



HAL
open science

Apports de la paléogénétique à l'étude des helminthes gastro-intestinaux anciens

Nathalie Côté

► **To cite this version:**

Nathalie Côté. Apports de la paléogénétique à l'étude des helminthes gastro-intestinaux anciens. Ecologie, Environnement. Université de Franche-Comté, 2015. Français. NNT : 2015BESA2013 . tel-01456919

HAL Id: tel-01456919

<https://theses.hal.science/tel-01456919>

Submitted on 6 Feb 2017

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

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

Université de Bourgogne Franche-Comté

U.F.R SCIENCES ET TECHNIQUES

ÉCOLE DOCTORALE ENVIRONNEMENT-SANTÉ

Laboratoire Chrono-environnement (UMR UFC/CNRS 6249/USC INRA)

Thèse en vue de l'obtention du titre de

DOCTEUR DE L'UNIVERSITÉ DE BOURGOGNE FRANCHE-COMTÉ

SPÉCIALITÉ SCIENCES DE LA VIE ET DE L'ENVIRONNEMENT

**APPORTS DE LA PALÉOGÉNÉTIQUE À L'ÉTUDE DES
HELMINTHES GASTRO-INTESTINAUX ANCIENS**

Présentée et soutenue publiquement par

Marie Louise Nathalie Côté

le 16 décembre 2015

Membres du jury :

Olivier Dutour (Professeur, École Pratique des Hautes Études)	Rapporteur
Michel Drancourt (Professeur, Université de la Méditerranée)	Rapporteur
Eva-Maria Geigl (Directrice de recherche CNRS, Université Paris Diderot)	Présidente du jury
Thierry Grange (Directeur de recherche CNRS, Université Paris Diderot)	Examineur
Francis Raoul (MCF, HDR, Université de Bourgogne Franche-Comté, Besançon)	Examineur
Catherine Thèves (Chargée de recherche CNRS, Université Paul Sabatier, Toulouse)	Examinatrice
Matthieu Le Bailly (MCF, Université de Bourgogne Franche-Comté, Besançon)	Co-directeur de thèse
Nicolas Capelli (MCF, HDR, Université de Bourgogne Franche-Comté, Besançon)	Directeur de thèse

Remerciements

Je tiens tout d'abord à remercier le Professeur Olivier Dutour (Université de Bordeaux) et le Professeur Michel Drancourt (Université de la Méditerranée) pour avoir accepté d'être rapporteurs de cette thèse. Je remercie aussi le Docteur Catherine Thèves (Université Paul Sabatier) et le Docteur Francis Raoul (Université de Bourgogne Franche-Comté) pour leur participation comme membres du jury et examinateurs.

Matthieu, je sais que tu avais envie depuis longtemps d'introduire cet aspect paléogénétique pour la paléoparasitologie. Je te remercie de m'avoir donné la chance de participer à ce beau projet. Ta disponibilité et ta patience pour répondre à mes questions sur les parasites et la paléoparasitologie (est-ce un œuf ? Non, c'est une bulle d'air!) m'ont permis de découvrir cette discipline totalement nouvelle pour moi mais aussi de m'initier à l'archéologie.

Nicolas, tu as accepté d'être le directeur de cette thèse de paléogénétique. Malgré une thématique plutôt éloignée de tes sujets de recherches de prédilection, je te remercie de l'intérêt que tu as démontré au cours de ces trois dernières années.

Malgré la distance, vous m'avez fait confiance tous les deux pour toutes les étapes de cette thèse et avez toujours été présents et disponibles pour des relectures, des discussions et aide administrative. Faire cette thèse sur deux universités aurait pu être beaucoup compliqué sans votre souplesse et compréhension.

Eva-Maria et Thierry, vous avez accepté de m'accueillir dans votre laboratoire pour ce qui devait être six mois... et qui s'est transformé en trois ans ! Bien sûr, sans vous, cette thèse n'aurait pas été possible. Votre enthousiasme et dévouement aux sciences sont contagieux, et vous avez su créer autour de vous une équipe dynamique et motivée.

À tous les archéologues, qui sont trop nombreux pour être nommés ici : merci de nous avoir confié vos précieux vestiges pour des analyses paléoparasitologiques et paléogénétiques. Plus particulièrement, je voudrais remercier Corinne Thevenet de l'Université de Nanterre et de l'Inrap pour m'avoir initié aux fouilles archéologiques et d'avoir prélevé « en direct » du sédiment pour cette thèse. Cette fouille a été l'occasion de rencontrer d'autres spécialistes, tel Lamys Hachem, qui a pris le temps de m'expliquer les rudiments de l'archéozoologie, ou Michael Ilett, avec à qui j'ai découvert mon premier squelette.

Je remercie aussi tous les biologistes, aussi très nombreux, qui ont généreusement accepté de donner de l'ADN ou des spécimens de parasites, pour la mise au point de cette étude. Jenny Knapp, de l'Université de Franche-Comté, m'a entre autre mis en contact avec plusieurs chercheurs en parasitologie. Je la remercie aussi pour ses conseils sur la purification de l'ADN d'*Echinococcus*.

Les congrès sont l'occasion de créer des contacts au niveau scientifique et humain. Je salue les chercheurs que j'ai pu rencontrer à Londres, à Lund, à Bordeaux et à Besançon. Je suis heureuse de mentionner en particulier Annelise Binois, doctorante en archéozoologie de l'Université Paris 1, pour sa contribution avec des échantillons provenant d'un contexte historique très intéressant... qui n'ont malheureusement pas donné les résultats escomptés !

Au cours de ces trois années, j'ai aussi été entourée de collègues merveilleux, qui m'ont aidé pour mes expériences ou avec qui j'ai partagé des projets (et des cafés). Mélanie, même si nous n'avons pas partagé de café, je te remercie de ton aide au quotidien, toi l'experte M Plex ! Tu as une telle patience, un tel calme, tu es un modèle pour moi, ne change pas ! Ce fut un plaisir de partager ces moments au labo avec toi, mais aussi en dehors du labo, de voir le développement de ta petite famille. Je pense aussi à Silvia, qui a toujours été disponible pour m'expliquer le fonctionnement des nombreux appareils que j'ai utilisés. C'est aussi toi qui m'a initié aux « règles » du labo-fossile et qui a partagé ma détresse lorsque le robot faisait des siennes ! Ton dynamisme est précieux, ce fut très agréable de travailler avec toi. Andy, continue de parler français, tu te débrouilles comme un chef ! Merci pour ton aide avec la langue anglaise, le spiking de l'ADN de Lambda, de caribou...Diyendo, cher Mario, tu m'as tellement fait rire au cours de ces années, merci pour ces beaux moments de discussions plus ou moins scientifiques, et de découvertes de série (Shameless !) Et merci pour ton aide sur Illumina lors de mes derniers jours au labo. Giulia qui, malgré toute la meilleure volonté, n'arrive pas à me faire aligner plus de trois mots à la suite en italien ! Il ne nous reste plus qu'à continuer à manger des gelato. Je suis très heureuse de t'avoir rencontré et découvert ton coin de pays. Je sais que nous resterons en contact, peu importe où la vie nous portera. Que mon collègue de bureau Olivier soit remercié pour sa patience et sa bonne humeur, et pour m'avoir enduré pendant tout ce temps ! Merci aussi d'avoir accepté de partager certains de tes échantillons et de m'avoir expliqué tout ce que je sais sur le MiSeq (même si l'appareil a planté de manière inusitée l'unique fois que je l'ai utilisé...). Finalement, même s'il est maintenant parti vers de nouveaux horizons, Julien a amplement contribué au succès de ce projet, en gérant avec brio le séquençage, mais aussi en transmettant son savoir-faire avant son départ.

J'espère que ton dictionnaire d'expressions québécoises continuera de s'enrichir ! Que du succès dans tes nouveaux projets, Julien !

Je pense aussi à tous ceux qui ont fait un séjour plus ou moins long dans l'équipe : Clémentine, Damien et Samantha, qui ont obtenu avec succès leur master recherche. Bon début d'aventure à Sam, pour cette thèse qui sera couronnée de succès, j'en suis certaine ! Claudio, bon courage pour la suite, collègue temporaire de bureau !

Je remercie les autres doctorants de l'Institut Jacques Monod pour ces moments passés ensemble, sur le canapé bleu, à la cuisine du 5^{ème} ou une fois par mois, sur la terrasse du 3^{ème} étage.

Outre les membres du labo, je remercie les amis que j'ai côtoyés au cours de ces trois dernières années. Les Iranian girls de Paris, avec qui j'ai passé de superbes moments. Je remercie entre autre Sara, qui a eu le même parcours que moi depuis la maîtrise à Besançon, et qui a toujours eu une oreille attentive. J'ai une pensée pour Chou, ses lettres manuscrites, ses boucles d'oreilles artisanales. Mes trois séjours à la Cité Universitaire m'ont permis de rencontrer des gens merveilleux, originaires d'un peu partout sur la planète, et partager un repas, un verre ou faire un jogging avec vous a été très enrichissant.

Merci aussi à mon coloc Thomas pour l'initiation à l'escalade, m'avoir aidé à insérer correctement une légende de figure et pour les apéro-Bescherelle. Et aussi à Jef qui a installé en urgence un ordinateur après que j'ai cassé mon Dell en arrivant au Canada ! Autres canadiens à remercier : ma cousine Jo qui a imprimé ce manuscrit pour que je puisse le corriger, et Steph, qui m'a hébergé durant ma « retraite » rédaction.

Je souhaite aussi remercier Christophe, qui a été présent dès le début de cette aventure, et bien avant. Bien que nos vies aient maintenant pris des chemins séparés, tu as continué à croire en moi, et cela fut important pour moi.

Finalement, je remercie bien évidemment ma famille pour leur soutien infaillible et leurs encouragements depuis le début de mes études (il y a des années !). Peu importe les décisions que j'ai prises, vous avez toujours été présents sans juger mes choix. Et malgré la distance, je sais que j'ai toujours pu compter sur vous. Et non, maman, je ne dissèque pas de souris ! Cricri, ma petite sœur chérie. Je ne sais pas ce que j'aurais fait sans toi ces derniers mois. Merci d'être là, tout le temps.

Sommaire

Liste des figures	13
Liste des tableaux	15
Liste des abréviations	17
INTRODUCTION GÉNÉRALE	21
PARTIE 1 : ÉTUDE BIBLIOGRAPHIQUE	27
Chapitre I : Une introduction à la paléogénétique	29
I. Introduction générale : de la paléontologie à la paléogénétique	29
II. La paléogénétique	31
III. Bilan du chapitre I	47
Chapitre II : La parasitologie	49
I. Généralités	49
II. Les parasites étudiés	53
III. Bilan du chapitre II	77
Chapitre III : La paléoparasitologie	79
I. Historique	79
II. La paléoparasitologie morphologique	83
III. La paléogénétique pour l'étude des helminthes gastro-intestinaux	88
IV. Bilan du chapitre III	99
PARTIE 2 : MATÉRIEL ET MÉTHODE	101
Chapitre I : Les sites archéologiques	105
I. Les sites préhistoriques	105
II. Les sites antiques en France	115
III. Les sites médiévaux français	119
IV. Les sites modernes	126
Chapitre II : Introduction méthodologique	133
I. L'ADN et le sédiment	133
II. Le génotypage des helminthes gastro-intestinaux anciens par une approche à haut-débit	137

Chapitre III : A new high-throughput approach to genotype ancient human	
gastrointestinal parasites	149
I. La réplication des expériences sur de l'ADN ancien sédimentaire	149
II. A new high-throughput approach to genotype ancient human gastrointestinal parasites	153
III. La publication scientifique	157
PARTIE 3 : RÉSULTATS ET DISCUSSION	189
Chapitre I : Les sites préhistoriques	193
I. La Draga, Espagne	193
II. Arbon-Bleiche III, Suisse	195
III. Passel, France	196
IV. Menneville, France	197
V. Les mines de sel de Chehr Abad, Iran	198
Chapitre II : Les sites antiques	201
I. L'Amphithéâtre de Metz : les résultats de la paléoparasitologie moléculaire	201
II. Jeu de Paume, Beauvais : les résultats de la paléoparasitologie moléculaire	201
III. Mas de Vignoles, Nîmes : les résultats de la paléoparasitologie moléculaire	201
IV. Discussion et conclusions sur les sites antiques	202
Chapitre III : Les sites médiévaux français	205
I. Rue Saint-Laurent, Beauvais : les résultats de la paléoparasitologie moléculaire	205
II. Place des Véens, Auxerre : les résultats de la paléoparasitologie moléculaire	205
III. Les Turqueries, Calais: les résultats de la paléoparasitologie moléculaire	205
IV. Discussion et conclusions sur ces trois sites médiévaux	206
V. Teteghem, France	208
Chapitre IV : Les sites modernes	211
I. Le Fort Saint-Sébastien, France	211
II. Baillif, Guadeloupe	213
III. Les sites alsaciens de la Première Guerre Mondiale	214
Tableau récapitulatif des résultats	215
Conclusions générales	219
CONCLUSIONS ET PERSPECTIVES	221

VALORISATION SCIENTIFIQUE	227
I. Autres publications	229
II. Communications orales	229
III. Communications affichées	230
RÉFÉRENCES BIBLIOGRAPHIQUES	231
Résumé	251
Abstract	253

Liste des figures

FIGURE 1 : DOMMAGES POUVANT AFFECTER UNE MOLECULE D'ADN	32
FIGURE 2 : PROFIL MAPDAMAGE	35
FIGURE 3 : REPRESENTATION DU LABORATOIRE DE HAUT-CONFINEMENT DE L'INSTITUT JACQUES MONOD.....	44
FIGURE 4 : LE CYCLE DE VIE DES TÆNIA SP.	55
FIGURE 5 : LE CYCLE DE VIE D'ÆCHINOCOCCUS SP.	59
FIGURE 6 : LE CYCLE DE VIE DU DIPHYLLOBOTHRUM SP.	63
FIGURE 7 : LE CYCLE DE VIE DE FASCIOLA SP.	65
FIGURE 8 : LE CYCLE DE VIE DE DICROCOELIUM DENDRITICUM.	67
FIGURE 9 : LE CYCLE DE VIE D'ASCARIS SP.	69
FIGURE 10 : TROIS MODELES DE TRANSMISSION D'ASCARIS DANS LES POPULATIONS PORCINES ET HUMAINES.	70
FIGURE 11 : LE CYCLE DE VIE D'ENTEROBIUS VERMICULARIS.	73
FIGURE 12 : LE CYCLE DE VIE DE TRICHURIS TRICHIURA.....	74
FIGURE 13 : CLE D'IDENTIFICATION D'ÆUFS D'HELMINTHES HUMAINS.....	81
FIGURE 14 : UN ÆUF D'ASCARIS LUMBRICOIDES.....	83
FIGURE 15 : UN ÆUF DE TRICHURIS TRICHIURA	84
FIGURE 16 : DES ÆUFS D'ENTEROBIUS VERMICULARIS	84
FIGURE 17 : DES EMBRYOPHORES DE TÆNIADES.....	85
FIGURE 18 : UN ÆUF DE DIPHYLLOBOTHRUM SP.	86
FIGURE 19 : DES ÆUFS DE FASCIOLA HEPATICA	86
FIGURE 20 : UN ÆUF DE DICROCOELIUM SP.....	87
FIGURE 21 : EVOLUTION DES DONNEES DISPONIBLES DANS GENBANK	90
FIGURE 22 : RECONSTITUTION DU SITE DE LA DRAGA	105
FIGURE 23 : LE LAC DE CONSTANCE (BODENSEE).....	108
FIGURE 24 : CARTE DES SITES ARCHEOLOGIQUES DE LA VALLEE DE L' AISNE	111
FIGURE 25 : UNE SEPULTURE HUMAINE DE MENNEVILLE.....	112
FIGURE 26 : LA LOCALISATION DU SITE DE CHEHR ABAD	113
FIGURE 27 : LA MOMIE NUMERO 5 DU SITE DE CHEHR ABAD	114
FIGURE 28 : LA MAQUETTE DE LA CITE DIVODURUM	116
FIGURE 29 : LES SITES DE JEU DE PAUME.....	117
FIGURE 30 : LE SITE D'AUXERRE	121
FIGURE 31 : LA FOSSE CONTENANT LES SQUELETTES DE MOUTONS, TETEGHEM	124
FIGURE 32 : UNE AQUARELLE DU FORT DE SAINT-SEBASTIEN	127
FIGURE 33 : INHUMATION D'UN SQUELETTE DU CIMETIERE DE BAILLIF.....	129
FIGURE 34 : LE SQUELETTE D'UN SOLDAT ALLEMAND RETROUVE A CARSPACH	130
FIGURE 35 : REPRESENTATION DES STRUCTURES SECONDAIRES QUE PEUVENT FORMER LES AMORCES.....	139
FIGURE 36 : ANALYSE PAR QPCR DE LA FORMATION DE STRUCTURES SECONDAIRES PAR LES AMORCES	140
FIGURE 37 : DETAILS SUR LES PUCES ION TORRENT PGM ACTUELLEMENT DISPONIBLES.....	142

FIGURE 38 : LES PREMIERES ETAPES DE CREATION D'UNE BANQUE DE SEQUENÇAGE ION TORRENT.....	144
FIGURE 39 : LE PRINCIPE DE LA PCR EN EMULSION ION TORRENT.....	146
FIGURE 40 : L'HETEROGENEITE DU NOMBRE SEQUENCES POUR DES ECHANTILLONS DE TYPE SEDIMENTAIRES	150

Liste des tableaux

TABLEAU 1 : DEGRADATION DES AMPLICONS DE D. MELANOGASTER APRES CONTAMINATION ARTIFICIELLE.....	42
TABLEAU 2 : RESULTATS COMPARATIFS DE MICROSCOPIE ET DE PALEOGENETIQUE POUR LE SITE DE LA DRAGA	193
TABLEAU 3 : RESULTATS COMPARATIFS DE MICROSCOPIE ET DE PALEOGENETIQUE POUR LE SITE DE MENNEVILLE	197
TABLEAU 4 : RESULTATS COMPARATIFS DE MICROSCOPIE ET DE PALEOGENETIQUE DU SITE DE CHEHR ABAD.	199
TABLEAU 5 : RESULTATS COMPARATIFS DE MICROSCOPIE ET PALEOGENETIQUE POUR LES SITES ANTIQUES.	202
TABLEAU 6 : RESULTATS COMPARATIFS DE MICROSCOPIE ET DE PALEOGENETIQUE POUR LES SITES MEDIEVAUX.....	206
TABLEAU 7 : RESULTATS COMPARATIFS DE MICROSCOPIE ET PALEOGENETIQUE POUR LE SITE DE TETEGHEM.....	208
TABLEAU 8 : RESULTATS COMPARATIFS DE MICROSCOPIE ET PALEOGENETIQUE POUR LE SITE DE ST-GERMAIN-EN-LAYE.	211
TABLEAU 9 : RECAPITULATIF DES RESULTATS MICROSCOPIQUES ET GENETIQUES DES SITES ANALYSES.	215

Liste des abréviations

A : adénine

AD : *anno domini* (Après Jésus-Christ ou après le début de l'ère commune)

ADN : acide désoxyribonucléique

ADNa : acide désoxyribonucléique ancien

ARNr : acide ribonucléique ribosomique

Atp : adénosine triphosphate synthase

BP : *before present* (calculé à partir de 1950)

BC : *before christ* (Avant Jésus-Christ ou avant l'ère commune)

C : cytosine

Cox : cytochrome oxydase

Cytb : cytochrome B

EDTA : *ethylenediaminetetracetic acid*

G : guanine

Inrap : Institut national de recherches archéologiques préventives

ITS : *internal transcribed spacer*

Mb : méga-bases

NAD : nicotinamide adénine dinucléotide déshydrogénase

Pb : paires de bases

PCR : *polymerase chain reaction*

qPCR : *quantitative polymerase chain reaction*

SNP : *single-nucleotide polymorphism*

T : thymine

It's a matter of prescience
No, not the science fiction kind
It's all about ignorance, and greed, and miracles for the blind
The media parading, disjointed politics
Founded on petrochemical plunder and we're its hostages
We can't do nothing and think someone else will make it right
Kyoto now

Gurewitz et Graffin

INTRODUCTION GÉNÉRALE

Introduction générale

En touchant plusieurs milliards de personnes sur terre, les parasitoses gastro-intestinales dues à des helminthes sont une préoccupation majeure identifiée comme telle par l'Organisation Mondiale de la Santé. Ces infections regroupent des parasitoses liées aux contaminations fécales (*Ascaris* sp, *Trichuris trichiura*), à l'ingestion de viande ou poisson crus (*Taenia* sp, *Diphyllobothrium* sp.) ou à des contaminations accidentelles, après consommation de végétaux de bords de rivage (*Fasciola hepatica* par exemple). Dans d'autres cas, comme pour le parasite *Enterobius vermicularis*, la vie en collectivité permet une transmission facilitée. Par exemple, dans les crèches, le partage des jouets et leur portage à la bouche en font un lieu privilégié pour sa transmission. Alors que l'état sanitaire global, les niveaux de santé et d'hygiène se sont améliorés au cours des siècles derniers, du moins dans les pays industrialisés, 350 000 personnes sont sévèrement atteintes par des vers du genre *Ascaris* sp., et 60 000 décès par an seraient directement associés à ce parasite (CDC, 2015).

La caractérisation de l'évolution préhistorique et historique des relations entre l'Homme et les parasites au cours du développement des sociétés humaines ouvre de nouvelles perspectives sur ces problématiques de santé publique. La paléoparasitologie est l'étude des restes parasitaires préservés dans des échantillons archéologiques. Cette discipline permet de mieux comprendre l'état de santé des populations passées et l'impact des facteurs environnementaux et anthropologiques qui ont facilité la transmission des parasites en s'interrogeant sur leur occurrence et sur la fréquence de chacune des espèces au cours du temps. En effet, les œufs libérés dans le milieu par la forme adulte des helminthes, après leur passage dans le tractus digestif, peuvent être préservés au cours du temps. Leur coque, en partie composée de chitine, les rend résistant aux processus de dégradation chimique et physique. Ainsi, des œufs d'helminthes ont pu être retrouvés dans des coprolithes, dans des latrines, sur des organes de momies, dans le sédiment prélevé sous des squelettes, dans des couches d'habitation... Les paléoparasitologues procèdent à l'identification des taxons présents dans les échantillons archéologiques par une analyse microscopique des caractéristiques morphologiques et micrométriques des œufs. Mais le diagnostic au niveau de l'espèce n'est possible que pour une minorité d'helminthes car la plupart des œufs au sein d'une même famille ou d'un même genre présentent des caractéristiques identiques.

Or la connaissance plus précise des espèces apporte des informations importantes sur leur répartition et leur transmission. Il est par ailleurs essentiel de pouvoir distinguer les parasites infectant les êtres humains de ceux infectant les animaux. Par exemple, des œufs de Tæniidae peuvent aussi bien être émis par un renard infecté par *Echinococcus* sp. que par un humain infecté par un *Tania* sp. Mais même dans le cas d'une origine anthropique indiscutable (des œufs préservés dans un coprolithe retrouvé sur une momie humaine par exemple), il ne sera pas possible de discriminer entre *T. solium*, un ver solitaire qui se contracte après consommation de porc cru, et *T. saginata* associé à l'ingestion de bœuf cru.

En 2001, la paléoparasitologie s'est orientée vers la voie de la biologie moléculaire lorsque les premières séquences anciennes d'ADN ont été obtenues pour le parasite *Ascaris* sp. Après avoir isolé une centaine d'œufs, deux courts fragments de gènes mitochondriaux ont été séquencés à partir d'un coprolithe médiéval belge. Au cours des 14 années suivantes, un peu plus d'une dizaine de travaux ont été publiés en paléoparasitologie moléculaire. La plupart de ces travaux proposaient une approche commune : après observation microscopique de un ou plusieurs taxons de parasites, ceux-ci sont ciblés par PCR, puis séquencés par la technique classique de Sanger. Ces travaux ont bénéficié de l'enrichissement des banques de données génétiques publiques en séquences d'ADN modernes de parasites, bien que les données disponibles représentent encore mal la diversité des espèces et des origines géographiques.

Ma thèse s'inscrit dans le cadre d'un projet collaboratif entre le Laboratoire de Chrono-environnement, hébergeant l'unique équipe de paléoparasitologie française, dirigée par Matthieu Le Bailly, et l'équipe « Épigénome et Paléogénome », sous la direction des paléogénéticiens Eva-Maria Geigl et Thierry Grange de l'Institut Jacques Monod de Paris. Grâce à l'expertise en biologie moléculaire et paléogénétique de ces derniers, nous y avons développé une approche de génotypage qui permet la caractérisation d'une grande diversité des parasites à partir d'échantillons environnementaux complexes, combinant la sensibilité et la spécificité de la PCR au séquençage à haut-débit. Nous avons appliqué cette approche à la paléoparasitologie et nous avons obtenu des résultats pour des échantillons archéologiques allant du Néolithique à la Première Guerre Mondiale. Nous avons pu déterminer au niveau de l'espèce les helminthes qui avaient été identifiés par microscopie au niveau du genre ou de la famille mais aussi diagnostiquer certains taxons dont les œufs sont facilement dégradés, tel *Enterobius vermicularis*. La méthode, que nous avons développée et adaptée pour les parasites, a été testée systématiquement de manière à explorer ses limites et identifier les conditions à respecter pour assurer la fiabilité maximale de la détection génétique des helminthes anciens.

Nous avons exposé dans un article, soumis à *Plos One* (*A new high-throughput approach to genotype ancient human gastrointestinal parasites* (Côté *et al.*, 2016)) les performances de la méthode et décrit ses conditions d'utilisation optimale.

Le manuscrit de thèse introduira dans une première partie d'étude bibliographique, les domaines de la paléogénétique, de la parasitologie et de la paléoparasitologie. Dans une seconde partie de matériel et méthode, je présenterai les sites archéologiques étudiés et la méthodologie que nous avons mise au point pour l'étude d'helminthes gastro-intestinaux anciens. Cette section sera complétée par la publication scientifique « *A new high-throughput approach to genotype ancient human gastrointestinal parasites* ». Une troisième partie regroupera les résultats paléogénétiques obtenus, qui seront ensuite discutés. Je conclurai par un chapitre sur les perspectives associées à la paléoparasitologie moléculaire.

PARTIE 1 : ÉTUDE
BIBLIOGRAPHIQUE

Chapitre I : Une introduction à la paléogénétique

I. Introduction générale : de la paléontologie à la paléogénétique

Les traces de la vie ancienne sont depuis longtemps une source de fascination pour les Hommes. Entre l'Antiquité, où les fossiles ont été à l'origine du mythe des dragons, et le 19^{ème} siècle, avec la formalisation des idées sur la modification des espèces par Lamarck et Darwin, ces restes furent le sujet de spéculations nombreuses. Les fossiles¹ furent les premiers objets macroscopiques d'étude de l'évolution, car leurs formes pouvaient rappeler celle d'animaux ou de plantes contemporaines.

Les premiers textes sur la paléontologie proviennent de Thalès et de Xénophane, décrivant des fossiles dont l'origine provenait du lent déplacement des mers. Théophraste propose une explication différente, qui persista pendant 2000 ans, la *vis plastica* : les forces plastiques souterraines façonnent les fossiles retrouvés (de Wit, 1992). Au 9^{ème} siècle, en Chine, Shen Kuo observa des fossiles marins dans des strates montagneuses, et compris avec raison que ces montagnes avaient été par le passé au niveau de la mer. Il associa également les différences dans la composition de la biodiversité végétale passée et actuelle à des changements climatiques. Plus tard, en Europe, le naturaliste Bernard Palissy fut condamné à être pendu et brûlé après avoir proposé en 1576 que la disparition des bivalves, sans homologues actuels, était liée à une sur-pêche (Buffetaut, 1998).

La paléontologie, au sens scientifique du terme, est créée par George Cuvier, spécialiste de l'anatomie comparée, à la fin du 18^{ème} siècle. Bien malgré lui, alors qu'il développe la théorie du catastrophisme, il contribuera à l'émergence des idées sur la théorie de l'évolution. Selon lui, les espèces disparaissent lors de catastrophes locales ou générales, et de nouvelles espèces apparaissent *ex nihilo*. Au début du 19^{ème} siècle, Jean-Baptiste Lamarck, considéré comme le fondateur de la biologie, ou du moins, de la propagation de l'utilisation du terme, est titulaire d'une chaire du Muséum d'histoire naturelle de Paris.

¹ Le terme fossile est utilisé ici dans son sens le plus large, c'est-à-dire un reste (os, coquille, dents) ou un moulage dans une roche sédimentaire

Lamarck est un naturaliste français qui travaille de manière intensive sur la classification des invertébrés et il est à l'origine du concept de transformisme et de la thèse de la transmission des caractères acquis (Lamarck, 1815-1822). Henri-Marie de Blainville, médecin et zoologue, succède à Lamarck en 1830 à la chaire du Muséum et occupe deux ans plus tard la chaire d'anatomie comparée, laissée vacante par le décès de Cuvier. Il se consacra principalement à l'Histoire naturelle des mollusques et à la classification des mammifères. En 1822, en collaboration avec son éditeur, il crée le terme « paléontologie » (De Blainville, 1822).

Les fossiles, qu'ils soient sous forme de restes tels des ossements, des dents, des graines de végétaux, ou des moulages dans une roche sédimentaire, furent donc les premiers matériaux analysés par ces biologistes. Au cours des 19 et 20^{èmes} siècles, les restes fossiles étudiés sont de plus en plus nombreux et variés. La paléobotanique ne se restreint plus à l'analyse des parties aériennes ou des fruits, mais tente de reconstruire les paléoenvironnements et les paléoclimats en mettant en évidence des pollens anciens (palynologie), des graines ou semences anciennes (carpologie) (Renault-Miskovsky, 2002). Les archéozoologues deviennent de plus en plus spécialisés : archéomalacologie, archéoictyologie, archéoentomologie et la paléoparasitologie, qui sont respectivement l'étude des mollusques, des poissons, des insectes et des parasites. Le recours à des instruments d'optique donnant accès au monde microscopique permet à des disciplines telles la paléoparasitologie de se développer (voir le chapitre III pour les détails sur les débuts de la paléoparasitologie)

Ce sont aussi les avancées technologiques à la fin du 20^{ème} siècle dans le domaine de la biologie moléculaire qui ont permis l'essor d'une nouvelle discipline, la paléogénétique. Celle-ci s'imposa rapidement comme étant complémentaire aux études d'archéozoologie, d'archéobotanique, à l'anthropologie et à la paléopathologie. Dans la suite de ce chapitre seront exposées les caractéristiques des molécules d'ADNa (ADN ancien) et leur préservation au cours du temps. Nous exposerons ensuite les facteurs pouvant introduire des difficultés particulières lorsque nous travaillons sur du matériel ancien : la présence d'inhibiteurs de réaction enzymatique et la présence de molécules contaminantes. Nous terminerons en exposant les précautions qui devraient être mise en place lors de toutes études en paléogénétique.

II. La paléogénétique

La paléogénétique a pour objectif l'étude des molécules d'ADN préservées dans les restes d'organismes. Cette discipline a pu se développer avec la mise au point des techniques de biologie moléculaire d'extraction et de purification des acides nucléiques, puis de la PCR (*Polymerase Chain Reaction*) (Mullis *et al.*, 1986).

En 1984, la première séquence d'ADNa est obtenue à partir d'un échantillon de muscle prélevé sur un équidé éteint depuis 1883 et conservé dans un musée (Higuchi *et al.*, 1984). Rapidement, les travaux dans le domaine permirent de constater que les molécules d'ADN analysées possèdent des caractéristiques biochimiques les distinguant de leurs équivalents modernes : elles sont courtes, présentent des sites abasiques, des liaisons intermoléculaires atypiques et des modifications des pyrimidines, conséquences de lésions oxydatives (Paabo, 1989).

A. La dégradation post-mortem de l'ADN

Dès la mort de l'organisme, divers processus se mettent en route rapidement, entraînant la dégradation des molécules d'ADN en courts fragments et en modifiant les nucléotides. Lors du séquençage, ces bases modifiées pourront être, par erreur, considérées comme des mutations, et dans d'autres cas, elles empêcheront la progression des enzymes sur le brin d'ADN qui sert de matrice.

Ces processus peuvent être de types endogènes, c'est-à-dire que les événements *post-mortem* dans l'organisme contribuent à fragmenter et à endommager les biomolécules. Ils peuvent être aussi exogènes, par l'action de bactéries saprophytes et/ou liés aux conditions du milieu de préservation. Nous décrirons dans un premier temps l'impact des enzymes sur les molécules d'ADN, puis les différentes dégradations chimiques qu'elles peuvent subir au cours du temps. Celles-ci sont résumées sur la Figure 1.

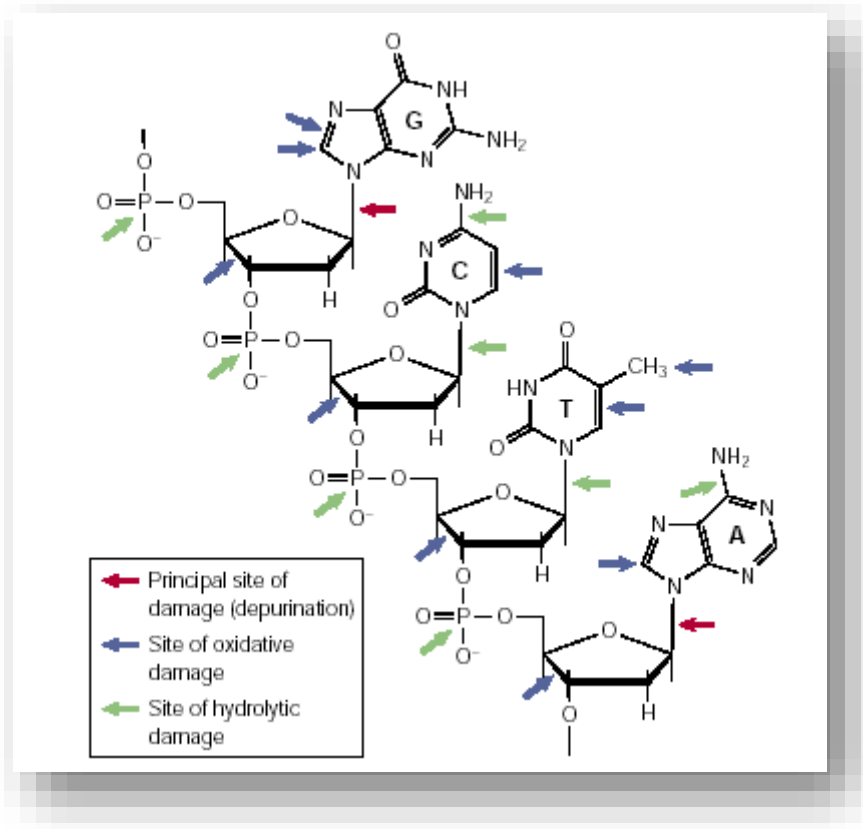


Figure 1 : dommages pouvant affecter une molécule d'ADN après la mort d'un organisme (d'après (Lindahl, 1993)).

1) Les dégradations enzymatiques

Habituellement contenues les lysosomes, des DNases de type II sont libérées au moment du décès de l'organisme. Leur effet sur les acides nucléiques est facilité car ces derniers seront rendus plus accessibles par l'action des protéases qui les libèrent de leurs interactions avec des protéines protectrices (Friedberg *et al.*, 1995). Des nucléases d'origine environnementale provenant des divers organismes saprophytes peuvent aussi contribuer à la fragmentation des biomolécules.

2) Les dégradations chimiques

i. L'oxydation

Les guanines sont les principales cibles de cette modification chimique entraînant leur transformation en une 8-hydroxyguanine. Cette base modifiée sera lue comme une thymine, la base complémentaire insérée sera une adénine. Ce phénomène entraînera la conclusion fautive d'une mutation de type transversion. Les cytosines peuvent être oxydées en 5-hydroxy-5-méthylhydantoin ou 5-hydroxyhydantoin, molécules ne pouvant pas être lues par l'enzyme *Taq* polymérase, bloquant ainsi la réaction d'amplification. Elles peuvent aussi s'oxyder en 5-hydroxycytosine, provoquant des transitions C → T (Hoss *et al.*, 1996).

La structure désoxyribose-phosphate peut aussi être la cible de l'oxydation, ce qui fragmentera la molécule d'ADN (Paabo, 1989).

ii. Les pontages

Parmi les différentes molécules organiques pouvant s'associer entre elles, nous pouvons mentionner la réaction de condensation entre un groupement amine des acides nucléiques ou de protéines et le groupement carbonyle d'un sucre, donnant une molécule appelée « produit de Maillard ». Bien que l'ADN ainsi séquestré ne soit pas disponible pour les réactions enzymatiques, ce produit est réversible après traitement par le bromure de N-phénylthiazolium (PTB, N-phenacylthiazolium bromide) (Vasan *et al.*, 1996).

iii. L'hydrolyse

Cette réaction chimique en présence d'eau est favorisée par un pH acide et de faibles forces ioniques. Elle peut intervenir à différents endroits de la molécule d'ADN et entraîner des dommages, qui peuvent avoir divers types de conséquences sur les analyses génétiques.

a) Génération de sites abasiques

L'hydrolyse peut toucher les liaisons N-glycosidiques entre les bases azotées et le désoxyribose, entraînant la formation d'un site abasique (dépuration ou dépyrimidination). En fragilisant le squelette de la molécule, le brin d'ADNa peut subir une cassure. Au moment d'une réaction de PCR, face à ce type de site, la polymérase aura tendance à ajouter aléatoirement une base, avec une préférence pour une adénine.

b) Désamination

Les cytosines sont les bases le plus affectées par la perte du groupement amine (Lindahl, 1996), et ce dommage est le plus fréquemment retrouvé dans les molécules d'ADNa (Hofreiter *et al.*, 2001). La cytosine désaminée correspondant à une molécule d'uracile, une adénine sera incorporée sur le brin néosynthétisé, résultant en des mutations $C \rightarrow T$ ou $G \rightarrow A$. La désamination d'une adénine donne quant à elle une hypoxanthine, provoquant des mutations $A \rightarrow G$, alors que la désamination d'une guanine en xanthine, sera lu comme une cytosine, sans conséquence sur le nucléotide intégré.

Le séquençage à haut-débit, sans étape préalable d'amplification par PCR, a montré que les cytosines désaminées se retrouvent en proportions élevées sur les deux extrémités de la molécule d'ADNa (Green *et al.*, 2009). La *Figure 2* montre la fréquence (axe des y) des quatre nucléotides incorporés en 5' (partie gauche de la figure) et en 3' (partie de droite) sur une banque d'ADNa génomique. Nous observons effectivement une augmentation de la fréquence des T en 5' (en rouge) et des A en 3', comparativement aux autres nucléotides, indiquant la présence de nucléotides désaminés aux extrémités.

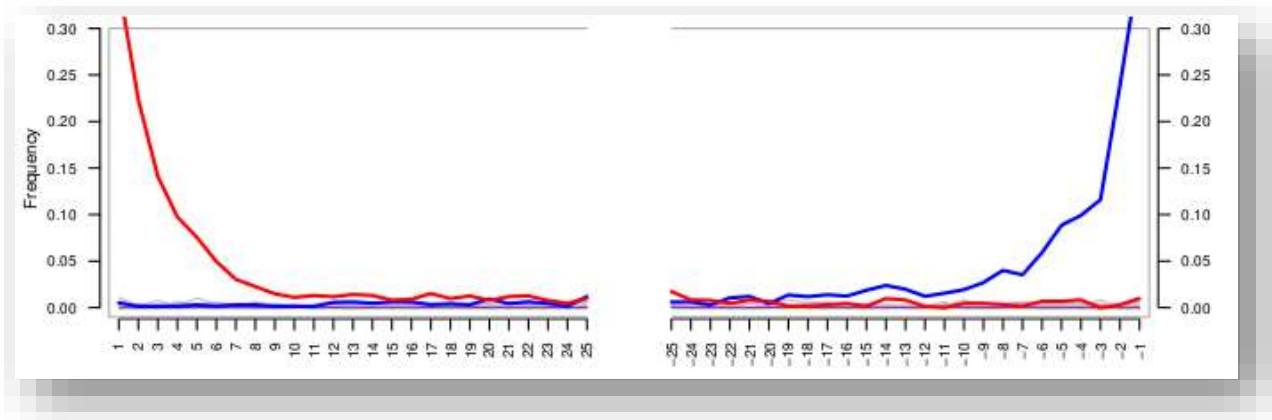


Figure 2 : profil MapDamage (Ginolhac *et al.*, 2011) d'une banque d'ADNa. Les analyses ont été effectuées par D. Massilani à partir des données de séquençage shotgun Illumina d'un extrait de *Bos sp* d'environ 8700 BP. L'augmentation de la fréquence des T en 5' (en rouge) et des A en 3' (en bleu) indique la présence de cytosines désaminées aux extrémités de la molécule.

B. Les facteurs environnementaux influençant la préservation de l'ADN

La préservation de l'ADN est un événement gouverné par plusieurs facteurs physico-chimiques. Ces facteurs dépendent de l'environnement dans lequel l'organisme est décédé ainsi que du déroulement du processus de dégradation/fossilisation.

Parmi les processus de dégradations décrits précédemment, les enzymes bactériennes affectent la préservation de l'ADN. Les facteurs ayant un impact sur leurs actions pourront donc contribuer à mieux préserver les biomolécules. Une température basse diminuera ainsi l'activité des nucléases et l'efficacité des processus d'autolyse. Si les tissus sont rapidement congelés au moment de la mort, et le reste sur une longue période, les microorganismes, environnementaux ou commensaux, auront plus difficilement accès aux cellules, ce qui contribuera également à une meilleure préservation de l'ADN. Ainsi, les séquences et génomes d'organismes les plus anciens obtenus à ce jour proviennent d'environnements polaire ou de sols de type *permafrost*. Le plus ancien génome obtenu est celui d'un équidé de 560-780 000 retrouvé au Yukon, Canada (Orlando *et al.*, 2013). Des travaux menés sur des carottes glaciaires de 400 000 ans du Groenland ont montré que le sud de la région était couverte d'une forêt de conifères peuplée d'insectes (Willerslev *et al.*, 2007). D'autres études ont permis l'analyse, par des approches de métabarcoding, de la végétation et du régime alimentaire d'herbivores arctiques sur les 50 000 dernières années (Willerslev *et al.*, 2014), ou encore de la biodiversité des mycètes de 16-32 000 ans en Sibérie (Bellemain *et al.*, 2013).

A l'opposé, les déserts chauds et secs, à sols basiques, ont tendance à ralentir les phénomènes d'hydrolyse. Les grottes, qui sont des microenvironnements à la température généralement stable, isolées des radiations ultra-violettes, présentent des bonnes conditions pour la préservation de l'ADN. La grotte de Sima de los Huesos, dans les montagnes de Atapuerca en Espagne, a livré des ossements d'ours des cavernes (*Ursus deningeri*) et d'hominidés datés du Pléistocène moyen (~300 000 ans). Ceux-ci étaient assez bien préservés pour permettre des analyses génétiques (Dabney *et al.*, 2013 ; Meyer *et al.*, 2014).

L'activité des procaryotes aérobies et les mécanismes oxydatifs de dégradation sont ralentis en milieu hypoxique. Les tourbières, tel le site de Windover aux États-Unis d'Amérique (Hauswirth *et al.*, 1994) est un exemple de milieu anoxique ayant permis une bonne préservation de molécules d'ADN.

C. Les inhibiteurs de réactions enzymatiques

Les inhibiteurs des réactions enzymatiques (nous décrivons ici les réactions d'amplification par PCR) sont des substances hétérogènes, co-extraits avec les biomolécules. Ces substances, bien que pouvant être différentes selon l'origine de l'échantillon (sédiments, fèces, sang, urine, plantes...) ont en commun d'agir sur les composants du mélange réactionnel (ADN ou *Taq* polymérase principalement) et d'inhiber l'amplification des molécules cibles. Ce phénomène peut avoir un impact aussi bien sur l'analyse d'échantillons modernes qu'anciens, bien que les conséquences sur ces derniers soient plus importantes, du fait de la faible quantité de matériel de départ. La présence d'inhibiteurs, ou le fait de ne pas interpréter des résultats de PCR ou de séquençage en prenant en considération leur présence potentielle, peut entraîner des conclusions de faux-négatifs.

Ces composés inhibiteurs peuvent être des substances humiques (acides humiques et fluviqes) régulièrement co-extraits avec des échantillons provenant d'environnements sédimentaires ou aquatiques (Tsai et Olson, 1992). Il peut aussi s'agir d'éléments liés aux cultures cellulaires (hémoglobine, polyanéthol sulfonate de sodium) (Akane *et al.*, 1994), de l'urée (Khan *et al.*, 1991), des molécules issues de la purification de tissus végétaux, telles les polysaccharides ou les polyphénols (Demeke et Adams, 1992), de restes de solutions introduites par les expériences d'extraction ou de purification de l'ADN (détergent, EDTA, ions, phénol, éthanol...) (Kreader, 1996).

Plusieurs mécanismes ont été proposés pour expliquer l'effet de ces substances sur l'inhibition de l'amplification de l'ADN. Certaines, comme l'EDTA, peuvent chélater des ions magnésium, nécessaires à l'activité de la *Taq*. La présence de cations manganèse, souvent libérés après la solubilisation de la partie minérale des fossiles, peut empêcher la progression de l'enzyme en se liant à l'ADN. Les cations divalents peuvent aussi avoir un impact sur la fidélité de la polymérisation par la *Taq* polymérase (Eckert et Kunkel, 1990). Dans d'autres cas, l'inhibition se produit car le composé entre en interaction avec l'enzyme. Les polyphénols pourraient, par l'entremise de ses anneaux phénol, former une structure secondaire avec la *Taq* et empêcher son activité (Mayr *et al.*, 2005).

Des méthodes ont été testées par le passé pour tenter d'éliminer les inhibiteurs au moment de la purification des acides nucléiques. Ces approches, généralement coûteuses, n'ont pas été probantes car elles ont entraîné en même temps une perte des molécules biologiques (Kreader, 1996). Les stratégies développées consistent donc plutôt à utiliser des agents stabilisateurs d'enzymes.

La BSA (*Bovine Serum Albumin*) est utilisée pour diminuer les effets de ces inhibiteurs, bien qu'elle ait habituellement peu d'effet sur les molécules co-extraites avec les plantes (Wei *et al.*, 2008). Une concentration de 0.05-0.1 µg/µl de BSA serait suffisante pour lever une inhibition, alors qu'une quantité plus importante (3-10 µg/µl) empêche l'amplification par PCR. Les lysines de cette protéine pourraient se lier aux groupements carbonyles des molécules inhibitrices, ce qui préviendrait l'interaction de ces dernières avec la *Taq* polymérase (Huang *et al.*, 2004). La BSA stabilise aussi l'enzyme en empêchant une dénaturation de surface (Malhotra *et al.*, 1998).

Plusieurs entreprises commercialisant des kits de purification de l'ADN ont mis au point des solutions qui permettraient de se débarrasser de certains inhibiteurs. La société QIAGEN commercialise le QIAmp® DNA Stool Mini Kit qui inclut une étape d'élimination des inhibiteurs par l'utilisation d'une résine InhibitEX, dont la composition n'est pas précisée. Son efficacité a été testée par qPCR sur 16 extraits d'ADNa, dans lesquels une quantité connue d'ADN moderne a été ajoutée (Champlot, 2010). Il a été remarqué que selon l'échantillon et/ou le site archéologique, la résine pouvait soit retarder l'amplification de l'ADN moderne, soit améliorer son amplification, soit n'avoir aucun effet. D'autres kits d'élimination des inhibiteurs ont aussi été testés et se sont avérés inefficaces sur du matériel archéologique.

D. Les contaminations

La contamination d'un échantillon par de l'ADN exogène² est un problème important pour les analyses en paléogénétique. Comme il a été décrit au début de ce chapitre, les molécules d'ADN sont soumises à des processus physico-chimiques qui entraînent leur dégradation en de courts fragments dont les nucléotides sont souvent modifiés. En résultent des difficultés méthodologiques pour les analyses génétiques et des résultats pouvant être faussés par la conclusion erronée de la présence de polymorphismes, due aux processus diagénetiques. À cela s'ajoute l'amplification préférentielle des polymérase pour un substrat moderne, dont le nombre de molécules sera plus élevé que le nombre de molécules anciennes. Cette situation est donc qualitativement et quantitativement en défaveur des biomolécules anciennes, qui dans le cadre d'expériences de PCR, risquent d'être amplifiées tardivement, comparativement à de l'ADN exogène moderne. Il est donc important de bien comprendre les différentes sources de contaminations pour pouvoir mettre en œuvre des mesures permettant de les minimiser.

1) Les contaminations au moment des fouilles archéologiques

La première source de contamination avec de l'ADN moderne exogène est celle associée à la découverte d'un échantillon archéologique : l'ADN des fouilleurs peut se retrouver sur l'échantillon. Ce problème concerne principalement les travaux d'anthropologie, pour lesquels les molécules endogènes peuvent être mélangées avec l'ADN des archéologues.

Mais il existe d'autres contextes où cette contamination pourrait être problématique, par exemple pour la détermination génétique de l'origine de coprolithes. Il est parfois difficile, par exemple, de distinguer un excrément fossilisé de chien de celui d'un humain. L'ADN permettant de les discriminer provient des cellules de l'épithélium intestinal et est distribué en gradient descendant de l'extérieur vers l'intérieur des fèces. La partie externe, la plus riche en ADN, est aussi celle pouvant être manipulée par les fouilleurs.

Des contaminations inter-échantillons peuvent aussi se produire, si deux vestiges sont entreposés ou emballés ensembles, en milieu humide.

² Nous utiliserons le terme « exogène » par opposition à celui d'endogène, qui lui décrit l'ADN du ou des organisme(s) qui sont étudiés dans le cadre d'un projet. Une contamination par de l'ADN exogène pourra aussi bien être de l'ADN environnemental, de l'ADN provenant d'une expérience précédente, présent dans des réactifs ou de l'ADN de l'expérimentateur.

2) *Les contaminations intrinsèques*

Selon le type d'échantillon et le projet, l'ADN endogène peut être mélangé à l'ADN exogène, sans que ce type de contamination soit évitable. Un prélèvement de sédiment pour un projet de paléoparasitologie sera nécessairement contaminé par des virus et une multitude d'organismes, naturellement présents : bactéries, levures, champignons, plantes, pollens ... Dans certains cas, il serait possible de purifier une plus grande proportion d'ADN endogène en adaptant les méthodes d'extraction afin de préserver l'intégrité des parois de bactéries et/ou des champignons.

3) *Les contaminations au laboratoire*

Lorsqu'un échantillon arrive au laboratoire, il subira divers traitements pouvant l'exposer à plusieurs sources de contamination, décrites dans les paragraphes suivants.

i. *Les contaminations inter-échantillons*

Ce genre de contamination peut survenir si les surfaces de travail ne sont pas nettoyées de manière adéquate entre différentes séries d'expériences. Le développement de méthodes d'analyses à haut-débit, permettant de traiter simultanément jusqu'à 96 échantillons en parallèle, peut aussi générer ce type de contamination. L'ouverture de plaques de PCR doit se faire précautionneusement pour éviter de générer des microgouttelettes qui peuvent passer d'un puits à un autre. L'utilisation d'un robot-pipetteur présente l'avantage d'un gain de temps et d'argent, mais il doit être programmé d'une manière à minimiser les possibilités de contaminations croisées.

ii. Les contaminations « carry-over »

Ce type de contamination se caractérise par la dissémination d'amplicons dans le laboratoire pouvant intervenir lors de l'ouverture de tubes de PCR consécutive à des réactions d'amplification. Ces molécules, qui sont le produit de couples d'amorces qui peuvent être utilisés fréquemment, peuvent correspondre à la cible ancienne que l'on cherche à détecter. Ces amplicons, présents en nombre plus élevé que les molécules anciennes, auront plus de chance d'être réamplifiés par la suite.

iii. Les contaminations des réactifs et du matériel de laboratoire

Toutes les solutions, poudres, enzymes, tubes ou embouts de pipettes, produits dans des usines, restent sous le contrôle d'êtres humains, qui peuvent les contaminer, et ceci malgré la stérilisation par l'irradiation des matériaux en plastique. Un exemple, tiré des sciences criminelles, illustre bien le genre de conséquences dramatiques que peut avoir ce type de contamination. Entre 1993 et 2008, l'ADN d'une femme fut retrouvé sur plusieurs scènes de crime, en Allemagne, en Autriche et en France. En 2009, un journal allemand révéla que cet ADN appartenait en fait à une employée de l'entreprise fabriquant des cotons tiges de prélèvement utilisés par les services de police européens (Ferracci-Porri, 2009).

De nombreux agents stabilisateurs sont fabriqués à partir d'animaux comme le bœuf, le porc ou le cheval, telles la BSA ou la HSA (*Horse Serum Albumin*). Des séquences de plusieurs animaux, domestiques ou sauvages, ont été identifiées dans de nombreux réactifs utilisés pour les réactions de PCR (Leonard *et al.*, 2007). Les enzymes *Taq* polymérases contiennent une telle quantité d'ADN bactérien que l'analyse de faible quantité de cible procaryote, même en dehors d'un contexte paléogénétique, peut également être problématique (Bottger, 1990 ; Corless *et al.*, 2000).

Certaines marques de colonnes de purification de l'ADN basées sur le système de membrane de silice ont montré être source de contamination par des parvovirus (Rosseel *et al.*, 2014) et des rongeurs (Erlwein *et al.*, 2011 ; Zheng *et al.*, 2011).

E. La prévention des contaminations

En ayant identifié les sources potentielles de contaminations, il est possible d'en éliminer certaines, ou du moins de les minimiser.

1) Prélèvements des échantillons archéologiques

Cette partie de la prévention des contaminations doit être élaborée en collaboration avec les archéologues. Il est donc nécessaire de mener des campagnes d'information et de sensibilisation auprès des professionnels et étudiants, que ce soit au moment de conférences, d'enseignements universitaires, de formations professionnelles ou lors de visites sur des sites archéologiques.

Ainsi, il est préférable que les échantillons soient prélevés avec des gants, puis entreposés dans un sachet contenant du sédiment du site de prélèvement. Le sachet ne sera pas fermé hermétiquement pour minimiser la croissance des moisissures et bactéries, et devrait être conservé au frais. En stockant rapidement l'échantillon à -20°C, l'ADN endogène sera mieux préservé (Pruvost *et al.*, 2007).

