SNARE ciblée n'est pas la même. Malgré les différences présentes entre les différents sérotypes de BoNT, l'architecture de la toxine est la même pour tous (Binz et al., 2010; Montecucco & Schiavo, 1995; Pirazzini et al., 2017).

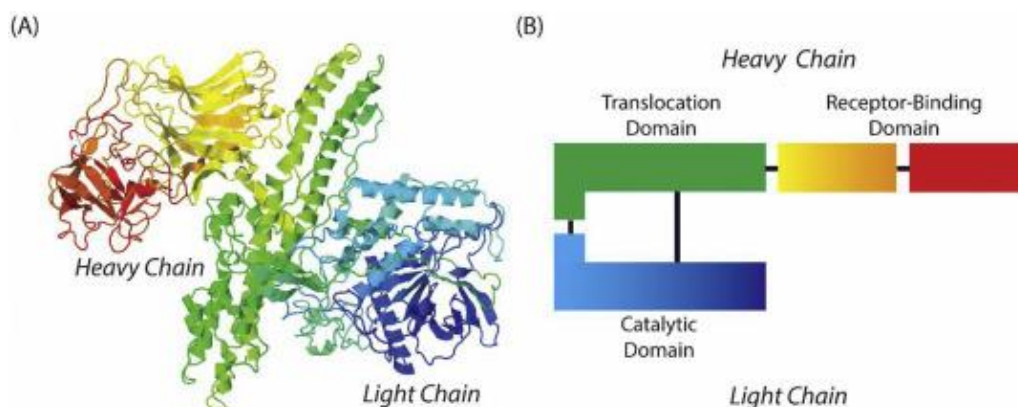


Figure 2 : A) Structure cristallographique de la neurotoxine botulique (BoNT). La chaîne légère (domaine catalytique) est représentée en bleu tandis que la chaîne lourde est représentée en vert (domaine de translocation H<sub>n</sub>), jaune et rouge (domaine de liaison H<sub>c</sub>)

B) Représentation schématique de la neurotoxine botulique (même code couleur que la figure A).

Source : Lim et al.,2019.

La toxine est d'abord synthétisée par la bactérie sous la forme d'une seule chaîne protéique de 150 kDa. Ce précurseur sera ensuite clivé par des protéases pour former la chaîne légère et la chaîne lourde reliées par le pont disulfure. La toxine est produite avec d'autres protéines appelées protéines non toxiques associées (NAPs). L'ensemble de ces protéines va former le *precursor toxin complex* (PTC).

Le gène *bont* est toujours positionné à côté du gène *ntnh* codant la protéine non toxique non hémagglutinante (NTNH). Cette protéine de 140 kDa, dont la structure est similaire à celle de BoNT, fait partie des NAPs et va former avec la BoNT un hétérodimère nommé PTC fonctionnel minimal (M-PTC). NTNH va avoir un rôle de protection de BoNT pour faire face aux conditions rencontrées au niveau de l'estomac et du tractus digestif, comme le pH et la présence d'enzymes digestives.

Les autres protéines associées (ha17, ha33, ha70) sont des protéines hémagglutinantes qui semblent faciliter la transcytose de BoNT au travers de la barrière intestinale. Ces protéines vont former avec BoNT et NTNH le complexe L-PTC (Figure 3) (Lee et al., 2013). La composition du complexe protéique peut différer entre sérotypes mais aussi entre sous-types (Gu et al., 2012; Kalb et al., 2017; Pirazzini et al., 2017).

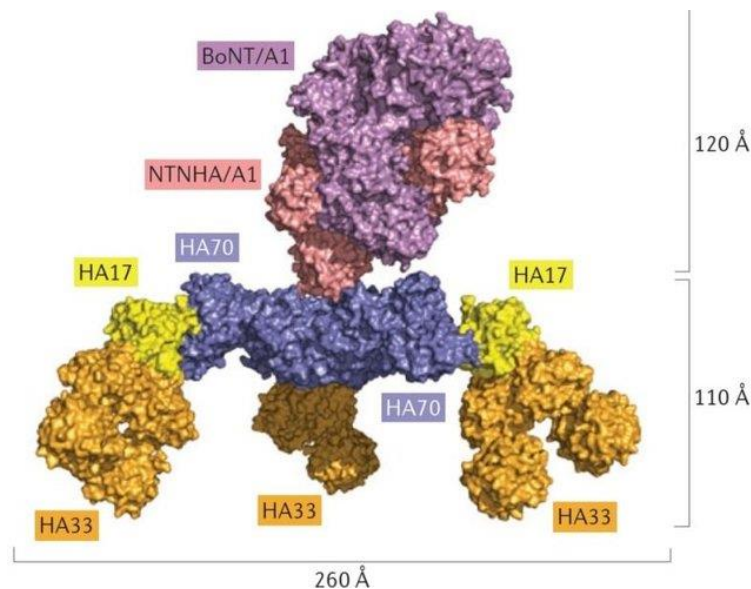


Figure 3 : Structure du precursor toxin complex (PTC). En violet, la toxine botulique (ici BoNT/A1), en rose, la protéine non hémagglutinante NTNHA (ici NTNHA/A1), en bleu, jaune et orange les protéines hémagglutinantes (ici, respectivement HA70, HA17 et HA33).

Source : Lee et al., 2013.

### 3.2. Mécanisme d'action de la toxine

Le botulisme se manifeste avec une paralysie flasque. Elle est causée par l'action de la BoNT qui empêche la libération de l'acétylcholine au niveau de la jonction neuromusculaire. Dans une configuration où la toxine botulique n'est pas présente dans l'organisme, l'arrivée d'un potentiel d'action va entraîner l'ouverture des canaux  $\text{Ca}^{2+}$  présents sur la cellule nerveuse, entraînant l'entrée d'ions  $\text{Ca}^{2+}$  dans le cytosol. Cette entrée d'ions va provoquer la fusion des vésicules synaptiques contenant l'acétylcholine, un neuromédiateur synthétisé dans la fibre nerveuse, avec la membrane plasmique. La fusion des membranes de la vésicule et de la cellule nerveuse va être possible grâce à la formation d'un complexe de plusieurs protéines SNARE. La protéine SNARE synaptobrevine est présente sur la vésicule synaptique tandis que les protéines SNAP-25 et syntaxine sont présentes sur la membrane de la cellule nerveuse. L'acétylcholine est ensuite libérée au niveau de la fente synaptique située entre la cellule nerveuse et la cellule musculaire. Cette dernière va recevoir l'information lorsque les molécules du neuromédiateur vont se fixer au niveau des récepteurs post-synaptiques, permettant ainsi la contraction des muscles (Figure 4) (Arnon et al., 2001; Montecucco & Schiavo, 1995).

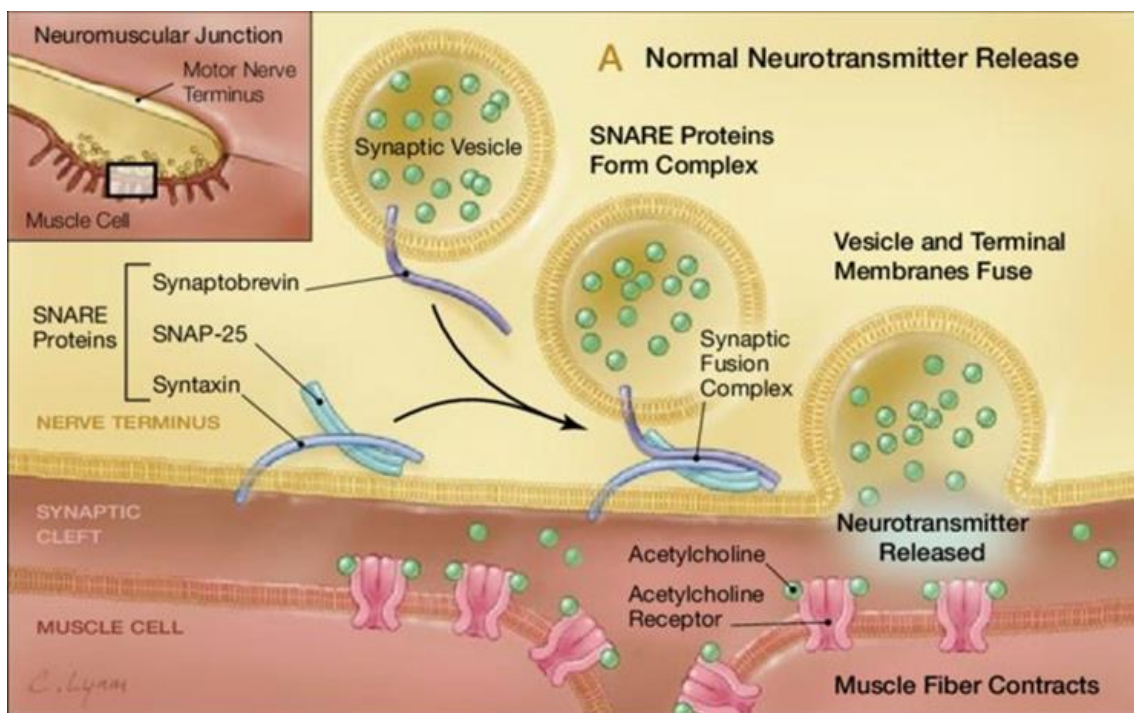


Figure 4 : Fonctionnement normal d'une cellule nerveuse lors de la transmission du signal nerveux aux cellules musculaires.

Source : Arnon et al., 2001.

La chaîne lourde de la toxine botulique va se fixer au niveau de la synapse puis être internalisée. Après sa translocation dans le cytosol, la chaîne légère va reconnaître et cliver la protéine SNARE cible spécifique du type toxinique de la toxine (Types B, D, F et G pour la synaptobrevine, types A, C et E pour SNAP-25 et type C pour la syntaxine). Cela va empêcher la formation du complexe SNARE et donc la fusion entre la vésicule synaptique et la membrane de la cellule nerveuse. Sans libération d'acétylcholine dans la fente synaptique, le signal nerveux ne peut pas être transmis aux cellules musculaires, le muscle ne peut donc pas se contracter (Figure 5) (Arnon et al., 2001; Poulain & Humeau, 2003).

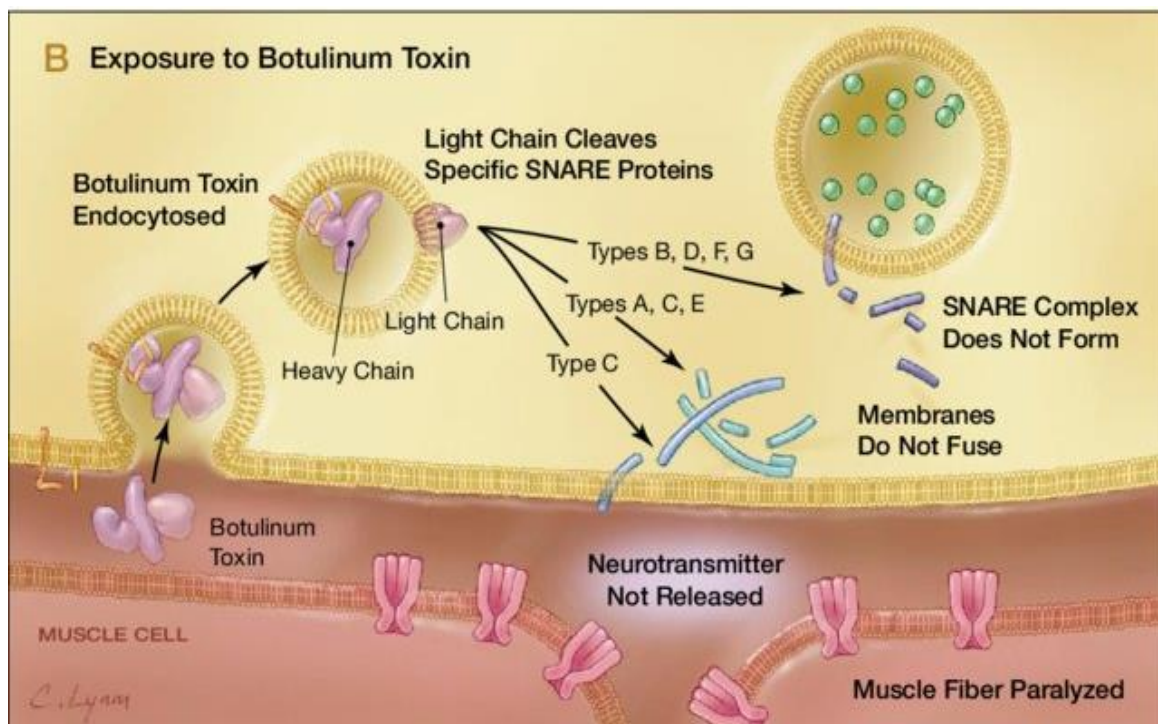


Figure 5 : Mécanisme d'action de la neurotoxine botulique. Action de la toxine sur la libération de l'acétylcholine après son internalisation dans la cellule nerveuse.

Source : Arnon et al., 2001.

### 3.3. Types et sous-types de BoNT

Comme évoqué précédemment, plus de 40 sous-types de toxines botuliques ont déjà été caractérisés (Tableau 2) (Peck et al., 2017). La structure de la toxine botulique ainsi que la propriété toxicologique restent les mêmes pour l'ensemble des types toxiques (Peck et al., 2017). Les différentes BoNT se distinguent entre elles de par leur immunogénicité à certains anticorps anti-BoNT. Le sérotype d'une toxine est donc défini en la confrontant à un sérum contenant des anticorps dirigés contre une BoNT d'un certain type. Par exemple, si la toxine est neutralisée en présence d'anticorps dirigés contre la BoNT A, cela confirme que la toxine est du type A. En plus de ces

différences immunologiques, les BoNT vont toucher des hôtes différents et vont cibler des protéines différentes dans la cellule nerveuse.

Tableau 2 : Liste des différentes espèces bactériennes capables de produire la toxine botulique et leurs types et sous-types toxiniques associés. Source : Peck et al., 2017

<b>Espèces</b>	<b>Types toxiniques</b>	<b>Sous-types de BoNT</b>
<i>Clostridium botulinum</i> Group I	A, B, F et X	A1; A2; A3; A4; A5; A6; A7; A8 B1; B2; B3; B5; B6; B7; B8 F1; F2; F3; F4; F5; H-F/A; X
<i>Clostridium botulinum</i> Group II	B, E et F	B4 E1; E2; E3; E6; E7; E8; E9; E10; E11; E12 F6
<i>Clostridium botulinum</i> Group III	C et D	C ; D ; C/D ; D/C
<i>Clostridium argentinense</i> Group IV	G	G
<i>Clostridium butyricum</i>	E	E4 ; E5
<i>Clostridium baratii</i>	F	F7
<i>Enterococcus faecium</i>	En/J	En/J

Les différences entre les types toxiniques vont se situer au niveau de leur composition en acides aminés (Figure 6) (Hill & Smith, 2013; Hill et al., 2015). Le séquençage des différents gènes *bont* suivi de leur comparaison a montré des différences en acides aminées allant de 37,2 à 69,6% (Hill & Smith, 2013). Au sein même des types, il existe des différences importantes permettant de les discriminer en sous-types distincts. Ces différences peuvent être mineures (1 à 7% entre les sous-types du sérotype B) ou majeures (3 à 36% entre les sous-types du sérotype F) entre les toxines d'un même type (Hill et al., 2015). Elles peuvent entraîner des changements de caractéristiques plus ou moins importants selon la région mutée (toxicité, affinité avec les récepteurs cibles ou encore stabilité de la toxine). Depuis 2017, de nouvelles règles de nomenclature ont été mises en place pour la classification des nouveaux sous-types. Les sous-types doivent être déterminés à partir de la séquence en acides aminées. Pour être considéré comme un nouveau sous-type, la composition en acides aminés de la toxine doit différer d'au moins 2,6% par rapport à celles des autres sous-types déjà caractérisés (Peck et al., 2017).

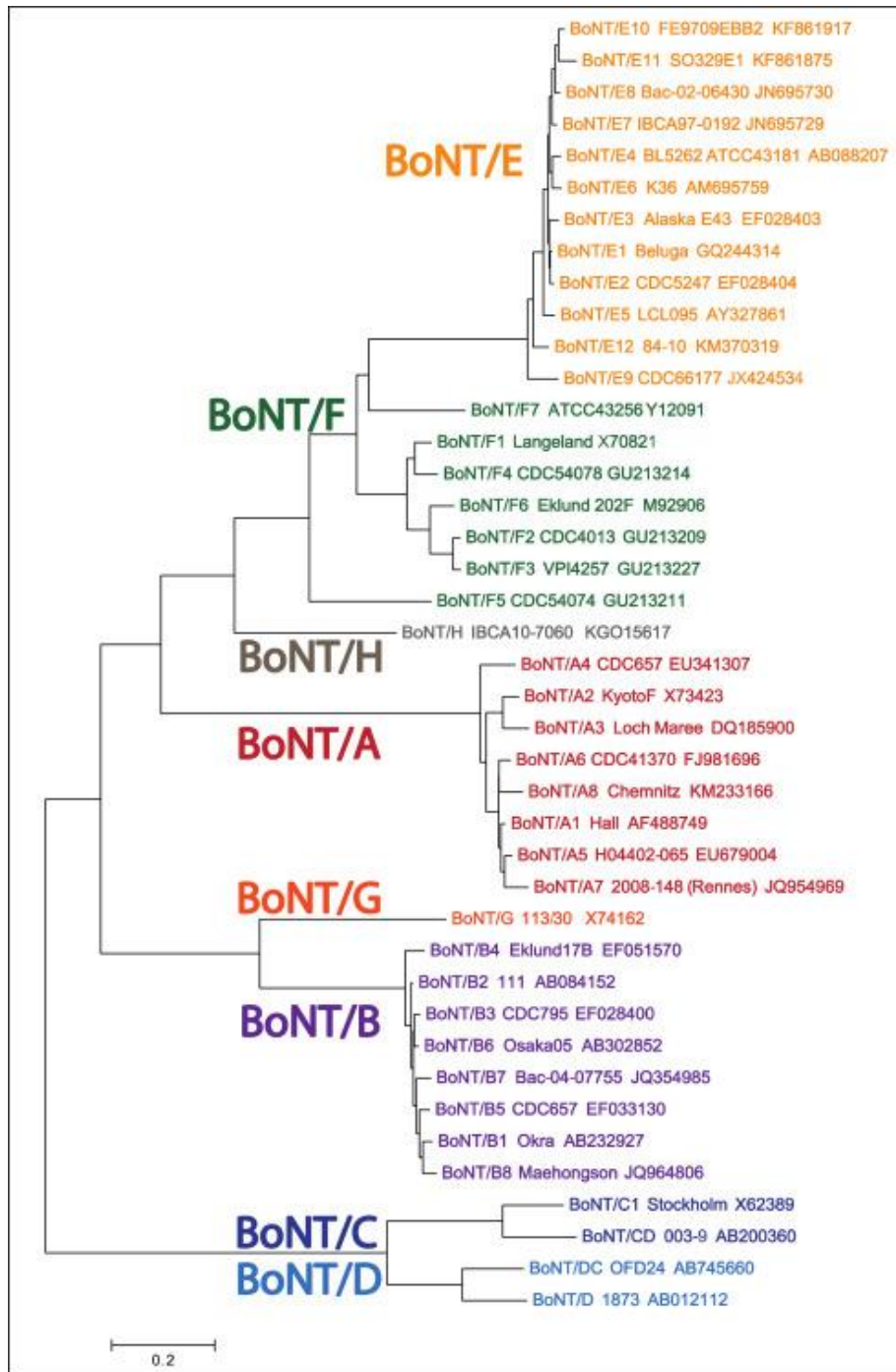


Figure 6 : Dendrogramme représentant l'ensemble des sous-types de BoNT basé sur leur différence au niveau de la composition en acides aminés de la toxine.

Source : Hill et al., 2015.

Le gène codant pour la toxine peut se situer à différents endroits du génome de la bactérie, que ce soit sur le chromosome (BoNTs A, B, E et F) ou sur des éléments génétiques mobiles (EGM) comme des plasmides (BoNTs A, B, E, F, G, En et X) ou des bactériophages (BoNTs C, D et leurs mosaïques C/D et D/C) (Hill et al., 2015; Zhang et al., 2018; Zhang et al., 2017). La présence d'une même toxine chez plusieurs espèces de clostridies différentes (i.e. BoNT E chez *C. botulinum* groupe II et *C. butyricum*) ainsi que leur localisation sur des EGM suggèrent que des événements de transferts horizontaux de gènes sont intervenus au cours de leur évolution (Hill et al., 2015).

Une spécificité de certaines souches de *Clostridium botulinum* est la présence de deux ou plusieurs gènes *bont* dans leur génome (Barash & Arnon, 2014; Gonzalez-Escalona et al., 2014). Les souches contenant deux gènes *bont* sont dites bivalentes et vont produire les deux toxines à des proportions égales ou différentes. Pour les souches où la production est différente, la toxine produite en plus grande quantité sera notée en majuscule tandis que la deuxième, moins abondante, sera notée en minuscule (i. e. Ab, Af, Ba, Bf) (Barash & Arnon, 2014; Giménez, 1984). Il existe aussi des souches bivalentes où une seule toxine est produite, l'autre gène n'étant pas exprimé. C'est notamment le cas de plusieurs souches du groupe I produisant BoNT A tout en possédant un gène *bont b* dit « silencieux » car aucune production de BoNT B n'est détectée dans les conditions testées (Hutson et al., 1996). Ces souches sont annotées A(B) dans ce cas de figure. Enfin, une souche possédant 3 gènes *bont* différents (un gène *bont a* et deux gènes *bont f*) a aussi été caractérisée dans la littérature (Dover et al., 2013).

Les souches de *Clostridium botulinum* du groupe III produisent des toxines botuliques mosaïques de type C/D et D/C en plus des types C et D. Le gène codant pour la toxine de type C/D est composé à deux tiers du gène *bont c* (la chaîne légère et le domaine de translocation H<sub>n</sub>) et à un tiers du gène *bont d* (domaine de liaison H<sub>c</sub>) (Figure 7) (Nakamura et al., 2012). Pour le gène codant pour la toxine D/C, deux tiers du gène sont similaires au gène *bont d* (la chaîne légère et le domaine de translocation H<sub>n</sub>) tandis que le dernier tiers est similaire à *bont c* (domaine de liaison H<sub>c</sub>) (Figure 7) (Moriishi et al., 1996a; Moriishi et al., 1996b; Nakamura et al., 2012). Ces sous-types mosaïques n'ont été décrits pour la première fois qu'en 1996 (Moriishi et al., 1996a).

Les spécificités de ces toxines seront présentées plus en détail dans la partie traitant du botulisme animal (cf partie II).

Une autre toxine mosaïque atypique, initialement nommée BoNT H dans une première étude, a été découverte dans une souche bivalente (IBCA10-7060) produisant majoritairement BoNT B. Une seconde étude a permis de déterminer que la structure de cette toxine présente des régions ayant des



similarités avec les sérotypes A et F. En effet la chaîne légère correspond à celle retrouvée pour BoNT F5 tandis que le domaine de liaison Hc correspond à celui retrouvé pour BoNT A1 (Barash & Arnon, 2014; Maslanka et al., 2016).

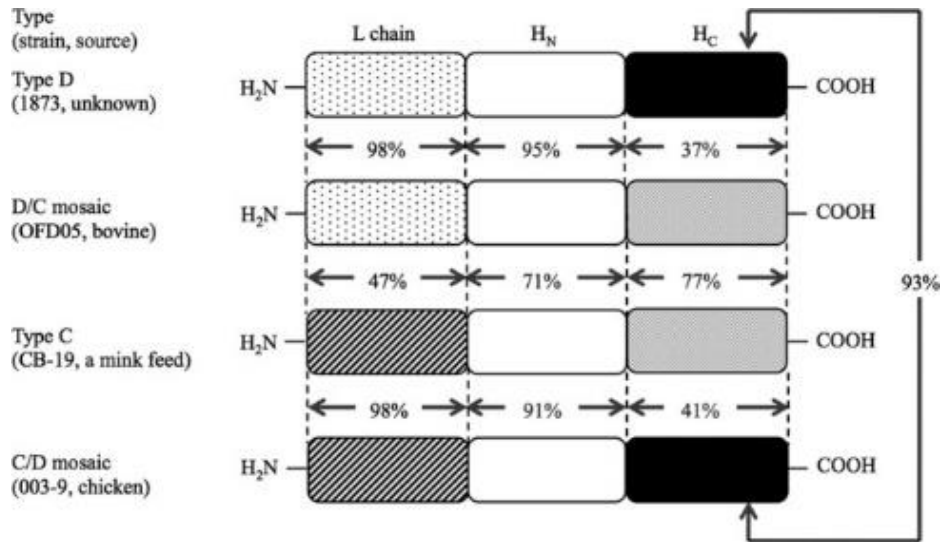


Figure 7 : Représentation schématique de la structure des neurotoxines botuliques produites par les souches de *Clostridium botulinum* du groupe III (C, D, C/D et D/C). Pour chaque région (LC, Hn et Hc), un % d'homologie en acides aminés est noté entre chaque type. Au-dessus de 70 % d'homologie, la région est représentée avec un même remplissage.

Source : Nakamura et al., 2012.

Il existe 7 types de toxines botuliques historiques (A-G), 9 décrites dans la littérature à ce jour (A-G, H, X) et plus de 40 sous-types. Il existe également des toxines mosaïques (C/D, D/C, F/A) et une souche peut être capable de produire plusieurs toxines différentes. Chaque type toxinique va avoir une cible différente au niveau de la synapse, cependant l'activité est identique pour l'ensemble des types toxiques qui vont empêcher la libération de l'acétylcholine dans la fente synaptique.

#### 4. Le botulisme chez l'humain

Le botulisme chez l'Homme est une maladie rare mais très grave, à déclaration obligatoire en Europe, pouvant causer la mort si non traitée dans les temps (Sobel, 2005). Les types toxiques associés au botulisme chez l'Homme sont les neurotoxines de type A, B, E et F. Il existe plusieurs formes de botulisme humain (alimentaire, infantile, ingestion de spores chez des patients présentant des facteurs de risque tels que les dysbioses, par blessure, par inhalation et iatrogène) selon l'origine de contamination. La forme la plus fréquente est le botulisme alimentaire ou l'intoxication alimentaire (Mazuet et al., 2018; Popoff, 2018). Cette forme est liée à la consommation d'un aliment dans lequel *C. botulinum* a pu croître et produire de la neurotoxine botulique. Après ingestion, la toxine va passer dans la circulation sanguine au niveau du jéjunum ou du duodénum. Ce sont souvent des aliments préparés à la maison (conserves de viandes, légumes, charcuteries notamment) qui vont être mis en cause même si des aliments industriels périmés ou mal conservés ont déjà causés des intoxications (King, 2008; Pingeon et al., 2011).

Les cas de botulisme humain C et D sont très rares (Anses, 2021). En effet, les toxines C et D n'ont été associées qu'à 2 foyers confirmés de botulisme humain et 11 suspects entre 1950 et 2006 au niveau mondial (Anses, 2021).

Le botulisme se manifeste chez l'humain par une atteinte des nerfs crâniens, suivie d'une paralysie flasque bilatérale des muscles pouvant induire une insuffisance respiratoire grave et le décès. L'hospitalisation est nécessaire : surveillance (prévention d'une insuffisance respiratoire) puis administration d'une antitoxine heptavalente. Si le diagnostic et le traitement sont mis en place rapidement, la guérison se fait sans séquelles. La létalité est faible.

Le botulisme chez l'Homme est une maladie rare mais très grave. Il est associé principalement aux *C. botulinum* des groupes I et II (toxines A, B, E et F). Les cas de botulisme humain associés aux toxines produites par les souches de *C. botulinum* du groupe III (C, D, C/D ou D/C) sont extrêmement rares.

## II. Le Botulisme animal

Le botulisme touche également les espèces animales sauvages ou domestiques (Tableau 3) entraînant des conséquences à la fois écologiques mais aussi économiques. Les principaux foyers de botulisme animal confirmés en France concernent les espèces aviaires et les bovins (Le Marechal et al., 2016). En Europe, la perte économique moyenne lors d'un épisode de botulisme dans un cheptel avicole de 30 000 oiseaux est de 20 000 euros. Pour un cheptel de 60 bovins, elle s'élève à 90 000 euros (Braun, 2006; Souillard et al., 2014). Dans ce chapitre, les différentes spécificités du botulisme animal pour ces deux groupes seront détaillées.

Tableau 3 : Liste des principaux animaux affectés par le botulisme et les types de BoNT associés pour chaque hôte.

Source : Le Marechal et al., 2016.

Hôte affecté	Types toxiques impliqués
Volaille	C, D, C/D, D/C, E
Oiseaux sauvages	C, D, C/D, D/C, E, A
Bovin	D, D/C, C/D, C, A, B
Poisson	E
Animaux à fourrure	C, C/D, A, E
Cheval	C, D, B, A

### 1. Botulisme aviaire

Le botulisme aviaire est considéré comme la maladie la plus importante en terme de mortalité chez les oiseaux sauvages (Rocke, 2006). Le botulisme concerne à la fois l'avifaune sauvage et les volailles. En avifaune sauvage, le botulisme peut avoir un impact écologique très important avec des cas pouvant atteindre jusqu'à plusieurs dizaines de milliers d'oiseaux morts (Vidal et al., 2013). Cela peut également impacter les plans de sauvegarde des espèces protégées (Anza et al., 2016; Vidal et al., 2013; Work et al., 2010). En élevage, les épisodes de botulisme vont entraîner des pertes économiques majeures pour l'éleveur et la filière (Le Marechal et al., 2016). En Europe, elle est considérée comme une maladie émergente (Skarin et al., 2013).

Le botulisme a été détecté chez 264 espèces aviaires provenant de 39 familles d'oiseaux (Rocke, 2006). Une grande partie de ces 264 espèces sont des oiseaux migrateurs comme les anatidés ou les oiseaux côtiers (Le Marechal et al., 2016; Rocke, 2006). En exploitation avicole, le botulisme va

principalement toucher le poulet de chair, la dinde ou le faisan, même si des foyers ont été recensés en élevage de canards, poules pondeuses, pintades, oies et cailles (Ventujol et al., 2017).

### **1.1. Types toxiques associés et sensibilité**

Les souches responsables du botulisme aviaire sont les souches de *C. botulinum* du groupe III produisant les neurotoxines de type C, D, C/D et D/C. Plus rarement, des épisodes impliquant les neurotoxines E et A ont été rapportés (Yule et al., 2006). En conditions expérimentales, les espèces aviaires sont sensibles à toutes les toxines botuliques lorsqu'elles sont administrées par voie intraveineuse même si cette sensibilité va dépendre du type toxinique injectée et de l'espèce aviaire testée (Tableau 4) (Anses, 2021; Rasetti-Escargueil et al., 2019). Les oiseaux vont être globalement plus sensibles aux toxines de type C et C/D et vont être plus résistants à la toxine de type D. Les poulets, dindes, faisans et paons vont être plus sensibles à la toxine C que les canards (Tableau 4) (Anses, 2021; Rasetti-Escargueil et al., 2019). Des exceptions existent comme les vautours ou autres charognards qui présentent une résistance à la toxine de type C (Kalmbach, 1939).

Tableau 4 : Toxicité des différentes neurotoxines botuliques selon l'espèce aviaire testée. Les valeurs sont exprimées en multiple de souris LD50/kg selon le mode d'inoculation (IV : intraveineuse, VO : Voie orale, IP : intrapéritonéale). Aucune valeur n'a été rapportée pour la BoNT D/C.

Tableau adapté de Rasetti-Escargueil et al, 2019 et Anses, 2021.

<b>BoNT</b> <b>Espèce</b>	A	B	C	C/D	D	E	F
<b>Poulet</b>	10 (IV) 1 million/kg (VO)	20 000 (IV)	16 000 (IV) 26 000 (IV)/animal	5600/animal (IV)	>320 000 (IV)	100 (IV)	640 000 (IV)
<b>Canard</b>		15 000 (IV) >10 millions/animal (VO)	500- 76 000 45- 80 000 (VO) 19 000-320 000/animal (VO) 96 000/animal (IP) 320 000/kg (VO)			2,5 millions/animal (IV) 200 000/animal (VO)	
<b>Paon</b>	170 (IV)	33 000 (IV)	2700 (IV)		>320 000 (IV)	170 (IV)	640 000 (IV)
<b>Faisan</b>	44-170 440 000/kg (VO)	88 000 (IV)	70 (IV)		>320 000 (IV)	440 (IV) 440 000/kg (VO)	640 000 (IV)
<b>Dinde</b>	20 (IV) 200 000/kg (VO)	40 000 (IV)	320 (IV)		>320 000 (IV)	200 (IV) 200 000/kg (VO)	640 000 (IV)

## 1.2. Données de surveillance du botulisme aviaire

L'Anses a publié un rapport en juillet 2021 dans lequel une compilation des données du CNR et du LNR sur la période 2009-2019 a permis d'obtenir une image des épisodes de botulisme animal diagnostiqués en France (Anses, 2021).

Ces données montrent que le botulisme aviaire représentait 77,5% des 592 foyers botuliques recensés chez différentes espèces animales, 41,7% provenant de la filière avicole et 35,8% provenant de l'avifaune sauvage. Que ce soit en élevage ou en faune sauvage, le type toxinique le plus représenté est le type C/D. En faune sauvage, la toxine de type C/D a été identifiée dans la quasi-totalité des cas confirmés entre 2013 et 2019 (Le Marechal et al., 2019). L'ensemble de ces résultats est en concordance avec ce qui est observé en Europe où le botulisme aviaire est essentiellement causé par des souches de *C. botulinum* de type C/D (Anza et al., 2014a). Les types toxiniques D/C et D sont aussi détectés en élevage de dindes. Le botulisme aviaire suit un rythme saisonnier, avec une survenue des cas plus importante lors du troisième trimestre (80% des cas en faune sauvage et 47,3% des foyers en élevage). Cette augmentation peut s'expliquer par une augmentation de la température, favorable à la croissance de *C. botulinum* du groupe III et à la production de toxine (Anses, 2021).

Il est à noter qu'avant 2010, le diagnostic biologique se faisait principalement avec la méthode de séroneutralisation sur souris. Cette méthode ne permettant pas de distinguer les formes mosaïques des non-mosaïques, les foyers de botulisme impliquant les types D et C détectés avant 2010 sont probablement surévalués. Depuis la mise en place des techniques de détection par PCR (Woudstra et al., 2012), les formes mosaïques sont devenues majoritaires.

Entre 2007 et 2008, une recrudescence des épisodes de botulisme dans les élevages de volailles a été observée dans plusieurs pays d'Europe (Suède, France, Norvège, Danemark), pour des raisons qui sont encore à ce jour inexplicables (Skarin et al., 2013). En France, en 2007, 121 épisodes ont ainsi été recensés en élevage d'après les données du réseau national d'observations épidémiologiques en aviculture (RNOEA) contre 20 en moyenne les années précédentes d'après le rapport de l'AFSSA en 2002. Le nombre de foyers est redescendu à une vingtaine en moyenne les années suivantes selon les données des CNR et LNR (Le Bouquin et al., 2017). Cette recrudescence n'a pas été observée en faune sauvage avec une moyenne de 14 cas détectés chaque année entre 2005 et 2019 (Anses, 2021).

Les espèces aviaires sont sensibles à la toxine E. Des épisodes de botulisme de type E ont été recensés en France à la fin des années 90 dans des élevages avicoles avec 5 épisodes chez des poulets de chair en France entre 1997 et 2000, le dernier épisode confirmé remontant à 2002 selon le rapport de l'AFSSA sur le botulisme animal en 2002. L'apparition d'épisodes impliquant ce type toxinique à la fin des années 90 a suscité des inquiétudes du fait de la sensibilité de l'Homme à ce type toxinique.

La toxine est ainsi systématiquement recherchée lors d'une suspicion de botulisme aviaire. Le gène codant la toxine E a été détecté dans 3 cas en faune sauvage en 2018 de manière concomitante avec le type C/D. Ce type toxinique est retrouvé dans les régions septentrionales (Nord de l'Europe et de l'Amérique) lors de cas de botulisme impliquant des oiseaux sauvages piscivores (Lan Chun et al., 2015). Des épisodes de type E sont recensés annuellement autour des Grands Lacs aux Etats-Unis et au Canada (Chipault et al., 2015).

### **1.3. Signes cliniques**

Les signes cliniques du botulisme chez les espèces aviaires se manifestent par une paralysie flasque ascendante qui commence par les pattes puis progresse vers les ailes, le cou et enfin les paupières. Cette paralysie se traduit par des difficultés à se déplacer et un comportement léthargique des oiseaux avec les yeux fermés. La paralysie du cou va se traduire par une difficulté à relever la tête, les ailes sont pendantes. Des diarrhées peuvent aussi être observées. Ces signes cliniques sont associés à des problèmes respiratoires qui constituent une des principales causes de mortalité en élevage. En faune sauvage, la noyade est une des causes principales de mort chez les oiseaux atteints de botulisme du fait de l'incapacité des oiseaux à relever la tête lorsqu'ils sont sur l'eau (Le Marechal et al., 2016).

### **1.4. Botulisme aviaire en élevage**

#### **1.4.1. Pathogénèse**

Il est admis que la production de la toxine botulique a lieu *in vivo* chez les poulets (Popoff, 1989). En effet la quantité de toxines nécessaire à l'apparition des symptômes est telle qu'il n'est pas possible d'en retrouver autant dans l'environnement des animaux. Des essais *in vivo* ont permis de démontrer que les signes cliniques de botulisme sont observés chez les volailles après ingestion de *C. botulinum* et production de la toxine botulique au niveau des caeca (Miyazaki & Sakaguchi, 1978). Il a également été démontré que les animaux ne présentent des signes cliniques que s'ils ont accès aux fientes, suggérant que les animaux ingèrent de nouveau la toxine et *C. botulinum* via la litière par coprophagie (Hyun & Sakaguchi, 1989). Un schéma présentant la pathogénèse est présenté dans la figure 8 (Anses, 2021).

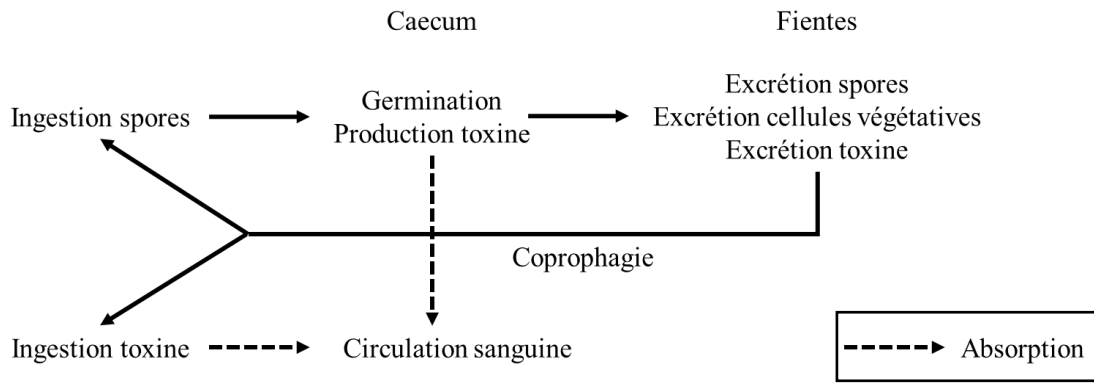


Figure 8 : Pathogénèse du botulisme chez les volailles.

Source : Adaptée du rapport d'expertise collective sur le botulisme (Anses, 2021)

#### 1.4.2. Véhicules et sources de contamination

Les véhicules et les sources de contamination à l'origine des épisodes de botulisme en élevages avicoles sont très rarement identifiés. Dans la littérature scientifique, il a été démontré que la litière (balles de riz) constitue un véhicule de contamination (Bano et al., 2013; Fillo et al., 2021). La comparaison génomique des souches isolées de poulets présentant des signes cliniques de botulisme et d'une souche isolée d'une balle de riz stockée dans l'élevage mais non encore utilisée a montré une clonalité entre les souches. Les balles de riz ont été identifiées comme positives pour *C. botulinum* de type C/D à deux reprises dans deux élevages différents. Les auteurs de cette étude recommandent l'analyse des balles de riz avant livraison auprès des éleveurs pour garantir l'absence de contamination (Fillo et al., 2021). L'origine de la contamination de la balle de riz n'a pas été investiguée dans ces études.

L'aliment est considéré comme le véhicule de contamination la plus probable même si aucun élément ne permet d'étayer cette hypothèse dans la littérature. L'eau d'abreuvement constitue également un véhicule potentiel (Anses, 2021).

Les trois véhicules de contamination que sont l'aliment, la litière et l'eau peuvent être contaminés soit par un cadavre lors de leur récolte ou de leur stockage ou au cours de l'élevage ; soit par un transfert de contamination via la manipulation par du matériel contaminé ou par des fèces d'animaux sauvages.

#### 1.4.3. Mesures préventives et traitement

En élevage de volaille, une attention particulière devra être portée au niveau du stockage et de la manipulation des aliments, de la litière et de l'eau d'abreuvement utilisés dans l'exploitation. Des



mesures de biosécurité doivent être mises en place au niveau du sas sanitaire. Les cadavres d'animaux doivent être ramassés régulièrement. Le matériel et les différents bâtiments de l'exploitation doivent faire l'objet d'opérations de nettoyage et de désinfection fréquentes. De plus, les opérations de désinsectisation et dératissage sont essentielles (Le Marechal et al., 2012). D'autres mesures peuvent être mises en place en cas de survenue d'un épisode comme la création d'une « infirmerie » dans le bâtiment pour séparer les animaux sains et les animaux atteints, l'élimination des animaux malades et un ramassage plus fréquent des cadavres.

Dans le cas où un épisode de botulisme est confirmé dans un élevage, un traitement aux  $\beta$ -lactamines peut être administré par voie orale à l'ensemble des animaux de l'élevage. Des vaccins existent mais ils sont rarement utilisés que ce soit en élevage ou dans la faune sauvage (Anniballi et al., 2013). Il n'y a pas d'autorisation de mise sur le marché (AMM) en France.

## **1.6. Botulisme aviaire en faune sauvage**

### **1.6.1. Pathogénèse**

La pathogénèse du botulisme en faune sauvage est encore méconnue à ce jour. Deux cycles d'infection sont décrits dans la littérature. Les oiseaux peuvent soit s'intoxiquer en ingérant directement la neurotoxine (Holmes, 2019), soit s'infecter en ingérant *C. botulinum*, la production de toxine se déroulant ensuite *in vivo* (Rocke, 2006). Bien que l'importance de ces deux types de cycles ne soit pas encore connue, il est probable qu'ils coexistent (Skarin et al., 2015).

### **1.6.2. Réservoirs de *C. botulinum***

Les cas de botulisme en faune sauvage vont être initiés par la présence de *C. botulinum* dans l'environnement (sol, eau, sédiments marins et d'eau douce) ou dans les organismes présents comme les algues, plantes et invertébrés mais aussi sur certains vertébrés comme les poissons et les oiseaux (Espelund & Klaveness, 2014; Reed & Rocke, 1992). Ces réservoirs biotiques peuvent devenir toxiques lorsqu'ils rentrent en décomposition, devenant des substrats propices à la germination, croissance et production de toxine par *C. botulinum* (Reed & Rocke, 1992). Les oiseaux vont s'infecter/s'intoxiquer en se nourrissant de plantes, algues et invertébrés (planctons, asticots) où des cellules végétatives, spores et neurotoxines se sont accumulées (Chun et al., 2013). Les oiseaux piscivores vont aussi pouvoir se contaminer en se nourrissant de poissons contaminés par *C. botulinum*, notamment les souches produisant la neurotoxine de type E (Yule et al., 2006).

Un scénario de contamination secondaire décrit dans la littérature mettant en jeu des invertébrés est appelé cycle cadavre-asticot (carcass-maggot cycle). Ce scénario a lieu lorsque les oiseaux

consomment des asticots présents sur des cadavres d'autres oiseaux contaminés. Quand un oiseau porteur de *C. botulinum* meurt, le cadavre devient un environnement favorable à la germination des spores et la production de BoNT. Les mouches nécrophages vont être attirées par le cadavre et vont y pondre des œufs. Les asticots vont alors se développer et, en se nourrissant, vont héberger une quantité importante de toxine et/ou de bactéries (généralement de type C ou C/D). Les autres oiseaux vont ensuite s'intoxiquer/s'infecter en consommant les asticots contaminés (Figure 9) (Espelund & Klaveness, 2014; Friend et al., 1999; Rocke, 2006).

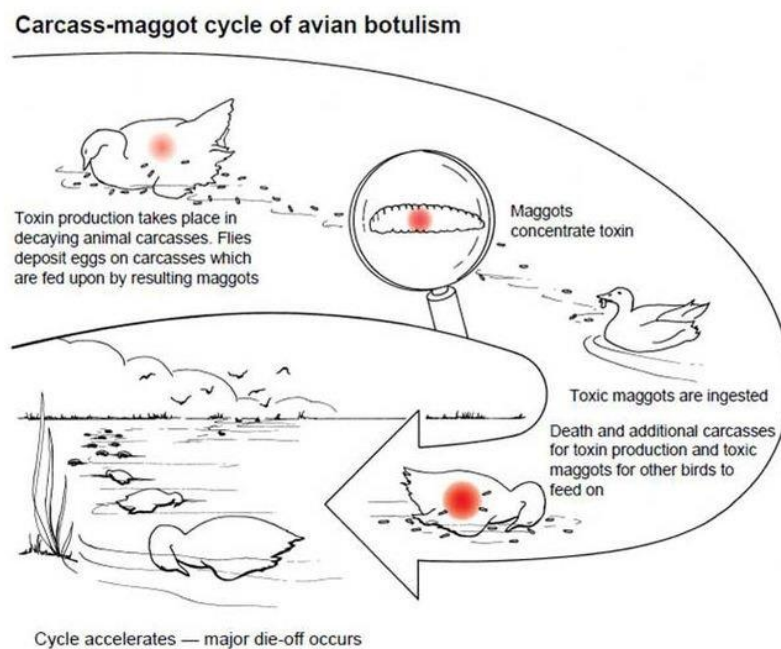


Figure 9 : Cycle de contamination « carcass-maggot » des oiseaux sauvages aquatiques.

Source : Friend et al., 1999

Le cycle épidémiologique de *C. botulinum* dans un point d'eau met en jeu beaucoup d'acteurs qui ont tous une importance (Figure 10) (Espelund & Klaveness, 2014). L'hypothèse la plus probable de l'apparition d'un foyer de botulisme pourrait être la mort d'animaux porteurs de la bactérie, mort du botulisme ou d'une autre raison. La possibilité de l'apparition d'un foyer via l'ingestion par les oiseaux d'algues ou de plantes contaminées reste cependant encore à explorer (Espelund & Klaveness, 2014).

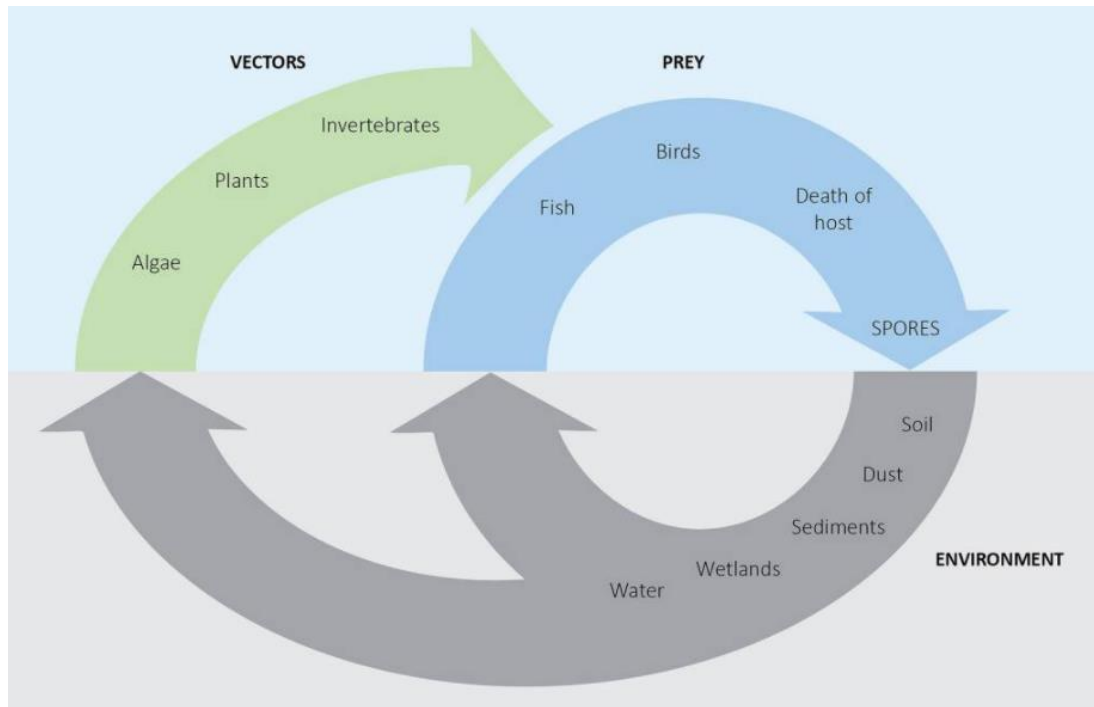


Figure 10 : Représentation schématique du cycle épidémiologique de *C. botulinum* dans l'écosystème d'eau douce.

Source : Espelund et al., 2014

L'augmentation de la température et la baisse du niveau de l'eau, notamment dans les lacs et les étangs va entraîner une baisse du niveau de l'eau par assèchement mais aussi une augmentation de la température et du pH de l'eau ainsi qu'une diminution de son potentiel redox (Friend et al., 1999). Couplé à la présence de matières organiques en décomposition (cadavres d'animaux, accumulation d'algues et plantes) servant de substrats à la bactérie, les conditions vont être propices à la germination, prolifération et production de toxine par *C. botulinum* (Figure 11). Les contaminations vont donc surtout avoir lieu lors des périodes chaudes (fin du printemps, été, début de l'automne) car les conditions sont plus propices à la croissance de la bactérie (Ventujol et al., 2017).

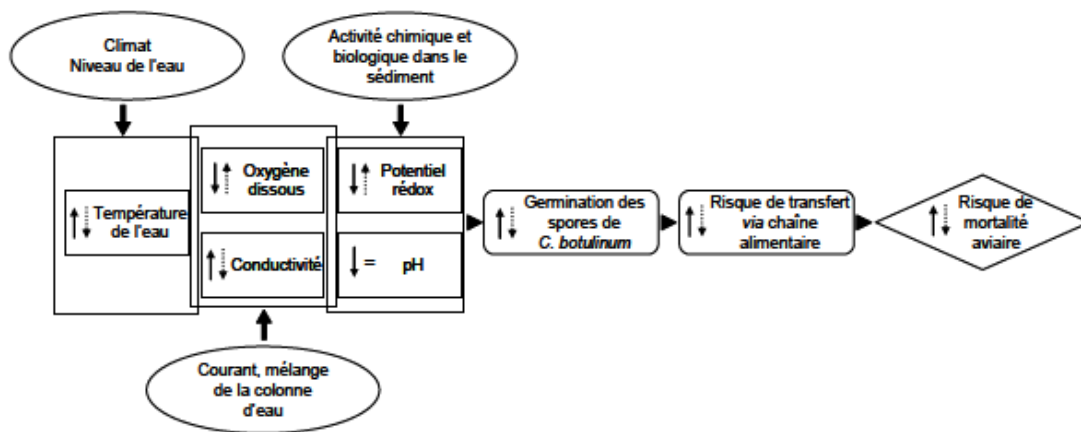


Figure 11 : Influence des conditions environnementales sur les pics de mortalité lors d'un évènement de botulisme aviaire.

Source : Le Marechal 2015 Lettre SAGIR N°182

### 1.6.3. Mesures préventives et traitement

La présence ubiquitaire de *C. botulinum* dans l'environnement rend la prévention des foyers de botulisme aviaire dans la faune sauvage compliquée. Des mesures peuvent être mises en place telles que l'enlèvement rapide des cadavres d'oiseaux morts afin de prévenir la prolifération d'asticots contaminés et donc de limiter le nombre d'oiseaux affectés lors de l'apparition d'un foyer de botulisme (Anza et al., 2014b; Evelsizer et al., 2010). Une autre mesure pour limiter les sources de contaminations est d'enlever les algues ou plantes en décomposition puisque ce sont des zones où l'oxygène est peu présent et sont donc propices à la prolifération de la bactérie (Espelund & Klaveness, 2014)

Il est possible de soigner les oiseaux atteints de botulisme en les nourrissant et en les abreuvant dans un centre de soin par exemple. Une très bonne récupération peut être observée en particulier chez les oiseaux aquatiques. Au contraire, une efficacité limitée est rapportée pour les limicoles, goélands, grèbes et foulques (Locke & Friend, 1989). L'un des enjeux est également de prévenir les contaminations croisées entre la faune sauvage et les élevages, même si aucune donnée n'est actuellement disponible à ce sujet dans la littérature.

## 1.7. Portage sain

Une des hypothèses pour expliquer l'initiation d'un épisode de botulisme aviaire est le portage asymptomatique de *C. botulinum* par les oiseaux et le développement de la bactérie suite à la mort de l'animal (Dohms, 2008). Cependant, peu d'études ont exploré cette hypothèse. Une première étude publiée en 1974 où des investigations ont été menées sur 100 litières provenant d'élevage n'ayant pas connu d'épisode de botulisme n'a pas permis la détection de *C. botulinum* (Roberts, 1974). En Norvège, le même résultat a été obtenu lors de l'investigation de 100 lots de poulets de chair provenant d'élevages ayant connu ou non un épisode de botulisme (Hardy & Kaldhusdal, 2013).

D'autres études ont au contraire permis de détecter la bactérie sur des animaux sains. Une étude de l'Anses dont le but était d'évaluer la prévalence de *C. botulinum* dans des élevages de poulets de chair a permis la détection de la bactérie dans 1 seul élevage, uniquement sur des prélèvements d'environnement, les prélèvements d'organes étant négatifs (Le Marechal et al., 2012). Une autre étude de l'Anses menée par le LNR sur 5 élevages sans historiques de botulisme avait permis la détection de *C. botulinum* sur les prélèvements d'organes d'une dinde et sur le prélèvement de fumier de l'élevage associé (Données LNR non publiées). Enfin, une étude menée en Suède a permis de détecter *C. botulinum* de type C/D dans 11% des élevages sains (Blomqvist et al., 2009).

Ces résultats discordants dans la littérature pourraient s'expliquer par la faible présence de la bactérie, qui serait alors en dessous des seuils de détection des méthodes (Le Bouquin et al., 2017; Popoff, 1989)

Au LNR, des enquêtes épidémiologiques menées lors d'épisodes de botulisme bovin ont permis de mettre en évidence du portage sain de *C. botulinum* de type D/C dans les ateliers de volaille de l'élevage ou situés à proximité de celui-ci. Cependant, l'absence d'outils de typage rend actuellement impossible la validation des hypothèses de contamination croisée liée à un portage sain de volaille (Souillard et al., 2017).

Dans la faune sauvage, une étude a été effectuée en Autriche sur des fèces collectés dans des zones où des foyers de botulisme ont été recensés et dans des zones sans historique de botulisme. Il a été mis en évidence la présence de *C. botulinum* dans 83% des prélèvements provenant des zones avec un historique et 80% pour les autres zones (Zechmeister et al., 2005). A contrario, une étude récente sur la prévalence de *C. botulinum* chez la sarcelle et le canard colvert a été menée sur 4 sites en France en 2021. Sur ses 4 sites, 2 ont un historique de botulisme. Au total, 200 oiseaux ont été étudiés (171 oiseaux prélevés avec écouvillon cloacaux et 29 oiseaux chassés avec détection dans différents organes). Concernant les 171 oiseaux écouvillonnés, *C. botulinum* a été détecté sur

seulement 1 oiseau provenant d'un site avec historique de botulisme et sur aucun prélèvement provenant des sites sans historique. Pour les 29 oiseaux chassés dans un site avec historique, *C. botulinum* a été retrouvé en particulier dans le caecum. Cette étude n'a pas permis de réfuter l'hypothèse de l'absence de portage sain de *C. botulinum* chez les oiseaux sauvages en dehors des épisodes de botulisme. Les résultats sont cependant très différents de ceux observés dans l'étude autrichienne et suggère l'absence d'un portage sain chez des oiseaux présents sur des sites sans historique de botulisme (Meloni et al., 2021).

## **2. Botulisme chez les bovins**

### **2.1. Types toxiques associés et sensibilité**

Les épisodes de botulisme dans les élevages bovins sont généralement associés aux toxines de type C, D et leurs mosaïques C/D et D/C produites par les souches de *C. botulinum* du groupe III (Heider et al., 2001; Le Marechal et al., 2019; Nakamura et al., 2010). Plus rarement, des épisodes de type A et B ont aussi été rapportés (Prévot et al., 1950; Kelch et al., 2000; Schocken-Iturrino et al., 1990).

### **2.2. Données de surveillance du botulisme chez le bovin**

D'après la compilation des données des foyers de botulisme animal en France publiée dans le rapport de l'Anses en 2021, 20,3 % des foyers botuliques recensés chez les animaux sur la période 2009-2019 concernent des bovins avec une dizaine de foyers recensés chaque année (Données LNR/CNR). Actuellement, les épisodes de botulisme bovin détectés par le LNR sont uniquement associés au type toxinique D/C (88 % des épisodes entre janvier 2017 et janvier 2020) suivi du type C (12 %). Comme le botulisme aviaire, avant 2010, les souches de type D/C étaient sous-évaluées car le diagnostic du botulisme animal se faisait avec la méthode de séroneutralisation sur sérum.

### **2.3. Pathogénèse**

La pathogénèse du botulisme chez le bovin est peu connue. L'ingestion de la toxine botulique préformée est considérée comme la voie de contamination principale, cependant d'autres voies de contamination existent comme l'ingestion de spores de *C. botulinum* ou la contamination via une plaie profonde. En élevage, les bovins vont se contaminer lorsqu'ils vont s'abreuver ou se nourrir avec de l'eau ou des aliments préalablement contaminés (Doutre, 1969; Hogg et al., 2008; Le Marechal et al., 2019). La part et le rôle réels de chaque voie de contamination (ingestion de toxines, ingestion de toxines préformées, de formes végétatives et spores ou ingestion de la bactérie seule) reste à explorer.

## 2.4. Sources de contamination

Les deux principales sources de contamination dans le cas d'épisodes de botulisme bovin rapportées dans la littérature sont :

Tout d'abord, la contamination via du fumier de volailles symptomatiques ou porteuses saines de *C. botulinum*. La contamination peut avoir lieu lors de l'épandage du fumier de volailles ou par exemple lors de la manipulation de l'aliment avec du matériel contaminé qui a aussi servi à la manipulation du fumier, notamment dans les exploitations qui élèvent à la fois des volailles et des bovins (Relun et al., 2017; Souillard et al., 2017). La présence de cadavres dans l'aliment, l'eau d'abreuvement, pâtures, fourrages, ensilage ou enrubannage où *C. botulinum* va pouvoir croître et produire de la toxine botulique est la deuxième source la plus rapportée (Galey et al., 2000; Le Marechal et al., 2019; Myllykoski et al., 2009; Wilson et al., 1995).

D'autres sources de contamination moins fréquentes existent. Des épisodes de bovins datant des années 80 et 90 font état de litières de volailles contaminées identifiées comme sources de contamination en tant qu'aliment aux bovins (Appleyard & Mollison, 1985; Hogg et al., 1990; Jean et al., 1995; Neill et al., 1989; Smart et al., 1987). Cette source est cependant devenue beaucoup plus rare aujourd'hui.

## 2.5. Signes cliniques

L'apparition des signes cliniques peut survenir de quelques heures après l'ingestion de la toxine et/ou de la bactérie, à quelques semaines après l'exposition. Plusieurs formes vont exister selon la sensibilité de l'animal et la dose de toxine présente dans l'organisme (Le Marechal et al., 2016) :

- La forme suraigüe est la forme la plus sévère de la maladie puisque l'animal va rapidement présenter une paralysie flasque et un décubitus latéral. La mort survient seulement quelques heures après l'apparition des premiers signes cliniques.
- La forme aigüe va durer plusieurs jours et va débiter par des problèmes digestives (constipation, colique), des déplacements difficiles, baisse de la tonicité, une production de lait moindre pour les vaches laitières, difficultés pour se nourrir. La progression de la maladie vers la tête va s'accompagner de nouveaux signes cliniques comme des plus grandes difficultés à se nourrir (problème au niveau de la mastication et de la déglutition), paralysie de la langue qui va pendre, forte salivation, mydriase. Enfin, la paralysie va se généraliser au niveau abdominal et des pattes avec des difficultés

respiratoires et un décubitus sterno-abdominal. Le bovin meurt soit par asphyxie liée aux problèmes respiratoires ou suite à une mauvaise déglutition.

- La forme subaigüe est plus lente et les symptômes sont similaires à la forme aigüe. Le bovin peut soit guérir en quelques semaines, soit mourir quelques jours après l'apparition des signes cliniques selon l'évolution de ceux-ci.

La survenue d'un épisode de botulisme dans un élevage de bovin se traduit par un taux de mortalité généralement très élevé avec une distribution biphasique. En effet, une partie des bovins de l'élevage présente une forme suraigüe ou aigüe de la maladie avec des pertes élevées dans l'élevage dans les jours qui suivent l'apparition des premiers symptômes. Après 2 ou 3 semaines, une réapparition des signes cliniques est observée sur les bovins qui ont survécu à la première vague (Le Marechal et al., 2019).

## **2.6. Portage sain**

Les bovins peuvent, comme les espèces aviaires, être porteurs asymptomatiques de *C. botulinum*. Ici, peu d'études ont été menées pour déterminer la prévalence de portage sain de *C. botulinum* par les bovins. Elles font état de la présence de *C. botulinum* de type A, B, D, E et F dans le rumen, l'intestin et les fèces de bovins ne présentant pas de signes cliniques (Dahlenborg et al., 2003; Fohler et al., 2016; Seyboldt et al., 2015). Une autre étude a été menée avec comme objectif l'évaluation de l'occurrence de *C. botulinum* C, D, C/D, D/C et E chez des vaches saines élevées dans des exploitations proches de zones avec des historiques de botulisme chez des volailles. Les résultats ont montré que *C. botulinum* type C/D a été retrouvé dans des fèces de vaches saines d'une exploitation élevant volailles et bovins peu de temps après la survenue d'un épisode de botulisme dans l'atelier avicole. Les bovins ont probablement été contaminés indirectement au cours de l'épisode et sont devenus porteurs intermittents de la bactérie (Souillard et al., 2015).

Dans un élevage qui a été affecté par un épisode de botulisme, une étude de l'Anses a rapporté la présence de *C. botulinum* de type C un mois après l'épisode sur des prélèvements de bovins qui se remettaient de la maladie (Le Marechal et al., 2019).

## **2.7. Mesures préventives et traitement.**

Comme pour les volailles, des mesures préventives peuvent être mises en place dans les élevages de bovins afin de prévenir les contaminations. Une stricte application de la biosécurité est indispensable à la prévention d'un épisode. Les conditions de stockage des aliments ou l'entretien des canalisations d'eau ou traitement de l'eau de boisson doivent garantir l'absence de contamination



par la faune sauvage. Une bonne gestion des cadavres des animaux est nécessaire pour éviter la croissance de *C. botulinum* et la production de toxine. Le nettoyage et la désinfection des bâtiments d'élevage et du matériel utilisé pour la manipulation de l'aliment et de l'eau doivent être faits méticuleusement et fréquemment. Pour les exploitations mixtes avec des ateliers de volailles et de bovins, des mesures de biosécurité supplémentaires sont nécessaires pour prévenir des contaminations croisées, avec une vigilance particulière concernant le matériel.

Comme pour les espèces aviaires, des vaccins existent mais aucun ne dispose d'une AMM en France. Seul l'ULTRAVAC dispose d'une ATU permettant aux vétérinaires sanitaires de procéder à la vaccination d'urgence en cas de forte suspicion de botulisme dans un élevage bovin.

### **3. Réglementation concernant le botulisme animal**

Le botulisme animal a été règlementé en France à partir de 2006 en tant que maladie à déclaration obligatoire (MDO) pour les animaux sauvages et les bovins et maladie réputée contagieuse (MRC) pour les volailles (décrets n° 2006-179 et n° 2006-178 du 17 février 2006). En 2013, un nouveau texte est rentré en application et le botulisme animal a été classé comme danger sanitaire de catégorie 1 pour toutes les espèces sensibles (Arrêté du 29 juillet 2013 relatif à la définition des dangers sanitaires de première et deuxième catégorie pour les espèces animales).

Une nouvelle réglementation est entrée en application en 2021. Le botulisme animal ne fait pas partie de la liste des maladies relevant de l'*Animal Health Law* (AHL-règlement UE 2016/429 du parlement et du conseil) qui est entrée en application en avril 2021 et définit les mesures de surveillance, prévention et lutte. Cependant, chaque état peut définir une liste des maladies d'intérêt national pour lesquelles des mesures pourraient être mises en place. Le devenir réglementaire du botulisme animal en France n'est pas encore finalisé.

Chaque année, environ 60 foyers de botulisme animal sont confirmés en France, principalement chez les espèces aviaires (filière avicole et avifaune sauvage) et chez les bovins. Des variations sont observées certaines années sans que les facteurs à l'origine de ces variations n'aient été identifiés pour le moment. Une saisonnalité du botulisme aviaire a été observée, notamment chez l'avifaune sauvage, avec une augmentation significative du nombre de foyers botuliques au cours du troisième trimestre.

Les souches de *C. botulinum* responsables du botulisme animal en France ces 10 dernières années sont pour la quasi-totalité des souches du groupe III, espèce bactérienne faisant partie du groupe *C. novyi sensu lato*. Les types toxiques les plus retrouvés sont les formes mosaïques, avec une majorité de foyers de botulisme aviaire associée au type C/D et une majorité de foyers de botulisme bovin associée au type D/C.

A ce jour, la pathogénèse du botulisme animal est encore méconnue, notamment chez les bovins. Les sources de contamination à l'origine des épisodes peuvent être suspectées lors des enquêtes épidémiologiques ou via la détection de *C. botulinum* dans les échantillons. Pour mieux comprendre l'épidémiologie du botulisme animal en particulier et l'identification des sources de contamination, il est nécessaire de pouvoir caractériser les souches impliquées dans les épisodes afin de tracer la contamination.

La prochaine partie de l'introduction sera consacrée à la présentation du groupe d'espèces *Clostridium novyi sensu lato* dont fait partie *C. botulinum* du groupe III et des différents outils de typage qui ont déjà été développés avant le démarrage de mes travaux de thèse.

### III. *Clostridium novyi sensu lato*

Dans les années 80, une proximité génétique entre *C. novyi*, *C. haemolyticum* et *C. botulinum* de type C a été mise en évidence via des comparaisons des homologies ADN-ADN (Nakamura et al., 1983). Une étude basée sur le séquençage du 16S rDNA a ensuite montré que les souches de *C. botulinum* du groupe III sont plus étroitement liées phylogénétiquement à *C. novyi* et *C. haemolyticum* qu'aux souches de *C. botulinum* des groupes I et II (Sasaki et al., 2001).

Ces résultats ont été confirmés par le séquençage complet de génomes de *C. botulinum* du groupe III, *C. novyi* et *C. haemolyticum* (Skarin et al., 2011). Ces 3 espèces étant génétiquement liées, il a été proposé de les regrouper sous le terme collectif de *C. novyi sensu lato* (Skarin et al., 2011). Sur la base de l'analyse comparative des génomes de souches de ces 3 espèces (Figure 12), différentes lignées ont été définies (Ia, Ib, II, III, IV) (Skarin & Segerman, 2014). La diversité génétique de ces souches semble liée à leurs EGM. En effet, les plasmides et les gènes codant pour les toxines peuvent être détectés dans les différentes lignées.

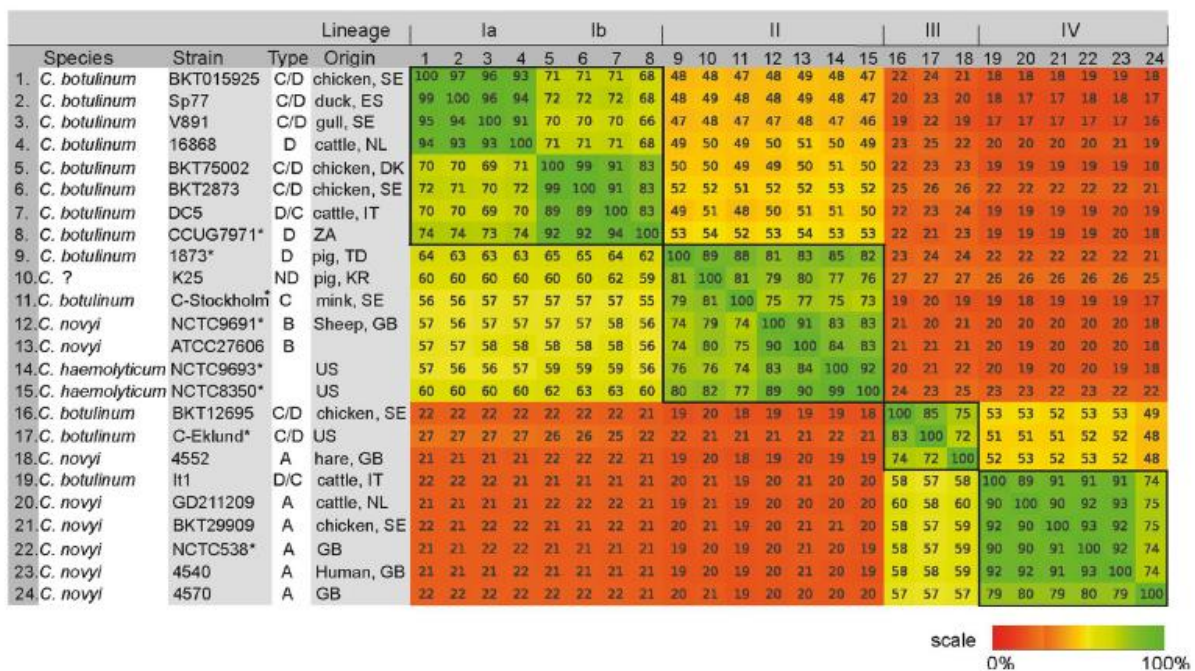


Figure 12 : Heat-map présentant la similitude moyenne des génomes complets de *C. novyi sensu lato* et leur regroupement par lignée en utilisant le logiciel Gegendes.

Source : Skarin et al., 2014

Ces trois espèces bactériennes, malgré leur proximité génétique, causent trois maladies différentes :

- Le botulisme animal pour *C. botulinum* du groupe III (cf. chapitre sur le botulisme animal).
- La gangrène gazeuse chez l'homme et l'animal et l'hépatite infectieuse nécrosante chez le mouton pour *C. novyi* (Nakamura et al., 1983).
- L'hémoglobininurie bacillaire chez les ruminants pour *C. haemolyticum* (Nakamura et al., 1983).

Compte tenu de ce contexte, l'ensemble du projet de thèse portera sur l'étude du groupe d'espèces *C. novyi sensu lato*. Dans ce chapitre, les différentes caractéristiques propres à ce groupe d'espèces seront détaillées, en commençant par l'aspect physiologique puis l'aspect génomique. Enfin les différents outils utilisés pour typer les souches de *C. novyi sensu lato* seront présentés dans une dernière partie.

## **1. Physiologie des souches de *C. novyi sensu lato*.**

### **1.1. Caractéristiques microbiologiques**

*C. novyi sensu lato* se présente sous la forme d'un bacille à coloration de Gram positif d'une taille allant de 1,6 à 22 µm de long et de 0,5 à 2 µm de large (Cato et al., 1986). Ces bactéries ont la capacité de sporuler ce qui va leur conférer une forte résistance dans l'environnement. Les spores formées ont une morphologie ovalaire et vont se former dans la partie terminale ou subterminale de la bactérie en la déformant. La présence de flagelle tout autour de la membrane plasmique lui confère une motilité avec une ciliature péritriche.

### **1.2. Caractéristiques macroscopiques**

Sur milieu gélosé, les colonies de *C. novyi sensu lato* peuvent se présenter sous plusieurs formes. Elles sont généralement de tailles moyennes (entre 3 et 5 mm) et peuvent être rondes ou irrégulières. Les souches de *C. novyi sensu lato* produisent la lipase, ce qui va donner un aspect iridescent à la surface des colonies à la lumière (Figure 13). Elles produisent généralement la lécithinase, qui va induire l'apparition d'une zone opaque autour de la colonie sur des géloses à base d'œuf (Milieu Egg Yeast Agar (EYA) par exemple). L'activité lécithinase est variable selon les souches et il a été rapporté que certaines souches ne présentaient pas d'activité lécithinase (Oguma et al., 1986). Enfin, elles possèdent un pouvoir hémolytique, visible sur les géloses à base de sang, et sont capables d'hydrolyser la gélatine (Oguma et al., 1986).

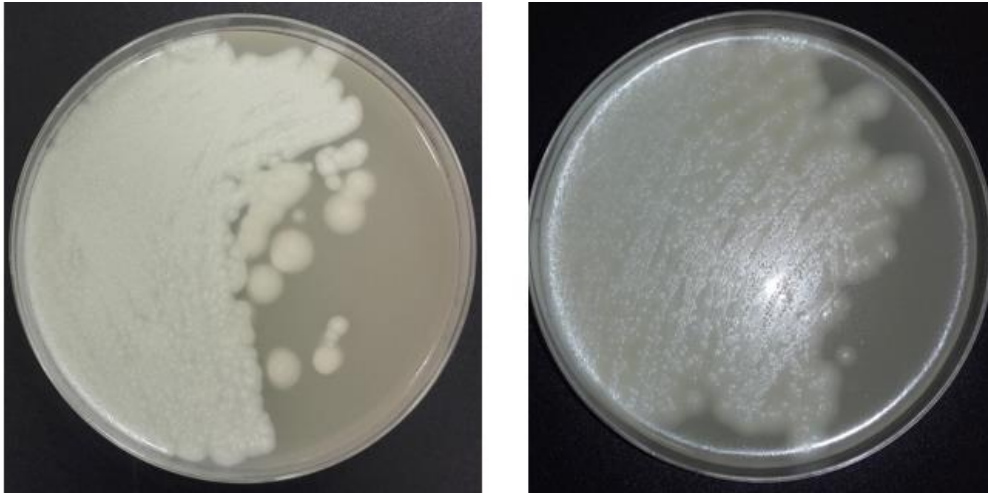


Figure 13 : Géloses EYA ensemencées avec des souches de *C. botulinum* du groupe III. L'utilisation d'une lumière sur la gélose de droite permet de faire ressortir l'aspect iridescent à la surface des colonies.

Source : Images prises lors de la thèse

### 1.3. Métabolisme, sporulation et toxinogénèse

Le type respiratoire des souches de *C. novyi sensu lato* est anaérobie stricte. Les souches ont un métabolisme chimioorganotrophe fermentaire, elles vont puiser de l'énergie en oxydant des composés organiques comme des glucides et des acides aminés. Elles ont toutes la capacité de fermenter le glucose et le ribose. Selon les souches, elles peuvent aussi fermenter d'autres sucres (Fructose, galactose, mélibiose, maltose, mannose), produire de l'hydrogène sulfuré H<sub>2</sub>S et de l'indole (Oguma et al., 1986). Les produits résultant de la fermentation de ces composés sont principalement des acides gras et des composés organiques volatiles donnant une odeur caractéristique lors de la culture de ces bactéries.

La température optimale de croissance de ce groupe d'espèces se situe entre 37 et 40°C et elles peuvent croître à des températures allant de 15°C à 45°C. La croissance est inhibée lorsque l'environnement est à un pH inférieur à 5.1 ou lorsque le taux en NaCl est supérieur à 5% (Popoff, 2018). Elles sont très sensibles à la présence d'oxygène.

L'efficacité de la toxinogénèse est plus importante lorsque la bactérie se trouve dans des conditions optimales à sa croissance (températures, pH, concentration en NaCl, présence de nutriments). La synthèse des neurotoxines botuliques va principalement se produire lors de la fin de la phase exponentielle et le début de la phase stationnaire (Connan et al., 2013). Le gène *bont* se trouvant sur un bactériophage pour les *C. botulinum* du groupe III, une souche de *C. novyi sensu lato* peut perdre ou gagner la faculté de synthétiser la toxine botulique, la perte ou le gain du phage pouvant

se produire dans l'environnement naturel ou en conditions de laboratoire (Eklund et al., 1974; Eklund et al., 1972; Eklund et al., 1971).

Lorsque les conditions ne sont plus propices à la survie ou à la croissance de *C. novyi sensu lato* (manque de nutriments, quorum sensing par exemple), les cellules végétatives vont initier un processus de sporulation (Gauvry et al., 2017). La spore est une forme permettant de résister à des environnements défavorables à la croissance de *C. novyi sensu lato* mais c'est aussi une forme permettant sa dissémination.

La germination des spores est quant à elle induite lorsque les conditions redeviennent favorables à la croissance de *C. novyi sensu lato*. Pour les souches de *C. botulinum* du groupe III, les cadavres d'animaux sont des environnements propices à la sporulation, la germination mais aussi la production de toxine. En effet, ce type d'environnement se trouve en anaérobiose, à un faible potentiel rédox, un pH proche de 7 et possède une grande quantité de nutriments (Espelund & Klaveness, 2014; Rocke, 2006).

#### **1.4. Isolement des souches**

L'ensemble des caractéristiques physiologiques de *C. novyi sensu lato* présentées précédemment fait que l'isolement des souches est très difficile à effectuer, d'une part à cause des conditions anaérobies nécessaires à la culture de ces souches, d'autre part du fait de l'absence de trait caractéristique permettant de faire de la culture sélective (Franciosa et al., 1996; Skarin et al., 2010). Ces bactéries croissent difficilement en cultures mixtes, que ce soit sur milieu liquide ou gélosé, car elles sont peu compétitrices vis-à-vis des autres espèces présentes dans leur environnement (Skarin et al., 2010). Un autre facteur limitant est la perte d'EGM (dont le phage présentant le gène *bont* pour les souches toxiques) lors des cultures successives en laboratoire (Franciosa et al., 1996).

A ce jour, aucun milieu sélectif n'a été développé pour ce groupe d'espèces. L'isolement des souches se fait essentiellement en répétant successivement des cultures sur gélose à partir d'un prélèvement brut enrichi avec un milieu de culture liquide TPGY (Tryptose Peptone Glucose Yeast) ou FCMM (Forced Cooked Meat Medium). Les milieux gélosés utilisés sont principalement les milieux EYA (Egg Yeast Agar), Mc Clung Toabe Agar ou Blood Agar Base No.2 (BAB2) (Bano et al., 2015; Skarin et al., 2010). Le but de ces repiquages successifs est d'obtenir une souche pure en sélectionnant des colonies isolées présentant les caractéristiques morphologiques propres aux souches de *C. novyi sensu lato* tout en éliminant la flore annexe (Skarin et al., 2010).

Une méthode d'isolement par immunocapture a été développée. Elle consiste à utiliser des billes magnétiques portant des anticorps ciblant spécifiquement des spores de *C. botulinum* du groupe III, permettant ainsi d'aller récupérer les bactéries d'intérêt dans le bouillon d'enrichissement avant de procéder à l'isolement sur milieu gélosé. Avec cette méthode, 15 souches de *C. novyi sensu lato* (dont 7 *C. botulinum* C/D) ont pu être isolées à partir de 46 échantillons (Anza et al. 2014). Plus récemment, 4 souches de *C. botulinum* de type D et 1 souche de type C/D furent isolées en utilisant un protocole similaire avec un taux de succès relativement faible (25 % pour le type C/D, 0 % pour le type C, 45 % pour le type D (Park et al., 2022).

Au sein du groupe *C. novyi sensu lato*, il existe des différences biochimiques entre les souches (concernant les sucres, activités lipase et lécithinase par exemple). Les souches sont anaérobies strictes, peu compétitrices par rapport à d'autres espèces bactériennes. Il n'y a pas de milieu sélectif pour isoler les souches de ce groupe. Les méthodes d'isolement utilisées actuellement sont fastidieuses et présente un taux de succès faible.

## 2. Caractéristiques des génomes de *C. novyi sensu lato*

L'avènement des nouveaux outils de biologie moléculaire a permis de mettre en évidence que les souches de *C. botulinum* du groupe III étaient plus proches génétiquement des souches des espèces *C. novyi* et *C. haemolyticum* que des souches de *C. botulinum* du groupe I et II (Sasaki et al., 2001; Skarin et al., 2011; Skarin & Segerman, 2014). Comme évoqué précédemment, le terme *Clostridium novyi sensu lato* a été proposé en 2011 pour regrouper les espèces *C. novyi*, *C. botulinum* du groupe III et *C. haemolyticum* (Skarin et al., 2011). A la date du 31/12/2021, 118 génomes de *C. novyi sensu lato* étaient disponibles dans les bases de données publiques.

### 2.1. Chromosome

Le génome des souches de *C. novyi sensu lato* est composé d'un ADN chromosomique circulaire d'une taille se situant entre 2,2 et 2,8 millions de paires de base (en se référant aux 5 génomes complets disponibles dans la littérature) (Skarin et al., 2011; Woudstra et al., 2021). Comme pour l'ensemble des génomes du genre *Clostridium*, le contenu en GC de l'ADN chromosomique est bas (~27-28 %). Le chromosome a une taille plus petite que ceux des souches de *C. botulinum* du groupe I et II, *C. argentinense* du groupe IV et *C. butyricum* du groupe V (comprises entre 3,6 et 4,5 Mbp selon le groupe). Cette caractéristique indique que les souches de *C. novyi sensu lato* ont une tolérance plus faible vis-à-vis des éléments génétiques redondants ou excessifs (Skarin et al., 2011).

Les séquences d'ADN codantes représentent 86% du chromosome. Les gènes sont majoritairement groupés sur le brin d'ADN répliqué dans le sens 3'-5' (*leading strand*) qui est enrichi en guanine. Peu de gènes sont acceptés dans le brin retardé (*lagging strand*), indiquant qu'une efficacité de transcription élevée est cruciale pour les souches de *C. novyi sensu lato* (Skarin et al., 2011).

## 2.2. Plasmidome

En plus de l'ADN chromosomique, des éléments génétiques mobiles (EGM) comprenant des phages et des plasmides viennent s'ajouter au contenu génomique. Ces éléments extrachromosomiques vont former ce qui est appelé le plasmidome (Skarin et al., 2011). Comme pour le chromosome, les EGM retrouvés chez *C. novyi sensu lato* présentent un GC % bas (~26-28%). Le plasmidome va avoir un rôle sur la pathogénicité d'une souche puisque les gènes de virulence vont se retrouver sur les EGM. Par exemple, le gène *bont* codant la neurotoxine botulique C, D, C/D ou D/C est retrouvé sur un bactériophage d'environ 200 000 paires de base. Ce bactériophage, qui a des caractéristiques proches du chromosome (bas GC %, gènes majoritairement présents sur le brin sens), semble évolué selon des mécanismes similaires (Skarin et al., 2011).

Le plasmidome va aussi jouer un rôle majeur dans la diversité des génomes chez *C. novyi sensu lato*. Un grand nombre de plasmides et de phages présents chez *C. novyi sensu lato* ont été caractérisés et il a été rapporté que, selon les souches, la composition du mobilome peut-être différente (Woudstra et al., 2018; Woudstra et al., 2016). L'analyse de 24 génomes de *C. novyi sensu lato* a permis d'identifier 61 EGM différents répartis en 13 catégories différentes (nommés de PG1 à PG13). PG1 regroupe les phages portant le gène *bont* et un EGM présent dans la souche *C. haemolyticum* NTCTC9693 qui en est dépourvu. Le gène codant la toxine C3 est aussi présent dans cette catégorie d'EGM (Skarin & Segerman, 2014). En 2021, de nouveaux EGM ont été caractérisés, ajoutant 10 nouvelles catégories aux 13 déjà établies dans l'étude précédente (nommées de PG14 à PG23). La plupart des EGM nouvellement caractérisés sont des prophages (Fillo et al., 2021).

## 2.3. Transferts horizontaux de gènes

La diversité du plasmidome retrouvée chez *C. novyi sensu lato* démontre l'existence de transferts d'EGM dans ce groupe d'espèces. Ce phénomène, essentiel pour que la bactérie s'adapte à l'évolution de son environnement, est fréquent entre des espèces génétiquement proches et occasionnel entre des espèces éloignées (Skarin & Segerman, 2011).



Le fait qu'une souche puisse perdre sa capacité à produire la neurotoxine botulique démontre que des EGM peuvent être perdus ou acquis. Des études ont déjà démontré que l'EGM contenant le gène codant pour la neurotoxine botulique pouvait être transféré dans une souche non toxique de *C. novyi* et que le phénomène inverse pouvait aussi se produire (Eklund et al., 1974; Eklund et al., 1972; Eklund et al., 1971). De plus, une souche de *C. botulinum* produisant la neurotoxine de type C peut perdre le bactériophage puis, lors d'un transfert d'EGM, récupérer un nouveau bactériophage contenant le gène codant pour la neurotoxine de type D, changeant ainsi le type de toxine produite (Eklund & Poysky, 1974).

Ces phénomènes de transferts horizontaux de gènes peuvent être accélérés par la présence de séquences d'insertion (IS) (Skarin & Segerman, 2011). Une caractéristique des génomes de *C. novyi sensu lato* est la présence d'un grand nombre de ces séquences d'insertion sur le chromosome et sur les EGM. Les génomes ont donc probablement été soumis à une forte activité des transposons, que ce soit au niveau du chromosome ou des EGM. Des groupes d'EGM (PG8-9-10-11) partagent du matériel génétique qui serait la conséquence de l'échange d'un module contenant des gènes de structures phagiques via des phages non apparentés à ces EGM. Ce module phagique est présent dans les 4 groupes mais pas sur l'ensemble des EGM. Le gène codant pour la toxine alpha a été trouvé dans plusieurs génomes de souches de *C. novyi* sur des EGM des groupes PG9 et 10. Le gène retrouvé sur l'EGM du groupe PG9 est proche de deux éléments mobiles : une transposase et une recombinase, ce qui est une preuve que ces échanges de matériel génétique sont favorisés par des séquences d'insertion et des enzymes recombinases (Skarin & Segerman, 2014). Les nouvelles catégories caractérisées en 2021 montrent des signes d'échanges de matériel génétique avec des catégories de la première étude. En effet, les catégories PG14 et PG20 montrent des similarités avec les catégories PG10 et PG7 (Fillo et al., 2021). Il est probable que les EGM aient fortement contribué à l'évolution constante du génome des souches de *C. novyi sensu lato*, ce qui pourrait expliquer les grandes différences génétiques retrouvées dans ce groupe d'espèces (Skarin et al., 2011).

#### **2.4. Répartition des souches en 4 lignées**

La comparaison des génomes de *C. novyi sensu lato* a permis de définir 4 lignées différentes annotées de I à IV, avec deux sous-lignées pour la lignée I (Ia et Ib) (Skarin & Segerman, 2014). Les génomes séquencés depuis cette classification initiale confirment la pertinence de cette proposition et son adéquation à *C. novyi sensu lato* (Fillo et al., 2021; Woudstra et al., 2016). La lignée I ne contient que des souches productrices de toxine botulique, les souches de la sous-lignée Ia produisant principalement la BoNT C/D et les souches de la sous-lignée Ib produisant principalement la BoNT

D/C, même si quelques souches produisant d'autres types toxiques peuvent être retrouvées dans ces deux sous-lignées. Les souches de la lignée II vont regrouper les trois espèces de *C. novyi sensu lato*. Les souches de *C. botulinum* du groupe III retrouvées dans cette lignée produisent principalement la BoNT C. Enfin, très peu de génomes sont disponibles pour les deux dernières lignées. La lignée III présente deux souches de *C. botulinum* de type C/D et une souche de *C. novyi* mais le faible nombre de souches présentes ne permet pas de définir une tendance sur le contenu de cette lignée. Enfin, pour les 6 souches de la lignée IV, ce sont majoritairement des souches de *C. novyi* type A qui sont retrouvées (Skarin & Segerman, 2014).

## 2.5. Système CRISPR-Cas

Comme évoqué précédemment, une forte diversité est retrouvée au niveau du plasmidome de *C. novyi sensu lato*. Dans une étude précédente, il a été démontré que les souches de *C. botulinum* de type C/D possèdent plus d'EGM que les souches de type D/C. La comparaison de 17 génomes de souches de type C/D et D/C isolées en France a mis en évidence une différence au niveau des systèmes CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)-Cas. Trois systèmes CRISPR-Cas fonctionnels ont ainsi été détectés dans les souches de type D/C tandis que pour les souches de type C/D, un seul système a été détecté. De plus, le système retrouvé ne présente pas tous les gènes *cas* et ne semble pas fonctionnel (Woudstra et al., 2016). Ces résultats pourraient expliquer le plus faible nombre d'EGM détecté dans les souches D/C pourvues de ces systèmes fonctionnels.

Le système CRISPR-Cas est un système de défense présent chez les bactéries et les archées pour faire face à l'invasion d'EGM. Ce système se compose de deux parties distinctes : les gènes *cas* et le locus CRISPR (Figure 14a). Selon la classe et le type de système, la composition en gènes *cas* va être différente (Figure 14b). Le locus CRISPR va quant à lui être composé d'une succession de séquences, appelées spacers, correspondant à une partie du génome des EGM invasifs, séparés entre eux par la même séquence répétée (Figures 14a et 15). A ce jour, les systèmes CRISPR-Cas se divisent en 2 classes, en 6 types et plus de 25 sous-types (Makarova et al., 2020).

