

ONIRIS - ECOLE NATIONALE VETERINAIRE, AGROALIMENTAIRE ET DE
L'ALIMENTATION

ANNEE 2020

**OUTILS DE CARACTERISATION MOLECULAIRE DES
COMMUNAUTES DE NEMATODES GASTRO-
INTESTINAUX CHEZ LES CAPRINES**

THESE

pour le

diplôme d'Etat de

DOCTEUR VETERINAIRE

présentée et soutenue publiquement

le 18/12/2020

devant

la Faculté de Médecine de Nantes

par

Laurie, Marie, Yvette BEDOUE

Née le 07/09/1996 à Tours (37)

JURY

Président : Monsieur Yves MAUGARS, Professeur de la Faculté de Médecine de Nantes

Rapporteur : Madame Suzanne BASTIAN, Maître de Conférences à ONIRIS

Assesseur : Monsieur Albert AGOULON, Maître de Conférences à ONIRIS

ONIRIS - ECOLE NATIONALE VETERINAIRE, AGROALIMENTAIRE ET DE
L'ALIMENTATION

ANNEE 2020

**OUTILS DE CARACTERISATION MOLECULAIRE DES
COMMUNAUTES DE NEMATODES GASTRO-
INTESTINAUX CHEZ LES CAPRINES**

THESE

pour le

diplôme d'Etat de

DOCTEUR VETERINAIRE

présentée et soutenue publiquement

le 18/12/2020

devant

la Faculté de Médecine de Nantes

par

Laurie, Marie, Yvette BEDOUE

Née le 07/09/1996 à Tours (37)

JURY

Président : Monsieur Yves MAUGARS, Professeur de la Faculté de Médecine de Nantes

Rapporteur : Madame Suzanne BASTIAN, Maître de Conférences à ONIRIS

Assesseur : Monsieur Albert AGOULON, Maître de Conférences à ONIRIS



Liste des enseignants-chercheurs d'ONIRIS

Département **BPSA** Biologie, Pathologie et Sciences de l'Aliment

Responsable : Hervé POULIQUEN - Adjoint : Emmanuel JAFFRES

Nutrition et Endocrinologie	Patrick NGUYEN (Pr)
Pharmacologie et Toxicologie	Jean-Claude DESFONTIS (Pr) Martine KAMMERER (Pr) Yassine MALLEM (Pr) Hervé POULIQUEN (Pr) Antoine ROSTANG (MCC)
Physiologie fonctionnelle, cellulaire et moléculaire	Jean-Marie BACH (Pr) Julie HERVE (MC) Lionel MARTIGNAT (Pr) Grégoire MIGNOT (MC)
Histologie et anatomie pathologique	Jérôme ABADIE (MC) Marie-Anne COLLE (Pr) Laetitia JAILLARDON (MC) Frédérique NGUYEN (MC)
Pathologie générale, microbiologie et immunologie	Hervé SEBBAG (MC)
Biochimie alimentaire industrielle	Clément CATANEO (MC) Joëlle GRUA (MC) Laurent LE THUAUT (MC) Carole PROST (Pr) Thierry SEROT (Pr) Florence TEXIER (MC)
Microbiotech	Géraldine BOUE (MC) Nabila HADDAD (MC) Emmanuel JAFFRES (MC) Mathilde MOSSER (MC) Hervé PREVOST (Pr) Raouf TAREB (MCC)

SAESP Santé des Animaux d'Élevage et Santé Publique

Responsable : Alain CHAUVIN - Adjoint : Raphaël GUATTEO	
Elevage, nutrition et santé des animaux domestiques	Nathalie BAREILLE (Pr) François BEAUDEAU (Pr) Ségolène CALVEZ (MC) Christine FOURICHON (Pr) Aurélien MADOUASSE (MC) Lucile MARTIN (Pr) Nora NAVARRO-GONZALEZ (MCC)
Infectiologie	Albert AGOULON (MC) Suzanne BASTIAN (MC) Alain CHAUVIN (Pr) François MEURENS (Pr) Emmanuelle MOREAU (PR) Carole PEROZ (MC) Nadine RAVINET (MC) Nathalie RUVOEN-CLOUET (Pr)
Médecine des animaux d'élevage	Sébastien ASSIE (MC) Catherine BELLOC (Pr) Isabelle BREYTON (MC) Christophe CHARTIER (Pr) Raphaël GUATTEO (Pr) Mily LEBLANC MARIDOR (MC) Anne RELUN (MC)
Hygiène et qualité des aliments	Jean-Michel CAPPELIER (Pr) Eric DROMIGNY (MC HDR) Michel FEDERIGHI (Pr) Bruno LE BIZEC (Pr) Catherine MAGRAS (Pr) Marie-France PILET (Pr) Fanny RENOIS-MEURENS (MC)

DSC Sciences cliniques

Responsable : Catherine IBISCH - Adjoint : Olivier GAUTHIER	
Anatomie comparée	Eric BETTI (MC) Claude GUINTARD (MC) Margarida RIBEIRO DA SILVA NEUNLIST (MCC)
Pathologie chirurgicale et anesthésiologie	Eric AGUADO (MC HDR) Olivier GAUTHIER (Pr) Eric GOYENVALLE (MC HDR) Caroline TESSIER* (MC) Gwénola TOUZOT-JOURDE (MC)
Dermatologie, parasitologie des carnivores et des équidés, mycologie	Patrick BOURDEAU (Pr) Emmanuel BENSIGNOR (Pr Ass)
Médecine interne, imagerie médicale et législation professionnelle vétérinaire	Nora BOUHSINA (MCC) Nicolas CHOUIN (MC) Anne COUROUCE (Pr) Jack-Yves DESCHAMPS (Pr) Amandine DRUT (MC) Marion FUSELLIER-TESSON (MC) Catherine IBISCH (MC) Françoise ROUX (Pr) Odile SENECAT (MC)
Biotechnologies et pathologie de la reproduction	Djemil BENCHARIF (MC HDR) Lamia BRIAND (MC HDR) Jean-François BRUYAS (Pr) Francis FIENI (Pr)

GPA Génie des procédés alimentaires

Responsable : Olivier ROUAUD - Adjoint : Sébastien CURET-PLOQUIN

Lionel BOILLEREAUX (Pr) Marie DE LAMBALLERIE (Pr) Francine FAYOLLE (Pr) Vanessa JURY (MC) Alain LEBAIL (Pr) Jean-Yves MONTEAU (MC HDR) Laurence POTTIER (MC) Cyril TOUBLANC (MC)	Sébastien CURET PLOQUIN (MC) Dominique DELLA VALLE (MC HDR) Michel HAVET (Pr) Emilie KORBEL (MC) Catherine LOISEL (MC) Olivier ROUAUD (Pr) Eve-Anne NORWOOD (MCC)
---	---

MSC Management, statistiques et communication

Responsable : Michel SEMENOU - Adjointe : Pascal BARILLOT (MC)

Mathématiques, Statistiques, Informatique	Véronique CARIOU (MC) Philippe COURCOUX (MC) El Mostafa QANNARI (Pr) Michel SEMENOU (MC) Chantal THORIN (Pr Ag.) Evelyne VIGNEAU (Pr)
Economie, gestion	Pascal BARILLOT (MC) Ibrahima BARRY (MCC) Florence BEAUGRAND (MC) Sibylle DUCHAINE (MC) Sonia EL MAHJOUB (MC) Jean-Marc FERRANDI (Pr) Samira ROUSSELIERE (MC)
Langues et communication	Marc BRIDOU (PLPA) David GUYLER (Ens. cont.) Shaun MEEHAN (Ens. cont.) Linda MORRIS (PCEA) Ian NICHOLSON (Ens. Cont.)

BTs : **Laurence Freret (PCEA)** Christophe Caron (PLPA), Pascale Fleury(PCEA), Virginie Magin (Ens. Cont.), Françoise Bricet (IAE).

Professeurs émérites : Poncelet

Guide de lecture des tableaux précédents :Pr : Professeur, Pr. AG : Professeur agrégé, MC : maître de Conférences, MCC : MC contractuel, PLPA : Professeur Lycée Professionnel Agricole, PCEA : Professeur Certifié Enseignement Agricole, IE : Ingénieur d'Etudes ; IAE : Ingénieur de l'Agriculture et de l'Environnement ; ens. cont.: enseignant contractuel; HDR : Habilité à Diriger des Recherches

* Vétérinaire spécialiste d'une spécialité européenne, américaine ou française

La reproduction d'extraits de cette thèse est autorisée avec mention de la source. Toute reproduction partielle doit être fidèle au texte utilisé. Cette thèse devra donc être citée en incluant les éléments bibliographiques suivants :

- Nom et prénoms de l'auteur : Bedouet Laurie
- Année de soutenance : 2020
- Titre de la thèse : Outils de caractérisation moléculaires de communautés de Nématodes gastro-intestinaux chez les Caprinés
- Intitulé du diplôme : Thèse de doctorat vétérinaire
- Université de soutenance : Faculté de Médecine de Nantes
- Ecole de soutenance : Oniris : Ecole Nationale Vétérinaire, Agroalimentaire et de L'alimentation Nantes Atlantique
- Nombre de pages : 90 p.

Remerciements

A **Monsieur Yves MAUGARS**, Professeur de la Faculté de Médecine de Nantes,
Pour m'avoir fait l'honneur d'accepter la présidence de ce jury,
Hommages respectueux.

A **Monsieur Albert AGOULON**, Maître de Conférences à ONIRIS,
Pour avoir accepté de faire partie de ce jury,
Sincères remerciements.

A **Madame Suzanne BASTIAN**, Maître de Conférences à ONIRIS,
Pour votre investissement et vos nombreux conseils pour ce travail, remarquables par nos
longs échanges Skype ; pour m'avoir fait découvrir de nouveaux horizons numériques avec
votre maîtrise de R ; pour n'avoir pas seulement été ma tutrice de thèse mais également ma
conseillère d'orientation lorsqu'en j'ai eu besoin,
Très sincères remerciements.

A ma famille,

A mes parents, Maman, Papa,

Pour m'avoir permis de suivre mon chemin, aussi tortueux qu'il semble se présenter ; pour votre soutien dans les moments les plus difficiles ; pour m'avoir transmis cette passion pour les nouveaux horizons et donné l'opportunité de voyager à l'autre bout du monde (promis Maman je reviendrais de temps en temps à la maison) ; pour m'avoir laissé développer mes talents artistiques et culinaires (mais non Papa je ne mettrais pas plus de sucre dans mes brownies).

A mon grand frère, Florian,

Pour notre complicité née depuis nos jeux game boy et les dimanches matin devant les dessins animés ; pour tes glaces en hiver ; pour ta bonne humeur matinale ; pour ton soutien.

A mes grand-parents, Mamie Jeanette, Papi Jean-Claude,

Pour ces longues après-midis de pêche, parfois infructueuses ; pour ces weekends sous le soleil de Noirmoutier, à déguster des crevettes ; pour ton humour Papi, pour ton extrême gentillesse Mamie ; et surtout, pour votre soutien à tous les deux, que je sais sans faille depuis le début.

A mes grand-parents, Mamie Josette, Papi Georges,

Pour m'avoir écouté pendant de longs moments le soir, raconter mes journées de stage puis mes journées de travail ; pour suivre avec application mes conseils avisés de vétérinaire ; pour ses cagettes remplies à ras bord de légumes et de figues ; pour ta créativité Mamie, pour ton sens du partage Papi ; et surtout, pour votre soutien à tous les deux que je sais également sans faille depuis le début (sauf peut-être quand il s'agit de goûter mes recettes ...).

Au reste de ma famille, Mélanie, Mahélie, Kassy, Raphaël, Tata Nathalie, Tonton Marc, Tata Cathy,

Pour tous nos moments partagés en famille, autour d'un bon repas très souvent.

A mon cousin, Julien

J'espère que tu as trouvé la paix où tu es maintenant.

To Aytuğ,

My love,

For being that gentle and caring person, for your cooking skills way beyond my expectations ("Sushi master"), for making me laugh when I'm moody or sad, for trusting me when I don't, for being part of my life even if it is miles away, and because this is the way,
I love you

A mes amis de toujours,

Aux Demdems, Jade et Olivier,

Pour tous ces m&m's et ces schoko-bons mangeaient en apéritif ; pour vos innombrables blagues ; pour cette amitié qui est née avant qu'on ne vienne au monde Jade et moi.

A mes patates douces, mes pandas éclatants,

Coline,

Pour ton déhanché endiablé sur la piste de danse ; pour ce petit bout de femme que tu es, si rafraîchissant.

Emma,

Pour cette complicité depuis notre binôme de TP de physique-chimie en seconde.

Noémie,

Pour être cette amie aussi attentive et compréhensive ; pour nos étés Toureangeaux à deux.

Juliette,

Pour ta créativité débordante, ta gentillesse et tes pâtisseries.

Philippine,

Pour nos soirées nantaises et nos escapades nazairiennes ; pour nos tisanes à la camomille et au miel de lavande ; pour nos petits-déjeuners à midi et nos dîners à minuit en vacances.

Zoé,

Twaaaa Zozo ; pour cette inimitié qui s'est transformée en grande amitié.

Louis,

Pour ta palette de maquillage bien plus fournie que la mienne, pour les paillettes que tu mets sur mes yeux ; pour les chips au poulet rôti ; pour cette personne affirmée que tu es devenu.

A vous sept,

Pour nos nouveaux ans décalés, nos vacances improvisés et nos soirées rocambolesques ; pour nos verres place Plum' (en espérant qu'ils reviennent très vite) ; pour votre soutien et nos conversations réconfortantes ; pour nos bientôt dix années d'amitié et les futures à venir.

A mes amis de prépa,

A Anne-Charlotte, Baptiste, Mathieu, Camille, Mona, Lisa, Tanguy, Nico,

Pour ces deux années d'entraide et de jolis moments,

Le temps nous a un peu éloigné mais j'espère que nous pourrons y remédier.

A Camille J.,

Pour ta gentillesse et ta douceur,

Je te souhaite tout le meilleur pour la suite.

Et à ceux que je n'ai pas cité mais avec qui j'ai partagé ces deux années de durs labeurs.

A mes amis de véto,

A la « Team Groupe 6 », les « 5A dans l'instalove »,

Charlotte,

Pour avoir été à mes côtés, littéralement, du début à la fin de cette aventure ; pour toutes ces soirées où j'ai bien cru perdre l'ouïe (en plus de la mémoire) ; pour ton soutien et ton oreille attentive ; pour ton gâteau à la noix de coco.

Mélanie,

Pour m'avoir initié au « vert » ; pour notre passion commune pour le yahourt glacé, la cuisine et les haltères ; pour nos nombreuses soirées à regarder des films ou des séries, pour nos après-midis créatives.

Lola,

Pour avoir été ma partenaire de voyage sur les terres Malgaches ; pour nos balades en mer à la recherche des baleines ou à photographier des poissons, pour nos longues soirées à jouer au solitaire ou lire des romans improbables, pour nos weekends au Wifi bar, pour notre attente dans les aéroports, pour cette expérience incroyable.

Céline,

Pour ce grand clown que tu es, la crack de musculation que tu es devenue et la grande vétérinaire phyto que tu deviendras ; pour tes astuces écolos et ta passion pour la cuisine (et manger en général, passion partagée).

Camille P.,

Pour ta persévérance dans la voie que tu t'es choisie ; toi aussi tu nous as manqué pendant cette dernière année mais c'était pour mieux se retrouver.

Diane,

Pour notre passion commune pour la culture Argentine, j'espère qu'un jour nous aurons l'occasion d'y retourner ensemble ; pour ces quelques mois de colocation riche en émotions.

Joséphine,

Pour ta simplicité et tes compétences en communication, pour les frites à ton anniversaire.

Claire,

Pour nos covoyages du weekend, pour ton grain de folie qui s'est libéré.

A Géraldine,

Pour notre petite colocation au « Bol du moine », pour m'avoir fait découvrir ma passion pour les manèges à sensation, pour ce voyage au pays des cigognes et des bretzels (et de la choucroute, de la tarte au linz, de la tarte flambée ...), pour nos retours de soirée (même si je ne me souviens pas des trajets en vélo), pour notre symbiose culinaire.

A Valentin,

Pour ces *before* et ces *after* improbables ; pour nos conversations entourés de cagettes de légumes et d'œufs.

A Amandine,

Pour avoir été ma partenaire de voyage sur les terres argentines, pour nos aventures en stop mémorables, pour nos longues balades à couper le souffle et les jambes, pour nos débuts en tant que vétérinaire de faune sauvage.

Et à ceux que je n'ai pas cité mais avec qui j'ai partagé ces cinq dernières années.

A mes amis d'ailleurs,

Por la guardia B, **Nico S., Bicho, Rocio, Seba, Pol, Luci, Flor, Chinchi, Marinu, Francisco, Sixto, Lucho, Lucas T., Nico P., Ivi, Lucas F., Milton,**

Por haberme descubrir la vida Argentina, el mate, la cumbia, los boludos, los asados, las tortas fritas, las empanadas ... tantas cosas que me hacen esperar una vuelta en vuestro país a ver ustedes de nuevo.

Por **Mati y Daiana,** mis tutores,

Por vuestra voluntad de integrarme y iniciar a la vida Argentina desde el principio, por todos los momentos de intercambio,
Estan los bienvenidos a Francia.

Por mis amigos de la RAE2 :

Clara, mi compañera de habitación,

Por tu risa comunicativa y tu alegría de vivir, por tu deseo de aprender,
Estoy segura que la vida reservate sorpresas y encuentras, la primera sera yo en tu hermoso pueblo lo espero.

Tamara,

Por tu fuerza y tu locura, por tu amor de la vida latina y tu pasión por el viaje,
Sé que vas volver una grande veterinaria,

Sandra,

Por todos los momentos compartidos al rededor de una vaca abierta o alimentando un ternero,
por los proximos en Madrid,
Te deseo lo mejor en tu nueva vida.

Canelle, Aude, Marie-Nour y Adri,

Por todos los momentos juntos en nuestra segunda casita,
Vos deseo lo major para todas y sobre todo que nos encontramos de nuevo.

Por todas las otras personas que encontré durante este hermoso viaje y esta experiencia inolvidable.

Table des matières

Liste des enseignants-chercheurs d'ONIRIS	2
Remerciements.....	6
Table des matières.....	11
Table des figures	13
Table des tableaux.....	14
Table des sigles	15
Introduction.....	16
PARTIE 1 : CARACTERISATION MOLECULAIRE DES ESPECES DE NEMATODES GASTRO-INTESTINAUX.....	18
I. Les méthodes d'identification des nématodes gastro-intestinaux	19
A. L'identification morphologique	19
1. Le principe de la taxonomie	19
2. Les limites des critères morphologiques d'identification.....	20
3. Notions d'espèces polymorphiques et cryptiques	23
B. L'identification moléculaire	24
1. L'ADN comme empreinte digitale moléculaire	24
2. La génétique de la spéciation et l'avènement du <i>DNA barcoding</i>	25
II. Les marqueurs génétiques d'identification taxonomique	28
A. Les séquences nucléaires.....	28
1. L'ADN ribosomique.....	28
2. Les séquences non codantes	29
3. Les séquences codantes	30
B. Les séquences mitochondriales	31
1. Les particularités de l'ADNm	31
2. Les gènes mitochondriaux d'intérêt	33
CONCLUSION PARTIELLE.....	34
III. Les méthodes de séquençage de l'ADN	35
A. Le choix d'une méthode	35
B. L'amplification de séquences d'ADN.....	35
1. Récupération de l'ADN.....	35
2. Le principe de la PCR	36
3. Les paramètres de la PCR	38
4. Les méthodes basées sur le principe de la PCR	39
C. Le séquençage	41

1. Les méthodes de séquençage.....	41
2. Repérer des erreurs dans une séquence de nucléotides	42
Conclusion de la partie I	45
PARTIE 2 : ANALYSE BIO-INFORMATIQUE DE SEQUENCES NUCLEOTIDIQUES DE NEMATODES GASTRO-INTESTINAUX	46
I. Introduction.....	47
II. Matériels et méthodes	48
A. Inventaire des espèces de nématodes gastro-intestinaux.....	48
B. Récupération des données de GenBank	48
C. Extraction et bilan des données sur les marqueurs génétiques.....	48
D. Stratégies d'identification moléculaire.....	49
E. Simulation d'une PCR <i>in silico</i>	49
III. Résultats	52
A. Espèces de nématodes gastro-intestinaux et bilan des séquences disponibles.....	52
B. Bilan des données récupérées.....	52
C. Stratégies d'identification moléculaire.....	53
D. Simulation d'une PCR <i>in silico</i>	54
IV. Discussion	56
Conclusion de la partie 2.....	62
CONCLUSION GENERALE.....	63
Bibliographie.....	65
Annexe 1 : Organisation des données et métadonnées GenBank.	74
Annexe 2 : Tableau des espèces de nématodes gastro-intestinaux des Caprinés – nombre de séquences dans Genbank.	75
Annexe 3 : Extrait de l'inventaire des séquences GenBank constitué dans cette étude et publications associées.....	77
Annexe 4 : Nombre de séquences de marqueurs génétiques disponibles pour 40 espèces de Nématodes gastro-intestinaux de Caprinés, recensées dans cette étude.	79
Annexe 5 : Tableau récapitulatif des séquences des amorces et sondes publiées pour la région ITS-2 des nématodes gastro-intestinaux et leur spécificité attendue.	80
Annexe 6 : Tableau des résultats de la PCR <i>in silico</i>	84
Annexe 7 : Tableau des capacités de détection attendues dans les publications d'amorces, confrontées aux résultats de la PCR <i>in silico</i>	86
Annexe 8 : Variabilité interspécifique des séquences pour les positions des amorces NC1 et NC2.	88

Table des figures

Figure 1 : Exemple de la classification taxonomique de l'espèce <i>Haemonchus contortus</i> , inspiré de (5).....	19
Figure 2 : Exemples d'œufs d'helminthes présents chez les Ruminants (6)	20
Figure 3 : Schéma des caractéristiques observables des larves de stade L3 (6).	21
Figure 4 : Observations au microscope de larves L3 de <i>Chabertia sp.</i> (A), <i>Haemonchus sp.</i> (B) et <i>Trichostrongylus sp.</i> (C) (1).....	22
Figure 5 : Observations microscopiques de la bourse copulatoire et des spicules de <i>Trichostrongylus axei</i> (A), <i>Trichostrongylus colubriformis</i> (B) et <i>Trichostrongylus vitrinus</i> (C) (8)	23
Figure 6 : Schéma de l'assortiment indépendant des chromosomes homologues lors de la méiose (16).	25
Figure 7 : Schéma de l'organisation des séquences nucléaires ribosomiques des Eucaryotes	28
Figure 8 : Graphique des variations observées au sein des gènes ribosomiques (22)	30
Figure 9 : Agencement du génome mitochondrial de <i>Marshallagia marshalli</i> . (55)	32
Figure 10 : Chronologie des différentes méthodes d'identification des Nématodes gastro-intestinaux	35
Figure 11 : Schéma du principe de la PCR (86)	37
Figure 12 : Schéma d'un dimère d'amorces	38
Figure 13 : Schéma du principe de la méthode de <i>Duplex Sequencing</i> (100)	44
Figure 14 : Sites de fixation des amorces NC1 et NC2 sur l'ADN ribosomique des Nématodes	54

Table des tableaux

Tableau 1 : Numéro d'accèsion des séquences constituant le lot de séquences pour la simulation de la PCR <i>in silico</i>	50
Tableau 2 : Classification des sources bibliographiques en fonction de la stratégie appliquée pour l'identification d'espèces ou de genres de nématodes gastro-intestinaux.....	53
Tableau 3 : Tableau dynamique des résultats de la PCR, à la manière d'un résultat de puce microfluidique, inspiré de (124).....	60

Table des sigles

ADN : *acide désoxyribonucléique*

ADNr : *ADN ribosomique*

ARN : *acide ribonucléique*

ARNr : *ARN ribosomique*

ARNt : *ARN de transfert*

ATP : *adénosine triphosphate*

BLAST : *Basic Local Alignment Search Tool*

cox : *cytochrome oxydase*

ddNTP : *didésoxyribonucléotide triphosphate*

dNTP : *désoxyribonucléotide triphosphate*

DOI : *Digital Object Identifier*

ETS : *External-Transcribed Spacer*

IGS : *InterGenic Spacer*

indel : *insertion ou délétion*

ITS : *Internal-Transcribed Spacer*

MT-PCR : *PCR multiplexe en tandem*

nad4 : *4ème sous-unité de la nicotinamide déshydrogénase*

NCBI : *National Center of Biology Information*

NGS : *Next Generation Sequencing*

numts : *nuclear mitochondrial pseudogenes*

PCR : *Polymerase Chain Reaction*

PCR-DGGE : *PCR-Denaturing Gradient Gel Electrophoresis*

PCR-RFLP : *PCR - Restriction Fragment Length Polymorphism*

qPCR : *PCR quantitative ou PCR en temps réel*

SNP : *Single Nucleotide Polymorphism*

Introduction

Le point de départ de cette thèse provient d'une étude des espèces de nématodes gastro-intestinaux réalisée chez les Caprinés de la réserve zoologique de la Haute-Touche (Indre, 36) (1). L'objectif était de caractériser les nématodes digestifs infestant les dix espèces ou races de Caprinés du parc pour améliorer les plans de traitements antiparasitaires de ces animaux. Les méthodes de coproscopie et de coproculture ont permis d'identifier plusieurs genres de nématodes sans être capable d'identifier jusqu'au rang de l'espèce. Par ailleurs, certains genres restaient indifférenciables morphologiquement et ont été regroupés sous le nom de « strongles digestifs ». Ces derniers comportent des espèces ayant un pouvoir pathogène élevé chez les Caprinés, telles que *Haemonchus contortus* ou *Teladorsagia circumcincta* ; qu'il est donc important de savoir mettre en évidence pour mettre en œuvre les traitements adaptés.

Les nématodes gastro-intestinaux sont des vers ronds d'importance majeure chez les Ruminants. Ils sont à l'origine de nombreuses maladies responsables de pertes économiques importantes dans le monde entier. Les animaux se contaminent par l'intermédiaire des pâtures où se trouvent les formes infestantes des parasites (œufs, larves), elles-mêmes issues des fèces d'animaux infectés. La présence et l'abondance dans le milieu extérieur des nématodes gastro-intestinaux sont influencées par de nombreux facteurs comme le climat, la présence d'autres espèces animales domestiques ou sauvages, les méthodes d'élevage ... Certains nématodes ont la capacité d'entrer en hypobiose afin d'échapper à des conditions climatiques peu favorables ce qui rend leur détection plus difficile. Les manifestations cliniques de gastro-entérites parasitaires sont très diverses : anémie, avortement, diminution de la prise alimentaire, diarrhée, perte de poids ... avec, à certaines périodes, des pics de mortalité. Mais les formes cliniques sont minoritaires par rapport aux infections sub-cliniques, responsables de pertes de productivité et donc plus difficiles à mettre en évidence. Le diagnostic clinique des nématodes gastro-intestinaux est par conséquent peu spécifique (2).

La gestion des infestations par les nématodes gastro-intestinaux est principalement basée sur l'utilisation d'anthelminthiques. Cependant, même avec la mise en place de stratégies de traitement, ce type de contrôle est coûteux et souvent peu efficace. D'autant que l'utilisation excessive des anthelminthiques a conduit à l'apparition à travers le monde de résistances à ces molécules. C'est le cas notamment des trois genres de parasites les plus importants économiquement chez les Petits Ruminants : *Teladorsagia sp.*, *Haemonchus sp.* et *Trichostrongylus sp.* ; pour lesquels des populations multi-résistantes ont été décrites (3).

Le développement de méthodes alternatives aux anthelminthiques est tributaire d'une meilleure connaissance des espèces. L'identification des espèces présentes apparaît donc comme un élément clé dans la réalisation d'études épidémiologiques, pathogéniques, et la mise en place de moyens de prévention et de lutte ciblés contre les nématodes gastro-intestinaux. Les méthodes d'identification employées couramment s'appuient sur les caractéristiques morphologiques des différents stades de développement des parasites (œufs, larves, adultes). Cependant, il n'est pas toujours possible d'identifier l'espèce ou même le genre à partir de ces méthodes qui requièrent notamment du temps et un niveau d'expertise élevé. L'enjeu réside également dans l'identification de plusieurs espèces simultanément, étant donné que les situations de co-infestations, c'est-à-dire d'infestations par plusieurs espèces de parasites en même temps, sont très fréquentes (2).

Depuis plusieurs années maintenant, les approches moléculaires se sont multipliées afin de développer des méthodes capables de décrire la diversité et l'abondance relative des nématodes dans

un environnement donné. Elles s'appuient sur l'extraction puis la comparaison de l'information génétique contenue dans l'ADN (acide désoxyribonucléique) de chaque parasite. Une standardisation de ces méthodes est nécessaire afin d'en faire des outils fiables et pratiques d'étude de la biodiversité. Des bases de données regroupant les séquences nucléotidiques de divers organismes se sont développées afin de rendre accessible l'information récoltée à travers le monde.

L'objectif de cette thèse est de savoir si suffisamment d'outils et de données moléculaires sont déjà disponibles pour pouvoir identifier et décrire de manière robuste et standardisée des communautés de parasites, en prenant l'exemple des nématodes gastro-intestinaux chez les Caprinés, et, le cas échéant, identifier où sont les manques. La première partie aborde le principe de la caractérisation moléculaire des taxons et en particulier à l'échelle de l'espèce. La seconde partie présente une étude *in silico* des séquences nucléotidiques de nématodes gastro-intestinaux des Caprinés, disponibles dans la base de données GenBank au premier semestre de 2020.

PARTIE 1 : CARACTERISATION MOLECULAIRE DES ESPECES DE NEMATODES GASTRO-INTESTINAUX

I. Les méthodes d'identification des nématodes gastro-intestinaux

A. L'identification morphologique

1. Le principe de la taxonomie

La taxonomie est une discipline scientifique qui vise à nommer et à classer les diverses formes d'êtres vivants (4). On utilise la nomenclature binomiale, établie par Carl von Linné au XVIII^{ème} siècle, pour désigner un organisme. Cette nomenclature reprend le nom du genre puis le nom de l'espèce auquel l'organisme appartient. Il existe des catégories de classements plus vastes de façon à former une classification hiérarchique : les genres semblables sont rassemblés en familles, les familles en ordre, les ordres en classes, les classes en embranchements ou phylums, les phylums en règnes, et les règnes en domaine. Au fur et à mesure, des niveaux intermédiaires, comme les sous-ordres ou les super-familles, ont été créés afin d'affiner cette classification (figure 1).