2) La décontamination des surfaces

Des travaux menés par l'équipe « Épigénome et Paléogénome » (Champlot *et al.*, 2010) ont permis de mesurer l'efficacité de plusieurs solutions de décontamination des surfaces. Après avoir contaminé artificiellement des gants en latex avec une quantité connue de courts fragments d'ADN, ils ont été soumis à divers traitements décontaminants avant de mesurer leurs efficacités par qPCR (Tableau 1).

Tableau 1 : pourcentage de dégradation des amplicons de *Drosophila melanogaster* après contamination artificielle de gants en latex. CoPA est une solution de Cuivre-bis-(phénantroline)-sulfate en présence d'eau oxygénée. Les UV correspondent à une irradiation à 256 nm pendant une heure à 10 centimètres de distance (Champlot *et al.*, 2010).

Traitements	Eau de javel	CoPA	UV	DNA/RNA away	Détergent	Douche d'air	Contrôle
% dégradation	99.4	99.5	95.6	62.3	70.9	-0.8	50.3

Les résultats de cette étude ont permis de mettre en place un protocole où les surfaces sont décontaminées à l'eau de javel avant et après utilisation. Le matériel à risque de corrosion (embouts de pipette, porte-échantillons de cryobroyeur ...) est traité par une combinaison de *RNA away* (Life Technologies) associée à une exposition à courte distance aux UV. Ces données montrent aussi l'inefficacité de la décontamination par l'installation d'un système de lampes UV au plafond des pièces des laboratoires à haut-confinement.

3) La décontamination des réactifs

Les réactifs de biologie moléculaire, que ce soient les *Taq* polymérases, les tampons commerciaux, les nucléotides, les oligonucléotides, les mélanges réactionnels pour qPCR, les agents stabilisateurs ... sont souvent contaminés par de l'ADN pouvant interférer avec l'analyse d'une très faible quantité d'ADN dégradé. Champlot *et al.* 2010 ont analysé par qPCR plusieurs milliers de contrôles négatifs en utilisant des couples d'amorces pour *Bos* sp. Jusqu'à 66% des contrôles négatifs étaient contaminés lors des différentes conditions réactionnelles testées (enzymes utilisées, mélange réactionnel d'amplification fabriqué au laboratoire...).

L'effort de décontamination des réactifs dépendra des espèces ciblées par le projet de paléogénétique. Dans le cas des études de paléoparasitologie, les échantillons sont intrinsèquement contaminés par de l'ADN bactérien, viral ... du fait de leur origine sédimentaire tandis que la probabilité que les réactifs soient contaminés avec de l'ADN de parasites est très faible. La décontamination des réactifs présenterait dans ce cas peu d'intérêt. Mais les frontières entre les différents champs de la paléopathologie ne sont pas franches : un kyste retrouvé sur un squelette humain peut être signe d'une infection parasitaire ou bactérienne. Si l'on est intéressé de déterminer l'origine du kyste, il serait nécessaire de décontaminer les réactifs utilisés.

i. La décontamination par exposition aux UV

Les solutions tampons ou ioniques commerciales ainsi que la BSA/HSA, le glycérol et les détergents peuvent être irradiés à 254 nm ; chaque face exposée cinq minutes ; dans des tubes en polypropylène laissant passer les UV, à une distance de 1 cm de la lampe (1.2 joules/cm²) (Champlot *et al.*, 2010). L'enzyme *Taq* polymérase et les oligonucléotides ne doivent pas être irradiés par les rayonnements UV, car cela entraîne une réduction de l'amplification de l'ordre de 10⁵ (Sarkar et Sommer, 1990).

ii. Traitement à la DNase

La *Taq* polymérase et les nucléotides peuvent être traités à la DNase thermo-sensible ciblant les molécules double-brin. L'enzyme décontaminante utilisée est inactivée entre 50-55°C, ce qui est bien en dessous de la température d'activation de la *Taq*. Les oligonucléotides formant des structures secondaires double-brin de type épingle à cheveux, ne peuvent pas être décontaminés par cette méthode car ils seraient aussi dégradés. Il est à noter qu'à partir d'une certaine concentration, la DNase peut inhiber la réaction d'amplification. Les détails de ce protocole de décontamination sont disponibles dans (Champlot *et al.*, 2010).

iii. Traitement au bromure d'éthidium monoazoture

Cette méthode est aussi basée sur l'élimination de l'ADN double-brin en utilisant un agent intercalant, le bromure d'éthidium monoazoture (EMA). Cet agent photo-actif nécessite une exposition à une lumière de longueur d'onde élevée (supérieure à 400 nm) pour activer le groupement azoture qui va se lier de manière covalente aux molécules d'ADN présentes dans la solution réactionnelle. Ces molécules ne pourront donc pas participer à la réaction d'amplification subséquente (Patel *et al.*, 2012 ; Rueckert et Morgan, 2007). Les oligonucléotides formant des structures double-brin ne pourront pas être décontaminés par cette méthode puisqu'ils ne seront plus capables de se lier aux molécules-cibles, mais tous les composants d'une solution réactionnelle de PCR ainsi que les différentes enzymes utilisées pour générer une banque de séquençage peuvent être décontaminés à l'EMA.

F. Le laboratoire de haut-confinement

Bénéficier d'un accès à un laboratoire de haut-confinement, dont l'utilisation est restreinte aux biologistes moléculaires formés, est indispensable et un facteur important pour la génération de résultats fiables. Les objectifs d'une telle structure sont doubles :

- Séparer les expériences menées sur des échantillons modernes (que nous définirons comme postérieurs au 20^{ème} siècle) des expériences effectuées sur des échantillons anciens.
- Concernant les échantillons anciens : minimiser les risques de contaminations inter-échantillons et utiliser des réactifs et matériels décontaminés de manière appropriée.

1) Le laboratoire moderne

Comme pour tout laboratoire de biologie moléculaire, les postes de préparation, extraction et purification des acides nucléiques sont séparés du poste de préparation des réactions enzymatiques (PCR, banque de séquençage). Le poste d'analyse post-PCR est situé dans un local indépendant, où les utilisateurs se doivent de porter des sur-chaussures, blouses jetables, gants, et de respecter la règle « Tout ce qui rentre en salle post-PCR n'en ressort plus ». Séparé physiquement du laboratoire moderne, le laboratoire à haut-confinement de l'Institut Jacques Monod été prévu dès les plans initiaux de construction (Figure 3).



Figure 3 : représentation schématique du laboratoire de haut-confinement de l'Institut Jacques Monod.

2) *La première pièce : le sas*

On accède à cette première pièce uniquement en portant des vêtements lavés à l'eau javellisée et strictement dédiés aux expériences de paléogénétique. La pression dans le sas est supérieure à la pression atmosphérique normale. On y entrepose du matériel et des vestiges archéologiques. Les expérimentateurs y enfilent une première combinaison jetable, une première paire de gants, des chaussures javelissables en plastique, des sur-chaussures, un masque et une charlotte. Tout le matériel entrant destiné aux expériences sera abondamment vaporisé de manière homogène à l'eau de javel. Le sas est séparé de la première pièce du laboratoire par une porte qui ne peut être ouverte qu'au moment où la porte d'entrée est fermée.

3) *La seconde pièce : la préparation des échantillons*

Les échantillons à analyser seront préparés (découpage des os, pesée des sédiments...) dans cette pièce, avant d'être réduit en poudre par cryobroyage. Une surpression par rapport au sas y est maintenue, pour éviter que des contaminants provenant du sas puissent y pénétrer.

4) *La troisième pièce : l'extraction et la purification de l'ADN*

En pénétrant dans cette pièce, une seconde combinaison jetable et une seconde paire de gants doivent être enfilée. Cette pièce, dont la pression est supérieure à la précédente, est munie d'un poste de sécurité microbiologique (PSM). C'est dans cette enceinte que sont fabriqués les tampons nécessaires à l'extraction de l'ADN et pour procéder à la purification des molécules sur des colonnes de silices. Des aliquotes des différents échantillons pourront ensuite être préparés avec l'objectif de les caractériser qualitativement (tests d'inhibition) ou quantitativement (dosage *Qbit*® ou *Nanodrop*®).

5) *La quatrième pièce : la préparation des réactions enzymatiques*

Lorsqu'il est nécessaire d'accéder à cette pièce, la combinaison utilisée dans la troisième pièce doit être retirée, et une tenue dédiée doit être enfilée, les gants et sur-chaussures doivent également être changés. Les mélanges réactionnels pour les amplifications par PCR ou pour les banques de séquençage de type *shot-gun* sont préparés dans cette dernière pièce, dont la pression est la plus élevée.

Mentionnons que les réactifs, qu'ils soient enzymatiques ou chimiques, utilisés dans ce laboratoire confiné proviennent toujours des stocks issus directement du fournisseur. Ils ne sont en aucun cas des aliquotes de solutions déjà utilisées dans le laboratoire moderne. Il est aussi utile de signaler que toutes les expériences sont effectuées dans des enceintes de travail dédiées à une étape expérimentale spécifique, nettoyées à l'eau de javel et irradiées à très courte distance par des rayons UV germicides après manipulation. Les équipements en contact avec les échantillons et les réactifs, comme les tubes de cryobroyeurs et les pipettes, sont décontaminés soit avec l'eau de javel, soit avec du *RNase away* selon leur sensibilité à la corrosion, et irradiés aux UV.

III. Bilan du chapitre I

Au cours des trente dernières années, la paléogénétique a connu un essor rapide et important. Parallèlement à la mise en place de nouvelles technologies de séquençage, présentant des coûts moins élevés, les travaux de recherche se sont multipliés pour tenter de répondre à des questions aussi diverses et variées que la domestication d'animaux tels les chevaux (Ludwig *et al.*, 2009) ou les bœufs (Geigl, 2011), sur la composition de la biodiversité des paléoenvironnements (Bellemain *et al.*, 2013), sur l'analyse des régimes alimentaires d'animaux maintenant disparus (Willerslev *et al.*, 2014 ; Wood *et al.*, 2012) ou sur l'évolution des pathogènes dans les sociétés humaines (*Yersinia pestis* (Bos *et al.*, 2011) ou *Mycobacterium tuberculosis* (Harkins *et al.*, 2015)).

Du fait des caractéristiques intrinsèques de l'ADNa, qui est composé de courtes molécules dégradées, modifiées biochimiquement et noyées dans de l'ADN environnemental, les échantillons archéologiques sont sujets à des contaminations, dès le début des fouilles. Pour minimiser celles introduites en laboratoire, il est hautement préférable que les analyses de paléogénétique soient menées par des équipes spécialisées ayant accès à un laboratoire de haut-confinement. C'est ce que nous avons souhaité mettre en œuvre dans le cadre de cette thèse par cette collaboration entre le laboratoire Chrono-environnement et l'Institut Jacques Monod. De plus, les méthodes de biologie moléculaire doivent être adaptées au matériel ancien, ce qui a nécessité l'expertise de l'équipe de « Paléogénome et Épigénome » au sein de laquelle ont été réalisés tous les travaux de cette thèse. Les membres de cette équipe ont partagé avec générosité au cours de ces trois années, leur expérience et compétences en paléogénétique, ainsi que leur temps, ce qui a permis de rendre fructueux ce projet.

Chapitre II : La parasitologie

I. Généralités

A. *Définitions du parasitisme*

En 1875, le naturaliste Pierre-Joseph Van Beneden décrit trois grandes classes de relations biologiques pouvant exister entre des organismes d'espèces différentes (Van Beneden, 1875) :

- *Les Commensaux : ceux qui s'attachent à un animal, non pour vivre à ses dépens, mais pour profiter des restes à sa table,*
- *Les Mutualistes : ceux qui, vivant exclusivement des excréments naturels des animaux, jouent sur leur peau ou leur muqueuse le même rôle que les chiens de Constantinople jouent dans les rues de cette ville, c'est-à-dire qu'ils exécutent un véritable travail de voirie, ils font en un mot, la toilette de leur hôte,*
- *Les Parasites : ceux qui ont besoin, pour vivre, des humeurs qui entretiennent la propre vie de leur hôte.*

Cette définition du parasitisme n'est pas partagée par Jean-Pierre Mégnin et le comte de Saint-Fargeau, puisque ce dernier donne cette définition : « Le parasite est celui qui vit au dépens d'autrui en mangeant son bien et non sa nourrice même » (Mégnin, 1895). Mégnin propose de séparer les parasites inoffensifs des parasites dangereux.

Jusqu'au milieu du 20^{ème} siècle, le parasitisme est décrit comme une relation de symbiose non-mutuelle entre deux organismes, au sein de laquelle un parasite se nourrit, s'abrite et se reproduit au détriment d'un hôte. Contrairement à la notion de mutualisme, cette relation est bénéfique uniquement pour le parasite, aux dépens de l'organisme parasité. C'est ce deuxième point qui la différencie du commensalisme (interaction biologique non-mutuelle et ne causant pas de dommage à l'hôte) et de la prédation (organisme tuant une proie pour se nourrir ou nourrir sa progéniture).

Actuellement, la notion de parasitisme est plus nuancée, en particulier si on se base sur les définitions médicales du terme. Par exemple, un traité de microbiologie médicale de 1973 décrit un parasite comme :

(...) un organisme qui réside sur ou dans un autre organisme vivant, afin de trouver l'environnement et les substances nutritives dont il a besoin pour croître et se reproduire. Ceci n'implique pas qu'un parasite doive nuire à son hôte. Au contraire, les parasites qui ont le plus de succès atteignent un équilibre avec l'hôte, qui assure la survie, la croissance et la propagation de l'un et l'autre. Ainsi, la majorité des interactions hôte-parasite n'ont pas la maladie pour résultat : l'infection demeure latente ou subclinique (Jawertz et al., 1973).

Cette définition se rapproche beaucoup du commensalisme, et les notions exposées dans un manuel de protozoologie médicale (Wéry, 1995) va plus loin en décrivant plusieurs types de relations hôte-parasite :

- Le parasite commensal, qui vit au dépend d'un organisme sans lui causer de tort. Mais dans certaines circonstances, cette relation pacifique cesse et le parasite devient pathogène (infection opportuniste).
- Le parasite symbiotique, qui vit au dépend de l'hôte mais en l'aidant pour certaines fonctions métaboliques, nutritives ou de défense.
- Le parasite pathogène qui entraîne par sa présence des désordres métaboliques, des lésions organiques ou des réactions immunologiques.

Cette classification rejoint sur certains points celle de Mégnin, mais introduit la notion qu'un changement environnemental, métabolique ou immunitaire peut intervenir et transformer une interaction neutre ou positive en une interaction dommageable pour l'un des deux organismes associés.

B. Les types de parasites

Les parasites peuvent tout d'abord être classés selon leur taille. Ils peuvent être macroscopiques, comme c'est le cas pour les vers ou helminthes, tels *Schistosoma* spp., l'agent causal de la bilharziose ou le *Tania solium*, le ver solitaire intestinal. Ils peuvent aussi être microscopiques, par exemple des protistes tels les Apicomplexa intracellulaires *Toxoplasma gondii* ou les *Plasmodium* ssp., responsables de la malaria chez l'humain.

Les parasites peuvent aussi être classés selon leur localisation chez l'hôte. Certains sont des ectoparasites : ils vivent sur la surface corporelle d'un autre être vivant. Parmi eux, mentionnons plusieurs espèces d'arthropodes hématophages telle les *Ixodida* (tiques) ou la famille des Siphonaptera (puces). Ces ectoparasites peuvent être aussi des plathelminthes microscopiques, comme c'est le cas pour les *Dactylogyrus* qui s'accrochent aux branchies des poissons (Koskivaara *et al.*, 1991) ou de *Longidorus africanus* qui est quant à lui un nématode ectoparasite de plantes (Cohn, 1970). Mais la plupart des parasites connus sont des endoparasites. Ce sont des organismes dont le cycle de vie se déroule à l'intérieur d'un ou plusieurs hôtes. Tous les parasites décrits dans la section B) de ce chapitre sont des endoparasites.

Les parasitoïdes sont des ecto- ou endoparasites dont l'issue finale de la relation avec l'hôte se termine par la mort (ou la stérilisation) de ce dernier. Les braconides (insectes hyménoptères) pondent leurs œufs dans une chenille, qui sera dévorée lors de l'éclosion des larves. Dans cet exemple, des particules pseudo-virales codant des protéines contrôlant le système immunitaire de la chenille sont également injectées en même temps que les œufs (Bezier *et al.*, 2009).

C. Des exemples de contrôle du parasite sur l'hôte

Certains parasites ont besoin d'un seul hôte définitif, sur/dans lequel l'adulte émet des œufs, qui vont éclore, et se développer en un nouvel adulte. Ce type de cycle de vie est appelé monoxène. Dans d'autres cas, le cycle de vie est plus complexe et les différents stades du parasite (pontes des œufs, et développement des différents stades larvaires avant d'arriver à l'adulte) nécessitent un ou plusieurs hôtes intermédiaires. Les cycles de vie sont alors qualifiés de dixéniques ou trixéniques. Le parasite doit pouvoir rencontrer les bons organismes et environnements pour survivre et se reproduire, et dans ce cadre, des mécanismes évolutifs de contrôle des hôtes sont alors mis en place.

Schistocephalus solidus est un cestode dont l'hôte final est un oiseau ichtyophage. Ses œufs sont émis avec les fèces de l'oiseau, qui se disséminent dans l'eau, milieu dans lequel se développera une larve coracidium. Celle-ci sera ingérée par le premier hôte intermédiaire, un copépode, dans lequel la larve de second stade se développera. Ce copépode sera ensuite consommé par une épinoche et la larve plérocercœide va croître dans son abdomen.

Ce poisson sera ensuite consommé par un oiseau. L'épinoche adopte normalement certains comportements pour éviter la prédation, dont la fuite pour se cacher, suivi d'une période d'immobilité. Lorsqu'il est infecté par *S. solidus*, le poisson perd ces comportements, augmentant ainsi les risques d'être capturé par un oiseau, permettant de terminer le cycle de vie du parasite (Barber et Scharsack, 2010). Ce comportement serait associé à une diminution des sécrétions endocrines de norépinephrine au niveau du télencéphale et une augmentation du ratio acide 5-hydroxyindoleacétique : 5-hydroxytryptamine, causées par la présence du cestode (Barber, 2013).

Spinochordodes tellinii est un parasite dont le stade juvénile se déroule dans une sauterelle ou un criquet, le ver devant quitter ensuite cet hôte pour rejoindre un plan d'eau, où il se reproduira et libérera des œufs. Ces insectes étant terrestre, *S. tellinii* influence leur comportement en les poussant à se lancer dans de l'eau, où ils se noieront. Ce comportement surviendra uniquement si un plan d'eau est à proximité. Une étude de protéomique a identifié un facteur sécrété uniquement dans le système nerveux central des insectes infectés et pouvant être impliqué dans la réponse géotactique. Par ailleurs, des protéines du parasite, sécrétées uniquement durant la phase de manipulation du comportement, ont aussi été identifiées. Certaines, de la famille des Wnt (*wingless integration site*), pourraient agir par mimétisme sur le système nerveux central de l'insecte pour le diriger vers un milieu aqueux (Biron *et al.*, 2005).

De nombreux autres exemples sont connus parmi lesquels celui de *Dicrocoelium dendriticum*, manipulant le comportement diurne d'un de ses hôtes intermédiaires, qui sera exposé dans la section suivante.

II. Les parasites étudiés

Dans cette section, les espèces d'helminthes gastro-intestinaux étudiés dans le cadre de cette thèse seront décrites. Leurs principales caractéristiques biologiques, incluant leur mode de reproduction et de nutrition, seront mentionnées, ainsi que les pathologies qu'ils peuvent causer chez les êtres humains. Pour chaque genre, le cycle de vie sera exposé, suivi d'une description des génomes mitochondriaux et, s'ils ont été étudiés, des génomes nucléaires.

A. *Les Plathelminthes*

Cet embranchement d'animaux protostomiens (super-embranchement Platyzoa), est composé de 4 classes : Les Turbellaria, les Monogenea, les Trematoda et les Cestoda. Ce sont des acœlomates secondaires, c'est-à-dire qu'ils ont perdu leur cœlome au cours de l'évolution. Ces vers plats ne possèdent qu'une seule cavité corporelle (parfois absente) et une unique ouverture, servant pour les apports alimentaires et l'élimination des déchets. L'appareil respiratoire est absent, les échanges gazeux s'effectuent par diffusion à travers le tégument. La plupart de ces vers sont hermaphrodites, et se reproduisent par fécondation croisée (*Echinococcus*) ou autofécondation (*Tænia*, *Echinococcus*). Les Turbellaria sont des organismes libres aquatiques alors que les trois autres classes sont des parasites des vertébrés.

Seules les classes des cestodes et des trématodes seront évoquées pour lesquelles des taxons ont été étudiés

1) *Les cestodes*

Les parasites adultes sont composés d'un scolex, qui porte les éléments permettant de s'accrocher à l'hôte, et d'un strobile, générant une succession de proglottis chacun possédant un ensemble complet d'organes reproducteurs mâles et femelles. Le nombre de proglottis varient selon les genres : jusqu'à 2500 chez *Tænia saginata*, trois pour *Echinococcus granulosus*. Le système digestif est absent : les aliments prédigérés par l'hôte sont absorbés par le tégument possédant des microtiches (microvillosités). Ce sont des osmotrophes. Les proglottis postérieurs, les plus âgés, sont matures pour la fonction reproductrice femelle. Ils sont fécondés par les proglottis antérieurs plus jeunes, de fonction mâle.

i. Les Tæniidae

Ces vers plats ont un cycle de vie nécessitant deux mammifères : un hôte intermédiaire, qui est le plus souvent un ongulé domestique ou sauvage (accidentellement l'humain, causant une impasse biologique pour la suite du cycle de vie du parasite), et un hôte final carnivore.

a) Les *Tænia* sp. (Linné, 1758)

Trois espèces de *Tænia*, peuvent infecter l'être humain : (1) *T. saginata*, dont la transmission s'effectue en consommant du bœuf cru ou mal cuit, principalement en Europe de l'Est, Russie, Afrique de l'Est et Amérique latine, (2) *T. solium* pour lequel la contamination s'effectue en consommant du porc cru ou mal cuit, surtout en Amérique latine, Europe de l'Est, Afrique sub-saharienne et en Asie, (3) *T. asiatica*, dont la majorité des cas se retrouve en Corée, Chine, Taiwan, Indonésie, et en Thaïlande, après ingestion de viande ou d'abats de porc.

Un téniasis est souvent asymptomatique, mais *T. saginata* entraîne parfois certains symptômes car le ver est plus long (10 mètres, *versus* trois mètres pour les deux autres parasites). Une contamination par un ténia se traduit par des douleurs abdominales, une perte d'appétit parfois accompagnée d'une perte de poids. Le signe le plus visible est le passage de proglottis (anneaux) par l'anus dans les fèces, dispersant par la même occasion les œufs. Une cysticercose peut se développer lors d'une infection à *T. solium* : l'Homme devient l'hôte intermédiaire en consommant les œufs matures, via des aliments souillés ou par auto-contamination. Des larves de *Tænia* se développent et s'enkystent dans le cerveau, les muscles et éventuellement d'autres tissus. Dans ce cas, le parasite se retrouve dans une impasse biologique et ne peut pas poursuivre son cycle de vie (CDC, 2015).

Le Praziquantel (mode d'action encore peu connu) et le Niclosamide (inhibiteur de la fixation du glucose) sont les traitements proposés en cas d'infection par un cestode de type *Tænia* (Vidal, 2015).

- Le cycle de vie dixénique

L'hôte intermédiaire (porc dans le cas du *T. solium* ou *T. asiatica*, bœuf pour le *T. saginata*) est contaminé par la consommation d'œufs embryonnés ou de proglottis gravides de l'environnement. Les œufs se développent en oncosphères qui éclosent et traversent la paroi intestinale pour aller se loger dans les muscles. À ce stade, l'oncosphère se développe en cysticerques, qui peuvent survivre pendant des années dans le tissu musculaire. L'humain se contamine en ingérant cette forme larvaire, lors de la consommation de viande crue ou mal cuite. En deux mois, la larve se développe en ver adulte dans le petit intestin, et peut y rester pendant des années. L'adulte produit des proglottis, segments d'anneaux qui deviennent gravides, se détachent et sont expulsés avec les fèces (Figure 4). Un proglottis grvide contiennent jusqu'à 100 000 œufs pour *T. saginata*, et jusqu'à 50 000 pour *T. solium* (CDC, 2015).

Certaines espèces de *Tænia* peuvent contaminer d'autres animaux. *T. pisiformis* contamine le chien ou le chat (l'hôte intermédiaire est le lapin ou le lièvre), *T. multiceps* provoque une cénurose (tournis) du mouton.

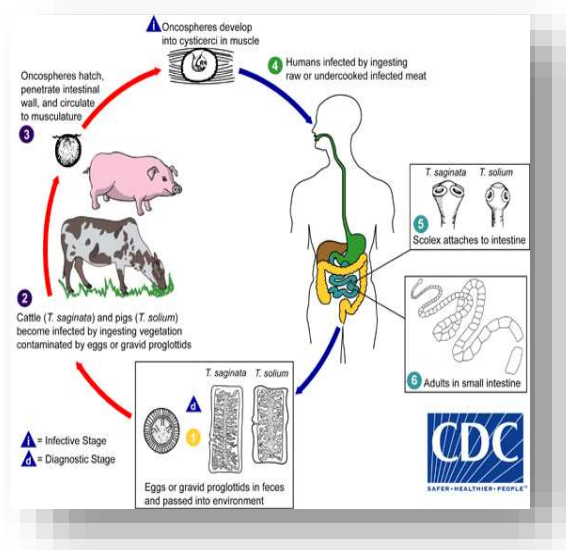


Figure 4 : le cycle de vie des *Tænia* sp.

- Les génomes mitochondriaux

Plusieurs génomes mitochondriaux de *Taenia* ont été séquencés (*T. saginata*, *T. solium*, *T. asiatica*, *T. pisiformis*, *T. laticollis*, *T. regis*, *T. hydatigena*, *T. crassiceps*, *T. twitchelli*, *T. martis*, *T. arctos*, *T. serialis*, *T. madoquae*, *T. multiceps*, *T. ovis*, et *T. crocutae*) (NCBI - janvier 2015) et leur taille est en moyenne 13 500 nucléotides.

Les génomes des *Taenia* humains possèdent 12 gènes codant des protéines, deux ARN ribosomiques (petite et grande sous-unité ribosomique) et 22 ARN de transfert (Jeon *et al.*, 2005 ; Jeon *et al.*, 2007). Le gène *Atp8* (adénosine triphosphate synthase) est absent, comme chez tous les cestodes et la majorité des nématodes séquencés jusqu'à présent. Ce gène est aussi absent chez le chordé *Halocynthia roretzi* et le mollusque *Mytilus edulis* (Nakao *et al.*, 2002), ce qui signifierait que cette délétion s'est produite indépendamment, plusieurs fois au cours de l'évolution de ces animaux. Les génomes des *Taenia* sont caractérisés par un taux élevé d'A+T (70-71%). La taille des protéines est la même chez les trois espèces, sauf pour Nad1 et Nad4 (nicotinamide adenine dinucleotide) de *T. solium* et Cox2 (cytochrome oxydase 2) de *T. asiatica*. La différence de séquences pour les gènes codants des protéines entre *T. saginata* et *T. asiatica* est de 4,6%, et de 11% entre *T. saginata* et *T. solium*.

- Le génome nucléaire de *Taenia solium*

Le génome de 122 Mb (mégabases), est composé de 9,9% de séquences répétées et de 12 490 gènes. Cette composition est similaire à celui d'autres génomes de cestodes séquencés (*Echinococcus granulosus*, *E. multilocularis* et *Hymenolepis microstoma*) et l'ordre des gènes est globalement conservés entre ces quatre espèces (Tsai *et al.*, 2013). L'étude de ces génomes a permis de mettre en évidence certaines caractéristiques importantes pour comprendre les relations entre l'hôte et le parasite, et les mécanismes d'adaptation de ce dernier, liés à son mode de vie :

- Une adaptation métabolique au parasitisme : l'analyse de ces génomes a montré une réduction des capacités métaboliques et une augmentation de la capacité d'absorption des nutriments. Les hydrates de carbone, leur principale source d'énergie, peuvent être obtenus par phosphorylation oxydative, ou par deux voies anaérobiques, la fermentation lactique et la dismutation du malate, processus au cours duquel ce dernier est oxydé en pyruvate par une enzyme malique (Braeckman *et al.*, 2008).

L'absence de voie métabolique des acides gras et du cholestérol oblige les cestodes à utiliser les transporteurs d'acide gras et des enzymes d'élongation des lipides de l'hôte. Plusieurs gènes associés au peroxyosome, organe impliqué dans l'oxydation des acides gras, tels *Pex1*, *Pex3* et *Pex11* (*Peroxisomal Biogenesis factors*) sont absents.

- Une augmentation du nombre de copies des HSP70 (Heat Shock Protein) cytosoliques : ces chaperons moléculaires contribuent à la bonne conformation de protéines pouvant être dénaturées au cours de stress cellulaires. Les HSP ont été hautement conservées au cours de l'évolution (Mayer, 2013). Trente-deux copies ont été retrouvées chez *T. solium*, alors que six sont présentes chez la drosophile et deux chez l'Homme. Les copies non-canoniques présentent un niveau d'expression constitutive peu élevé aux différents stades de développement, contrairement aux gènes canoniques.

- Des cibles thérapeutiques : contrairement à leurs hôtes mammifères, les cestodes, tout comme les trématodes (voir les paragraphes suivants), ne possèdent qu'une seule copie du gène du *Cytochrome P450*. Leur capacité à oxyder des molécules stéroïdes ou xénobiotiques s'en trouve réduite. La connaissance des génomes des cestodes permet de développer de nouvelles cibles thérapeutiques, ou d'adapter des traitements préexistants. Une vingtaine de protéines susceptibles d'être affectées par des composants pharmacologiques, tel un amide hydrolase des acides gras, pouvant être ciblée par le Propofol, ou une protéine impliquée dans la biosynthèse des purines, pouvant être ciblée par le Ribavirin, ont été identifiées.

b) *Echinococcus granulosus* et *E. multilocularis*

Cet autre Tæniadé est responsable de zoonoses et peut causer chez l'être humain soit une échinococcose alvéolaire (*E. multilocularis*), soit une échinococcose cystique, aussi appelée hydatidose (*E. granulosus*).

E. granulosus est surtout retrouvé en Afrique, Europe, Asie, Moyen-Orient et en Amérique. Les hôtes intermédiaires sont les moutons principalement, puis les chèvres, porcs, vaches, chevaux, cervidés et chameaux. Les hôtes définitifs sont des canidés, qui sont contaminés par la consommation d'organes des herbivores. Trois millions de personnes sont actuellement infectées par ce parasite (Zheng *et al.*, 2013).

Une échinococcose hydatique, causée par les larves, peut se présenter sous deux formes :

- Dans le cas d'un kyste hydatique du foie, une longue période de latence, pendant laquelle le kyste se développe, est suivie d'une période d'état, avec des signes de souffrance biliaire. L'évolution se fait par une surinfection de voies biliaires, une rupture du kyste ou bien sa calcification. Une guérison spontanée est possible.
- Un kyste hydatique pulmonaire est souvent secondaire à une infection hépatique. La période de latence est alors plus courte, et l'obstruction d'une bronche se produit rapidement. Lors de la période « kyste malade non-rompu », des crises d'urticaire, une toux accompagnée d'hémoptysie et d'une fébricule sont les symptômes les plus courants. La rupture du kyste s'accompagne de vomique (expectoration subite et abondante de sérosités, de pus et de sang) pouvant contenir des granulations. Sans traitement chirurgical, les alvéoles pulmonaires subiront une fibrosclérose et un affaissement.

Dans les deux situations, le kyste est constitué de quatre structures :

- 1) L'adventice : la structure la plus externe, non-parasitaire, constituée par le parenchyme de l'organe refoulé par le kyste,
- 2) La membrane anhiste : la paroi externe du kyste, composée de plusieurs couches d'une substance proche de la chitine,
- 3) La membrane proligère : une couche cellulaire syncytiale,
- 4) Les vésicules proligères, à l'intérieur : contiennent les protoscolex (larves), qui baignent dans un liquide hydatique. Une rupture du kyste entraîne une libération des larves pouvant donner des localisations secondaires.

E. multilocularis a quant à lui une forte prévalence dans les latitudes nordiques d'Europe, d'Asie et d'Amérique. Les hôtes intermédiaires et définitifs sont respectivement les rongeurs et les renards, les chiens, les coyotes et les chats. Lorsqu'il touche l'humain, il peut causer une échinococcose alvéolaire, le plus souvent à localisation hépatique. Après une période de latence ne présentant pas de signe pathognomonique³, une période d'état s'installe.

³ Pathognomonique : se dit d'un signe clinique ou symptôme qui est caractéristique d'une seule maladie et dont la présence permet d'assurer un diagnostic certain.

Un ictère et une hépatomégalie sont les deux symptômes les plus fréquents. La pathologie peut aussi se manifester par un tableau clinique simulant un cancer métastatique. Sans traitement, la mort est l'issue finale.

- Le cycle de vie dixénique

Un hôte intermédiaire ingère des embryophores, qui éclosent dans le petit intestin et libère un oncosphère (ou embryon hexacanthe), qui va migrer par le système circulatoire sanguin vers des organes (le foie et les poumons en particulier). L'oncosphère s'y loge et se développe en kyste, où il produira des kystes-filles et des larves hydatides. Un phénomène de polyembryonie est observable : un embryon donnera plusieurs larves. L'hôte définitif s'infecte en consommant les organes contaminés de l'hôte intermédiaire. Dans le système digestif du canidé, la larve se fixe à la muqueuse intestinale et se développe en ver adulte de deux à sept millimètres en 32 à 80 jours. Les embryophores sont éliminés dans les fèces et souillent les pâturages (Figure 5).

L'Homme peut accidentellement être hôte intermédiaire en s'infectant par l'ingestion d'embryophores, via le pelage d'animaux domestique ou en consommant des baies ramassées au sol et souillées d'excréments de l'hôte définitif par exemple (CDC, 2015).

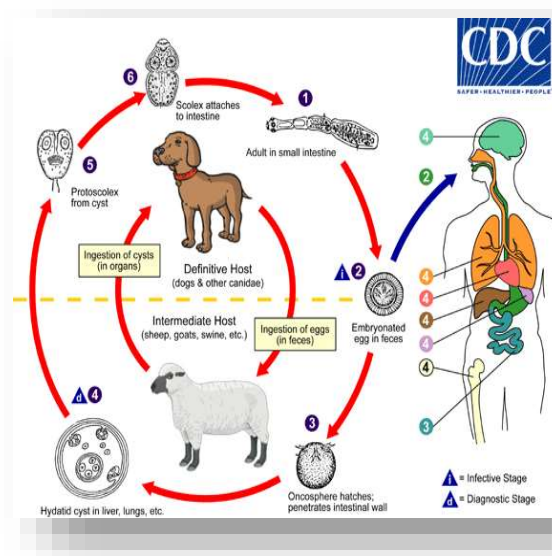


Figure 5 : le cycle de vie d'*Echinococcus* sp.

- Les génomes mitochondriaux

Avant l'obtention de données génétiques, seules cinq espèces d'*Echinococcus* étaient considérées comme taxonomiquement valides : *E. granulosus*, *E. multilocularis*, *E. vogeli*, *E. oligarthrus* et *E. shiquicus*. Les autres représentaient des variants ou souches d'*E. granulosus* (10 génotypes, G1 à G10). Cette conclusion était basée sur leurs différences morphologiques, les hôtes infectés, leurs cycles de vie, les caractéristiques des kystes hydatides et d'autres données biochimiques. Les progrès en génétique depuis le début des années 2000 ont permis de mettre en évidence d'importantes variations intra-spécifiques, poussant à reconsidérer la taxonomie de l'espèce. Ainsi, les 10 génotypes sont maintenant répartis au sein des espèces suivantes : *E. granulosus sensu stricto* (G1, G2, G3), *E. equinus* (G4), *E. ortleppi* (G5), *E. felidis* et *E. canadensis* (G6 à G10) (Thompson et McManus, 2002). Les génomes de ces espèces sont disponibles dans les banques de données du NCBI (janvier 2015).

Les génomes d'*E. multilocularis* (13 738 nucléotides) (Nakao *et al.*, 2002) et d'*E. granulosus* (13 588 nucléotides) (Le *et al.*, 2002) possèdent des gènes identiques à ceux retrouvés dans les génomes des *Taenia* sp.. De même, les gènes codant des protéines sont dans le même ordre avec une proportion de A+T relativement similaire (69-67%).

- Les génomes nucléaires

Les génomes de ces deux espèces d'*Echinococcus* possèdent entre 10 300 et 11 200 gènes répartis sur neuf chromosomes. Selon les études, la taille du génome d'*E. granulosus* se situe entre 114,9 Mb et 151,6 Mb. La première estimation concernant le génome d'*E. multilocularis* est de 115 Mb. Les séquences répétées représentent environ 30% des génomes, et le taux de G+C est de 42% chez *E. granulosus*, allant jusqu'à 49% au niveau des exons (Tsai *et al.*, 2013 ; Zheng *et al.*, 2013). Tout comme pour le génome nucléaire de *T. solium*, nous pouvons mentionner quelques caractéristiques importantes associées au parasitisme :

- Des adaptations métaboliques au parasitisme : en plus des adaptations métaboliques communes aux vers plats (déjà présentées dans la section sur le génome nucléaire de *T. solium*), *E. multilocularis* ne possède pas les enzymes nécessaires pour la biosynthèse de la proline et de la sérine. Une perte de la capacité de synthèse *de novo* des pyrimidines, des purines et de plusieurs acides aminés, a aussi été mise en évidence lors de travaux sur le génome d'*E. granulosus*, suggérant une dépendance pour les nutriments fournis par l'hôte. Ces déficits enzymatiques semblent être compensés par de nombreux gènes codant des protéases (219 gènes) et des transporteurs d'acides aminés ou protéines (68 gènes). Par ailleurs, l'incapacité à produire du cholestérol *de novo* est compensée par la présence de stérol O-acyltransférase et de cholestérol estérase permettant d'utiliser les molécules de l'hôte pour libérer des acides gras.
- La morphologie des stades parasitaires : l'étude de génomique comparative des deux espèces d'*Echinococcus* a permis de retrouver des variations génétiques, cohérentes avec des variations morphologiques. La couche acellulaire laminée externe, retrouvée au moment du stade larvaire métacestode, permet de distinguer ces deux représentants du genre. Ces variations se retrouvent aussi au niveau des gènes de la famille des apomucines, impliqués dans la structure de ce type de membrane, ainsi qu'au niveau des gènes de galactosyltransférases, protéines modifiant les apomucines par des résidus galactoses.
- Une augmentation du nombre de copies de HSP70 cytosoliques : tout comme chez *T. solium*, une augmentation du nombre de copies des *HSP70* est retrouvée chez le genre *Echinococcus* (22 copies). Huit copies subtélomériques au niveau du chromosome huit ont été identifiées dans le génome d'*E. multilocularis*. D'autres protéines associées au stress, telles 13 protéines de stress universelles (USP) présentent en nombre variable chez les plathelminthes, ont aussi été retrouvées dans le génome d'*E. granulosus*. Ce renforcement des mécanismes de défense, absent chez d'autres groupes de parasites tels les nématodes, pourraient être liés à des adaptations au stress causé par l'environnement hostile du tractus digestif : variations de pH entre l'estomac et l'intestin, présence de diverses protéases de l'hôte et de dérivés réactifs de l'oxygène ou ROS (*Reactive Oxygen Species*).

- Des cibles thérapeutiques : l'identification de nouvelles protéines a permis d'établir 1082 cibles thérapeutiques potentielles, dont 150-200 pourraient être ciblées par des composés chimiques déjà existants. Parmi celles-ci, mentionnons un groupe de protéines ayant des fonctions associées à la modification post-traductionnelle de type prénylation, pouvant être ciblé par la Lovastatine, déjà connu pour son effet anti-prolifératif chez les *Schistosoma* et *Trypanosoma* (Urbina *et al.*, 1993).

ii. Les Diphyllbothriidae : les *Diphyllbothrium* sp. (Cobbold, 1858)

Les *Diphyllbothrium* sont un genre de ver plat, aussi appelé Bothriocéphale ou Tænia du poisson, pouvant parasiter les mammifères ichtyophages. Chez l'Homme, l'espèce parasitant le plus fréquemment est *Diphyllbothrium latum*, mais 14 espèces peuvent être impliquées, incluant *D. nihonkaiense*, *D. pacificum* ou *D. dendriticum* (Chen *et al.*, 2014 ; Pastor-Valle *et al.*, 2014 ; Wicht *et al.*, 2010). Les différents membres de ce genre sont difficiles à distinguer au niveau morphologique, mais des données écologiques et génétiques permettent de conclure que ce sont effectivement des espèces différentes (Nakao *et al.*, 2007).

L'incidence de cette parasitose, associée à la consommation de poissons d'eau douce crus ou peu cuits, est importante dans l'hémisphère Nord, au Chili et en Ouganda (Dupouy-Camet et Peduzzi, 2004) et une recrudescence des infections a été récemment rapportée en Europe. En effet, de nombreux cas sont décrits dans des zones non-endémiques, en France, en Suisse et en Finlande par exemple (Wicht *et al.*, 2010). Vingt millions de personnes seraient actuellement porteurs de ce parasite.

La bothriocéphalose est le plus souvent asymptomatique, mais peut s'accompagner d'un inconfort digestif, de diarrhées, de vomissements et d'une perte de poids. Des complications comme une anémie pernicieuse due à une déficience en vitamine B12, ou des occlusions intestinales, sont parfois observées mais restent rares.

- Le cycle de vie trixénique

Un mammifère ichtyophage émet des fèces contaminées avec des œufs, qui atteignent la maturité après un passage de 18 à 20 jours dans de l'eau douce entre 15-25°C. Un coracidium est libéré et ingéré par un copépode du genre *Cyclops* sp. ou *Diaptomus* sp., dans lequel se développera une larve procercoïde. Le crustacé est consommé par un petit poisson, le second hôte intermédiaire. La larve quitte alors le copépode pour migrer dans les muscles du poisson et se transformer en larve pleurocercoïde (sparganum). Même si la larve a atteint son stade infectieux, l'être humain consomme rarement ces petits poissons. La contamination s'effectue plutôt par la consommation d'espèces prédatrices ayant mangé ces poissons, telles la truite, la perche ou le doré, qui deviennent à leur tour porteuses du sparganum. Un adulte mature hermaphrodite se développe ensuite dans l'intestin de l'animal piscivore, produisant un ver pouvant atteindre 20 mètres de longueur. Jusqu'à un million d'œufs immatures peuvent être relâchés dans les fèces via les proglottis (Figure 6) (CDC, 2015).

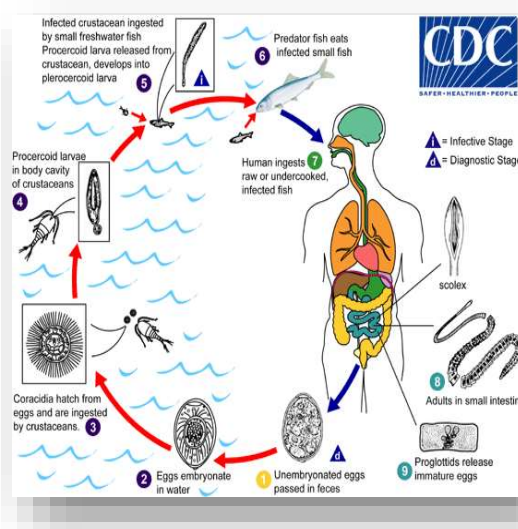


Figure 6 : le cycle de vie du *Diphyllobothrium* sp.

- Les génomes mitochondriaux

Les génomes de *Diphyllobothrium latum* et *D. nihonkaiense* sont composés de 13 720 et 13 747 pb, et possèdent les mêmes caractéristiques que celui des *Taenia* et des *Echinococcus*, en terme de gènes, d'ordre des gènes et de taux d'A+T. La comparaison gène à gène montre des variations de la séquence des gènes codant des protéines allant de 6,9% (*Cox2*) à 15,2% (*Atp6*) confirmant le statut d'espèces distinctes (Nakao *et al.*, 2007).

Les génomes de *D. pacificum* et de *D. dendriticum* ne sont pas encore disponibles, mais plusieurs travaux de génétiques sont en cours pour étudier les relations entre ces deux espèces. L'analyse des séquences des ARNr 18S, et des ITS1 et ITS2 (*Internal transcribed spacer*) n'a pas été discriminante vis-à-vis des espèces *D. latum*, *D. dendriticum*, *D. ditremum* et *D. nihonkaiense*. Ces séquences sont aussi de mauvais marqueurs du polymorphisme intra-spécifique. Les gènes *Cox1* et *Cytb* (cytochrome b) ont permis non seulement de distinguer ces espèces mais aussi de mettre en évidence différents haplogroupes pour l'espèce *D. latum* (Wicht *et al.*, 2008). Le gène *Cox1* a d'ailleurs été étudié pour des analyses médicales, confirmant un diagnostic de diphyllbothriose à *D. nihonkaiense*, alors que l'observation des proglottis expulsés suggéraient un *D. latum* (Chen *et al.*, 2014).

2) Les trématodes

Ces plathelminthes possèdent des ventouses buccales et ventrales sur la forme adulte pour pouvoir se maintenir dans les tissus de l'hôte et présentent une régression des organes sensoriels. Les vers sont hermaphrodites, avec ou sans autofécondation. Parmi les espèces pathogènes chez l'Homme, nous pouvons mentionner *Fasciola* sp. et *Dicrocoelium* sp., qui seront décrits plus en détail, mais aussi *Clonorchis* sp. ou *Schistosoma* sp., agent causal de la bilharziose. Une infection par ces douves pourra être traitée par du Triclabendazole (un inhibiteur des microtubules) ou du Praziquantel.

i. Les Fasciolidae : *Fasciola* sp. (Linné, 1758)

Fasciola hepatica et *F. gigantica* sont des douves infectant principalement le bétail, retrouvées dans plus de 50 pays et touchant deux millions de personnes. La première espèce est retrouvée dans les zones tempérées alors que la seconde est plutôt associée aux zones tropicales (Shu *et al.*, 2012). Leurs ressemblances au niveau de la morphologie des œufs et des adultes rendent ces deux espèces difficiles à distinguer. Cette parasitose peut causer une distomatose, caractérisée par des troubles gastro-intestinaux lors de la migration de la larve de l'intestin vers les voies biliaires. Cette phase d'invasion cause des nausées et des vomissements, des douleurs abdominales ainsi qu'une fièvre intermittente.

Lorsque la forme adulte du parasite est développée, les manifestations cliniques de la phase d'état peuvent être discrètes, mais une inflammation ou une obstruction des voies biliaires sont possibles, pouvant entraîner une stase biliaire avec ictère (CDC, 2015).

- Le cycle de vie dixénique

Les œufs sont non-embryonnés à la ponte dans des fèces de ruminants. Ils doivent séjourner en eau douce bien oxygénée, à une température supérieure à 10°C, avant d'éclore au printemps. Un miracidium est libéré et doit trouver son hôte, une lymnée (gastéropode), en 24 heures. Il pénètre alors son tégument et se transforme en sporocyste puis en une larve cercaire pouvant nager pour aller se fixer sur des végétaux et s'y enkyster sous forme de métacercaire. Les hôtes définitifs sont contaminés par consommation de cette végétation, et les êtres humains plus particulièrement en mangeant du cresson ou d'autres végétaux de bords de rives ou poussant dans l'eau. À ce stade, la métacercaire migre à travers l'intestin puis le foie pour se développer en douve adulte au niveau des voies biliaires et du tissu hépatique, dont elle se nourrit (Figure 7) (CDC, 2015). Une contamination est aussi possible après consommation de foie de mouton ou de bovin parasités.

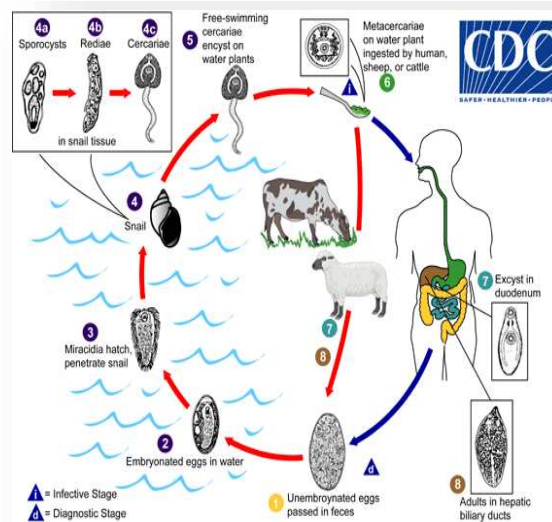


Figure 7 : le cycle de vie de Fasciola sp.

- Les génomes mitochondriaux

Les génomes de *Fasciola* sp. (14 478 pb pour *F. gigantica* et 14 462 pour *F. hepatica*) possèdent les mêmes gènes que les autres plathelminthes décrits et un taux de A+T de 62,7%. La différence interspécifique pour l'ensemble des gènes codant des protéines est de 11,6% (Liu *et al.*, 2014a).

Plusieurs travaux ciblant des gènes mitochondriaux (*Nad1*, *Cox1*) (Tadayon *et al.*, 2015; Wannasan *et al.*, 2014) ou nucléaires (*ITS1-2*, *ARNr5.8s*) (Shu *et al.*, 2012) ont montré que la distinction des deux espèces au niveau moléculaire était plutôt aisée.

ii. Les Dicrocoeliidae : *Dicrocoelium* sp. (Dujardin, 1845)

La petite douve du foie infecte le bétail, en particulier les moutons et les chèvres. Une dicrocoeliose peut causer chez l'humain une distomatose, si une fourmi contaminée est ingérée accidentellement. Cette parasitose est exceptionnelle chez l'Homme, moins de 15 cas ont été documentés (Gentilini, 2012). Le ver adulte hermaphrodite d'environ un centimètre colonise les canaux hépatiques.

- Le cycle de vie trixénique

Les ruminants sont les hôtes définitifs les plus fréquents, même si d'autres herbivores, des carnivores et les humains peuvent aussi être contaminés. Des œufs embryonnés sont dispersés en même temps que les fèces de l'hôte définitif. Un escargot, tel *Zebrina* sp. ou *Cionella* sp., les ingère, permettant l'éclosion de miracidiums. Ils migreront à travers la paroi intestinale vers les tissus vasculaires, où ils se transformeront en sporocystes, qui se dirigeront dans la glande digestive. Une larve cercaire sera libérée et se dirigera vers la chambre respiratoire, où elle sera déposée dans l'environnement avec le mucus émit par l'escargot. Cette substance devra être ingérée par une fourmi (*Formica* ssp.). La cercaire évolue alors en larve métacercaire après un passage dans l'intestin vers l'hémocoèle. Après avoir mangé une fourmi infectée, l'hôte définitif sera contaminé par la métacercaire qui va s'enkyster dans le petit intestin avant de migrer par le conduit biliaire pour donner la douve adulte.

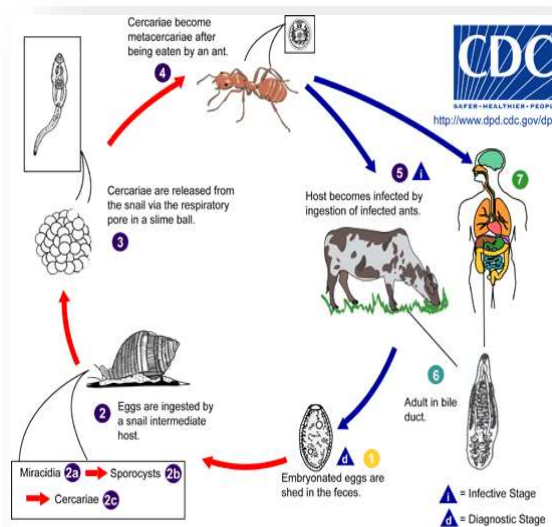


Figure 8 : le cycle de vie de *Dicrocoelium dendriticum*.

Le parasitisme de la fourmi s'accompagne d'une modification de son comportement : quelques cercaires iront s'enkyster dans le ganglion sous-œsophagien, qui contrôle le nerf mandibulaire. Les fourmis infectées ont un comportement diurne normal, mais resteront accrochées par leur mandibule à l'extrémité des brins d'herbe à la tombée du jour. Le parasite manipule le comportement de son hôte intermédiaire pour augmenter ses chances d'être mangé par un herbivore (Figure 8) (CDC, 2015).

- Les génomes mitochondriaux

Les génomes des espèces *Dicrocoelium dendriticum* et *D. chinensis* ont été séquencés en 2014, et sont similaires à celui des autres génomes de cestodes en terme de taille (14 884 et 14 917 bp respectivement), de composition, d'ordre des gènes et biais de la composition en A+T (Liu *et al.*, 2014a). En comparant les génomes complets, les deux espèces présentent 11,8% de variations nucléotidiques pour les gènes codant des protéines et 11,4% de variations au niveau des séquences en acides aminés.

B. Les Némathelminthes

Ces vers ronds sont des protostomiens ecdysozoaires, c'est-à-dire des animaux dont le développement s'effectue par mues cuticulaires. Cet embranchement regroupe des organismes pseudocoelomates, avec un tube digestif ouvert sur une bouche et un anus. Le système nerveux est composé simplement d'un ganglion périoesophagien et de deux cordons nerveux. La forme adulte femelle est plus grosse que le mâle. La respiration s'effectue par la cuticule, et il y a une absence de système circulatoire et respiratoire.

1) *Les nématodes*

Certaines estimations portent à 100 millions le nombre d'espèces composant ce phylum, celles-ci colonisant le sol, le milieu marin ou l'eau douce, et menant une vie libre (*Caenorhabditis elegans*) et/ou parasite des animaux (Oxyuridea) ou des plantes (*Tylenchus* sp.) (Poinar *et al.*, 2010).

Les nématodes regroupent des helminthes transmis par le sol par contamination oro-fécale et affecteraient jusqu'à 2 milliards de personnes dans le monde. Ce problème de santé publique majeur est une source de préoccupations pour l'OMS, qui publie régulièrement des rapports sur l'avancée de l'éradication de ces parasitoses (OMS, 2013).