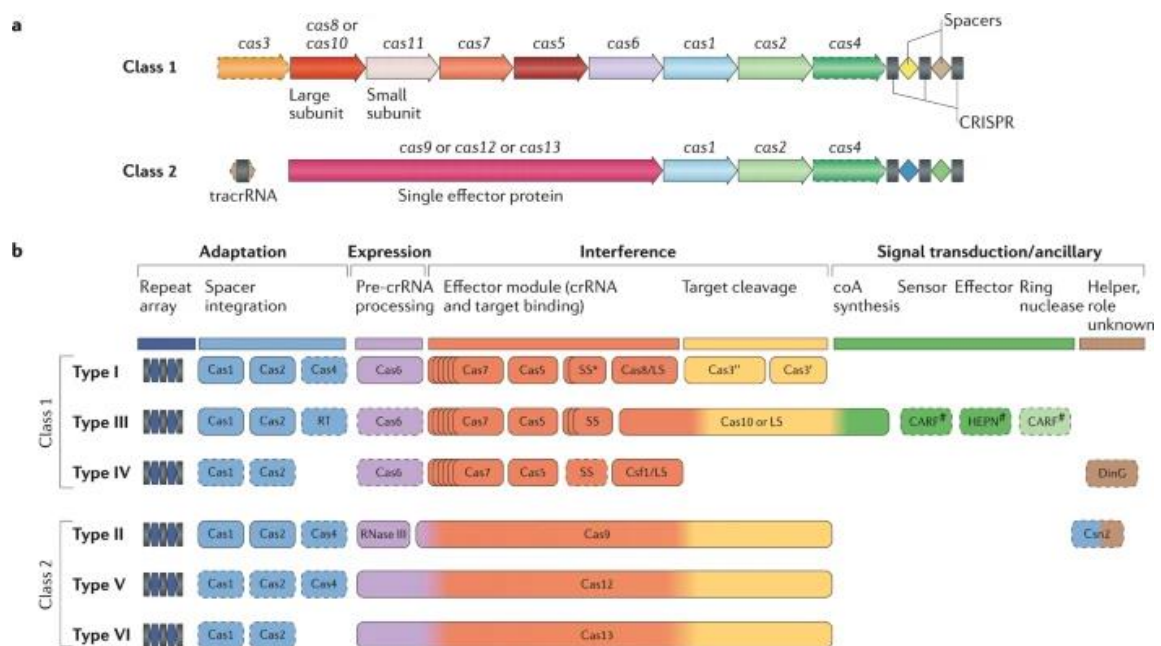


Figure 14 : Liste des différentes classes et types de systèmes CRISPR-Cas et rôles des différents gènes *cas*.

Source : Makarova et al., 2020

Lors de l'arrivée d'un EGM dans le cytoplasme, le module d'adaptation va venir capturer une petite séquence de l'élément invasif puis l'intégrer au locus CRISPR entre deux séquences répétées. Le locus CRISPR va ensuite être transcrit et maturer en plusieurs petits ARNcr complémentaires de la séquence de chaque spacer présent dans le locus CRISPR. Ces ARNcr vont servir de guide lorsque l'EGM se présente de nouveau dans la cellule bactérienne afin que celui-ci soit dégradé par une protéine Cas (Figure 15) (Semenova & Severinov, 2017). L'absence de certains gènes *cas* dans le système le rend non fonctionnel. Par exemple, en absence des gènes *cas1* et *cas2* composant le module d'adaptation, le système ne va pas pouvoir acquérir de nouveaux spacers. Il peut cependant compenser ce manque si un autre système fonctionnel possédant *cas1* et *cas2* est présent dans le génome.

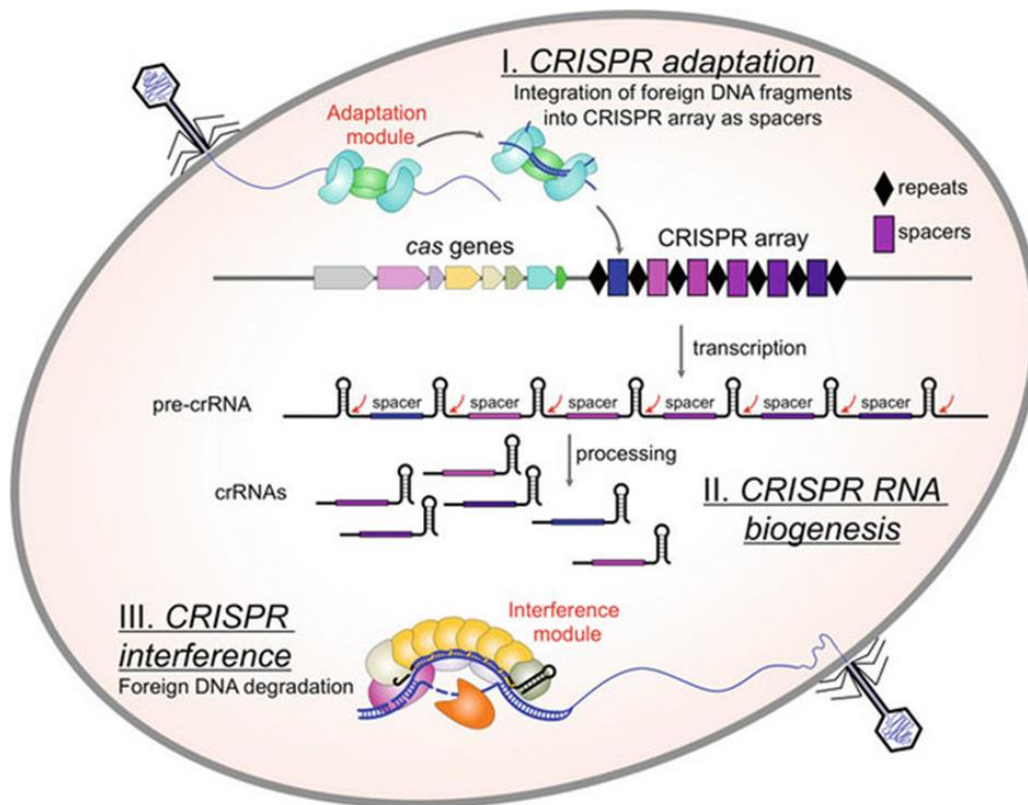


Figure 15 : Fonctionnement d'un système CRISPR-Cas.

Source : Semenova and Severinov., 2017.

Une première étude portant sur l'analyse des systèmes CRISPR-Cas de l'ensemble des souches de plusieurs groupes de *C. botulinum* a été publiée en 2017 (Negahdaripour et al., 2017). Les résultats de cette étude montrent que des systèmes CRISPR-Cas sont retrouvés dans 80% des 20 génomes étudiés. La présence fréquente de systèmes CRISPR-Cas sur des plasmides a aussi été rapportée. Pour *C. botulinum* du groupe III, des EGM provenant de 8 souches ont été étudiés ainsi que le chromosome d'une souche (BKT015925, seul génome complet disponible à l'époque de l'étude). Les systèmes CRISPR-Cas ont été retrouvés sur le prophage p1 contenant le gène codant la neurotoxine botulique ou le plasmide p2 de 6 des 8 souches testées (Negahdaripour et al., 2017). Une étude plus récente sur l'analyse des systèmes CRISPR-Cas de 241 souches de *C. botulinum* du groupe I et *C. sporogenes* a observé une présence de système dans 83% des souches avec une localisation sur le chromosome et sur les EGM. Les types I-B et III (A, B et D) sont les plus retrouvés. Les systèmes de type I-B retrouvés sur les plasmides ne possèdent pas tous les gènes *cas* et semblent non fonctionnels, comme pour les systèmes retrouvés sur les génomes de *C. botulinum* de type C/D (Wentz et al., 2022; Woudstra et al., 2016).

Les génomes de *C. novyi sensu lato* présentent un pourcentage en GC faible. Ils se répartissent en 4 lignées et présentent une très grande diversité au niveau des éléments génétiques mobiles (plasmides et phages). La diversité des EGM et la présence de nombreux éléments d'insertion dans les génomes suggèrent l'existence de transferts horizontaux fréquents. Des systèmes CRISPR-Cas ont par ailleurs été identifiés dans les génomes de *C. botulinum*, à la fois au niveau du chromosome mais également au niveau de certains EGM avec des différences entre les génomes des souches. Les premiers résultats de la littérature semblent montrer que les souches ayant plus d'EGM présentent des systèmes CRISPR-Cas incomplets et potentiellement non fonctionnels. Ces systèmes seraient incapables d'acquérir de nouveaux spacers ce qui va avoir comme conséquence une invasion plus fréquente d'EGM étrangers. Cependant, le panel de souches testées n'étant pas exhaustif, ces hypothèses sont encore à confirmer.

### **3. Présentation des outils et méthodes de typage existants**

La disponibilité croissante des génomes ces dernières années a permis l'identification de nouvelles cibles potentielles pertinentes pour typer les souches du groupe *C. novyi sensu lato*.

La présence d'un plasmidome important chez ce groupe d'espèces font des EGM des cibles d'intérêt pour discriminer les souches. Ces bactéries possédant des flagelles tout autour de leur membrane, cibler les gènes codant pour cette structure peut s'avérer pertinent. Des outils de typages déjà existants pour de nombreux pathogènes seront aussi décrits dans ce paragraphe.

Parmi les cibles potentielles encore non testées dans les études disponibles dans la littérature scientifique, l'utilisation des systèmes CRISPR-Cas comme cible pour typer les souches semble pertinente au vu de la forte proportion de génomes présentant ce type de système dans l'ensemble des groupes de *C. botulinum* (Negahdaripour et al., 2017; Wentz et al., 2022).

Les méthodes développées pour le typage des souches de *C. novyi sensu lato* sont présentées dans le Tableau 5 et seront détaillées dans cette partie.

Tableau 5 : Liste des méthodes de typage existantes pour la caractérisation de *C. novyi sensu lato*

Méthode	Nom complet	Cible(s)	Principe	Souche pure	Paragraphe	Références
PFGE	Pulsed-Field Gel Electrophoresis	Génome complet	Enzymes de restriction puis gel d'électrophorèse en champ pulsé	Oui	3.1.1	Skarin 2010-2014-2015 Nakamura 2010
AFLP	Amplified Fragment Length Polymorphism	Régions non ciblées	Enzymes de restriction puis PCR puis gel d'électrophorèse	Oui	3.1.2	Myllykoski 2009
RAPD	Random Amplified Polymorphic DNA	Régions non ciblées	PCR puis gel d'électrophorèse	Oui	3.2.1	/
MLVA	Multi Loci VNTR Analysis	Régions VNTR	PCR puis séparation électrophorèse capillaire et séquençage	Non	3.2.2	Auricchio 2019
MLST	Multi Loci Sequence Typing	Gènes ciblés	PCR puis séquençage	Oui	3.2.3	Scalfaro 2019
FliC	PCR de détection du gène <i>fliC</i>	Gènes ciblés	PCR	Non	3.2.4	Woudstra 2015 Fillo 2021
EGM	PCR de détection des éléments génétiques mobiles	Gènes ciblés	PCR	Non	3.2.5	Woudstra 2018
WGS	Whole genome sequencing	Génome complet	Séquençage	Oui	3.3.1	Skarin 2011-2014 Woudstra 2015-2017 Woudstra 2016-2021 Fillo 2021
MALDI-TOF	Matrix Assisted Laser Desorption Ionisation-Time Of Flight	Protéines	Obtention de spectres par spectrométrie de masse	Oui	3.3.2	Grosse-Herrenthey 2008 Bano 2017

### **3.1. Méthodes de caractérisation utilisant des enzymes de restrictions**

#### **3.1.1. Pulse Field Gel Electrophoresis (PFGE)**

##### **3.1.1.1. Principe de la méthode**

L'électrophorèse en champ pulsé (PFGE) est une méthode de typage utilisée pour l'étude de nombreux pathogènes. Le principe de la méthode est de couper l'ADN génomique à l'aide d'enzymes de restriction puis de faire migrer les fragments par électrophorèse afin d'obtenir un profil appelé pulsotype. En électrophorèse classique, les grandes molécules d'ADN (> 50 kbp) sont impossibles à séparer même en changeant au maximum la résolution du gel d'agarose. Grâce à l'électrophorèse en champs pulsés, l'orientation du courant électrique va être alternée au cours de la migration de l'ADN sur le gel, permettant aux grands fragments de pouvoir passer les mailles du gel.

##### **3.1.1.2. Historique**

La méthode a été présentée pour la première fois en 1984 par les chercheurs David C. Schwartz et Charles R. Cantor sur la levure *Saccharomyces cerevisiae*. Cette nouvelle méthode a été développée dans le but de séparer des plus grandes molécules d'ADN avec une meilleure résolution que le gel d'électrophorèse classique (Schwartz & Cantor, 1984).

##### **3.1.1.3. Utilisation de la PFGE pour le typage de *C. botulinum***

Pour *C. botulinum*, la PFGE a d'abord été développée pour les souches des groupes I et II (Hielm et al., 1998; Leclair et al., 2006; Lin & Johnson, 1995; Umeda et al., 2009) puis adaptée pour *C. botulinum* du groupe III (Skarin et al., 2010). Elle a été appliquée dans différentes études pour comparer des souches isolées au sein d'un même épisode ou isolées de différents pays. Dans une première étude, elle fut utilisée sur 13 souches de *C. botulinum* du groupe III et a permis de discriminer des souches de type D/C responsables de botulisme bovin au Japon (Nakamura et al., 2010). Elle fut ensuite utilisée sur des souches de type C responsables de botulisme aviaire en Suède et Norvège où des similarités ont pu être observées entre les souches de type C des deux pays (Figure 16) (Skarin et al., 2010). Enfin, elle fut utilisée sur des souches de type C/D responsables de botulisme aviaire en Espagne et en Suède. Dans cette étude, les souches provenant de Suède se sont regroupées avec les souches provenant d'Espagne (Figure 17), indiquant une forte proximité génomique des souches de type C/D en Europe. Des différences au niveau des EGM présents dans ces souches furent relevées, mais la méthode n'a pas permis de différencier les deux plasmides uniques retrouvés car il n'y avait aucun site de restriction sur ces EGM (Anza et al., 2014a).

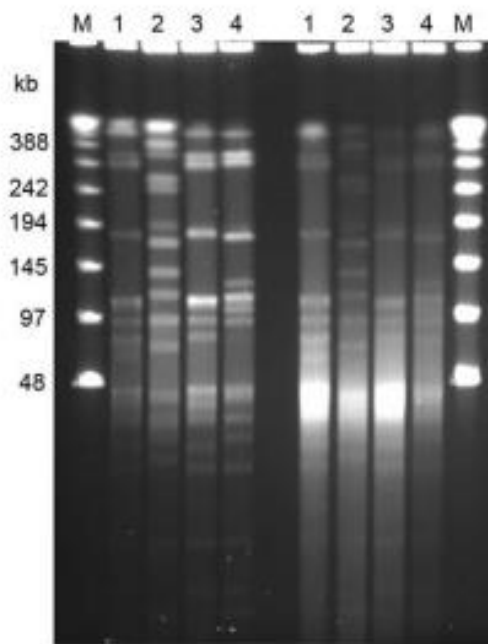


Figure 16 : Gel d'électrophorèse en champ pulsé testé sur 4 souches de *C. botulinum* type C responsables de botulisme aviaire (4 premiers puits : fixation des bactéries au formaldéhyde, 4 derniers puits : sans formaldéhyde, M : Marqueur de taille, 1 : Souche 08-BKT015925, 2 : Souche 07-BKT002873, 3 : Souche 07-C6N, 4 : Souche 07-V891).

Source : Skarin et al., 2010.

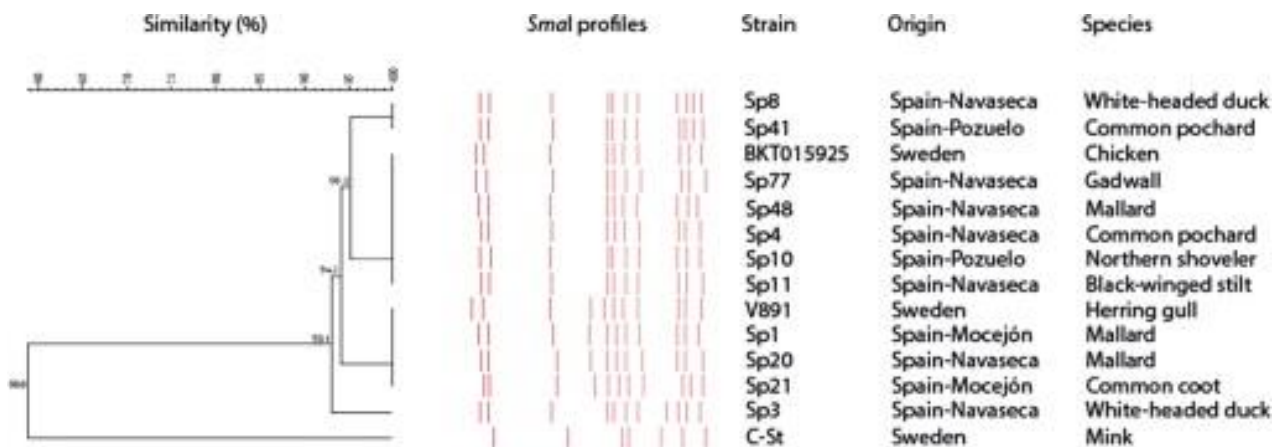


Figure 17 : Dendrogramme des pulsotypes de 14 souches isolées de *C. botulinum* du groupe III.

Source : Anza et al., 2014.



#### **3.1.1.4. Limites**

La PFGE est une méthode laborieuse et comporte certaines limites. Tout d'abord, il faut disposer d'une souche pure pour pouvoir appliquer la méthode, ce qui est, comme évoqué précédemment, difficile pour *C. novyi sensu lato* (cf III.1.4). De plus, la production de DNases endogènes est un phénomène commun à certaines espèces bactériennes du genre *Clostridium*, qui va entraîner une dégradation de l'ADN au cours de l'extraction (Chaudhuri et al., 1996; Hielm et al., 1998; Lin & Johnson, 1995). Les espèces bactériennes de *C. novyi sensu lato* étant des bacilles Gram positif capables de sporuler, la présence de spores composées de plusieurs tuniques très résistantes et de cellules végétatives avec une paroi en peptidoglycane plus épaisse rend la lyse bactérienne plus difficile, ce qui va avoir un impact sur la quantité d'ADN extrait (Le Marechal et al., 2018; Rohde et al., 2019). Le manque de site de restriction sur certains EGM peut aussi entraîner des biais dans l'interprétation des résultats (Anza et al., 2014a).

### **3.1.2. Amplified Fragment-Length Polymorphism (AFLP)**

#### **3.1.2.1. Principe de la méthode**

La méthode AFLP se base sur le polymorphisme de longueur de fragments amplifiés aléatoirement dans le génome. Le principe de cet outil de typage est de digérer l'ADN génomique bactérien par deux enzymes de restriction, puis de ligaturer les produits de la digestion avec des oligonucléotides. Une PCR est ensuite effectuée pour amplifier les produits de digestion, les amorces ciblant les oligonucléotides ajoutés aux extrémités des fragments. La dernière étape est d'effectuer une migration des fragments sur gel d'agarose afin d'obtenir un profil (Vos et al., 1995).

#### **3.1.2.2. Utilisation de la méthode pour *C. novyi sensu lato***

Cette méthode a été développée et utilisée pour des souches de *C. botulinum* du groupe I et II (Types A, B, E et F) responsables de botulisme et sur des souches de *C. novyi* responsables d'autres maladies chez l'humain (Keto-Timonen et al., 2005; McLauchlin et al., 2002). La méthode AFLP a aussi été utilisée sur 174 souches de *C. botulinum* représentatives de chaque type toxinique dans une étude visant à examiner la diversité génétique présente chez ces bactéries. Les résultats montrent que cette méthode de typage permet de grouper les souches par groupe et par toxine produite (Figure 18) (Hill et al., 2007). Enfin, pour les souches de *C. botulinum* du groupe III impliquées dans des épisodes de botulisme animal, une étude a permis de caractériser des souches au cours d'une enquête épidémiologique sur un épisode de botulisme bovin en Finlande (Myllykoski et al., 2009).

### 3.1.2.3. Limites

Comme pour la PFGE, une souche pure est requise pour pouvoir appliquer la méthode et la production de DNases endogènes par les souches de *C. botulinum* peut dégrader l'ADN et affecter les résultats, même si les fragments obtenus par digestion pour cette technique sont plus petits que ceux obtenus en PFGE (Vos et al., 1995).

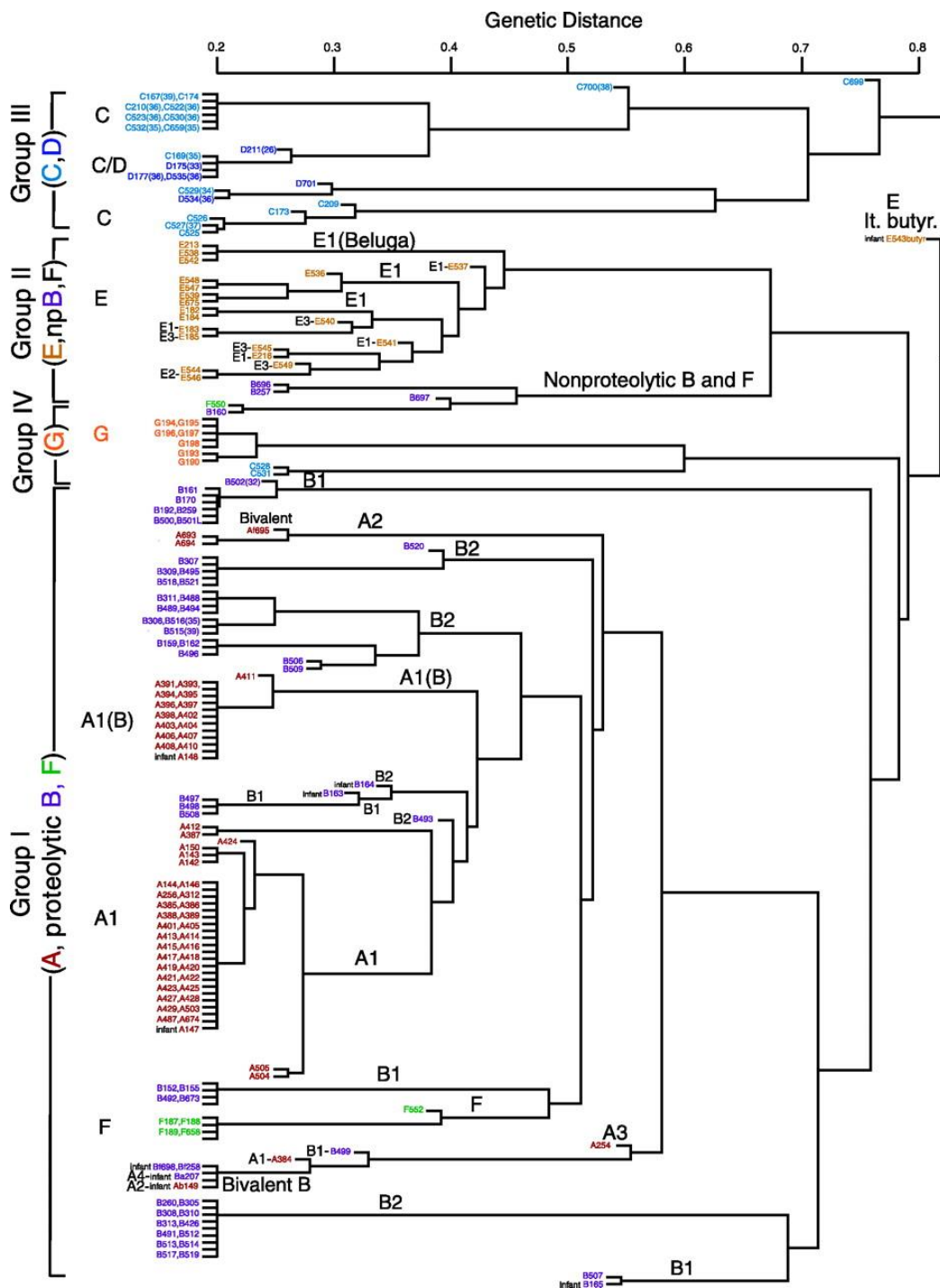


Figure 18 : Dendrogramme de 174 souches de *C. botulinum* basé sur les résultats obtenus avec la méthode AFLP.

Source : Hill et al., 2007.

Les méthodes PFGE et AFLP permettent de discriminer les souches. Leur application dans le contexte du botulisme animal présente deux inconvénients majeurs : il est nécessaire de disposer de la souche pure isolée et la présence de DNase peut entraîner des difficultés méthodologiques pour certaines souches.

D'une manière générale, les outils de typage basés sur des enzymes de restriction que sont la PFGE et l'AFLP se sont révélés être des outils discriminants et reproductibles mais la disponibilité des séquenceurs de 2<sup>ème</sup> et 3<sup>ème</sup> génération à un tarif très compétitif rend ces approches moins attractives aujourd'hui.

### **3.2. Méthodes de caractérisation basées sur l'outil PCR**

La multiplication du nombre de génomes dans les bases de données publiques ces 20 dernières années a permis de mettre en évidence des régions génomiques d'intérêt pour le développement d'outils de typage. Ce chapitre détaillera l'ensemble des outils de typage utilisés chez *C. novyi sensu lato* basés sur des PCR ciblant ces séquences d'intérêt.

#### **3.2.1. Random Amplified Polymorphic DNA (RAPD)**

##### **3.2.1.1. Principe de la méthode**

L'outil de typage RAPD se base sur l'amplification aléatoire de différentes régions du génome bactérien avec des amorces courtes aléatoires. Le polymorphisme va alors être la conséquence du nombre de sites pouvant s'hybrider avec les amorces. Les produits amplifiés sont ensuite soumis à une migration sur gel d'électrophorèse afin d'obtenir un profil caractéristique de chaque génome testé. La comparaison des profils obtenus va ainsi permettre de typer les souches.

##### **3.2.1.2. Utilisation chez *C. botulinum***

Cette technique a essentiellement été testée dans le cadre d'études visant à typer des souches de *C. botulinum* de type E (Hannett et al., 2011; Leclair et al., 2006).

##### **3.2.1.3. Limites**

Comme pour les outils de typage PFGE et AFLP, une souche pure est nécessaire pour utiliser cette méthode de typage. La dégradation de l'ADN par les DNases produites par les souches de *C. botulinum* peut avoir un impact sur les profils obtenus, biaisant de ce fait les résultats. De plus, le fait que les amplifications soient aléatoires rend cette technique peu reproductible.

## **3.2.2. Multi Loci VNTR Analysis (MLVA)**

### **3.2.2.1. Généralités**

La MLVA est un outil de typage qui se base sur l'amplification et le séquençage de plusieurs loci du génome bactérien appelés VNTR, pour Variable Number of Tandem Repeats. Ce sont des régions du génome présentant une succession de courtes séquences répétées en tandem, dont le nombre va être variable selon les souches. En effet, ce sont des régions instables présentant des erreurs apparues au cours des différentes étapes de modification de l'ADN (Réplication, transcription, recombinaison, réparation), entraînant l'addition ou la délétion d'une ou plusieurs répétitions au niveau de ces VNTR. Le génotypage des espèces bactériennes par la méthode MLVA va donc reposer sur la variation naturelle du nombre de motifs répétés pour un nombre de locus VNTR déterminé au cours du développement de la méthode pour une espèce donnée.

### **3.2.2.2. Historique**

Les régions VNTR font partie de la famille des minisatellites, caractérisée pour la première fois sur l'ADN humain en 1980 (Wyman & White, 1980). Ces minisatellites ont d'abord été utilisés en médecine légale pour réaliser des empreintes génétiques lors d'enquêtes criminelles ou pour effectuer des tests de paternité. Ces régions génomiques sont présentes chez tous les organismes vivants, eucaryotes et procaryotes. Chez les bactéries, l'utilisation des VNTR pour le typage est intervenue suite au séquençage de *Haemophilus influenzae* et de l'identification de cibles d'intérêt (van Belkum et al., 1997). Elle a ensuite été développée pour *Mycobacterium tuberculosis* (Frothingham & Meeker-O'Connell, 1998) avant d'être nommée MLVA quelques années plus tard lors du développement de la méthode pour *Bacillus anthracis* (Keim et al., 2000).

### **3.2.2.3. Principe de la méthode**

La première étape de la MLVA consiste à déterminer des régions VNTR susceptibles de présenter un intérêt pour le typage des souches d'une espèce bactérienne. Une fois les cibles choisies, des amorces PCR sont conçues aux extrémités des régions ciblées. Après amplification par PCR, les tailles des amplicons obtenues sont déterminées par électrophorèse (gel d'agarose classique ou électrophorèse capillaire). Un séquençage des produits amplifiés permet de valider le nombre de répétitions présents dans chaque locus VNTR étudié. Après détermination du nombre de répétitions pour l'ensemble des régions VNTR testées, chaque souche va se voir attribuer un profil MLVA qui sera un code composé de N chiffres, N étant le nombre de loci VNTR étudiés. Le chiffre sera quant

à lui égal au nombre de répétitions présent dans le locus VNTR. Un dendrogramme est ensuite construit pour comparer les profils MLVA des souches étudiées.

#### **3.2.2.4. Utilisation de la méthode MLVA chez *C. botulinum***

La méthode MLVA a été développée dans un premier temps pour les souches de *C. botulinum* du groupe I de type A en 2008 avec un schéma comportant 10 locus VNTR. Cette méthode a permis de caractériser 59 souches de *C. botulinum* type A et 3 souches bivalentes Ba en 38 génotypes. Pour les souches de type A, elle a permis de les discriminer en 4 sous-types (BoNT/A1-A4). Des 38 génotypes, 30 proviennent du sous-type BoNT/A1, démontrant que cette méthode est discriminante pour des souches d'un même sous-type (Macdonald et al., 2008).

En 2011, un nouveau schéma MLVA de 15 locus a été développé pour les souches de *C. botulinum* du groupe I. Sur ces 15 locus VNTR, 10 provenaient de l'étude de 2008. Les cinq nouveaux locus VNTR ont été ajoutés pour que la méthode puisse être appliquée sur les sérotypes B et F. Ce nouveau schéma a été utilisé sur 89 souches de *C. botulinum* du groupe I et 5 souches de *C. sporogenes* (espèce modèle non toxique). Au total, 89 génotypes distribués en 7 clusters principaux ont pu être identifiés suite à cette analyse MLVA. A noter que 28 souches testées ne sont pas regroupées avec ces 7 clusters. Ce nouveau schéma, en plus de pouvoir inclure plus de souches du groupe I, est plus discriminant que son prédécesseur. Il a été utilisé par d'autres équipes pour typer d'autres souches de *C. botulinum* du groupe I (Raphael et al., 2014; Umeda et al., 2013).

En complément, un autre schéma utilisant seulement 7 des 15 locus a aussi été testé pour faciliter l'analyse et abaisser son coût. Le dendrogramme obtenu à partir de ce schéma de 7 locus est très proche de celui obtenu avec 15 locus. Ce schéma semble cependant sous-estimer la distance évolutive entre les groupes, le rendant de fait moins discriminant. Il reste un bon outil pour discriminer les souches au niveau du cluster mais pas au niveau de la souche (Fillo et al., 2011).

### **3.2.3. Multi Loci Sequence Typing (MLST)**

#### **3.2.3.1. Généralités**

La MLST est un outil de typage basé sur l'étude des séquences de plusieurs gènes de ménage présents dans le génome bactérien. Ces gènes, en évoluant, vont présenter un polymorphisme qui va permettre de discriminer les souches au sein d'une même espèce bactérienne.

#### **3.2.3.2. Principe de la méthode**

Après la sélection des gènes de ménages et le design d'amorces PCR pour chaque cible, le principe de la méthode repose sur l'amplification des régions des gènes de ménage sélectionnés puis le séquençage des produits d'amplification. En comparant les séquences obtenues pour chaque cible, un profil allélique, nommé *ST* pour *Sequence Typing*, est attribué à chaque souche.

#### **3.2.3.3. Utilisation chez *C. novyi sensu lato***

La MLST a d'abord été développée pour les souches de *C. botulinum* du groupe I et II. Pour le groupe I, la méthode se base sur 7 gènes de ménage (*aroE*, *mdh*, *aceK*, *appB*, *rpoB*, *recA* et *hsp*) tandis que pour le groupe II, la méthode se base sur 15 gènes de ménage (*oppB*, *pta*, *recA*, *tuf*, *rpoB*, *pyc*, *mutL*, *gyrB*, *lepA*, *ilvD*, *guaA*, *atpD*, *trpB*, *16S*, *23S*) (Jacobson et al., 2008; Macdonald et al., 2011).

Récemment, un schéma MLST a été développé pour les souches de *C. botulinum* du groupe III. Celui-ci se base sur 13 gènes de ménage (*arpD*, *dnaN*, *efTu*, *GroEL*, *gyrB*, *lepA*, *pheS*, *proC*, *rpL10*, *ropB*, *secD*, *secF*, *spo0A*) et a été testé sur 60 souches de *C. novyi sensu lato* provenant de plusieurs pays d'Europe (Scalfaro, 2019).

#### **3.2.3.4. Limites**

Il a été rapporté que la capacité de discrimination des souches obtenue avec la MLST est plus basse que celle obtenue avec l'outil MLVA, que ce soit avec les souches du groupe I comme celles sur groupe III (Fillo et al., 2011; Macdonald et al., 2011; Scalfaro, 2019). Contrairement à la RAPD, cette technique utilise des cibles de PCR prédéfinies, il est donc nécessaire d'avoir une souche pure pour cette technique. De plus, les gènes de ménages sont des régions où les mutations sont rares, ce qui va rendre difficile le typage de souches génétiquement proches.

### 3.2.4. Comparaison des souches de *C. botulinum* sur la base du gène *fliC*

#### 3.2.4.1. Généralités

Le flagelle est une structure présente chez les bactéries qui leur confère la mobilité. C'est un filament de forme hélicoïdale ancré au niveau de la membrane plasmique bactérienne par un corps basal. Une troisième partie nommée crochet va se situer entre le filament et le corps basal. Il va faire le lien entre les deux structures et transmettre le mouvement du corps basal au filament (Imada, 2018; Macnab, 2004; Minamino & Imada, 2015). Selon l'espèce bactérienne, le nombre et la localisation du ou des flagelles sont différents, ce qui va avoir un impact sur la ciliature de la bactérie :

- Présence d'un flagelle polaire → Ciliature monotriche
- Présence de plusieurs flagelles polaires → Ciliature lophotriche
- Présence d'un flagelle à chaque pôle de la bactérie → Ciliature amphitriche
- Présence de flagelles tout autour de la bactérie → Ciliature péritriche

La flagelline est une protéine majeure dans la constitution du flagelle. Elle va donner au filament flagellaire sa forme hélicoïdale, essentielle pour accomplir sa fonction. Ce filament est composé de plusieurs sous-unités de flagelline (Nakamura & Minamino, 2019).

La protéine se divise en 3 parties :

- Deux régions conservées aux extrémités N-terminale et C-terminale responsables de la sécrétion et la polymérisation du flagelle.
- Une région centrale hypervariable qui va représenter la partie exposée de l'organelle. C'est cette région qui contribue à l'antigénicité de la protéine. Au niveau du gène, la composition en nucléotides et la taille de cette région centrale vont différer au sein d'une même espèce.

#### 3.2.4.2. Utilisation chez *C. novyi sensu lato*

Les différentes espèces de *C. botulinum* présentent une ciliature péritriche du fait de la présence de flagelles tout autour de la bactérie. Pour les souches de *C. botulinum* du groupe III, la flagelline est codée par le gène *fliC* présent sur le chromosome bactérien. Des outils de typage ciblant le gène *fliC* semblent donc pertinents pour discriminer ces souches. D'autres gènes codant cette protéine ont déjà été utilisés pour le génotypage d'autres espèces pathogènes (Yadav et al., 2018) dont *C. botulinum* des groupes I et II (avec les gènes *flaA* et *flaB*) (Paul et al., 2007; Woudstra et al., 2013).

En ce qui concerne le groupe d'espèces *C. novyi sensu lato*, 7 types de gènes *fliC* ont été caractérisés jusqu'à présent dans les génomes disponibles et nommés de *fliC*-I à *fliC*-VII. Une

première étude sur la diversité du gène de flagelline *fliC* présent chez plusieurs souches a permis via le séquençage et la comparaison de 68 séquences de gènes *fliC* de grouper les gènes en 4 types classés de *fliC*-I à *fliC*-IV. L'analyse du gène *fliC* d'une des souches (D-4947, N° Accession KM496392.1) n'a pas permis de le classer dans un des types. Le gène a donc été classé comme *fliC*-V (Woudstra et al., 2015). En 2021, une seconde étude sur les gènes *fliC* de 60 souches de *C. botulinum* du groupe III nouvellement isolées en Italie a permis de mettre en évidence deux nouveaux types de gène *fliC* nommés *fliC*-VI et *fliC*-VII. Les séquences des 7 types de *fliC* ont été comparées et une matrice de similarité permet de voir les résultats de ces comparaisons (Tableau 6) (Fillo et al., 2021).

Tableau 6 : Matrice de similarités des 7 types de gène *fliC* caractérisés chez *C. botulinum* du groupe III.

Source : adaptée de Fillo et al., 2021.

<i>fliC</i> type	<i>fliC</i> -I	<i>fliC</i> -II	<i>fliC</i> -III	<i>fliC</i> -IV	<i>fliC</i> -V	<i>fliC</i> -VI	<i>fliC</i> -VII
<i>fliC</i> -I	/						
<i>fliC</i> -II	84.86	/					
<i>fliC</i> -III	92.66	86.93	/				
<i>fliC</i> -IV	46.79	47.48	47.48	/			
<i>fliC</i> -V	83.94	86.01	84.86	47.25	/		
<i>fliC</i> -VI	50.69	50.92	51.38	57.11	49.31	/	
<i>fliC</i> -VII	61.24	61.47	61.93	60.78	59.86	82.8	/

En ce qui concerne la première étude, des systèmes PCR ont été développés pour ces 5 types de *fliC* et des tests ont été effectués sur 560 échantillons (organes et environnement) provenant d'épisodes de botulisme animal survenus entre 2006 et 2014 dans plusieurs pays d'Europe et en Nouvelle-Calédonie (Woudstra et al., 2015).

La plupart des souches de type C/D ont un gène *fliC* de type I, montrant l'absence de variabilité au niveau du gène *fliC* pour ces souches. Pour les souches de type D/C, 3 types de gène *fliC* ont été retrouvés avec une majorité de *fliC*-IV et des échantillons positives pour *fliC*-I et II. Cela suggère que la diversité des gènes *fliC* est plus importante chez les souches de type D/C que chez celles de type C/D. Les gènes *fliC*-II ont majoritairement été retrouvés dans des génomes de souches de *C. botulinum* de type C dans les deux études. Les gènes *fliC*-III et *fliC*-V n'ont été détectés dans aucun échantillon. Ceci peut s'expliquer par le fait que ces gènes ont été retrouvés sur des génomes de souches plus anciennes provenant de continents différents (*fliC*-III : Eklund 1970 - USA et *fliC*-V : D-4947 – Afrique). De plus, les échantillons positifs pour le type D/C originaires de Nouvelle-Calédonie sont négatifs pour les gènes *fliC* testés, soulevant l'hypothèse qu'un autre gène *fliC* puisse



être présent dans ces souches et renforçant l'hypothèse que la localisation géographique joue un rôle dans la diversité des gènes *fliC*. Les deux gènes *fliC*-VI et *fliC*-VII caractérisés dans la deuxième étude ont été retrouvés dans le génome de 4 souches provenant du même épisode de botulisme bovin de type D/C, confortant l'hypothèse d'une variabilité plus importante au niveau des gènes *fliC* chez les souches de type D/C par rapport aux souches de type C/D. Des fragments de ces gènes sont aussi présents sur le génome de deux souches de type C/D phylogénétiquement proches isolées en Suède (BKT2873 et BKT75002) (Fillo et al., 2021; Skarin & Segerman, 2014; Woudstra et al., 2015).

En conclusion, la diversité du gène *fliC* retrouvée chez *C. botulinum* du groupe III est moins importante que ce qui avait été rapporté pour les gènes *flaA* et *flaB* présents dans les génomes des souches de *C. botulinum* du groupe I et II. Pour les souches géographiquement proches, notamment celles de type C/D provenant d'Europe, l'utilisation de l'analyse des gènes *fliC* comme outil de typage ne semble pas pertinente, l'ensemble de ces souches arborant le gène *fliC*-I. Une diversité du gène *fliC* plus importante est retrouvée chez les souches de *C. botulinum* de type D/C de lignée Ib même si la grande majorité des souches porte un gène *fliC*-IV (Woudstra et al., 2013; Woudstra et al., 2015). Cette cible ne semble donc pas appropriée pour le typage des souches de *C. novyi sensu lato*.

### **3.2.5. Développement de PCR ciblant des EGM**

#### **3.2.5.1. Généralités**

Certaines bactéries arborent, en plus de leur chromosome bactérien, des EGM qui font partie intégrante de leur génome. L'ensemble de ces EGM, appelé plasmidome, va leur permettre d'acquérir de nouveaux gènes, qui vont leur donner de nouvelles caractéristiques (facteurs de virulence, résistance aux antibiotiques par exemple).

Comme évoqué précédemment, ce plasmidome est un acteur majeur du génome du groupe d'espèces *C. novyi sensu lato*. En effet, de nombreux EGM variés ont été caractérisés. Ces EGM sont sujets à des transferts horizontaux, ce qui en fait un contributeur majeur de la diversité génétique présente dans ce groupe d'espèces (Skarin & Segerman, 2011, 2014).

#### **3.2.5.2. Utilisation chez *C. novyi sensu lato***

L'analyse des génomes complets a permis de mettre en évidence que la combinaison des EGM diffère d'une souche à l'autre, même entre des souches génétiquement proches (Skarin et al., 2011). L'utilisation des EGM comme cibles semble donc pertinente pour le typage des souches de *C. novyi sensu lato*.

Des PCR ont été développées pour cibler une partie des gènes de toxines retrouvés au niveau des différents EGM lors d'une étude dont le but était d'analyser le contenu en EGM de souches pures et d'échantillons naturellement contaminés (Woudstra et al., 2018).

- Pour le prophage p1, le gène ciblé et celui de la toxine botulique, les amorces PCR ayant été conçues pour chaque BoNT présent chez *C. botulinum* groupe III dans une étude précédente (Woudstra et al., 2012).
- Pour le plasmide p2, le gène ciblé est celui de l'alpha-toxine.
- Pour le plasmide p3, deux régions du gène de la toxine C2 sont ciblées (C2-I et C2-II).
- Pour le prophage p4, c'est une partie du gène codant la toxine putative CbC4\_7002.
- Pour le plasmide p5, le gène ciblé est un dérivé de la toxine epsilon, avec deux régions (I et II).
- Pour le prophage p6, une partie du gène codant la bactériocine est ciblé.

Au total, 110 extraits d'ADN de souches pures de *C. botulinum* du groupe III ont été testés et 11 profils de plasmidome ont été retrouvés. Pour les souches de type C, 2 profils présentent soit le plasmide p3 et le prophage p6 (13/20) soit le plasmide p3 seul (7/20). Trois profils sont retrouvés pour les souches de type D : aucun des EGM testés (1/5), le plasmide p3 seul (1/5) ou accompagné du plasmide p2 (3/5). Pour les 60 souches de types C/D étudiées, 4 profils de plasmidome sont retrouvés : p2-p3 (12/60), p2-p3-p4 (2/60), p2-p3-p4-p5 (14/60) et p2-p3-p5 (32/60). Aucune souche de type C/D ne présente le prophage p6. Enfin, pour les souches de type D/C, la plupart des souches ne contiennent aucun des marqueurs génétiques des EGM testés (21/23) tandis qu'une souche présente le profil p3-p6 et une autre souche présente le profil p2-p3 (Figure 19) (Woudstra et al., 2018).

bont type	Origin	nb of strains	p2	p3	p3	p4	p5	p5	p6
			alphatox	C2-I	C2-II	hyptox	epsilonI	epsilonII	bacteriocin
C	Pasteur, Sweden, The Netherlands, Germany	13	-	+	+	-	-	-	+
C	Pasteur, Italy, Brazil	7	-	+	+	-	-	-	-
D	Africa	1	-	+	+	-	-	-	-
D	The Netherlands	3	+	+	+	-	-	-	-
C/D	Japan, The Netherlands, Italy, France, Germany	12	+	+	+	-	-	-	-
C/D	Sweden, The Netherlands	2	+	+	+	+	-	-	-
C/D	Sweden, France, Italy	14	+	+	+	+	+	+	-
C/D	Italy, France, The Netherlands, Germany	32	+	+	+	-	+	+	-
D/C	Pasteur	1	-	+	+	-	-	-	+
D/C	Italy	1	+	+	+	-	-	-	-
D/C	Pasteur, Italy, Japan, France, Brazil, The Netherlands, New Caledonia	21	-	-	-	-	-	-	-

Figure 19 : Résultats des différents profils de plasmidome obtenus sur 110 extraits d'ADN positifs à différentes toxines de *C. botulinum* groupe III.

Source : Woudstra et al., 2018.

### 3.2.5.3. Conclusion

Cette étude a mis en lumière une différence entre les plasmidomes des souches de type C/D et ceux des souches de type D/C. Les EGM recherchés sont plus présents sur les souches de type C/D, au contraire des souches de type D/C qui ne présentent majoritairement pas ces EGM. Comme évoqué précédemment, une des hypothèses avancées dans l'étude serait la présence d'un système CRISPR-Cas fonctionnel pour les souches de type D/C. En effet, dans une étude antérieure, une analyse des régions CRISPR-cas a révélé que les souches de type D/C possèdent jusqu'à trois systèmes CRISPR-Cas complets (donc théoriquement fonctionnels) tandis que les souches de type C/D présentent un système partiel, donc possiblement inactif (Woudstra et al., 2016). Cette approche présente un inconvénient majeur : la stabilité de ces EGM, en particulier en conditions de laboratoire. Il est connu que le prophage p1 est instable en conditions de laboratoire, la stabilité des autres EGM est inconnue. La stabilité de l'élément ciblé par une méthode de typage est un critère essentiel pour développer une méthode de typage (van Belkum et al., 2007), cet aspect n'a pas été abordé dans cette étude et devra être pris en compte dans le futur sur un ou plusieurs EGM sont retenus pour mettre au point un schéma de typage.

### 3.2.6. PCR ciblant les systèmes CRISPR-Cas

Les généralités sur les systèmes CRISPR-Cas ont déjà été présentées précédemment (Cf III.2.5). Des méthodes de typage basées sur les répétitions et les spacers composant le locus CRISPR ont déjà été développées pour certaines espèces ou genres bactériens telles que *Mycobacterium tuberculosis* (Kamerbeek et al., 1997), *Campylobacter* (Kovanen et al., 2014), *Salmonella* (Xie et al., 2017) ou *Clostridioides difficile* (Andersen et al., 2016).

La méthode, nommée initialement spoligotypage, a été développée pour la première fois en 1997 pour le typage de *Mycobacterium tuberculosis*. Un seul système CRISPR-Cas avec des répétitions conservées est retrouvé dans le génome de cette espèce. Un nombre très limité de spacers a été identifié et répertorié (une cinquantaine recensée sur plus de 1000 souches) (Jansen et al., 2002). Ceux-ci sont utilisés pour classer les souches en fonction des spacers identifiés. Chez *Salmonella*, 2 systèmes CRISPR-Cas ont été détectés dans l'ensemble des souches (Fabre et al., 2012; Fricke et al., 2011; Jansen et al., 2002). L'utilisation des spacers pour discriminer les souches a montré que cette méthode est tout aussi fiable et discriminante que d'autres outils de typage déjà utilisés (PFGE ou MLVA). Enfin, chez *Campylobacter jejuni*, il a été montré que cette méthode de typage n'est pas adaptée car, après analyse de 184 souches, 26% d'entre elles ne pouvaient pas être typées car il n'y avait pas de locus CRISPR ou seulement une répétition présente (Schouls et al., 2003). Il a par contre

été proposé d'utiliser cette méthode en seconde intention lorsque des souches présentent des profils identiques avec les outils MLST ou ALFP, méthodes de référence pour le typage de *Campylobacter jejuni*. Une étude plus récente a confirmé que l'analyse des régions CRISPR est pertinente pour comparer des ST sur des isolats qui présentent des liens épidémiologiques (Kovanen et al., 2014).

La présence d'un ou plusieurs systèmes CRISPR-Cas dans la plupart des génomes de *C. botulinum* du groupe III étudiés dans une précédente étude soulève la question de l'utilisation de ces systèmes pour développer un outil de typage (Negahdaripour et al., 2017).

### **3.3. Autres méthodes de caractérisation**

#### **3.3.1. Analyse comparative du génome complet**

Le développement des outils de séquençage de deuxième et troisième génération a permis de rendre accessible à un faible coût le séquençage des génomes des souches bactériennes. Avec l'augmentation du nombre de génomes séquencés disponibles dans les bases de données publiques, la comparaison des génomes entre eux permet d'avoir un aperçu plus complet des différences génétiques présentes entre des souches d'une même espèce, approche plus exhaustive que celles ciblant spécifiquement certaines parties du génome. Les études utilisant l'analyse des génomes complets de *C. botulinum* et *C. novyi sensu lato* en tant qu'outil de typage sont présentées dans ce chapitre.

##### **3.3.1.1. Comparaison par Alignement fragmenté des génomes**

Le premier outil bio-informatique qui a été utilisé pour l'analyse des génomes de *C. novyi sensu lato* est la comparaison des génomes par alignement fragmenté proposé par l'outil Gegenees (Ågren et al., 2012). Le principe est de fragmenter le génome entier ou des régions spécifiques (EGM par exemple) en petits fragments d'environ 200 pb puis d'utiliser un algorithme BLAST (Basic Local Alignment Search Tool) pour les aligner avec les fragments des autres génomes entiers ou régions sélectionnées pour la comparaison. Une matrice de similarité sous forme de « heat-map » est alors calculée avec l'ensemble des scores obtenus lors des alignements BLAST. Cet outil a permis de classer *C. novyi sensu lato* en plusieurs lignées lors de l'étude des génomes de 24 souches comprenant les trois espèces du groupe mais aussi de comparer et grouper les différents EGM caractérisés lors de l'étude de ces génomes (Figures 12 et 20) (Skarin & Segerman, 2014). Cet outil a été utilisé dans l'ensemble des publications qui portent sur la comparaison génomique de *C. novyi sensu lato* (Fillo et al., 2021; Skarin & Segerman, 2014; Woudstra et al., 2017; Woudstra et al., 2016; Woudstra et al., 2021).

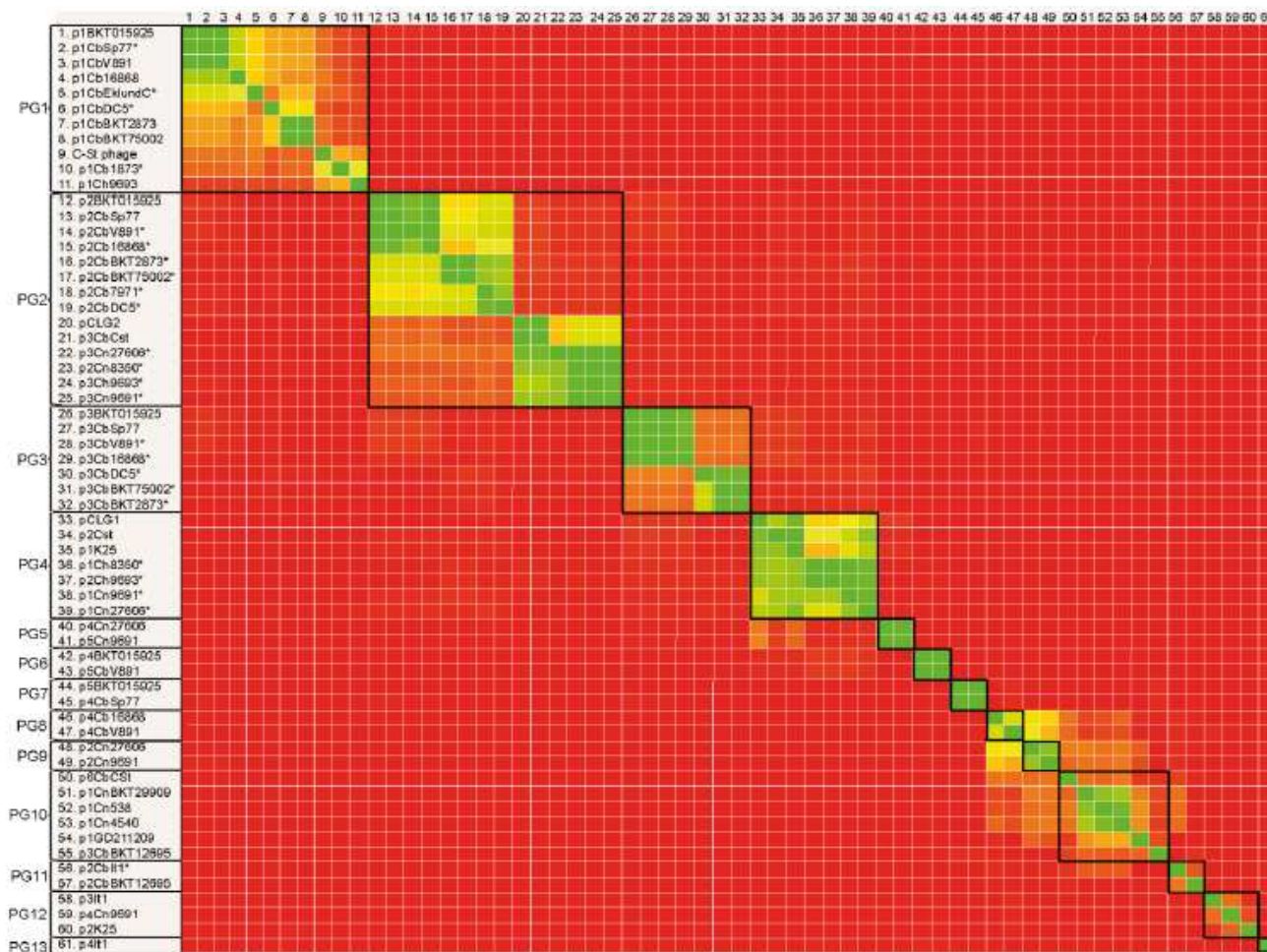


Figure 20 : Matrice de similarités des 61 éléments génétiques mobiles caractérisés chez *C. novyi sensu lato*. Treize groupes de plasmides ont été identifiés grâce à ces résultats. Les EGM avec une étoile sont incomplets.

Source : Skarin et al., 2014

### 3.3.1.2. Analyse des Single Nucleotide Polymorphism (SNP)

Un single nucleotide polymorphism, ou SNP, est une variation d'une seule paire de base dans le génome de deux souches de la même espèce. L'analyse comparative des génomes en se basant sur les SNP s'est révélée être très discriminante pour des souches de *C. botulinum* du groupe I, plus discriminante que d'autres méthodes de typage (PFGE et MLST) qui n'ont pas permis leur discrimination (Gonzalez-Escalona et al., 2014). L'utilisation des SNP pour comparer des souches de *C. botulinum* du groupe III provenant d'épisodes de botulisme animal survenus en France et d'une autre étude (Skarin & Segerman, 2014) a permis de les différencier en trois groupes distincts (Figure 21a) (Woudstra et al., 2016). Les souches arborant le gène codant la neurotoxine de type C/D forment un cluster différent de ceux présentant le gène de la neurotoxine D/C, ce qui est en concordance avec les résultats obtenus lors de l'utilisation de l'outil Gegenees lors de la définition des lignées (Lignée

Ia pour la plupart des souches C/D et Ib pour D/C) (Skarin & Segerman, 2014). Le cluster intégrant les souches de type C/D est séparé en deux groupes. Un de ces groupes comprend la souche de référence BKT015925 provenant de Suède, la souche SP77 provenant d'Espagne et 11 souches françaises. Ces souches, très proches génétiquement, ont fait l'objet d'une analyse séparée pour mieux les différencier (Figure 21b). Cette nouvelle analyse a permis de distinguer les 13 souches. Cependant, aucune corrélation n'a pu être mise en évidence entre la distance phylogénétique de ces 13 souches et leur répartition géographique, leur origine ou la date de l'épisode. Encore une fois, l'hypothèse d'une dissémination en Europe d'une souche clonale de type C/D a été suggérée lors de cette étude, renforçant les résultats déjà obtenus avec la PFGE (Anza et al., 2014a).

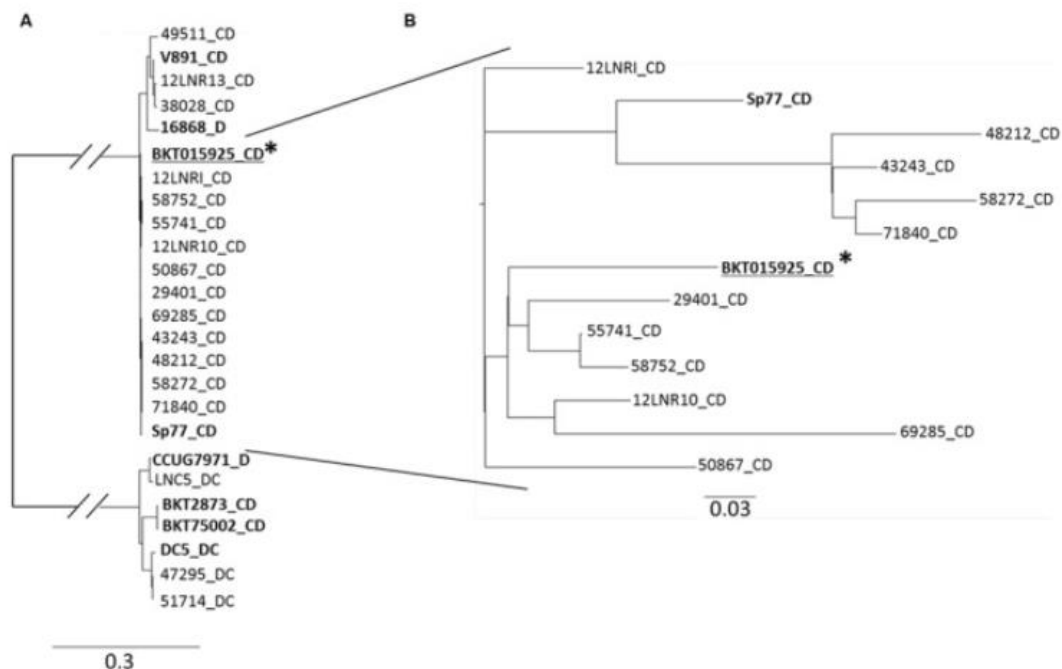


Figure 21 : A) Arbre phylogénétique obtenu après analyse des SNP de plusieurs génomes de *C. botulinum* du groupe III. Les souches en gras proviennent de l'étude Skarin et Segerman., 2014. La souche BKT015925 annotée avec une étoile sert de référence pour l'analyse des SNP.

B) Deuxième arbre regroupant uniquement le groupe de 13 souches de type C/D très proches phylogénétiquement dans l'arbre A.

Source : Woudstra et al., 2016

### **3.3.2. Matrix Assisted Laser Desorption Ionization - Time of Flight Mass Spectrometry (MALDI-TOF MS)**

#### **3.3.2.1. Généralités sur la spectrométrie de masse**

La spectrométrie de masse est une technique d'analyse permettant de détecter et identifier des protéines et molécules d'intérêt en utilisant le rapport masse sur charge pour les séparer. En 1975, l'identification bactérienne est pour la première fois évoquée comme application possible de la spectrométrie de masse (Anhalt & Fenselau, 1975). L'optimisation de la méthode avec le MALDI-TOF (Matrix Assisted Laser Desorption Ionization - Time of Flight Mass Spectrometry) en fait une méthode d'identification bactérienne très sensible et spécifique (Karas & Hillenkamp, 1988). Elle s'est démocratisée depuis une dizaine d'années et est utilisée dans de nombreux laboratoires d'analyses. En plus d'être peu coûteuse, cette technique d'identification s'avère fiable et plus rapide que les techniques d'identification classiques.

#### **3.3.2.2. Principe**

Le principe de l'identification de genre ou d'espèce d'une colonie par spectrométrie de masse est de déterminer les masses moléculaires de différentes protéines bactériennes, principalement des protéines ribosomales, protéines de choc froid (*Cold Shock Proteins*) et protéines de liaison à l'ADN [203]. Les molécules et protéines vont d'abord être ionisées puis elles vont être séparées dans un analyseur en fonction de leur rapport masse/charge ( $m/z$ ). Selon l'instrument utilisé, la source d'ionisation et l'analyseur ne seront pas les mêmes.

L'instrument MALDI-TOF est un spectromètre de masse qui va coupler une source d'ionisation laser assistée par une matrice (MALDI) et un analyseur de temps de vol (TOF) afin de séparer et identifier les protéines d'une bactérie. Après extraction des protéines, l'extrait est mélangé à une matrice puis soumis à l'action d'un faisceau laser afin d'être vaporisé, formant ainsi des ions de masses différentes sous forme gazeuse. Les ions vont ensuite être accélérés par un champ électrique afin d'être séparés en fonction de leur masse moléculaire en passant par un tube qui va permettre la mesure du temps de vol (en anglais Time Of Flight ou TOF). La vitesse des ions va dépendre du rapport masse sur charge ( $m/z$ ), le temps de vol dans le tube est donc déterminé afin de mesurer ce ratio pour chaque protéine. Une fois l'échantillon analysé, une empreinte spectrale est obtenue. Les empreintes spectrales sont variables entre les espèces avec des pics spécifiques du genre, de l'espèce et même de la sous-espèce. L'empreinte spectrale est ensuite comparée à une base de données afin d'obtenir un score qui, selon sa valeur, va permettre de confirmer ou infirmer le genre et l'espèce de la souche testée.

### 3.3.2.3. Utilisation du MALDI-TOF pour l'identification et le typage de *C. novyi sensu lato*

Pour *C. botulinum*, la spectrométrie de masse a été utilisée en première intention comme alternative au test sur souris pour la détection des toxines botuliques via le développement de l'Endopep-MS (Anses, 2021).

La spectrométrie de masse a également été évaluée en tant que méthode d'identification et pour typer les souches. Dans une première étude, l'utilisation du MALDI-TOF pour l'identification bactérienne s'est révélée fiable pour le genre *Clostridium* où une étude a permis de différencier les empreintes spectrales de 31 espèces différentes de ce genre dont *C. botulinum* du groupe III, *C. novyi* et *C. haemolyticum* (Grosse-Herrenthey et al., 2008; Sandrin et al., 2013). Le dendrogramme construit avec les résultats de cette étude montre des similarités avec le dendrogramme construit avec la comparaison des séquences des gènes ARNr 16S (Grosse-Herrenthey et al., 2008).

Une autre étude se concentrant principalement sur *C. novyi sensu lato* et son typage a été réalisée sur 81 souches de *C. botulinum* du groupe III provenant d'épisodes de botulisme animal. Ces souches ont été soumises à une analyse sur MALDI-TOF puis, après obtention des empreintes spectrales, comparées aux empreintes spectrales d'autres clostridies déjà présentes dans les bases de données (Figure 22). Les résultats de cette étude montrent que les souches de *C. novyi sensu lato* testées ne sont pas distinguables au niveau de l'espèce puisque les souches de référence *C. novyi* ATCC 25758 et *C. haemolyticum* ATCC 9650 se regroupent avec des souches de *C. botulinum* du groupe III. Cependant, elles sont plus éloignées d'autres souches de *C. botulinum* du groupe III, ce qui est en accord avec les résultats obtenus lors de la comparaison génomique de ces espèces qui avait permis la caractérisation des 4 lignées (Bano et al., 2017; Skarin & Segerman, 2014).



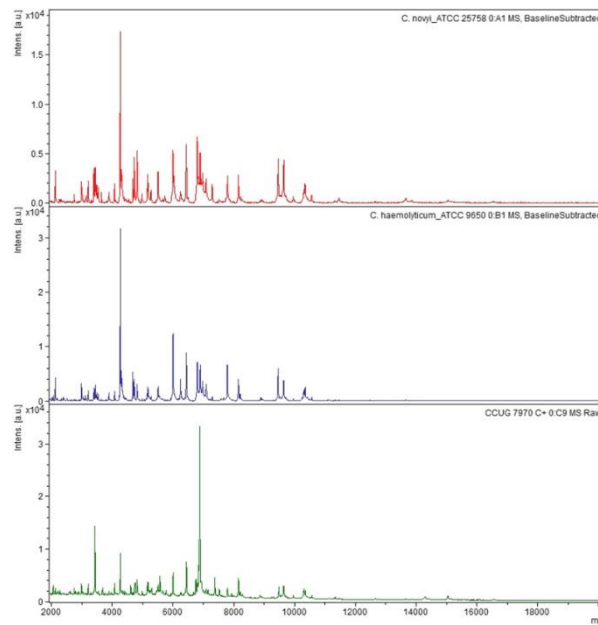


Figure 22 : Empreintes spectrales obtenues suite à l'analyse au MALDI-TOF de trois échantillons de *C. novyi sensu lato* (*C. novyi* ATCC 25758, *C. haemolyticum* ATCC 9650 et *C. botulinum* type C CCUG 7970).

Source : Bano et al., 2017

La comparaison des souches de *C. botulinum* isolées d'épisodes de botulisme animal par MALDI-TOF a permis de les répartir en 14 groupes lors de cette étude. Les neuf premiers groupes (I à IX) contiennent des souches de *C. botulinum* produisant les toxines C/D et D/C. Avec le groupe X comprenant les 2 souches de références de *C. novyi* et *C. haemolyticum* et le groupe XI comprenant 2 souches de type C provenant d'un même épisode de botulisme bovin, le cluster I-IX forme un plus gros cluster qui va se distinguer des groupes XII et XIII qui vont être composés uniquement de souches produisant la toxine C provenant de différents épisodes de botulisme bovin. Enfin, le groupe XIV est plus éloigné des autres groupes avec seulement 15,7% de similarité et va comprendre 2 souches de type C provenant du même épisode de botulisme bovin (Bano et al., 2017).

L'ajout de nouvelles empreintes spectrales provenant d'autres souches de *C. novyi sensu lato* permettrait d'améliorer la résolution lors de la comparaison des souches de ce groupe d'espèces. Bien que prometteuse, cette approche nécessite de disposer d'une souche pure. Cependant, si la souche pure est disponible, il semble plus intéressant de la séquencer afin de disposer du génome et d'obtenir ainsi une quantité d'informations plus importante qu'avec le spectre MALDI-TOF.

### 3.4. Conclusion

Différentes approches ont été explorées pour caractériser les souches du groupe *C. novyi sensu lato* (Tableau 5). Les approches testées ces dernières années se basent plutôt sur des outils PCR (gènes de la flagelline, Plasmidome, MLVA, MLST), tenant ainsi compte des difficultés rencontrées pour l'isolement des souches du groupe *C. novyi sensu lato*. Les méthodes plus archaïques comme la PFGE ou l'AFLP ne présentent plus trop d'intérêt car elles sont trop fastidieuses (besoin d'une souche pure, méthodes plus longues) et très peu discriminantes. Les avantages des outils PCR ciblés sont leurs spécificités et le fait qu'ils soient utilisables sur des échantillons bruts, permettant donc de s'affranchir de l'étape d'isolement.

L'augmentation du nombre de génomes dans les bases de données publiques a permis l'utilisation d'outils de typage plus complet et plus discriminant puisqu'ils prennent en compte l'ensemble du génome et peuvent discriminer les souches à la paire de base près (Analyse comparative par fragmentation, SNP). Ces outils sont actuellement les méthodes de typage de référence pour un grand nombre de bactéries pathogènes. Cependant, pour *C. novyi sensu lato*, le nombre de génomes disponibles est encore bien trop faible pour que ces outils soient utilisés en routine dans les laboratoires. La mise en place de protocoles fiables allant de l'isolement des souches à l'assemblage des génomes est donc plus que nécessaire pour pouvoir optimiser ce type d'outils bio-informatiques dans le futur.

Enfin l'outil MALDI-TOF est aujourd'hui une référence pour l'identification bactérienne. En ce qui concerne le typage, le fait qu'il faut avoir une souche pure pour l'utiliser le rend moins intéressant en le comparant à du séquençage de génome complet.

## Objectifs de la thèse

Une recrudescence du botulisme animal a été observée au niveau mondial à la fin des années 2000 sans que les facteurs à l'origine de cette augmentation n'aient été clairement identifiés (Skarin et al., 2013). Au niveau français, une soixantaine de foyers de botulisme animal sont confirmés chaque année (30 en élevages avicoles, 20 en faune sauvage et 10 en élevages bovins sur la période 2009-2019) (Anses, 2021). Cette maladie touche en particulier les volailles, l'avifaune sauvage et les bovins, pouvant entraîner de fortes mortalités et des pertes économiques importantes (Braun, 2006; Souillard et al., 2014). Malgré ce contexte, l'épidémiologie du botulisme animal est encore méconnue et les données d'épidémiologie moléculaire relatives aux souches impliquées dans le botulisme animal sont peu abondantes dans la littérature scientifique (Anses, 2021). Le faible nombre de génomes séquencés disponibles aujourd'hui illustre bien cette problématique (Fillo et al., 2021; Skarin et al., 2011; Skarin & Segerman, 2014; Woudstra et al., 2017; Woudstra et al., 2015b; Woudstra et al., 2021). Cette faible disponibilité de données relatives à l'épidémiologie moléculaire des souches s'explique principalement par le fait que les souches de *C. botulinum* du groupe III sont difficiles à isoler. L'absence de méthode d'isolement fiable, robuste et applicable en routine constitue actuellement un verrou méthodologique à la caractérisation des souches de *C. botulinum* du groupe III et plus globalement à la compréhension de l'épidémiologie du botulisme animal. Différentes techniques de génotypage ont été testées dans le cadre du botulisme animal mais aucune n'a été identifiée jusqu'à présent comme méthode de référence. L'absence d'outils de typage pour caractériser les souches impliquées dans les épisodes de botulisme animal est actuellement un frein pour la compréhension de l'épidémiologie du botulisme, la gestion et la prévention des épisodes. De plus, l'avènement du séquençage des génomes avec la disponibilité à bas coût des technologies de séquençage de deuxième et troisième génération ouvre de nouvelles perspectives mais nécessite de disposer d'une souche pure.

Le premier objectif de ces travaux de thèse vise à développer la méthode d'isolement des souches de *C. botulinum* du groupe III impliquées dans le botulisme animal. Une fois la méthode développée, l'optimisation des étapes permettant le séquençage des souches (extraction d'ADN, préparation des bibliothèques) sera effectuée. La finalité est de proposer une approche applicable en routine, qui permettra de venir en appui aux investigations épidémiologiques menées sur le terrain.

Le second objectif de la thèse est de tester, d'identifier et d'évaluer des méthodes de typage adaptées au contexte particulier du botulisme animal, en particulier des méthodes s'affranchissant de l'étape d'isolement des souches. La méthode MLVA développée en Italie sera évaluée sur des

souches du laboratoire ainsi que sur des échantillons bruts d'épisodes de botulisme animal au cours desquels des enquêtes épidémiologiques ont eu lieu. Les génomes déjà disponibles dans les bases de données publiques seront analysés afin d'identifier des cibles d'intérêt pour le développement d'outils de typage. La finalité est de disposer d'un outil discriminant permettant de tracer les souches impliquées dans les épisodes de botulisme animal.

Ce projet de thèse s'est déroulé au laboratoire de l'Anses de Ploufragan-Plouzané-Niort au sein de l'Unité Hygiène et Qualité des Produits Avicoles et Porcins (HQPAP) et sous Assurance de la Qualité. Un Plan de Management de la Qualité spécifique au projet a été rédigé et appliqué selon les préconisations de la norme ISO 10 006 « Systèmes de management de la qualité - Lignes directrices pour le management de la qualité dans les projets » et de la norme NF EN ISO/IEC 17025 « Exigences générales concernant la compétence des laboratoires d'étalonnages et d'essais ». Deux audits internes et la rédaction d'un Plan Management de la Qualité du projet ont été réalisés pendant le projet de thèse (Annexe 1).

## **Partie I**

### **Isolement des souches de *Clostridium botulinum* du groupe III**

## 1. Contexte de l'étude et présentation de l'article

Comme évoqué dans l'introduction, l'épidémiologie du botulisme animal est encore méconnue malgré l'évolution des outils de biologie moléculaire. Un des facteurs expliquant cette situation est la difficulté à isoler les souches du groupe *C. novyi sensu lato* dont fait partie *C. botulinum* du groupe III. En effet, il n'existe pas de milieu sélectif pour isoler ces bactéries. La présence d'une flore annexe dense dans les prélèvements rend difficile l'identification des colonies en l'absence de milieu sélectif et l'isolement d'une souche pure sans contaminant constitue un réel défi. De plus, le repiquage successif des souches en conditions de laboratoire peut provoquer la perte d'éléments génétiques mobiles dont le bactériophage qui porte les gènes codant pour les différentes toxines botuliques C, D, C/D et D/C. Le faible nombre de génomes de *C. novyi sensu lato* disponibles dans les bases de données publiques illustre bien ce problème d'isolement.

Le premier objectif de ce projet de thèse est de développer et d'optimiser un protocole d'isolement fiable et facile à utiliser pour isoler les souches à partir des échantillons reçus au laboratoire et, à terme, permettre d'augmenter le nombre de génomes. Le protocole développé au cours de cette étude prend en compte la capacité de résistance des spores de *C. novyi sensu lato* à certains protocoles d'extraction d'ADN. Un de ces protocoles est le kit InstaGene™ Matrix (Bio-Rad, Marne-La-Coquette, France) qui se base sur une matrice composée de billes Chelex microscopiques chargées négativement qui vont capter les ions métalliques dont fait partie  $Mg^{2+}$ , cofacteur des enzymes dont certaines sont responsables de la dégradation de l'ADN. La lyse bactérienne est quant à elle provoquée par un choc thermique à 95 °C pendant 8 minutes. A la fin de l'extraction, les spores qui ont résisté à ce protocole vont se retrouver dans le culot d'extraction. Après ensemencement de 100 µl du culot d'extraction sur une gélose non sélective EYA (Egg Yolk Agar), la majeure partie de la flore bactérienne étant éliminée au cours de l'extraction, il était possible d'isoler et identifier des colonies de *C. novyi sensu lato*.

Cette méthode a été testée sur des échantillons contaminés artificiellement avec des spores de *C. botulinum* du groupe III de chaque type toxinique (C, D, C/D et D/C) ainsi que sur 22 enrichissements (échantillons naturellement contaminés) de la collection du LNR botulisme aviaire représentatifs des cas de botulisme aviaire diagnostiqués entre 2014 et 2019 en France. Sur ces 22 enrichissements, l'isolement de *C. novyi sensu lato* a été possible pour 21 enrichissements, dont 14 sans la perte du gène codant pour la toxine botulique.

Cette étude a été valorisée par une publication dans le journal *Toxins* en Janvier 2020 et deux posters scientifiques présentées en ligne lors des Journées scientifiques et doctorales de l'Anses

(JSDA) en 2020 et lors du World One Health Congress (WOHC) qui a eu lieu du 30 Octobre au 3 Novembre 2020 (Annexe 2).

De nouvelles souches ont été isolées au cours du projet de thèse via l'utilisation de ce protocole, en particulier des souches de type D responsables d'épisodes de botulisme dans des élevages de dindes entre 2014 et 2020 ainsi qu'une souche de type C/D responsable de botulisme dans un élevage de dindes. Des essais ont également été menés pour isoler des souches de *C. botulinum* à partir d'autres matrices, comme des échantillons d'environnement ou d'aliment. La méthode développée ici n'a cependant pas permis d'isoler des souches pures à partir des matrices complexes.

Grâce à ces résultats, un protocole a été rédigé et est maintenant disponible pour les activités de recherche et de référence de l'unité HQPAP (Annexe 3). Une dernière étape d'extraction d'ADN a été ajoutée à la fin du protocole d'isolement dans l'optique de séquencer directement le génome de la souche après son isolement.

En facilitant l'étape d'isolement des souches, cette méthode ouvre de nouvelles perspectives pour l'étude de l'épidémiologie du botulisme animal et le développement d'outils de typage, en permettant notamment le séquençage de nouveaux génomes à partir des souches pures isolées via cette nouvelle méthode. Cependant, d'autres étapes doivent encore être optimisées pour obtenir de l'ADN de bonne qualité et des données de séquences exploitables (extraction d'ADN, séquençages Illumina et MinION). L'optimisation de ces étapes est présentée dans la partie séquençage de ce manuscrit.

2. Article n°1

**Article N°1**

Development of An Innovative and Quick Method for  
the Isolation of *Clostridium botulinum* Strains  
Involved in Avian Botulism Outbreaks


**Le Gratiot T**, Poezevara T, Rouxel S, Houard E, Mazuet C, Chemaly M, Le Maréchal C.

Toxins, 2020; 12(1):42. <https://doi.org/10.3390/toxins12010042>



Article

# Development of An Innovative and Quick Method for the Isolation of *Clostridium botulinum* Strains Involved in Avian Botulism Outbreaks

Thibault Le Gratiet <sup>1,2</sup>, Typhaine Poezevara <sup>1</sup>, Sandra Rouxel <sup>1</sup>, Emmanuelle Houard <sup>1</sup>,  
Christelle Mazuet <sup>3</sup>, Marianne Chemaly <sup>1,\*†</sup> and Caroline Le Maréchal <sup>1,\*†</sup> 

<sup>1</sup> Unit of Hygiene and Quality of Poultry and Pork Products, French Agency for Food, Environmental and Occupational Health & Safety (ANSES), BP 53, 22440 Ploufragan, France; thibault.legratiet@anses.fr (T.L.G.); Typhaine.POEZEVARA@anses.fr (T.P.); Sandra.ROUXEL@anses.fr (S.R.); Emmanuelle.HOUARD@anses.fr (E.H.); Marianne.CHEMALY@anses.fr (M.C.)

<sup>2</sup> UFR of Life Sciences and Environment, University of Rennes 1, 35 000 Rennes, France

<sup>3</sup> National Reference Center for Anaerobic Bacteria and Botulism, Institut Pasteur, 25-28 rue du Docteur Roux, 75724 Paris, France; christelle.mazuet@pasteur.fr

\* Correspondence: caroline.lemarechal@anses.fr; Tel.: +33-(2)-96-01-8533; Fax: +33-2-96-01-8538

† Share senior co-authorship.

Received: 20 December 2019; Accepted: 9 January 2020; Published: 10 January 2020



**Abstract:** Avian botulism is a serious neuroparalytic disease mainly caused by a type C/D botulinum neurotoxin produced by *Clostridium botulinum* group III, one of the entwined bacterial species from the *Clostridium novyi sensu lato* genospecies. Its isolation is very challenging due to the absence of selective media and the instability of the phage carrying the gene encoding for the neurotoxin. The present study describes the development of an original method for isolating *C. botulinum* group III strains. Briefly, this method consists of streaking the InstaGene matrix extraction pellet on Egg Yolk Agar plates and then collecting the colonies with lipase and lecithinase activities. Using this approach, it was possible to isolate 21 *C. novyi sensu lato* strains from 22 enrichment broths of avian livers, including 14 toxic strains. This method was successfully used to re-isolate type C, D, C/D, and D/C strains from liver samples spiked with five spores per gram. This method is cheap, user-friendly, and reliable. It can be used to quickly isolate toxic strains involved in avian botulism with a 64% success rate and *C. novyi sensu lato* with a 95% rate. This opens up new perspectives for *C. botulinum* genomic research, which will shed light on the epidemiology of avian botulism.

**Keywords:** *Clostridium botulinum*; isolation; botulism

**Key Contribution:** This study presents a very simple and efficient protocol to isolate strains involved in avian botulism outbreaks. No such method existed previously, preventing large-scale epidemiological investigations from being conducted. This method opens up new avenues for investigating avian botulism outbreaks.

## 1. Introduction

Botulism is a flaccid paralytic disease that can affect both humans and animals. Clinical signs are caused by the action of botulinum neurotoxins (BoNTs) on the neuromuscular junction at peripheral nerves [1,2], blocking the fusion of acetylcholine vesicles and thus preventing transmission of nerve impulses [3]. BoNTs are metalloproteases composed of a heavy chain that mediates binding to receptors and translocation across endosomal membranes and a light chain that cleaves one of the three “Soluble N-ethylmaleimide-Sensitive-Factor Attachment Protein Receptors” (SNARE), namely

vesicle-associated membrane protein (VAMP), synaptosomal-associated protein 25 (SNAP-25), and syntaxin-1A/1B [2,4]. Proteolysis of SNARE by BoNTs prevents the fusion of synaptic vesicles with the neuronal presynaptic plasma membrane leading to the failure of neurotransmission [1].

BoNTs are currently represented by at least seven serotypes and more than 40 subtypes [5]. They are produced during vegetative growth by spore-forming and anaerobic Gram-positive bacteria of the genus *Clostridium* [1], in particular by *Clostridium botulinum*. BoNT-like genes have also been computationally identified in non-clostridial species such as *Weissella*, *Enterococcus*, and *Chryseobacterium* [6,7].

*C. botulinum* is divided into four distinct groups (I to IV). Most outbreaks of animal botulism are caused by BoNT C and D or a chimeric fusion of C and D, i.e., C/D or D/C, produced by *C. botulinum* group III strains [8]. Animal botulism is considered an emerging disease in Europe [9], especially in livestock where an outbreak can induce high mortality rates (up to 100% in a flock), resulting in large economic losses [9,10]. Moreover, considering spore persistence in the environment [11], outbreaks frequently recur in affected flocks [12]. Botulism can also have a huge impact on wildlife species, in which massive outbreaks have been reported worldwide [13] and yearly [10]. In Europe, avian botulism is frequently associated with BoNT C/D [14–16], while bovine botulism is associated with BoNT D/C or C [14,16]. Several factors may explain why BoNT responsible for botulism outbreaks varies among animal species. BoNT C/D is more lethal for chickens than BoNT C, which may explain why BoNT C/D and not BoNT C is related to avian botulism [17]. It has also been suggested that type C and C/D strains could occupy different ecological niches: mammals for BoNT C (mink, horses, cattle) and birds for BoNT C/D [16]. Chickens are considered naturally resistant to type D botulism but can serve as a reservoir [18], while bovines appear to be highly sensitive to BoNT D/C [19].

Despite being reported worldwide in the literature since the early twentieth century, animal botulism has only been partially studied and is still difficult to control. Efforts have been made to develop diagnostic tools [14,20–24] but many aspects of the disease still need to be addressed, notably the epidemiology of animal botulism, which is still poorly understood, and the development of molecular epidemiological tools that are sorely lacking. While epidemiological tools are available for *C. botulinum* group I [25,26], few have been developed for *C. botulinum* group III. One of the reasons behind this situation is the difficulty in isolating *C. botulinum* group III and the absence of any consensual and reliable isolation method for this pathogen.

*C. botulinum* group III grow poorly in mixed culture and on agar plates. It is an obligatory anaerobic bacteria and there are no known traits suitable for selective cultivation [27]. Another problem in isolating toxigenic strains is the unstable lysogeny of the phage carrying the gene encoding for BoNT C, D, C/D, and D/C [28,29] that is easily lost under laboratory conditions [14], especially after several transfers in culture media [30]. The mechanisms behind the interactions between the phage carrying the *bont* gene and *C. botulinum* group III or *Clostridium novyi sensu lato* and factors inducing its loss are currently unknown.

The conventional method used to isolate *C. botulinum* group III consists of repeatedly culturing the sample in pre-reduced Tryptone, Peptone, Glucose Yeast extract (TPGY) [27] or Fortified-Cooked Meat Medium (F-CMM) [30–32] and then streaking the enrichment broth on non-selective media such as Egg Yolk Agar (EYA) [32], Blood Agar Base No.2 (BAB2) with 5% defibrinated sheep blood and 2.5% agarose [32], or a McClung Toabe Agar plate [27]. Colonies positive for lipase (an iridescent sheen is visible on the surface of the colonies) and lecithinase (revealed as an opaque precipitate surrounding the colony) have to be collected and screened individually to detect the presence of BoNT genes or the colony's ability to produce BoNTs. Lecithinase activity varies between strains, however, and some isolates have been reported as negative for lecithinase activity [33]. A positive lipase reaction alone does not distinguish *C. botulinum* GIII from other *C. botulinum* types or from other *Clostridium* species [30]. Moreover, it is extremely difficult to distinguish toxigenic from nontoxigenic strains, although it is highly recommended to collect the smallest colonies [32]. This method is time-consuming, requires specific skills and lots of practice, and a simpler and more efficient isolation method is much needed [27]. Anza et al. recently developed a method based on immunomagnetic separation using

beads coated with antibodies targeting *C. botulinum* group III strains that allowed the isolation of both toxigenic and nontoxigenic strains. This approach, however, involves extra costs and ethical concerns for the production of antibodies. Alternative strategies are needed for the isolation of strains involved in avian botulism outbreaks.

*C. botulinum* group III is closely related to *Clostridium novyi* and *Clostridium haemolyticum* [26,34]. These bacterial species are historically defined by their ability to cause a certain disease i.e., botulism, black disease, and bacillary hemoglobinuria, respectively. Genomic and mass spectrometric comparison of these three bacterial species have shown that they are very similar or even indistinguishable [16,29,34]. The “genospecies” name of *C. novyi sensu lato* was proposed to include these entwined species [29]. The differences between them are related to their plasmid content and phages—notably, genes encoding for toxins—which upon horizontal movement can transfer disease-causing properties [34]. *Bont* gene translocation from a *C. botulinum* group III strain to a *C. novyi* strain, for instance, has already been demonstrated [31]. BoNT is responsible for the clinical signs observed in botulism. The usual objective when studying botulism is therefore, quite reasonably, to isolate BoNT-producing strains. However, considering the very close relatedness of toxic and nontoxic strains (i.e., strains with or without the *bont* gene), it would appear beneficial to isolate *C. novyi sensu lato* strains instead of focusing only on *C. botulinum* group III when studying animal botulism. This strategy was applied by Anza et al. who considered both toxic and non-toxic *C. novyi sensu lato* during the isolation of clostridial strains from samples collected from different botulism outbreaks [35].

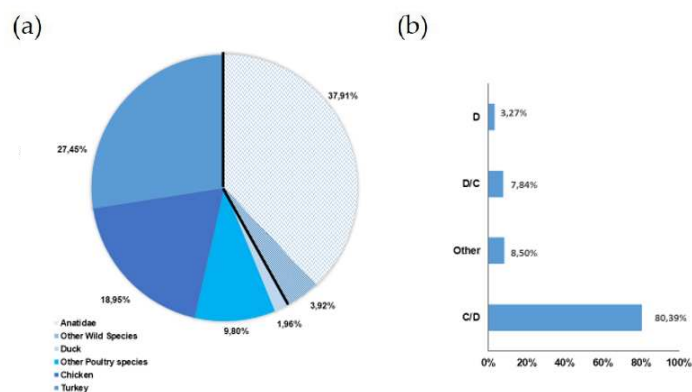
This article presents a new user-friendly, quick, and reliable method for isolating both *C. botulinum* group III and *C. novyi sensu lato* strains involved in avian botulism outbreaks. This method opens new perspectives for the exploration of many topics, including the monitoring of strain dissemination or, more generally, strain tracking.

## 2. Results

### 2.1. Selection of Samples for Evaluating the Isolation Protocol

Between April 2013 and October 2019, 153 samples were confirmed as being positive for *Clostridium botulinum* group III by the French National Reference Laboratory (NRL) for avian botulism.

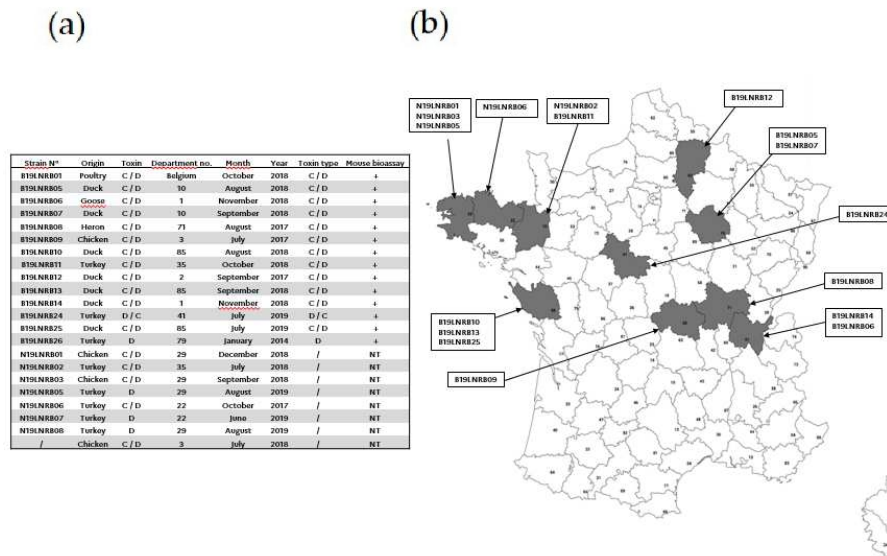
As described in Figure 1a, 58.16% of avian botulism outbreaks concerned poultry production and 41.84% concerned wild avian species. *C. botulinum* group III mainly affects three avian families or species: 37.91% were wild Anatidae, 27.45% were turkeys, and 18.95% were chickens. The remaining 15.68% were other avian species either from poultry production (11.76%) or the wildlife sector (3.92%).



**Figure 1.** Overview of avian botulism outbreaks diagnosed by the French National Reference Laboratory (NRL) between 2013 and 2019. (a) Distribution of avian species (poultry species (58.2%) are represented by colors and wild birds (41.8%) by symbols); (b) BoNT types detected in the liver (C/D: detection of the *bont* C/D gene, D/C: detection of *bont* D/C, D: detection of *bont* D, other: simultaneous detection of several *bont* genes (out of C, C/D, D/C, and D)).

As shown in Figure 1b, 80.39% of these samples were of type C/D, 7.84% were of type D/C (only in turkeys), 3.27% were of type D (only in turkeys), and 8.50% indicated the presence of several *bont* genes at the same time. Indeed, some samples were simultaneously positive for more than one type of *C. botulinum* group III. No pure type C outbreak was detected over this period. It is noteworthy that for two outbreaks in wild birds, a signal was detected simultaneously for types C/D and E in 2018. Our study did not evaluate the isolation of *C. botulinum* type E.

Twenty-two positive enrichment broths (Figure 2a) were selected from the samples collected between 2013 and 2019 to represent the outbreaks diagnosed in France over this period, i.e., various BoNT types, various regions (as depicted in Figure 2b), as well as various avian species.