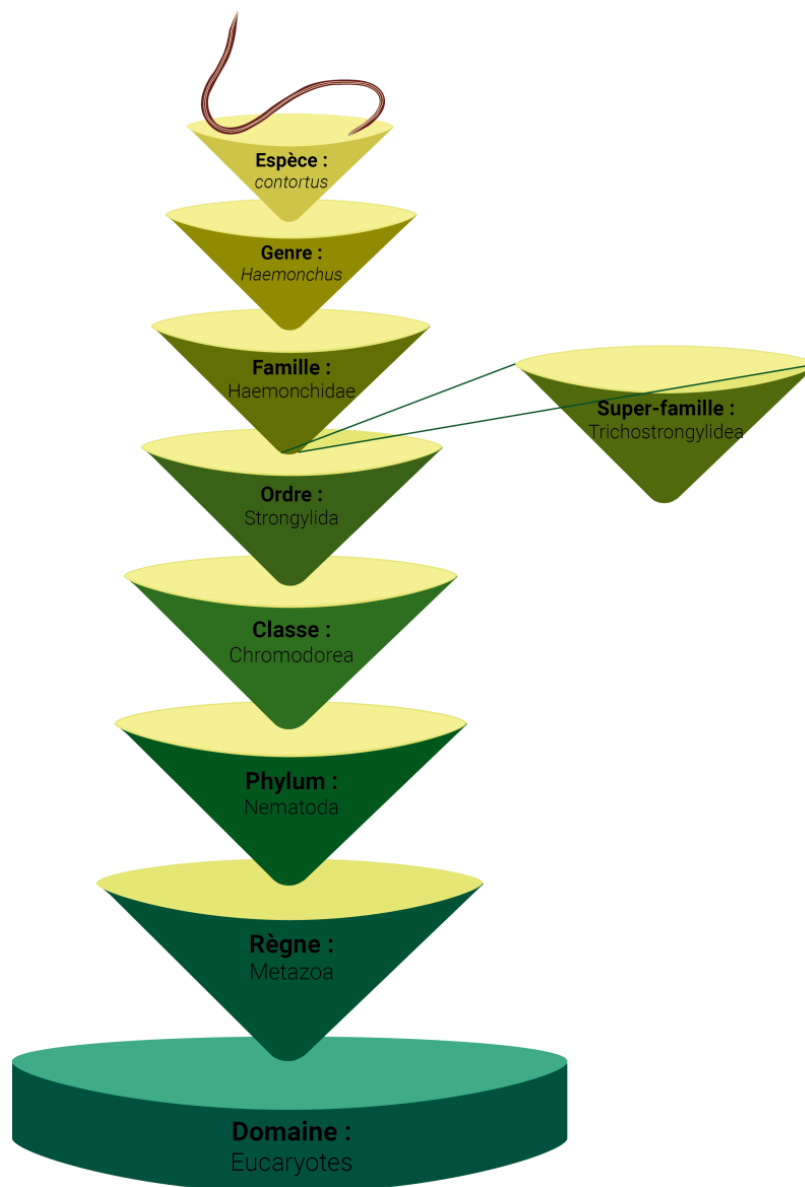


Figure 1 : Exemple de la classification taxonomique de l'espèce *Haemonchus contortus*, inspiré de (5)

Cette classification est établie selon des degrés d'homologie entre organismes. Des organismes qui partagent des caractéristiques communes sont dits apparentés. Plus le nombre de ces caractéristiques est élevé, plus ces organismes seront placés ensemble dans un rang taxonomique ou taxon inférieur, c'est-à-dire dans un même genre ou une même espèce. Par exemple, *Haemonchus contortus* et *Ostertagia circumcincta* sont assez semblables pour être classés dans la même famille : les Haemonchidae ; mais suffisamment différents pour appartenir à des genres distincts : respectivement *Haemonchus sp.* et *Ostertagia sp.*.

2. Les limites des critères morphologiques d'identification

D'un point de vue pratique, la classification des nématodes est principalement basée sur les caractères morphologiques des organismes. Il s'agit d'une méthode assez peu coûteuse à mettre en œuvre mais présentant plusieurs inconvénients.

a) Observation microscopique des œufs

L'un des premiers caractères discriminants des nématodes est la morphologie des œufs. Ces derniers sont observés par des méthodes de coproscopie à partir de fèces d'animaux. Il en existe plusieurs : par flottation, par sédimentation. Les œufs sont identifiés au microscope selon leur taille, leur forme, leur couleur ou encore l'épaisseur de leur paroi (figure 2).

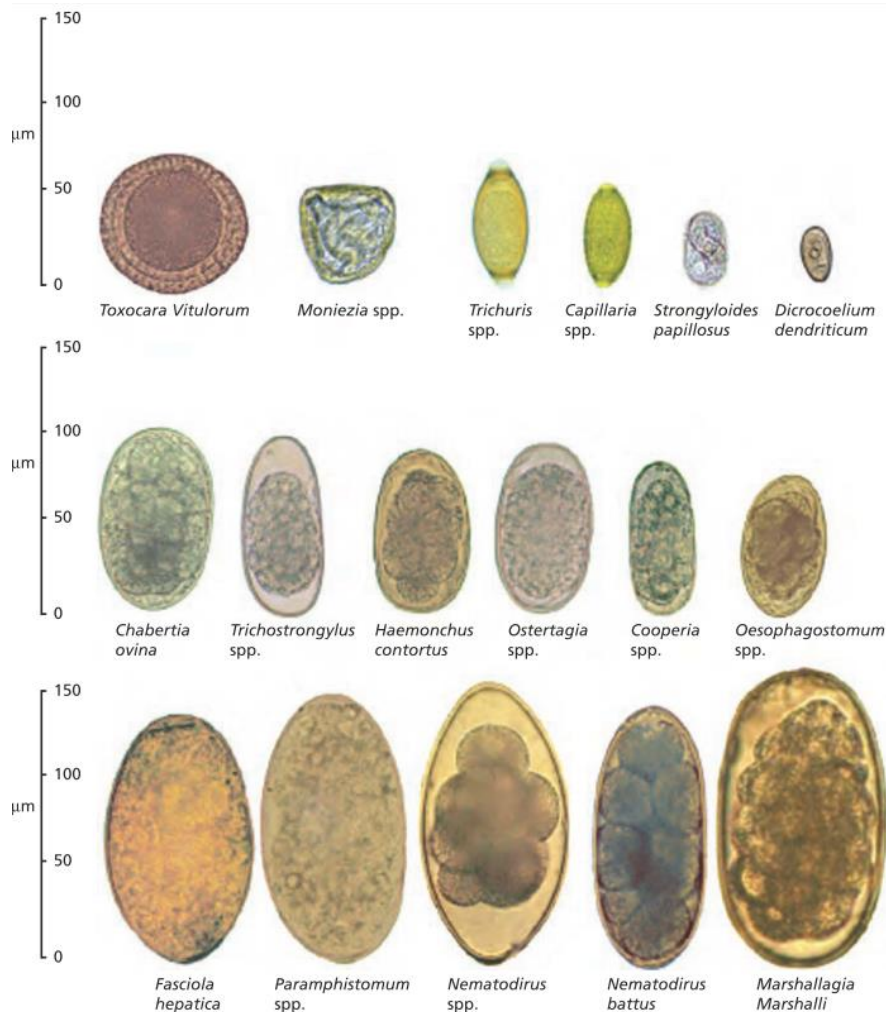


Figure 2 : Exemples d'œufs d'helminthes présents chez les Ruminants (6)

Cependant, chez les nématodes, il n'est pas toujours possible d'identifier l'espèce ni même le genre à partir de la morphologie des œufs. Divers genres de nématodes sont d'ailleurs regroupés et désignés sous le nom de « strongles digestifs » (*Haemonchus sp.*, *Ostertagia sp.*, *Marshallagia sp.*, *Cooperia sp.*, *Nematodirus sp.*, *Trichostrongylus sp.*, *Bunostomum sp.*, *Oesophagostomum sp.* et *Chabertia sp.*) puisque leur différenciation n'est pas possible sur la seule observation de leurs œufs. Ces derniers sont ovales, de taille variable (60 – 125 par 30 – 60 µm), à paroi fine, et contiennent 4 à 32 cellules ou blastomères. Seuls les genres *Nematodirus sp.* et *Marshallagia sp.* sont différenciables. Les œufs de *Marshallagia sp.* sont similaires à ceux des autres strongles mais sont de taille plus importante : 150 – 180 par 80 – 120 µm. Les œufs de *Nematodirus sp.* sont également de grande taille : 150 – 210 par 70 – 120 micromètres ; mais la morula, constituée de huit blastomères, est bien distincte de la paroi. L'espèce *Nematodirus battus* est reconnaissable sur la seule morphologie de son œuf dont la forme est plus allongée et dont la morula occupe l'ensemble de l'espace interne de l'œuf (figure 2) (1).

b) Observation microscopique des larves

L'identification peut être précisée en observant les caractères morphologiques de stades de développement plus avancés. C'est le cas du stade larvaire L3. Les larves sont obtenues après coproculture : il s'agit de faire éclore les œufs et permettre le développement jusqu'au stade désiré. Les larves sont ensuite récoltées et peuvent être observées au microscope. Les caractéristiques alors utilisées pour l'identification sont la taille, la morphologie de l'extrémité crâniale et de la gaine caudale (*tail sheath* dans la figure 3) et le nombre de cellules intestinales (*intestinal cells* dans la figure 3, figure 4) (6).

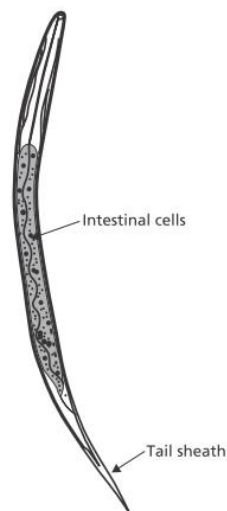


Figure 3 : Schéma des caractéristiques observables des larves de stade L3 (6).

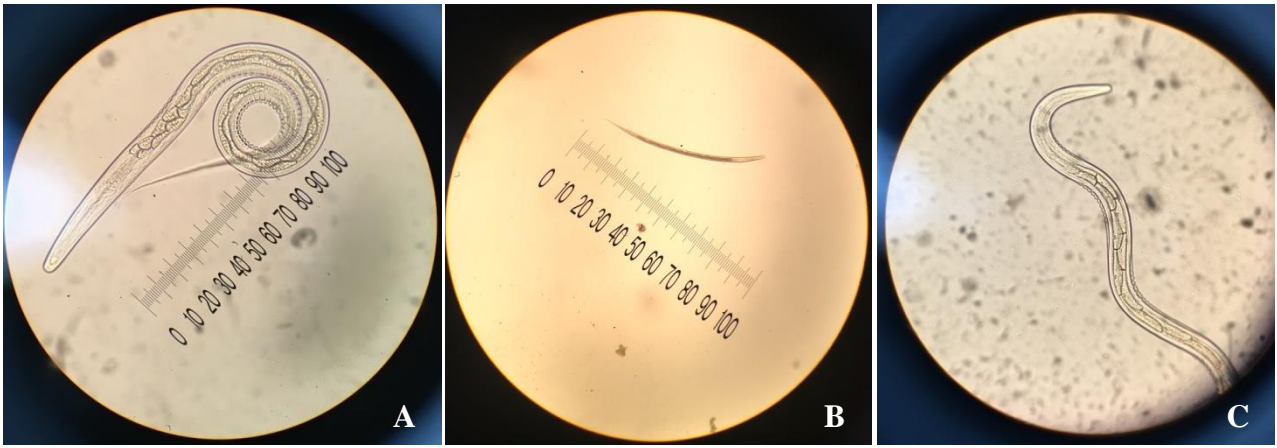


Figure 4 : Observations au microscope de larves L3 de *Chabertia sp.* (A), *Haemonchus sp.* (B) et *Trichostrongylus sp.* (C) (1)

A partir de ces caractéristiques, il est possible de distinguer les nématodes à l'échelle du genre. Néanmoins les genres *Teladorsagia sp./Ostertagia sp.* et *Trichostrongylus sp.* restent difficilement différenciables. Il n'est également pas possible de déterminer l'espèce au sein des genres.

En général, une incubation à 27°C pendant sept à dix jours est recommandée pour le développement de la plupart des larves de nématodes. L'identification morphologique à partir du stade larvaire est donc une méthode qui demande du temps. De plus, les conditions de culture optimales varient d'une espèce à une autre. Ceci est notamment à prendre en considération lorsque l'objectif de l'étude est de déterminer l'abondance d'une espèce : l'application de conditions de culture identiques va favoriser le développement d'une espèce par rapport à une autre et biaiser les résultats. Cette méthode permet donc de donner des indications sur l'identité des espèces ou genres présents mais n'est pas adaptée pour leur quantification. De plus, la composition du milieu de culture (présence de moisissures, disponibilité en oxygène, pH), l'âge de l'hôte, la conservation des échantillons avant culture, sont d'autres sources potentielles de variations des résultats (7).

c) Observation microscopique des vers adultes

Enfin l'observation des caractéristiques des vers au stade adulte permet un degré de précision encore plus important pour l'identification des nématodes. Les parasites sont récoltés par filtration du contenu du tube digestif des animaux au cours d'une autopsie. Premièrement, l'espèce hôte et l'organe autopsié donne des indications quant à l'identification. En effet, à partir des connaissances déjà acquises sur les différentes espèces de parasites, on peut relier un organe à un ou plusieurs parasites. Par exemple, les genres *Haemonchus sp.* et *Teladorsagia sp.* sont habituellement des parasites de l'abomasum. Les genres *Trichostrongylus sp.* et *Nematodirus sp.* sont communément présents dans l'intestin grêle (8). Afin d'éviter la migration des vers au sein du tube digestif, il est important de ligaturer les différentes portions de ce dernier lors de l'autopsie des animaux. Au stade adulte, les caractères morphologiques permettant la distinction taxonomique sont : la capsule buccale, l'œsophage, la position de la vulve chez la femelle mais principalement la bourse copulatoire et les spicules du mâle (figure 5). Cette méthode d'identification nécessite donc d'avoir des organismes intègres pour pouvoir repérer les critères cités ci-dessus et surtout un niveau d'expertise assez élevé. Ce dernier inconvénient est commun à l'ensemble des stades de développement des nématodes pour lesquels les similitudes entre genres et espèces sont très nombreuses et limite une standardisation de cette méthode.

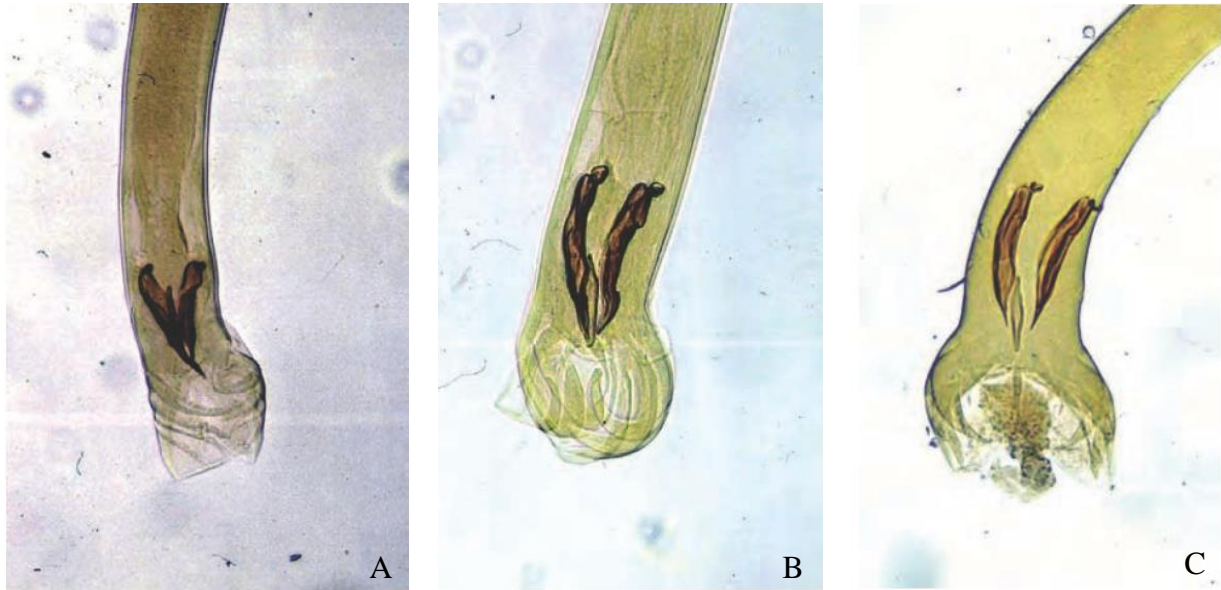


Figure 5 : Observations microscopiques de la bourse copulatoire et des spicules de *Trichostrongylus axei* (A), *Trichostrongylus colubriformis* (B) et *Trichostrongylus vitrinus* (C) (8)

3. Notions d'espèces polymorphiques et cryptiques

Une espèce polymorphique est une espèce comprenant plusieurs morphes différents, c'est-à-dire qu'elle comprend des individus de morphologies différentes mais de génomes similaires. C'est le cas notamment des espèces appartenant au genre *Teladorsagia sp.* Initialement, *T. circumcincta*, *T. trifurcata* et *T. davtiani* ont été définis comme trois espèces distinctes sur la base de différences morphologiques au niveau de l'appareil génital mâle. Or plusieurs observations amènent à penser que *T. trifurcata* et *T. davtiani* ne seraient en réalité que des morphotypes de *T. circumcincta*. En effet les trois morphotypes mâles sont toujours observés conjointement dans une même population, avec *T. circumcincta* majoritaire, alors que les femelles présentent toute la même morphologie (9). L'analyse du marqueur génétique ITS-2 des trois morphotypes ne révèle d'ailleurs pas de différence significative, soutenant ainsi l'hypothèse selon laquelle il s'agirait de la même espèce (10). Les genres *Ostertagia sp.* (11) et *Cooperia sp.* (12) constituent d'autres exemples d'espèces polymorphiques parmi les nématodes gastro-intestinaux.

A l'inverse, il existe des espèces présentant des morphologies similaires mais qui sont distinctes sur le plan génétique : il s'agit d'espèces cryptiques. Un exemple provient encore une fois du genre *Teladorsagia sp.* pour lequel l'espèce *T. borearcticus* a été définie distinctement de *T. circumcincta* sur la base d'un spectre d'hôte et d'une localisation géographique différents ainsi que sur des différences dans la séquence codant la 4^{ème} sous-unité de la nicotinamide déshydrogénase (*nad4*) de l'ADN mitochondrial (ADNm) (13). Au sein du genre *Haemonchus sp.*, le statut des espèces *H. contortus* et *H. placei* est controversé. Bien que présentant un spectre d'hôte différent : *H. contortus* est plutôt spécifique des Ovins alors que *H. placei* est spécifique des Bovins ; ces deux parasites ne présentent que peu de différences morphologiques. Le spectre d'hôte suffit-il à distinguer les deux espèces ? La comparaison de leur génome a permis de confirmer leur statut comme deux espèces distinctes mais de morphologie proche (14,15).

Enfin, des phénomènes d'hybridation ont été décrits chez plusieurs genres de nématodes gastro-intestinaux. Outre la transmission possible de gènes de pathogénicité et de résistance aux traitements antiparasitaires, l'échange de gènes impliqués dans la spécificité d'hôte peut perturber la caractérisation d'une espèce par son spectre d'hôte. Des expériences d'hybridation entre espèces appartenant aux genres *Cooperia sp.* et *Ostertagia sp.* ont montré que très souvent, les descendants de tels accouplements présentent un mélange de caractères morphologiques appartenant aux deux espèces parentales (9).

Ainsi, l'identification des espèces de nématodes gastro-intestinaux basée uniquement sur des caractères morphologiques peut être à l'origine d'erreurs taxonomiques. L'ensemble des inconvénients de l'identification morphologique a donc encouragé les scientifiques à rechercher d'autres méthodes plus fiables, plus rapides, applicables pour n'importe quel stade physiologique du parasite et ne nécessitant pas un niveau d'expertise élevé, c'est-à-dire capables d'être standardisées.

B. L'identification moléculaire

Quelle information l'ADN apporte-t-il sur le rang taxonomique d'un individu ?

1. L'ADN comme empreinte digitale moléculaire

L'ADN est le support physique de l'information génétique. Cette information réside dans l'enchaînement des nucléotides qui constitue la molécule d'ADN. Les gènes sont des séquences particulières de nucléotides qui codent pour la fabrication d'autres molécules, comme les protéines, nécessaires au fonctionnement cellulaire et donc de l'organisme. L'ADN est d'abord transcrit en acide ribonucléique (ARN) qui est ensuite traduit en une séquence d'acides aminés formant la protéine.

L'ADN est en grande partie contenu dans le noyau de la cellule. Il est formé de deux longues chaînes, appelés brins, formant une double hélice. Deux doubles hélices forment un chromosome. Les organismes diploïdes, dont les nématodes, possèdent deux paires de chromosomes. Cette organisation permet la réplication de l'ADN, la division cellulaire mais également la transmission entre génération de l'information génétique. En effet, chez les espèces à reproduction sexuée, lors de la formation des gamètes, chaque cellule fille reçoit un brin d'ADN ou chromatide : c'est la méiose (figure 6).

La reproduction sexuée est une source de diversité génétique. Lors de la métaphase de la méiose I, les paires de chromosomes homologues s'alignent aléatoirement sur la plaque équatoriale, de sorte que deux combinaisons de chromosomes sont possibles pour une cellule fille issue de la méiose I (figure 6).

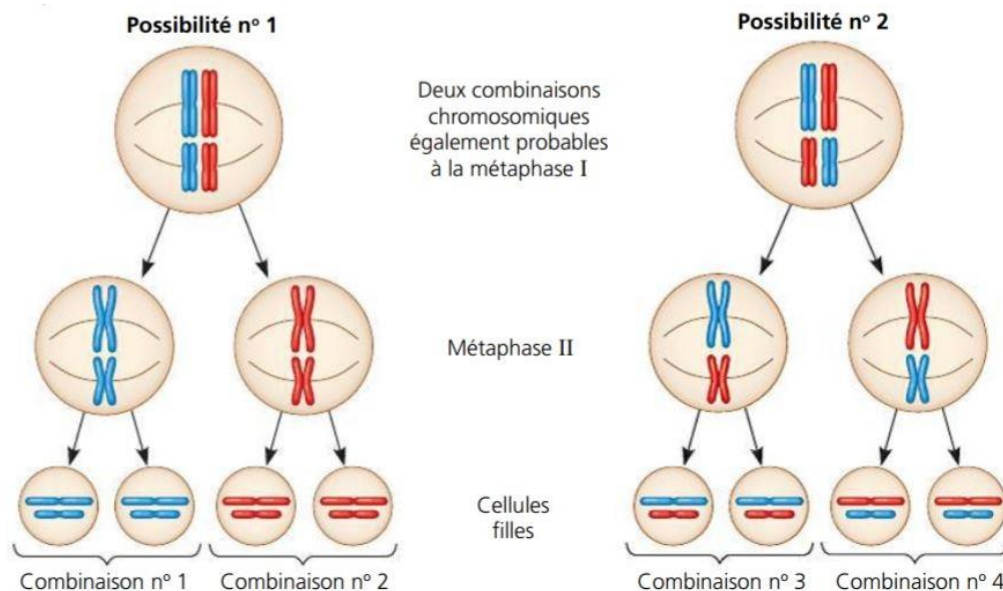


Figure 6 : Schéma de l'assortiment indépendant des chromosomes homologues lors de la méiose (16).

Plus tôt au cours de la prophase de la méiose I, lorsque les chromosomes s'apparient par paire homologue, des échanges de segments d'ADN peuvent avoir lieu entre deux chromatides non sœurs, c'est-à-dire n'appartenant pas au même chromosome. Ce mécanisme est appelé enjambement ou « crossing-over » et est à l'origine de nouvelles combinaisons d'allèles. Enfin, la nature aléatoire de la fécondation est un facteur supplémentaire de diversité génétique.

Les mutations constituent la première source de diversité. Il s'agit d'erreurs dans la réplication de la séquence de nucléotides de l'ADN devenues permanentes. Il en existe plusieurs types (substitution, insertion, délétion) ayant des conséquences plus ou moins importantes selon qu'elles modifient ou non la séquence en acide-aminés d'une protéine. Leur occurrence est rare du fait des mécanismes de réparation de l'ADN qui existe au sein de la cellule chez les Eucaryotes. Leur fréquence est estimée à 1 par 100 000 gènes pour les règnes animal et végétal. Néanmoins lorsqu'une mutation se produit dans un gamète ou dans une cellule productrice de gamètes et qu'elle échappe aux mécanismes de réparation, elle est susceptible d'être transmise aux générations futures. La sélection naturelle s'occupe ensuite de perpétuer ou non cette mutation. En effet, une mutation est plus à même d'être transmise aux générations suivantes si elle confère un avantage sélectif aux organismes qui la possèdent (16).

2. La génétique de la spéciation et l'avènement du *DNA barcoding*

L'ADN est un bon support d'études puisqu'il s'agit d'une molécule stable biologiquement et que l'on retrouve dans la quasi-totalité des tissus biologiques. Dans la partie précédente, il a été démontré que plusieurs mécanismes étaient à l'origine de la diversité génétique, de sorte que l'ADN d'un organisme est unique et constitue son identité moléculaire. L'ADN constitue également une molécule informative sur l'histoire évolutive d'un individu. En effet, en comparant les séquences ADN, on peut définir plusieurs échelles de diversité génétique : individuelle, population, espèce, genre ... résultant de l'évolution de chaque organisme qui compose cette échelle. C'est ce qui établit la base moléculaire de la classification des êtres vivants.

Plusieurs phénomènes évolutifs influent la diversité génétique des individus. Tout d'abord, les événements de migration des populations en association avec le succès reproducteur des individus participent aux échanges de gènes, aussi appelé flux génétique. Il en résulte une homogénéisation du patrimoine génétique des populations et donc une diminution de la diversité génétique. C'est d'ailleurs lorsqu'il y a une interruption de ce flux génétique, principalement du fait d'un isolement géographique des populations ou par l'occupation de niches écologiques différentes, qu'une espèce se constitue. Les deux populations évoluent chacune de leur côté, accumulant des mutations, jusqu'à atteindre un degré de divergence tel que leur reproduction n'est plus compatible (17,18). De plus, l'étude des séquences génétiques permet de révéler des différences que la simple étude du phénotype ne peut mettre en évidence : c'est le cas des espèces cryptiques évoquées précédemment (13,15,19). Ainsi, deux organismes ayant un nombre élevé de mutations entre leur séquence ADN sont peu apparentés, c'est-à-dire qu'ils sont éloignés dans leur histoire évolutive. En revanche, plus le pourcentage de similitude entre les séquences est élevé, plus les organismes peuvent être classés dans de bas rangs taxonomiques (genre, espèce) suggérant une histoire évolutive commune récente. Chaque rang taxonomique est ainsi caractérisé par un certain degré de similitude ou de divergence. Une espèce regroupe donc des organismes ayant un degré important de similitude entre leur séquence ADN.

Pour certains, ce degré de similitude est tel qu'une séquence de nucléotides pourrait servir de marqueur génétique d'identité d'une espèce. Ainsi, deux organismes qui partagent la même séquence nucléotidique pour un gène donné appartiendrait à la même espèce. C'est le principe du *DNA barcoding* introduit par Paul D. N. Hebert et son équipe en 2003 pour le règne animal (20) et que l'on pourrait traduire par « code-barre génétique ». Un code-barre est une suite de huit, dix ou treize chiffres (voire même quinze à dix-huit) dont l'application la plus commune est l'identification de produits pour le commerce. Cette suite de chiffres offre la possibilité de générer jusqu'à 30 000 milliards de combinaisons différentes. De la même manière, une séquence génétique, disons de 30 paires de bases, possèdent jusqu'à $1e+18$ combinaisons possibles. Cependant, du fait de pressions fonctionnelles, certains nucléotides ou séquences de nucléotides sont conservés à la même position entre espèces différentes, parfois même très éloignées, tout comme il existe des variations intraspécifiques. Pour le premier point, Hebert et al. suggèrent de ne s'intéresser qu'à des gènes codant des protéines. En effet, dans ce type de séquences, le nucléotide en troisième position de chaque codon est souvent moins conservé entre espèces et introduit donc de la diversité génétique. Cette particularité s'explique par la dégénérescence du code génétique : une mutation en troisième position d'un codon ne modifie pas forcément l'acide-aminé transcrit et a donc peu de conséquences fonctionnelles (20). Le deuxième point en revanche est plus préoccupant. Il suppose que deux individus n'ayant pas un degré de similitude de 100% entre leur séquence peuvent néanmoins appartenir à la même espèce. Mais alors, à partir de quel degré de similitude entre séquences peut-on parler d'espèce ?

Il n'existe a priori pas de valeur référence de pourcentage de similitude dans les séquences permettant d'attester l'appartenance de deux individus à une même espèce. Avant toute affiliation taxonomique moléculaire, il convient de déterminer la variation génomique pour chaque gène d'intérêt. Ceci d'autant plus que tous les gènes n'évoluent pas à la même vitesse, ce qui signifie que certains gènes accumulent plus de mutations que d'autres pour un même intervalle de temps. On parle d'horloge moléculaire. La vitesse d'évolution d'un gène dépend de son importance, c'est-à-dire s'il code pour une fonction essentielle à la survie de l'organisme ou non. Pour les séquences essentielles,

la majorité des modifications seront nuisibles et seulement une minorité seront neutres, c'est-à-dire sans conséquence. Ces gènes évoluent lentement et sont donc plus conservés entre individus. Ils permettront donc principalement de distinguer des individus à des rangs taxonomiques tels que l'ordre ou la famille. Au contraire, pour des séquences ayant une importance moindre, le nombre de mutations nuisibles sera plus faible par rapport au nombre de mutations neutres et le gène correspondant évoluera donc plus rapidement. Ces gènes seront plus à même de distinguer des individus ayant une histoire évolutive proche, autrement dit des individus pouvant être classés au rang du genre, de l'espèce ou même de la population (5).

Les études de communautés d'organismes se limitent rarement à l'identification d'une seule espèce. Le développement de nouvelles technologies de séquençage plus productives en termes de nombres et de longueurs de séquences a permis l'analyse de la diversité des espèces présentes dans un environnement donné. Le principe du *DNA barcoding* associé à ces nouvelles technologies de séquençage constitue le *metabarcoding*. Généralement, pour assigner une espèce à une séquence ADN, cette dernière est comparée avec des séquences bien connues issues de banques de données. Un arbre phylogénétique peut également être généré à partir des séquences connues et non connues : l'identification de l'espèce est alors fonction de la position dans l'arbre (21).

Le choix du marqueur génétique apparaît donc de grande importance. En effet, il doit être assez spécifique pour pouvoir identifier une seule espèce à la fois mais tout en ayant un certain degré de conservation pour pouvoir être retrouvé chez toutes les espèces à analyser. Toutes les régions de l'ADN sont-elles appropriées pour servir de marqueur génétique ?

II. Les marqueurs génétiques d'identification taxonomique

L'ADN est en grande partie contenu dans le noyau des cellules. Mais une petite partie est également présente dans les organites cytoplasmiques que sont les mitochondries chez les animaux. Il existe donc deux sources potentielles de marqueurs génétiques d'identification.

Comme nous l'avons vu précédemment, toutes les régions de l'ADN nucléaire et mitochondrial sont susceptibles d'accumuler des mutations mais certaines régions le sont d'avantage, comme les régions non-codantes ou les introns, parce qu'elles ne sont pas soumises à des pressions fonctionnelles. Au contraire, les gènes sont associés à une fonction particulière et sont donc peu propices à l'accumulation de mutations spontanées puisqu'une altération de la fonction en question peut être délétère pour la survie de l'organisme. Par conséquent, quelles régions de l'ADN peuvent servir de marqueur génétique d'identification ?