Les traitements médicamenteux consistent en la prise d'Albendazole, de Mebendazole ou de Flubendazole, qui inhibent l'incorporation de la tubuline dans le cytosquelette, ou de Pyrantel, qui entraîne un blocage neuromusculaire par une activation prolongée des récepteurs à l'acétylcholine (Vidal, 2015).

i. Les Ascarididae : *Ascaris lumbricoides* (Linnaeus 1758) et *A. suum* Goeze, 1782)

Ce genre de ver rond, pouvant atteindre jusqu'à 35 cm de longueur, peut infecter respectivement l'être humain ou le porc. Selon l'OMS, près de 1,5 milliard de personnes sont parasitées à travers le monde, dont 350 millions seraient gravement atteints, causant environ 60 000 décès par année (Appelt *et al.*, 2014). Les cas sont plus fréquents dans les régions tropicales et subtropicales, surtout lorsque les conditions d'hygiène sont inadéquates.

- Le cycle de vie monoxène

Le ver adulte vit dans le petit intestin de son hôte, où la femelle peut pondre jusqu'à 200 000 œufs par jour. Ceux-ci sont libérés dans l'environnement avec les fèces, et doivent y séjourner au moins 18 jours selon les conditions du sol, pour devenir infectieux. Le cycle se poursuit lorsqu'un hôte, humain ou porc, consomme des aliments ou de l'eau souillés par des œufs. La larve peut alors éclore dans l'intestin, le traverse pour emprunter le système porte et la circulation systémique vers les poumons. Elle arrive à maturation après environ deux semaines, remonte l'arbre bronchique vers la gorge, pour y être avalée et atteindre le petit intestin, où elle se développera en ver adulte (Figure 9) (CDC, 2015).

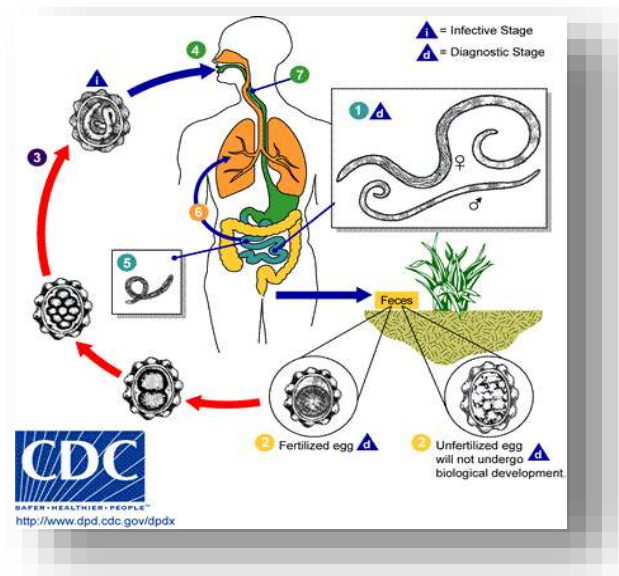


Figure 9 : le cycle de vie d'*Ascaris* sp.

- *Ascaris lumbricoides* et *A. suum* : taxonomie

Les deux espèces du genre, *Ascaris lumbricoides*, infectant les humains, et *A. suum*, infectant les porcs, sont sujettes à un débat concernant leur taxonomie. Leur similarité morphologique, tant au niveau du ver adulte que des œufs, ainsi que leurs caractéristiques biologiques ont maintenu le débat sur la nécessité de définir deux espèces, ou si le genre ne serait pas plutôt constitué de variants infectant indifféremment les deux hôtes potentiels (Leles *et al.*, 2012). Trois modèles sont proposés pour expliquer la transmission du parasite, dans les régions où les humains et les porcs sont infectés (Figure 10).

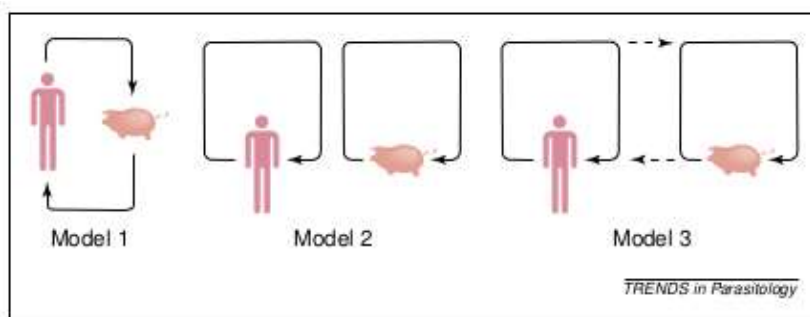


Figure 10 : trois modèles possibles de transmission d'*Ascaris* dans les populations porcines et humaines. Dans le modèle 1, les humains et les porcs sont deux hôtes alternatifs que le parasite peut infecter pour survivre. Dans le modèle 2, les parasites sont spécialisés pour un hôte en particulier et vivent en sympatrie. Le modèle 3 expose deux populations d'*Ascaris* spécialisées pour un hôte, mais avec des infections croisées possibles (Anderson, 2001).

Depuis le début du 20^{ème} siècle, plusieurs études ont été menées pour tenter d'apporter un éclairage sur le genre. Concernant les analyses morphologiques, des variations notamment au niveau des crêtes denticulées ont été observées entre les *Ascaris* du porc et ceux de l'humain. Cependant, ces différences ne sont pas toujours visualisées lors d'études indépendantes, ce point étant probablement lié au fait que les spécimens étudiés provenaient de régions géographiques distinctes. Ces études ne permettent pas d'associer les variations observées à la biogéographie des parasites ou au fait qu'ils infectent un hôte en particulier (Galvin, 1968 ; Anderson, 2001). Les analyses microscopiques des œufs ne permettent pas non plus de distinguer les espèces.

Des expériences de contaminations expérimentales ont montré que des œufs retrouvés chez le porc pouvaient infecter l'humain, et inversement. Ceci montre que les vers peuvent accomplir un cycle de vie complet, indépendamment du type d'hôte. D'autres expériences ont, à l'opposé, conclu que des génotypes du parasite, retrouvés dans certaines populations humaines étudiées, était incapable d'engendrer une infection chez le porc. Dans tous les cas, ces conclusions ne reflètent pas tout à fait les réels contextes de transmission, car les quantités d'œufs utilisées pour les infections artificielles étaient beaucoup plus élevées qu'en milieu naturel (Loreille et Bouchet, 2003).

Au niveau moléculaire, plusieurs marqueurs génétiques ont été analysés. En utilisant comme cible nucléaire *ITS1*, cinq génotypes ont été déterminés dans des populations chinoises, G1 étant associé principalement à des hôtes humains et G3 aux hôtes porcins. Les trois autres génotypes étaient retrouvés indifféremment chez les deux hôtes (Peng *et al.*, 2003). Au Brésil, la même cible génétique a montré une variabilité interindividuelle élevée, ne permettant pas d'associer un génotype à un hôte en particulier (Leles *et al.*, 2010).

Au niveau mitochondrial, les gènes *Cox1* et *Nad1* ont permis d'établir 26 haplogroupes d'*Ascaris* dans les populations chinoises, dont un seul est retrouvé chez les deux hôtes (Peng *et al.*, 2005). La Chine étant considérée comme une zone endémique pour les ascaridoses, ces résultats suggèrent que les cycles de transmission dans ce type de région sont hôte-spécifique. Cependant, l'étude de microsatellites de 258 *Ascaris* chinois a montré des contaminations croisées ainsi que des hybrides entre les vers retrouvés chez l'Homme et ceux retrouvés chez le porc (Zhou *et al.*, 2012).

Une analyse plus exhaustive a été réalisée sur le gène *Cox1*, pour étendre les haplogroupes établis précédemment dans les populations chinoises. Cinq cent trente-six vers d'*Ascaris* prélevés chez l'Homme ou le porc, provenant de 32 locations différentes, réparties dans 11 pays, ont été génotypés (Betson *et al.*, 2014). Soixante-quinze haplogroupes ont été identifiés, dont 43 nouveaux. H1 et H3 sont des haplogroupes les plus fréquemment trouvés chez les vers humains ; H7, H28, H52, H64 sont souvent identifiés chez le porc, H64 étant le seul spécifique à des *Ascaris* prélevés chez le porc. Les haplogroupes retrouvés en Europe (zone non - endémique) correspondent à une situation du modèle 3 de la Figure 10, car les vers humains sont porteurs d'haplogroupes typiques des vers du porc, indiquant une origine zoonotique. À l'opposé, en Afrique, les vers des porcs et ceux des humains sont porteurs d'haplogroupes distincts, indiquant une barrière aux échanges génétiques (modèle 2).

La combinaison des résultats d'haplogroupes mitochondriaux et nucléaires semblent suggérer qu'un seul marqueur n'est pas suffisant pour définir une ou deux espèces d'*Ascaris*. Par ailleurs, ces résultats confirment que le mode de transmission diffère dans les régions endémiques ou non, suggérant que le parasite est peut-être actuellement dans une phase d'adaptation, suite à un changement d'hôte qui a pu se produire par le passé (Leles *et al.*, 2012).

- Les génomes mitochondriaux

La composition en gènes, ainsi que leur ordre, est le même chez les deux espèces et correspond à celui décrit précédemment pour les cestodes. Le taux en A+T est élevée (71,7 - 71,8%). L'identité de séquence nucléique entre les espèces *Ascaris lumbricoides* et *A. suum* est de 98,1% alors que celle entre des génomes d'*A. suum* américain et chinois est de 98,5% (Liu *et al.*, 2012b ; Okimoto *et al.*, 1992). Ces pourcentages indiquent que la variabilité « interspécifique » est aussi élevée que la variabilité liée à la géographie.

- Le génome nucléaire

En 2011, 273 Mb d'une ébauche du génome nucléaire d'*Ascaris suum* a été publié avec une couverture de 80X (Jex *et al.*, 2011). Ce génome possède un contenu en GC de 37,9%, et présente peu de séquences répétées, comparativement aux autres génomes de métazoaires publiés. Il est composé d'environ 18 500 gènes codant des protéines, dont une portion importante de peptidases, permettant de dégrader et pénétrer les tissus de l'hôte, et de molécules pouvant moduler la réponse immunitaire. L'analyse des voies métabolique a prédit l'existence de 225 « *choke points* », liés à des enzymes dont l'absence entrainerait une toxicité cellulaire ou une privation en métabolites essentiels. Parmi eux a été identifié une inosine-5'-monophosphate déshydrogénase qui possède plusieurs inhibiteurs connus, tels les analogues de l'acide mycophénolique. Cette enzyme pourrait être ciblée dans le développement de traitement ascaricide.

ii. Les Oxyridae : *Enterobius vermicularis*

Ces petits nématodes de deux à treize millimètres de long, de la classe des Chromadorea, sont les parasites les plus répandus dans les pays industrialisés occidentaux et infectent uniquement les êtres humains. Lors d'une oxyurose, les femelles quittent l'intestin pour l'anus pour y pondre des œufs, causant à l'hôte des démangeaisons pouvant empêcher le sommeil. Le traitement à l'Albendazole ou au Flubendazole est habituellement efficace.

- Le cycle de vie

La contamination se produit lorsque des objets portant des œufs infectieux sont touchés, puis les mains portées à la bouche. Une auto-contamination est aussi possible, après contact avec les mains de la région péri-anale. Les œufs étant de très petite taille (50-20 µm), ils peuvent être aéroportés et inhalés. Environ un mois plus tard, les larves éclosent dans le petit intestin et les adultes colonisent le gros intestin. Les femelles migrent vers l'anus durant la nuit pour y pondre des œufs, qui deviennent infectieux en moins de six heures (Figure 11) (CDC, 2015).

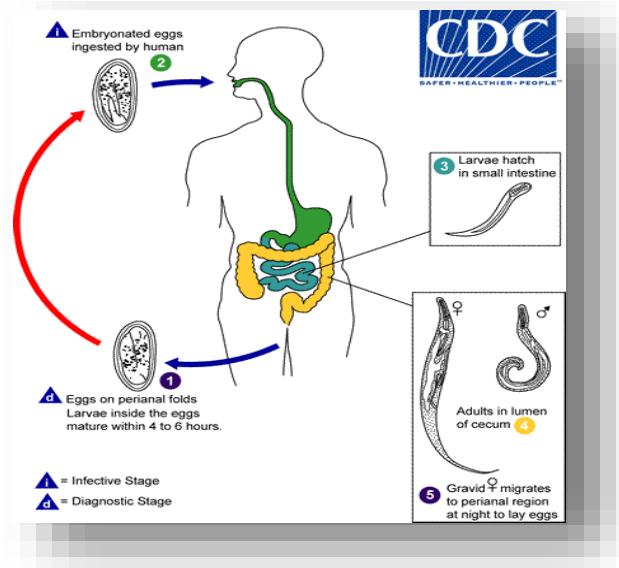


Figure 11 : le cycle de vie d'*Enterobius vermicularis*.

- Le génome mitochondrial

Cette espèce possède un génome de 14 010 pb, encodant les mêmes gènes que les autres nématodes décrits. Parmi les 12 gènes codant des protéines, huit utilisent TTG comme codon d'initiation de la traduction, le codon le plus fréquent chez les nématodes, ATT, n'est pas prédit comme étant utilisé. Ce même profil se retrouve aussi chez *Anisakis simplex*, un parasite des poissons marins (Kang *et al.*, 2009).

iii. Les Trichuridae : *Trichuris trichiura* (Linné, 1771)

Le genre *Trichuris* est composé d'une vingtaine d'espèces, dont les plus importantes sont *T. trichiura* (infectant les humains), *T. suis* (parasite du porc), *T. vulpis* (parasite des canidés, pouvant aussi toucher l'être humain), *T. campanula* (pouvant affecter les chats) et *T. muris* (parasite des rongeurs). Ce nématode est l'un des trois principaux helminthes, avec *Ascaris* sp. et les Ankylostomes, pouvant infecter les êtres humains par une transmission via le sol. Sept cents millions de personnes en sont affectées à travers le monde et environ 10 000 en décèdent chaque année.

L'Albendazole et le Mebendazole, les traitements classiques, n'éradiqueraient que le tiers des parasitoses traitées, avec une réinfection très fréquentes (Jia *et al.*, 2012). Le genre *Trichuris* a développé une niche particulière dans le système digestif puisqu'il colonise l'épithélium du colon, un des tissus au renouvellement cellulaire le plus rapide (CDC, 2015).

Une infection par *T. trichiura* peut être asymptomatique, persistante pendant des années, ou bien légère avec une simple colite. Dans d'autres cas, elle peut se présenter sous la forme d'un tableau dysentérique ou anémique.

- Le cycle de vie monoxénique

Le cycle de vie de *Trichuris trichiura* est similaire à celui d'*Ascaris* sp., ce qui explique qu'une co-infestation des deux parasites est fréquemment observée. Les œufs non-embryonnés se retrouvent dans le sol avec les fèces et deviennent infectieux en 15-30 jours. Après ingestion d'aliments contaminés par des œufs ou portage des mains souillées à la bouche (péril fécal), les larves sont libérées dans le petit intestin, avant de passer vers le gros intestin et coloniser la muqueuse intestinale. Les femelles commenceront à pondre 3 000 à 20 000 œufs par jour après deux mois d'infection (Figure 12) (CDC, 2015).

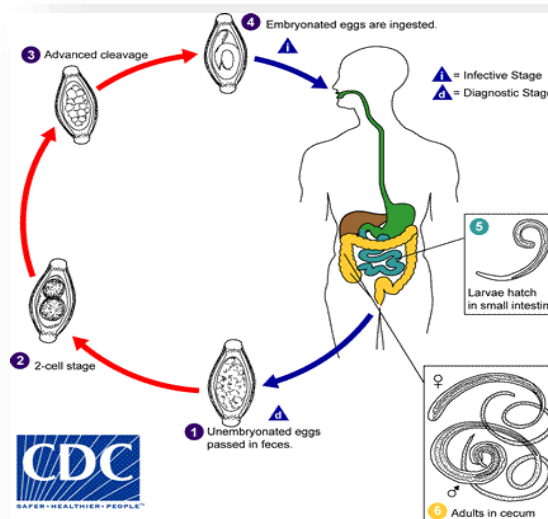


Figure 12 : le cycle de vie de *Trichuris trichiura*.

- Génome mitochondrial

Les séquençages des génomes de *Trichuris trichiura* et *T. suis* ont été menés simultanément, de manière à pouvoir les comparer et appuyer ou non les précédentes hypothèses phylogénétiques concluant en la distinction des espèces. L'analyse des *ITS* (Cutillas *et al.*, 2009 ; Liu *et al.*, 2014b) ayant montré une grande variabilité intra-spécifique au sein du genre, ce marqueur nucléaire n'est pas approprié pour une étude phylogénétique. Contrairement au genre *Ascaris*, les données génomiques et génétiques montrent que ces deux parasites sont bien deux espèces de *Trichuris* distinctes (Liu *et al.*, 2012a).

Les comparaisons de la séquence en nucléotides des gènes codant des protéines et de la séquence en acides aminés de *T. trichiura* et *T. suis* montrent des variations allant respectivement de 25,4%-37,4% et 13,6-62,5%, Cox1 étant la protéine la mieux conservée et Atp8, celle présentant le plus de variations. Ces différences entre deux espèces d'un même genre sont plus élevées que ce que l'on retrouve habituellement chez les nématodes : *Ancylostoma caninum* et *A. duodenale* présentent entre 4,9 et 10% de variation au niveau de la séquences en acides aminés (Jex *et al.*, 2009 ; Hu *et al.*, 2002).

Les génomes de ces deux *Trichuris* présentent également la particularité de coder le gène *Atp8*, présent uniquement chez les nématodes du groupe paraphylétique des adenophorean (Lavrov et Brown, 2001).

- Le génome nucléaire

Les génomes nucléaires des espèces *Trichuris trichiura*, *T. suis* et *T. muris* ont été décryptés en 2014. Dans le cas du parasite du porc, un assemblage de 80 Mb avec une couverture de 140X a été publié (Jex *et al.*, 2014), 75,2 Mb et 85 Mb ont été séquencées pour les parasites de l'Homme et de la souris (Foth *et al.*, 2014). Leurs analyses ont mis en évidence :

- Des adaptations au parasitisme : l'étude de ces génomes et du transcriptome de *T. muris*, à apporter des nouvelles connaissances quant à l'implication de la région antérieure du parasite dans l'ancrage à long terme dans la muqueuse intestinale, l'immuno-modulation de la réponse de l'hôte et la nutrition du parasite. Environ 75 chymotrypsines de la famille des sérine-protéases seraient présentes chez ces espèces, dont les trois quarts seraient surexprimées dans la partie antérieure du ver adulte.

Parmi celles-ci, les deux tiers seraient des protéines secrétées, ce qui est cohérent avec les hypothétiques fonctions précédemment mentionnées pour cette région du ver.

- Des cibles thérapeutiques : de nouvelles cibles thérapeutiques ont pu être identifiées, dont 29 gènes essentiels pouvant déjà être ciblés par la pharmacologie existante. Parmi ceux-ci, mentionnons une adénosine désaminase, un récepteur au LDL (*low density lipoprotein*) et une histone desacétylase pouvant être visés respectivement par un antiviral, un antibiotique et une statine.
- L'immunomodulation de la réponse de l'hôte : ce phénomène était connu avant l'étude des génomes nucléaires, en particulier chez *T. suis*, sans que les mécanismes moléculaires et cellulaires impliqués ne soient décrits. Des médecins ont proposé de traiter des maladies telles les colites ulcéreuses ou les maladies de Crohn par une infection artificielle à *T. suis* après avoir 1/ étudié les données immunologiques obtenues à partir d'animaux de laboratoire infectés expérimentalement par *T. suis*, 2/ constaté que les maladies intestinales inflammatoires d'origine auto-immune sont peu fréquentes dans les pays où le taux d'helminthiase est élevé (Elliott et Weinstock, 2012 ; Summers *et al.*, 2005). Un modèle biologique sur les effets immunologiques de *T. suis* a ensuite été proposé lors de l'analyse des données génomiques. En se basant sur les gènes surexprimés durant le stade larvaire et au niveau de la partie antérieure du ver (stichosome), plusieurs effets anti-inflammatoires ont été présentés. De nombreuses molécules entrent en jeu pour inhiber les réponses immunitaires telles la différenciation de lymphocytes T pro-inflammatoires, la production de monoxyde d'azote par les macrophages, la stimulation d'interleukines anti-inflammatoires et en limitant la destruction des tissus de l'hôte par la sécrétion d'inhibiteurs de protéases. Une diminution des symptômes associés à certaines pathologies digestives a pu être observée après ingestion d'œufs infectieux lors d'étude en double-aveugle (Summers *et al.*, 2005). Cependant, la communauté médicale soulève diverses questions concernant la sécurité de ce type de traitement : même si *T. suis* ne parasite pas l'être humain, il existe des risques de colonisation extra-digestive, dont les risques peuvent être importants pour un patient (Hsu *et al.*, 2005 ; Van Kruiningen et West, 2005).

III. Bilan du chapitre II

Les helminthes gastro-intestinaux sont dans nos sociétés modernes, une source de préoccupations pour la santé publique. Leurs impacts sur les populations peuvent être importants dans les pays en voie de développement, où les maladies associées à une contamination oro-fécale sont nombreuses. Une infestation parasitaire associée à d'autres pathologies et à la malnutrition peut engendrer un mauvais état de santé général, surtout chez les enfants.

Les mesures d'hygiène à mettre en place pour prévenir les infestations parasitaires sont nombreuses et représentent un investissement financier important : installer des toilettes et traiter les eaux usées, ne consommer que de l'eau potable, ne pas utiliser d'engrais de fertilisation d'origine humaine, traiter les animaux infestés... Ces mesures permettent en même temps de diminuer la transmission d'autres maladies, telles l'hépatite A ou le choléra. Bien que les plus récentes statistiques de l'OMS montrent une augmentation de la proportion de la population mondiale ayant accès à l'eau potable (près de deux milliards de personnes en plus par rapport à 1990), 185 millions de personnes consomment de l'eau de surface pour leur besoin quotidien. Si l'on rajoute à ces données les 2,5 milliards de personnes n'ayant pas accès à des installations sanitaires améliorées (OMS et UNICEF, 2013), des efforts pour enrayer ces parasitoses restent encore à faire au niveau des politiques nationales et internationales.

Chapitre III : La paléoparasitologie

I. Historique

La paléoparasitologie est une des disciplines fondatrices de la paléopathologie, dont un des pionniers est Sir Armand Ruffer. Après avoir contracté la diphtérie en s'inoculant un sérum, Ruffer décida d'aller se reposer en Égypte, et y accepta un poste de professeur de bactériologie. Probablement inspiré par ce séjour dans ce pays à l'histoire longue et passionnante, il développe la paléopathologie, qu'il décrit comme « la science des maladies qui peuvent être démontrées sur les restes humains et animaux des temps anciens » (Ruffer, 1913). Il s'intéresse tout particulièrement à l'étude de tissus provenant de momies. En 1910, il effectue la première observation de paléoparasitologie en identifiant par microscopie des œufs de *Schistosoma haematobium* au niveau des reins d'une momie de la XX^e dynastie (1186-1069 BC). Par la suite, il identifie une tuberculose chez une momie d'un prêtre d'Amon, en observant les signes du mal de Pott ainsi qu'un abcès du muscle psoas. Au cours de sa carrière, il diagnostiqua de nombreuses pathologies congénitales (pied varus-équin), infectieuses (tuberculose, poliomyélite, ostéomyélite), liées aux conditions de vie (artériosclérose, rachitisme, caries dentaires) ou à l'âge (goutte, rhumatismes) (Ruffer, 1921).

Durant le 20^{ème} siècle, la paléoparasitologie continue de se développer avec comme objectif principal la mise en évidence microscopique des restes de parasites dans des contextes archéologiques historiques et préhistoriques. Le plus souvent, les marqueurs qui sont retrouvés sont des œufs d'helminthes gastro-intestinaux, plus rarement des restes de larves ou de vers adultes. Les œufs, forme de dissémination émise par les vers, possèdent une coque composée en partie de chitine et d'un nombre variable de couches lipidiques. Ces molécules les rendent relativement résistants aux processus de dégradation diagénétique et taphonomique. La composition de la paroi et/ou de l'enveloppe varie d'un taxon à l'autre, ce qui signifie qu'avec des conditions de préservation identiques, l'identification de certains parasites est plus ou moins probable ou facile. Les œufs d'*Enterobius vermicularis* par exemple, sont fragiles et rarement retrouvés en contexte archéologique. À l'opposé, les Ascaridae et Trichuridae, dont la coque est plus épaisse, sont fréquemment identifiés.

Il est par ailleurs nécessaire de prendre en compte le cycle de vie des parasites dans l'interprétation des taxons qui ne sont retrouvés que rarement. Si l'éclosion de la larve, a lieu précocement après le dépôt dans l'environnement, on peut s'attendre à une sous-évaluation de la présence de ce taxon en raison de la destruction de l'œuf ou de la perte des caractéristiques le rendant identifiables. C'est notamment le cas pour les Ancylostomides ou les Trichostrongylides (Bouchet *et al.*, 2003a). L'identification microscopique des taxons se base sur l'analyse des caractéristiques morphologiques et micrométriques des œufs (Figure 13). Le diagnostic au niveau de l'espèce n'est possible que pour une minorité des parasites étudiés, car la plupart des œufs au sein d'un même genre, voire d'une même famille, présentent des caractéristiques identiques.

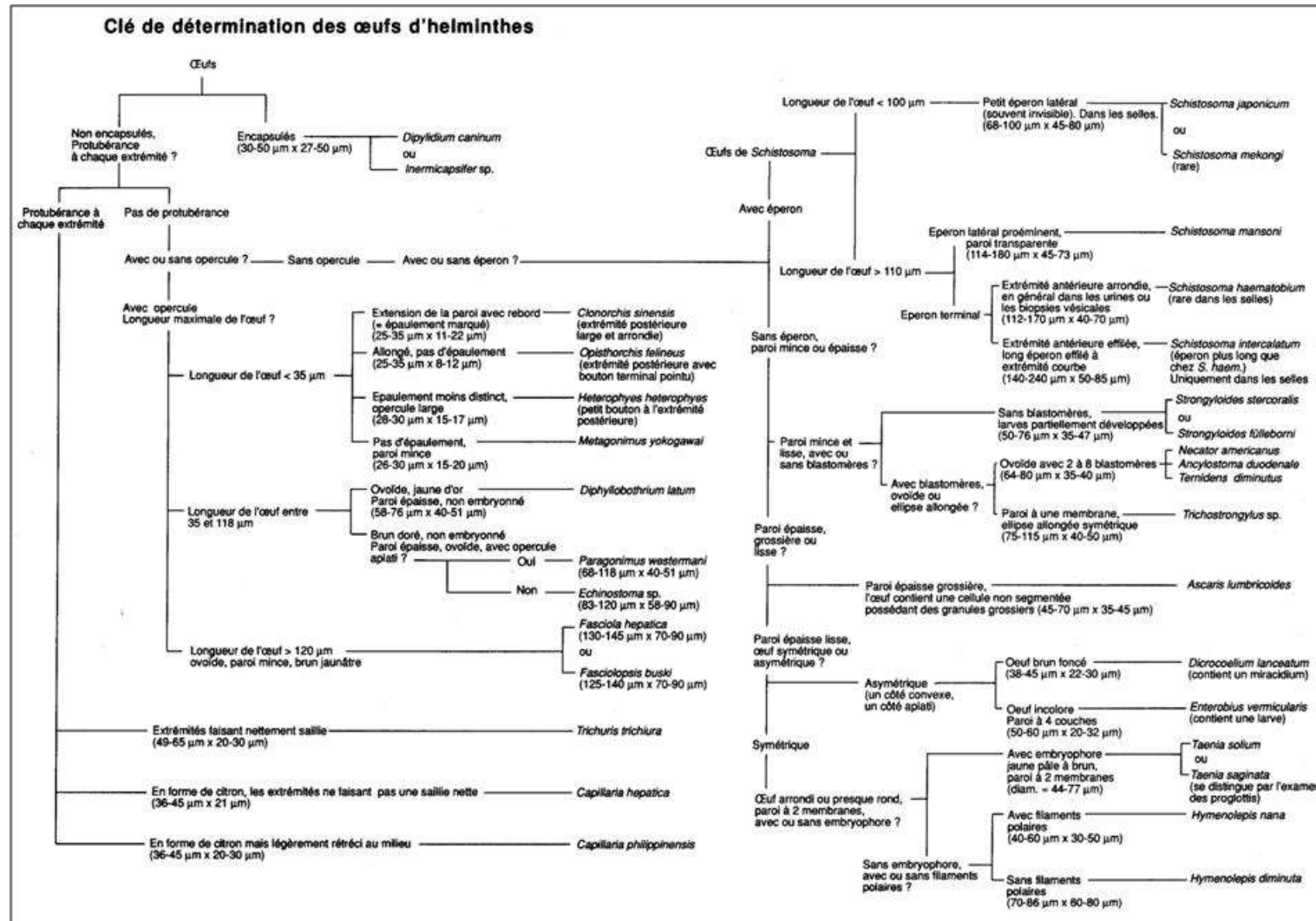


Figure 13 : clé d'identification d'œufs d'helminthes humains (Chiodini et al., 1993). Cette table comporte une erreur car elle décrit l'œuf de *Dicrocoelium* comme n'étant pas operculé.

Dans ce contexte, les paléoparasitologues ont été à l'affut des avancés en biologie moléculaire et en immunologie afin d'affiner leurs analyses. La méthode ELISA (*Enzyme-Linked ImmunoSorbent Assay*) a permis de mettre en évidence des antigènes de *S. haematobium* dans des tissus d'une momie égyptienne de 1198 BC (Deelder *et al.*, 1990). Moins de 10 ans plus tard, une publication rapportait l'identification d'ADN de *Trypanosoma cruzi*, provenant de restes humains momifiés au Nord du Chili et datant de 9000 ans (Guhl *et al.*, 1999). Ce résultat, basé sur l'interprétation d'une électrophorèse sur gel, n'était pas accompagné d'un séquençage de l'ADN, mais d'autres équipes ont poursuivi l'étude de ce parasite responsable de la maladie de Chagas par hybridation moléculaire (Ferreira *et al.*, 2000) puis par séquençage (Madden *et al.*, 2001). La même année, les premières séquences de l'ADN mitochondrial de l'helminthe gastro-intestinal *Ascaris* sont obtenues (Loreille *et al.*, 2001).

Que les analyses soient menées classiquement par microscopie ou par un diagnostic moléculaire, les informations obtenues en paléoparasitologie sont nombreuses et diverses. Elle permet tout d'abord d'obtenir des connaissances d'ordre sanitaire et médical. Comme il a été décrit dans le chapitre II, certaines parasitoses sont associées au péril fécal, c'est-à-dire à une transmission oro-fécale via la consommation d'aliments ou d'eau souillés. Au niveau paléo-ethnologique, cette discipline permet de mieux comprendre les modes de vie des populations anciennes. Un régime alimentaire comportant du poisson cru étant plutôt associé à un risque de contracter une diphyllobothriose, alors que la consommation de porc mal cuit peut transmettre un *Tænia solium*. Sur une échelle temporelle et géographique plus large, la répartition des parasitoses ou leur apparition chez l'être humain peuvent être des indicateurs d'une révolution profonde. Le Néolithique, avec la mise en place de l'agriculture et la domestication des animaux, est une période qui a pu favoriser des changements d'hôtes et augmenter la quantité de parasites présents. En effet, un plus grand nombre d'espèces avec un nombre croissant d'individus localisés sur un site restreint, a pu favoriser la transmission de nouveaux parasites aux humains et aux animaux (voir (Zammit, 2005) pour les concepts de promiscuité Homme/animal et (Dittmar, 2009)).

II. La paléoparasitologie morphologique

Dans cette section seront décrits la morphologie des œufs d'helminthes gastro-intestinaux étudiés dans le chapitre II.

A. Ascaris sp.

Comme mentionné précédemment, les œufs des deux espèces *Ascaris lumbricoides* et *A. suum* ne sont pas distinguables. Ces œufs, de 45-70 µm x 35-45 µm, sont sans protubérance, ni opercule ou éperon. La paroi est épaisse et mamelonnée, et ses trois couches de lipides sont observables. Si les conditions environnementales sont favorables, les œufs peuvent rester infectieux pendant six ans (VanBuskirk *et al.*, 2009). Ceux provenant d'un porc restent généralement viables 40 jours à une température comprise entre -18 et -27°C (Hindiyeh, 2004).

Notons que, dans une perspective plus large que la parasitologie humaine, la plus ancienne trace d'œufs d'Ascariidés remonte au Crétacé. Un coprolithe appartenant à un dinosaure (Iguanodon) a été analysé et une infestation par ce nématode a été déterminée (Poinar et Boucot, 2006).

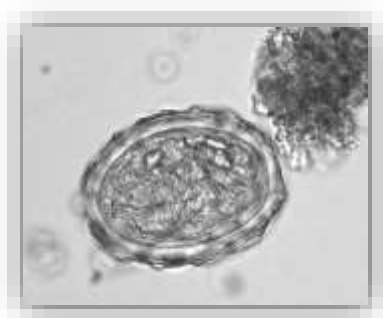


Figure 14 : un œuf d'*Ascaris lumbricoides* provenant de la cavité abdominale d'un soldat de la Première Guerre Mondiale. Carspach, France (Le Bailly *et al.*, 2014a).

B. Trichuris trichiura

Les œufs de *Trichuris trichiura* et *T. suis* sont considérés comme non-distinguables. Ils présentent des protubérances aux extrémités saillantes et leur taille est de 49-65 µm x 20-30 µm. La coque est épaisse et leur forme de citron les rend facilement reconnaissables.

La plus ancienne infestation humaine actuellement rapportée remonte au Néolithique (5200-5300 BP), avec la momie italienne Ötzi qui était porteur de ce parasite (Aspöck *et al.*, 1996).



Figure 15 : un œuf de *Trichuris trichiura*, provenant de la cavité abdominale d'un soldat de la Première Guerre Mondiale. Carspach, France (Le Bailly *et al.*, 2014a).

C. Enterobius vermicularis

Les œufs d'*Enterobius vermicularis* sont incolores. Ils sont ovalaires et asymétriques. Leur paroi est épaisse et lisse, présentant quatre couches. La taille des œufs est approximativement de 55 µm x 30 µm.

Ces œufs sont rarement retrouvés dans des vestiges européens avec une seule occurrence répertoriée par (Herrmann, 1988) alors qu'ils le sont plus fréquemment dans des coprolithes d'Amérique (de Araujo *et al.*, 1985 ; Iniguez *et al.*, 2003b).



Figure 16 : des œufs d'*Enterobius vermicularis* (CDC).

D. Les Taeniidae

Les œufs des *Tenia* pouvant infecter les humains sont non-distinguables des œufs du genre *Echinococcus*, parasite des carnivores. Les œufs symétriques sont presque ronds, avec une paroi épaisse et lisse à deux membranes. Ils sont délimités par une membrane qui entoure l'embryophore. Celle-ci est très fragile et est détruite au cours de la taphonomie, c'est donc uniquement l'embryophore qui est retrouvé dans les échantillons archéologiques, de 32 à 36 µm de diamètre.

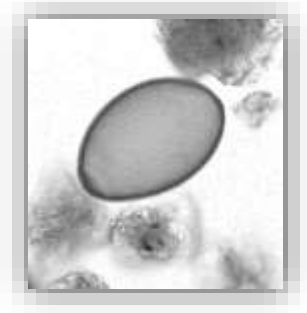


Figure 17 : des embryophores de *Teniadés* (Nezamabadi *et al.*, 2013b).

Des cas de transport à longues distances d'œufs, par l'intermédiaire d'oiseaux ou d'insectes, ont été rapportés, suite à des cas d'infestations de moutons par le Ténia du chien en région insulaire, en absence d'hôte canin (Torgerson *et al.*, 1995).

E. Diphylobothrium sp.

La taille de ces œufs avec opercule est comprise entre 60-70 µm x 40-60 µm. Ils présentent une coque mince, avec un épaissement appelé mucron, au pôle opposé à l'opercule. Les œufs peuvent présenter des variations de l'ornementation, plus ou moins ponctuées. Elles peuvent permettre d'orienter le diagnostic vers certaines espèces de *Diphylobothrium* en particulier (*D. dendriticum* ou *D. latum*) mais ce critère est insuffisant pour identifier les 14 espèces qui peuvent infecter l'humain.



*Figure 18 : un œuf de Diphyllbothrium sp. de 3350-3100 BC
(Le Bailly et Bouchet, 2013).*

F. Fasciola sp.

Ces œufs ovoïdes operculés, à paroi mince, sont sans protubérance, non-encapsulés mais présentent des épaissements au niveau des pôles. Ils mesurent entre 130-200 μm x 60-105 μm , cette large fourchette de taille permet de distinguer les espèces *Fasciola hepatica* de *F. gigantica*, cette dernière présentant des œufs de taille plus importante.



Figure 19 : des œufs de Fasciola hepatica (CDC).

G. Dicrocoelium sp.

Ces œufs ovoïdes sont non-encapsulés, sans protubérance, avec un opercule large et aplati. La paroi, épaisse et lisse, est asymétrique. L'œuf brun foncé mesure 38-45µm x 22-30 µm.



Figure 20 : un œuf de Dicrocoelium sp. (Le Bailly et Bouchet, 2010).

III. La paléogénétique pour l'étude des helminthes gastro-intestinaux

A. Généralités

Le développement de la paléogénétique a rapidement laissé entrevoir des perspectives pour l'étude des pathogènes anciens. Limitée au cours du 20^{ème} siècle, comme nous l'avons vu, par des observations microscopiques pour les helminthes gastro-intestinaux, l'identification des autres maladies infectieuses se basaient sur des analyses morphologiques et macroscopiques. Par exemple, des atteintes osseuses pouvant être associées à des bactéries telles des *Staphylococcus* sp. , des *Mycobacterium* sp. ou des *Treponema* sp. ; des mégaviscères, tel un mégacôlon, pouvant être observés dans des cas de maladie de Chagas, causée par le protozoaire *Trypanosoma cruzi*.

La première étude de paléogénétique portant sur des pathogènes s'est intéressée à l'identification par PCR de la bactérie *M. tuberculosis* préservée dans des squelettes montrant des signes de mal de Pott (Spigelman et Lemma, 1993). Les travaux se sont ensuite succédés pour mettre en évidence d'autres pathogènes, tels *Yersinia pestis*, l'agent causal de la peste (Drancourt *et al.*, 1998 ; Bos *et al.*, 2011, Tran *et al.*, 2013), ou *Bartonella* sp., dont *B. quintana*, transmise par les poux, est responsable de la fièvre des tranchées (Drancourt *et al.*, 2005). Le virus de la variole a aussi pu être mis en évidence sur des momies sibériennes du 17^{ème}-18^{ème} siècles (Biagini *et al.*, 2012).

L'intégration d'études génétiques permet une meilleure compréhension de l'occurrence et de la répartition des pathologies ayant affectées les populations anciennes et offre la possibilité d'ancrer la paléopathologie dans le concept de pathocénose. Proposé par Grmek, il définit ce concept en 1969:

Les états pathologiques au sein d'une population déterminée, dans le temps et dans l'espace, forment un ensemble, que nous nommons pathocénose ; la fréquence et la distribution de chaque maladie dépendent, en plus de divers facteurs endogènes [infectiosité, virulence, vecteurs] et écologiques [climat, urbanisation, promiscuité], de la fréquence et de la distribution de toutes les autres maladies ; la pathocénose tend vers un état d'équilibre, ce qui est particulièrement sensible dans une situation écologique stable (Grmek, 1969).

En permettant un diagnostic spécifique et simultané de plusieurs pathologies infectieuses, la génétique permet d'apporter cet aspect intégratif décrit par Grmek, en opposition à un procédé analytique, où chaque maladie serait décrite individuellement.

En effet, grâce à la paléogénomique ou la métagénomique, il devient possible d'étudier plusieurs pathologies au sein d'une population ou de suivre leur évolution au cours du temps. Des explications sur des changements d'occurrence de maladies ou leur apparition, en lien avec des facteurs écologiques ou associés à des changements dans les habitudes alimentaires ou sanitaires, peuvent aussi être proposées. L'intégration des approches moléculaires, morphologiques et microscopiques apparaît donc pertinente dans l'optique de mieux décrire les états pathologiques des populations du passé.

B. La paléoparasitologie moléculaire

Les analyses génétiques appliquées aux helminthes gastro-intestinaux sont un apport important pour l'étude des parasitoses ayant affectées les populations passées. La microscopie ne permet pas de déterminer plusieurs taxons au niveau de l'espèce, une information importante pour établir les états de santé et les modes de vie des populations anciennes. Par ailleurs, les processus taphonomiques peuvent affecter les caractéristiques morphologiques, ce qui peut rendre l'identification moins précise, voire impossible. L'obtention de séquences génétiques permet :

- 1- de confirmer la présence de taxons observés par microscopie,
- 2- d'établir précisément les espèces présentes,
- 3- de mettre en évidence d'autres taxons qui n'auraient pas été identifiés par microscopie,
- 4- de suivre l'évolution d'un génome de parasite au cours du temps.

1) Les prérequis aux analyses moléculaires

L'identification des espèces présentes dans un échantillon archéologique nécessite de comparer les séquences anciennes d'ADN obtenues à des séquences modernes de références. Ces dernières sont obtenues la plupart du temps à partir de spécimens adultes pour lesquels l'espèce sera identifiable. L'alignement comparatif des séquences modernes et anciennes permettra donc de diagnostiquer une espèce de parasites.

Au cours de leurs travaux de biologie moléculaire, les chercheurs qui obtiennent des séquences nucléotidiques peuvent les rendre publique en les soumettant à une banque de données telles GenBank, EMBL ou DDBJ (Benson *et al.*, 2005 ; Kanz *et al.*, 2005 ; Tateno *et al.*, 2002). Le nombre de séquences disponibles est actuellement très important (Figure 21). Une augmentation exponentielle est observable à partir du début des années 2000, et concomitante à la mise au point du séquençage de nouvelle génération (*Next-generation sequencing, NGS*).

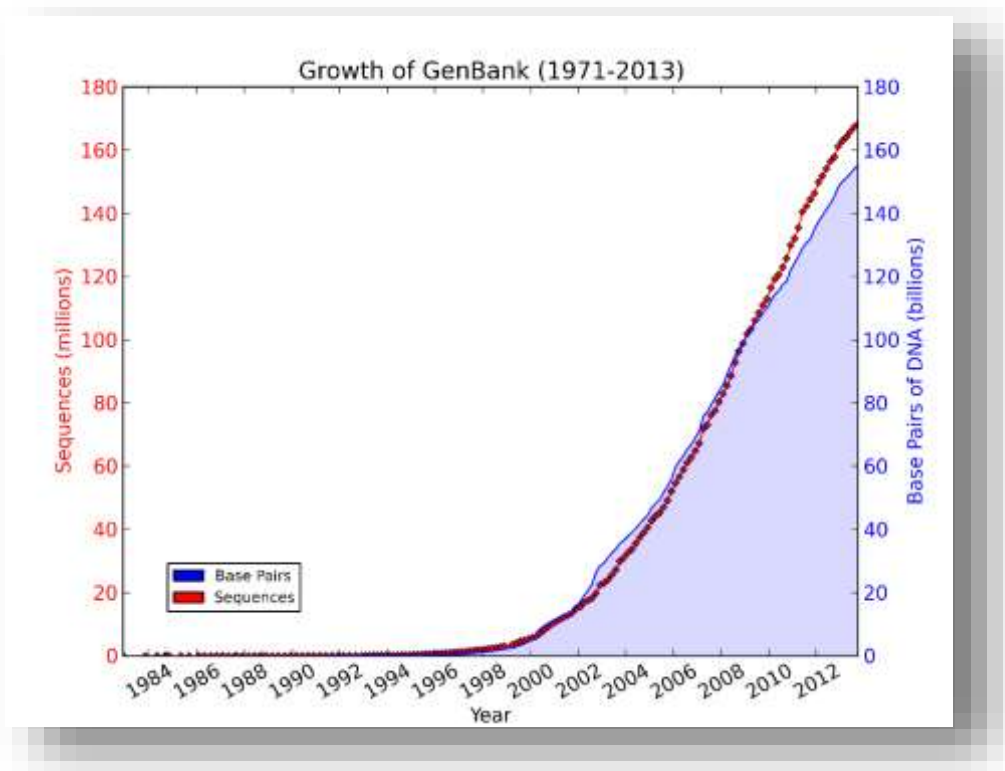


Figure 21 : évolution des données disponibles dans GenBank entre 1984 et 2012 (Mark A. Pauley, University of Nebraska).

Le nombre de séquences de références modernes ne devraient donc pas constituer un frein pour les analyses génétiques en paléoparasitologie, mais dans les faits, tous les taxons ne sont pas représentés de manière similaire dans les banques. Par exemple, début 2015, 344 séquences différentes d'ADN mitochondrial étaient disponibles pour le genre *Ascaris*, 473 pour les espèces de *Tenia saginata*, *T. asiatica* et *T. solium*, alors que pour les espèces de *Capillaria* pouvant affecter l'être humain (*C. hepatica* et *C. aerophila*), 28 séquences avaient été déposées dans GenBank (Benson *et al.*, 2005).

Ces différences peuvent s'expliquer entre autre par l'importance du parasite au niveau de la santé publique et des pertes économiques pouvant être engendrées au niveau de l'industrie agricole. Les ascaridioses touchant 1 500 millions de personnes à travers le monde (OMS, 2002) et affectant l'industrie porcine, cette parasitose est naturellement plus étudiée que les capillarioses n'affectant que sporadiquement l'humain.

Idéalement, les séquences de références modernes devraient être le reflet de la diversité biogéographique, puisque des variations génétiques associées à la géographie peuvent être observées. Pour que cela puisse être possible, il faudrait que les spécimens de parasites soient faciles à obtenir et que de nombreux pays s'investissent dans des projets de recherche les concernant. Ce n'est pas le cas, et actuellement, de nombreuses séquences publiées sont uniquement d'origine asiatique (Chine, Inde...), ou sud-américaine (Brésil, Mexique...) (Hancock *et al.*, 2001 ; Anantaphruti *et al.*, 2013 ; Leles *et al.*, 2009 ; Wang *et al.*, 2013 ; Arizono *et al.*, 2009).

2) *Les travaux de paléoparasitologie moléculaire sur les helminthes gastro-intestinaux*

Les 14 publications à ce jour ayant donné lieu à la production de séquences anciennes d'helminthes gastro-intestinaux humains sont basées sur une approche similaire. Après une analyse microscopique des œufs de parasites, un à trois taxons observés sont ciblés par PCR et confirmés par un séquençage de type Sanger. Seuls deux études de biologie moléculaire ont été menés dans le but de cibler des espèces non-observées par microscopie. La première s'est intéressée au fait que le parasite cosmopolite *Enterobius vermicularis* n'est que rarement observé dans les échantillons anciens, alors qu'il est fréquent dans les populations actuelles (Iniguez *et al.*, 2003b). La caractérisation du parasite reste cependant discutable car les auteurs n'ont pas séquencé les amplicons obtenus. Le second travail a cherché à mettre en évidence l'association classique de *Trichuris trichiura* et *Ascaris* sp., alors que seuls des œufs de premiers helminthes ont été observés dans les échantillons archéologiques étudiés (Leles *et al.*, 2008).

Dans cette partie nous exposerons et analyserons les résultats de biologie moléculaire obtenus sur les helminthes gastro-intestinaux. Seuls les travaux ayant donné lieu à la publication de séquences d'ADN seront mentionnés.

Ceux discutant d'amplicons ayant donné des « séquences non-spécifiques » ne sont pas pris en considération, ainsi que ceux dont les analyses sont basées sur des approches d'hybridations moléculaires ou par la technique de RAPD (*Random Amplified Polymorphic DNA*) (Iniguez *et al.*, 2003a ; Jaeger et Iniguez, 2014). En effet, tous les échantillons de paléoparasitologie (sédiment, coprolithe ...) peuvent être considérés comme des échantillons environnementaux, c'est-à-dire qu'ils sont composés d'une minorité d'ADN endogène, noyé dans une grande quantité d'ADN bactérien, viral, provenant de levures, de champignons... Même si la spécificité des amorces utilisées est vérifiée en interrogeant les banques de données, ces dernières sont loin d'être exhaustives. Les séquences d'ADN de toute la biodiversité terrestre n'y étant pas encore toutes répertoriées. Pour cette raison, il est nécessaire de procéder au séquençage des amplicons obtenus, même si ceux-ci présentent, après électrophorèse, une bande de taille attendue car elles peuvent correspondre à des produits non-spécifiques.

Nous terminerons en exposant quelques cas sur des helminthes affectant les animaux, et sur des parasites de type protistes (*Trypanosoma* sp., *Leishmania* sp., *Plasmodium* sp., *Toxoplasma* sp.) et sur un ectoparasite commun, le pou (*Pediculus humanus*).

i. *Ascaris* sp.

Cet helminthe fut le premier ciblé par des analyses paléogénétiques (Loreille *et al.*, 2001). Cent quatre œufs prélevés sur un coprolithe belge (1400 AD) ont permis d'obtenir des fragments des gènes *Cytb* et de l'*ARNr 18S*. Peu de séquences modernes de référence étaient disponibles à cette époque, les auteurs ont eux-mêmes séquencé une partie du gène du *Cytb* d'*Ascaris lumbricoides*. Les séquences anciennes pour ce gène présentaient trois substitutions par rapport au parasite de l'Homme et deux par rapport à celui du porc. La séquence ancienne obtenue pour le gène de l'*ARNr 18S* était parfaitement identique à l'unique séquence de référence disponible à ce moment du parasite *A. suum* (numéro d'accès U94367). En 2015, les banques de données contiennent plusieurs séquences de l'*ARNr 18S* d'*A. lumbricoides*, présentant aussi 100% d'identité avec les séquences anciennes de l'étude de Loreille *et al.* (Soe *et al.*, 2015).

Une équipe brésilienne étudia les co-infestations *T. trichiura* / *Ascaris* sur des vestiges précolombiens (8800 BP) d'Amérique du Sud (Leles *et al.*, 2008). L'origine anthropique des coprolithes est suggérée par le contexte archéologique et la présence d'œufs d'*E. vermicularis* et *T. trichiura*.

Le même fragment de *Cytb* a été étudié par PCR et séquençage, et les conclusions, basées sur la position nucléotidique 38 de l'amplicon (une cytosine chez *A. suum*, une thymine chez *A. lumbricoides*), ne sont plus d'actualité car d'autres séquences sont maintenant disponibles dans les banques de données (Liu *et al.*, 2012b), indiquant que cette position n'est pas spécifique à l'une ou l'autre des deux espèces d'*Ascaris*.

D'autres fragments du gène *ARNr 18S* ont été obtenus à partir de coprolithes humains espagnols du 18^{ème} siècle. Les amplicons étaient identiques à 100% à la séquence de référence d'*A. suum* (numéro d'accès AF036587), confirmant que ce gène ne permet pas de distinguer les espèces du genre. L'origine humaine de l'échantillon a été confirmée par l'obtention d'une séquence d'ADN mitochondrial d'haplogroupe K (Gijon Botella *et al.*, 2010).

Des œufs prélevés directement sur la surface osseuse du bassin d'un corps emballé dans du tissu ont aussi été analysés. Ce squelette, d'origine coréenne, a été retrouvé dans une tombe médiévale correspondant à l'époque de la dynastie Joseon (1392-1910 AD). Des œufs d'*Ascaris* en association avec *Trichuris trichiura* ont été observés en microscopie, seul le premier de ces parasites a été étudié par une approche moléculaire. Les amorces utilisées pour amplifier les séquences de gènes *Cytb* et d'ARN 18S sont celles publiées dans l'étude de Loreille *et al.* de 2001. La séquence obtenue pour le fragment du gène *Cytb* est identique à celle obtenue à partir des œufs prélevés sur le coprolithe belge (Oh *et al.*, 2010b).

Une cinquième étude a mis en évidence des œufs d'*Ascaris* sp., *Trichuris trichiura* et *Fasciola* sp. dans des échantillons environnementaux provenant d'un village Viking danois du 9^{ème} siècle AD (Soe *et al.*, 2015). Les chercheurs ont isolé les œufs et ont choisi une nouvelle cible pour la paléogénétique en parasitologie, le gène *Cox1*, en couvrant un fragment de 383 pb avec trois couples d'amorce chevauchants. Cette approche est basée entre autre sur des travaux précédents de cette équipe, qui ont décrit 75 haplogroupes pour ce gène (Betson *et al.*, 2014). Du fait d'un échantillonnage non-exhaustif, lié à la difficulté d'obtenir dans une même région géographique des vers provenant du porc et de l'humain, ces haplogroupes ne recouvrent probablement pas toute la diversité mondiale. Mais ces données permettent toutefois de mieux comprendre la répartition et la diversité génétique actuelle du parasite. Dans le cas des échantillons danois, l'haplogroupe 7 a pu être identifié. Celui-ci se retrouve principalement dans les populations modernes de porcs au Danemark et au Royaume-Uni (Betson *et al.*, 2014), mais aussi chez l'être humain en Chine (Peng *et al.*, 2005). Au vue de l'origine environnementale de l'échantillon et de l'haplogroupe obtenu, aucune conclusion basée sur les analyses génétiques ne peut être portée quant à l'origine porcine ou humaine du parasite retrouvé.

Ces études nous montrent qu'un fragment court du gène *ARNr 18S* ne permet pas de distinguer les espèces d'*Ascaris*. Par ailleurs, Soe *et al.* ont généré 23 nouvelles séquences modernes de ce gène (numéro d'accèsion KM079631-654), et leur analyse a permis de confirmer qu'il est effectivement un mauvais candidat pour étudier la phylogénie et la taxonomie de ce genre. Ces études permettent aussi de constater que, par l'enrichissement des banques de données avec des nouvelles séquences de référence, ce qui pouvait être considéré comme du polymorphisme ancien, serait plutôt attribuable à un manque de représentativité biogéographique des échantillons de références. Par exemple, les variations dans la séquence du gène du *Cytb* obtenu dans l'échantillon belge et coréen correspondent en fait à 100% à la séquence moderne d'*A. lumbricoides* publiée par (Park *et al.*, 2011).

ii. *Trichuris trichiura*

Bien que ce parasite soit observé fréquemment dans les échantillons archéologiques, seulement trois travaux ont permis son identification au niveau moléculaire. Des prélèvements paléoparasitologiques ont été effectués sur une seconde tombe coréenne construite sous la dynastie Joseon, mettant aussi en évidence l'association des parasites *Trichuris* / *Ascaris*. Deux fragments chevauchants de 189 et 187 pb ciblant le gène *ARNr 18S* ont permis de conclure à une similarité de 100% avec les références modernes du parasite *T. trichiura* (Oh *et al.*, 2010a).