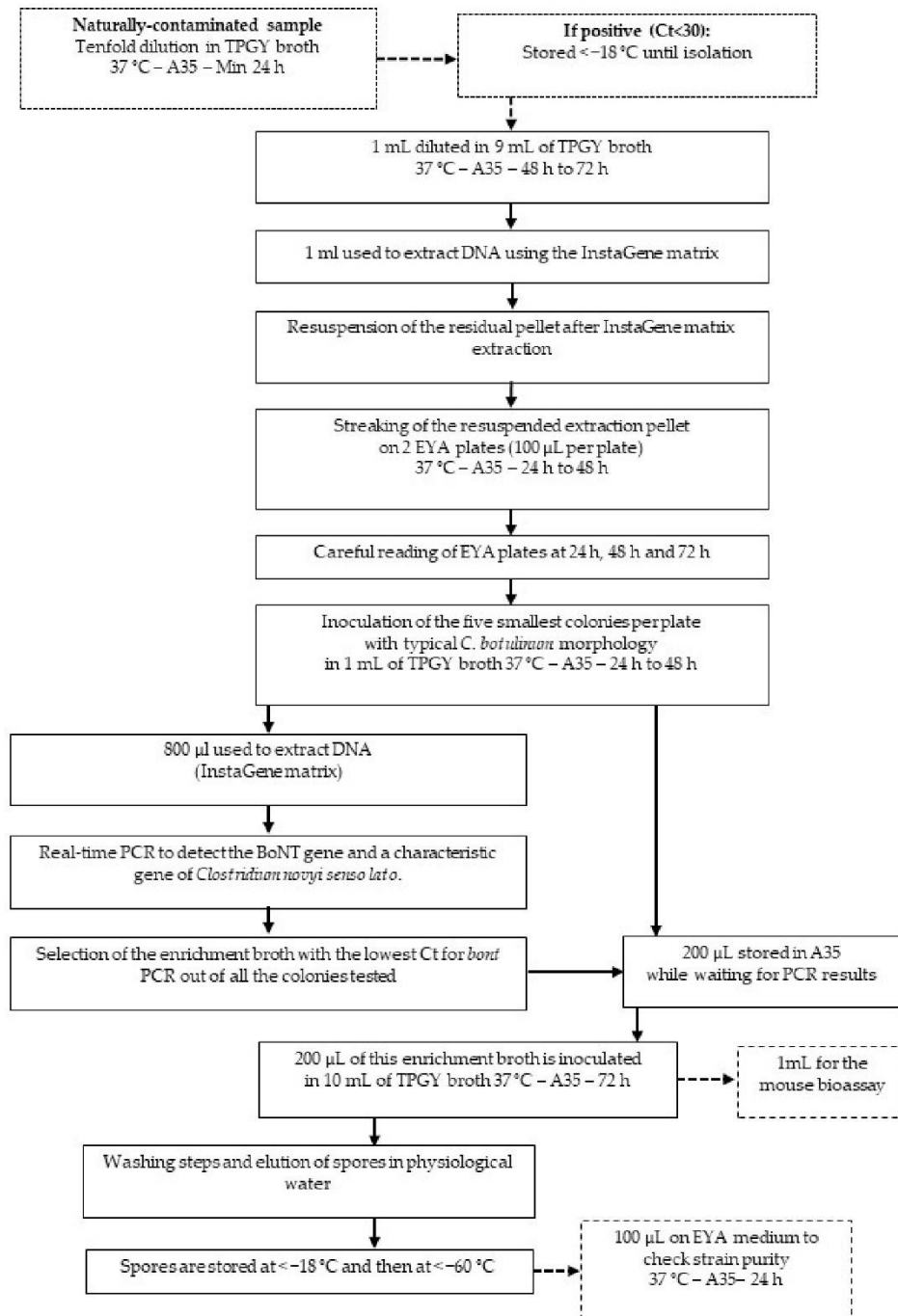


**Figure 2.** Data regarding the enrichment broths tested in our study to evaluate the isolation method. (a) Overview of the data (strain reference number, animal species, BoNT type identified using PCR, number of the *département* (French administrative unit) in which the outbreak occurred, month of the outbreak, year of the outbreak, BoNT type of the strain isolated, results of the mouse bioassay (+: positive, -: negative, NT: Not Tested), (b) Map showing the distribution of the outbreaks and enrichment broths included in this study.

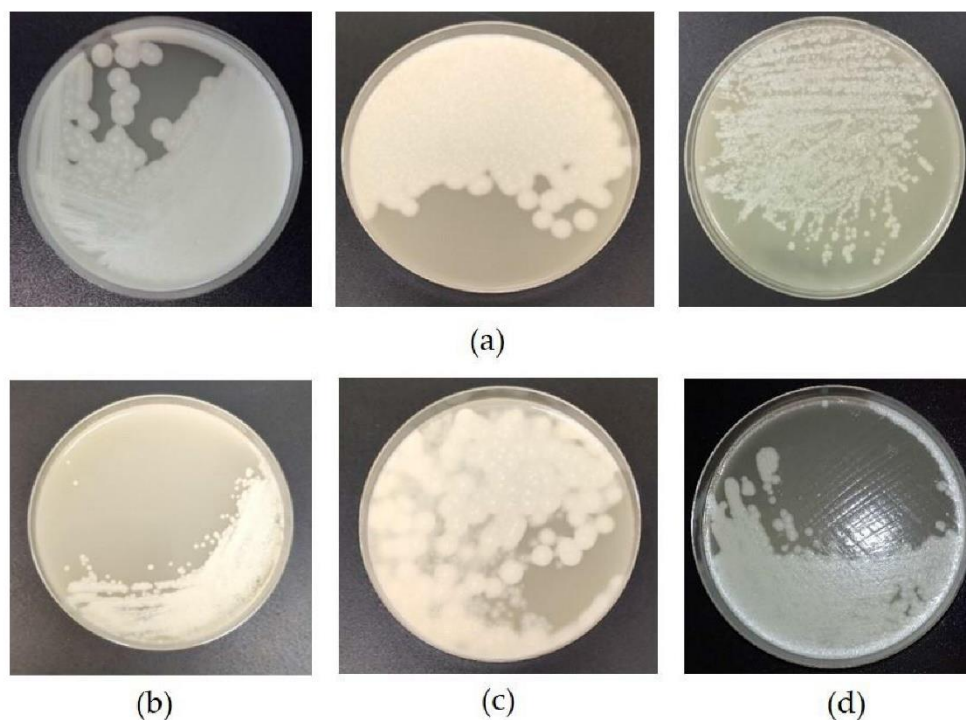
### 2.2. Evaluation of the Method Using Naturally Contaminated Samples

An original protocol was designed to isolate *C. botulinum* group III strains (Figure 3) and evaluated using the 22 naturally contaminated samples listed in Figure 2 and selected as previously explained.

Using this approach, it was possible to isolate toxic or non-toxic *C. novyi sensu lato* strains from 21 enrichment broths (examples of strains plated on EYA during the isolation process are shown in Figure 4) with a success rate of 95%. Fourteen toxic strains (64% of all the strains isolated) were successfully isolated from the tested enrichment broths. The isolates were still toxic after the isolation process by PCR and using the mouse bioassay. The remaining strains (32%) were positive for *C. novyi sensu lato*, showing that the method is appropriate for isolating strains in this clostridial group.



**Figure 3.** Workflow of *Clostridium novyi sensu lato* isolation from samples positive either for type C, C/D, D, D/C *bont* genes using real-time PCR. (A35: anaerobic station A35, Ct: Cycle threshold detected using real-time PCR).



**Figure 4.** Pictures of colonies of a BoNT type C/D (a), D (b), D/C (c) and nontoxic (d) strains plated on Egg Yolk Agar (EYA) after isolation according to the protocol depicted in Figure 3. Note the differences in colony morphology depending on the strain.

### 2.3. Confirmation of the Method Using Spiked Samples

Besides the assay with the method depicted in Figure 3 using naturally contaminated samples, samples spiked with a known initial spore contamination level and of a known BoNT type were used to confirm that the method is indeed appropriate for isolating *C. botulinum* group III strains involved in avian botulism outbreaks. A method had been previously developed and validated for detecting *C. botulinum* in avian liver using PCR with a view of diagnosing avian botulism [24]. Its limit of detection was determined as five spores per gram of liver. In our study, broiler livers were spiked with five spores per gram of *C. botulinum* type C, D, C/D, and D/C. The isolation process depicted in Figure 3 was successfully used to re-isolate the strains used to spike the livers (Table 1). This result proves that this protocol can be applied to isolate *C. botulinum* group III strains, whatever the BoNT type, with an initial concentration as low as five spores per gram of liver.

**Table 1.** PCR detection of colonies isolated from liver samples spiked with five spores of *C. botulinum* group III per gram of liver.

Spore Reference	BoNT Type	<i>C. novyi sensu lato</i> Group Detection after Isolation According to Figure 3	<i>ntnh</i> Detection after Isolation According to Figure 3
CIP 109 785	C	+	+
CIP 105 256	D	+	+
B17LNRB4	C/D	+	+
B19LNRB3	D/C	+	+

### 3. Discussion

It is considered very difficult and thus a real challenge to isolate toxigenic *C. botulinum* group III [14,27,30]. This study presents a new method for the isolation of toxic and non-toxic *C. novyi sensu lato* strains from avian livers tested positive for *bont* gene using real-time PCR after their enrichment in TPGY.

Using the conventional isolation method, namely streaking a broth positive for BoNT-producing *C. botulinum* group III on non-selective agar plates, a high proportion of isolates are reported to be nontoxigenic [27,30]. Franciosa et al. reported the isolation of 11 strains from eight different samples by selecting at least ten colonies per plate, revealing a low rate of toxigenic strains on plates (less than 14%) [30]. In our study, the number of toxigenic colonies (i.e., positive for the *bont* gene using real-time PCR) out of all the colonies collected on plates and positive for *C. novyi sensu lato* varies from 25% to 100% depending on the strains.

When tested on naturally contaminated samples, the approach developed by Anza et al. using beads coated with antibodies managed to isolate 15 strains from 46 samples [35]. All these strains were positive for the *C. novyi sensu lato* gene and only eight were toxigenic [35]. The method we developed, described herein, managed to isolate 21 strains from 22 samples, 14 of these strains being toxigenic. Our cheaper, user-friendly method was thus more effective in isolating toxigenic strains.

Our approach is based on the ability of *C. botulinum* GIII to survive the InstaGene extraction protocol, which seems to eliminate most of the other bacterial contaminants that usually disturb the isolation process. The InstaGene matrix has previously been shown to fail to extract intracellular DNA from *Bacillus cereus* spores [36]. Moreover, the same study demonstrated that the majority of *B. cereus* spores appeared little or unaffected by the InstaGene matrix extraction protocol. Our results seem to indicate that this is also the case for *C. botulinum* GIII, which appears to survive the InstaGene matrix extraction protocol very well.

Liver was the only matrix tested here because it is used for routine confirmation of avian botulism in our laboratory [15,24]. Using the conventional method, isolation was reported to be easier from liver than from caeca [27]. No comparisons between the method's efficiency according to the matrices used were made during this study, but the ability of *C. botulinum* to survive the InstaGene extraction protocol might also make it possible to use other matrices successfully.

The InstaGene matrix is made up of negatively charged microscopic beads that trap the metal ions required as catalysts or cofactors in enzymatic reactions. This prevents extracted DNA being degraded by enzymes, which are blocked by the unavailability of catalysts and cofactors. Cell disruption is performed by incubation at 95 °C for eight min. Few studies have been conducted to evaluate the thermal resistance of *C. botulinum* group III isolates, but the available data show that type C spores of terrestrial strains can resist 101 °C for 14 min and marine strains up to 93 °C for 18 min, although variations in thermal resistance between isolates have been reported [37].

In order to take into account the potential diversity of *C. botulinum* group III strains and evaluate the method's efficiency in isolating strains during outbreaks, a variety of enrichment samples from the collection of the French National Reference Laboratory (NRL) for avian botulism were selected so as to be as representative as possible of the avian botulism outbreaks diagnosed in France. The authors therefore compiled an overview of the avian botulism outbreaks diagnosed since 2013 and the enrichment broths were selected on this basis so as to cover, insofar as possible, the avian species, geographical areas and BoNT types involved. As in most European countries [14,16,35,38], BoNT type C/D (80.39%) was the most frequently detected in avian botulism outbreaks. Type D/C (7.84%), type D (3.27%), and samples positive for several or all primers (C, D, C/D, and D/C) (8.50%) were detected far more scarcely. Type D/C had been reported in studies conducted in Italy, where it had been found either in the intestinal contents of a stork [16] or from coot, duck, or heron samples [38]. All the type D/C-positive samples in our study were detected in turkeys and not in wild birds. Samples positive for all primers had also been reported in some other avian outbreaks [38]. It appears that type D is uncommon in avian botulism outbreaks and has not been reported in other available

studies [16,21,22,38]. No type C BoNTs were detected either in the samples analyzed during our study or in previously published studies that distinguished type C/D from type C BoNTs [14,16].

Poultry and wild birds can both be infected by botulism. Broilers (18.95%) and turkeys (27.45%) were the main poultry species, while mallards and more generally Anatidae (37.91%) were the main wild birds. This result is in accordance with the avian species recorded as being affected by botulism in other countries [10,13,14,16,39] except for turkeys, for which case reports are rare in the literature [40].

#### 4. Conclusions

The isolation method developed in this study managed to successfully isolate 95% of *C. novyi sensu lato* strains, including 64% of toxigenic strains when testing type C/D, D, D/C, and C samples. Moreover, it is quicker, cheaper, and more user-friendly than previous methods.

It provides an efficient, reliable, and easy-to-use protocol for isolating *C. botulinum* group III strains and opens up new avenues to better explore strain diversity, track strains during an outbreak, and improve knowledge of this bacterium's physiology.

#### 5. Materials and Methods

##### 5.1. Samples

An overview of the outbreaks diagnosed by the French NRL for avian botulism was compiled to be able to select representative samples of situations encountered in the field in order to develop and evaluate the method. All outbreaks confirmed by the French NRL (clinical signs and a positive real-time PCR detection in samples collected from animals [24]) between 2013 and 2019 were listed and are presented in Figure 1.

Twenty-two naturally contaminated livers were used in this study and are listed in Table 1. These samples were collected from animals in a context of suspected avian botulism in various regions of France, as well as in Belgium, between July 2017 and June 2019.

According to a previously optimized method [24], livers ( $\leq 25$  g) were ten-fold diluted in TPGY and incubated for at least 24 h at 37 °C in an anaerobic chamber (A35, Don Whitley distributed by Biomérieux, Bruz, France). DNA was extracted using 1 mL of this enrichment broth and BoNT genes were detected using real-time PCR. All the samples used in this study were positive for types C/D, D/C, or D. Enrichment broths were stored at below  $-18^{\circ}\text{C}$  until isolation.

##### 5.2. Confirmation of the Method Using Spiked Samples

Spore solutions from *C. botulinum* type C (strain CIP-109 785, from the collection housed at the Institut Pasteur, Paris, France), type D (strain CIP-105 256, Institut Pasteur, Paris, France), type C/D (B17LNRB4, ANSES, Ploufragan, France), and type D/C (B18LNRB3, ANSES, Ploufragan, France) were prepared and titrated using the five-tube most probable number (MPN) method as previously described [41]. Samples of chicken liver (5 g) that tested negative for the presence of *C. botulinum* using real-time PCR were inoculated with 5 MPN spores of each strain per g of liver. Spiked livers were then processed as described above for naturally contaminated samples.

##### 5.3. InstaGene DNA Extraction

Enrichment broths (1 mL) were gently thawed, ten-fold diluted in TPGY, and incubated for at least 72 h in an anaerobic chamber (A35, Don Whitley distributed by Biomérieux, Bruz, France) at 37 °C. One milliliter of this enrichment broth was then incubated at 80 °C for 15 min to inactivate BoNT. After centrifuging at  $12,850\times g$  for one minute, the pellet was resuspended in 200  $\mu\text{L}$  of InstaGene matrix (Bio-Rad, Marne-La-Coquette, France) and incubated at 56 °C for 30 min then at 95 °C for 8 min. A final centrifugation at  $12,850\times g$  for 3 min was performed to separate microbeads (in the residual pellet) from the DNA extract (in the supernatant).



#### 5.4. Real-Time PCR

The real-time PCR, primers, and probes in this study were used according to Woudstra et al. [42]. Real-time PCR with a Bio-Rad CFX96 thermal cycler (Bio-Rad, Marne-La-Coquette, France) was used to detect the *C. botulinum* BoNT gene and a gene characteristic of *C. novyi sensu lato* [42]. Each assay was performed in a total volume of 20 µL that contained 5 µL DNA template, 10 µL PerfeCTaqPCR ToughMix (Quantabio, distributed by VWR, Fontenay-sous-Bois, France), and a final concentration of 600 nM for primers and 400 nM for probes. The thermal profile consisted of 5 min at 95 °C, followed by 45 cycles of denaturation at 95 °C for 15 s and annealing/elongation at 55 °C for 30 s. Each PCR run included positive and negative controls for each target. A sample was considered positive when the Ct (Cycle threshold) was below 38.

#### 5.5. Isolation on Egg Yolk Agar

The isolation protocol is fully depicted in Figure 3. The pellet obtained after DNA extraction using InstaGene matrix (Bio-Rad, Marne-la-Coquette, France) was homogenized using a vortex and 100 µL per plate was streaked over a pre-reduced Egg Yolk Agar (EYA) medium comprising 4% Blood Agar Base No 2 (Oxoid, Dardilly, France), 1.5% Bacto Agar (VWR, Fontenay-sous-Bois, France), and extemporaneously-prepared 10% egg yolk emulsion. Five colonies per plate positive for lipase and lecithinase activities [27] were individually collected after 24 h, 48 h, and 72 h of incubation at 37 °C in an anaerobic chamber, and cultured each time in 1 mL of TPGY broth at 37 °C in the anaerobic chamber (A35 Don Whitley distributed by Biomérieux, Bruz, France) for at least 24 h. DNA was extracted from 800 µL of each culture using the InstaGene matrix (Bio-Rad, Marne-la-Coquette, France) and the *ntnh* gene and a gene characteristic of *C. novyi sensu lato* [42] was detected using real-time PCR as previously described [24]. The remaining 200 µL was stored in the anaerobic chamber until PCR results were obtained.

#### 5.6. Strain Storage

PCR results were analyzed so as to select the enrichment broth with the lowest Ct (i.e., presenting the greatest homogeneity) for both PCRs. The remaining 200 µL of the colony enrichment broth with the lowest Ct for the BoNT gene (*ntnh*) and the gene characteristic of *C. novyi sensu lato* were incubated in 10 mL of TPGY broth at 37 °C in the anaerobic chamber for at least 72 h. Supernatant was discarded after centrifugation (except 1 mL was stored for the mouse bioassay for BoNT detection) and the pellet was washed three times with physiological water. Finally, the pellet was resuspended in physiological water and stored first at a temperature below −18 °C and then at a temperature below −80 °C. Strain purity was checked by streaking 100 µL of the final suspension on pre-reduced EYA plates. After incubation for 24 h at 37 °C under anaerobic conditions, the EYA plates were examined to evaluate contamination.

#### 5.7. Detection of BoNT Production by Strains Following Isolation

The detection of BoNT in the 24-h enrichment culture of isolated strains was based on the mouse lethality assay. The tests were performed in accordance with European Directive 2010/63/EU on the protection of animals used for scientific purposes (laboratory animal use agreement no. 2013-0116). One mL of enrichment broth was collected, centrifuged, filtered, and five-fold diluted in 50 mM phosphate buffer (pH 6.5) containing 1% gelatin. A volume of 0.5 mL was injected intraperitoneally into Swiss mice weighing 20–22 g (Charles River Laboratories, l'Arbresle, France). The mice were observed for up to 4 days for the presence of typical clinical signs (pinching of the waist, labored breathing, and paresis) and euthanized immediately after observation of such signs.

**Author Contributions:** Conceptualization, T.L.G., C.M., M.C., C.L.M.; methodology, T.L.G., C.M., M.C., C.L.M.; validation, T.L.G., C.M., M.C., C.L.M.; formal analysis, T.L.G., C.M., M.C., C.L.M.; investigation: T.L.G., T.P., S.R., E.H., C.M., M.C., C.L.M.; resources T.L.G., T.P., S.R., E.H., C.M., M.C., C.L.M.; writing—original draft preparation: T.L.G., C.L.M., M.C.; writing—review and editing: T.L.G., T.P., S.R., E.H., C.M., M.C., C.L.M.; supervision, CLM, MC; project administration, C.L.M.; funding acquisition, C.L.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by ANIHWA, grant ANIBOTNET and the Côtes-d’Armor Departmental Council. T.L.G. is the recipient of a PhD grant from ANSES and “Région Bretagne”.

**Acknowledgments:** C.L.M. was the recipient of a short-term mission grant funded by the MedVetNet Association (2016\_STM\_3) for a training period on the isolation of *C. botulinum* group III strains at the IZSve laboratory in Italy under Luca Bano. The authors are grateful to Luca Bano for this initial training and his advice on the isolation process.

**Conflicts of Interest:** The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Tehran, D.A.; Pirazzini, M. Novel Botulinum Neurotoxins: Exploring Underneath the Iceberg Tip. *Toxins* **2018**, *10*, 190. [[CrossRef](#)]
2. Pirazzini, M.; Azarnia Tehran, D.; Zanetti, G.; Rossetto, O.; Montecucco, C. Hsp90 and Thioredoxin-Thioredoxin Reductase enable the catalytic activity of Clostridial neurotoxins inside nerve terminals. *Toxicon* **2018**, *147*, 32–37. [[CrossRef](#)]
3. Lindstrom, M.; Korkeala, H. Laboratory diagnostics of botulism. *Clin. Microbiol. Rev.* **2006**, *19*, 298–314. [[CrossRef](#)]
4. Zhang, S.; Masuyer, G.; Zhang, J.; Shen, Y.; Lundin, D.; Henriksson, L.; Miyashita, S.I.; Martinez-Carranza, M.; Dong, M.; Stenmark, P. Identification and characterization of a novel botulinum neurotoxin. *Nat. Commun.* **2017**, *8*, 14130. [[CrossRef](#)]
5. Peck, M.W.; Smith, T.J.; Anniballi, F.; Austin, J.W.; Bano, L.; Bradshaw, M.; Cuervo, P.; Cheng, L.W.; Derman, Y.; Dorner, B.G.; et al. Historical Perspectives and Guidelines for Botulinum Neurotoxin Subtype Nomenclature. *Toxins* **2017**, *9*, 38. [[CrossRef](#)]
6. Doxey, A.C.; Mansfield, M.J.; Montecucco, C. Discovery of novel bacterial toxins by genomics and computational biology. *Toxicon* **2018**, *147*, 2–12. [[CrossRef](#)]
7. Mansfield, M.J.; Doxey, A.C. Genomic insights into the evolution and ecology of botulinum neurotoxins. *Pathog. Dis.* **2018**, *76*. [[CrossRef](#)]
8. Le Marechal, C.; Woudstra, C.; Fach, P. Botulism. In *Clostridial Diseases in Animals*; Uzal, F.A., Songer, J.G., Prescott, J.F., Popoff, M.R., Eds.; John Wiley & Sons, Inc.: Ames, IA, USA, 2016; pp. 303–330.
9. Skarin, H.; Tevell Åberg, A.; Woudstra, C.; Hansen, T.; Löfström, C.; Koene, M.; Bano, L.; Hedeland, M.; Anniballi, F.; De Medici, D.; et al. The workshop on animal botulism in Europe. *Biosecur. Bioterror.* **2013**, *11*, S183–S190. [[CrossRef](#)] [[PubMed](#)]
10. Ventujol, A.; Decors, A.; Le Maréchal, C.; Toux, J.Y.; Allain, V.; Mazuet, C.; Bayon-Auboyer, M.-H.; Le Bouquin, S.; Souillard, R. Le botulisme aviaire en France: Etude des cas signalés dans la faune sauvage et dans les élevages par deux réseaux de surveillance entre 2000 et 2013. *Epidémiologie et Santé Animale* **2017**, *72*, 85–102.
11. Notermans, S.; Dufrenne, J.; Oosterom, J. Persistence of *Clostridium botulinum* type B on a cattle farm after an outbreak of botulism. *Appl. Environ. Microbiol.* **1981**, *41*, 179–183. [[CrossRef](#)] [[PubMed](#)]
12. Souillard, R.; Woudstra, C.; Le Marechal, C.; Dia, M.; Bayon-Auboyer, M.H.; Chemaly, M.; Fach, P.; Le Bouquin, S. Investigation of *Clostridium botulinum* in commercial poultry farms in France between 2011 and 2013. *Avian Pathol.* **2014**, *43*, 458–464. [[CrossRef](#)] [[PubMed](#)]
13. Rocke, T.E. The global importance of avian botulism. In *Waterbirds Around the World*; G.C. Boere, C.A.G., Stroud, D.A., Eds.; The Stationery Office: Edinburgh, UK, 2006; pp. 422–426.
14. Woudstra, C.; Skarin, H.; Anniballi, F.; Fenicia, L.; Bano, L.; Drigo, I.; Koene, M.; Bayon-Auboyer, M.H.; Buffereau, J.P.; De Medici, D.; et al. Neurotoxin gene profiling of *Clostridium botulinum* types C and D native to different countries within Europe. *Appl. Environ. Microbiol.* **2012**, *78*, 3120–3127. [[CrossRef](#)] [[PubMed](#)]

15. Le Marechal, C.; Ballan, V.; Rouxel, S.; Bayon-Auboyer, M.H.; Baudouard, M.A.; Morvan, H.; Houard, E.; Poezevara, T.; Souillard, R.; Woudstra, C.; et al. Livers provide a reliable matrix for real-time PCR confirmation of avian botulism. *Anaerobe* **2016**, *38*, 7–13. [[CrossRef](#)] [[PubMed](#)]
16. Bano, L.; Drigo, I.; Tonon, E.; Pascoletti, S.; Puiatti, C.; Anniballi, F.; Auricchio, B.; Lista, F.; Montecucco, C.; Agnoletti, F. Identification and characterization of *Clostridium botulinum* group III field strains by matrix-assisted laser desorption-ionization time-of-flight mass spectrometry (MALDI-TOF MS). *Anaerobe* **2017**, *48*, 126–134. [[CrossRef](#)]
17. Takeda, M.; Tsukamoto, K.; Kohda, T.; Matsui, M.; Mukamoto, M.; Kozaki, S. Characterization of the neurotoxin produced by isolates associated with avian botulism. *Avian Dis.* **2005**, *49*, 376–381. [[CrossRef](#)]
18. Popoff, M. Revue sur l'épidémiologie du botulisme bovin en France et analyse de sa relation avec les élevages de volailles. *Revue Scientifique et Technique de l'Office international des Epizooties* **1989**, *8*, 129–145. [[CrossRef](#)]
19. Nakamura, K.; Kohda, T.; Shibata, Y.; Tsukamoto, K.; Arimitsu, H.; Hayashi, M.; Mukamoto, M.; Sasakawa, N.; Kozaki, S. Unique biological activity of botulinum D/C mosaic neurotoxin in murine species. *Infect. Immun.* **2012**, *80*, 2886–2893. [[CrossRef](#)]
20. Bjornstad, K.; Tevell Aberg, A.; Kalb, S.R.; Wang, D.; Barr, J.R.; Bondesson, U.; Hedeland, M. Validation of the Endopep-MS method for qualitative detection of active botulinum neurotoxins in human and chicken serum. *Anal. Bioanal. Chem.* **2014**, *406*, 7149–7161. [[CrossRef](#)]
21. Anniballi, F.; Auricchio, B.; Delibato, E.; Antonacci, M.; De Medici, D.; Fenicia, L. Multiplex real-time PCR SYBR Green for detection and typing of group III *Clostridium botulinum*. *Vet. Microbiol.* **2012**, *154*, 332–338. [[CrossRef](#)]
22. Anniballi, F.; Auricchio, B.; Woudstra, C.; Fach, P.; Fiore, A.; Skarin, H.; Bano, L.; Segerman, B.; Knutsson, R.; De Medici, D. Multiplex real-time PCR for detecting and typing *Clostridium botulinum* group III organisms and their mosaic variants. *Biosecur. Bioterrori.* **2013**, *11*, S207–S214. [[CrossRef](#)]
23. Hansbauer, E.M.; Skiba, M.; Endermann, T.; Weisemann, J.; Stern, D.; Dorner, M.B.; Finkenwirth, F.; Wolf, J.; Luginbuhl, W.; Messelhauser, U.; et al. Detection, differentiation, and identification of botulinum neurotoxin serotypes C, CD, D, and DC by highly specific immunoassays and mass spectrometry. *Analyst* **2016**, *141*, 5281–5297. [[CrossRef](#)] [[PubMed](#)]
24. Le Marechal, C.; Rouxel, S.; Ballan, V.; Houard, E.; Poezevara, T.; Bayon-Auboyer, M.H.; Souillard, R.; Morvan, H.; Baudouard, M.A.; Woudstra, C.; et al. Development and Validation of a New Reliable Method for the Diagnosis of Avian Botulism. *PLoS ONE* **2017**, *12*, e0169640. [[CrossRef](#)] [[PubMed](#)]
25. Anniballi, F.; Fillo, S.; Giordani, F.; Auricchio, B.; Tehran, D.A.; di Stefano, E.; Mandarino, G.; De Medici, D.; Lista, F. Multiple-locus variable number of tandem repeat analysis as a tool for molecular epidemiology of botulism: The Italian experience. *Infect. Genet. Evol.* **2016**, *46*, 28–32. [[CrossRef](#)] [[PubMed](#)]
26. Giordani, F.; Fillo, S.; Anselmo, A.; Palozzi, A.M.; Fortunato, A.; Gentile, B.; Azarnia Tehran, D.; Ciammaruconi, A.; Spagnolo, F.; Pittiglio, V.; et al. Genomic characterization of Italian *Clostridium botulinum* group I strains. *Infect. Genet. Evol.* **2015**, *36*, 62–71. [[CrossRef](#)] [[PubMed](#)]
27. Skarin, H.; Lindberg, A.; Blomqvist, G.; Aspan, A.; Baverud, V. Molecular characterization and comparison of *Clostridium botulinum* type C avian strains. *Avian Pathol.* **2010**, *39*, 511–518. [[CrossRef](#)]
28. Sakaguchi, Y.; Hayashi, T.; Kurokawa, K.; Nakayama, K.; Oshima, K.; Fujinaga, Y.; Ohnishi, M.; Ohtsubo, E.; Hattori, M.; Oguma, K. The genome sequence of *Clostridium botulinum* type C neurotoxin-converting phage and the molecular mechanisms of unstable lysogeny. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 17472–17477. [[CrossRef](#)]
29. Skarin, H.; Hafström, T.; Westerberg, J.; Segerman, B. *Clostridium botulinum* group III: A group with dual identity shaped by plasmids, phages and mobile elements. *BMC Genom.* **2011**, *12*. [[CrossRef](#)]
30. Franciosa, G.; Fenicia, L.; Caldiani, C.; Aureli, P. PCR for detection of *Clostridium botulinum* type C in avian and environmental samples. *J. Clin. Microbiol.* **1996**, *34*, 882–885. [[CrossRef](#)]
31. Eklund, M.W.; Poysky, F.T.; Meyers, J.A.; Pelroy, G.A. Interspecies conversion of *Clostridium botulinum* type C to *Clostridium novyi* type A by bacteriophage. *Science* **1974**, *186*, 456–458. [[CrossRef](#)]
32. Bano, L.; Drigo, I.; Tonon, E.; Berto, G.; Tavella, A.; Woudstra, C.; Capello, K.; Agnoletti, F. Evidence for a natural humoral response in dairy cattle affected by persistent botulism sustained by non-chimeric type C strains. *Anaerobe* **2015**. [[CrossRef](#)]

33. Oguma, K.; Yamaguchi, T.; Sudou, K.; Yokosawa, N.; Fujikawa, Y. Biochemical classification of *Clostridium botulinum* type C and D strains and their nontoxigenic derivatives. *Appl. Environ. Microbiol.* **1986**, *51*, 256–260. [[CrossRef](#)] [[PubMed](#)]
34. Skarin, H.; Segerman, B. Plasmidome interchange between *Clostridium botulinum*, *Clostridium novyi* and *Clostridium haemolyticum* converts strains of independent lineages into distinctly different pathogens. *PLoS ONE* **2014**, *9*, e107777. [[CrossRef](#)] [[PubMed](#)]
35. Anza, I.; Skarin, H.; Vidal, D.; Lindberg, A.; Baverud, V.; Mateo, R. The same clade of *Clostridium botulinum* strains is causing avian botulism in southern and northern Europe. *Anaerobe* **2014**, *26*, 20–23. [[CrossRef](#)] [[PubMed](#)]
36. Brauge, T.; Faille, C.; Inglebert, G.; Dubois, T.; Morieux, P.; Slomianny, C.; Midelet-Bourdin, G. Comparative evaluation of DNA extraction methods for amplification by qPCR of superficial vs intracellular DNA from *Bacillus* spores. *Int. J. Food Microbiol.* **2018**, *266*, 289–294. [[CrossRef](#)] [[PubMed](#)]
37. Segner, W.P.; Schmidt, C.F.; Boltz, J.K. Minimal growth temperature, sodium chloride tolerance, pH sensitivity, and toxin production of marine and terrestrial strains of *Clostridium botulinum* type C. *Appl. Microbiol.* **1971**, *22*, 1025–1029. [[CrossRef](#)] [[PubMed](#)]
38. Woudstra, C.; Le Marechal, C.; Souillard, R.; Anniballi, F.; Auricchio, B.; Bano, L.; Bayon-Auboyer, M.H.; Koene, M.; Mermoud, I.; Brito, R.B.; et al. Investigation of *Clostridium botulinum* group III's mobilome content. *Anaerobe* **2017**, *49*, 71–77. [[CrossRef](#)]
39. Jang, I.; Kang, M.S.; Kim, H.R.; Oh, J.Y.; Lee, J.I.; Lee, H.S.; Kwon, Y.K. Occurrence of avian botulism in Korea during the period from June to September 2012. *Avian Dis.* **2014**, *58*, 666–669. [[CrossRef](#)]
40. Popp, C.; Hauck, R.; Gad, W.; Hafez, H.M. Type C botulism in a commercial Turkey farm: A case report. *Avian Dis.* **2012**, *56*, 760–763. [[CrossRef](#)]
41. Lindberg, A.; Skarin, H.; Knutsson, R.; Blomqvist, G.; Baverud, V. Real-time PCR for *Clostridium botulinum* type C neurotoxin (BoNTC) gene, also covering a chimeric C/D sequence—Application on outbreaks of botulism in poultry. *Vet. Microbiol.* **2010**, *146*, 118–123. [[CrossRef](#)]
42. Woudstra, C.; LeMarechal, C.; Souillard, R.; Bayon-Auboyer, M.H.; Anniballi, F.; Auricchio, B.; De Medici, D.; Bano, L.; Koene, M.; Sansonetti, M.H.; et al. Molecular gene profiling of *Clostridium botulinum* group III and their detection in naturally contaminated samples originating from various European countries. *Appl. Environ. Microbiol.* **2015**. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

## **Partie II**

**Évaluation d'une méthode MLVA comme outil de typage et de traçabilité  
des souches de *C. novyi sensu lato***

## **1. Développement de la méthode MLVA pour la caractérisation des souches de *C. botulinum* du groupe III**

Le développement de la méthode MLVA pour le typage moléculaire de *Clostridium botulinum* du groupe III a été initié lors du projet européen ANIBOTNET (2016-2019). Ce projet, financé par le programme Animal Health and Welfare ERA-Net (ANIHWA, 2012-2016), se déclinait en 5 tâches avec comme objectifs l'amélioration des connaissances relatives aux souches de *C. botulinum* du groupe III et le développement d'outils permettant d'améliorer la surveillance, le contrôle et la prévention du botulisme animal. Un des objectifs de ce projet était de développer et d'évaluer des schémas MLVA et MLST comme outils de typage. Le Centre National de Reference (CNR) italien pour le botulisme basé à l'Istituto Superiore di Sanita (ISS) était en charge du développement de ces outils sur la base de leur expérience précédente de développement d'un outil MLVA pour le typage des souches de *C. botulinum* du groupe I (Fillo et al., 2011). Dans ce paragraphe, les travaux menés par l'équipe de l'ISS pour le développement de la méthode MLVA seront présentés. Ces travaux ont été menés en amont de mon projet de thèse.

### **1.1. Identification des régions VNTR d'intérêt et design des amorces**

La première étape du développement de la méthode MLVA est la recherche de régions VNTR d'intérêt à partir d'un génome complet de référence de *C. botulinum* du groupe III. A partir de l'analyse du génome de la souche de type C/D BKT015925 et à l'aide de la base de données des répétitions en tandem (<http://minisatellites.i2bc.paris-saclay.fr/>), 27 régions VNTR ont été identifiées. Sur ces 27 régions VNTR, seules 9 ont été sélectionnées pour faire partie du schéma MLVA du fait de leur pouvoir discriminant (Tableau 7a). Ces 9 régions sont situées sur le chromosome du génome de référence. Les amorces et conditions PCR spécifiques pour chaque région VNTR ont ensuite été optimisées (Tableau 7a).

### **1.2. Évaluation de la méthode MLVA**

Des systèmes PCR ciblant ces 9 VNTR ont été mis au point puis ont ensuite été évaluées sur un panel de 65 souches de *C. novyi sensu lato* (62 souches de *C. botulinum* du groupe III dont 4 de type C, 2 de type D, 11 de type D/C, 40 de type C/D et 5 dépourvus du gène de la toxine, 2 souches de *C. novyi* dont un de type A et un de type B et une souche de *C. haemolyticum*). Les souches pures ont été isolées à partir de différentes matrices pour être le plus représentatif possible de la diversité de ce groupe et de différents pays. Le nombre de répétitions en tandem a ensuite été déterminés pour chaque région VNTR après amplification par PCR en point final suivie par une étape d'électrophorèse capillaire en utilisant le kit Agilent DNA 12 000 (Agilent, Les Ulis, France) et l'appareil Agilent 2

100 Bioanalyzer (Agilent, Les Ulis, France). Un séquençage des amplicons a également été effectué pour valider que la taille du fragment d'ADN amplifié corresponde bien au nombre de répétitions présentes déterminées à partir de l'électrophorèse capillaire. Pour chaque échantillon, un code MLVA à 9 chiffres correspondant aux nombres de répétitions de chaque VNTR a ainsi été obtenu (Auricchio et al., 2019). Les codes ont ensuite été importés dans le logiciel BioNumerics 7.5, où chaque nombre de répétitions d'un VNTR est rentré comme valeur de caractère, et un dendrogramme a été calculé en utilisant un coefficient catégoriel et un algorithme de regroupement UPGMA (Unweighted pair group method with arithmetic mean) (Figure 23) (Auricchio et al., 2019). Une table de correspondance entre la taille de l'amplicon déterminée par l'électrophorèse capillaire et le nombre de répétitions effectif en séquençage a été élaborée de façon à pouvoir déduire rapidement le code MLVA de chaque souche testée sans passer de manière systématique par l'étape de séquençage (Tableau 7b). Cette étape reste cependant nécessaire dès lors que la taille déterminée via électrophorèse capillaire ne correspond pas aux valeurs attendues et indiquées dans ce tableau.

Tableau 7 : a) Liste des amorces pour l'amplification des 9 régions VNTR de la méthode MLVA développée pour le typage des souches de *Clostridium botulinum* du groupe III. b) Tableau de correspondance des tailles des amplicons selon le nombre de répétitions pour chaque VNTR étudié.

Source: Standard Operating Procedure on MLVA for *C. botulinum* group III, Anniballi et al

VNTR	Amorce sens (5'-3')	Amorce antisens (5'-3')	Séquence de la répétition (5'-3')
MLVA_ANB_14	CATATTTCCGCAGCCGTT	TGTGTTGAAGGCATGAATCC	GCAAGGATTCATGCTTTCAAC
MLVA_ANB_15	TGGATCATGTGAAAATGCAAG	CAGCTCCTGCCTTCGCTA	AAGTAATAAAA
MLVA_ANB_16	GAAGCTGTAGGGCGTAGACAA	TGTCCTGGCATCAATTCTGA	AAAGTAAAGTAAAGGAAAGT
MLVA_ANB_17	TTTCTTCGTGCAATCCCTCT	GAAAATGCAAGGATGCTAAGA	ATCTTTTTCTTT
MLVA_ANB_18	TGTGAACCTGCCAACCT	GGTGCACATTGACAACAGC	CATCATAAA
MLVA_ANB_19	TTATCCGGGGTTCCCATT	CAACTTGATACCGGCCCC	GAAGAAATA
MLVA_ANB_20	GGACGCAACCTCAGAGAATG	ACTTCTTCGTCCATAGCTCCA	GTTTAAATAATGAAG
MLVA_ANB_23	GGTTTAGATGCTAGAGGGTTTGA	AACGTGCTTCAAAGGCAGT	TTTTAATATCCTTTA
MLVA_ANB_24	CTGTTTCTCCCCTTTGTGTCA	GGCGTTAAAGGAAGCGGTAG	CTTTAAAGCTTTACTAAA

VNTR	Taille du VNTR (pb)	Taille de l'amplicon (pb)						
		1 répétition	2 répétitions	3 répétitions	4 répétitions	5 répétitions	6 répétitions	7 répétitions
MLVA_ANB_14	21	108	129 (118-139)	150 (140-161)	171 (162-180)	192	213 (203-224)	234 (236-245)
MLVA_ANB_15	11	208 (202-213)	219 (215-226)	230 (227-235)	241	252	263	274
MLVA_ANB_16	20	222 (212-230)	242 (240-254)	262 (256-270)	282	302 (290-311)	322	342
MLVA_ANB_17	12	222	234 (230-240)	246 (242-250)	258 (254-263)	270	282	294
MLVA_ANB_18	9	164	173 (171-177)	182 (179-186)	191 (189-195)	200 (196-204)	209	218
MLVA_ANB_19	9	156 (152-160)	165 (161-170)	174 (172-178)	183 (179-186)	192 (188-191)	201	210
MLVA_ANB_20	15	243 (243-250)	258 (251-260)	273	288	303	318	333
MLVA_ANB_23	15	235	250 (244-258)	265 (259-273)	280 (274-282)	295	310	325
MLVA_ANB_24	18	161	179 (178-187)	197 (190-201)	215	233	251	269



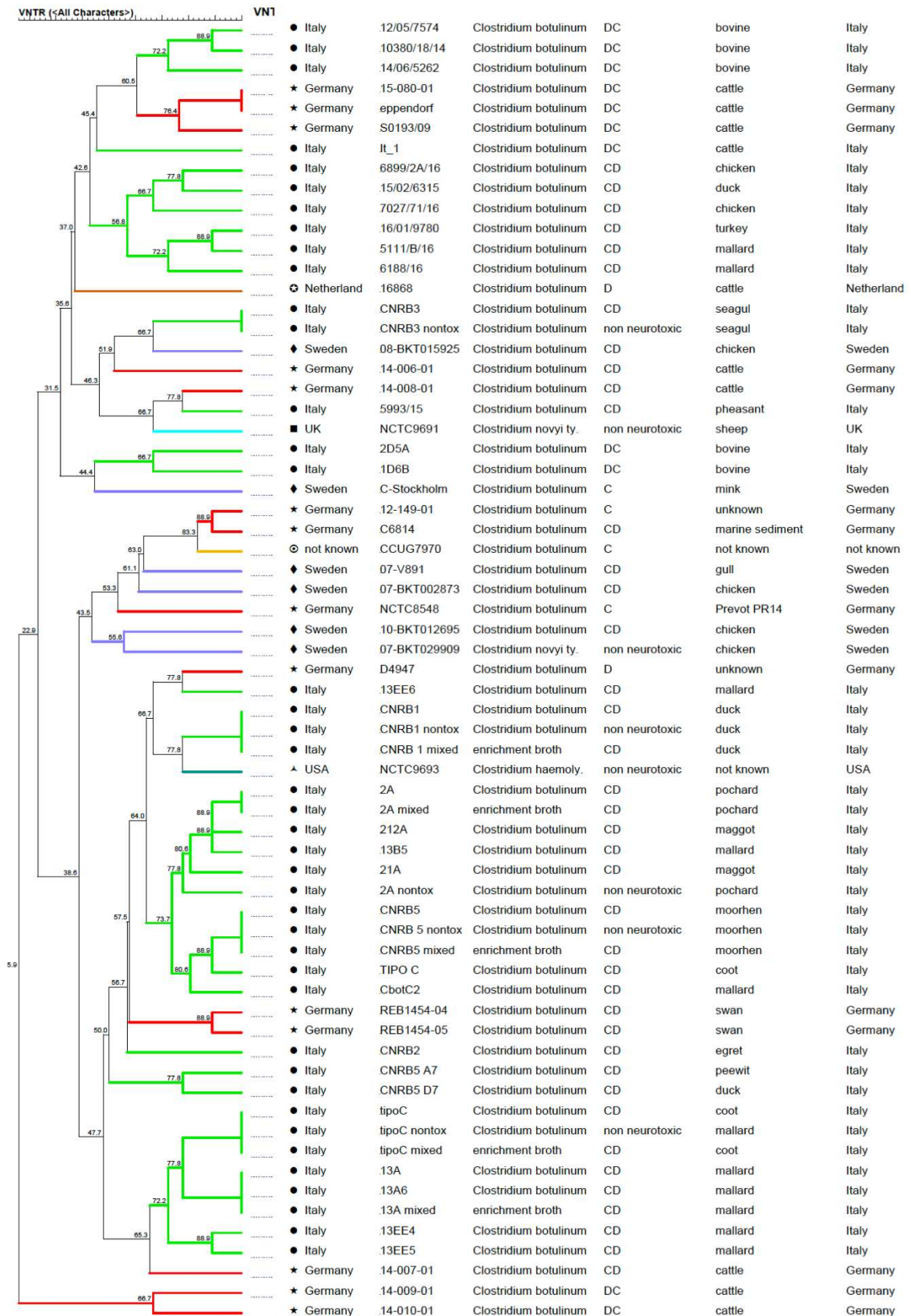


Figure 23 : Dendrogramme calculé à partir des codes MLVA des échantillons testés lors du développement de la méthode à l'ISS.

Source : Auricchio et al., 2019.

La méthode MLVA a été choisie au départ dans le cadre du projet ANIBOTNET car elle permet de travailler non seulement avec des souches pures mais également avec des extraits d'ADN bruts (Chen et al., 2011), ce qui présente un intérêt majeur dans le cas de *C. botulinum* GIII pour lequel l'isolement des souches est une étape compliquée.

Au début de ma thèse, l'ISS avait donc développé cette méthode MLVA, l'avait testé sur une collection de souches pures représentatives des épisodes de botulisme animal en Europe, sur des extraits d'ADN bruts issus de la collection du LNR français et avait rédigé un protocole mis à disposition des partenaires du projet (Annexe 4).

Ces travaux ont été présentés lors d'un workshop sur le botulisme en Mars 2019 à Maisons-Alfort (Auricchio et al., 2019) mais n'ont pas encore été publiés dans un journal à comité de lecture.

### **1.3. Transfert de la méthode au laboratoire de Ploufragan-Plouzané-Niort**

Au cours de mon projet de thèse, un séjour de recherche à l'ISS de Rome a été organisé du 23 Septembre au 2 Octobre 2019 pour effectuer une formation sur la méthode MLVA au CNR italien. Lors de cette formation, les différentes étapes de la MLVA ont été passées en revue (Enrichissement, extraction d'ADN, PCR, électrophorèse capillaire) ainsi que le traitement des données sur le logiciel BioNumerics et la construction d'un dendrogramme. Ce séjour a fait l'objet d'une bourse de mobilité attribuée par l'association MEDVETNET et d'une seconde bourse de mobilité attribuée par l'école doctorale EGAAL.

Compte tenu de la réglementation MOT et des difficultés découlant de cette réglementation pour les échanges de souches ou d'ADN, il n'a pas été envisagé de réaliser un essai de comparaison entre les laboratoires à partir du même panel de souches. Le LNR et l'ISS ont très peu de souches en commun dans leur collection. Sans cet aspect réglementaire, le transfert de méthode aurait été fait via la comparaison des résultats obtenus pour un panel de souches et aurait grandement simplifié la mise en place de la méthode au laboratoire.

## **2. Évaluation de la méthode MLVA sur des échantillons du LNR botulisme aviaire**

### **2.1. Changement de kit d'électrophorèse capillaire**

Des premiers tests ont été effectués dans le cadre de mon projet de thèse lors du transfert de la méthode au sein du LNR botulisme aviaire avec le kit d'électrophorèse capillaire Agilent DNA 12 000 (Agilent, Les Ulis, France). Ce kit est recommandé pour la séparation des fragments ayant une taille comprise entre 100 et 12 000 paires de base avec un taux d'erreur d'environ 15 %. Les tailles

attendues pour les amplicons des 9 régions VNTR sont comprises entre 150 et 300 paires de base. Le kit Agilent DNA 1000 (Agilent, Les Ulis, France) est recommandé pour des fragments d'ADN d'une taille située entre 25 et 1000 paires de base avec un taux d'erreur d'environ 10 %. Ce kit apparaît donc plus adapté compte tenu de la taille attendue des amplicons ciblées dans notre schéma MLVA et a donc été retenu pour réduire la marge d'erreur dans la suite des essais.

## **2.2. Évaluation de la méthode sur différents types de souches de *C. novyi sensu lato***

La méthode MLVA a été testée sur des souches isolées au cours de la thèse avec le nouveau protocole d'isolement dans le but de l'évaluer dans les conditions du laboratoire (équipements présents dans l'unité HQPAP, réactifs, effet manipulateur). Au total, 20 souches isolées au cours la thèse ont été testées (15 de type C/D, 4 de type D et 1 de type D/C). Cette sélection de souches est représentative des épisodes de botulisme aviaire en France entre 2014 et 2019 (Le Gratiet et al., 2020). Après amplification par PCR des 9 régions VNTR, les amplicons ont ensuite été séquencés pour vérifier que le nombre de répétitions déterminées par électrophorèse capillaire correspondait bien au nombre réel de répétitions présentes dans le génome.

Malgré le nombre faible de souches testées, les résultats de typage des 20 souches sélectionnées montrent une répartition des souches en fonction de leur type toxinique (Figure 1 de l'article présenté au JRA page 131). Ce résultat concorde avec ceux obtenus lors d'analyses génomiques qui aboutissent globalement à la répartition des souches en lignée en fonction de leurs types toxiniques (avec des exceptions pouvant être expliquées par des transferts horizontaux de gènes, dont celui codant pour la toxine botulique présent dans un phage (Skarin & Segerman, 2014)).

La reproductibilité est un des critères à évaluer lors du développement d'une méthode de typage (van Belkum et al., 2007). Les résultats des analyses MLVA obtenus pour les 20 souches testées ont montré des amplifications non reproductibles d'un essai à l'autre et un effet laboratoire pour les VNTR 20 et 23. L'effet « laboratoire » a pu être mis en évidence via la représentation des résultats dans un dendrogramme. Sous l'influence des résultats obtenus pour les VNTR 20 et 23, les souches se regroupaient par laboratoire. Ces deux cibles ont donc été éliminées.

Pour les souches de type C/D françaises testées ici, une amplification a été détectée pour les 7 régions VNTR. Les pourcentages de similarité entre ces souches de type C/D sont compris entre 52,2% et 100%, avec des profils MLVA proches de celui du génome de référence de type C/D BKT015925. Aucun lien entre les profils MLVA des souches françaises et leur origine géographique n'a pu être mis en évidence. Un regroupement des souches selon le pays d'origine semble cependant observé, les souches provenant d'Italie étant groupées séparément des souches provenant de France.

Cependant, l'année d'isolement des souches italiennes n'étant pas précisée, aucune conclusion ne peut être tirée sur cette tendance. De plus, les quelques souches en provenance d'Allemagne ou de Suède s'intercalent avec les souches françaises et italiennes, montrant la nécessité d'inclure un plus grand nombre de souches originaires d'autres pays pour pouvoir tirer des conclusions.

Pour les souches de type D/C françaises, seuls 3 ou 4 régions VNTR ont pu être amplifiées sur les 7 analysées. L'absence d'amplification peut être expliquée par l'absence de la région VNTR dans le génome de ces souches ou la survenue d'une ou plusieurs mutations au niveau de la région ciblée par les amorces. Le génome servant de référence au développement du schéma MLVA provenant d'une sous-lignée différente de ces souches (sous-lignée Ia pour BKT015925, les souches de type D/C sont majoritairement de la sous-lignée Ib), ceci pourrait expliquer pourquoi certaines régions n'amplifient pas correctement. Cependant, peu de souches de ce type toxinique ont été testées dans l'étude. En analysant les génomes de type D/C séquencés en 2021 (Fillo et al., 2021), il s'est avéré que le VNTR19 qui ne présente pas d'amplification dans notre étude est bien présent sur les génomes. Cependant, plusieurs mutations au niveau de la séquence ciblée par l'amorce sont observées, ce qui valide l'hypothèse d'une absence d'amplification due à des amorces non adaptées à ces souches. Le VNTR17 n'est quant à lui pas présent sur les génomes de type D/C analysés. La méthode MLVA actuelle basée sur les 7 VNTR est donc plutôt adaptée aux souches de type C/D et moins au type D/C.

Enfin, pour les souches de type D testées au cours de mon projet de thèse, il n'y avait aucune amplification pour les 7 régions VNTR analysées. Ces résultats suggèrent que ces souches sont très éloignées génétiquement de la souche utilisée pour le développement de la méthode et que ces 7 cibles VNTR ne sont sans doute pas adaptées pour la caractérisation de ces souches de type D. Ces souches ont été impliquées dans des épisodes de botulisme dans des élevages de dindes en France entre 2014 et 2020 (7 épisodes au total). L'émergence de ce type d'épisodes interpelle, d'autant plus que la littérature ne rapporte aucun épisode de botulisme de type D pour cette espèce animale. Des études complémentaires sont donc nécessaires pour mieux caractériser ces souches de type D associées à des épisodes de botulisme dans des élevages de dindes.

Ces résultats ont été présentées sous la forme d'une communication orale et d'un article lors des Journées de la Recherche Avicole et Palmipèdes à foie gras (JRA) qui se sont déroulées le 9 et 10 Mars 2022 à Tours.

### **2.3. Application de la méthode à l'étude d'épisodes de botulisme animal**

Pour appliquer cet outil de typage et évaluer son intérêt pour l'étude d'épisodes de botulisme animal en conditions réelles, deux épisodes de botulisme animal au cours desquels un suivi

longitudinal de plusieurs mois a été mis en place par le LNR ont été sélectionnés. L'objectif dans ces deux cas était de tracer la contamination dans le temps et d'identifier la source de contamination en travaillant principalement avec des extraits d'ADN bruts.

### **2.3.1. Exemple 1 d'application : Épisode de botulisme de type D/C**

#### **2.3.1.1. Présentation de l'épisode et du suivi dans le temps**

Le premier épisode retenu pour l'évaluation de la MLVA est un épisode de botulisme bovin de type D/C survenu en 2018 dans un élevage mixte comprenant un atelier bovin et un atelier volaille. La description de l'épisode et de l'ensemble des investigations menées sont présentées en détail dans un article scientifique publié en 2021 dans le journal *Frontiers in Microbiology*.

Brièvement, l'épisode a débuté dans l'atelier bovin comprenant 92 vaches laitières. Le diagnostic a été établi sur la base des signes cliniques puis confirmé via la détection de *C. botulinum* de type D/C par PCR en temps réel (Le Marechal, 2019). L'épisode s'est ensuite étendu à l'atelier des génisses, aboutissant à la mort de 97 animaux sur les 117 présents sur l'élevage soit un taux de mortalité de 79 %.

Afin d'identifier la source de contamination de cet épisode, une enquête épidémiologique a été menée dans l'exploitation. Au cours de l'épisode, 40 prélèvements d'environnement ont été collectés et ont permis la mise en évidence de *C. botulinum* type D/C dans 16 d'entre eux (fumier de volaille, certains composants de la ration, le godet du tracteur servant à la manipulation du fumier de volaille et des aliments et dans les bâtiments de stockage de l'aliment, le bac d'équarrissage, le lisier de bovin, la mélangeuse, loquette du bâtiment bovin, le rideau extérieur et de l'intérieur du grillage du bâtiment d'élevage de volaille). Cette première série de résultats a permis d'émettre l'hypothèse suivante : le fumier de volaille positif pour *C. botulinum* type D/C aurait contaminé le godet du tracteur (lui-même détecté positif). Ce dernier a été ensuite utilisé pour préparer la ration des bovins (3 composants de la ration détectés positifs) puis distribués aux animaux ; les refus de ration des vaches laitières ont été donnés aux génisses, expliquant le décalage dans le temps au niveau de l'apparition des signes cliniques. La contamination du fumier de volaille pourrait s'expliquer par un portage sain des animaux de ce lot.

Suite à cet épisode, un chantier de nettoyage et désinfection a été mis en place au niveau de l'atelier volaille. Le premier contrôle a montré une persistance de *C. botulinum* au niveau du bâtiment. Compte tenu du contexte et de ces résultats, un suivi des lots successifs de poulets de chair a été mis

en place dans l'élevage pour évaluer la contamination dans le temps. Ce suivi a permis de mettre en évidence une contamination du matériel du couvoir avant l'entrée des poulets de chair dans l'élevage.

Une enquête épidémiologique a donc été menée ensuite au couvoir et a permis de détecter *C. botulinum* de type D/C dans 6 sur les 58 échantillons prélevés.

#### **2.3.1.2. Utilisation de la MLVA.**

Sur les 278 prélèvements effectués au cours de cette étude épidémiologique, 65 se sont révélés positifs pour *C. botulinum* de type D/C. La chronologie des événements et les résultats de détection ont permis d'émettre des hypothèses au niveau de l'origine de la contamination qui a abouti à l'épisode de botulisme bovin et sur les voies de dissémination et de persistance. Cependant, sans typage des souches, il n'est pas possible de confirmer ces hypothèses. Le schéma MLVA développée à l'ISS a donc été appliqué à cet épisode pour comparer les échantillons positifs.

L'isolement des souches à partir des échantillons positifs n'a pas été possible, les extraits d'ADN bruts ont donc été utilisés. Une sélection a été opérée au niveau de l'ensemble des échantillons afin d'être aussi représentatif que possible de l'ensemble des investigations menées sur le terrain. Un total de treize échantillons a ainsi été retenu. La méthode MLVA a été utilisée selon le protocole établi par le CNR botulisme italien (Standard Operating Procedure MLVA *C. botulinum* Group III, cf Annexe 4). Les analyses ont été réalisées par électrophorèse capillaire et séquençage.

La MLVA a permis de mettre en évidence deux profils différents au sein des treize échantillons indiquant que deux souches différentes ont été détectés au cours de l'investigation. Le fait que 10 échantillons (provenant du fumier de volaille, du contenu du rumen, du godet et du méteil) possèdent le même profil MLVA indique que l'hypothèse la plus probable pour expliquer la contamination est celle d'un transfert de *C. botulinum* depuis le fumier de volailles vers les bovins via un transfert par le godet du tracteur. La présence d'échantillons provenant du couvoir avec le même profil MLVA démontre que la même souche circule dans l'exploitation et dans le couvoir. Les investigations menées ne permettent pas de déterminer si la contamination initiale s'est opérée depuis l'élevage ou depuis le couvoir.

Un poster sur cet épisode et l'utilisation de la MLVA a été présenté lors du congrès international One Health EJP Annual Scientific Meeting 2021 qui a eu lieu du 9 au 11 Juin 2021 (Annexe 2)

## Article N°2

Asymptomatic carriage of *C. botulinum* type D/C in broiler flocks as the source of contamination of a massive botulism outbreak on a dairy cattle farm.

Souillard, R., Grosjean, D., **Le Gratiot, T.**, Poezevara, T., Rouxel, S., Balaine, L., Macé, S., Martin, L., Anniballi, F., Chemaly, M., Le Bouquin, S., & Le Maréchal, C.

Frontiers in microbiology, 2021, 12, 679377. <https://doi.org/10.3389/fmicb.2021.679377>



# Asymptomatic Carriage of *C. botulinum* Type D/C in Broiler Flocks as the Source of Contamination of a Massive Botulism Outbreak on a Dairy Cattle Farm

Rozenn Souillard<sup>1</sup>, Daniel Grosjean<sup>2</sup>, Thibault Le Gratiot<sup>3</sup>, Typhaine Poezevara<sup>3</sup>, Sandra Rouxel<sup>3</sup>, Loïc Balaine<sup>1</sup>, Sabrina Macé<sup>3</sup>, Laure Martin<sup>3</sup>, Fabrizio Anniballi<sup>4</sup>, Marianne Chemaly<sup>3</sup>, Sophie Le Bouquin<sup>1</sup> and Caroline Le Maréchal<sup>3\*</sup>

<sup>1</sup> ANSES, French Agency for Food, Environmental and Occupational Health Safety, Epidemiology, Health and Welfare Unit, Ploufragan, France, <sup>2</sup> DDCSPP de la Meuse, Departmental Authority in Charge of Veterinary Services for Meuse Department, Bar-le-Duc, France, <sup>3</sup> ANSES, French Agency for Food, Environmental and Occupational Health Safety, Hygiene and Quality of Poultry and Pig Products Unit, Ploufragan, France, <sup>4</sup> Department of Veterinary Public Health and Food Safety, National Reference Centre for Botulism, Istituto Superiore di Sanità, Rome, Italy

## OPEN ACCESS

### Edited by:

Javier Carballo,  
University of Vigo, Spain

### Reviewed by:

Krzysztof Skowron,  
Nicolaus Copernicus University  
in Toruń, Poland  
Frederique Pasquali,  
University of Bologna, Italy

### \*Correspondence:

Caroline Le Maréchal  
caroline.lemarechal@anses.fr

### Specialty section:

This article was submitted to  
Food Microbiology,  
a section of the journal  
Frontiers in Microbiology

Received: 11 March 2021

Accepted: 07 June 2021

Published: 29 June 2021

### Citation:

Souillard R, Grosjean D,  
Le Gratiot T, Poezevara T, Rouxel S,  
Balaine L, Macé S, Martin L,  
Anniballi F, Chemaly M, Le Bouquin S  
and Le Maréchal C (2021)  
Asymptomatic Carriage of  
*C. botulinum* Type D/C in Broiler  
Flocks as the Source  
of Contamination of a Massive  
Botulism Outbreak on a Dairy Cattle  
Farm. *Front. Microbiol.* 12:679377.  
doi: 10.3389/fmicb.2021.679377

In winter 2018, a massive type D/C cattle botulism outbreak occurred on a mixed dairy and broiler farm in France. An investigation was conducted based on the hypothesis of asymptomatic carriage in poultry. We set out to identify the source of contamination of the dairy cattle and to monitor the contamination of broilers over time, including the hatchery delivering chicks to the farm. Environmental samples were collected on the farm during the cattle outbreak ( $n = 40$ ), after the outbreak for three successive broiler flocks ( $n = 128$ ), and once in the hatchery delivering the chicks ( $n = 58$ ). These samples were analyzed using real-time PCR after an enrichment step to detect *Clostridium botulinum* type D/C. The results showed contamination in the manure from the broilers raised just before the onset of the cattle outbreak ( $5 + /5$ ), as well as in some of the components of the cattle ration ( $3 + /17$ ). This latter contamination is likely due to the use of the same tractor bucket to remove litter from the poultry house and to prepare the cattle ration on the same day. Contamination monitoring over several months revealed continuous asymptomatic carriage in the broilers ( $4 + /20$  and  $17 + /20$  cloacal swabs in 2 successive flocks), a persistence of *C. botulinum* type D/C in the ventilation system of the poultry house ( $8 + /14$ ), and contamination of the equipment coming from the hatchery used for delivering the chicks ( $3 + /18$ ). Further investigations conducted in the hatchery demonstrated contamination in the hatchery by *C. botulinum* type D/C ( $6 + /58$ ). Comparison of samples using a multilocus variable number tandem repeat analysis showed the same profile for samples collected on broilers, cattle and in the hatchery. This study highlighted the crucial role of the implementation of biosecurity measures in mixed farms to avoid cross-contamination between production units given the potential asymptomatic carriage of poultry. This study also revealed the contamination of the poultry hatchery. Further investigations are required to better understand the role of hatcheries in the epidemiology of animal botulism.

**Keywords:** botulism, cattle, poultry, epidemiology, hatchery, investigation, MLVA, PCR



## INTRODUCTION

Botulism is a severe neurological disease caused by botulinum neurotoxins (BoNT) that prevent the release of acetylcholine at synaptic junctions and result in progressive symmetrical flaccid paralysis of muscles. There are nine different BoNTs (A, B, C, D, E, F, G, H, or H/A, or F/A, X) (Rasetti-Escargueil et al., 2020) and more than 40 subtypes have been described (Peck et al., 2017). Human botulism is a rare disease is mainly caused by BoNTs A, B, E, and to a lesser extent, F (Rasetti-Escargueil et al., 2019). Botulism is more common in animals than in humans and results in high mortality rate, raising significant animal welfare and economic concerns (Annibaldi et al., 2013; Relun et al., 2017; Rasetti-Escargueil et al., 2019). Avian botulism is generally associated with BoNT C/D, whereas bovine botulism is more frequently associated with BoNT D/C, and to a lesser extent, BoNT C (Woudstra et al., 2012; Bano et al., 2017; Le Gratiat et al., 2020). Considering the serious consequences of botulism on bovine and avian species, a better understanding of this disease — particularly in terms of potential mechanisms of transmission — is crucial to improve prevention and management of animal botulism outbreaks in an efficient manner.

Poultry litter has been considered as a major source of contamination for cattle botulism outbreaks via contact or close proximity (Popoff, 1989; Payne et al., 2011; Relun et al., 2017). Cross-contamination from poultry to cattle has been widely reported in the literature (Senturk and Cihan, 2007; Payne et al., 2011; Ramírez-Romero et al., 2014; Relun et al., 2017; Souillard et al., 2017). Poultry is indeed considered as a reservoir and source of amplification of type D *Clostridium botulinum* and its toxin (Popoff and Argente, 1996). Poultry litter can be used as fertilizer, animal bedding or even as feed supplements (Payne et al., 2011). Surprisingly, no data is available on the prevalence of *C. botulinum* or BoNTs in poultry litter or more generally in healthy poultry. Recently, healthy carriage on a poultry farm was suspected as a source of two cattle botulism outbreaks due to a transfer of poultry manure to the cattle farms (Souillard et al., 2017). Consequently, asymptomatic carriage of *C. botulinum* in poultry can occur, which may represent a reservoir of *C. botulinum* (Rasetti-Escargueil et al., 2019). Animals are either resistant to some BoNT types or *C. botulinum* carriage occurs at low bacterial loads in the digestive tract (Rasetti-Escargueil et al., 2019). These low levels may explain the failure to detect *C. botulinum* carriage in previous studies, as they may be below the limit of detection of available methods (Popoff, 1989).

In 2018, a large BoNT type D/C cattle botulism outbreak occurred on a farm with both dairy and broiler production units in eastern France. Based on the hypothesis of asymptomatic *C. botulinum* carriage in poultry, the objectives of this study were (i) to identify the source of cattle botulism contamination using epidemiological investigations and strain tracking and (ii) to monitor the contamination of broilers over time including at the hatchery delivering chicks to the farm.

## MATERIALS AND METHODS

### Case History and Diagnosis of Botulism on the Farm

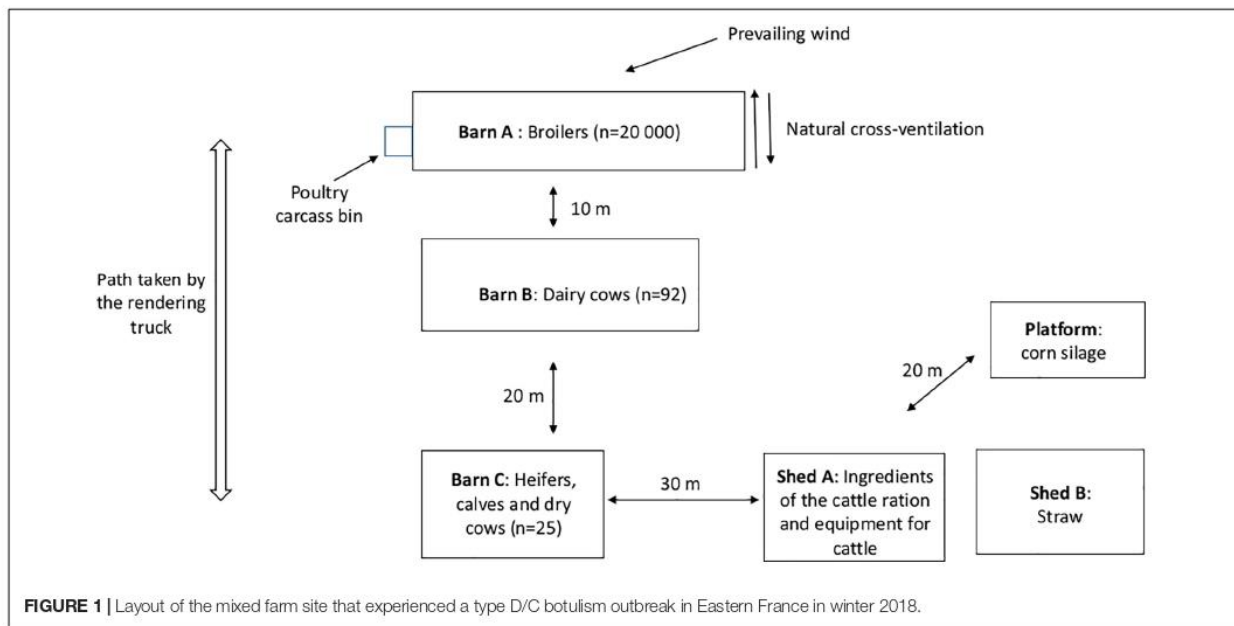
A botulism outbreak occurred in winter 2018 on a farm with dairy and broiler production in Meuse department, located in eastern France. **Figure 1** shows the different barns of the farm. Ninety-two dairy cows were being housed in Barn B and 25 heifers, calves and dry cows in Barn C. These two barns are separated by about 20 m. During winter, cattle are kept indoors and, beginning early spring, grazed on pasture. The Barn A poultry house is separated by about 10 m from Barn B. It is a “Louisiane”-style barn of 1000 m<sup>2</sup> with windows, natural and transversal ventilation, and a dirt floor. In this house, a flock of 20 000 broilers (Flock No. 1) was reared in January 2018. Moreover, the farm has two sheds: Shed A for the storage of the cattle ration ingredients and Shed B for straw storage. In addition, corn silage is also stored on a platform near Shed B. A second site for cattle production is located about 10 km away from this area.

On January 29, the first clinical signs of paralysis were observed in dairy cows in Barn B and mortality started (**Table 1**). On 30 January, samples were collected on two cows that died after showing clinical signs suggestive of botulism. The liver, gall bladder, rectal contents, ruminal contents, feces, and intestines were frozen before shipment to the laboratory. *C. botulinum* types C, D, C/D, and D/C in these samples were screened for as previously described (Le Maréchal et al., 2019). Ruminal contents from the two investigated cows as well as rectal contents from one cow were positive for *C. botulinum* type D/C using real-time PCR after enrichment in trypticase peptone-glucose-yeast extract (TPGY) broth. None of the samples were positive when using fortified-cooked meat medium (F-CMM) with a thermal treatment (70°C, 10 min). On 31 January, signs of paralysis and mortality also occurred in the heifers in Barn C. No clinical signs and no mortality were observed on the second site 10 km away. Vaccination using Ultravac<sup>®</sup> Botulinum (Zoetis, France) was performed on February 2 with a second injection on March 5 in both cattle barns.

Two phases of death were observed on the farm. Mortality was high within the first week (January 29 to February 5) with 68 dead cows in Barn B and 8 heifers in Barn C. A second mortality phase was observed from February 26 to March 5 with 15 cows and 1 heifer. Out of the 117 cows present on the farm at the beginning of the outbreak, 92 cows died, indicating a mortality rate of 79%. All animals with clinical signs eventually died (naturally or humanely euthanized), no recovery was reported.

### Epidemiological Investigation and Collection of Environmental Samples on the Farm

An epidemiological investigation was conducted on the farm. A questionnaire was filled out with the farmer and the veterinarian to describe the history and chronology of the cases. It included questions on the overall management of cattle and poultry (farming practices and manure management, etc.) and



the movements of material, personnel, and vehicles between the poultry and dairy production areas.

During the outbreak, environmental samples ( $n = 40$ ) were collected on the farm on three different days (January 30, February 5, and February 23) to investigate the source of contamination of the cattle (Table 1): 17 samples of the cattle ration stored in Shed A ( $n = 14$ ) and on the platform ( $n = 3$ ), 5 samples of the broiler manure from Flock No. 1 stored in a field since January 26, 3 samples in Barn B, 5 samples in Shed A and 10 samples in Barn A and its surroundings.

After the end of the outbreak, samples ( $n = 128$ ) were also collected for several months in the poultry house to monitor the contamination of broilers over time (Table 2): 9 swabs after cleaning and disinfecting the house with 7 swabs inside the house and 2 swabs outside the house; 55 swabs on the day of chick delivery to the farm, including 11 swabs inside the house; 2 swabs outside the house, 20 cloacal swabs on chicks, 18 swabs on hatchery-origin material, and 4 swabs on the material returned to the hatchery; and finally 64 swabs at the end of the rearing period, with 4 swabs inside the poultry house and 60 cloacal swabs of broilers on the three following broiler flocks (Flocks No. 3, 4, and 5). It was not possible to sample the second flock of broilers reared just after Flock No. 1. Swabs used for sample collection in the farm were obtained from Sodibox (Nevez, France). Cloacal swabs were collected by veterinarians in accordance with the European and French regulation on farmed animal protection.

### Epidemiological Investigation and Collection of Environmental Samples in the Hatchery

An epidemiological investigation was also conducted at the hatchery providing the chicks. A questionnaire was filled

out with the hatchery owner to collect information on the hatchery operations (from egg hatching to the departure of chicks, cleaning and disinfection, waste management, and vehicle movements on the site, biosecurity measures, etc.). A breeder house was also located on a site near the hatchery. At the hatchery site, samples ( $n = 58$ ) were collected (swabs taken on the walls, floor and equipment): 41 inside the hatchery (egg receiving room, incubation and hatcher rooms, chick sorting, and departure rooms), 9 outside the hatchery (truck platform, waste containers, central area and breeder house surroundings) and 8 in annex rooms (trolley storage rooms and refrigerated egg holding rooms).

### Culture Conditions, DNA Extraction, and Real-Time PCR

Enrichment of 226 samples, DNA extraction and PCR were performed as previously described (Le Maréchal et al., 2019).

### Multilocus Variable-Number of Tandem-Repeat Analysis (MLVA)

Thirteen DNA extracts positive for type D/C (with a Ct below 35) were selected for MLVA analysis using nine conventional PCRs (Supplementary Table 1), one for each variable number tandem repeat (VNTR) locus as described previously in Auricchio et al. (2019) and Scalfaro et al. (in preparation)<sup>1</sup>. The MLVA PCR mixture contained 10  $\mu$ L of HotStarTaq Master mix (Qiagen, Courtaboeuf, France), 1  $\mu$ L of forward and reverse primers (10  $\mu$ M for a final concentration of 500 nM), 7  $\mu$ L of water and 1  $\mu$ L of template DNA. Amplification consisted of the following

<sup>1</sup> Scalfaro, C., Vicenza, T., Le Maréchal, C., Bano, L., Bilei, S., Chemaly, M., et al. (in preparation). Multilocus variable-number of tandem-repeat analysis as a tool for *Clostridium botulinum* group III sub-typing. *Microorganisms*

**TABLE 1 |** Chronology of the events on the farm affected by the cattle botulism outbreak type D/C and detection of *C. botulinum* type D/C during the outbreak.

Chronology of the outbreak events	22/01/18	26/01/18	27-28/01/18	29/01/18	30/01/18	31/01/18	02/02/18	05/02/18	23/02/18	05/03/18
	Departure of broilers (Flock No 1 Barn A) for slaughter	Removal of broiler litter using a tractor bucket and storage of manure in a field Transfer of the ingredients of the cattle ration (Shed A) with the same bucket in a mixing wagon to feed cows.	Heifers fed with leftovers from the cattle ration Cleaning and disinfection of the broiler house (Barn A)	Detection of paralysis and first cow mortalities (Barn B)	Diagnosis of botulism: <i>C. botulinum</i> type D/C in ruminal contents (S3) (2 + /2)* and rectal content (1 + /2)*	Paralysis and heifer mortalities (Barn C)	Vaccination of all animals (including calves, dry cows) with Ultravac Botulinum (first injection)			Vaccination with Ultravac Botulinum (second injection)
Detection of <i>C. botulinum</i> type D/Con the farm* (n = 40)					Cattle ration Shed A (n = 3) Wrapped grass 1 + /3			<b>Cattle ration</b> <b>Shed A (n = 11)</b> <u>Meslin (S1)</u> <b>1 + /3</b> Rape <b>1 + /3</b> Brewery grains 0 + /3 Corn 0 + /2 <b>Cattle ration (n = 3)</b> Corn silage 0 + /3 <b>Broiler manure stored since 26/01 (n = 5)</b> <u>Manure (S2) in a field 5 + /5</u>	<b>Barn B (n = 3)</b> Liquid manure <b>1 + /1</b> Swab of the floor of the stall <b>1 + /1</b> Swab of cattle feed table 0 + /1 <b>Shed A (n = 5)</b> Swab of the floor of the shed 0 + /1 Swab of the bucket <b>1 + /1</b> Swab of the tarpaulin covering rape <b>1 + /1</b> Swab of the rape storage area 0 + /1 Swab in the mixing wagon <b>1 + /1</b> <b>Barn A (n = 10)</b> <u>Inside the house</u> Swab of the ventilation system <b>2 + /4</b> Swab of the floor of the house 0 + /1 Swab of the shower? room 0 + /1 Feed from silo 0 + /1 <u>Outside the house</u> Swab of the surroundings 0 + /1 Swab of the carcass bin <b>1 + /1</b> Swab of the path 0 + /1	

\*Number of positive samples/number of samples collected. Samples tested by MLVA are underlined and the sample code in **Figure 2** is indicated in **parentheses**. Samples detected positive are in **bold**.

**TABLE 2** | Detection of *C. botulinum* type D/C in the poultry house after the cattle outbreak ( $n = 129$ ) using swab samples.

After extensive disinfection of the poultry house $n = 9$		Upon chick delivery $n = 55$	Upon broiler departure $n = 64$
Flock No 1*	21/03/18 Inside the house $n = 7$ Ventilation system (S5) <b>2 + /4</b> Floor of the house 0 + /1 Changing room 0 + /1 Feed silo 0 + /1 Carcasses bin (S4) <b>1 + /1</b> Surroundings 0 + /1	22/05/18 Inside the house $n = 7$ Outside the house $n = 2$ Ventilation system <b>2 + /4</b> Floor of the house 0 + /1 Changing room 0 + /1 Feed silo 0 + /1 Surroundings 0 + /1 Carcasses bin 0 + /1	13/06/18 Inside the house $n = 1$ Broilers $n = 20$ Cbcacal swabs <b>4 + /20</b> (S6 and S7)
Flock No 2*		ND	ND
Flock No 3*		22/05/18 Inside the house $n = 7$ Outside the house $n = 2$ Inside the house $n = 2$ Material from the hatchery $n = 2$	13/06/18 Inside the house $n = 1$ Broilers $n = 20$ Cbcacal swabs <b>4 + /20</b> (S6 and S7)
Flock No 4*		9/07/18 Inside the house $n = 2$ Material from the hatchery $n = 2$	6/08/18 Inside the house $n = 3$ Broilers $n = 20$ Cbcacal swabs <b>17 + /20</b> (S11 to S13)
Flock No 5*		27/08/18 Inside the house $n = 2$ Material from the hatchery $n = 16$ Chicks from the hatchery $n = 20$ Material returned to the hatchery $n = 4$	24/09/18 Broilers $n = 20$ Cbcacal swabs 0 + /20

ND, no data.

\*Number of positive samples/number of samples collected (positive samples are in bold). Samples tested by MLVA are underlined and the sample code in Figure 2 is indicated in parentheses.

cycle program: 95°C for 15 min, 35 cycles of 30 s at 94°C, of 1 min at 55°C, of 1 min at 72°C, and one final cycle at 72°C for 5 min.

Multilocus variable-number of tandem-repeat analysis typing was performed on a T100 ThermalCycler (Bio-Rad, Marnes-la-Coquette, France). PCR products were analyzed by capillary electrophoresis using a 2100 bioanalyzer system (Agilent, Les Ulis, France) to determine the number of repeats for each VNTR locus, deduced from each PCR product size. Fragment length and the corresponding number of repeat units was also checked by sequencing on a 3500 Series Genetic Analyzer (Applied Biosystems, Thermo Fischer Scientific, Illkirch-Graffenstaden, France).

The number of repeats obtained from each VNTR locus was imported into Bio-Numerics version 7.5 (Applied Math, Sint-Martens-Latem, Belgium) as character values. A dendrogram was calculated using the categorical coefficient and the UPGMA clustering algorithm to compare DNA extracts from samples tested in our study and available profiles from a previous study (Auricchio et al., 2019) or from other animal botulism outbreaks (Supplementary Table 2).

## RESULTS

### Chronology of the Events in the Mixed Farm

The chronology of the events is given in Table 1. Broiler Flock No. 1 was slaughtered on January 22, one week before the onset of the first clinical signs in cows. During their rearing period, an unusual event took place at 15 days of age: one side of the house collapsed due to a storm and caused the death of 120 broilers. The total mortality of this flock was 923 out of 20 000 broilers (4.6%). Otherwise, nothing unusual was reported regarding this flock nor for previous ones.

On January 26, poultry litter was removed from the house using a tractor bucket and manure was then stored in a field away from the farm (5 km from the farm). The same bucket was used later on the same day (after a quick rinse using a pressure cleaner with cold water) to prepare the cattle ration. Each ingredient of the cattle ration was taken one by one using the bucket and transferred to the mixing wagon to feed the cows later in the evening. Rape was taken first, then wrapped grass and meslin (all stored in Shed A as illustrated in Figure 1), then corn silage and finally brewery grains. Two days later, heifers were fed with the leftover dairy cattle ration. The day following the removal of poultry litter, Barn A was cleaned and disinfected as usual, with lime spread on the floor of the house and in the surroundings and a quaternary ammonium disinfectant sprayed (Spectragen®, Synthèse Élevage, France) in the house.

On January 29, three days after the distribution of the ration, paralysis and the first mortalities were noticed in the dairy cows in Barn B and the diagnosis of botulism type D/C was confirmed the next day. Two days later, on January 31, the same clinical signs occurred in heifers in Barn C.

The same farm personnel and equipment used for poultry and cattle were employed for both production units. Noteworthy, the poultry carcass bin was located next to the changing room of the poultry house and the rendering truck must drive by the cattle barns to reach the poultry house to collect poultry carcasses from the bin (Figure 1).

At the second cattle site 10 km away, no clinical signs were observed. This site was handled by the same farm personnel, using different equipment.

### Detection of *C. botulinum* Type D/C on the Farm

During the outbreak, 40 samples were collected at different times to investigate the source of contamination of cattle on the farm (Table 1). *C. botulinum* type D/C was detected in the five samples collected on February 5 from Flock No. 1 manure stored in a field since January 26 as well as three ingredients of the cattle ration (number of positive samples/number of samples tested, 3 + /17), i.e., meslin, wrapped grass, and rape. Wrapped grass was sampled at the same time as the samples for botulism diagnosis on January 30, whereas meslin and rape were sampled on February 5, several days after the initiation of botulism outbreak on the farm. Investigations conducted on February 23 revealed two positive samples in Barn B (liquid manure and swab of the floor), three positive samples in Shed A (swabs from the bucket, the mixing wagon and a rape tarpaulin) and three positive samples in the poultry house (two swabs from the ventilation system and one swab from the carcass bin). Given this detection of *C. botulinum* type D/C in the poultry house, extensive decontamination was implemented on March 19, with a strong cleaning agent (Decagen® detergent, Synthèse Élevage, France), with disinfection (Virugen®, Synthèse Élevage, France) in the house and quicklime and caustic soda on the floor) and another disinfection using formaldehyde, with both procedures being carried out following the safety measures recommended when using such products.

After the outbreak, the 128 samples collected to monitor the contamination of broilers over time (Table 2) revealed the detection of *C. botulinum* type D/C in the ventilation system of the poultry house (8 + /14) (after the extensive disinfection of the house: 2 + /4; upon chick delivery of the following flocks, 4 + /8; and at the end of the rearing period, 2 + /2), contamination of some equipment coming from the hatchery (3 + /18) (2 swabs on chick box bottom and 1 swab of the egg transport trolleys and trolley wheels) and a continuous healthy carriage of broilers for several months detected in two consecutive flocks (4 + /20 and 17 + /20 cloacal swabs in two flocks and 1 swab of the floor litter). Moreover, the carcass bin was still detected positive after the extensive disinfection of the house.

### Detection of *C. botulinum* Type D/C in the Poultry Hatchery

Detection of *C. botulinum* at the hatchery is presented in Table 3. *C. botulinum* was detected in 6 out of the 58 samples: in the machine that cuts the article chick box, in the container of sludge washing water, on the loading platform, in the storage rooms

**TABLE 3** | Detection of *C. botulinum* type D/C at the hatchery delivering chicks on the farm ( $n = 58$ ).

	Sampling location	Detection of <i>C. botulinum</i> type D/C
Inside the hatchery * $n = 41$	Egg receiving room	0 + /4
	Incubation room	0 + /12
	Transfer (to hatcher) room	0 + /2
	Hatcher room	0 + /11
	Changing room	0 + /3
	Chick sorting room	<b>1 + /5</b> Swab of the machine cutting the article box containing the chicks
	Chick departure room	0 + /4
vOutside the hatchery * $n = 9$	Trucks platforms	<b>1 + /3</b> Swab of the loading platform
	Waste containers	<b>1 + /3</b> Swab of the container of sludge washing water
	Central area	0 + /1
	Breeder house surroundings	0 + /2
Annex room* $n = 8$	Wash room (uncleaned egg trolleys)	<b>1 + /3</b> Swab of the room of the uncleaned egg trolleys (S9)
	Clean room (cleaned egg trolleys)	<b>2 + /3</b> Swab of the room of the cleaned egg trolleys and swab of the trolley wheels (S10)
	Egg holding (cooler) room	0 + /2

\*Number of positive samples/number of samples collected. Samples tested by MLVA are underlined and the sample code in Figure 2 is indicated in parentheses. Samples detected positive are in bold.

of the uncleaned and cleaned egg trolleys, and on the wheels of the cleaned egg trolleys of hatching eggs. In addition, the epidemiological investigation revealed that the rendering truck entering the site to collect hatchery waste (silo of feathers and shells, breeder carcasses) passes hatchery vehicles or equipment (chick or egg delivery trucks, egg trolleys wheeled outside to be stored in annex rooms).

### MLVA Analysis

Thirteen samples positive for *C. botulinum* type D/C were selected among the 291 samples analyzed in our study. Samples were selected so as to obtain an overview of the different investigated areas included in our study and to be as representative as possible of the encountered samples and situations, and finally to compare and evaluate the relatedness of strains detected in the different positive samples. Unfortunately, MLVA results obtained with samples initially detected positive for type D/C with a late Ct (above 35) using real-time PCR were not interpretable (such as for example the swab on the bucket).

Samples selected for MLVA analysis are indicated in Tables 1–3 and detailed number of detected repeats for each VNTR locus in Supplementary Table 3.

A dendrogram was generated based on the VNTR repeat unit profiles (Figure 2). Two different profiles were detected among the selected samples. Three cloacal swabs collected in Flock No. 4 presented a profile different from the other samples (76.8% of similarities between the two groups). These results show that at least two different strains were detected during these investigations with the detection in Flock No. 4 broilers with a profile different from the one detected during the cattle outbreak. However, no differences could be detected between the other investigated samples (Flock No. 1 manure, ruminal contents from a dead cow, cloacal swabs collected in Flock No. 3, samples from the hatchery), demonstrating the presence of the same strain in all these samples.

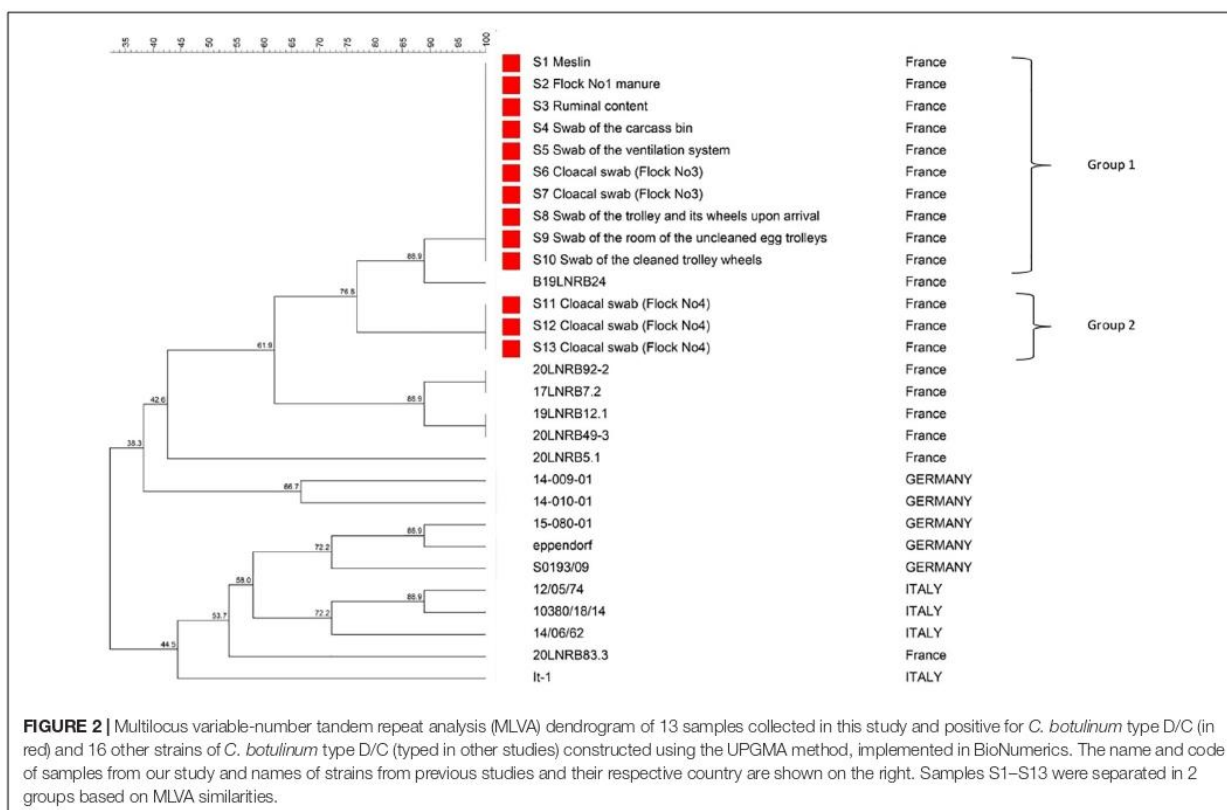
### DISCUSSION

An overview of the detection of *C. botulinum* type D/C on the farm (poultry and cattle barns) and in the hatchery as well as the reference of samples that were compared using MLVA is provided in Figure 3. This figure also presents the most likely scenario of the contamination pathways suggested by our investigations.