A. Les séquences nucléaires

1. L'ADN ribosomique

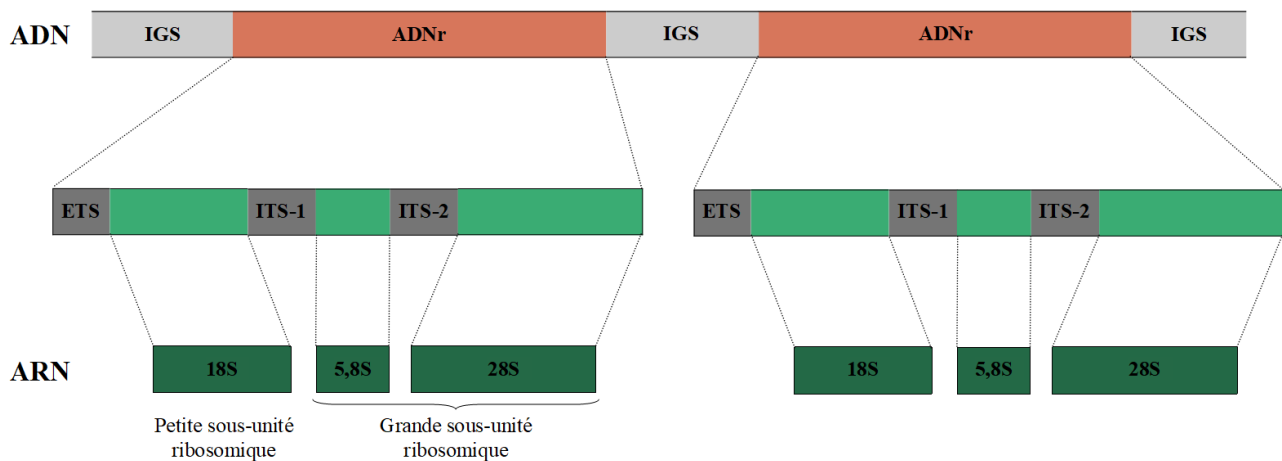


Figure 7 : Schéma de l'organisation des séquences nucléaires ribosomiques des Eucaryotes

L'ADN ribosomique (ADNr) correspond aux séquences d'ADN codant les ARN ribosomiques (ARNr) qui, associés à des protéines, forment le corps des ribosomes, acteurs de la machinerie moléculaire. Cet ADN est constitué par une répétition en tandem de domaines intergéniques IGS (InterGenic Spacer), des ETS (External-Transcribed Spacer), ITS (Internal-Transcribed Spacer) et les ADNr de la petite sous-unité ribosomique (ADNr 18S) et de la grande sous-unité ribosomique (ADNr 28S et 5,8S) (figure 7). La présence d'exemplaires multiples de ces séquences permet à la cellule de produire la quantité adéquate d'ARN ribosomiques afin de répondre à ses besoins importants en ribosomes. Cela constitue également un intérêt pratique dans la détection de la séquence au sein d'un échantillon biologique puisqu'elle est naturellement présente en quantité importante et peut donc être facilement amplifiée.

Alors qu'il serait logique de penser que chaque copie évolue indépendamment des autres, accumulant des mutations différentes, des études ont prouvé le contraire. Ces séquences répétées et dispersées au sein du génome sont plutôt bien conservées, avec un faible taux de variation intragénomique, suggérant une évolution dite concertée de ces séquences. Ce processus ferait intervenir le mécanisme d'homogénéisation par crossing-over inégaux (ou recombinaison inégale) et de

conversion de gènes (22). De plus, la longueur de la matrice d'ADNr est à peu près similaire chez les espèces de nématodes : de 6 500 à 7 500 bases ; permettant ainsi la comparaison de ces séquences par alignement (23).

2. Les séquences non codantes

ETS et ITSs correspondent à des introns. Ce sont des séquences non codantes situées au sein d'un transcrit primaire qui sont enlevées du transcrit au cours de la maturation (4). Contrairement aux régions codantes 18S, 5.8S et 28S, elles seraient donc soumises à une pression de sélection moins importante, résultant en un nombre plus élevé de mutations au sein de ces séquences et donc à un certain niveau de diversité génétique entre des organismes. Quel niveau taxonomique chez les nématodes gastro-intestinaux les séquences ETS et ITS permettent-elles de discriminer ?

Une étude destinée à distinguer les espèces au sein du genre *Haemonchus sp.* a mis en évidence l'existence de 2 types dominants et différents en longueur de séquences ETS chez *H. contortus*, absents chez *H. placei*. Il en résulte que l'utilisation d'amorces spécifiques couvrant cette région dans une réaction de polymérisation en chaîne, *polymerase chain reaction* en anglais (PCR) permettrait de distinguer ces deux espèces (14).

Les séquences ITS-1 et ITS-2 ont été plus largement utilisées comme marqueur génétique. Un de leur avantage est qu'elles sont entourées de régions conservées (séquences 18S et 5.8S, 5.8S et 28S respectivement), facilitant ainsi leur isolement par l'emploi d'amorces universelles. La longueur d'ITS-1 est légèrement plus importante (364 à 522 paires de bases) que celle d'ITS-2 (215 à 484 paires de bases) (23). Au sein des espèces *Ostertagia ostertagi* et *Ostertagia lyrata*, la séquence d'ITS-1 présente par ailleurs la particularité d'être plus longue (823 paires de base) par rapport aux autres nématodes, du fait de la présence d'une répétition en tandem de deux séquences (11). L'intérêt diagnostique de cette particularité a par la suite été exploité avec succès en l'utilisant comme marqueur spécifique afin d'identifier ces deux espèces au sein d'un mélange de parasites issus de genres différents (24,25).

Des études similaires cherchant à caractériser les séquences ITS-1 et ITS-2 afin d'en déterminer l'intérêt comme marqueur génétique ont été conduites pour d'autres genres de nématodes : *Cooperia sp.* (12,26,27), *Haemonchus sp.* (15), *Trichostrongylus sp.* (28–31), *Nematodirus sp.* (32–34), *Teladorsagia sp.* (10), *Chabertia sp.* et *Oesophagostomum sp.* (35,36), *Marshallagia sp.* (37).

Comme mentionné précédemment, ces séquences ne codent pas pour une fonction en particulier mais il a été montré chez des levures qu'elles seraient tout-de-même impliquées dans les processus de maturation des ARNr (38,39). En effet, des modifications structurales dans les régions ITS peuvent être à l'origine d'anomalie de la structure secondaire, aboutissant à des ARNr non fonctionnels (22). Il en résulte que ces séquences sont certes plus soumises à variation que les séquences codantes adjacentes mais conservent néanmoins un certain niveau d'homogénéité intraspécifique (40).

Aucune preuve de variation dans les séquences nucléotidiques de différents stades biologiques d'un même individu n'a été mise en évidence (30,32,35,41). Ceci est en faveur de l'utilisation de ses séquences pour n'importe quel type d'échantillons, sans avoir besoin d'isoler les œufs, les larves ou les individus adultes.

Trois types de variations ont par ailleurs été mis en évidence au sein des séquences ITS et également étudiés pour d'autres types de séquences. Tout d'abord, la variation interspécifique correspond aux différences entre individus d'espèces différentes, appartenant au même genre ou non. La variation intraspécifique correspond à la variation qui existe entre individus de la même espèce mais pouvant être issus de régions géographiques différentes. Pour certains scientifiques l'existence d'une telle variation signifie que l'homogénéisation de ses séquences est incomplète et limite leur usage dans les analyses phylogénétiques puisqu'un individu ne peut pas être représentatif de son espèce (42). Pour d'autres, tant que le niveau de variation intraspécifique est inférieur au niveau de variation interspécifique, les séquences conservent leur intérêt comme outils d'étude moléculaire (30). Cette hypothèse s'accorde avec les études pour lesquelles l'espèce des parasites analysés a clairement été déterminée sur des critères morphologiques notamment. Mais cela ne peut s'appliquer pour des études cherchant à statuer sur l'appartenance ou non à la même espèce de deux individus morphologiquement proches ; d'autant que des variations au sein des copies d'ADNr d'un même individu ont été identifiées, correspondant à de la variation intra-individuelle ou intra-génomique (10,15,30,31). Les variations relevées dans les séquences entre deux individus ne peuvent donc être uniquement attribuées à de la variation interspécifique et pourraient tout aussi bien relevées de la variation intraspécifique ou intra-génomique de ses séquences.

3. Les séquences codantes

Les séquences codantes du génome ribosomique, constituées par l'ADNr 18S, 5.8S et 28S, sont plutôt conservées du fait de leur implication dans la machinerie cellulaire, ce qui représente un intérêt pour leur comparaison entre différents individus.

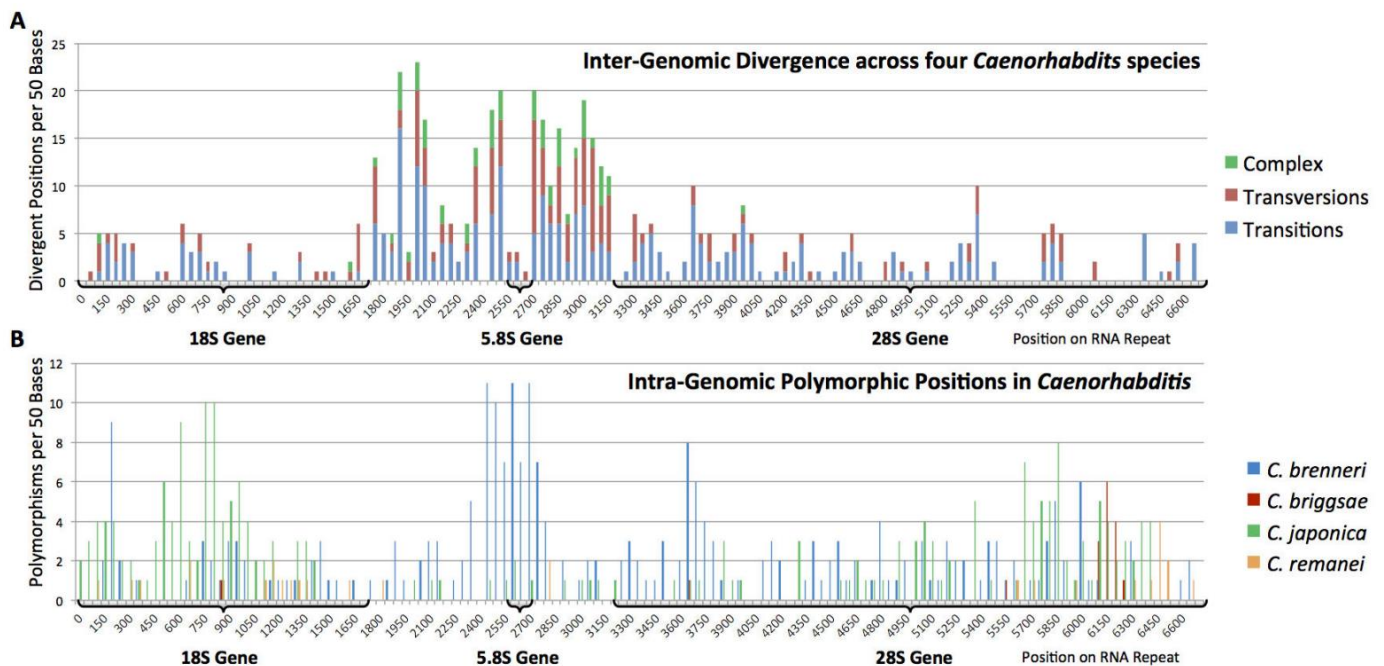


Figure 8 : Graphique des variations observées au sein des gènes ribosomiques (22)

La figure 8 présente la variabilité des séquences ribosomiques entre espèces appartenant au même genre : *Caenorhabditis* sp. (A) ; et la variabilité au sein des copies d'ADNr d'un même individu pour plusieurs espèces (B). Le polymorphisme interspécifique (A) semble plutôt localisé au niveau des séquences ITS alors que le polymorphisme intraspécifique ou intra-génomique (B) est plutôt localisé au niveau des séquences codantes 18S, 5.8S et 28S. Le faible niveau de variation interspécifique de ces dernières les rend peu propices à leur utilisation comme marqueur génétique.

La conservation fonctionnelle de ces séquences est telle qu'elle ne permet pas la différenciation des individus à des rangs inférieurs tels que le genre ou l'espèce.

Par exemple, dans le genre *Haemonchus sp.*, la comparaison des séquences nucléotidiques de l'ADNr 18S de trois espèces (*H. contortus*, *H. placei* et *H. similis*) a montré 100% d'identité, signifiant ainsi qu'il n'est pas possible de distinguer ces trois espèces à partir de cette séquence. De même, la comparaison avec d'autres espèces de genre (*Ostertagia ostertagi*) ou même de sous-famille (*Nematodirus battus*) différents ne montre que très peu de variation (1.0 et 2.0% de divergence respectivement) (14).

Une étude portant sur la séquence 5.8S de 18 espèces de nématodes appartenant à des super-familles et des familles différentes a montré un très faible niveau de variation (de 0 à 5,8% de différence). En particulier, parmi les huit espèces appartenant à la super-famille des *Trichostrongylidea*, sept présentaient des séquences identiques ; *Trichostrongylus tenuis* différait des autres séquences pour trois positions. Ce grand pourcentage de similarité révèle l'intérêt limité de cette séquence pour les études phylogénétiques ou taxonomiques des nématodes au rang de la super-famille et pour les rangs inférieurs. Néanmoins, les pourcentages de différences avec d'autres espèces telles que *Caenorhabditis elegans*, *Heterodera avenae* et *Heterodera schachtii*, sont suffisamment élevés (12,3 à 15,6% ; 30 à 31,3 % respectivement) pour attester de l'utilité de cette séquence dans les études au rang taxonomique de l'ordre (43).

L'ADNr 28S est constitué de deux types de séquences : des segments structuraux relativement conservés, et des segments hautement divergents appelés domaines D (44). Des différences significatives entre espèces au sein de ces domaines variables ont été mises en évidence, seulement dans le cas où ces espèces ne sont pas trop éloignées phylogénétiquement (45).

Les séquences ribosomiques codantes semblent plus appropriées pour les études phylogénétiques (45–48). En effet, alors que les séquences ITS présentent certes l'avantage d'être hautement divergentes entre taxons et d'être entourées de séquences conservées, il est difficile de les comparer par alignement du fait des variations intraspécifiques à la fois dans la taille et dans la séquence. Au contraire, la longueur du gène 18S est relativement constante et présente également la particularité de contenir à la fois des régions conservées et des régions divergentes (49) tout comme le gène 28S.

B. Les séquences mitochondriales

1. Les particularités de l'ADNm

Les mitochondries sont les organites cellulaires où se déroulent les mécanismes de la respiration cellulaire et de la production énergétique chez la plupart des organismes eucaryotes. Ces organites présentent la particularité de posséder leur propre système génétique. Même si la grande majorité des protéines mitochondriales sont codées par le génome nucléaire, il persiste toujours un génome mitochondrial codant notamment pour des protéines de la chaîne respiratoire, des ARNr et des ARN de transfert (ARNt). Une des hypothèses pouvant expliquer la persistance de ce système génétique distinct est le fait que la plupart des protéines non-ribosomiques codées par les mitochondries sont de nature hautement hydrophobe. Leur transport à travers la membrane mitochondriale depuis le cytoplasme serait alors trop coûteux en énergie (50).

L'ADNm est circulaire. Sa taille est relativement conservée chez les nématodes, le nombre total de paires de bases étant compris entre 13 519 (chez *Nematodirus spathiger*) et 15 221 (chez *Mecistocirrus digitalis*). Il code pour douze protéines : les trois sous-unités de la cytochrome oxydase (*cox1-3*), six sous-unités de la nicotinamide déshydrogénase (*nad1-6*) ainsi que la sous-unité 4L (*nad4L*), le cytochrome b et la sixième sous-unité de l'adénosine triphosphate (ATP) synthase (*atp6*) ; deux ARNr (la petite et la grande sous-unités ribosomiques) et vingt-deux ARNt. Il comprend également deux régions non codantes (notées NC) (51–56). L'agencement des gènes le long de la molécule d'ADN correspond à l'ordre de transcription des gènes. Cet ordre n'est pas le même pour tous les organismes du grand phylum des Nématodes. L'agencement présenté dans la figure 9 et correspondant à l'agencement du génome de *Marshallagia marshalli*, est commun à l'ordre des Strongylidés. Il s'agit de l'arrangement GA2 (57).

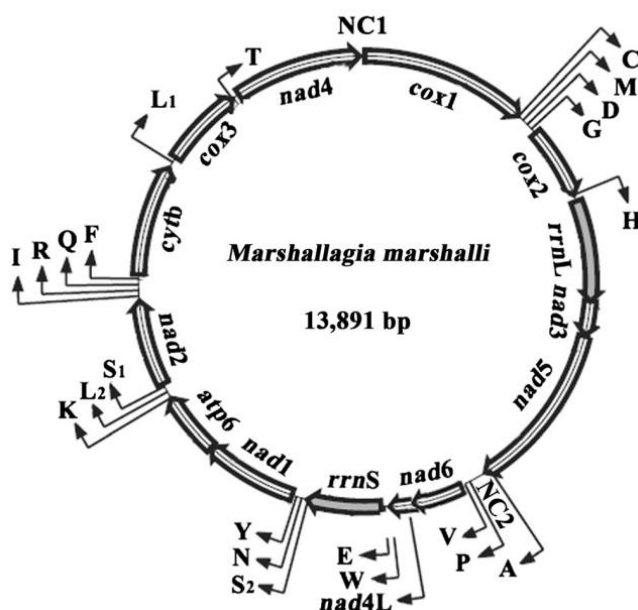


Figure 9 : Agencement du génome mitochondrial de *Marshallagia marshalli*. (55)

La nomenclature des gènes reprend la nomenclature standard. Seuls les gènes des 22 ARNt sont représentés par la première lettre de l'acide-aminé correspondant, avec un numéro permettant de différencier les deux ARNt spécifiques de la leucine et de la sérine.

Dans le règne animal, l'ADNm est haploïde, ce qui signifie que le génome mitochondrial est présent en un seul exemplaire, contrairement à l'ADN nucléaire qui est diploïde. De plus, l'ADNm est principalement issu des mitochondries maternelles du fait des processus de maturation des spermatozoïdes et de l'oocyte qui éliminent les mitochondries paternelles. On parle d'hérédité maternelle (58). Il en résulte que les variations observées dans les séquences nucléotidiques mitochondriales sont en majorité issues de différences dans la fréquence et/ou la pression sélective des mutations, et dans une moindre mesure, des différences de généalogie (59). Du fait du caractère haploïde et de la transmission maternelle, la quantité d'ADNm au sein d'une population est donc inférieure à la quantité d'ADN nucléaire. L'ADNm est par conséquent plus sujet à la dérive génétique et est donc plus variable que l'ADN nucléaire (60).

De plus, les mutations au sein de l'ADNm sont d'autant plus favorisées que la réplication de l'ADNm est moins fidèle et qu'il n'y a pas de système de réparation comme pour le génome nucléaire. Cela signifie que le génome mitochondrial évolue plus rapidement que le génome nucléaire. Il

constitue donc un bon support pour l'étude de phénomènes évolutifs récents comme l'apparition d'une nouvelle espèce. En particulier, l'ADNm serait plus à même de distinguer des espèces cryptiques, autrement dit des espèces ayant divergé il y a peu de temps. La variation interspécifique est bien plus importante que la variation intraspécifique permettant une délimitation plus fiable des espèces (61,62).

2. Les gènes mitochondriaux d'intérêt

Le gène *cox1* code pour la sous-unité 1 du complexe cytochrome c oxydase qui intervient dans les réactions d'oxydo-réduction de la chaîne respiratoire mitochondriale. La séquence nucléotidique de ce gène possède des régions relativement conservées qui en font un excellent support pour la construction d'amorces universelles, qu'elles soient communes à plusieurs phylums (63) ou spécifiques des nématodes (64), permettant ainsi la récupération de cette séquence même chez des espèces inconnues. De plus, contrairement aux autres gènes mitochondriaux, la variabilité de la séquence en acide-aminés du gène *cox1* est plus faible (59) du fait de son importance fonctionnelle. Il en résulte qu'il est possible, à partir de l'étude des substitutions dans la séquence en acide-aminés, de déterminer le phylum ou l'ordre d'un organisme non identifié, avant de déterminer ses rangs taxonomiques inférieurs par l'étude des variations dans sa séquence nucléotidique (20). Associés aux caractéristiques propres à l'ADNm que sont un taux d'évolution élevé et une héritabilité maternelle, le gène *cox1* a été choisi comme marqueur génétique de référence pour le *DNA barcoding* (65,20). Cependant, bien que le gène *cox1* soit capable de différencier des espèces entre elles, voire même des populations de parasites (51), ce résultat ne peut être généralisé à l'ensemble des nématodes. En effet, pour la sous-famille des *Ostertagiinae* par exemple, la divergence interspécifique au sein du gène *cox1* n'est pas suffisante pour permettre de délimiter correctement toutes les espèces (66).

Les autres gènes mitochondriaux fréquemment employés sont les gènes *nad*, en particulier le gène *nad4*, qui codent pour des sous-unités de la nicotinamide déshydrogénase. Il s'agit d'une enzyme également impliquée dans la chaîne respiratoire mais dont la séquence en acides aminés est moins conservée que celle des cytochromes oxydases (62). Il en résulte que la séquence en nucléotides est beaucoup plus variable et est donc plus appropriée pour les études à l'échelle de la population (61) en conjonction avec d'autres marqueurs comme ITS-2 par exemple. Il a d'ailleurs été montré, pour l'espèce *Haemonchus contortus* notamment, une diversité génétique très importante au sein des populations de parasites associée à une faible différenciation selon leur origine géographique à l'échelle du continent, indiquant un flux génique important entre ces populations. Au contraire, les populations issues de continents différents montrent plus de différenciation génétique. Une explication à cette homogénéisation génétique serait le mouvement des hôtes entrepris par l'homme. D'un point de vue épidémiologique, la conséquence de ce flux génique est la propagation plus rapide de gènes tels que les gènes de résistance aux anthelminthiques (67-72). Chez les animaux sauvages, étant donné qu'il y a moins de mouvements et moins de pression sélective liée aux anthelminthiques, les populations de parasites seraient plus isolées génétiquement et donc plus différenciables (73). La séquence *nad4* a également servi pour déterminer les relations entre espèces étroitement apparentées telles que *H. contortus* et *H. placei* (19) ou *H. contortus* et *Ashworthius sidemi*, un nématode hématophage également parasite des cervidés sauvages (74) ; ou encore pour statuer sur l'identité d'espèces cryptiques avec notamment la différenciation de *Teladorsagia boreoarcticus* par rapport au complexe cryptique *T. circumcincta/T. trifurcata* (68).

CONCLUSION PARTIELLE

Les marqueurs génétiques aujourd'hui employés sont issus de deux sources : l'ADN nucléaire, et plus particulièrement l'ADNr ; et l'ADNm. Le choix de l'un ou de l'autre s'appuie sur des analyses de variabilités intra- et interspécifique des séquences nucléotidiques correspondantes. Du fait des propriétés inhérentes à l'ADNm, ce dernier évolue plus rapidement que l'ADN nucléaire et possède donc une variabilité intraspécifique assez importante. Il est donc plus propice pour les études à l'échelle de la population, autrement dit entre des individus ayant une histoire évolutive très récente. Parmi les gènes mitochondriaux, le gène *cox1* a largement été employé parce qu'il possède tout de même un certain degré de conservation lié à la fonction de la protéine codée. Un autre gène mitochondrial souvent utilisé est le gène *nad4* pour lequel le degré de variabilité intraspécifique est assez élevé. Parmi l'ADN nucléaire, les études se sont essentiellement concentrées sur l'ADNr parce qu'il est universel, présents en large quantité dans le noyau et possède un degré de variabilité intraspécifique relativement faible comparé au degré de variabilité interspécifique. Les séquences 18S, 5.8S et 28S, parce qu'il s'agit de séquences codantes, restent très conservées, même entre individus d'espèces différentes. Les séquences ITS, au contraire, subissent moins de pression de sélection. La variabilité au sein de ses séquences permet de distinguer des individus au rang de l'espèce.

Face à ces niveaux de variabilités, pouvant différer entre espèces, il est conseillé de ne pas utiliser un seul marqueur génétique mais de les combiner. De plus, les données moléculaires ne peuvent seules résoudre la taxonomie des nématodes. Une approche intégrative incluant la morphologie et l'écologie des espèces est recommandée.

III. Les méthodes de séquençage de l'ADN

Le terme séquençage ici correspond à l'ensemble des processus et techniques depuis l'extraction de l'ADN jusqu'à la détermination de sa séquence nucléotidique.

A. Le choix d'une méthode

Il y a plusieurs paramètres à prendre en compte dans le choix des méthodes de séquençage. Tout d'abord, il y a la nature du résultat recherché. En effet, de nombreuses applications du séquençage de l'ADN sont possibles, ne se limitant pas à la simple identification du rang taxonomique d'un organisme. Des outils de diagnostic ou encore l'identification de facteurs de pathogénicité ou de résistance aux médicaments sont également développés à partir de ses méthodes. Le résultat obtenu peut être de nature qualitative ou bien quantitative, lorsque l'on recherche l'abondance d'une espèce dans un milieu par exemple. Il peut s'agir de s'intéresser à une espèce ou un groupe d'espèces en particulier ou bien étudier la diversité d'une population. Le nombre d'échantillons à analyser est également à prendre en considération, de même que le temps et le coût (75).

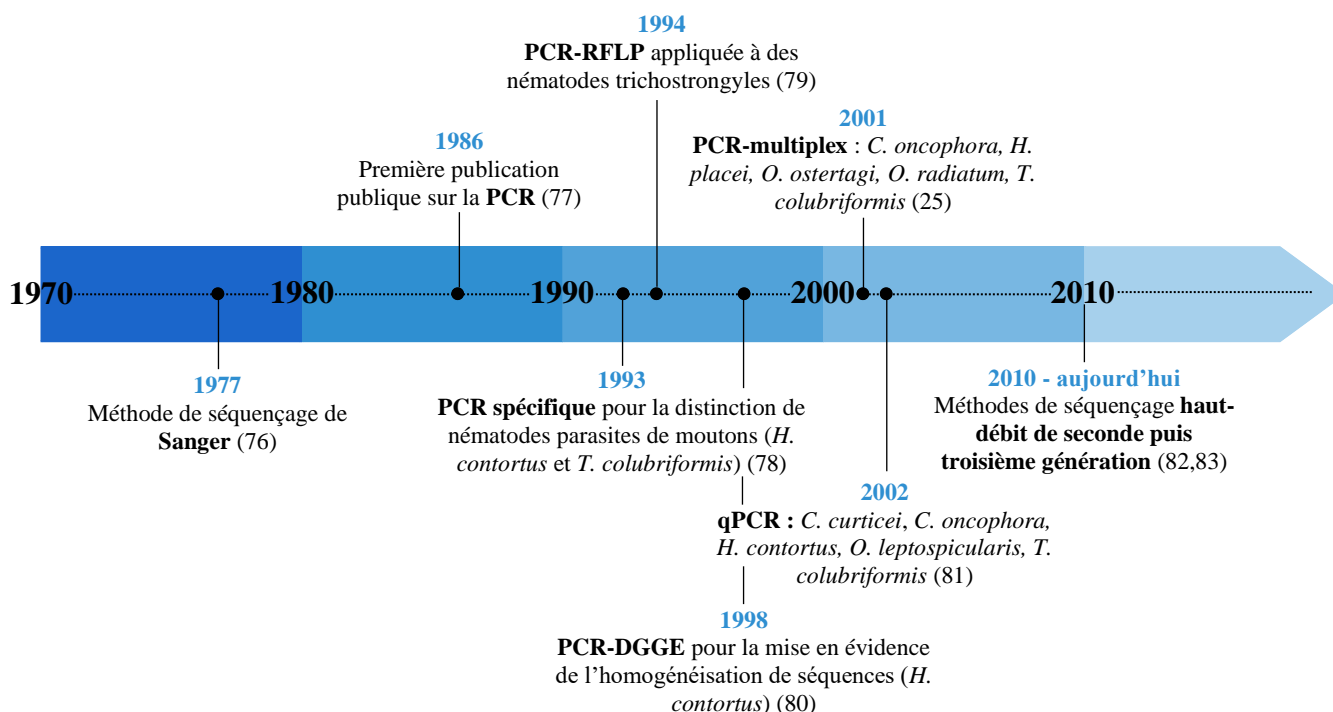


Figure 10 : Chronologie des différentes méthodes d'identification des Nématodes gastro-intestinaux

Dans cette thèse, nous nous limiterons aux méthodes appliquées à l'identification des nématodes et plus particulièrement des nématodes gastro-intestinaux des Caprinés (figure 10).

B. L'amplification de séquences d'ADN

1. Récupération de l'ADN

L'étape préalable à toute analyse moléculaire est la récupération du matériel génétique à partir des échantillons récoltés.

Outre l'élimination des débris cellulaires et autres contaminants, la purification des échantillons nécessite également l'élimination des inhibiteurs de la PCR. Ces derniers sont naturellement présents et peuvent agir par inactivation de l'ADN polymérase, enzyme nécessaire à la réplication de l'ADN ; par dégradation ou capture des acides nucléiques, ou bien par interférence

avec l'étape de lyse cellulaire (84). Une première étape de broyage peut être obtenue par ajout d'un mélange de sodium dodecyl sulfate et de protéinase K (55) ou par broyage mécanique. Par la suite, différents kits de purification sont disponibles, utilisant divers procédés : la chromatographie ionique (les colonnes Nucleobond® de chez Macherey-Nagel, le QIAGEN Genomic-tip® de chez Quiagen) ou la précipitation des protéines (le Wizard® Genomic DNA Purification Kit de chez Promega). Une séparation par phase est également possible avec le phénol-chloroforme. Du fait de la grande résistance de la cuticule des nématodes aux actions chimiques, enzymatiques et physiques, des adaptations aux protocoles classiques peuvent être nécessaires. Par exemple, le temps d'incubation avec des protéinases K peut être augmenté, les étapes de centrifugation peuvent être répétées, une étape supplémentaire de précipitation des protéines peut être ajoutée. Enfin une dernière étape de purification consiste en l'ajout d'isopropanol qui va permettre la précipitation et donc la concentration de l'ADN (85).

La quantité de matériel génétique récupérée doit être suffisamment importante pour permettre une bonne spécificité et sensibilité des analyses ultérieures.

2. Le principe de la PCR

La proportion d'ADN codant est souvent infime par rapport à l'ADN non-codant, ce qui implique que la séquence des gènes d'intérêt est souvent noyée dans le reste des séquences nucléotidiques. D'autre part, la quantité d'ADN dans un échantillon est souvent trop faible pour pouvoir être analysée telle quelle. Les scientifiques ont donc mis au point des techniques permettant d'augmenter le nombre de copies de ces gènes : c'est le principe du clonage. Parmi ces techniques, la PCR permet l'amplification *in vitro* de n'importe quel segment spécifique d'une ou plusieurs molécules d'ADN pouvant être présentes en faible quantité. Cette technique automatisée est rapide et efficace puisqu'elle permet la production de milliards de copies en quelques heures alors que les méthodes précédentes pouvaient prendre plusieurs jours. Le principe de la PCR repose sur la répétition cyclique d'une répllication ciblée de l'ADN associée à des variations de température (figure 11).