Des observations microscopiques sur des échantillons environnementaux provenant d'une fosse à purin localisée dans un village médiéval tchèque ont été menées, et ont mis en évidence la présence d'œufs de plusieurs helminthes gastro-intestinaux : *T. trichiura*, *Ascaris* sp., *Tenia* sp., *Diphyllobothrium latum*, *Dicrocoelium* sp., *Fasciola hepatica*, mais aussi de la famille des Opisthorchiidae et des Heterophyidae (Myskova *et al.*, 2014). Les études génétiques ciblaient principalement le champignon pathogène *Encephalitozoon* ssp., le parasite unicellulaire *Cryptosporidium* ssp. et le protozoaire du genre *Giardia* ssp. mais l'ADN de *Trichuris* fut aussi analysé. La séquence du gène *ARNr 18S* de cet helminthe présentait « une identité » avec la séquence de référence moderne DQ118536, sans mentionner si celle-ci était de 100%.

Les échantillons danois décrits dans la partie précédente ont aussi été analysés génétiquement pour ce parasite (Soe *et al.*, 2015). Un fragment de 281 pb du gène de l'ARNr *18S* a été amplifié en utilisant deux couples d'amorces chevauchant. Après séquençage, il présente 100% d'identité avec des séquences de références modernes de *T. trichiura* et six variations par rapport à *T. suis*.

iii. *Fasciola* sp.

Des fragments d'ADNa correspondant à la grande douve du foie ont été mis en évidence dans les travaux de (Soe *et al.*, 2015) par l'amplification des *ITS1* (*Internal Transcribed Spacer*). L'unique amplicon séquencé présentait une identité de 100% avec *Fasciola hepatica* et deux variations par rapport à *F. gigantica*.

iv. *Enterobius vermicularis*

Les études sur ce parasite ont porté exclusivement sur des coprolithes du continent américain (Brésil, Chili et Etats-Unis d'Amérique) de 4000 BC à 900 AD (Iniguez *et al.*, 2006). Après avoir effectué une PCR nichée sur un fragment de 498 pb de l'espaceur des gènes ARNr *5S*, les auteurs ont mis en évidence, sur deux clones différents, trois variations nucléotidiques absentes des références modernes. Etant donné qu'il n'y a, à ce jour, que sept séquences de références, toutes provenant d'oxyures d'Amérique, il semble impossible de se prononcer sur un variant géographique ou ancien.

v. Autres helminthes gastro-intestinaux

Clonorchis sinensis, aussi appelé douve de Chine, n'a pas été étudié dans le cadre de cette thèse. Il est principalement retrouvé en Extrême-Orient et le corpus analysé ne comporte pas d'échantillon provenant de cette région.

Lors de fouilles archéologiques d'un site chinois de 167 BC, des œufs de cette douve ont été observés dans la vésicule biliaire d'un homme enterré (Liu *et al.*, 2007). Le séquençage de l'*ITS2* n'a pas montré de différence comparativement aux séquences modernes. La séquence de l'*ITS1* comportait quant à elle 15 et 16 variations nucléotidiques par rapport aux séquences chinoises modernes.

Des échantillons du site archéologique coréen de Waegwan, dont les analyses parasitologiques avaient permis de mettre en évidence les parasites *Ascaris* et *Trichuris* (décrit ci-dessus), ont aussi été analysés pour ce trématode, après que ses œufs aient été observés (Shin *et al.*, 2013). La tombe étudiée contenait le squelette d'un homme dont le pedigree était connu et décédé en 1685. Le gène *Cox1* et les *ITS2* et *ITS1* ont été séquencés, les deux premières cibles présentent 100% d'identité avec les séquences de références modernes. Deux séquences différentes pour l'*ITS1* ont été obtenues, provenant probablement de copies génomiques différentes, se distinguant par une insertion et sept délétions nucléotidiques. Chacune de ces séquences correspondent à 100% aux séquences de références modernes présentes dans les banques de données et à 95-96% à la séquences publiée par (Liu *et al.*, 2007).

L'analyse des séquences anciennes et modernes indique un taux de variation intra-spécifique différent entre les *ITS1* et 2. Les travaux de (Lee et Huh, 2004) avait déjà montré que l'*ITS1* de différentes souches moderne de *Clonorchis sinensis* présentait des délétions (jusqu'à 22 nucléotides), ce qui n'était pas le cas pour l'*ITS2* ou l'ARNr18S.

Le dernier cas d'helminthe gastro-intestinal humain présenté est celui du trématode *Echinostoma* sp., dont la morphologie des œufs ne permet pas une détermination de l'espèce. Après observation microscopique d'un coprolithe brésilien retrouvé sur un corps momifié et datant de 560 BP, des œufs du parasite ont été identifiés (Sianto *et al.*, 2005). Les analyses de paléogénétique ont déterminé l'espèce *E. paraensei* en séquençant le gène *Cox1* (Leles *et al.*, 2014). La momie présentait aussi un mégacôlon pouvant évoquer une maladie de Chagas, une infection par le parasite *Trypanosoma cruzi* a aussi été recherchée (voir le point 2- du paragraphe suivant) (Fernandes *et al.*, 2008). Le parasite *Echinostoma* étant actuellement endémique en Asie du Sud-Est, en Extrême-Orient et en Amérique du Sud, il n'a pas été étudié dans le cadre de nos travaux.

Des helminthes gastro-intestinaux n'affectant pas les humains ont aussi été le sujet d'analyses paléogénétiques. C'est le cas de l'étude portant sur les parasites affectant le Moa, un oiseau néozélandais maintenant éteint, qui ont été analysés par une approche de *barcoding* moléculaire (Wood *et al.*, 2013). Des amorces universelles, amplifiant un fragment de l'ARNr 18S des invertébrés, ont été utilisées et ont permis la détection :

- d'Apicomplexa, quelques séquences correspondants à des *Cryptosporidium* spp. et au sous-ordre des Eimeriorina,
- des nématodes de la super famille des Heterakoidae ou de la famille des Trichostrongylidae,
- des trématodes de l'ordre des Echinostomida.

Cette approche permet donc un criblage large des espèces présentes et les résultats pourraient ensuite être utilisés pour cibler plus finement les taxons présents.

vi. Autres parasites

i. Les protistes

Trypanosoma cruzi, causant la maladie de Chagas et affectant les populations des Amériques, est considéré comme un parasite « souvenir » : celui-ci aurait franchi la barrière des espèces lorsque les humains sont arrivés sur ce territoire il y a 13 000 ans (Darling et Donoghue, 2014). Ces organismes de la classe des Kinetoplastida possèdent des kinétoplastes, un ADN organisé en réseau dans une mitochondrie géante. Ces molécules, présentes en nombreuses copies par cellule, ont été ciblées pour identifier *T. cruzi* dans les échantillons archéologiques.

Jusqu'en 2008, la PCR et l'hybridation moléculaire ont été utilisées pour montrer la présence du parasite dans divers échantillons sud-américains datés entre 9000 BP-1400 AD (Guhl *et al.*, 1999 ; Aufderheide *et al.*, 2004 ; Lima *et al.*, 2008). À l'exception du dernier travail cité, les conditions expérimentales incluaient un contrôle positif d'ADN moderne et ne mentionnait pas la présence de laboratoire de confinement normalement dédié à la paléogénétique. La première séquence d'ADN obtenue à partir d'une momie brésilienne (540 AD) présentant un mégacôlon, un des signe de la maladie de Chagas, a permis d'identifier le génotype I (Fernandes *et al.*, 2008), infectant actuellement les animaux et humain du bassin de l'Amazonie.

Le parasite unicellulaire *Leishmania donovani*, une des 21 espèces causant des leishmanioses viscérales chez l'humain, touche les populations des régions tropicales et subtropicales (CDC, 2015). Des échantillons d'Égypte ancienne (3500-2800 BC et du Moyen-Empire de 2050-500 BC) et de Nubie (actuellement la République du Soudan) (550AD-750 AD) ont été analysés par PCR (Zink *et al.*, 2006). Des amplicons ont été obtenus pour 12% des échantillons nubiens et 9% des échantillons du Moyen-Empire, résultats confirmés par séquençage. Du fait du commerce entre les deux pays, de la présence du pathogène en Nubie à l'ère chrétienne et de l'absence de celui-ci avant le Moyen-Empire, les auteurs ont conclu que le parasite aurait été introduit en Égypte par les Nubiens. Cette hypothèse pourrait être confirmée par la mise en évidence d'ADN de *Leishmania* en Nubie dès 2050 BC.

Le dernier exemple de parasite unicellulaire abordé est celui de l'agent du paludisme ou malaria, *Plasmodium* sp., un protozoaire infectant les hépatocytes et les érythrocytes. Cette zoonose cause environ 500 000 décès par an (CDC, 2015). Des chercheurs ont analysé des échantillons d'origine égyptienne de 4000 ans. Après avoir obtenus de nombreux amplicons de tailles différentes et des séquences non-spécifiques pour l'ARNr *18S* (cible classique pour étudier la malaria) les auteurs ont analysé la séquence du gène *Pfcr* (*Plasmodium falciparum* Chloroquine Resistance Transporter). Parmi les 91 échantillons testés, seuls deux se sont avérés positifs (2,2%) (Nerlich *et al.*, 2008). Ce résultat soulève des doutes quant à la spécificité des tests immunologiques effectués sur d'autres échantillons égyptiens de la même époque, où 40% des momies analysées et 92% des momies présentant des signes osseux d'anémie furent déclarées atteintes de malaria (Miller *et al.*, 1994 ; Rabino Massa *et al.*, 2000).

Finalement, une approche de type métagénomique appliquée à des momies égyptiennes (806 BC-124 AD) et des squelettes boliviens précolombiens a montré une co-infection par *Plasmodium* sp. et par *Toxoplasma* sp. chez une des momies (Khairat *et al.*, 2013).

c) *Pediculus humanus*

Cet arthropode hématophage est un ectoparasite dont le cycle de vie se déroule au niveau des téguments ou de la chevelure humaine. Des momies péruviennes de 1025 AD ont été retrouvées avec de longs cheveux tressés sur lesquels des poux ont été identifiés (Raoult *et al.*, 2008). L'ADN a été extrait et les fragments des gènes du *Cytb* et *Cox1* ont permis de définir un phylotype A, que les auteurs ont pu comparer avec sa répartition actuelle.

IV. Bilan du chapitre III

L'étude des helminthes gastro-intestinaux affectant les populations anciennes s'est développée au cours du 20^{ème} siècle. Par l'utilisation de la microscopie optique, les paléoparasitologues peuvent identifier divers taxons par l'observation de certaines caractéristiques morphologiques et micrométriques, plus ou moins spécifiques aux familles et genre de parasites. Il est par exemple impossible de distinguer les genres *Echinococcus* et *Taenia*, de la famille des Taeniidae, ou les diverses espèces de *Taenia* pouvant infester les humains (*T. saginata*, *T. solium* et *T. asiatica*).

Au début du 21^{ème} siècle, les développements méthodologiques et technologiques dans le domaine de la biologie moléculaire et de la paléogénétique permettent de préciser les diagnostics établis par la microscopie classique. En comparant des séquences d'ADNa à des séquences modernes, il est possible de distinguer la plupart des taxons au niveau des espèces. Cette approche est soumise à la présence dans les banques de données d'une quantité suffisamment importantes de séquences de références, provenant de plusieurs régions géographiques, pour pouvoir établir des alignements génétiques comparatifs pertinents. Ce n'est pas encore le cas pour de nombreux genre d'helminthes gastro-intestinaux.

*PARTIE 2 : MATÉRIEL ET
MÉTHODE*

Le laboratoire Chrono-environnement héberge l'unique équipe française travaillant sur la paléoparasitologie, animée par Matthieu Le Bailly. Il a pu, au fil des années, entrer en contact avec des nombreux archéologues et anthropologues, qui lui ont confié des vestiges. Ses travaux ont donné lieu à de nombreuses publications où les œufs d'helminthes gastro-intestinaux sont décrits par microscopie et mis en relation avec le contexte historique (Le Bailly *et al.*, 2006 ; Le Bailly *et al.*, 2012 ; Nezamabadi *et al.*, 2013b). De part une volonté d'amener la discipline à évoluer avec les techniques disponibles au 21^{ème} siècle, ce projet de thèse a été mis en place grâce à des soutiens locaux et régionaux. Les travaux de génétique réalisés s'inscrivent dans une optique de valorisation du corpus conservés au laboratoire. Ils représentent une occasion unique d'analyser différents types d'échantillons (sédiments, coprolithes issus de latrines, de couches d'occupation, de fosse à purin...), provenant de différentes régions géographiques et sur une période chronologique large (du Néolithique jusqu'à la Première Guerre Mondiale).

Ces travaux ont été réalisés grâce à l'expertise de biologie moléculaire et de paléogénomique de l'équipe « Épigénome et Paléogénome » de l'Institut Jacques Monod, Université Paris Diderot, dirigée par les Dr. Eva-Maria Geigl et Thierry Grange, et aux nouvelles techniques qui y ont récemment été développées. Leur accueil m'a permis d'acquérir de nombreuses compétences dans le domaine de l'ADNa et de côtoyer les autres chercheurs et étudiants de leur équipe, élargissant ainsi ma formation et mes connaissances en génétique et paléogénétique.

Dans cette partie, nous décrirons dans un premier chapitre les sites archéologiques pour lesquels des vestiges ont fait l'objet d'analyses génétiques. Lorsque ces informations sont disponibles, nous mentionnerons le contexte de la fouille, les détails du site, les modalités de prélèvements et les taxons observés par microscopie. Les observations par microscopie ont été effectuées par M. Le Bailly et les étudiants de son équipe, selon la technique décrite dans (Dufour et Le Bailly, 2013).

Dans le cadre de ce projet sur la paléoparasitologie moléculaire, les échantillons archéologiques étudiés sont de nature sédimentaire (coprolithes, sédiments sous des squelettes, provenant de couches d'habitation, ou de structures en creux de type fossés, puits, latrines...). La préservation des biomolécules dans ce contexte sera décrite dans le second chapitre. La méthode de génotypage « *aMPlex Torrent* » qui a été développée au sein de l'équipe « Épigénome et Paléogénome » pour l'étude d'échantillons archéologiques de composition complexe sera aussi présentée.

Les expériences préliminaires sur les échantillons de paléoparasitologie seront décrites et les résultats, montrant une possibilité d'hétérogénéité dans la détection des taxons, seront expliqués, dans le but d'introduire les adaptations de la méthode « *aMPlex Torrent* » dans ce contexte particulier. Celles-ci ont donné lieu à la publication d'un article scientifique (Côté *et al.*, 2016), qui sera présenté dans un dernier chapitre.

Chapitre I : Les sites archéologiques

I. Les sites préhistoriques

A. *La Draga, Espagne*

Ce site lacustre du Néolithique ancien (cardial final, 7250-7050 BP) est localisé autour du lac de Banyoles, en Catalogne (Figure 22). Il fut découvert en 1990 lors de l'aménagement d'un parc à proximité du plan d'eau. Le site s'étend sur 8 000 mètres carrés, mais seulement 284 mètres carrés ont été fouillés, dont une partie se trouvait sous le niveau de l'eau. Le site était inondé depuis le Moyen-Âge, des pompes ont été utilisées pour abaisser le niveau des eaux et permettent le travail des archéologues. Le sol archéologique, d'une épaisseur de 20 à 30 centimètres, est recouvert d'une tourbe qui s'est formée après l'abandon du site. Ces points ont contribué à la bonne préservation du matériel archéologique puisque le contexte humide de l'habitat a favorisé la préservation de divers restes organiques, tels des végétaux, des charbons, de la faune et des outils (Tarrus *et al.*, 1994).



Figure 22 : reconstitution du site de La Draga (Photo tirée de <http://www.arqueoxarxa.cat>).

1) *Le contexte archéologique*

Les fouilles ont montré la présence de 30 foyers, deux plateformes dallées, une zone de dépotoir et plusieurs trous de poteaux. La disposition des cabanes, par l'étude de la composition des bois des poteaux, n'avait pas encore été étudiée au moment de la publication du rapport de fouille. Les informations concernant leur répartition au sein du site et la proximité avec le lac sont donc manquantes.

L'étude des foyers a permis d'émettre l'hypothèse de la torréfaction des céréales et la cuisson d'aliments, incluant de la viande, car des restes de faunes y ont été découverts. Les habitants de La Draga étaient des agriculteurs et pasteurs. Ils cultivaient diverses céréales (trois types de blé et de l'orge) mais aussi des légumineuses (fèves, petits pois) et des baies provenant d'arbres. Une quantité importante de restes animaux (9 410 ossements) ont été retrouvés, dont 96% des restes déterminés sont attribuables à des animaux domestiques. Un canidé domestique a été identifié, les autres ossements correspondent à des bœufs, moutons, chèvres et porcs, en des proportions similaires. Plusieurs espèces d'animaux sauvages sont aussi présentes en nombre plus restreint : aurochs, sangliers, divers cervidés, renards, loups, blaireaux et lapins. La chasse représente un faible part dans les activités des habitants du site.

L'étude des profils d'abattage indique que les animaux domestiques étaient pour la plupart tués au moment de leur optimum au niveau de la quantité et de la qualité de leur viande, signifiant que l'alimentation était la principale raison de leur élevage. Par ailleurs, des chèvres et moutons âgés ont aussi été retrouvés, suggérant qu'ils pouvaient être source de produits autres que la viande, telle le lait et la laine. Finalement, les analyses paléopathologiques ayant mis en évidence des signes de lésions osseuses chez les bœufs, ceux-ci ont pu aussi être employés pour des travaux de ferme (Antolín *et al.*, 2014).

2) *Les résultats de paléoparasitologie microscopique*

Trois cents échantillons de sédiments provenant des couches d'habitations émergés ont été envoyés au laboratoire Chrono-environnement. Sur l'ensemble des vestiges, principalement des œufs d'*Ascaris* sp. et de *Trichuris* sp. ont été identifiés. Quelques œufs de *Capillaria* sp., de *Paramphistomum* sp. et de la famille des Taeniidae ont aussi été mis en évidence.

B. Arbon-Bleiche III, Suisse

Ce site est situé sur le pourtour du lac de Constance, un ensemble de plans d'eau dans le Nord des Alpes où plusieurs sites préhistoriques sont regroupés. Les sites lacustres néolithiques présentés sur la Figure 23 ont fait l'objet d'une analyse paléoparasitologique. Ceux-ci présentent un intérêt archéologique et anthropologique important du fait qu'ils ont été le lieu d'une transition entre deux groupes de cultures différentes, les Pfyn et les Horgen. Par ailleurs, cette transition est contemporaine à une dépression climatique, géologiquement répertoriée, appelé le Piora 2. Celle-ci se caractérise par un refroidissement rapide du climat et une augmentation des précipitations.

Les sites du lac de Constance offrent donc une opportunité rare d'étudier au niveau paléoparasitologique :

- la rencontre entre deux populations, entraînant l'échange de parasites inexistantes dans l'une ou l'autre des populations,
- l'impact d'un changement environnemental sur l'accessibilité aux ressources naturelles : une population qui augmente sa consommation de poisson ou la pratique de la chasse des animaux juvénile pour compenser les pertes céréalières due aux chutes des températures, ne sera pas infectée par les mêmes parasites (Le Bailly *et al.*, 2007).

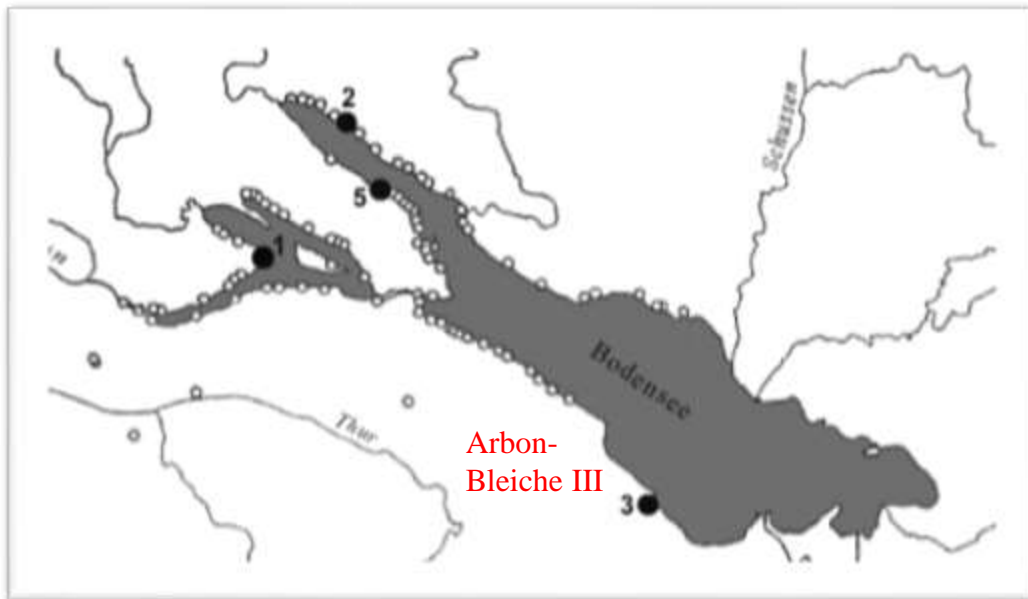


Figure 23 : le lac de Constance (Bodensee). Les divers sites néolithiques étudiés en paléoparasitologie sont indiqués. 1) Hornstaad-Hörnle I, 2) Sipplingen, 3) Arbon-Bleiche III, 4) Torwiesen II, 5) Wallhausen-Ziegelhütte, 6) Seekirch-Stockwiesen. Les points blancs sont d'autres sites préhistoriques de la région (Le Bailly et al., 2007).

1) Le contexte archéologique

Dans le cadre de cette thèse, seul le site d'Arbon-Bleiche III a été étudié. Ce village suisse, localisé au sud du lac de Constance, fut occupé de 3384 à 3370 BC. Ce site, dont l'occupation n'a duré que 15 ans, correspond à une période de co-existence entre les cultures néolithiques Pfyn (3900-3380 BC) et Horgen (3380-2900 BC). Il fait par ailleurs partie d'un des cinq villages contemporains à Ötzi, l'Homme des glaces dont les restes fossilisés ont été retrouvés à 100 km au Sud-Est.

Des restes de 27 habitations bien préservées sous l'eau ont été retrouvés sur une zone d'environ 1100 m². De nombreux coprolithes, conservés dans les couches d'habitation, supposés être d'origine anthropique par l'absence d'esquilles d'os, ont été récupérés. L'immersion a permis la préservation de céramiques, dont un des usages suggérés étaient la préparation et la conservation des aliments. Une croûte de matière organique a pu être observée dans certaines de ces céramiques, formée de restes de céréales et d'autres plantes, ainsi que des ossements de poissons.

L'analyse de la faune présente sur le site a révélé la présence d'animaux domestiqués (vaches, cochon, chèvres moutons et chiens) et sauvages (ours, sangliers, grenouille et cerfs). La faune aquatique se compose quant à elle de Cyprinidae, de Salmonidae, de brochets et de perches. Les plantes sont majoritairement représentées par des céréales (blé et orge) mais aussi par du lin et pavot à opium et divers plantes sauvages (noisettes, framboises, prunes et mûres). L'alimentation des animaux domestiqués a aussi pu être décrite et se composait de sapin, de mûres, d'aulnes et de noisetiers (Spangenberg *et al.*, 2006).

2) *Les résultats de paléoparasitologie microscopique*

Les analyses microscopiques de l'ensemble des sédiments et des coprolithes provenant des couches d'habitation, ensevelies sous l'eau depuis plusieurs milliers d'années, ont mis en évidence la présence de *Fasciola* sp., *Tænia* sp., *Capillaria* sp., *Trichuris* sp., *Diphyllobothrium* sp., *Dioctophyma* sp. et *Opisthorchis* sp. (Le Bailly *et al.*, 2007). Ces trois derniers helminthes se contractent après consommation de poissons crus d'eau douce.

C. Passel, France

Ce site situé dans l'Oise, constitué d'un immense monument architectural composé d'une enceinte, est daté entre 6250 et 5650 BP.

1) *Le contexte archéologique*

Ce site important, sous influence des groupes culturels néolithiques du Chasséen septentrional et du Michelsberg, expose la complexification sociale de l'époque. En effet, l'édification de cette construction a sûrement été le fruit de la collaboration de plusieurs villages. Sa fonction n'est pas connue pour l'instant, mais elle a pu être érigée dans un but défensif, commémoratif, religieux ou politique. L'enceinte est composée d'une palissade et de quelques fossés. Le mobilier détritique est abondant, composé de restes de céramiques, de silex et de faunes. Les dépôts intentionnels sont quant à eux des vases, des ossements humains et animaux (bovidés, cervidés) (Inrap, 2015).

2) *Les résultats de paléoparasitologie microscopique*

Quatorze échantillons de sédiments provenant du fossé périphérique ont été prélevés. Les analyses de paléoparasitologie ont permis d'identifier des œufs de *Trichuris* sp. et Tœniidae, qui sont les deux taxons les plus abondants (entre quatre et 164 œufs). Quelques œufs de *Capillaria* sp., *Paramphistomum* sp. et *Dicrocoelium* sp., avec moins de quatre œufs par échantillons ont été identifiés.

D. Menneville, France

Menneville est une petite ville de Picardie, dans la vallée de l'Aisne. Cette vallée a fait l'objet de fouilles amatrices et préventives depuis plusieurs siècles. En 1764, des ouvriers firent la découverte aux alentours de Soissons de huit squelettes humains. Après la Deuxième Guerre Mondiale, des sablières s'implantent dans la vallée de l'Aisne, entraînant la découverte de nombreux sites archéologiques et laissant présager de la richesse et de l'importance de la région. À partir de 1962, la prospection aérienne est utilisée pour repérer diverses structures, principalement des fossés circulaires. S'en suivent différentes fouilles programmées de sauvetage, concernant des sites néolithiques à Cys-la-Commune ou Vailly-sur-Aisne ou une nécropole gauloise à Pernant. A partir des années 70, les archéologues Soudsky et Boureux ont souhaité développer des chantiers-écoles pour le Néolithique et ont repéré par prospection aérienne plusieurs sites de périodes différentes dans la vallée (Bailloud *et al.*, 1982)(Figure 24).

Une collaboration entre les responsables du site archéologique de Menneville et l'équipe « Épigénome et Paléogénome » m'a permis de rencontrer l'archéo-anthropologue Corinne Thevenet. Suite à une discussion concernant mon intérêt pour la transition Néolithique et les échanges de parasites Hommes ↔ Animaux, elle m'a proposé de participer aux fouilles archéologiques de ce site. J'ai donc eu l'opportunité d'acquérir des compétences en archéologie sur site et de prélever, en collaboration avec les archéologues et archéozoologues responsables, des échantillons pour mon travail de thèse. Ce chantier accueillait par ailleurs de nombreux étudiants en archéologie, ce fut l'occasion de les sensibiliser aux problématiques de la paléogénétique, au niveau de l'échantillonnage, de la conservation des prélèvements et des possibilités de contaminations par de l'ADN moderne.

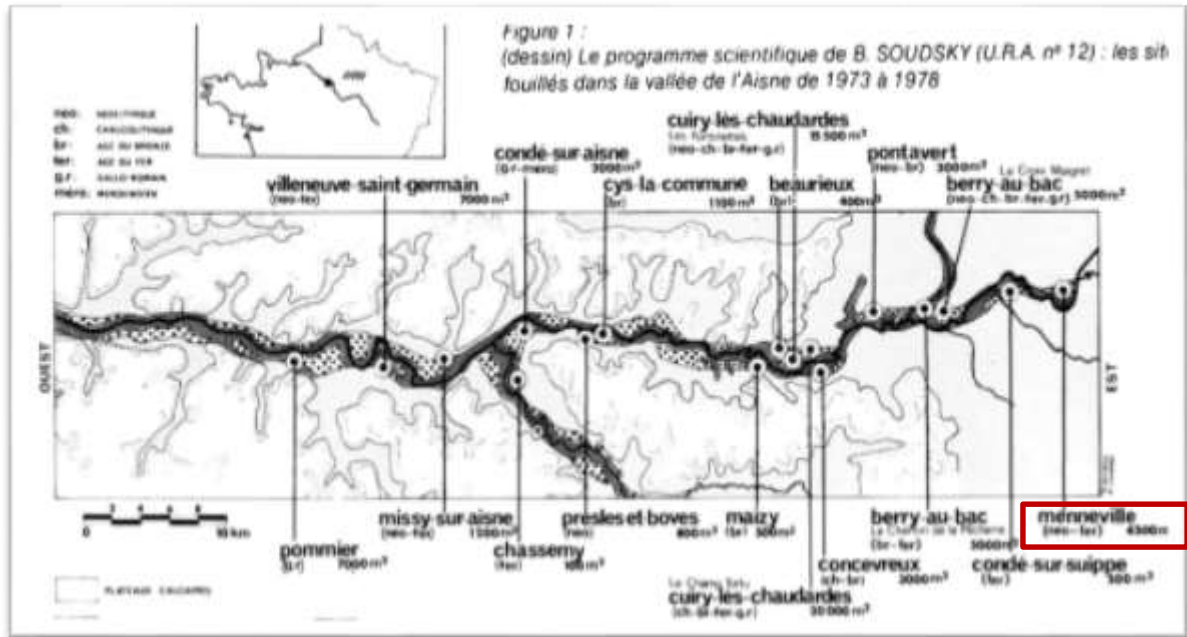


Figure 24 : carte des sites archéologiques de la Vallée de l’Aisne fouillés depuis les années 1970 (Bailloud et al., 1982). Le site de Menneville est encadré en rouge.

1) Le contexte archéologique

Le site de Menneville, lieu-dit Derrière le Village, est un village néolithique du rubané récent (5500-4700 BC) entouré d’un fossé à fond pointu, garni de quelques poteaux. Les fouilles de la fin des années 1970 et celles des années 2000 ont mis en évidence huit maisons, 17 sépultures et plusieurs squelettes animaux dans le fossé (Coudart et Demoule, 1982). Les sépultures, localisées sur les bords sud des maisons, représentent principalement des enfants. Le fossé d’enceinte de 1020 m de circonférence et d’une profondeur allant jusqu’à 1 mètre 80, contenait 15 inhumations. Des restes de faune, composés de bovidés, d’ovins ou de caprins, leurs étaient associés. Les os de bovins étaient présents en morceaux débités, pouvant représenter des restes de repas, alors que ceux des caprinés étaient des squelettes entiers en connexion, correspondant plutôt à un geste de dépôt. Les restes de faunes associés aux maisons étaient différents de ceux retrouvés dans les fossés : vaches, taureaux, moutons, chèvres, porcs (animaux domestiqués), cerfs, chevreuil, sangliers, aurochs (faune sauvage) (ASAVA, 2014).

Au cours de l’été 2014, une sépulture contenant un squelette féminin a été découverte dans le fossé entourant le village (Figure 25). J’étais présente sur le site lorsque C. Thevenet a prélevé plusieurs grammes de sédiments à l’intérieur du bassin et un contrôle, autour du crâne.

L'échantillonnage a eu lieu lorsque les os étaient encore emboîtés dans le sédiment. Lors du dégagement du squelette et la poursuite de la fouille sur la région connexe, des restes de faune ont été mis en évidence. Des ossements de bœufs, de porcs et de moutons ou chèvres ont été identifiés par l'archéozoologue Lamys Hachem. Le porc était localisé sous le crâne du squelette humain.



Figure 25 : une sépulture humaine de Menneville montrant un squelette entouré d'ossement animaux (sous les pieds, sous les genoux et au niveau du rachis cervical) (crédit photo : UMR 8215, Trajectoires, Maison Archéologie & Ethnologie).

2) Les résultats de la paléoparasitologie microscopique

Les analyses microscopiques ont permis d'identifier des œufs d'ascaridés minéralisés provenant de l'échantillon prélevé sous le crâne. Les sédiments provenant du bassin ce sont avérés négatifs.

E. Les mines de sel de Chehr Abad, Iran

Cette région, localisée sur les plateaux Nord-Ouest de l'Iran, est située à proximité d'une route d'échange avec les plateaux centraux. Entouré de buttes et steppes semi-arides de climat continental, le site a montré une bonne préservation des restes organiques (momies, coprolithes) ainsi que des textiles.

1) Le contexte archéologique

Entre 1993 et 2004, des mineurs travaillant dans une zone du village de Hamzehlu de la province de Zanzan (Figure 26) ont retrouvé deux momies humaines, ainsi que différents outils de travail, incluant des bottes de cuir, des couteaux et une meule à pierre. Les fouilles archéologiques ont débuté en 2005 et ont entraîné la découverte de momies supplémentaires. Elles correspondent à trois individus de sexe masculin, qui seraient décédés lors de l'effondrement des galeries de la mine de sel où ils travaillaient (Figure 27). Cette mine a été intensément exploitée pendant la moitié du premier millénaire BC. Les analyses au carbone 14 effectuées sur les vêtements suggèrent que les individus seraient décédés entre 2286-1545 BP (Ramaroli *et al.*, 2010). Au total, deux momies proviendraient de la période Sassanide et trois de la période Achéménide.



Figure 26 : la localisation du site de Chehr Abad , encadré en rouge (Ramaroli *et al.*, 2010).



Figure 27 : la momie numéro 5 du site de Chehr Abad (photo A. Aali).

La découverte de ce site a soulevé des questions fondamentales quant à l'exploitation du sel, lié au contexte socio-économique de l'époque et aux territoires participant à ce type de commerce.

2) *Les résultats de la paléoparasitologie microscopique*

Un échantillon provenant de la momie numéro 5 datée de 2286 +/- 28 BP a été obtenu à partir de sédiments attachés au squelette au niveau des os pelviens et des têtes fémorales. Trois autres échantillons de sédiments du site d'occupation de la mine ont aussi été prélevés et sont datés de 2500-1500 BP.

Des œufs de Taeniidae ont été mis en évidence dans ces quatre prélèvements. Du fait de l'association d'au moins un des prélèvements à une momie humaine, il est possible de préciser qu'ils appartiennent au genre *Taenia* sp. Cependant, il n'est pas possible de déterminer l'espèce. Pour les échantillons provenant des couches d'occupation, des œufs d'*Enterobius vermicularis*, d'*Ascaris* sp., de *Trichuris* sp. et un œuf du parasite équin *Oxyuris equi* ont été identifiés (Nezamabadi *et al.*, 2013b).

II. Les sites antiques en France

A. *L'Amphithéâtre de Metz*

1) *Le contexte archéologique*

La ville de Metz s'est construite sur les vestiges de *Divodurum*, dont l'occupation débute au 1^{er} siècle avant notre ère (Figure 28). Les travaux associés à l'extension de la gare SNCF ont amené la prescription de fouille de ce site archéologique de 11 hectares, montrant des signes d'occupation de l'Antiquité jusqu'à l'époque moderne. La zone de l'Amphithéâtre, fouillée entre 2006 et 2008, s'étend sur 17 000 mètres carrés.

Datant de la fin du 1^{er} siècle de notre ère, l'Amphithéâtre de Metz, un des plus grands de l'époque puisqu'il peut accueillir 25 000 spectateurs, est localisé dans la partie sud de la ville antique, légèrement en périphérie du centre. Il serait entre autre associé à des activités funéraires, puisque des tombes de plomb dans des sarcophages en pierre ont été retrouvées. Après les attaques par les Alamans en 253, la ville se fortifie en laissant l'Amphithéâtre extra-muros. L'histoire antique de la ville se termine en 451 par les attaques des Huns.

Des fosses du 2-4^{ème} siècle, à visée détritique, montrent une accumulation de faune, principalement des bœufs et des animaux sauvages. Un fossé beaucoup plus profond, probablement à visée défensive, est retrouvé au sud de l'Amphithéâtre.



Figure 28 : la maquette de la cité Divodurum. L'Amphithéâtre est visible à gauche.
(Source inconnue, disponible sous <http://www.miroirdutemps.fr>).

2) Les résultats de paléoparasitologie microscopique

Un coprolithe retrouvé dans un puits a été analysé par microscopie. Il provient d'un mammifère contaminé par du *Diphyllobothrium* sp. et du *Fasciola* sp. Les sédiments échantillonnés sur un autre puits présentaient des œufs d'*Ascaris* sp., *Trichuris* sp. et *Tenia* sp.

B. Jeu de Paume, Beauvais

1) Le contexte archéologique

Le quartier du Jeu de Paume de la ville de Beauvais fut découvert lors des fouilles menées en prévision de la construction d'un centre commercial (Figure 29). Ce site, dont l'occupation remonte au 1^{er} siècle de notre ère et s'est poursuivie pendant 300 ans, était fortement structuré autour des bâtiments publics. Les habitations, en bois puis en pierre, étaient munies d'un puits. Ces derniers, une fois asséchés, étaient utilisés pour entasser des déchets. Des coquilles d'huitres y ont été retrouvées, attestant de certaines habitudes alimentaires. Associé aux maisons était construit un système d'évacuation des eaux usées.

Parmi les activités artisanales, notons la production de colle à partir de collagène bovin, révélée par la présence d'esquilles d'os (Fémolant, 1999 ; Desachy, 1991).

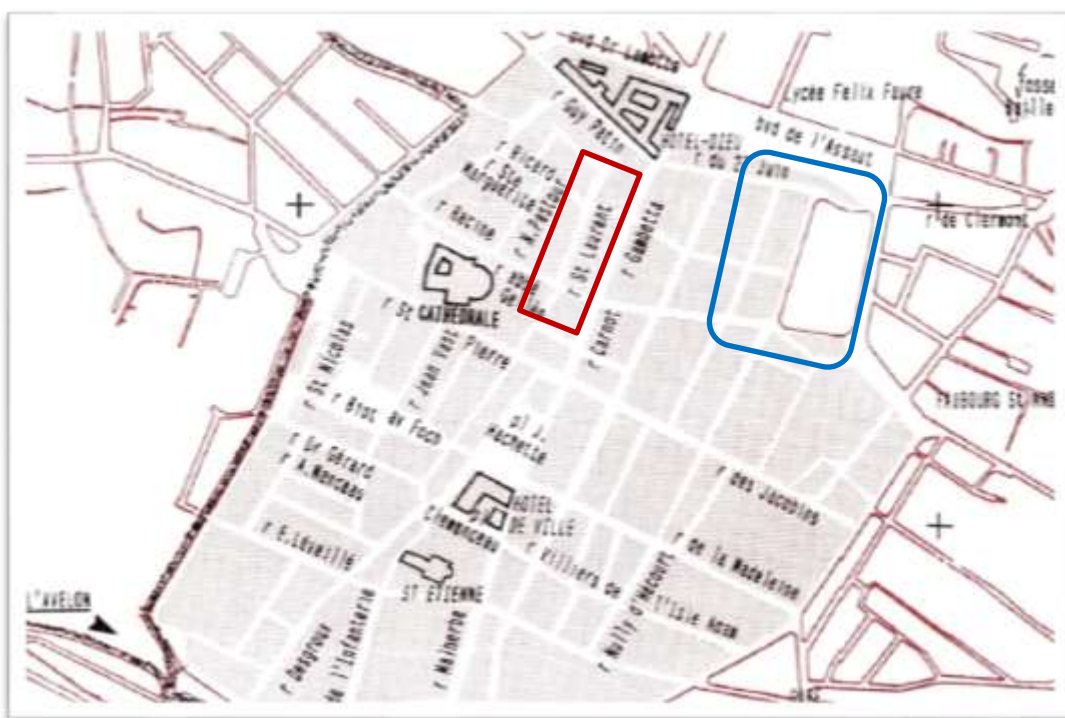


Figure 29 : les sites de Jeu de Paume (en bleu) et rue Saint-Laurent (site médiéval, en rouge), Beauvais (Desachy, 1991).

2) Les résultats de la paléoparasitologie microscopique

Les analyses microscopiques ont montré que les sédiments provenant des latrines contenaient des œufs d'*Ascaris* sp., *Trichuris* sp., *Capillaria* sp., *Dicrocoelium* sp. et *Fasciola* sp.

C. Mas de Vignoles, Nîmes

1) Le contexte archéologique

Des opérations d'archéologie préventives ont mis à jour une occupation du Néolithique ancien jusqu'au Moyen-Âge. Du mobilier a été retrouvé dans des puits comblés, comprenant entre autre des tessons et des ossements animaux. La section du site analysé dans le cadre de ce projet est datée de 600 AD par le matériel céramique. Des fossés délimitant des petites parcelles sont présents sur le site et témoignent de sa vocation agraire. Des traces de la présence de cultures céréalières, de chanvre et d'élevage de bovins, porcins et ovo-caprins ont été retrouvées (Tardy, 2015).

2) Les résultats de la paléoparasitologie microscopique

Des œufs d'*Ascaris* sp., *Trichuris* sp., *Capillaria* sp., *Dicrocoelium* sp., *Paramphistomum* sp. et *Fasciola* sp. ont été mis en évidence dans le sédiment provenant des puits du site.

III. Les sites médiévaux français

A. *Rue Saint-Laurent, Beauvais*

1) *Le contexte archéologique*

A la fin du haut Moyen-Âge, la ville de Beauvais prend de l'ampleur par l'ingérence des Rois Philippe Ier et Louis V dans la gestion de la frontière normande et la lutte contre les vassaux. Elle obtient le statut de « Commune » en 1122, ce qui engendra des tensions par rapport aux seigneurs, car les habitants sont soumis au droit féodal. Malgré tout, Beauvais est une ville riche de par sa production commerciale, incluant la vente de draps. La vie s'organise autour des paroisses, celle de St-Laurent étant une des premières mises en place.

Lors des fouilles archéologiques, quelques habitats ont été recensés, fabriqué en bois et en torchis sur un soubassement de pierre. Il subsiste aussi des dépotoirs et fosses d'aisance, des latrines ont été retrouvées rue des Jacobins. Les malades sont pris en charge à l'Hôtel-Dieu, implantée près de la porte St-Laurent (Figure 29) (Fémolant, 1999).

2) *Les résultats de la paléoparasitologie microscopique.*

Les deux latrines (12-14^{ème} siècle) analysées ont mis en évidence une quantité importante d'œufs de *Trichuris* sp., *Ascaris* sp., *Dicrocoelium* sp., *Tenia* sp. et *Fasciola* sp.

B. Place des Véens, Auxerre

1) Le contexte archéologique

L'îlot nord de la Place a été fouillé de manière préventive en 2010, en amont d'un projet immobilier. L'occupation du site débute au 3^{ème} siècle de notre ère et est marquée par des activités de production de chaux. Au Moyen-Âge, ces structures sont abandonnées et sur leur ancien emplacement sont creusés deux fossés. Un premier servira à l'écoulement des eaux alors que le second deviendra un plan d'eau stagnante alimenté par des eaux usées. Celui-ci a eu une fonction de merderon, de statut public, entretenu et utilisé par toute la communauté. Les analyses carpologiques du site montre une accumulation de matière organique provenant d'activités artisanales de confection du vin, mais aussi de rejet provenant de latrines, mélangeant rejets d'origine animale et humaine. Cette fosse ferait donc partie d'un véritable plan d'aménagement sanitaire de la ville, dirigée par les autorités locales (Aumard *et al.*, 2013).

L'analyse des dépôts de faune montre que ceux-ci se déroulent en deux phases : aux 12-13^{ème} siècles, des vaches sont retrouvées, puis au 14^{ème} siècle, des bœufs sont majoritairement recensés. Cela peut s'expliquer par la révolution technique agricole entraînant une augmentation de la superficie des champs cultivés, et l'utilisation intensive des bœufs pour les travaux. Le 14^{ème} siècle semblerait être aussi être la période d'introduction de la charcuterie, qui se distingue de la boucherie par l'utilisation des matières grasses et abats (Mercier, 2013).

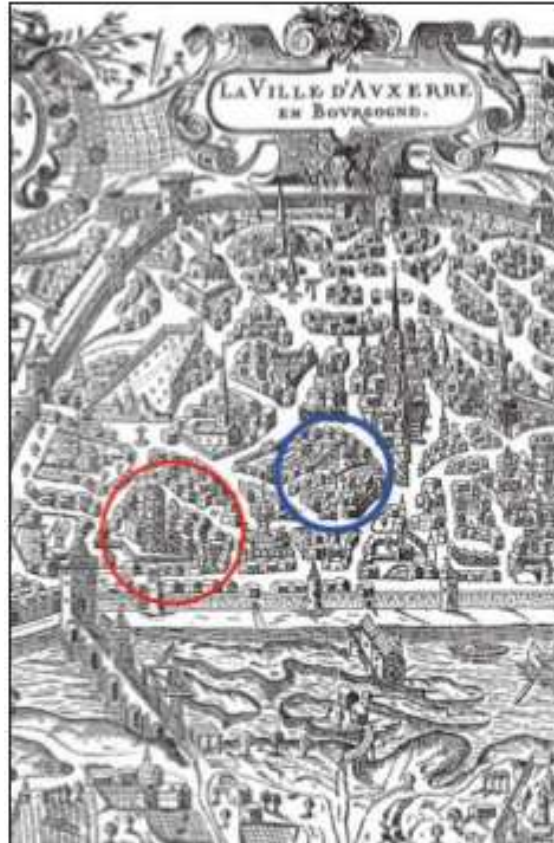


Figure 30 : le site d'Auxerre. La Place des Véens est indiquée en bleue (Aumard et al., 2013).

2) Les résultats de la paléoparasitologie microscopique

L'analyse paléoparasitologique par microscopie des sédiments provenant de deux puits datant des 11-12^{ème} siècles ont montré la présence d'œufs de *Trichuris* sp., *Ascaris* sp., *Dicrocoelium* sp., *Tænia* sp. et *Fasciola* sp.

C. *Les Turqueries, Calais*

1) *Le contexte archéologique*

Les fouilles de ce site, débutées en 2012, ont été réalisées par la communauté d'agglomération de Cap Calaisis. Ce site, dont l'occupation débute au Premier Moyen-Âge (8-10^{ème} siècle), est localisé sur un ancien marais. Initialement, les archéologues ont émis l'hypothèse d'un gisement conchylicole mais après analyses, il présenterait plutôt des aspects liés à des activités agro-pastorales. La question d'une consommation locale de coquillages ou de leur commercialisation fait partie des hypothèses qui sont explorées.

Parmi les 500 structures mises à jour, aucune trace d'habitation n'est retrouvée, uniquement des enclos associés à des fosses, liées à l'activité du site. Outre un abondant mobilier (céramiques, meule, pilier de four en terre cuite ...) des restes de faunes ont été découverts. Le site est orienté vers l'élevage bovin et à sa transformation. Les types de coquillage retrouvés sont plutôt associés à une culture d'eau douce (Moriceau, 2015).

2) *Les résultats*

Treize coprolithes ont été prélevés pour les analyses paléoparasitologiques et paléogénétiques. L'origine humaine n'est pas confirmée, car morphologiquement, ils peuvent être confondus avec des coprolithes canins. Les parasites identifiés sont *Ascaris* sp., *Trichuris* sp., *Capillaria* sp. et *Dioctophyma* sp. Ce dernier taxon peut être contracté par tout mammifère consommant du poisson, mais le fait qu'il ait été identifié sur un coprolithe avec l'association typique *Ascaris* / *Trichuris*, penche pour une origine humaine. On ne peut cependant pas exclure, de par leur comportement coprophage, qu'un chien, contaminé par *Dioctophyma*, ait mangé des fèces humaines contaminées avec les deux précédents parasites.

D. Teteghem, France

L'étude de ce site médiéval s'inscrit dans un contexte différent des autres cas exposés. Le colloque du 20th *European Paleopathology Association meeting* en Suède (26-30 août 2014) fut l'occasion de rencontrer l'archéozoologue Annelise Binois (Université de Paris I Panthéon-Sorbonne), avec qui j'ai discuté d'une problématique associée à la paléoparasitologie dans le contexte de sa thèse de doctorat.

1) Le contexte archéologique

Téteghem est une commune du Nord-Pas-de-Calais, et le lieu-dit de Carline 3 a fait l'objet d'une prescription de fouille préventive par l'Inrap avant la construction de lotissement d'habitations. Le diagnostic archéologique établi en 2008 a révélé la présence d'une occupation rurale aux 13-14^{ème} siècles. Un ensemble de bâtiments, associés à des celliers abandonnés, a été retrouvé en lien avec des fosses contenant des squelettes de chevaux, de vaches et de moutons. Les fouilles ont permis d'établir une chronologie de l'occupation agricole du site, qui a débuté vers le 12^{ème} siècle, avec une intensification des pratiques (exploitation des tourbes, argiles, élevage) au cours du 14^{ème} siècle, puis d'un abandon du site au siècle suivant (Direction Générale des Patrimoines, 2013).

Une fosse contenant six moutons enterrés lors d'un dépôt unique a été découverte, pouvant indiquer une *mortalité de masse* (Figure 31). Cette expression regroupe des notions « d'augmentation rapide, soudaine et transitoire du taux de mortalité dans une population donnée, animale en l'occurrence. Le résultat en est naturellement le décès brutal de plusieurs individus sur un temps relativement court » (Binois, sous-presse).

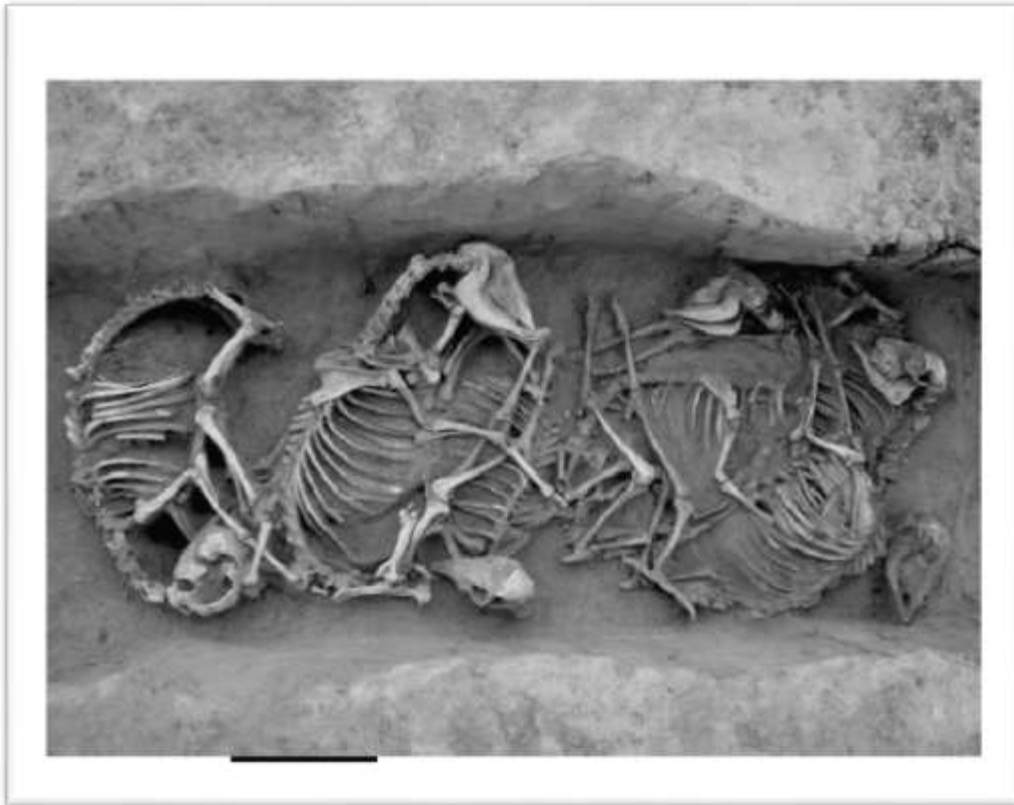


Figure 31 : la fosse contenant les squelettes de moutons, Tétéghem (Photographie A. Binois).

Plusieurs causes peuvent être évoquées dans le cadre d'une mortalité de masse :

- Les intempéries sont les causes les plus fréquentes de mortalité de masse chez les ovins. Les hivers froids et rigoureux affectent principalement les agneaux.
- Les épizooties : une clavelé ou variole ovine (due au virus *Variola ovina*), *Fasciola hepatica* ou des maladies du charbon (causée par les bactéries *Bacillus anthracis* ou *Clostridium chauvei*) sont les principaux pathogènes pouvant être mis en cause.
- Les attaques de prédateurs, les loups pouvant tuer plusieurs moutons sur une courte période, ou causer chez les animaux un état de panique tel qu'ils décèdent par asphyxie en se regroupant.
- Les causes accidentelles : foudroiement, incendies, empoisonnement...

Une étude par critères a été appliquée au site de Tétéghem et lui a permis de proposer l'hypothèse d'une mortalité liée à une infection par *Fasciola hepatica*, que nous avons décidé de tester dans le cadre de nos travaux de paléogénétique.

2) *Les résultats de la paléoparasitologie microscopique*

Plusieurs kilos de sédiments ont été prélevés au niveau du bassin, et éventuellement dans la région abdominale et thoracique, pour chacun des six moutons inhumés. Du sédiment du fond de la fosse, après retrait des squelettes, et un échantillon à 10 centimètres de profondeur sous la fosse ont aussi été prélevés. Un contrôle négatif en dehors de la fosse, et des sédiments d'une fosse postérieure au 13^{ème} siècle (échantillon ST73) ont été récupérés pour analyse comparative.

Au niveau des squelettes de moutons, un très faible nombre d'œufs (un ou deux) d'*Ascaris* sp., de *Trichuris* sp. de *Fasciola* sp. et de *Capillaria* sp. ont été observés. Aucun œuf n'a été retrouvé dans le fond ou en-dehors de la fosse. Des œufs de type Ascaridés ont été identifiés dans l'échantillon ST73.

IV. Les sites modernes

A. *Le Fort Saint-Sébastien, France*

Ce fort se situe dans la plaine d'Achères, au nord de la forêt de Saint-Germain-en-Laye dans les Yvelines. Des fouilles préventives par l'Inrap ont été menées, préalablement à l'extension d'une usine d'épuration des eaux. Celles-ci ont révélé la présence du Fort, ainsi que des occupations datant du Néolithique et de l'Âge du Bronze. Seuls les échantillons associés au Fort font partie du corpus étudiés.