### Identification of the Source of Cattle Contamination Through Epidemiological Investigations

Poultry manure has been reported to be a major source for cattle botulism outbreaks (Payne et al., 2011; Ramirez-Romero et al., 2014; Relun et al., 2017; Souillard et al., 2017). Poultry manure is also incriminated in this study, based on the chronology of the outbreak events, and our results showing positive samples (particularly the poultry manure samples) as well as MLVA profiles similar between the manure and the cows that suffered from botulism (Group 1, Figure 2). The positive manure samples from poultry Flock No. 1 (5 + /5) as well as the detection of *C. botulinum* in three samples collected within the poultry house (2 swabs of the ventilation system and 1 swab of the carcass bin) strongly suggest that *C. botulinum* type D/C was present in the poultry house when Flock No. 1 was present.

The tractor bucket used to remove broiler litter from the poultry house and then shortly after used to prepare the cattle ration was also positive for *C. botulinum* type D/C and may have transferred the contamination from poultry manure to the ration ingredients (3 + /17) and to the mixing wagon, which was positive as well. The three ingredients of the ration first taken with the bucket were all detected positive (rape, wrapped grass, and meslin). The MLVA results of the meslin moreover showed the same profile as Flock No. 1 manure and ruminal contents from a dead cow. The succession of these events seems therefore to have resulted in the contamination of dairy cattle fed with this ration, thereby initiating the outbreak. Ration leftovers were then



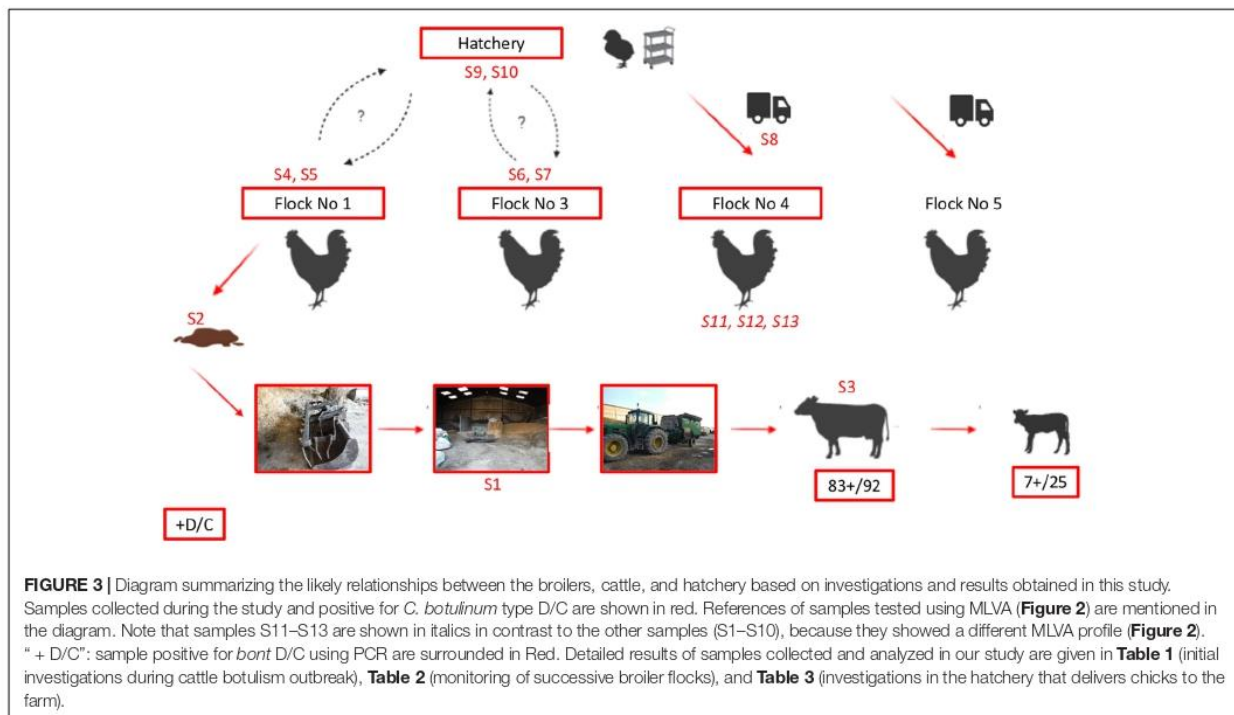
**FIGURE 2 |** Multilocus variable-number tandem repeat analysis (MLVA) dendrogram of 13 samples collected in this study and positive for *C. botulinum* type D/C (in red) and 16 other strains of *C. botulinum* type D/C (typed in other studies) constructed using the UPGMA method, implemented in BioNumerics. The name and code of samples from our study and names of strains from previous studies and their respective country are shown on the right. Samples S1–S13 were separated in 2 groups based on MLVA similarities.

distributed to heifers resulting in the onset of the disease in these animals as well. It is noteworthy that these samples were collected after the outbreak, except for the wrapped grass, therefore cross-contamination secondary to the outbreak cannot be completely ruled out. However, regarding the timing of the events and the analysis results, this seems unlikely. MLVA results also strongly support this scenario, the same profile being detected in Flock No. 1 manure, ruminal contents from a cow, ration ingredients as well as in samples collected in the poultry house after the Flock No. 1 rearing period.

An unusual event that occurred during the Flock No. 1 growth period may explain the manure contamination. The sudden death of 120 broilers due to the collapse of one side of the house during the rearing period may have increased the risk of carcasses being left in the litter despite daily carcass removal by the farmer. Carcasses as a decaying organic matter harboring high amounts of protein substrates and anaerobic conditions is known to support the growth of *C. botulinum* and BoNT production (Annibaldi et al., 2013). Any stress factor in broilers can disturb the balance of the intestinal ecosystem (Tsiouris, 2016); therefore this event may have induced stress in broilers thereby providing favorable conditions for *C. botulinum* growth by modifying intestinal balance. The speed of the onset of clinical signs, 2 days after the distribution of contaminated ration, may be compatible with intoxication of the cattle by ingestion of preformed BoNTs. This hypothesis was supported by the detection of vegetative cells only, and not spores in the ruminal contents of two cows

sampled at the beginning of the outbreak. Ruminal contents were only positive when analyzed using TPGY and not F-CMM, which includes a heat treatment.

In addition to this scenario, several biosecurity failures were also highlighted during the investigation. First, the storage conditions of the cattle ration ingredients were not appropriate, because they were stored uncovered, exposing them to potential contamination by wild birds or rodents. Providing safe, high-quality, and properly stored feed to animals is one of the key measures to minimize the risk of botulism (Annibaldi et al., 2013). Rodents may be a reservoir of a variety of pathogens, in particular *C. botulinum* (Popoff, 1995; Popp et al., 2012; Skarin et al., 2013; Souillard et al., 2014). Moreover, the location of the rendering container (carcass bin), testing positive on two occasions, next to the poultry house can be a source of poultry contamination. Its location also obliges the rendering truck to drive alongside the cattle barns. Given the risk generated by carcasses in regard to botulism and the ability of spores to persist in the environment (Popoff, 1995), the management of the carcasses on a farm is a major critical point for the prevention of the disease. Third, equipment is also shared between the dairy and poultry production units, which can result in cross-contamination between the units. Finally, close proximity of poultry and dairy cows (less than 10 m between both barns with transversal ventilation in the poultry house) may also be a source of *C. botulinum* dissemination via dust and wind. Dust inside the ventilation system in the poultry house is frequently



contaminated after a botulism outbreak (Souillard et al., 2014). Windborne transmission of spore- or BoNT- contaminated material has also been suggested (Hogg et al., 2008).

Since this outbreak, specific measures have been implemented to prevent recurrence of the disease: animals are vaccinated yearly, feed storage area has been reorganized, each component being now separated and protected; the bucket used to load feed in the mixing-wagon is now dedicated only to this activity, the tractor used for poultry manure is no longer used for cows feed; boots and clothes used for poultry and bovines by the farmer are now separated and dedicated to one activity. No new case has been reported since then.

### Monitoring of Broilers Contamination Over Time

Few data are available regarding healthy carriage of *C. botulinum* in poultry and it remains a current issue to better understand the onset of animal botulism (Rasetti-Escargueil et al., 2019). Our monitoring of broilers on this farm using cloacal swabs revealed continuous healthy carriage of *C. botulinum* in successive broiler flocks over several months. After the first detection of *C. botulinum* in the poultry house in January, the following broiler flocks remained healthy carriers of *C. botulinum* at least until August as demonstrated by positive cloacal swabs detected at the end of the rearing period for two flocks.

At least two scenarios can be considered to explain this persistent contamination of broilers over time on the farm. First, our monitoring showed persistence of *C. botulinum* type D/C in the house, particularly in the ventilation system (8 + /14), despite

the disinfection operations conducted between each flock. This detection is likely due to the high resistance of spores in the environment that can survive for several years (Notermans et al., 1981; Wobeser et al., 1987), particularly in the critical areas of the poultry house that are difficult to disinfect, such as the ventilation systems, as reported previously (Souillard et al., 2014, 2016). Consequently, this environmental persistence of *C. botulinum* in the poultry house is a source of recontamination and may have resulted in healthy carriage in successive broiler flocks. A second scenario that may explain *C. botulinum* carriage in broilers for several months is the re-introduction of *C. botulinum* via chick delivery, suggested by the detection of *C. botulinum* on equipment from the hatchery. Two swabs collected on chick box bottoms for Flock No. 4 and one swab collected on trolleys and their wheels for Flock No. 5 were already positive for *C. botulinum* type D/C upon their arrival on the farm, before any contact with the farm. The same MLVA profile was detected for sample S8 (see **Figure 2**) as previous samples collected in the farm (**Table 2** and **Figure 3**). Considering these findings, the question of a potential contamination of the hatchery thus arose and investigations were implemented in the hatchery delivering chicks to the farm to better explore this hypothesis.

### Evaluation of Hatchery Contamination

*Clostridium botulinum* type D/C was detected in 6 of the 58 samples collected in the hatchery: on materials (machine cutting the article box containing the chicks and wheels of trolleys transporting the hatching eggs), in the annex room (room where trolleys of hatching eggs are stored) and on



outside surroundings (in the container of sludge washing water and on the platform where hatching eggs are unloaded). Contaminated areas were either located outside or in areas that are recognized as difficult to clean and disinfect. One sample positive for *C. botulinum* type D/C tested using MLVA showed the same profile (Group 1) as samples collected during the outbreak, demonstrating that the same strain was detected in both places.

During the hatchery visit and based on detection of *C. botulinum* type D/C in some samples, several risk factors were identified. First, similar vehicle routes were highlighted, in which rendering trucks and hatchery vehicles or equipment, for example, pass each other. Failures in cleaning and disinfection operations were also pointed out, as illustrated by the detection of *C. botulinum* type D/C on the system used to cut article boxes during chick delivery or in the room used to store cleaned trolleys or on wheels of cleaned trolleys.

In poultry breeding, the hatchery occupies a central position by being in daily contact with breeder farms to collect eggs and with broiler farms to deliver chicks. Hatcheries can serve as a reservoir and source of pathogenic microorganisms and via the movement of vehicles (delivery trucks), people or equipment (trays, trolleys, and chick boxes, etc.) can facilitate the dissemination of microorganisms (McMullin, 2009; Osman et al., 2018). Consequently, the contamination detected in our study in the hatchery can be also explained by this permanent, exchange of potentially healthy carriers of *C. botulinum* between the hatchery and broiler farms. It was not possible to identify the initial source of contamination, i.e., hatchery or poultry farms, here, because the hatchery was only taken into account after the botulism case investigation had begun.

Despite the hygiene and biosecurity procedures in the hatcheries, infectious agents can contaminate hatcheries by being transported on or within eggs, on hatchery personnel, on trolleys and trays, or as airborne contaminants (Wales and Davies, 2020). A wide range of microorganisms, such as *Salmonella*, *E. coli*, *Pseudomonas*, *Staphylococci*, *Mycoplasma*, or *Aspergillus*, can be detected in hatcheries and disseminated to chicks and subsequently to farms (Qureshi, 2002; McMullin, 2009). To the best of our knowledge, there is currently no data available regarding the contamination of hatcheries by *C. botulinum*. The risks of hatchery contamination arise from the hygiene of hatching eggs, the multiple exchanges between farms, and also the management of the hatchery, involving vehicles, people and equipment all along the process from the arrival of the hatching eggs to the delivery of chicks to customer farms (Qureshi, 2002; McMullin, 2009). In our study, *C. botulinum* type D/C was not detected in any of the samples collected in the incubation ( $n = 12$ ) and hatching rooms ( $n = 11$ ). Moreover, no asymptomatic carriage was detected in chicks delivered on the farm using cloacal swabs on Flock No. 5. The hatching process is considered as the most risky step for microbial dissemination (McMullin, 2009; Kim and Kim, 2010). This does not seem to be the case here based on detection results. No vertical transmission of *C. botulinum* from breeders to chicks was demonstrated here or in previous studies. The exact role of hatcheries in the epidemiology of animal botulism requires further investigation.

## MLVA as a Useful Tool for Epidemiological Investigations

Genotyping methods such as pulsed-field gel electrophoresis (PFGE) and amplified fragment length polymorphism (AFLP) have been successfully used in the context of animal botulism outbreaks for tracking strains (Myllykoski et al., 2009; Skarin et al., 2015). However such approaches require isolation of pure strains, thus limiting the number of samples that can be analyzed. Despite improvements regarding *C. botulinum* group III isolation (Le Gratiet et al., 2020), it is still time-consuming and phages encoding BoNT are easily lost. MLVA has been available for *C. botulinum* group I and II for many years (Fillo et al., 2011; Umeda et al., 2013; Anniballi et al., 2016). It has been recently developed for *C. botulinum* group III subtyping (Auricchio et al., 2019). MLVA presents a major advantage because it does not require the isolation of the strain and can be directly used on DNA extracted from the initial samples (Kahlisch et al., 2010; Chen et al., 2011; Pereyre et al., 2012; Pailhoriès et al., 2015). Our study demonstrates the usefulness of the MLVA approach as a subtyping tool intended for tracing and tracking *C. botulinum* group III, in particular for investigations during animal botulism outbreaks.

In our study, profiles of samples encoded S1 to S10 were shown to be identical (Group 1). Three cloacal swabs collected on broilers from Flock No. 4 presented a different profile from the other samples collected during the study (76.8% similarity between Groups 1 and 2). The origin of this second profile was not explained here. Unfortunately, MLVA results from samples collected at the beginning of the Flock No. 4 rearing period were not interpretable (DNA amount was too low in these samples as shown by a late Ct obtained for these samples). Therefore, it was not possible to determine if this contamination could be linked to the hatchery or not. Samples from the hatchery (samples S8, S9, S10) were indeed all part of Group 1.

## CONCLUSION

This study demonstrates that broilers can be healthy carriers of *C. botulinum* type D/C that can lead to cattle contamination and in the initiation of a botulism outbreak. This also shows that this contamination can last several months, spanning successive flocks. This study also shows that the environment of hatcheries can be contaminated by *C. botulinum* and may become a source of introduction of *C. botulinum* in poultry farms via chick delivery. As illustrated by a massive cattle botulism outbreak due to cross-contamination between poultry and cows, this study highlights the major importance of the implementation of appropriate biosecurity measures in mixed farms to avoid cross-contamination between the production units involving equipment (specific material or disinfection), vehicles (separate travel routes) and personnel (shoes and clothes in shower room). Further investigations are now required to evaluate *C. botulinum* contamination occurrence and level in hatcheries so as to better understand their potential role in the epidemiology of botulism.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

## ETHICS STATEMENT

Samples were collected by veterinarians for diagnostic purposes on dead animals (death due to the botulism outbreak). Cloacal swabs were collected on broilers in the farm by veterinarians for official analyses. This article presents non-experimental clinical veterinary studies. Animals on farms were treated in accordance with the European and French regulation on farmed animal protection. Samples were collected in accordance with Regulation EC/1099/2009 and Directive 2010/63/EU. Written informed consent was obtained from the owners for the participation of their animals in this study.

## AUTHOR CONTRIBUTIONS

CL, RS, SL, DG, and MC designed the study. CL, TP, SR, SM, and LM analyzed the samples for detection of *C. botulinum*. RS and CL analyzed the data and drafted the manuscript. DG, RS, and LB collected the epidemiological data and samples on field. TL and FA were in charge of the MLVA analysis. SL, FA, and MC provided their expertise feedback. All authors reviewed the draft and contributed significantly to the final manuscript.

## REFERENCES

- Anniballi, F., Fillo, S., Giordani, F., Auricchio, B., Tehran, D. A., Di Stefano, E., et al. (2016). Multiple-locus variable number of tandem repeat analysis as a tool for molecular epidemiology of botulism: the Italian experience. *Infect. Genet. Evol.* 46, 28–32. doi: 10.1016/j.meegid.2016.10.014
- Anniballi, F., Fiore, A., Löfström, C., Skarin, H., Auricchio, B., Woudstra, C., et al. (2013). Management of animal botulism outbreaks: from clinical suspicion to practical countermeasures to prevent or minimize outbreaks. *Bioterrorism* 11, S191–S199. doi: 10.1089/bsp.2012.0089
- Auricchio, B., Scalfaro, C., Vicenza, T., Bano, L., Bilei, S., Chemaly, M., et al. (2019). *Multilocus Variable-Number of Tandem-Repeat Analysis as a tool for Clostridium Botulinum Group III Subtyping ANIBOTNET Workshop Botulism, Paris, 28-29 march 2019*.
- Bano, L., Drigo, I., Tonon, E., Pascoletti, S., Puiatti, C., Anniballi, F., et al. (2017). Identification and characterization of *Clostridium botulinum* group III field strains by matrix-assisted laser desorption-ionization time-of-flight mass spectrometry (MALDI-TOF MS). *Anaerobe* 48, 126–134. doi: 10.1016/j.anaerobe.2017.08.004
- Chen, S., Li, J., Saleh-Lakha, S., Allen, V., and Odumeru, J. (2011). Multiple-locus variable number of tandem repeat analysis (MLVA) of *Listeria monocytogenes* directly in food samples. *Int. J. Food Microbiol.* 148, 8–14. doi: 10.1016/j.ijfoodmicro.2011.04.014
- Fillo, S., Giordani, F., Anniballi, F., Gorgao, O., Ramišse, V., Vergnaud, G., et al. (2011). *Clostridium botulinum* group I strain genotyping by 15-locus multilocus variable-number tandem-repeat analysis. *J. Clin. Microbiol.* 49, 4252–4263. doi: 10.1128/JCM.05396-11
- Hogg, R., Livesey, C., and Payne, J. (2008). Diagnosis and implications of botulism. *In Practice* 30, 392–397. doi: 10.1136/inpract.30.7.392
- Kahlisch, L., Henne, K., Draheim, J., Brettar, I., and Höfle, M. G. (2010). High-resolution in situ genotyping of *Legionella pneumophila* populations

## FUNDING

This work was funded by the French Agency for Food, Environmental and Occupational Health & Safety (ANSES), the Côtes d'Armor Departmental Council, and the ANIBOTNET project (ANIHWA call). TL is the recipient of a Ph.D. grant from ANSES and the Bretagne Regional Council, and the recipient of a short stay grant funded by the MedVetNet Association for a training period regarding the MLVA method at the ISS laboratory in Italy under Fabrizio Anniballi's supervision.

## ACKNOWLEDGMENTS

The authors would like to thank the farmers and veterinarians for their participation in the study; Concetta Scalfaro (ISS) for her help in MLVA analysis; and Valérie Rose (ANSES, Ploufragan Laboratory) and François-Xavier Briand (ANSES, Ploufragan Laboratory) for their help and expertise in the sequencing PCR products for MLVA analysis. **Figure 3** was created with BioRender.com.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2021.679377/full#supplementary-material>

in drinking water by multiple-locus variable-number tandem-repeat analysis using environmental DNA. *Appl. Environ. Microbiol.* 76, 6186–6195. doi: 10.1128/aem.00416-10

- Kim, J. H., and Kim, K. S. (2010). Hatchery hygiene evaluation by microbiological examination of hatchery samples. *Poultry Sci.* 89, 1389–1398. doi: 10.3382/ps.2010-00661
- Le Gratiot, T., Poezevara, T., Rouxel, S., Houard, E., Mazuet, C., Chemaly, M., et al. (2020). Development of an innovative and quick method for the isolation of *Clostridium botulinum* strains involved in avian botulism outbreaks. *Toxins (Basel)* 12:42. doi: 10.3390/toxins12010042
- Le Maréchal, C. L., Hulin, O., Macé, S., Chuzeville, C., Rouxel, S., Poëzevara, T., et al. (2019). A case report of a botulism outbreak in beef cattle due to the contamination of wheat by a roaming cat carcass: from the suspicion to the management of the outbreak. *Animals* 9:1025. doi: 10.3390/ani9121025
- McMullin, P. F. (2009). Hygiene and microbiological control in hatcheries. *Avian Biol. Res.* 2, 93–97. doi: 10.3184/175815509X431885
- Myllykoski, J., Lindström, M., Keto-Timonen, R., Söderholm, H., Jakala, J., Kallio, H., et al. (2009). Type C bovine botulism outbreak due to carcass contaminated non-acidified silage. *Epidemiol. Infect.* 137, 284–293. doi: 10.1017/S0950268808000939
- Notermans, S., Dufrenne, J., and Oosterom, J. (1981). Persistence of *Clostridium botulinum* type B on a cattle farm after an outbreak of botulism. *Appl. Environ. Microbiol.* 41, 179–183. doi: 10.1128/aem.41.1.179-183.1981
- Osman, K. M., Kappel, A. D., Elhadidy, M., Elmougy, F., El-Ghany, W. A. A., Orabi, A., et al. (2018). Poultry hatcheries as potential reservoirs for antimicrobial-resistant *Escherichia coli*: a risk to public health and food safety. *Sci. Rep.* 8:5859. doi: 10.1038/s41598-018-23962-7
- Pailhoriès, H., Buzelé, R., Picardeau, M., Robert, S., Mercier, E., Mereghetti, L., et al. (2015). Molecular characterization of *Leptospira* sp by multilocus variable number tandem repeat analysis (MLVA) from clinical samples: a case report. *Int. J. Infect. Dis.* 37, 119–121. doi: 10.1016/j.ijid.2015.06.026

- Payne, J. H., Hogg, R. A., Otter, A., Roest, H. I., and Livesey, C. T. (2011). Emergence of suspected type D botulism in ruminants in England and Wales (2001 to 2009), associated with exposure to broiler litter. *Vet. Rec.* 168:640. doi: 10.1136/vr.d1846
- Peck, M. W., Smith, T. J., Anniballi, F., Austin, J. W., Bano, L., Bradshaw, M., et al. (2017). Historical perspectives and guidelines for botulinum neurotoxin subtype nomenclature. *Toxins* 9:38. doi: 10.3390/toxins9010038
- Pereyre, S., Charron, A., Hidalgo-Grass, C., Touati, A., Moses, A. E., Nir-Paz, R., et al. (2012). The spread of *Mycoplasma pneumoniae* is polyclonal in both an endemic setting in France and in an epidemic setting in Israel. *PLoS One* 7:e38585. doi: 10.1371/journal.pone.0038585
- Popoff, M. (1989). Revue sur l'épidémiologie du botulisme bovin en France et analyse de sa relation avec les élevages de volailles. *Rev. Sci. Tech. Off. Int. Epiz.* 8, 129–145. doi: 10.20506/rst.8.1.404
- Popoff, M. R. (1995). "Ecology of neurotoxicogenic strains of clostridia," in *Clostridial Neurotoxins. Current Topics in Microbiology and Immunology*, ed. C. Montecucco (Berlin: Springer-Verlag), 1–29. doi: 10.1007/978-3-642-85173-5\_1
- Popoff, M. R., and Argente, G. (1996). Le botulisme animal est-il une menace pour l'homme? *Bull. l'Académie Vétérinaire France* 69, 373–382. doi: 10.4267/2042/63881
- Popp, C., Hauck, R., Gad, W., and Hafez, H. M. (2012). Type C botulism in a commercial Turkey farm: a case report. *Avian Dis.* 56, 760–763. doi: 10.1637/10274-061212-Case.1
- Qureshi, A. (2002). Hatchery sanitization and chick mortality. *World Poultry* 18, 24–25.
- Ramírez-Romero, R., Ramírez-Hernández, C., García-Márquez, L. J., Macedo-Barragán, R. J., Martínez-Burnes, J., and López-Mayagoitia, A. (2014). Bovine diseases causing neurological signs and death in Mexican feedlots. *Trop. Anim Health Prod.* 46, 823–829. doi: 10.1007/s11250-014-0572-y
- Rasetti-Escargueil, C., Lemichez, E., and Popoff, M. R. (2019). Public health risk associated with botulism as foodborne zoonoses. *Toxins (Basel)* 12:17. doi: 10.3390/toxins12010017
- Rasetti-Escargueil, C., Lemichez, E., and Popoff, M. R. (2020). Human botulism in France, 1875–2016. *Toxins (Basel)* 12:338. doi: 10.3390/toxins12050338
- Relun, A., Dorso, L., Douart, A., Chartier, C., Guatteo, R., Mazuet, C., et al. (2017). A large outbreak of bovine botulism possibly linked to a massive contamination of grass silage by type D/C *Clostridium botulinum* spores on a farm with dairy and poultry operations. *Epidemiol. Infect.* 145, 3477–3485. doi: 10.1017/s0950268817002382
- Senturk, S., and Cihan, H. (2007). Outbreak of botulism in a dairy herd in Turkey. *Ir. Vet. J.* 60, 481–484. doi: 10.1186/2046-0481-60-8-481
- Skarin, H., Lindgren, Y., and Jansson, D. S. (2015). Investigations into an outbreak of botulism caused by *Clostridium botulinum* Type C/D in laying hens. *Avian Dis.* 59, 335–340. doi: 10.1637/10861-051214-Case
- Skarin, H., Tevell Åberg, A., Woudstra, C., Hansen, T., Löfström, C., Koene, M., et al. (2013). The workshop on animal botulism in Europe. *Biosec. Bioterrorism* 11, S183–S190. doi: 10.1089/bsp.2012.0076
- Souillard, R., Le Marechal, C., Ballan, V., Mahe, F., Chemaly, M., and Le Bouquin, S. (2017). A bovine botulism outbreak associated with a suspected cross-contamination from a poultry farm. *Vet. Microbiol.* 208, 212–216. doi: 10.1016/j.vetmic.2017.07.022
- Souillard, R., Le Marechal, C., Ballan, V., Rouxel, S., Leon, D., Balaine, L., et al. (2016). Investigation of a type C/D botulism outbreak in free-range laying hens in France. *Avian Pathol.* 46, 195–201. doi: 10.1080/03079457.2016.1240355
- Souillard, R., Woudstra, C., Le Marechal, C., Dia, M., Bayon-Auboyer, M. H., Chemaly, M., et al. (2014). Investigation of *Clostridium botulinum* in commercial poultry farms in France between 2011 and 2013. *Avian Pathol.* 43, 458–464. doi: 10.1080/03079457.2014.957644
- Tsiouris, V. (2016). Poultry management: a useful tool for the control of necrotic enteritis in poultry. *Avian Pathol.* 45, 323–325. doi: 10.1080/03079457.2016.1154502
- Umeda, K., Wada, T., Kohda, T., and Kozaki, S. (2013). Multi-locus variable number tandem repeat analysis for *Clostridium botulinum* type B isolates in Japan: comparison with other isolates and genotyping methods. *Infect. Genet. Evol.* 16, 298–304. doi: 10.1016/j.meegid.2013.02.022
- Wales, A., and Davies, R. (2020). Review of hatchery transmission of bacteria with focus on *Salmonella*, chick pathogens and antimicrobial resistance. *World's Poultry Sci. J.* 76, 517–536. doi: 10.1080/00439339.2020.1789533
- Wobeser, G., Marsden, S., and Macfarlane, R. J. (1987). Occurrence of toxigenic *Clostridium botulinum* type C in the soil of wetlands in Saskatchewan. *J. Wildlife Dis.* 23, 67–76. doi: 10.7589/0090-3558-23.1.67
- Woudstra, C., Skarin, H., Anniballi, F., Fenicia, L., Bano, L., Drigo, I., et al. (2012). Neurotoxin gene profiling of *Clostridium botulinum* types C and D native to different countries within Europe. *Appl. Environ. Microbiol.* 78, 3120–3127.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Souillard, Grosjean, Le Gratiet, Poezevara, Rouxel, Balaine, Macé, Martin, Anniballi, Chemaly, Le Bouquin and Le Marechal. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

## 2.3.2. Exemple 2 d'application : Épisode de botulisme de type C/D

### 2.3.2.1. Présentation de l'épisode et du suivi dans le temps

Le deuxième cas d'étude retenu pour l'évaluation de la MLVA en conditions réelles est un élevage de dindes ayant rencontré des épisodes successifs de botulisme entre Avril 2019 et Août 2020. Cette exploitation avicole comprend 4 bâtiments d'élevage. Après l'apparition de signes cliniques chez des dindes (Figure 24) dans 3 des 4 bâtiments de l'élevage, le botulisme a été suspecté. Le premier épisode a été confirmé par PCR sur des foies de dindes du bâtiment 2 en Avril 2019 puis en Mai 2019 dans le bâtiment 4 et en Juin 2019 dans le bâtiment 1. Le type toxinique C/D a été détecté. Des taux de mortalité très élevés ont été rapportés dans les 3 bâtiments (100 % des mâles dans 2 bâtiments, 100 % des femelles dans le troisième). Le bâtiment 3 n'a pas été touché.



Figure 24 : Dindes présentant des signes cliniques dans un des trois bâtiments affectés.

Source : J. Boutant

Suite à la confirmation du botulisme dans l'exploitation, une enquête épidémiologique a été menée et *C. botulinum* de type C/D a été détecté dans 25 prélèvements sur 43. Cette enquête a permis de suspecter en particulier le blé qui servait à l'alimentation des animaux des 3 bâtiments atteints. En Septembre 2019 et en Août 2020, deux récurrences de botulisme C/D ont été confirmées dans l'exploitation.

### 2.3.2.2. Utilisation de la MLVA pour tracer la contamination.

Pour valider l'hypothèse émise sur l'origine de contamination du premier épisode et la persistance de la souche dans l'exploitation entraînant les deux récurrences, la méthode MLVA a été utilisée pour caractériser les souches présentes dans des échantillons représentatifs de l'enquête épidémiologiques. Au total, 9 échantillons positifs (Tableau 8) ont été sélectionnés en fonction de leur intérêt dans l'enquête et de la charge bactérienne estimée lors de la PCR de détection ( $Ct < 30$ ). Six échantillons de foies provenant des trois épisodes ont été inclus dans l'étude (3 foies du premier

épisode provenant des bâtiments 2 et 4, un foie du deuxième épisode et deux foies du troisième épisode). En ce qui concerne les prélèvements d'environnement de l'exploitation, un échantillon provenant du tas de blé stocké dans le hangar ainsi que deux échantillons d'environnement ont aussi été testés (une chiffonnette de la mangeoire du bâtiment 4 prélevée lors du premier épisode et une chiffonnette de la chaîne d'alimentation prélevée lors du troisième épisode).

Pour l'amplification des régions VNTR, contrairement à la première étude terrain, seulement 7 des 9 régions ciblées ont été évaluées. Les résultats obtenus après électrophorèse capillaire et séquençage des amplicons indiquent que le même profil MLVA est retrouvé pour les 9 échantillons évalués dans cette étude (Figure 1 article JRA page 131, en bleu).

Tableau 8 : Liste des échantillons testés pour la deuxième étude terrain et leurs codes MLVA correspondants.

Échantillon	VNTR14	VNTR15	VNTR16	VNTR17	VNTR18	VNTR19	VNTR24
<b>A- Foie épisode 1 bâtiment 2</b>	3	3	2	3	6	4	2
<b>B- Foie épisode 1 bâtiment 4</b>	3	3	2	3	6	4	2
<b>C- Foie épisode 1 bâtiment 4</b>	3	3	2	3	6	4	2
<b>D- Blé</b>	3	3	2	3	6	4	2
<b>E- Mangeoire</b>	3	3	2	3	6	4	2
<b>F- Foie épisode 2</b>	3	3	2	3	6	4	2
<b>G- Foie épisode 3</b>	3	3	2	3	6	4	2
<b>H- Foie épisode 3</b>	3	3	2	3	6	4	2
<b>I- Chaîne d'alimentation</b>	3	3	2	3	6	4	2

Dans cette étude, la MLVA a permis de montrer que le même profil est retrouvé dans l'ensemble des échantillons testés suggérant que le blé est bien la source de contamination initiale à l'origine de l'épisode et que la même souche a persisté au cours des épisodes. Ces résultats ont été valorisés dans la même présentation et le même un article que l'évaluation de la méthode sur les souches pures lors des 14<sup>ème</sup> journées de la recherche avicole et palmipèdes à foie gras (JRA) qui ont eu lieu les 9 et 10 Mars 2022 à Tours.

### 3. Calcul du pouvoir discriminant de la méthode MLVA

Le pouvoir discriminant d'une méthode de typage est sa faculté à distinguer deux souches non liées. Un indice de discrimination particulièrement utilisé pour évaluer l'outil de typage MLVA est l'indice Hunter-Gaston (HGDI) (Hunter et Gaston., 1988) qui se calcule comme ceci :

Équation 1 : Indice Hunter-Gaston

$$D = 1 - \frac{1}{N(N-1)} \sum_{j=1}^S n_j(n_j - 1)$$

Avec  $D$  correspondant à l'indice de discrimination, compris entre 0 et 1,  $N$  le nombre total de souches évaluées,  $S$  le nombre de profils MLVA différents retrouvés,  $n_j$  le nombre de souches présentes dans le profil MLVA  $j$ . Pour qu'une méthode soit considérée comme discriminante, la valeur de  $D$  doit être comprise entre 0,90 et 0,99 (Hunter et Gaston, 1988). Pour notre évaluation, les résultats de typage de 53 souches ont été pris en compte :

- 30 souches provenant des résultats de l'ISS, en ne prenant en compte que celles retrouvées dans des espèces aviaires.
- 20 souches isolées au cours de la thèse.
- 3 souches provenant des deux épisodes présentés précédemment (2 profils pour le cas de type D/C, 1 profil pour le cas de type C/D). Malgré la présence de plusieurs échantillons avec le même profil MLVA, seul un échantillon est retenu par profil pour éviter d'avoir un biais dans le calcul car l'origine des échantillons est la même.

L'indice HGDI a été calculé à partir des profils MLVA obtenus avec 7 régions VNTR, en éliminant les régions VNTR 20 et 23. L'indice HGDI calculé sur la base des résultats des 53 souches pour les 7 VNTR est de  $D = 0,98 \pm 0,018$  avec un intervalle de confiance à 95% (Grundmann et al., 2001). Cet indice étant situé entre 0,95 et 0,99, la méthode MLVA peut être considérée comme très discriminante. L'indice a aussi été calculé pour chacune des 7 régions VNTR individuellement. Les 7 indices HGDI calculés se situent entre 0,41 et 0,80, ce qui explique l'intérêt d'avoir le plus de cibles possibles lors du développement d'un schéma MLVA (Tableau 9).

Tableau 9 : Indices de discrimination Hunter et Gaston (HGDI) calculés pour chaque VNTR et pour le schéma MLVA avec 7 VNTR (un intervalle de confiance à 95%).

	VNTR14	VNTR15	VNTR16	VNTR17	VNTR18	VNTR19	VNTR24	Σ VNTR
HGDI	0,80 ± 0,03	0,71 ± 0,05	0,41 ± 0,15	0,73 ± 0,05	0,79 ± 0,05	0,61 ± 0,10	0,60 ± 0,06	0,98 ± 0,02

### **3. Stabilité de la méthode**

La stabilité d'une méthode de typage est un critère important lors de l'évaluation d'un outil. Une méthode est dite stable lorsque la cible n'évolue pas trop rapidement lors des étapes de repiquages et sous-cultures de la souche en laboratoire (van Belkum et al., 2007). Pour notre méthode, cela veut dire que le nombre de répétitions de chaque région VNTR ne doit pas évoluer au cours de ces étapes. Pour s'assurer de la stabilité de la méthode, la souche de référence de type C/D BKT015925 a été mise en culture puis repiquée à 10 reprises (1 mL de culture bactérienne dans 9 mL de milieu TPGY pour chaque repiquage, culture de 24 heures minimum à 37 °C en conditions anaérobies). Une extraction d'ADN a ensuite été réalisée sur la première et la dernière culture puis les extraits d'ADN ont été soumis à la méthode MLVA et les deux profils obtenus ont été comparés. Les résultats de ce test ont montré que les profils MLVA des deux extraits sont identiques, confirmant que la méthode est stable pour les 7 régions ciblées de *C. botulinum* du groupe III.

### **4. Conclusion et perspectives sur la méthode MLVA**

Le développement d'une méthode MLVA pour caractériser des souches de *C. botulinum* du groupe III est une avancée majeure pour valider les hypothèses émises lors d'enquêtes épidémiologiques. L'isolement des souches étant encore très difficile voire impossible pour les échantillons provenant de matrices complexes, disposer d'une méthode de typage ne nécessitant pas cette étape d'isolement est essentiel. Les résultats obtenus au cours de cette thèse et lors du développement de la MLVA par l'ISS montrent que cette méthode est bien adaptée aux souches de la lignée Ia (type C/D), avec une amplification PCR pour l'ensemble des 7 régions VNTR du schéma MLVA (stabilité des cibles, reproductibilité, pouvoir discriminant, concordance épidémiologique). Pour les souches des autres lignées, la méthode ne permet pas l'obtention d'une amplification de manière systématique pour toutes les cibles, en particulier pour les souches de type D pour lesquelles aucune amplification n'a été obtenue. La recherche d'autres cibles plus adaptées sur des génomes complets des lignées pour lesquelles les 7 VNTR ne semblent pas appropriés (Woudstra et al., 2021) pourrait permettre de pallier ce problème.

Ces travaux de thèse ont permis le transfert de la méthode MLVA depuis l'ISS au laboratoire de Ploufragan et cette approche pourra être utilisée très prochainement en routine dans le cadre des analyses du LNR botulisme aviaire.

La suite des travaux de thèse a porté sur l'exploitation des génomes de *C. novyi sensu lato* et notamment des régions VNTR.

## 5. Publications en lien avec la méthode MLVA

Publication pour les Journées de la Recherche Avicole et Palmipèdes à Foie gras  
**CARACTERISATION DE *CLOSTRIDIUM BOTULINUM* DANS LE CADRE D'EPISODES  
DE BOTULISME AVIAIRE EN FRANCE : EVALUATION D'UN NOUVEL OUTIL**

**Le Gratiet Thibault<sup>1</sup>, Anniballi Fabrizio<sup>2</sup>, Boutant Jérémy<sup>3</sup>, Mellal Mokrane<sup>4</sup>, Rouxel Sandra<sup>1</sup>,  
Poezevara Typhaine<sup>1</sup>, Houard Emmanuelle<sup>1</sup>, Rose Valérie<sup>1</sup>, Briand François-Xavier<sup>5</sup>, Chemaly  
Marianne<sup>1</sup>, Le Maréchal Caroline<sup>1</sup>**

<sup>1</sup>ANSES PLOUFRAGAN-PLOUZANE-NIORT, UNITE HYGIENE ET QUALITE DES PRODUITS AVICOLES ET PORCINS - BP53 - 22440  
PLOUFRAGAN, FRANCE

<sup>2</sup>ISTITUTO SUPERIORE DI SANITA - CNR BOTULISME ITALIEN - ROME, ITALIE

<sup>3</sup>SOCAVET, 75 boulevard de Penthièvre, 22600 LOUDEAC, FRANCE

<sup>4</sup>VET&SPHERE - RESEAU CRISTAL, ZI DE TIRPEN, 56140 MALESTROIT, FRANCE

<sup>5</sup>ANSES PLOUFRAGAN-PLOUZANE-NIORT, UNITE VIROLOGIE, IMMUNOLOGIE, PARASITOLOGIE AVIAIRE ET CUNICOLES - BP53 - 22440  
PLOUFRAGAN

[caroline.lemarechal@anses.fr](mailto:caroline.lemarechal@anses.fr)

### RÉSUMÉ

Le botulisme aviaire est une maladie neuroparalytique sévère, dû à l'action de la toxine botulique produite par *Clostridium botulinum*, touchant à la fois les oiseaux sauvages et d'élevage. L'absence d'outils de typage aboutit à une méconnaissance de l'épidémiologie du botulisme aviaire. L'isolement des souches de *C. botulinum* est difficile, ce qui empêche d'appliquer directement les outils qui existent en routine pour d'autres espèces bactériennes et nécessite le développement d'outils dédiés et adaptés. Une méthode de typage nommée Multi Loci VNTR (Variable Number of Tandem Repeats) analysis (MLVA) a ainsi été développée récemment pour caractériser *C. botulinum* dans le cadre d'épisodes de botulisme animal directement à partir des échantillons, sans étape d'isolement. Cette méthode basée sur le polymorphisme de régions répétées (VNTR) consiste à amplifier 7 cibles différentes par PCR puis à déterminer pour chaque cible le nombre de répétitions. L'objectif de notre étude était de caractériser *C. botulinum* dans le contexte d'épisodes de botulisme aviaire en France via cette méthode MLVA. Vingt souches pures représentatives de la collection du LNR botulisme aviaire ont été sélectionnées et comparées par MLVA. Les résultats montrent que les souches de type C/D, type le plus fréquent en aviaire, se regroupent ensemble avec un pourcentage de similarité compris entre 52,2 et 100%, que ces souches soient associées à un foyer en élevage ou un cas en faune sauvage. Notre étude n'a pas permis la mise en évidence de lien entre la répartition des souches selon leur profil MLVA et l'année, le département ou l'espèce d'origine. La méthode a également été testée dans le cadre d'investigations menées dans un élevage de dindes présentant des récurrences sur des lots successifs. Le même profil a été détecté au cours des 3 épisodes différents, suggérant la persistance de la souche au cours du temps. Cette étude montre que la MLVA constitue un outil de typage prometteur pour explorer l'épidémiologie moléculaire du botulisme aviaire, que ce soit à partir de souches isolées ou d'échantillons du terrain.

### ABSTRACT

#### **Typing of *C. botulinum* involved in avian botulism outbreaks in France: Evaluation of a new tool.**

Avian botulism is a severe neuroparalytic disease affecting both poultry and wild birds. This disease is due to the action of the botulinum neurotoxin produced by *Clostridium botulinum*. Unavailability of typing tools results in a poor understanding of the epidemiology of avian botulism. Indeed, *C. botulinum* isolation is very difficult, preventing the use of common typing tools. The development of typing tools dedicated to the study of *C. botulinum* is needed. A typing tool named Multi Loci VNTR Analysis (MLVA) has been recently developed to characterize *C. botulinum* within the frame of animal botulism, directly useable on samples without requiring strain isolation. This method is based on the polymorphism of repeated regions (VNTR). Seven VNTR regions are amplified using PCR and sized to determine the number of repeats of each region. The aim of this study was to characterize *C. botulinum* involved in avian botulism outbreaks in France using MLVA. Twenty isolated strains representative of the French NRL collection were selected and compared. Results showed that French type C/D strains, the most common neurotoxin type in avian botulism clustered together with similarities ranging from 58.2% to 100%, including both wild bird and poultry strains. No link was highlighted in our study between strain distribution and factors such as year of the outbreak, location or bird species. MLVA was also directly tested on samples collected in a turkey farm where recurrent botulism outbreaks occurred in successive flocks. The same MLVA profile was detected on samples collected from 3 different outbreaks, suggesting the persistence of the strain over time in the farm. This study showed that MLVA is a promising typing tool for the exploration of the molecular epidemiology of avian botulism, either on isolated strains or on field samples.



## INTRODUCTION

Le botulisme est une maladie neuroparalytique sévère causée par l'action de la neurotoxine botulique produite par des bactéries du genre *Clostridium*, notamment *Clostridium botulinum*, bactérie ubiquitaire à coloration de Gram positif, anaérobie stricte et sporulée. Neuf types de toxine botulique sont décrites (classés de A à G, H et X) et plus de 40 sous-types ont été caractérisés à ce jour (Peck et al., 2017).

De nombreuses espèces animales sont sensibles au botulisme dont les espèces aviaires. En France, environ 25 foyers de botulisme aviaire ont été recensés annuellement sur la période 2009-2019 (Anses 2021), la toxine majoritairement identifiée étant la toxine C/D (Anses 2021). Des épisodes associés aux toxines C, D, D/C ou E sont également rapportés mais à une fréquence beaucoup plus faible. En élevage, les épisodes de botulisme entraînent des pertes économiques majeures (Souillard et al. 2014). Les foyers de botulisme en avifaune sauvage vont pour leur part avoir des conséquences écologiques importantes, certains épisodes pouvant affecter jusqu'à plusieurs dizaines de milliers d'oiseaux sauvages (Rocke 2006). Malgré ce contexte, l'épidémiologie du botulisme aviaire est encore méconnue et les données d'épidémiologie moléculaire relatives aux souches impliquées dans le botulisme animal sont quasi inexistantes à ce jour. Ceci s'explique principalement par le fait que ces souches sont difficiles à isoler. Différents facteurs expliquent cette difficulté notamment l'absence de milieu sélectif. Cela constitue actuellement un verrou méthodologique à la caractérisation des souches de *C. botulinum* responsables des épisodes de botulisme animal.

L'approche classique pour typer une souche bactérienne est d'isoler la souche puis de la caractériser ensuite. Une stratégie pour pallier la difficulté rencontrée pour isoler les souches de *C. botulinum* est l'utilisation d'outils permettant de s'affranchir de cette étape d'isolement.

La méthode de typage appelée Multi Loci VNTR (Variable Number of Tandem Repeats) Analysis (MLVA) est basée sur le polymorphisme de plusieurs VNTR, des régions du génome présentant des séquences répétées. Cette méthode consiste en la détermination du nombre de répétitions au niveau de différents VNTR. Elle présente l'avantage majeur d'être utilisable directement à partir des ADN bruts. Cette méthode est en effet basée sur des amplifications par PCR, permettant de travailler directement sur les échantillons (Chen et al. 2011). La MLVA est couramment utilisée avec succès pour le typage de nombreuses espèces bactériennes pathogènes (Van Belkum. 2007), dont les souches de *C. botulinum* impliquées dans le botulisme humain (Fillo et al. 2011).

Une méthode MLVA a été développée pour typer *C. botulinum* dans le contexte d'épisodes de botulisme

animal dans le cadre d'un projet précédent (ANIBOTNET) et testée sur un panel de souches européennes (Auricchio et al, 2019).

L'objectif de notre étude est de tester cette méthode MLVA à la fois sur des souches isolées en France mais également sur des échantillons du terrain collectés dans le cadre d'investigations épidémiologiques menées dans un élevage avicole.

## 1. MATERIELS ET METHODES

### 1.1. Sélection des souches et échantillons

#### 1.1.1 Souches pures de *C. botulinum*

Vingt souches de *C. botulinum* de la collection du LNR botulisme aviaire représentatives des épisodes de botulisme aviaire français survenus entre 2014 et 2019 (années, types toxiques, espèces affectées, ratio élevage/faune sauvage, localisation) ont été sélectionnées pour être caractérisées selon la méthode MLVA qui a été développée pour *C. botulinum* de type C, D, C/D et D/C. Ces souches ont été isolées en 2019 dans le cadre d'une étude précédente (Le Gratiet et al., 2020).

#### 1.1.2 Echantillons collectés dans le cadre du suivi épidémiologique d'un épisode de botulisme

La méthode a par ailleurs été testée afin d'évaluer la possibilité de tracer la contamination par *C. botulinum* au cours d'un épisode de botulisme via l'utilisation de l'approche MLVA. Pour cela, un élevage de dindes ayant connu des épisodes successifs de botulisme a été retenu. Le premier épisode a été confirmé en laboratoire en avril 2019 via l'analyse de foies d'animaux atteints qui a permis de mettre en évidence *C. botulinum* de type C/D par PCR. Sur les 4 bâtiments que comprend l'élevage, 3 ont été atteints avec des taux de mortalité élevés (100% des mâles dans 2 bâtiments et 100% des femelles dans le troisième). L'apparition des signes cliniques a été observée à 55 jours dans un bâtiment et 82 et 93 jours dans les 2 autres bâtiments. Pour notre étude, 3 foies provenant de 2 bâtiments ont été inclus (Echantillons A, B, C).

Une enquête épidémiologique a ensuite été menée dans cet élevage pour identifier la source de contamination et a permis de détecter *C. botulinum* de type C/D dans un tas de blé (5 positifs sur 8 échantillons prélevés sur l'ensemble du tas) stocké dans un hangar de l'élevage et utilisé par l'éleveur en complément de l'aliment commercial dans les 3 bâtiments atteints. Ce blé n'était pas utilisé dans le quatrième bâtiment de l'élevage, non atteint pendant l'épisode. Un échantillon de blé provenant du tas de blé positif pour *C. botulinum* type C/D (Echantillon D) et une chiffonnette prélevée au niveau d'une mangeoire d'un des bâtiments atteints (Echantillon E) ont été inclus dans notre étude.

L'élevage a ensuite connu deux récurrences de botulisme en septembre 2019 et août 2020. Des foies prélevés sur des dindes présentant des signes cliniques lors de ces deux épisodes (Echantillons F, G et H) ainsi qu'une

chiffonnette de la chaîne d'alimentation d'un des bâtiments atteint de botulisme lors du dernier épisode (Echantillon I) ont été inclus dans l'étude.

Les 9 échantillons sélectionnés présentent un signal précoce en PCR (Ct inférieur à 30), afin de disposer d'une quantité d'ADN suffisante pour obtenir une amplification exploitable des régions VNTR testées.

## 1.2. Typage de *C. botulinum* par MLVA

Le typage par MLVA a été mis en œuvre tel que décrit précédemment (Souillard et al. 2021) en se basant sur l'analyse de 7 VNTR au lieu de 9 (les VNTR 20 et 23 présentaient en effet des amplifications non reproductibles et laboratoires dépendants (données non présentées) et n'ont donc pas été utilisées dans le cadre de cette étude). Brièvement, 7 VNTR ont été amplifiées par PCR. La taille des amplicons a ensuite été déterminée par électrophorèse capillaire et le nombre de répétitions a été déterminée par séquençage Sanger des amplicons. Un code MLVA à 7 chiffres (correspondant au nombre de répétitions obtenus pour chaque VNTR) a été attribué à chaque échantillon. Les codes ont ensuite été importés dans le logiciel Bio Numerics 7.5 (Applied Maths, Sint-Martens-Latem, Belgique) pour la construction d'un dendrogramme construit en utilisant un coefficient catégoriel et l'algorithme UPGMA (Unweighted Pair Group Method with Arithmetic mean). Les profils MLVA de *C. botulinum* en provenance d'autres pays publiés précédemment (Auricchio et al. 2019) et d'échantillons typés dans le cadre d'une autre enquête épidémiologique (Souillard et al. 2021) ont été ajoutés dans le dendrogramme.

L'indice de discrimination de Hunter-Gaston (HGDI) pour le schéma MLVA a été calculé (Hunter and Gaston, 1988) avec un intervalle de confiance à 95% (Grundmann et al., 2001).

## 2. RESULTATS ET DISCUSSION

### 2.1 Comparaison de la sélection des souches pures par MLVA

Le premier objectif de notre étude était de comparer par MLVA une sélection de souches françaises de *C. botulinum* avec les autres souches déjà typées précédemment en utilisant la méthode MLVA (Auricchio et al. 2019, Souillard et al. 2021). L'indice HGDI calculé est 0.98 (95% CI 0,96 – 0.99), > 95%, valeur recommandée pour les schémas de typage.

Le type toxinique le plus fréquemment associé au botulisme aviaire est le type C/D, les types D et D/C sont également détectés mais à une fréquence beaucoup plus faible (Anses 2021). Le profil de 4 souches de type D et 3 de D/C a été inclus dans le dendrogramme. Bien que le nombre de souches de type D et D/C soit faible, une tendance des souches à se regrouper par type toxinique est observée (Figure 1). Cette répartition des souches en fonction de leur type toxinique par la méthode MLVA est en accord avec les

analyses des génomes publiées dans la littérature qui montrent une répartition des souches dans des lignées différentes en fonction de leur type toxinique (Skarin et al. 2014 ; Fillo et al. 2021). Compte tenu du nombre restreint de souches, ce résultat reste cependant à confirmer via l'analyse d'un plus grand nombre de souches.

Pour *C. botulinum* de type C/D, les souches françaises présentent un pourcentage de similarité compris entre 52,2% et 100%. Elles se répartissent en 6 groupes de 1 à 6 souches avec un pourcentage de similarité de 100%. Ces 6 groupes comprennent des souches provenant de départements différents, ne permettant pas de mettre en évidence un lien entre un profil MLVA et une origine géographique.

A première vue, une tendance au regroupement des souches de type C/D par pays semble se dessiner dans la figure 1 sur la base de l'analyse des souches isolées en France et en Italie. Cependant, peu de souches d'autres pays que la France et l'Italie ont été typées par MLVA pour le moment. De plus, les quelques souches en provenance d'Allemagne et de Suède s'intercalent dans les groupes de souches d'origine française ou italienne, suggérant que des souches présentant des profils MLVA identiques peuvent être retrouvées malgré une origine géographique différente. Cette proximité de souches isolées de pays différents avait déjà été suggérée via la comparaison par PFGE (Pulse Field Gel Electrophoresis, autre méthode de typage bactérien très classique) de souches isolées en Suède et en Espagne pour des épisodes survenus en 2007, 2008, 2010 et 2011 (Anza et al. 2014). Dans notre étude, l'année de survenue de l'épisode de botulisme n'est pas disponible pour l'ensemble des souches, ne permettant pas d'évaluer une éventuelle différence liée à l'année d'isolement.

Le typage d'un nombre plus important de souches en provenance d'autres pays que la France et l'Italie est nécessaire pour explorer l'effet « origine géographique » et « date d'isolement » dans la comparaison des souches par MLVA.

Des souches isolées de cas de botulisme en faune sauvage et de foyers en élevages se regroupent : un groupe de 10 souches (6 de faune sauvage et 4 d'élevages ; cf. encadré rouge dans la Figure 1) présente ainsi un pourcentage de similarité d'au moins 73% et un groupe de 2 souches (B19LNRB07 isolée de poulet et N19LNRB01 de canard colvert) présente un pourcentage de similarité de 100%. Un troisième groupe de 3 souches présente le même profil MLVA (B19LNRB11 provenant d'un élevage et B19LNRB06 et B19LNRB13 provenant de cas en faune sauvage). La proximité génétique des souches impliquées dans les cas en faune sauvage et les épisodes en élevage a été rapportée récemment dans une étude menée via une comparaison génomique de souches isolées en Italie (Fillo et al. 2021). Aucun lien épidémiologique n'a cependant pu être mis en évidence entre ces cas (années et départements différents). Des analyses complémentaires sont nécessaires pour comparer des

souches isolées de mêmes zones géographiques et sur une même période afin d'évaluer et déterminer les risques de transmission entre la faune sauvage et les élevages.

## **2.2 Utilisation de la MLVA pour identifier la source de contamination et tracer la contamination lors d'épisodes successifs de botulisme dans un élevage avicole.**

Le second objectif de notre étude était d'évaluer la possibilité d'utiliser la MLVA pour tracer la contamination via l'application de cette approche à l'étude d'un cas concret. Pour cela, 9 échantillons collectés dans le cadre d'épisodes successifs de botulisme de type C/D dans un élevage de dindes ont été sélectionnés. L'analyse MLVA de ces 9 échantillons a permis d'obtenir le même profil dans tous les échantillons (Figure 1 en bleu), profil différent des profils obtenus pour les autres souches typées jusqu'alors, suggérant la présence de la même souche ou d'une souche très proche dans ces 9 échantillons. L'enquête épidémiologique menée dans cet élevage avait montré que le blé stocké dans l'élevage et distribué aux animaux constituait la source la plus probable de contamination à l'origine du premier épisode. Le même profil MLVA a été détecté dans les foies des animaux atteints, le blé stocké et la chiffonnette réalisée au niveau de la mangeoire dans un des bâtiments, supportant cette hypothèse. L'origine de la contamination lors des épisodes de botulisme aviaire est rarement identifiée, l'aliment est cependant souvent suspecté (Anses 2021). Notre étude semble donc confirmer l'aliment comme source de contamination à l'origine de l'initiation d'un épisode.

Les échantillons collectés lors des 2 autres épisodes survenus la même année et l'année suivante présentent le même profil que les échantillons du premier épisode, suggérant l'implication de la même souche ou d'une

souche très proche dans l'ensemble des épisodes de botulisme dans cet élevage. Les analyses de détection de *C. botulinum* de type C/D réalisées après les opérations de nettoyage et désinfection via le chiffonnage de surfaces dans et aux abords des bâtiments depuis la survenue du premier épisode en avril 2019 montrent une persistance de la contamination dans l'élevage, pouvant expliquer les récurrences observées. La forte résistance de *C. botulinum* dans l'environnement est largement décrite dans la littérature et des récurrences sont fréquemment rapportées (Souillard et al. 2014).

## **CONCLUSION**

Cette étude montre que la MLVA est un outil de typage adapté à l'étude de l'épidémiologie moléculaire de *C. botulinum* dans le cadre d'épisodes de botulisme aviaire en particulier pour permettre de tracer une souche ou d'identifier les sources de contamination sans passer par l'étape longue et fastidieuse d'isolement.

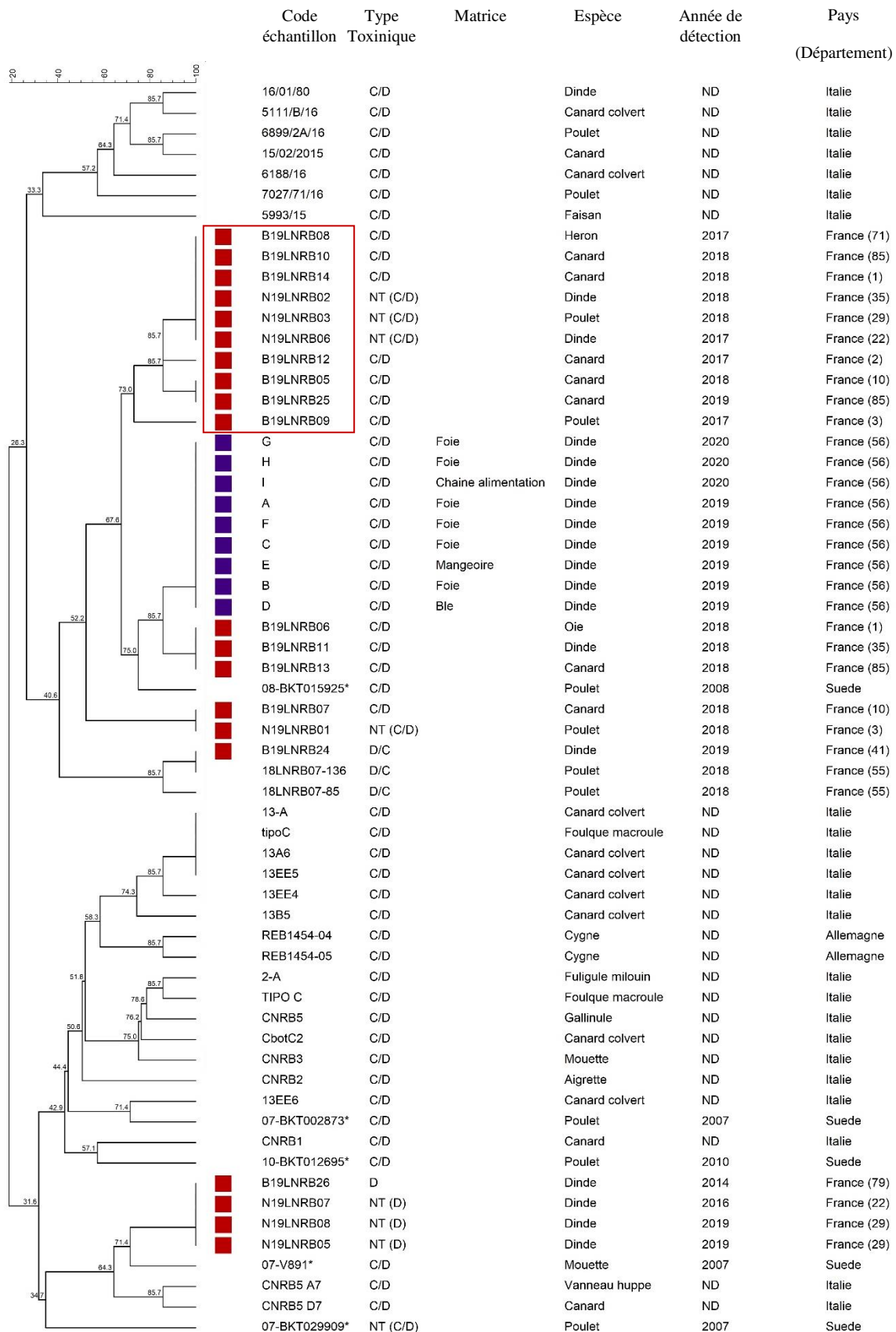
L'utilisation de cet outil à l'échelle européenne, que ce soit en élevage ou en faune sauvage pourrait permettre la mise en évidence de liens épidémiologiques potentiels entre les épisodes de botulisme.

## **REMERCIEMENTS**

Cette étude a reçu le soutien financier du conseil départemental des Côtes d'Armor. Thibault Le Gratiot a été le lauréat d'une bourse de mobilité de l'association MedVetNet et de l'école doctorale EGAAL pour un séjour au Centre National de Référence pour le botulisme italien basé à l'Istituto Superiore di Sanità de Rome qui a permis le transfert de la méthode MLVA au laboratoire de l'Anses dans le cadre de sa thèse financée par la région Bretagne et l'Anses.

## **REFERENCES BIBLIOGRAPHIQUES**

- Anses (2021). (Saisines 2019-SA-0112 à 2019-SA-0115). Maisons-Alfort : Anses, 170 p.
- Anza, I., 2014. *Anaerobe*, 26 (1), 20-23
- Auricchio, B., 2019. ANIBOTNET Workshop Botulism, Paris, 28-29 march 2019.
- Chen, S., 2011. *Int J Food Microbiol*, 148 (1), 8-14
- Fillo, S., 2011. *J Clin Microbiol*, 49 (12), 4252-4263.
- Fillo, S., 2021. *Microorganisms*, 9 (11), 2347
- Grundmann H., 2001. *J. Clin. Microbiol.* 39, 4190-4192.
- Hunter, P. R., 1988. *J. Clin. Microbiol.* 26, 2465-2466.
- Rocke, T.E., 2006. *Waterbirds around the world*, 422-426.
- Skarin, H., & Segerman, B. 2014. *PLoS One*, 9 (9), e107777.
- Souillard, R., 2014. *Avian Pathol*, 43 (5), 458-464
- Souillard, R., 2021. *Frontiers in Microbiology*, 12 (1753)
- Van Belkum, A., 2007. *FEMS Immunol Med Microbiol*, 49 (1), 22-27
- Vidal, D., 2013. *Appl Environ Microbiol*, 79 (14), 4264-4271.
- Woudstra, C., 2012. *Appl Environ Microbiol*, 78 (9), 3120-3127



**Figure 1.** Dendrogramme représentant les profils MLVA (7 VNTR) de souches pures de *C. botulinum* impliquées dans des épisodes de botulisme aviaire en France (■), d'échantillons bruts provenant de 3 épisodes de botulisme successifs dans un élevage de dindes (■) et de profils publiés précédemment (Auricchio et al. 2019 ; Souillard et al. 2021) (Construit avec le logiciel de BioNumerics 7.5.)

Encadré rouge : 10 souches de *C. botulinum* de type C/D avec un % de similarité > 73%. \* = Génome disponible.  
 NT = Non toxique ND = Non déterminée

# **PARTIE III**

**Optimisation des étapes du séquençage et analyses comparatives des  
génomés de *C. novyi sensu lato***

L'évaluation de la MLVA a montré que l'utilisation d'un schéma MLVA commun à l'ensemble des souches de *C. novyi sensu lato* n'est pas adaptée compte tenu de la variabilité d'amplification selon les types toxiques des souches ou la lignée concernée. Développer des outils de typage spécifiques de chaque lignée semble être la solution la plus appropriée pour pouvoir discriminer avec plus de pertinence ces souches, notamment pour les lignées III et IV pour lesquelles le schéma MLVA développé ne permet pas d'obtenir d'amplification. Pour atteindre cet objectif, augmenter le nombre de génomes disponibles est nécessaire pour déterminer des cibles spécifiques des lignées avec l'assurance que celles-ci se trouvent sur l'ensemble des souches de la lignée. Dans ce chapitre, les travaux menés dans le cadre du projet de thèse sur l'optimisation des étapes d'extraction d'ADN et de préparation des bibliothèques en vue du séquençage du génome seront développés. Une deuxième partie sera consacrée à la comparaison et l'analyse des génomes disponibles dans les bases de données et des génomes séquencés au cours de cette thèse dans le but de trouver des cibles pertinentes pour le développement d'outils de typage de *C. novyi sensu lato*.

## **1. Optimisation des étapes menant au séquençage du génome complet**

### **1.1. Comparaison des kits d'extraction d'ADN**

#### **1.1.1. Contexte de l'étude**

Seulement une centaine de génomes du groupe *C. novyi sensu lato* sont disponibles dans les bases de données publiques pour le moment. Parmi ces génomes, seuls 5 sont des génomes complets. Disposer de génomes complets est pourtant beaucoup plus informatif pour l'étude des génomes bactériens que les génomes non fermés. Par ailleurs, concernant le groupe *C. novyi sensu lato*, cela permet de disposer de manière fiable de la séquence et de l'organisation des EGM. Deux approches sont actuellement disponibles pour obtenir le génome complet d'une souche : l'utilisation de la technologie PacBio (Woudstra et al., 2021) ou la combinaison du séquençage Illumina (« short reads ») et MinIon (« long » reads ») (Gonzalez-Escalona & Sharma, 2020). Cette dernière approche tire profit des avantages et inconvénients des deux technologies de séquençage. L'Illumina permet d'obtenir de manière fiable les séquences et le MinIon qui présente un taux d'erreur important mais permet d'obtenir des séquences longues sert de squelette pour aligner les contigs obtenus en Illumina. Dans les deux cas, il est nécessaire de disposer d'un ADN extrait de bonne qualité et en quantité suffisante. Selon la méthode de séquençage utilisée, les exigences concernant la qualité et la quantité de matériel génétique ne sont pas les mêmes. En effet, pour le séquençage « long-reads », il faut une grande quantité d'ADN (de 1 à 5 µg selon le séquenceur utilisé) et de l'ADN non fragmenté tandis que pour un séquençage « short-reads », la quantité d'ADN est moindre (1 à 500 ng pour les génomes

de petite taille d'après Illumina) et il y a moins d'exigences concernant la fragmentation de l'ADN. *C. botulinum* étant une bactérie à coloration de Gram positif et pouvant sporuler, l'étape de lyse bactérienne constitue un « défi » et aboutit à un rendement d'extraction généralement assez faible ne permettant pas l'obtention d'une quantité d'ADN compatible avec les contraintes du séquençage. De plus, la libération fréquente de DNases au cours de cette lyse va induire une dégradation de l'ADN (Skarin et al., 2010), aboutissant à l'obtention d'un ADN fragmenté. Il est donc nécessaire d'avoir un protocole d'extraction permettant d'obtenir de l'ADN en quantité et de qualité suffisante pour permettre le séquençage de la souche et l'exploitation des données obtenues.

Dans la littérature scientifique, les différentes études qui ont mené au séquençage de génomes de *C. novyi sensu lato* ont fait appel à plusieurs kits d'extraction sans qu'un consensus ne soit établi sur celui qui est le plus optimal pour le séquençage short-reads et/ou long-reads. Au cours de la thèse, une comparaison de plusieurs protocoles d'extraction sélectionnés à partir de la littérature portant sur *C. novyi sensu lato* et d'autres bactéries Gram positif sporulantes a été réalisée afin de sélectionner les protocoles d'extraction les plus pertinents pour permettre le séquençage des souches. Un article a été rédigé sous la forme d'une short-note à soumettre (des résultats complémentaires seront ajoutés à cet article et l'anglais sera révisé par le service traduction de l'Anses avant soumission) :

### 1.1.2. Article N°3

## Comparison of five DNA extraction protocols for the sequencing of *Clostridium novyi sensu lato* strains.

Thibault Le Gratiet<sup>1</sup>, Caroline Le Maréchal<sup>1</sup>, Marianne Chemaly<sup>1</sup>

<sup>1</sup>ANSES, French Agency for Food, Environmental and Occupational Health Safety, Hygiene and Quality of Poultry and Pig Products Unit, Ploufragan, France

\* **Correspondence: Caroline Le Maréchal** caroline.lemarechal@anses.fr

**Key words:** *Clostridium novyi sensu lato*, DNA extraction, long-read sequencing

### ABSTRACT

Availability of whole genome sequencing technologies make sequencing affordable in a short time. However, extracting large quantities of DNA is particularly challenging with *Clostridium novyi sensu lato* due to production of DNases that can be released during DNA extraction and to its cell wall and spore structure that can be difficult to disrupt. The objective of our study was to evaluate DNA extraction protocols to select the most appropriate one for long-read whole genome sequencing.

Five DNA extraction protocols and two culture conditions were tested on a *Clostridium novyi sensu lato* strain. DNA yield and quality were evaluated. Results showed that two methods (one commercial DNA extraction kit and one in-house protocol) yielded DNA quality and quantity suitable for long-read (MinIon) sequencing, opening new perspectives for long-read genome sequencing of *Clostridium novyi sensu lato*.

### INTRODUCTION

*Clostridium novyi sensu lato* is a collective “genospecies” name covering *Clostridium botulinum* group III which is responsible for animal botulism, *Clostridium novyi* that is responsible for gas gangrene and infectious necrotic hepatitis in sheep and *Clostridium haemolyticum* which is responsible for bacillary hemoglobinuria (Skarin & Segerman, 2014). They are Gram positive, anaerobic, rod shaped, spore-forming bacteria that can be found ubiquitously in soils and marine sediments. This genotypic name was proposed because of their closed relationship demonstrated using whole genome comparison, MALDI-TOF MS, 16S rRNA similarity and DNA-DNA hybridizations (Bano et al., 2017; Nakamura et al., 1983; Sasaki et al., 2001; Skarin et al., 2011).



Despite improvements in the general field of whole genome sequencing (WGS) over the past years, only 118 genomes of *C. novyi sensu lato* are currently available in public databases. The challenging isolation process to isolate pure strains of *C. novyi sensu lato* may explain this low number of available genomes. There is actually no available selective media for the isolation of *C. novyi sensu lato*. Moreover strict anaerobic conditions are required, *C. haemolyticum* is indeed considered as one of the most fastidious and oxygen-sensitive anaerobes known (Hatheway, 1990). Numerous mobile genetic elements through a large plasmidome have been detected in *C. novyi sensu lato* and can be easily lost during cultivation (Fillo et al., 2021; Sakaguchi et al., 2005; Skarin et al., 2011; Skarin & Segerman, 2011, 2014). Lately, a new isolation protocol has been developed (Le Gratiet et al., 2020), opening new perspectives in the sequencing of this genospecies.

Among the 118 available genomes, only 5 are closed genomes (Woudstra et al., 2021). Working with fragmented assemblies could be problematic to track mobile genetic elements or to determine the exact location of some genes especially when they are flanked by repetitive insertion sequences (Bouso & Planet, 2019; Wick et al., 2017). Long-read sequencing seems very promising to overcome these drawbacks. Nevertheless, it requires high DNA quality and quantity, higher than for PCR, conventional typing or even for short-read sequencing like Illumina. The DNA extraction method should so take into account this first feature. MinION sequencing, one of the most commonly available techniques for long-read sequencing requires high DNA quantities (at least 400–1000 ng of high molecular weight DNA (average fragment size of more than 30 kb)) of high level quality (with low contamination measured through the ratios A260/280 and A260/230) (nanoporetech.com).

DNA extraction could be divided into 3 steps: cell disruption, nucleic acid release, removal of contaminants (Schultzhaus et al., 2019). Extracting high quality and quantity of DNA from *C. novyi sensu lato* could be challenging. Some clostridial species, including *C. botulinum* produce endogenous DNases resulting in DNA degradation (Hielm et al., 1998; Kristjánsson et al., 1994; Skarin et al., 2010). Clostridia are moreover spore-forming Gram-positive bacteria, and are therefore particularly difficult to lyse. Appropriate methods achieving cell lysis without DNase release is so expected to get high DNA quality suitable for WGS. No study has been published yet to establish one reliable protocol to perform DNA extraction for WGS purpose in the case of *C. novyi sensu lato*. The available studies that have sequenced *C. novyi sensu lato* strains have used different approaches (Fillo et al., 2021; Saeb et al., 2017; Skarin et al., 2011; Skarin & Segerman, 2014; Woudstra et al., 2017; Woudstra et al., 2021; Woudstra et al., 2015) and only one focused on long-read sequencing.

The objective of our study was to compare the efficiency of five DNA extraction protocols selected from previous studies on *C. novyi sensu lato*, *C. botulinum* Groups I and II and other Gram

positive spore-forming bacterial species (Knüpfer et al., 2020) to obtain DNA suitable for short reads and long-reads sequencing.

## **2. MATERIAL AND METHODS**

### **2.1. Bacterial strain and growth conditions**

A non-toxic *C. novyi sensu lato* strain (N17LNRB01) isolated from a duck liver initially positive for *C. botulinum* type C/D during a routine diagnosis (first analysis conducted in 2014) was selected for this study. Isolation was performed through successive plating on Egg Yolk Agar.

A suspension of N17LNRB01 spores was prepared as described previously (Hyun & Sakaguchi, 1989). Spores were stored in sterile water at 5°C. One mL of the spore suspension was ten-fold diluted in pre-reduced Tryptone Peptone Glucose Yeast (TPGY) broth in triplicates under anaerobic conditions at 37°C (A35, Don Whitley distributed by Biomérieux, Bruz, France). Two different incubation times were tested during this study: 16 hours corresponding to the end of exponential phase and 16 hours followed by a subculture of 6 hours by diluting 1 mL of the 16 hours culture in 49 mL of fresh TPGY broth so as to reach the beginning or the middle of the exponential phase.

After incubation, aliquots of 1 mL of enrichment broths were stored at -18°C until DNA extraction. The same cultures were used to be able to compare the different DNA extraction methods.

### **2.2. DNA extraction**

DNA extraction was performed using 4 commercial kits according to manufacturers' instructions and 1 in-house protocol.

The following commercial kits were evaluated: QiAmp DNA mini kit (Qiagen, Courtaboeuf, France), DNeasy Ultraclean Microbial kit (Qiagen, Courtaboeuf, France), MasterPure GRAM positive DNA purification kit (Epicentre, CA, USA), and MasterPure Complete DNA and RNA purification kit (Epicentre, CA, USA). They will be referred as QA, DN, MPG and MPC in the remainder of the article (Table 1).

The in-house protocol was performed as previously described (Woudstra et al., 2021). Briefly, cell were lysed in a buffer containing lysozyme (final concentration: 20 mg/ml) after an incubation of 10 minutes at 37°C. Proteins were then digested through a proteinase K treatment at 65 °C during 5 minutes and precipitated by this addition of ammonium acetate and finally pelleted by centrifugation. DNA was precipitated from this supernatant after addition of cold isopropanol, then collected and washed using ethanol 70. After air-drying, DNA was suspended in 100 µl Tris 10 mM.

All DNA were extracted using 1 mL of the six cultures. Extracted DNA were stored at -18 °C until further use.

### **2.3. DNA yield and quality**

DNA extract concentrations were assessed using a Qubit 2.0 fluorometer (Invitrogen, Thermo Fisher Scientific, Illkirch, France) and the Invitrogen Qubit dsDNA HS assay kit (Invitrogen, Thermo Scientific) according to the manufacturer's instructions. When DNA concentration was higher than 100 ng/μl, the DNA extract was ten-fold diluted to be at range of the assay.

The  $A_{260/280}$  and  $A_{260/230}$  absorbance ratios were measured using a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific) to evaluate DNA quality and purity.

### **2.4. DNA integrity**

Integrity of the extracted DNA was assessed using a 2100 Bioanalyzer system (Agilent, Les Ulis, France) with the Agilent DNA 12000 kit (Agilent). One microliter of each DNA extract was used for this assay according to the manufacturer instructions. Only DNA extracts with a DNA concentration above 10 ng/μl were tested.

### **2.5. Test of the absence of spores in the DNA extracts**

*C. novyi sensu lato* spores can be very resistant to some DNA extraction protocols and spores might remain viable after the extraction (Le Gratiet et al., 2020). Twenty μL of DNA extract were plated on EYA (Egg Yolk Agar) and incubated under anaerobic conditions at 37°C during 48 hours to evaluate the DNA extract contamination by residual spores.

## **3. RESULTS**

### **3.1. Comparison of the DNA extraction protocols**

Table 1 presents the characteristics of the different DNA extraction protocols tested in this study. Except for the DN kit which uses bead-beating for cell disruption, all protocols rely on physical (through vortexing and centrifugation) and chemical actions (Lysozyme, protease or cell lysis solution from the manufacturer's) for the lysis of vegetative cells. Regarding the DNA purification step, the DN and the QA kits use a silica-based column while the three other protocols are based on precipitation using isopropanol followed by washing steps using ethanol. From a practical point of view, commercial kits are user friendly while the in-house protocol requires extra time and dedicated caution for reagents preparation before DNA extraction. Duration of DNA extractions ranges between 50 and 150 minutes, QA being the most time-consuming one because of longer incubation steps and DN being the fastest one. All kits were in the same range of prices.

### **3.2. DNA Yield and quality**

Before the DNA yield measurements, potential contamination of DNA extracts incubated EYA was evaluated by plating DNA extracts on EYA. No *C. novyi sensu lato* colony was detected on EYA plates after incubation under anaerobic conditions during this study, confirming the absence of detectable viable spores in DNA extracts.

DNA yield measured using the Qubit 2.0 fluorometer (Figure 1) showed great variations depending on the DNA extraction protocol used. The in-house protocol extracted the highest quantity of DNA with a culture of 16 hours followed by an additional subculture of 6 hours allowing collection of 3140 ng followed by the MPC kit that allowed the collection of 2504 ng. DNA yield when using the other protocols was below 340 ng.

The purity of each DNA extract was measured by determining the absorbance ratios using the Nanodrop 1000 spectrophotometer (Figure 1). The first ratio is the DNA  $A_{260/280}$  ratio that evaluates the contamination of the DNA sample by proteins. DNA should have a  $A_{260/280}$  ratio above 1.8 (Bouso & Planet, 2019). Here in our study, all  $A_{260/280}$  ratios were above 1.76, suggesting low protein contamination in most samples. The second ratio commonly measured to evaluate the quality of DNA is the  $A_{260/230}$  ratio (evaluating the presence of other impurities) that should range between 1.8 and 2.2 (Yang et al., 2022). Among the methods tested here, ratios were within these orders of magnitude for only one kit, the MPC with an average  $A_{260/280}$  of 2,15 for both incubation times and an average  $A_{260/230}$  of 2,05 with a 16 hours incubation and 2,08 with a 16+6 hours incubation.  $A_{260/230}$  ratios were out of this range for the other protocols, suggesting presence of contaminants in DNA extracts. The low  $A_{260/230}$  ratios observed for QA and DN may be linked to the low DNA yields obtained using these kits.

### **3.3. Incubation time**

Two different culture durations were tested here: 16 hours or 6 hours after an overnight culture. DNA yield was higher for the most efficient kits after a subculture of 6 hours than when using a 16 hours culture (respectively 2504 and 2247 ng with the MPC kit, 3140 and 1527 ng with the in-house protocol).

### **3.4. DNA integrity**

Based on the DNA extraction yield, integrity of DNA extracts were only evaluated for MPC and the in-house method, which were the most efficient protocols. An average size of approximately 2900 base pairs (bp) was detected for MPC, while sizes of the DNA fragments were widely distributed from 2500 to more 12 000 bp for the in-house protocol.

## 4. DISCUSSION

The low availability of genomic data for the genospecies *C. novyi sensu lato*, especially closed genomes is a challenge to perform comparative genomics analyses and study mobile genetic elements. Selection of the right DNA extraction protocol is a crucial step toward WGS, especially for long-read sequencing. Indeed, without high quality DNA, sequencing-library preparations will be affected resulting in coverage depth variation (van Heesch et al., 2013). Quality and quantity can be altered during the bacterial lysis and the DNA purification steps due to several biological features of *C. novyi sensu lato* strains. For Gram positive and spore-forming bacteria, bacterial lysis is known to be difficult because of the thick layer of peptidoglycan of their vegetative cells (Cabeen & Jacobs-Wagner, 2005) and the complex wall structure of their spores that give them resistance to enzymes, chemical and mechanical actions (Knüpfer et al., 2020). In this study, five DNA extraction protocols were compared to assess their suitability for both short and long reads sequencing.

Results of this comparison showed that the in-house protocol from Woudstra et al. 2021 and the commercial kit MPC achieved high DNA quantity (above 0.5 µg), which is compatible with MinION sequencing (Bouso & Planet, 2019) and other WGS technologies. Woudstra et al. 2021 retrieved better integrity of DNA than MPC but lower  $A_{260/230}$  ratio. Low  $A_{260/230}$  ratio can cause osmotic imbalances in the flow cell (Bouso & Planet, 2019). Extra washing steps should perhaps be added to improve this ratio before sequencing. DNA quantities from the 3 other DNA extraction protocols were too low for WGS.

MPC kit showed one of the highest extracted DNA yield, which has already reported previously through the comparison of other DNA extraction kits for DNA extraction of *Yersinia pestis* and the spore forming bacteria *Bacillus anthracis* for example (Dauphin et al., 2010; Knüpfer et al., 2020) and successful compatibility with MinION sequencing (Anwar et al., 2020; Biggel et al., 2022).

Surprisingly, despite being specifically designed for Gram-positive bacteria and using the same principle as the MPC kit, the MPG kit did not show the expected results with low DNA yield, not suitable for WGS.

As previously reported (Woudstra et al., 2021), it seems that a sub-culture of 6 hours after an overnight culture of spores is the most suitable cultivation way before DNA extraction, based on DNA yield obtained in our study for both methods.

## 5. CONCLUSION

This study shows that two methods could be considered for sequencing of *C. novyi sensu lato* using MinION: MasterPure Complete DNA and RNA purification kit as well as the in-house method developed in Woudstra et al. 2021.

## 6. Acknowledgments:

This study was supported by the “Institut Carnot AgriFood Transition” (TRACK’SPORE project), the French National Research Agency, and the Bretagne Regional Council of Région Bretagne. Thibault Le Gratiet is the recipient of a PhD grant from ANSES and “Région Bretagne.”

The authors are grateful to Cédric Woudstra for his help and advice on the protocol he developed for the extraction of DNA for long-read sequencing.

## 7. Conflicts of Interest:

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## 8. REFERENCES

- Anwar, M. Z., Zervas, A., Hansen, L. H., Barker, G., Anesio, A. M., & Jacobsen, C. S. (2020). Complete Genome and Plasmid Sequences of *Salmonella enterica* subsp. *enterica* Serovar Enteritidis PT1, Obtained from the *Salmonella* Reference Laboratory at Public Health England, Colindale, United Kingdom. *Microbiology resource announcements*, *9*(2). doi:10.1128/mra.01064-19
- Bano, L., Drigo, I., Tonon, E., Pascoletti, S., Puiatti, C., Anniballi, F., et al. (2017). Identification and characterization of *Clostridium botulinum* group III field strains by matrix-assisted laser desorption-ionization time-of-flight mass spectrometry (MALDI-TOF MS). *Anaerobe*, *48*, 126-134. doi:10.1016/j.anaerobe.2017.08.004
- Biggel, M., Zurfluh, K., Hoehn, S., Schmitt, K., Frei, A., Jans, C., et al. (2022). Complete Genome Sequence of *Hafnia paralvei* Isolate AVS0177, Harboring *mcr-9* on a Plasmid. *Microbiology resource announcements*, *11*(1), e00966-00921. doi:doi:10.1128/MRA.00966-21
- Bouso, J. M., & Planet, P. J. (2019). Complete nontuberculous mycobacteria whole genomes using an optimized DNA extraction protocol for long-read sequencing. *BMC Genomics*, *20*(1), 793. doi:10.1186/s12864-019-6134-y
- Cabeen, M. T., & Jacobs-Wagner, C. (2005). Bacterial cell shape. *Nature Reviews Microbiology*, *3*(8), 601-610. doi:10.1038/nrmicro1205
- Dauphin, L. A., Stephens, K. W., Eufinger, S. C., & Bowen, M. D. (2010). Comparison of five commercial DNA extraction kits for the recovery of *Yersinia pestis* DNA from bacterial suspensions and spiked environmental samples. *J Appl Microbiol*, *108*(1), 163-172. doi:https://doi.org/10.1111/j.1365-2672.2009.04404.x
- Fillo, S., Giordani, F., Tonon, E., Drigo, I., Anselmo, A., Fortunato, A., et al. (2021). Extensive Genome Exploration of *Clostridium botulinum* Group III Field Strains. *Microorganisms*, *9*(11), 2347. doi:10.3390/microorganisms9112347

- Hatheway, C. L. (1990). Toxigenic clostridia. *Clinical microbiology reviews*, 3(1), 66-98. doi:doi:10.1128/CMR.3.1.66
- Hielm, S., Björkroth, J., Hyytiä, E., & Korkeala, H. (1998). Genomic analysis of *Clostridium botulinum* group II by pulsed-field gel electrophoresis. *Appl Environ Microbiol*, 64(2), 703-708. doi:10.1128/AEM.64.2.703-708.1998
- Hyun, S. H., & Sakaguchi, G. (1989). Implication of coprophagy in pathogenesis of chicken botulism. *Nihon Juigaku Zasshi*, 51(3), 582-586. doi:10.1292/jvms1939.51.582
- Knüpfer, M., Braun, P., Baumann, K., Rehn, A., Antwerpen, M., Grass, G., et al. (2020). Evaluation of a highly efficient DNA extraction method for *Bacillus anthracis* endospores. *Microorganisms*, 8(5), 763. doi:10.3390/microorganisms8050763
- Kristjánsson, M., Samore, M. H., Gerding, D. N., DeGirolami, P. C., Bettin, K. M., Karchmer, A. W., et al. (1994). Comparison of restriction endonuclease analysis, ribotyping, and pulsed-field gel electrophoresis for molecular differentiation of *Clostridium difficile* strains. *J Clin Microbiol*, 32(8), 1963-1969. doi:10.1128/jcm.32.8.1963-1969.1994
- Le Gratiet, T., Poezevara, T., Rouxel, S., Houard, E., Mazuet, C., Chemaly, M., et al. (2020). Development of an innovative and quick method for the isolation of *Clostridium botulinum* strains involved in avian botulism outbreaks. *Toxins*, 12(1), 42.
- Nakamura, S., Kimura, I., Yamakawa, K., & Nishida, S. (1983). Taxonomic relationships among *Clostridium novyi* Types A and B, *Clostridium haemolyticum* and *Clostridium botulinum* type C. *J Gen Microbiol*, 129(5), 1473-1479. doi:10.1099/00221287-129-5-1473
- Saeb, A. T., Abouelhoda, M., Selvaraju, M., Althawadi, S. I., Mutabagani, M., Adil, M., et al. (2017). The Use of Next-Generation Sequencing in the Identification of a Fastidious Pathogen: A Lesson From a Clinical Setup. *Evolutionary bioinformatics online*, 12, 1176934316686072-1176934316686072. doi:10.1177/1176934316686072
- Sakaguchi, Y., Hayashi, T., Kurokawa, K., Nakayama, K., Oshima, K., Fujinaga, Y., et al. (2005). The genome sequence of *Clostridium botulinum* type C neurotoxin-converting phage and the molecular mechanisms of unstable lysogeny. *Proceedings of the National Academy of Sciences of the United States of America*, 102(48), 17472-17477. doi:10.1073/pnas.0505503102
- Sasaki, Y., Takikawa, N., Kojima, A., Norimatsu, M., Suzuki, S., & Tamura, Y. (2001). Phylogenetic positions of *Clostridium novyi* and *Clostridium haemolyticum* based on 16S rDNA sequences. *International Journal of Systematic and Evolutionary Microbiology*, 51(3), 901-904. doi:https://doi.org/10.1099/00207713-51-3-901
- Schultzhause, J. N., Taitt, C. R., Orihuela, B., Smerchansky, M., Schultzhause, Z. S., Rittschof, D., et al. (2019). Comparison of seven methods for DNA extraction from prosomata of the acorn barnacle, *Amphibalanus amphitrite*. *Analytical Biochemistry*, 586, 113441. doi:https://doi.org/10.1016/j.ab.2019.113441
- Skarin, H., Hafstrom, T., Westerberg, J., & Segerman, B. (2011). *Clostridium botulinum* group III: a group with dual identity shaped by plasmids, phages and mobile elements. *BMC Genomics*, 12(1), 185. doi:10.1186/1471-2164-12-185
- Skarin, H., Lindberg, A., Blomqvist, G., Aspán, A., & Båverud, V. (2010). Molecular characterization and comparison of *Clostridium botulinum* type C avian strains. *Avian Pathol*, 39(6), 511-518. doi:10.1080/03079457.2010.526923
- Skarin, H., & Segerman, B. (2011). Horizontal gene transfer of toxin genes in *Clostridium botulinum*: Involvement of mobile elements and plasmids. *Mob Genet Elements*, 1(3), 213-215. doi:10.4161/mge.1.3.17617
- Skarin, H., & Segerman, B. (2014). Plasmidome interchange between *Clostridium botulinum*, *Clostridium novyi* and *Clostridium haemolyticum* converts strains of independent lineages into distinctly different pathogens. *PLOS ONE*, 9(9), e107777. doi:10.1371/journal.pone.0107777
- van Heesch, S., Mokry, M., Boskova, V., Junker, W., Mehon, R., Toonen, P., et al. (2013). Systematic biases in DNA copy number originate from isolation procedures. *Genome biology*, 14(4), R33-R33. doi:10.1186/gb-2013-14-4-r33

- Wick, R. R., Judd, L. M., Gorrie, C. L., & Holt, K. E. (2017). Completing bacterial genome assemblies with multiplex MinION sequencing. *Microb Genom*, 3(10), e000132. doi:10.1099/mgen.0.000132
- Woudstra, C., Brito, R. B., Júnior, A. A. F., Silva, R. O. S., Lobato, F. C. F., & Fach, P. (2017). Draft Genome Sequences of Five Brazilian *Clostridium botulinum* Group III Type D/C Strains. *Genome Announc*, 5(20), e00349-00317. doi:doi:10.1128/genomeA.00349-17
- Woudstra, C., Le Maréchal, C., Souillard, R., Bayon-Auboyer, M.-H., Mermoud, I., Desoutter, D., et al. (2016). New Insights into the Genetic Diversity of *Clostridium botulinum* Group III through Extensive Genome Exploration. *Frontiers in microbiology*, 7, 757-757. doi:10.3389/fmicb.2016.00757
- Woudstra, C., Maklin, T., Derman, Y., Bano, L., Skarin, H., Mazuet, C., et al. (2021). Closing *Clostridium botulinum* Group III Genomes Using Long-Read Sequencing. *Microbiology resource announcements*, 10(22), e0136420. doi:10.1128/MRA.01364-20
- Woudstra, C., Le Maréchal, C., Souillard, R., Bayon-Auboyer, M.-H., Mermoud, I., Desoutter, D., et al. (2015). Draft Genome Sequences of 17 French *Clostridium botulinum* Group III Strains. *Genome Announc*, 3(5), e01105-01115. doi:doi:10.1128/genomeA.01105-15
- Yang, S., Johnson, M. A., Hansen, M. A., Bush, E., Li, S., & Vinatzer, B. A. (2022). Metagenomic sequencing for detection and identification of the boxwood blight pathogen *Calonectria pseudonaviculata*. *Scientific Reports*, 12(1), 1399. doi:10.1038/s41598-022-05381-x



Table 1: List of the five DNA extraction protocols used in this study. \*: Price on manufacturer web page accessed March 2022 (without shipping costs) \*\* DNeasy Blood and Tissue kit and QiaAmp DNA mini kit are similar for the DNA extraction of Gram positive bacteria.