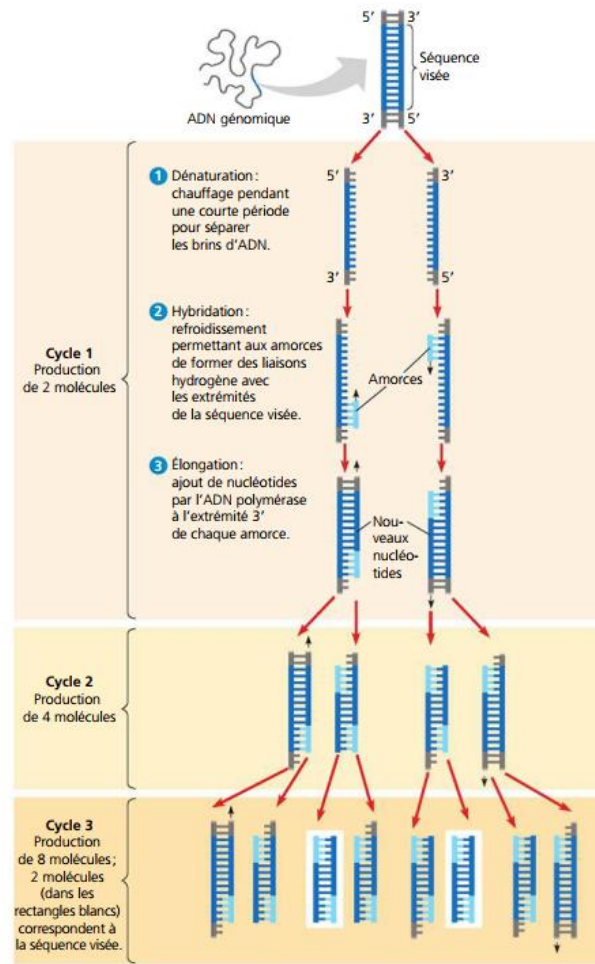


Figure 11 : Schéma du principe de la PCR (86)

Un cycle ne dure que quelques minutes et est composé de trois étapes :

- une première étape de dénaturation au cours de laquelle le mélange réactionnel contenant les molécules d'ADN est chauffé afin de séparer les brins d'ADN ;
- un refroidissement permet ensuite la fixation d'amorces sur l'ADN. Ces amorces correspondent à de courtes molécules d'ADN monocaténaire (15 à 30 nucléotides) dont la séquence est complémentaire de l'extrémité 3' de la séquence d'intérêt ;
- une ADN polymérase permet ensuite la synthèse du reste de la séquence de nucléotides. Cette ADN polymérase présente la particularité d'être thermostable puisque dans le cas contraire, elle serait dénaturée en même temps que l'ADN à la première étape.

Les produits obtenus sont ensuite séparés par électrophorèse en fonction de leur charge électrique et de leur taille et sont visualisés sous ultra-violet après ajout de bromure d'éthidium.

A la fin de chaque cycle, le nombre de molécules d'ADN est doublé. Ce n'est qu'au bout du troisième cycle que des molécules correspondant exactement à la séquence d'intérêt sont créées. Ainsi, après 33 cycles, le mélange contient environ un milliard de copies de la séquence d'intérêt (86).

3. Les paramètres de la PCR

La température optimale pour chaque étape est indiquée par le fabricant du thermocycleur utilisé. Elle est généralement supérieure à 90°C pour l'étape de dénaturation, entre 55°C et 60°C pour l'étape d'hybridation et autour de 70°C pour l'étape d'élongation (87).

La spécificité de la PCR réside dans le choix des amorces : la séquence doit être assez grande pour pouvoir se fixer sur l'ADN mais assez courte pour ne s'intéresser qu'à la séquence d'intérêt. En général, la longueur des amorces n'excède pas 30 nucléotides. Pour certaines applications, l'amorce doit correspondre exactement à la séquence d'intérêt. Mais pour d'autres applications, pour lesquelles la séquence recherchée n'est pas entièrement connue ou pour l'identification d'éventuelles mutations, l'amorce peut ne pas être exactement identique et même intentionnellement différente. Il est préférable que les différences se localisent au niveau de l'extrémité 5' de l'amorce et non au niveau de l'extrémité 3' afin de ne pas perturber la fixation de l'ADN polymérase. Outre la spécificité avec la séquence d'intérêt, un des paramètres à prendre en compte dans la détermination de la séquence nucléotidique de l'amorce est la conformation de la structure secondaire. En effet, la succession de bases pyriques ou pyrimidiques est à l'origine de boucles dans la structure secondaire pouvant rendre la fixation à l'ADN plus incertaine. D'autre part, lorsque des amorces différentes sont présentes dans le mélange réactionnel, il faut veiller à ce qu'aucune ne puisse s'hybrider avec une autre et être à l'origine de produits d'amplification non souhaités. Cela peut avoir lieu notamment lorsque les extrémités 3' de deux amorces sont complémentaires, l'ADN polymérase entreprenant alors l'élongation de chaque amorce jusqu'à la fin de l'autre comme présenté dans la figure 12 (87). La composition en nucléotides de la séquence de l'amorce et sa longueur vont également influencer la température d'hybridation (88,87).

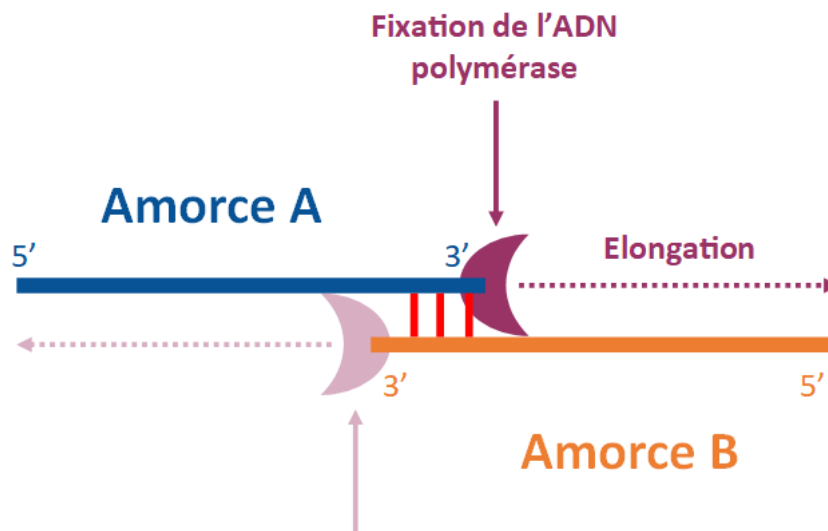


Figure 12 : Schéma d'un dimère d'amorces

Plusieurs paramètres peuvent être sources d'erreurs intrinsèques à la technique de PCR. Tout d'abord, il apparaît évident que plus la quantité de matériel génétique au départ est faible, plus le risque de générer des erreurs est important. De même, plus le nombre de cycles est élevé, plus il est possible de multiplier une erreur commise lors de cycles précédents. Le choix de l'ADN polymérase est également important puisqu'elles ne présentent pas toutes la même fiabilité, illustrée par le taux d'erreur par nucléotide par cycle. Certaines possèdent d'ailleurs la capacité de relire dans le sens 3'-

5' la séquence de nucléotides nouvellement synthétisée, de repérer des erreurs d'appariement et de les corriger par excision du nucléotide mal apparié : on parle de « correction d'épreuve ». Cette aptitude diminue bien évidemment le taux d'erreur possible. La composition du milieu réactionnel est également importante : la concentration des quatre désoxyribonucléotides, qui doit être équivalente ; le pH et la concentration en ions magnésium. Enfin, la composition en nucléotides de la séquence d'intérêt influe sur le type de mutation possible. En effet la probabilité de mauvais appariements entre deux bases puriques est plus élevée qu'entre deux bases pyrimidiques ou entre une base purique et une base pyrimidique (87,89).

4. Les méthodes basées sur le principe de la PCR

La méthode de la PCR a par la suite été couplée à d'autres techniques afin de mieux caractériser les produits obtenus tout en gardant la capacité de travailler à partir de petites quantités de matériel génétique.

La technique de PCR - Restriction Fragment Length Polymorphism (PCR-RFLP) ou polymorphisme de longueur des fragments de restriction fut l'une des premières à être développée. Des produits de PCR obtenus à partir d'espèces ou de populations différentes sont digérés par des enzymes de restriction spécifiques, c'est-à-dire dont le site de restriction est connu pour être localisé dans une séquence d'intérêt, comme la séquence d'ADN nucléaire ITS-2 (79). Les fragments alors obtenus sont séparés selon leur taille par électrophorèse. Ainsi les variations dans les séquences ADN, et donc dans la position des sites de restriction, sont mises en évidence par des profils électrophorétiques différents.

Néanmoins, l'un des inconvénients de la technique de PCR-RFLP est l'existence de variations intraspécifiques au sein des sites de restriction, à l'origine d'erreur d'interprétation des fragments. Diverses méthodes ont alors été développées pour mettre en évidence ces mutations, parmi lesquelles se trouve la PCR-Denaturing Gradient Gel Electrophoresis (PCR-DGGE) ou électrophorèse sur gel en gradient dénaturant. Les produits de PCR sont introduits dans un gel de polyacrylamide et soumis à un gradient d'agents dénaturants à haute température. La position sur le gel à partir de laquelle la molécule d'ADN commence à être dénaturée dépend de sa composition en nucléotides et non de sa longueur. Ainsi des fragments ayant la même taille mais qui diffèrent de par leur séquence en nucléotides, même pour une seule base, ne seront pas localisés à la même position sur le gel (80).

En parallèle de ces méthodes se développe la technique de la PCR en temps réel ou PCR quantitative (qPCR). L'un des intérêts de cette technique par rapport aux autres est que toutes les étapes peuvent être réalisées dans un « tube fermé » car elle ne nécessite pas d'étape supplémentaire de caractérisation des produits d'amplification pouvant être source de contamination. Le principe est le suivant : un marqueur fluorescent est incorporé à la synthèse de chaque produit PCR, permettant ainsi de mesurer en temps réel le nombre de copies synthétisées à chaque cycle. Le nombre de cycles de PCR nécessaire pour atteindre un certain seuil de fluorescence est relevé. Il s'agit de ce l'on nomme la valeur « Ct » pour « *Crossing threshold* » qui signifie « seuil de franchissement ». Cette valeur est liée à la quantité d'ADN initiale par la formule : $quantité\ d'ADN\ seuil = quantité\ d'ADN\ initiale \times 2^{Ct}$; puisque la quantité d'ADN double à chaque cycle en théorie. Le résultat peut alors être quantitatif lorsque l'on s'intéresse à un échantillon ou bien semi-quantitatif lorsque plusieurs échantillons sont analysés, pouvant inclure un échantillon dont la quantité initiale d'ADN est connue. Les composés fluorescents utilisés tels que le *SYBR Green I* par exemple sont capables de se fixer sur n'importe quelle molécule d'ADN double brin du mélange, c'est-à-dire aussi

bien la séquence cible amplifiée que les éventuels dimères d'amorces ou des séquences non cibles. Pour pallier à ce problème, une analyse des courbes de fusion peut être réalisée. Il s'agit d'une étape supplémentaire à la PCR au cours de laquelle le mélange est chauffé. Les amplicons, alors sous forme double brin, sont tous fixés par le composé fluorescent *SYBR Green*. Au fur et à mesure que la température augmente, l'ADN se dénature et libère le *SYBR Green* qui n'émet alors plus de fluorescence. La température de dénaturation d'un brin d'ADN dépend de la composition en nucléotides de sa séquence. Ainsi la courbe de fusion, qui représente la fluorescence en fonction de la température, est spécifique d'une molécule d'ADN. Des courbes de référence peuvent être établies pour des espèces dont la séquence nucléotidique est connue (90). Une autre alternative peut être l'utilisation de sondes fluorescentes spécifiques. En effet, ces sondes correspondent à de courtes séquences nucléotidiques qui, contrairement aux composés fluorescents, s'hybrident spécifiquement avec la séquence ADN souhaitée. Des sondes peuvent donc être utilisées pour mesurer l'amplification d'une séquence ADN d'une espèce ou d'un genre en particulier (81).

Les méthodes décrites précédemment permettent la différenciation d'une seule espèce à la fois, si bien que pour analyser un mélange de parasites, situation la plus fréquente de co-infection, il est nécessaire de réaliser des réactions séparées, ce qui augmente considérablement le temps de d'analyse. La méthode de PCR multiplexe utilise une combinaison d'amorces spécifiques différentes pour identifier simultanément plusieurs espèces dans un même échantillon. L'efficacité et la fiabilité de cette technique dépendent de la concentration en amorces et en ADN : plus la concentration en ADN est faible par rapport à la concentration en amorces, plus les erreurs d'amplification sont fréquentes. De plus, l'affinité des amorces et la taille des fragments amplifiés sont variables en fonction des espèces, ce qui peut être à l'origine d'un biais d'amplification (25,41,81). Du fait de l'existence de variations intra- et interspécifiques ainsi qu'intra-individuelles, il est recommandé de tester la spécificité propre à chaque amorce utilisée dans le format multiplex, avant application dans une région géographique différente (90). Autrement dit, il serait nécessaire de créer un lot d'amorces spécifiques d'une région donnée et non réutiliser celles provenant d'études menées dans d'autres pays (91).

Sur le principe de la PCR multiplexe, une autre méthode a été développée : la PCR multiplexe en tandem (MT-PCR). Le protocole expérimental se décompose en deux étapes : une première étape d'amplification de l'ADN en présence du mélange d'amorces (PCR multiplexe standard) mais limitée à 10-15 cycles ; puis une seconde étape d'amplification de chaque séquence cible dans des réactions distinctes. Le nombre limité de cycles de la première étape permet d'enrichir le milieu en séquences cibles tout en limitant les interactions entre amorces pouvant être à l'origine de la production de séquences artéfactuelles ou de biais d'amplification. De plus, une étape de dilution précède la seconde étape d'amplification afin de diminuer la quantité d'inhibiteurs de PCR. La MT-PCR aspire donc à limiter les risques d'amplifications non spécifiques par cette deuxième étape de réactions en tandem. Un autre argument avancé est l'utilisation d'un nombre limité de marqueurs de fluorescence. En effet, dans une PCR multiplexe classique, la différenciation des séquences cibles est permise par l'utilisation de sondes émettant une fluorescence de longueurs d'onde différentes. Or de telles sondes sont coûteuses et requiert l'utilisation de thermocycleurs à multiples canaux. Très souvent le nombre de canaux de longueurs d'onde différentes est de quatre à six, limitant ainsi le nombre d'espèces ou de genres pouvant être identifié en même temps (92,93).

Enfin, une dernière technique assez récente se base sur la présence de sites polymorphiques uniques entre des séquences. Une première PCR cible une portion de l'ADN qui possède des régions

conservées parmi plusieurs genres de nématodes gastro-intestinaux d'intérêt (région ITS-2). Puis une méthode de « mini-sequencing SNaPshot » est appliquée. Elle consiste à hybrider une seule amorce à l'intérieur des séquences obtenues lors de la première étape et à allonger ces amorces d'un seul nucléotide. Le mélange réactionnel contient des nucléotides A T C et G marqués chacun d'un fluorochrome différent. Les fragments, très courts (séquence d'une amorce allongée d'un nucléotide), sont séparés par électrophorèse selon leur poids puis identifiés grâce à la fluorescence. Au final, les fragments ne diffèrent entre eux que d'un seul nucléotide qui correspond à un site de polymorphisme unique. *In silico*, l'alignement de séquences de référence avait permis de repérer ces sites où seul un nucléotide diffère. Les amorces sont ensuite créées pour s'hybrider à gauche de ces sites. Cette méthode est particulièrement intéressante lorsque la séquence complète de l'amplicon n'est pas recherchée. Cela réduit la quantité de données à traiter. De plus, il est possible d'appliquer cette méthode en format multiplex, c'est-à-dire en mélangeant différentes amorces. Dans ce cas, la présence de deux amorces de même taille peut être problématique. Mais il est possible de rajouter, au moment de la fabrication des amorces, une queue poly-C (fragment composé uniquement de cytosines) qui ne s'hybridera pas sur la cible mais qui fournira un profil de migration différent, permettant ainsi de distinguer les deux fragments.

C. Le séquençage

Le terme séquençage est cette fois-ci employé pour désigner les techniques permettant de déterminer la séquence nucléotidique.

1. Les méthodes de séquençage

Historiquement, la méthode de Sanger (76) est une méthode capitale puisqu'elle a permis les premiers séquençages de génomes complets, dont celui de *Caenorhabditis elegans* en 1998 (94). Le principe repose sur l'utilisation de didésoxyribonucléotides triphosphates (ddNTP). Une séquence nucléotidique est constituée d'une succession de nucléotides liés entre eux par une liaison phosphodiester. Les ddNTP correspondent à des nucléotides modifiés pour lesquels il manque un groupement hydroxyle OH au niveau du carbone 3'. Lorsqu'un ddNTP est incorporé dans un fragment, l'absence du groupement OH empêche la formation d'une liaison phosphodiester et stoppe la synthèse. On parle de méthode de terminaison de chaîne. Quatre réactions de séquençage sont ainsi menées en parallèle dans quatre tubes distincts, chacun contenant la même amorce qui se lie à l'extrémité 5' du brin d'ADN que l'on souhaite séquençer, les quatre désoxyribonucléotides triphosphates (dNTP) et un seul type de ddNTP marqué par une fluorescence. Au cours du séquençage, l'incorporation d'un ddNTP intervient de manière aléatoire. Au final, dans un même tube, toutes les copies d'ADN synthétisées sont interrompues par le même nucléotide. Les fragments obtenus sont ensuite séparés selon leur taille sur un gel de polyacrylamide. Un détecteur de fluorescence permet de déterminer la nature du nucléotide qui termine chaque fragment et ainsi reconstituer la séquence ADN d'intérêt (95).

Les progrès technologiques sont tels qu'il existe désormais des méthodes bien plus performantes de séquençage. Il s'agit des nouvelles technologies de séquençage à très haut débit ou « next-generation high-throughput DNA sequencing technologies » (NGS). Elles répondent aux besoins d'analyser un très grand nombre de séquences ADN simultanément avec des techniques simples, relativement rapides et peu coûteuses. Les fragments synthétisés sont assez courts : de 30 à 250 paires de bases. Ces techniques s'appliquent notamment à l'étude de la variabilité génétique et au polymorphisme de nucléotide simple (SNP pour *Single Nucleotide Polymorphism* en anglais) au

DNA barcoding et aux études de métagénomique ou métagénétique. Ces dernières consistent à étudier l'ADN d'une communauté d'espèces, soit un mélange d'espèces, directement prélevées dans leur milieu naturel et à en comprendre l'abondance relative en lien avec les interactions interspécifiques mais aussi avec leur environnement.

On a tout d'abord les technologies dites de seconde génération, la méthode Sanger étant considérée comme la technologie de première génération. Chacune d'entre elles est définie par une combinaison de méthodes de préparation du substrat, c'est-à-dire des brins d'ADN à séquencer ; de séquençage et d'imagerie ; et enfin d'analyse des données. La préparation du substrat consiste à fixer sur un support solide ou plateforme, soit des molécules uniques d'ADN soit de multiples répliques obtenus par amplification par PCR. L'avantage principale des molécules uniques d'ADN est qu'elles ne font pas appel à la technique de PCR, évitant ainsi les biais d'amplification et l'incorporation d'erreurs pouvant être considérées à tort comme des mutations. L'étape suivante consiste à séquencer les substrats précédents et à déterminer la séquence de nucléotides par des techniques d'imagerie. Le principe général de cette étape repose sur l'incorporation de nucléotides ou de sondes longues de quelques nucléotides seulement, identifiables grâce à un fluorophore. A chaque cycle, la chaîne de nucléotides est rallongée d'un seul nucléotide ou d'une seule sonde à la fois et dont la composition est déterminée grâce à la fluorescence émise. L'élongation de la chaîne peut être interrompue puis reprise grâce à l'utilisation de terminateurs réversibles. Le procédé s'inspire des ddNTP de la méthode de Sanger mais dans le cas des NGS, il y a une étape supplémentaire de clivage éliminant la fonction de blocage et permettant ainsi la fixation d'un nouveau nucléotide et donc la poursuite de l'élongation. Ensuite, selon l'objectif recherché mais aussi le coût, le travail et le temps disponibles, les séquences sont alignées avec des séquences de référence ou bien assemblées *de novo*, c'est-à-dire pour reconstituer le génome (82).

Il existe aujourd'hui des technologies dites de troisième génération, encore peu utilisées. Les avantages principaux par rapport aux technologies précédentes sont un rendement plus important, l'acquisition de longues séquences ADN (3 000 à 15 000 paires de bases voire plus de 100 000 paires de bases) limitant les *gaps* ou lacunes lors de la reconstitution des génomes entiers, une meilleure précision permettant la détection de rares variants, et enfin un moindre coût de mise en œuvre. De plus, l'une des avancées majeures de ces technologies est que la synthèse de la molécule d'ADN n'est pas interrompue pour déterminer la nature de chaque nucléotide incorporé mais le décryptage se fait en temps réel. En contrepartie, le taux d'erreurs est assez élevé (95) et les plateformes adaptées pour ce type de technologie sont encore peu nombreuses. De plus, il n'y a pas encore d'application pour les nématodes gastro-intestinaux.

Ces technologies sont donc capables de générer une quantité très importante de données génétiques dont le plus difficile est, au final, le traitement bio-informatique et le stockage. C'est pourquoi la méthode de Sanger n'est pas entièrement abandonnée au profit de ces nouvelles technologies et est encore mise en œuvre lorsque peu de compétences bio-informatiques sont requises ; un petit nombre de séquences étant plus facile à analyser (75).

2. Repérer des erreurs dans une séquence de nucléotides

Comment savoir si le changement de nucléotides identifié est bien issu d'une mutation et non d'une erreur de séquençage ? Pour être capable d'opérer une telle distinction, il faut connaître les sources d'erreurs possibles et leur probabilité d'occurrence. Les sources d'erreurs sont multiples et

peuvent provenir de la séquence elle-même ou bien de la technique, que ce soit pour l'amplification ou le séquençage.

La séquence elle-même peut être source d'erreur. Des phénomènes d'intégration de séquences d'ADNm dans le génome nucléaire ont été décrit chez de nombreuses espèces Eucaryotes. On parle de pseudogènes mitochondriaux nucléaires ou numts (*nuclear mitochondrial pseudogenes* en anglais). Ces copies non fonctionnelles de l'ADNm peuvent avoir divergé de leur séquence initiale. Or étant susceptibles d'être co-amplifiées avec de l'ADNm orthologue, elles peuvent être à l'origine d'une surestimation de la variation de l'ADNm (96). La fréquence et la taille des numts ne semblent pas corrélées à la taille du génome mitochondrial ou nucléaire. De plus, il a été montré chez *Caenorhabditis elegans* que ce phénomène était plutôt rare voire absent (97). Néanmoins, ce résultat ne peut être généralisé à l'ensemble du phylum des Nématodes puisque des numts ont été identifiés au sein d'individus de la famille des *Ostertagiinae* (66). Les numts peuvent être identifiés par la présence d'indels (insertions ou délétions) et de codons stop et par un biais de composition de la séquence (pourcentage en base AT supérieur à la moyenne par exemple). Cependant l'identification de ces numts n'est pas toujours possible et évidente et demande du temps, ce qui va à l'encontre des principes du *barcoding* comme méthode rapide à mettre en œuvre (96).

Un autre phénomène susceptible d'impacter la variabilité de l'ADNm est la contamination par des bactéries symbiontes. Ces bactéries, présentes à l'intérieur des cellules, sont transmises par les femelles à leur progéniture dans les cellules embryonnaires et sont capables d'induire un biais de sélection. En effet, la sélection agit en faveur de la transmission des cellules comportant ces symbiontes et favorise donc la transmission de l'haplotype ou ensemble des gènes associés à ces cellules. Lorsque l'infection est bien établie au sein d'une population, l'haplotype associé à l'ADNm original est complètement remplacé par l'haplotype de l'ADNm lié au symbionte. Les conséquences d'un tel biais peuvent être une augmentation ou une diminution de la diversité de l'ADNm. Dans le cas d'une infection récente ou de la propagation d'une infection, la variabilité de l'ADNm a tendance à diminuer du fait de l'homogénéisation des haplotypes. Au contraire, lorsque différentes souches ou espèces de symbiontes infectent une même population, la diversité génétique peut s'accroître du fait de la présence de différents haplotypes. Cela peut donc entraîner des erreurs d'interprétation de la variabilité des séquences génétiques en considérant à tort deux populations comme deux espèces différentes ou bien comme une seule et même espèce. Cette contamination par des bactéries symbiontes est répandue chez les Arthropodes et pourrait être généralisée à l'ensemble des invertébrés, dont les Nématodes (66). La mise en évidence d'une telle contamination ne peut être envisagée par PCR du fait de la diversité des espèces bactériennes pouvant être impliquées. Seule une dissection des ovaires des femelles puis une observation au microscope permettrait d'attester l'absence ou la présence de telles bactéries. Encore une fois, cette vérification demande du temps (98).

Les sources d'erreurs intrinsèques à la technique de PCR abordées au paragraphe II.B.3. correspondent aux erreurs liées au paramétrage de la technique. L'expérimentateur est donc capable de corriger ou du moins limiter ces erreurs en adaptant les paramètres. Pour rappel, il s'agit de la quantité de matériel génétique au départ, la composition en nucléotides de la séquence d'intérêt et de l'amorce, le choix de l'ADN polymérase, le nombre de cycles et la composition du milieu réactionnel. Très souvent, une erreur d'incorporation par l'ADN polymérase est écartée par le fait que le taux de différences observé est supérieur au taux d'erreurs de la polymérase utilisée. Les paramètres de la PCR influencent également la probabilité de générer des produits dits chimériques. Ces produits

résultent de l'élongation incomplète d'un fragment au cours d'un cycle. Ce fragment s'hybride ensuite avec un autre brin matrice au cours du cycle suivant. La séquence du produit obtenu est alors issue de deux brins matrice distincts. De plus, les modifications de température au cours de la PCR peuvent être à l'origine d'une désamination de la cytosine et ainsi induire une substitution des bases C en bases T (89,99).

Les technologies de séquençage de nouvelles générations cumulent les erreurs liées à l'amplification des séquences lors de la préparation des plateformes et les erreurs lors du séquençage et de l'analyse par imagerie. Ce niveau d'hétérogénéité artificiel fixe donc une limite assez haute et en-dessous de laquelle les rares variants ne peuvent être identifiés, ce qui est problématique lors de la distinction d'espèces relativement apparentées.

Une étape supplémentaire visant à déterminer la séquence nucléotidique de quelques produits finaux est suggérée (87,89). Or il apparaît difficile de distinguer une erreur de séquençage d'une mutation par la simple comparaison de plusieurs produits. En effet si l'erreur d'incorporation intervient au cours des premiers cycles de la PCR, le nombre de produits finaux arborant cette erreur sera très important. L'argument selon lequel plus la fréquence d'une différence à un même site est élevée, plus la probabilité que ce soit une mutation réelle est élevée ne sera pas valable. Une solution serait de comparer des produits issus de PCR distinctes. Dans ce cas, la co-occurrence de variations aux mêmes emplacements peut s'expliquer par une mutation réelle et non une erreur de séquençage (34). Une autre solution développée est le séquençage indépendant des deux brins de la molécule d'ADN, distinguables grâce à des séquences types à chaque extrémité : il s'agit de la méthode de *Duplex Sequencing* (figure 13) (100).

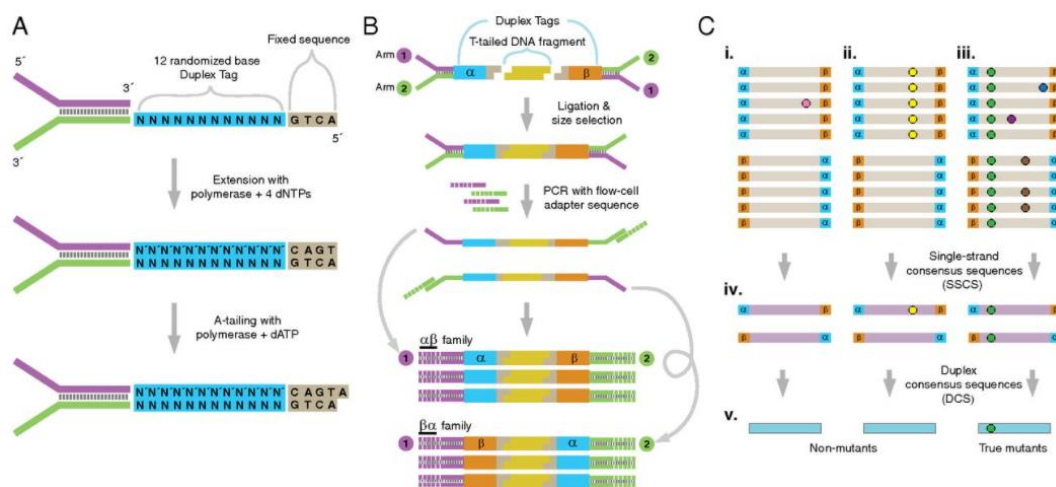


Figure 13 : Schéma du principe de la méthode de *Duplex Sequencing* (100)

Il existe encore d'autres méthodes ayant été développées récemment et dont l'objectif est d'identifier les rares variants génétiques, c'est-à-dire ceux dont la fréquence est assez faible pour être confondue avec des erreurs de séquençage (101–103).

D'autre part, les numts ou la présence d'une bactérie symbionte peuvent concerner un nombre variable d'individus dans une population. Leur impact dépendra alors de la fréquence d'échantillonnage. Une faible taille d'échantillon est plus à même de sélectionner ce type d'erreur ou au contraire de passer à côté. Une grande taille d'échantillon présente l'avantage d'inclure ce type d'erreur mais des vérifications sont alors nécessaires pour les identifier.

Conclusion de la partie I

L'identification des espèces de nématodes gastro-intestinaux par l'étude de leurs caractères morphologiques présentent de nombreux inconvénients dont le temps et le niveau d'expertise nécessaires, ce dernier rendant la méthode difficilement standardisable ; l'existence d'espèces polymorphiques ou cryptiques, la possibilité d'hybridation.

L'ADN apparaît au contraire comme une molécule stable et dont la séquence renferme un certain nombre d'informations. Ainsi les études moléculaires offrent de nouvelles perspectives en matière d'identification et de classification taxonomique en développant des marqueurs génétiques d'identification. Plusieurs gènes ont été étudiés afin de déterminer leur potentiel comme marqueur génétique d'espèce. Parmi l'ADN nucléaire, les régions ITS-1 et ITS-2 présentent les taux de variabilité inter- et intraspécifiques les plus intéressants pour distinguer des espèces entre elles. Le niveau de variation plus important des séquences mitochondriales, telles que *cox1* et *nad4*, est plus approprié pour distinguer des individus ayant une histoire évolutive assez récente. Autrement dit ces séquences sont plus appropriées pour l'étude des populations. Néanmoins, la mise en évidence de variations intraspécifiques mais aussi intra-individuelles limitent l'emploi de ces séquences comme marqueur unique d'identité. Pour une identification plus fiable, l'utilisation de plusieurs marqueurs génétiques est recommandée.

De plus, l'interprétation des différences entre séquences peut être erronée. En effet, les technologies permettant de récupérer la séquence nucléotidique, que ce soit l'amplification par PCR ou le séquençage *sensu stricto*, peuvent être sources d'erreurs par l'incorporation incorrectes de nucléotides. Ainsi une différence entre des séquences sera interprétée à tort comme une mutation. De plus, la séquence au sein de l'organisme même peut être erronée par l'intégration de séquences non homologues (cas des pseudogènes mitochondriaux nucléaires) ou bien non représentative de l'espèce ou même de la population du fait de biais de transmission entre générations (cas des infections par des bactéries symbiontes). Toutes ces sources d'erreurs possibles suggèrent l'association de la méthode moléculaire avec d'autres méthodes, morphologiques par exemple, pour une identification plus fiable des espèces.