1) *Le contexte archéologique*

Cette fortification militaire a été érigée sous Louis XIV entre 1669 et 1670. Un premier fort a été édifié en 1669 et représente une structure de 600 mètres par 380 mètres (Figure 32). La seconde structure, construite l'année suivante, se présente sous la forme d'une enceinte semi-circulaire. Ces fortifications de terre, entourées de fossés de sept mètres de largeur et trois mètres de profondeur, étaient dédiées à la simulation de siège et de prise de places fortes. Le Roi, habitant jusqu'en 1682 au Château de Saint-Germain, a choisi cette localisation pour pouvoir être à proximité de ses troupes et avoir un certain contrôle sur leur entraînement (Hurard et Hanotaux, 2012).



Figure 32 : une aquarelle du Fort de Saint-Sébastien (Archives municipales d'Herblay).

Cette découverte est importante pour l'archéologie française, car il est l'unique représentant de la poliorcétique, l'art du siège, au 17^{ème} siècle. La structure a été utilisée pendant deux années pour entraîner 30 000 soldats en vue du siège de Maastricht (Pays-Bas, 1673) (Inrap, 2012). A la fin de l'entraînement des soldats et leur départ en campagne, le fort est rasé et les terres sont remises en culture.

Les fouilles ont montré la présence de bâtiments de cantonnement et de campement pour les troupes associées à diverses structures révélant le quotidien des soldats, leur mode de vie, et leur alimentation, principalement composé de pain, bœuf (restes de faune sur le site) et de poisson en conserve. À proximité des écuries étaient installées des latrines-tonneaux, creusées jusqu'à cinq mètres de profondeur (Hurard, 2012 ; 2013).

2) Les résultats de la paléoparasitologie microscopique

Des prélèvements de matières organiques ont été effectués dans le fond de quatre latrines-tonneaux et un contrôle négatif en dehors d'une latrine. Concernant l'ensemble du site, des œufs de *Trichuris trichiura*, *Ascaris* sp., *Dicrocoelium* sp. et *Fasciola hepatica* ont été identifiés. Aucun œuf n'a été retrouvé dans le contrôle et dans la latrine 4329.

B. Baillif, Guadeloupe

Baillif fut l'un des premiers lieux de peuplements européens à s'établir sur l'île de la Guadeloupe, lorsque des pères dominicains vinrent s'y installer à partir de 1636. Un projet de contournement routier entre Basse-Terre et Baillif a amené la prescription en 2010 d'un diagnostic archéologique avant le début des travaux.

1) Le contexte archéologique

Un cimetière de l'époque coloniale utilisé entre 1816 et 1846 AD et composé de sépulture en cercueil, a été retrouvé sous les maisons des rues fouillées. Cent quatre-vingt-treize sépultures furent découvertes, organisées selon les pratiques funéraires chrétiennes de l'époque (cercueil en terre, corps sur le dos, la tête orientée vers l'ouest) (Figure 33). Seule la section du site concernée par les aménagements routiers est fouillée, les archéologues estiment que plus d'un millier de sépultures pourraient être présentes dans le cimetière.

Les corps inhumés sont exclusivement adultes, majoritairement de sexe masculin (bien que plusieurs os du bassin aient été trop endommagés pour préciser le sexe). Leur état sanitaire semble plutôt bon et aucune lésion de type traumatique n'est observée. Des traces d'actes médicaux sont cependant à noter : amputation d'une jambe et découpage du sommet du crâne, pouvant être le signe d'une autopsie (Romon, 2010).

Le statut de l'utilisation du cimetière reste à préciser : était-ce un site dédié aux habitants de la paroisse, aux esclaves des pères dominicains de la sucrerie voisine ou bien pour les soldats de l'hôpital militaire de Basse-Terre ? L'absence de signes magico-religieux dans les cercueils et le bon état de santé apparent des inhumés portent à croire que la dernière hypothèse serait la bonne (Kacki et Romon, 2013).



*Figure 33 : inhumation d'un squelette du cimetière de Baillif (Crédit photo C. Samuelian)
(Le Bailly et al., 2014b).*

2) Les résultats de paléoparasitologie microscopique

Aucun œuf de parasite n'a été mis en évidence dans les échantillons de sédiments prélevés sous les squelettes. Cependant, étant donné le contexte historique, une étude en immunologie a été réalisée, qui a permis de montrer la présence d'*Entamoeba histolytica*, une amibe pathogène humaine responsable de dysenterie, chez 16% des individus testés (Le Bailly et al., 2014b). Ce résultat correspond à sept individus testés positivement, avec au moins deux réplifications positives des expériences.

C. Les sites alsaciens de la Première Guerre Mondiale

1) Le contexte archéologique

Le site d'Entzheim-Geispolsheim, dans le Bas-Rhin a été découvert lors de travaux exploratoires préalables à des aménagements urbains. Des traces d'occupation depuis le Néolithique ont pu être mis en évidence. Datant de la Première Guerre Mondiale, des tranchées allemandes, dédiées à la défense de la ville de Strasbourg, ont aussi été trouvées. Une latrine, composée de quatre tonneaux, était installé en association avec la structure (Le Bailly *et al.*, 2012).

À Carspach, dans le Haut-Rhin, des sources historiques et un monument commémoratif mentionnait la présence d'une galerie allemande souterraine de la même époque. Mesurant 125 mètre, elle servait de refuge lors des attaques ennemies. Le 18 mars 1918, une attaque des Français obligea les soldats allemands à se réfugier dans la galerie, qui s'effondra et entraîna la mort de 34 d'entre eux. Au cours des fouilles préalables aux aménagements urbains, les squelettes de 21 soldats furent découverts (Figure 34) (Le Bailly *et al.*, 2014a).



Figure 34 : le squelette d'un soldat allemand retrouvé à Carspach dans une galerie effondrée (Crédit photo M. Landolt) (Le Bailly et al., 2014a).

2) *Les résultats de paléoparasitologie microscopique*

Concernant le premier site, des prélèvements paléoparasitologiques ont été effectués sur les fonds des tonneaux. Des œufs d'*Ascaris* sp., *Trichuris* sp. et de Taeniidae furent identifiés.

Sur le site de Carspach, du sédiment de la région abdominale provenant de trois squelettes a été analysé. Les trois mêmes taxons ont été identifiés, ainsi que des œufs de Capillariidae.

Chapitre II : Introduction méthodologique

I. L'ADN et le sédiment

Lorsque nous réfléchissons à la composition du sol, la présence de molécules d'ADN n'est pas ce que nous évoquons en premier lieu. Pourtant, jusqu'à 10% du phosphate sédimentaire proviendrait de l'ADN, preuve de son abondance (Levy-Booth *et al.*, 2007). Les destins de la molécule, une fois son cycle extracellulaire environnemental débuté, peuvent être multiples et variés, selon divers facteurs physiques, chimiques et biologiques. Nous décrirons dans cette section les trois principales voies que peuvent suivre l'ADN en contact avec un milieu sédimentaire.

A. La dégradation de l'ADN en ses composants de bases

Comme nous l'avons vu dans la section sur les dégradations enzymatiques de la partie I, l'ADN est rapidement clivé en courts fragments après la mort d'un organisme. Dans le sol, les microorganismes sont les principales sources de DNase I (Blum *et al.*, 1997) qui peuvent dégrader l'ADN jusqu'au nucléotide. La cinétique de la réaction est principalement influencée par la température et la présence de certaines particules sédimentaires pouvant protéger les biomolécules, comme les argiles. Les nucléotides peuvent être ensuite recyclés par les procaryotes, mais ils sont peu solubles dans les lipides des membranes cellulaires et leur passage par transport passif à travers ces membranes est donc difficile. Ils pourront subir une hydrolyse, qui donnera des nucléosides solubles, et cette réaction peut se poursuivre jusqu'à l'obtention des bases purines ou pyrimidines et du désoxyribose. Ces constituants peuvent réintégrés le compartiment intracellulaire des organismes et participer à la synthèse métabolique *de novo* ou au recyclage des nucléotides. Ils peuvent aussi être dégradés en PO_4^{3-} , CO_2 ou NH_4 qui seront réabsorbés par les organismes pour entrer dans divers processus métaboliques ou poursuivre leur cycle nutritifs dans le sol (Levy-Booth *et al.*, 2007).

B. La transformation bactérienne

Les bactéries compétentes peuvent intégrer, par transformation génétique, cet ADN extracellulaire en provenance du sol et ensuite l'exprimer. Cependant, ces procaryotes sont souvent en dormance ou sont limitées dans leur phase de croissance exponentielle, freinant ainsi les processus de transformation.

Ce phénomène est une source d'inquiétude quant à l'utilisation des plantes transgéniques. En effet, les chloroplastes génétiquement modifiés sont présents en plusieurs copies par cellules et ont une homologie de séquences élevées avec les génomes procaryotes, augmentant la probabilité de transformation.

C. Persistance à long terme de l'ADN dans le sol

Nous allons maintenant aborder les différents processus qui entraînent la préservation de l'ADN à long terme dans le sol. Celle-ci, médiée par des facteurs physiques, chimiques et biologiques, intervient grâce à deux éléments : les particules sédimentaires et les substances humiques interagissent aussi bien avec les molécules d'ADN qu'avec les nucléases en les immobilisant, diminuant ainsi la dégradation enzymatique de l'ADN.

Ces processus peuvent en partie expliquer comment des biomolécules anciennes provenant d'échantillons archéologiques sédimentaires ont pu être analysées. Par exemple, une approche de *barcoding* moléculaire a contribué à la compréhension des paléoenvironnements du Groenland il y a 400 000 ans (Willerslev *et al.*, 2007). Une étude portant sur des sédiments marins de 32 000 ans a permis d'étudier la paléodiversité de mycètes dans le permafrost arctique (Bellemain *et al.*, 2013). Une reconstruction de paléoenvironnements végétaux entre 10 400 et 1 300 BP a enfin été menée par une approche de métagénomique à partir de sédiments lacustres scandinaves (Parducci *et al.*, 2013).

i. Les substances humiques

Ces substances, qui correspondent aux produits de dégradations des acides aminés, lipides, polysaccharides...ont été décrites dans le chapitre I de la partie 1. Elles peuvent soit inhiber les réactions enzymatiques en laboratoire lorsqu'elles sont co-extraites avec l'ADN, soit participer à la préservation de ces molécules lorsqu'elles sont encore en milieu sédimentaire.

En système artificiel, il a été montré que 2 mg d'acide humique pouvait se lier à 30 µg d'ADN de *Bacillus subtilis*, réaction favorisée par un faible pH. Au cours de cette expérience, 70-80% de l'ADN n'était pas élué après deux lavages au NaCl 0,1M, avec un tampon d'extraction classique ou avec de l'eau. La protection de l'ADN contre les nucléases a été évaluée par le taux de transformants obtenus en présence et en absence d'acide humique. La présence de 1 ng de DNase I inhibe 96-99% de la transformation bactérienne par de l'ADN libre, alors que ce taux chute à 17-26% en présence d'acide humique (Crecchio et Stotzky, 1998).

ii. Les particules minérales

L'ADN peut aussi établir des liaisons avec les particules minérales du sol, le nombre de sites présents et la disponibilité des cations déterminant la quantité de molécules qui peut être liée. La notion de surface spécifique permet d'évaluer théoriquement la capacité d'adsorption d'un sol : le sable à une faible surface spécifique de 0,0485 m²/g, alors que les argiles de type kaolinite et montmorillonite possèdent des valeurs respectives de 10,97 m²/g (Slater *et al.*, 2006) et 812 m²/g (Petersen *et al.*, 1996). La proportion en argile aura donc un impact direct sur la quantité d'ADN adsorbé : l'adsorption de 20 µg d'ADN sera de 40% supérieur sur du limon argileux (composé à 6,2% d'argile) que sur du sable (0,1% d'argile) (Blum *et al.*, 1997).

En plus de la surface spécifique, la densité en charge de surface des argiles affecte aussi l'adsorption. Les différents types d'argiles possèdent un nombre variable de charges négatives en surface, des cations doivent être impliqués pour que les molécules d'ADN, aussi chargées négativement, puissent être adsorbées. Bien que la kaolinite ait une surface d'échange spécifique moins élevée, dans un environnement pauvre en cations divalents, elle adsorbait plus de molécules que la montmorillonite (Levy-Booth *et al.*, 2007). Notons que ces ions peuvent jouer le même rôle pour la liaison aux acides humiques.

La localisation de l'ADN au sein de la structure argileuse peut aussi varier en fonction du pH. Lorsqu'il est supérieur à 5, la molécule se localise en surface sur les charges négatives par l'intermédiaire des cations. Si le pH diminue, l'ajout d'un ion H^+ aux groupements hydroxyles en surface de l'argile localisera l'ADN plutôt dans sa structure interne. Par ailleurs, dans ces sols acides, l'ADN n'est plus chargé et les interactions pourront donc éventuellement aussi s'établir en surface même en absence de cations.

Finalement, la composition même de la molécule d'ADN peut jouer un rôle sur l'adsorption aux argiles : les purines entrent plus facilement en interaction avec ces particules sédimentaires, ainsi que les molécules doubles brins, de masse moléculaire élevée (Franchi *et al.*, 2003).

II. Le géotypage des helminthes gastro-intestinaux anciens par une approche à haut-débit

Les projets associant archéologie et génétique requièrent souvent l'analyse d'un nombre importants d'échantillons. En effet, selon le type d'étude, il peut être nécessaire :

- d'étudier un grand nombre d'échantillons provenant d'un site archéologique unique,
- d'analyser plusieurs sites archéologiques correspondant à une même période,
- d'effectuer une étude rétrospective sur l'occupation d'un site au cours du temps.

Plusieurs centaines d'échantillons peuvent ainsi être analysés, sans connaissance préalable sur la préservation de l'ADN. Il est important de pouvoir proposer, lors d'une collaboration avec des archéologues ou des anthropologues, une méthode d'analyse efficace, peu coûteuse et permettant d'étudier simultanément un grand nombre d'échantillons. C'est dans l'optique de proposer une approche respectant ces contraintes, ainsi que celles associées aux caractéristiques de l'ADNa, que la méthode « *aMPlex Torrent* » a été développée au sein de l'équipe « Épigénome et Paléogénome ».

« *aMPlex Torrent* » associe la spécificité et le pouvoir de détection de la PCR au séquençage à haut-débit *Ion Torrent*[®]. Cette technologie de séquençage permet de traiter simultanément jusqu'à 96 échantillons, par l'utilisation d'étiquettes moléculaires (les *barcodes*). Une fois optimisée, cette méthode s'avère être peu dispendieuse et offre l'avantage de donner des résultats en moins de trois jours. Cette méthode étant décrite dans la publication du chapitre suivant, elle ne sera ici que brièvement présentée.

A. Les PCR multiplex

Une PCR multiplex consiste en l'utilisation de plusieurs couples d'amorces dans une seule réaction de PCR. Ces amorces peuvent cibler l'ADN d'espèces différentes, diverses régions génétiques d'une unique espèce ou bien les deux à la fois, selon les objectifs du projet.

i. La mise au point des PCR multiplex

La mise au point de ce type d'expérience nécessite un effort particulier, du fait de la présence de plusieurs amorces pouvant interagir entre elles, mais aussi parce, dans le cas de l'ADNa, elles ciblent des molécules courtes et dégradées.

Dans un premier temps, les amorces doivent être créées en alignant les séquences de références modernes entre elles. Il est ainsi possible de sélectionner des régions génétiques courtes, mais présentant un polymorphisme nucléotidique (*single-nucleotide polymorphism, SNP*) qui permettra de distinguer les espèces d'un même genre. Dans la mesure du possible, il est préférable d'aligner plusieurs séquences provenant de régions géographiques variées, car la présence d'un polymorphisme intra-spécifique important reste possible. Ces amorces doivent ensuite être testées par qPCR, pour s'assurer de leur efficacité, qui doit être supérieure à 90%. De plus, seuls les couples ne formant pas de dimères avant 40 cycles pourront être inclus dans la suite des mises au point. La création des amorces doit bien sûr respecter les critères de bases, concernant leur taille, leur contenu en G+C, les températures d'hybridation et l'absence de structure secondaire en épingle à cheveux (*Hairloop*) (Chen *et al.*, 2002).

Dans un second temps, en se basant sur des prédictions *in silico* (logiciel PriDimer Check, (Qu *et al.*, 2009)), les amorces sont regroupées dans une ou plusieurs PCR multiplex, qui devront aussi être optimisées car elles peuvent être influencées par (Elnifro *et al.*, 2000) :

- la formation de dimères entre amorces hétérologues,
- la concentration en magnésium,
- la concentration des différentes amorces,
- l'amplification préférentielle de certaines cibles génétiques.

Des qPCR permettent i/ d'analyser l'efficacité des PCR multiplex, ii/ de s'assurer que des dimères, entre amorces d'une même PCR multiplex, ne seront pas amplifiés avant l'ADN-cible.

ii. L'impact des structures secondaires des amorces sur l'efficacité des PCR

L'optimisation des PCR multiplex peut être longue et difficile car les phénomènes de formation de structures secondaires ou d'appariement entre les différentes amorces sont des événements que l'on souhaite éviter (Figure 35). Lors de la formation de structure de type épingle à cheveux, les amorces ne seront plus disponibles pour l'amplification. De plus, si des dimères d'amorces sont amplifiés, ceux-ci entrent en compétition avec les molécules-cibles pour l'utilisation des réactifs du mélange réactionnel (Wei *et al.*, 2008). Ces phénomènes sont d'origine stochastique et peuvent se produire plus ou moins tôt lors d'une PCR. La qPCR est donc la meilleure approche pour évaluer leur impact.

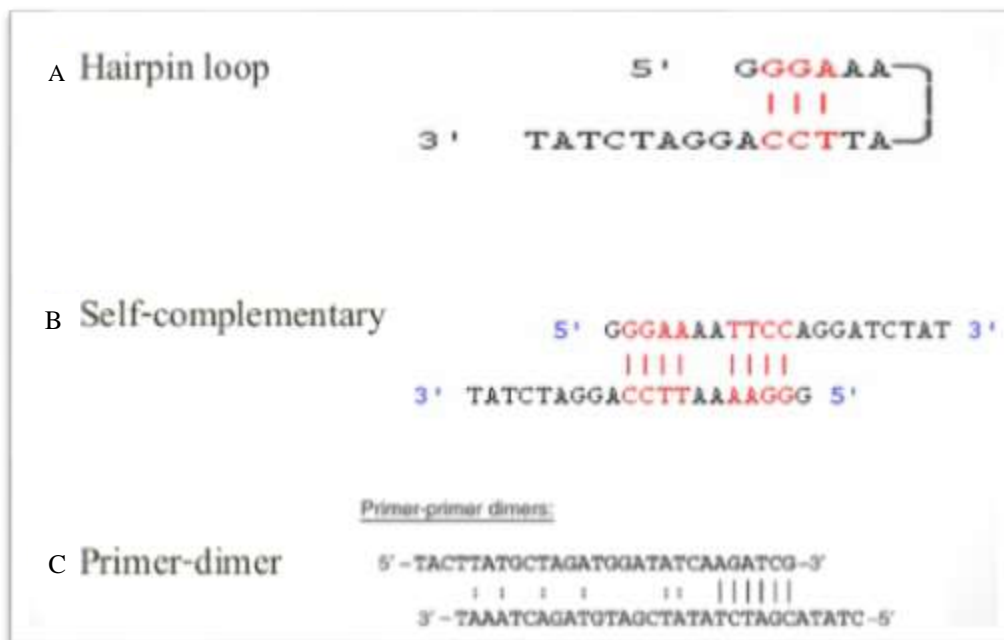


Figure 35 : représentation des trois types de structures secondaires que peuvent former les amorces. A) Hairpin loop (épingle à cheveux) : repliement de l'oligonucléotide sur lui-même, s'il existe une zone de complémentarité interne. B) Self-complementary : une amorce présentant une zone de complémentarité avec elle-même pourra s'auto-apparier. C) Primer-dimer : hybridation d'oligonucléotides entre eux, que ce soit dans un couple (sens+anti-sens) ou entre deux couples d'amorces différents (sens du couple 1 + anti-sens du couple 2 par exemple) (adaptée de http://bioweb.uwlax.edu/GenWeb/Molecular/seq_anal/primer_design/primer_design.htm).

Ces structures peuvent être en partie évitées lors de l'analyse *in silico* des amorces créées et des PCR multiplex. Leur formation peut aussi être étudiée par qPCR en analysant les courbes de fusion des amplifications qui peuvent se produire dans les contrôles négatifs.

Ainsi, seules les amorces générant tardivement des dimères et dont la température de fusion sera différente de celui des cibles à amplifier pourront être retenues pour la suite des expériences. Un exemple d'un couple d'amorces se comportant de manière attendue est présenté à la Figure 36.

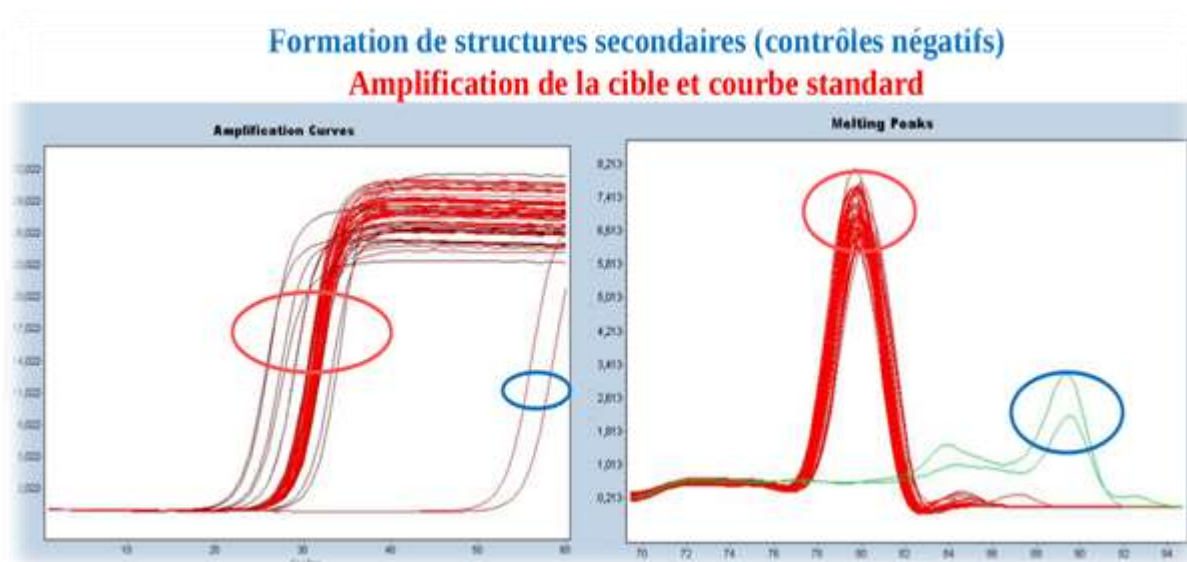


Figure 36 : analyse par qPCR de la formation de structures secondaires par les amorces (entourée en bleue) comparée à l'amplification de la cible et d'une gamme standard (entourée en rouge) (Courtoisie de T. Grange).

B. Le séquençage par la technologie Ion Torrent ®

La seconde partie de la méthode « *aMPlex Torrent* » est le séquençage des amplicons par la technologie d'*Ion Torrent Personal Genome Machine*® (PGM). Mise sur le marché en 2010, cette méthode dite de « nouvelle-génération » offre une précision de séquençage de 99,7% sur 250 pb (Merriman *et al.*, 2012).

i. Présentation de la technologie

Le principe de séquençage est basé sur la libération d'ions H^+ au moment de la formation de la liaison phosphodiester entre deux désoxyribonucléotides. Lors de la polymérisation du brin d'ADN complémentaire, les nucléotides sont envoyés par flux au sein d'un microenvironnement (un puits localisé sur une puce de séquençage) où est présent une molécule d'ADN simple brin.

S'il y a liaison de ce nucléotide avec le nucléotide précédent de la molécule en cours de synthèse, la libération du proton entrainera une variation de pH au sein du microenvironnement. Cette variation sera d'autant plus intense que plusieurs nucléotides identiques seront liés à la suite. C'est ce changement de pH qui permettra au séquenceur de déterminer l'ordre d'incorporation des nucléotides dans la molécule d'ADN.

En 2015, trois tailles de puces de séquençage sont disponibles sur le PGM : 314, 316 et 318. Elles se distinguent par leur nombre de puits, qui sera corrélé au nombre de séquences finales obtenues. La Figure 37 propose un résumé des caractéristiques pour les trois types de puces. Mentionnons que la durée de séquençage d'une puce 314 est d'environ deux heures.

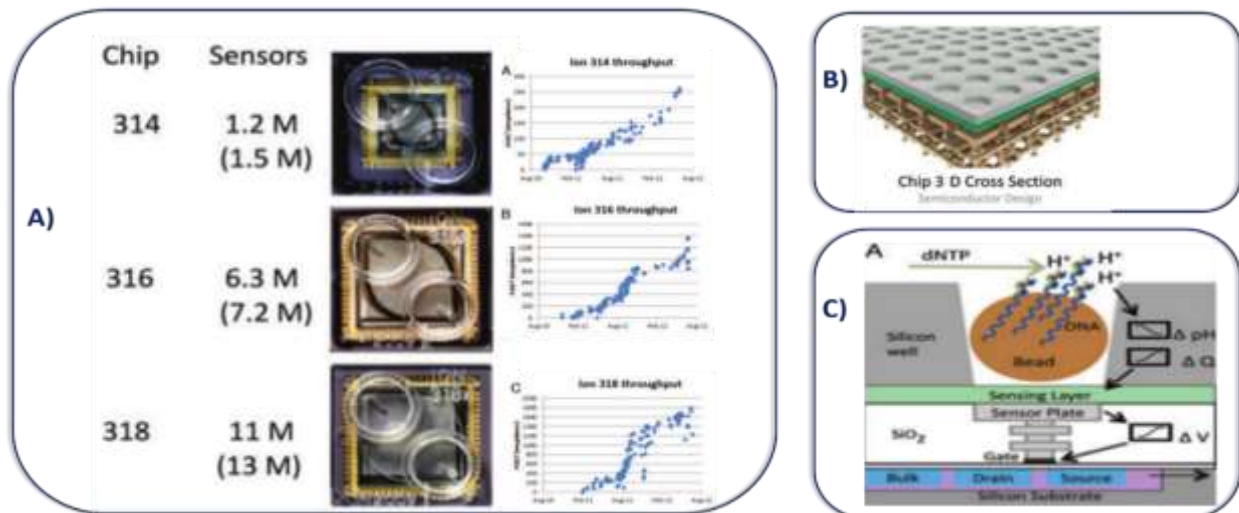


Figure 37 : détails sur les puces Ion torrent PGM actuellement disponibles. A) Pour chaque puce, le nombre de millions de puits adressables est mentionné, suivi d'une photo de la puce. La dernière colonne représente l'évolution au cours du temps du nombre de mégabases pouvant être séquencées B) Représentation schématique des puits contenus dans les puces, qui sont les microenvironnements où se déroulent le séquençage C) Représentation schématique du séquençage au cours duquel les nucléotides liés libèrent des ions H^+ qui sont détectés par un capteur au fond du puits (tiré de (Merriman *et al.*, 2012)).

Les caractéristiques de cette technologie sont bien adaptées aux projets de paléogénétique où la taille visée des amplicons est généralement inférieure à 100 pb. Cependant, comme les autres technologies de séquençage, *Ion Torrent* possède une erreur typique : la détection du nombre de nucléotides composant un homopolymère. Lorsqu'un même nucléotide est intégré plusieurs fois à la suite sur le brin en cours de synthèse, le nombre de protons libérés augmente et la capacité des senseurs à interpréter la variation de pH pour attribuer le bon nombre de nucléotides diminue. Au cours des dernières années, ce biais a été pris en compte et intégré dans les logiciels d'identification des bases. Le taux d'erreur par base pour un homopolymère de cinq nucléotides est passé de 4,5% à 0,7% (Merriman *et al.*, 2012). Généralement, il est plus facile d'interpréter ces erreurs de type insertions/délétions, que les erreurs de types substitutions, générées par le séquençage *Illumina* par exemple.

ii. La préparation d'une librairie de séquençage

De la préparation des amplicons obtenus après les PCR multiplex jusqu'au séquençage, plusieurs étapes sont nécessaires afin de créer une banque de séquençage *Ion Torrent*.

a) La réparation des extrémités

Lors des PCR, l'utilisation de l'enzyme *Taq Fast Start DNA polymerase (Roche)* ajoute un A à la fin des amplicons, générant des extrémités cohésives. Celles-ci sont complétées pour obtenir des extrémités franches, puis phosphorylées en 5' par les enzymes T4 ADN polymérase et T4 kinase polynucléotidique. Cette étape est nécessaire pour poursuivre avec la ligature des adaptateurs étiquetés (*barcodes*).

b) La ligature des adaptateurs

L'objectif de cette étape est de lier à une extrémité de l'amplicon un adaptateur A et à l'autre extrémité, un adaptateur P1. Ces oligonucléotides seront nécessaires pour la PCR en émulsion, mais aussi parce que l'adaptateur A possède une des 96 étiquettes spécifiques à l'échantillon. L'enzyme utilisée est une T4 ADN ligase, qui catalyse la formation d'une liaison phosphodiester entre une extrémité d'ADN 5' phosphorylée et l'hydroxyle en 3' d'une autre molécule. Les adaptateurs étant à extrémités cohésives non-phosphorylées, il en résultera sur le brin ligaturé deux brèches qui devront être réparées. Ces deux premières parties du protocole sont schématisées à la Figure 38 B) et C).

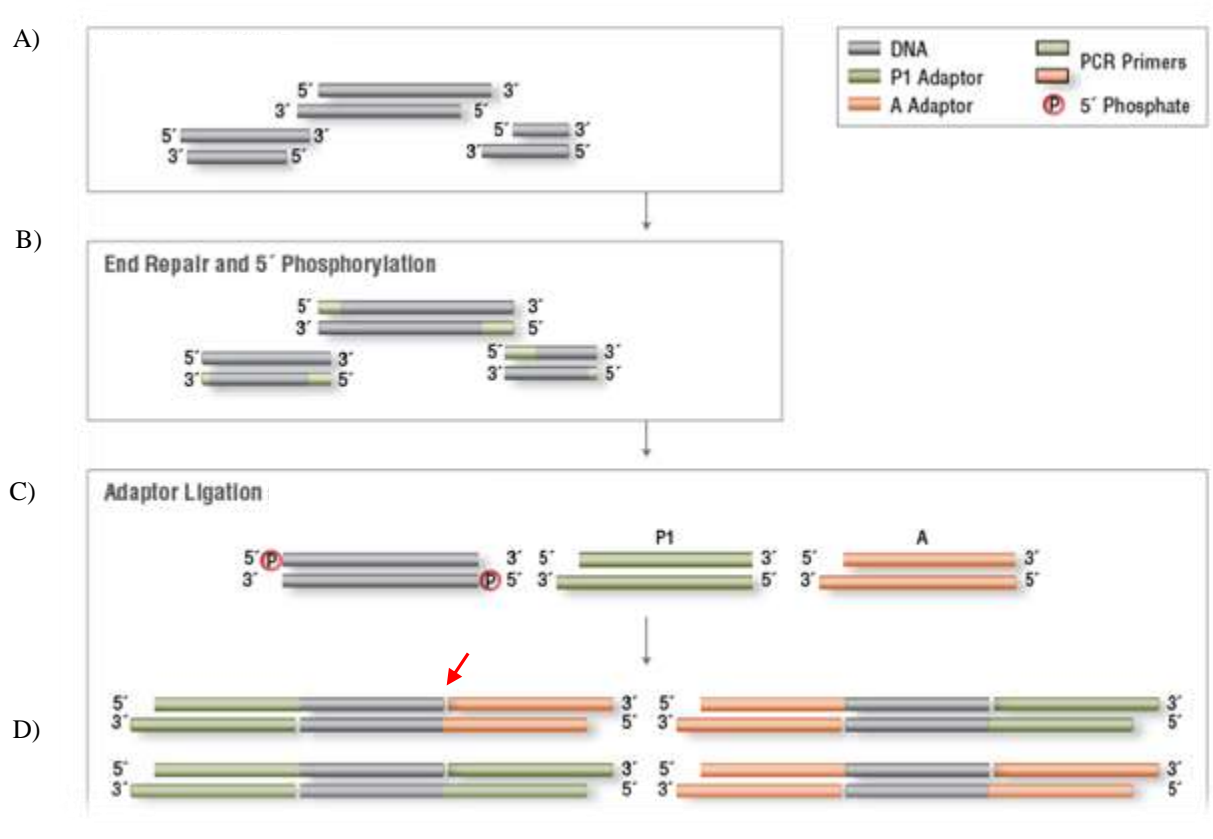


Figure 38 : les premières étapes de création d'une banque de séquençage Ion Torrent. A) Les amplicons obtenus avec leurs extrémités cohésives B) La réparation des extrémités et la phosphorylation en 5' C) Ligature des adaptateurs P1 et A D) Les quatre types de fragments pouvant être obtenus après ligature. La brèche à combler est indiquée par une flèche rouge (figure tirée de www.neb.com).

c) La sélection de taille

Afin de poursuivre avec une banque ne contenant que des amplicons qui ont été ligaturés aux deux extrémités avec les adaptateurs, nous procédons à une étape de sélection de taille. Cela permet aussi d'éliminer les petits fragments comme les dimères d'adaptateurs, les concatémères de divers produits (adaptateurs, amplicons...) et les longs fragments qui pourraient correspondre à des amplicons non-spécifiques. Cette étape est réalisée par électrophorèse sur gel, sur un appareil soit de type *E-gel* (Invitrogen) soit de type *LabChip XT* (*Calliper*) si la fourchette de taille est large.

d) La réparation de brèches

Nous avons vu que la ligature des adaptateurs non-phosphorylés générait une brèche (Figure 38 D). Un mélange d'ADN polymérase ayant une activité exonucléasique 5'→3' est utilisé pour synthétiser le brin complémentaire après déplacement des nucléotides de l'adaptateur. Lors de cette étape, la banque est aussi amplifiée, en utilisant les oligonucléotides A et P1 comme amorces de PCR. Cela permet d'augmenter la proportion de fragments ayant été ligaturés correctement avec les deux adaptateurs, car ceux ayant été ligaturés aux deux extrémités par un adaptateur A ou un adaptateur P1 ne seront pas amplifiés (Figure 38 D), les deux exemples du bas de la figure)

e) La caractérisation de la banque

La quantification de la banque générée est effectuée de trois manières différentes : par qPCR, par dosage au *Qbit* et par *Bioanalyzer*. Ce dernier appareil permet, via son système d'électrophorèse en micro-capillaire, un contrôle quantitatif et qualitatif de la banque. Il est ainsi possible de contrôler aussi bien la taille de la banque, que le bon déroulement de chacune des étapes précédentes. Nous pouvons donc vérifier la taille et la quantité d'amplicons après la PCR multiplex, l'efficacité de la ligature mais aussi estimer si le nombre de cycles d'amplification effectués lors de la réparation de brèches est suffisant pour poursuivre avec la PCR en émulsion.

f) La PCR en émulsion

Lors de cette étape, les divers fragments d'ADN composants la banque sont amplifiés par PCR sur une particule (*ISP, Ion Sphere Particle*). Tous les réactifs nécessaires à l'amplification sont présents dans une gouttelette d'huile, et la banque est diluée de manière à ce qu'idéalement, un seul fragment se localise dans chaque gouttelette. La Figure 39 expose le principe de la PCR en émulsion.

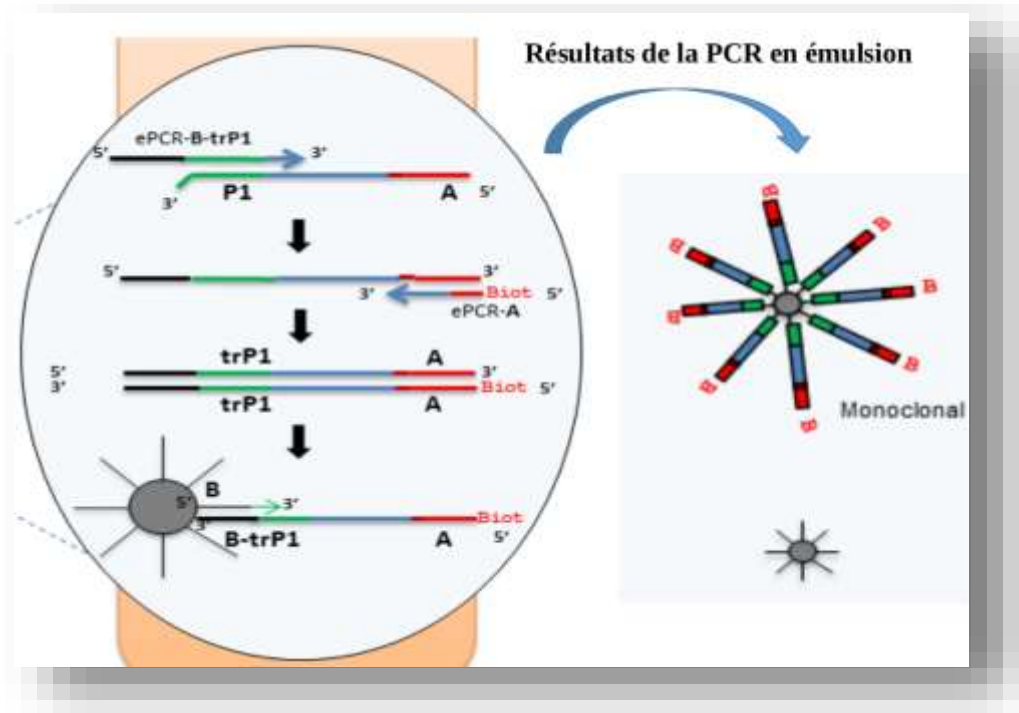


Figure 39 : le principe de la PCR en émulsion Ion Torrent. Sur la partie de gauche est exposée l'amplification des fragments de la banque sur une bille et l'insertion de la biotine, qui permettra lors de l'étape suivante, l'enrichissement par des billes de streptavidine. À droite est représenté un ISP monoclonal (tous les amplicons proviennent d'un unique fragment de départ) et un ISP vide, qui sera éliminé lors de l'enrichissement (figure tirée de <http://www.biorigami.com>).

Les particules sont recouvertes d'oligonucléotides complémentaires aux adaptateurs P1. L'utilisation d'amorces A biotinylées permet d'obtenir des billes recouvertes du même fragment de départ (amplification clonale).

g) L'enrichissement des billes recouvertes d'ADN

Le principe du séquençage est en partie basé sur le remplissage d'une puce dont le fond est recouvert de puits où iront se loger les billes. Pour que ces puits soient occupés au maximum par des ISP recouverts de fragments d'ADN et pour éliminer celles qui n'ont pas amplifié de molécules, une étape d'enrichissement par des billes de streptavidine magnétique est réalisée. Le remplissage de la puce de séquençage peut ensuite être effectué. Le séquençage débute puis le logiciel intégré au serveur de l'appareil procède à l'alignement des séquences obtenues sur des séquences de référence.

C. L'analyse des résultats

Les technologies de séquençage de nouvelle-génération permettent d'obtenir des millions de séquences correspondant à 96 échantillons maximum. Pour analyser ces résultats de manière rapide et reproductible, nous avons mis en place un script informatique, exécutable sous le système d'exploitation *Linux*. Il permet d'utiliser différents outils bio-informatiques pour traiter, compter et explorer la variabilité des séquences obtenues. Ce script est exposé au chapitre suivant avec la publication scientifique.

Chapitre III : A new high-throughput approach to genotype ancient human gastrointestinal parasites

Après avoir présenté la méthode « *aMPlex Torrent* » au chapitre précédent, nous discuterons dans cette section des expériences qui nous ont amené à constater des fluctuations dans la détection des espèces présentes dans les échantillons analysés par cette approche. L'origine de cette hétérogénéité sera ensuite discutée avant de présenter comment nous avons intégré ces considérations pour rendre la méthode proposée encore plus robuste.

I. La réplication des expériences sur de l'ADN ancien sédimentaire

Au cours des travaux de cette thèse, nous avons effectué un certain nombre de réplifications au niveau de la purification de l'ADN sédimentaire ainsi que des PCR multiplex, qui ont ensuite été séquencées individuellement. Cela nous a amené à constater des variations dans la détection des espèces de parasites présents. Lorsque cette variabilité est retrouvée entre différentes purifications, elle serait associée à une hétérogénéité de la répartition des molécules biologiques dans le sédiment. Lorsqu'elle est plutôt observée entre les réplifications de PCR sur un même extrait, la variabilité serait alors liée à la sensibilité de la détection en contexte de multiplexage en raison d'une faible quantité d'ADN de départ. Ces deux situations sont illustrées à la Figure 40. On y observe le nombre de séquences par amplicons obtenus à partir d'un échantillon pour lequel l'ADN a été extrait et purifié deux fois. Sur chacun de ces deux extraits, trois réplifications de PCR multiplex ont été réalisées, puis chacune ont été séquencées individuellement. On peut remarquer que :

- entre deux purifications, un couple d'amorces permettra ou pas d'obtenir des séquences : par exemple, le couple Entero2 a amplifié sa cible uniquement lors de la deuxième purification.
- pour une même purification, le nombre de séquences obtenues varient : par exemple, pour le couple d'oligonucléotide Fas3, les réplifications de PCR de la purif1 ont permis d'obtenir 30 séquences ou aucune.

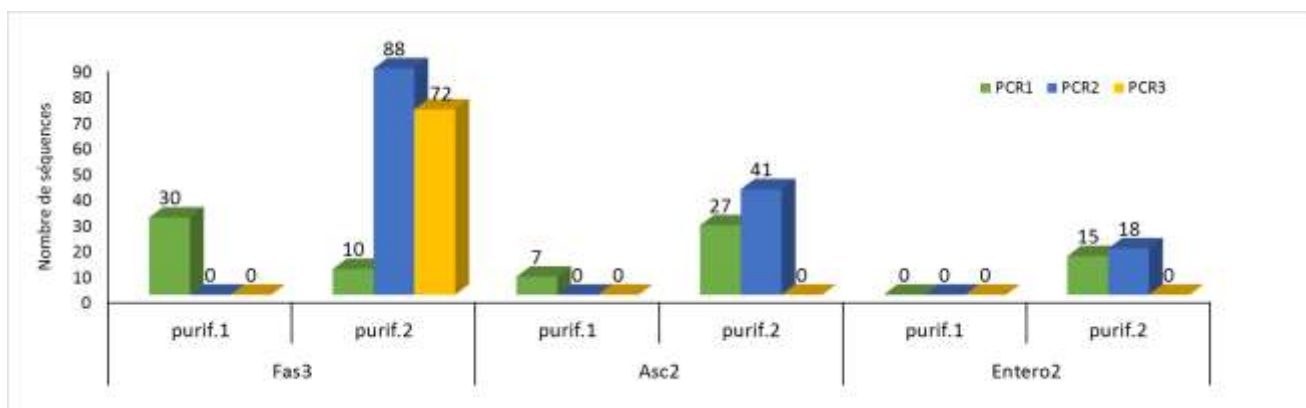


Figure 40 : l'hétérogénéité du nombre de séquences obtenues pour des échantillons de type sédimentaires. Le nombre de séquences obtenu avec la méthode « aMPlex Torrent » pour un échantillon archéologique. Deux réplifications de purifications de l'ADN (purif1 et purif 2) et trois réplifications de PCR multiplex ont été effectuées.

Ces variations peuvent entraîner des difficultés quant aux conclusions que nous pouvons porter sur la présence effective d'une espèce en générant de fausses conclusions négatives.

A. L'hétérogénéité du sédiment

Comme nous l'avons présenté sur la Figure 40, l'extraction et la purification de l'ADN peuvent entraîner des variations dans les taxons détectés au cours des réplifications des procédures. Cela s'explique par le fait que l'ADN des parasites, qu'il soit préservé dans les œufs ou par des interactions avec les particules sédimentaires, n'est pas réparti de manière uniforme au sein d'un échantillon. Il faudrait probablement de considérer deux extraits d'ADN sédimentaire comme deux échantillons différents.

Une solution pour diminuer l'hétérogénéité de l'ADN distribué dans le sédiment est d'augmenter la quantité de matériel traité. Actuellement, les méthodes commerciales d'extraction et de purification de l'ADN sédimentaire permettent de traiter un maximum de 10 grammes de sédiment. Puisque celui-ci est composé d'une grande quantité d'ADN environnemental, traiter plus de 10 grammes pourrait entraîner une saturation de la membrane de silice qui permet de purifier l'ADN.

Il serait donc préférable de multiplier les réplifications d'extraction, avec comme conséquence négative une augmentation considérablement des coûts expérimentaux. Par ailleurs, nous n'avons pas toujours accès à une quantité suffisante d'échantillon archéologique pour pouvoir procéder ainsi.

B. Les PCR avec un faible nombre de molécules-cibles de départ

Une des caractéristiques de l'ADN ancien est de présenter une faible quantité de molécules endogènes, modifiées biochimiquement et diluées avec de l'ADN environnemental. Malgré la sensibilité de la PCR, ces caractéristiques peuvent influencer son efficacité, qui peut être fluctuantes lors de la réplification d'une même expérience.

Le choix des amorces les plus efficaces suivi de la mise au point minutieuse des conditions de PCR multiplex, le tout analysé par qPCR, nous permettent de proposer une approche moléculaire optimale pour la détection des molécules rares. Cependant, il est difficile de la tester en conditions réelles, c'est-à-dire sur un extrait d'ADN ancien. De plus, les molécules des différents parasites ciblés ne sont pas présentes en quantité équivalente dans un extrait d'ADN, ce qui affecte l'efficacité des PCR multiplex (Wei *et al.*, 2008).

Deux processus ont été proposés pour expliquer ces phénomènes de variations observées lors des PCR (Wagner *et al.*, 1994 ; Polz et Cavanaugh, 1998) :

- *PCR selection* : c'est le processus par lequel la réaction va favoriser certaines cibles nucléotidiques en raison de leur structure primaire ou secondaire, de la composition, de la stabilité et de la spécificité des amorces, ou de la composition des régions voisines des sites d'hybridation des amorces. Dans le cadre de ce projet qui est basé sur l'utilisation d'amorces visant à distinguer différentes espèces d'helminthes sur la base de polymorphismes, il est possible que certains taxons soient amplifiés avec plus ou moins d'efficacité. Les molécules anciennes étant fragmentées et des régions très courtes sont ciblées, les biais de composition de ces séquences peuvent aussi affecter la probabilité de leur préservation, car la dégradation de l'ADN est aussi affectée par sa séquence primaire, certains nucléotides et dinucléotides étant plus fragiles que d'autres.

- *PCR drift* : ce phénomène regroupe les évènements stochastiques qui ont lieu lors des premiers cycles de l'amplification. Dans ce cas, les biais introduits ne seront pas reproductibles lors des réplifications de PCR. La formation de dimères d'amorces fait partie de ces évènements aléatoires, pouvant se produire plus ou moins précocement au moment des PCR.

Une solution a été proposée pour agir sur le *PCR drift* lorsque l'on cherche à amplifier plusieurs cibles à partir de quelques picogrammes d'ADN. Il s'agit notamment d'augmenter le nombre de réplifications de PCR et de les regrouper avant de poursuivre avec les expériences suivantes. Cela avait été entre autre proposé par (Chandler *et al.*, 1997), après avoir constaté des fluctuations entre la PCR et la RFLP (*Restriction Fragment Length Polymorphisms*) pour la détection de quelques picogrammes d'ADN de procaryotes.

C'est en tenant compte de ces différentes considérations que nous avons adapté la méthode « *aMPlex Torrent* » aux spécificités de ce projet sur la paléoparasitologie moléculaire. Dans la partie suivante, nous décrivons comment nous les avons intégrés à notre méthodologie pour rendre encore plus fiable la détection d'une grande diversité d'helminthes gastro-intestinaux.

II. A new high-throughput approach to genotype ancient human gastrointestinal parasites

La méthode « *aMPlex Torrent* » combine la sensibilité et la spécificité de la PCR au haut-débit du séquençage de nouvelle génération, tout en permettant d'analyser jusqu'à 96 échantillons simultanément par l'utilisation de *barcodes* moléculaires. Nous décrivons dans la publication les résultats génétiques obtenus pour 16 espèces d'helminthes gastro-intestinaux humains provenant de vestiges de 7200 ans BP jusqu'à la Première Guerre Mondiale, et les comparons aux données obtenues par l'observation microscopique des œufs d'helminthes.

La méthode de PCR multiplexée dans le cadre de la paléogénétique a été par le passé appliquée pour amplifier plusieurs fragments cibles d'une même espèce. Puisqu'elle nécessite une faible quantité d'ADN au départ, il est possible d'obtenir par des amplicons chevauchants, des fragments contigus de plusieurs kilobases. Cette approche a permis entre autre d'obtenir les génomes mitochondriaux complets du mammouth laineux (*Mammuthus primigenius*) (Krause *et al.*, 2006) et de l'ours des cavernes (*Ursus spelaeus*) (Stiller *et al.*, 2009). Notre approche diffère de ces travaux car elle permet de cibler des espèces différentes sur des régions génétiques permettant de les distinguer. Théoriquement, seules les conditions de la PCR multiplex, c'est-à-dire les possibilités de mélanger entre elles des amorces, limitent le nombre d'espèces pouvant être ciblé.

Au cours des expériences, plusieurs échantillons, présentés dans la partie 2 de ce manuscrit et la table 1 de la publication, ont été analysés et pour lesquels nous avons obtenu des résultats génétiques. L'analyse de ces résultats nous a amené à faire deux constats.

Le premier est que la détection d'un taxon est plus fiable en utilisant deux couples d'amorces pour le génotypage. En effet, bien que toutes les amorces aient été testées préalablement dans de mêmes conditions sur de l'ADN moderne, en contexte de PCR multiplex et en présence d'un mélange d'ADN environnemental, leur efficacité peut être fluctuante. Ce phénomène est exposé sur la figure 2 de la publication. Chaque panneau concerne un taxon, détecté par deux couples d'amorces (le plus efficace étant en noir, le moins efficace en gris). En regardant un cas extrême, le cas de l'échantillon Ca (fig2 c), jusqu'à 20 000 séquences ont été obtenus avec le couple Dicro22, le couple Dicro6.1 ayant donné entre 12 et 495 séquences. Pour l'échantillon H, le premier couple a donné entre 0 et 51 séquences, alors que Dicro6.1 n'a pas permis de détecter l'espèce.

Cette différence d'efficacité peut probablement être corrélée à la quantité d'ADN-cible de départ : lorsque celle-ci est importante, un couple d'amorces pourrait être suffisant pour déterminer la présence d'une espèce. Dans tous les cas, il est préférable d'utiliser au moins deux couples d'amorces pour augmenter la fiabilité de détection par un des deux couples.

Le second constat associé à l'analyse de ces résultats est que le nombre de réplifications généralement admis et habituellement réalisé, ne permet pas toujours de détecter à chaque fois toutes les espèces présentes de façon reproductible. La Figure 40 du présent manuscrit exposait déjà la variabilité observée lors de la réplification des PCR multiplex. Celle-ci est aussi décrite sur la figure 2 de la publication. On y remarque que selon les échantillons et les couples d'amorces, un même taxon n'est pas nécessairement détecté lors de chaque PCR. Par exemple, *Trichuris trichiura* de l'échantillon G et *Dicrocoelium dendriticum* de l'échantillon H, sont détectés dans moins de 50% des cas avec le couple d'amorce le plus efficace.

Nous avons comparé les espèces détectées après séquençage individuel de chaque réplification de PCR, à celles détectées par un mélange reprenant un certain volume de chacune de ces réplifications (Pr1-5 sur la figure 2, et voir la fig. 1 qui décrit la procédure). Dans le cas des échantillons avec peu d'ADN, où seul un couple d'amorces permet d'obtenir des résultats, séquencer uniquement le Pr1-5 (que nous appellerons aussi « *pool* ») permettrait d'assez bien caractériser l'échantillon. Dans les deux exemples mentionnés précédemment, *T. trichiura* de l'échantillon G et *D. dendriticum* de l'échantillon H ont été détectés dans une réplification de PCR sur cinq et dans le mélange Pr1-5. Cette approche semble performante pour évaluer la présence des espèces dans un échantillon et diminuer les coûts d'analyses en permettant d'augmenter le nombre d'échantillons analysés lors d'un séquençage. Si une sensibilité plus grande est requise (par exemple, s'il existe une discordance entre les analyses génétiques et les observations microscopiques), il est toujours possible de reprendre le séquençage sur les réplifications individuelles de PCR. Il faut donc faire un compromis entre fiabilité et coût de la détection. Il nous faut alors garder en tête que ce compromis peut ensuite être ajusté selon l'importance de l'échantillon et de la fiabilité de sa caractérisation.

Dans une troisième partie de la publication, nous montrons que cette approche permet aussi de détecter des nouveaux variants génétiques. Ceux-ci peuvent être des variants anciens, absents des populations actuelles, ou bien des variants encore présents aujourd'hui, mais pas encore identifiés et déposés dans les banques de données. Deux exemples sont exposés sur la fig. 3. Tout d'abord, nous avons identifié deux nouveaux *SNP* sur le gène *NadI* d'*Ascaris sp.* dans un échantillon de La Draga (7250 BP).

Nous avons comparé les séquences anciennes obtenues aux séquences modernes pour montrer qu'il existe dans les populations actuelles, autant chez le parasite retrouvé chez le porc que chez l'être humain, une diversité génétique importante. Ensuite, nous exposons les nouveaux variants retrouvés en Europe pour l'espèce *Trichuris trichiura* sur les copies du gène de la grande sous-unité ribosomique. Ceux-ci sont aussi comparés aux références modernes, données provenant uniquement de spécimens de l'Équateur et de la Chine. Puisqu'il existe peu de diversité moderne pour effectuer cette comparaison, il est d'autant plus difficile de conclure sur l'origine des variations que nous avons identifiée.

Nous exposons dans les données supplémentaires un flux d'analyses bio-informatiques, le séquençage à haut-débit génère une quantité importante de données dont le traitement est facilité par l'utilisation d'un script. Celui que nous avons mis en place comporte plusieurs étapes permettant de traiter les fichiers d'alignement au format bam (*binary alignment/map*) générés par le programme TMAP (*Torrent Mapping Alignment Program*), de compter le nombre de séquences obtenues pour chacun des amplicons (logiciel featureCounts, (Liao *et al.*, 2014)), d'identifier les variants obtenus en comparant à une séquence de référence moderne (logiciel Mpileup et VarScan, (Koboldt *et al.*, 2012 ; Li *et al.*, 2009)) et d'obtenir une séquence consensus par échantillon. Pour faciliter la compréhension et l'utilisation du script, nous fournissons les fichiers que nous avons générés pour les analyses. Ils peuvent donc être utilisés tels quels, ou modifiés pour d'autres projets.