<b>DNA extraction kit</b>	<b>Short name used in the article</b>	<b>Cell disruption method</b>	<b>DNA Purification method</b>	<b>Elution volume (µl)</b>	<b>Completion Time (minutes)</b>	<b>Average Cost (€/sample)*</b>	<b>References</b>
Qiagen QIAamp DNA Mini Kit	QA	Physical and Chemical lysis (Lysozyme 20 mg/ml, 30°-37°C and 15°- 95°C)	Silica column	200	150	4,24	(Woudstra et al., 2016)**
Qiagen DNeasy UltraClean Microbial kit	DN	Bead beating	Silica column	50	50	4,92	(Knüpfer et al., 2020)
Lucigen MasterPure Gram Positive DNA purification kit	MPG	Physical and Chemical lysis (Lysozyme 30 000 U/µl, 30°-37°C and Cell lysis solution 15°- 65°C)	Isopropanol precipitation	35	90	2,90	(Fillo et al., 2021)
Lucigen MasterPure Complete DNA and RNA purification kit	MPC	Physical and Chemical lysis (Cell lysis solution, 15°- 65°C)	Isopropanol precipitation	35	90	2	(Knüpfer et al., 2020)
In-House protocol	In-house	Physical and Chemical lysis (Lysozyme: 20 mg/ml, 10°-37°C)	Isopropanol precipitation	100	75	/	(Woudstra et al., 2021)

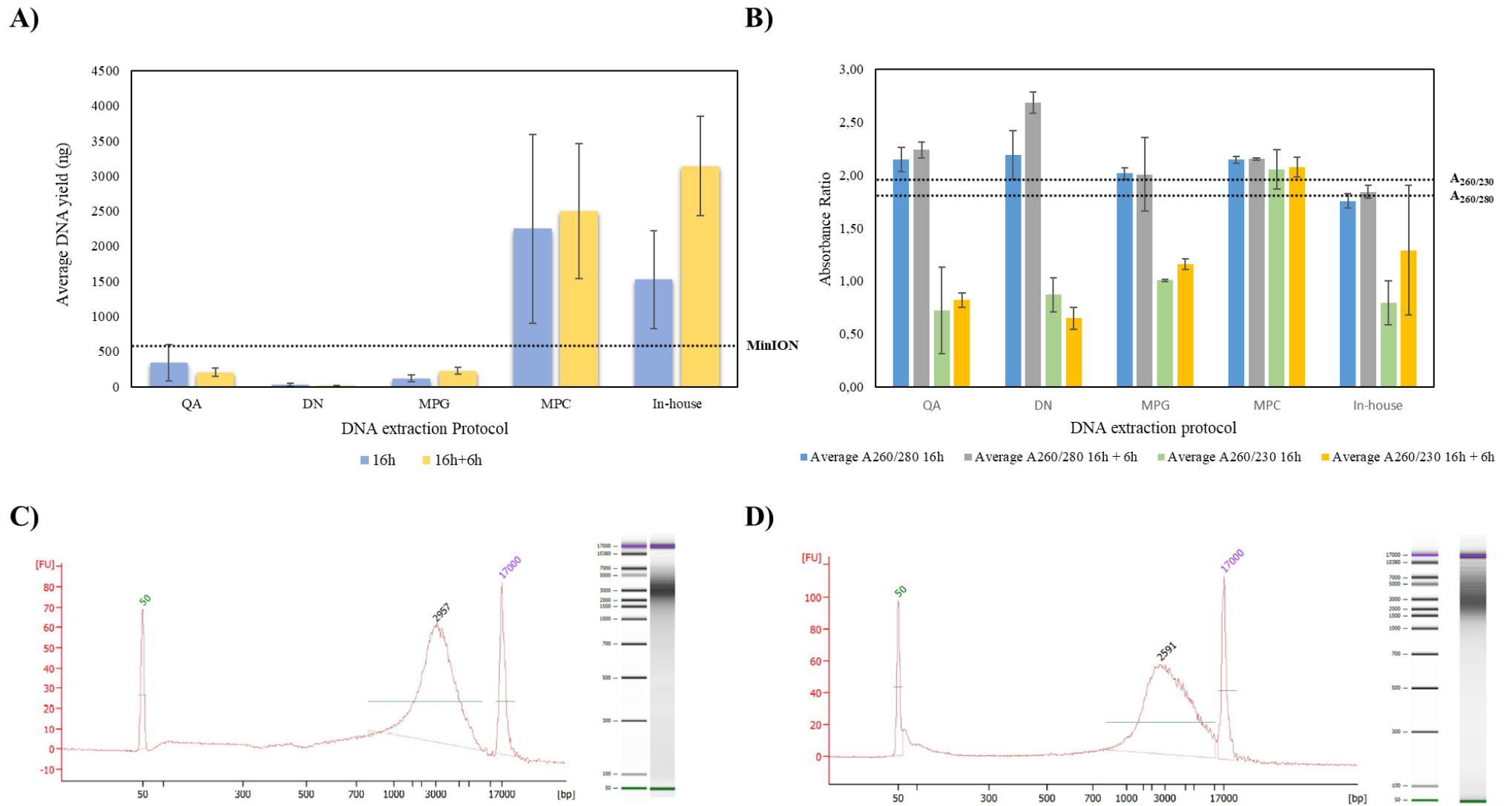


Figure 1: Quantity, quality and integrity of the extracted DNA for 5 different DNA extraction protocols.

A) Average DNA yield obtained with Qubit B) Average absorbance ratios obtained with NanoDrop C) Capillary electrophoresis of a DNA extracted with MPC D) Capillary electrophoresis of a DNA extracted with the in-house protocol.

### 1.1.3. Conclusions générales de l'étude et perspectives

Cette étude montre que le protocole d'extraction d'ADN utilisé a un impact sur la qualité et la quantité d'ADN obtenues. Il est nécessaire d'adapter certaines étapes au cas particulier de *C. novyi sensu lato*, notamment au niveau de l'étape de lyse cellulaire, à la fois pour extraire suffisamment d'ADN mais également pour prévenir la libération ou l'action de DNases. Il est ainsi nécessaire de déterminer la durée de culture optimale pour avoir une proportion majoritaire de cellules végétatives dans la culture lors de l'extraction, de modifier le temps d'incubation lors de la lyse bactérienne pour prévenir l'action des DNases à 37 °C et de disposer d'une quantité de bactéries suffisantes au départ pour obtenir la quantité d'ADN souhaitée.

Dans cette étude, deux protocoles se sont révélés compatibles avec le séquençage. Le kit « MasterPure Complete DNA and RNA purification » permet d'obtenir une grande quantité d'ADN de bonne qualité ( $A_{260/280}$  et  $A_{260/230}$ ) mais la taille des fragments d'ADN est en moyenne inférieure à 3000 pb. Le protocole développé dans l'étude menée par Woudstra et al. (C. Woudstra et al., 2021) permet quant à lui d'obtenir une plus grande quantité d'ADN de bonne qualité (sauf le ratio  $A_{260}/A_{230}$ ) avec des tailles de fragments d'ADN comprises entre 2500 pb et 12000 pb. Le kit MasterPure Complete DNA and RNA purification peut être utilisé pour préparer de l'ADN en vue d'un séquençage Illumina. Ce kit et le protocole développé par Woudstra et al. 2021 semblent compatibles avec un séquençage long-reads, ce qui reste à confirmer avec le séquençage MinION de la souche N17LNRB01.

L'objectif final est de coupler les données de séquençage du MinION et les résultats de séquençage short-reads déjà obtenus afin d'obtenir un génome complet circularisé (chromosome et EGM) selon l'approche présentée dans une étude récente sur *C. botulinum* du groupe I (Gonzalez-Escalona & Sharma, 2020).

## 1.2. Préparation des bibliothèques (Séquençage short-reads)

Lors du séquençage avec la technologie Illumina, l'étape suivant l'extraction d'ADN et son dosage est la préparation des bibliothèques. Cette étape se divise en deux parties. Pendant la première partie, l'ADN extrait est fragmenté puis amplifié par PCR. Des séquences index adaptatrices sont ensuite insérées aux extrémités des fragments d'ADN pour que les séquences de la souche puissent être identifiées lors du séquençage. En effet, plusieurs souches sont séquencées simultanément lors d'un séquençage, des séquences index différentes vont donc être attribuées à chaque souche pour pouvoir les identifier lors de l'exploitation des résultats. La deuxième partie consiste à normaliser les bibliothèques : chaque échantillon est quantifié par fluorescence puis la taille moyenne des

fragments est déterminée. La molarité des échantillons est ensuite calculée à partir de la concentration en ADN et de la taille moyenne des fragments. Les échantillons sont ensuite dilués pour être tous à la même concentration. Enfin, après une dernière étape de dilution et de dénaturation, les échantillons sont mélangés et le séquençage est lancé.

Pour les souches de *C. novyi sensu lato* isolées au cours de ma thèse, le séquençage des souches de *C. botulinum* a été effectué à la plateforme Identypath de l'Anses de Maisons-Alfort ou au CNR des bactéries anaérobies et du botulisme de l'Institut Pasteur de Paris car elles détiennent les autorisations MOT nécessaires à la manipulation du matériel génétique de *C. botulinum*.

La préparation des bibliothèques a été faite avec le kit Illumina Nextera XT (Illumina, CA, USA). Ce kit utilise une fragmentation par tagmentation : un transposon coupe à la fois l'ADN génomique et insère les séquences adaptatrices (Adey et al., 2010). Suite à la fragmentation et à l'amplification par PCR, les résultats de quantification (réalisée au Qubit et au BioAnalyzer) ont révélé que les concentrations d'ADN étaient trop faibles (Tableau 10) pour avoir une couverture de séquençage suffisante et que les tailles des fragments n'étaient pas uniformes. Une des hypothèses émises pour expliquer ce résultat est le pourcentage en GC bas caractéristique de *C. botulinum*, cas de figure pour lequel le kit Nextera XT ne semble pas adapté. Une étude de 2019 avait déjà montré que ce kit n'était pas adapté aux génomes présentant des pourcentages en GC trop élevés ou trop bas, les résultats de séquençage révélant une couverture inégale et basse des génomes (Sato et al., 2019). Cette même étude a montré que pour le kit Illumina Nextera DNA Flex, le biais de couverture lié au pourcentage en GC bas n'était pas retrouvé (Sato et al., 2019). Dans notre étude, ce kit a été utilisé en seconde intention sur 7 extraits d'ADN et les résultats ont montré qu'il fonctionne mieux que le kit Illumina Nextera XT, les concentrations d'ADN étant plus importantes (Tableau 10) et les tailles des fragments plus uniformes lors des analyses.

Tableau 10 : Quantités d'ADN amplifiés selon le kit de préparation des bibliothèques utilisé

Souches	Nextera XT (ng/μl)	Nextera DNA Flex (ng/μl)
B19LNRB01	0,486	16,9
B19LNRB07	0,5	25
B19LNRB27	0,166	22
B19LNRB12	0,67	22,2
B19LNRB13	0,736	20,4
19LNRB22-4	0,426	18,4
N19LNRB03	0,556	21,4

Pour les souches non toxiques du groupe *C. novyi sensu lato* isolées au cours de la thèse, le séquençage a été effectué par la plateforme de séquençage iGenSeq basée à l'Institut du Cerveau de Paris (ICM) via la plateforme de séquençage de l'unité de Génétique Virale et Biosécurité (GVB) de l'Anses de Ploufragan. Cette plateforme utilise le kit Nextera XT pour la préparation des bibliothèques. Cependant, les couvertures obtenues étaient plus importantes que celles obtenues lors du séquençage effectué à la plateforme IdentityPath. Le même kit d'extraction d'ADN a été utilisé dans les deux cas. La profondeur de séquençage était plus importante à l'ICM (qui dispose de deux séquenceurs haut débit NovaSeq 6000 et NextSeq 500) qu'à la plateforme IdentityPath (qui dispose d'un séquenceur moyen débit MiSeq) lors du séquençage des souches, ce qui pourrait avoir compensé les éventuels problèmes de rendement lors de la préparation des bibliothèques. Aucun problème n'a été signalé par l'ICM à ce jour au niveau de l'étape de préparation des bibliothèques. Le kit de préparation des bibliothèques utilisé dans les articles de la littérature pour le séquençage des différents groupes de *C. botulinum* varie selon les études (Tableau 11). Pour les souches de *C. novyi sensu lato*, c'est le kit Nextera XT qui est utilisé dans la plupart des études (Fillo et al., 2021; Skarin & Segerman, 2014; Woudstra et al., 2017; Woudstra et al., 2016). Pour les souches de *C. botulinum* des groupes I et II, les plus récentes études de séquençage de génome short-reads n'utilisent pas le kit Nextera XT pour la préparation des bibliothèques mais les kits Nextera DNA flex, Illumina TruSeq® DNA ou Roche KAPA HyperPlus (Dover et al., 2021; Gladney et al., 2021; Gonzalez-Escalona & Sharma, 2020).

Tableau 11 : Exemples de kits utilisés pour la préparation des bibliothèques pour le séquençage Illumina de *C. novyi* et différents groupes de *C. botulinum*. Le kit avec une \* est utilisé pour une fragmentation mécanique.

Espèce ou groupe testé	Nextera XT	Nextera DNA flex	Autre kit	Référence
<i>C. novyi sensu lato</i>	X			(Skarin & Segerman, 2014)
<i>C. novyi sensu lato</i>	X			(Woudstra et al., 2016)
<i>C. novyi sensu lato</i>	X			(Woudstra et al., 2017)
<i>C. novyi sensu lato</i>	X			(Fillo et al., 2021)
<i>C. novyi sensu lato</i> ( <i>C. haemolyticum</i> )			TruSeq® DNA kit (Illumina)*	(Lagneaux et al., 2021)
<i>C. botulinum</i> groupe I		X		(Gonzalez-Escalona, 2020)
<i>C. botulinum</i> Type B			KAPA HyperPlus kit (Roche)	(Dover et al., 2021)
<i>C. botulinum</i> Type A		X		(Gladney et al., 2021)

Sur la base de ces résultats et des données disponibles dans la littérature, l'utilisation du kit Illumina Nextera DNA Flex semble plus adaptée au séquençage des génomes du groupe *C. novyi sensu lato*.

## 2. Analyses comparatives des génomes de *C. novyi sensu lato*

### 2.1. Analyses comparatives de l'ensemble du génome

#### 2.1.1. Comparaison de la taille des génomes par lignée

La taille des 118 génomes de *C. novyi sensu lato* a été comparée selon les lignées (Tableau 12). Les résultats montrent que la taille moyenne des génomes de la lignée Ia et Ib (respectivement 3,07 et 3,13 Mb) est plus grande que celles des lignées II, III et IV (respectivement 2,62, 2,84 et 2,49 Mb). Dans les bases de données, seulement 5 génomes ont été séquencés complètement (Skarin et al., 2011; Woudstra et al., 2021). D'après l'analyse des 5 génomes complets disponibles (1 de lignée Ia, 2 de lignée Ib et 2 de lignée II), une différence de taille est retrouvée au niveau du chromosome qui est plus grand pour les souches de lignée Ia et Ib (2,78 Mb pour BKT015925, 2,84 Mb pour BKT2873 et 2,88 Mb pour 3859/11) comparé aux souches de lignée II (2,20 Mb pour 1873 et 2,30 Mb pour C-Stockholm) (Skarin & Segerman, 2011; Woudstra et al., 2021). D'après les études effectuées sur le plasmidome, le nombre d'EGM qui varie en fonction des lignées influe également sur la taille du

génomique (Skarin et al., 2011; Skarin & Segerman, 2011, 2014; Woudstra et al., 2018; Woudstra et al., 2016).

Ces premières observations restent à confirmer, le nombre de génomes disponibles étant encore assez faible, notamment pour les lignées III et IV.

Tableau 12 : Tailles moyennes des génomes (chromosome et EGM) de *C. novyi sensu lato* en fonction des lignées.

Lignée	Ia	Ib	II	III	IV	Total
Nombre de génomes disponibles	47	34	27	3	7	118
Taille moyenne du génome (Mb)	3,08 (± 0,08)	3,13 (± 0,18)	2,62 (± 0,12)	2,84 (± 0,11)	2,49 (± 0,09)	2,94 (± 0,26)
Taille max (Mb)	3,20	3,47	2,87	2,96	2,61	3,47
Taille mini (Mb)	2,82	2,57	2,30	2,80	2,32	2,30

### 2.1.2. Séquençage de souches isolées pendant le projet de thèse

Sur les 35 isolats séquencés au cours du projet (18 et 17 souches isolées pendant et après le développement de la méthode d'isolement, respectivement), seuls 6 génomes sont exploitables. Pour les autres génomes, la taille obtenue varie entre 3.5 Mb et 8 Mb, ce qui est supérieure à la taille maximale retrouvée chez *C. novyi sensu lato*, suggérant la présence de contaminants dans les séquences obtenues.

En analysant les données plus précisément, les données de séquençage de certaines souches présentent des séquences d'autres clostridies (*C. perfringens* et *C. difficile* entre autres). Des enrichissements de ces souches ont été effectués puis ont été ensemencés sur des milieux gélosés sélectifs pour *C. difficile* (Milieu chromID *C. difficile*, Biomérieux) et *C. perfringens* (milieu Tryptone Sulfite Cyclo-sérine ou TSC) pour mettre en évidence une éventuelle contamination bactérienne lors de l'étape d'isolement. Cependant, aucune colonie n'a pu être observée après plusieurs jours d'incubation en conditions anaérobies, ne confirmant pas l'hypothèse de la contamination à cette étape. Une contamination lors de l'étape d'extraction d'ADN peut également être suspectée, celle-ci n'a pas été explorée pour le moment.

Pour d'autres souches, les résultats semblent indiquer la présence simultanée de plusieurs souches de *C. novyi sensu lato*. Les outils bio-informatiques disponibles pour le moment à l'Anses ne nous ont pas permis de distinguer les génomes des différentes souches, ces génomes étant trop proches pour réassigner les séquences à chacune des souches. Les outils utilisés récemment dans le

cadre des investigations pour tracer des souches d'*Escherichia coli* après un séquençage « Shotgun » (métagénomique) pourraient être testés dans le cas de cette étude (Saltykova et al., 2020).

### 2.1.3. Analyses kmer

Le logiciel Gegenees utilisé lors des différentes comparaisons de génomes de *C. novyi sensu lato* (Fillo et al., 2021; Skarin & Segerman, 2014; Woudstra et al., 2016) étant devenu obsolète sur les dernières versions de Windows, un nouvel outil bio-informatique développé à l'Anses avant ce projet a été utilisé (en collaboration avec l'unité GVB et un des développeurs de l'outil) afin de comparer les génomes de *C. novyi sensu lato* séquencés au cours de cette thèse avec ceux déjà disponibles (<https://github.com/afelten-Anses/QuickPhylo>). Ce workflow utilise Mash (Ondov et al., 2016) afin de calculer une estimation de la distance de Jaccard à partir de kmers extraits depuis les génomes. À partir des distances calculées, un arbre phylogénétique est créé à l'aide de l'algorithme de Neighbor-Joining. Cet arbre est ensuite enraciné par une approche de Midpoint.

Au total, l'outil a été utilisé pour comparer 125 génomes de *C. novyi sensu lato* (118 disponibles dans les bases de données publiques et les 7 génomes exploitables parmi les 35 génomes séquencés au cours de la thèse). Les 7 génomes séquencés lors de la thèse proviennent d'une souche de *C. botulinum* de type C/D (B19LNRB28) impliquée dans un épisode de botulisme en faune sauvage en 2018, de 2 souches de *C. botulinum* de type D (B20LNRB05 et B20LRB20) impliquées dans des épisodes de botulisme dans des élevages de dinde en 2015 et 2020 et de 4 souches de *C. novyi* où le phage contenant le gène *bont* a été perdu au cours de l'isolement. Trois de ces souches étaient initialement porteuses du gène *bont* de type D et ont été impliquées dans des épisodes de botulisme dans des élevages de dindes (N20LNRB01 et N19LNRB07 en 2019, N20LNRB17 en 2015). Enfin, la souche N17LNRB01 a été isolée avant la thèse par le LNR botulisme aviaire et provient d'un épisode de botulisme en faune sauvage en 2014 où la toxine détectée initialement était de type C/D. La comparaison de l'ensemble de ces génomes a permis la construction d'un arbre phylogénétique (Figure 25).



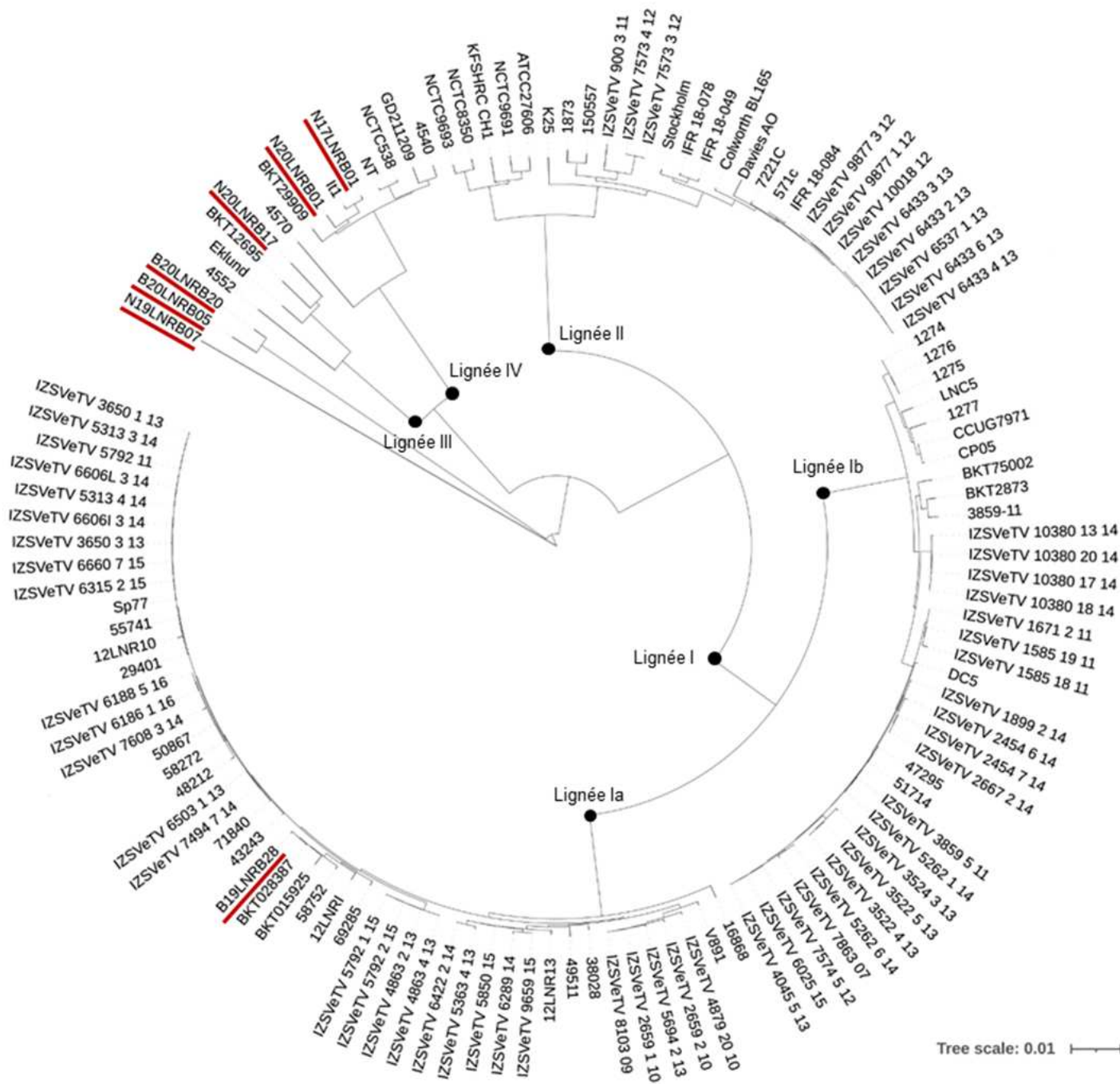


Figure 25 : Arbre phylogénétique présentant la distribution des souches de *C. novyi sensu lato* sur la base d'une analyse kmer des génomes. Les souches soulignées en rouge sont celles qui ont été séquencées au cours de la thèse.

### 2.1.3.1. Résultats de la comparaison des génomes

L'arbre phylogénétique obtenu à partir de l'analyse comparative des kmers extraits des génomes de *C. novyi sensu lato* montre une répartition des souches très similaire à celle obtenue via l'utilisation de Gegenees (Fillo et al., 2021; Skarin & Segerman, 2014), avec un groupement des souches par lignée (I à IV). Pour les souches séquencées au cours de la thèse, B19LNRB28 fait partie de la lignée Ia. Les souches N20LNRB17 et N20LNRB01 qui possédaient la toxine de type D font respectivement parties de la lignée IV et III, ce qui semble indiquer que les souches de *C. botulinum* du type D associées à des épisodes de botulisme dans des élevages de dindes en France n'ont pas toutes une origine commune. La souche N17LNRB01 fait quant à elle partie de la lignée IV.

Les souches B20LNRB05, B20LNRB20, qui sont aussi des souches possédant le gène *bont D*, ne se groupent avec aucune des lignées connues de *C. novyi sensu lato* malgré une taille de génome proche des tailles attendues pour ce groupe d'espèces (respectivement 2.6 et 2.7 Mb). La souche N19LNRB07 n'a pas été groupée dans une lignée, mais la taille de son génome est plus grande qu'attendue (3,6 Mb). La première hypothèse qui expliquerait ces résultats est la présence de séquences d'autres espèces bactériennes liées à un problème lors de l'isolement de la souche. C'est probablement le cas pour la souche N19LNRB07 qui présente un génome plus grand que la taille moyenne des souches de *C. novyi sensu lato* (2,94 Mb). Une autre hypothèse pour les souches B20LNRB05 et B20LNRB20 pourrait être l'appartenance de ces souches de *C. botulinum* à une lignée différente des 4 déjà définies et qui n'a pas encore été caractérisée à ce jour. Au cours de la thèse, ces souches de *C. botulinum* de type D ont été soumises aux différents systèmes PCR décrits dans la littérature et ciblant les gènes de flagellines ou les EGM (Woudstra et al., 2018; Woudstra et al., 2015) pour essayer de déterminer leur origine. Dans le chapitre précédent, la MLVA a ainsi été testée sur ces souches et les résultats ont montré l'absence d'amplification pour les 7 VNTR. Des PCR développées lors d'études précédentes (Woudstra et al., 2018; Woudstra et al., 2015) ont aussi été effectuées sur ces souches pour voir si les gènes *fliC* ou des EGM déjà caractérisés sont présents. Cependant, ces PCR n'ont pas permis l'obtention d'une amplification. Seules les PCR de détection du gène *bont D* (Woudstra et al., 2012) et d'un gène spécifique de *C. novyi sensu lato* (Woudstra et al., 2015) ont permis d'obtenir une amplification caractéristique, renforçant l'hypothèse que ce sont des souches de *C. botulinum* différentes des souches des autres lignées. Cependant, des études complémentaires sont nécessaires pour pouvoir confirmer ou infirmer cette hypothèse et explorer cette hypothèse de l'existence d'une lignée V dans le groupe *C. novyi sensu lato*.

### 2.1.3.2. Séquençage d'isolats impliqués dans un épisode de botulisme aviaire

Dans le cadre de l'étude présentée précédemment (cf. 2.3.2 du chapitre sur la MLVA), la comparaison par MLVA de 8 enrichissements positifs pour *C. botulinum* de type C/D a montré la présence d'un même profil dans tous les échantillons. Un isolement a été effectué à partir de 8 enrichissements provenant des différents épisodes de botulisme de l'étude selon la méthode développée au cours de la thèse dans l'optique de séquencer le génome des 8 isolats puis de les comparer pour vérifier s'il s'agit bien d'une seule et même souche comme le suggèrent les résultats de comparaison par MLVA.

Les isolats toxigènes ont été séquencés au CNR des bactéries anaérobies et du botulisme de l'Institut Pasteur (laboratoire avec autorisation MOT) tandis que les souches ayant perdu le gène *bont* au cours de l'isolement ont été séquencées à l'ICM. De plus, deux souches toxigènes ont été isolées en parallèle en Allemagne<sup>1</sup>. Les données de séquences obtenues ont été comparées aux génomes de la base de données avec l'outil bio-informatique présenté précédemment (analyse K-mer).

L'analyse comparative des génomes a révélé que les souches isolées en Allemagne font parties de la lignée Ia (Figure 26, soulignées en vert), confirmant les résultats de la MLVA où les profils des 9 échantillons se trouvent dans un cluster avec la souche BKT015925 (cf chapitre 5 partie MLVA – Figure 1 de l'article page 131).

Concernant les souches isolées dans le cadre de mon projet de thèse à Ploufragan, les 4 souches non toxigènes se regroupent avec des souches de la lignée IV qui est composée majoritairement de souches de *C. novyi* (Figure 26, soulignées en bleu). Les 4 souches toxigènes présentent quant à elles des tailles trop grandes pour être exploitables (8 Mb au lieu de 3,08 Mb en moyenne pour une souche de lignée Ia) et ne sont pas représentées sur la figure 26. Ce résultat suggère la présence de plusieurs souches dans l'échantillon de départ. Le génome des souches non toxique de la lignée IV et de la souche isolée en Allemagne ont été recherchés dans les séquences des 4 souches toxigènes mais aucune concordance n'a été mise en évidence. Le gène de la toxine C et non de la toxine C/D a par ailleurs été identifié dans les séquences obtenues.

Ces résultats mettent en évidence la présence d'au moins deux souches de *C. novyi sensu lato* dans les échantillons provenant de cet élevage : une appartenant à la lignée Ia (*C. botulinum* type C/D, d'après les données de séquençage du RKI) et l'autre à la lignée IV (*C. novyi*, d'après les données de séquences de l'ICM). Les résultats de séquençage des 4 souches toxiques isolées à Ploufragan

---

<sup>1</sup> Par Martin Dorner du Robert Koch Institute (RKI) dans le cadre d'une collaboration.

suggèrent la présence d'autres souches de *C. novyi sensu lato* dans ces échantillons. Ces deux espèces présentant la même morphologie sur gélose EYA, il est possible que la présence de ces différentes souches n'ait pas été identifiée au moment de l'isolement.

En recherchant les régions VNTR ciblées par la méthode MLVA avec l'outil BLAST dans les génomes de souches de la lignée IV isolées dans le cadre de cette étude, une seule des 7 cibles a été retrouvée (VNTR16) avec des mutations au niveau de la séquence qui sert pour une des deux amorces. Ce résultat explique qu'un seul profil ait été mis en évidence via l'analyse MLVA et que la présence d'une seconde souche dans l'échantillon n'ait pas été mise en évidence.

Ce résultat montre une limite à la méthode d'isolement qui n'avait pas été identifiée lors de la mise au point de la méthode. La présence de plusieurs souches de *C. novyi sensu lato* dans l'organe de l'animal complique en effet cette étape d'isolement et l'obtention d'une souche pure. L'équipe du Robert Koch Institute a réussi à isoler la souche toxigène à partir d'un enrichissement en bouillon F-CMM en faisant des repiquages successifs sur gélose. Dans le cadre du projet de thèse, seul le bouillon TPGY a été testé. Il serait intéressant de tester la méthode avec le bouillon FCMM et comparer son efficacité pour parvenir à l'isolement des souches de *C. botulinum* du groupe III.

Par ailleurs, dans le cadre de l'étude menée spécifiquement pour cet épisode de botulisme, il serait intéressant de déterminer la proportion de chaque souche (celle de la lignée Ia et lignée IV) dans l'échantillon de départ puis au fur et à mesure du processus d'isolement. Pour mener cette exploration, il faudrait disposer d'un outil moléculaire permettant de déterminer la lignée des souches présentes dans l'échantillon.

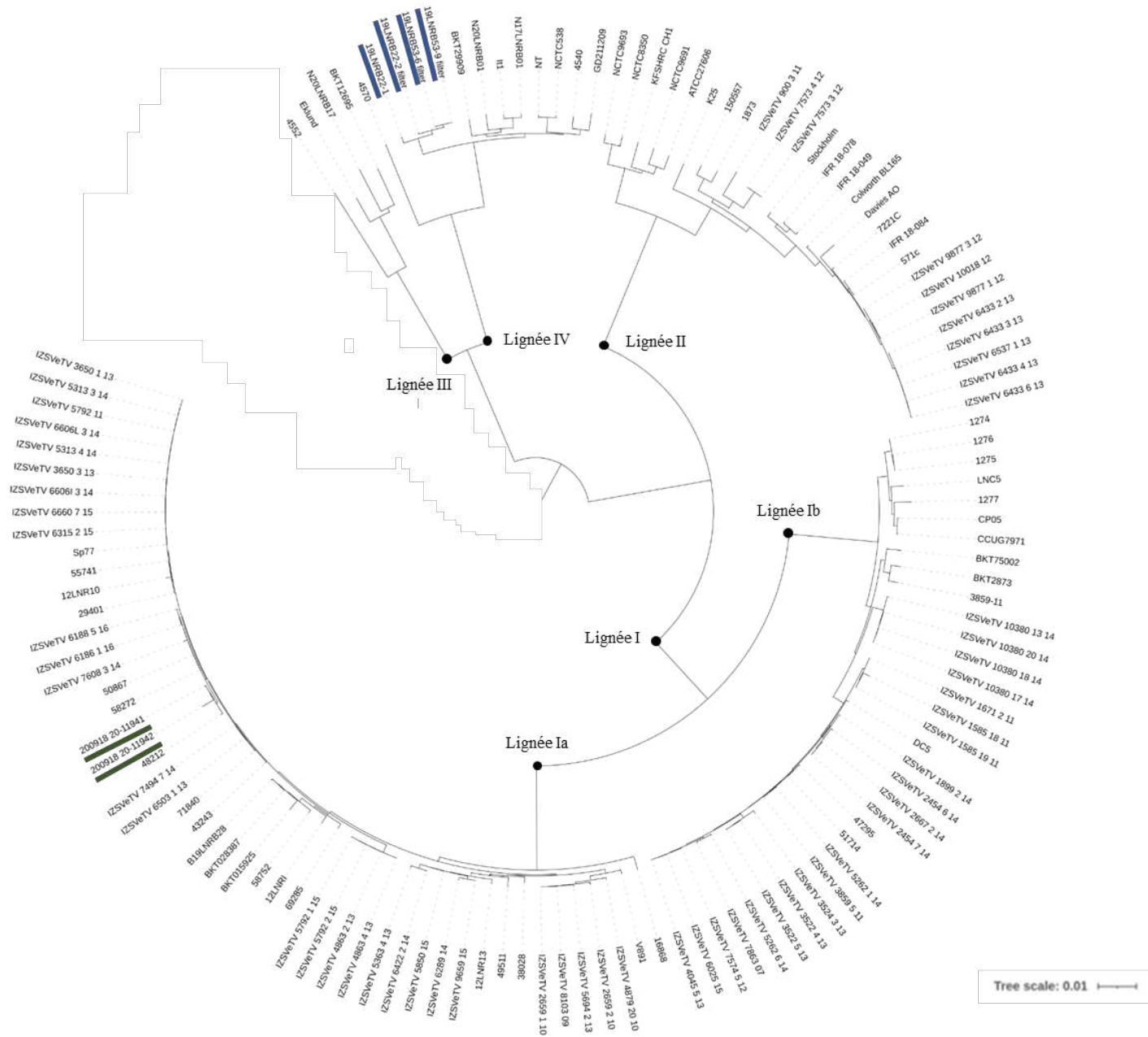


Figure 26 : Arbre phylogénétique présentant la distribution des souches de *C. novyi sensu lato* sur la base d'une analyse kmer des génomes. Concernant l'étude, les souches soulignées en vert sont les souches séquencées en Allemagne, les souches soulignées en bleu sont les souches non toxigènes séquencées à l'ICM.

### 2.1.3.3. Recherche de cibles d'intérêt pour déterminer la lignée d'une souche

Après l'analyse comparative des 125 génomes de *C. novyi sensu lato*, une recherche de gènes spécifiques de chaque lignée a été entreprise dans le but de développer des PCR permettant d'attribuer directement une lignée à la souche à partir d'un enrichissement brut. Des travaux antérieurs de l'Anses menés sur les souches de *C. botulinum* du groupe I et II avaient montré que les gènes codant pour les flagellines constituaient une cible pertinente pour discriminer les souches (Woudstra et al., 2013). Une étude menée en 2015 par l'Anses avait ciblé le gène *fliC* pour le typage des souches de *C. botulinum* du groupe III (Woudstra et al., 2015). Des systèmes PCR avaient été mis au point, ce qui avait permis d'évaluer la pertinence de cette cible pour le typage de souches à partir d'échantillons du terrain. Une autre étude menée en Italie en 2021 a aussi recherché les gènes *fliC* sur 60 génomes de *C. botulinum* du groupe III à partir des PCR développées en 2015 (Fillo et al., 2021).

Les résultats ont montré que l'ensemble des échantillons positifs pour le type C/D (les souches présentant ce type sont principalement issues de la lignée Ia) étaient positifs pour le gène *fliC*-I. Pour les échantillons positifs pour le type D/C (souches majoritairement présentes dans la lignée Ib), 5 types de gènes *fliC* (I, II, IV et VI et VII) ont été retrouvés avec une majorité de *fliC*-IV. Ce sont les seules souches de *C. novyi sensu lato* possédant le gène *fliC*-IV. Les souches de type C (majoritairement de la lignée II) ont uniquement le gène *fliC*-II dans leur génome. Les autres gènes *fliC* ont été peu détectés dans les échantillons testés (Fillo et al., 2021; Woudstra et al., 2015). Cette approche ne se révélait donc pas pertinente pour discriminer les souches puisque la répartition obtenue coïncidait avec celle des types toxiques, avec comme exception la diversité de gènes *fliC* retrouvés dans les souches de la lignée Ib qui empêche de lier un gène *fliC* à une lignée spécifique.

Dans le cadre de la thèse, les gènes recherchés pour le développement de PCR spécifiques de lignée sont les gènes codant pour les flagellines de manière globale, pas uniquement le gène *fliC*. Les premiers résultats obtenus montrent qu'il existe un ou plusieurs gènes codant la flagelline spécifiques de la lignée (Figure 27) :

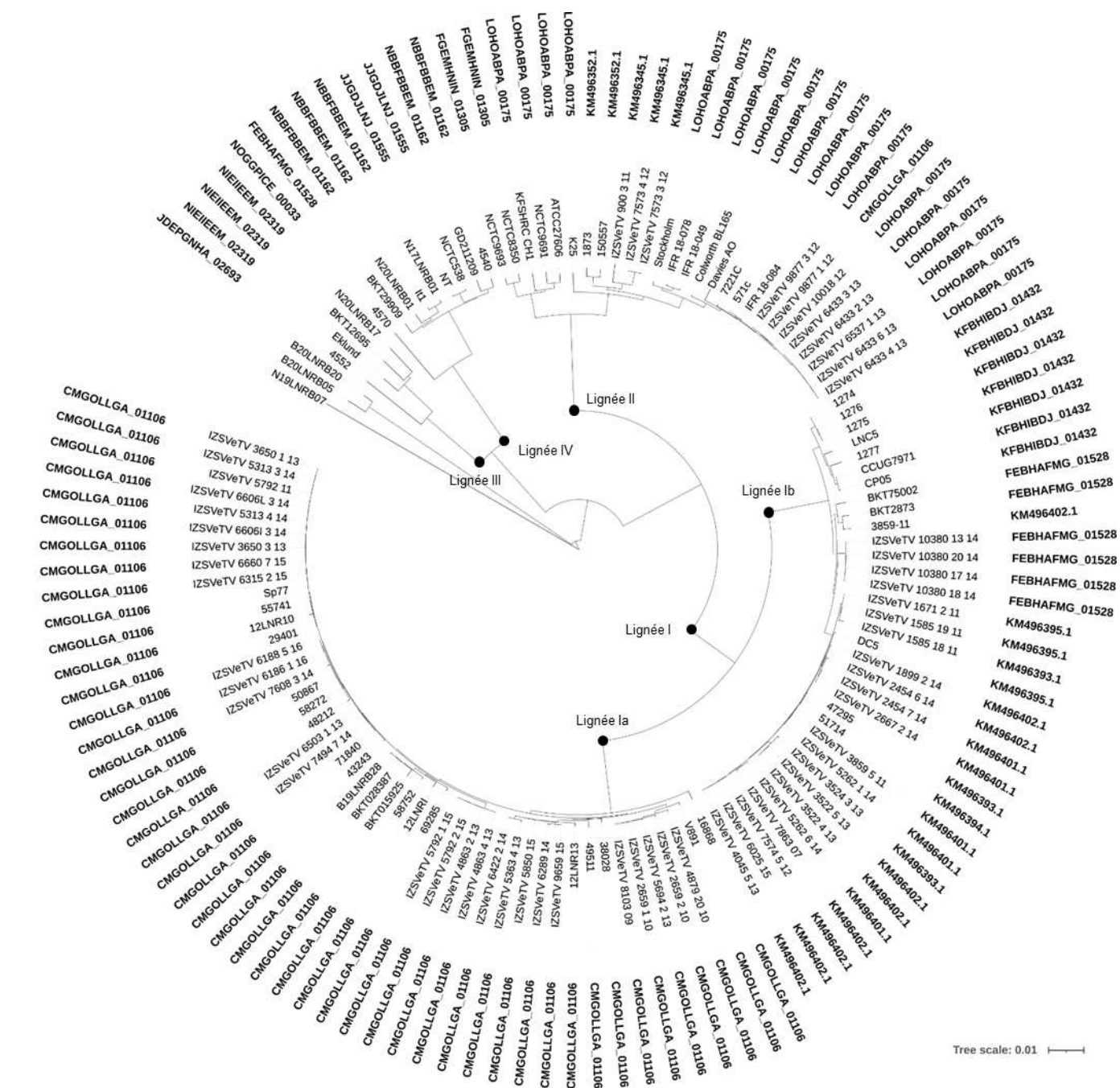


Figure 27 : Arbre phylogénétique présentant la distribution des souches de *C. novyi sensu lato* sur la base d'une analyse kmer des génomes. Pour chaque lignée, un ou plusieurs gènes de flagelline spécifiques ont été retrouvés dans les génomes.

- Pour la lignée Ia, le gène annoté CMGOLLGA codant pour une flagelline est retrouvé spécifiquement dans l'ensemble des 47 génomes de la lignée.
- Pour la lignée Ib, trois cibles différentes sont retrouvées: 5 séquences correspondantes au gène *fliC-IV* annotés KM496 très similaires génétiquement (similarités > 99%) sont retrouvés sur 21 des 34 génomes de la lignée. Les gènes FEBHAFMG (6/34) et KFBHIBDJ (7/34) sont retrouvés sur les autres génomes en plus des gènes *fliC* déjà identifiés dans les autres études.
- Pour la lignée II, le gène annoté LOHOABPA est retrouvé spécifiquement dans 19 des 27 génomes. Des séquences correspondantes au gène *fliC-II* sont retrouvés dans la plupart des génomes. Le gène CMGOLLGA retrouvé chez les souches de lignée Ia est retrouvé sur un des génomes. Enfin, le gène FGEMHIN est retrouvé sur les deux derniers génomes.
- Pour la lignée III, le gène annoté NIEIIEEM est retrouvé sur 3 des 4 génomes. Le deuxième gène retrouvé (JDEPGNHA) sur la dernière souche est complètement différent.
- Pour la lignée IV, c'est le gène annoté NBBFBBEM qui est retrouvé dans 5 des 9 génomes. Deux autres génomes arborent le gène de flagelline JJGDJLNJ, qui à 94% de similarité avec NBBFBBEM. Les deux derniers génomes possèdent deux autres gènes très différents, FEBHAFMG qui est retrouvé dans la lignée Ib et NOGGPICE.

Suite à ces résultats, des amorces ont été dessinées dans le but de développer des systèmes PCR de détection de ces cibles spécifiques. Les cibles choisies pour le développement des PCR sont : CMGOLLGA pour la lignée Ia, FEBHAFMG pour Ib, LOHOABPA pour la lignée II, NIEIIEEM pour la lignée III et NBBFBBEM pour la lignée IV. Ces systèmes PCR sont actuellement en cours de développement. Si les résultats sont prometteurs, des PCR supplémentaires seront aussi développées pour les lignées où plusieurs cibles spécifiques sont détectées (notamment la lignée Ib).

L'objectif à terme serait de pouvoir définir directement la lignée de la souche à partir de l'échantillon de départ en effectuant une ou plusieurs PCR (un gène *fliC* et un ou plusieurs gènes spécifiques de la lignée si le gène *fliC* seul ne suffit pas). Une fois cette lignée définie, un outil de typage spécifique de la lignée détectée serait alors utilisé pour qu'il soit le plus adapté à la souche étudiée. Nous avons vu précédemment que l'outil MLVA est plutôt adapté aux souches de lignée Ia. Avec le développement d'outils de typage spécifiques des lignées, l'objectif serait de développer pour chaque lignée un outil approprié pour ne plus se retrouver dans une situation où



les cibles recherchées ne sont pas détectées (comme ça a été le cas avec les souches de type D lors de l'évaluation de la MLVA).

## **2.2. Analyses des systèmes CRISPR-Cas**

En parallèle de l'optimisation des différentes étapes de séquençage et de l'analyse comparative des génomes complets, une étude a été lancée sur l'analyse des systèmes CRISPR-Cas présents chez les souches de *C. novyi sensu lato*. Elle a pour objectif d'évaluer leur présence dans le groupe d'espèces, de déterminer leurs caractéristiques et d'explorer l'origine des spacers ciblés par ces systèmes de défense, dans l'optique de mieux comprendre l'interaction des MGE avec *C. novyi sensu lato*. Enfin, à partir des résultats obtenus lors de cette étude, l'utilisation des systèmes CRISPR-Cas comme cibles pour le développement d'un outil de typage sera évaluée.

### **2.2.1. Présentation de l'étude**

Comme évoqué dans l'introduction du manuscrit, les EGM (bactériophages, plasmides) représentent une part importante du génome des souches de *Clostridium novyi sensu lato* (13,5 % du génome de la souche BKT015925 par exemple) (Skarin & Segerman, 2011). Le plasmidome est très diversifié au sein du groupe *C. novyi sensu lato* (Skarin & Segerman, 2014) et certaines caractéristiques comme la production de toxine botulique chez les souches de *C. botulinum* du groupe III proviennent de ces EGM.

Pour mieux comprendre le lien entre ces systèmes CRISPR-Cas et la présence des EGM au niveau des souches de *C. novyi sensu lato*, les 58 génomes disponibles dans les bases de données publiques au moment de l'étude (démarrage en mars 2020) ont été analysés à l'aide de différents outils bio-informatiques afin de détecter et caractériser ces systèmes avec un focus sur les spacers pour permettre l'identification des EGM ciblés par ces systèmes CRISPR-Cas, reflétant les EGM rencontrés par les souches.

Sur les 58 génomes étudiés, 55 présentent un ou plusieurs systèmes CRISPR-Cas qui sont retrouvés à la fois sur le chromosome mais aussi sur certains EGM. Le type I-B est le système le plus fréquent et est retrouvé dans l'ensemble des lignées de *C. novyi sensu lato*. Pour les souches de lignée Ia composées majoritairement de *C. botulinum* de type C/D, ce système est incomplet et localisé sur un EGM, confortant les résultats retrouvés dans une étude précédente (Woudstra et al., 2016). Des résultats similaires ont aussi été observés sur des souches de *C. botulinum* du groupe I et *C. sporogenes* dans une autre étude (Wentz et al., 2022). Cinq autres types de systèmes CRISPR-Cas

sont retrouvés dans les génomes (I-D, II-C, III-B, III-D et V-U). Des gènes *cas14a* spécifiques des systèmes de type V-U ont pu être identifiés dans deux cas :

- à côté des systèmes I-B présents sur le phage portant le gène *bont* des souches de la lignée Ia.
- dans 12 génomes de la lignée II au sein d'un élément intégratif dont le contenu génétique présente une structure similaire à celle d'un PICI (Phage Inducible Chromosomal Island). Ce PICI putatif est associé à un locus CRISPR ainsi qu'à un système de restriction modification.

Les PICIs sont une famille d'îlots de pathogénicité contribuant au transfert horizontal de gènes, à l'adaptation à l'hôte et à la virulence chez les bactéries à Gram positif. Ils sont aussi très répandus parmi les bactéries à coloration de Gram négatif (Fillol-Salom et al., 2018). C'est la première fois que ce type d'élément est décrit chez *C. botulinum*.

En ce qui concerne la composition en spacers des locus CRISPR, 4320 spacers ont été identifiés au niveau des 55 génomes pour lesquels un système CRISPR-Cas a été détecté. Après élimination des doublons (pour ne conserver uniquement qu'un seul exemplaire de chaque spacer retrouvé), l'origine de ces séquences a été déterminée pour 2045 spacers uniques. Seulement 474 spacers correspondent à au moins une séquence d'EGM dans les bases de données avec au moins 80% d'homologie et les 1503 correspondances retrouvées au total proviennent de 776 EGM uniques (48,8% de plasmides et 51,2% de phages). Une grande proportion des spacers identifiés sont originaires du groupe d'espèces *C. novyi sensu lato* (20,2%). Ce résultat confirme l'existence d'échanges fréquents de ces EGM au sein du groupe d'espèces.

Cette étude a fait l'objet d'un article scientifique qui a été publié le 13 Septembre 2021 dans le journal *Frontiers in Microbiology* et d'un poster présenté au 16<sup>ème</sup> Congrès National de la Société Française de Microbiologie qui a eu lieu à Nantes du 22 au 24 Septembre 2021 (Annexe 2).

2.2.2. Article N°4

**Article N°4**

Exploration of the Diversity of Clustered Regularly  
Interspaced Short Palindromic Repeats-Cas Systems in  
*Clostridium novyi sensu lato*

**Le Gratiot T**, Le Maréchal C, Devaere M, Chemaly M and Woudstra C.

Front. Microbiol., 13 September 2021. <https://doi.org/10.3389/fmicb.2021.711413>



# Exploration of the Diversity of Clustered Regularly Interspaced Short Palindromic Repeats-Cas Systems in *Clostridium novyi sensu lato*

## OPEN ACCESS

### Edited by:

Theresa Smith,  
Northern Arizona University,  
United States

### Reviewed by:

Raul Raya,  
Centro de Referencia para  
Lactobacilos (CERELA), Consejo  
Nacional de Investigaciones  
Científicas y Técnicas (CONICET),  
Argentina  
María Carolina Quecine,  
University of São Paulo, Brazil

### \*Correspondence:

Caroline Le Marechal  
caroline.lemarechal@anses.fr

†These authors have contributed  
equally to this work and share first  
authorship

### Specialty section:

This article was submitted to  
Evolutionary and Genomic  
Microbiology,  
a section of the journal  
Frontiers in Microbiology

Received: 18 May 2021

Accepted: 16 August 2021

Published: 13 September 2021

### Citation:

Le Gratiet T, Le Marechal C,  
Devaere M, Chemaly M and  
Woudstra C (2021) Exploration of the  
Diversity of Clustered Regularly  
Interspaced Short Palindromic  
Repeats-Cas Systems in *Clostridium  
novyi sensu lato*.  
Front. Microbiol. 12:711413.  
doi: 10.3389/fmicb.2021.711413

Thibault Le Gratiet<sup>1,2†</sup>, Caroline Le Marechal<sup>1\*†</sup>, Marie Devaere<sup>1</sup>, Marianne Chemaly<sup>1</sup> and Cédric Woudstra<sup>3</sup>

<sup>1</sup>Hygiene and Quality of Poultry and Pig Products Unit, ANSES, French Agency for Food, Environmental and Occupational Health Safety, Ploufragan, France, <sup>2</sup>UFR of Life Sciences and Environment, University of Rennes, Rennes, France, <sup>3</sup>Department of Veterinary and Animal Sciences, University of Copenhagen, Frederiksberg, Denmark

Classified as the genospecies *Clostridium novyi sensu lato* and distributed into four lineages (I–IV), *Clostridium botulinum* (group III), *Clostridium novyi*, and *Clostridium haemolyticum* are clostridial pathogens that cause animal diseases. *Clostridium novyi sensu lato* contains a large mobilome consisting of plasmids and circular bacteriophages. Here, we explored clustered regularly interspaced short palindromic repeats (CRISPR) arrays and their associated proteins (Cas) to shed light on the link between evolution of CRISPR-Cas systems and the plasmid and phage composition in a study of 58 *Clostridium novyi sensu lato* genomes. In 55 of these genomes, types I-B (complete or partial), I-D, II-C, III-B, III-D, or V-U CRISPR-Cas systems were detected in chromosomes as well as in mobile genetic elements (MGEs). Type I-B predominated (67.2%) and was the only CRISPR type detected in the Ia, III, and IV genomic lineages. Putative type V-U CRISPR Cas14a genes were detected in two different cases: next to partial type-IB CRISPR loci on the phage encoding the botulinum neurotoxin (BoNT) in lineage Ia and in 12 lineage II genomes, as part of a putative integrative element related to a phage-inducible chromosomal island (PICI). In the putative PICI, Cas14a was associated with CRISPR arrays and restriction modification (RM) systems as part of an accessory locus. This is the first time a PICI containing such locus has been detected in *C. botulinum*. Mobilome composition and dynamics were also investigated based on the contents of the CRISPR arrays and the study of spacers. A large proportion of identified protospacers (20.2%) originated from *Clostridium novyi sensu lato* (p1\_Cst, p4\_BKT015925, p6\_Cst, CWou-2020a, p1\_BKT015925, and p2\_BKT015925), confirming active exchanges within this genospecies and the key importance of specific MGEs in *Clostridium novyi sensu lato*.

**Keywords:** CRISPR-Cas, type V-U CRISPR, Cas14a, phage-inducible chromosomal island, protospacers, mobilome, *Clostridium botulinum* group III

## INTRODUCTION

*Clostridium novyi sensu lato* is a genospecies, containing the closely related species *Clostridium botulinum* (group III), *Clostridium novyi*, and *Clostridium haemolyticum* (Skarin et al., 2011; Skarin and Segerman, 2014). This proposed genospecies (Skarin et al., 2011) is based on genomic comparisons: their chromosome is highly conserved and the differences between these three species can be attributed to the content of their mobile genetic elements (MGEs; Skarin and Segerman, 2011). These three clostridia are ubiquitous, anaerobic, Gram-positive, and spore-forming pathogens that can affect both humans and animals. Botulinum neurotoxins (BoNTs) produced by *C. botulinum* group III are responsible for animal botulism. Alpha, beta, and gamma toxins produced by *Clostridium novyi* cause gas gangrene in humans and animals and black disease in animals, mostly in sheep. The beta toxin produced by *C. haemolyticum* is responsible for bacillary hemoglobinuria affecting ruminants (Skarin and Segerman, 2014). These virulence genes are all carried on MGEs (plasmids and phages) as part of their mobilome. The *Clostridium novyi sensu lato* mobilome is highly diverse: More than 60 plasmids or circular bacteriophages (called phages hereafter) categorized into 13 groups were detected upon sequencing only 24 genomes (Skarin and Segerman, 2014). Considering the few genome sequences available in the public databases, exploring further *Clostridium novyi sensu lato* MGE diversity is difficult. Yet, clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated proteins (Cas), together making up the CRISPR-Cas system, which records MGEs past encounters, could help to shed light on a putative larger mobilome within *Clostridium novyi sensu lato*.

Found in bacteria and archaea, CRISPR-Cas is a prokaryotic adaptive immunity system that interferes with invading phages and plasmids (Koonin and Makarova, 2019). CRISPR systems are composed of CRISPR arrays, divided into repeats and spacers (short variable DNA sequences), and *cas*-associated genes. Spacers, which are of foreign origin (protospacers), are integrated into the CRISPR locus, thereby acting as a memory of previous infections. Analysis of these spacers can give clues as to the mobile elements encountered in the past by a given organism (McGinn and Marraffini, 2019), thus providing information on the identity of horizontally transferable elements and their interaction with their bacterial hosts in their environment. The immune function of CRISPR is carried out by Cas proteins in three successive steps (adaptation, expression, and interference; Briner et al., 2015; Makarova et al., 2020; Pan et al., 2020). Briefly, during the adaptation stage, a part of the foreign DNA (i.e., protospacers) is incorporated into CRISPR arrays. Then, these CRISPR arrays are transcribed to generate short processed CRISPR RNA (crRNA). The last stage consists of crRNA guiding Cas proteins to their DNA target and finally cleavage of the target. With the ever-increasing number of available bacterial genomes, the classification of CRISPR-Cas systems is constantly evolving (Makarova et al., 2020). Currently, two classes, six types, and more than 20 subtypes of CRISPR-Cas systems have been described based on *cas* gene content (McGinn and Marraffini, 2019; Makarova et al., 2020).

Up to now, only one study has explored *in silico* *C. botulinum* CRISPR-Cas systems, with an analysis of 20 genomes (mainly draft sequences) demonstrating 80% prevalence of CRISPR loci in *C. botulinum* (Negahdaripour et al., 2017). Most of the studied strains contained several CRISPR loci, with one-fifth located on plasmids, although CRISPR loci are commonly found on circular chromosomes in bacteria (Negahdaripour et al., 2017). Once again, this particularity attests to the importance of the *C. botulinum* mobilome. Only eight strains of *C. botulinum* group III (including seven BoNT type C/D and one type D, from genomic lineages Ia, Ib, III, and IV) were included in this study and led to the detection of CRISPR loci and CRISPR arrays, demonstrating the existence of CRISPR systems in *C. botulinum* group III strains (Negahdaripour et al., 2017). However, with only a few genome sequences analyzed, the full diversity of the CRISPR content of *C. botulinum* group III has not been fully explored. Further investigations are therefore required to better characterize the CRISPR-Cas systems and their spacers memory in *C. botulinum* group III genomes and more generally in the *Clostridium novyi sensu lato* genospecies. The objective of our study was thus to explore the CRISPR-Cas systems in the *Clostridium novyi sensu lato* genospecies to evaluate their presence, determine their characteristics, and explore the protospacer origins to gain insight into the MGEs interacting in this taxon.

## MATERIALS AND METHODS

### Genome Selection

Fifty-eight genomes (Table 1) available in the GenBank database (National Center for Biotechnology Information, NCBI; <https://www.ncbi.nlm.nih.gov/>) as of January 2021 were retrieved to explore the CRISPR-Cas systems in the *Clostridium novyi sensu lato* genospecies. Seven of them were complete genomes (BKT015925, C-Stockholm, BKT2873, 1873, 3859/11, 150557, and NT) and the remaining 51 were draft genomes: 45 belong to *C. botulinum* (BoNT/C, C/D, /D, and D/C), 10 belong to *Clostridium novyi*, and three belong to *C. haemolyticum*, showing high variability in sampling year, location, country, sample origin, and genomic lineage. The *Clostridium novyi sensu lato* genospecies, defined according to a previously published classification (Skarin and Segerman, 2011), fall into four different lineages, based on genomic comparisons using Gegenees (Agren et al., 2012), a software tool using a fragmented alignment approach to compare bacterial genomes (Skarin and Segerman, 2014).

### CRISPR-Cas Systems Identification

The CRISPRCasFinder server (<https://crisprcas.i2bc.paris-saclay.fr/>, University of Paris-Saclay, France) was used with default parameters (Couvin et al., 2018) to identify CRISPR loci, as well as to determine the presence and content of *cas* genes. The genetic environment of solitary CRISPR arrays with no predicted CRISPR locus was further explored using the HHpred server (<https://toolkit.tuebingen.mpg.de/tools/hhpred/>, Max Planck

Institute, Germany; Zimmermann et al., 2018) with default parameters to search for undetected *cas* genes. Most of the available genomes were draft sequences, which may have resulted in incomplete or partial CRISPR-Cas loci identification.

### In silico Analysis of the Spacer Content in CRISPR Arrays

Clustered regularly interspaced short palindromic repeats arrays were identified using the CRISPR Recognition tool (CRT), software designed to detect CRISPR arrays (Bland et al., 2007). Settings were chosen so as to have at least three repeats in the CRISPR array with a repeat length between 19 and 38 base pairs (bp) and a spacer length between 19 and 48 bp. The spacer content of each CRISPR array was then extracted manually. Identical spacers were detected in several genomes and were removed to keep only one copy of each spacer that will be called “unique spacer” in the rest of the study. Complementary spacers in the reverse strand or harboring one or two mismatches were considered as unique spacers. Unique spacers were then screened with a Basic Local Alignment Search Tool (BLAST) search to identify protospacers using the CRISPRTarget tool ([http://crispr.otago.ac.nz/CRISPRTarget/crispr\\_analysis.html](http://crispr.otago.ac.nz/CRISPRTarget/crispr_analysis.html), University of Otago, New Zealand; Biswas et al., 2013). Only matches in the GenBank-Phage and Refseq-plasmid databases with at least 80% homology were considered. In addition, only the available complete (not draft) genomes were used in these databases. Parameters for the initial BLAST screen were gap open = -10 and extend = -2, minimum BLAST score = 21 with nucleotide match = 1 and mismatch = -2, E value = 1, and Word = 7. Species targeted, target name and type, location in the genome, and gene function (for the main MGEs detected) were reported. Protospacer-adjacent motif (PAM) sequences associated with identified protospacers were reported when identified. In these databases, some MGEs had different names although they were identical (e.g., pCN2 is identical to pCLG2). Unnamed plasmids were designated with a “p\_” followed by their strain name and number.

### Anti-CRISPR Genes Search

Anti-CRISPR proteins were searched on *Clostridium novyi sensu lato* plasmidome using AcRanker,<sup>1</sup> a machine learning system for direct identification of anti-CRISPR proteins (Eitzinger et al., 2020).

## RESULTS

### Overview of CRISPR-Cas Types Detected in *Clostridium novyi sensu lato*

Six different CRISPR-Cas system types belonging to class 1 and 2 CRISPR loci were identified in 55 out of the 58 *Clostridium novyi sensu lato* strains included in our study, located either on the chromosome or on MGEs (BoNT phages, plasmids p2BKT015925, pCLG2/pCN2; **Table 1**; **Figure 1**). Twenty to

209 spacers were detected per strain. *In silico*-identified PAM sequences associated with type I-B, I-D, and II-C CRISPRs matched with canonical PAMs (Kieper et al., 2018; **Figure 1**). Type I-B CRISPR predominated and was detected in all lineages (**Figure 1**). Type I-D was found in 10 *C. botulinum* BoNT/D and D/C strains and in two *Clostridium novyi* strains (lineages I-b and II). Type III-B was detected in four *C. botulinum* BoNT D/C and one *Clostridium novyi* strains (lineage I-b; **Table 1**). The type III-B CRISPR-Cas locus appeared to be devoid of the Cas proteins involved in spacer acquisition (Cas1 and Cas2) and located in close vicinity to the type I-B CRISPR locus (among the complete available genomes). *Clostridium novyi* strain 150557 (lineage II) seems to possess a type III-D CRISPR locus. Among class 2 CRISPR loci, canonical type II-C was present in seven *C. botulinum* BoNT/D and D/C strains (lineage Ib) and type V-U in 12 *C. botulinum* BoNT/C, *Clostridium novyi*, and *C. haemolyticum* strains (lineage II).

Clustered regularly interspaced short palindromic repeats-Cas genes were found either in the chromosome or in MGEs (in which case, they were located on BoNT phages and plasmids, as illustrated in **Figure 2**). When located on MGEs, the CRISPR loci were of type I-B. Some strains harbored up to three different type I-B CRISPRs (located in the chromosome, BoNT phages, and plasmids). They varied among MGEs. For example, the type I-B CRISPR locus from p1\_BKT015925 was different from that of p2\_BKT015925, whereas none were detected for p1\_Eklund phage and p1\_3859/11 phage carried instead a bacteriophage exclusion system (BREX; Goldfarb et al., 2015). Of the available p1 BoNT phage and plasmid sequences, two type I-B CRISPR-Cas loci with major differences were identified (**Figure 2**): one shared by p1\_Stockholm, p1\_16868 BoNT phages, and the p2\_BKT015925 and pCLG2/pCN2 plasmids and a second one shared by the p1\_BKT015925 and p1\_BKT2873 and p1\_1873 BoNT phages. Therefore, two different CRISPR-Cas loci were acquired independently in these MGEs. CRISPR loci were also shared between plasmids and phages: we detected a related CRISPR locus in the p1\_16868 phage and p2\_BKT015925 plasmid (**Figure 2**). Furthermore, the CRISPR loci were unrelated to chromosomal type I-B CRISPR loci (data not shown) and exhibited no link with the bacterial genomic lineages.

These CRISPR loci located on MGEs were mostly incomplete, except in the p1\_16868 BoNT phage, which contain an entire type I-B CRISPR (effector and adaptation module; Koonin et al., 2017). Incomplete CRISPR loci had lost Cas4, Cas1, and Cas2, which constitute the adaptation module (in MGEs p1\_BKT015925, p2\_BKT015925, p1\_BKT2873, and p1\_1873), as well as the Cas3 endonuclease that cleaves the DNA target in the effector module (in the p1\_Stockholm and pCLG2/pCN2 MGEs). Noteworthy, transposases interrupting the CRISPR loci were detected in p1\_BKT015925, p1\_BKT2873, p1\_1873, and p1\_Stockholm. In p1\_BKT015925, up to five transposases were detected in close vicinity to CRISPR genes. Additionally, they showed similarities with TnpB and Cas14a [one of 376 amino acids (a.a.) with a 100% HHPRED probability of being Cas14a, E-value 3.6e-43, score 327.87; another of 470 a.a. with a Cas14a probability of 100%, E-value 2.9e-40, and

<sup>1</sup><http://acranker.pythonanywhere.com/>

**TABLE 1** | Strain metadata and clustered regularly interspaced short palindromic repeats (CRISPR)-CRISPR-associated proteins types.

Strain	Species	Estimated genome Size (Mbp)	Location	Year	Origin	BoNT type	Lineage <sup>a</sup>	GenBank access number	CRISPR type	CRISPR location	Number of spacers in the genome
Eklund	<i>Clostridium botulinum</i>	2.96	United States	/	/	C	III	ABDQ000000000.1	I-B	chr	20
K25	<i>Clostridium</i> sp.	2.60	South-Korea	2012	Pig	NT	II	JENU000000000.1	I-B, V-U	/	91
Stockholm	<i>Clostridium botulinum</i>	2.70	Sweden	1946	Mink	C	II	CP063816-CP063821	I-B, V-U	chr, p1, pCLG2/pCN2	85
IFR 18/084	<i>Clostridium botulinum</i>	2.61	/	/	/	C	II	SXEF000000000	I-B, V-U	chr	118
Colworth BL165	<i>Clostridium botulinum</i>	2.49	United Kingdom	1970	/	C	II	SWNS000000000	I-B, V-U	chr	135
571C	<i>Clostridium botulinum</i>	2.62	United States	/	/	C	II	SWUK000000000	I-B, V-U	chr	121
IFR 18/049	<i>Clostridium botulinum</i>	2.41	/	/	/	(C)	II	SWNT000000000	I-B, V-U	chr	91
7221C	<i>Clostridium botulinum</i>	2.68	/	/	/	C	II	SXDK000000000	I-B, V-U	chr	133
Davies AO	<i>Clostridium botulinum</i>	2.56	/	/	/	C	II	SXEV000000000	I-B, V-U	chr	119
IFR 18/078	<i>Clostridium botulinum</i>	2.46	/	/	/	(C)	II	SXES000000000	I-B, V-U	chr	51
12LNR10	<i>Clostridium botulinum</i>	3.04	France	2012	Turkey	C/D	1a	LGVQ000000000	I-B	p1, p2	83
12LNR13	<i>Clostridium botulinum</i>	3.07	France	2012	Chicken	C/D	1a	LGVT000000000	I-B	chr, p1	81
12LNR1	<i>Clostridium botulinum</i>	3.00	France	2012	Duck	C/D	1a	LGVR000000000	I-B	p1, p2	73
29401	<i>Clostridium botulinum</i>	3.04	France	2008	Chicken	C/D	1a	LGVP000000000	I-B	p1, p2	76
38028	<i>Clostridium botulinum</i>	3.12	France	2008	Chicken	C/D	1a	LGVO000000000	I-B	p1, p2	75
43243	<i>Clostridium botulinum</i>	3.00	France	2009	Guinea Fowl	C/D	1a	LGVU000000000	I-B	p1, p2	72
48212	<i>Clostridium botulinum</i>	3.00	France	2008	Duck	C/D	1a	LGVS000000000	I-B	p1, p2	90
49511	<i>Clostridium botulinum</i>	3.09	France	2008	Chicken	C/D	1a	LHYP000000000	I-B	chr, p1, p2	65
50867	<i>Clostridium botulinum</i>	3.08	France	2008	Chicken	C/D	1a	LHYQ01000025	I-B	p1, p2	84
55741	<i>Clostridium botulinum</i>	3.04	France	2008	Turkey	C/D	1a	LHYR000000000	I-B	p1, p2	83
58272	<i>Clostridium botulinum</i>	3.07	France	2008	Chicken	C/D	1a	LHYS000000000	I-B	p1, p2	87
58752	<i>Clostridium botulinum</i>	2.82	France	2008	Chicken	(C/D)	1a	LHYT000000000	I-B	p2	45
69285	<i>Clostridium botulinum</i>	2.94	France	2008	Chicken	(C/D)	1a	LHYU000000000	I-B	p2	45
71840	<i>Clostridium botulinum</i>	3.00	France	2008	Chicken	C/D	1a	LHYV000000000	I-B	p1, p2	84
BKT2873	<i>Clostridium botulinum</i>	3.19	Sweden	2007	Chicken	C/D	1b	CP063965-CP063968	I-B	chr, p1	80
BKT12695	<i>Clostridium botulinum</i>	2.75	Sweden	2010	Chicken	(C/D)	III	JENP000000000	I-B	chr	68
BKT015925	<i>Clostridium botulinum</i>	3.20	Sweden	2008	Chicken	C/D	1a	CP002410.1	I-B	p1, p2	83
BKT028387	<i>Clostridium botulinum</i>	2.83	Sweden	2007	Chicken	(C/D)	1a	AESB000000000	I-B	p2	45
BKT75002	<i>Clostridium botulinum</i>	3.14	Denmark	2010	Chicken	C/D	1b	JENS000000000	I-B	chr, p1	80

(Continued)

TABLE 1 | Continued

Strain	Species	Estimated genome Size (Mbp)	Location	Year	Origin	BoNT type	Lineage <sup>a</sup>	GenBank access number	CRISPR type	CRISPR location	Number of spacers in the genome
It1	<i>Clostridium botulinum</i>	2.50	Italy	/	Bovine	(D/C?)	IV	JENO00000000	/	/	0
Sp77	<i>Clostridium botulinum</i>	3.06	Spain	2011	Duck	C/D	la	JENQ00000000	I-B	p1, p2	77
V891	<i>Clostridium botulinum</i>	3.17	Sweden	2007	Gull	C/D	la	AESC00000000	I-B	chr, p1	71
1873	<i>Clostridium botulinum</i>	2.57	Chad	1958	Ham	D	II	CP063822-CP063828	I-B	chr, p1, pCLG2/pCN2	22
16868	<i>Clostridium botulinum</i>	3.08	Netherlands	2001	Bovine	D	la	JENR00000000	I-B	p1, p2	91
CCUG7971	<i>Clostridium botulinum</i>	2.81	South Africa	1926	/	(D)	Ib	JDRZ00000000	I-D, II-C	/	26
3859/11	<i>Clostridium botulinum</i>	2.89	Italy	2011	Bovine	D/C	Ib	CP063959-CP063964	I-B, I-D, III-B	chr, p2	69
1274	<i>Clostridium botulinum</i>	2.94	Brazil	/	Vaccine strain	D/C	Ib	MVY00000000	I-D, II-C	/	70
1275	<i>Clostridium botulinum</i>	2.94	Brazil	/	Vaccine strain	D/C	Ib	MVZ00000000	I-D, II-C	/	70
1276	<i>Clostridium botulinum</i>	2.91	Brazil	/	Vaccine strain	D/C	Ib	MVJA00000000	I-D, II-C	/	70
1277	<i>Clostridium botulinum</i>	2.57	Brazil	/	Vaccine strain	D/C	Ib	MVJB00000000	II-C	/	32
47295	<i>Clostridium botulinum</i>	3.18	France	2009	Bovine	D/C	Ib	LHYX00000000	I-B, I-D, III-B	/	43
51714	<i>Clostridium botulinum</i>	3.18	France	2009	Bovine	D/C	Ib	LHY00000000	I-B, I-D, III-B	/	48
CP05	<i>Clostridium botulinum</i>	2.88	Brazil	/	Vaccine strain	D/C	Ib	MVJC00000000	I-D, II-C	/	46
DC5	<i>Clostridium botulinum</i>	3.32	Italy	/	Bovine	D/C	Ib	JDRY00000000	I-B, I-D, III-B	/	85
LNC5	<i>Clostridium botulinum</i>	2.89	New Caledonia	2013	Bovine	D/C	Ib	LHYW00000000	I-D, II-C	/	39
4540	<i>Clostridium novyi</i>	2.50	Great Britain	2000	Human	NT	IV	JENL00000000	/	/	0
4552	<i>Clostridium novyi</i>	2.80	Great Britain	2000	/	NT	III	JENJ00000000	I-B	/	126
4,570	<i>Clostridium novyi</i>	2.32	Great Britain	2001	/	NT	IV	JDRX00000000	I-B	/	200
BKT29909	<i>Clostridium novyi</i>	2.46	Sweden	2007	Chicken	NT	IV	JENM00000000	/	/	4
GD211209	<i>Clostridium novyi</i>	2.46	Netherlands	2009	Bovine	NT	IV	JENN00000000	/	chr	25
NCTC538	<i>Clostridium novyi</i>	2.52	Great Britain	1920	Human	NT	IV	JENK00000000	I-B	/	94
ATCC27606	<i>Clostridium novyi</i>	2.61	Germany	/	/	NT	II	JENW00000000	I-B, I-D, III-B	/	112
NCTC9691	<i>Clostridium novyi</i>	2.64	Great Britain	1955	Sheep	NT	II	JENV00000000	I-B, I-D, V-U	/	77
150557	<i>Clostridium novyi</i>	2.30	South Korea	2015	Pig	NT	II	NZ_CP029458	I-B, III-D, V-U	chr, pCLG2/pCN2	78
NT	<i>Clostridium novyi</i>	2.55	/	/	/	NT	IV	NC_008593.1	I-B	chr	96
NCTC9693	<i>Clostridium haemolyticum</i>	2.61	United States	1955	/	NT	II	JENX00000000	I-B	/	39
KFSHRC_CH1	<i>Clostridium haemolyticum</i>	2.87	Saudi Arabia	2014	Human	NT	II	NZ_LSZB01000000	I-B, V-U	/	209

(Continued)



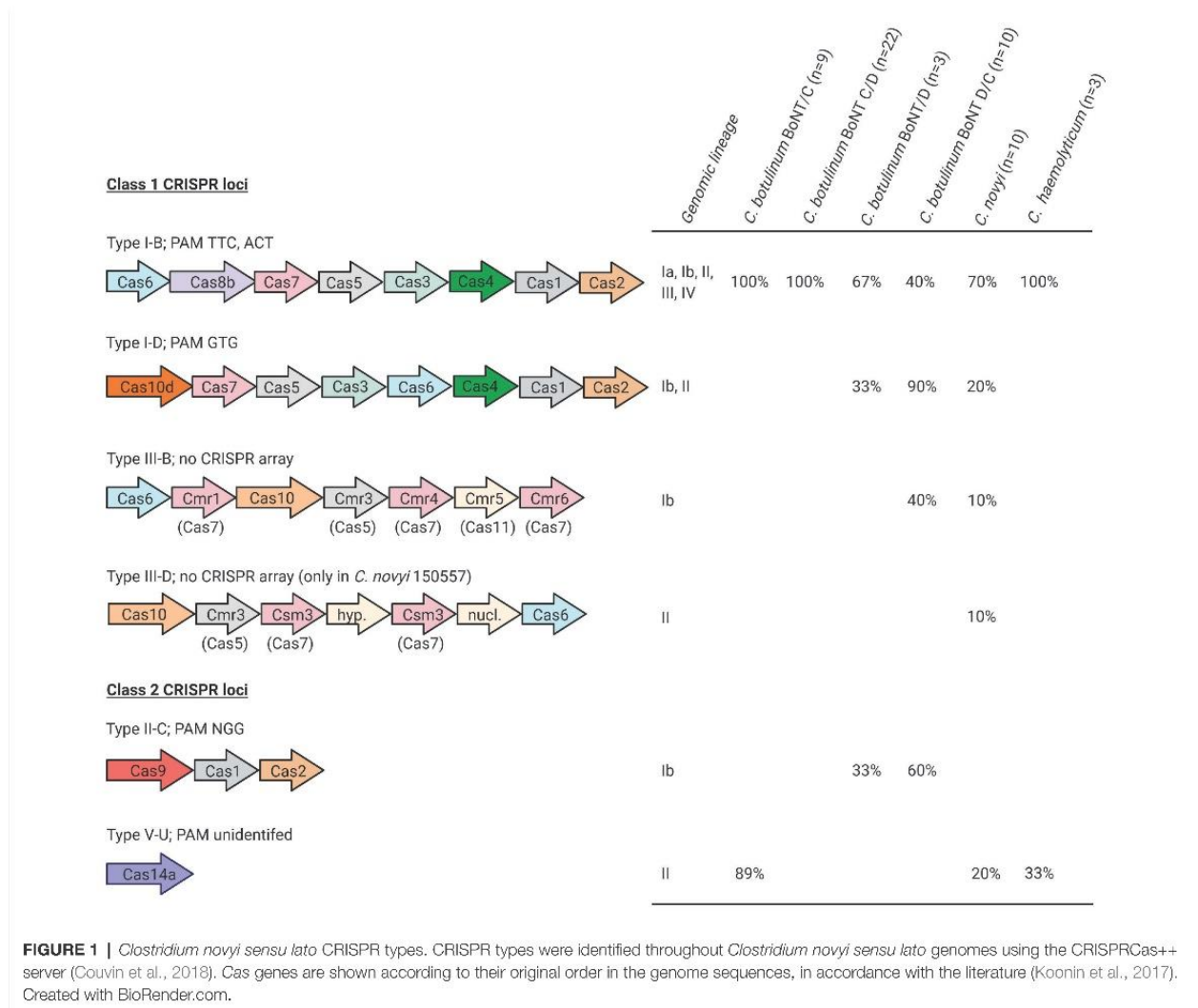
**TABLE 1 |** Continued

Strain	Species	Estimated genome Size (Mbp)	Location	Year	Origin	BoNT type	Lineage <sup>a</sup>	GenBank access number	CRISPR type	CRISPR location	Number of spacers in the genome
NCTC8350	<i>Clostridium haemolyticum</i>	2.47	United States	1946	/	NT	II	NZ_JDSA000000000	I-B	/	38

A slash indicates that no information could be retrieved. For the *Clostridium botulinum* strains, the botulinum neurotoxin (BoNT) type is indicated in parentheses when the BoNT phage has been lost and is not present in the genome sequences. The *Clostridium novyi* and *Clostridium haemolyticum* strains are non-toxicogenic (NT). Complete genomes are indicated in gray, all others are draft sequences.

<sup>a</sup>According to Skarin et al. (2011).

0, incomplete or lost BoNT phage p1; /, no information retrieved; NT, non-BoNT-producing strain.

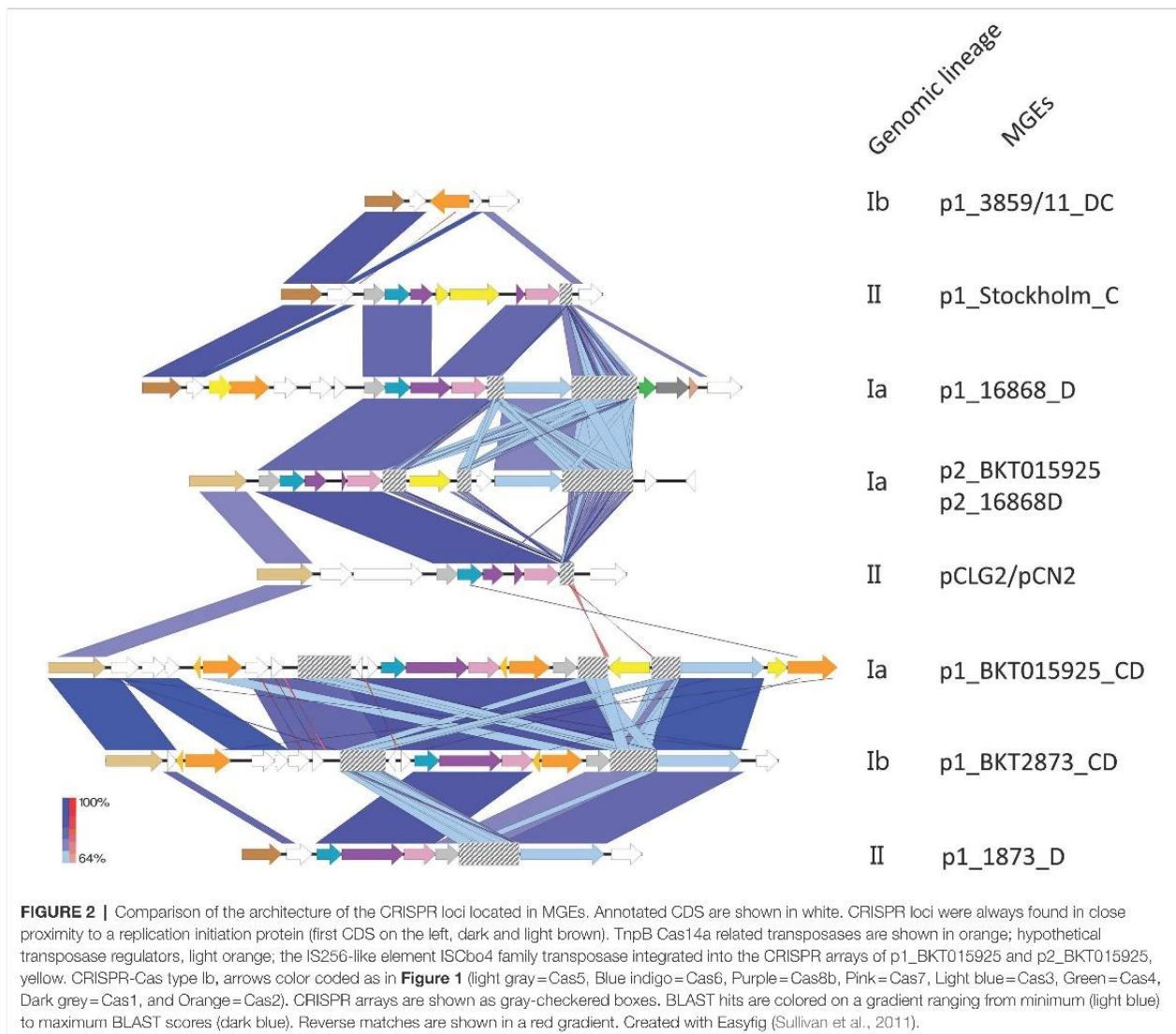


**FIGURE 1 |** *Clostridium novyi sensu lato* CRISPR types. CRISPR types were identified throughout *Clostridium novyi sensu lato* genomes using the CRISPRCas++ server (Couvin et al., 2018). Cas genes are shown according to their original order in the genome sequences, in accordance with the literature (Koonin et al., 2017). Created with BioRender.com.

score 342.39]. Detection of TnpB transposases next to Cas proteins (in BKT015925, BKT2873, 1873, and C-Stockholm located on the *bont* phage) raises the question of their potential role in the CRISPR-Cas system and possible connection with the evolution of their CRISPR loci.

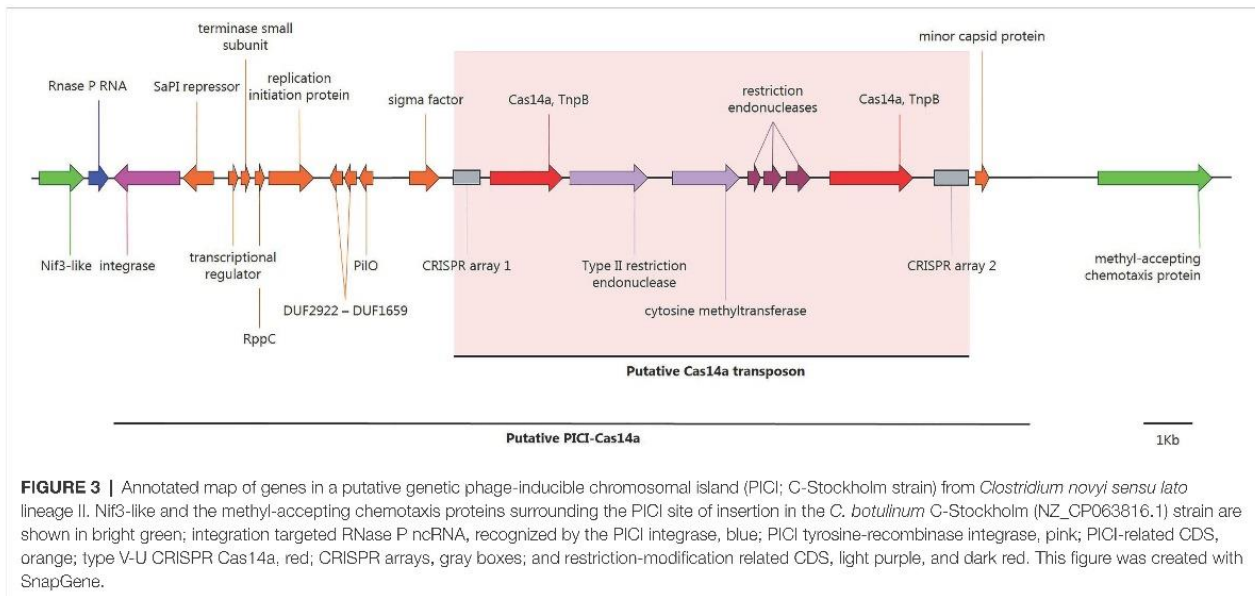
### A Putative Phage-Inducible Chromosomal Island Found in *Clostridium novyi sensu lato*

Two other putative TnpB transposases (TnpB\_IS605) were detected on the chromosome of *Clostridium novyi sensu lato* genomic lineage II, next to solitary CRISPR arrays (Table 1;



**Figure 3**). HHPRED predicted that the TnpB transposases were homologous to class 2 type V-U CRISPR Cas14a with a probability of 100%, E-value  $3.1e-38$ – $1.6e-46$ , and score 327.01–367.87 (present in 12 strains of *C. botulinum* BoNT/C, *Clostridium novyi*, *C. haemolyticum*, and *Clostridium* sp.; **Table 1**). The two putative chromosomal *cas14a* genes encode homologous proteins (35.6% homology at the C terminus between the two genes) conserved among 11 strains, but associated with distinct CRISPR arrays with different repeats (gttgagaatcaacaaaggatatttaagc and gtttagttaaactatgtgaaatgtaaat) and diverse spacers. Strain *Clostridium novyi* 150557 had only one putative *cas14a* gene and one CRISPR array. Noteworthy, three different types of restriction modification (RM) systems were identified among the 12 strains, located between the two identified *cas14a* genes (**Supplementary Figure S1**). Furthermore, *cas14a* genes were found to be part of an

integrative element splitting the RNA component of RNase P (*rnpB*; **Figure 3**; **Supplementary Figure S1**). The detailed analysis of this integrative element revealed a genetic content similar to a phage-inducible chromosomal island (PICI), with the presence of a *Staphylococcus aureus* pathogenicity island (SaPI) repressor homolog, a terminase small subunit, a phage interference redirecting packaging protein (*rppC*) and a minor capsid protein, which are characteristic of PICIs (Penadés and Christie, 2015; **Figure 3**). This putative PICI was detected in 13 strains (**Supplementary Figure S1**). In the *Clostridium novyi* NCTC 9693 strain, it was not integrated in *rnpB*, but located 30 kb away, upstream from a type I-D CRISPR-Cas system. Furthermore, in *C. botulinum* BoNT/C strain IFR 18/078, this PICI was split by another larger integrative element, which appeared to be plasmid pCN1 from the C-Stockholm strain (**Supplementary Figure S1**). This putative



PICl is the first reported in *C. botulinum*; further investigations are therefore required to confirm its nature and role.