PARTIE 2 : ANALYSE BIO-INFORMATIQUE DE SEQUENCES NUCLEOTIDIQUES DE NEMATODES GASTRO-INTESTINAUX

I. Introduction

Le phylum des Nématodes est l'un des phylums les plus étudiés avec environ 25 000 à 26 000 espèces décrites. Pourtant, cela ne représenterait qu'un quart des espèces présentes sur Terre (104). Les Nématodes, et en particulier les nématodes gastro-intestinaux, se retrouvent très souvent en situation de co-infection au sein d'un même hôte. Ainsi les études sur la biologie et l'épidémiologie de ces parasites nécessitent d'être capable de pouvoir distinguer les espèces entre elles au sein d'un même échantillon. Etant données les difficultés inhérentes aux méthodes morphologiques, biochimiques et/ou immunologiques d'identification, une approche moléculaire semble prometteuse. La première partie de cette thèse a permis d'aborder les marqueurs moléculaires ou génétiques ainsi que les outils capables d'identifier des espèces de nématodes gastro-intestinaux.

Les données moléculaires sont regroupées dans des bases de données dont la plus importante est « GenBank », gérée par le *National Center of Biology Information* (NCBI). Il s'agit d'une base de données publique de séquences de nucléotides accessibles à tous (www.ncbi.nlm.nih.gov). Elle est en lien avec d'autres bases de données : la European Molecular Biology Laboratory Nucleotide Sequence Database (EMBL-Bank) et la DNA Data Bank of Japan (DDBJ) ; de façon à fournir des données uniformes et complètes au niveau mondial. D'autres bases de données plus spécifiques aux nématodes existent telles que celle créée par le laboratoire Blaxter de l'Université d'Edinburg accessible via le site internet : www.nematodes.org/ ; *WormBase* (www.wormbase.org/) ou encore *NemaBiome* (www.nemabiome.ca/). Toutes ces bases de données ont pour objectif de concentrer les informations, projets de recherche, outils d'exploitation en rapport avec les Nématodes. Néanmoins, il s'agit souvent de bases isolées les unes des autres de sorte que chacune peut comporter des données différentes. Ainsi GenBank apparaît comme une banque plus généraliste mais où se retrouve toutes les informations de chacune des banques spécialistes.

L'objectif de l'étude menée sur les nématodes gastro-intestinaux des Caprinés de la réserve zoologique de la Haute Touche était de pouvoir identifier les espèces de parasites en cause pour améliorer les plans de traitements antiparasitaires. Un protocole d'autopsie parasitaire avait été proposé afin de collecter des nématodes adultes et ainsi tenter d'identifier jusqu'au rang de l'espèce. Pour aller encore plus loin et en employant une approche moléculaire de caractérisation, il serait intéressant de récupérer les séquences génétiques de ces parasites et de les comparer avec des séquences issues des bases de données mentionnées précédemment. Cela permettrait notamment de déterminer l'espèce pour laquelle l'identification morphologique reste difficile, même au stade adulte. De plus, des amorces pourraient être construites à partir des séquences récupérées et des séquences homologues dans les bases de données et ainsi créer un outil de diagnostic spécifique des communautés de nématodes gastro-intestinaux des Caprinés de la réserve zoologique de la Haute Touche.

Une étape préliminaire à ce travail et qui constitue l'objectif de cette seconde partie est de déterminer si les données et outils moléculaires disponibles aujourd'hui sont en nombre suffisant pour pouvoir décrire la diversité d'espèces au sein d'une telle communauté de parasites nématodes. Pour cela, un inventaire des données moléculaires disponibles dans GenBank est réalisé pour les espèces identifiées comme appartenant au groupe des nématodes gastro-intestinaux des Caprinés. Dans un second temps, les marqueurs génétiques utiles et les stratégies d'identification moléculaire appliquées sont déterminées. Enfin la spécificité de ces outils est vérifiée *in silico*.

II. Matériels et méthodes

A. Inventaire des espèces de nématodes gastro-intestinaux

L'inventaire des espèces de parasites retenues a été fait à partir de six références bibliographiques (8,105–109). Une espèce est sélectionnée lorsqu'elle remplit les critères suivants :

- espèce infestant un hôte appartenant au groupe des Caprinés domestiques et sauvages parmi lesquels *Capra hircus* (la Chèvre), *Ovis aries* (le Mouton), *Rupicapra rupicapra* (le Chamois), *Ovis musimon* (le Mouflon Corse), *Capra ibex* (le Bouquetin des Alpes), *Ammotragus lervia* (le Mouflon à manchettes).
- espèce considérée comme appartenant au groupe des nématodes gastro-intestinaux c'est-à-dire appartenant au phylum des Nématodes et dont la localisation chez l'hôte correspond à un organe du tube digestif : rumen, reticulum, omasum, abomasum, intestin grêle, gros intestin, caecum.

B. Récupération des données de GenBank

Pour chaque taxon, les informations suivantes sont récoltées à partir de la base de données « Taxonomy » de GenBank et répertoriées dans un tableau Excel : le « ID taxonomy », la classification taxonomique (noms scientifiques de famille, genre, espèce) et le nombre de séquences de nucléotides disponibles dans GenBank. La date de la recherche est également mentionnée puisque le nombre de séquences nucléotidiques est susceptible de varier dans le temps du fait de nouvelles publications. Pour une meilleure compréhension, ce tableau est par la suite désigné sous le nom de « tableau des espèces de nématodes gastro-intestinaux des Caprinés ».

Pour chaque espèce dont le nombre de séquences disponibles est relativement faible, les informations suivantes sont récupérées manuellement pour chaque séquence, à partir de la base de données « Nucleotide » de GenBank et avec l'affichage « GenBank » : le numéro d'accession, le nom du gène ou du transcrit ou de la protéine correspondante en précisant si la séquence est complète ou partielle (champ DEFINITION), la longueur en paires de bases, l'espèce-hôte où elle a été détectée (champ FEATURES > source > host), le pays où le parasite a été récolté (champ FEATURES > source > country), le « Digital Object Identifier » (DOI) de la publication correspondante dans le cas où elle a été publiée et est disponible (champ JOURNAL). Une analyse de la publication est parfois nécessaire afin de récupérer les informations manquantes. Cette table est par la suite désignée sous le nom « table des séquences GenBank ».

L'organisation ainsi que la signification des données GenBank sont présentées en annexe 1.

C. Extraction et bilan des données sur les marqueurs génétiques

La première partie de cette thèse a permis d'identifier les régions codantes ou non de l'ADN pouvant avoir un intérêt comme marqueur génétique d'identification. Elles appartiennent soit à l'ADNr : 18S, ITS-1, 5.8S, ITS-2, 28S ; soit à l'ADNm : *cox1* et *nad4*. L'extraction et le bilan des données sont réalisés avec le logiciel R, et en particulier les packages de la suite *tidyverse*, par exemple le package *stringr*, qui facilitent l'analyse des données de type « caractère » (110,111). Les séquences incluant les marqueurs génétiques cités précédemment sont dénombrées pour chaque espèce à partir de la table des séquences GenBank.

D. Stratégies d'identification moléculaire

Par la suite, on s'intéresse particulièrement à la région ITS-2. En effet, au II.A.2. de la revue bibliographique de la première partie de cette thèse, on a montré qu'au sein de cette région, la variation intraspécifique était suffisamment faible et la variation interspécifique suffisante pour permettre l'identification moléculaire au rang taxonomique de l'espèce.

Dans les publications portant sur ce marqueur, nous avons recensé les amorces et sondes internes utilisées, ainsi que leur spécificité attendue.

E. Simulation d'une PCR *in silico*

Afin de vérifier la spécificité des amorces employées pour l'identification moléculaire des nématodes gastro-intestinaux, une réaction de PCR *in silico* est réalisée. Un lot de séquences ITS-2 est constitué à partir de la table des séquences GenBank, qui est facilement manipulable par l'application de filtres sur les données avec le logiciel Excel. Une séquence seulement est choisie par espèce pour limiter le nombre de données à traiter. Les espèces concernées correspondent aux espèces pour lesquelles des amorces ont été identifiées lors de l'analyse bibliographique menée dans la partie précédente. D'autres espèces, pour lesquelles il n'y a pas d'amorces spécifiques mais dont l'ensemble des séquences a été récupéré, complètent la liste pour vérifier la spécificité des amorces ciblant un genre particulier ou le phylum des Nématodes. Un premier lot de 29 séquences est constitué avec des séquences incluant la région ITS-2 mais aussi les régions adjacentes 5.8S et 28S pour être sûr que la séquence de la région ITS-2 soit bien complète. Cependant les résultats de la PCR *in silico* sont peu concluants : un certain nombre d'amorces ne s'hybride pas (résultats non montrés). Un second lot de séquences est constitué avec des séquences plus longues allant de l'ADNr 18S à l'ADNr 28S (18S-28S) et incluant donc les séquences complètes des régions ITS-2 mais aussi ITS-1 et 5.8S. Le tableau 1 regroupe les numéros d'accèsion des 37 séquences constituant ce deuxième lot. Lorsque plusieurs séquences 18S-28S sont disponibles pour une même espèce, la séquence la plus longue est choisie, en supposant qu'elle contient plus d'informations. Dans le cas contraire où il n'y a pas de séquence 18S-28S disponible pour l'espèce, le choix est reporté sur d'autres séquences précisées entre parenthèses dans le tableau 1, en privilégiant toujours les séquences les plus longues.

Tableau 1 : Numéro d'accèsion des séquences constituant le lot de séquences pour la simulation de la PCR *in silico*

Le « p » mentionné dans l'intitulé de certaines séquences signifie « partielle ». Par exemple, la séquence « ITS-2p/28p » correspond à la séquence IST-2-28S pour laquelle l'extrémité 5' ne comprend pas l'ensemble de la séquence ITS-2, de même pour l'extrémité 3' avec la séquence 28S. « NM » signifie que le caractère partiel ou complet de la séquence n'est pas précisé. L'astérisque indique les espèces pour lesquelles la séquence a directement été cherchée dans GenBank, en dehors du processus d'inventaire.

Phylum ou Genre ou Espèce	Numéro d'accèsion de la séquence
<i>Haemonchus contortus</i> *	KX829170.1
<i>Haemonchus longistipes</i>	AB682685.1
<i>Ostertagia sp.</i>	<i>O. lyrata</i> : KX929995.1 (ITS-1p/5.8/ITS-2/28p) <i>O. trifurcata</i> : KC295419.1
<i>Ostertagia ostertagi</i> *	KT428385.1
<i>Ostertagia leptospicularis</i>	KC998723.1 (5.8/ITS-2/28p)
<i>Teladorsagia circumcincta</i> *	JF680984.1
<i>Marshallagia marshalli</i>	KT428384.1
<i>Cooperia sp.</i>	<i>C. pectinata</i> : MH267785.1 (ITS-1p/5.8/ITS-2/28p)
<i>Cooperia curticei</i>	JF680982.1 (ITS-1p/5.8/ITS-2/28p)
<i>Cooperia punctata</i>	AB534602.1
<i>Cooperia surnabada</i>	AJ000032.1 (ITS-2 NM)
<i>Cooperia oncophora</i> *	AB238690.1
<i>Nematodirus sp.</i>	<i>N. davtiani alpinus</i> : AJ239113.1 <i>N. helvetianus</i> : AF194142.1
<i>Nematodirus filicollis</i>	AF194140.1
<i>Nematodirus spathiger</i>	AF194144.1
<i>Nematodirus battus</i>	JF345079.1
<i>Trichostrongylus sp.</i>	<i>T. capricola</i> : JF276022.1 (ITS-2p) <i>T. longispicularis</i> : KY355070.1 (ITS-2p/28p) <i>T. probolurus</i> : JQ925866.1 (ITS-1p/5.8/ITS-2/28p) <i>T. rugatus</i> : Y14818.1 (ITS-2 NM)
<i>Trichostrongylus axei</i>	JQ889794.1 (ITS-1p/5.8/ITS-2/28p)
<i>Trichostrongylus colubriformis</i>	JF680985.1
<i>Trichostrongylus vitrinus</i>	JF680986.1
<i>Bunostomum trigonocephalum</i>	MG182023.1
<i>Oesophagostomum asperum</i>	AB971665.1
<i>Oesophagostomum venulosum</i>	HQ283349.1
<i>Oesophagostomum columbianum</i>	AJ006150.1 (ITS-2 NM)
<i>Chabertia ovina</i>	KF913471.1
<i>Trichuris sp.</i>	<i>T. discolor</i> : LC320148.1 <i>T. globulosa</i> : LN651156.1 (ITS-2p) <i>T. skrjabini</i> : AJ489248.1 (ITS-1p/5.8/ITS-2)
<i>Trichuris ovis</i>	KF266766.1
<i>Nematodes</i>	<i>Camelostrongylus mentulatus</i> : KY930444.1 (5.8p/ITS2/28p) <i>Parabronema skrjabini</i> : KP670202.1 <i>Spiculoptera boehmi</i> : AJ577460.1 (ITS-2 NM)

La PCR *in silico* est réalisée avec l'outil « primersearch » de la plateforme bio-informatique « Galaxy » accessible via le lien internet suivant : <https://usegalaxy.org/> (112,113). Les paires d'amorces sont testées une par une sur le lot de séquences 18S-28S. Un pourcentage de « mismatch » de 10% est autorisé. Il s'agit du nombre de nucléotides différents autorisés entre l'amorce et la séquence cible. Pour des amorces de 20 paires de bases de longueur, 10% de mismatch correspondent à un à deux nucléotides de différence.

La spécificité des sondes n'a pas été vérifiée par manque de temps.

III. Résultats

A. Espèces de nématodes gastro-intestinaux et bilan des séquences disponibles

Au total, 57 espèces de nématodes gastro-intestinaux ont été retenues à partir de la revue bibliographique. Les informations disponibles pour chacune sont présentées dans le tableau des espèces de nématodes gastro-intestinaux des Caprinés en annexe 2. Le nombre de séquences disponibles notamment varie entre 0 et 88 238 avec un total de 176 582 séquences. La répartition du nombre de ces séquences est la suivante :

- 0 séquence disponible pour 11 espèces
- 1 séquence disponible pour 4 espèces
- entre 2 et 10 séquences disponibles pour 9 espèces
- entre 11 et 100 séquences disponibles pour 18 espèces
- entre 101 et 500 séquences disponibles pour 8 espèces
- au-delà de 500 séquences disponibles pour 6 espèces : *Cooperia oncophora*, *Haemonchus contortus*, *Ostertagia ostertagi*, *Teladorsagia circumcincta*, *Strongyloides papillosus*, *Gongylonema pulchrum*.

Pour un quart des espèces, le nombre de séquences disponibles est nul ou égal à un. Pour la moitié des espèces, le nombre de séquences disponibles est inférieur ou égal à 20. Et pour les trois quarts des espèces, le nombre de séquences disponibles est inférieur à 78. La moyenne est de 3 098 séquences environ. Cependant, l'écart-type est très élevé (environ 13 500), ce qui signifie que les valeurs sont très dispersées autour de la moyenne.

B. Bilan des données récupérées

Parmi les 6 espèces dont le nombre de séquences disponibles est supérieur à 500 (de 961 pour *C. oncophora* à 88 238 séquences pour *T. circumcincta*), seules les informations pour les séquences de *Cooperia oncophora* ont été recueillies. Devant le temps important alloué à cette tâche, les cinq autres espèces sont écartées du bilan quantitatif. Au total, les informations détaillées de 3 628 séquences réparties chez 40 espèces de nématodes gastro-intestinaux ont été récupérées manuellement. Un extrait de la table des séquences GenBank est présenté en annexe 3.

Les séquences d'ADNr couvraient totalement ou partiellement la région codante, mais le champ DEFINITION des métadonnées de Genbank n'est pas standardisé, ce qui rend le dénombrement des différents gènes difficile. Avec l'aide de fonctions d'analyse de texte de la suite logicielle R {tidyverse}, le nombre de séquences de chaque marqueur génétique a été déterminé pour chaque espèce, le tableau bilan est présenté en annexe 4. La région ITS-2 est le marqueur présent dans le plus grand nombre de séquences (1050) parmi 38 espèces de nématodes sur les 40 étudiées. Puis ce sont les régions adjacentes 28S et 5.8S qui sont les plus fréquentes avec 545 séquences parmi 33 espèces et 544 séquences parmi 32 espèces respectivement. La région ITS-1 est présente dans 452 séquences parmi 31 espèces. Enfin, la séquence 18S est présente dans 202 séquences parmi 25 espèces de nématodes. En comparaison avec l'ADN nucléaire, les régions de l'ADNm sont présentes dans un nombre assez proche de séquences : 420 pour *cox1* et 403 pour *nad4* ; mais réparties chez un nombre plus faible d'espèces : 21 pour *cox1* et seulement 9 pour *nad4*. La moyenne du nombre de séquences par marqueur génétique varie de 5 pour la région 18S à 27 pour la région ITS-2 avec, pour tous les marqueurs, un écart-type supérieur à la moyenne indiquant une dispersion des valeurs élevée.

C. Stratégies d'identification moléculaire

Les différents articles utilisant les marqueurs moléculaires peuvent être regroupés selon trois grandes stratégies d'identification, basées sur des techniques de PCR. Pour l'exemple de la région ITS-2, les sources bibliographiques employant l'une ou l'autre stratégie sont citées dans le tableau 2.

Tableau 2 : Classification des sources bibliographiques en fonction de la stratégie appliquée pour l'identification d'espèces ou de genres de nématodes gastro-intestinaux.

Stratégies	Stratégie 1	Stratégie 2	Stratégie 3
Sources bibliographiques	Hoste et al., 1993 (28) Hoste et al., 1995 (30) Stevenson et al., 1995 (15) Stevenson et al., 1996 (10) Newton et al., 1998a,b,c (12,32,35) Gasser et al., 1998 (80) Heise et al., 1999 (26) Avramenko et al., 2015 (114) Paguem et al., 2020 (115)	Schnieder et al., 1999 (116) Dallas et al., 2000 (37) von Samson - Himmelstjerna et al., 2002 (81) Wimmer et al., 2004 (41) Bott et al., 2009 (90) Learmount et al., 2009 (117) Demeler et al., 2013 (118) Santos et al., 2019 (119)	Roeber et al., 2012 (92) McNally et al., 2013 (120) Bisset et al., 2014 (91)

La stratégie 1 consiste à réaliser une PCR avec des amorces génériques, c'est-à-dire situées dans des régions conservées de l'ADN et entourant une région hypervariable. La séquence du produit d'amplification est obtenue par séquençage puis comparée à des banques de séquences dont l'appartenance taxonomique est connue.

Lorsque cette stratégie est appliquée pour identifier un individu unique, la méthode de Sanger, décrite en III.C.1. de la partie 1, permet le séquençage simple du produit amplifié. L'algorithme de comparaison des séquences employé par la suite est l'algorithme BLAST (*Basic Local Alignment Search Tool*). Ce dernier compare une à une les séquences disponibles dans la base de données avec la séquence d'intérêt et renvoie un pourcentage d'identité et des scores d'homologie pour chaque comparaison. La séquence choisie pour l'identification est alors la séquence la plus proche. Il faut rester prudent car plusieurs taxons peuvent avoir 100% d'homologie selon la région étudiée. Plus le fragment amplifié et séquencé est long, plus on a de chance de détecter du polymorphisme, des substitutions, insertions ou délétions, typiques de telle ou telle espèce.

Lorsque cette stratégie est appliquée à un mélange d'individus d'espèces différentes, les technologies NGS ou « haut-débit » permettent le séquençage de ces multiples fragments d'ADN. Le résultat est sous forme d'un fichier pouvant contenir des milliers de séquences différentes qui doivent être chacune comparées à une base de données pour les « assigner » à la séquence la plus proche sur le même principe que décrit précédemment. Les éléments à comparer étant beaucoup plus nombreux, cela requiert des algorithmes spéciaux tels que *mothur* par exemple (<http://mothur.org/>).

La stratégie 2 s'apparente à la méthode de la PCR en temps réel décrite en III.B.4 de la partie 1 et correspond à l'amplification d'une région hypervariable et à son identification par l'hybridation de composés spécifiques. Une première option consiste à employer des paires d'amorces spécifiques et le produit amplifié est caractérisé dans un second temps par l'analyse des courbes de fusion, qui varient selon la composition en nucléotides. Une seconde option consiste à employer des sondes

marquées par un fluorochrome de couleur spécifique et qui se fixent à l'intérieur de la région hypervariable sans initier de synthèse d'ADN comme c'est le cas avec les amorces. L'accumulation du produit d'amplification est détectée par augmentation de l'intensité du signal fluorescent et indique donc la présence de la séquence cible dans le mélange de départ. La couleur de la fluorescence renseigne ensuite sur l'identité de la séquence.

La stratégie 3 est basée sur la technique de la PCR multiplexe décrite en III.C.1 qui emploie plusieurs paires d'amorces spécifiques en simultanée. Cette stratégie permet notamment d'analyser des échantillons complexes où sont présents des mélanges de parasites. Cependant, outre la spécificité des amorces pour un taxon particulier, il faut aussi vérifier qu'elles ne sont pas capables de s'hybrider entre elles, surtout lors de la distinction d'espèces assez proches. Il est possible d'insérer intentionnellement quelques erreurs dans la séquence de l'amorce de sorte qu'elle est toujours capable de s'hybrider avec sa séquence cible mais est plus réfractaire à se fixer à des séquences non cibles, incluant les autres amorces du mélange.

Les amorces génériques utilisées dans les publications adoptant une stratégie 1 correspondent aux amorces NC1 et NC2 (121). NC1 se fixe au niveau de l'extrémité 3' de la région 5.8S et NC2 au niveau de la région 5' de la région 28S (figure 14).

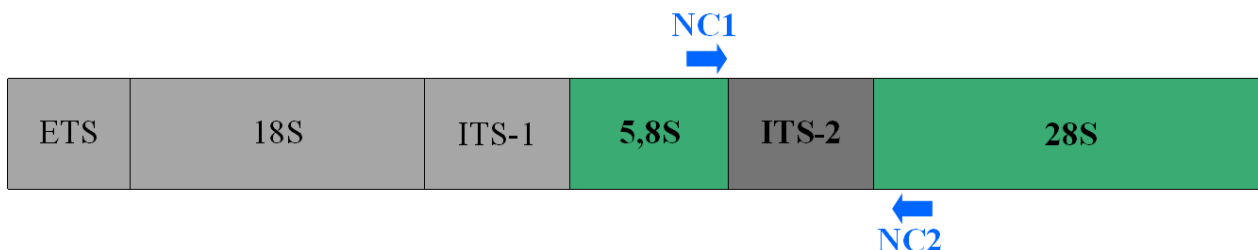


Figure 14 : Sites de fixation des amorces NC1 et NC2 sur l'ADN ribosomique des Nématodes

Le tableau présenté en annexe 5 récapitule les séquences des amorces et sondes employées dans les publications pour lesquelles une stratégie 2 ou 3 a été identifiée. Les 130 séquences publiées se partagent en 7 paires d'amorces et 6 sondes spécifiques de genre, et 28 paires d'amorces et 54 sondes spécifiques d'espèce.

D. Simulation d'une PCR *in silico*

Des PCR ont été simulées sur un jeu de 37 séquences d'espèce (14 genres différents) (Tableau 1), avec les 7 paires d'amorces spécifiques de genre et les 28 paires spécifiques d'espèce (spécificité attendue, au vu de la publication d'origine) (annexe 5).

L'outil « primersearch » sur la plateforme Galaxy (112,113) permet de déterminer si la séquence des paires d'amorces est présente dans le lot de séquences indiqué. Si la séquence est présente, la position du premier nucléotide hybridé de l'amorce sur la séquence cible est précisée, aussi bien dans le sens 5'-3' pour l'amorce « forward » que 5'-3' pour l'amorce « reverse » sur le brin complémentaire (résultats non montrés). Les paires d'amorces sont désignées par leur nom abrégé correspondant dans le tableau de l'annexe 5. La longueur du produit amplifié est également donnée et a été reportée dans un tableau présenté en annexe 6.

Le nombre de PCR « positives » est déterminé en comptant le nombre de produits d'amplification obtenus pour chaque paire d'amorce, sans tenir compte de leur spécificité. Il s'étend

de 0 à 11 pour Tela_13 et est en moyenne de 2 avec un écart-type de 2. Seulement 4 paires d'amorces n'ont aucun produit d'amplification : Haem_1, Tric_26, Tric_27 et Oeso_31.

Le tableau en annexe 7 confronte la spécificité attendue de chaque paire d'amorce (selon sa publication d'origine) avec les résultats de la PCR *in silico*. Les spécificités attendues et leur publication d'origine sont recensées dans le tableau en annexe 5 en précisant si la paire d'amorce est considérée comme spécifique d'un phylum, d'un genre ou d'une espèce. La paire d'amorces NC1_NC2 ne s'est hybridée qu'avec quelques séquences alors qu'elle est supposée être générique pour le phylum des Nématodes auquel appartiennent toutes les séquences testées. Parmi les sept paires d'amorces spécifiques d'un genre, la paire Haem_1 ne s'est hybridée avec aucune des deux espèces d'*Haemonchus sp.* ; et la paire Oste_6 s'est hybridée avec l'espèce *Marshallagia marshalli* alors qu'elle n'appartient pourtant pas au genre *Ostertagia sp.* Pour les paires d'amorces spécifiques d'une espèce, 13 paires ne s'hybrident qu'avec la séquence de leur espèce cible. Les 15 autres paires s'hybrident soit avec leur espèce cible et d'autres espèces non cibles, en particulier la paire Tela_13 supposée être spécifique de *Teladorsagia circumcincta* mais qui s'hybride avec 10 autres espèces différentes ; soit avec des espèces non cibles uniquement, comme c'est le cas de la paire Coop_19.

IV. Discussion

L'étude de communautés d'organismes nécessite l'accès à des données en nombre suffisant pour pouvoir caractériser l'ensemble des espèces ou rangs taxonomiques supérieurs (genre, famille ...) constituant cette communauté. La caractérisation moléculaire passe par l'étude de la séquence des molécules d'ADN.

Le dénombrement des séquences nucléotidiques disponibles dans GenBank pour les 57 espèces de nématodes gastro-intestinaux sélectionnées montrent de grandes disparités. En effet, il y a peu d'espèces sur-représentées avec plus de 500 séquences disponibles et beaucoup d'espèces sous-représentées avec seulement une seule voire aucune séquence disponible. La sur-représentation de certaines espèces peut s'expliquer par leur pouvoir pathogène important, connu et étudié depuis de nombreuses années maintenant, comme c'est le cas de *Haemonchus contortus*. Très souvent, ces espèces de parasites sont également importantes pour leur pouvoir pathogène chez d'autres espèces d'hôtes d'intérêt plus important par rapport aux Caprinés. C'est le cas de *Cooperia oncophora*, parasite d'intérêt majeur pour l'élevage de Bovins ; ou encore *Gongylonema pulchrum*, préoccupant en santé publique car il est capable d'infecter l'Homme. Les espèces sous-représentées correspondent parfois à des espèces pour lesquelles la classification taxonomique n'est pas certaine et est encore en discussion. C'est le cas notamment des espèces appartenant à un groupe d'espèces polymorphiques ou cryptiques. L'espèce *Teladorsagia davtiani*, par exemple, est identifiée comme une espèce à part entière alors qu'il ne pourrait s'agir que d'un morphotype de *T. circumcincta*, tout comme *T. trifurcata* (9).

Au 15 octobre 2020, GenBank compterait 220 millions de séquences de nucléotides dont un peu plus de 4,3 millions concernant les Nématodes, d'après le rapport accessible sur le site de GenBank via le lien internet suivant : <https://ftp.ncbi.nih.gov/genbank/gbrel.txt>. La difficulté majeure rencontrée pour réaliser l'inventaire des séquences est le manque d'harmonisation dans l'intitulé des séquences déposées (champ DEFINITION). GenBank est une base de données publique alimentée par des soumissions directes de séquences par des scientifiques ou par des lots de séquences multiples issues de projets particuliers de séquençage (Whole-Genome Shotgun, Environmental sample sequence ...). Dans le cas des soumissions directes, certaines vérifications d'éventuelles contaminations de la séquence par des vecteurs (plasmide) ou de la véracité des données fournies sont réalisées afin d'assurer la qualité de l'information : l'outil BankIt par exemple utilise une variante de l'algorithme BLAST nommé « Vecscreen » pour vérifier la contamination par des vecteurs (122). Cependant, le nombre de séquences est tel que ces vérifications ne peuvent être fiables à 100%. De plus, l'intitulé de la séquence reste libre au dépositaire et ne contient très souvent pas seulement le nom du gène ou du transcrypt correspondant mais d'autres informations comme l'espèce, le produit de traduction, le numéro de l'isolat, etc. Les bases de données plus spécifiques telles que *Nemabiome*, *WormBase* ou la base de données du laboratoire Blaxter de l'Université d'Edinburg ; regroupent certes un plus petit nombre de séquences par rapport à GenBank mais les métadonnées qui accompagnent ces séquences apparaissent plus fiables.

Ce travail visait à déterminer le marqueur génétique le plus approprié pour caractériser les communautés de nématodes à l'échelle de l'espèce, notamment sur le critère d'un nombre important de séquences disponibles, comme référence de comparaison, pour les espèces connues. Les résultats du nombre de séquences disponibles par espèce incluant un marqueur génétique potentiel, identifié dans la première partie de cette thèse, confirment les données bibliographiques : la région ITS-2 est le marqueur le plus étudié. Plus généralement, les marqueurs issus de l'ADN nucléaire sont plus

étudiés en comparaison avec les marqueurs de l'ADNm pour lesquels le nombre d'espèces chez qui on les retrouve est inférieur. L'étendue pour chaque marqueur indique là encore une grande disparité entre les espèces. Pour ITS-2 par exemple, le nombre de séquences s'étend de 0 pour *Ostertagia kolchida* et *Skryabinema ovis* à 191 pour *Trichostrongylus colubriformis*.

Les variabilités intra- et interspécifique de la région ITS-2 sont telles que cette région est principalement employée comme marqueur génétique d'identification au rang taxonomique de genre et surtout d'espèce. Pour pouvoir identifier les nématodes gastro-intestinaux présents dans un échantillon multi-espèces, trois stratégies d'identification ont été définies. Les stratégies 2 et 3 supposent que les espèces recherchées sont connues au préalable : telle sonde ou telle paire d'amorces est utilisée pour identifier une espèce en particulier. Il peut s'agir de l'espèce la plus pathogène par exemple. La stratégie 1, au contraire, peut être décrite comme une approche sans *a priori* : une paire d'amorces générique amplifie une région dont le séquençage et la comparaison avec des séquences de référence dans un second temps permettent de déterminer l'identité de l'espèce. La première approche, avec *a priori*, est plus rapide, moins compliquée à interpréter mais moins informative : il est possible de passer à côté d'espèces qui ne seraient pas incluses dans les cibles des amorces choisies. La seconde approche, sans *a priori*, est plus riche en informations mais plus chère car elle nécessite un séquençage haut débit, et est plus longue à mettre en œuvre du fait des calculs bio-informatiques complexes lors de la comparaison avec les banques de séquences de référence. Néanmoins elle sera plus appropriée pour décrire l'ensemble des espèces présentes dans un échantillon, comprenant également les espèces minoritaires et/ou exotiques. Ainsi, la première approche s'applique plutôt à l'étude des communautés de parasites de la faune domestique puisque ces parasites sont généralement assez bien caractérisés. La seconde approche est plus adaptée pour l'étude des communautés de parasites de la faune sauvage et exotique pour lesquelles la probabilité qu'une espèce de parasite rare soit présente est plus élevée. L'identification de cette espèce ne sera possible que s'il existe des séquences de référence ou des séquences proches dans les banques de données, d'où l'intérêt de constituer un inventaire des séquences disponibles.