Nous concluons en précisant que nous avons obtenu les premières séquences anciennes pour les genres *Tænia* sp. et *Diphyllobothrium* sp., et pu identifier le parasite *Enterobius vermicularis* dans plusieurs échantillons européens. Ce parasite est pour l'instant déterminé en microscopie principalement dans des échantillons d'Amérique, avec une unique occurrence en Europe dans un site allemand médiéval (Bouchet *et al.*, 2003b).

Nous suggérons de procéder aux analyses paléoparasitologiques par une approche complémentaire de microscopie et de génétique, pour pouvoir détecter un maximum d'espèces. Cependant, l'approche « *aMPlex Torrent* » peut aussi être utilisée seule, en particulier car sa mise en place dans un laboratoire de biologie moléculaire relève d'une expertise largement distribuée, alors que les compétences en paléoparasitologie sont rares dans le domaine de l'archéologie.

III. La publication scientifique

RESEARCH ARTICLE

A New High-Throughput Approach to Genotype Ancient Human Gastrointestinal Parasites

Nathalie M. L. Côté^{1,2}, Julien Daligault¹, Mélanie Pruvost¹, E. Andrew Bennett¹, Olivier Gorgé¹, Silvia Guimaraes¹, Nicolas Capelli², Matthieu Le Bailly², Eva-Maria Geigl^{1‡*}, Thierry Grange^{1‡*}

1 Institut Jacques Monod, CNRS, University Paris Diderot, UMR 7592, Epigenome and Paleogenome group, 15 rue Hélène Brion, 75205, Paris, Cedex 13, France, **2** University of Bourgogne Franche-Comte, CNRS UMR 6249 Chrono-environment, 16 route de Gray, 25030, Besançon cedex, France

‡ TG and EMG are joint senior authors on this work.

* eva-maria.geigl@ijm.fr (EMG); thierry.grange@ijm.fr (TG)



 OPEN ACCESS

Citation: Côté NML, Daligault J, Pruvost M, Bennett EA, Gorgé O, Guimaraes S, et al. (2016) A New High-Throughput Approach to Genotype Ancient Human Gastrointestinal Parasites. PLoS ONE 11(1): e0146230. doi:10.1371/journal.pone.0146230

Editor: Michael Knapp, University of Otago, NEW ZEALAND

Received: August 5, 2015

Accepted: December 14, 2015

Published: January 11, 2016

Copyright: © 2016 Côté et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: Nathalie M.L. Côté was supported by a PhD fellowship from the Franche-Comte region. The financial support for the project was obtained from the French National Research Center (CNRS, TG, EMG), the Faculty of Sciences and Techniques of the Franche-Comte University, Besançon (MLB, NC), and the Franche-Comte region, France (MLB, NC). The sequencing facility of the Jacques Monod Institute, Paris, is supported by grants to TG from the University Paris Diderot, the "Fondation pour la Recherche Médicale" (DGE20111123014), and the

Abstract

Human gastrointestinal parasites are good indicators for hygienic conditions and health status of past and present individuals and communities. While microscopic analysis of eggs in sediments of archeological sites often allows their taxonomic identification, this method is rarely effective at the species level, and requires both the survival of intact eggs and their proper identification. Genotyping via PCR-based approaches has the potential to achieve a precise species-level taxonomic determination. However, so far it has mostly been applied to individual eggs isolated from archeological samples. To increase the throughput and taxonomic accuracy, as well as reduce costs of genotyping methods, we adapted a PCR-based approach coupled with next-generation sequencing to perform precise taxonomic identification of parasitic helminths directly from archeological sediments. Our study of twenty-five 100 to 7,200 year-old archeological samples proved this to be a powerful, reliable and efficient approach for species determination even in the absence of preserved eggs, either as a stand-alone method or as a complement to microscopic studies.

Introduction

Infections with gastrointestinal helminths are a major public health concern. Soil-transmitted helminthiasis affect about 2 billion people worldwide [1,2]. Estimates suggest that the nematode *Ascaris lumbricoides* infects nearly 1.2 billion people and *Trichuris trichiura* 800 million. *A. lumbricoides* infections, caused by ingesting food or soil contaminated with their eggs, result in roughly 60,000 deaths per year [1,2]. The eggs of the tapeworms (Cestoda) *Taenia sp.* and *Diphyllobothrium sp.*, found in undercooked meat and fish respectively, develop larval stages that infect the muscles of the intermediate host [3]. *Echinococcus* infection can cause life-threatening cysticercoses. The flukes *Fasciola hepatica* and *Dicrocoelium sp.*, helminths of the class of Trematoda, infect the liver of various mammals, including humans [4]. Liver infection

“Région Ile-de-France” (grant 11015901). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

by flukes in cattle and sheep causes severe economic losses [5]. In general, food- and soil-transmitted helminths produce a wide range of symptoms that include intestinal manifestations (diarrhea, abdominal pain), impaired physical growth, and general malaise and weakness that may affect working and learning capacities.

As high as these numbers are, the percentage of the human population infected with gastrointestinal helminths in the past was likely to have been even higher due to poor hygiene and insufficient medical treatment, but little is known of the past infection rates of various parasites. Owing to the impact of parasites on public health described above, there is a genuine interest of archaeologists and historians to evaluate the health status of ancient societies and populations through the analysis of the prevalence of parasites in archaeological sites. Indeed, gastrointestinal helminths release their eggs into the intestine, which then pass into the environment to resume their life cycle. These eggs are sometimes preserved over time in the sediments of archaeological sites and can be found in mummified tissues, in and under human and animal skeletons in burials, but also in latrines, cesspits and coprolites [6].

Taxonomic assignment of eggs is achieved through microscopic analysis using visually detectable differences in their size, shape, color, and opercula (for review, see [7]). This morphological approach has revealed interesting insights into the health status of past populations and individuals, their diet, and their contact with wild and domestic animals (e.g., [6,8,9]). Nevertheless, taphonomic processes, i.e., the processes by which biological material decays, can lead to the total degradation of eggs, giving rise to false-negative results [10]. Its outreach is also limited since morphological criteria often do not allow the discrimination between species or even genera (e.g., *Taenia sp.* and *Echinococcus sp.*). Assignment of archeological material to the species level, however, will permit a new level of anthropological and biological investigations.

The genetic analysis of DNA from these and other parasites preserved in eggs or sediments can overcome some of these limitations and yield results at a higher level of resolution. A paleogenetic approach would be restricted only by poor DNA preservation, which must be taken into account in the experimental design. Several studies have analyzed archeological remains using a polymerase chain reaction (PCR) approach targeting DNA fragments from 150 bp to 600 bp long (e.g., [11]). The prerequisite of such studies is that the present-day genetic diversity of the parasites is known and similar to that of the past. Moreover, these studies limit their PCR primer design to the parasites previously identified in the sample and still required microscopic identification of parasite eggs, thus excluding the identification of parasites that had escaped visual detection. Metagenomic studies would allow the identification of highly divergent or even unknown parasites [12], but the costs of such studies are often too high to be routinely used to answer archeological questions.

In order to improve the utility of PCR-based parasite identification studies, we have developed a cost-effective high-throughput approach tailored to the needs of archeologists, i.e., the analysis of relatively large numbers of archaeological remains without the need of prior microscopic analysis. This method uses optimized DNA extraction techniques from sediments and adapts the “*aMPlex Torrent*” approach to the analysis of ancient parasite DNA preserved in archaeological sediments of various origins (such as latrines, cesspits, human burials, etc). The *aMPlex Torrent* approach is based on multiplex PCR optimized for highly degraded ancient DNA molecules followed by next generation sequencing of the amplicons on the Ion Torrent platform to efficiently and inexpensively genotype large numbers of archaeological samples [13]. We designed PCR primers to target different taxa of human parasites that are commonly found in archeological sites, taking into account both the short fragment length of ancient DNA molecules and the genetic diversity of the taxa analyzed. Primers were optimized within four multiplex PCRs to allow the identification of 16 species of human tapeworms,

roundworms, pinworms and flukes in up to 96 samples simultaneously. The results obtained with this approach on a variety of ca. 100 to 7,200 year-old samples from various archeological and taphonomic contexts were compared with those obtained with the microscopic approach. This study reveals the power of the genetic approach and reports the detection of parasite DNA even in the absence of corresponding eggs. This new approach for the genetic identification of parasites in archeological sites enables researchers to trace back parasite lineages and better understand their evolution. Additionally, the higher resolution analyses possible through genotyping can clarify transmission events from population migrations and contacts between populations, the conquest of new environments, demographic changes and the domestication of animals during the Neolithic, and overall health status of past individuals and populations.

Materials and Methods

No permits were required for the described study, in particular soil sampling, which complied with all relevant regulations.

Archeological samples

Sediments from archeological sites have been collected from a wide variety of archeological and taphonomic contexts, geographic areas and periods dating from the Iberian Neolithic (ca. 7,200 BP) to the last century (First World War) (see [Table 1](#)).

Microscopic analysis

Prior to genetic analysis, all samples have been analyzed according to standard paleoparasitological protocols at the laboratory of paleoparasitology of the University of Franche-Comté [14]. For each sample, five grams of sample were rehydrated for one week in a mixed solution of 0.5% tri-sodium phosphate (TSP) and 5% glycerinated water. The samples were then crushed in a mortar and subjected to an ultrasound bath for one minute, before finally being filtered in a column composed of sieves with 315, 160, 50, and 25 μm meshes. Since eggs varied in size between 30–160 μm in length and 15–90 μm in width [15], residues from the two last sieves (50 and 25 μm) were transferred to PVC tubes. Ten slides (22x22 mm) were prepared from each fraction and analyzed under the light microscope, amounting to about 5% of the recovered fraction.

Extraction and purification of ancient DNA

The pre-PCR procedures were carried out in the high containment ancient DNA analysis facility with positive air pressure of the Jacques Monod Institute using strict experimental procedures as described [16]. Two to ten grams of sediment were ground to fine powder using a Freezer Mill (6750, Spex Certiprep, Metuchen, NJ). The powder was then purified using PowerMax[®] Soil DNA Isolation Kit (MO BIO Laboratories, Inc. Carlsbad, CA) following the instructions of the manufacturer. Control mock extractions were performed by replacing sediment suspension with 10 mL of distilled water, with one mock extraction for 9 sample extractions. DNA in the 5 mL extract eluted from the MO BIO column was subsequently purified and concentrated on a QIAquick spin column (Qiagen, Hilden, Germany) using 15 mL QG buffer (Qiagen) with 5 mL isopropanol as binding buffer, and PE buffer (Qiagen) for washing. Final elution was performed twice with 45 μL of EB buffer (Qiagen) to which 0.05% Tween-20 was added. To ensure that negative results were not caused by the presence of PCR inhibitors, PCR inhibition of each extract was tested via quantitative PCR (qPCR) on modern genomic DNA, as described [17]. Inhibition in all analyzed samples was found to be low or absent.

Table 1. List of samples analyzed with results of both genetic (DNA sequences) and microscopic analyses (Eggs).

Sample	Location	Date	Type of sample	Context	DNA sequences	Eggs
Aa	Eastern France	20th c. CE	sediment	Barrels (latrine used by soldiers)	<i>Ascaris</i> , <i>Trichuris trichiura</i> , <i>Fasciola hepatica</i> , <i>Taenia asiatica</i>	<i>Ascaris</i> , <i>Trichuris</i> , <i>Taeniidae</i>
Ab	Eastern France	20th c. CE	sediment	Barrels (latrine used by soldiers)	<i>Ascaris</i> , <i>Trichuris trichiura</i> , <i>Fasciola hepatica</i>	<i>Ascaris</i> , <i>Trichuris</i>
B	Guadeloupe	19th c. CE	sediment	burial	<i>Trichuris trichiura</i>	0
Ca	Northern France	1669–1671	sediment	Latrines	<i>Dicrocoelium dendriticum</i> , <i>Diphyllobothrium</i>	<i>Trichuris</i>
Cb	Northern France	1669–1671	sediment	Latrines	<i>Ascaris</i> , <i>Trichuris trichiura</i> , <i>Taenia saginata</i> , <i>Dicrocoelium dendriticum</i> , <i>Fasciola hepatica</i>	<i>Ascaris</i> , <i>Trichuris</i> , <i>Dicrocoelium</i> , <i>Fasciola</i>
Cc	Northern France	1669–1671	sediment	Latrines	<i>Ascaris</i> , <i>Trichuris trichiura</i> , <i>Taenia solium</i> , <i>Dicrocoelium dendriticum</i> , <i>Fasciola hepatica</i>	<i>Ascaris</i> , <i>Trichuris</i> , <i>Dicrocoelium</i> , <i>Fasciola</i>
Da	Northern France	12–14th c. CE	sediment	Latrines	<i>Ascaris</i> , <i>Trichuris trichiura</i> , <i>Taenia solium</i> , <i>Enterobius vermicularis</i>	<i>Ascaris</i> , <i>Trichuris</i> , <i>Dicrocoelium</i> , <i>Taeniidae</i> , <i>Fasciola</i>
Db	Northern France	12–14th c. CE	sediment	Latrines	<i>Ascaris</i> , <i>Trichuris trichiura</i> , <i>Fasciola hepatica</i> , <i>Enterobius vermicularis</i>	<i>Ascaris</i> , <i>Trichuris</i> , <i>Fasciola</i> , <i>Dicrocoelium</i>
E	France	13th c. CE	sediment	Mix of sediments around coxal bones	<i>Ascaris</i> , <i>Dicrocoelium dendriticum</i>	<i>Ascaris</i> , <i>Trichuris</i> , <i>Fasciola</i>
F	Northern France	11–12th c. CE	sediment	Cesspit	<i>Ascaris</i> , <i>Trichuris trichiura</i> , <i>Enterobius vermicularis</i> , <i>Dicrocoelium dendriticum</i> , <i>Fasciola hepatica</i>	<i>Ascaris</i> , <i>Trichuris</i> , <i>Dicrocoelium</i> , <i>Fasciola</i> , <i>Taeniidae</i>
G	Northern France	8–10th c. CE	coprolite	Ditch	<i>Fasciola hepatica</i>	<i>Ascaris</i> , <i>Trichuris</i>
H	Southern France	600–800 CE	sediment	Cesspit	<i>Ascaris</i> , <i>Trichuris trichiura</i> , <i>Dicrocoelium dendriticum</i> , <i>Fasciola hepatica</i> , <i>Enterobius vermicularis</i>	<i>Ascaris</i> , <i>Trichuris</i> , <i>Dicrocoelium</i> , <i>Fasciola</i>
Ia	Eastern France	284–476 CE	sediment	Cesspit-latrine	<i>Ascaris</i> , <i>Taenia solium</i> , <i>Trichuris trichiura</i> , <i>Enterobius vermicularis</i>	<i>Ascaris</i> , <i>Taeniidae</i> , <i>Trichuris</i>
Ib	Eastern France	284–476 CE	sediment	Cesspit-latrine	<i>Ascaris</i> , <i>Trichuris trichiura</i>	<i>Fasciola</i> , <i>Diphyllobothrium</i>
J	Northern France	150 BCE	sediment	Latrines	<i>Ascaris</i> , <i>Dicrocoelium dendriticum</i> , <i>Trichuris trichiura</i> , <i>Taenia solium</i>	<i>Ascaris</i> , <i>Dicrocoelium</i> , <i>Fasciola</i> , <i>Trichuris</i>
Ka	Iran	2500–1500 BP	sediment	Occupational level	<i>Ascaris</i> , <i>Taenia saginata</i> , <i>Dicrocoelium dendriticum</i> , <i>Trichuris trichiura</i> , <i>Enterobius vermicularis</i>	<i>Taeniidae</i> , <i>Trichuris</i> , <i>Enterobius</i>
Kb	Iran	2500–1500 BP	sediment	Occupational level	<i>Ascaris</i> , <i>Dicrocoelium dendriticum</i> , <i>Trichuris trichiura</i>	<i>Ascaris</i> , <i>Taeniidae</i> , <i>Trichuris</i>
La	Northern France	ca 5000 BCE	sediment	Below skeleton in coxal region	<i>Ascaris</i>	0
Lb	Northern France	ca 5000 BCE	sediment	Under skull	<i>Ascaris</i>	0
M	Northern France	6250–5650 BP	sediment	Peripheral ditch	0	<i>Trichuris</i> , <i>Taeniidae</i> , <i>Dicrocoelium</i>
Na	Spain	7250–7050 BP	sediment	Occupational level	<i>Ascaris</i> , <i>Trichuris trichiura</i> , <i>Taenia saginata</i> , <i>Dicrocoelium dendriticum</i> , <i>Enterobius vermicularis</i>	<i>Ascaris</i> , <i>Trichuris</i> , <i>Taeniidae</i>
Nb	Spain	7250–7050 BP	sediment	Occupational level	0	<i>Ascaris</i> , <i>Trichuris</i> , <i>Taeniidae</i>
Nc	Spain	7250–7050 BP	sediment	Occupational level	0	<i>Ascaris</i> , <i>Trichuris</i>
Nd	Spain	7250–7050 BP	sediment	Occupational level	0	<i>Ascaris</i> , <i>Trichuris</i> , <i>Taeniidae</i>
Ne	Spain	7250–7050 BP	sediment	Occupational level	0	<i>Ascaris</i> , <i>Trichuris</i> , <i>Taeniidae</i>

doi:10.1371/journal.pone.0146230.t001

Primer design and testing

We focused on species of parasites common in archeological samples for which modern reference DNA sequences were available. We designed primer pairs to target tapeworms (*Taenia saginata*, *T. solium*, *T. asiatica*, *Echinococcus granulosus*, *E. multilocularis*, *Diphyllobothrium latum*, *D. dendriticum*, *D. nihonkaiense*), roundworms (*Ascaris lumbricoides*, *A. suum*), a whipworm (*Trichuris trichiura*), a pinworm (*Enterobius vermicularis*), and flukes (*Fasciola hepatica*, *F. gigantica*, *Dicrocoelium dendriticum*, *D. chinensis*). To select for amplicons containing single nucleotide polymorphisms (SNPs) that would allow the identification of human gastro-intestinal helminths in a species-specific manner, we designed primers based on the analysis of 3,352 mitochondrial and nuclear DNA sequences present in GenBank[®] at the end of 2013. Aligned sequences were used to design primers hybridizing to conserved sequences and flanking one or several diagnostic SNPs for each genus with the exception of *Ascaris lumbricoides* and *A. suum*, the distinction of which cannot be achieved using short DNA fragments based on currently available DNA sequences [18,19]. Primer design was performed using the program primer3 within the software Geneious 6.1 [20]. For primer optimization, the software Oligo 6 (Molecular Biology Insights, Inc) was used to select primers with a Td of 65° to 67°C (nearest neighbor method) and for which there was no predicted 3'-dimer having a ΔG below -1.6 kcal/mol. In this way, we designed 82 primer pairs amplifying fragments between 52 and 121 bp and tested them for efficiency and dimer formation using qPCR in a LightCycler 480[®] (Roche Applied Science, Penzberg, Germany). PCR amplification was carried out at the genomic platform of the Institut Jacques Monod in a 384-well plate using 5 μ L of SYBR Green I Master Mix 1x pre-heated to 37°C and containing 0.5 μ M of each primer and 2 μ L of DNA, distributed by an epMotion[®] 5070 liquid handling robot (Eppendorf, Hamburg, Germany). A cycling program of 8 min of activation at 95°C, followed by 60 cycles at 95°C for 10 sec and 60°C for 1 min was used, with a melting curve established at the end of the run through one melting step (95°C for 5 sec, then 60°C for 1 min and a temperature increase of 0.11°C/sec with continuous fluorescence measurement).

The performance of the primer pairs was tested *in silico* and *in vitro* and primer pairs were required to have an efficiency superior to 90% and not to yield primer-dimers before 40 cycles. Sixteen primer pairs that satisfied these criteria, the two best pairs per each species (S1 Table), were selected. The resulting amplicons were between 52 to 113 bp and included species-specific SNPs (S2 Table).

Multiplex PCR

We made use of *in silico* prediction for possible dimer formation in multiplex PCRs with the software PriDimerCheck (http://biocompute.bmi.ac.cn/MPprimer/primer_dimer.html). Based on the output of the program, 16 primer pairs were divided into 4 different multiplex reactions to minimize the predicted stability of the possible 3'-annealed dimers. Each multiplex was optimized by testing different MgCl₂ concentrations (3 mM, 4 mM and 5 mM) and primer concentrations (0.08 μ M, 0.15 μ M, 0.2 μ M, 0.25 μ M and 0.3 μ M). Each multiplex reaction was mixed in the high containment laboratory of the Institut Jacques Monod using a 4 μ L mixture of each targeted DNA at 2.5 pg/ μ L, in a total volume of 30 μ L of 50 mM Tris/HCl pH 8.3 (at 25°C), 10 mM KCl, 5 mM (NH₄)₂SO₄, 3–5 mM MgCl₂, 1 mg/mL BSA (bovine serum albumin), 0.08–0.3 μ M of each primer, 0.25 mM dA/G/CTPs and 0.50 mM dUTP, 2 U of Fast Start DNA polymerase (Roche Applied Science, Penzberg, Germany), and 0.5 U of cod uracil-N-Glycosylase (UNG, Arcticzymes, Tromsø, Norway). MgCl₂ solution, reaction buffer and BSA were decontaminated by exposing the solutions in UV-pervious tubes (Qubit[®], Life Technologies, Carlsbad, CA) to UV light for 10 min at a short distance as described [21]. A negative control

with no template was performed for each multiplex reaction, processed in the same way as the samples throughout the whole experimental procedure. The cycling program consisted of 15 min at 37°C (carry-over contamination prevention through digestion by UNG of dUTP-labeled amplicons), 95°C for 10 min (inactivation of UNG and activation of the Fast Start DNA polymerase), followed by 40 cycles at 95°C for 10 sec and 60°C for 1 min and a final extension step at 72°C for 10 min. Cycling was performed in a laboratory of molecular genetics of the Institut Jacques Monod separated from its high containment laboratory.

After the multiplex PCR amplification, simplex qPCR analyses of the amplification products were performed at the genomic platform of the Institut Jacques Monod using each specific primer pair individually, to measure the production of each PCR product and of primer-dimers that might have formed. Each simplex qPCR was pipetted in yet another laboratory of the Institut Jacques Monod, under negative air pressure acting as a DNA trap, as described above using 2 µL of the multiplex product diluted to 1/50th. In order to minimize carry-over contamination, the dilution of the PCR product was performed in a dedicated laboratory, physically separated from the high containment laboratory, the modern DNA facility, the post-PCR facility and the genomic platform. These steps established the optimal conditions for each multiplex using qPCR (S3 Table).

When applied to ancient samples, each multiplex PCR was performed using 4 µL of extract using the optimal conditions established as described above and reported in S3 Table. Thus, when five PCR replicates were performed, 20 µL in total (about a quarter of the purified extract) were analyzed with each primer pair.

Ion torrent library

Ion torrent sequencing adapters were ligated to the replicates of multiplex PCRs in the genomic platform of the Institut Jacques Monod using a Tecan Freedom Evo 100 robot equipped with a 4-channel liquid handling arm using disposable tips, a gripper to move objects, a double thermo-block and an automated vacuum solid phase extraction. First, PCR products were converted to blunt ended fragments with 5'-phosphates using NEBNext[®] End-Repair Module (New England Biolabs, Ipswich, MA) (1 X NEBNext buffer, 0.1 µL End Repair enzyme mix, 10–30 µL PCR products, for a total volume of 50 µL with an incubation at 25°C for 30 min). Prior to ligation, products were purified with the NucleoSpin 96 PCR clean-up kit (Macherey-Nagel, Düren, Germany) and eluted in 50 µL. Ligation of Ion Torrent barcodes and adapters was performed as follows: 20 µL of purified end-repaired products were incubated with 0.66 µM of barcoded-adapters and 0.66 µM of P1 adapter, 1 X NEBNext ligation buffer and 1 µL of NEBNext Quick ligase for 30 min at 16°C. After the ligation, 60 µL of binding buffer (i.e., NT buffer, Macherey-Nagel) were added and the samples were pooled before purification on a single silica column. The pooled bar-coded PCR products were size-selected using the Caliper LabChip XT (Perkin-Elmer, Waltham, MA) to select the PCR products ligated to the adapters. The selected products were then subjected to nick repair and amplification with the OneTaq Hot Start DNA polymerase (NEB) in a 40 µL reaction volume containing Ion Torrent primers A and P1 (0.5 µM), incubated for 20 min at 68°C (nick repair step) and then amplified using the following program: initial denaturation at 94°C for 5 min, (94°C for 15 sec, 60°C for 15 sec, 68°C for 40 sec) for 6 cycles, final elongation at 68°C for 5 min. Products were finally purified using the QIAquick PCR Purification Kit (Qiagen). Size distribution and concentration of the libraries were assessed with an Agilent 2100 Bioanalyzer (Agilent Technologies Inc., Santa Clara, CA). Emulsion PCR and Ion Sphere Particle enrichment were conducted with the Ion OneTouch System[™] (Life Technologies) using the Ion OneTouch[™]200 Template kit v2 DL according to the manufacturer's protocol. The DNA libraries were sequenced on the Ion Torrent Personal Genome Machine (PGM) Sequencer using the Ion PGM 200 Sequencing Kit and Ion 314 or 316 semiconductor sequencing chips (Life Technologies). It

should be noted that the described procedure can be easily adapted to Illumina sequencing technology by using the appropriate adapters if a PGM is not available. The lower sequencing costs of the PGM with Ion 314 chips for up to 200 bp-long sequences, and the superior ability of the PGM to analyze fragments very heterogeneous in size, make it a preferred choice for most applications where the 314 chips provides enough sequencing reads to reliably characterize the PCR products.

Data analysis

Using the software provided in the Torrent suite on the Ion Torrent server, sequencing reads were demultiplexed and mapped with the Torrent Mapping Alignment Program (TMAP) to a reference fasta sequence created by concatenating the various PCR product reference sequences, each separated by a series of Ns. The sequence corresponding to the primers described herein is provided in [S1 File](#). Following mapping, the .bam and .bai files were grouped in a single directory on a Unix-based computer to be further processed. First, files were renamed to allocate simple descriptive names, using bash commands described in the introduction of the analysis bash script in [S4 File](#). The analysis bash script is dependent on the following programs (and tested with the versions indicated between brackets): featureCounts (Subread package 1.4.6) [22], samtools and bcftools (1.2) [23], and VarScan 2 (2.3.7) [24]. featureCounts uses a .gff file describing the features of the fasta file used for mapping. We provide a .gff file in [S3 File](#) corresponding to the PCR products and primers described herein to count the proper PCR products (identified as `misc_feature`), as well as the reads that contain either the forward (identified as `primer_bind`) or the reverse (identified as `primer_bind_reverse`) primers. This should enable the user to detect possible primer dimers, which will have a higher read count for primers than for amplicons. The resulting `CountAmplicons.count`, `CountUpperPrimer.count`, and `CountLowerPrimer.count` files in the count directory report the number of reads obtained for each feature of each sample that was processed by the script, thus allowing identification of the parasite genera detected in each sample. “Samtools mpileup” is used in combination with Varscan to generate a list of the variants detected in each sample. This list is found in the resulting .varscan files in the Varscan directory and allows the identification of parasites at the species level using the information recapitulated in [S2](#) and [S3](#) Tables. Finally, the combination of the “samtools mpileup”, “bcftools call”, and `vcftools.pl vcf2fq` is used to generate a consensus sequence from the concatenated PCR products obtained for each sample. The consensus sequences in fasta format found in the Fasta directory within the Consensus directory can be imported into a software environment for multiple alignments to analyze the data in more detail. We used Geneious [20] to import both the consensus fasta and the initial bam files. We performed a multiple sequence alignment of the fasta files in Geneious using MAFFT 7 [25], with manual editing when necessary. The various outputs were verified by analyzing the initial sequencing reads within the bam files displayed in Geneious.

Results and Discussion

Principle of the experimental strategy

In order to genotype as impartially as possible a large number of more or less well preserved ancient parasite species within a complex DNA mixture containing large quantities of environmental DNA, such as sediments and coprolites, a screening and typing method is needed that satisfies the following criteria: (1) it must take into account the known genetic diversity of the targeted parasites, and (2) it must perform equally well with parasites that are genetically very different. At present, parasites, and in particular gastrointestinal helminths, are best characterized at the level of their mitochondrial DNA. Since mitochondrial DNA also has the highest copy number, the likelihood that rare preserved molecules are detected in an ancient sample is higher than for nuclear DNA. Unfortunately, a large number of parasites cannot be identified

at the species level using only a single universal primer pair. This is particularly true when ancient and degraded short DNA molecules are to be analyzed. Thus, primers were designed to fulfill the following conditions: (i) DNA sequences shorter than 80 bp were preferentially targeted, (ii) the primer binding sites must be conserved within a taxonomic order, family or at least genus, and (iii) the DNA sequence between the primers must allow discrimination between species. These constraints argue in favor of an analysis strategy based on a multiplex PCR approach targeting the different branches of the main parasites observed in archeological sites. Therefore, we adapted to the genotyping of ancient gastrointestinal parasites the aMPlex Torrent method which we have recently developed and which combines the sensitivity of the PCR and the efficiency of next-generation sequencing (NGS) [13]. The principle of the “*parasite aMPlex Torrent*” method presented here is described in Fig 1A. Moreover, this method allows the simultaneous analysis of many samples without requiring prior microscopic analysis, which is particularly adapted for archeological projects, and even more so for the analysis of gastrointestinal helminths, which are often preserved in sediments as mixtures of several species.

We designed primer pairs amplifying various short regions of the mitochondrial DNA of the genera *Taenia*, *Echinococcus*, *Diphyllobothrium*, *Trichuris*, *Ascaris*, *Enterobius*, *Fasciola* and *Dicrocoelium* to discriminate 16 species that have been frequently observed in archaeological samples. To increase the reliability of this approach, two diagnostic genomic regions were targeted for each parasite genus. The best two primer pairs were selected for each genus, i.e., those that provide the lowest likelihood of dimer formation and thus the highest sensitivity, and those that performed best in the multiplex PCRs following the strategy described in Fig 1 and in the Materials & Methods section. After ligating sequencing adapters, sequencing was performed on the pooled multiplex PCRs containing sample-specific barcodes on the Ion Torrent sequencing platform [13].

Performance of the aMPlex Torrent approach for the genotyping of ancient parasites

The performance of the method was evaluated on a variety of archeological samples from various geographical and cultural origins and taphonomic contexts dating from ca. 100 to 7,200 years ago (Table 1). DNA of each targeted parasite was detected even in samples several thousand years old (see Table 1). This demonstrates that the method is powerful enough to detect rare DNA molecules in the complex DNA mixtures of environmental samples. The reliable detection of rare DNA molecules, however, is not trivial and constitutes a general difficulty in ancient DNA research. Indeed, when analyzing poorly preserved samples, separate PCR assays from the same sample may or may not contain one of the few targeted DNA molecules present in the total extract, rendering reproduction of the results difficult. Moreover, stochastic formation of primer dimers, which can be unpredictable and vary in their appearance over the course of the PCR, can dominate the reaction and reduce sensitivity (S1 Fig). Indeed, if primer-dimers form early during the PCR, lower quantities of target DNA molecules may never be amplified to detectable levels. In the case of multiplex PCR, the situation is even more complex since the probability of heterologous dimer formation is increased exponentially when multiple primer pairs are used simultaneously. As a consequence, the likelihood that DNA molecules are detected in a reliable manner in every PCR is the highest for the most abundant DNA molecules that have been analyzed with the most efficient primer pairs. The rarer the DNA molecules and the less efficient the primer pairs are, the lower the probability will be to detect them reproducibly, as can be seen for all analyzed genera in Fig 2A, 2B and 2C. For each genus, the analysis of three samples containing distinct DNA quantities was deduced from the DNA

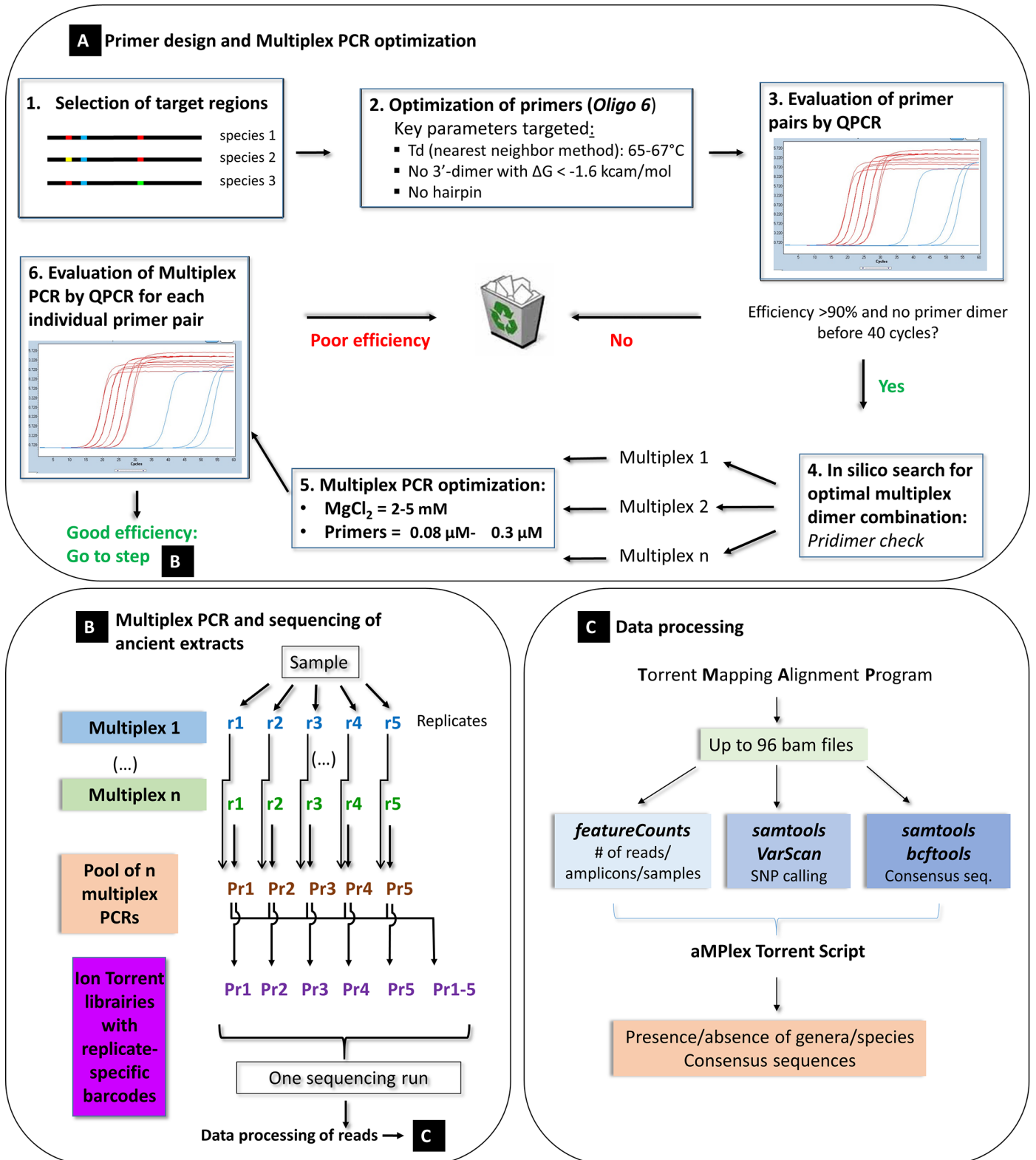


Fig 1. Schematic representation of the workflow of the aMPlex Torrent method for the genotyping of ancient gastrointestinal helminth parasites.

doi:10.1371/journal.pone.0146230.g001

sequences obtained, ordered according to decreasing numbers of sequences (reads). The samples Ia (Fig 2A and 2B) and Cb (Fig 2C) each provide a high number of reads for each primer pair in almost all replicates (r1-r5). Among the two primer pairs detecting each genus, the more efficient one yielded a higher number of reads (Trich4, Asc4, Dicro22). Other samples contained presumably less DNA (Cb in Fig 2A and 2B; Ca in Fig 2C), yielding a lower number of reads per replicate and not all replicates yielding reads. Accordingly, the less efficient primers (Trich3, Asc2, Dicro6.1) not only yielded fewer reads but these also were found in fewer replicates (e.g., r5 for Cb in Fig 2A and r1, r3 and r4 for Ca in Fig 2C). Samples that contained DNA at the lower limit of detection (G in Fig 2A, H in Fig 2B, H1 in Fig 2C) yielded only few reads in only a subset of replicates and only with the more efficient primer pairs.

In order to achieve maximal sensitivity of parasite DNA detection, it is thus advisable to perform several replicates to identify situations where the DNA quantity may not be sufficient to ensure detection with a single PCR. We consider that three to five replicates are a manageable number for the study of archaeological samples. We compared both strategies, (1) sequencing independently five replicates and (2) pooling these five replicates prior to barcoding for the analysis of a larger number of samples per sequencing run, which reduces the costs of the analysis. For most samples and primer pairs where detection was not achieved in all replicates, the pool of these replicates was able to summarize the result with a single barcode (Pr1-5 in Fig 2). At the extreme lower limit of DNA detectability, i.e., in situations where a small number of reads (<10) was detected in only one out of five replicates, the pool did not always yield a sequence (data not shown). This could be attributed to stochastic fluctuations preventing detection of very low DNA copy numbers that are the result of pooling, which dilutes the sample five-fold. However, because it is difficult to completely exclude contamination issues in this kind of experimental set-up, we recommend refraining from considering samples with such low read counts as positive.

In summary, sequencing costs and sample preparation must be weighed against the potential gain in detection power. Although five independent PCRs provide a more direct estimate of the parasite DNA present in a sample, it is likely that the less expensive pooled PCR approach should be satisfactory for most archaeological situations. The number of reads obtained from each sample may reflect either the initial quantity of parasite DNA and/or its preservation.

Genotyping of ancient parasites can reveal unknown past genetic diversity

While the assay was developed to genotype parasite species based on known present-day diversity, the short diagnostic DNA fragments targeted may nevertheless reveal yet undescribed genetic diversity that informs us incidentally on the evolution of these parasites at different time points and geographic locations. Indeed, we could observe in certain samples novel genetic variations within the DNA fragments analyzed. For example, for *Ascaris lumbricoides* and *A. suum*, although most positive samples showed a sequence for fragment Asc4 that was identical to the major present-day sequence described, we found in two ancient samples, dated around 7,000 BP, sequences that diverged by two to four SNPs within this 57-nt long fragment (Fig 3A). Also notable, for *Trichuris trichiura*, the sequences that we identified in all European archeological samples were markedly distinct from the currently known genetic diversity, which has been only characterized in China and Central America (Fig 3B). These novel sequences of human-specific *Trichuris trichiura* were clearly distinct from those found in related parasites of domestic animals. Furthermore, the sequences of parasites found in domestic animals diverge markedly in the primer binding sites and thus should be poorly amplified with the primers used (Fig 3C and S2 Fig). Both analysis of *T. trichiura* DNA sequences using

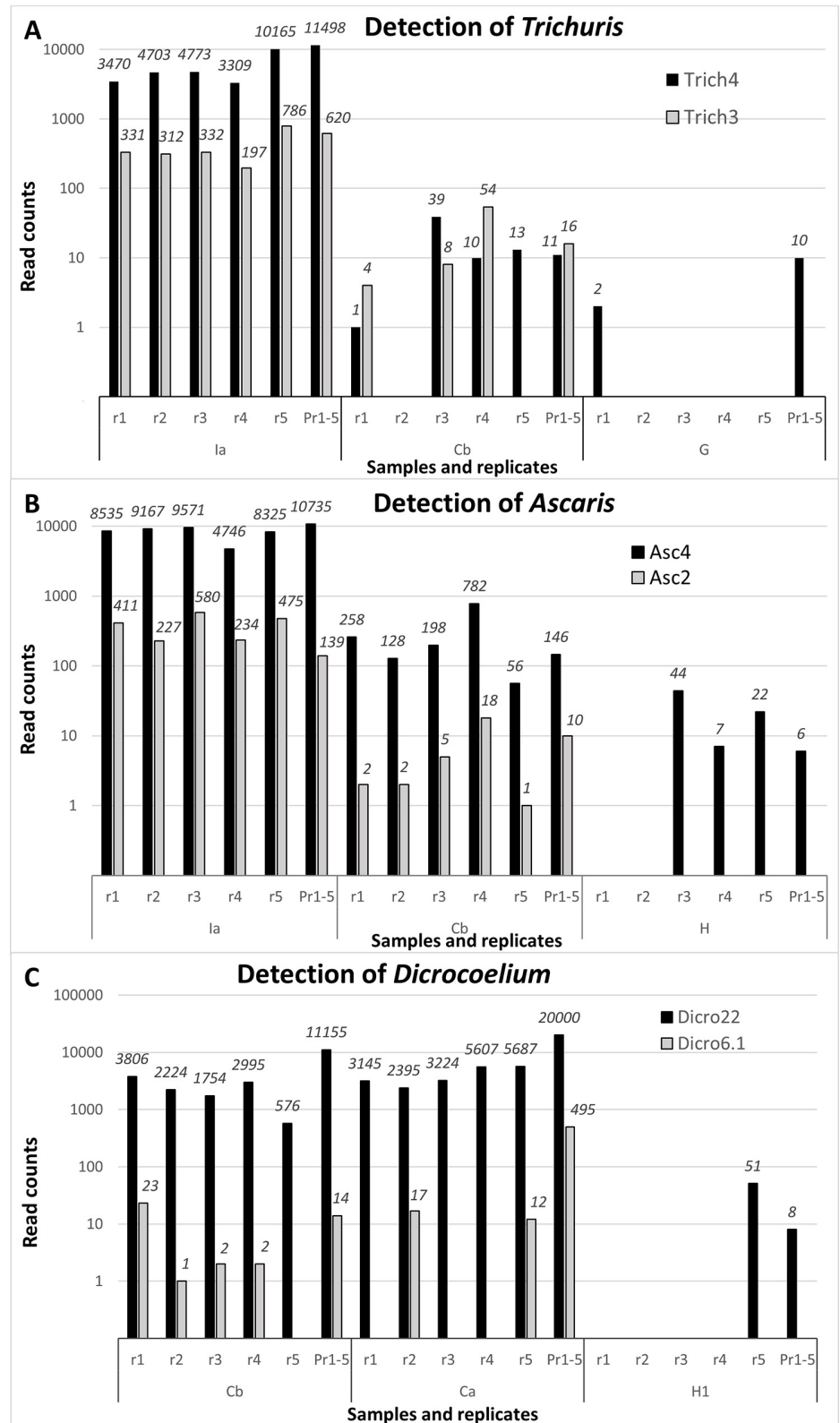


Fig 2. Comparison between the numbers of reads obtained with various primer pairs and samples. Different samples (Ia, Ca, Cb, G, H, H1) in five independent replicates (r1 to r5), as well as in the pool of these five replicates (pool) are represented. **A.** Primer pairs Trich4 and Trich3 that both target *Trichuris trichiura*. **B.** Primer pairs Asc4 and Asc2 that target *Ascaris lumbricoides* and *A. suum*. **C.** Primer pairs Dicro22 and Dicro6.1 that target *Dicrocoelium dendriticum* and *D. chinensis*. Black: the more efficient of the two primer pairs, Gray: the less efficient.

doi:10.1371/journal.pone.0146230.g002

the median joining network and phylogenetic analysis based on the Maximum Likelihood method reveal that the ancient European DNA sequences are affiliated with the Asian and American DNA sequences and are more basal with respect to those from the domestic animals (Fig 3C and 3D). These preliminary observations raise the possibility that these human parasites may have originated from animals, with the ancient European form being closer to the initial transmission event. More data on the genetic diversity of this parasitic species in the human host is needed to clarify its evolutionary history and better pinpoint the timing and location of possible transmission event(s).

Comparison between genotyping and microscopic analysis

In order to explore how well a PCR-based approach compares with the traditional microscopic analysis of eggs, we compared the ability to detect seven genera in our complete sample set using both microscopy and genotyping (Table 1). For each genus, both approaches detect identically either the presence or the absence of the corresponding parasite in about three quarters of the cases (Table 2). Overall, the aMPlex Torrent approach generated a taxonomic specificity that could not be achieved with the microscopic identification method. In 14% of the samples parasite DNA was detected in the absence of microscopically detectable eggs. This value raised as high as 20 and 24% in the case of the genera *Dicrocoelium* and *Enterobius*, respectively, suggesting that unidentifiable egg or larval debris can still contain DNA and that morphological and DNA preservation can be uncoupled, presumably influenced by both the properties of the eggs and the taphonomic conditions in the burial site. Notwithstanding the higher taxonomic specificity of the PCR-based approach, DNA was not detected in 15% of our samples despite the presence of identifiable eggs. This result may have been caused by the uneven distribution of eggs in the soil samples analyzed with either method, in particular because in some of the samples microscopic detection relied on the identification of only one or a few eggs per sample. This effect can be minimized with the genotyping approach, where the proportion of extract analyzed in a single PCR is equal to 10 slides analyzed by microscopy. Indeed, the scalability of genotyping can be used to increase the likelihood of detecting rare or less well-preserved material, overcoming this limitation of the reliability of morphology-based methods. Apart from increasing the starting material used for the DNA extracts, performing PCR replicates can ease the screening of a greater proportion of a given sample, i.e., by five times when five PCR replicates are performed. Still, the disparity we observe between microscopic and genetic results may be traced to a combination of the limitations of each method. For microscopy, degraded eggs may show morphological similarities with closely related taxa, and may thus be misidentified or belong to sub-taxa which have yet to be characterized genetically. For genotyping, DNA may remain undetectable by PCR either because the eggs may not contain larvae or DNA might not always be preserved within morphologically identifiable eggs. It should be noted that in this study, even though one of four 7,200 year-old samples (Na in Table 1) yielded sequences for five different species, DNA was not often detected in the samples that were older than 5,000 years. It is well-established that there is a high extent of variability in DNA preservation in old skeletal samples, in particular linked to variations of the burial conditions (e.g., [26]), and it is likely that such a variability in DNA preservation also occurs among parasite eggs.

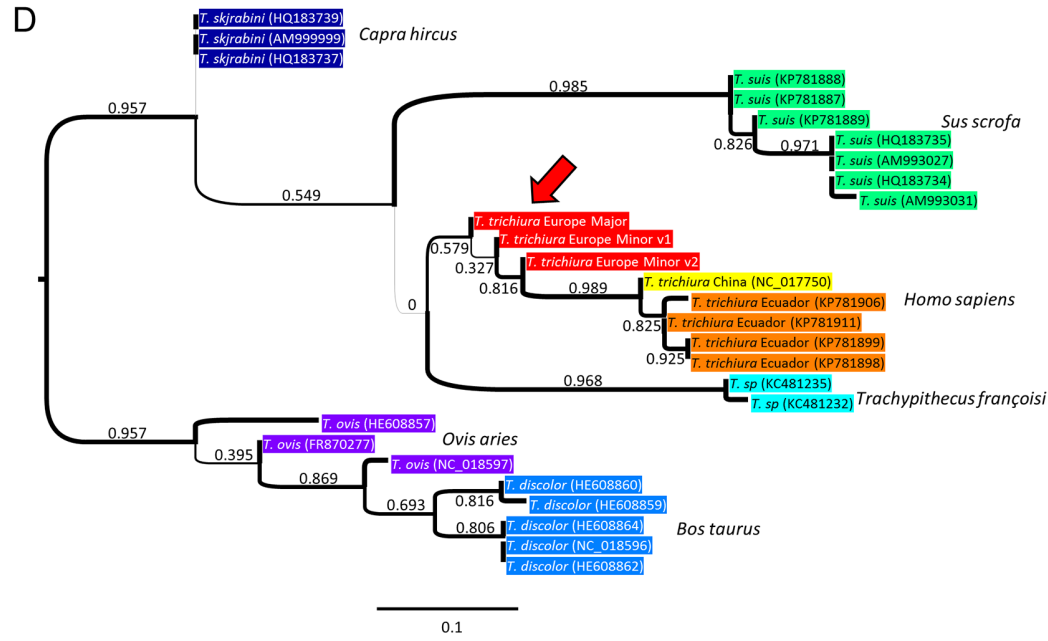
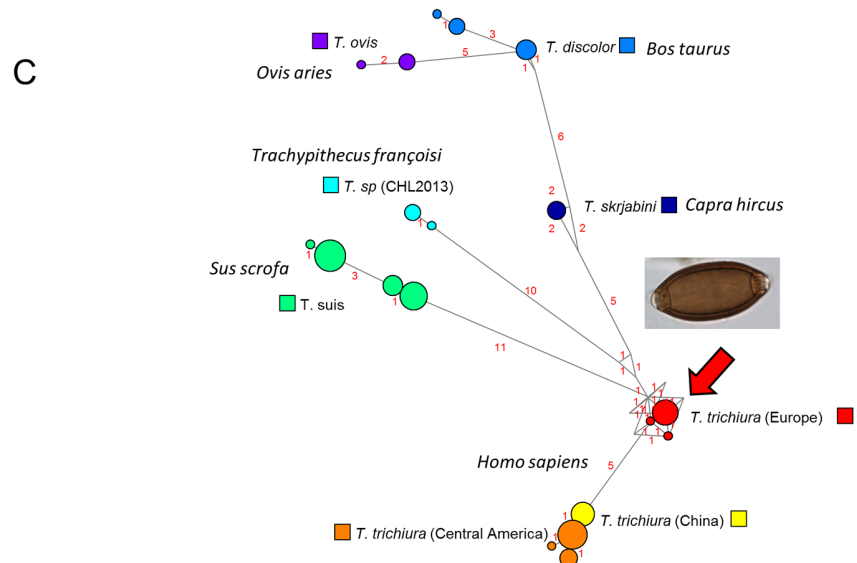
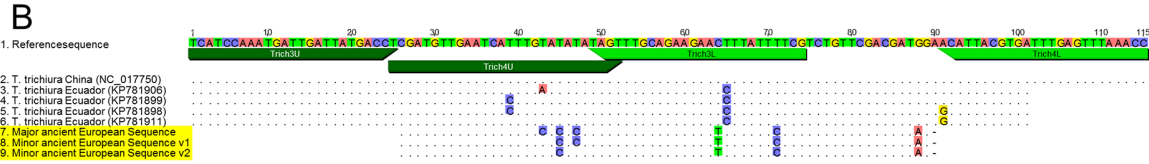
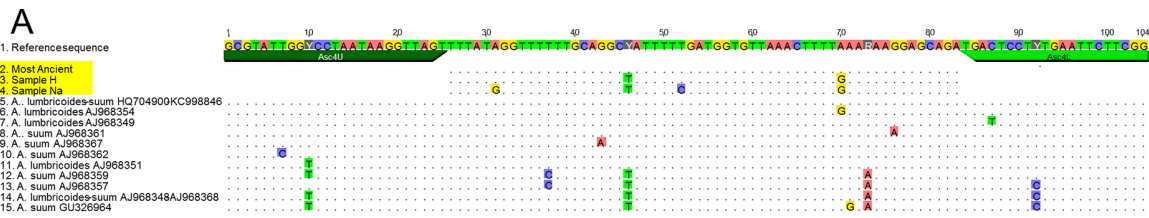


Fig 3. Genotyping reveals past genetic diversity of parasites. **A.** Comparison of ancient and modern genetic diversity of *Ascaris lumbricoides* and *A. suum* on the Asc4 PCR fragment. The ancient sequences are highlighted in yellow. When multiple sequences were identical, a single sequence was represented. The variants of the *Ascaris* sequences were found in samples H and Na, as indicated. **B.** Comparison of ancient and modern genetic diversity of *Trichuris trichiura* on the overlapping Trich3 and Trich4 PCR fragments. The variants v1 and v2 of the *Trichuris* sequences were found in samples Da and Ia, respectively. **C.** Median joining network [28] showing the genetic distance separating the human and animal *Trichuris* species, each represented with a different color, and the central position occupied by the sequence of the ancient European human parasite, indicated by a red arrow. The name of the host is indicated in italics. The number of SNPs distinguishing each sequence is indicated in red alongside the corresponding connecting link. The combined sequences, without primers, of the Trich3 and Trich4 fragments were used. **D.** Maximum likelihood (ML) tree showing the phylogenetic relationships between the human and animal *Trichuris* sequences. The tree was constructed using PHYML 3.0 [29] and the Shimodaira-Hasegawa-like branch test (SH) was used to evaluate the statistical support of the nodes, which are indicated along the corresponding branches of the tree. The most distant animal *Trichuris* sequences (*T. ovis* & *T. discolor*) were used as an outgroup to reveal the relationships between ancient European and modern Asian and American *T. trichiura* sequences. The other represented items are as described in panel C.

doi:10.1371/journal.pone.0146230.g003

Interestingly, the highest number of the proportion of eggs for which no DNA sequences were obtained was observed for the genus *Trichuris*, followed by *Ascaris* and *Taenia*, suggesting either (i) that a particular characteristic of the eggs of these species may make them less suitable for longer term DNA preservation, (ii) that the larvae may have abandoned the eggs in particular environmental conditions, or (iii) that the eggs belonged to animal-borne parasites containing genetic variability beyond that assessed with the primer pairs used. Thus, for the majority of samples, direct, simultaneous, high-throughput genotyping of soil extracts gave similar results but with a higher level of taxonomic clarification than microscopic identification. Both approaches, however, were found to produce low levels of false negatives for different reasons. To sum up, while for many archeological sites the aMPlex approach alone can produce more specific taxonomic data for many samples with greater efficiency than microscopic egg identification, in order to reduce the rate of false negatives due to the limitations of DNA degradation, this method may be used in tandem with traditional microscopy for samples found to contain severely degraded DNA (see also [27]).

Conclusions

We adapted the aMPlex Torrent approach [13] to the reliable and efficient detection of 16 different species of gastrointestinal parasites simultaneously in multiple ancient samples, such as sediments from occupational levels and burials of archeological sites, ancient latrines, and coprolites. This is the first time genotyping has been applied to such a wide range of parasites and archaeological sites. These gastrointestinal helminths are the most commonly detected by microscopy

Table 2. Comparison of performance of the aMPlex Torrent genotyping and the microscopic approaches to identify parasites in archeological samples up to 7,200 years old.