### Identification of Protospacers to Trace Past Encounters of *Clostridium novyi sensu lato* With MGEs

The spacers of *Clostridium novyi sensu lato* CRISPR arrays were also investigated. Spacer arrays indeed represent the memory of the past encounters with MGEs and, by tracing back the spacer origin (the protospacer); they can be used to better understand the mobilome composition/interaction among *Clostridium novyi sensu lato* strains. Spacer sequences were extracted to build a representative database from the 58 strains used in this study (Table 1). We recovered 4,320 spacers, which mostly belonged to *C. botulinum* BoNT C/D (35.8%), *C. botulinum* BoNT/C (20.2%), and *Clostridium novyi* (18.8%), proportional to the number of strains used in our study (Figure 4A; Supplementary Table S1). Shared identical spacers were investigated first (Supplementary Table S2). Interestingly, the distribution of strains based on shared spacers was very similar to the distribution previously obtained using Gegenees (Skarin and Segerman, 2014) or based on SNP phylogeny profiling (Woudstra et al., 2016). For example, 17 *C. botulinum* BoNT C/D strains from lineage Ia shared more than 90% of their spacers (Supplementary Table S2). Analysis of the spacers in strains BKT2873 and BKT75002 (99% whole genome similarity using Gegenees; Skarin and Segerman, 2014) revealed 100% identical spacer content (Supplementary Table S2).

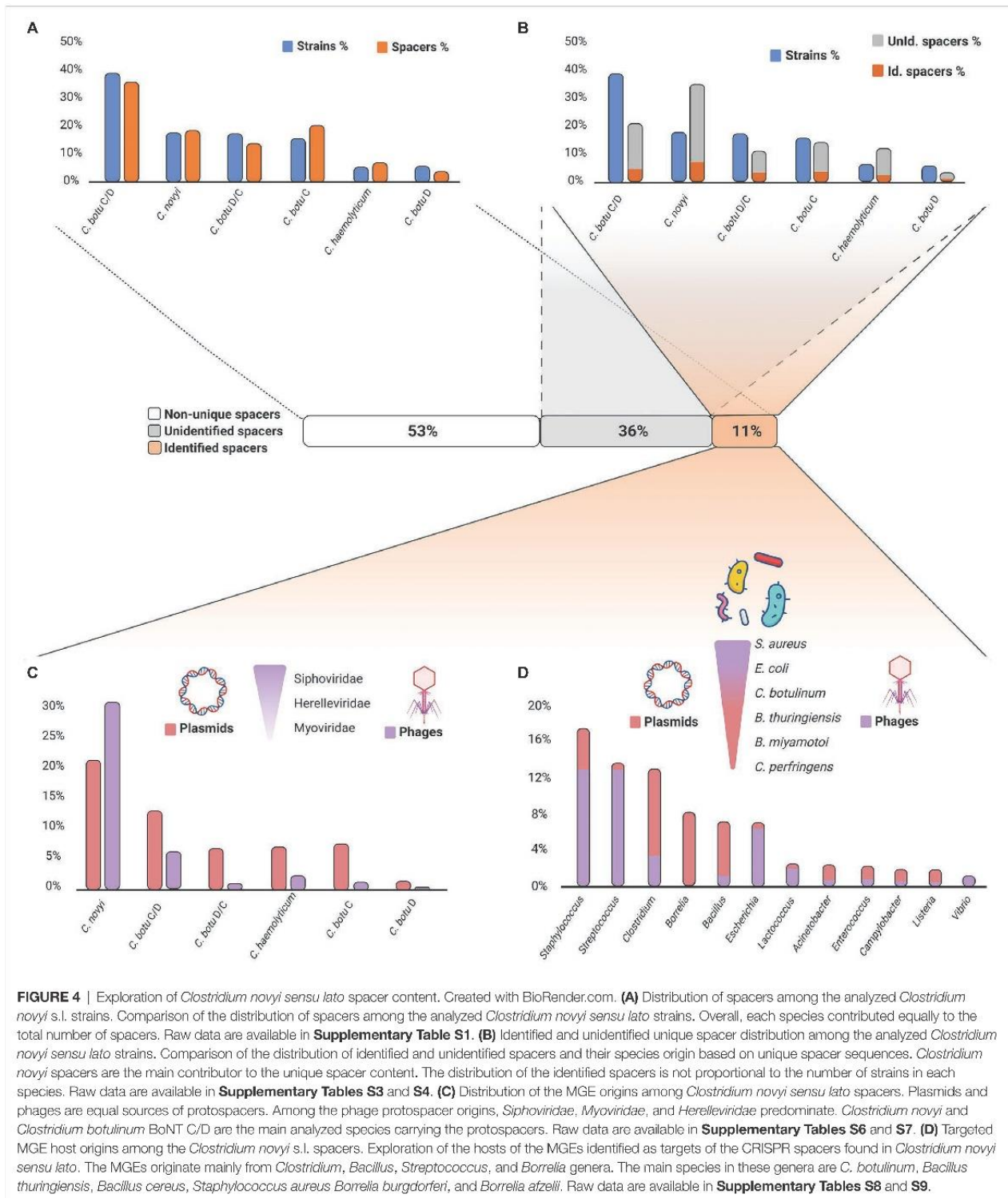
After exploring the shared spacers, we focused on the analysis of the 2,045 unique spacers (47.3%; Supplementary Table S3). Among the unique spacers, the distribution was no longer proportional to the number of strains in each analyzed species, with *Clostridium novyi* spacers now representing up to 34.6% of the unique spacer content and *C. botulinum* BoNT C/D

spacers accounting for 20.6% (Supplementary Table S3). This pattern is in concordance with our above observation on spacers widely shared in the *C. botulinum* BoNT C/D lineage Ia, resulting in a lower number of unique spacers, with many of them being present multiple times in genetically related strains.

We then screened the unique spacers via a BLAST search using the CRISPRtarget tool (Biswas et al., 2013) to determine their origin (the protospacers) and found 474 spacers matching with at least 80% homology in the GenBank-Phage and Refseq-plasmid databases (Figure 4B; Supplementary Table S4). The remaining 1,571 unidentified spacers (Supplementary Table S5) belong to unknown or uncharacterized MGEs, the so-called “spacer dark matter” (McGinn and Marraffini, 2019).

A BLAST search on the 474 spacers returned 1,503 protospacer hits in the Refseq-Plasmid and GenBank-Phage databases (Figure 4C; Supplementary Table S6). Most of the identified protospacers originally belonged to spacers from *Clostridium novyi* (51.7%, with 20.9% located in plasmids and 30.8% in phages) and *C. botulinum* BoNT C/D (18.8%, 12.6% located in plasmids, and 6.1% in phages; Supplementary Table S7). The high number of protospacers identified relative to the number of spacer sequences reflects the acquisition of protospacers from homologous genes present in different or related MGEs.

The 1,503 BLAST hits retrieved from the identified protospacers corresponded to 776 unique MGEs (Supplementary Table S8), composed of 379 plasmids (48.8%) and 397 phages (51.2%). Phages were mostly temperate *Siphoviridae* and *Herelleviridae* (among the most common occurring phage species; Supplementary Table S8). The phages ranged in size from 497,513 bp for the largest (*Myoviridae* phage G) to 17,637 bp for the smallest (*Podoviridae* vB\_SauP\_phiAGO1.9). MGEs originating from the *Staphylococcus*, *Streptococcus*, *Clostridium*, *Bacillus*, *Escherichia*, and *Borrelia* bacterial genera represented 66.8% of the unique MGEs (Figure 4D; Supplementary Table S9).



Interestingly, the bacterial genera varied in the type of MGE contribution (plasmid or phage). *Clostridium*, *Bacillus*, and *Borrelia* mainly contributed plasmids (32.3%), whereas

*Staphylococcus*, *Streptococcus*, and *Escherichia* mainly contributed phages (32.5%; **Supplementary Table S9**). At the species level, *S. aureus*, *Escherichia coli*, *C. botulinum*, *Bacillus thuringiensis*,

**TABLE 2 |** The score value represents the percentage of homology between the spacer and its target (Biswas et al., 2013).

**Table 2A: Spacer BLAST scores of the main targeted bacterial species.**

Genus	Protospacer species origin	% BLAST results	Average score	St. dev.	Min.	Max.
<i>Clostridium</i>	<i>Clostridium botulinum</i>	28%	28.76	4.97	21	38
<i>Staphylococcus</i>	<i>Staphylococcus aureus</i>	11%	21.65	0.80	21	26
<i>Streptococcus</i>	<i>Streptococcus thermophilus</i>	7%	24.38	1.16	22	26
<i>Escherichia</i>	<i>Escherichia coli</i>	4%	22.25	0.54	22	24
<i>Borrelia</i>	<i>Borrelia miyamotoi</i>	2%	22.19	0.68	21	23
<i>Bacillus</i>	<i>Bacillus thuringiensis</i>	2%	23.84	1.95	21	28
<i>Clostridium</i>	<i>Clostridium perfringens</i>	2%	23.38	2.88	21	32
<i>Lactococcus</i>	<i>Lactococcus lactis</i>	2%	23.50	2.15	21	30
<i>Clostridium</i>	<i>Clostridium difficile</i>	1%	23.00	0.82	22	24
<i>Staphylococcus</i>	<i>Staphylococcus epidermidis</i>	1%	22.25	0.83	21	23
<i>Acinetobacter</i>	<i>Acinetobacter baumannii</i>	1%	22.67	0.94	22	24
<i>Streptococcus</i>	<i>Streptococcus pyogenes</i>	1%	23.27	2.21	21	28
<i>Bacillus</i>	<i>Bacillus cereus</i>	1%	22.86	1.28	21	25
<i>Streptococcus</i>	<i>Streptococcus agalactiae</i>	1%	23.87	0.88	22	26
<i>Streptococcus</i>	<i>Streptococcus mitis</i>	1%	24.36	2.38	22	28

**Table 2B: Spacer BLAST scores of the main targeted *C. botulinum* mobile genetic elements (MGEs).**

Mobile genetic element	Protospacer species origin	% BLAST results	Average score	St. dev.	Min.	Max.
p1_Cst	<i>Clostridium botulinum</i>	6%	31.36	4.98	21	38
p4_BKT015925	<i>Clostridium botulinum</i>	5%	29.42	4.43	21	38
p6_Cst	<i>Clostridium botulinum</i>	4%	30.57	4.69	22	37
CWou-2020a	<i>Clostridium botulinum</i>	3%	29.03	5.33	22	38
p1_BKT015925	<i>Clostridium botulinum</i>	2%	30.46	3.71	24	36
p2_BKT015925	<i>Clostridium botulinum</i>	1%	27.58	3.66	22	34

The higher the score, the better the homology, indicating a putative recent encounter. A: *Clostridium botulinum* is the main target of the spacers with up to 28% of the BLAST results, together with the highest average score. This illustrates the active dissemination of *C. botulinum* mobile genetic elements (MGEs) throughout *Clostridium novyi sensu lato*. B: The most encountered *C. botulinum* MGEs are p1\_Cst > p4\_BKT015925 > p6\_Cst, accounting for 15% of the BLAST results. Other *C. botulinum* MGEs (CWou-2020a, p1\_BKT015925, and p2\_BKT015925) account for 6% of the BLAST results, showing less frequent encounters. The average BLAST scores are roughly equivalent among the *C. botulinum* MGEs, suggesting concomitant dissemination.

**TABLE 3 |** Spacers targeting the restricted mobilome.

MGE name	MGE type	Localization of spacers	BoNT-producing strains	% of all unique spacers	% of BoNT+ – strains containing spacers	Number of spacers/strain
p1_Cst, p1_BKT015925	BoNT phage	97.8% chromosomal, 2.2% MGEs	+ –	4.2% 3.8%	84.1% 50.0%	1–13 2–16
p4_BKT015925, p6_Cst	Prophage/plasmid hybrid	93.1% chromosomal, 6.9% MGEs	+ –	5.2% 3.7%	65.9% 78.6%	1–11 2–14
CWou-2020a	Prophage	100% chromosomal	+ –	1.2% 1.3%	47.7% 50.0%	1–5 1–8
p2_BKT015925	Plasmid	100% chromosomal	+ –	0.5% 0.3%	25.0% 28.6%	1–4 1

Origin of spacers specifically targeting the restricted mobilome p1\_Cst, p1\_BKT015925, p4\_BKT015925, p6\_Cst, CWou-2020a, and p2\_BKT015925 were analyzed. BoNT+ strains refer to *C. botulinum* BoNT C, C/D, D, and D/C; BoNT– strains refer to *Clostridium novyi*, *C. haemolyticum*, and *Clostridium* sp. Distribution of spacers with regard to the entire BLAST results, to the proportion of BoNT +/- strains and the number of spacers targeting this restricted mobilome are given. BoNT, botulinum neurotoxin; MGE, mobile genetic element; BoNT-, *C. novyi*, *C. haemolyticum*, and *Clostridium* sp.; BoNT+, others.

*Streptococcus thermophilus*, *Clostridium perfringens*, *Lactococcus lactis*, *Staphylococcus epidermidis*, *Acinetobacter baumannii*, *Streptococcus pyogenes*, *Borrelia miyamotoi*, *Clostridium difficile*, *Bacillus cereus*, *Streptococcus agalactiae*, and *Streptococcus mitis* were the main sources of MGEs (Supplementary Table S9B).

### A Limited Number of *C. Botulinum* MGEs Predominate *Clostridium novyi sensu lato* Spacer Sequences

Protospacers identified from MGEs belonging to *C. botulinum* represented 27.6% of all BLAST results, with the highest average BLAST score of 28.8 (Table 2A), therefore suggesting that they are the most recent and frequently encountered ones. Of these MGEs, p1\_Cst, p4\_BKT015925, p6\_Cst, C\_Wou-2020a, p1\_BKT015925, and p2\_BKT015925 represented up to 20.2% of the BLAST results (hereafter called the “restricted mobilome”), showing a roughly similar average BLAST score, suggesting parallel and concomitant dissemination (Table 2B). Although these MGEs belong to different groups of mobile elements (p1\_Cst and p1\_BKT015925 are pseudolysogenic BoNT phages; Skarin et al., 2011, p4\_BKT015925 and p6\_Cst have the characteristics of prophage-plasmid hybrids; Pfeifer et al., 2021, CWou-2020a is a circular prophage and p2\_BKT015925 is a plasmid; Skarin et al., 2011), they appeared to be genuinely recognized by the CRISPR systems. In a few cases, CRISPR protection was incomplete (e.g., against p4\_BKT015925), because the strain had spacers as well as the immunized MGE as part of their genome (Woudstra et al., 2017). This may suggest that these MGEs persist even in the presence of spacers matching with 100% homology, using an unidentified strategy (screening for anti-CRISPR using AcRanker was negative, data not shown; Eitzinger et al., 2020). On the contrary, full CRISPR protection was also observed, e.g., strains that had acquired protospacers against p2\_BKT015925 were devoid of the corresponding plasmid (Woudstra et al., 2017).

In *Clostridium novyi sensu lato*, BoNT-producing strains showed slightly more encounters with the restricted mobilome with 11.1% of BLAST results, whereas non-BoNT-producing strains accounted for 9.1% (Table 3). The localization of spacers targeting the restricted mobilome was mainly chromosomal, but also found in MGEs (Table 3; Supplementary Table S6), indicating that MGEs containing CRISPR arrays may be involved in the protection against these elements. Phage-plasmid hybrids p4\_BKT015925 and p6\_Cst were the main source of acquired protospacers (8.9%), with 65.9% of BoNT-producing, and 78.6% of non-BoNT-producing strains being immunized. A high number of protospacers was detected in some strains, with up to 16 spacers for *C. haemolyticum* KFSHRC\_CH1 targeting the BoNT phages (Table 3; Supplementary Table S6). The protospacers in these MGEs originated mainly from hypothetical genes and intergenic regions (39.4%) with high BLAST scores (Supplementary Table S10). Other protospacers were located in phage-related genes (e.g., capsid, tail, baseplate, terminase, and holin), but also plasmid-partition and replication genes (e.g., ParM, ParB, and replication initiation protein; Supplementary Table S10).

### Spacer Sequences in *C. botulinum* Group III MGE CRISPR Arrays

The presence of a type I-B CRISPR-Cas system associated with CRISPR arrays was detected in BoNT phages p1\_Cst, p1\_1873, p1\_BKT015925, p1\_BKT2873, p1\_16868, and plasmids p2\_BKT015925 and p?CLG2/pCN2 (Figure 2; Supplementary Table S11). From the 171 spacers extracted (11.4% of the unique spacers), surprisingly none were shared among any of the different MGEs. Protospacer identification was possible for 28.6%, yielded 133 BLAST hits at an equivalent ratio between phages and plasmids, and belonged mostly to phages from *Escherichia* and *Enterococcus* (41.4%) and to plasmids from *Clostridium* and *Bacillus* (40.6%). *Clostridium botulinum* MGEs (26.3%), with BLAST scores ranging from 22 to 36, predominated, especially p4\_BKT015925, p6\_Cst, p1\_Cst, p1\_BKT015925, CWou-2020a, p5\_BKT015925, and p2\_BKT015925 (18.0%; Supplementary Table S11), in accordance with our previous results (Table 2B). Interestingly, protospacers from p4\_BKT015925 were found to have been frequently acquired by CRISPR arrays located on chromosomes or on MGEs, suggesting that *C. botulinum* MGEs have also protected themselves against this potentially invasive element. In addition, BoNT phage acquisition of protospacers from BoNT phages was also detected (p1\_16868 vs. p1\_BKT015925; p1\_BKT015925 vs. p1\_Cst; and p1\_BKT2873 vs. p1\_BKT015925), suggesting that BoNT phages to use the CRISPR system to target other BoNT phages.

### PICI-Related Spacer Content

A type V-U CRISPR was detected in a putative PICI element present in 12 *Clostridium novyi sensu lato* strains from lineage II, containing two distinctive CRISPR arrays (namely CRISPR 1 and CRISPR 2). In the CRISPR arrays, CRISPR 1 contained only 7–9 spacers, and CRISPR 2 had 9–28 spacers. Interestingly, the main spacer diversity was found in *Clostridium novyi*, *C. haemolyticum*, and *Clostridium* sp., which did not share any spacers (Supplementary Table S12). Among the eight *C. botulinum* BoNT/C strains, the PICI sequences were highly conserved (as illustrated in Supplementary Figure S1 for the PICI-Cas14a elements) and had almost exactly the same CRISPR arrays (differences in 1–2 spacers), suggesting that they are highly phylogenetically related. This close relationship was confirmed by comparing whole genome sequences (data not shown), therefore suggesting that all *C. botulinum* BoNT/C strains having a PICI are clonal. When considering the spacers extracted from the two CRISPR arrays, 51.8% were unique. Of the unique spacers, 32.2% gave BLAST results and could be attributed to a protospacer (Supplementary Table S12). Among the identified protospacers, 18 with the best BLAST score (25–36) belonged to the above-identified restricted mobilome (Table 2B). Interestingly, one spacer located in the *C. botulinum* BoNT/C Stockholm strain in the PICI CRISPR array one targets its own BoNT phage (BLAST score 30) as well as plasmid pCLG2/pCN2 (BLAST score 28; Supplementary Table S12). Other identified spacers were mainly acquired by the CRISPR arrays from PICIs in *Clostridium novyi* (7), *C. haemolyticum* (5), and *Clostridium* sp. strains (5), which did

not possess these MGEs as part of their genomes (**Supplementary Table S12**).

## DISCUSSION

### *Clostridium novyi sensu lato* Species Harbor a Range of CRISPR-Cas Systems

The first objective of our study was to investigate the diversity and composition of *Clostridium novyi sensu lato* CRISPR-Cas systems, which had not been fully explored yet (Woudstra et al., 2016; Negahdaripour et al., 2017). We detected a high prevalence of CRISPR-Cas systems (complete or partial) throughout the 58 strains included in our study, with only three strains lacking CRISPR-Cas systems (*C. botulinum* It1, *Clostridium novyi* BKT29909, and *Clostridium novyi* 4540, all lineage IV). Accordingly, a CRISPR prevalence of 80% was detected in *C. botulinum* strains, in a study including only eight *Clostridium novyi sensu lato* genomes (Negahdaripour et al., 2017). This prevalence is much higher than the rate of 46% reported for bacteria (Grissa et al., 2007) or other clostridia species such as *C. perfringens*, which showed a CRISPR prevalence of 53.15% (Long et al., 2019). The location of some CRISPR-Cas systems on MGEs may play a role in this high prevalence. Some strains, in particular from lineage Ia, only have CRISPR-Cas systems in MGEs and none in their chromosome.

In our study, the number of CRISPR loci varied between one and three per strain, which seems to be common in pathogenic bacteria (Cady et al., 2011; Shariat et al., 2013; Yin et al., 2013; Karah et al., 2015), and belonged to class 1 and 2. Class 1 CRISPR-Cas systems (types I-B, I-D, III-B, and III-D) were identified in the *Clostridium novyi sensu lato* strains investigated in our study, similar to previous studies that have shown that the type I CRISPR-Cas system is the most common type found in clostridia (Makarova et al., 2020) and that type III-B, followed by type I-B, is the most commonly identified types in *C. botulinum* groups I and II (Negahdaripour et al., 2017). It is noteworthy that only one type of CRISPR-Cas system (I-B) was detected in lineages Ia, III, and IV, but several types were detected in lineages Ib and II. Previous studies have shown that the number of MGEs is higher in *C. botulinum* BoNT C/D (lineage Ia) than in BoNT D/C (lineage Ib) containing several different CRISPR types (Woudstra et al., 2016, 2017). In *B. cereus*, strains with active CRISPR-Cas systems have fewer MGEs than strains with partial CRISPR-Cas systems or without any CRISPR-Cas systems (Zheng et al., 2020). Likewise, CRISPR-Cas systems may play a barrier role against horizontal gene transfer by preventing the acquisition of MGEs, resulting in a lower occurrence of plasmids and phages in strains from lineage Ib, and the opposite in lineage Ia strains.

Type I-B CRISPR loci (either complete or incomplete) were the most common CRISPR type detected in *Clostridium novyi sensu lato*. Type I-B seems to be typical of clostridia. It has indeed been detected in *C. perfringens* (Long et al., 2019) as well as in *C. chauvoei* (Thomas et al., 2017), *C. tetani* (Cohen et al., 2017), and in all *C. difficile* genomes (Hargreaves et al.,

2014; Andersen et al., 2016). Here, several type I-B CRISPR arrays were also located on MGEs such as the BoNT phages and the plasmid p2\_BKT015925. The localization in prophages has also been reported in *C. difficile* (Hargreaves et al., 2014; Boudry et al., 2015). Prophage localization of CRISPR arrays may play a role in preventing infection by related competing phages (Sorek et al., 2008). It may also play a role in the mobilome dynamics in *Clostridium novyi sensu lato*, in particular in lineage Ia for which the localization of a partial type I-B CRISPR in the pseudodysogenic BoNT phage was preponderant, whereas no CRISPR locus was detected chromosomally. Noteworthy, type I-B CRISPR loci located on BoNT phages in BoNT C/D strains was incomplete, lacking the adaptation modules (Cas1, Cas2, and Cas4). In a previous study, putative type I-related operons, derivatives of subtype I-F or subtype I-B encoding interference *cas* genes, but not associated with *cas1*, *cas2*, or *cas3* genes, have been detected (Makarova et al., 2015). They are located either on plasmids or associated with transposon-related genes (Makarova et al., 2015). Here, they were located in plasmids and in the BoNT phages sometimes associated with transposases. The functionality of these CRISPRs was not experimentally investigated in our study, but the presence of various spacers in CRISPR arrays suggests they are active. The absence of the adaptation module (Cas1, Cas2, and Cas4) would only be detrimental to the acquisition of new spacers, and it would not prevent the effector complex from performing its function, except in the absence of Cas3. Without Cas3, the effector complex is not able to cleave its target – but only bind –, which may result in transcription interference (CRISPRi; Hawkins et al., 2015). The presence of complete CRISPR-IB systems in the chromosome for some strains may also compensate for the missing Cas proteins, allowing these incomplete systems to be fully functional (e.g., in the *C. botulinum* group III Stockholm strain). Transposases showing similarities with TnpB and Cas14a and interrupting the CRISPR loci were detected in some strains. This is intriguing because Cas14a has been shown to be capable of both transposase- and CRISPR-related activities (Shmakov et al., 2017). Further investigations are required to determine the exact role of these transposases.

There may be other biological significance for these incomplete CRISPR-Cas systems located on MGEs, as roles other than immunological protection have been revealed for CRISPR-Cas systems, such as signal transduction and gene regulation (Bikard et al., 2012). Considering that some CRISPR locus located on MGEs can lose their endonuclease, the cascade complex would only be able to bind its target, therefore promoting steric transcriptional inhibition (Zhang et al., 2021). Further experimental investigations are required to better understand the role and mechanisms of these incomplete type I-B systems located in MGEs in the biology of *Clostridium novyi sensu lato*.

In addition to type I-B, other CRISPR-Cas system subtypes were found in *Clostridium novyi sensu lato* strains, except *C. botulinum* BoNT C/D (lineage Ia) and strains from lineages III and IV. Type I-D was found in *C. botulinum* BoNT/D (33%) and D/C (90%) and in *Clostridium novyi* (20%; lineages Ib and II, **Figure 1**). Considering the low number of available

BoNT D/C genomes (currently 10 genomes in public databases), it is difficult to draw conclusions and establish generalities. However, based on the analysis of these genomes, a complete type I-D system appears to be preponderant mostly in *C. botulinum* BoNT D/C strains (lineage Ib). This CRISPR-Cas system is a hybrid of type I and type III systems (Lin et al., 2020; McBride et al., 2020), most often reported in cyanobacteria and archaea (Lin et al., 2020). So far, little is known about this subtype and its biological properties (Lin et al., 2020).

Type II-C was the second-most common CRISPR type found in *C. botulinum* BoNT D/C strains (60%) in lineage Ib. This class II CRISPR system is detected in diverse bacterial species in various environments, especially in pathogenic and commensal bacteria (Mir et al., 2018). A possible role in virulence for type II-C CRISPR-Cas systems has been suggested (Mir et al., 2018), but there is no experimental evidence for such function in *Clostridium novyi sensu lato*. Because *C. botulinum* genome engineering has been successful using CRISPR type II-B Cas9 (Mertaaja et al., 2021), type II-C can also be explored as a native alternative for genome modification using a mini-plasmid strategy (Pyne et al., 2016). This strategy may be advantageous, as modification of clostridia is cumbersome.

Finally, CRISPR type III was also found in *Clostridium novyi sensu lato* and is represented in numerous archaea but is less frequent in bacteria (Koonin et al., 2017), except in *C. botulinum* groups I and II where type III-B has been reported to be preponderant (Negahdaripour et al., 2017). Type III systems degrade both DNA and RNA (Pyenson and Marraffini, 2017; Tamulaitis et al., 2017; Terns, 2018). Here in our study, type III-B and III-D systems were detected in four strains (lineages Ib and II). They lacked *cas1* and *cas2* genes, which seems to be common in type III-B, C, and D systems (Makarova et al., 2015), especially in genomes also containing type I CRISPR-Cas loci. Due to their plasticity in crRNA selection, type III-B interference machinery can use type I spacers to counter infection from bacteriophages that have escaped the type I defenses through PAM mutations (Silas et al., 2017). Here, the four genomes in which these CRISPR systems were detected also had type I-B and type I-D systems, which suggest such a mechanism for these strains.

Lastly, two putative type V-U CRISPR *cas14a* genes encoding homologous proteins associated with two distinct CRISPR arrays were detected in 81.25% of *Clostridium novyi sensu lato* strains belonging to genomic lineage II, and not in other lineages. Type V-U CRISPR targets single-stranded DNA (ssDNA) sequences without requiring restrictive signal sequences (Khan et al., 2019). The ability of Cas14a to specifically target ssDNA suggests a role in defense against ssDNA viruses or MGEs that propagate through ssDNA intermediates, which may be an advantage in some ecosystems containing ssDNA viruses (Harrington et al., 2018), such as soils (Kim et al., 2008), water (Angly et al., 2006; Rosario et al., 2009), or human and farm-animal feces (Sikorski et al., 2013). Here no ssDNA viruses were identified among the spacers located in the two CRISPR arrays from the putative type V-U CRISPR Cas14a systems with a BLAST match (**Supplementary Table S12**). This absence does not exclude the possibility of finding such ssDNA targets

among the spacers with no BLAST matches, but there may be other roles for these putative Cas14a systems. For example, the localization of CRISPR-arrays close to the Cas14a genes may suggest that Cas14a plays a role in the spacer acquisition step.

The presence of this putative type V-U CRISPR only in strains from genomic lineage II suggests a specific mechanism of horizontal transfer for this lineage. Moreover, it was included in a putative PICI. PICIs are a recently discovered family of pathogenicity islands that contribute to horizontal gene transfer, host adaptation, and virulence in Gram-positive cocci, but are also widespread among Gram-negative bacteria (Fillol-Salom et al., 2018). These highly evolved molecular parasites can hijack phage-packaging systems, thus allowing its transfer to other bacterial cells (Crestani et al., 2020). In doing so, PICIs protect the host resident population against specific phages. Although the insertion region was present in all *Clostridium novyi sensu lato* genomic lineages (data not shown), only lineage II possessed the putative PICI-Cas14a element, which suggests a narrow host-range dissemination. Furthermore, the presence of diverse spacer sequences in the PICI CRISPR arrays indicates that the system actively acquires protospacers from invading MGEs. The high protospacer identification BLAST score suggests recent acquisition, especially in the identified restricted mobilome. Interestingly, all *C. botulinum* BoNT/C strains having a PICI were clonal with highly conserved PICI sequences; despite being isolated from different countries and years (1946 in Sweden for strain Stockholm and 1970 in United Kingdom for strain Colworth BL165 for example). This might suggest a high stability of these genomes, including the PICI sequences but more genomic data of *C. botulinum* BoNT/C lineage II strains are required to better explore and understand this result.

Noteworthy, an RM system was associated with the putative *cas14a* genes that may add another layer of defense for the PICI host. Transfer of genes encoding non-essential or accessory functions by PICIs has been previously reported (Penadés and Christie, 2015), but to our knowledge, this is the first time that CRISPR and RM systems have been encountered in a PICI. Further investigations are required to study this putative PICI element and its role in *Clostridium novyi sensu lato* lineage II ecology in more detail.

### Protospacer Origins Reveal a Handful of MGEs That Are Widespread Among *Clostridium novyi sensu lato* Species

The second objective of our study was to investigate the spacers found in the CRISPR arrays of *Clostridium novyi sensu lato*. Their analysis revealed a high correlation with strain distribution previously obtained using other tools, in particular the distribution of genetic lineages (Skarin and Segerman, 2014; Woudstra et al., 2016). Strains previously identified as being closely related by genomic comparison indeed harbor the same or almost the same spacers (Skarin and Segerman, 2014; Woudstra et al., 2016). Genotyping spacers thus holds promise for identifying common ancestors among strains, which would allow strain tracking with high resolution and accuracy or



highlight potential relatedness between isolates. This approach has been used for several pathogens such as *Campylobacter* (Kovanen et al., 2014), *Salmonella* (Xie et al., 2017), or *C. difficile* (Andersen et al., 2016).

A large proportion of the identified unique spacers originated from *Clostridium novyi* and *C. botulinum* BoNT C/D strains. However, the current number of available genomes of *Clostridium novyi sensu lato* is quite low (58) and the lineages, BoNT types and species are unequally represented. Detection of most of the unique spacers in *Clostridium novyi* and *C. botulinum* BoNT C/D genomes may thus be linked to the high number of strains or high diversity observed in both species. Among the few previously analyzed genomes (Skarin and Segerman, 2014), *Clostridium novyi* and *C. botulinum* BoNT C/D were distributed throughout three different lineages, whereas *C. botulinum* BoNT/D strains were classified into two lineages and *C. botulinum* type C, type D/C and *C. haemolyticum* in only one single lineage. This may explain why the detection of unique spacers was higher in *Clostridium novyi* and *C. botulinum* BoNT C/D than in the other genomes. However, more genomes are needed to confirm this observation. The high diversity of spacers detected in *Clostridium novyi* genomes despite the low number of available genomes (10) may also be linked to the high diversity of strains included in our study. They were isolated between 1920 and 2015, from five different countries and five different animal species.

Among the spacers whose origin was identified, plasmids and phages were found in equivalent proportions (respectively 51.2 and 48.8% of all MGEs), as already previously reported in *C. botulinum* group I, II, and III (Negahdaripour et al., 2017). This distribution is different from what is commonly reported in prokaryotes, where 85% of spacers usually map to phages (McGinn and Marraffini, 2019), including in other clostridia species such as *C. perfringens* (Long et al., 2019). In addition, some MGEs of *Clostridium novyi sensu lato* contain proteins characteristic of phages as well as plasmids (e.g., p1\_BKT015925, p1\_Cst, p4\_BKT015925, and p6\_Cst), resulting in misleading annotations in public databases, which may have artificially inflated the proportion of spacers from plasmids.

The exploration of CRISPR spacer sequences allowed the identification of their protospacers and thus past encounters of *Clostridium novyi sensu lato* with MGEs. The hosts from which the targeted MGEs originate provided partial information on the environmental microbiome composition in which *Clostridium novyi sensu lato* thrived and on the bacterial species with which MGEs are commonly exchanged. The main bacterial species from which the MGEs targeting *Clostridium novyi sensu lato* originated included *S. aureus*, *E. coli*, *C. botulinum*, *B. thuringiensis*, *B. miyamotoi*, *C. perfringens*, *S. epidermidis*, and *S. thermophilus* (38.1% of the total BLAST results). BLAST scores were lower for *Borrelia* than for *Clostridium* and *Bacillus* reflecting putative distant matches. *Clostridium botulinum* was the primary source of the main MGEs detected in this study, with the best BLAST scores, reflecting active dissemination of their MGEs throughout *Clostridium novyi sensu lato*.

The analysis of protospacers also provides information on the ability of MGEs to disseminate successfully within a bacterial

species. Of the 776 unique MGEs identified from the identified protospacers, only six of them represented up to 20.2% of the total BLAST results with the best BLAST scores, namely, phage-plasmid hybrids (p4\_BKT015925, p6\_Cst), BoNT pseudolysogenic phages (p1\_Cst, p1\_BKT015925), prophages (*C. Wou-2020a*), and a plasmid (p2\_BKT015925) originating from *C. botulinum*. Considering the large mobilome of *Clostridium novyi sensu lato* (Skarin and Segerman, 2014), it may be surprising that only a handful of MGEs are responsible for such large protospacer acquisition. This may imply that these elements are highly transferable throughout *Clostridium novyi sensu lato*. Although spacers targeting p1\_Cst, p1\_BKT015925 were detected in 84.1% of *C. botulinum* group III strains and 50% of non-toxic *Clostridium novyi sensu lato* strains (Table 3), no spacer targeting the BoNT phage from strains 1873, BKT2873, or 3859 was detected, suggesting no encounters between available sequenced strains and these MGEs. Half of the non-toxic strains did not have any spacers targeting p1\_Cst, p1\_BKT015925, and neither did six out of the seven lineage IV strains, suggesting no encounter between these strains and p1\_Cst and p1\_BKT015925, perhaps due to their presence in different ecological niches. This also suggests that these non-toxic strains do not have CRISPR-Cas systems trained against the BoNT phage from the genomes currently available in the database and may therefore potentially be infected by these MGEs. Movements of plasmids and toxin genes across lineage boundaries have been reported in *Clostridium novyi sensu lato* strains (Skarin and Segerman, 2014).

### Limited Acquisition of Mobilome Protospacers Results in Variable Protection Outcomes

Regarding spacers from CRISPR arrays located either in the chromosome or in MGEs, two main scenarios were encountered: (1) the genome containing the spacer(s) is devoid of the MGE carrying the protospacer; (2) the genome containing the spacer(s) also contains the protospacer MGE. A good example of the first situation is the plasmid p2\_BKT015925. Protospacers from p2\_BKT015925 were indeed only present in strains devoid of the plasmid. In that case, it seems that the CRISPR system efficiently prevents the acquisition of p2\_BKT015925 and plays its expected immunological role, demonstrating the active part CRISPR-Cas systems play to regulate the presence or absence of a specific MGE in *Clostridium novyi sensu lato* strains. This was also the case for prophage CWou-2020a, which was only detected in strains that had not acquired CWou-2020a protospacers. Interestingly, several *C. botulinum* BoNT C/D strains had a CWou-2020a homologous prophage (NZ\_AESB01000027) present in their genome, but only strains that had not acquired CWou-2020a protospacers, with incomplete type I-B CRISPR system located on MGEs. The absence of spacers targeting CWou-2020a in strains harboring NZ\_AESB01000027 may be related to the mutual exclusion of both prophages, either by integrating in the same location or by encoding multiple infection exclusion proteins (Bondy-Denomy et al., 2016), therefore avoiding the need for a CRISPR system defending specifically against CWou-2020a.

The second situation was observed for p1\_BKT015925, p4\_BKT015925 as well as p6\_BKT015925 and may

be considered a paradox, because both targets and the corresponding spacers were detected at the same time. Induction of autoimmunity and elimination of the targeted MGE should be expected from such situations. Previous studies have reported the detection of self-targeting spacers (targeting host chromosomal DNA) in many bacterial species (Wimmer and Beisel, 2020). The exact impact of self-targeting is unknown and may be minor. Moreover, some studies have reported examples in which a self-targeting spacer can be tolerated, for example through the activation of effective DNA repair or through mutations or even deletions of targets from the genome (Wimmer and Beisel, 2020). Alternatively, these self-targeting spacers may have biological roles other than immunity such as evolution or RNA degradation (Wimmer and Beisel, 2020). For example, three spacers from *C. botulinum* BoNT C/D strains localized in the chromosome target p1\_BKT015925 and had this BoNT phage within their genome. Interestingly, these spacers specifically target variable regions of p1\_BKT015925, suggesting that this may be part of a mechanism involved in the prevention of superinfection by several related BoNT phages at the same time. The existence of such a mechanism to prevent superinfection by multiple BoNT phages is also supported by the detection of protospacers from BoNT phages and located on other BoNT phages. For example, spacer 16868\_D\_90 (present in strains *C. botulinum* BoNT D 16868, BoNT D/C 1274, 1275, and 1276) located on the BoNT D and D/C phages target BoNT C/D phage p1\_BKT015925 on the CRISPR Cas7 gene (with four mismatches). The detection of spacers targeting p4\_BKT015925 found on BoNT phage p1\_BKT015925 and plasmid p2\_BKT015925 suggests this kind of competition between MGEs (Supplementary Tables S10 and S11).

## CONCLUSION

In the present study, the exploration of the CRISPR-Cas systems in the *Clostridium novyi sensu lato* genospecies found six different CRISPR-Cas system types (mainly I-B) belonging to class 1 and 2 CRISPR loci, protecting against up to 776 MGEs, but only a handful of them being widespread. Although the CRISPR mechanism and CRISPR-MGE interaction have been largely investigated, there is much less information on CRISPR systems used by MGEs targeting other MGEs, whose functions remain largely uncharacterized (Faure et al., 2019). This study highlighted that CRISPR-Cas systems are numerous in *Clostridium novyi sensu lato* strains, and may be present either in the bacterial chromosome or in MGEs. The components carrying CRISPR-Cas systems seemed to have been recruited as anti-MGE systems and for inter-MGE conflicts, to protect mainly against a restricted number of MGEs. The role and function of these CRISPR-Cas systems in the bacterial life cycle and more generally in pathogenesis, in particular when located on MGEs, now need to be explored. Furthermore, studying this restricted mobilome may prove very useful for biocontrol strategies against the pathogenic *Clostridium novyi sensu lato* species.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

## AUTHOR CONTRIBUTIONS

CW designed and coordinated the study. CM is the recipient of a grant that funded this study. TG, CW, and MD performed the bioinformatics analysis. TG, CM, MD, MC, and CW participated in data analysis and interpretation of the results. CM, MC, and CW supervised the study. CW, TG, and CM wrote the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

This study was supported by the “Institut Carnot AgriFood Transition” (TRACK'SPORE project), the French National Research Agency, and the Bretagne Regional Council of Région Bretagne. TG is the recipient of a PhD grant from ANSES and “Région Bretagne.”

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2021.711413/full#supplementary-material>

**Supplementary Table S1** | *Clostridium novyi sensu lato* total spacer content.

**Supplementary Table S2** | *Clostridium novyi sensu lato* duplicated spacer content.

**Supplementary Table S3** | *Clostridium novyi sensu lato* unique spacer content.

**Supplementary Table S4** | *Clostridium novyi sensu lato* unique identified spacer origin.

**Supplementary Table S5** | *Clostridium novyi sensu lato* unique unidentified spacer origin.

**Supplementary Table S6** | BLAST results on *Clostridium novyi sensu lato* unique spacer origins.

**Supplementary Table S7** | Proportion of MGEs among the BLAST results on the recognized spacers.

**Supplementary Table S8** | Composition of MGE host genus origins.

**Supplementary Table S10** | Main targeted mobile genetic element (MGE) genes.

**Supplementary Table S11** | Identification of mobile genetic element (MGE) spacer (protospacer) origin.

**Supplementary Table S12** | Identification of phage-inducible chromosomal island (PICl) spacer (protospacer) origin.

## REFERENCES

- Agren, J., Sundström, A., Häfström, T., and Segerman, B. (2012). Gegences: fragmented alignment of multiple genomes for determining phylogenomic distances and genetic signatures unique for specified target groups. *PLoS One* 7:e39107. doi: 10.1371/journal.pone.0039107
- Andersen, J. M., Shoup, M., Robinson, C., Britton, R., Olsen, K. E., and Barrangou, R. (2016). CRISPR diversity and microevolution in *Clostridium difficile*. *Genome Biol. Evol.* 8, 2841–2855. doi: 10.1093/gbe/evw203
- Angly, F. E., Felts, B., Breitbart, M., Salamon, P., Edwards, R. A., Carlson, C., et al. (2006). The marine viromes of four oceanic regions. *PLoS Biol.* 4:e368. doi: 10.1371/journal.pbio.0040368
- Bikard, D., Hatoum-Aslan, A., Mucida, D., and Marraffini, L. A. (2012). CRISPR interference can prevent natural transformation and virulence acquisition during in vivo bacterial infection. *Cell Host Microbe* 12, 177–186. doi: 10.1016/j.chom.2012.06.003
- Biswas, A., Gagnon, J. N., Brouns, S. J., Fineran, P. C., and Brown, C. M. (2013). CRISPRTarget: bioinformatic prediction and analysis of crRNA targets. *RNA Biol.* 10, 817–827. doi: 10.4161/rna.24046
- Bland, C., Ramsey, T. L., Sabree, E., Lowe, M., Brown, K., Kyrpides, N. C., et al. (2007). CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC Bioinformatics* 8:209. doi: 10.1186/1471-2105-8-209
- Bondy-Denomy, J., Qian, J., Westra, E. R., Buckling, A., Guttman, D. S., Davidson, A. R., et al. (2016). Prophages mediate defense against phage infection through diverse mechanisms. *ISME J.* 10, 2854–2866. doi: 10.1038/ismej.2016.79
- Boudry, P., Semenova, E., Monot, M., Datsenko, K. A., Lopatina, A., Sekulovic, O., et al. (2015). Function of the CRISPR-Cas system of the human pathogen *Clostridium difficile*. *mBio* 6, e01112–e01115. doi: 10.1128/mBio.01112-15
- Briner, A. E., Lugli, G. A., Milani, C., Duranti, S., Turroni, F., Guéimonde, M., et al. (2015). Occurrence and diversity of CRISPR-Cas systems in the genus *Bifidobacterium*. *PLoS One* 10:e0133661. doi: 10.1371/journal.pone.0133661
- Cady, K. C., White, A. S., Hammond, J. H., Abendroth, M. D., Karthikeyan, R. S. G., Lalitha, P., et al. (2011). Prevalence, conservation and functional analysis of *Yersinia* and *Escherichia* CRISPR regions in clinical *Pseudomonas aeruginosa* isolates. *Microbiology* 157, 430–437. doi: 10.1099/mic.0.045732-0
- Cohen, J. E., Wang, R., Shen, R.-F., Wu, W. W., and Keller, J. E. (2017). Comparative pathogenomics of *clostridium tetani*. *PLoS One* 12:e0182909. doi: 10.1371/journal.pone.0182909
- Couvin, D., Bernheim, A., Toffano-Nicoche, C., Touchon, M., Michalik, J., Néron, B., et al. (2018). CRISPRCasFinder, an update of CRISPRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. *Nucleic Acids Res.* 46, W246–W251. doi: 10.1093/nar/gky425
- Crestani, C., Forde, T. L., and Zadoks, R. N. (2020). Development and application of a prophage integrase typing scheme for group B *Streptococcus*. *Front. Microbiol.* 11:1993. doi: 10.3389/fmicb.2020.01993
- Eitzinger, S., Asif, A., Watters, K. E., Iavarone, A. T., Knott, G. J., Doudna, J. A., et al. (2020). Machine learning predicts new anti-CRISPR proteins. *Nucleic Acids Res.* 48, 4698–4708. doi: 10.1093/nar/gkaa219
- Faure, G., Shmakov, S. A., Yan, W. X., Cheng, D. R., Scott, D. A., Peters, J. E., et al. (2019). CRISPR-Cas in mobile genetic elements: counter-defence and beyond. *Nat. Rev. Microbiol.* 17, 513–525. doi: 10.1038/s41579-019-0204-7
- Fillo-Salom, A., Martínez-Rubio, R., Abdulrahman, R. F., Chen, J., Davies, R., and Penadés, J. R. (2018). Phage-inducible chromosomal islands are ubiquitous within the bacterial universe. *ISME J.* 12, 2114–2128. doi: 10.1038/s41396-018-0156-3
- Grissa, I., Vergnaud, G., and Pourcel, C. (2007). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* 8:638. doi: 10.1186/1471-2105-8-172
- Goldfarb, T., Sberro, H., Weinstock, E., Cohen, O., Doron, S., Charpak-Amikam, Y., et al. (2015). BREX is a novel phage resistance system widespread in microbial genomes. *EMBO J.* 34, 169–183. doi: 10.15252/embj.201489455
- Hargreaves, K. R., Flores, C. O., Lawley, T. D., and Clokie, M. R. J. (2014). Abundant and diverse clustered regularly interspaced short palindromic repeat spacers in *Clostridium difficile* strains and prophages target multiple phage types within this pathogen. *mBio* 5, e01045–e01013. doi: 10.1128/mBio.01045-13
- Harrington, L. B., Burstein, D., Chen, J. S., Paez-Espino, D., Ma, E., Witte, I. P., et al. (2018). Programmed DNA destruction by miniature CRISPR-Cas14 enzymes. *Science* 362:839. doi: 10.1126/science.aav4294
- Hawkins, J. S., Wong, S., Peters, J. M., Almeida, R., and Qi, L. S. (2015). Targeted transcriptional repression in bacteria using CRISPR interference (CRISPRi). *Methods Mol. Biol.* 1311, 349–362. doi: 10.1007/978-1-4939-2687-9\_23
- Karah, N., Samuelsen, Ø., Zarrilli, R., Sahl, J. W., Wai, S. N., and Uhlin, B. E. (2015). CRISPR-cas subtype I-fb in *Acinetobacter baumannii*: evolution and utilization for strain subtyping. *PLoS One* 10:e0118205. doi: 10.1371/journal.pone.0118205
- Khan, M. Z., Haider, S., Mansoor, S., and Amin, I. (2019). Targeting plant ssDNA viruses with engineered miniature CRISPR-Cas14a. *Trends Biotechnol.* 37, 800–804. doi: 10.1016/j.tibtech.2019.03.015
- Kieper, S. N., Almendros, C., Behler, J., McKenzie, R. E., Nobrega, F. L., Haagsma, A. C., et al. (2018). Cas4 facilitates PAM-compatible spacer selection during CRISPR adaptation. *Cell Rep.* 22, 3377–3384. doi: 10.1016/j.celrep.2018.02.103
- Kim, K.-H., Chang, H.-W., Nam, Y.-D., Roh, S. W., Kim, M.-S., Sung, Y., et al. (2008). Amplification of uncultured single-stranded DNA viruses from rice paddy soil. *Appl. Environ. Microbiol.* 74:5975. doi: 10.1128/AEM.01275-08
- Koonin, E. V., and Makarova, K. S. (2019). Origins and evolution of CRISPR-Cas systems. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 374:20180087. doi: 10.1098/rstb.2018.0087
- Koonin, E. V., Makarova, K. S., and Zhang, F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. *Curr. Opin. Microbiol.* 37, 67–78. doi: 10.1016/j.mib.2017.05.008
- Kovanen, S. M., Kivistö, R. I., Rossi, M., and Hänninen, M. L. (2014). A combination of MLST and CRISPR typing reveals dominant *Campylobacter jejuni* types in organically farmed laying hens. *J. Appl. Microbiol.* 117, 249–257. doi: 10.1111/jam.12503
- Lin, J., Fuglsang, A., Kjeldsen, A. L., Sun, K., Bhoobalan-Chitty, Y., and Peng, X. (2020). DNA targeting by subtype I-D CRISPR-Cas shows type I and type III features. *Nucleic Acids Res.* 48, 10470–10478. doi: 10.1093/nar/gkaa749
- Long, J., Xu, Y., Ou, L., Yang, H., Xi, Y., Chen, S., et al. (2019). Diversity of CRISPR/Cas system in *Clostridium perfringens*. *Mol. Gen. Genomics* 294, 1263–1275. doi: 10.1007/s00438-019-01579-3
- Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., et al. (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 13, 722–736. doi: 10.1038/nrmicro3569
- Makarova, K. S., Wolf, Y. I., Iranzo, J., Shmakov, S. A., Alkhnbashi, O. S., Brouns, S. J. J., et al. (2020). Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* 18, 67–83. doi: 10.1038/s41579-019-0299-x
- Mcbride, T. M., Schwartz, E. A., Kumar, A., Taylor, D. W., Fineran, P. C., and Fagerlund, R. D. (2020). Diverse CRISPR-Cas complexes require independent translation of small and large subunits from a single gene. *bioRxiv* [Preprint]. doi: 10.1101/2020.04.18.045682
- McGinn, J., and Marraffini, L. A. (2019). Molecular mechanisms of CRISPR-Cas spacer acquisition. *Nat. Rev. Microbiol.* 17, 7–12. doi: 10.1038/s41579-018-0071-7
- Mertaogja, A., Nowakowska, M. B., Mascher, G., Heljanko, V., Groothuis, D., Minton, N. P., et al. (2021). CRISPR-Cas9-based toolkit for *clostridium botulinum* group II spore and sporulation research. *Front. Microbiol.* 12:617269. doi: 10.3389/fmicb.2021.617269
- Mir, A., Edraki, A., Lee, J., and Sontheimer, E. J. (2018). Type II-C CRISPR-Cas9 biology, mechanism, and application. *ACS Chem. Biol.* 13, 357–365. doi: 10.1021/acscchembio.7b00855
- Negahdaripour, M., Nezafat, N., Hajighahramani, N., Rahmatbadi, S. S., and Ghasemi, Y. (2017). Investigating CRISPR-Cas systems in *clostridium botulinum* via bioinformatics tools. *Infect. Genet. Evol.* 54, 355–373. doi: 10.1016/j.meegid.2017.06.027
- Pan, M., Nethery, M. A., Hidalgo-Cantabrana, C., and Barrangou, R. (2020). Comprehensive mining and characterization of CRISPR-Cas systems in *Bifidobacterium*. *Microorganisms* 8:720. doi: 10.3390/microorganisms8050720
- Penadés, J. R., and Christie, G. E. (2015). The phage-inducible chromosomal islands: a family of highly evolved molecular parasites. *Annu. Rev. Virol.* 2, 181–201. doi: 10.1146/annurev-virology-031413-085446
- Pfeifer, E., Moura De Sousa, J. A., Touchon, M., and Rocha, E. P. C. (2021). Bacteria have numerous distinctive groups of phage-plasmids with conserved phage and variable plasmid gene repertoires. *Nucleic Acids Res.* 49, 2655–2673. doi: 10.1093/nar/gkab064

- Pyenson, N. C., and Marraffini, L. A. (2017). Type III CRISPR-Cas systems: when DNA cleavage just isn't enough. *Curr. Opin. Microbiol.* 37, 150–154. doi: 10.1016/j.mib.2017.08.003
- Pyne, M. E., Bruder, M. R., Moo-Young, M., Chung, D. A., and Chou, C. P. (2016). Harnessing heterologous and endogenous CRISPR-Cas machineries for efficient markerless genome editing in *clostridium*. *Sci. Rep.* 6:25666. doi: 10.1038/srep25666
- Rosario, K., Nilsson, C., Lim, Y. W., Ruan, Y., and Breitbart, M. (2009). Metagenomic analysis of viruses in reclaimed water. *Environ. Microbiol.* 11, 2806–2820. doi: 10.1111/j.1462-2920.2009.01964.x
- Shariat, N., Dimarzio, M. J., Yin, S., Dettinger, L., Sandt, C. H., Lute, J. R., et al. (2013). The combination of CRISPR-MVLST and PFGE provides increased discriminatory power for differentiating human clinical isolates of *salmonella enterica subsp. enterica* serovar Enteritidis. *Food Microbiol.* 34, 164–173. doi: 10.1016/j.fm.2012.11.012
- Shmakov, S., Smargon, A., Scott, D., Cox, D., Pyzocha, N., Yan, W., et al. (2017). Diversity and evolution of class 2 CRISPR-Cas systems. *Nat. Rev. Microbiol.* 15, 169–182. doi: 10.1038/nrmicro.2016.184
- Sikorski, A., Argüello-Astorga, G. R., Dayaram, A., Dobson, R. C., and Varsani, A. (2013). Discovery of a novel circular single-stranded DNA virus from porcine faeces. *Arch. Virol.* 158, 283–289. doi: 10.1007/s00705-012-1470-0
- Silas, S., Lucas-Elio, P., Jackson, S. A., Aroca-Crevillén, A., Hansen, L. L., Fineran, P. C., et al. (2017). Type III CRISPR-Cas systems can provide redundancy to counteract viral escape from type I systems. *eLife* 6:27601. doi: 10.7554/eLife.27601
- Skarin, H., Hafström, T., Westerberg, J., and Segerman, B. (2011). *Clostridium botulinum* group III: A group with dual identity shaped by plasmids, phages and mobile elements. *BMC Genomics* 12:185. doi: 10.1186/1471-2164-12-185
- Skarin, H., and Segerman, B. (2011). Horizontal gene transfer of toxin genes in *clostridium botulinum*: involvement of mobile elements and plasmids. *Mob. Genet. Elem.* 1, 213–215. doi: 10.4161/mge.1.3.17617
- Skarin, H., and Segerman, B. (2014). Plasmidome interchange between *clostridium botulinum*, *clostridium novyi* and *clostridium haemolyticum* converts strains of independent lineages into distinctly different pathogens. *PLoS One* 9:e107777. doi: 10.1371/journal.pone.0107777
- Sorek, R., Kunin, V., and Hugenholtz, P. (2008). CRISPR—a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat. Rev. Microbiol.* 6, 181–186. doi: 10.1038/nrmicro1793
- Sullivan, M. J., Petty, N. K., and Beatson, S. A. (2011). Easyfig: a genome comparison visualizer. *Bioinformatics* 27, 1009–1010. doi: 10.1093/bioinformatics/btr039
- Tamulaitis, G., Venclovas, Č., and Siksnys, V. (2017). Type III CRISPR-Cas immunity: major differences brushed aside. *Trends Microbiol.* 25, 49–61. doi: 10.1016/j.tim.2016.09.012
- Terns, M. P. (2018). CRISPR-based technologies: impact of RNA-targeting systems. *Mol. Cell* 72, 404–412. doi: 10.1016/j.molcel.2018.09.018
- Thomas, P., Semmler, T., Eichhorn, I., Lübke-Becker, A., Werckenthin, C., Abdel-Ghil, M. Y., et al. (2017). First report of two complete *clostridium* chauvoei genome sequences and detailed in silico genome analysis. *Infect. Genet. Evol.* 54, 287–298. doi: 10.1016/j.meegid.2017.07.018
- Wimmer, F., and Beisel, C. L. (2020). CRISPR-Cas systems and the paradox of self-targeting spacers. *Front. Microbiol.* 10:3078. doi: 10.3389/fmicb.2019.03078
- Woudstra, C., Le Marechal, C., Souillard, R., Anniballi, F., Auricchio, B., Bano, L., et al. (2017). Investigation of *clostridium botulinum* group III's mobilome content. *Anaerobe* 49, 71–77. doi: 10.1016/j.anaerobe.2017.12.009
- Woudstra, C., Le Marechal, C., Souillard, R., Bayon-Auboyer, M. H., Mermoud, I., Desoutter, D., et al. (2016). New insights into the genetic diversity of *clostridium botulinum* group III through extensive genome exploration. *Front. Microbiol.* 7:757. doi: 10.3389/fmicb.2016.00757
- Xie, X., Hu, Y., Xu, Y., Yin, K., Li, Y., Chen, Y., et al. (2017). Genetic analysis of *salmonella enterica serovar Gallinarum biovar Pullorum* based on characterization and evolution of CRISPR sequence. *Vet. Microbiol.* 203, 81–87. doi: 10.1016/j.vetmic.2017.02.010
- Yin, S., Jensen, M. A., Bai, J., Debroy, C., Barrangou, R., and Dudley, E. G. (2013). The evolutionary divergence of Shiga toxin-producing *Escherichia coli* is reflected in clustered regularly interspaced short palindromic repeat (CRISPR) spacer composition. *Appl. Environ. Microbiol.* 79, 5710–5720. doi: 10.1128/AEM.00950-13
- Zhang, R., Xu, W., Shao, S., and Wang, Q. (2021). Gene silencing through CRISPR interference in bacteria: current advances and future prospects. *Front. Microbiol.* 12:635227. doi: 10.3389/fmicb.2021.635227
- Zheng, Z., Zhang, Y., Liu, Z., Dong, Z., Xie, C., Bravo, A., et al. (2020). The CRISPR-Cas systems were selectively inactivated during evolution of *Bacillus cereus* group for adaptation to diverse environments. *ISME J.* 14, 1479–1493. doi: 10.1038/s41396-020-0623-5
- Zimmermann, L., Stephens, A., Nam, S. Z., Rau, D., Kübler, J., Lozajic, M., et al. (2018). A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.* 430, 2237–2243. doi: 10.1016/j.jmb.2017.12.007

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Le Gratiot, Le Marechal, Devaere, Chemaly and Woudstra. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

### 2.2.3. Utilisation des systèmes CRISPR-Cas pour le développement d'un outil de typage

L'objectif de la thèse étant de développer des outils de typage, la possibilité de développer un ou des outils de typage à partir des locus CRISPR-Cas présents chez *C. novyi sensu lato* a été étudiée.

De manière générale, il est possible de développer une méthode de typage basée sur les systèmes CRISPR-Cas si : les systèmes CRISPR-Cas sont présents chez l'ensemble des souches étudiées et si les systèmes détectés sont conservés au sein des souches de façon à proposer un outil applicable à l'ensemble des souches en ciblant une zone commune.

Concernant *C. novyi sensu lato*, 6 systèmes partiels ou complets ont été retrouvés sur 55 des 58 génomes testés, soit 95 % de prévalence. Une approche basée sur l'analyse de ces séquences ne pourra donc pas être applicable sur l'ensemble des souches de *C. novyi sensu lato*. Cependant, l'absence de système n'a été mise en évidence que sur certains génomes de la lignée IV. Un outil pourrait donc être applicable pour les 4 autres lignées (Ia, Ib, II et III).

Les systèmes CRISPR-Cas identifiés dans ces 55 génomes présentent une forte variabilité. Pas moins de 6 systèmes différents ont ainsi été identifiés, avec un nombre de systèmes variant de 0 à 3 selon la souche. De plus, les séquences des répétitions sont elles aussi très variables avec plus de 30 séquences différentes retrouvées. Cette variabilité dans les séquences des répétitions montre que celles-ci ne constituent pas une cible pertinente pour le développement d'une méthode de typage applicable à l'ensemble des souches de *C. novyi sensu lato* comme cela a pu être développé pour *Mycobacterium tuberculosis* ou *Salmonella* (Fabre et al., 2012; Kamerbeek et al., 1997).

En conclusion, l'analyse des génomes de *C. novyi sensu lato* montre une diversité importante que ce soit au niveau des systèmes CRISPR-Cas ou des séquences des répétitions au contraire d'autres espèces bactériennes comme *M. tuberculosis* ou *Salmonella* pour lesquelles des outils de typage basés sur les systèmes CRISPR-Cas ont été développés (Fabre et al., 2012; Fricke et al., 2011; Jansen et al., 2002; van Embden et al., 2000). Le développement d'une méthode de typage ciblant les systèmes CRISPR-Cas pour le typage de *C. novyi sensu lato* ne semble donc pas approprié. Cependant, cette approche pourrait être envisagée pour du sous-typage au sein de certaines lignées (I et III) comme cela a été suggéré pour *Campylobacter* (Kovanen et al., 2014). Pour cela, il serait nécessaire de pouvoir identifier au préalable la lignée à laquelle appartiennent les souches afin de proposer et d'utiliser un outil adapté à chacune d'entre elles. Le développement de systèmes PCR pour identifier la lignée d'une souche est actuellement en cours suite à l'identification de cibles dans le cadre de ce projet de thèse.

## Discussion et Perspectives

Bien que rapporté dans la littérature scientifique depuis plus d'un siècle, le botulisme animal est une maladie pour laquelle de nombreux aspects restent encore à étudier, tels que l'identification des facteurs de risque à l'origine de l'initiation d'un épisode, la compréhension de la pathogénèse ou bien encore le rôle de la contamination environnementale dans le cycle de cette maladie (Anses, 2021). Des épisodes de botulisme animal sont rapportés annuellement à l'échelle mondiale, cependant aucune étude n'a porté à ce jour sur l'étude de liens potentiels entre ces épisodes à une échelle globale, en particulier peu d'études ont été menées jusqu'à présent sur la comparaison des souches de *C. botulinum* impliquées dans les foyers de botulisme animal. Cet élément est pourtant essentiel pour mieux comprendre les mécanismes et proposer des mesures de prévention ou de gestion. A ce jour, une centaine de génomes sont disponibles dans les bases de données publiques, ce qui s'explique notamment par l'absence de méthode simple et robuste pour isoler les souches de *C. botulinum* du groupe III impliquées dans le botulisme animal.

Cette absence de méthode d'isolement constitue actuellement un verrou méthodologique qui n'a pas permis le développement d'outils de typage comme cela a pu être le cas pour d'autres bactéries pathogènes. Disposer d'outils pour permettre de tracer les souches de *C. botulinum* impliquées dans les épisodes de botulisme animal et de les caractériser au niveau moléculaire est primordial pour progresser dans la compréhension de cette maladie.

Les premières études menées sur les génomes de *C. botulinum* du groupe III ont démontré une proximité génétique avec deux autres espèces de clostridies, *C. novyi* et *C. haemolyticum*, au point d'être difficilement distinguables (Sasaki et al., 2001; Skarin et al., 2011). Ces similarités ont eu pour conséquence la création du groupe d'espèces *Clostridium novyi sensu lato*. Au sein du groupe *C. novyi sensu lato*, 4 lignées ont été définies à l'aide d'une analyse comparative des génomes de ce groupe (Skarin & Segerman, 2014).

Ainsi, ce travail de thèse avait pour objectif d'élaborer et évaluer des approches permettant de typer ces souches de *C. botulinum* du groupe III liées à des épisodes de botulisme animal en tenant compte de ce contexte particulier.

## 1. Développement d'une méthode d'isolement des souches de *C. botulinum* du groupe III

La première étape a consisté à améliorer la méthode d'isolement des souches. Une méthode rapide et facile à utiliser a pu être mise au point pour isoler les souches de *C. botulinum* du groupe III à partir des foies d'origine aviaire. Notre approche s'est basée sur la capacité de résistance des spores au cours de l'extraction d'ADN réalisée avec le kit InstaGene. Cette étape d'extraction permet d'éliminer une grande partie des bactéries présentes dans l'enrichissement de départ. Ce kit a été choisi car il avait déjà été démontré qu'il ne permettait pas l'extraction d'ADN intracellulaire des spores de l'espèce *Bacillus cereus* (Brauge et al., 2018). La méthode d'isolement a été évaluée sur un panel d'enrichissements représentatifs des épisodes de botulisme aviaire survenus en France et confirmés par le LNR botulisme aviaire entre 2013 et 2019. Pour s'assurer que la méthode fonctionne sur l'ensemble des souches de *C. botulinum* du groupe III, des foies ont été artificiellement contaminés avec 4 types toxiniques (C, D, C/D et D/C). Sur les 22 enrichissements naturellement contaminés testés au cours de l'étude, 21 souches ont pu être isolées. Le gène codant pour la toxine a été conservé sur 14 de ces 21 souches, ce qui est un taux plus élevé qu'avec les méthodes d'isolement utilisées auparavant (Anza et al., 2014a; Franciosa et al., 1996). Il est à noter que la méthode ne fonctionne que sur des enrichissements de foies. En ce qui concerne les échantillons d'environnement, la difficulté pour isoler les souches s'explique par une forte proportion d'espèces bactériennes résistantes à l'extraction (ce qui a pour conséquence l'impossibilité d'obtenir des colonies isolées sur les géloses) et par une présence de *C. botulinum* en plus faible quantité que dans les organes. L'utilisation d'autres milieux gélosés pour la culture de *C. novyi sensu lato* (Milieux BAB2, Fastidious Anaerobe Agar ou McClung Toabe agar par exemple) pourrait être évaluée pour déterminer si l'isolement pourrait être facilité par rapport au milieu EYA.

Le développement de cette nouvelle méthode d'isolement est une avancée pour explorer la diversité des souches du groupe d'espèces *C. novyi sensu lato*. L'obtention d'un plus grand nombre de souches isolées permettra, à l'aide du séquençage de génomes complets, de mieux tracer les souches lors des épisodes de botulisme animal en plus d'améliorer nos connaissances sur la physiologie de la bactérie. L'isolement des souches pourrait cependant encore être plus facilité en développant un milieu de culture sélectif, notamment pour les échantillons d'environnement pour lesquels l'isolement n'est pas optimal pour le moment. Pour ce faire, l'étude des génomes pour trouver des caractéristiques spécifiques des souches de ce groupe d'espèces (gènes de résistance à certains antibiotiques ou des voies métaboliques spécifiques) est nécessaire. D'autres outils basés sur l'immunocapture, comme cela a déjà été fait pourraient être explorés (Anza et al., 2014a). La combinaison de l'immunocapture et de la cytométrie en flux a par exemple permis d'isoler avec

succès des souches d'espèces bactériennes présentes dans le microbiote intestinal comme *Faecalibacterium praunitzii* (Bellais et al., 2022). L'immunocapture qui nécessite la production d'anticorps ciblant la bactérie d'intérêt n'a pas été envisagée dans un premier temps du fait de la réglementation MOT et des difficultés d'expédition des souches vers un prestataire pour obtenir les anticorps.

Une alternative à cette étape d'isolement pour l'étude des souches de *C. botulinum* est le développement de systèmes PCR suffisamment spécifiques pour s'affranchir de l'étape d'isolement. Un premier outil nommé MLVA a donc été évalué au cours de cette thèse.

## **2. Évaluation de l'outil de typage MLVA**

Au cours des dernières décennies, plusieurs outils de typage ont été utilisés pour tracer les souches responsables de botulisme animal (Anza et al., 2014a; Hill et al., 2007; Myllykoski et al., 2009). Juste avant le début de ma thèse, un schéma MLVA a été développé pour typer les souches de *C. novyi sensu lato* par le CNR italien (Auricchio et al., 2019). Cet outil est déjà utilisé pour le typage d'un grand nombre de bactéries pathogènes dont les souches de *C. botulinum* des groupes I et II (Fillo et al., 2011; Macdonald et al., 2008; Umeda et al., 2013). Dans le cadre de ma thèse, j'ai eu l'opportunité d'effectuer un séjour de recherche de 10 jours à l'Istituto Superiore di Sanita de Rome dans l'équipe de Fabrizio Anniballi pour être formé à la méthode MLVA et pouvoir la transférer au laboratoire à mon retour. J'ai ensuite évalué la méthode sur les souches isolées au laboratoire et sur des échantillons bruts collectés dans le cadre de 2 épisodes de botulisme animal. Les résultats obtenus ont montré que la méthode est applicable pour les souches de la lignée Ia (majoritairement composée de souches de *C. botulinum* du groupe III de type C/D) avec une amplification de l'ensemble des cibles et partiellement pour celles de la lignée Ib (majoritairement composée de souches de *C. botulinum* du groupe III de type D/C) avec une amplification d'une partie des cibles. Sur les 9 régions VNTR sélectionnées initialement pour le schéma MLVA, l'amplification n'était pas systématique pour deux d'entre elles, avec un manque de reproductibilité au sein des souches ou entre les échantillons. Il a donc été décidé de supprimer ces 2 VNTR parmi les cibles de la méthode. Par ailleurs, une absence d'amplification de certaines voire de toutes les cibles VNTR a été observée pour les souches des lignées II, III et IV. Cela peut s'expliquer, selon la cible, soit par l'absence du VNTR dans le génome de ces souches soit par des mutations au niveau des régions ciblées par les amorces PCR.

En conclusion, la méthode MLVA s'est révélée extrêmement intéressante pour tracer des souches lors d'un épisode de botulisme et pour valider les hypothèses émises sur l'origine de



contamination lorsque l'épisode est lié à une souche de la lignée Ia. Le schéma MLVA évalué est ainsi applicable pour les souches de la lignée Ia. Pour les autres souches, il serait intéressant de construire de nouveaux schémas spécifiques de chaque lignée en modifiant les séquences des amorces des régions VNTR déjà existantes ou en recherchant de nouvelles cibles à l'aide des génomes disponibles dans les bases de données. L'étude menée sur un épisode de type D/C a permis l'amplification systématique de 3 des 7 VNTR (Souillard et al., 2021). La disponibilité de nouveaux génomes suite à la publication de Fillo et al. en 2021, notamment de génomes des lignées Ib et II, ouvre de nouvelles perspectives pour ces deux lignées (Fillo et al., 2021). La recherche d'autres VNTR spécifiques à ces deux lignées constitue une piste à explorer pour le typage de ces souches.

Ces résultats montrent que le développement d'un outil de typage PCR discriminant et commun à l'ensemble des lignées ne semble pas adapté du fait des trop fortes disparités retrouvées entre les souches de lignées différentes et ne semble donc pas être la solution à privilégier. Des cibles spécifiques à chaque lignée doivent être définies pour pouvoir, dans un premier temps, déterminer la lignée des souches étudiées. Pour ce faire, une analyse comparative des génomes est nécessaire pour sélectionner ces cibles. Cependant, le manque de génome dans les bases de données, notamment pour les lignées II, III et IV, peut avoir comme conséquence un manque de spécificité des cibles lors de la sélection, le panel étant trop petit pour être représentatif des souches présentes naturellement. Afin d'obtenir un nombre de génomes plus important, les différentes étapes en amont du séquençage doivent être optimisées, à commencer par l'extraction d'ADN qui s'est révélée être plus difficile que prévue au cours du projet de thèse.

### **3. Sélection des protocoles d'extraction d'ADN et séquençage du génome**

Les bactéries sporulées Gram positif dont fait partie les trois espèces du groupe *C. novyi sensu lato* sont connues pour être difficiles à lyser. En forme végétative, la couche de peptidoglycane des bactéries à coloration de Gram positif (Cabeen & Jacobs-Wagner, 2005) est plus difficile à rompre que celles des espèces à coloration de Gram négatif. En forme d'endospore, la bactérie est encore plus résistante et peut même survivre à toute les étapes de certains protocoles d'extraction d'ADN (Knüpfner et al., 2020; Le Gratiet et al., 2020). Pour les souches de *C. novyi sensu lato*, la libération d'enzymes DNases pouvant fragmenter l'ADN extrait lors de la lyse bactérienne est une limite supplémentaire à l'obtention d'un rendement d'extraction acceptable pour du séquençage de génome. Toutes ces difficultés, en plus de l'isolement des souches, expliquent le faible nombre de génomes disponibles dans les bases de données publiques. En France, la réglementation sur les MOT rajoute des contraintes supplémentaires (peu de laboratoires autorisés à séquencer les souches de *C.*

*botulinum* du groupe III, coût du transport, délai administratif d'obtention des autorisations de cession). Dans les études ayant permis le séquençage de génomes de souches de *C. novyi sensu lato*, différents protocoles ont été utilisés pour extraire l'ADN (Fillo et al., 2021; Skarin et al., 2011; Skarin & Segerman, 2014; Woudstra et al., 2017; Woudstra et al., 2021).

Lors de ce projet, une comparaison de 5 protocoles d'extraction d'ADN a été effectuée dans le but de déterminer lesquels sont adaptés pour effectuer du séquençage de génome short-reads et/ou long-reads. Les résultats de cette étude montrent que 2 protocoles testés permettent d'obtenir une quantité d'ADN suffisante pour du séquençage short-reads : le MasterPure Complete DNA and RNA purification Kit et un protocole interne développé dans une étude où le séquençage du génome complet de 4 souches de *C. botulinum* du groupe III a été effectué (Woudstra et al., 2021).

Dans notre étude, le kit MasterPure Complete DNA and RNA purification et le protocole interne permettent d'obtenir un meilleur rendement à partir d'un enrichissement de 6 h. Une étude où 20 protocoles d'extraction ont été comparés sur la bactérie sporulée *Bacillus anthracis*, le kit MasterPure Complete DNA and RNA purification s'est également révélé être le protocole avec le plus haut rendement d'extraction d'ADN à partir de cellules végétatives. Ce kit n'est cependant pas adapté pour extraire l'ADN de spores bactériennes (Knüpfer et al., 2020).

Concernant les kits où le rendement n'est pas suffisant pour le séquençage de génome, une optimisation des protocoles est possible. Pour le MasterPure Gram Positive purification kit, pourtant spécifique des bactéries à coloration de Gram positif, le rendement insuffisant peut s'expliquer par une incubation à 37 °C trop longue lors de l'étape de lyse, qui pourrait laisser le temps aux enzymes DNases libérées de dégrader l'ADN. Le volume de culture à extraire peut aussi être augmenté pour avoir plus de cellules végétatives à lyser et donc extraire plus d'ADN. Enfin, le volume d'élution final va aussi déterminer la concentration d'ADN finale. Avec le kit QIAamp DNA mini kit, le volume d'élution de 200 µl est trop élevé pour obtenir une concentration d'ADN suffisante pour le séquençage de génome alors que la quantité extraite est adaptée pour du séquençage short-reads. En ce qui concerne la taille des fragments d'ADN extrait, donnée importante pour pouvoir faire le séquençage long-reads, les résultats indiquent que le protocole interne est plus adapté pour ce type de séquençage, en plus d'avoir déjà été utilisé dans une précédente étude (Woudstra et al., 2021).

En ce qui concerne la préparation des bibliothèques, la faible proportion de bases GC des génomes de *C. novyi sensu lato* peut entraîner des difficultés lors de l'étape de tagmentation avec le kit standard Nextera XT utilisé pour le séquençage Illumina. En effet, la taille des fragments obtenue n'est pas uniforme et la quantité d'ADN amplifiée peut se révéler insuffisante pour obtenir une bonne

couverture lors du séquençage. L'utilisation du kit Nextera DNA flex semble plus adaptée aux souches de *C. novyi sensu lato* et plus globalement aux génomes à bas GC%.

Plusieurs séries de séquençage ont été effectuées au cours de cette thèse, avec des résultats mitigés. Des séquences contaminantes (provenant d'autres clostridies) ont été retrouvées dans les données de séquençage et les génomes assemblés ont une taille plus grande que la taille attendue. Une des hypothèses expliquant ce résultat pourrait être la présence de plusieurs souches de *C. novyi sensu lato* qui auraient été « isolées » en même temps avec le nouveau protocole d'isolement (des souches commensales en plus de la souche responsable du botulisme par exemple). La présence naturelle de souches du groupe *C. novyi sensu lato* n'a pas été étudiée pour le moment, cet élément n'est donc pas connu. La présence de souches multiples n'a pas été détectée au moment de l'étape d'isolement, ceci étant probablement lié à des morphologies très similaires des différentes souches sur milieu gélosé et n'a été révélée que via les données de séquençage. La détection de souches de la lignée Ia et de la lignée IV lors de l'étude de cas (paragraphe 2.1.3.2 de la partie sur l'optimisation des étapes du séquençage et les analyses comparatives des génomes de *C. novyi sensu lato*) semble aller dans le sens de cette hypothèse de co-infection par au moins une souche toxique et une souche non-toxique, retrouvées au niveau des foies des animaux symptomatiques. Le rôle de la présence de plusieurs souches au cours du processus d'infection reste à élucider et ouvre de nouvelles pistes de recherche dans la compréhension de la pathogénèse du botulisme aviaire. L'existence possible de cette co-infection n'avait pas été mise en évidence via la MLVA, les souches de la lignée IV n'amplifiant pas les cibles PCR du schéma MLVA actuel. Seule la présence des souches de la lignée Ia avait donc pu être mise en évidence.

Parmi les résultats de séquences non exploitables, la présence de séquences d'autres clostridies a été mise en évidence. Cependant la recherche de ces autres clostridies dans les stocks glycérolés des souches séquencées n'a pas permis de les détecter. Une contamination ultérieure lors de l'étape d'extraction par exemple pourrait être possible et à explorer.

En conclusion, il reste encore beaucoup de travail à effectuer pour que le séquençage de génomes d'une souche de *C. novyi sensu lato* puisse se faire en routine comme c'est le cas pour de nombreux pathogènes tels que *Salmonella* ou *Escherichia coli*.

Les innovations permanentes, notamment dans le domaine de la métagénomique, pourraient dans les prochaines années permettre d'obtenir des génomes sans passer par l'étape d'isolement des souches via le séquençage Shotgun qui commence à être étudié pour certaines bactéries pathogènes (Buytaers et al., 2020; Buytaers et al., 2021; Saltykova et al., 2020). Comme pour les outils PCR,

l'approche métagénomique qui s'affranchit de l'étape d'isolement (étape qui constitue actuellement l'un des problèmes majeurs dans l'obtention des séquences) constitue l'une des pistes les plus prometteuses pour l'étude moléculaire des souches impliquées dans les épisodes de botulisme animal.

#### 4. Recherche de cibles d'intérêt pour le typage de *C. novyi sensu lato*

L'évaluation de la méthode MLVA a révélé que l'utilisation d'un outil commun à toutes les lignées de *C. novyi sensu lato* n'est pas pertinente au vue des différences génomiques trop importantes entre les souches pour proposer une cible adaptée. Le développement d'outils de typage spécifiques pour chaque lignée est donc la clef pour discriminer les souches de manière fiable. Avant de trouver des cibles pertinentes pour chaque lignée, le développement de systèmes PCR permettant de déterminer directement la lignée de la souche à typer est nécessaire pour pouvoir sélectionner les outils de typage à utiliser.

D'après les études sur les gènes *fliC* et le contenu en EGM réalisées avant et pendant cette thèse (Fillo et al., 2021; Skarin & Segerman, 2014; Woudstra et al., 2018; Woudstra et al., 2015), certaines caractéristiques sont retrouvées dans les lignées :

- Les souches de *C. botulinum* possédant le gène *bont C/D* sont principalement des souches de la lignée Ia et possèdent tous le gène *fliC-I*. Elles possèdent toutes des EGM des catégories 2 et 3 (PG 2 et PG 3) et sont les seules à posséder ceux de la catégorie 7 (PG7).
- Les souches de *C. botulinum* possédant le gène *bont D/C* sont principalement des souches de la lignée Ib. Les gènes *fliC-I, II, IV, VI* et *VII* sont retrouvées dans cette lignée. Les gènes *fliC-IV, VI* et *VII* sont cependant retrouvés spécifiquement dans la lignée Ib. Cette lignée est la seule à avoir des EGM de la catégorie 14 (PG 14).
- Les souches de *C. botulinum* possédant le gène *bont C* sont principalement des souches de la lignée II qui possèdent le gène *fliC-II*. Les EGM des catégories 4 et 23 (PG4 et PG23) sont uniquement retrouvées dans cette lignée.

Pour les lignées III et IV, le faible nombre de génomes présents dans les bases de données (respectivement 3 et 7 génomes) ne permet pas de mettre en lumière des caractéristiques spécifiques de ces lignées.

Malgré ces tendances, les événements de transferts horizontaux du bactériophage portant le gène *bont* entre souches de *C. novyi sensu lato*, la perte potentielle d'EGM et la présence de plusieurs gènes *fliC* retrouvées dans la lignée Ib ne permettent pas l'utilisation de ces cibles pour valider avec certitude la lignée de la souche à tester. L'analyse comparative des 118 génomes disponibles lors de

cette thèse a permis de trouver un ou plusieurs gènes de flagelline spécifiques de chaque lignée. Des systèmes PCR ont été dessinés et sont actuellement en cours d'évaluation. L'objectif final serait d'arriver à un protocole d'une ou plusieurs PCR permettant de déterminer la lignée de la souche avec précision à partir de l'extrait d'ADN brut.

## 5. Analyse des systèmes CRISPR-Cas

Dans une étude précédente visant à étudier le contenu en EGM de *C. botulinum* du groupe III, les systèmes CRISPR-Cas ont été étudiés sur des souches de type C/D et D/C (Woudstra et al., 2016). Des systèmes incomplets ont été retrouvés sur les souches de type C/D et des systèmes complets chez les souches de type D/C (Woudstra et al., 2016). L'utilisation de ces systèmes en tant que cible pour du typage a été testé pour les souches de type D/C génétiquement très proches et a permis de les différencier (Woudstra et al., 2016). Pour les souches de type C/D, cette approche n'a cependant pas été évoquée (Woudstra et al., 2016). Dans une autre étude, il a été montré que les systèmes CRISPR-cas sont présents dans 80 % des génomes de *C. botulinum* testés et qu'ils sont retrouvés dans les EGM chez *C. botulinum* du groupe III (Negahdaripour et al., 2017). Au cours de la thèse, les systèmes CRISPR-Cas ont été recherchés et étudiés sur l'ensemble des génomes de *C. novyi sensu lato* disponibles dans le but de déterminer la prévalence de ces systèmes dans le groupe d'espèces, d'explorer la diversité du contenu en spacers des souches et d'évaluer si ces régions pouvaient être utilisées pour le développement d'un outil de typage.

Les résultats de l'étude ont révélé la présence des systèmes CRISPR-Cas dans 95 % des souches de *C. novyi sensu lato* étudiées. Une étude similaire réalisée sur 241 souches de *C. botulinum* du groupe I et *C. sporogenes* avait montré une prévalence des systèmes de 83 % (Wentz et al., 2022). Ceci suggère une forte présence de ces régions chez *C. botulinum* en général malgré la forte diversité génétique présente entre les différents groupes. Pour les souches de *C. novyi sensu lato* possédant des locus CRISPR, le nombre de systèmes présent varie entre 1 et 3. Six types de systèmes différents sont retrouvés à la fois sur le chromosome et sur certains EGM. Les systèmes de classe 1 (I-B, I-D, III-B et III-D) sont les plus présents, un résultat déjà observé chez les clostridies de manière générale (Makarova et al., 2020). Pour les souches des lignées I, II et IV seul le système de type I-B est identifié. Pour les autres lignées, une plus grande diversité de systèmes est observée avec des types I-B, I-D, III-B et II-C pour la lignée Ib et I-B, I-D, III-D et V-U pour la lignée II. Les séquences des « repeats » sont également très variables (plus de 30 séquences différentes de « repeats » détectées dans les 58 génomes).

Le contenu en spacers et les EGM ciblés par les systèmes est très diversifié. Au total, 2045 spacers uniques ont été détectés. A titre de comparaison, pour l'espèce *M. tuberculosis*, seulement 50 spacers uniques sont présents pour plus de 1000 souches analysées (van Embden et al., 2000). Il n'y a qu'une petite proportion de spacers dont la cible a pu être identifiée (474 spacers) correspondant à 776 EGM uniques différents. La distribution entre phages et plasmides est proche du 50/50, comme observé dans l'étude faite avant la thèse (Negahdaripour et al., 2017). Cette proportion n'est pas retrouvée chez d'autres procaryotes ou la proportion de phages ciblés se trouve aux alentours de 85% (Long et al., 2019; McGinn & Marraffini, 2019). Le résultat obtenu lors de cette étude pourrait être expliqué par une mauvaise annotation de certains plasmides qui seraient en réalité des phages. Sur les 776 EGM uniques identifiés, seuls six d'entre eux représentent 20,2 % des résultats de BLAST. Il s'agit d'EGM retrouvés couramment chez *C. novyi sensu lato* (2 bactériophages contenant le gène *bont*, 2 hybrides plasmide-phage p4 et p6, le plasmide p2 et le prophage C\_Wou-2020a). Cette observation suggère que des transferts horizontaux arrivent fréquemment entre souches de *C. novyi sensu lato* pour ces 6 éléments, ce qui a déjà été établi de manière générale dans une autre étude (Skarin & Segerman, 2014). Concernant la présence des EGM dans les génomes possédant un spacer le ciblant, deux cas de figures existent. Pour le premier, le spacer ciblant l'EGM est uniquement retrouvé dans les génomes ne possédant pas l'EGM. Dans notre étude, c'est le cas avec le plasmide p2 et le prophage C\_Wou2020a. L'autre cas de figure est la présence de l'EGM et du spacer le ciblant dans le génome. Cela a été observé pour les EGM p1, p4 et p6. Ce résultat paradoxal a déjà été observé dans une autre étude sur plusieurs espèces bactériennes mais avec un auto-ciblage du chromosome bactérien et non d'un EGM (Wimmer & Beisel, 2020). Dans cette étude, il est expliqué que l'impact de cet auto-ciblage est mineur car compensé par réparation de l'ADN après le clivage de la zone ciblée (Wimmer & Beisel, 2020). Une autre hypothèse émise est la présence de ce spacer pour éviter une surinfection de la souche avec deux EGM de même catégorie. Cependant, ces études ont été réalisées uniquement *in silico*. Des expérimentations *in vitro* sont nécessaires pour mieux comprendre la co-habitation de ces spacers et de l'EGM. Par exemple, il serait intéressant de transférer un EGM de même catégorie dans la souche possédant déjà l'EGM et le spacer le ciblant pour voir si le nouvel EGM transféré est éliminé ou persiste après plusieurs sous-cultures.

La présence d'un PICI ou Phage Inducible Chromosomal Islands a été observée sur 12 génomes de la lignée II près d'un système CRISPR-Cas de type V-U. C'est la première fois que ce type d'EGM est retrouvé chez *C. botulinum*, alors qu'il est surtout retrouvé chez des bactéries à coloration Gram positif (Fillol-Salom et al., 2018). Dans le PICI détecté chez *C. novyi sensu lato*, un gène *cas14a* est associé à un locus CRISPR et à un système de modification restriction (3 types différents retrouvés

sur les 12 génomes), un autre système de défense bactérien ciblant l'ADN exogène. Des recherches supplémentaires sont nécessaires pour étudier cet élément PICI et son rôle dans l'écologie de la lignée II de *Clostridium novyi sensu lato*.

En ce qui concerne l'utilisation des systèmes CRISPR-cas comme cibles pour le développement d'outils de typage communs à l'ensemble des souches de *C. novyi sensu lato*, les observations laissent penser que ce n'est pas possible au vue de la diversité retrouvée au niveau des types, des répétitions de séquences et des spacers sur les 58 génomes. Cependant, pour les souches de la lignée Ia qui sont très proches génétiquement, les séquences de répétitions communes observées entre certaines souches peuvent potentiellement servir comme cible pour typer les souches de cette lignée. Des répétitions communes ont aussi été observées dans les lignées III et IV, même s'il faut garder à l'esprit que le nombre de génomes étudiés pour ces lignées est plus faible. Cet outil, s'il se révèle discriminant, pourrait servir, au même titre que la MLVA, à discriminer ces souches qui sont beaucoup retrouvées en Europe depuis maintenant plus d'une dizaine d'années (Anza et al., 2014a; Skarin et al., 2013; Woudstra et al., 2016).