Le choix de l'une ou l'autre des stratégies repose sur l'hypothèse que les amorces ou sondes employées sont soit génériques, comme c'est le cas de NC1 et NC2, soit spécifiques. L'inventaire des amorces et sondes dans les publications du tableau 2 montre qu'il n'y en a pas de publiées pour toutes les espèces susceptibles de se retrouver dans des échantillons multi-spécifiques issus de Caprinés. C'est le cas pour trois genres : *Camelostrongylus sp.*, *Parabronema sp.* et *Spiculoptergia sp.* ; et pour 15 espèces. Leur identification dans un mélange ne sera possible qu'avec l'utilisation des amorces génériques suivie d'un séquençage. La simulation d'une PCR *in silico* permet de vérifier la spécificité attendue associée à chaque paire d'amorces. La spécificité est une des caractéristiques intrinsèques d'un test diagnostique. Pour la technique de PCR, il s'agit de la capacité d'un test à n'amplifier la séquence que d'une espèce ou d'un genre en particulier. L'autre caractéristique intrinsèque d'un test diagnostique et qui n'est pas étudiée ici est la sensibilité. Cette dernière correspond à la capacité d'un test à détecter toutes les séquences de l'espèce cible dans un mélange. Elle dépend de la quantité d'ADN présente dans le mélange initial : on parle de détectabilité ; ainsi que de la présence effective de la séquence complémentaire à l'amorce.

Les résultats de la PCR *in silico* montrent en premier que la paire d'amorces NC1_NC2 échoue à amplifier l'ensemble des espèces appartenant pourtant toute au phylum des Nématodes. Pour les espèces pour lesquelles il n'y avait pas de séquence 18S-28S disponible, l'absence d'hybridation de la paire d'amorces NC1_NC2 peut s'expliquer par le fait que la séquence choisie en remplacement était plus courte et ne couvrait donc pas le site de fixation des amorces, ce qui est notamment le cas

pour l'espèce *Cooperia surnabada* par exemple, représentée par une séquence ITS-2 seule. Mais même dans le cas où il s'agit bien de séquences 18S-28S, on peut avoir des variations. La figure en annexe 8 est un alignement de six séquences 18S-28S de diverses espèces de nématodes gastro-intestinaux. L'alignement de ses séquences a été obtenu avec l'outil BLAST de GenBank (123) sur la séquence 18S-28S de *Haemonchus contortus* dont le numéro d'accèsion est le suivant : KX829170. Les cinq séquences ont été sélectionnées parmi les séquences ayant un pourcentage d'identité élevé avec la séquence de *H. contortus* et appartenant à une espèce de nématode gastro-intestinal. On observe que pour quatre séquences, les trois supplémentaires de *H. contortus* et celle de *T. circumcincta*, le site de fixation de NC2 est absent ou incomplet alors qu'il s'agit de séquences spécifiées comme étant des séquences 18S-28S dans GenBank.

Ensuite, pour les amorces spécifiques de genre ou d'espèce, trois situations sont mises en évidence. La première correspond à une paire d'amorces qui ne s'hybride avec aucune des séquences pour lesquelles elle est désignée comme spécifique (paire Haem_1). La seconde situation, majoritaire, correspond au cas où la paire d'amorces ne s'hybride pas exclusivement avec la séquence du genre ou de l'espèce cible. Enfin la troisième situation correspond au cas où la paire d'amorce s'hybride exclusivement avec une espèce non cible (paire Coop_19). Au final, seules quelques amorces semblent réellement spécifiques en n'amplifiant que la séquence de leur espèce cible. Comment expliquer le manque de spécificité majoritaire des autres amorces ?

Il pourrait s'agir d'un problème lors de la détermination de la séquence de l'amorce : un nombre pas assez important de séquences a été pris en compte pour obtenir une séquence spécifique de l'espèce seulement. Une étape de vérification *in silico* de la spécificité pourrait aussi être ajoutée lors du processus de création des amorces. Ainsi avant application sur un échantillon de terrain, une simulation de PCR *in silico*, au cours de laquelle la séquence de l'amorce serait confrontée à un très grand nombre de séquences, pourrait mettre en évidence les défauts de spécificité de l'amorce.

Ce manque de spécificité peut également s'expliquer par l'existence de variabilité intraspécifique dans toutes les séquences ADN qui pourrait grandement influencer l'hybridation des amorces de même que les sondes. Or dans cette simulation, on s'affranchit de cette variabilité en ayant choisi qu'une seule séquence pour représenter une espèce dans le lot de test. Comme évoqué précédemment et dans l'idéal, pour vérifier la spécificité *in silico* de chacune des amorces, une comparaison BLAST de leur séquence avec l'ensemble des séquences de GenBank serait nécessaire. Mais cette méthode serait très chronophage. Une autre alternative serait de limiter l'utilisation des amorces à un hôte et une localisation géographique. En effet, les amorces seraient alors spécifiques non pas d'une espèce mais d'une population au sein de cette espèce, pour laquelle la variabilité intraspécifique est plus faible. Cette alternative nécessite la création de paires d'amorces pour chaque communauté d'organismes à étudier, ce qui peut représenter un certain coût.

Mais l'existence de variation intraspécifique dans les séquences ADN contribue à un problème beaucoup plus large : une séquence nucléotidique unique suffit-elle pour caractériser une espèce ? Les résultats de la recherche bibliographique menée dans la partie I de cette thèse suggèrent plutôt de combiner les marqueurs génétiques. Il serait donc intéressant de réaliser un inventaire des amorces disponibles pour d'autres marqueurs et vérifier si une combinaison de ses marqueurs serait plus à même de caractériser une espèce malgré les variations intraspécifiques. Il paraît également difficile de déterminer un seuil de variation intraspécifique au-delà duquel les individus ne pourraient plus être considérés comme appartenant à la même espèce, tant ce niveau de variation diffère entre espèces. D'autre part, comme vu au III.C.2. de la partie 1, les sources d'erreurs de séquençage sont multiples et peuvent être considérées à tort comme de la variabilité génétique. C'est pourquoi certaines publications conseillent de combiner non pas les marqueurs génétiques mais les approches

d'identification. Très souvent, l'approche moléculaire est combinée avec l'approche morphologique (13,45,66).

Lorsqu'il s'agit d'étudier des communautés d'organismes, très souvent la méthode mise en place est en format multiplex afin d'identifier simultanément plusieurs espèces. Le risque majoritaire dans ce cas est l'hybridation potentielle des amorces entre elles au cours de la PCR. Or la simulation entreprise ici n'inclut pas la séquence des autres amorces dans le lot des séquences test. La formation de dimère d'amorces est problématique dans le sens où elle induit une compétition entre amorces pour l'hybridation sur leur séquence cible. Plusieurs techniques existent pour lesquelles les amorces ne sont pas utilisées ensembles dans les réactions de PCR, tout en permettant la caractérisation de plusieurs espèces en simultanée. La technique de la PCR multiplexe en tandem (MT-PCR) décrite au III.B.4. de la partie 1 a été notamment développée dans cet objectif. Cette méthode permet la réalisation d'une centaine à quelques centaines de réactions de PCR en simultanée. La technologie des puces microfluidiques offre des perspectives beaucoup plus importantes (124) mais n'a encore jamais été employée pour la caractérisation des communautés de nématodes gastro-intestinaux. Les puces renferment un ensemble de micro-canaux connectés entre eux, de manière à réaliser une fonction voulue comme une réaction de PCR, et reliés à l'extérieur par des entrées et des sorties percées à travers la puce. Les puces 46.46 ou 96.96 commercialisées par la compagnie américaine Fluidigm (www.fluidigm.com, San Francisco, Californie, Etats-Unis) permettent de réaliser 2 304 ou 9 216 réactions de PCR simultanément sur 48 ou 96 échantillons testés pour 48 ou 96 paires d'amorces respectivement. Outre la réalisation d'un grand nombre de réactions de PCR en même temps, cette technologie permet de réduire le nombre de manipulations puisque tout le processus se déroule dans le dispositif, les produits d'amplification étant directement détectés sur la puce. De plus, le coût de l'analyse est significativement réduit de par l'utilisation de faibles quantités de réactifs et la rapidité dans l'enchaînement des étapes. Cette technologie des puces microfluidiques n'est cependant applicable que pour les stratégies 2 et 3 pour lesquelles les espèces cibles des amorces employées sont connues. Les résultats obtenus pour la PCR *in silico* peuvent être représentés à la manière d'une puce microfluidique (tableau 3).

Tableau 3 : Tableau dynamique des résultats de la PCR, à la manière d'un résultat de puce microfluidique, inspiré de (124)

Les cases colorées correspondent aux espèces pour lesquelles un produit d'amplification simulé a été obtenu, ce qui signifie que les séquences des amorces étaient présentes, à 1 ou 2 bases près. Les cases grisées correspondent aux espèces pour lesquelles il n'y avait pas de produit d'amplification.

Genre	espèce	sous-espèce	NCL_NC2	Haem_1	Haem_2	Haem_3	Haem_4	Haem_5	Oste_6	Oste_7	Oste_8	Oste_9	Tela_10	Tela_11	Tela_12	Tela_13	Tela_14	Mars_15	Coop_16	Coop_17	Coop_18	Coop_19	Nema_20	Nema_21	Nema_22	Tric_23	Tric_24	Tric_25	Tric_26	Tric_27	Tric_28	Buno_29	Oeso_30	Oeso_31	Char_32	Char_33	Tric_34	Tric_35				
			Haemonchus	contortus																																						
Haemonchus	longistipes																																									
Ostertagia	lyrata																																									
Ostertagia	trifurcata																																									
Teladorsagia	circumcincta																																									
Ostertagia	ostertagi																																									
Ostertagia	leptospicularis																																									
Marshallagia	marshalli																																									
Cooperia	pectinata																																									
Cooperia	curticei																																									
Cooperia	punctata																																									
Cooperia	sumabada																																									
Cooperia	oncophora																																									
Nematodirus	davtiani	alpinus																																								
Nematodirus	helvetianus																																									
Nematodirus	filicollis																																									
Nematodirus	spathiger																																									
Nematodirus	battus																																									
Trichostrongylus	capricola																																									
Trichostrongylus	longispicularis																																									
Trichostrongylus	probolurus																																									
Trichostrongylus	rugatus																																									
Trichostrongylus	axei																																									
Trichostrongylus	colubriformis																																									
Trichostrongylus	vitrinus																																									
Bunostomum	trigonocephalum																																									
Oesophagostomum	asperum																																									
Oesophagostomum	venulosum																																									
Oesophagostomum	columbianum																																									
Chabertia	ovina																																									
Trichuris	discolor																																									
Trichuris	globulosa																																									
Trichuris	skrjabini																																									
Trichuris	ovis																																									
Camelostromylus	mentulatus																																									
Parabronema	skrjabini																																									
Spiculopteragia	boehmi																																									

Les stratégies employant des amorces spécifiques restent néanmoins limitées lorsqu'il s'agit de caractériser précisément des communautés d'organismes. Une autre approche possible, et déjà évoquée au I.B.2. de la partie 1, est l'approche du *metabarcoding*. Grâce aux nouvelles technologies de séquençage haut débit, cette approche permet l'analyse précise et détaillée de la composition en espèces d'échantillons complexes. Elle a déjà été appliquée pour l'étude des communautés de nématodes gastro-intestinaux des bovins mais s'étend également aujourd'hui aux ovins, bisons et chevaux. Elles consistent à séquencer la région ITS-2 de l'ensemble des molécules d'ADN présentes dans un mélange. Ces séquences sont ensuite comparées à des séquences références afin d'identifier l'espèce à laquelle elle appartient. Cette approche applique donc une stratégie de type 1. Sa particularité est qu'elle génère une quantité considérable de données permettant de déterminer précisément la composition en espèces d'un échantillon. Le résultat de cette approche est le projet *Nemabiome* accessible via l'adresse internet suivante : <https://www.nemabiome.ca/index.html>. Il comprend une base de données de séquences ITS-2 de plus de 1400 espèces de Nématodes, avec 2 609 séquences disponibles pour l'ordre des Strongyloides. Le projet *Nemabiome* propose également des protocoles standardisés pour la collecte et la préparation des échantillons, le séquençage et les analyses bio-informatiques pour assigner de nouvelles séquences à des espèces déjà référencées. Le

fragment considéré est généré avec des amorces génériques NC1 et NC2, légèrement modifiées. Depuis sa publication en 2015 (114), ce projet a été cité cinquante fois dans d'autres publications scientifiques. La dernière publication en date du deuxième semestre de 2020 porte notamment sur l'application du *nemabiome metabarcoding* pour améliorer le diagnostic de résistance aux anthelminthiques. Le *metabarcoding* est ici associé au dénombrement des œufs dans les fèces de brebis pour déterminer l'impact qualitatif et quantitatif d'un traitement anthelminthique (125). Un autre champ d'application où le *metabarcoding* présente un intérêt particulier est l'étude du parasitisme de la faune sauvage pour laquelle l'assemblage d'espèces de nématodes gastro-intestinaux n'est pas toujours connu à l'avance (126).

Conclusion de la partie 2

L'inventaire des séquences nucléotidiques dans la base de données de GenBank permet de mettre en évidence qu'un certain nombre de parasites n'y sont pas ou très peu représentés, avec une seule à zéro séquence disponible. D'autres espèces, au contraire, possèdent un nombre tellement important de séquences que la récupération de ses données manuellement n'est pas envisageable du fait du temps dépensé. Des outils d'exploitation de ces lots de données apparaissent nécessaires, d'autant plus avec l'avènement du *metabarcoding*, qui permet de traiter un nombre considérable de données en même temps. En plus du développement de tels outils, une harmonisation des annotations de séquences serait utile afin de pouvoir standardiser la méthode.

Parmi les marqueurs génétiques d'identification identifiés dans la première partie de cette thèse, la région ITS-2 possède les niveaux de variations inter- et intraspécifique les plus appropriés pour la caractérisation moléculaire des nématodes gastro-intestinaux des Caprinés au rang taxonomique de l'espèce. Les différentes stratégies d'identification exploitant ce marqueur emploient soit des amorces génériques, ciblant le rang taxonomique du phylum des Nématodes, soit des amorces spécifiques du genre ou de l'espèce. De telles amorces spécifiques ne sont pas disponibles pour toutes les espèces de nématodes gastro-intestinaux, de sorte que leur identification dans un mélange ne sera possible qu'avec l'utilisation des amorces génériques suivie d'un séquençage. La vérification *in silico* du spectre de parasites ciblés par les amorces montre un manque de spécificité majoritaire de ces amorces qui ne s'hybrident pas exclusivement avec la séquence de leur espèce cible. Un défaut de vérification lors de la création des séquences des amorces, l'existence d'une variabilité intraspécifique dans toute séquence nucléotidique et la possibilité d'erreurs de séquençage sont des hypothèses pouvant expliquer le manque de spécificité. La première pourrait être évitée en mettant en place, lors de la construction des amorces, une simulation de PCR *in silico* comme réalisée dans cette thèse mais avec un nombre plus important de séquences qui pourrait inclure l'ensemble des séquences disponibles de GenBank. Les deux dernières hypothèses mènent à combiner les marqueurs génétiques voire à combiner les approches d'identification dans une approche intégrative afin d'avoir une caractérisation optimale de l'espèce.

La caractérisation de communautés d'organismes nécessite très souvent de pouvoir identifier simultanément des espèces. La PCR multiplexe a été développée dans cet objectif mais avec un risque très élevé de formation de dimères d'amorces. De nouvelles technologies existent aujourd'hui afin d'éviter ce problème. Outre la PCR multiplexe en tandem dont le principe a été développé dans la première partie, la technologie des puces microfluidiques pourrait être employée. Enfin, le *metabarcoding*, grâce aux technologies de séquençage haut débit et la standardisation du protocole, permet la caractérisation plus précise des communautés. En effet, les insertions, délétions et substitutions présentes sur la longueur du brin séquencé donnent beaucoup plus d'informations que les seuls sites d'hybridation des amorces de PCR et des sondes internes.

CONCLUSION GENERALE

Les nématodes gastro-intestinaux représentent un groupe de parasites d'importance économique mondiale, tout particulièrement chez les Caprinés. Très souvent un même hôte est co-infesté par plusieurs espèces. La difficulté de leur identification morphologique est un frein pour la mise en place de traitements antiparasitaires optimaux qui dépendent d'une meilleure connaissance des dynamiques d'infestation par les espèces les plus pathogènes et des spécificités d'espèce d'hôte. Une approche combinée avec des méthodes moléculaires d'identification permettrait de surmonter cette difficulté.

L'ADN constitue un support physique universel et stable de l'information génétique. En comparant l'ADN de plusieurs individus, il est possible de définir des degrés de similitude ou de différence permettant de les classer en rangs taxonomiques. Le principe de l'identification moléculaire repose sur le fait qu'à une espèce sont associées des séquences caractéristiques de nucléotides, localisées au sein de régions de l'ADN plus ou moins variables d'une espèce à l'autre. Ainsi, les gènes *cox1* et *nad4* de l'ADN mitochondrial possèdent un niveau de variation interspécifique et surtout intraspécifique relativement élevé, ce qui les rend plus appropriées pour l'étude de populations d'organismes ayant un fort degré de similitude entre eux. Parmi les régions de l'ADN nucléaire, les marqueurs les plus utilisés sont les gènes codant pour l'ARN ribosomique, puisqu'il est commun à tous les organismes et est présent sous forme de répétitions conservées au sein du génome, avec une variabilité intraspécifique relativement faible. Les séquences inter-géniques non codantes ITS-1 et ITS-2 permettent de différencier au rang de l'espèce, contrairement aux séquences codantes 18S, 5.8S et 28S, qui sont très conservées et ne présentent des mutations caractéristiques qu'entre genres ou même entre rangs taxonomiques plus élevés.

L'accès à l'information génétique est permis grâce aux méthodes d'amplification et de séquençage des molécules d'ADN. Les différentes techniques basées sur le principe de la PCR amplifient l'ADN afin d'en obtenir une quantité suffisante pour être caractérisée. Le séquençage *sensu stricto* fournit la séquence nucléotidique d'une molécule d'ADN. Néanmoins, les technologies actuelles ne sont pas fiables à 100% et des erreurs peuvent être interprétées à tort comme des mutations. Les approches d'identification morphologique et moléculaire sont complémentaires et peuvent être regroupées dans une approche intégrative, qu'il convient de standardiser pour la rendre pratique à mettre en œuvre et comparable d'une étude à l'autre.

La caractérisation des nématodes gastro-intestinaux des Caprinés du parc de la Haute-Touche s'est heurtée à la difficulté de l'identification morphologique simultanée des espèces présentes, pour en étudier la spécificité d'espèce hôte et les fluctuations saisonnières. En préalable à une caractérisation moléculaire des communautés présentes, le présent travail est un inventaire des données génétiques déjà disponibles dans la banque de données Genbank. Il montre une grande disparité d'information génétique disponible, pour les différentes espèces de nématodes gastro-intestinaux. Une revue bibliographique sur les marqueurs génétiques utilisables pour ces taxons a permis de mettre en évidence plusieurs stratégies d'identification, par PCR suivie ou non de séquençage. Une simulation *in silico* des paires d'amorces publiées à ce jour, sur des séquences de référence par espèce, met en évidence un manque de spécificité pour le marqueur ribosomique ITS-2, le plus couramment utilisé pour l'identification. L'intégration d'une étape de PCR *in silico* dans le processus de sélection des amorces est ainsi intéressante pour vérifier leur spécificité, au vu des connaissances existantes, avant application sur des échantillons de terrain. Cette démarche permet aussi de compléter la panoplie d'amorces, pour détecter et différencier des séquences caractéristiques

d'un hôte et d'une localisation géographique donnée. Les technologies de PCR, de séquençage à haut-débit et les méthodes bio-informatiques permettent aujourd'hui l'analyse simultanée d'un grand nombre de données, rendant plus accessible la caractérisation moléculaire des communautés d'organismes.

Bibliographie

1. Massou J. Cortèges de nématodes digestifs associés à diverses espèces de Caprinés au sein d'une réserve zoologique [Thèse de doctorat vétérinaire]. [Oniris : Ecole Nationale Vétérinaire, Agroalimentaire et de l'Alimentation Nantes Atlantique]: Faculté de Médecine de Nantes; 2019.
2. Tariq KA. A Review of the Epidemiology and Control of Gastrointestinal Nematode Infections of Small Ruminants. *Proc Natl Acad Sci India Sect B Biol Sci.* juin 2015;85(2):693-703.
3. Jackson F, Bartley D, Bartley Y, Kenyon F. Worm control in sheep in the future. *Small Rumin Res.* oct 2009;86(1-3):40-5.
4. Reece JB, Campbell NA. *Campbell biologie.* Saint-Laurent; Montréal: ERPI ; Pearson; 2012.
5. Reece JB, Campbell NA. Chapitre 26 : La phylogénèse et l'arbre de la vie. In: *Campbell biologie.* Saint-Laurent; Montréal: ERPI ; Pearson; 2012. p. 619-42.
6. Laboratory Diagnosis of Parasitism. In: *Veterinary Parasitology* [Internet]. John Wiley & Sons, Ltd; 2015 [cité 10 déc 2019]. p. 259-312. Disponible sur: <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119073680.ch4>
7. Roeber F, Kahn L. The specific diagnosis of gastrointestinal nematode infections in livestock: Larval culture technique, its limitations and alternative DNA-based approaches. *Vet Parasitol.* oct 2014;205(3-4):619-28.
8. Veterinary Helminthology. In: *Veterinary Parasitology* [Internet]. John Wiley & Sons, Ltd; 2015 [cité 10 déc 2019]. p. 1-109. Disponible sur: <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119073680.ch1>
9. Wyrobisz A, Kowal J, Nosal P. Insight into species diversity of the Trichostrongylidae Leiper, 1912 (Nematoda: Strongylida) in ruminants. *J Helminthol.* nov 2016;90(6):639-46.
10. Stevenson LA, Gasser RB, Chilton NB. The ITS-2 rDNA of *Teladorsagia circumcincta*, *T. trifurcata* and *T. davtiani* (Nematoda: Trichostrongylidae) indicates that these taxa are one species. *Int J Parasitol.* oct 1996;26(10):1123-6.
11. Zarlenga DS, Hoberg EP, Stringfellow F, Lichtenfels JR. Comparisons of Two Polymorphic Species of *Ostertagia* and Phylogenetic Relationships within the *Ostertagiinae* (Nematoda: Trichostrongyloidea) Inferred from Ribosomal DNA Repeat and Mitochondrial DNA Sequences. *J Parasitol.* août 1998;84(4):806.
12. Newton L. Genetic evidence indicating that *Cooperia surnabada* and *Cooperia oncophora* are one species. *Int J Parasitol.* févr 1998;28(2):331-6.
13. Hoberg EP, Monsen KJ, Kutz S, Blouin MS. Structure, biodiversity, and historical biogeography of nematode faunas in holarctic ruminants: Morphological and molecular diagnoses for *Teladorsagia boreoarcticus* n. sp. (Nematoda: Ostertagiinae), a dimorphic cryptic species in muskoxen (*Ovibos moschatus*). *J Parasitol.* 1 oct 1999;85(5):910-34.
14. Zarlenga DS, Stringfellow F, Nobary M, Lichtenfels JR. Cloning and Characterization of Ribosomal RNA Genes from Three Species of *Haemonchus* (Nematoda: Trichostrongyloidea) and Identification of PCR Primers for Rapid Differentiation. *Exp Parasitol.* févr 1994;78(1):28-36.

15. Stevenson LA, Chilton NB, Gasser RB. Differentiation of *Haemonchus placei* from *H. contortus* (Nematoda: Trichostrongylidae) by the ribosomal DNA second internal transcribed spacer. *Int J Parasitol.* avr 1995;25(4):483-8.
16. Reece JB, Campbell NA. Chapitre 13 : La méiose et les cycles de développement sexués. In: *Campbell biologie*. Saint-Laurent; Montréal: ERPI ; Pearson; 2012. p. 281-98.
17. Ahmed M, Singh M, Bera A, Bandyopadhyay S, Bhattacharya D. Molecular basis for identification of species/isolates of gastrointestinal nematode parasites. *Asian Pac J Trop Med.* août 2011;4(8):589-93.
18. Reece JB, Campbell NA. Chapitre 23 : L'évolution des populations. In: *Campbell biologie*. Saint-Laurent; Montréal: ERPI ; Pearson; 2012. p. 543-64.
19. Blouin MS, Yowell CA, Courtney CH, Dame JB. *Haemonchus placei* and *Haemonchus contortus* are distinct species based on mtDNA evidence. *Int J Parasitol.* 1997;27(11):1383-7.
20. Hebert PDN, Cywinska A, Ball SL, deWaard JR. Biological identifications through DNA barcodes. *Proc R Soc Lond B Biol Sci.* 7 févr 2003;270(1512):313-21.
21. Coissac E, Riaz T, Puillandre N. Bioinformatic challenges for DNA metabarcoding of plants and animals: BIOINFORMATIC FOR DNA METABARCODING. *Mol Ecol.* avr 2012;21(8):1834-47.
22. Bik HM, Fournier D, Sung W, Bergeron RD, Thomas WK. Intra-Genomic Variation in the Ribosomal Repeats of Nematodes. Robinson-Rechavi M, éditeur. *PLoS ONE.* 11 oct 2013;8(10):e78230.
23. Chilton NB. The use of nuclear ribosomal DNA markers for the identification of bursate nematodes (order Strongylida) and for the diagnosis of infections. *Anim Health Res Rev.* déc 2004;5(2):173-87.
24. Zarlenga DS, Gasbarre LC, Boyd P, Leighton E, Lichtenfels JR. Identification and semi-quantitation of *Ostertagia ostertagi* eggs by enzymatic amplification of ITS-1 sequences. *Vet Parasitol.* juin 1998;77(4):245-57.
25. Zarlenga DS, Barry Chute M, Gasbarre LC, Boyd PC. A multiplex PCR assay for differentiating economically important gastrointestinal nematodes of cattle. *Vet Parasitol.* juin 2001;97(3):201-11.
26. Heise M, Epe C, Schnieder T. Differences in the second internal transcribed spacer (ITS-2) of eight species of gastrointestinal nematodes of ruminants. *J Parasitol.* juin 1999;85(3):431-5.
27. Ramünke S, de Almeida Borges F, von Son-de Fernex E, von Samson-Himmelstjerna G, Krücken J. Molecular marker sequences of cattle *Cooperia* species identify *Cooperia spatulata* as a morphotype of *Cooperia punctata*. Chiang T-Y, éditeur. *PLOS ONE.* 6 juill 2018;13(7):e0200390.
28. Hoste H, Gasser RB, Chilton NB, Mallet S, Beveridge I. Lack of intraspecific variation in the second Internal Transcribed Spacer (ITS-2) of *Trichostrongylus colubriformis* ribosomal DNA. *Int J Parasitol.* déc 1993;23(8):1069-71.
29. Gasser RB, Hoste H. Genetic markers for closely-related parasitic nematodes. *Mol Cell Probes.* oct 1995;9(5):315-9.

30. Hoste H, Chilton NB, Gasser RB, Beveridge I. Differences in the second internal transcribed spacer (Ribosomal DNA) between five species of *Trichostrongylus* (Nematoda: Trichostrongylidae). *Int J Parasitol.* janv 1995;25(1):75-80.
31. Hoste H, Chilton NB, Beveridge I, Gasser RB. A comparison of the first internal transcribed spacer of ribosomal DNA in seven species of *Trichostrongylus* (Nematoda: Trichostrongylidae)fn1fn1Nucleotide sequence data of the ITS-1 of the seven species have been deposited in the EMBL, GenBank and DDJB Databases (GenBank accession numbers Y15875-Y15881). *Int J Parasitol.* août 1998;28(8):1251-60.
32. Newton LA, Chilton NB, Beveridge I, Gasser RB. Differences in the second internal transcribed spacer of four species of *Nematodirus* (Nematoda: Molineidae). *Int J Parasitol.* 1 févr 1998;28(2):337-41.
33. Audebert F, Durette-Desset M-C, Chilton NB. Internal transcribed spacer rDNA can be used to infer the phylogenetic relationships of species within the genus *Nematodirus* (Nematoda: Molineoidea)1Note: Nucleotide sequence data have been deposited in the EMBL, GenBank and DDJB databases. GenBank accession numbers: AJ251569, AJ251570, AJ251571 and AJ251572.1. *Int J Parasitol.* 1 févr 2000;30(2):187-91.
34. Nadler SA, Hoberg EP, Hudspeth DSS, Rickard LG. RELATIONSHIPS OF *NEMATODIRUS* SPECIES AND *NEMATODIRUS BATTUS* ISOLATES (NEMATODA: TRICHOSTRONGYLOIDEA) BASED ON NUCLEAR RIBOSOMAL DNA SEQUENCES. *J Parasitol.* juin 2000;86(3):588-601.
35. Newton LA, Chilton NB, Beveridge I, Gasser RB. Systematic relationships of some members of the genera *Oesophagostomum* and *Chabertia* (Nematoda: Chabertiidae) based on ribosomal dna sequence datafn1fn1Note: Nucleotide sequence data reported in this paper are available in the embl, GenBankTM and DDJB databases under the accession numbers Y10790, Y11735, Y11736, Y11733, Y10789, AJ006149 and AJ006150. *Int J Parasitol.* nov 1998;28(11):1781-9.
36. Zhao GH, Hu B, Song JK, Jia YQ, Li HM, Wang CR, et al. Characterization of *Oesophagostomum asperum* and *O. columbianum* by internal transcribed spacers of nuclear ribosomal DNA. *J Helminthol.* mars 2014;88(1):74-81.
37. Dallas JF, Irvine RJ, Halvorsen O, Albon SD. Identification by polymerase chain reaction (PCR) of *Marshallagia marshalli* and *Ostertagia gruehneri* from Svalbard reindeer. *Int J Parasitol.* juin 2000;30(7):863-6.
38. Musters W, Boon K, van der Sande CA, van Heerikhuizen H, Planta RJ. Functional analysis of transcribed spacers of yeast ribosomal DNA. *EMBO J.* déc 1990;9(12):3989-96.
39. van der Sande CAFM, Kwa M, van Nues RW, van Heerikhuizen H, Raué HA, Planta RJ. Functional analysis of internal transcribed spacer 2 of *Saccharomyces cerevisiae* ribosomal DNA. *J Mol Biol.* févr 1992;223(4):899-910.
40. Gasser RB, Newton SE. Genomic and genetic research on bursate nematodes: significance, implications and prospects. *Int J Parasitol.* avr 2000;30(4):509-34.
41. Wimmer B, Craig BH, Pilkington JG, Pemberton JM. Non-invasive assessment of parasitic nematode species diversity in wild Soay sheep using molecular markers. *Int J Parasitol.* 1 avr 2004;34(5):625-31.