Genus	DNA/eggs correlated			DNA/eggs uncorrelated		
	No DNA, no eggs	DNA & eggs present	Correlated fraction	DNA, no eggs	Eggs, no DNA	Uncorrelated fraction
<i>Ascaris</i>	3	13	64%	4	5	36%
<i>Trichuris</i>	2	13	60%	2	8	40%
<i>Taenia/Echinococcus</i>	11	5	64%	3	6*	36%
<i>Dicrocoelium</i>	12	5	68%	5	3	32%
<i>Fasciola</i>	13	5	72%	3	4	28%
<i>Diphyllobothrium</i>	23	0	92%	1	1	8%
<i>Enterobius</i>	18	1	76%	6	0	24%
Total	82	42	75%	24	27	25%

The numbers of samples in each category are represented, as well as the proportion of samples, in %, for which the results obtained with both methodologies are correlated or not.

* Microscopic determination cannot distinguish between *Taenia* and *Echinococcus*.

doi:10.1371/journal.pone.0146230.t002

in archeological samples [6,7]. Our method allowed for the first time the detection of DNA sequences of ancient *Diphyllobothrium sp.* and *Taenia sp.* While comparative analyses applied to highly degraded material revealed false negatives with both genotyping and microscopy methods, the presented method was powerful enough to detect even highly degraded remains for which no eggs were observed. Thus, we were able to detect *Enterobius vermicularis* in seven samples while up to now its past occurrence has been reported only once in Europe using microscopy, in an ancient German medieval site [10]. We detected *E. vermicularis* DNA in samples as old as 7,200 years, indicating that this parasite was already frequent in Western Europe at this time. Moreover, the new genotyping approach allows systematic assignment at the species level for a number of genera, which is rarely possible using microscopy. Finally, it also enables the discovery of ancient genetic diversity not yet described in modern samples. Thus, the aMPlex Torrent method can be used either as a complement to conventional microscopic analyses, or independently to characterize the biodiversity of gastrointestinal parasites in archeological sites. Compared to the microscopy approach, the genotyping approach provides a much more precise identification, improving the investigative power of paleoparasitological studies. The ability to screen many ancient samples for DNA preservation opens the way to future paleogenomic studies that could shed light on the evolution of parasitism, in particular to study the exchange of parasites between humans and animals following animal domestication.

Supporting Information

S1 Fig. Examples of simplex PCR with primer pairs giving rise (b,c) or not (a) to primer dimers. Standard curves performed in duplicates with serial 5-fold dilution of reference DNA are colored in brown, whereas the no template controls (NTC) are represented in red when they give rise to amplification products, and in blue when they do not. a-b) amplification phase of the qPCR, c) melting curve phase of the experiment displayed in panel b. In this latter experiment, four of the six NTCs which give rise to primer dimers have various distinct melting temperatures (T_m). The T_m of the dimer is usually different from that of the product, but it happens sometimes that these values are similar, as shown here for one of the dimers. In such cases, only electrophoretic analyses can distinguish between dimers and PCR products. Three of these dimers were generated at a Ct between 33 and 36, which is similar to the cycles of the PCR where products corresponding to rare initial molecules typical of ancient samples are also detected. A primer pair with such properties is thus not recommended for the detection of ancient DNA molecules. Results obtained with an optimal primer pair have been displayed in panel a) for comparison. (TIF)

S2 Fig. Comparison of the ancient and modern genetic diversity of *Trichuris* species as revealed through the combination of the overlapping Trich3 and Trich4 PCR fragments. The divergence of the primers from sequences of animal species can be seen. Most mismatches between the primers and modern and ancient sequences correspond to G-T mismatch, which are the least destabilizing mismatches (e.g, [1])

1. Pan S, Sun X, Lee JK (2006) DNA stability in the gas versus solution phases: a systematic study of thirty-one duplexes with varying length, sequence, and charge level. *J Am Soc Mass Spectrom* 17: 1383–1395.

(TIF)

S1 File. ref_parasites.fasta

(FASTA)

S2 File. ref_parasites2.fasta

(FASTA)

S3 File. ref_parasites.gff

(GFF)

S4 File. aMPlexmpileup.sh

(SH)

S5 File. *Ascaris* and *Trichuris* sequences

(FASTA)

S6 File. All supplementary information in one file.

(PDF)

S1 Table. Optimal primer pairs targeting gastrointestinal helminths.

(DOCX)

S2 Table. Sequences of the insert of the PCR products showing the variability found in published sequences within and between species. When several related sequences were available, only the positions differing with the reference sequence are indicated, whereas identical positions are represented by dots. The IUPAC nucleotide ambiguity code was used.

(DOCX)

S3 Table. Experimental conditions for the various multiplex PCRs. Optimal MgCl₂, and primer concentrations used for each individual multiplex PCR are shown.

(DOCX)

S4 Table. Location of the diagnostic SNPs for species discrimination to interpret the varscan data files when using the ref_parasites fasta file for mapping the reads.

(DOCX)

Acknowledgments

We thank Coralie Martin, Robert Nicol, Abdelillah Brahim, Solène Le Gal, Teresa Garate, Keesom Eom, Martha Betson, Akira Ito, Minoru Nakao, Joong-Ki Park, Bruno Gottstein, Jenny Knapp, Jan Brabec, Ralf Bialek, Angela Di Cesare and Cristina Cutillas for providing modern parasite samples. We thank Maria Saña, Michaël Landolt, Stéphane Buttner, Hervé Pomarède, Christophe Tardy, Franck Gama, Benjamin Dufour, Tristan Moriceau, Séverine Hurard, Nicolas Cayol, Thomas Romon, Sacha Kacki, Annelise Binois, Corinne Thévenet, Thomas Stöllner, Abozfal Aali, Nicole Boenke, Masoud Nezamabadi, Jean-Marc Fémolant, Sébastien Lefèvre, and Tetsuya Yanagida for providing archeological samples.

Author Contributions

Conceived and designed the experiments: TG NMLC EMG. Performed the experiments: NMLC JD MLB MP OG SG. Analyzed the data: NMLC TG EMG. Contributed reagents/materials/analysis tools: MLB EMG NMLC TG MP NC. Wrote the paper: TG EMG EAB NMLC.

References

1. Bethony J, Brooker S, Albonico M, Geiger SM, Loukas A, Diemert D, et al. (2006) Soil-transmitted helminth infections: ascariasis, trichuriasis, and hookworm. *Lancet* 367: 1521–1532. PMID: [16679166](#)
2. Brooker S (2010) Estimating the global distribution and disease burden of intestinal nematode infections: Adding up the numbers—A review. *International journal for Parasitology* 40: 1137–1144. doi: [10.1016/j.ijpara.2010.04.004](#) PMID: [20430032](#)
3. Dorny P, Praet N, Deckers N, Gabriel S (2009) Emerging food-borne parasites. *Vet Parasitol* 163: 196–206. doi: [10.1016/j.vetpar.2009.05.026](#) PMID: [19559535](#)

4. CDC (2015) Parasites. Atlanta: Centers for Disease Control and Prevention. <http://www.cdc.gov/parasites/>
5. Mehlhorn H, Aspöck H (2008) Encyclopedia of parasitology, 3rd ed. Berlin; New York: Springer.
6. Bouchet F, Guidon N, Dittmar K, Harter S, Ferreira LF, Chaves SM, et al. (2003) Parasite remains in archaeological sites. *Mem Inst Oswaldo Cruz* 98 Suppl 1: 47–52. PMID: [12687762](#)
7. Sianto L, Chame M, Silva CS, Goncalves ML, Reinhard K, Fugassa M, et al. (2009) Animal helminths in human archaeological remains: a review of zoonoses in the past. *Rev Inst Med Trop Sao Paulo* 51: 119–130. PMID: [19551285](#)
8. Le Bailly M, Landolt M, Bouchet F (2012) First World War German soldier intestinal worms: an original study of a trench latrine in France. *J Parasitol* 98: 1273–1275. doi: [10.1645/GE-3200.1](#) PMID: [22924925](#)
9. Le Bailly M, Landolt M, Mauchamp L, Dufour B (2014) Intestinal parasites in First World War German soldiers from "Kilianstollen", Carspach, France. *PLoS One* 9: e109543. doi: [10.1371/journal.pone.0109543](#) PMID: [25333988](#)
10. Bouchet F, Harter S, Le Bailly M (2003) The state of the art of paleoparasitological research in the Old World. *Mem Inst Oswaldo Cruz* 98 Suppl 1: 95–101. PMID: [12687768](#)
11. Iniguez AM, Reinhard K, Carvalho Goncalves ML, Ferreira LF, Araujo A, Paulo Vicente AC (2006) SL1 RNA gene recovery from *Enterobius vermicularis* ancient DNA in pre-Columbian human coprolites. *Int J Parasitol* 36: 1419–1425. PMID: [16950265](#)
12. Cano RJ, Rivera-Perez J, Toranzos GA, Santiago-Rodriguez TM, Narganes-Storde YM, Chanlatte-Baik L, et al. (2014) Paleomicrobiology: revealing fecal microbiomes of ancient indigenous cultures. *PLoS One* 9: e106833. doi: [10.1371/journal.pone.0106833](#) PMID: [25207979](#)
13. Guimaraes S, Pruvost M, Daligault J, Stoetzel E, Côté N, Bennett EA, et al. (submitted) A cost-effective high-throughput metabarcoding approach powerful enough to genotype ~44,000 year-old rodent remains from Northern Africa. *Molecular Ecology Resources*.
14. Dufour B, Le Bailly M (2013) Testing new parasite egg extraction methods in paleoparasitology and an attempt at quantification. *International Journal of Paleopathology* 3: 199–203.
15. Ash LR, Orihel TC (2007) *Ash & Orihel's Atlas of Human Parasitology*. Chicago, Illinois, USA: American Society for Clinical Pathology Press.
16. Bennett EA, Massilani D, Lizzo G, Daligault J, Geigl EM, Grange T (2014) Library construction for ancient genomics: single strand or double strand? *Biotechniques* 56: 289–300. doi: [10.2144/000114176](#) PMID: [24924389](#)
17. Pruvost M, Schwarz R, Correia VB, Champlot S, Braguier S, Morel N, et al. (2007) Freshly excavated fossil bones are best for amplification of ancient DNA. *Proc Natl Acad Sci U S A* 104: 739–744. PMID: [17210911](#)
18. Betson M, Nejsum P, Bendall RP, Deb RM, Stothard JR (2014) Molecular epidemiology of ascariasis: a global perspective on the transmission dynamics of *Ascaris* in people and pigs. *J Infect Dis* 210: 932–941. doi: [10.1093/infdis/jiu193](#) PMID: [24688073](#)
19. Leles D, Gardner SL, Reinhard K, Iniguez A, Araujo A (2012) Are *Ascaris lumbricoides* and *Ascaris suum* a single species? *Parasit Vectors* 5: 42. doi: [10.1186/1756-3305-5-42](#) PMID: [22348306](#)
20. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28: 1647–1649. doi: [10.1093/bioinformatics/bts199](#) PMID: [22543367](#)
21. Champlot S, Berthelot C, Pruvost M, Bennett EA, Grange T, Geigl EM (2010) An efficient multistrategy DNA decontamination procedure of PCR reagents for hypersensitive PCR applications. *PLoS One* 5.
22. Liao Y, Smyth GK, Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30: 923–930. doi: [10.1093/bioinformatics/btt656](#) PMID: [24227677](#)
23. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25: 2078–2079. doi: [10.1093/bioinformatics/btp352](#) PMID: [19505943](#)
24. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. (2012) VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 22: 568–576. doi: [10.1101/gr.129684.111](#) PMID: [22300766](#)
25. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30: 772–780. doi: [10.1093/molbev/mst010](#) PMID: [23329690](#)

26. Pruvost M, Schwarz R, Bessa Correia V, Champlot S, Grange T, Geigl E- M (2008) DNA diagenesis and palaeogenetic analysis: critical assessment and methodological progress. *Palaeogeogr Palaeoclimatol Palaeoecol* 266: 211–219.
27. Cleeland LM, Reichard MV, Tito RY, Reinhard KJ, Lewis CM Jr. (2013) Clarifying Prehistoric Parasitism from a Complementary Morphological and Molecular Approach. *J Archaeol Sci* 40: 3060–3066. PMID: [23645967](#)
28. Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16: 37–48. PMID: [10331250](#)
29. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 59: 307–321. doi: [10.1093/sysbio/syq010](#) PMID: [20525638](#)

Supplementary data for:

A new high-throughput approach to genotype ancient human gastrointestinal parasites

Nathalie M. L. Côté, Julien Daligault, Mélanie Pruvost, E. Andrew Bennett, Olivier Gorgé, Silvia Guimaraes, , Nicolas Capelli, Matthieu Le Bailly, Eva-Maria Geigl*, Thierry Grange*

About the Authors

N. M. L. Côté, J. Daligault, M. Pruvost, O. Gorgé, S. Guimaraes, E. A. Bennett E.-M. Geigl, T. Grange

Institut Jacques Monod, CNRS, University Paris Diderot, UMR 7592, Epigenome and Paleogenome group, 15 rue Hélène Brion, 75205 Paris cedex 13, France.

N. M. L. Côté, M. Le Bailly, N. Capelli

University of Bourgogne Franche-Comte, CNRS UMR 6249 Chrono-environment, 16 route de Gray, 25030, Besançon cedex, France.

*equal contribution

Corresponding Authors

eva-maria.geigl@ijm.fr; thierry.grange@ijm.fr

Primer name	Organism Gene	Size of amplicon (bp)	Primer sequences
Tae23	<i>Taenia</i> Cytochrome oxidase 1 (cox1)	73	U : GAG GTT TTA GGT TCW TAT GGT T L : CAT TAT GAG AAG CYA CAG GAC T
Tae32	<i>Taenia</i> Cytochrome b (cytb)	113	U : ATA TGG CTC GTG CTT TGT ATT ATT C L: AGG TAA TAT ATA TCC WGT AAA AGC CTC
Echino5	<i>Echinococcus</i> Phosphoenolpyruvate carboxykinase (pepck, nuclear)	109	U : CAC ATG TTG CTG AAG GCG TTA L : AAT GTA CGG TCG TTC ATG CAG
Echino23	<i>Echinococcus</i> cox1	99	U: TTT TGA TCC GTT AGG TGG T L: CCA AAT CCA GGC ARA ATC
Diphyllo2	<i>Diphyllobothrium</i> cox1	52	U: GTT GTG TGG GGG CAT CAT A L:CAG CCG TCT TTA CAT CTA AAC C
Diphyllo23	<i>Diphyllobothrium</i> cox1	67	U: TTT AYG GGT TGT TAT TTG CT L: ATA TGA TGC CCC CAC ACA
Asc2	<i>Ascaris</i> cytb	74	U: GCC AAA GCA CCA TCA TTA GAA TA L : GGT ATG GTT TTG GGT TTT CAG A
Asc4	<i>Ascaris</i> NADH dehydrogenase 1 (nad1)	104	U : GCG TAT TGG YCC TAA TAA GGT TAG T L :CCG AAG AAT TCA RAG GAG TCA
Trich3	<i>Trichuris trichiura</i> Large ribosomal subunit (LSU)	74	U :TCA TCC AAA TGA TTG ATT ATG ACC T L :CGA AAA TAA AGT TCT TCT GCA AAC TA
Trich4	<i>Trichuris trichiura</i> LSU	91	U: TCG ATG TTG AAT CAT TTG TAT ATA TAG T L:GGT TTA AAC TCA AAT CAC GTA ATG T
Dicro22	<i>Dicrocoelium</i> Internal transcribed spacer 2 (nuclear)	76	U : TAC ACA CAC CTA GTT ATC AGA CAG L: ACA GAC CGC GCA TAA ATA
Dicro6.1	<i>Dicrocoelium</i> nad1	77	U: TAA GTA TAA GTT TST GGT TTC TCA GTT TC L: AAG CWA CCA AAA TCA TCA AAA ACA
Fas2	<i>Fasciola</i> cox1	69	U : GTT GAT TGG GGG KTT TGG TA L : CGA GGC AAA TTC AAA TCA GG
Fas3	<i>Fasciola</i> rRNA18s (nuclear)	85	U : AAC CTG CGG AAG GAT CAT TA L: GCA AAT TTT TAT CGC ATG ACA
Entero2	<i>Enterobius vermicularis</i> SL1RNA (nuclear)	56	U: TTT ATT TCC AAG CCA CAG ACT CA L: AAT TTC TCG TTC CGG CTC AG
Entero4	<i>Enterobius vermicularis</i> cox1	53	U: CTG TGC CRA CTG GGG TAA AG L : TCC CCC TAT CAA AGT CAA CAA C

S1 Table. Optimal primer pairs targeting gastrointestinal helminths.

Côté et al, Genotyping of ancient parasites, Supplementary data

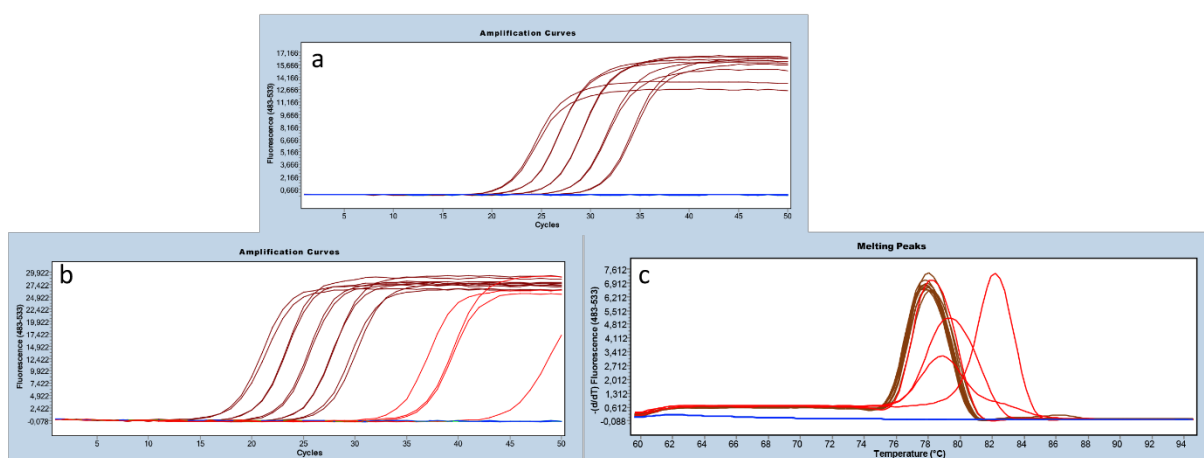
Genus	species	PCR product 1	PCR product 2
<i>Taenia</i>	<i>solium</i>	Tae23 CATCTAGCTGTTTGGTGAATTTTTTATG T.....G.....	Tae32 AAGTTATAGAAAGAAGGGTGTGTGAAATGTTGGGTTTATTTTATTTTATTGGTTATGGGAGC.....A.....G.....
	<i>saginata</i>	.G....T...AT..RG..... .G.A...T...AT...G.....A.....A.....A.....G...G.....A.....T..
	<i>asiatica</i>T...AT...G.....A.....A.....A.....G...G.....A.....T..
<i>Echino-coccus</i>	<i>granulosus</i>	Echino5 AAGGCATCCTTGCCCAATGGATCTCGCCAAAAGATCTTGACGAATCCCTTAGGGATCGCTTCCCTGG	Echino23 GGTGATCCTATTTTGTTC AACATATGTTTGGTTTTTGGCCATCCTGAGGTTTATGTGTTA.....H.....A.....Y.....T...T.....A.....C.....
	<i>multilo-cularis</i>Y.A.....T...G.....T..R.....T...G.....G..
<i>Diphyllo-bothrium</i>	<i>latum</i>	Diphyllo2 TGTTTACGGTG	Diphyllo23 ATGTTTCTATAGTTTGTAGGTAGAGT
	<i>dendri-ticum</i>A..AY.....C..C..G....T..
	<i>nihon-kaiense</i>A...C.....T..
<i>Ascaris</i>	<i>lumbri-coides and suum</i>	Asc2 TTTTGACTGGTACTTTTTTGGCTTTTTATGT.....R.....AC.....G.....	Asc4 TTTATAGGTTTTTTCAGGCCATTTTTGATGGTGTAAACTTTTAAAGAARGAGCAGAY.....T.....A..G.....
<i>Trichuris</i>	<i>trichiura</i>	Trich3 CGATGTTGAATCATTGTATATA	Trich4 TTGCAGAAGAACTTTATTTTCGTCTGTTTCGACGATGGA
<i>Dicro-coelium</i>	<i>dendri-ticum</i>	Dicro22 GTGGAGATGTGTCTACGGAGTCGTGGCT	Dicro6.1 GGGATTGGTTGTCTTGGTTGGTT
	<i>chinensis</i>A.....

<i>Fasciola</i>	<i>hepatica</i>	Fas2 ATTATTTGTTGCC YT TATTGTTAGG T ATTCY.....G...	Fas3 CCTGAAAATCTACTCT C ACACAAGCGATACACGTGTGACCGTCA
	<i>gigantica</i>	.Y.....A.....T..GC.T..G.....T..GC.T..G.....A..A..T..GC.T..... T
<i>Enterobius</i>	<i>vermicularis</i>	Entero2 CTGATGTTTCATGT	Entero4 GTTTTTAGTTG

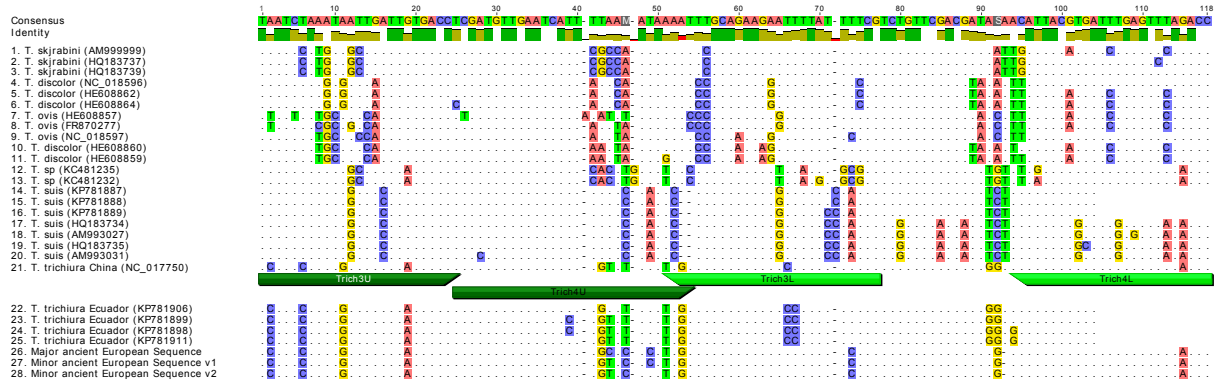
S2 Table. Sequences of the insert of the PCR products showing the variability found in published sequences within and between species. When several related sequences were available, only the positions differing with the reference sequence are indicated, whereas identical positions are represented by dots. The IUPAC nucleotide ambiguity code was used.

Multiplex 1 4 mM MgCl ₂		Multiplex 2 5 mM MgCl ₂		Multiplex 3 4mM MgCl ₂		Multiplex 4 5 mM MgCl ₂	
primers	[μ M]	primers	[μ M]	primers	[μ M]	primers	[μ M]
Tae32	0.15	Tae23	0.2	Trich3	0.25	Dicro6.1	0.2
Echino5	0.08	Trich4	0.08	Entero4	0.08	Asc4	0.2
Diphyllo2	0.08	Dicro22	0.08			Fas3	0.08
		Entero2	0.08			Fas2	0.08
		Diphyllo23	0.08			Echino23	0.08
						Asc2	0.08

S3 Table. Experimental conditions for the various multiplex PCRs. Optimal MgCl₂, and primer concentrations used for each individual multiplex PCR are shown.



S1 Fig. Examples of simplex PCR with primer pairs giving rise (b,c) or not (a) to primer dimers. Standard curves performed in duplicates with serial 5-fold dilution of reference DNA are colored in brown, whereas the no template controls (NTC) are represented in red when they give rise to amplification products, and in blue when they do not. a-b) amplification phase of the qPCR, c) melting curve phase of the experiment displayed in panel b. In this latter experiment, four of the six NTCs which give rise to primer dimers have various distinct melting temperatures (T_m). The T_m of the dimer is usually different from that of the product, but it happens sometimes that these values are similar, as shown here for one of the dimers. In such cases, only electrophoretic analyses can distinguish between dimers and PCR products. Three of these dimers were generated at a Ct between 33 and 36, which is similar to the cycles of the PCR where products corresponding to rare initial molecules typical of ancient samples are also detected. A primer pair with such properties is thus not recommended for the detection of ancient DNA molecules. Results obtained with an optimal primer pair have been displayed in panel a) for comparison.



S2 Fig. Comparison of the ancient and modern genetic diversity of *Trichuris* species as revealed through the combination of the overlapping Trich3 and Trich4 PCR fragments. The divergence of the primers from sequences of animal species can be seen. Most mismatches between the primers and modern and ancient sequences correspond to G-T mismatch, which are the least destabilizing mismatches (e.g., [1]).

1. Pan S, Sun X, Lee JK (2006) DNA stability in the gas versus solution phases: a systematic study of thirty-one duplexes with varying length, sequence, and charge level. *J Am Soc Mass Spectrom* 17: 1383-1395.

NN

3. gff file for mapping and downstream analyses

S3 File. ref_parasites.gff

##gff-version 3

##source-version geneious 6.1.8

```

ref_parasitesGeneious misc_feature      85   89   .   .   .   gene_id "Taenia23";
ref_parasitesGeneious misc_feature     229  232  .   .   .   gene_id "Taenia32";
ref_parasitesGeneious misc_feature     385  395  .   .   .   gene_id "Echinococcus5";
ref_parasitesGeneious misc_feature     538  546  .   .   .   gene_id "Echinococcus23";
ref_parasitesGeneious misc_feature     668  672  .   .   .   gene_id "Diphyllobothrium2";
ref_parasitesGeneious misc_feature     782  784  .   .   .   gene_id "Diphyllobothrium23";
ref_parasitesGeneious misc_feature     898  902  .   .   .   gene_id "Ascaris2";
ref_parasitesGeneious misc_feature    1036 1044  .   .   .   gene_id "Ascaris4";
ref_parasitesGeneious misc_feature    1185 1187  .   .   .   gene_id "Trichuris3";
ref_parasitesGeneious misc_feature    1320 1328  .   .   .   gene_id "Trichuris4";
ref_parasitesGeneious misc_feature    1450 1452  .   .   .   gene_id "Dicrocoelium22";
ref_parasitesGeneious misc_feature    1571 185   .   .   .   gene_id "Dicrocoelium61";
ref_parasitesGeneious misc_feature    1684 1704  .   .   .   gene_id "Fasciola2";
ref_parasitesGeneious misc_feature    1809 1813  .   .   .   gene_id "Fasciola3";
ref_parasitesGeneious misc_feature    1937 1938  .   .   .   gene_id "Enterobius2";
ref_parasitesGeneious misc_feature    2044 2047  .   .   .   gene_id "Enterobius4"
ref_parasitesGeneious primer_bind      51   72   .   +   .   gene_id "Tae23U";
ref_parasitesGeneious primer_bind     174  198   .   +   .   gene_id "Tae32U";
ref_parasitesGeneious primer_bind     337  357   .   +   .   gene_id "Echino5U";
ref_parasitesGeneious primer_bind     496  514   .   +   .   gene_id "Echino23U";
ref_parasitesGeneious primer_bind     645  663   .   +   .   gene_id "Diphyllo2U";
ref_parasitesGeneious primer_bind     747  766   .   +   .   gene_id "Diphyllo23U";
ref_parasitesGeneious primer_bind     864  236   .   +   .   gene_id "Asc2U";
ref_parasitesGeneious primer_bind     988  1012  .   +   .   gene_id "Asc4U";
ref_parasitesGeneious primer_bind    1142 1166  .   +   .   gene_id "Trich3U";
ref_parasitesGeneious primer_bind    1266 1292  .   +   .   gene_id "Trich4U";
ref_parasitesGeneious primer_bind    1407 1430  .   +   .   gene_id "Dicro22U";
ref_parasitesGeneious primer_bind    1533 1561  .   +   .   gene_id "Dicro61U";
ref_parasitesGeneious primer_bind    1660 1679  .   +   .   gene_id "Fas2U";
ref_parasitesGeneious primer_bind    1779 1798  .   +   .   gene_id "Fas3U";
ref_parasitesGeneious primer_bind    1914 1936  .   +   .   gene_id "Entero2U";
ref_parasitesGeneious primer_bind    2020 2039  .   +   .   gene_id "Entero4U";
ref_parasitesGeneious primer_bind_reverse 102  123   .   -   .   gene_id "Tae23L";
ref_parasitesGeneious primer_bind_reverse 261  286   .   -   .   gene_id "Tae32L";
ref_parasitesGeneious primer_bind_reverse 425  445   .   -   .   gene_id "Echino5L";
ref_parasitesGeneious primer_bind_reverse 577  594   .   -   .   gene_id "Echino23L";
ref_parasitesGeneious primer_bind_reverse 675  696   .   -   .   gene_id "Diphyllo2L";
ref_parasitesGeneious primer_bind_reverse 796  813   .   -   .   gene_id "Diphyllo23L";
ref_parasitesGeneious primer_bind_reverse 915  937   .   -   .   gene_id "Asc2L";
ref_parasitesGeneious primer_bind_reverse 1071 1091  .   -   .   gene_id "Asc4L";
ref_parasitesGeneious primer_bind_reverse 1190 1215  .   -   .   gene_id "Trich3L";
ref_parasitesGeneious primer_bind_reverse 1332 1356  .   -   .   gene_id "Trich4L";
ref_parasitesGeneious primer_bind_reverse 1459 1482  .   -   .   gene_id "Dicro22L";
ref_parasitesGeneious primer_bind_reverse 1586 1609  .   -   .   gene_id "Dicro61L";
ref_parasitesGeneious primer_bind_reverse 1709 1728  .   -   .   gene_id "Fas2L";
ref_parasitesGeneious primer_bind_reverse 1843 1863  .   -   .   gene_id "Fas3L";
ref_parasitesGeneious primer_bind_reverse 1950 1969  .   -   .   gene_id "Entero2L";
ref_parasitesGeneious primer_bind_reverse 2051 2072  .   -   .   gene_id "Entero4L";

```

4. Analyses script

S4 File. aMPlexmpileup.sh

#Before launching the script, the bam files from the Ion Torrent can be renamed with a more convenient name:

```
ls -l *.{bam,bai} | awk '{print $9}' > oldNames.txt
```

#Open the oldNames.txt in a text editor, (eventually in Excel) and type in a second column (or separated by a tab) the new names for each file, with a .bam or .bai (as appropriate) extension in the file name.

#Choose the name according to your objective, usually we use the sample name and replicate number.

Create a text file with the two columns: RenameOldNew.txt Be careful to have Linux type Line Feed, in particular, change them if you have been using Excel

Change both BAM and bai file names with the following command

```
cat RenameOldNew.txt | while read old new; do mv $old $new; done
```

When copying the script, the fasta and the gff files from this Word document, it is essential to use the Linux way of handling the newline instruction (use either NotePad++ or the dos2unix command)

#Save the script starting from the following line and name it as aMPlexmpileup.sh

```
#! /bin/bash
```

Script to analyze amplicons from the aMPlex Torrent workflow

It uses as input files the BAM files resulting from the mapping to the reference sequence with the Ion Torrent Mapping Alignment Program (TMAP). It is necessary to use a fasta file with the same name in the first line as that used for mapping.

It requires in the working directory fasta and gff files from the reference sequence, with each amplicon being named using the attribute misc_feature. It also requires the software featureCounts, samtools, bcftools, and VarScan.

A symbolic link to the VarScan.v2.3.7.jar should be in the same folder.

Modify as desired within the script the names of the fasta and gff files (highlighted).

Run the script as follows to keep a log file (here Report_aMPlex_script.txt) with standard error and standard output together: sh aMPlexmpileup.sh >>Report_aMPlex_script.txt 2>&1

Create a symbolic link to the VarScan jar file using the proper address where you stored it (here ~/Software/VarScan.v2.3.7/VarScan.v2.3.7.jar)

```
ln -s ~/Software/VarScan.v2.3.7/VarScan.v2.3.7.jar ./VarScan.v2.3.7.jar
```

Convert BAM to SAM

```
for file in *.bam
```

```
do
```

```
echo $file
```

```
samtools view -h $file > $file".sam"
```

```
done
```

Clean extension tags

```
for file in *.bam.sam
```

```
do
```

```
mv "$file" "${file%.bam.sam}.sam"
```

```
done
```

Determine the number of reads per amplicon

```
featureCounts -a ref_parasites.gff -t misc_feature -R -o CountAmplicons.count *.sam
```

Determine the number of reads that includes the primers (to assess primer dimers (by subtracting the number of Amplicon reads) and other artefacts)

```
featureCounts -a ref_parasites.gff -t primer_bind -R -o CountUpperPrimers.count *.sam
```

```
featureCounts -a ref_parasites.gff -t primer_bind_reverse -R -o CountLowerPrimers.count *.sam
```

```

# Identify variants. These lines identify the variants with each sample analyzed in a sequencing run in a
different file.
# Mpileup generates a pileup of read bases using the alignments to the reference sequence.
# Varscan detects variation based on the reference sequence and produces a table with the percentage of
reads with each SNP.
# To permit species identification from the varscan files, discriminative SNP in the reference sequence
should be indicated using the degeneracy IUPAC code (e.g., Y, R, W, ...).

for file in *.bam
do
samtools mpileup -f ref_parasites.fasta $file | java -jar VarScan.v2.3.7.jar mpileup2snp >$file.varscan
done

# Clean extension tags
for file in *.bam.varscan
do
mv "$file" "${file%.bam.varscan}.varscan"
done

# Produce consensus sequences in the fastq format, assuming that the vcfutils.pl script is located in the
following path: /usr/share/samtools/vcfutils.pl (or modify script accordingly)
# Because degeneracy in the reference decreases the accuracy of the consensus sequences obtained, a
sequence without degeneracy is provided with a different file name (here ref_parasites2.fasta) but with
the same name in the first line of the fasta file.

for file in *.bam
do
#use this command if you have installed version 1.x (at least up to 1.2) of samtools and bcftools
samtools mpileup -uf ref_parasites2.fasta $file | bcftools call -c | vcfutils.pl vcf2fq > $file.fq
#If you have an older version of samtools installed (0.1.19 and earlier), you should use this version
instead (uncomment it, and comment the previous line)
#samtools mpileup -uRf ref_parasites2.fasta $file | bcftools view -cg - | /usr/share/samtools/vcfutils.pl
vcf2fq > $file.fq
done

# Clean up extension tags
for file in *.bam.fq
do
mv "$file" "${file%.bam.fq}.fq"
done

# Clean up files
mkdir samfile
mkdir countread
mkdir Varscan
mkdir Consensus
mkdir featureCount
mkdir BAMs
mv *.sam samfile
mv *.count countread
mv *.summary countread
mv *.varscan Varscan
mv *.fq Consensus
mv *.featureCounts featureCount

```

```
mv *.bam BAMs
mv *.bai BAMs

# Convert the consensus fastq files in the Consensus folder in fasta

# Recover the sequence from the fastq file (single sequence file)
cd Consensus
for file in *.fq
do
sed -r '/^+/q' $file | sed -r 's/^+//g' | sed -r '@/d' > $file.seq
done

# Clean extension tags
for file in *.fq.seq
do
mv "$file" "${file%.fq.seq}.seq"
done

# Create a fasta file by using the name of the file as the name of the sequence
for file in *.seq
do
echo \>$file | sed 's/.seq//' >tempfile
cat $file >>tempfile
mv tempfile $file.fa
done

# Clean extension tags
for file in *.seq.fa
do
mv "$file" "${file%.seq.fa}.fa"
done

# Final clean up
rm *.seq
mkdir Fasta
mv *.fa Fasta
```

Genus	species	Diagnostic SNP
<i>Taenia</i>	<i>solium</i>	74:A, 80:C, 85:T, 86:G, 207:G, 220:G, 229:T, 236:A, 241:A, 250:G, 259:A
	<i>saginata</i>	74:G, 80:T, 85:A, 86:T, 207:A, 220:A, 229:A, 236:G, 241:G, 250:A, 259:T
	<i>asiatica</i>	74:A, 80:T, 85:A, 86:T, 207:A, 220:A, 229:A, 236:G, 241:G, 250:A, 259:T
<i>Echinococcus</i>	<i>granulosus</i>	562:A,C,T
	<i>multilocularis</i>	562:G
<i>Diphyllobothrium</i>	<i>latum</i>	671:G, 781:T, 784:T, 787: A, 793: A
	<i>dendriticum</i>	671:A, 781:C, 784:C, 787: G, 793: T
	<i>nihonkaiense</i>	671:A, 781:T, 784:C, 787: A, 793: T
<i>Dicrocoelium</i>	<i>dendriticum</i>	1578:G
	<i>chinensis</i>	1578:A
<i>Fasciola</i>	<i>hepatica</i>	1696:A, 1697:T, 1815:C
	<i>gigantica</i>	1696:G, 1697:C, 1815:T

S4 Table. Location of the diagnostic SNPs for species discrimination to interpret the varscan data files when using the ref_parasites fasta file for mapping the reads.

S5 File. *Ascaris* and *Trichuris* sequences

>Ascaris_Asc4_Most_Ancient_samples

GCGTATTGGYCCTAATAAGGTTAGTTTTATAGGTTTTTGCAGGCCATTTTTGATGGTGTAAACTTTTAAAGAA
GGAGCAGATGACTCCTYTGAATTCTTCGG

>Ascaris_Asc4_Sample_H

GCGTATTGGYCCTAATAAGGTTAGTTTTATAGGTTTTTGCAGGCTATTTTTGATGGTGTAAACTTTTGAAGAA
GGAGCAGATGACTCCTYTGAATTCTTCGG

>Ascaris_Asc4_Sample_Na

GCGTATTGGYCCTAATAAGGTTAGTTTTATGGGTTTTTGCAGGCTATTTTCGATGGTGTAAACTTTTGAAGAA
GGAGCAGATGACTCCTYTGAATTCTTCGG

>Trichuris_Trich3-4_Major_ancient_European_Sequence

TCATCCAAATGATTGATTATGACCTCGATGTTGAATCATTTGCACACATAGTTTGCAGAAGAATTTTATTCTCGT
CTGTTTCGACGATAGACATTACGTGATTTGAGTTTAAACC

>Trichuris_Trich3-4_Minor_ancient_European_Sequence_v1

TCATCCAAATGATTGATTATGACCTCGATGTTGAATCATTTGTACACATAGTTTGCAGAAGAATTTTATTCTCGT
CTGTTTCGACGATAGACATTACGTGATTTGAGTTTAAACC

>Trichuris_Trich3-4_Minor_ancient_European_Sequence_v2

TCATCCAAATGATTGATTATGACCTCGATGTTGAATCATTTGTACATATAGTTTGCAGAAGAATTTTATTCTCGT
CTGTTTCGACGATAGACATTACGTGATTTGAGTTTAAACC

*PARTIE 3 : RÉSULTATS ET
DISCUSSION*

Dans cette partie, les résultats des analyses génétiques sont exposés, en précisant s'ils ont été obtenus par la méthode décrite dans la publication scientifique « *pool* de PCR » ou par l'approche classique de séquençage de trois répliquions de PCR. Pour chaque site, un tableau, mentionnant le nombre maximal de séquences obtenues par espèces, complètera la description des résultats. Une discussion, par site ou par grande période (Antiquité, Moyen-Âge) sera ensuite proposée, pour décrire l'apport du génotypage moléculaire à la caractérisation des sites archéologiques. Un tableau récapitulatif des résultats de paléoparasitologie microscopique et moléculaire conclura cette partie.

Chapitre I : Les sites préhistoriques

I. La Draga, Espagne

A. *Les résultats de la paléoparasitologie moléculaire*

Cinq échantillons ont fait l'objet d'analyses de paléogénétique, et un seul a généré des résultats. Les premières analyses génétiques de l'échantillon P91 par la méthode classique des trois réplifications de PCR multiplex séquencées individuellement n'avaient permis d'obtenir que sept séquences d'*Ascaris* sp. et neuf d'*Enterobius vermicularis*. Cependant, en appliquant la méthodologie proposée dans au chapitre précédent, nous avons pu identifier, après extraction et purification d'un nouvel extrait d'ADN, des séquences provenant des parasites *Taenia saginata*, *Enterobius vermicularis* et *Ascaris* sp. Nous avons également retrouvé des séquences de *Dicrocoelium dendriticum* et de *Trichuris trichiura*. Malgré l'application des deux approches sur les quatre autres vestiges, aucune séquence n'a été obtenue pour ce site.

Le Tableau 2 présente les résultats comparatifs entre la microscopie et le séquençage, pour l'échantillon P91.

Tableau 2 : résultats comparatifs de microscopie et de paléogénétique pour l'échantillon P91 de La Draga. La méthode classique correspond au séquençage des trois réplifications de PCR. La méthode des « pool » est celle décrite dans la publication scientifique. Le genre Capillaria n'était pas ciblé génétiquement par notre approche.

Échantillons	Œufs intacts identifiés	ADN obtenu	Nombre de séquences méthode classique	Nombre de séquences méthode « pool »
La Draga (LaDragP91)	<i>Ascaris</i> sp.	<i>Ascaris</i> sp.	7	70
	<i>Trichuris</i> sp.	<i>Trichuris trichiura</i>	0	7
	Taeniidae	<i>Taenia saginata</i>	0	9
	<i>Capillaria</i> sp.	<i>Enterobius vermicularis</i>	9	19
		<i>Dicrocoelium dendriticum</i>	0	17

B. Discussion et conclusions

Un seul échantillon de La Draga (P91), parmi les cinq analysés par paléogénétique, a donné des résultats au séquençage. Ceci peut être dû à l'ancienneté du site (jusqu'à 7250 BP) et à des conditions peut-être moins propices à la préservation de l'ADN dans des couches d'occupation recouvertes par le niveau du lac pendant des milliers d'années.

Étant donné le contexte d'échantillonnage (couches d'habitation), les sédiments sont probablement un mélange de matières provenant des humains et des nombreux animaux vivant sur le site. Les œufs de Taeniidae peuvent provenir de *Taenia* ou d'*Echinococcus*, puisque des canidés, étaient présents sur le site. Les analyses génétiques nous ont permis d'établir la présence de *Taenia saginata*, à savoir l'espèce utilisant le bœuf comme hôte intermédiaire et l'être humain comme hôte définitif.

Le séquençage confirme aussi la présence d'*Ascaris* sp. bien que nous ne pouvons pas conclure si les hôtes infectés étaient humains ou porcins. Il peut y avoir un mélange de parasites ayant touché les deux espèces, étant donnée l'origine de l'échantillon. Cette situation est probable, car les animaux vivaient en contact étroit avec les humains, ce qui favorisait les échanges de parasites. Cependant, cette étude ciblait le gène *Nad1* et *Cytb*, qui ne permettent pas de déterminer les haplogroupes du parasite, de préciser ces hypothèses ni d'en émettre de nouvelle sur la présence de plusieurs souches sur le site .

Associées à l'identification de *Trichuris trichiura*, les analyses confirment que les habitants du site étaient touchés par le péril fécal, leur nourriture et leur eau étant contaminées avec des matières fécales d'origine animale et/ou anthropique.

Les œufs fragiles d'*Enterobius vermicularis* n'avaient pas été retrouvés intacts dans les échantillons observés en microscopie mais nous avons identifié ce parasite par notre approche génétique. Jusqu'à ce jour, un unique site européen a identifié ces œufs sur un vestige allemand médiéval (Herrmann, 1988) et aucun travaux de paléogénétique n'avait ciblé ce parasite sur des échantillons issus du Vieux Continent. Nous pouvons conclure avec nos résultats que les infestations par ce nématode touchaient déjà l'Homme au début du Néolithique en Europe de l'Ouest.

Finalement, le fait de trouver de l'ADN de *Dicrocoelium* sp. en absence d'œuf nous place dans une situation difficile quant à la conclusion d'une infection des humains. Nous avons montré que les œufs peuvent être absents de l'échantillon observé en microscopie alors que l'ADN a été préservé. Mais de par le cycle trixénique de ce parasite, incluant entre autre une fourmi parmi les hôtes intermédiaire, on ne peut pas exclure qu'un arthropode soit à l'origine de la présence de cet ADN. Par ailleurs, si des œufs sont présents sur le site, mais non-observés dans les prélèvements effectués, ils pourraient aussi provenir aussi des ruminants présents dans le village.

II. Arbon-Bleiche III, Suisse

A. Les résultats de la paléoparasitologie moléculaire

Aucune séquence de parasite n'a pu être mise en évidence dans les échantillons analysés. Étant donné la faible quantité de matériel disponible, des réplifications d'extraction et purification de l'ADN n'ont pas pu être effectuées.

B. Discussion et conclusions

Après le site de La Draga, Arbon-Bleiche III est le deuxième site lacustre dont le prélèvement des vestiges ensevelis n'a pas permis de mener de manière concluante des analyses génétiques. Dans ce contexte, la paléogénétique ne permet pas de préciser les conclusions déjà proposée via les observations microscopiques, à savoir la présence de plusieurs parasites pouvant être associés à l'ichtyophagie (*Diphyllbothrium* sp., *Dioctophyma* sp. et *Opistorchis* sp). Les autres taxons peuvent avoir aussi bien infecté les habitants du village que les animaux domestiques ou sauvages des environs.

L'absence de résultat génétique est probablement due à un problème de préservation de l'ADN, soit du fait des conditions du site, soit lié aux conditions de conservation post-fouille des échantillons, les analyses de microscopie optique ayant été effectuées entre 2002 et 2005.

III. Passel, France

A. Les résultats de paléoparasitologie moléculaire

Les analyses ont débuté par le séquençage d'un vestige pour lequel trois taxons de parasites avaient pu être identifiés en microscopie. Aucune séquence d'helminthes n'a pu être obtenue avec la méthode « *aMPlex Torrent* ». Nous avons utilisé lors de cette étape 10 grammes de sédiments.

Dans l'objectif de mieux appréhender le potentiel de préservation de l'ADN du site et vu le nombre important d'échantillons à analyser, nous avons choisi de procéder de la manière suivante : deux grammes de chacun des 14 prélèvements ont été regroupés de manière à créer trois échantillons. Cinq répliques de PCR et le « *pool* » ont été séquencés pour chacun de ces trois extraits. Aucune séquence n'a été obtenue.

B. Discussion et conclusions

Ce site, le troisième plus ancien que nous ayons analysé par une approche génétique, n'a pas permis de séquencer d'amplicons provenant d'helminthes gastro-intestinaux. Il est possible que ce résultat soit lié à une mauvaise préservation de l'ADN du fait de l'ancienneté du site. Cependant, le site de Menneville (présenté dans le paragraphe suivant) est daté de 5500-4700 BC et des séquences d'helminthes ont été obtenues, ainsi que sur le site de La Draga (7200 BP). Des conditions géochimiques et environnementales particulières au site de Passel ont pu contribuer à une mauvaise préservation des molécules biologiques.

Pour confirmer ces résultats négatifs, il serait préférable de traiter individuellement chacun des 14 prélèvements par la méthode « *aMPlex Torrent* ». En débutant la purification avec 10 grammes de sédiments pour chaque vestige biologique (plutôt que deux grammes), nous augmentons la probabilité de récupérer des molécules provenant d'helminthes. Le « *pool* » des répliques de PCR pourraient ensuite être séquencés, ce qui nous permettrait de constater si de l'ADN a pu être effectivement préservé. Selon les résultats, il est ensuite possible de procéder au séquençage individuel de chaque réplique, pour mieux apprécier la diversité de parasites présents.

IV. Menneville, France

A. *Les résultats de la paléoparasitologie moléculaire*

Des analyses génétiques ont été effectuées sur les échantillons de sédiment associés au squelette humain. Des séquences du parasite *Ascaris* sp. ont été identifiées au niveau du bassin et de la tête. Au niveau du vestige provenant du bassin, des séquences de *Dicrocoelium dendriticum* ont aussi été identifiées.

Le Tableau 3 récapitule les résultats obtenus pour le site de Menneville.

Tableau 3 : résultats comparatifs de microscopie et de paléogénétique pour le site de Menneville. La méthode des « pool » est celle décrite dans la publication scientifique.

Échantillons	Œufs intacts identifiés	ADN obtenu	Nombre de séquences méthode « pool »
Mdv-bassin	Aucun	<i>Ascaris</i> sp.	>1000
		<i>Dicrocoelium dendriticum</i>	17
Mdv-crâne (contrôle)	Ascaridés minéralisés	<i>Ascaris</i> sp.	300

B. *Discussion et conclusions*

Concernant les séquences de *Dicrocoelium* qui ont été identifiées en absence d'œufs observés, le même constat que pour le site de La Draga s'impose : il est impossible de conclure que le parasite infestait la femme inhumée car les séquences peuvent provenir d'un des deux autres hôtes intermédiaires du parasite (la fourmi et l'escargot).

Ascaris sp. étant un parasite pouvant infecter les cochons et les humains, l'association physique des squelette de ces deux espèces explique certainement l'obtention de séquences du parasites dans le bassin et sous le crâne. Cette découverte du même parasite sur deux squelettes humain et animal est particulièrement intéressante dans ce contexte Néolithique. En effet, cela soulève la question de l'origine de la contamination pour chacune des deux espèces.

Dans le cas présent, l'étude comparative de l'haplogroupe (basé sur les travaux de Betson et al. 2014) du parasite trouvé sous le bassin à celui retrouvé associé au porc pourrait apporter des informations importantes sur l'histoire du genre. Est-ce que le porc et la femme partageait un parasite de même haplogroupe ? Si c'est le cas, ceci soutiendrait l'hypothèse que les deux espèces peuvent être deux hôtes alternatifs à *Ascaris*, une contamination croisée étant déjà possible au Néolithique. Dans le cas contraire, cela pourrait signifier qu'à cette période, en France, chacun des deux hôtes possibles présentait un haplogroupe spécifique. D'autres questions peuvent être soulevées : qu'en est-il de la diversité des haplogroupes au Néolithique ? Comment cette diversité évolue-t-elle au cours du temps ? Quelques-unes de ces questions pourraient trouver réponse par une étude ciblée du gène *Cox1* du Néolithique jusqu'à aujourd'hui.

V. Les mines de sel de Chehr Abad, Iran

A. *Les résultats de la paléoparasitologie moléculaire*

Deux échantillons (ChehP81 et ChehP85) provenant des couches d'occupation du site ont été analysés par l'approche de paléogénétique. Les parasites *Trichuris trichiura*, *Tania saginata*, *Enterobius vermicularis*, *Ascaris* sp. et *Dicrocoelium dendriticum* ont été mis en évidence. Un bilan des résultats est proposé au Tableau 4.

Tableau 4 : résultats comparatifs de microscopie et de paléogénétique du site de Chehr Abad. La méthode classique correspond au séquençage des trois réplifications de PCR. La méthode des « pool » est celle décrite dans la publication scientifique.

échantillons	Œufs intacts identifiés	ADN obtenu	Nombre de séquences méthode classique	Nombre de séquences méthode « pool »
Cheh-P81	<i>Taeniidae</i>	<i>Taenia saginata</i>	0	>1000
	<i>Trichuris</i> sp.	<i>T. trichiura</i>	14	7
	<i>Enterobius vermicularis</i>	<i>E. vermicularis</i>	0	>1000
Cheh-P85	<i>Ascaris</i> sp.	<i>Ascaris</i> sp.	5	6
	<i>Taeniidae</i>	<i>Dicrocoelium dendriticum</i>	0	5

B. Discussion et conclusions

Les résultats des analyses confirment le mauvais état sanitaire dans lequel vivaient les mineurs. Les oxyures, qui infectent principalement les enfants dans le monde actuel, ont été retrouvés chez ces humains adultes. La présence d'*Ascaris* en association avec *Trichuris trichiura* suggère la consommation d'aliments ou d'eau souillés par des matières fécales.

Concernant leur régime alimentaire, nous pouvons conclure, par l'identification génétique de *Taenia saginata*, que c'est la consommation de viande de bœuf crue ou mal cuite qui explique la présence des embryophores de Ténia retrouvés.

L'obtention de séquences de *Dicrocoelium* sp., en absence d'œuf, ne permet pas de conclure que celles-ci proviennent du parasite ayant infecté un humain, et non pas des restes d'une fourmi présente dans l'échantillon. Sur le même site, mais non-analysés dans le cadre des travaux exposés ici, des œufs de *Dicrocoelium* sp. ont été observés dans des coprolithes d'herbivores (Nezamabadi *et al.*, 2013a). Cela suggère que des ruminants (chèvres ou moutons) étaient les hôtes les plus probables du parasite, sans que cela exclue toute dicrocoeliose humaine.

La présence de parasite typique des équidés indique que des chevaux ou ânes étaient présents sur le site, peut-être pour aider aux travaux de la mine ou pour le transport du sel extrait de la mine.

Nous n'avons pour l'instant que peu de connaissances sur le mode de vie de ces travailleurs et donc de l'origine de ces infestations. Étaient-ils tous logés sur place, ou bien habitaient-ils un village à proximité ? Provenaient-ils tous du même village ? Ces informations seraient nécessaires à la compréhension de la transmission des multiples parasitoses dont les mineurs étaient atteints.

Chapitre II : Les sites antiques

I. L'Amphithéâtre de Metz : les résultats de la paléoparasitologie moléculaire

De l'ADN des parasites *Trichuris trichiura*, *Ascaris* sp., *Taenia solium* et *Enterobius vermicularis* ont été identifiés dans les deux vestiges analysés par l'approche de paléogénétique.

II. Jeu de Paume, Beauvais : les résultats de la paléoparasitologie moléculaire

Les analyses de paléogénétique d'un des échantillons ont mis en évidence la présence de *Trichuris trichiura*, *Ascaris* sp. et *Dicrocoelium dendriticum*.

III. Mas de Vignoles, Nîmes : les résultats de la paléoparasitologie moléculaire

Les parasites *Fasciola hepatica*, *Ascaris* sp., *Diphyllobothrium latum*, *Trichuris trichiura*, *Dicrocoelium dendriticum* et *Enterobius vermicularis* ont été identifiés par l'approche paléogénétique.