Ces travaux de thèse ont permis de proposer une méthode de typage pour les souches de la lignée Ia de type C/D avec un protocole d'isolement des souches et une méthode MLVA. Ces outils pourront être utilisés en routine très rapidement par le LNR botulisme aviaire. Pour les souches des autres lignées, des travaux complémentaires sont nécessaires pour proposer des outils adaptés à chacune de ces lignées. Les perspectives de ces travaux de thèse sont de:

- Finaliser l'évaluation des systèmes PCR (cibles identifiées au cours de la thèse et ciblant différents gènes de flagelline) pour identifier la lignée de la ou des souches dans un échantillon terrain.
- Evaluer la possibilité de combiner le séquençage « short reads » et « long reads » pour obtenir un génome complet de *C. novyi sensu lato*.
- Suite à la détection d'au moins deux souches distinctes dans l'ensemble des échantillons lors de l'étude de cas :
  - Explorer la fréquence des co-infections par des souches des différentes lignées dans la collection du LNR.
  - Explorer le rôle de cette co-infection dans la pathogénèse du botulisme animal.
  - Explorer le rôle de cette co-infection dans la physiologie de cette bactérie et dans la stabilité du phage codant pour la toxine.
- Evaluer le potentiel du séquençage Shotgun pour permettre de tracer les souches impliquées dans le botulisme animal.

## Bibliographie

- Adey, A., Morrison, H. G., Asan, Xun, X., Kitzman, J. O., Turner, E. H., et al. (2010). Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. *Genome biology*, *11*(12), R119. doi:10.1186/gb-2010-11-12-r119
- Ågren, J., Sundström, A., Håfström, T., & Segerman, B. (2012). Gegenees: Fragmented alignment of multiple genomes for determining phylogenomic distances and genetic signatures unique for specified target groups. *PLOS ONE*, *7*(6), e39107. doi:10.1371/journal.pone.0039107
- Andersen, J. M., Shoup, M., Robinson, C., Britton, R., Olsen, K. E., & Barrangou, R. (2016). CRISPR Diversity and Microevolution in *Clostridium difficile*. *Genome Biol Evol*, *8*(9), 2841-2855. doi:10.1093/gbe/evw203
- Anhalt, J. P., & Fenselau, C. (1975). Identification of bacteria using mass spectrometry. *Analytical Chemistry*, *47*(2), 219-225. doi:10.1021/ac60352a007
- Anniballi, F., Fiore, A., Löfström, C., Skarin, H., Auricchio, B., Woudstra, C., et al. (2013). Management of animal botulism outbreaks: from clinical suspicion to practical countermeasures to prevent or minimize outbreaks. *Biosecur Bioterror*, *11 Suppl 1*, S191-199. doi:10.1089/bsp.2012.0089
- Anses. (2021). *Clostridium botulinum* : mise à jour des connaissances sur les différentes formes des types C, D, mosaïque C/D et D/C et E. (saisines 2019-SA-0112 à 2019-SA-0115). *Maisons-Alfort : Anses*, 170p.
- Anza, I., Skarin, H., Vidal, D., Lindberg, A., Baverud, V., & Mateo, R. (2014)a. The same clade of *Clostridium botulinum* strains is causing avian botulism in southern and northern Europe. *Anaerobe*, *26*, 20-23. doi:10.1016/j.anaerobe.2014.01.002
- Anza, I., Vidal, D., Feliu, J., Crespo, E., & Mateo, R. (2016). Differences in the vulnerability of waterbird species to botulism outbreaks in mediterranean wetlands: an assessment of ecological and physiological factors. *Appl Environ Microbiol*, *82*(10), 3092-3099. doi:10.1128/AEM.00119-16
- Anza, I., Vidal, D., & Mateo, R. (2014)b. New insight in the epidemiology of avian botulism outbreaks: necrophagous flies as vectors of *Clostridium botulinum* type C/D. *Environ Microbiol Rep*, *6*(6), 738-743. doi:10.1111/1758-2229.12197
- Appleyard, W. T., & Mollison, A. (1985). Suspected bovine botulism associated with broiler litter waste. *Vet Rec*, *116*(19), 522. doi:10.1136/vr.116.19.522
- Arnon, S. S., Schechter, R., Inglesby, T. V., Henderson, D. A., Bartlett, J. G., Ascher, M. S., et al. (2001). Botulinum toxin as a biological weapon: medical and public health management. *JAMA*, *285*(8), 1059-1070. doi:10.1001/jama.285.8.1059
- Aureli, P., Fenicia, L., Pasolini, B., Gianfranceschi, M., McCroskey, L. M., & Hatheway, C. L. (1986). Two cases of type E infant botulism caused by neurotoxicogenic *Clostridium butyricum* in Italy. *J Infect Dis*, *154*(2), 207-211. doi:10.1093/infdis/154.2.207
- Auricchio, B., Scalfaro, C., Vicenza, T., Bano, L., Bilei, S., Chemaly, M., et al. (2019). Multilocus Variable Number of Tandem Repeats Analysis (MLVA) as a tool for *Clostridium botulinum* group III subtyping. *Poster presentation Workshop botulism, Maisons-Alfort*.
- Bano, L., Drigo, I., Tonon, E., Agnoletti, F., Giovanardi, D., & Morandini, E. (2013). Rice hulls as a possible source of *Clostridium botulinum* type C spores for poultry. *Vet Rec*, *173*(17), 427-428. doi:10.1136/vr.f6521
- Bano, L., Drigo, I., Tonon, E., Berto, G., Tavella, A., Woudstra, C., et al. (2015). Evidence for a natural humoral response in dairy cattle affected by persistent botulism sustained



- by non-chimeric type C strains. *Anaerobe*, 36, 25-29.  
doi:10.1016/j.anaerobe.2015.09.007
- Bano, L., Drigo, I., Tonon, E., Pascoletti, S., Puiatti, C., Anniballi, F., et al. (2017). Identification and characterization of *Clostridium botulinum* group III field strains by matrix-assisted laser desorption-ionization time-of-flight mass spectrometry (MALDI-TOF MS). *Anaerobe*, 48, 126-134. doi:10.1016/j.anaerobe.2017.08.004
- Barash, J. R., & Arnon, S. S. (2014). A novel strain of *Clostridium botulinum* that produces type B and type H botulinum toxins. *J Infect Dis*, 209(2), 183-191.  
doi:10.1093/infdis/jit449
- Bellais, S., Nehlich, M., Ania, M., Duquenoy, A., Mazier, W., van den Engh, G., et al. (2022). Species-targeted sorting and cultivation of commensal bacteria from the gut microbiome using flow cytometry under anaerobic conditions. *Microbiome*, 10(1), 24.  
doi:10.1186/s40168-021-01206-7
- Benecke, R., Jost, W. H., Kanovsky, P., Ruzicka, E., Comes, G., & Grafe, S. (2005). A new botulinum toxin type A free of complexing proteins for treatment of cervical dystonia. *Neurology*, 64(11), 1949-1951. doi:10.1212/01.Wnl.0000163767.99354.C3
- Bengtson, I. A. (1922). Preliminary note on a toxin-producing anaerobe isolated from the larvae of *Lucilia caesar*. *Public Health Reports (1896-1970)*, 164-170.
- Binz, T., Sikorra, S., & Mahrhold, S. (2010). Clostridial neurotoxins: mechanism of SNARE cleavage and outlook on potential substrate specificity reengineering. *Toxins*, 2(4), 665-682. doi:10.3390/toxins2040665
- Blasi, J., Chapman, E. R., Link, E., Binz, T., Yamasaki, S., De Camilli, P., et al. (1993)a. Botulinum neurotoxin A selectively cleaves the synaptic protein SNAP-25. *Nature*, 365(6442), 160-163. doi:10.1038/365160a0
- Blasi, J., Chapman, E. R., Yamasaki, S., Binz, T., Niemann, H., & Jahn, R. (1993)b. Botulinum neurotoxin C1 blocks neurotransmitter release by means of cleaving HPC-1/syntaxin. *The EMBO journal*, 12(12), 4821-4828.
- Blomqvist, G., Skarin, H., Lindberg, A., Båverud, V., & Engstrom, B. (2009). Surveillance for *Clostridium botulinum* type C or C/D in Swedish broilers. *Proceedings of the 16th World Veterinary Poultry Congress 415 Marrakesh, Morocco*.
- Brauge, T., Faille, C., Inglebert, G., Dubois, T., Morieux, P., Slomianny, C., et al. (2018). Comparative evaluation of DNA extraction methods for amplification by qPCR of superficial vs intracellular DNA from *Bacillus* spores. *Int J Food Microbiol*, 266, 289-294. doi:10.1016/j.ijfoodmicro.2017.12.012
- Braun, U. (2006). Botulism in cattle. *Schweiz Arch Tierheilkd*, 148(7), 331-339.  
doi:10.1024/0036-7281.148.7.331
- Broussard, L. A. (2001). Biological agents: weapons of warfare and bioterrorism. *Mol Diagn*, 6(4), 323-333. doi:10.1054/modi.2001.29155
- Burke, G. S. (1919). The Occurrence of *Bacillus botulinus* in Nature. *Journal of Bacteriology*, 4(5), 541-553. doi:10.1128/jb.4.5.541-553.1919
- Buytaers, F. E., Saltykova, A., Denayer, S., Verhaegen, B., Vanneste, K., Roosens, N. H. C., et al. (2020). A Practical Method to Implement Strain-Level Metagenomics-Based Foodborne Outbreak Investigation and Source Tracking in Routine. *Microorganisms*, 8(8). doi:10.3390/microorganisms8081191
- Buytaers, F. E., Saltykova, A., Mattheus, W., Verhaegen, B., Roosens, N. H. C., Vanneste, K., et al. (2021). Application of a strain-level shotgun metagenomics approach on food samples: resolution of the source of a *Salmonella* food-borne outbreak. *Microb Genom*, 7(4). doi:10.1099/mgen.0.000547
- Cabeen, M. T., & Jacobs-Wagner, C. (2005). Bacterial cell shape. *Nature Reviews Microbiology*, 3(8), 601-610. doi:10.1038/nrmicro1205

- Carruthers, A., Kiene, K., & Carruthers, J. (1996). Botulinum A exotoxin use in clinical dermatology. *J Am Acad Dermatol*, 34(5 Pt 1), 788-797. doi:10.1016/s0190-9622(96)90016-x
- Cato, E. P., Georges, W. L., & Finegold, S. M. (1986). Genus *Clostridium*. In: *Sneath PHA, Mair NS, Sharpe ME, et al. eds. Bergey's Manual of Systematic Bacteriology, Vol. 2. Baltimore: Williams & Wilkins*, 1141-1200.
- Cenciarelli, O., Riley, P. W., & Baka, A. (2019). Biosecurity threat posed by botulinum toxin. *Toxins*, 11(12), 681. doi:10.3390/toxins11120681
- Chaudhuri, P., Singh, S., & Qari, S. (1996). DNase of *Clostridium septicum* : An obstacle in plasmid isolation. *Indian Journal of Animal Sciences*, 66(7), 653-656.
- Chen, S., Li, J., Saleh-Lakha, S., Allen, V., & Odumeru, J. (2011). Multiple-locus variable number of tandem repeat analysis (MLVA) of *Listeria monocytogenes* directly in food samples. *Int J Food Microbiol*, 148(1), 8-14. doi:10.1016/j.ijfoodmicro.2011.04.014
- Chipault, J. G., White, C. L., Blehert, D. S., Jennings, S. K., & Strom, S. M. (2015). Avian botulism type E in waterbirds of Lake Michigan, 2010–2013. *Journal of Great Lakes Research*, 41(2), 659-664. doi:10.1016/j.jglr.2015.03.021
- Chun, C. L., Ochsner, U., Byappanahalli, M. N., Whitman, R. L., Tepp, W. H., Lin, G., et al. (2013). Association of toxin-producing *Clostridium botulinum* with the macroalga *Cladophora* in the Great Lakes. *Environ Sci Technol*, 47(6), 2587-2594. doi:10.1021/es304743m
- Collins, M. D., Lawson, P. A., Willems, A., Cordoba, J. J., Fernandez-Garayzabal, J., Garcia, P., et al. (1994). The phylogeny of the genus *Clostridium* : proposal of five new genera and eleven new species combinations. *Int J Syst Bacteriol*, 44(4), 812-826. doi:10.1099/00207713-44-4-812
- Connan, C., Denève, C., Mazuet, C., & Popoff, M. R. (2013). Regulation of toxin synthesis in *Clostridium botulinum* and *Clostridium tetani*. *Toxicon*, 75, 90-100.
- Cruz-Morales, P., Orellana, C. A., Moutafis, G., Moonen, G., Rincon, G., Nielsen, L. K., et al. (2019). Revisiting the Evolution and Taxonomy of Clostridia, a Phylogenomic Update. *Genome Biol Evol*, 11(7), 2035-2044. doi:10.1093/gbe/evz096
- Dahlenborg, M., Borch, E., & Rådström, P. (2003). Prevalence of *Clostridium botulinum* types B, E and F in faecal samples from Swedish cattle. *Int J Food Microbiol*, 82(2), 105-110. doi:10.1016/s0168-1605(02)00255-6
- Diao, M. M., André, S., & Membré, J. M. (2014). Meta-analysis of D-values of proteolytic *Clostridium botulinum* and its surrogate strain *Clostridium sporogenes* PA 3679. *Int J Food Microbiol*, 174, 23-30. doi:10.1016/j.ijfoodmicro.2013.12.029
- Dohms, J. E. (2008). Botulism. *Diseases of poultry* (eds. Y.M. Saif, A.M. Fadly, J.R. Glisson, L.R. McDougald, L.K. Nolan and D.E. Swayne) 12th edn. Blackwell Publishing Ames, 879-885.
- Doutre, M. (1969). Fréquence au Sénégal du botulisme animal d'origine hydrique. *Revue d'élevage et de médecine vétérinaire des pays tropicaux*, 22(1), 29-31. doi:doi: 10.19182/remvt.7659
- Dover, N., Barash, J. R., Bell, J. M., Sylvester, M. D., & Arnon, S. S. (2021). Closed Genome Sequence of *Clostridium botulinum* Strain IBCA10-7060 Type Bh. *Microbiology resource announcements*, 10(34), e0038321. doi:10.1128/mra.00383-21
- Dover, N., Barash, J. R., Hill, K. K., Davenport, K. W., Teshima, H., Xie, G., et al. (2013). *Clostridium botulinum* strain Af84 contains three neurotoxin gene clusters: bont/A2, bont/F4 and bont/F5. *PLOS ONE*, 8(4), e61205. doi:10.1371/journal.pone.0061205
- Dykes, J. K., Lúquez, C., Raphael, B. H., McCroskey, L., & Maslanka, S. E. (2015). Laboratory investigation of the first case of botulism caused by *Clostridium butyricum*

- type E toxin in the United States. *J Clin Microbiol*, 53(10), 3363-3365.  
doi:10.1128/JCM.01351-15
- Eklund, M. W., & Poysky, F. T. (1974). Interconversion of type C and D strains of *Clostridium botulinum* by specific bacteriophages. *Applied microbiology*, 27(1), 251-258. doi:10.1128/am.27.1.251-258.1974
- Eklund, M. W., Poysky, F. T., Meyers, J. A., & Pelroy, G. A. (1974). Interspecies conversion of *Clostridium botulinum* type C to *Clostridium novyi* type A by bacteriophage. *Science*, 186(4162), 456-458. doi:10.1126/science.186.4162.456
- Eklund, M. W., Poysky, F. T., & Reed, S. M. (1972). Bacteriophage and the toxigenicity of *Clostridium botulinum* type D. *Nature New Biology*, 235(53), 16-17.  
doi:10.1038/newbio235016a0
- Eklund, M. W., Poysky, F. T., Reed, S. M., & Smith, C. A. (1971). Bacteriophage and the toxigenicity of *Clostridium botulinum* type C. *Science*, 172(3982), 480-482.  
doi:doi:10.1126/science.172.3982.480
- Erbguth, F. J. (2008). From poison to remedy: the chequered history of botulinum toxin. *J Neural Transm (Vienna)*, 115(4), 559-565. doi:10.1007/s00702-007-0728-2
- Espelund, M., & Klaveness, D. (2014). Botulism outbreaks in natural environments - an update. *Frontiers in microbiology*, 5, 287-287. doi:10.3389/fmicb.2014.00287
- Evelsizer, D. D., Clark, R. G., & Bollinger, T. K. (2010). Relationships between local carcass density and risk of mortality in molting mallards during avian botulism outbreaks. *J Wildl Dis*, 46(2), 507-513. doi:10.7589/0090-3558-46.2.507
- Fabre, L., Zhang, J., Guigon, G., Le Hello, S., Guibert, V., Accou-Demartin, M., et al. (2012). CRISPR Typing and Subtyping for Improved Laboratory Surveillance of *Salmonella* Infections. *PLOS ONE*, 7(5), e36995. doi:10.1371/journal.pone.0036995
- Fillo, S., Giordani, F., Anniballi, F., Gorgé, O., Ramiisse, V., Vergnaud, G., et al. (2011). *Clostridium botulinum* group I strain genotyping by 15-locus multilocus variable-number tandem-repeat analysis. *J Clin Microbiol*, 49(12), 4252-4263.  
doi:10.1128/jcm.05396-11
- Fillo, S., Giordani, F., Tonon, E., Drigo, I., Anselmo, A., Fortunato, A., et al. (2021). Extensive Genome Exploration of *Clostridium botulinum* Group III Field Strains. *Microorganisms*, 9(11), 2347. doi:10.3390/microorganisms9112347
- Fillol-Salom, A., Martínez-Rubio, R., Abdulrahman, R. F., Chen, J., Davies, R., & Penadés, J. R. (2018). Phage-inducible chromosomal islands are ubiquitous within the bacterial universe. *Isme j*, 12(9), 2114-2128. doi:10.1038/s41396-018-0156-3
- Fohler, S., Discher, S., Jordan, E., Seyboldt, C., Klein, G., Neubauer, H., et al. (2016). Detection of *Clostridium botulinum* neurotoxin genes (A–F) in dairy farms from Northern Germany using PCR: A case-control study. *Anaerobe*, 39, 97-104.  
doi:https://doi.org/10.1016/j.anaerobe.2016.03.008
- Franciosa, G., Fenicia, L., Caldiani, C., & Aureli, P. (1996). PCR for detection of *Clostridium botulinum* type C in avian and environmental samples. *J Clin Microbiol*, 34(4), 882-885. doi:10.1128/jcm.34.4.882-885.1996
- Fricke, W. F., Mammel Mark, K., McDermott Patrick, F., Tartera, C., White David, G., LeClerc, J. E., et al. (2011). Comparative genomics of 28 *Salmonella enterica* Isolates: Evidence for CRISPR-mediated adaptive sublineage evolution. *Journal of Bacteriology*, 193(14), 3556-3568. doi:10.1128/JB.00297-11
- Friend, M., Franson, J. C., & Ciganovich, E. A. (1999). *Field manual of wildlife diseases: general field procedures and diseases of birds*: US Geological Survey.
- Frothingham, R., & Meeker-O'Connell, W. A. (1998). Genetic diversity in the *Mycobacterium tuberculosis* complex based on variable numbers of tandem DNA repeats.

- Microbiology (Reading, England)*, 144 ( Pt 5), 1189-1196. doi:10.1099/00221287-144-5-1189
- Galey, F. D., Terra, R., Walker, R., Adaska, J., Etchebarne, M. A., Puschner, B., et al. (2000). Type C botulism in dairy cattle from feed contaminated with a dead cat. *J Vet Diagn Invest*, 12(3), 204-209. doi:10.1177/104063870001200302
- Gauvry, E., Mathot, A.-G., Leguérinel, I., Couvert, O., Postollec, F., Broussolle, V., et al. (2017). Knowledge of the physiology of spore-forming bacteria can explain the origin of spores in the food environment. *Research in microbiology*, 168(4), 369-378.
- Geiges, M. L. (2002). The history of botulism. *Curr Probl Dermatol*, 30, 77-93. doi:10.1159/000060681
- Gimenez, D., & Ciccarelli, A. (1970). Another type of *Clostridium botulinum*. *Zentralblatt für Bakteriologie, Parasitenkunde, Infektionskrankheiten und Hygiene, Abt. I (Originale)*, 215(2), 221-224.
- Giménez, D. F. (1984). *Clostridium botulinum* subtype Ba. *Zentralbl Bakteriolog Mikrobiol Hyg A*, 257(1), 68-72.
- Gladney, L., Halpin, J. L., & Lúquez, C. (2021). Genomic Characterization of Strains From a Cluster of Infant Botulism Type A in a Small Town in Colorado, United States. *Frontiers in microbiology*, 12, 688240-688240. doi:10.3389/fmicb.2021.688240
- Gonzalez-Escalona, N., & Sharma, S. K. (2020). Closing *Clostridium botulinum* Group I Genomes Using a Combination of Short- and Long-Reads. *Frontiers in microbiology*, 11, 239. doi:10.3389/fmicb.2020.00239
- Gonzalez-Escalona, N., Thirunavukkarasu, N., Singh, A., Toro, M., Brown, E. W., Zink, D., et al. (2014). Draft genome sequence of bivalent *Clostridium botulinum* strain IBCA10-7060, encoding botulinum neurotoxin B and a new FA mosaic type. *Genome Announc*, 2(6). doi:10.1128/genomeA.01275-14
- Gonzalez-Escalona, N., Timme, R., Raphael, B. H., Zink, D., & Sharma, S. K. (2014). Whole-genome single-nucleotide-polymorphism analysis for discrimination of *Clostridium botulinum* group I strains. *Appl Environ Microbiol*, 80(7), 2125-2132. doi:10.1128/AEM.03934-13
- Grenda, T., Grenda, A., Krawczyk, P., & Kwiatek, K. (2021). Botulinum toxin in cancer therapy—current perspectives and limitations. *Applied Microbiology and Biotechnology*. doi:10.1007/s00253-021-11741-w
- Grosse-Herrenthey, A., Maier, T., Gessler, F., Schaumann, R., Böhnelt, H., Kostrzewa, M., et al. (2008). Challenging the problem of clostridial identification with matrix-assisted laser desorption and ionization-time-of-flight mass spectrometry (MALDI-TOF MS). *Anaerobe*, 14(4), 242-249. doi:10.1016/j.anaerobe.2008.06.002
- Grundmann, H., Hori, S., & Tanner, G. (2001). Determining confidence intervals when measuring genetic diversity and the discriminatory abilities of typing methods for microorganisms. *J Clin Microbiol*, 39(11), 4190-4192. doi:doi:10.1128/JCM.39.11.4190-4192.2001
- Gu, S., Rumpel, S., Zhou, J., Strotmeier, J., Bigalke, H., Perry, K., et al. (2012). Botulinum neurotoxin is shielded by NTNHA in an interlocked complex. *Science*, 335(6071), 977-981. doi:10.1126/science.1214270
- Gunnison, J. B., Cummings, J. R., & Meyer, K. F. (1936). *Clostridium Botulinum* Type E. *Experimental Biology and Medicine*, 35(2), 278-280. doi:10.3181/00379727-35-8938p
- Hall, J. D., McCroskey, L. M., Pincomb, B. J., & Hatheway, C. L. (1985). Isolation of an organism resembling *Clostridium barati* which produces type F botulinum toxin from an infant with botulism. *J Clin Microbiol*, 21(4), 654-655. doi:10.1128/jcm.21.4.654-655.1985

- Hannett, G. E., Stone, W. B., Davis, S. W., & Wroblewski, D. (2011). Biodiversity of *Clostridium botulinum* type E associated with a large outbreak of botulism in wildlife from Lake Erie and Lake Ontario. *Appl Environ Microbiol*, *77*(3), 1061-1068. doi:10.1128/aem.01578-10
- Hardy, S. P., & Kaldhusdal, M. (2013). Type C and C/D toxigenic *Clostridium botulinum* is not normally present in the intestine of healthy broilers. *Vet Microbiol*, *165*(3-4), 466-468. doi:10.1016/j.vetmic.2013.03.022
- Heider, L. C., McClure, J. T., & Leger, E. R. (2001). Presumptive diagnosis of *Clostridium botulinum* type D intoxication in a herd of feedlot cattle. *The Canadian veterinary journal = La revue veterinaire canadienne*, *42*(3), 210-212.
- Hielm, S., Björkroth, J., Hyytiä, E., & Korkeala, H. (1998). Genomic analysis of *Clostridium botulinum* group II by pulsed-field gel electrophoresis. *Appl Environ Microbiol*, *64*(2), 703-708. doi:10.1128/AEM.64.2.703-708.1998
- Hill, K. K., & Smith, T. J. (2013). Genetic diversity within *Clostridium botulinum* serotypes, botulinum neurotoxin gene clusters and toxin subtypes. *Curr Top Microbiol Immunol*, *364*, 1-20. doi:10.1007/978-3-642-33570-9\_1
- Hill, K. K., Smith, T. J., Helma, C. H., Ticknor, L. O., Foley, B. T., Svensson, R. T., et al. (2007). Genetic diversity among botulinum neurotoxin-producing clostridial strains. *Journal of Bacteriology*, *189*(3), 818-832. doi:10.1128/JB.01180-06
- Hill, K. K., Xie, G., Foley, B. T., & Smith, T. J. (2015). Genetic diversity within the botulinum neurotoxin-producing bacteria and their neurotoxins. *Toxicon*, *107*, 2-8. doi:https://doi.org/10.1016/j.toxicon.2015.09.011
- Hogg, R., Livesey, C., & Payne, J. (2008). Diagnosis and implications of botulism. *In practice*, *30*(7), 392-397.
- Hogg, R. A., White, V. J., & Smith, G. R. (1990). Suspected botulism in cattle associated with poultry litter. *Vet Rec*, *126*(19), 476-479.
- Holmes, P. (2019). Avian botulism - a recurring paralytic disease of wild UK waterbirds. *Vet Rec*, *185*(9), 261-262. doi:10.1136/vr.15417
- Hutson, R. A., Zhou, Y., Collins, M. D., Johnson, E. A., Hatheway, C. L., & Sugiyama, H. (1996). Genetic characterization of *Clostridium botulinum* type A containing silent type B neurotoxin gene sequences. *Journal of Biological Chemistry*, *271*(18), 10786-10792. doi:https://doi.org/10.1074/jbc.271.18.10786
- Hyun, S. H., & Sakaguchi, G. (1989). Implication of coprophagy in pathogenesis of chicken botulism. *Nihon Juigaku Zasshi*, *51*(3), 582-586. doi:10.1292/jvms1939.51.582
- Imada, K. (2018). Bacterial flagellar axial structure and its construction. *Biophysical reviews*, *10*(2), 559-570. doi:10.1007/s12551-017-0378-z
- Jacobson, M. J., Lin, G., Whittam, T. S., & Johnson, E. A. (2008). Phylogenetic analysis of *Clostridium botulinum* type A by multi-locus sequence typing. *Microbiology (Reading, England)*, *154*(Pt 8), 2408-2415. doi:10.1099/mic.0.2008/016915-0
- Jansen, R., Embden, J. D. A. v., Gaastra, W., & Schouls, L. M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology*, *43*(6), 1565-1575. doi:https://doi.org/10.1046/j.1365-2958.2002.02839.x
- Jean, D., Fecteau, G., Scott, D., Higgins, R., & Quessy, S. (1995). *Clostridium botulinum* type C intoxication in feedlot steers being fed ensiled poultry litter. *The Canadian veterinary journal = La revue veterinaire canadienne*, *36*(10), 626-628.
- Kakinuma, H., Maruyama, H., Takahashi, H., Yamakawa, K., & Nakamura, S. (1996). The first case of type B infant botulism in Japan. *Acta Paediatr Jpn*, *38*(5), 541-543. doi:10.1111/j.1442-200x.1996.tb03542.x

- Kalb, S. R., Baudys, J., Smith, T. J., Smith, L. A., & Barr, J. R. (2017). Characterization of hemagglutinin negative botulinum progenitor toxins. *Toxins*, 9(6). doi:10.3390/toxins9060193
- Kalmbach, E. R. (1939). American vultures and the toxin of *Clostridium botulinum*. *Journal of the American Veterinary Medical Association*, 94, 187-191.
- Kamerbeek, J., Schouls, L., Kolk, A., van Agterveld, M., van Soolingen, D., Kuijper, S., et al. (1997). Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol*, 35(4), 907-914. doi:10.1128/jcm.35.4.907-914.1997
- Karas, M., & Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem*, 60(20), 2299-2301. doi:10.1021/ac00171a028
- Keim, P., Price, L. B., Klevytska, A. M., Smith, K. L., Schupp, J. M., Okinaka, R., et al. (2000). Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. *J Bacteriol*, 182(10), 2928-2936. doi:10.1128/jb.182.10.2928-2936.2000
- Kelch, W. J., Kerr, L. A., Pringle, J. K., Rohrbach, B. W., & Whitlock, R. H. (2000). Fatal *Clostridium botulinum* toxicosis in eleven Holstein cattle fed round bale barley haylage. *J Vet Diagn Invest*, 12(5), 453-455. doi:10.1177/104063870001200511
- Kerner, J. (1820). *Neue Beobachtungen über die in Württemberg so häufig vorkommenden tödtlichen Vergiftungen durch den Genuß geräucherter Würste: Osiander*.
- Kerner, J. (1822). *Das Fettgift oder die Fettsäure und ihre Wirkungen auf den thierischen Organismus: Cotta*.
- Keto-Timonen, R., Nevas, M., & Korkeala, H. (2005). Efficient DNA fingerprinting of *Clostridium botulinum* types A, B, E, and F by amplified fragment length polymorphism analysis. *Appl Environ Microbiol*, 71(3), 1148-1154. doi:10.1128/AEM.71.3.1148-1154.2005
- King, L. A. (2008). Two severe cases of botulism associated with industrially produced chicken enchiladas, France, August 2008. *Eurosurveillance*, 13(37), 18978. doi:https://doi.org/10.2807/ese.13.37.18978-en
- Knüpfer, M., Braun, P., Baumann, K., Rehn, A., Antwerpen, M., Grass, G., et al. (2020). Evaluation of a highly efficient DNA extraction method for *Bacillus anthracis* endospores. *Microorganisms*, 8(5), 763. doi:10.3390/microorganisms8050763
- Kovanen, S. M., Kivistö, R. I., Rossi, M., & Hänninen, M. L. (2014). A combination of MLST and CRISPR typing reveals dominant *Campylobacter jejuni* types in organically farmed laying hens. *J Appl Microbiol*, 117(1), 249-257. doi:10.1111/jam.12503
- Lagneaux, A. S., Hénard, S., Diancourt, L., Stein, E., Perez, P., Mathieu, P., et al. (2021). *Clostridium haemolyticum* Infection: A Cause of Hemolytic Anemia in a Patient with Bone Marrow Necrosis. *Microorganisms*, 9(8), 1568. doi:10.3390/microorganisms9081568
- Lamanna, C., Eklund, H. W., & Mc, E. O. (1946). Botulinum toxin (type A); including a study of shaking with chloroform as a step in the isolation procedure. *J Bacteriol*, 52, 1-13. doi:10.1128/jb.52.1.1-13.1946
- Lan Chun, C., Kahn, C. I., Borchert, A. J., Byappanahalli, M. N., Whitman, R. L., Peller, J., et al. (2015). Prevalence of toxin-producing *Clostridium botulinum* associated with the macroalga *Cladophora* in three Great Lakes: growth and management. *Sci Total Environ*, 511, 523-529. doi:10.1016/j.scitotenv.2014.12.080

- Lawson, P. A., Citron, D. M., Tyrrell, K. L., & Finegold, S. M. (2016). Reclassification of *Clostridium difficile* as *Clostridioides difficile*. *Anaerobe*, 40, 95-99. doi:<https://doi.org/10.1016/j.anaerobe.2016.06.008>
- Le Bouquin, S., Le Maréchal, C., Ballan, V., Rouxel, S., Leon, D., Balaine, L., et al. (2017). Investigation d'un cas de botulisme aviaire dans un élevage de poules pondeuses plein-air. *Bulletin épidémiologique, santé animale et alimentation*, 78, 13-17.
- Le Gratiet, T., Poezevara, T., Rouxel, S., Houard, E., Mazuet, C., Chemaly, M., et al. (2020). Development of an innovative and quick method for the isolation of *Clostridium botulinum* strains involved in avian botulism outbreaks. *Toxins*, 12(1), 42.
- Le Marechal, C., Fourour, S., Ballan, V., Rouxel, S., Souillard, R., & Chemaly, M. (2018). Detection of *Clostridium botulinum* group III in environmental samples from farms by real-time PCR using four commercial DNA extraction kits. *BMC Res Notes*, 11(1), 441. doi:10.1186/s13104-018-3549-5
- Le Marechal, C., Hulin, O., Mace, S., Chuzeville, C., Rouxel, S., Poezevara, T., et al. (2019). A case report of a botulism outbreak in beef cattle due to the contamination of wheat by a roaming cat carcass: from the suspicion to the management of the outbreak. *Animals : an open access journal from MDPI*, 9(12), 1025. doi:10.3390/ani9121025
- Le Maréchal, C., Poëzevara, T., Rouxel, S., Ballan, V., Houard, E., Souillard, R., et al. (2019). Le botulisme aviaire en France : Bilan des épisodes diagnostiqués par le LNR Botulisme Aviaire depuis 2013. *Treizièmes Journées de la Recherche Avicole et Palmipèdes à Foie Gras*, 621-625.
- Le Maréchal, C., Woudstra, C., & Fach, P. (2016). Botulism. In *Clostridial Diseases of Animals* (pp. 303-330).
- Le Maréchal, C., Woudstra, C., & Souillard, R. (2012). Botulisme aviaire : un test de détection PCR en temps réel. *Filières avicoles*, 757, 64-65.
- Leclair, D., Pagotto, F., Farber, J. M., Cadieux, B., & Austin, J. W. (2006). Comparison of DNA fingerprinting methods for use in investigation of type E botulism outbreaks in the Canadian Arctic. *J Clin Microbiol*, 44(5), 1635-1644. doi:10.1128/JCM.44.5.1635-1644.2006
- Lee, K., Gu, S., Jin, L., Le, T. T. N., Cheng, L. W., Strotmeier, J., et al. (2013). Structure of a bimodular botulinum neurotoxin complex provides insights into its oral toxicity. *PLoS pathogens*, 9(10), e1003690-e1003690. doi:10.1371/journal.ppat.1003690
- Lew, M. F., Brashear, A., & Factor, S. (2000). The safety and efficacy of botulinum toxin type B in the treatment of patients with cervical dystonia: summary of three controlled clinical trials. *Neurology*, 55(12 Suppl 5), S29-35.
- Lim, C. Y., Granger, J. H., & Porter, M. D. (2019). SERS detection of *Clostridium botulinum* neurotoxin serotypes A and B in buffer and serum: Towards the development of a biodefense test platform. *Analytica Chimica Acta: X*, 1, 100002. doi:<https://doi.org/10.1016/j.acax.2018.100002>
- Lin, W.-J., & Johnson, E. A. (1995). Genome analysis of *Clostridium botulinum* type A by pulsed-field gel electrophoresis. *Appl Environ Microbiol*, 61(12), 4441-4447.
- Lin, W. J., & Johnson, E. A. (1995). Genome analysis of *Clostridium botulinum* type A by pulsed-field gel electrophoresis. *Appl Environ Microbiol*, 61(12), 4441-4447. doi:10.1128/aem.61.12.4441-4447.1995
- Lindström, M., Kiviniemi, K., & Korkeala, H. (2006). Hazard and control of group II (non-proteolytic) *Clostridium botulinum* in modern food processing. *Int J Food Microbiol*, 108(1), 92-104. doi:<https://doi.org/10.1016/j.ijfoodmicro.2005.11.003>
- Locke, L. N., & Friend, M. (1989). *Avian botulism: geographic expansion of a historic disease* (13.2.4). Retrieved from Washington, DC: <http://pubs.er.usgs.gov/publication/70006740>

- Long, J., Xu, Y., Ou, L., Yang, H., Xi, Y., Chen, S., et al. (2019). Diversity of CRISPR/Cas system in *Clostridium perfringens*. *Mol Genet Genomics*, 294(5), 1263-1275. doi:10.1007/s00438-019-01579-3
- Macdonald, T. E., Helma, C. H., Shou, Y., Valdez, Y. E., Ticknor, L. O., Foley, B. T., et al. (2011). Analysis of *Clostridium botulinum* serotype E strains by using multilocus sequence typing, amplified fragment length polymorphism, variable-number tandem-repeat analysis, and botulinum neurotoxin gene sequencing. *Appl Environ Microbiol*, 77(24), 8625-8634. doi:10.1128/AEM.05155-11
- Macdonald, T. E., Helma, C. H., Ticknor, L. O., Jackson, P. J., Okinaka, R. T., Smith, L. A., et al. (2008). Differentiation of *Clostridium botulinum* serotype A strains by multiple-locus variable-number tandem-repeat analysis. *Appl Environ Microbiol*, 74(3), 875-882. doi:10.1128/AEM.01539-07
- Macnab, R. M. (2004). Type III flagellar protein export and flagellar assembly. *Biochim Biophys Acta*, 1694(1-3), 207-217. doi:10.1016/j.bbamcr.2004.04.005
- Makarova, K. S., Wolf, Y. I., Iranzo, J., Shmakov, S. A., Alkhnbashi, O. S., Brouns, S. J. J., et al. (2020). Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat Rev Microbiol*, 18(2), 67-83. doi:10.1038/s41579-019-0299-x
- Mansfield, M. J., Adams, J. B., & Doxey, A. C. (2015). Botulinum neurotoxin homologs in non-*Clostridium* species. *FEBS Lett*, 589(3), 342-348. doi:10.1016/j.febslet.2014.12.018
- Maslanka, S. E., Lúquez, C., Dykes, J. K., Tepp, W. H., Pier, C. L., Pellett, S., et al. (2016). A Novel Botulinum Neurotoxin, Previously Reported as Serotype H, Has a Hybrid-Like Structure With Regions of Similarity to the Structures of Serotypes A and F and Is Neutralized With Serotype A Antitoxin. *J Infect Dis*, 213(3), 379-385. doi:10.1093/infdis/jiv327
- Mazuet, C., Jourdan Da Silva, N., Sautereau, J., & Popoff, M. R. (2018). Le botulisme humain en France, 2013-2016. *Bulletin épidémiologique hebdomadaire 2018, Santé publique France*, 3, 46-54.
- McGinn, J., & Marraffini, L. A. (2019). Molecular mechanisms of CRISPR-Cas spacer acquisition. *Nat Rev Microbiol*, 17(1), 7-12. doi:10.1038/s41579-018-0071-7
- McLauchlin, J., Salmon, J. E., Ahmed, S., Brazier, J. S., Brett, M. M., George, R. C., et al. (2002). Amplified fragment length polymorphism (AFLP) analysis of *Clostridium novyi*, *C. perfringens* and *Bacillus cereus* isolated from injecting drug users during 2000. *J Med Microbiol*, 51(11), 990-1000. doi:10.1099/0022-1317-51-11-990
- Meloni, E., Le Maréchal, C., Millot, F., Payne, A., Mazuet, C., Guillemain, M., et al. (2021). Estimation of *Clostridium botulinum* prevalence in wild waterfowl in France. *Poster for the 35th Congress of the International Union of Game Biologists, Budapest, Hungary*.
- Meyer, K., & Gunnison, J. (1929). South African Cultures of *Clostridium botulinum* and *parabotulinum*. XXXVII: With a Description of *Cl. botulinum* Type D, n. sp. *J Infect Dis*, 106-118.
- Minamino, T., & Imada, K. (2015). The bacterial flagellar motor and its structural diversity. *Trends in Microbiology*, 23(5), 267-274. doi:https://doi.org/10.1016/j.tim.2014.12.011
- Miyazaki, S., & Sakaguchi, G. (1978). Experimental botulism in chickens: the cecum as the site of production and absorption of botulinum toxin. *Jpn J Med Sci Biol*, 31(1), 1-15. doi:10.7883/yoken1952.31.1
- Moller, V., & Scheibel, I. (1960). Preliminary report on the isolation of an apparently new type of *Cl. botulinum*. *Acta Pathol Microbiol Scand*, 48, 80. doi:10.1111/j.1699-0463.1960.tb04741.x



- Montecucco, C., & Schiavo, G. (1995). Structure and function of tetanus and botulinum neurotoxins. *Q Rev Biophys*, 28(4), 423-472. doi:10.1017/s0033583500003292
- Moriishi, K., Koura, M., Abe, N., Fujii, N., Fujinaga, Y., Inoue, K., et al. (1996b). Mosaic structures of neurotoxins produced from *Clostridium botulinum* types C and D organisms. *Biochim Biophys Acta*, 1307(2), 123-126. doi:10.1016/0167-4781(96)00006-1
- Moriishi, K., Koura, M., Fujii, N., Fujinaga, Y., Inoue, K., Syuto, B., et al. (1996a). Molecular cloning of the gene encoding the mosaic neurotoxin, composed of parts of botulinum neurotoxin types C1 and D, and PCR detection of this gene from *Clostridium botulinum* type C organisms. *Appl Environ Microbiol*, 62(2), 662-667. doi:10.1128/aem.62.2.662-667.1996
- Myllykoski, J., Lindström, M., Keto-Timonen, R., Söderholm, H., Jakala, J., Kallio, H., et al. (2009). Type C bovine botulism outbreak due to carcass contaminated non-acidified silage. *Epidemiol Infect*, 137(2), 284-293. doi:10.1017/s0950268808000939
- Naik, P. P. (2021). Utilities of Botulinum Toxins in Dermatology and Cosmetology. *Clinical, cosmetic and investigational dermatology*, 14, 1319-1330. doi:10.2147/CCID.S332247
- Nakamura, K., Kohda, T., Shibata, Y., Tsukamoto, K., Arimitsu, H., Hayashi, M., et al. (2012). Unique biological activity of botulinum D/C mosaic neurotoxin in murine species. *Infection and immunity*, 80(8), 2886-2893. doi:10.1128/IAI.00302-12
- Nakamura, K., Kohda, T., Umeda, K., Yamamoto, H., Mukamoto, M., & Kozaki, S. (2010). Characterization of the D/C mosaic neurotoxin produced by *Clostridium botulinum* associated with bovine botulism in Japan. *Vet Microbiol*, 140(1-2), 147-154. doi:10.1016/j.vetmic.2009.07.023
- Nakamura, S., Kimura, I., Yamakawa, K., & Nishida, S. (1983). Taxonomic relationships among *Clostridium novyi* Types A and B, *Clostridium haemolyticum* and *Clostridium botulinum* type C. *J Gen Microbiol*, 129(5), 1473-1479. doi:10.1099/00221287-129-5-1473
- Nakamura, S., & Minamino, T. (2019). Flagella-Driven Motility of Bacteria. *Biomolecules*, 9(7), 279. doi:10.3390/biom9070279
- Negahdaripour, M., Nezafat, N., Hajjigharamani, N., Rahmatabadi, S. S., & Ghasemi, Y. (2017). Investigating CRISPR-Cas systems in *Clostridium botulinum* via bioinformatics tools. *Infect Genet Evol*, 54, 355-373. doi:10.1016/j.meegid.2017.06.027
- Neill, S. D., McLoughlin, M. F., & McIlroy, S. G. (1989). Type C botulism in cattle being fed ensiled poultry litter. *Vet Rec*, 124(21), 558-560. doi:10.1136/vr.124.21.558
- Oguma, K., Yamaguchi, T., Sudou, K., Yokosawa, N., & Fujikawa, Y. (1986). Biochemical classification of *Clostridium botulinum* type C and D strains and their nontoxigenic derivatives. *Appl Environ Microbiol*, 51(2), 256-260. doi:10.1128/aem.51.2.256-260.1986
- Ondov, B. D., Treangen, T. J., Melsted, P., Mallonee, A. B., Bergman, N. H., Koren, S., et al. (2016). Mash: fast genome and metagenome distance estimation using MinHash. *Genome biology*, 17(1), 132. doi:10.1186/s13059-016-0997-x
- Park, H. Y., Lee, K., Jung, S. C., & Cho, Y. S. (2022). Prevalent toxin types of *Clostridium botulinum* in South Korean cattle farms. *Vet Anim Sci*, 15, 100239. doi:10.1016/j.vas.2022.100239
- Paul, C. J., Twine, S. M., Tam, K. J., Mullen, J. A., Kelly, J. F., Austin, J. W., et al. (2007). Flagellin diversity in *Clostridium botulinum* groups I and II: a new strategy for strain identification. *Appl Environ Microbiol*, 73(9), 2963-2975. doi:10.1128/aem.02623-06

- Peck, M. W. (2009). Biology and genomic analysis of *Clostridium botulinum*. *Adv Microb Physiol*, 55, 183-265, 320. doi:10.1016/s0065-2911(09)05503-9
- Peck, M. W., Smith, T. J., Anniballi, F., Austin, J. W., Bano, L., Bradshaw, M., et al. (2017). Historical Perspectives and Guidelines for Botulinum Neurotoxin Subtype Nomenclature. *Toxins*, 9(1), 38. doi:10.3390/toxins9010038
- Pingeon, J. M., Vanbockstael, C., Popoff, M. R., King, L. A., Deschamps, B., Pradel, G., et al. (2011). Two outbreaks of botulism associated with consumption of green olive paste, France, September 2011. *Eurosurveillance*, 16(49), 20035. doi:doi:https://doi.org/10.2807/ese.16.49.20035-en
- Pirazzini, M., Rossetto, O., Eleopra, R., & Montecucco, C. (2017). Botulinum Neurotoxins: Biology, Pharmacology, and Toxicology. *Pharmacol Rev*, 69(2), 200-235. doi:10.1124/pr.116.012658
- Popoff, M. R. (1995). Ecology of Neurotoxicogenic Strains of Clostridia. In C. Montecucco (Ed.), *Clostridial Neurotoxins: The Molecular Pathogenesis of Tetanus and Botulism* (pp. 1-29). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Popoff, M. R. (1989). Revue sur l'épidémiologie du botulisme bovin en France et analyse de sa relation avec les élevages de volailles. *Rev Sci Tech*, 8(1), 129-145. doi:10.20506/rst.8.1.404
- Popoff, M. R. (2018). *Botulinum Toxins, Diversity, Mode of Action, Epidemiology of Botulism in France*: IntechOpen.
- Portinha, I. M., Douillard, F. P., Korkeala, H., & Lindstrom, M. (2022). Sporulation Strategies and Potential Role of the Exosporium in Survival and Persistence of *Clostridium botulinum*. *International journal of molecular sciences*, 23(2), 754. doi:10.3390/ijms23020754
- Poulain, B., & Humeau, Y. (2003). Le mode d'action des neurotoxines botuliques : aspects pathologiques, cellulaires et moléculaires. *Annales de Réadaptation et de Médecine Physique*, 46(6), 265-275. doi:https://doi.org/10.1016/S0168-6054(03)00114-4
- Prévo, A. R., Huet, M., Tardieux, P. (1950). Étude de vingt-cinq foyers récents de botulisme animal *Bulletin de l'Académie Vétérinaire de France*, 103(9), 481-487.
- Raphael, B. H., Shirey, T. B., Lúquez, C., & Maslanka, S. E. (2014). Distinguishing highly-related outbreak-associated *Clostridium botulinum* type A(B) strains. *BMC Microbiology*, 14(1), 192. doi:10.1186/1471-2180-14-192
- Rasetti-Escargueil, C., Lemichez, E., & Popoff, M. R. (2019). Public Health Risk Associated with Botulism as Foodborne Zoonoses. *Toxins*, 12(1), 17. doi:10.3390/toxins12010017
- Reed, T. M., & Rocke, T. E. (1992). The Role of Avian Carcasses in Botulism Epizootics. *Wildlife Society Bulletin (1973-2006)*, 20(2), 175-182.
- Relun, A., Dorso, L., Douart, A., Chartier, C., Guatteo, R., Mazuet, C., et al. (2017). A large outbreak of bovine botulism possibly linked to a massive contamination of grass silage by type D/C *Clostridium botulinum* spores on a farm with dairy and poultry operations. *Epidemiol Infect*, 145(16), 3477-3485. doi:10.1017/s0950268817002382
- Roberts, T. A., Aitken, I.D. (1974). Botulism in birds and mammals in Great Britain and an assessment of the toxicity of *Clostridium botulinum* type C toxin in the domestic fowl. *Spore Research 1973. Academic Press, London*, 1-9.
- Rocke, T. (2006). Waterbirds around the world. *The Stationery Office*, 422-426.
- Rohde, M., Fischetti, V. A., Novick, R. P., Ferretti, J. J., Portnoy, D. A., Braunstein, M., et al. (2019). The Gram-Positive Bacterial Cell Wall. *Microbiology Spectrum*, 7(3), 7.3.10. doi:doi:10.1128/microbiolspec.GPP3-0044-2018

- Rowe, F. J., & Noonan, C. P. (2017). Botulinum toxin for the treatment of strabismus. *The Cochrane database of systematic reviews*, 3(3), CD006499-CD006499. doi:10.1002/14651858.CD006499.pub4
- Saltykova, A., Buytaers, F. E., Denayer, S., Verhaegen, B., Piérard, D., Roosens, N. H. C., et al. (2020). Strain-Level Metagenomic Data Analysis of Enriched *In Vitro* and *In Silico* Spiked Food Samples: Paving the Way towards a Culture-Free Foodborne Outbreak Investigation Using STEC as a Case Study. *International journal of molecular sciences*, 21(16), 5688.
- Sandrin, T. R., Goldstein, J. E., & Schumaker, S. (2013). MALDI TOF MS profiling of bacteria at the strain level: A review. *Mass Spectrometry Reviews*, 32(3), 188-217. doi:10.1002/mas.21359
- Sasaki, Y., Takikawa, N., Kojima, A., Norimatsu, M., Suzuki, S., & Tamura, Y. (2001). Phylogenetic positions of *Clostridium novyi* and *Clostridium haemolyticum* based on 16S rDNA sequences. *International Journal of Systematic and Evolutionary Microbiology*, 51(3), 901-904. doi:https://doi.org/10.1099/00207713-51-3-901
- Sato, M. P., Ogura, Y., Nakamura, K., Nishida, R., Gotoh, Y., Hayashi, M., et al. (2019). Comparison of the sequencing bias of currently available library preparation kits for Illumina sequencing of bacterial genomes and metagenomes. *DNA Research*, 26(5), 391-398. doi:10.1093/dnares/dsz017
- Scalfaro, C., et al. (2019). Genetic characterization of *Clostridium botulinum* group III using Multilocus Sequence Typing. *Poster presentation Workshop botulism, Maisons-Alfort*.
- Schiavo, G., Benfenati, F., Poulain, B., Rossetto, O., Polverino de Laureto, P., DasGupta, B. R., et al. (1992). Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteolytic cleavage of synaptobrevin. *Nature*, 359(6398), 832-835. doi:10.1038/359832a0
- Schiavo, G., Santucci, A., Dasgupta, B. R., Mehta, P. P., Jontes, J., Benfenati, F., et al. (1993). Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct COOH-terminal peptide bonds. *FEBS Lett*, 335(1), 99-103. doi:10.1016/0014-5793(93)80448-4
- Schocken-Iturrino, R. P., Avila, F. A., Berchielli, S. C., & Nader Filho, A. (1990). First case of type A botulism in zebu (*Bos indicus*). *Vet Rec*, 126(9), 217-218.
- Schouls, L. M., Reulen, S., Duim, B., Wagenaar, J. A., Willems, R. J., Dingle, K. E., et al. (2003). Comparative genotyping of *Campylobacter jejuni* by amplified fragment length polymorphism, multilocus sequence typing, and short repeat sequencing: strain diversity, host range, and recombination. *J Clin Microbiol*, 41(1), 15-26. doi:10.1128/jcm.41.1.15-26.2003
- Schwartz, D. C., & Cantor, C. R. (1984). Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell*, 37(1), 67-75. doi:https://doi.org/10.1016/0092-8674(84)90301-5
- Scott, A. B. (2004). Development of botulinum toxin therapy. *Dermatol Clin*, 22(2), 131-133, v. doi:10.1016/s0733-8635(03)00019-6
- Seddon, H. (1922). Bulbar paralysis in cattle due to the action of a toxicogenic bacillus, with a discussion on the relationship of the condition to forage poisoning (botulism). *Journal of Comparative Pathology and Therapeutics*, 35, 147-190.
- Segner, W. P., & Schmidt, C. F. (1971). Heat Resistance of Spores of Marine and Terrestrial Strains of *Clostridium botulinum* Type C. *Applied microbiology*, 22(6), 1030-1033. doi:doi:10.1128/am.22.6.1030-1033.1971
- Sellin, L. C. (1981). The action of botulinum toxin at the neuromuscular junction. *Med Biol*, 59(1), 11-20.

- Semenova, E., & Severinov, K. (2017). Interdependencies Between the Adaptation and Interference Modules Guide Efficient CRISPR-Cas Immunity. In P. Pontarotti (Ed.), *Evolutionary Biology: Self/Nonsel Evolution, Species and Complex Traits Evolution, Methods and Concepts* (pp. 51-62). Cham: Springer International Publishing.
- Seyboldt, C., Discher, S., Jordan, E., Neubauer, H., Jensen, K. C., Campe, A., et al. (2015). Occurrence of *Clostridium botulinum* neurotoxin in chronic disease of dairy cows. *Vet Microbiol*, 177(3-4), 398-402. doi:10.1016/j.vetmic.2015.03.012
- Skarin, H., Hafstrom, T., Westerberg, J., & Segerman, B. (2011). *Clostridium botulinum* group III: a group with dual identity shaped by plasmids, phages and mobile elements. *BMC Genomics*, 12(1), 185. doi:10.1186/1471-2164-12-185
- Skarin, H., Lindberg, A., Blomqvist, G., Aspán, A., & Båverud, V. (2010). Molecular characterization and comparison of *Clostridium botulinum* type C avian strains. *Avian Pathol*, 39(6), 511-518. doi:10.1080/03079457.2010.526923
- Skarin, H., Lindgren, Y., & Jansson, D. S. (2015). Investigations into an Outbreak of Botulism Caused by *Clostridium botulinum* Type C/D in Laying Hens. *Avian Dis*, 59(2), 335-340. doi:10.1637/10861-051214-Case
- Skarin, H., & Segerman, B. (2011). Horizontal gene transfer of toxin genes in *Clostridium botulinum*: Involvement of mobile elements and plasmids. *Mob Genet Elements*, 1(3), 213-215. doi:10.4161/mge.1.3.17617
- Skarin, H., & Segerman, B. (2014). Plasmidome interchange between *Clostridium botulinum*, *Clostridium novyi* and *Clostridium haemolyticum* converts strains of independent lineages into distinctly different pathogens. *PLOS ONE*, 9(9), e107777. doi:10.1371/journal.pone.0107777
- Skarin, H., Tevell Åberg, A., Woudstra, C., Hansen, T., Löfström, C., Koene, M., et al. (2013). The workshop on animal botulism in Europe. *Biosecur Bioterror*, 11 Suppl 1, S183-190. doi:10.1089/bsp.2012.0076
- Smart, J. L., Jones, T. O., Clegg, F. G., & McMurtry, M. J. (1987). Poultry waste associated type C botulism in cattle. *Epidemiol Infect*, 98(1), 73-79. doi:10.1017/s0950268800061720
- Smith, T., Williamson Charles, H. D., Hill, K., Sahl, J., Keim, P., Relman David, A., et al. (2018). Botulinum Neurotoxin-Producing Bacteria. Isn't It Time that We Called a Species a Species? *mBio*, 9(5), e01469-01418. doi:10.1128/mBio.01469-18
- Sobel, J. (2005). Botulism. *Clinical Infectious Diseases*, 41(8), 1167-1173. doi:10.1086/444507
- Souillard, R., Grosjean, D., Le Gratiot, T., Poezevara, T., Rouxel, S., Balaine, L., et al. (2021). Asymptomatic Carriage of *C. botulinum* Type D/C in Broiler Flocks as the Source of Contamination of a Massive Botulism Outbreak on a Dairy Cattle Farm. *Frontiers in microbiology*, 12, 679377. doi:10.3389/fmicb.2021.679377
- Souillard, R., Le Maréchal, C., Ballan, V., Mahé, F., Chemaly, M., & Le Bouquin, S. (2017). A bovine botulism outbreak associated with a suspected cross-contamination from a poultry farm. *Veterinary Microbiology*, 208, 212-216. doi:https://doi.org/10.1016/j.vetmic.2017.07.022
- Souillard, R., Maréchal, C. L., Hollebecque, F., Rouxel, S., Barbé, A., Houard, E., et al. (2015). Occurrence of *C. botulinum* in healthy cattle and their environment following poultry botulism outbreaks in mixed farms. *Veterinary Microbiology*, 180(1), 142-145. doi:https://doi.org/10.1016/j.vetmic.2015.07.032
- Souillard, R., Woudstra, C., Le Maréchal, C., Dia, M., Bayon-Auboyer, M. H., Chemaly, M., et al. (2014). Investigation of *Clostridium botulinum* in commercial poultry farms in France between 2011 and 2013. *Avian Pathol*, 43(5), 458-464. doi:10.1080/03079457.2014.957644

- Suen, J. C., Hatheway, C. L., Steigerwalt, A. G., & Brenner, D. J. (1988). *Clostridium argentinense* sp. nov.: A Genetically Homogeneous Group Composed of All Strains of *Clostridium botulinum* Toxin Type G and Some Nontoxigenic Strains Previously Identified as *Clostridium subterminale* or *Clostridium hastiforme*. *International Journal of Systematic and Evolutionary Microbiology*, 38(4), 375-381. doi:https://doi.org/10.1099/00207713-38-4-375
- Tatu, L., & Feugeas, J. P. (2021). Botulinum Toxin in WW2 German and Allied Armies: Failures and Myths of Weaponization. *European Neurology*, 84(1), 53-60. doi:10.1159/000512812
- Ting, P. T., & Freiman, A. (2004). The story of *Clostridium botulinum*: from food poisoning to Botox. *Clin Med (Lond)*, 4(3), 258-261. doi:10.7861/clinmedicine.4-3-258
- Tréhard, H., Poujol, I., Mazuet, C., Blanc, Q., Gillet, Y., Rossignol, F., et al. (2016). A cluster of three cases of botulism due to *Clostridium baratii* type F, France, August 2015. *Euro Surveill*, 21(4). doi:10.2807/1560-7917.Es.2016.21.4.30117
- Umeda, K., Seto, Y., Kohda, T., Mukamoto, M., & Kozaki, S. (2009). Genetic characterization of *Clostridium botulinum* associated with type B infant botulism in Japan. *J Clin Microbiol*, 47(9), 2720-2728. doi:10.1128/JCM.00077-09
- Umeda, K., Wada, T., Kohda, T., & Kozaki, S. (2013). Multi-locus variable number tandem repeat analysis for *Clostridium botulinum* type B isolates in Japan: comparison with other isolates and genotyping methods. *Infect Genet Evol*, 16, 298-304. doi:10.1016/j.meegid.2013.02.022
- van Belkum, A., Melchers, W. J., Ijsseldijk, C., Nohlmans, L., Verbrugh, H., & Meis, J. F. (1997). Outbreak of amoxicillin-resistant *Haemophilus influenzae* type b: variable number of tandem repeats as novel molecular markers. *J Clin Microbiol*, 35(6), 1517-1520. doi:10.1128/jcm.35.6.1517-1520.1997
- van Belkum, A., Tassios, P. T., Dijkshoorn, L., Haeggman, S., Cookson, B., Fry, N. K., et al. (2007). Guidelines for the validation and application of typing methods for use in bacterial epidemiology. *Clin Microbiol Infect*, 13 Suppl 3, 1-46. doi:10.1111/j.1469-0691.2007.01786.x
- van Embden, J. D., van Gorkom, T., Kremer, K., Jansen, R., van Der Zeijst, B. A., & Schouls, L. M. (2000). Genetic variation and evolutionary origin of the direct repeat locus of *Mycobacterium tuberculosis* complex bacteria. *J Bacteriol*, 182(9), 2393-2401. doi:10.1128/jb.182.9.2393-2401.2000
- Ventujol, A., Decors, A., Le Maréchal, C., Toux, J., Allain, V., Mazuet, C., et al. (2017). Le botulisme aviaire en France: Etude des cas signalés dans la faune sauvage et dans les élevages par deux réseaux de surveillance entre 2000 et 2013. *Epidémiol Santé Animale*, 72, 85-102.
- Vidal, D., Anza, I., Taggart, M. A., Pérez-Ramírez, E., Crespo, E., Hofle, U., et al. (2013). Environmental factors influencing the prevalence of a *Clostridium botulinum* type C/D mosaic strain in nonpermanent Mediterranean wetlands. *Appl Environ Microbiol*, 79(14), 4264-4271. doi:10.1128/aem.01191-13
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., et al. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic acids research*, 23(21), 4407-4414. doi:10.1093/nar/23.21.4407
- Wentz, T. G., Muruvanda, T., Lomonaco, S., Thirunavukkarasu, N., Hoffmann, M., Allard, M. W., et al. (2017). Closed Genome Sequence of *Chryseobacterium piperi* Strain CTM/ATCC BAA-1782, a Gram-Negative Bacterium with Clostridial Neurotoxin-Like Coding Sequences. *Genome Announc*, 5(48), e01296-01217. doi:doi:10.1128/genomeA.01296-17

- Wentz, T. G., Tremblay, B. J. M., Bradshaw, M., Doxey, A. C., Sharma, S. K., Sauer, J.-D., et al. (2022). Endogenous CRISPR-Cas Systems in Group I *Clostridium botulinum* and *Clostridium sporogenes* Do Not Directly Target the Botulinum Neurotoxin Gene Cluster. *Frontiers in microbiology*, *12*. doi:10.3389/fmicb.2021.787726
- Wilson, R. B., Boley, M. T., & Corwin, B. (1995). Presumptive botulism in cattle associated with plastic-packaged hay. *J Vet Diagn Invest*, *7*(1), 167-169. doi:10.1177/104063879500700131
- Wimmer, F., & Beisel, C. L. (2020). CRISPR-Cas Systems and the Paradox of Self-Targeting Spacers. *Frontiers in microbiology*, *10*. doi:10.3389/fmicb.2019.03078
- Work, T. M., Klavitter, J. L., Reynolds, M. H., & Bleher, D. (2010). Avian botulism: a case study in translocated endangered Laysan Ducks (*Anas laysanensis*) on Midway Atoll. *J Wildl Dis*, *46*(2), 499-506. doi:10.7589/0090-3558-46.2.499
- Woudstra, C., Brito, R. B., Júnior, A. A. F., Silva, R. O. S., Lobato, F. C. F., & Fach, P. (2017a). Draft Genome Sequences of Five Brazilian *Clostridium botulinum* Group III Type D/C Strains. *Genome Announc*, *5*(20), e00349-00317. doi:10.1128/genomeA.00349-17
- Woudstra, C., Lambert, D., Anniballi, F., Medici, D. D., Austin, J., & Fach, P. (2013). Genetic Diversity of the Flagellin Genes of *Clostridium botulinum* Groups I and II. *Appl Environ Microbiol*, *79*(13), 3926-3932. doi:10.1128/AEM.00686-13
- Woudstra, C., Le Maréchal, C., Souillard, R., Anniballi, F., Auricchio, B., Bano, L., et al. (2018). Investigation of *Clostridium botulinum* group III's mobilome content. *Anaerobe*, *49*, 71-77. doi:10.1016/j.anaerobe.2017.12.009
- Woudstra, C., Le Maréchal, C., Souillard, R., Bayon-Auboyer, M.-H., Mermoud, I., Desoutter, D., et al. (2016). New Insights into the Genetic Diversity of *Clostridium botulinum* Group III through Extensive Genome Exploration. *Frontiers in microbiology*, *7*, 757-757. doi:10.3389/fmicb.2016.00757
- Woudstra, C., Le Maréchal, C., Souillard, R., Bayon-Auboyer, M. H., Anniballi, F., Auricchio, B., et al. (2015). Molecular gene profiling of *Clostridium botulinum* group III and its detection in naturally contaminated samples originating from various European countries. *Appl Environ Microbiol*, *81*(7), 2495-2505. doi:10.1128/aem.03915-14
- Woudstra, C., Maklin, T., Derman, Y., Bano, L., Skarin, H., Mazuet, C., et al. (2021). Closing *Clostridium botulinum* Group III Genomes Using Long-Read Sequencing. *Microbiology resource announcements*, *10*(22), e0136420. doi:10.1128/MRA.01364-20
- Woudstra, C., Skarin, H., Anniballi, F., Fenicia, L., Bano, L., Drigo, I., et al. (2012). Neurotoxin gene profiling of *Clostridium botulinum* types C and D native to different countries within Europe. *Appl Environ Microbiol*, *78*(9), 3120-3127. doi:10.1128/aem.07568-11
- Wyman, A. R., & White, R. (1980). A highly polymorphic locus in human DNA. *Proceedings of the National Academy of Sciences*, *77*(11), 6754-6758. doi:10.1073/pnas.77.11.6754
- Xie, X., Hu, Y., Xu, Y., Yin, K., Li, Y., Chen, Y., et al. (2017). Genetic analysis of *Salmonella enterica* serovar Gallinarum biovar Pullorum based on characterization and evolution of CRISPR sequence. *Veterinary Microbiology*, *203*, 81-87. doi:https://doi.org/10.1016/j.vetmic.2017.02.010
- Yadav, R., Yadav, J., Maherchandani, S., & Kashyap, S. K. (2018). Typing of *Campylobacter jejuni* isolated from poultry on the basis of flaA-RFLP by various restriction enzymes. *Veterinary and Animal Science*, *6*, 1-5. doi:https://doi.org/10.1016/j.vas.2018.06.003

- Yule, A. M., Barker, I. K., Austin, J. W., & Moccia, R. D. (2006). Toxicity of *Clostridium botulinum* type E neurotoxin to Great Lakes fish : Implications for avian botulism. *Journal of Wildlife Diseases*, 42(3), 479-493. doi:10.7589/0090-3558-42.3.479
- Zechmeister, T. C., Kirschner, A. K., Fuchsberger, M., Gruber, S. G., Suess, B., Rosengarten, R., et al. (2005). Prevalence of botulinum neurotoxin C1 and its corresponding gene in environmental samples from low and high risk avian botulism areas. *Altex*, 22(3), 185-195.
- Zengiski, A. C. S., Basso, I. B., Cavalcante-Leão, B. L., Stechman-Neto, J., Santos, R. S., Guariza-Filho, O., et al. (2021). Effect and longevity of botulinum toxin in the treatment of gummy smile: a meta-analysis and meta-regression. *Clin Oral Investig.* doi:10.1007/s00784-021-04223-w
- Zhang, S., Lebreton, F., Mansfield, M. J., Miyashita, S. I., Zhang, J., Schwartzman, J. A., et al. (2018). Identification of a Botulinum Neurotoxin-like Toxin in a Commensal Strain of *Enterococcus faecium*. *Cell Host Microbe*, 23(2), 169-176.e166. doi:10.1016/j.chom.2017.12.018
- Zhang, S., Masuyer, G., Zhang, J., Shen, Y., Lundin, D., Henriksson, L., et al. (2017). Identification and characterization of a novel botulinum neurotoxin. *Nature communications*, 8, 14130. doi:10.1038/ncomms14130

## ANNEXES

Annexe 1 : Plan de Management Qualité de la Thèse et rapport final des deux audits effectués lors de la thèse.

Plan de Management Qualité de la Thèse

« Typage des souches de *Clostridium botulinum* du groupe III : développement d'outils et application à l'épidémiologie du botulisme animal. »

Marianne CHEMALY (Directrice de thèse)

Caroline LE MARECHAL (Co-encadrante de thèse/Chef de projet)

Thibault LE GRATIET (Doctorant)

Révision 01 du 04/10/2021

Révision 01 04/10/2021



## PLAN

INTRODUCTION .....	3
Le Système de management de la qualité .....	3
Les acteurs du projet .....	4
La conduite du projet .....	5
a- Planning et étapes de revues .....	5
b- Traçabilité .....	5
c- Audit Interne .....	6
d- Qualité des résultats.....	6
e- Processus de management des risques .....	6
f- Gestion du MOT .....	7
g- Fin du projet : revue de résultats .....	8
LES ÉQUIPEMENTS.....	8
LES CONSOMMABLES .....	8
LES LOCAUX .....	9
FINANCEMENT DE LA THÈSE .....	9
GESTION DES ÉCARTS .....	9
COMMUNICATIONS SUR LE PROJET .....	9
ANNEXES.....	10

## INTRODUCTION

L'Anses (Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail), créée le 1<sup>er</sup> Juillet 2010 de la fusion de l'Afssa (Agence française de sécurité sanitaire des aliments) et l'Afsset (Agence française de sécurité sanitaire de l'environnement et du travail), est un établissement public placé sous la tutelle des ministères chargés de la Santé, de l'Agriculture, de l'Environnement, du Travail et de la Consommation. L'Agence assure des missions de veille, d'expertise et de recherche contribuant à assurer la sécurité sanitaire humaine dans le domaine du travail, de l'environnement et de l'alimentation. Elle va notamment s'assurer de la protection de la santé des végétaux, de la protection de la santé et du bien-être animal et la sécurité des aliments.

L'Agence organise son fonctionnement en s'appuyant sur un système de management de la qualité, fondé sur une approche par processus, afin d'assurer la réalisation de ses différentes activités dans une organisation d'ensemble cohérente et efficace. Ce système s'appuie sur la norme ISO 9001 pour laquelle l'Anses est certifiée.

Les projets de thèse de l'Agence sont conduits selon la norme FD ISO 10006 (Système de management de la qualité - Lignes directrices pour le management de la qualité dans les projets), mis en œuvre par la direction, les chefs d'unité, les responsables du projet et le SMQ. Selon cette norme, un projet est un processus unique qui consiste en un ensemble d'activités coordonnées et maîtrisées comportant des dates de début et de fin, dont l'objectif final se doit conforme à des exigences spécifiques. Pour le management des thèses, sa gestion correspond à la mise en œuvre du projet tel qu'il est validé par les instances de l'Anses.

Le projet de thèse « Typage des souches de *Clostridium botulinum* du groupe III : développement d'outils et application à l'épidémiologie du botulisme animal » est co-financé par la région Bretagne et l'Anses. Le projet commence le 2 Janvier 2019 pour une durée de 3 ans et 3 mois dans l'unité Hygiène et Qualité des Produits Avicoles et Porcins (HQPAP) de l'Anses Ploufragan-Plouzané-Niort. Compte tenu du contexte sanitaire, le projet a été prolongé de 3 mois par la région Bretagne.

L'unité HQPAP est gérée selon la norme NF EN ISO/IEC 17025 (Exigences générales concernant la compétence des laboratoires d'étalonnages et d'essais), qui établit les exigences générales de compétence pour effectuer des essais et des étalonnages. Cette norme prouve que le laboratoire a la capacité de fournir des résultats d'étalonnages et d'essais respectant le niveau de qualité nécessaire pour satisfaire les clients. L'unité HQPAP est accréditée par le COFRAC, le laboratoire se référant au document LAB GTA 59 (Analyses microbiologiques des produits et environnements agroalimentaires), au document LAB GTA 36 (Essais et analyses en bactériologie animale) et au domaine BIOMOL-SA (Analyse de biologie moléculaire en santé animale).

### Le Système de management de la qualité

Le système de management de la qualité (SMQ) correspond à l'ensemble des activités mises en place par l'organisme pour atteindre ses objectifs liés à son plan qualité. Il est indispensable à la maîtrise et à l'amélioration de l'ensemble des processus effectués par l'organisme. En effet, un SMQ va permettre une optimisation des activités, générant une amélioration continue des performances et donc des résultats obtenus. Le service est composé d'un responsable qualité, son adjoint, un responsable métrologie et d'une secrétaire. Chaque unité de recherche a un correspondant qualité qui s'assure que la politique qualité est assurée au sein de l'unité. Un correspondant métrologie est aussi présent et va être responsable du suivi des contrôles des appareils de mesure.

Au laboratoire de l'Anses de Ploufragan-Plouzané-Niort, ce système est mis en œuvre pour l'ensemble des 8 unités de recherche selon les objectifs définis par la direction, sur la base de la norme NF EN ISO/IEC 17025. La documentation qualité comprend le Manuel Qualité, les procédures, les modes opératoires et les supports d'enregistrement. La procédure P.ESS.P1 (PPN/PG/0017) décrit la gestion d'un projet à l'Anses Ploufragan-Plouzané-Niort, dont les projets de thèse.

La rédaction du Plan Management Qualité (PMQ) de ce projet de thèse fait partie intégrante du SMQ suivi par l'unité HQPAP. Ce document retranscrit l'organisation des travaux effectués lors de la thèse dans le but d'assurer la qualité des résultats obtenus.

### Les acteurs du projet

Pour la gestion d'une thèse, l'équipe comprend des scientifiques confirmés connaissant les méthodes utilisées ainsi que toute personne ou structure ayant des compétences indispensables à la mise en œuvre et à la réussite du projet. L'ensemble des acteurs intervenant dans le projet de thèse est défini par l'organigramme fourni en annexe 5. Les agents Anses intervenant dans le projet sont qualifiés et sensibilisés à l'assurance qualité. Leurs responsabilités formalisées sont définies par l'organigramme. Des fiches personnelles sont remplies par chaque agent de l'unité où leurs compétences sont renseignées.

Le chef de projet est le directeur ou co-directeur de la thèse, il est titulaire d'une HDR ou d'une expérience reconnue équivalente par le directeur de thèse. Pour ce projet, Caroline LE MARECHAL tient le rôle de chef de projet. Elle valide les travaux du doctorant en accord avec le directeur de thèse et s'assure que le système de management de la qualité du projet de thèse est mis en œuvre. Le directeur du projet de thèse est Marianne CHEMALY, son rôle est, selon la norme FD ISO 10006, l'élaboration et la mise en œuvre d'un projet correspondant à une série d'objectifs novateurs, dont la réalisation peut impliquer plusieurs entités. Le directeur s'assure du bon déroulement de l'apprentissage du doctorant.

Le comité de pilotage de la thèse regroupe des experts qui ont pour rôle de faire le point sur l'avancée du projet de thèse, de repérer d'éventuelles difficultés tout en proposant des solutions pour y remédier, et de vérifier que la thèse peut être soutenue en respectant les règles en vigueur. Il se réunira périodiquement au cours duquel le doctorant présentera ses résultats.

Ce comité est composé de :

- Mme Marianne Chemaly, Directeur de thèse, titulaire HDR
- Mme Caroline Le Maréchal, encadrante de thèse
- Mme Sergine Even, chargée de recherche INRA UMR STLO à Rennes, titulaire HDR
- Mme Christelle Mazuet, scientifique, Institut Pasteur de Paris, Responsable CNR des bactéries anaérobies et du botulisme.
- Mr Patrick Fach, directeur de recherche, titulaire HDR, Plateforme IDENTYPATH, laboratoire de sécurité des aliments, Anses Maisons-Alfort.
- Mr Cedric Woudstra, chercheur postdoctoral, Sécurité sanitaire des aliments et zoonoses, Université de Copenhague

Mr Eric Chorin, responsable du SMQ, est le représentant de la direction pour le suivi du management qualité de la thèse.

Mme Laetitia Bonifait est la correspondante qualité de l'unité Hygiène et Qualité des Produits Avicoles et Porcins.

## La conduite du projet

### a- Planning et étapes de revues

Les tâches du projet de thèse ont été prévues au départ comme décrit ci-dessous :

- Isolement de souches à partir d'échantillons collectés lors d'épisodes de botulisme.
- Typage des souches par les méthodes MLVA et MLST.
- Étude du plasmidome.
- Analyses comparative des génomes entiers et recherche de cibles pour le typage.
- La métagénomique au service de l'épidémiologie moléculaire.

Ce planning a été préparé avant le début de la thèse par le chef du projet. L'ensemble des acteurs a pris connaissance de ce planning :

- Le directeur de thèse et le chef du projet lors de la mise en place du projet.
- Le doctorant au moment de postuler et lors de son arrivée au laboratoire.
- Les membres du comité de thèse lors de la première réunion.

Il est susceptible d'être modifié en fonction de l'avancée du projet, des résultats obtenus ou en concertation avec le comité de thèse. Des étapes de revue, organisées tout au long du projet, ont pour objectif de faire le point sur l'avancée de la thèse. Elles peuvent mettre en lumière des écarts par rapport aux objectifs fixés au début du projet. Le planning, le contenu ainsi que le PMQ peuvent ainsi faire l'objet de modifications.

Les différents types d'étapes de revue sont :

- Les réunions de comités de thèse, au nombre de 3, prévues à la fin de chaque année de thèse. Elles permettent de faire le point sur l'avancement du projet et de décider des directions à prendre. Suite à chaque réunion, un compte-rendu est écrit par le doctorant et transmis aux différents membres de ce comité et archivé au laboratoire.

- Les réunions de travail entre le doctorant, le directeur de thèse et le chef de projet sont organisées périodiquement (tous les deux mois ou lorsque des résultats importants sont obtenus). Lors de ces réunions, les résultats sont présentés, ainsi que l'ensemble des événements liés à la thèse (formations suivies, participation à des congrès, obligations de l'école doctorale). Les comptes-rendus de ces réunions sont rédigés par le doctorant et consignés dans son cahier de réunion.

Les propositions de modifications du projet par le doctorant au directeur de thèse et au chef de projet sont discutées au cours des réunions de travail, et inversement. Le directeur de thèse et le chef de projet décident des modifications à effectuer. Si un désaccord persiste, le comité de suivi de thèse tranche en faveur d'un des deux parties ou propose un compromis.

### b- Traçabilité

Les résultats obtenus au cours de la thèse sont consignés dans un cahier de laboratoire par le doctorant. Ce dernier est vérifié et signé régulièrement par le chef de projet. Il relate de l'ensemble des manipulations effectuées (modes opératoires, procédures), ainsi que du cheminement intellectuel emprunté par le doctorant. La gestion des cahiers de laboratoire et de réunion de l'unité se fait à l'aide d'un fichier Excel qu'on retrouve sur O:\06-QUALITE\CODES PROJETS -CAHIERS DE LABO. Un autre cahier de laboratoire est utilisé pour les compte rendus des réunions qui ont lieu au cours du projet. Il

contient toutes les prises de décision et les thèmes abordés au cours des réunions périodiques entre le doctorant, le chef de projet et le directeur de thèse, mais aussi les réunions exceptionnelles (résultats importants, collaborations avec des personnes extérieures, comité de suivi individuel). Les cahiers sont classés dans le bureau du doctorant (salle 347).

Les résultats obtenus sont stockés sur le répertoire O:\ dans le dossier réservé au doctorant ainsi que dans le répertoire H:\ dédié aux personnes habilitées selon la LISTE0035 (PPN/FSE/1025).

#### c- Audit Interne

Un audit interne, effectué par le responsable qualité du laboratoire, est effectué un an après le début de la thèse. Cet audit permet de vérifier que la thèse se déroule dans le respect des engagements du présent Plan Management de la Qualité.

En effet, d'après la norme NF EN ISO/IEC 17000 de 2005, un audit interne est un processus systématique indépendant et documenté permettant d'obtenir des enregistrements, énoncés des faits ou autres informations pertinentes, et de les évaluer de manière objective pour déterminer dans quelles mesures les exigences sont respectées.

Au terme de cet audit, un rapport est rédigé et transmis aux membres du comité de thèse. Une révision du Plan Management de la Qualité peut avoir lieu si un écart est détecté.

#### d- Qualité des résultats

Pour l'ensemble des résultats du projet, il est nécessaire d'inclure des témoins négatifs et positifs afin de valider les résultats finaux. Des témoins positifs et négatifs sont inclus dans tous les essais réalisés au sein de l'unité HQPAP, tels que les enrichissements pour la détection de *Clostridium botulinum*, ou encore les PCR.

#### e- Processus de management des risques

D'après la norme FD ISO 10006 le terme « risque » considère, comme le terme « incertitude », les aspects tant positifs que négatifs. Le management des risques du projet traite de manière préventive des incertitudes tout au long du projet.

Ces risques peuvent empêcher l'aboutissement du projet. Une liste des différents risques avec actions préventives adaptées est détaillée ci-dessous :

- ✓ Une panne de l'enceinte anaérobie entrainerait un arrêt de l'isolement des souches de *Clostridium botulinum* et donc l'impossibilité de poursuivre certains projets nécessitant de travailler avec une souche isolée. (Action préventive : contrat de maintenance, acquisition prévue en 2020 d'un générateur d'atmosphère anaérobie en jarre)
- ✓ Difficultés dans le transfert de la méthode MLVA au sein du laboratoire (matériel différent, nouvelle méthode non utilisée jusqu'à présent au laboratoire) (Action préventive : formation de Thibault à l'ISS pour garantir ce transfert, contact avec Qiagen pour tester le même appareil qu'à l'ISS)
- ✓ Problème de dégradation des ADN lors de la préparation des ADN pour le séquençage des souches (bactéries difficiles à lyser et productrices de DNase) (Actions préventives : vérification de la qualité de l'ADN avant envoi, appui technique avec ISS et plateforme IdentityPath pour améliorer la méthode d'extraction - Sélection d'un kit d'extraction

d'ADN adapté à *Clostridium botulinum* du groupe III après une étude comparative de plusieurs kits).

- ✓ Survie de *C. botulinum* au processus d'extraction d'ADN, contamination des extraits d'ADN, non conforme à la réglementation MOT (action préventive : qualification des protocoles d'extraction d'ADN (vérification systématique de la stérilité des ADN avant utilisation).
- ✓ Difficultés dans l'isolement des souches du fait de l'absence d'une méthode robuste, impossibilité d'avoir un nombre suffisant de souches pures pour effectuer une évaluation des méthodes de typage liées aux génomes complets (Action préventive : Tester des méthodes de typage qui ne nécessitent pas l'isolement des souches comme la MLVA ou du typage avec des données de métagénomique).
- ✓ Il est impératif de disposer des autorisations de cession pour expédier les ADN à la plateforme IDENTYPATH pour le séquençage des souches. L'autorisation doit être renouvelée annuellement, un retard dans l'obtention de cette autorisation entraînerait des délais pour l'envoi des ADN à séquencer. (Action préventive : Prévoir la demande de cession au moins deux mois avant la date programmée d'expédition, demande de cession permanente).
- ✓ L'impossibilité de séquencer les souches isolées par la plateforme IDENTYPATH car aucune disponibilité ou la plateforme n'est pas fonctionnelle (Action préventive : effectuer le séquençage avec le matériel présent sur le site de Ploufragan en respectant les autorisations de l'ANSM liées aux MOT – exemple : emprunt du MinION présent dans l'unité GVB).
- ✓ Mise en place d'un confinement suite à une pandémie, impossibilité d'effectuer les manipulations pour certaines tâches du projet (Action préventive : demande d'extension de la durée de la thèse équivalente à la durée du confinement auprès des financeurs et de l'école doctorale).
- ✓ Difficultés d'approvisionnement pour certains consommables (Plaques PCR, pointes à filtre...) (Action préventive : anticipation par la gestionnaire des commandes de l'unité HQPAP, révision de certains protocoles pour réduire la consommation des réactifs et consommables sous tension)

Cette liste pourra être mise à jour lors d'une révision du Plan Management de la Qualité ou lors de réunions organisées à l'arrivée d'un événement critique mis en évidence par l'un des acteurs du projet. Si le point critique ne peut être résolu, ceci pourra conduire à une révision du projet de thèse.

#### f- Gestion du MOT

*Clostridium botulinum* (*C. botulinum*) fait partie de la liste des MOT (Micro-organismes et toxines hautement pathogènes) de l'arrêté du 30 juin 2010 fixant la liste des micro-organismes et toxines. De ce fait, des procédures sont mises en place pour être en conformité avec la réglementation, en particulier l'arrêté des bonnes pratiques du 11 juin 2013. Seules les personnes habilitées et déclarées auprès de l'ANSM sont autorisées à manipuler *C. botulinum*. La procédure P.HQPAP.PER.H1 (PPN/PS/0187 : Habilitation du personnel de l'unité) décrit les modalités d'habilitation et de maintien d'habilitation.

Le fait de manipuler un MOT entraîne des contraintes supplémentaires notamment au niveau administratif par rapport à d'autres pathogènes. Par exemple, le transport de tout ou partie de *C. botulinum* nécessite au préalable l'obtention d'une autorisation auprès de l'ANSM. Les mouvements des tubes sont suivis et tracés de façon à pouvoir connaître l'état des stocks à n'importe quel moment.

g- Fin du projet : revue de résultats

La fin du projet se traduit par la rédaction du manuscrit de thèse suivie d'une soutenance devant un jury amené à valider les résultats obtenus et le mémoire de thèse. Un bilan de management qualité de la thèse sera aussi rédigé par le responsable Qualité pour les membres du jury. Toute modification ou révision de document concernant le projet (fiche de définition du projet de thèse et plan de management qualité) aboutira à l'archivage des anciennes versions.

## LES ÉQUIPEMENTS

Les équipements de l'unité HQPAP sont gérés selon sur la base de la norme NF EN ISO/IEC 17025. L'inventaire et le suivi des équipements sont formalisés par une gestion planifiée et maîtrisée. Un code est attribué pour chaque matériel. Un dossier de vie de l'appareil regroupant une fiche d'inventaire, une fiche de suivi de matériel et la documentation technique du fabricant. Une personne est désignée responsable pour chaque appareil, suppléée par le responsable de la salle où se trouve l'appareil.

Une vérification métrologique est effectuée sur les équipements de mesure et d'essai, dans le but de maîtriser en permanence l'aptitude à l'emploi des équipements. Deux types de vérifications sont mises en place dans l'unité :

- Les vérifications de type interne : les agents de l'unité opèrent les vérifications de matériels soumis à des contrôles métrologiques en collaboration avec le correspondant métrologie de l'unité. Ces contrôles peuvent avoir lieu de façon initiale, périodique (micropipettes, thermocycleurs) et de routine avant leur utilisation par l'agent (balances et pH-mètres). Une centrale de mesure des températures est également présente au sein du laboratoire de Ploufragan-Plouzané afin de suivre la variation des températures au sein des enceintes thermostatées (réfrigérateurs, congélateurs, surgélateurs, étuves).
- Les vérifications de type externe : certains matériels nécessitent des opérations de maintenance et/ou de réparations qui sont assurées par des prestataires de services. Ces opérations sont généralement réalisées dans le cadre de contrat d'entretien et visent à prévenir des anomalies de fonctionnement.

## LES CONSOMMABLES

Les consommables sont gérés dans chaque unité sur la base de la norme NF EN ISO/IEC 17025. Il existe deux types de consommables utilisés dans l'unité HQPAP :

- Les consommables acquis à l'extérieur (kits, produits chimiques, matériels plastiques)
- Les consommables fabriqués au sein de l'unité (milieux de culture)

La gestion des consommables est détaillée et se fait selon la procédure P.HQPAP.CON.G1 (PPN/PS/0198: Gestion du stock des consommables). Pour les consommables fabriqués au sein de l'unité tels que les milieux de culture, la date de fabrication, de péremption et le numéro d'autoclave utilisé s'il y a lieu sont précisés.

Pour assurer la traçabilité, un numéro de lot interne est attribué en plus du numéro de lot fabricant. Les stocks sont gérés en prenant en compte des dates de péremption et la qualité des

consommables. La qualité des consommables est évaluée à leur arrivée par le personnel présent à la réception.

Le choix des fournisseurs des produits et matériels se fait en tenant compte de la certification des fournisseurs et ou des fabricants, le prix des consommables, le délai de livraison et la qualité des consommables dans la qualité d'un résultat, d'un essai ou d'une étude selon les procédures P.MAT.G1 (PPN/PG/0025 : Achats de matériels scientifiques et techniques) et P.CON.G1 (PPN/PG/0010 : Achat de consommables). L'évaluation des fournisseurs est effectuée tous les ans selon la procédure générale P.ACT.E1 (PPN/PG/0006 : Evaluation des fournisseurs). Le résultat de ces évaluations est utilisé en tant que critère pertinent dans la sélection des fournisseurs.

## LES LOCAUX

L'unité HQPAP est installée au site des Croix à Ploufragan, dans le bâtiment F14. L'accès au bâtiment est limité au moyen d'un badge permettant l'ouverture des portes d'accès. La circulation dans les locaux est décrite dans la procédure P.HQPAP.LOC.C1 (PPN/PS/0254 : Circuit des personnes dans l'unité HQPAP). L'accès dans les laboratoires se fait avec un retrait des vêtements extérieurs et changement de chaussures. Un registre permet d'identifier les visiteurs autorisés à pénétrer dans les locaux. *C. botulinum* étant un organisme MOT, l'accès à certaines zones est restreint aux personnes habilitées à manipuler cet organisme. Le bâtiment est équipé d'une climatisation et d'un système de ventilation d'air filtré avec maintien d'un niveau de pression négative par rapport à l'extérieur.

## FINANCEMENT DE LA THÈSE

Le projet a une durée de 3 ans et 3 mois à partir du 02 Janvier 2019. La thèse est financée à 50% par la région Bretagne et à 50% par l'Anses. Le projet Track'Spore (financement AgriFood Transition) assure le financement d'une partie des travaux de la thèse.

## GESTION DES ÉCARTS

La non-conformité est, selon la norme NF EN ISO 9000, la non satisfaction d'une exigence du système qualité et/ou du client. A l'Anses, toute non-conformité est suivie de l'enregistrement d'une fiche de progrès détaillant l'évènement survenu. Une investigation est ensuite effectuée pour déterminer les causes de la non-conformité. Des mesures curatives, correctives et/ou préventives avec évaluations sont mises en places selon la procédure générale Anses ANSES/PG/0004.

La non-confirmation d'une hypothèse ou d'un résultat attendu est notée dans le cahier de manipulations. Une investigation est menée par le doctorant pour mettre la lumière sur les raisons de cette non-confirmation, en lien avec le directeur de thèse et le chef de projet. Elle peut provoquer une revue de projet selon la nature de l'évènement.

## COMMUNICATIONS SUR LE PROJET

Au cours de la thèse, le doctorant doit participer à des congrès, nationaux ou internationaux. Lors de ces congrès, il sera amené à communiquer sur les résultats obtenus. Ses communications peuvent se faire sous la forme de présentation orale ou poster. A minima, une publication scientifique devra aussi être soumise dans un journal à comité de lecture. Des réunions avec le comité de thèse seront organisées périodiquement au cours duquel le doctorant présentera l'avancée de ses travaux de recherche. Le doctorant présentera aussi ses résultats au chef de projet et au directeur de thèse tous les deux mois ou lors de résultats inattendus.