42. Nadler SA, Hoberg EP, Hudspeth DSS, Rickard LG. RELATIONSHIPS OF NEMATODIRUS SPECIES AND NEMATODIRUS BATTUS ISOLATES (NEMATODA: TRICHOSTRONGYLOIDEA) BASED ON NUCLEAR RIBOSOMAL DNA SEQUENCES. *J Parasitol.* juin 2000;86(3):588-601.
43. Chilton NB, Hoste H, Hung G-C, Beveridge I, Gasser RB. The 5.8S rDNA sequences of 18 species of bursate nematodes (order Strongylida): Comparison with rhabditid and tylenchid nematodes. *Int J Parasitol.* janv 1997;27(1):119-24.
44. Chilton NB, Huby-Chilton F, Gasser RB. First complete large subunit ribosomal RNA sequence and secondary structure for a parasitic nematode: phylogenetic and diagnostic implications. *Mol Cell Probes.* févr 2003;17(1):33-9.
45. Gouÿ de Bellocq J, Ferté H, Depaquit J, Justine J-L, Tillier A, Durette-Desset M-C. Phylogeny of the Trichostrongylina (Nematoda) Inferred from 28S rDNA Sequences. *Mol Phylogenet Evol.* juin 2001;19(3):430-42.
46. Carreno RA, Nadler SA. PHYLOGENETIC ANALYSIS OF THE METASTRONGYLOIDEA (NEMATODA: STRONGYLIDA) INFERRED FROM RIBOSOMAL RNA GENE SEQUENCES. *J Parasitol.* oct 2003;89(5):965-73.
47. Chilton NB, Huby-Chilton F, Gasser RB, Beveridge I. The evolutionary origins of nematodes within the order Strongylida are related to predilection sites within hosts. *Mol Phylogenet Evol.* juill 2006;40(1):118-28.
48. Callejón R, Nadler S, De Rojas M, Zurita A, Petrášová J, Cutillas C. Molecular characterization and phylogeny of whipworm nematodes inferred from DNA sequences of cox1 mtDNA and 18S rDNA. *Parasitol Res.* nov 2013;112(11):3933-49.
49. Floyd R, Abebe E, Papert A, Blaxter M. Molecular barcodes for soil nematode identification. *Mol Ecol.* avr 2002;11(4):839-50.
50. Alberts B. *Biologie moléculaire de la cellule.* Paris: Flammarion médecine-sciences; 2000.
51. Van Der Veer M, De Vries E. A single nucleotide polymorphism map of the mitochondrial genome of the parasitic nematode *Cooperia oncophora*. *Parasitology.* avr 2004;128(4):421-31.
52. Jex AR, Hu M, Littlewood DTJ, Waeschenbach A, Gasser RB. Using 454 technology for long-PCR based sequencing of the complete mitochondrial genome from single *Haemonchus contortus* (Nematoda). *BMC Genomics.* 2008;9(1):11.
53. Jex AR, Hall RS, Littlewood DTJ, Gasser RB. An integrated pipeline for next-generation sequencing and annotation of mitochondrial genomes. *Nucleic Acids Res.* 1 janv 2010;38(2):522-33.
54. Zhao G-H, Jia Y-Q, Cheng W-Y, Zhao W, Bian Q-Q, Liu G-H. Characterization of the complete mitochondrial genomes of *Nematodirus oiratianus* and *Nematodirus spathiger* of small ruminants. *Parasit Vectors.* 2014;7(1):319.
55. Sun M-M, Han L, Zhang F-K, Zhou D-H, Wang S-Q, Ma J, et al. Characterization of the complete mitochondrial genome of *Marshallagia marshalli* and phylogenetic implications for the superfamily Trichostrongyloidea. *Parasitol Res.* janv 2018;117(1):307-13.

56. Ahmad A, Yang X, Zhang T, Wang C, Zhou C, Yan X, et al. Characterization of the Complete Mitochondrial Genome of *Ostertagia trifurcata* of Small Ruminants and its Phylogenetic Associations for the Trichostrongyloidea Superfamily. *Genes*. 31 janv 2019;10(2):107.
57. Hu M, Chilton N, Gasser R. The Mitochondrial Genomics of Parasitic Nematodes of Socio-Economic Importance: Recent Progress, and Implications for Population Genetics and Systematics. In: *Advances in Parasitology* [Internet]. Elsevier; 2003 [cité 3 avr 2020]. p. 133-212. Disponible sur: <https://linkinghub.elsevier.com/retrieve/pii/S0065308X03560031>
58. Anderson TJC, Komuniecki R, Komuniecki PR, Jaenike J. Are mitochondria inherited paternally in *Ascaris*? *Int J Parasitol*. août 1995;25(8):1001-4.
59. Lynch M, Jarrell PE. A method for calibrating molecular clocks and its application to animal mitochondrial DNA. *Genetics*. déc 1993;135(4):1197-208.
60. Anderson TJC, Blouin MS, Beech RN. Population Biology of Parasitic Nematodes: Applications of Genetic Markers. In: *Advances in Parasitology* [Internet]. Elsevier; 1998 [cité 18 août 2020]. p. 219-83. Disponible sur: <https://linkinghub.elsevier.com/retrieve/pii/S0065308X0860425X>
61. Blouin MS, Yowell CA, Courtney CH, Dame JB. Substitution bias, rapid saturation, and the use of mtDNA for nematode systematics. *Mol Biol Evol*. 1 déc 1998;15(12):1719-27.
62. Blouin MS. Molecular prospecting for cryptic species of nematodes: mitochondrial DNA versus internal transcribed spacer. *Int J Parasitol*. mai 2002;32(5):527-31.
63. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol*. oct 1994;3(5):294-9.
64. Prosser SeanWJ, Velarde-Aguilar MariaG, León-Règagnon Virginia, Hebert PaulDN. Advancing nematode barcoding: A primer cocktail for the cytochrome c oxidase subunit I gene from vertebrate parasitic nematodes. *Mol Ecol Resour*. févr 2013;n/a-n/a.
65. Hebert PDN, Ratnasingham S, de Waard JR. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc R Soc Lond B Biol Sci* [Internet]. 7 août 2003 [cité 20 août 2020];270(suppl_1). Disponible sur: <https://royalsocietypublishing.org/doi/10.1098/rsbl.2003.0025>
66. Lv J, Zhang Y, Feng C, Yuan X, Sun D, Deng J, et al. Species discrimination in the subfamily Ostertagiinae of Northern China: assessment of DNA barcode in a taxonomically challenging group. *Parasitol Res*. mars 2016;115(3):987-96.
67. Troell K, Engström A, Morrison DA, Mattsson JG, Höglund J. Global patterns reveal strong population structure in *Haemonchus contortus*, a nematode parasite of domesticated ruminants. *Int J Parasitol*. 1 oct 2006;36(12):1305-16.
68. Gharamah AA, Azizah MNS, Rahman WA. Genetic variation of *Haemonchus contortus* (Trichostrongylidae) in sheep and goats from Malaysia and Yemen. *Vet Parasitol*. 10 sept 2012;188(3):268-76.
69. Yin F, Gasser RB, Li F, Bao M, Huang W, Zou F, et al. Genetic variability within and among *Haemonchus contortus* isolates from goats and sheep in China. *Parasit Vectors*. 25 sept 2013;6:279.

70. Hussain T, Periasamy K, Nadeem A, Babar ME, Pichler R, Diallo A. Sympatric species distribution, genetic diversity and population structure of *Haemonchus* isolates from domestic ruminants in Pakistan. *Vet Parasitol.* 15 déc 2014;206(3-4):188-99.
71. Shen D, Wang J, Zhang D, Peng Z, Yang T, Wang Z, et al. Genetic diversity of *Haemonchus contortus* isolated from sympatric wild blue sheep (*Pseudois nayaur*) and sheep in Helan Mountains, China. *Parasit Vectors* [Internet]. 19 sept 2017 [cité 5 sept 2020];10. Disponible sur: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5606089/>
72. Laosutthipong C, Eardmusic S. Genetic characterization of *Haemonchus contortus* from slaughtered goats in Cha-am District, Phetchaburi Province, Thailand. 2019;8.
73. Blouin MS, Yowell CA, Courtney CH, Dame JB. Host Movement and the Genetic Structure of Populations of Parasitic Nematodes. *Genetics.* nov 1995;141(3):1007-14.
74. Lehrter V, Jouet D, Liénard E, Decors A, Patrelle C. *Ashworthius sidemi* Schulz, 1933 and *Haemonchus contortus* (Rudolphi, 1803) in cervids in France: integrative approach for species identification. *Infect Genet Evol.* 1 déc 2016;46:94-101.
75. Seesao Y, Gay M, Merlin S, Viscogliosi E, Aliouat-Denis CM, Audebert C. A review of methods for nematode identification. *J Microbiol Methods.* juill 2017;138:37-49.
76. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci.* 1 déc 1977;74(12):5463-7.
77. Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. 1986. *Biotechnol Read Mass.* 1992;24:17-27.
78. Roos MH, Grant WN. Species-specific PCR for the parasitic nematodes *Haemonchus contortus* and *Trichostrongylus colubriformis*. *Int J Parasitol.* mai 1993;23(3):419-21.
79. Gasser RB, Chilton NB, Hoste H, Stevenson LA. Species identification of trichostrongyle nematodes by pcr-linked rflp. *Int J Parasitol.* avr 1994;24(2):291-3.
80. Gasser RB, Zhu X, Chilton NB, Newton LA, Nedergaard T, Guldberg P. Analysis of sequence homogenisation in rDNA arrays of *Haemonchus contortus* by denaturing gradient gel electrophoresis. *Electrophoresis.* oct 1998;19(14):2391-5.
81. von Samson-Himmelstjerna G, Harder A, Schnieder T. Quantitative analysis of ITS2 sequences in trichostrongyle parasites. *Int J Parasitol.* nov 2002;32(12):1529-35.
82. Metzker ML. Sequencing technologies — the next generation. *Nat Rev Genet.* janv 2010;11(1):31-46.
83. Schadt EE, Turner S, Kasarskis A. A window into third-generation sequencing. *Hum Mol Genet.* 15 oct 2010;19(R2):R227-40.
84. Rådström P, Knutsson R, Wolffs P, Lövenklev M, Löfström C. Pre-PCR Processing : Strategies to Generate PCR-Compatible Samples. *Mol Biotechnol.* 2004;26(2):133-46.
85. Seesao Y, Audebert C, Verrez-Bagnis V, Merlin S, Jérôme M, Viscogliosi E, et al. Monitoring of four DNA extraction methods upstream of high-throughput sequencing of Anisakidae nematodes. *J Microbiol Methods.* juill 2014;102:69-72.

86. Reece JB, Campbell NA. Chapitre 20 : La biotechnologie. In: Campbell biologie. Saint-Laurent; Montréal: ERPI ; Pearson; 2012. p. 459-94.
87. Kramer MF, Coen DM. Enzymatic Amplification of DNA by PCR: Standard Procedures and Optimization. *Curr Protoc Cytom.* juill 2006;37(1):A.3K.1-A.3K.15.
88. Wu DY, Ugozzoli L, Pal BK, Qian J, Wallace RB. The Effect of Temperature and Oligonucleotide Primer Length on the Specificity and Efficiency of Amplification by the Polymerase Chain Reaction. *DNA Cell Biol.* avr 1991;10(3):233-8.
89. Eckert KA, Kunkel TA. DNA polymerase fidelity and the polymerase chain reaction. *Genome Res.* 1 août 1991;1(1):17-24.
90. Bott NJ, Campbell BE, Beveridge I, Chilton NB, Rees D, Hunt PW, et al. A combined microscopic-molecular method for the diagnosis of strongylid infections in sheep. *Int J Parasitol.* sept 2009;39(11):1277-87.
91. Bisset SA, Knight JS, Bouchet CLG. A multiplex PCR-based method to identify strongylid parasite larvae recovered from ovine faecal cultures and/or pasture samples. *Vet Parasitol.* 24 févr 2014;200(1):117-27.
92. Roeber F, Jex AR, Campbell AJD, Nielsen R, Anderson GA, Stanley KK, et al. Establishment of a robotic, high-throughput platform for the specific diagnosis of gastrointestinal nematode infections in sheep. *Int J Parasitol.* déc 2012;42(13-14):1151-8.
93. Roeber F, Jex AR, Gasser RB. Next-Generation Molecular-Diagnostic Tools for Gastrointestinal Nematodes of Livestock, with an Emphasis on Small Ruminants. In: *Advances in Parasitology* [Internet]. Elsevier; 2013 [cité 26 sept 2020]. p. 267-333. Disponible sur: <https://linkinghub.elsevier.com/retrieve/pii/B9780124077058000045>
94. The C. elegans Sequencing Consortium. Genome Sequence of the Nematode C. elegans: A Platform for Investigating Biology. *Science.* 11 déc 1998;282(5396):2012-8.
95. Jaspard E. Methode sequencage sequencing NGS next generation high throughput technologies assemblage genome transcriptome genomique fonctionnelle vegetale RNA seq Enseignement et recherche Biochimie - Universite Angers Emmanuel Jaspard biochimej [Internet]. [cité 1 oct 2020]. Disponible sur: <http://biochimej.univ-angers.fr/Page2/COURS/9ModulGenFoncVeg/5MethEtudGenFonc/1MethodeSEQUENGAGE/1SEQUENGAGE.htm>
96. Song H, Buhay JE, Whiting MF, Crandall KA. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proc Natl Acad Sci.* 9 sept 2008;105(36):13486-91.
97. Richly E. NUMTs in Sequenced Eukaryotic Genomes. *Mol Biol Evol.* 12 févr 2004;21(6):1081-4.
98. Hurst GDD, Jiggins FM. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proc R Soc B Biol Sci.* 7 août 2005;272(1572):1525-34.
99. Potapov V, Ong JL. Examining Sources of Error in PCR by Single-Molecule Sequencing. Kalendar R, éditeur. *PLOS ONE.* 6 janv 2017;12(1):e0169774.

100. Schmitt MW, Kennedy SR, Salk JJ, Fox EJ, Hiatt JB, Loeb LA. Detection of ultra-rare mutations by next-generation sequencing. *Proc Natl Acad Sci.* 4 sept 2012;109(36):14508-13.
101. Kinde I, Wu J, Papadopoulos N, Kinzler KW, Vogelstein B. Detection and quantification of rare mutations with massively parallel sequencing. *Proc Natl Acad Sci.* 7 juin 2011;108(23):9530-5.
102. Hoang ML, Kinde I, Tomasetti C, McMahon KW, Rosenquist TA, Grollman AP, et al. Genome-wide quantification of rare somatic mutations in normal human tissues using massively parallel sequencing. *Proc Natl Acad Sci.* 30 août 2016;113(35):9846-51.
103. Preston JL, Royall AE, Randel MA, Sikkink KL, Phillips PC, Johnson EA. High-specificity detection of rare alleles with Paired-End Low Error Sequencing (PELE-Seq). *BMC Genomics.* déc 2016;17(1):464.
104. Hugot J-P, Baujard P, Morand S. Biodiversity in helminths and nematodes as a field of study: an overview. *Nematology.* 2001;3(3):199-208.
105. Chartier C, Reche B. Gastrointestinal helminths and lungworms of French dairy goats: Prevalence and geographical distribution in Poitou-Charentes. *Vet Res Commun.* sept 1992;16(5):327-35.
106. Zaffaroni E, Teresa Manfredi M, Citterio C, Sala M, Piccolo G, Lanfranchi P. Host specificity of abomasal nematodes in free ranging alpine ruminants. *Vet Parasitol.* juin 2000;90(3):221-30.
107. Torina A, Dara S, Marino AMF, Sparagano O a. E, Vitale F, Reale S, et al. Study of Gastrointestinal Nematodes in Sicilian Sheep and Goats. *Ann N Y Acad Sci.* 2004;1026(1):187-94.
108. Mayo E, Ortiz J, Martínez-Carrasco C, Garijo MM, Espeso G, Hervías S, et al. First description of gastrointestinal nematodes of Barbary sheep (*Ammotragus lervia*): the case of *Camelostrongylus mentulatus* as a paradigm of phylogenetic and specific relationship between the parasite and its ancient host. *Vet Res Commun.* sept 2013;37(3):209-15.
109. Zanzani S, Gazzonis A, Di Cerbo A, Varady M, Manfredi M. Gastrointestinal nematodes of dairy goats, anthelmintic resistance and practices of parasite control in Northern Italy. *BMC Vet Res.* 2014;10(1):114.
110. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the Tidyverse. *J Open Source Softw.* 21 nov 2019;4(43):1686.
111. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing [Internet]. Vienna, Austria: R fondation for Statistical Computing; 2020. Disponible sur: <https://www.R-project.org/>
112. Rice P, Longden I, Bleasby A. EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet.* juin 2000;16(6):276-7.
113. Blankenberg D, Taylor J, Schenck I, He J, Zhang Y, Ghent M, et al. A framework for collaborative analysis of ENCODE data: Making large-scale analyses biologist-friendly. *Genome Res.* 1 juin 2007;17(6):960-4.
114. Avramenko RW, Redman EM, Lewis R, Yazwinski TA, Wasmuth JD, Gilleard JS. Exploring the Gastrointestinal “Nemabiome”: Deep Amplicon Sequencing to Quantify the Species

- Composition of Parasitic Nematode Communities. PLoS ONE [Internet]. 2 déc 2015 [cité 3 déc 2019];10(12). Disponible sur: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4668017/>
115. Paguem A, Abanda B, Ngwasiri NN, Eisenbarth A, Renz A, Streit A, et al. Host specificity and phylogeny of Trichostrongylidae of domestic ruminants in the Guinea savannah of the Adamawa plateau in Cameroon. *Vet Parasitol Reg Stud Rep*. juill 2020;21:100412.
 116. Schnieder T, Heise M, Epe C. Genus-specific PCR for the differentiation of eggs or larvae from gastrointestinal nematodes of ruminants. *Parasitol Res*. 24 sept 1999;85(11):895-8.
 117. Learmount J, Conyers C, Hird H, Morgan C, Craig BH, von Samson-Himmelstjerna G, et al. Development and validation of real-time PCR methods for diagnosis of *Teladorsagia circumcincta* and *Haemonchus contortus* in sheep. *Vet Parasitol*. déc 2009;166(3-4):268-74.
 118. Demeler J, Ramünke S, Wolken S, Ianiello D, Rinaldi L, Gahutu JB, et al. Discrimination of Gastrointestinal Nematode Eggs from Crude Fecal Egg Preparations by Inhibitor-Resistant Conventional and Real-Time PCR. *PLOS ONE*. 19 avr 2013;8(4):e61285.
 119. Santos MC, Amarante MRV, Amarante AFT. Establishment of co-infection and hybridization of *Haemonchus contortus* and *Haemonchus placei* in sheep. *J Helminthol*. nov 2019;93(06):697-703.
 120. McNally J, Callan D, Andronicos N, Bott N, Hunt PW. DNA-based methodology for the quantification of gastrointestinal nematode eggs in sheep faeces. *Vet Parasitol*. déc 2013;198(3-4):325-35.
 121. Gasser RB, Chilton NB, Hoste H, Beveridge I. Rapid sequencing of rDNA from single worms and eggs of parasitic helminths. *Nucleic Acids Res*. 1993;21(10):2525-6.
 122. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, et al. GenBank. *Nucleic Acids Res*. 26 nov 2012;41(D1):D36-42.
 123. Zhang Z, Schwartz S, Wagner L, Miller W. A Greedy Algorithm for Aligning DNA Sequences. *J Comput Biol*. févr 2000;7(1-2):203-14.
 124. Michelet L, Delannoy S, Devillers E, Umhang G, Aspan A, Juremalm M, et al. High-throughput screening of tick-borne pathogens in Europe. *Front Cell Infect Microbiol* [Internet]. 29 juill 2014 [cité 15 nov 2020];4. Disponible sur: <http://journal.frontiersin.org/article/10.3389/fcimb.2014.00103/abstract>
 125. Queiroz C, Levy M, Avramenko R, Redman E, Kearns K, Swain L, et al. The use of ITS-2 rDNA nemabiome metabarcoding to enhance anthelmintic resistance diagnosis and surveillance of ovine gastrointestinal nematodes. *Int J Parasitol Drugs Drug Resist*. déc 2020;14:105-17.
 126. Aivelo T, Medlar A. Opportunities and challenges in metabarcoding approaches for helminth community identification in wild mammals. *Parasitology*. 2018;145(5):608-21.
 127. GenBank Sample Record [Internet]. [cité 25 nov 2020]. Disponible sur: <https://www.ncbi.nlm.nih.gov/Sitemap/samplerecord.html#Top>

Annexe 1 : Organisation des données et métadonnées GenBank.

Lorsqu'une séquence est soumise à GenBank, un certain nombre d'informations, renseignées par le dépositaire via des outils internet tels que BankIt ou Sequin, l'accompagne.

Un numéro d'accèsion est attribué à chaque donnée GenBank et sert de numéro unique de référence, aussi bien dans la banque de données que dans les publications. Il renvoie systématiquement à la dernière version disponible de la séquence. Les versions antérieures sont identifiées avec deux autres numéros uniques : l'identifiant *Accession.version* et le numéro GI. L'*Accession.version* correspond au numéro d'accèsion agrémenté du numéro de la version. Ainsi si la séquence de numéro d'accèsion XXXX a été modifiée 5 fois, l'*Accession.version* sera noté : XXXX.5. Le numéro GI est unique à chaque version (122).

Les métadonnées qui accompagnent une séquence correspondent aux données biologiques associées à cette séquence et sont accessibles dans la base de données « Nucleotide » de GenBank (127).

Catégorie	Description
LOCUS	Informations diverses dont la longueur de la séquence, le type de molécule (ADN génomique, ARN messenger, ARN ribosomique ...), la date de modification.
DEFINITION	Brève description de la séquence qui reprend notamment l'organisme source, le nom du gène ou de la protéine correspondante.
SOURCE	Information en format-libre qui peut reprendre l'organisme source, le type de molécule. ORGANISM : nom scientifique et classification taxonomique de l'organisme source (lien vers la base de données « Taxonomy » de NCBI).
REFERENCE	Publications des auteurs associées à la séquence. Certaines séquences ne sont pas rapportées dans une publication. Elles sont mentionnées comme étant "unpublished" (non publiées) ou "in press" (en cours de publication). AUTHORS : liste des auteurs de la publication. TITLE : titre de la publication ou titre préventif en cas de non publication. JOURNAL : nom du journal duquel est issue la publication. PUBMED : identifiant PMID qui renvoie à la publication dans la base de données PUBMED. Direct Submission : dernière citation de la catégorie REFERENCE qui donne des informations sur le dépositaire de la séquence.
FEATURES	Annotations des gènes et des produits de ces gènes, ainsi que des régions d'importance biologique Source : informations diverses sur la séquence dont le nom scientifique de l'organisme source, le type de molécule, l'organisme hôte, le numéro Taxonomy ID associé à la base de données « Taxonomy », l'organisme hôte du parasite, le pays dans lequel la séquence a été trouvée. Gene : région d'intérêt biologique identifiée comme étant un gène particulier. CDS : séquence codante et information sur la protéine codée dont la séquence en acide-aminés correspondante.
ORIGIN	Séquence de nucléotides.

Annexe 2 : Tableau des espèces de nématodes gastro-intestinaux des Caprinés - nombre de séquences dans Genbank.

Vert : espèces pour lesquelles le nombre de séquences nucléotidiques disponibles dans GenBank est nul.

Rouge : espèces pour lesquelles le nombre de séquences nucléotidiques disponibles dans GenBank est supérieur à 500.

Famille	Genre	espèce	ID Taxonomy GenBank	Nombre de séquences dans GenBank	Date de la recherche
Ancylostomatidae	Bunostomum	trigonocephalum	66065	152	04/06/2020
Haemonchidae	Camelostrongylus	mentulatus	53334	16	04/06/2020
Capillariidae	Capillaria	longipes		0	04/06/2020
Chabertiidae	Chabertia	ovina	63233	77	04/06/2020
Cooperidae	Cooperia	curticei	62055	24	18/12/2019
Cooperidae	Cooperia	oncophora	27828	961	26/12/2019
Cooperidae	Cooperia	punctata	96640	132	29/06/2020
Cooperidae	Cooperia	pectinata	420891	43	04/01/2020
Cooperidae	Cooperia	surnabada	62054	1	05/01/2020
Ancylostomatidae	Gaigeria	pachyscelis		0	04/06/2020
Gongylonematidae	Gongylonema	monnigi		0	04/06/2020
Gongylonematidae	Gongylonema	pulchrum	637853	24 435	04/06/2020
Gongylonematidae	Gongylonema	verrucosum		0	04/06/2020
Haemonchidae	Haemonchus	contortus	6289	48 071	16/03/2020
Haemonchidae	Haemonchus	longistipes	239092	48	05/01/2020
Haemonchidae	Marshallagia	marshalli	106947	34	30/06/2020
Haemonchidae	Marshallagia	occidentalis	146285	20	04/06/2020
Molineidae	Nematodirus	davtani alpinus	94039	1	04/06/2020
Molineidae	Nematodirus	battus	28839	24	12/12/2019
Molineidae	Nematodirus	filicollis	61840	10	12/12/2019
Molineidae	Nematodirus	spathiger	61841	62	12/12/2019
Molineidae	Nematodirus	helvetianus	61842	105	12/12/2019
Molineidae	Nematodirus	abnormalis		0	12/12/2019
Molineidae	Nematodirus	lanceolatus		0	02/06/2020
Chabertiidae	Oesophagostomum	columbianum	85896	19	04/06/2020
Chabertiidae	Oesophagostomum	venulosum	63234	32	15/02/2020
Chabertiidae	Oesophagostomum	asperum	1213747	165	15/02/2020
Chabertiidae	Oesophagostomum	multifoliatum		0	15/02/2020
Haemonchidae	Ostertagia	ostertagi	6317	7 414	05/01/2020
Haemonchidae	Ostertagia	lyrata	72296	5	07/01/2020
Haemonchidae	Ostertagia	leptospicularis	52815	42	07/01/2020
Haemonchidae	Ostertagia	trifurcata	109209	5	13/01/2020
Haemonchidae	Ostertagia (Skrjabinagia)	kolchida	263385	6	04/06/2020
Habronematidae	Parabronema	skrjabini	499165	30	04/06/2020
Oxyuridae	Skrjabinema	ovis	2109847	1	04/06/2020
Oxyuridae	Skrjabinema	caprae		0	02/06/2020
Oxyuridae	Skrjabinema	alata		0	02/06/2020

Famille	Genre	espèce	ID Taxonomy GenBank	Nombre de séquences dans GenBank	Date de la recherche
Haemonchidae	Spiculopteragia	spiculoptera	126399	80	04/06/2020
Haemonchidae	Spiculopteragia	boehmi	239091	1	15/02/2020
Strongyloididae	Strongyloides	papillosus	174720	4797	04/06/2020
Haemonchidae	Teladorsagia	circumcincta	45464	88 238	04/06/2020
Haemonchidae	Teladorsagia	davtiani	293654	45	15/02/2020
Haemonchidae	Teladorsagia	pinnata		0	02/06/2020
Haemonchidae	Teladorsagia	trifurcata	1213752	3	04/06/2020
Trichostrongylidae	Trichostrongylus	axei	40349	443	03/03/2020
Trichostrongylidae	Trichostrongylus	affinus		0	15/02/2020
Trichostrongylidae	Trichostrongylus	colubriformis	6319	305	30/06/2020
Trichostrongylidae	Trichostrongylus	vitrinus	40352	436	15/02/2020
Trichostrongylidae	Trichostrongylus	capricola	126400	4	15/02/2020
Trichostrongylidae	Trichostrongylus	falculatus		0	15/02/2020
Trichostrongylidae	Trichostrongylus	longispicularis	2025163	2	15/02/2020
Trichostrongylidae	Trichostrongylus	probolurus	53330	8	15/02/2020
Trichostrongylidae	Trichostrongylus	rugatus	53331	5	15/02/2020
Trichuridae	Trichuris	ovis	93034	187	04/06/2020
Trichuridae	Trichuris	discolor	483153	42	04/06/2020
Trichuridae	Trichuris	globulosa	1616775	11	04/06/2020
Trichuridae	Trichuris	skrjabini	198341	41	04/06/2020

Annexe 3 : Extrait de l'inventaire des séquences GenBank constitué dans cette étude et publications associées.

Le « p » mentionné dans la nature de certaines séquences signifie « partielle ». Par exemple, la séquence « ITS-2p/28p » correspond à la séquence IST-2-28S pour laquelle l'extrémité 5' ne comprend pas l'ensemble de la séquence ITS-2, de même pour l'extrémité 3' avec la séquence 28S. Le caractère complet ou partiel concerne la région centrale dans la séquence et non les extrémités. Par exemple : la séquence « 5.8p/ITS-2/18p » est indiquée comme complète puisque la région centrale ITS-2 est comprise entre deux autres régions.

« NM » signifie « non mentionné » : lorsque l'information est absente dans les annotations GenBank et dans la publication associée.

N° d'accession	Espèce de Nématode	Matériel génétique			Espèce hôte	Localisation géographique	Publication	Accès ?	Date de la recherche
		Nature	Séquence partielle ou complète ?	Longueur fragment (paires de bases)					
AJ239113.1	<i>Nematodirus davitiani alpinus</i>	ITS-2	NM	228	Italie	DOI: 10.1016/S0020-7519(99)00123-X	oui	12/12/2019	
KC998800.1	<i>Nematodirus filicollis</i>	18p/TTS-1/5.8p	Complète	714	Nouvelle-Zélande	DOI: 10.1016/j.vetpar.2013.12.002	oui	12/12/2019	
KC998801.1	<i>Nematodirus filicollis</i>	18p/TTS-1/5.8p	Complète	713	Nouvelle-Zélande	DOI: 10.1016/j.vetpar.2013.12.002	oui	12/12/2019	
KC998802.1	<i>Nematodirus filicollis</i>	18p/TTS-1/5.8p	Complète	713	Nouvelle-Zélande	DOI: 10.1016/j.vetpar.2013.12.002	oui	12/12/2019	
KC998749.1	<i>Nematodirus filicollis</i>	5.8p/TTS-2/28p	Complète	451	Nouvelle-Zélande	DOI: 10.1016/j.vetpar.2013.12.002	oui	12/12/2019	
KC998750.1	<i>Nematodirus filicollis</i>	5.8p/TTS-2/28p	Complète	451	Nouvelle-Zélande	DOI: 10.1016/j.vetpar.2013.12.002	oui	12/12/2019	
AF194139.1	<i>Nematodirus filicollis</i>	18p/TTS-1/5.8/TTS-2/28p	Complète	863	Etats-Unis	DOI: 10.1645/0022-3395(2000)086[0588:RONSAN]2.0.CO;2	oui	12/12/2019	
AF194140.1	<i>Nematodirus filicollis</i>	18p/TTS-1/5.8/TTS-2/28p	Complète	863	Etats-Unis	DOI: 10.1645/0022-3395(2000)086[0588:RONSAN]2.0.CO;2	oui	12/12/2019	
AY439024.1	<i>Nematodirus filicollis</i>	5.8p/TTS-2c	Complète	488	Ecosse	DOI: 10.1016/j.ijpara.2003.11.022	oui	12/12/2019	
AJ251572.1	<i>Nematodirus filicollis</i>	18p/TTS-1/5.8p	Complète	386	Allemagne & Australie	DOI: 10.1016/S0020-7519(99)00179-4	oui	12/12/2019	
Y14011.1	<i>Nematodirus filicollis</i>	ITS-2	NM	236	Allemagne & Australie	DOI: 10.1016/S0020-7519(97)00150-1	oui	12/12/2019	
KF266763.1	<i>Nematodirus battus</i>	18p/TTS-1/5.8/TTS-2/28p	Complète	835	Irlande	Non publié	non	12/12/2019	
JF345079.1	<i>Nematodirus battus</i>	18p/TTS-1/5.8/TTS-2/28p	Complète	960	Irlande	Non publié	non	12/12/2019	
AY295828.1	<i>Nematodirus battus</i>	<i>rml</i> génomique	Partielle	932	Etats-Unis	DOI: http://dx.doi.org/10.1645/GE-76R	oui	12/12/2019	
AY292799.1	<i>Nematodirus battus</i>	ARNr 28S	Partielle	881	Etats-Unis	DOI: http://dx.doi.org/10.1645/GE-76R	oui	12/12/2019	
AF194138.1	<i>Nematodirus battus</i>	18p/TTS-1/5.8/TTS-2/28p	Complète	859	Angleterre	DOI: 10.1645/0022-3395(2000)086[0588:RONSAN]2.0.CO;2	oui	12/12/2019	

Annexe 4 : Nombre de séquences de marqueurs génétiques disponibles pour 40 espèces de Nématodes gastro-intestinaux de Caprinés, recensées dans cette étude.

Genre	espèce	sous-espèce	18S	ITS1	5.8S	ITS2	28S	cox1	nad4
Bunostomum	trigonocephalum		3	4	9	8	8	38	0
Camelostrongylus	mentulatus		0	0	2	14	2	0	0
Chabertia	ovina		9	8	14	22	16	6	6
Cooperia	curticei		4	11	12	14	13	0	0
Cooperia	oncophora		33	126	41	140	36	32	100
Cooperia	pectinata		0	6	6	21	6	0	0
Cooperia	punctata		3	13	22	57	23	0	0
Cooperia	sumabada		0	0	0	1	0	0	0
Haemonchus	longistipes		16	16	6	27	11	6	1
Marshallagia	marshalli		3	7	16	29	24	2	0
Marshallagia	occidentalis		2	5	6	15	6	5	0
Nematodirus	battus		19	17	18	18	18	0	0
Nematodirus	davtiani	alpinus	0	0	0	1	0	0	0
Nematodirus	filicollis		6	6	9	6	4	0	0
Nematodirus	helvetianus		4	12	80	103	72	0	0
Nematodirus	spathiger		8	14	39	49	30	2	0
Oesophagostomum	asperum		23	44	44	44	23	22	0
Oesophagostomum	columbianum		0	8	8	9	0	5	0
Oesophagostomum	venulosum		4	4	7	19	15	1	0
Ostertagia	kolchida		0	0	0	0	0	0	6
Ostertagia	leptospicularis		6	6	10	10	8	1	21
Ostertagia	lyrata		0	2	1	2	2	2	0
Ostertagia	trifurcata		1	1	2	4	3	0	0
Parabronema	skrjabini		5	22	25	27	17	0	0
Skrjabinema	ovis		0	0	0	0	1	0	0
Spiculoptera	boehmi		0	0	0	1	0	0	0
Spiculoptera	spiculoptera		2	6	6	27	5	0	38
Teladorsagia	davtiani		0	1	1	1	1	2	42
Teladorsagia	trifurcata		1	2	2	2	2	1	0
Trichostrongylus	axei		7	10	17	56	24	188	186
Trichostrongylus	capricola		0	0	0	2	2	0	0
Trichostrongylus	colubriformis		10	21	57	190	117	18	0
Trichostrongylus	longispicularis		0	0	0	2	2	0	0
Trichostrongylus	probolurus		0	3	3	5	0	1	0
Trichostrongylus	rugatus		0	1	1	3	2	0	0
Trichostrongylus	vitrinus		6	9	14	40	29	0	0
Trichuris	discolor		15	25	25	25	14	3	0
Trichuris	globulosa		0	0	0	6	0	3	0
Trichuris	ovis		3	38	39	48	6	70	0
Trichuris	skrjabini		9	4	2	2	3	12	3
Total Résultat			202	452	544	1050	545	420	403
Moyenne			5	11	14	26	14	11	10
Ecart-type			7,2	21,2	18,0	39,1	21,8	31,8	33,7
Nombre d'espèces possédant au moins une séquence du marqueur			25	31	32	38	33	21	9

Annexe 5 : Tableau récapitulatif des séquences des amorces et sondes publiées pour la région ITS-2 des nématodes gastro-intestinaux et leur spécificité attendue.

Les séquences des amorces aussi bien *forward* que *reverse* sont données dans le sens 5'-3'. L'amorce *forward* s'hybride sur le brin principal de la molécule d'ADN alors que l'amorce *reverse* s'hybride sur le brin complémentaire.

La mention « sp. » dans la colonne « espèce » signifie que l'amorce ou la sonde est spécifique du genre.

Les paires d'amorces sont signalées en bleu.

Genre	espèce	Séquences de nucléotides		Source biblio.	Nom abrégé
		Forward	Reverse		
Haemonchus	sp.	GAGGGCTAATTTCAACATTG	CAAAATGATAAAAGAACATCGTC	41	Haem_1
Haemonchus	contortus	CCATATACTACAATGTGGCTAATTC	TACAAATGATAAAAGAACATCGTCGC	118	Haem_2
Haemonchus	contortus	TGGCGACGATGTTT		81	
Haemonchus	contortus	GCGAATAATTGAGATTTGACTTAGATAGAGAC	GCTCAGGTTGCAITATACAAAATGATAAAA	81	Haem_3
Haemonchus	contortus	CAAATGGCATTGTCTTTTAG	TTAGTTTCITTTTCTCCGGCT	37	Haem_4
Haemonchus	contortus	ATTTGACAACAATGTTGAAAT		119	
Haemonchus	contortus	CAAATGATAAAAAGAACATCGTCG		119	
Haemonchus	contortus	CAACGTGATGTTATGAAAATTTGTAACA		91	
Haemonchus	contortus	CATGTATGGCGACGATGTTCTT		91	
Haemonchus	contortus		TCATATCATTACAGGGATGTTACAATTT	91	
Haemonchus	contortus		CTTGAACTGAAATGGGAATTTGICT	91	
Haemonchus	contortus	TTCCCATTTTCAGTTCAAGAAC	AATCTCAATATTCGCTGAGTACAC	117	Haem_5
Haemonchus	contortus	ACAATTTTCATAAACAATCACGTTGC		117	
Haemonchus	contortus	TCAAGAACATATACATGCA		120	
Haemonchus	contortus	TGACATGTATGGCGACGAT		120	
Haemonchus	longistipes	CAAATGATAAAAAGAACATCGTCA		119	
Ostertagia	sp.	GTCGAAATGGTATTTAATTACT	TTAGTTTCITTTTCTCCGGCT	116	Oste_6
Ostertagia	sp.	CATATTCATTGAGTACATTCAAATAGTG		119	
Ostertagia	ostertagi	TAACATTTGTTAACGTTACTGAATGATACTG	ATATAAATGATACATCGAATATACAATAC	118	Oste_7
Ostertagia	ostertagi	ACTTTAATT-GTGATAAATTCCTCCATTT		119	
Ostertagia	ostertagi		CAATGTTAACGTCATGTTGCATTTCA	91	
Ostertagia	ostertagi		AGTACATTCAAATAGTGGTAATATATTTTCAG	91	
Ostertagia	leptospicularis	TTCCCATTTTCAGTTCAAGAAAACAC	TTCATTTGAGTACATTCAAAATAGTGATAATATA	118	Oste_8
Ostertagia	leptospicularis	GTCGAAATGGTATTTAATTACT		81	
Ostertagia	leptospicularis	ATGAAAACACTACTACAGTGTGGCTAACACA	TTCTTTGAACTGAAATGGGAATTAATCA	81	Oste_9
Ostertagia	leptospicularis	CATGCAACATAAACGTTAACATAATG		91	
Teladorsagia	sp.	ACTTTAATT-GTGATAAATTCCTCCATTA		119	
Teladorsagia	circumcincta	AACATATGCAACATGACGTACGACGG	ATGATACATTTGAAACATATATTACCATACATGCT	118	Tela_10
Teladorsagia	circumcincta	TATGCAACATGACGTACGACGG	TTAGTTTCITTTTCTCCGGCT	90	Tela_11
Teladorsagia	circumcincta	ATGCAACATGACGTACGACGG		91	
Teladorsagia	circumcincta	AAACTACTACAGTGTGGCTAACATA		91	
Teladorsagia	circumcincta		GTACATTCAAAATAGTAGCAATACCGC	91	

Genre	espèce	Séquences de nucléotides		Source biblio.	Nom abrégé
		Forward	Reverse		
Teladorsagia	circumcincta	TCACATTTTGTGTCACAAATTATC	CAGGAACGTTACGACGGTAAT	41	Tela_12
Teladorsagia	circumcincta	CTTAATGATCTCGCCTAGACG	TTTCATCGATACGCCGAATCG	117	Tela_13
Teladorsagia	circumcincta	TCACAATTAATCGATTAGCTTCAGC		117	
Teladorsagia	circumcincta	CTTAATGATCTCGCCTAGACG	TTATATGTTAGCCACACTGTAGTAG	117	Tela_14
Teladorsagia	circumcincta	TCACAATTAATCGATTAGCTTCAGC		117	
Teladorsagia	circumcincta	GTTCAAGAATAACATATGC		120	
Teladorsagia	circumcincta	CCGTCGTAACGTTCCCTGAAT		120	
Marshallagia	marshalli	TCCTGAATGATATGAATGTAATTACC	CAATACAAAATGATATAATTGAACATACAG	37	Mars_15
Cooperia	sp.	TAATGGCATTGTCTACATCT	ATGATAACGAAATACTACTATCT	116	Coop_16
Cooperia	curticei	CTAATGGCATTGTCTACATTG		81	
Cooperia	curticei	GTGTGGCTAGCGTTTTAACACTGT	TCTTGAACATAAATGGGATTTGTCAGAA	81	Coop_17
Cooperia	curticei	TTTAAACAGTGATAATAGATTTC		119	
Cooperia	curticei	TTGGTTCTGACAAATCCCT		119	
Cooperia	curticei	TAATGGCATTGTCTACATTGGTTC		91	
Cooperia	punctata	TTTAAACAGTGATAATAGATTTC		119	
Cooperia	punctata	TTTGTCTAATGGYATYTGCTACAC		91	
Cooperia	punctata		AGATTTCATATCAITTCAGAAAATGTTTCAC	91	
Cooperia	surabada	TTTTAAACAGTGATAATAGATTG		119	
Cooperia	oncophora	ATGGCATTGTCTACATCTGTTT	AAATGATAACGAAATACTACTATCTCCA	118	Coop_18
Cooperia	oncophora	CATTTCTGAATGATAICAATCTAATTATC	TTAGTTTCITTTTCTCCGCT	90	Coop_19
Cooperia	oncophora	TTTTAAACAGTGATAATAGATTG		119	
Cooperia	oncophora	TATAGTAATATGAACATTTCTGAATGATATC		91	
Cooperia	oncophora		CTATAACGGGATTTGTCAAAAACAGA	91	
Nematodirus	sp.	GATAGAATACAGCATGGAITGT	AGCAGTCGATTTGTTCCACAGC	116	Nema_20
Nematodirus	filicollis	TGTCGAAATGGTACTTGCTGTTA		91	
Nematodirus	filicollis		GGGATTGACTGTTACGATGTAA	91	
Nematodirus	filicollis	CAGTCAATCCCATTCAAGTGAA	GTTAAAGAGCAGGTCCCAGATC	41	Nema_21
Nematodirus	spathiger		CATTCAGGAGCTTTGACACTAAT	91	
Nematodirus	battus	CCGGTATACCCATTCAAGTAAGG	TACAGTCAGTCCCGGTTG	41	Nema_22
Trichostrongylus	sp.	TTCTACAGTGTGGCTAACTCT	TTGAAACATTATTCATACAT	116	Tric_23
Trichostrongylus	sp.	TCGAATGGTCATTGTCAA	TTAGTTTCITTTCTCCGCT	90	Tric_24
Trichostrongylus	sp.	CATACAGGGATATTAATGTCGTTA		119	

Genre	espèce	Séquences de nucléotides		Source biblio.	Nom abrégé
		Forward	Reverse		
Trichostrongylus	sp.	YAGTTYAAGAAAYAATAAYA		120	
Trichostrongylus	sp.	CCTGTATGATGTGAACGTGT		120	
Trichostrongylus	axei	CATACAGGGATATTAATGTCGTT		119	
Trichostrongylus	axei	CAAATATTGTGATAAATCCCAATTTAGTTT		91	
Trichostrongylus	axei	GATGTTAATGTTGAACGACATTAATATC		91	
Trichostrongylus	axei	AGGGATAATTAATGTCGTTCA	TGATAATTCCCAATTTAGTTT	41	Tric_25
Trichostrongylus	colubriformis	CTTACGTCTGGTTCAGGGTTG	ACTGAAATGGGAAATCATCACAATATTT	118	Tric_26
Trichostrongylus	colubriformis	ACAGTGTGGCTAACTC		81	
Trichostrongylus	colubriformis	CTTACGTCTGGTTCAGGGTTGTT	ACTGAAATGGGAAATCATCACAATATTT	81	Tric_27
Trichostrongylus	colubriformis	TTCCTGAACTGAAATGGGAATC		119	
Trichostrongylus	colubriformis		ACATCATACAGGAAACATTAATGTCA	91	
Trichostrongylus	colubriformis		ATTATTCCTTGAACTGAARTGGGAATC	91	
Trichostrongylus	vitrinus	ATGTGAACGTGTTGTCACCTGTTTA		91	
Trichostrongylus	vitrinus	CAATTATCAATTTGTATTTGCAACCTGAT		91	
Trichostrongylus	vitrinus	AGGAAACATTAATGTCGTTACA	CTGTTTGTGCGAATGGTTATTA	41	Tric_28
Bunostomum	trigonocephalum		CGTTTAGITTAAGTCGACAAAACGA	91	
Bunostomum	trigonocephalum	AGTAATTCCCATTC AAGAGCA	GCAAATAACATAAAATATCATTTGTT	41	Buno_29
Oesophagostomum	sp.	AACGCTTTGACAGTGACAAAC		119	
Oesophagostomum	asperum	TCACAAGCCACACACTGTAGTAAT		119	
Oesophagostomum	venulosum	TGAAATGAGACAACCGTAGTCG	TTAGTTTCCTTTCCCTCCGCT	90	Oeso_30
Oesophagostomum	venulosum		CGACTACGGTTGTCTCATTTCA	91	
Oesophagostomum	venulosum		ATACATGCATGCATACATCACATG	91	
Oesophagostomum	columbianum	TGTCGAAACGATGCTTGCTTT	TTAGTTTCCTTTCCCTCCGCT	90	Oeso_31
Oesophagostomum	radiatum	AACGCTTTGACAGTGACAAG		119	
Charbertia	ovina	GATGACCTCGTTGTCAACCGTG	TTAGTTTCCTTTCCCTCCGCT	90	Char_32
Charbertia	ovina	TCACAAGCCACACACTGTAGTAAC		119	
Charbertia	ovina	CCTTGGAGCAATCCAAGGTG		91	
Charbertia	ovina	CAGGACTAAGAAATGCTTTGG		91	
Charbertia	ovina	CATGTGTGATCCTCGTACTAGATAAGA	ATGAACCGTACACCGTTGTCA	41	Char_33
Trichuris	sp.	GCGGCAGTGTGGATCTGGCTG	GTGATCCGCCGTTCCGGAATG	118	Tric_34
Trichuris	ovis	GTCGGCTTTTACCCTCCTA	ACCAGGCACCTAGCTTCCA	41	Tric_35

Annexe 6 : Tableau des résultats de la PCR *in silico*.

Chaque cellule du tableau contient la taille en paires de bases du fragment amplifié par la simulation de PCR.

Genre	espèce	sous-espèce	NCL_NC2	Haem_1	Haem_2	Haem_3	Haem_4	Haem_5	Oste_6	Oste_7	Oste_8	Oste_9	Tela_10	Tela_11	Tela_12	Tela_13	Tela_14	Mars_15	Coop_16	Coop_17
Haemonchus	contortus		321	0	226	78	261	115	0	0	0	0	0	0	0	140	0	0	0	0
Haemonchus	longispines		0	0	234	0	0	0	0	0	0	0	0	0	0	140	0	0	0	0
Ostertagia	lyrata		328	0	0	0	0	0	268	124	0	0	0	0	0	133	0	0	0	0
Ostertagia	trifurcata		0	0	0	0	0	0	0	0	0	91	0	0	324	136	279	0	0	0
Teladorsagia	circumcincta		336	0	0	0	0	0	276	0	0	91	147	219	320	136	279	0	0	0
Ostertagia	ostertagi		328	0	0	0	0	0	268	124	0	0	0	0	0	133	0	0	0	0
Ostertagia	leptospicularis		0	0	0	0	0	0	0	0	106	91	0	0	0	136	0	0	0	0
Marshallagia	marshalli		326	0	0	0	0	0	266	0	0	91	0	0	0	0	279	110	0	0
Cooperia	pectinata		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cooperia	curticei		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	79
Cooperia	punctata		331	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	192	79
Cooperia	sumabada		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	192	79
Cooperia	oncophora		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	192	79
Nematodirus	daviani	alpinus	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Nematodirus	helveticus		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Nematodirus	filicollis		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Nematodirus	spathiger		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Nematodirus	battus		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Trichostrongylus	capricola		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Trichostrongylus	longispicularis		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Trichostrongylus	probolus		0	0	0	0	0	0	0	0	0	0	0	0	0	136	0	0	0	0
Trichostrongylus	rugatus		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Trichostrongylus	axei		0	0	0	0	0	0	0	0	0	0	0	0	0	136	0	0	0	0
Trichostrongylus	colubriformis		328	0	0	0	0	0	0	0	0	0	0	0	0	136	0	0	0	0
Trichostrongylus	vitrinus		328	0	0	0	0	0	0	0	0	0	0	0	0	136	0	0	0	0
Bunostomum	trigonocephalum		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Oesophagostomum	asperum		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Oesophagostomum	venulosum		348	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Oesophagostomum	columbianum		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Chabertia	ovina		325	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Trichuris	discolor		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Trichuris	globulosa		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Trichuris	skrjabinii		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Trichuris	ovis		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Camelostomylus	mentulatus		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Parabronema	skrjabinii		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Spiculoptera	boehmi		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Nombre de PCR positives			10	0	2	1	1	1	4	2	1	4	1	1	2	11	3	1	3	4

Genre	espèce	sous-espèce	Coop_18	Coop_19	Nema_20	Nema_21	Nema_22	Tric_23	Tric_24	Tric_25	Tric_26	Tric_27	Tric_28	Bumo_29	Oeso_30	Oeso_31	Char_32	Char_33	Tric_34	Tric_35
Haemonchus	contortus		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Haemonchus	longistipes		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ostertagia	lyrata		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ostertagia	trifurcata		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Teladorsagia	circumcincta		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ostertagia	ostertagi		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ostertagia	leptospicularis		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Marshallagia	marshalli		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cooperia	pectinata		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cooperia	curticei		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cooperia	punctata		192	180	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cooperia	surabada		192	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cooperia	oncophora		192	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Nematodirus	daviani	alpinus	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Nematodirus	helvetianus		0	0	211	145	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Nematodirus	flicollis		0	0	0	145	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Nematodirus	spathiger		0	0	211	145	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Nematodirus	battus		0	0	0	0	142	0	0	0	0	0	0	0	0	0	0	0	0	0
Trichostrongylus	capricola		0	0	0	0	0	218	0	0	0	0	104	0	0	0	0	0	0	0
Trichostrongylus	longispicularis		0	0	0	0	0	0	0	0	0	0	104	0	0	0	0	0	0	0
Trichostrongylus	probolus		0	0	0	0	0	0	0	0	0	0	104	0	0	0	0	0	0	0
Trichostrongylus	rugatus		0	0	0	0	0	218	0	0	0	0	104	0	0	0	0	0	0	0
Trichostrongylus	axei		0	0	0	0	0	0	0	73	0	0	0	0	0	0	0	0	0	0
Trichostrongylus	colubriformis		0	0	0	0	0	0	269	0	0	0	104	0	0	0	0	0	0	0
Trichostrongylus	virinus		0	0	0	0	0	0	0	0	0	0	104	0	0	0	0	0	0	0
Bunostomum	trigonocephalum		0	0	0	0	0	0	0	0	0	0	0	171	0	0	0	0	0	0
Oesophagostomum	asperum		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Oesophagostomum	venulosum		0	0	0	0	0	0	0	0	0	0	0	0	106	0	0	0	0	0
Oesophagostomum	columbianum		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Chabertia	ovina		0	0	0	0	0	0	0	0	0	0	0	0	0	0	163	158	0	0
Trichuris	discolor		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	308	0
Trichuris	globulosa		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	177
Trichuris	skrjabini		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	278	0
Trichuris	ovis		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	308	182
Camelostomylus	mentulatus		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Parabronema	skrjabini		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Spiculopteria	boehmi		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Nombre de PCR positives			3	1	2	3	1	2	1	1	0	0	6	1	1	0	1	1	3	2

Annexe 7 : Tableau des capacités de détection attendues dans les publications d'amorces, confrontées aux résultats de la PCR *in silico*.

Vert : espèces pour lesquelles la paire d'amorce est bien spécifique.

« Ø » signifie qu'il n'y a pas eu de produit d'amplification pour cette paire d'amorce.

Rouge : espèces pour lesquelles un produit de PCR n'était pas attendu, selon la publication d'origine, alors qu'un produit est obtenu *in silico*.

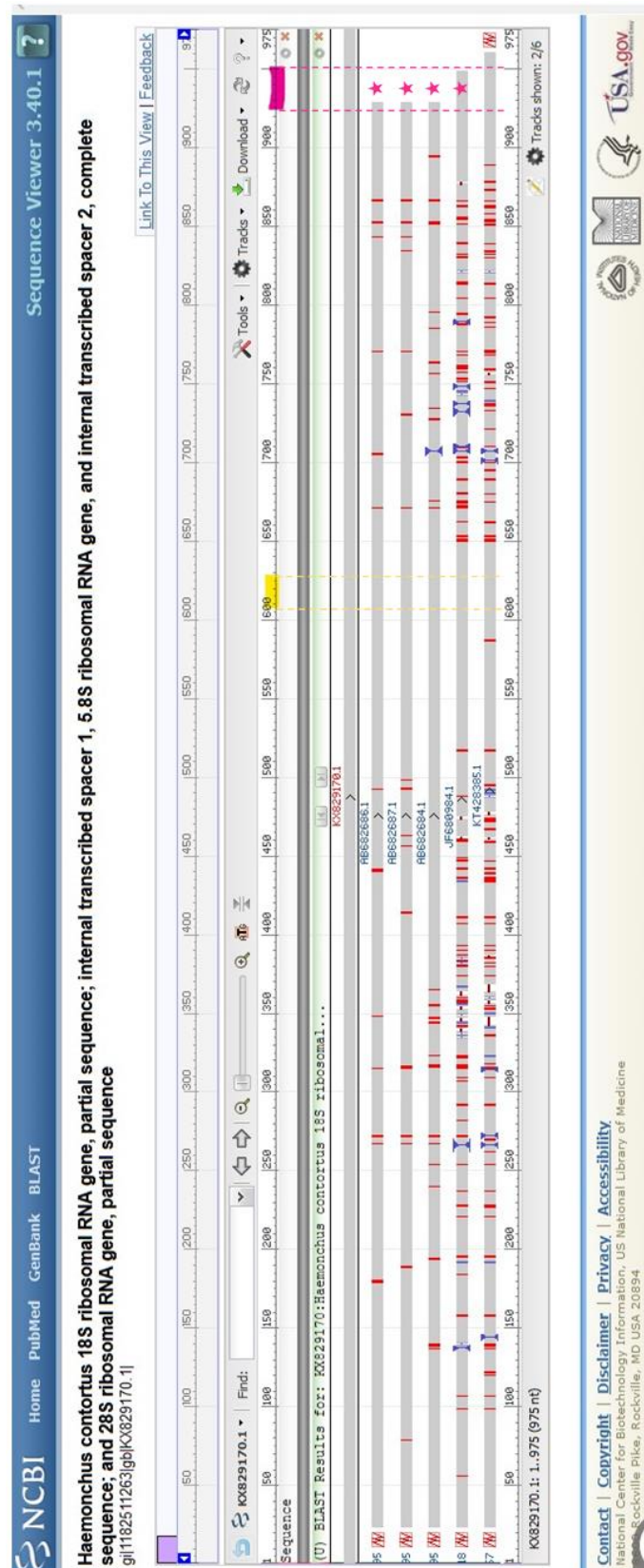
Nom abrégé	Hypothèse	Résultat	
NC1_NC2	Nématode	Phylum	H. contortus O. lyrata T. circumcincta O. ostertagi M. marshalli C. punctata T. colubriformis T. vitrinus O. venulosum C. ovina
Haem_1	Haemonchus sp.	Genre	Ø
Haem_2	Haemonchus contortus	Espèce	H. contortus H. longistipes
Haem_3	Haemonchus contortus	Espèce	H. contortus
Haem_4	Haemonchus contortus	Espèce	H. contortus
Haem_5	Haemonchus contortus	Espèce	H. contortus
Oste_6	Ostertagia sp.	Genre	O. lyrata T. circumcincta O. ostertagi M. marshalli
Oste_7	Ostertagia ostertagi	Espèce	O. lyrata O. ostertagi
Oste_8	Ostertagia leptospicularis	Espèce	O. leptospicularis
Oste_9	Ostertagia leptospicularis	Espèce	O. trifurcata T. circumcincta O. leptospicularis M. marshalli
Tela_10	Teladorsagia circumcincta	Espèce	T. circumcincta
Tela_11	Teladorsagia circumcincta	Espèce	T. circumcincta
Tela_12	Teladorsagia circumcincta	Espèce	O. trifurcata T. circumcincta
Tela_13	Teladorsagia circumcincta	Espèce	H. contortus H. longistipes O. lyrata O. trifurcata T. circumcincta O. ostertagi M. marshalli T. probolurus T. axei T. colubriformis T. vitrinus

Tela_14	Teladorsagia circumcincta	Espèce	O. trifurcata T. circumcincta M. marshalli
Mars_15	Marshallagia marshalli	Espèce	M. marshalli
Coop_16	Cooperia sp.	Genre	C. punctata C. surnabada C. oncophora
Coop_17	Cooperia curticei	Espèce	C. curticei C. punctata C. surnabada C. oncophora
Coop_18	Cooperia oncophora	Espèce	C. punctata C. surnabada C. oncophora
Coop_19	Cooperia oncophora	Espèce	C. punctata
Nema_20	Nematodirus sp.	Genre	N. helvetianus N. spathiger
Nema_21	Nematodirus filicollis	Espèce	N. helvetianus N. filicollis N. spathiger
Nema_22	Nematodirus battus	Espèce	N. battus
Tric_23	Trichostrongylus sp.	Genre	T. capricola T. rugatus T. colubriformis T. vitrinus
Tric_24	Trichostrongylus sp.	Genre	T. colubriformis
Tric_25	Trichostrongylus axei	Espèce	T. axei
Tric_26	Trichostrongylus colubriformis	Espèce	∅
Tric_27	Trichostrongylus colubriformis	Espèce	∅
Tric_28	Trichostrongylus vitrinus	Espèce	T. capricola T. longispicularis T. probolurus T. rugatus T. colubriformis T. vitrinus
Buno_29	Bunostomum trigonocephalum	Espèce	B. trigonocephalum
Oeso_30	Oesophagostomum venulosum	Espèce	O. venulosum
Oeso_31	Oesophagostomum columbianum	Espèce	∅
Char_32	Charbertia ovina	Espèce	C. ovina
Char_33	Charbertia ovina	Espèce	C. ovina
Tric_34	Trichuris sp.	Genre	T. discolor T. skrjabini T. ovis
Tric_35	Trichuris ovis	Espèce	T. globulosa T. ovis

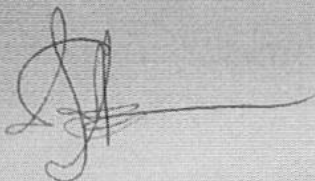
Annexe 8 : Variabilité interspécifique des séquences pour les positions des amorces NC1 et NC2.

Les séquences appartiennent aux espèces : *Haemonchus contortus* (AB682686.1, AB682687.1, AB682684.1), *Teladorsagia circumcincta* (JF680984) et *Ostertagia ostertagi* (KT428385.1).


L'emplacement de l'amorce NC1 et de l'amorce NC2 est spécifié par un rectangle jaune et un rectangle rose respectivement. Les étoiles roses indiquent les séquences pour lesquelles le site de fixation de NC2 est absent ou incomplet.



Vu : **L'enseignant Rapporteur**
De l'Ecole Nationale Vétérinaire,
Agroalimentaire et de l'Alimentation
Oniris

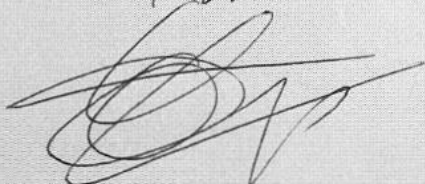


Vu : **La Directrice Générale**
De l'Ecole Nationale Vétérinaire,
Agroalimentaire et de l'Alimentation
Oniris
Laurence Deflesselle
Pour ordre et par délégation



Sandy LECOQ-ESPALLARGAS
Nantes, le 23/04/2020

Vu : **Le Président de la Thèse**
Professeur **Yves MAUCARS**



Vu : **Le Doyen de la Faculté de**
Médecine de Nantes
Professeur Pascale JOLLIET

Vu et permis d'imprimer

NOM : BEDOUE
Prénom : Laurie

OUTILS DE CARACTERISATION MOLECULAIRE DES COMMUNAUTES DE NEMATODES GASTRO-INTESTINAUX CHEZ LES CAPRINES

RESUME

Les infections des Caprinés par les nématodes gastro-intestinaux sont très souvent multiples. Les techniques de diagnostics morphologiques sont limitées pour identifier simultanément les espèces en cause. Le développement d'outils de caractérisation moléculaire de communautés de parasites nécessite la sélection d'un marqueur génétique d'identification ainsi que des données de référence en quantité suffisante. La région ITS-2 de l'ADN ribosomique apparaît comme la plus appropriée pour distinguer les nématodes au rang de l'espèce. L'inventaire des séquences nucléotidiques disponibles dans la base de données GenBank des nématodes gastro-intestinaux infestant les Caprinés illustre de grandes disparités entre les espèces. Trois stratégies de caractérisation moléculaire sont définies. Elles emploient soit des amorces génériques soit des amorces ciblant des espèces en particulier dont la spécificité est vérifiée par une simulation de PCR. Les limites concernant cette simulation *in silico* sont abordées et de nouvelles perspectives pour l'étude des communautés de nématodes gastro-intestinaux sont évoquées.

MOTS CLES

- Nématodes gastro-intestinaux
- Caprinés
- Marqueur génétique
- PCR
- GenBank
- Séquençage

JURY

Président : Monsieur le Professeur Yves MAUGARS

Rapporteur : Madame Suzanne BASTIAN, Maître de Conférences à ONRIS

Assesseur : Monsieur Albert AGOULON, Maître de Conférences à ONRIS

ADRESSE DE L'AUTEUR

Laurie BEDOUE
39 rue Jean Marie Boivin
37550 SAINT-AVERTIN

Nom de l'imprimeur
TOP OFFICE