## ANNEXES

Annexe 1 : Fiche de définition du projet de thèse

### PROJET DE THÈSE 2019-2021

#### « Typage des souches de *Clostridium botulinum* du groupe III : développement d'outils et application à l'épidémiologie du botulisme animal. »

Marianne CHEMALY (Directrice de thèse)

Caroline LEMARECHAL (Co-directrice de thèse/Chef de projet)

Thibault LE GRATIET (Doctorant)

Le botulisme est une maladie neuroparalytique affectant les hommes et les animaux. Elle se présente chez l'animal sous la forme d'une paralysie flasque ascendante qui entraîne le décès de l'animal. Elle va avoir un impact économique et environnemental important, puisqu'un épisode de botulisme peut entraîner la perte d'une grande partie du cheptel de l'élevage contaminé.

Cette pathologie est causée par l'action de la toxine botulique, produite principalement par l'espèce *Clostridium botulinum*, une bactérie à coloration de Gram positive, anaérobie stricte se présentant sous la forme de bacille au microscope. Elle a la capacité de sporuler ce qui lui confère la capacité de survivre dans l'environnement jusqu'à plusieurs années. Il existe 7 neurotoxines botuliques différentes historiquement caractérisées, désignées de la lettre A à G. D'autres toxines botuliques ont été décrites depuis mais non reliées à des épisodes cliniques de botulisme. Les clostridies capables de produire la toxine botulique vont ainsi être classées en 4 groupes selon des caractéristiques phénotypiques et les toxines produites. Les souches de *C. botulinum* du groupe III sont les principales souches responsables du botulisme animal. Les toxines botuliques C, D et leurs mosaïques C/D et D/C sont produites par les bactéries de ce groupe.

Les toxines botuliques sont des protéines composées d'une chaîne légère et d'une chaîne lourde (150 kDa environ). La chaîne lourde va jouer un rôle dans la liaison de la toxine aux cellules cibles ainsi que dans la translocation de la chaîne légère dans le cytoplasme. La toxine va agir en interférant sur le

mécanisme d'exocytose des vésicules synaptiques, inhibant la libération de l'acétylcholine dans la fente synaptique. Cette inhibition va avoir pour conséquence la paralysie flasque des muscles squelettiques de l'animal infecté. Il existe deux voies d'exposition à la toxine botulique :

- La première, qui est d'ingérer la toxine déjà formée dans l'eau, la nourriture ou sur les cadavres d'animaux.
- La deuxième est l'ingestion de spores qui, après germination, vont produire de la toxine botulique *in vivo*.

Le botulisme animal est saisonnier, avec une augmentation du nombre de foyers lors de l'été. En effet, la température augmente et se rapproche de la température optimale de croissance de *C. botulinum*. D'autres facteurs de risque sont à prendre en compte pour les éleveurs de volaille ou de bétail :

- Conditions d'hygiène des zones d'élevage
- Densité de l'élevage
- Conditions de stockage de la nourriture et du fumier
- Contamination par des insectes ou des cadavres d'oiseaux sauvages et animaux sauvages.
- Contamination par les eaux environnantes.

En Europe, les cas de botulisme animal sont principalement causés par la mosaïque C/D lorsqu'il s'agit de volailles, tandis que la mosaïque D/C est principalement responsable des cas bovins. L'épidémiologie du botulisme animal est peu connue du fait de l'absence d'outils de typage. Ceci s'explique notamment par le fait que les souches de *C. botulinum* du groupe III sont très difficiles à isoler. Il n'existe aucun milieu sélectif pour isoler efficacement *C. botulinum*, ce qui rend l'isolement des souches à partir de prélèvements du terrain délicat à réaliser et laborieux. De plus, le gène codant les toxines botuliques C ; D, C/D ou D/C est porté par un phage, instable en conditions de laboratoire, qui est souvent perdu lors du processus d'isolement ou lors du stockage des souches. L'analyse du peu de génomes séquencés montre que la diversité génétique se trouve en particulier au niveau des éléments génétiques mobiles, notamment du plasmidome.

Le développement d'outils de typage de *C. botulinum* de groupe III permettrait d'explorer l'épidémiologie du botulisme animal et améliorer la gestion et la prévention des épisodes. Dans le cadre du projet européen ANIBOTNET (2016-2019), coordonné par l'unité HQPAP du laboratoire Anses de Ploufragan-Plouzané-Niort, des méthodes MLVA (Multiple Loci VNTR Analysis) et MLVA-HRM (High Resolution Melting analysis) ont été développées ou sont en développement. Plusieurs génomes ont aussi été séquencés, venant s'ajouter à la base de données présente dans la littérature. Des études concernant les éléments génétiques mobiles présents chez *C. botulinum* ont été effectuées, dévoilant des cibles potentielles pour développer des outils de caractérisation moléculaire discriminant. Enfin, l'émergence des outils de métagénomique depuis plusieurs années pourraient être une piste pour développer des outils innovants ne nécessitant pas d'isolement des souches.

Dans le cadre du projet de thèse, différentes approches seront évaluées. Les objectifs de ce projet sont (i) le développement d'une méthode d'isolement pour faciliter l'obtention de souches pures, (ii) d'évaluer des méthodes classiques de caractérisation moléculaire comme la MLVA ou la MLVA-HRM, (iii) d'évaluer les cibles plasmidiques pour le typage moléculaire, (iv) de rechercher des cibles moléculaires intéressantes en comparant des génomes déjà présents dans la base de données et des génomes qui seront séquencés au cours de la thèse, (v) évaluer la métagénomique comme outil de typage. La finalité du projet de thèse est de proposer une approche de typage permettant de tracer les souches de *C. botulinum* du groupe III.

#### Tâche 1 : Isolement des souches

La difficulté à isoler *C. botulinum* étant un frein dans la compréhension de l'épidémiologie du botulisme animal, l'exploration de méthodes innovantes serait une solution pour faire des avancées. Une nouvelle méthode d'isolement a été développée et perfectionnée au cours de la première année de thèse et sera utilisée tout au long du projet pour isoler les souches de *C. botulinum* détectées par le LNR botulisme depuis sa création. Cette méthode ne fonctionne que pour les échantillons d'organes mais ne permet pas l'isolement de souches pures pour les échantillons d'environnement. Cette étude a fait l'objet d'un article scientifique publiée en Janvier 2020 dans le journal *Toxins* (<https://doi.org/10.3390/toxins12010042>). Le protocole est décrit dans la procédure MO.HQPAP.ESS.MA28 (Unité HQPAP : Isolement d'une souche de *C. botulinum* et mise en conservation – PPN/INS/0703). Une fiche de suivi d'isolement a aussi été mise en place dans la procédure ENR0138 (Unité HQPAP : Fiche d'isolement de souches de *C. botulinum* du groupe III - PPN/FSE/1206).

#### Tâche 2 : MLVA et MLVA-HRM

Ce sont 2 méthodes de typage classiques utilisées pour le typage des souches bactériennes. Elles ont été développées par le laboratoire national de référence du botulisme en Italie dans le cadre du projet ANIBOTNET.

La MLVA ne nécessite pas d'avoir une souche isolée et peut être utilisée à partir d'extraits d'ADN. Les méthodes développées par le partenaire du projet ANIBOTNET seront testées au cours de ce projet de thèse à partir des extraits d'ADN de la collection du LNR. Cela permettra d'évaluer le pouvoir discriminant de ces méthodes. Cette méthode a été utilisée dans le cadre de deux enquêtes épidémiologiques : Un épisode de botulisme bovin dans un élevage mixte bovin/volaille et un épisode de botulisme aviaire avec récurrences sur plusieurs années dans un élevage de volaille. Les extraits d'ADN positifs ont été analysés par MLVA et comparés pour déterminer si le profil MLVA est le même ou non.

-Pour l'épisode de botulisme bovin, le même profil MLVA a été retrouvé dans des prélèvements de bovin, de volaille et d'environnement dont le godet ayant servi à la manipulation du fumier de volaille et la ration des bovins, suggérant une contamination de la ration avec le godet. Cette enquête a fait l'objet d'un article scientifique qui a été publié le 29 Juin 2021 dans le journal *Frontiers in Microbiology* (<https://doi.org/10.3389/fmicb.2021.711413>).

-Pour l'épisode de botulisme aviaire, les profils MLVA retrouvés dans les échantillons prélevés au cours des différentes récurrences sont les mêmes, suggérant une persistance de la souche dans l'élevage. Un article scientifique est en cours de rédaction pour cette enquête.

Enfin, la méthode MLVA a été testée sur les extraits d'ADN des souches qui ont été isolées au cours du projet de thèse. Ces souches sont représentatives des épisodes de botulisme aviaire survenus en France de 2013 à 2019 (volaille et faune sauvage).

En collaboration avec le CNR botulisme italien (ISS), la méthode MLVA-HRM sera aussi évaluée au cours de la thèse pour la comparer à la méthode MLVA traditionnelle. Contrairement à la MLVA classique, cette méthode se base sur la température de fusion des séquences amplifiées pour la détermination du nombre de répétitions de la région. Cependant, le manque de diversité dans le nombre de répétitions des régions ciblées de nos échantillons testés ne nous permet pas de bien calibrer la méthode. La méthode MLVA classique a été privilégiée au cours de la thèse.

### Tâche 3 : Investigation du plasmidome

Les différentes études du génome des souches de *C. botulinum* du groupe III ont permis de montrer que celui-ci est très conservé au niveau chromosomique mais présente une variabilité au niveau du plasmidome (Skarin et al., 2011, Skarin and Segerman, 2014, Woudstra et al., 2016, Woudstra et al., 2018). Treize plasmides ont été identifiés dans une étude menée par le Statens Veterinärmedicinska Anstalt (SVA) (Skarin and Segerman, 2014).

Les souches de type C/D possèdent un grand nombre d'éléments génétiques mobiles tandis que les souches de type D/C n'en possèdent que très peu. L'hypothèse serait l'absence d'un système CRISPR-Cas complet pour les souches C/D, qui a pour conséquence une non dégradation du mobilome et donc une présence beaucoup plus importante d'éléments génétiques mobiles. Il serait donc pertinent d'explorer le mobilome chez les souches C/D pour trouver des cibles plasmidiques intéressantes qui permettrait le typage de *C. botulinum*. En ce qui concerne les souches D/C, l'étude du mobilome sera aussi à faire, cependant, une étude de la diversité des séquences CRISPR peut s'avérer nécessaire pour typer les souches génétiquement très proches (Woudstra et al., 2016). De plus le nombre de génomes disponibles pour ce type toxinique est faible et il sera nécessaire de disposer de plus de génomes pour explorer la piste du plasmidome pour ce type toxinique.

Une étude des systèmes CRISPR-Cas sur les génomes disponibles dans les bases de données a été effectuée au cours de la deuxième année de thèse pour comprendre la variabilité du plasmidome mais aussi pour déterminer si ces régions du génome peuvent servir à développer un outil de typage. La diversité des systèmes identifiées suggèrent que le développement d'un outil n'est pas pertinent. Cette étude a fait l'objet d'un article scientifique qui a été publiée le 13 Septembre 2021 dans le journal *Frontiers in Microbiology* (<https://doi.org/10.3389/fmicb.2021.711413>).

### Tâche 4 : Identification de cibles via analyse comparative des données WGS

Le nombre de génomes actuellement disponibles dans les banques de données publiques est relativement faible. Une étude menée sur *Clostridium botulinum* du groupe I a permis de mettre au point un séquençage de génome complet en couplant le séquençage short reads Illumina avec le séquençage long reads MinION. Cette approche n'a pas été testée sur les souches de *C. botulinum* du groupe III.

Le premier objectif de cette tâche est d'optimiser un protocole d'extraction d'ADN adapté à ces deux méthodes de séquençage pour pouvoir effectuer ce même assemblage couplé pour les génomes de *C. botulinum* du groupe III. L'extraction d'ADN de *C. botulinum* du groupe III est compliquée, en effet, la lyse bactérienne peut s'avérer difficile, entraînant une quantité d'ADN extrait insuffisante pour le séquençage. De plus, des DNases sont libérées lors de cette lyse, entraînant une fragmentation de l'ADN extrait, ce qui est rédhibitoire pour un séquençage de longs fragments. Une comparaison de 5 kits d'extraction d'ADN a été effectuée pour sélectionner le kit à le plus approprié pour *C. botulinum* du groupe III. Une fois le kit sélectionné, les ADN extraits seront séquencés par les deux méthodes de séquençage et un génome complet sera assemblé à l'aide d'outils bioinformatiques. Cette méthode sera d'abord utilisée pour séquencer une souche de référence du laboratoire puis sur des souches isolées au cours de la thèse.

En parallèle, à partir des données disponibles et du séquençage des souches isolées antérieurement par le LNR, une recherche de cibles pour le typage des souches sera réalisée. Les souches de *C. botulinum* du groupe III font partie d'une groupe plus large appelé *C. novyi sensu lato*. Les souches de ce groupe sont divisées en 4 lignées. Compte tenu des analyses comparatives des génomes de *C. novyi sensu lato*, il n'est pas envisageable de proposer un outil de typage adapté à l'ensemble des lignées. Une solution serait plutôt de proposer un outil adapté à chacune de ces lignées. Des travaux antérieurs au projet de thèse avaient montré que les gènes codant pour les flagellines pouvaient constituer une piste

intéressante. Avec l'appui de l'équipe de la plateforme de séquençage de l'unité GVB, les gènes codant pour les flagellines ont été recherchés dans les génomes du groupe *C. novyi sensu lato*. Ces analyses ont permis de trouver des régions spécifiques pour chaque lignée au niveau de ces gènes de flagellines. Des amorces PCR ont été désignées pour cibler une séquence de ces régions pour chaque lignée. Les systèmes PCR sont en cours d'évaluation/ Cette étape est financée par le projet Track' Spore (Tremplin Carnot AgriFood Transition) en collaboration avec l'ADRIA (Quimper).

#### Tâche 5 : Métagénomique

L'avancée technologique de la métagénomique depuis plusieurs années la rend plus accessible et moins coûteuse. Elle permettrait de s'affranchir de l'étape d'isolement de *C. botulinum* puisque l'ensemble de l'ADN présent dans le milieu est séquencé. Cette approche étant encore récente, il n'est pas encore démontré que le séquençage métagénomique puisse remplacer les méthodes classiques actuelles dans un futur proche. Deux méthodologies de métagénomique :

La première, la méthode « shotgun metagenomic », consiste à séquencer tous les fragments d'ADN présents dans l'échantillon, puis d'analyser les données à l'aide d'algorithmes complexes.

La deuxième, la méthode « high-throughput targeted-amplicon high », est plus ciblée puisqu'une cible prédéfinie va être amplifiée (l'ARN 16S pour l'étude des bactéries par exemple) puis ces produits PCR seront ensuite séquencés.

Cette tâche du projet prévue dans le programme initial de la thèse n'a pas pu être abordée faute de temps. Elle n'apparaîtra donc pas dans le manuscrit de thèse.

#### Tâche 6 : Rédaction de publications scientifiques et de la thèse

Une première publication scientifique a été rédigée et publiée pour présenter la méthode d'isolement mise en place au cours de la première année de Thèse. La publication a été soumise le 20 Décembre 2019, acceptée le 9 Janvier et publiée le 10 Janvier dans le journal *Toxins* (<https://doi.org/10.3390/toxins12010042>). Une deuxième publication scientifique a été rédigée et publiée pour présenter l'étude d'un épisode de botulisme bovin dans un élevage mixte volaille/bovin. Une partie de la publication présente l'utilisation de l'outil MLVA pour tracer la souche de type D/C impliquée dans l'épisode de botulisme. La publication a été soumise le 11 Mars 2021, acceptée le 7 Juin et publiée le 29 Juin dans le journal *Frontiers in Microbiology* (<https://doi.org/10.3389/fmicb.2021.679377>). Une troisième publication scientifique portant sur l'étude des systèmes CRISPR-Cas chez *Clostridium novyi sensu lato* a été rédigée et publiée. Elle a été soumise le 18 Mai 2021, acceptée le 16 Août et publiée le 13 Septembre dans le journal *Frontiers in Microbiology* (<https://doi.org/10.3389/fmicb.2021.711413>).

Des communications orales et des posters lors de congrès ou journées scientifiques ont aussi été effectués par le doctorant au cours de la thèse. Trois communications orales portant sur le projet de thèse ont été présentées au cours de la thèse : lors de la demi-journée d'échange sur le botulisme animal qui a eu lieu le 26 Novembre 2020, lors des journées scientifiques de l'école doctorale EGAAL qui ont eu lieu le 30 Juin et 1<sup>er</sup> Juillet 2021 et lors des Journées Scientifiques et Doctorales de l'Anses (JSDA) qui ont eu lieu en Septembre 2021.

Deux posters portant sur l'isolement des souches ont été présentés lors des Journées Scientifiques et Doctorales de l'Anses (JSDA) en Septembre 2020 et lors du World One Health Congress (WOHC) qui a eu lieu en ligne du 30 Octobre 2020 au 3 Novembre. Un poster portant sur l'outil de typage MLVA utilisé lors de l'enquête épidémiologique d'un épisode de botulisme bovin a été présenté au congrès EJP One Health qui a eu lieu en ligne du 9 au 11 Juin 2021. Un poster portant sur l'étude des systèmes CRISPR-Cas a été présenté lors du congrès de la SFM qui a eu lieu du 22 au 24 Septembre 2021 à Nantes.

Annexe 2 : Déclaration de la politique qualité du directeur du laboratoire de Ploufragan-Plouzané-Niort



## Déclaration de Politique Qualité du Laboratoire de Ploufragan-Plouzané-Niort

Le Laboratoire de Ploufragan-Plouzané-Niort fait siennes les politiques qualité du Directeur Général de l'Anses et du Directeur Général Délégué à la Recherche et la Référence.

Au sein du Laboratoire, l'objectif de la politique qualité est d'abord de garantir la fiabilité de la partie analytique de nos programmes de recherche et d'appui technique. Ce fonctionnement sous management de la qualité est essentiel pour les activités de référence mais la réflexion est étendue progressivement, à l'ensemble de nos missions. C'est d'ailleurs le sens de la Déclaration du Directeur général. L'Agence elle-même est certifiée ISO 9001 depuis 2013.

Le laboratoire de Ploufragan-Plouzané-Niort est engagé dans une politique d'accréditation depuis 1996. La norme ISO 17025 nous a incités à développer une politique d'amélioration continue de notre management qualité et une démarche de mesure de la satisfaction de nos clients. Notre position de laboratoire de référence dans plusieurs domaines nous a amenés à adapter cette dernière disposition à leur faible nombre. En conformité avec les évolutions introduites par la nouvelle version de la norme ISO 17025, le laboratoire :

- i) s'attache à garantir son impartialité (DPI révisées et analysées pour les responsables du laboratoire, des unités et des LNR / absence de participation aux études susceptibles d'être évaluées par d'autres entités de l'Anses / absence de responsabilité dans les décisions de gestion sanitaire prises par les tutelles / engagement par la présente déclaration à signaler les situations susceptibles de mettre en jeu l'impartialité de l'établissement) ;
- ii) veille à évaluer également ses prestataires de service internes (DTI par exemple) ;
- iii) s'est engagé dans une démarche d'analyse des risques et opportunités pour le laboratoire.

L'ensemble des mesures visant à garantir le fonctionnement du système d'assurance qualité du laboratoire ainsi que les évolutions décrites au paragraphe précédent, en cohérence avec les orientations nationales, est régulièrement analysé et si besoin révisé à l'occasion des réunions et de la revue de Direction annuelle du laboratoire.

En parallèle, les exigences réglementaires relatives aux Laboratoires Nationaux de Référence, se traduisent désormais par l'organisation d'EILA selon la norme ISO 17043. Comme pour les essais, le management qualité de l'organisation des EILA est destiné à donner encore plus confiance aux laboratoires de diagnostic qui ont généralement l'obligation d'y participer. L'unité VIPAC est accréditée au titre de la norme NF EN ISO 17043 sous le numéro 1-5619 depuis 2014. La majorité des OCIL fait l'objet d'audits internes selon cette norme.

Grâce à la compétence et l'expertise scientifique et technique des agents du laboratoire, nos activités dans les domaines de la santé et du bien-être des volailles, lapins, porcs, ruminants et poissons, d'une part, et de la qualité des produits issus de ces animaux (porcs, volailles et lapins), d'autre part, sont perçues par nos tutelles et nos partenaires professionnels comme faisant référence dans nos domaines

Liste 1050 version 02, (PPN/ERQ/0035) Page 1/3

de compétences (14 mandats de Laboratoire National de Référence, 4 mandats de laboratoire de référence de l'Organisation Mondiale de la Santé Animale, 1 mandat récemment obtenu de Centre de Référence de l'Union Européenne). Afin d'assurer encore et toujours leur fiabilité, leur rigueur et l'indépendance de leurs résultats, le développement du management qualité est indispensable.

Dès 2004, l'unité "Génétique Virale et Biosécurité" s'est engagée dans la voie d'une démarche « qualité en recherche » en s'appuyant à la fois sur les normes ISO 17025 et ISO 10006, démarche qui a depuis été étendue aux autres unités. Depuis la Revue de Direction de février 2005, l'essentiel des thèses de doctorat préparées au Laboratoire, est conduit sous management qualité, lequel doit contribuer à garantir la conduite des travaux dans le respect de la charte de déontologie des métiers de la recherche, notamment pour ce qui concerne la traçabilité de tous ces travaux, qui fait l'objet d'un axe d'amélioration continue au Laboratoire et à l'Agence.

La mise en œuvre de notre politique qualité concerne également progressivement le fonctionnement de nos services d'expérimentation animale (notamment dans le cadre du nouveau mandat de Centre de Référence de l'Union européenne pour les volailles et les petits animaux d'élevage) et de nos fonctions support. Elle a vocation à s'étendre progressivement à l'ensemble de nos activités et prestations. Cependant, ces évolutions ne peuvent se concevoir qu'avec les moyens humains et matériels (LIMS, logiciel LEILA pour la gestion des EILA, système Ennov de Gestion Electronique des Documents mis en place) adéquats et fonctionnels.

Conformément aux décisions du Directeur général, a débuté la mise en place effective de la gestion des risques majeurs. En lien avec la direction du Laboratoire, le Service Management Qualité est chargé d'accompagner les actions de suivi des risques et la remontée d'informations, ce qui correspond à la déclinaison locale de l'analyse des risques et opportunités demandée par la nouvelle version de la norme ISO17025.

Pour atteindre au mieux ces objectifs ambitieux en matière de politique qualité, le Laboratoire de Ploufragan-Plouzané-Niort s'est doté d'une organisation efficace avec un responsable qualité, rattaché au directeur et chargé de l'accompagner et de le conseiller dans la mise en œuvre de la politique qualité. Le responsable qualité dirige le Service Management Qualité, lequel est constitué d'une responsable métrologie, d'une secrétaire et d'un chargé de projet scientifique et technique à mi-temps (assistant de prévention pour le site de Ploufragan et chargé de la coordination HS intersite pour son autre mi-temps). Au sein de chaque unité et service, des correspondants qualité et métrologie sont nommés.

L'organisation du laboratoire a évolué fin 2018 avec i) la nomination d'un directeur adjoint pour les affaires générales (actuellement chef du SAFTI), ii) la fusion des deux unités d'épidémiologie pour constituer l'unité EPISABE et iii) la fusion du laboratoire de Niort avec le laboratoire de Ploufragan-Plouzané pour constituer un laboratoire unique, Ploufragan-Plouzané-Niort, au sein duquel l'équipe niortaise constitue une nouvelle unité dénommée « Pathologie et Bien Etre des Ruminants ». Un travail de mise en cohérence des systèmes qualité préexistants est en cours depuis novembre 2018 et a traité en priorité les chantiers prioritaires définis par la Direction du laboratoire, le responsable du SMQ, le chef d'unité et la correspondante qualité de l'unité PBER (gestion documentaire, gestion des travaux non-conformes, audits internes, revue de direction, fiches personnelles). Les évolutions résultant de ce travail de mise en cohérence seront décrites dans notre système documentaire à mesure qu'elles deviendront

## Déclaration de la politique de management du risque du laboratoire de Ploufragan- Plouzané-Niort

L'Anses laboratoire de Ploufragan-Plouzané-Niort utilise des micro-organismes pathogènes pour l'Homme et/ou les animaux pour mener à bien ses missions de référence et de recherche en sécurité des aliments et en santé animale. Parmi ces micro-organismes pathogènes, certains, détenus sur le site de Ploufragan, sont soumis à la réglementation Micro-Organismes et Toxines. Le laboratoire de Ploufragan-Plouzané-Niort doit donc satisfaire à ces obligations, en particulier celles de l'arrêté dit des « Bonnes Pratiques » du 11 juin 2013. Le présent document intitulé « Déclaration de la politique de management du risque du laboratoire de Ploufragan-Plouzané-Niort » a pour objectif de définir les lignes directrices et l'organisation documentaire en vue de garantir la sécurité et la sûreté biologique et de maîtriser les risques liés aux activités du laboratoire de Ploufragan-Plouzané-Niort.

La politique de sécurité et sûreté biologique du Laboratoire de Ploufragan-Plouzané-Niort est sous ma responsabilité, mon rôle étant de fixer les objectifs à atteindre et les moyens nécessaires à leurs mises en œuvre. Les objectifs pour cette année sont de pérenniser et poursuivre la mise en place des mesures exigées dans l'arrêté « bonnes pratiques » de manière à respecter les engagements pris auprès de l'ANSM dans le cadre des autorisations MOT détenues au laboratoire de Ploufragan-Plouzané-Niort.

Je délègue la mise en œuvre de cette politique à (aux) :

- Responsables du management du risque, placés directement sous mon autorité;
- Chefs d'unité concernés par des activités à risque biologique ou toxique ;
- Responsables de secteurs confinés de niveau biologique 2 et 3 (laboratoire et animalerie).

Chacun d'eux me rapporte directement toutes les difficultés ou problèmes rencontrés au quotidien. Ils communiquent aux personnels du laboratoire de Ploufragan-Plouzané-Niort la partie de la politique de sécurité et sûreté biologique qui leur est déléguée.

Pour mener à bien cette politique, un système de management du risque est mis en place dans les unités concernées, par les responsables du management du risque. Ce management du risque est organisé en lien avec le système de management de la qualité en place sur le site de Ploufragan-Plouzané-Niort et répondant aux exigences de la norme ISO 17 025 « Exigences générales des laboratoires d'analyse et d'essai » pour lesquelles le laboratoire de Ploufragan-Plouzané-Niort est accrédité par le COFRAC. Les personnels du laboratoire de Ploufragan-Plouzané-Niort s'engagent au respect strict des modes de fonctionnement et de bonnes pratiques appliqués à leur secteur de travail, et peuvent s'impliquer sans limite dans l'élaboration des documents relatifs à la sûreté et à la sécurité biologique.

En tant que directeur du laboratoire de Ploufragan-Plouzané-Niort, je m'engage à mettre en œuvre tous les moyens nécessaires pour :

- la protection du personnel, des prestataires, des visiteurs et de l'environnement vis-à-vis d'une exposition aux MOT stockés et/ou manipulés
- la réduction du risque à un niveau acceptable de dispersion intentionnelle
- la conformité aux exigences réglementaires applicables
- l'information du personnel et des tierces parties concernées
- la communication des obligations individuelles de chacun
- l'amélioration continue du système

Ploufragan, le 22 février 2019

Nicolas Etterradossi  
Directeur du Laboratoire de Ploufragan-Plouzané-Niort

  
Nicolas ETERRADOSSI

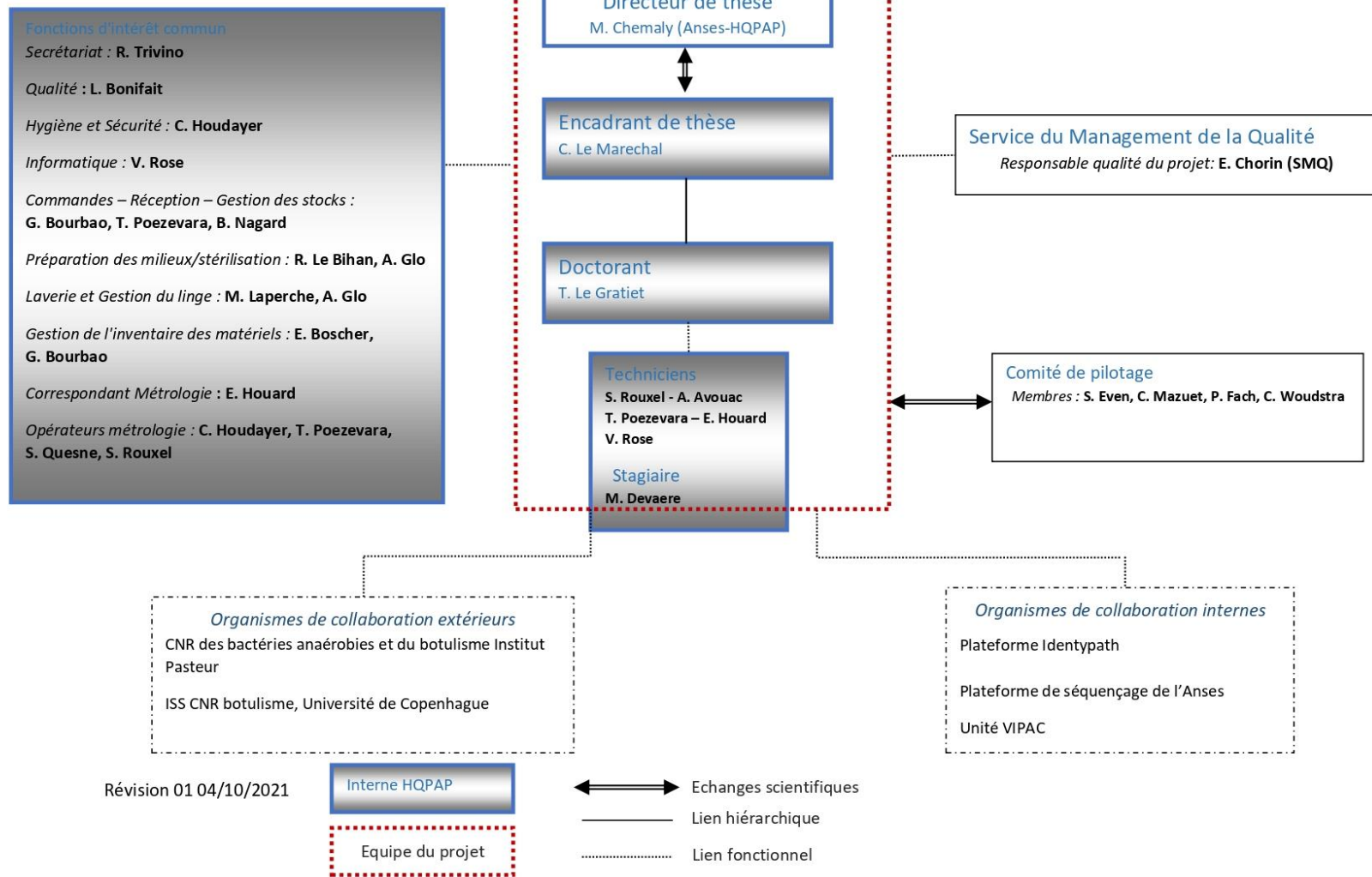


Annexe 3 : Tableau prévisionnel du déroulement de la thèse.

	Année 1												Année 2												Année 3												Année 4		
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3
Tâche 1 : Isolement des souches	[Yellow bar]												[Yellow bar]												[Yellow bar]												[Yellow bar]		
Tâche 2 : MLVA/MLVA-HRM	[Blue bar]												[Blue bar]												[Blue bar]												[Blue bar]		
Tâche 3 : Investigation du plasmidome	[Orange bar]												[Orange bar]												[Orange bar]												[Orange bar]		
Tâche 4 : Identification de cibles via analyse comparative des données WGS	[Green bar]												[Green bar]												[Green bar]												[Green bar]		
Tâche 5 : Métagénomique	[Red bar]												[Red bar]												[Red bar]												[Red bar]		
Tâche 6 : Rédaction de la thèse	[Purple bar]												[Purple bar]												[Purple bar]												[Purple bar]		

Révision 01 04/10/2021

Annexe 5 : organigramme de la thèse



Annexe 6 : Plan du bâtiment F14 (PPN/PS/0194)

Niveau 1 :



- Partie « Laboratoire de bactériologie »
- Salles concernées par l'accréditation

21

Niveau 3 :



- Partie « laboratoire de biologie moléculaire »
- Partie « bureaux »
- Salles concernées par l'accréditation

22

Laboratoire de Ploufragan-Plouzané-Niort

Service Management Qualité

Thèse préparée par Thibault LE GRATIET

**« Typage des souches de *Clostridium botulinum* du groupe III : développement d'outils et application à l'épidémiologie du botulisme animal. »**

**Attestation de management qualité de la thèse  
Par Eric CHORIN, chef du Service de Management Qualité**

Les travaux de thèse conduits par Thibault LE GRATIET se sont déroulés sous management de la qualité, dans l'unité « Hygiène et Qualité des Produits Avicoles et Porcins » (U HQPAP) à compter de Janvier 2019. Ce qui s'est concrétisé par la rédaction, dès le début des travaux et par le doctorant, d'un Plan de Management Qualité (PMQ) signé par lui-même, la Directrice de thèse, la Co-Encadrante de thèse et le Responsable qualité du Laboratoire.

Véritable engagement pour le management de la thèse, ce PMQ a deux normes pour référence :

- la norme NF EN ISO/CEI 17025 (Prescriptions générales concernant la compétence des laboratoires d'étalonnages et d'essais) appliquée dans ce projet, pour la gestion des ressources au sein de l'unité HQPAP ;
- la norme ISO 10006 (Systèmes de management de la qualité – lignes directrices pour le management de la qualité dans les projets) mise en œuvre pour la conduite du projet de thèse.

Le PMQ fait systématiquement le lien entre le sujet de la thèse, sa mise en œuvre et la formalisation de son déroulement. Avec ce document, ce sont des éléments d'organisation et de gestion des travaux de thèse qui ont été réfléchis, décidés et formalisés. Ce plan comprend notamment l'identification des « points critiques » dans le projet de thèse, les données de sortie attendues à chaque étape identifiée et des « revues de projet » conformément à la norme ISO 10006.

La mise en œuvre du PMQ

Comme pour toute démarche de management qualité, la conduite de la thèse s'est traduite par un suivi régulier de l'encadrement avec validation des travaux du doctorant. De nombreuses données ont été enregistrées conformément aux engagements en termes de traçabilité.

Le management de projet a donné lieu à des vérifications de l'application des dispositions d'organisation formalisées dans le PMQ. Cela s'est notamment traduit par deux audits internes de management de projet conduit par le Chef du Service Management Qualité les 10 Février 2020 et 07 Octobre 2021. Les principaux documents examinés lors de ces audits sont listés en annexe de cette attestation. Au cours de ces audits, aucune fiche de non-conformité n'a été notifiée : les dispositions du PMQ ont donc été parfaitement suivies par le doctorant.

Les exercices de traçabilité faits par l'auditeur se sont révélés conformes aux engagements de l'équipe de thèse. Plusieurs cahiers de laboratoires normalisés ont été utilisés, y compris pour les réunions de l'équipe de thèse. Les cahiers examinés sont extrêmement bien tenus et visés régulièrement par

l'encadrement. Le doctorant a aussi stocké ses nombreuses données dans un dossier électronique très bien structuré permettant de retrouver aisément l'ensemble des données brutes et des données analysées.

Les réunions prévues d'avancement du projet et de suivi de la doctorante ont effectivement été réalisées, les comptes rendus sont disponibles et permettent d'identifier le but de la réunion, les données d'entrée et de sortie, la validation des différentes étapes du projet, les éventuels besoins de réorientation des recherches...

Les audits ont permis d'identifier :

- le gros travail réalisé par l'équipe et en particulier par le doctorant avec 2 publications en 1<sup>er</sup> auteur et une troisième en tant que co-auteur;
- la qualité de la traçabilité des travaux conduits ;
- la rigueur du cheminement intellectuel du doctorant dans la conduite de son projet.

#### Bilan du management qualité de la thèse

Le déroulement de la thèse sous management qualité, a d'abord permis à Thibault LE GRATIET de découvrir deux référentiels qualité qui, dans l'esprit au moins, lui seront nécessaires pour sa carrière professionnelle et directement bénéfiques pour des fonctions éventuelles au sein d'un laboratoire ou dans le management d'un projet.

En second lieu, ce management qualité a dû lui être utile sur plusieurs points :

- la formalisation la plus complète possible du projet scientifique, en particulier au démarrage ;
- l'importance de clarifier les rôles respectifs des différents acteurs du projet mais aussi de s'intégrer dans leur agenda respectif ;
- l'intérêt de réfléchir, au moins au début du projet, aux éventuels points critiques et les actions préventives nécessaires pour les maîtriser. Dans le cas présent, cet aspect « point critique » a été parfaitement géré ;
- la nécessité de mettre en œuvre la traçabilité des ressources (équipements principaux suivis et vérifiés) afin de pouvoir analyser dans les meilleures conditions tout résultat inattendu ;
- l'importance de formaliser tous les résultats et les cheminements intellectuels.

Au final, je suis convaincu que Thibault LE GRATIET a su mener avec une grande rigueur son projet scientifique en ayant partagé dès le début de sa thèse avec son encadrement la démarche de management qualité.

Fait à Ploufragan le 28/03/2022



Eric CHORIN  
Chef du Service Management Qualité

-----

Annexe :

**Principaux documents examinés lors de l'audit du 10 Février 2020 :**

Il s'agit du 1<sup>er</sup> audit pour cette thèse. Quelques coquilles seront à corriger dans le rapport. La liste des points critiques du projet sera à établir.

Examen des documents du comité de thèse du 19/09/2019.

Formation Phylogénie à Maisons-Alfort (2 jours).

Formation éthique et recherche par l'école doctorale à Rennes.

Formation aux méthodes MLVA et MLST en Italie.

Examen du cahier des réunions X093845 : le cahier est bien tenu et visé régulièrement par Caroline.

30 souches ont été isolées pendant la 1<sup>ere</sup> année de la thèse. Cahier de labo X093843 : examen des manipulations. Examen d'un isolement du début à la fin. Les travaux ont permis de rédiger PPN/FSE/1206 : Isolement des souches de *C. botulinum* du groupe III et de réviser le MO.HQPAP.ESS.MA28 : isolement d'une souche de *C. botulinum* et mise en conservation. Examen de l'isollements du 14/10 /2019 et des PCR qui ont suivi. Les données sont correctement notées dans le cahier de manip. Les données informatiques sont enregistrées dans le dossier sécurisé du LNR botulisme. La nomenclature des fichiers informatique est respectée.

Examen de l'habilitation MOT : habilitation en date du 21/02/2019 suite à la lecture des documents le concernant et la réalisation de 8 échantillons connus (3 positifs sur 8) de façon satisfaisante. Le dossier d'habilitation est complet.

Les travaux avancent selon le rythme prévu. La 1<sup>ere</sup> publication est parue : Development of an innovative and quick method for the isolation of *Clostridium botulinum* strains involved in avian botulism outbreaks publiée dans Toxins le 10/01/2020.

Il n'y a pas eu de non confirmation ni de non-conformité ni de dérogation.

**Principaux documents examinés lors de l'audit du 07 Octobre 2021 :**

Traçabilité sur les manipulations MLVA dans le cadre d'un épisode de botulisme bovin et aviaire. Examen des résultats de PCR VNTR (nombre variable de répétitions en tandem), d'électrophorèse capillaire échantillon 18LNRB 07.142 (écouvillon cloacal de volaille) et de vérification du nombre de répétitions de l'amplicon par séquençage : on trouve bien la même séquence 4 fois comme attendu. On peut donc confirmer le profil MLVA pour la souche 142. Les données brutes et analysées sont stockées de façon sûre sur le réseau. Les cahiers de manip sont bien tenus et validés régulièrement. Une fiche de non-conformité a été rédigée dans Ennov.

Deux publiés acceptés en 1<sup>er</sup> auteur et une troisième en attente en co-auteur. Deux protocoles ont été rédigés sous forme de mode-opératoires et seront intégrés au système de gestion de la documentation sous Ennov.

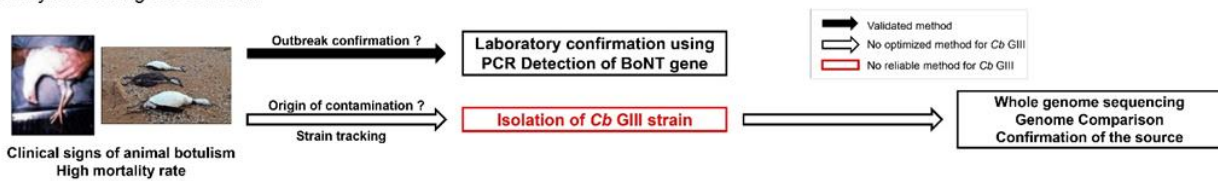
## Isolation of *Clostridium botulinum* strains involved in animal botulism outbreaks: a new rapid and reliable method



LE GRATIET Thibault, ROUXEL Sandra, POEZEVARA Typhaine, HOUARD Emmanuelle, CHEMALY Marianne, LE MARECHAL Caroline  
ANSES, Laboratory of Ploufragan-Plouzané-Niort, Unit of Hygiene and Quality of Poultry and Pork Products, France  
✉ thibault.legratiet@anses.fr

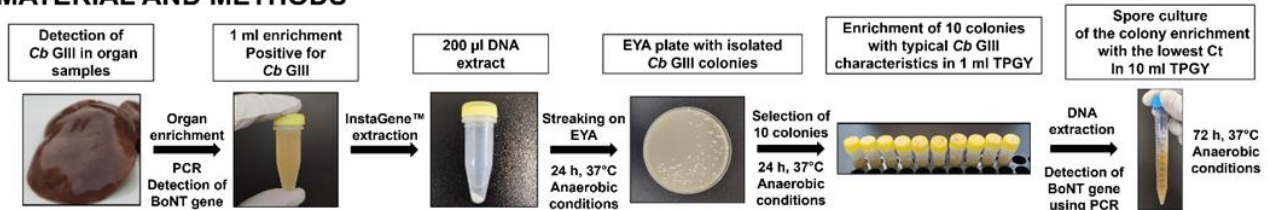
### INTRODUCTION

Animal botulism is a severe neuroparalytic disease mainly caused by the action of botulinum neurotoxins (BoNTs) C, D and mosaic variants C/D and D/C, produced by *Clostridium botulinum* Group III (*Cb* GIII) strains. This disease is a major environmental and economic concern due to the high mortality rate during an outbreak.



In order to improve the understanding of animal botulism, isolation of *Cb* GIII is required so as to genotype strains and be able to track down the contamination. However, isolation of *Cb* GIII is very challenging and no reliable method is available. There is indeed no selective medium for *Cb* GIII and the phage carrying the BoNT gene is commonly lost in laboratory conditions during isolation process. The objective of our study was to develop a new rapid and user-friendly method for the isolation of *Cb* GIII.

### MATERIAL AND METHODS



This study relies on the ability of *Cb* GIII to survive after DNA extraction with InstaGene™ kit because of its capacity to sporulate. Indeed, viable spores remain in the extract pellet. After streaking of the extract pellet on Egg Yolk Agar plate, colonies are cultured in 1 ml TPGY medium during 24 hours and a PCR is performed to detect BoNTs characteristic genes for each colony enrichment. Enrichment with the lowest Ct for these genes is added in the strain collection until further genomic analysis such as whole genome sequencing.

Naturally contaminated samples were selected so as to test samples representative of avian species (see Table below) and French regions (see map below in yellow) concerned by botulism outbreaks in poultry farms and wildlife. Broiler livers were spiked with 5 spores/g of *C. botulinum* type C, D, C/D, and D/C reference strains.

### RESULTS AND DISCUSSION

This isolation method was used to isolate 21 *Cb* GIII strains from samples collected during avian botulism outbreaks, with a 95% success rate. This is higher than previously reported in other studies. Furthermore, mouse bioassay showed that 66% of the isolated strains were still toxic after isolation. The method was successfully used to re-isolate the strains from spiked livers demonstrating the relevance of our protocol to isolate *Cb* GIII whatever the BoNT type.



Strain N°*	Origin	Matrix	BoNT	Mouse bioassay
B19LNR801	Poultry	Liver	C/D	+
B19LNR805	Duck	Liver	C/D	+
B19LNR806	Goose	Caeca	C/D	+
B19LNR807	Duck	Liver	C/D	+
B19LNR808	Heron	Liver	C/D	+
B19LNR809	Chicken	Liver	C/D	+
B19LNR810	Duck	Liver	C/D	+
B19LNR811	Turkey	Liver	C/D	+
B19LNR812	Duck	Liver	C/D	+
B19LNR813	Duck	Liver	C/D	+
B19LNR814	Duck	Liver	C/D	+
B19LNR824	Turkey	Liver	D/C	+
B19LNR825	Duck	Liver	C/D	+
B19LNR826	Turkey	Cloacal swab	D	+
N19LNR801	Chicken	Liver	C/D	NT*
N19LNR802	Turkey	Liver	C/D	NT*
N19LNR803	Chicken	Liver	C/D	NT*
N19LNR805	Turkey	Liver	D	NT*
N19LNR806	Turkey	Liver	C/D	NT*
N19LNR807	Turkey	Liver	D	NT*
N19LNR808	Turkey	Liver	D	NT*

\*: B=CbGIII; N= non toxic CbGIII; NT: Not Tested

### CONCLUSION AND PERSPECTIVES

This study provides a reliable, cheap and easy to use method to isolate *Clostridium botulinum* Group III strains. This breakthrough opens up new avenues to explore deeper into *Clostridium botulinum* Group III genomics and in a more general way to improve the understanding of the epidemiology of animal botulism.

<https://doi.org/10.3390/toxins12010042>

This poster received the support of the MedVetNet association and of the « Conseil départemental des Côtes d'Armor »



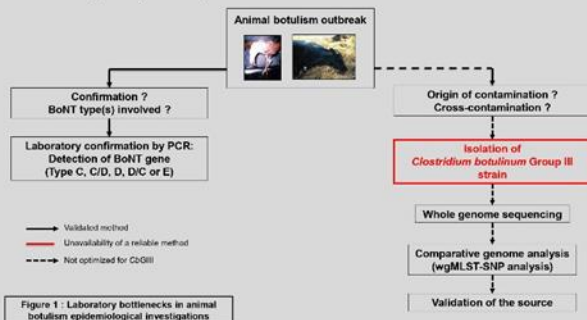
March 18-19, 2020

## Development of a reliable Method for the Isolation of *Clostridium botulinum* Group III Strains

LE GRATIET Thibault, ROUXEL Sandra, POEZEVARA Typhaine, HOUARD Emmanuelle, MAZUET Christelle<sup>1</sup>, CHEMALY Marianne, LE MARECHAL Caroline - Anses Ploufragan-Plouzané-Niort, Unité Hygiène et Qualité des Produits Avicoles et Porcins.

### STUDY CONTEXT

Animal botulism, a severe neuroparalytic disease, is due to the action of the botulinum neurotoxin (BoNT), mostly BoNT C/D, D/C, C and D, produced by *Clostridium botulinum* Group III (CbGIII).



Isolation of CbGIII strains is very challenging and no reliable method is available although this step is essential for genotyping and strain tracking. This study aims at developing an innovative method of CbGIII isolation.

### RESULTS AND DISCUSSION

In 2019, 21 strains including 65% toxigenic ones were isolated using this approach from 22 samples collected during avian botulism outbreaks, showing a 95% isolation success rate, much higher than previously reported methods.

Strain N°*	Origin	BoNT (initial sample)	Department N°	BoNT (strain)	Mouse bioassay
B19LNR01	Poultry	C / D	Belgium	C / D	+
B19LNR05	Duck	C / D	10	C / D	+
B19LNR06	Goose	C / D	1	C / D	+
B19LNR07	Duck	C / D	10	C / D	+
B19LNR08	Heron	C / D	71	C / D	+
B19LNR09	Chicken	C / D	3	C / D	+
B19LNR10	Duck	C / D	85	C / D	+
B19LNR11	Turkey	C / D	35	C / D	+
B19LNR12	Duck	C / D	2	C / D	+
B19LNR13	Duck	C / D	85	C / D	+
B19LNR14	Duck	C / D	1	C / D	+
B19LNR24	Turkey	D / C	41	D / C	+
B19LNR25	Duck	C / D	85	C / D	+
B19LNR26	Turkey	D	79	D	+
B19LNR01	Chicken	C / D	29	/	NT*
B19LNR02	Turkey	C / D	35	/	NT*
B19LNR03	Chicken	C / D	29	/	NT*
B19LNR05	Turkey	D	29	/	NT*
B19LNR06	Turkey	C / D	22	/	NT*
B19LNR07	Turkey	D	22	/	NT*
B19LNR08	Turkey	D	29	/	NT*
/	Chicken	C / D	3	/	NT*

\*: B=CbGIII; N= non toxic CbGIII; NT : Not Tested

Figure 3 : List of strains isolated in 2019

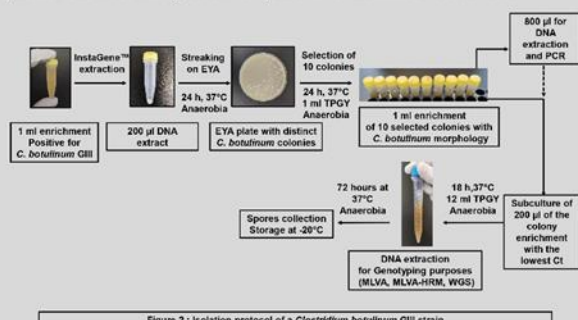
This method was also successfully used to re-isolate CbGIII from samples spiked with type C, D, C/D and D/C Cb.

Le Gratiel, T.; Poezevara, T.; Rouxel, S.; Houard, E.; Mazuet, C.; Chemaly, M.; Le Maréchal, C.

Development of An Innovative and Quick Method for the Isolation of *Clostridium botulinum* Strains Involved in Avian Botulism Outbreaks. *Toxins* 2020, 12, 42.

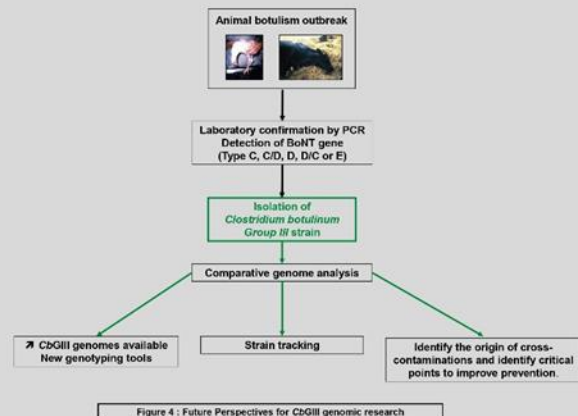
### MATERIAL AND METHODS

This method is based on the ability of CbGIII spores to survive to InstaGene™ matrix DNA extraction. Streaking of the final DNA extract suspension on Egg Yolk Agar (EYA) plates allows the growth of pure and isolated colonies.



Colonies are subcultured and *ntnh* gene (part of the BoNT phage) is detected using PCR to confirm the isolation of CbGIII strain. A final subculture of the newly isolated strain is performed for genotyping and storage purposes.

### CONCLUSION AND PERSPECTIVES



Improvement of the management and prevention of animal botulism outbreaks requires a better understanding of its epidemiology, in particular the characterization of CbGIII. Absence of a reliable method for the isolation of CbGIII results in a small number of published genomes, preventing from conducting epidemiological investigations. The availability of a rapid and efficient method opens new avenues in this field.



## USE OF MULTI-LOCUS VARIABLE NUMBER OF TANDEM REPEATS ANALYSIS (MLVA) TO IDENTIFY THE SOURCE OF CONTAMINATION OF A CATTLE BOTULISM OUTBREAK

Thibault Le Gratiot <sup>a\*</sup>, Fabrizio Annibaldi<sup>b</sup>, Rozenn Souillard<sup>a</sup>, Daniel Grosjean<sup>c</sup>, Typhaine Poezevara<sup>a</sup>, Sandra Rouxel<sup>a</sup>, Loïc Balaine<sup>a</sup>, Sabrina Macé<sup>a</sup>, Laure Martin<sup>a</sup>, Sophie Le Bouquin<sup>a</sup>, Marianne Chemaly<sup>a</sup>, Caroline Le Maréchal<sup>a</sup>

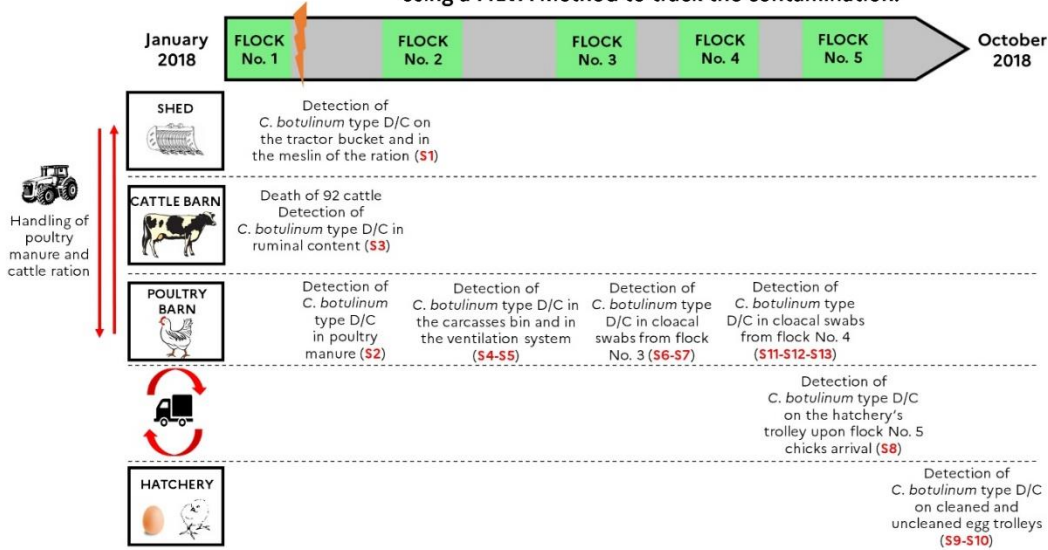
<sup>a</sup> : ANSES, French Agency for Food, Environmental and Occupational Health Safety, Ploufragan, France  
<sup>b</sup> : National Reference Centre for Botulism, Istituto Superiore di Sanità, 00161 Roma, Italy  
<sup>c</sup> : DDCSPP de la Meuse, Departmental authority in charge of veterinary services for Meuse department, France  
 \*thibault.legratiot@anses.fr

### INTRODUCTION

Animal botulism is a severe disease caused by the botulinum neurotoxins C, D, C/D and D/C produced by *Clostridium botulinum*.

In 2018, 92 cows died of type D/C botulism in a mixed farm of both cattle and poultry productions. *C. botulinum* type D/C was detected in the poultry carcass bin, poultry manure, cattle ration and tractor bucket as well as in successive poultry flocks after the cattle outbreak and in the hatchery delivering the chicks to the farm.

➤ The objective of our study was to compare a selection of positive samples collected during epidemiological investigations using a MLVA method to track the contamination.



**Figure 1 :** Diagram of the botulism outbreak and collection of samples over time.

### MATERIAL AND METHODS

Thirteen samples positive for *C. botulinum* type D/C were selected for MLVA (Figure 1).

Briefly, the MLVA method consists of the amplification of nine specific Variable Number of Tandem Repeats (VNTR) regions followed by the determination of the number of repeats for each target by capillary electrophoresis and sequencing.

A MLVA code was determined for each sample and a dendrogram was built using BioNumerics 7.5 software (Figure 2).

### RESULTS & DISCUSSION

As illustrated in the dendrogram (Figure 2), 10/13 samples clustered together demonstrating that the same profile was detected.

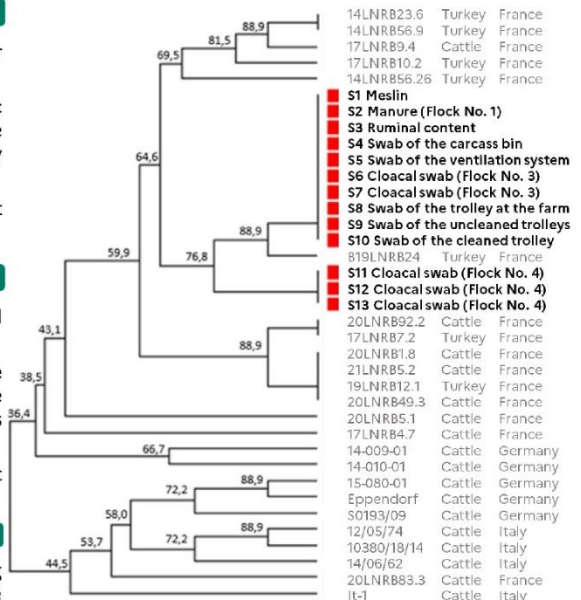
Profiles detected in the poultry manure (Flock No. 1) and in one of the affected cow were the same ones, confirming that poultry manure was the source of contamination. This MLVA comparison also showed hatchery was a reservoir of this *C. botulinum* type D/C profile.

A second profile was detected in poultry Flock No. 4. Its origin was not identified during the investigations.

### CONCLUSION

MLVA was successfully used to compare positive samples collected during this botulism outbreak. Coupled with epidemiological investigations in the farm, this typing tool confirmed that **poultry manure was the source of *C. botulinum* contamination.**

This example shows that **MLVA is a very powerful tool for *C. botulinum* tracking in the context of animal botulism outbreaks.**



**Figure 2 :** Dendrogram representing the 13 samples selected from the cattle outbreak and 24 other samples positive for *C. botulinum* type D/C.

A short term mission grant was attributed by the Med Vet Net association for a MLVA training at the Istituto Superiore di Sanità (ISS) in Roma, ITALY

## INTRODUCTION

- Le botulisme animal est une maladie neuroparalytique sévère causée par l'action de toxines botulique (BoNT) produites par les souches de *Clostridium botulinum* du groupe III qui appartiennent au groupe d'espèces *Clostridium novyi sensu lato* (groupe qui comprend aussi les souches de *C. novyi* et *C. haemolyticum*).
- La diversité génomique de *C. novyi sensu lato* est en particulier due aux éléments génétiques mobiles (EGM) très variés d'une souche à une autre. Un des mécanismes de défense contre l'invasion d'EGM chez les bactéries est le système CRISPR-Cas qui dégrade l'ADN étranger. Plusieurs types de système existent et sont définis selon leur répertoire de gènes cas.

**Objectif: Explorer les systèmes CRISPR-Cas chez *C. novyi sensu lato* afin de mieux comprendre la dynamique de son mobilome.**

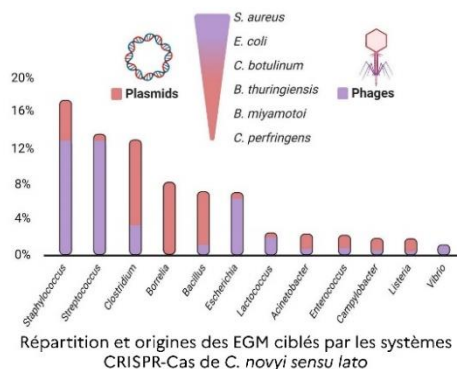
## MATÉRIELS ET MÉTHODES



## RÉSULTATS

### Identification et caractéristiques des systèmes CRISPR-Cas

- 6 types de systèmes CRISPR-Cas identifiés
- Retrouvés dans 55 des 58 génomes étudiés
- Type I-B prédominant suivi par I-D, V-U et II-C
- Présents sur le chromosome et sur les EGM
- Systèmes incomplets sur les EGM



Type de système CRISPR-Cas	<i>C. botulinum</i> BoNT C (n=9)	<i>C. botulinum</i> BoNT C/D (n=22)	<i>C. botulinum</i> BoNT D (n=3)	<i>C. botulinum</i> BoNT D/C (n=10)	<i>C. novyi</i> (n=10)	<i>C. haemolyticum</i> (n=3)
I-B	100%	100%	67%	40%	70%	100%
I-D			33%	90%	20%	
III-B				40%	10%	
III-D					10%	
II-C				60%		
V-U	89%				20%	33%

Répartition des systèmes CRISPR-Cas retrouvés chez 58 génomes de *C. novyi sensu lato*

### Analyse des spacers

- 4320 spacers extraits au cours de l'étude, 2045 uniques
- Contenu en spacers similaire pour les souches génétiquement proches

### Origine des EGM ciblés par les systèmes CRISPR-Cas

- 434 spacers identifiés – 1508 matches sur BLAST – 776 EGM uniques
- Répartition des EGM ciblés → 52,2% phages - 48,8% plasmides
- 66,8% des EGM ciblés → *Staphylococcus*, *Streptococcus*, *Clostridium*, *Bacillus*, *Escherichia* et *Borrelia*
- 6 EGM de *C. novyi sensu lato* = 20,2% des résultats de BLAST

## DISCUSSION ET CONCLUSION

Cette étude montre:

- Une forte prévalence et diversité des systèmes CRISPR-Cas dans le groupe *C. novyi sensu lato* avec une prédominance du type I-B, système également prévalent chez d'autres clostridies. Des études complémentaires sont nécessaires pour comprendre le rôle des systèmes incomplets retrouvés notamment au niveau des EGM.
- Les similarités retrouvées au niveau du contenu en spacers des souches génétiquement proches ouvrent de nouvelles perspectives dans le développement d'outils de typage des souches de *C. novyi sensu lato*.
- La forte proportion de spacers ciblant 6 EGM de *C. novyi sensu lato* suggèrent que ces éléments sont très dynamiques au sein de ce groupe d'espèces.

Annexe 3 : Fiche de suivi qualité du protocole d'isolement des souches de *C. botulinum* du groupe III.



**UNITE HQPAP : FICHE D'ISOLEMENT DE SOUCHES de  
C. BOTULINUM DU GROUPE III**

Manipulateur : \_\_\_\_\_

Code échantillon : \_\_\_\_\_

**Jour 1 : Remise en culture de l'enrichissement**

- Mettre du TPGY à régénérer puis le laisser refroidir. **Code TPGY :** \_\_\_\_\_
- Sortir le tube contenant l'enrichissement et le laisser décongeler (environ 20 minutes)
- Allumer le PSM et laisser le flux se mettre en place (15 minutes).
- Préparer sous le PSM le matériel nécessaire à la mise culture (sac de destruction, pipettes, cônes ...)
- Mettre en culture la souche à isoler. Avec une P1000, ensemercer 1 ml de l'enrichissement au fond du tube Falcon dans 9 ml de TPGY sans toucher les parois. Noter la sortie du tube dans l'ENR0101 (PPN/FSE/0908).

**Code tube Falcon :** \_\_\_\_\_

- Mettre un masque, une paire de gants et une charlotte en plus de la blouse. Une fois à l'intérieur de l'enceinte, mettre une nouvelle paire de gants.
- Incuber pendant 48h minimum dans l'enceinte anaérobie à 37°C.

**Date début incubation :** \_\_\_\_\_ **Heure :** \_\_\_\_\_

**Date fin incubation :** \_\_\_\_\_ **Heure :** \_\_\_\_\_

**Jour 2 : Incubation des EYA en conditions anaérobies**

- Incuber les EYA dans l'enceinte **Code EYA :** \_\_\_\_\_

**Jour 3 : Extraction Instagene**

**Date :** \_\_\_\_\_

**Code Instagene :** \_\_\_\_\_

- Si non incubé la veille, incuber les EYA dans l'enceinte (Minimum 4 h mais 24 h idéal, 2 boîtes).

**Code EYA :** \_\_\_\_\_

- Allumer le PSM et laisser le flux se mettre en place (15 minutes).
- Régler le bain sec à 80°C et le bain marie à 56°C.
- Préparer le matériel nécessaire à l'extraction Instagene sous le PSM (sac de destruction, vortex, pipettes...)
- Mettre un masque, une paire de gants et une charlotte en plus de la blouse. Une fois à l'intérieur de l'enceinte, mettre une nouvelle paire de gants.
- Fermer le tube d'enrichissement avant de le placer dans le sas. Nettoyer l'enceinte avec de l'Anioxyspray WS.
- Placer l'enrichissement sous le PSM.
- Avec une pipette 2 mL, après une délicate homogénéisation (1 pipetage aller-retour), prélever 1 ml de l'enrichissement dans un tube à vis pour l'extraction puis 1 ml pour le stockage au congélateur.
- Procéder à l'extraction Instagene en remplissant la fiche de suivi (ENR0111 - PPN/FSE/0979).
- Conserver 50 µL de l'extrait pour PCR (*Clostridium novyi sensu lato* et du gène ntnh).
- Vortexer le reste de l'extrait obtenu pendant 10 secondes.
- Dans l'enceinte, ensemercer 2 boîtes d'EYA avec chacune 100 µl de l'extrait d'ADN (1 boîte d'EYA incubé dans l'enceinte et 1 boîte d'EYA incubé dans l'enceinte dans un sachet Gaz Pack). Conserver le reste de l'extrait dans le congélateur.

**Code sachet +Gaz Pack :** \_\_\_\_\_

- Incuber pendant 48 h dans l'enceinte anaérobie à 37°C.
- Lecture de la boîte à 24 h et à 48 h.

## UNITE HQPAP : FICHE D'ISOLEMENT DE SOUCHES de C. BOTULINUM DU GROUPE III

**Jour 4 ou 5 : Repiquage des colonies**      Date J4 : \_\_\_\_\_      Date J5 : \_\_\_\_\_

J4    J5

- Régénérer du TPGY puis faire refroidir. Code TPGY : \_\_\_\_\_
- Aliquoter le TPGY à raison de 1,5 mL par tube Lock 2 ml (10 à 15 tubes par enrichissement à isoler)
- Code tube 2ml :**
- Mettre un masque, une paire de gants et une charlotte en plus de la blouse. Une fois à l'intérieur de l'enceinte, mettre une nouvelle paire de gants.
- Après lecture des boîtes, sélectionner une dizaine de colonies isolées dont la morphologie correspond à *C. botulinum*
- Repiquer les colonies dans les tubes 2 mL contenant 1,5 ml de milieu TPGY.
- Incuber 18 h minimum en enceinte anaérobie à 37°C.

J4 : Date mise en incubation : \_\_\_\_\_ Heure : \_\_\_\_\_ Date de sortie : \_\_\_\_\_ Heure : \_\_\_\_\_

J5 : Date mise en incubation : \_\_\_\_\_ Heure : \_\_\_\_\_ Date de sortie : \_\_\_\_\_ Heure : \_\_\_\_\_

### Jour 5 ou 6 : Extraction Instagene et PCR temps réel pour la détection de *C. botulinum*

J5    J6

- Allumer le PSM et laisser le flux se mettre en place (15 minutes).
- Régler le bain sec à 80°C et le bain marie à 56 °C
- Préparer le matériel nécessaire à l'extraction InstaGene pour le PSM (sac de destruction, vortex, pipettes...)
- Mettre un masque, une paire de gants et une charlotte en plus de la blouse. Une fois à l'intérieur de l'enceinte, mettre une nouvelle paire de gants.
- Après une délicate homogénéisation (1 pipetage aller-retour), prélever 800 µl de chaque enrichissement de colonie avec une P1000 dans un nouveau tube à vis, sortir les 800 µl à extraire de la station anaérobie, laisser les 200 µL restants dans l'enceinte en attente de résultat. Nettoyer l'enceinte avec de l'Anioxyspray WS.
- Placer les enrichissements sous le PSM.
- Procéder à l'extraction Instagene en remplissant la fiche de suivi (ENR0111 - PPN/FSE/0979).
- Procéder à la PCR temps réel en se référant à la fiche de suivi PCR (ENR0105 - PPN/FSE/1007) en ciblant GIII et CN.
- Régénérer du TPGY puis le faire refroidir. Code TPGY : \_\_\_\_\_ Code Falcon : \_\_\_\_\_
- Après lecture des résultats de PCR, repiquage de l'enrichissement dont le Ct est le plus précoce. Avec une P1000, prélever les 200 µl d'enrichissement conservé dans l'enceinte et les réensemencer dans un tube falcon de 15 ml contenant 10 ml de milieu TPGY en vue de la mise en collection de la souche isolée.
- Incuber 18 h minimum à 37°C en anaérobie et prélever 1 mL pour extraction d'ADN génomique pour séquençage (sauf si vendredi, dans ce cas repiquer 1 ml de la culture 72 h dans 9 ml de TPGY pour extraction d'ADN génomique)
- Incuber 72 h minimum à 37°C en anaérobie.

J5 : Date mise en incubation : \_\_\_\_\_ Heure : \_\_\_\_\_ Date de sortie : \_\_\_\_\_ Heure : \_\_\_\_\_

J6 : Date mise en incubation : \_\_\_\_\_ Heure : \_\_\_\_\_ Date de sortie : \_\_\_\_\_ Heure : \_\_\_\_\_

- Mise en collection selon le mode opératoire MO.HQPAP.ESS.MA28 (PPN/INS/0703).

Annexe 4: Standard Operating Procedure on Multiple-Locus Variable number of tandem repeat Analysis (MLVA) from *Clostridium botulinum* group III cultures.



ANIBOTNET Project

Animal botulism: innovative tools for diagnosis, prevention, control and epidemiological investigation

## Standard Operating Procedure on Multiple-Locus Variable number of tandem repeat Analysis (MLVA) from *Clostridium botulinum* group III cultures

Version	Written by	Verified by	Approved by
0	Fabrizio Anniballi Bruna Auricchio Stefano Bilei Paola De Santis		
1	Fabrizio Anniballi Bruna Auricchio Concetta Scalfaro		

Version: 1	MLVA from <i>Clostridium botulinum</i> group III cultures	1
------------	---	---



## TABLE OF CONTENTS

TABLE OF CONTENTS.....	2
PURPOSE .....	3
RESPONSIBILITY.....	3
SAFETY PRECAUTIONS (by CDC Botulism manual) .....	3
APPARATUS AND EQUIPMENT.....	3
REAGENTS .....	4
LIST OF PRIMERS .....	4
PROCEDURE .....	4
WARMING .....	5
APPENDIX 1 .....	5

Version: 1	MLVA from <i>Clostridium botulinum</i> group III cultures	2
------------	---	---



**Animal botulism: innovative tools for diagnosis, prevention, control and epidemiological investigation**

---

## PURPOSE

This Standard Operating Procedure (SOP) describes the procedures for carrying out the Multiple-Locus Variable number of tandem repeat Analysis (MLVA) from *Clostridium botulinum* group III cultures.

## RESPONSIBILITY

Managers and supervisors are responsible for making sure that technicians are properly trained and equipment and facility are maintained in good working order.

Laboratory personnel are responsible for reading and understanding this SOP and related documents and to perform these tasks in accordance with the SOPs.

## SAFETY PRECAUTIONS (by CDC Botulism manual)

Although type C, D and their mosaic toxins are not correlated to human botulism, some safety precautions should be taken in account. Therefore, all material suspected of containing botulinum toxins must be handled with cautions, and only experienced personnel should perform laboratory tests.

Laboratory coats, disposable surgical gloves, face shields or bench top Plexiglas shields should be used when handling specimens. All specimen containers, culture tubes, flasks, etc., containing botulinum toxin should be placed in leak-proof, unbreakable containers during storage or incubation to prevent accidental spills. If spills occur, the toxin can be neutralized by using a strong alkaline solution such as 0.1 M sodium hydroxide. *Clostridium botulinum* is inactivated with 1:10 dilution of household bleach. The appropriate disinfectant solutions must be in contact with the toxin or organism for 15 to 20 minutes to ensure inactivation. If the material is suspected to contain both toxin and organisms, the spill must be sequentially treated with bleach and sodium hydroxide.

Prior to commencing, read and understand the Material Safety Data Sheet (MSDS) for all hazardous chemicals used in this procedure.

## APPARATUS AND EQUIPMENT

- ✓ Capillary electrophoresis apparatus (Agilent Bioanalyzer 2100).
- ✓ Centrifuge tube, of capacities of 15 ml.
- ✓ DNA Sanger sequencer.
- ✓ Graduated pipettes, capable of taking volumes between 1 and 10 ml.
- ✓ Micro-centrifuge tubes, of capacities of 1.5 ml and 2.0 ml.
- ✓ Micro-centrifuge, for 2.0 ml tubes and with adjustable acceleration of up to 14 000 x g.
- ✓ Thin-walled PCR microtubes, 0.2 ml reaction tubes or multi-well PCR microplates.
- ✓ Thermal cycler.
- ✓ Pipette filter tips, for volumes between 1µl and 1 000 µl.
- ✓ Variable volume micropipettes, capable of taking volumes between 1µl and 1 000 µl.
- ✓ Bionumerics software for VNTRs analysis.

Version: 1	MLVA from <i>Clostridium botulinum</i> group III cultures	3
------------	---	---


**Animal botulism: innovative tools for diagnosis, prevention, control and epidemiological investigation**
**REAGENTS**

- ✓ Nuclease-free water.
- ✓ Primer set for each VNTRs.
- ✓ Qiagen HotStarTaq Master mix kit (Cat. 203443) or Qiagen Multiplex PCR kit (Cat.206143).
- ✓ Reagent for PCR-product sizing (Agilent DNA 7500 kit cat. 5067-1506 or DNA12000 kit cat 5067-1508).

**LIST OF PRIMERS**

N.	Name	Left Primer	Name	Right Primer
1	MLVA_ANB_1L	CTGATGTTAGCACGAACCTTGTTC	MLVA_ANB_1R	ACATGAGTACAGTCCAAGCAAA
2	MLVA_ANB_2L	TCCCACTACTGTTTTGCAGGA	MLVA_ANB_2R	CCAGAAAATCATCAATTCGGTGT
3	MLVA_ANB_3L	ACCTTCGTTATCCCATCTTTGT	MLVA_ANB_3R	AGAAAACCACCTCAAACCGT
4	MLVA_ANB_4L	ACAAGTGCCAGATGAGTGGT	MLVA_ANB_4R	TTCCAAGTCTCCACTTTTGT
5	MLVA_ANB_5L	CTCTCTGACGATACTTCCATTGA	MLVA_ANB_5R	AAGCCCCATTATACCGTGAG
6	MLVA_ANB_6L	ACAAGGCTTAGAGAGGTTCAAA	MLVA_ANB_6R	AAGCAGTCATATTTCTCGGTAGT
7	MLVA_ANB_7L	EGATGTTACCACTTGCTAAAACA	MLVA_ANB_7R	TCGCCTCTGAAAATCTACCA
8	MLVA_ANB_8L	GAATTTAGCCAAGAACAAGGA	MLVA_ANB_8R	TCCTGCACTTAAATCAGCTACC
9	MLVA_ANB_9L	ACTACCAATTTACGTGTGCTAGT	MLVA_ANB_9R	TTGTTGTGTAGCTGGACTTAGT
10	MLVA_ANB_10L	AGCCAATGGAGGAGGAAACA	MLVA_ANB_10R	TCGACTTCATCATCCCTTGTG
11	MLVA_ANB_11L	TGTTCTCCTAGCAATGGG	MLVA_ANB_11R	CCATTCTATTCGGTCAACAACT
12	MLVA_ANB_12L	CACCACAAGTAACGACCAA	MLVA_ANB_12R	TGTACGGTTGTTATCTTTAGGA
13	MLVA_ANB_13L	CCGTTGCTAGGTAGCATTCC	MLVA_ANB_13R	TTCCCGTACCATTACCTTT
14	MLVA_ANB_14L	CATATTTCCGAGCCGTT*	MLVA_ANB_14R	TGTGTTGAAGGCATGAATCC
15	MLVA_ANB_15L	TGGATCATGTGAAAATGCAAG	MLVA_ANB_15R	CAGCTCCTGCCTTCGCTA
16	MLVA_ANB_16L	GAAGCTGTAGGGCGTAGACAA	MLVA_ANB_16R	TGTCTGGCATCAATCTGA
17	MLVA_ANB_17L	TTTCTCGTGCAATCCCTCT	MLVA_ANB_17R	GAAAATGCAAGGATGCTAAGA
18	MLVA_ANB_18L	TGTGAACCTGCCAACCTT	MLVA_ANB_18R	GGTGACATTGACAACAGC
19	MLVA_ANB_19L	TTATCCGGGTTCCATT	MLVA_ANB_19R	CAACTTGATACCGGCCCC
20	MLVA_ANB_20L	GGACGCAACCTCAGAGAATG	MLVA_ANB_20R	ACTTCTCGTCCATAGTCCA
21	MLVA_ANB_21L	CCACTCCGGCGTATTACATC	MLVA_ANB_21R	GCTATGCCTACCAGAAAGTCA
22	MLVA_ANB_22L	CTTGAAAGTCTGTCCATGCA	MLVA_ANB_22R	CATAGCTGCTCCAGGGGTTA
23	MLVA_ANB_23L	GGTTTATAGTGTAGAGGGTTTGA	MLVA_ANB_23R	AACGTGCTTCAAAGGCAGT
24	MLVA_ANB_24L	CTGTTTCTCCCTTTGTGTCA	MLVA_ANB_24R	GGGTTAAAGGAAGCGGTAG
25	MLVA_ANB_25L	ETGTACCACCAGTTGATGCA	MLVA_ANB_25R	GGGCTAGGACAGTGTATGCT
26	MLVA_ANB_26L	AGTTTCTGAATCCATACCCATCA	MLVA_ANB_26R	GGAAAAGTCGCTACATGAGAGA
27	MLVA_ANB_27L	CTACTTCTTCTGTGACCTACA	MLVA_ANB_27R	TGGGAGCTATGAGTGAAGAATTG

\*Primers reported in red provides high discrimination and are suggested for typing.

**PROCEDURE**

- ✓ Extract DNA according to the specific SOP (Standard Operating Procedure on DNA extraction from *Clostridium botulinum* group III cultures).

Version: 1	MLVA from <i>Clostridium botulinum</i> group III cultures	4
------------	---	---





**Animal botulism: innovative tools for diagnosis, prevention, control and epidemiological investigation**

- ✓ Prepare the master mix as following described, considering a volume 10% greater than that required, to compensate for pipette calibration.

Component	Volume/reaction
HotStarTaq master mix*	12.5 µl
Left primer (10 mM) - final concentration 500 nM	0.125 µl
Right primer (10 mM) - final concentration 500 nM	0.125 µl
Template DNA	3 µl
Nuclease-free water	9.25 µl
<b>TOTAL VOLUME</b>	<b>25 µl</b>

\*the protocol was optimized also using Multiplex PCR kit Qiagen.

- ✓ Dispense the master mix in the PCR wells.
- ✓ Dispense the DNA in each well.
- ✓ Run the PCR according to the following thermal profile.

Phase	Time	Temperature
<b>Initial activation step</b>	15 min	95°C
<b>3-step cycling</b>		
denaturation	30 sec	94°C
annealing	60 sec	58°C
extension	60 sec	72°C
Number of cycles 35		
<b>Final extension</b>	5 min	72°C

- ✓ Size PCR product as described in the procedure reported in Appendix 1.
- ✓ Extract the VNTR (allele) for PCR product size using the Excel macro.
- ✓ Include the MLVA profile in the project database or analyse the data using Bionumerics or other suitable software.

## WARMING

If you during the PCR product sizing process you find new alleles, these need to be confirmed by means of DNA Sanger sequencer sizing and sequencing. The present procedure does not describe the sizing and sequencing by means of DNA Sanger sequencer. Please carry out sizing and sequencing using the procedures described by the DNA Sanger sequencer producer.

## APPENDIX 1

Agilent DNA 7500 and DNA 12000 kit guide.

Version: 1	MLVA from <i>Clostridium botulinum</i> group III cultures	5
------------	---	---

### Articles scientifiques publiés dans un journal à comité de lecture

**Le Gratiet, T.**; Poezevara, T.; Rouxel, S.; Houard, E.; Mazuet, C.; Chemaly, M.; Le Maréchal, C. Development of An Innovative and Quick Method for the Isolation of *Clostridium botulinum* Strains Involved in Avian Botulism Outbreaks. *Toxins* 2020, 12, 42. <https://doi.org/10.3390/toxins12010042>.

**Le Gratiet, T.**, Le Maréchal, C., Devaere, M., Chemaly, M., & Woudstra, C. (2021). Exploration of the Diversity of Clustered Regularly Interspaced Short Palindromic Repeats-Cas Systems in *Clostridium novyi sensu lato*. *Frontiers in microbiology*, 12, 711413. <https://doi.org/10.3389/fmicb.2021.711413>.

Souillard, R., Grosjean, D., **Le Gratiet, T.**, Poezevara, T., Rouxel, S., Balaine, L., Macé, S., Martin, L., Anniballi, F., Chemaly, M., Le Bouquin, S., & Le Maréchal, C. (2021). Asymptomatic Carriage of *C. botulinum* Type D/C in Broiler Flocks as the Source of Contamination of a Massive Botulism Outbreak on a Dairy Cattle Farm. *Frontiers in microbiology*, 12, 679377. <https://doi.org/10.3389/fmicb.2021.679377>.

### Publication sans comité de lecture

**Le Gratiet, T.**, Anniballi, F., Boutant, J., Mellal, M., Rouxel, S., Poezevara, T., Houard, E., Rose, V., Briand, F., Chemaly, M., Le Maréchal, C. Caractérisation de *Clostridium botulinum* dans le cadre d'épisodes de botulisme aviaire en France : Evaluation d'un nouvel outil. Article paru dans le cadre des 14<sup>ème</sup> Journées de la Recherche Avicole et Palmipèdes à Foie gras – 9 et 10 Mars 2022.

### Publication à soumettre

**Le Gratiet, T.**, Le Maréchal, C., Chemaly, M., Comparison of five DNA extraction protocols for the sequencing of *Clostridium novyi sensu lato* strains.

### Communications orales (\*Orateur)

**Le Gratiet, T\***. Développement d'outils moléculaires pour le typage des souches de *C. botulinum* du groupe III impliquées dans le botulisme aviaire. Journée d'échanges sur le botulisme animal le 26 Novembre 2020.

**Le Gratiet, T\***, Woudstra C., Anniballi F., Felten A., Rouxel S., Poezevara T., Houard E., Avouac A., Chemaly M., Le Maréchal C. Development of new typing tools for *Clostridium botulinum* group III strains responsible for animal botulism outbreaks. Journées scientifiques de l'école doctorale EGAAL le 01 Juillet 2021.

**Le Gratiet, T\***, Woudstra C., Anniballi F., Felten A., Rouxel S., Poezevara T., Houard E., Avouac A., Chemaly M., Le Maréchal C. Development of new molecular-based tools for the investigation of animal botulism outbreaks. Journées Scientifiques et Doctorales de l'Anses le 30 Septembre 2021.

**Le Gratiet, T\***. Développement d'outils de typage pour caractériser *Clostridium botulinum* du groupe III responsable du botulisme animal. Journée Avicole et Cunicole d'Information et d'Echanges le 09 Décembre 2021.

**Le Gratiet, Thibault**, Fabrizio Anniballi, Jérémy Boutant, Mokrane Mellal, Sandra Rouxel, Typhaine Poezevara, Emmanuelle Houard, Valérie Rose, François-Xavier Briand, Marianne Chemaly, and Caroline Le Maréchal\*. "Caractérisation de *Clostridium botulinum* dans le cadre d'épisodes de botulisme aviaire en France: évaluation d'un nouvel outil." Journées de la Recherche Avicole et des palmipèdes à foie gras à Tours les 9 et 10 mars 2022.

## Communications affichées

**Le Gratiet, T., Rouxel, S., Poezevara, T., Houard, E., Mazuet, C., Chemaly, M., Le Marechal, C.,** Development of a reliable method for the isolation of *Clostridium botulinum* Group III strains. Journées Scientifiques et Doctorales de l'Anses le 18 et 19 Mars 2020.

**Le Gratiet, T., Rouxel, S., Poezevara, T., Houard, E., Chemaly, M., Le Marechal, C.** Isolation of *Clostridium botulinum* strains involved in animal botulism outbreaks: a new rapid and reliable method. World One Health Congress 30 October-3 November 2020 – Edinburgh.

**Le Gratiet, T., Anniballi, F., Souillard, R., Grosjean, D., Poezevara, T., Rouxel, S., Balaine, L., Macé, S., Martin, L., Le Bouquin, S., Chemaly, M., Le Marechal, C.** Use of Multi Loci Variable Number of Tandem Repeats Analysis (MLVA) to identify the source of contamination of a cattle botulism outbreak. One Health EJP Annual Scientific Meeting 9-11 June 2021- Copenhagen.

**Le Gratiet, T., Le Marechal, C., Devaere, M., Chemaly, M., Woudstra C.** Investigation of *Clostridium novyi sensu lato* CRISPR-Cas systems. 16<sup>ème</sup> Congrès national de la Société Française de Microbiologie du 22 au 24 Septembre 2021 – Nantes.

## Bourses

Bourse MedVetNet lauréat 2019 1300 € : pour effectuer un séjour de recherche de 10 jours à l'Istituto Superiore di Sanità (Rome-Italie).

Bourse de mobilité de l'école doctorale EGAAL lauréat 2019 400 € : pour effectuer un séjour de recherche de 10 jours à l'Istituto Superiore di Sanità (Rome-Italie).

Bourse MedVetNet lauréat 2021 1000 € : pour le congrès World One Health Congress qui devait avoir lieu à Edimbourg du 30 Novembre au 3 Décembre qui a finalement eu lieu en ligne du fait des conditions sanitaires liées au Covid 19.

## Encadrement de stage

Marie Devaere, Master 2, Janvier – Juin 2022, Co-encadrement. Sujet de stage intitulé « Développement et évaluation d'outils pour le typage des souches de *C. botulinum* du groupe III ».

**Titre:** Typage des souches de *Clostridium botulinum* du groupe III : Développement d'outils et application à l'épidémiologie du botulisme animal.

**Mots clés :** *Clostridium botulinum*, botulisme animal, outils de typage, séquençage, CRISPR-Cas, MLVA.

**Résumé :** Le botulisme animal qui touche principalement les volailles, l'avifaune sauvage et les bovins se manifeste par une paralysie flasque causée par l'action des neurotoxines botuliques. Les toxines les plus fréquemment associées aux épisodes de botulisme animal sont celles de type C, D, C/D et D/C, produites par *Clostridium botulinum* du groupe III, une bactérie anaérobie stricte sporulante, regroupée avec les espèces *Clostridium novyi* et *Clostridium haemolyticum*, sous le terme *Clostridium novyi sensu lato*.

L'épidémiologie du botulisme animal est encore méconnue et les données d'épidémiologie moléculaire relatives aux souches impliquées dans le botulisme animal sont quasi inexistantes. Ceci s'explique principalement par l'absence d'outils de typage pour caractériser les souches du groupe *C. novyi sensu lato*, liée aux difficultés pour isoler les souches de ce groupe.

L'objectif de la thèse est de développer des outils de typage adaptés à ce contexte particulier pour permettre de mener des investigations épidémiologiques lors d'épisodes de botulisme animal et pouvoir à terme tracer les souches.

Dans un premier temps, une méthode d'isolement a été développée et validée sur des souches impliquées dans des épisodes de botulisme animal en France. Cette méthode a permis d'isoler 35 souches associées à des épisodes de botulisme aviaire.

Puis un outil de typage MLVA (Multi Loci Variable number of tandem repeats analysis) a été évalué sur la collection de souches isolées puis appliqué en conditions réelles dans le cadre d'investigations menées dans deux épisodes de botulisme animal. Une autre piste de typage a ensuite été explorée via l'étude des systèmes CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats – CRISPR associated protein). Les résultats ont révélé une prévalence de ces systèmes chez 95% des souches de *C. novyi sensu lato*, avec une grande diversité (6 types de système CRISPR-Cas). Cette grande diversité n'a pas permis de proposer un outil de typage basé sur les CRISPR-Cas.

Enfin, après une première étape d'optimisation des différentes étapes en amont du séquençage, les génomes de *C. novyi sensu lato* ont été comparés afin d'identifier de nouvelles cibles pertinentes.

Ces travaux ont permis de développer une approche permettant de séquencer les souches impliquées dans le botulisme aviaire (depuis l'isolement jusqu'au séquençage de l'isolat) et de mettre en place une méthode MLVA qui se révèle particulièrement adaptée au contexte du botulisme aviaire. Ils ont également permis d'identifier de nouvelles cibles très prometteuses qui permettront de proposer d'autres outils de typage dans le futur, applicables au botulisme animal de manière général.

**Title:** *Clostridium botulinum* group III typing: Development of tools and application to the epidemiology of animal botulism.

**Keywords:** *Clostridium botulinum*, animal botulism, typing tools, WGS, CRISPR-Cas, MLVA.

**Abstract:** Animal botulism, which affects mainly poultry, wildbirds and cattle, is a neuroparalytic disease caused by the action of botulinum neurotoxins type C, D, C/D and D/C. They are produced by *Clostridium botulinum* group III, an anaerobic, spore-forming bacteria belonging to the genospecies *Clostridium novyi sensu lato* alongside *Clostridium novyi* and *Clostridium haemolyticum*.

The epidemiology of animal botulism is still poorly understood and molecular typing data are missing. This is mainly due to the unavailability of typing tools for *C. novyi sensu lato* characterization, mostly because strain isolation is very difficult.

The objective of the PhD project is to develop typing tools adapted to this specific context to be able to perform epidemiological investigations during animal botulism outbreaks and strain tracking in the future.

First, an isolation method was developed and validated using strains involved in animal botulism outbreaks in France. This method allowed the isolation of 35 strains from avian botulism outbreaks.

Then, a MLVA (Multi Loci Variable number of tandem repeats analysis) typing tool was evaluated on the isolated strains and used in real conditions during the investigation of two animal botulism outbreaks. Another potential typing approach was then explored through the study of CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats – CRISPR associated protein) systems. The results revealed the presence of these systems in 95% of *C. novyi sensu lato* genomes, with a great diversity (6 different types of system). This high diversity prevents the development of a typing tool based on these CRISPR-Cas systems.

Finally, following the optimization of the different steps preceding sequencing, *C. novyi sensu lato* genomes were compared to identify new relevant targets.

This work allowed the development of an approach to sequence strains involved in avian botulism (from strain isolation to whole genome sequencing) and to set up an MLVA method that is particularly adapted to the study of avian botulism. It has also allowed the identification of new targets that are very promising for the development of new tools in the future, usable to investigate all animal botulism outbreaks.