IV. Discussion et conclusions sur les sites antiques

Selon les sites étudiés, la concordance entre les taxons observés et les ADN séquencés varie (Tableau 5).

Tableau 5 : résultats comparatifs de microscopie et paléogénétique pour les sites antiques français. La méthode classique correspond au séquençage des trois réplifications de PCR. La méthode des « pool » est celle décrite dans la publication scientifique.

Sites (échantillons)	Œufs intacts identifiés	ADN obtenu	Nombre de séquences méthode classique	Nombre de séquences méthode « pool »
Amphi- théâtre de Metz	<i>Fasciola</i> sp.	<i>Enterobius vermicularis</i>	10	53
	<i>Diphyllobothrium</i> sp.	<i>Diphyllobothrium latum</i>	9	0
	<i>Ascaris</i> sp.	<i>Ascaris</i> sp.	303	>1000
	<i>Trichuris</i> sp.	<i>Trichuris trichiura</i>	948	>1000
	Taeniidae	<i>Taenia solium</i>	0	>1000
Jeu de Paumes, Beauvais (Bv-JP42)	<i>Ascaris</i> sp.	<i>Ascaris</i> sp.	525	598
	<i>Trichuris</i> sp.	<i>Trichuris trichiura</i>	722	306
	<i>Dicrocoelium</i> sp.	<i>Dicrocoelium dendriticum</i>	235	>1000
	<i>Fasciola</i> sp.			
Mas de Vignoles (MdV-P6)	<i>Trichuris</i> sp.	<i>Trichuris trichiura</i>	0	317
	<i>Ascaris</i> sp.	<i>Ascaris</i> sp.	0	44
	<i>Fasciola</i> sp.	<i>Fasciola hepatica</i>	30	0
	<i>Dicrocoelium</i> sp.	<i>Dicrocoelium dendriticum</i>	0	>1000
	<i>Capillaria</i> sp.	<i>Diphyllobothrium latum</i>	0	100
		<i>Enterobius vermicularis</i>	23	0

Globalement, les populations françaises de l'Antiquité étaient parasitées par une grande diversité d'helminthes gastro-intestinaux. Les analyses mettent en évidence une nouvelle fois la présence d'oxyures (*Enterobius vermicularis*), alors qu'aucun œuf n'avait été observé par l'approche de microscopie. La présence de cette espèce indique des mauvaises conditions d'hygiène au quotidien, avec une absence de lavage des mains par exemple. Les contaminations ont pu se produire en touchant des objets souillés par des œufs, puis en portant les mains à la bouche ou lors de la préparation d'aliments.

Malgré la présence de système de gestion des eaux usées via des puits, sorte de latrines, les contaminations par péril fécale devaient être fréquentes. En effet, l'association *Trichuris / Ascaris*, de transmission oro-fécale, a été retrouvée sur les trois sites étudiés, tant au niveau de la détermination par microscopie optique que par la génétique.

Bien que des signes de la présence de bœufs aient été retrouvé à Beauvais, sous la forme de produit transformé (colle) et à Metz (restes de faune), c'est du *Tænia solium*, associé à la consommation de porc mal cuit ou cru, qui a été identifié. La consommation de viande de bœuf devait malgré tout être commune à l'époque et dans ces régions de France. Il est possible qu'en analysant un plus grand nombre d'échantillons, des œufs de Taeniidae soient plus fréquemment retrouvés et que les analyses génétiques identifient aussi du *T. saginata*, transmis par le bœuf et présentant un cycle de vie similaire au *T. solium*.

L'identification de *Diphyllobothrium* indique une consommation de poissons d'eau douce, mal cuits ou crus, par les populations occupant l'Amphithéâtre. L'espèce *D. latum* est celle infectant le plus fréquemment les êtres humains.

Finalement, il est difficile d'établir si les parasites *Fasciola hepatica* et *Dicrocoelium dendriticum* ont effectivement contaminé des humains. En effet, les vestiges analysés sont soit un mélange de sédiments provenant de puits ou de latrines, soit un coprolithe appartenant à un mammifère. Il est possible, étant donné la présence d'ovo-caprins sur les sites, que les œufs et les séquences d'ADN obtenus proviennent de ces animaux, qui sont les hôtes définitifs les plus fréquents pour ces deux parasites. Cependant, due à l'association dans un même coprolithe d'un parasite affectant les ruminants ou les consommateurs de végétaux (*Fasciola*) et d'un parasite affectant uniquement les mammifères monogastriques ichtyophages (*Diphyllobothrium*), l'origine humaine de cet échantillon est très probable.

Chapitre III : Les sites médiévaux français

I. Rue Saint-Laurent, Beauvais : les résultats de la paléoparasitologie moléculaire

Il a été possible de retrouver dans les deux vestiges analysés des séquences génétiques d'*Ascaris* sp. *Trichuris trichiura*, *Enterobius vermicularis* et *Tænia solium*.

II. Place des Véens, Auxerre : les résultats de la paléoparasitologie moléculaire

Les parasites *Dicrocoelium dendriticum*, *Ascaris* sp. *Fasciola hepatica* et *Trichuris trichiura* ont été identifiés dans l'échantillon analysé par l'approche génétique. Dans ce cas encore, la présence de *T. trichiura*, parasite humain, atteste de l'origine anthropique d'une partie du comblement du puits. Les autres taxons, bien que pouvant affecter l'être humain, peuvent aussi être d'origine animale.

III. Les Turqueries, Calais: les résultats de la paléoparasitologie moléculaire

Étant donné le nombre important de coprolithes à analyser, les vestiges ont été regroupés en trois sous-échantillons (environ deux grammes par coprolithe ont été utilisés) qui ont été séquencés individuellement. Les analyses paléogénétiques ont montré la présence de *Dicrocoelium dendriticum* et *Fasciola hepatica*. *Trichuris trichiura* n'a pas été identifié par l'approche génétique, soit du fait d'une faible quantité d'ADN préservé sur le site, soit parce qu'il s'agit d'une autre espèce (*T. suis* par exemple, affectant le porc). *Ascaris* sp. n'a pas non plus été retrouvé par cette approche. Enfin, les deux derniers taxons, *Dioctophyma* sp. et *Capillaria* sp., ne font pas partie des genres ciblés génétiquement car ils ne sont pas suffisamment renseignés dans les banques de données génétiques à l'heure actuelle.

IV. Discussion et conclusions sur ces trois sites médiévaux

Le Tableau 6 récapitule les résultats obtenus par la microscopie et la génétique. Concernant le site d'Auxerre, les analyses sont parfaitement concordantes. Pour les deux autres, les deux approches sont complémentaires, car certains taxons ont été retrouvés seulement avec l'un ou l'autre des méthodes : *Enterobius vermicularis* à Beauvais, *Trichuris* et *Ascaris* à Calais.

Tableau 6 : résultats comparatifs de microscopie et de paléogénétique pour les sites médiévaux français. La méthode classique correspond au séquençage des trois répliquions de PCR. La méthode des « pool » est celle décrite dans la publication scientifique. Les taxons *Capillaria* et *Dioctophyma* n'ont pas été ciblés par l'approche génétique.

Sites (échantillons)	Œufs intacts identifiés	ADN obtenu	Nombre de séquences méthode classique
Beauvais rue St- Laurent (P39)	<i>Ascaris</i> sp.	<i>Ascaris</i> sp.	977
	<i>Dicrocoelium</i> sp.	<i>Enterobius vermicularis</i>	110
	<i>Trichuris</i> sp.	<i>T. trichiura</i>	623
	<i>Fasciola</i> sp.		
Beauvais rue St- Laurent (P76)	<i>Ascaris</i> sp.	<i>Ascaris</i> sp.	>1000
	<i>Dicrocoelium</i> sp.	<i>Enterobius vermicularis</i>	209
	<i>Trichuris</i> sp.	<i>T. trichiura</i>	>1000
	Taeniidae <i>Fasciola</i> sp.	<i>Tenia solium</i>	5
Auxerre Place des Véens (P27)	<i>Ascaris</i> sp.	<i>Ascaris</i> sp.	233
	<i>Dicrocoelium</i> sp.	<i>D. dendriticum</i>	43
	<i>Trichuris</i> sp.	<i>T. trichiura</i>	25
	<i>Fasciola</i> sp.	<i>F. hepatica</i>	214
Calais	<i>Trichuris</i> sp.	<i>Fasciola hepatica</i>	13
	<i>Ascaris</i> sp.	<i>Dicrocoelium dendriticum</i>	7
	<i>Capillaria</i> sp.		
	<i>Dioctophyma</i> sp.		(analysé par la méthode des « pool »)

Malgré une amélioration de l'organisation sanitaire des villes françaises, les populations vivant entre le 8^{ème} et le 14^{ème} siècle sont atteintes de plusieurs parasitoses. Comme cela était le cas pour les sites antiques vu précédemment, l'association *Trichuris* / *Ascaris* est aussi retrouvée dans tous les sites médiévaux étudiés. Les conditions sanitaires ne permettaient donc pas encore d'éradiquer le péril fécal et les populations avaient probablement accès à de l'eau souillée par des œufs de ces deux nématodes.

La présence d'oxyures avait déjà été montrée après analyse des sites antiques français et nous constatons que les populations médiévales étaient aussi infestées. Les conditions d'hygiène au quotidien ne se sont sans doute pas améliorées durant cette époque, ce qui pourrait expliquer que ce parasite reste encore présent. Une meilleure prise en charge des malades dans les hôpitaux, comme à l'Hôtel-Dieu de Beauvais, peut aussi entraîner une contamination des soignants au moment des soins.

Bien que le développement de la charcuterie soit décrit sur le site d'Auxerre, nous n'y avons pas identifié la présence de *Taenia* sp. par microscopie ou par paléogénétique. Cette pratique alimentaire, basée sur l'utilisation des matières grasses et des abats, n'entraîne habituellement pas de contamination par ces vers plats. En effet, l'infestation des êtres humains se produit par consommation de tissus musculaires sous forme de viande mal cuite et infectée par des cysticerques. Seul le *T. asiatica* peut être transmis lors de la consommation d'abats, mais sa répartition semble être restreinte à l'Asie (Galan-Puchades et Fuentes, 2013). Le parasite *T. solium* a pu être identifié sur le site de Beauvais, signifiant la consommation de viande de porc mal cuite.

La présence d'œufs et/ou d'ADN de *Dicrocoelium dendriticum* ne permet pas de conclure que cette espèce avait effectivement infecté les humains. La présence d'œufs confirment qu'un mammifère était infecté, mais les échantillons proviennent de latrines et sont constitués d'un mélange de matière organique de plusieurs animaux. La même conclusion s'impose à propos de *Fasciola hepatica*, à l'exception du site de Calais, car les échantillons sont des coprolithes.

V. Teteghem, France

A. *Les résultats de la paléoparasitologie moléculaire*

Pour les analyses de biologie moléculaire, étant donné le nombre important de prélèvements effectués (au moins deux par squelette), auxquels s'ajoutent les contrôles, les sédiments provenant des squelettes ont été regroupés et homogénéisés avant leur traitement. Des séquences des parasites *Ascaris* sp. et *Echinococcus granulosus* ont été obtenus.

Le Tableau 7 récapitule les résultats obtenus pour ce site.

Tableau 7 : résultats comparatifs de microscopie et paléogénétique pour le site de Teteghem. Le taxon *Capillaria* n'a pas été ciblé par l'approche génétique.

Sites	Œufs intacts identifiés	ADN obtenu	Nombre de séquences méthode classique	Nombre de séquences méthode « pool »
Teteghem	<i>Ascaris</i> sp.	<i>Ascaris</i> sp.	0	160
	<i>Trichuris</i> sp.	<i>Echinococcus granulosus</i>	110	6
	<i>Fasciola</i> sp.			
	<i>Capillaria</i> sp.			

B. *Discussion et conclusions*

On constate une faible concordance entre les œufs identifiés et les séquences d'ADN obtenues, à l'exception du parasite *Ascaris* sp. L'ensemble des taxons identifiés suggèrent que le sédiment de la fosse a pu être mélangé à des matières fécales d'origine anthropique et/ou animale. La présence d'*Ascaris* et *Trichuris* signe la présence d'une infestation humaine ou porcine. L'absence de résultats génétiques pour *Trichuris* ne permet pas de distinguer *T. suis*, infectant le porc, de *T. trichiura*, touchant l'être humain. Le genre *Capillaria* parasite de nombreux animaux, incluant les mammifères.

Dans ce contexte, il apparaît délicat de corréler les hypothèses de mortalité de masse aux résultats que nous avons obtenus. Malgré l'absence de séquences nucléotidiques de *Fasciola* sp., des œufs ont été observés, mais en nombre si faible (un ou deux par échantillon), qu'il semble difficile de mettre en cause ce parasite dans une mortalité de masse des six ovidés.

Les moutons sont les réservoirs naturels pour *E. granulosus*. Ils sont l'hôte intermédiaire du parasite, qui va se développer sous forme de kystes dans les organes. Des mammifères carnivores se contamineront suite à la consommation de cette viande infestée, et libèreront dans leurs fèces des œufs, qui seront ingérés par les moutons lors du broutage. Bien qu'une échinococcose puisse être fatale aux ruminants, le temps variable d'incubation du développement de ce cestode, de quelque mois à des années, laisse à penser qu'il ne serait pas impliqué dans la mortalité de masse. Le fait d'avoir retrouvé de l'ADN en absence d'œufs suggère que certains moutons étaient tout de même probablement porteurs de ce parasite.

Chapitre IV : Les sites modernes

I. Le Fort Saint-Sébastien, France

A. Les résultats de la paléoparasitologie moléculaire

Les analyses génétiques ont confirmé la présence de *Dicrocoelium dendriticum*, *Ascaris* sp., *Fasciola hepatica* et *Trichuris trichiura*. Les espèces *Tænia saginata*, *T. solium* et *Diphyllobothrium nihonkaiense* ont par ailleurs aussi été retrouvées. Aucun parasite n'a été identifié dans le contrôle ni dans la latrine 4329. Le Tableau 8 résume les résultats obtenus.

Tableau 8 : résultats comparatifs de microscopie et paléogénétique pour le site de Saint-Germain-en-Laye. La méthode des « pool » est celle décrite dans la publication scientifique.

Échantillons	Oeufs intacts identifiés	ADN obtenu	Nombre de séquences méthode « pool »
Sgl4327	<i>Trichuris</i> sp.	<i>Diphyllobothrium nihonkaiense</i>	>1000
		<i>Dicrocoelium dendriticum</i>	17
Sgl4329	Aucun oeuf	Aucune séquence	
Sgl5270	<i>Ascaris</i> sp. <i>Trichuris</i> sp. <i>Fasciola</i> sp. <i>Dicrocoelium</i> sp.	<i>Ascaris</i> sp.	>1000
		<i>T. trichiura</i>	86
		<i>F. hepatica</i>	80
		<i>D. dendriticum</i> <i>Tænia solium</i>	>1000 278
Sgl7183	<i>Ascaris</i> sp. <i>Trichuris</i> sp. <i>Fasciola</i> sp. <i>Dicrocoelium</i> sp.	<i>Ascaris</i> sp.	782
		<i>T. trichiura</i>	>1000
		<i>F. hepatica</i>	19
		<i>D. dendriticum</i> <i>Tænia saginata</i>	>1000 233
Sgl-contrôle	Aucun œuf	Aucune séquence	

B. Discussion et conclusions

L'état de santé des soldats était une préoccupation des Rois de France depuis le 13^{ème} siècle, bien que des hôpitaux dédiés aux militaires ne fassent leur apparition que sous Louis XII, à Calais et Brouage. Louis XIV, surnommé le Roi-Soleil, est aussi parfois appelé le Roi-soldat, du fait de son intérêt pour l'entraînement de ses troupes et de l'état de santé des soldats au retour des combats. Dans cette optique, il fit construire en 1670 l'Hôtel royal des Invalides, pouvant prendre en charges 3000 infirmes. En 1708, un édit royal lança la création d'un Service de Santé des Armées, alors que les finances du royaume sont au plus bas. L'objectif de ce service est de créer un soutien à la santé, au plus près des armées (Moysan et Bernicot, 2008).

Les latrines du camp d'entraînement étant composées d'un mélange de matière organique (fèces et urine) émit par plusieurs personnes, sur une période de temps pouvant s'étendre sur les deux années d'occupation du site, nous ne pouvons pas exclure que certains individus puissent avoir été poly-parasités. On peut aussi supposer que les installations sanitaires et d'hébergement étaient fortuits, puisque le camp avait été érigé spécifiquement pour un entraînement ponctuel, en vue de la prise de Maastricht. Cet environnement a pu être propice à une transmission oro-fécale entre soldats des parasites *Ascaris* et *Trichuris*. Il n'est pas possible de se prononcer sur le moment de l'infestation par les autres parasites (avant l'arrivée au fort), mais la présence de restes de bœufs et la consommation de poissons en conserve ont pu être la source du *Taenia saginata* et *Diphyllobothrium nihonkaiense* respectivement.

La présence d'œufs et de séquences génétiques de *Dicrocoelium* signifie que les matières fécales de l'hôte définitif de cette douve ont pu se retrouver dans les latrines. Elles peuvent provenir d'un mouton, bien qu'aucun reste d'ovidés n'ait été identifié sur le site. Étant donnée la présence d'autres parasites de l'Homme dans ces latrines, il est enfin possible que ces œufs proviennent d'une pseudo-parasitose humaine, lors de leur passage, sans infestation, chez un soldat.

Les soldats du 17^{ème} siècle étaient particulièrement bien entraînés, et très peu mourraient sur les champs de bataille. Leur décès était plutôt attribuable à des gangrènes, consécutives à des blessures mal soignées, ou à des fièvres contagieuses, telle la fièvre typhoïde. Les analyses microscopiques et génétiques ayant révélé la présence de six parasites différents chez les soldats, nous pouvons supposer qu'ils n'étaient pas dans un état de santé optimal lors de leur départ au combat (Bois, non-daté).

II. Baillif, Guadeloupe

A. *Les résultats de paléoparasitologie moléculaire*

Un échantillon, positif pour *E. histolytica*, a été analysé par l'approche de paléogénétique et l'ADN du parasite *Trichuris trichiura* a été identifié. Après extraction et purification de l'ADN, trois réplifications des PCR multiplex ont été séquencés individuellement.

B. *Discussion et conclusions*

Les résultats de la paléoparasitologie sont en concordance avec les observations préliminaires du bon état sanitaire des personnes décédées. Nous avons identifié génétiquement une seule espèce de parasite, *Trichuris trichiura*, ce qui va dans le sens des conclusions posées par la microscopie et l'analyse anthropologique des personnes inhumées. Ces faits suggèrent le décès d'individus jeune *a priori* en bonne santé juste avant leur décès, qui a pu être causé de manière foudroyante par une dysenterie, le paludisme ou une fièvre jaune. Le fort pourcentage (16%) de diagnostic immunologique d'*Entamoeba histolytica* n'est pas surprenant dans la région, ce taux pouvant atteindre 36% dans certaines régions de l'Amazonie (Santos *et al.*, 2013).

Il pourrait être intéressant de poursuivre l'analyse génétique des helminthes gastro-intestinaux car un unique échantillon a été analysé, alors qu'une quarantaine de vestiges sont disponibles. Ces données permettraient de compléter les diagnostics immunologiques et d'avoir une meilleure vision de l'état de santé des soldats inhumés.

La mise en évidence d'ADN de *Trichuris* nous a convaincu que notre approche d'analyse d'échantillons, sans se baser sur la présence ou non d'œufs, est bien adaptée. En effet, il est possible que des œufs de *Trichuris* se trouvaient dans le sous-échantillon analysé par génétique mais absent du prélèvement analysé par microscopie. De même, tous les œufs ont pu être très mal préservés, donc non-identifiables, contrairement à l'ADN dont les molécules ont pu établir des interactions physico-chimiques avec les particules minérales du sol et ainsi être préservées.

III. Les sites alsaciens de la Première Guerre Mondiale

A. Les résultats de la paléoparasitologie moléculaire

Il a été possible d'identifier génétiquement les parasites *Ascaris* sp., *Trichuris trichiura*, *Fasciola hepatica* et *Taenia asiatica*.

B. Discussion et conclusions

L'association de parasites retrouvée sur les deux sites indique un mauvais état de santé des soldats et des conditions d'hygiène non-satisfaisantes durant la Grande Guerre. Les taxons diagnostiqués dans les latrines indiquent que leur utilisation était strictement dédiée à l'Homme, ce qui n'est pas surprenant vue le contexte historique. L'identification de *T. asiatica* est insolite car il se contracte habituellement après consommation de porc cru d'origine asiatique. Les soldats allemands n'avaient probablement pas accès à ce type d'aliment au cours des batailles, mais il est possible que l'infestation concernait un unique individu qui avait séjourné en Asie avant la guerre.

Les infections humaines à *Capillaria* sont rares et les parasitologues sont d'avis que retrouver des œufs dans des coprolithes peut être associé à du pseudo-parasitisme : les œufs sont ingérés et repassent directement dans les fèces sans causer de trouble. C'est ce qui semble s'être produit dans le cas du squelette dont l'abdomen contenait des œufs de Capillariidae. Les analyses paléoparasitologiques suggèrent que des rats, retrouvés sur le site, était porteur d'*Eucoleus gastricus*, et que les œufs auraient pu être accidentellement ingérés par le soldat (Le Bailly *et al.*, 2014a).

Tableau récapitulatif des résultats

Tableau 9 : récapitulatif des résultats microscopiques et génétiques des sites analysés.

Localisation et nom du site	Datation	Type d'échantillons	Nom des échantillons	Taxons observés en microscopie	Séquences d'ADN obtenues
Espagne La Draga	7250-7050 BP	Sédiments de couches d'habitation	LaDragP91 LaDragP25 LaDragP35 La DragP81 LaDragP94	<i>Ascaris</i> , Tæniidae, <i>Trichuris</i> , <i>Capillaria</i> <i>Trichuris</i> , <i>Ascaris</i> , Tæniidae <i>Trichuris</i> , <i>Ascaris</i> , <i>Capillaria</i> , Tæniidae, <i>Paramphistomum</i> <i>Trichuris</i> , <i>Ascaris</i> , <i>Capillaria</i> <i>Trichuris</i> , <i>Ascaris</i> , <i>Capillaria</i> , Tæniidae, <i>Paramphistomum</i>	<i>Ascaris</i> , <i>T. trichiura</i> , <i>T. saginata</i> , <i>E. vermicularis</i> , <i>D. dendriticum</i> } Aucune séquence
Suisse Arbon-Bleiche III	5334-5320 BP	Sédiments de couches d'habitation	Arb3146s	<i>Diphyllobothrium</i> , <i>Fasciola</i> , <i>Trichuris</i>	Aucune séquence
France Passel	6250-5650 BP	Sédiments de fossé d'enceinte	Pas1 à Pas14	<i>Trichuris</i> , Tæniidae, <i>Capillaria</i> , <i>Paramphistomum</i> , <i>Dicrocoelium</i>	Aucune séquence

Localisation et nom du site	Datation	Type d'échantillons	Nom des échantillons	Taxons observés en microscopie	Séquences d'ADN obtenues
France Menneville Derrière le village	~7000 BP	Sédiments en rapport avec un squelette humain	Mdv-bassin Mdv-crâne-ctrl	Aucun œuf identifiable Présence d'œufs très minéralisés de type Ascaridés	<i>Ascaris, Dicrocoelium dendriticum</i> <i>Ascaris</i>
Iran Chehr Abad	2500-1500 BP	Sédiments de couches d'habitation	ChehP85 ChehP81	Tæniidae, <i>Ascaris</i> Tæniidae, <i>Enterobius, Trichuris</i>	<i>Ascaris, D. dendriticum</i> <i>T. saginata, E. vermicularis, T. trichiura</i>
France, Metz, Amphithéâtre	284-476 AD	Coprolites	Mtz-US6420 Mtz-US12189	<i>Diphyllobothrium, Fasciola</i> <i>Ascaris, Trichuris, Tæniidae</i>	<i>Trichuris trichiura, Ascaris</i> <i>T. trichiura, E. vermicularis, T. solium, Ascaris, D. latum</i>
France Beauvais, Jeu de Paume	150 AD	Sédiments de latrines	Bv-JP-42	<i>Ascaris, Trichuris, Dicrocoelium, Fasciola</i>	<i>Ascaris, T. trichiura, D. dendriticum</i>
France Nîmes, Mas de Vignoles	600-800 AD	Sédiments provenant de puits	MdV-P6	<i>Trichuris, Ascaris, Fasciola, Dicrocoelium, Capillaria</i>	<i>T. trichiura, Ascaris, F. hepatica, D. dendriticum, Diphyllobothrium latum, Enterobius vermicularis</i>
France Beauvais, Rue St-Laurent	12-14 ^{ème} siècle AD	Sédiments de latrines	Bv-rsl-P39 Bv-rsl-P76	<i>Ascaris, Dicrocoelium, Trichuris, Fasciola</i> <i>Ascaris, Dicrocoelium, Trichuris, Fasciola, Tæniidae</i>	<i>Ascaris, T. trichiura, Enterobius vermicularis</i> <i>Ascaris, T. trichiura, T. solium, Enterobius vermicularis</i>

Localisation et nom du site	Datation	Type d'échantillons	Nom des échantillons	Taxons observés en microscopie	Séquences d'ADN obtenues
France Auxerre, Place de Véens	11-12 ^{ème} siècle AD	Sédiments provenant de puits	Aux-P27	<i>Trichuris, Fasciola, Ascaris, Dicrocoelium</i>	<i>T. trichiura, F. hepatica, Ascaris, D. dendriticum</i>
France Calais, Les Turqueries	8-10 ^{ème} siècle AD	Coprolithes provenant d'un fossé	13 échantillons Cal-X	<i>Ascaris, Trichuris, Capillaria, Dictyophyma</i>	<i>D. dendriticum, F. hepatica</i>
France Teteghem, Carline3	13 ^{ème} siècle AD	Sédiments de la région thoracique et abdominale de 6 moutons	Tet 1,6	<i>Fasciola, Ascaris, Trichuris, Capillaria</i>	<i>Ascaris, Echinococcus granulosus</i>
France Saint-Germain- en-Laye, Fort Saint-Sébastien	1669-1671	Sédiments provenant de tonneaux de latrines	Sgl-4329 Sgl-4329-ctrl Sgl-7183 Sgl-4327 Sgl-5270	Aucun œuf identifiable Aucun œuf identifiable <i>Ascaris, Trichuris, Dicrocoelium, Fasciola</i> <i>Trichuris</i> <i>Ascaris, Trichuris, Dicrocoelium, Fasciola</i>	Aucune séquence Aucune séquence <i>Ascaris, T. trichiura, D. dendriticum, F. hepatica, Taenia saginata</i> <i>Diphyllobothrium nihonkaiense, Dicrocoelium dendriticum</i> <i>Ascaris, T. trichiura, D. dendriticum, F. hepatica, Taenia solium</i>
Guadeloupe Baillif	19 ^{ème} siècle AD	Sédiments sous squelette humain	Bail-46-2	Aucun œuf identifiable	<i>Trichuris trichiura</i>

Localisation et nom du site	Datation	Type d'échantillons	Nom des échantillons	Taxons observés en microscopie	Séquences d'ADN obtenues
France Entzheim	Première Guerre Mondiale	Sédiments de tonneaux de latrines	Entz-P12 Entz-P8	<i>Ascaris, Trichuris</i> Tæniidae, <i>Ascaris, Trichuris</i>	<i>Ascaris, Tænia asiatica, Fasciola hepatica</i> <i>Ascaris, T. trichiura, Fasciola hepatica</i>

Conclusions générales

Les analyses paléogénétiques présentées dans ce chapitre nous ont permis d'obtenir des résultats pour des échantillons datés jusque 7200 BP. L'étude de divers types de sédiments, provenant de latrines, de couches d'habitation, de puits, ou sous des squelettes, ont précisé, confirmé ou complété les analyses microscopiques qui avaient été menées au préalable. Notre approche, cherchant à cibler une grande diversité d'helminthes gastro-intestinaux humains, sans se restreindre aux taxons identifiés en microscopie apparaît donc comme tout à fait appropriée. En effet, nous avons obtenu des résultats génétiques en absence d'œufs identifiables (site de Menneville, de la Guadeloupe), et compléter l'étude de la diversité présente sur un site lorsque des œufs de quelques taxons avaient été retrouvés (La Draga, Mas de Vignoles, Saint-Germain-en-Laye).

Les sites lacustres semblent être moins propices à la préservation de l'ADN. Des six échantillons analysés de La Draga et d'Arbon-Bleiche, un seul nous a permis d'obtenir des résultats après séquençage. Un séjour sous l'eau en climat méditerranéen (site espagnol) ou tempéré (site suisse) n'est probablement pas propice à la préservation de l'ADN sédimentaire, contrairement aux sites scandinaves présentés dans la partie II. Ces deux sites sont aussi parmi les plus vieux du corpus étudiés (7200 BP pour la Draga, 5300 BP pour Arbon-Bleiche) et, associés aux résultats négatifs pour le site de Passel (6250 BP), il semblerait que, tout comme pour les squelettes, la préservation de l'ADN pour des échantillons de plus de 5000 ans, est très variable (Pruvost *et al.*, 2007). Dans d'autres cas, comme pour le site de Menneville (~7000 BP), les conditions physico-chimiques du sédiment, et éventuellement le mode d'inhumation, ont permis de préserver l'ADN des helminthes gastro-intestinaux présents.

La composition et la structure de l'œuf semblent également avoir un impact sur la préservation de l'ADN. Ceux d'*Ascaris* sp., composés de trois couches épaisses de lipides et de chitine (Lysek *et al.*, 1985), sont régulièrement observés en microscopie et de l'ADN a pu être amplifié dans plusieurs vestiges archéologiques. La structure de ces œufs peut permettre une meilleure préservation de l'ADN, contrairement à des œufs plus fragiles, aussi observés moins fréquemment (*Taenia* sp., *Enterobius vermicularis* ...).

Dans d'autres cas, la présence d'un opercule (*Trichuris* sp. ou *Diphyllobothrium* sp. par exemple) pourrait être *a priori* un point négatif en offrant une moins bonne protection pour l'ADN présent dans l'œuf. Nous avons génétiquement mis en évidence le parasite *T. trichiura* dans dix sites archéologiques, ce qui signifie que l'œuf a tout de même pu offrir un bon milieu de préservation, ou que l'ADN se trouvait plutôt en interaction avec les particules sédimentaires. Le cas du *Diphyllobothrium* est différent : l'œuf mature éclot et libère la larve après un séjour d'une vingtaine de jours dans l'eau. Il est ainsi possible d'observer en microscopie des œufs vides et d'obtenir des résultats négatifs par la génétique, si aucun tissu ne reste présent après l'éclosion. C'est possiblement la situation du site lacustre d'Arbon-Bleiche.

L'analyse d'un nombre important d'échantillons français, du Néolithique au 20^{ème} siècle, ayant donné des résultats positifs pour le parasite *Ascaris* sp. nous amène maintenant à considérer une étude plus poussée du genre. En effet, ce contexte se prête assez bien à l'étude des haplogroupes du gène *Cox1*, déjà entamée sur des spécimens modernes. Même si les données génétiques modernes actuellement disponibles ne permettent pas de distinguer les espèces d'*Ascaris* affectant l'être humain et le porc, l'analyse de l'haplogroupe pourrait permettre de mieux comprendre l'impact de la domestication sur les infestations ayant touché l'Homme.

L'approche moléculaire adoptée lors de ce travail a également permis la mise en évidence de parasites rarement identifiés par l'approche de microscopie optique. C'est le cas en particulier de l'oxyure de l'Homme. Nous avons montré une présence ancienne, depuis 7200 BP en Europe de l'Ouest, du parasite *Enterobius vermicularis*, alors que ses œufs sont observés principalement dans des vestiges d'Amérique ou d'Asie. Ce résultat est un des points cruciaux de cette approche, qui montre à quel point il est important de coupler plusieurs méthodes d'études. Ainsi, nous avons pu démontrer, que même si les approches microscopiques et génétiques pouvaient fonctionner indépendamment l'une de l'autre, l'approche combinée permettait d'augmenter et de préciser la diversité des taxons mis en évidence.

CONCLUSIONS ET PERSPECTIVES

Conclusions et perspectives

Au cours des travaux de cette thèse, nous avons souhaité analyser par une approche de séquençage à haut-débit la diversité d'espèces d'helminthes gastro-intestinaux d'échantillons archéologiques de natures, de provenances et d'âges différents. Nous avons développé une approche performante, rapide et peu dispendieuse permettant de cribler jusqu'à 96 échantillons simultanément. Les résultats obtenus pour des sites jusqu'à 7200 BP nous ont convaincu que cette approche est complémentaire à la microscopie. Elle peut aussi se suffire à elle-même car nous avons montré que certains taxons avaient pu être détectés en absence d'œuf identifiable. La situation inverse est aussi vraie, car dans certains cas, malgré la présence d'œufs, nous n'avons pas obtenu les séquences d'ADN correspondantes. Cette situation n'est pas surprenante et est probablement due à l'hétérogénéité de la répartition et de la préservation de l'ADN et des œufs dans le sédiment, ainsi que la nature des conditions taphonomiques. C'est pour cette raison qu'il est préférable, lorsque cela est possible, d'effectuer les études de paléoparasitologie par ces deux approches.

Nous avons exposé dans une publication scientifique, présentée dans un chapitre de ce manuscrit, que les travaux de paléogénétique menés sur des extraits composés d'un mélange complexe d'ADN avec un faible nombre de molécules-cibles de départ, pouvaient engendrer de la variabilité quant à la détection des taxons présents. Nous avons proposé des mises au point méthodologiques pour tenir compte de ces aspects et présenter une manière robuste de génotyper les helminthes gastro-intestinaux présents dans des vestiges archéologiques.

La méthode « *aMPlex Torrent* », adaptée à l'analyse des helminthes gastro-intestinaux dans le cadre de cette thèse, présente l'avantage d'être très flexible. En utilisant les amorces appropriées, il est possible de diagnostiquer d'autres parasites en plus de ceux que nous avons présentés, humains ou animaux, tels *Clonorchis sinensis*, *Paramphistomum* sp. ou *Schistosoma* sp. Elle peut aussi être adaptée pour des projets totalement différents : l'étude de la biodiversité des rongeurs marocains (Guimaraes *et al.*, soumis), ou pour cibler sur une unique espèce, plusieurs régions génétiques d'intérêt (humain, chats, chevaux, ânes ...). Les seules limitations sont l'existence de séquences de références modernes et la mise au point des PCR multiplex permettant de regrouper entre-elles plusieurs amorces, sans perdre en efficacité.

Avec l'opportunité d'étudier la diversité des taxons d'helminthes présents dans un nombre important d'échantillons, il est ensuite possible de sélectionner ceux qui présentent un intérêt pour un projet plus particulier. Par exemple, il pourrait être envisageable d'étudier la modification du cycle de vie ou de transmission de parasites comme *Ascaris* sp. ou *Trichuris* sp. avant et après le Néolithique. Avec la mise en place de la domestication, une augmentation de la proximité des êtres humains et des animaux, dont les porcs, cette révolution a pu être une occasion pour ces helminthes d'infester de nouveaux hôtes, ou d'établir des échanges génétiques par introgression. Par exemple, en partageant une même niche écologique, les probabilités d'hybridation entre les *Ascaris* infectant les humains et ceux infectant les porcs ont pu augmenter. En sélectionnant des vestiges recouvrant ces périodes et ayant donné des résultats positifs par la méthode « *aMPlex Torrent* » (utilisée alors comme méthode de criblage à haut débit de la préservation de l'ADN de parasites) ce genre d'analyse plus ciblée peut être possible.

Les résultats exposés dans ce manuscrit et publiés dans l'article scientifique provenaient tous d'échantillons de type sédimentaires : latrines, puits, couches d'habitation, fossé d'inhumation ou entourant un village ou sédiments prélevés sous des squelettes... Il est possible que d'autres types de restes, tel des kystes calcifiés, puissent aussi être associés à une infection parasitaire. Par exemple, le parasite *Echinococcus granulosus* se développe en kystes hydatiques sur les poumons ou le foie, qui pourraient être retrouvés associés à un squelette. Ce genre de structure peut toutefois avoir d'autres d'origines infectieuses, tels un kyste tuberculeux, une brucellose, ou être liés à des processus physiopathologiques non-infectieux (Ortner et Putschar, 1981). Pour analyser ce genre de vestige, il serait alors préférable d'utiliser une approche de séquençage de nouvelle génération non-ciblée, générant une grande quantité de séquences (séquençage *shot gun*).

À ce jour, ce type de séquençage non-ciblé n'a pas permis d'identifier des helminthes gastro-intestinaux. Lors des travaux d'Appelt en 2014, des œufs d'*Ascaris* sp. et *Trichuris* sp. avaient été observés en microscopie et le séquençage direct d'amplicons avaient permis l'identification de l'un de ces deux parasites. Le séquençage *shot gun* de l'échantillon n'a cependant pas mis en évidence de séquences d'helminthes. Cette approche est probablement peu adaptée pour l'identification des parasites lorsqu'une faible quantité d'ADN est préservée et diluée dans l'ADN environnemental. Il faudrait alors utiliser une technologie de séquençage générant beaucoup plus de séquences, ce qui augmenterait considérablement les coûts. Par ailleurs, cette approche non-ciblée n'assurait pas d'obtenir des séquences permettant une identification au niveau de l'espèce.

Pour contrer cette limitation et amener la paléoparasitologie moléculaire à l'ère de la paléogénomique, l'enrichissement en séquences-cibles par une approche de capture par hybridation pourrait être utilisée. Via des appâts permettant d'enrichir en régions génétiques spécifiques à certaines espèces, il serait possible d'analyser des échantillons pour lesquels l'ADN a été faiblement préservé. Avec cette approche, il serait aussi possible de reconstituer de génomes nucléaires complets, ou du moins quelques gènes nucléaires, pouvant apporter des réponses à des questions concernant l'évolution des parasites. Par exemple, en s'intéressant à l'adaptation génétique des helminthes suite à des déplacements géographiques des populations-hôtes ou bien lors d'un changement d'espèce-hôte.

La paléoparasitologie est une discipline jeune, qui a pu bénéficier des avancées récentes en biologie moléculaire. Diverses approches, plus ou moins ciblées, peuvent être mise en place pour étudier les traces d'éventuels parasites préservées dans des échantillons archéologiques. Touchant les domaines de l'anthropologie, de la médecine et de la biologie de l'évolution, la paléoparasitologie offre la possibilité de mieux appréhender les pathologies qui affectaient les populations anciennes. Les infestations par des helminthes intestinaux étant associées aux conditions d'hygiène et aux habitudes alimentaires, l'inventaire des taxons présents sur un site permet aussi d'obtenir des informations précieuses sur le fonctionnement et l'évolution des sociétés.

VALORISATION SCIENTIFIQUE

Valorisation scientifique

I. Publications

- (1) **Côté N.M.L.**, *et al.* (2016) "A new high-throughput approach to genotype ancient human gastrointestinal parasites." *PLoS ONE*, no. 11 (1) : e0146230
- (2) **Côté M.L.N.** *et al.* "Paleogenetics and ancient parasites: the case of *Enterobius vermicularis* in Europe". Invitation de l'éditrice en chef de l'International Journal of Paleopathology.
- (3) Guimaraes S., Pruvost M., Daligault J., Stoetzel E., **Côté M.L.N.** *et al.* "A new metabarcoding approach powerful enough to genotype ~50-70,000 year-old rodent remains from Northern Africa." Soumis à Molecular Ecology Ressources.
- (4) **Côté N.** et Gorgé O. (2013). « La longue coévolution des organismes pathogènes. » *Biofutur*, vol. 32 (349):41-44.

II. Communications orales

- (1) **Côté N.** *et al.* « Étude paléogénétique des helminthes gastro-intestinaux ». 20ème colloque du Groupe des Méthodes Pluridisciplinaires Contribuant à l'Archéologie, 27-29 avril 2015, Besançon, France.
- (2) **Côté N.** *et al.* "Paleogenetics to study ancient parasitic infections". Séminaire interne de l'Institut Jacques Monod, 30 mars 2015.
- (3) **Côté N.** *et al.* "Paleogenetics and paleoparasitology". 3rd International Congress of Biomedical Sciences in Archeology, 6-9 novembre 2014, Bordeaux, France.

- (4) **Côté N. et al.** “Paleogenetics to study ancient parasitic infections”. 20th European Paleopathology Association meeting, 26-29 août 2014, Lund, Suède (Premier prix de la communication orale étudiante).
- (5) Pruvost M., Daligault J., Guimaraes S., **Côté N. et al.** “aMPlex_Torrent: a powerful and economical approach to study past biodiversity”. International Symposium on Biomolecular Archaeology. Basel, Switzerland, august 27-30 2014

III. Communications affichées

- (1) Grange T., Guimaraes S., **Côté N. et al.** “aMPlex Torrent”, a metabarcoding strategy relying on multiplex PCR and next generation sequencing adapted to the high-throughput analysis of ancient remains”. Congrès annuel de la Society for Molecular Biology and Evolution. Vienne, Autriche, 12-16 juillet 2015.
- (2) **Côté N. et al.** « Étude des parasites anciens par une approche paléogénétique ». 19ème forum des Jeunes Chercheurs, juin 13-14 2013, Dijon (Premier prix de la catégorie poster)

RÉFÉRENCES BIBLIOGRAPHIQUES

- Akane A., Matsubara K., Nakamura H. *et al.* (1994) "Identification of the heme compound copurified with deoxyribonucleic acid (DNA) from bloodstains, a major inhibitor of polymerase chain reaction (PCR) amplification." *J. Forensic Sci.*, no. 39 (2):362-72.
- Anantaphruti M., Thaenkham U., Kusolsuk T. *et al.* (2013) "Genetic Variation and Population Genetics of *Taenia saginata* in North and Northeast Thailand in relation to *Taenia asiatica*." *J. Parasitol. Res.*, no. 2013:310605.
- Anderson T.J. (2001) "The dangers of using single locus markers in parasite epidemiology: *Ascaris* as a case study." *Trends Parasitol.*, no. 17 (4):183-8.
- Antolín F., Buxo R., Jacomet S. *et al.* (2014) "An integrated perspective on farming in the early Neolithic lakeshore site of La Draga (Banyoles, Spain)." *Environmental Archaeology*, no. 19 (3):241-255.
- Appelt S., Armougom F., Le Bailly M. *et al.* (2014) "Polyphasic analysis of a middle ages coprolite microbiota, Belgium." *PLoS ONE*, no. 9 (2):e88376.
- Arizono N., Shedko M., Yamada M. *et al.* (2009) "Mitochondrial DNA divergence in populations of the tapeworm *Diphyllobothrium nihonkaiense* and its phylogenetic relationship with *Diphyllobothrium klebanovskii*." *Parasitol. Int.*, no. 58 (1):22-8.
- ASAVA. *Archéologie de la Vallée de l'Aisne*. Association pour le sauvetage archéologique de la Vallée de l'Aisne 2014. Disponible à <http://asava.info/accueil.html>.
- Aspöck H., Auer H. et Picher O. (1996) "*Trichuris trichiura* eggs in the Neolithic glacier mummy of the Alps." *Parasitology today*, no. 12 (7):255-256.
- Aufderheide A.C., Salo W., Madden M. *et al.* (2004) "A 9,000-year record of Chagas' disease." *Proc. Natl. Acad. Sci. U.S.A.*, no. 101 (7):2034-9.
- Aumard S., Büttner S., Cambou D. *et al.* (2013) "*La Place des Véens à Auxerre du IIe au XIVe siècle: "zone artisanale", marigot et cimetière (Yonne)"*, Dijon, vol. 30.
- Bailloud G., Bayle D., Beeching A. *et al.* (1982) "*L'archéologie dans la vallée de l'Aisne*", Revue archéologique de Picardie, vol. 1, Disponible à: http://www.persee.fr/web/revues/home/prescript/article/pica_1272-6117_1982_hos_1_1_1677.
- Barber I. (2013) "Sticklebacks as model hosts in ecological and evolutionary parasitology." *Trends Parasitol.*, no. 29 (11):556-66.
- Barber I. et Scharsack J.P. (2010) "The three-spined stickleback-*Schistocephalus solidus* system: an experimental model for investigating host-parasite interactions in fish." *Parasitology*, no. 137 (3):411-24.

- Bellemain E., Davey M.L., Kauserud H. *et al.* (2013) "Fungal palaeodiversity revealed using high-throughput metabarcoding of ancient DNA from arctic permafrost." *Environ. Microbiol.*, no. 15 (4):1176-89.
- Benson D.A., Karsch-Mizrachi I., Lipman D.J. *et al.* (2005) "GenBank." *Nucleic Acids Res.*, no. 33 (Database issue):D34-8.
- Betson M., Nejsum P., Bendall R.P. *et al.* (2014) "Molecular epidemiology of ascariasis: a global perspective on the transmission dynamics of *Ascaris* in people and pigs." *J. Infect. Dis.*, no. 210 (6):932-41.
- Bezier A., Annaheim M., Herbinier J. *et al.* (2009) "Polydnviruses of braconid wasps derive from an ancestral nudivirus." *Science*, no. 323 (5916):926-930.
- Biagini P., Theves C., Balaesque P. *et al.* (2012) "Variola virus in a 300-year-old Siberian mummy." *N. Engl. J. Med.*, no. 367 (21):2057-9.
- Binois A. "Approche méthodologique des mortalités de masse ovines en archéologie." In *Les dépôts d'ossements d'animaux en France, de la fouille à l'interprétation*. ed. Méniel & Auxiette, sous-presse.
- Biron D.G., Marche L., Ponton F. *et al.* (2005) "Behavioural manipulation in a grasshopper harbouring hairworm: a proteomics approach." *Proc. Biol. Sci.*, no. 272 (1577):2117-26.
- Blum S.A.E., Lorenz M.G. et Wackernagel W. (1997) " Mechanism of retarded DNA degradation and prokaryotic origin of DNases in nonsterile soils." *Systematic and Applied Microbiology*, no. 20:513-521.
- Bois J.-P. date inconnue. "*L'armée française de l'Ancien Régime au XVIIe et au XVIIIe siècle*". Conférence à la Société des Amis du vieux Blaye.
- Bos K.I., Schuenemann V.J., Golding G.B. *et al.* (2011) "A draft genome of *Yersinia pestis* from victims of the Black Death." *Nature*, no. 478 (7370):506-10.
- Bottger E.C. (1990) "Frequent contamination of *Taq* polymerase with DNA." *Clin. Chem.*, no. 36 (6):1258-9.
- Bouchet F., Guidon N., Dittmar K. *et al.* (2003a) "Parasite remains in archaeological sites." *Mem. Inst. Oswaldo Cruz*, no. 98 Suppl 1:47-52.
- Bouchet F., Harter S. et Le Bailly M. (2003b) "The state of the art of paleoparasitological research in the Old World." *Mem. Inst. Oswaldo Cruz*, no. 98 Suppl 1:95-101.

- Braeckman B.P., Houthoofd K. et Vanfleteren J.R. "Intermediary metabolism." In *WormBook*. ed. The C. elegans Research Community, 2008.
- Buffetaut E., *Histoire de la paléontologie*. ed. Presses universitaires de France, Paris, 1998.
- CDC. *Parasites* 2015. Disponible à <http://www.cdc.gov/parasites/>.
- Champlot S. 2010. *Phylogéographie au Pléistocène et à l'Holocène et domestication des équidés*. Thèse de Doctorat, Institut Jacques Monod, Université de Paris-Sud XI, Paris.
- Champlot S., Berthelot C., Pruvost M. *et al.* (2010) "An efficient multistrategy DNA decontamination procedure of PCR reagents for hypersensitive PCR applications." *PLoS ONE*, no. 5 (9).
- Chandler D.P., Fredrickson J.K. et Brockman F.J. (1997) "Effect of PCR template concentration on the composition and distribution of total community 16S rDNA clone libraries." *Mol. Ecol.*, no. 6 (5):475-82.
- Chen B.Y., Janes H.W. et Chen S. (2002) "Computer programs for PCR primer design and analysis." *Methods Mol. Biol.*, no. 192:19-29.
- Chen S., Ai L., Zhang Y. *et al.* (2014) "Molecular detection of *Diphyllobothrium nihonkaiense* in humans, China." *Emerging Infect. Dis.*, no. 20 (2):315-318.
- Chiodini P.L., Engbaek K., Heuck C.C. *et al.* (1993) "*Parasitologie médicale: techniques de base pour le laboratoire*", Organisation Mondiale de la Santé, Genève.
- Cohn E. (1970) "Observations on the feeding and symptomatology of xiphinema and longidorus on selected host roots." *J. Nematol.*, no. 2 (2):167-173.
- Corless C.E., Guiver M., Borrow R. *et al.* (2000) "Contamination and sensitivity issues with a real-time universal 16S rRNA PCR." *J. Clin. Microbiol.*, no. 38 (5):1747-52.
- Côté N.M.L., Daligault J., Pruvost M. *et al.* (2016) "A new high-throughput approach to genotype ancient human gastrointestinal parasites." *PLoS ONE*, no. 11 (1).
- Coudart A. et Demoule J.-P. (1982) "*Le site néolithique et chalcolithique de Menneville*", Revue archéologique de Picardie, vol. 1, Disponible à: http://www.persee.fr/web/revues/home/prescript/article/pica_1272-6117_1982_hos_1_1_1682.
- Crecchio C. et Stotzky G. (1998) "Binding of DNA on humic acids: effect of transformation on *Bacillus subtilis* and resistance to DNase." *Soil Biol. Biochem.*, no. 30 (8-9):1061-1067.

- Cutillas C., Callejon R., de Rojas M. *et al.* (2009) "*Trichuris suis* and *Trichuris trichiura* are different nematode species." *Acta Trop*, no. 111 (3):299-307.
- Dabney J., Knapp M., Glocke I. *et al.* (2013) "Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments." *Proc. Natl. Acad. Sci. U S A*, no. 110 (39):15758-63.
- Darling M.I. et Donoghue H.D. (2014) "Insights from paleomicrobiology into the indigenous peoples of pre-colonial America - a review." *Mem. Inst. Oswaldo Cruz*, no. 109 (2):131-9.
- Direction Générale des Patrimoines, *Bilan Scientifique de la Région Nord-Pas-de-Calais*, Ministère de la Culture et de la Communication, 2013.
- de Araujo A.J., Ferreira L.F., Confalonieri U.E. *et al.* (1985) "The finding of *Enterobius vermicularis* eggs in pre-Columbian human coprolites." *Mem. Inst. Oswaldo Cruz*, no. 80 (2):141-3.
- De Blainville H. (1822) "Analyse des principaux travaux dans les Sciences physiques, publiés dans l'année 1821." *Journal de Physique*, no. XCIV.
- de Wit H.C.D., *Histoire de la biologie du développement*. ed. Pudoc, Lausanne, 1992, vol. 3.
- Deelder A.M., Miller R.L., de Jonge N. *et al.* (1990) "Detection of schistosome antigen in mummies." *Lancet*, no. 335 (8691):724-5.
- Demeke T. et Adams R.P. (1992) "The effects of plant polysaccharides and buffer additives on PCR." *Biotechniques*, no. 12 (3):332-4.
- Desachy B. (1991) "*Le site archéologique de Beauvais (Oise)*", *Revue archéologique de Picardie*, vol. 3-4, Disponible à:
http://www.persee.fr/web/revues/home/prescript/article/pica_0752-5656_1991_num_3_1_1621.
- Dittmar K. (2009) "Old parasites for a new world: the future of paleoparasitological research. a review." *J. Parasitol.*, no. 95 (2):365-71.
- Drancourt M., Aboudharam G., Signoli M. *et al.* (1998) "Detection of 400-year-old *Yersinia pestis* DNA in human dental pulp: an approach to the diagnosis of ancient septicemia." *Proc. Natl. Acad. Sci. U S A*, no. 95 (21):12637-40
- Drancourt M., Tran-Hung L., Courtin J. *et al.* (2005) "*Bartonella quintana* in a 4000-year-old human tooth." *J. Infect. Dis.*, no. 191 (4):607-11.

- Dufour B. et Le Bailly M. (2013) "Testing new parasite egg extraction methods in paleoparasitology and an attempt at quantification." *International Journal of Paleopathology*, no. 3:199-203.
- Dupouy-Camet J. et Peduzzi R. (2004) "Current situation of human diphyllbothriasis in Europe." *Euro Surveill.*, no. 9 (5):31-35.
- Eckert K.A. et Kunkel T.A. (1990) "High fidelity DNA synthesis by the *Thermus aquaticus* DNA polymerase." *Nucleic Acids Res.*, no. 18 (13):3739-44.
- Elliott D.E. et Weinstock J.V. (2012) "Where are we on worms?" *Curr. Opin. Gastroenterol.*, no. 28 (6):551-6.
- Elnifro E.M., Ashshi A.M., Cooper R.J. *et al.* (2000) "Multiplex PCR: optimization and application in diagnostic virology." *Clin. Microbiol. Rev.*, no. 13 (4):559-70.
- Erlwein O., Robinson M.J., Dustan S. *et al.* (2011) "DNA extraction columns contaminated with murine sequences." *PLoS ONE*, no. 6 (8):e23484.
- Fémolant J.-M. (1999) "*Beauvais*", Revue archéologique de Picardie, vol. 16, Disponible à: http://www.persee.fr/web/revues/home/prescript/article/pica_1272-6117_1999_hos_16_1_2056.
- Fernandes A., Iniguez A.M., Lima V.S. *et al.* (2008) "Pre-Columbian Chagas disease in Brazil: *Trypanosoma cruzi* I in the archaeological remains of a human in Peruacu Valley, Minas Gerais, Brazil." *Mem. Inst. Oswaldo Cruz*, no. 103 (5):514-6.
- Ferracci-Porri M., *L'affaire du fantôme de Heilbronn*. ed. Normant, France, 2009.
- Ferreira L.F., Britto C., Cardoso M.A. *et al.* (2000) "Paleoparasitology of Chagas disease revealed by infected tissues from Chilean mummies." *Acta Trop.*, no. 75 (1):79-84.
- Foth B.J., Tsai I.J., Reid A.J. *et al.* (2014) "Whipworm genome and dual-species transcriptome analyses provide molecular insights into an intimate host-parasite interaction." *Nat. Genet.*, no. 46 (7):693-700.
- Franchi M., Ferris J.P. et Gallori E. (2003) "Cations as mediators of the adsorption of nucleic acids on clay surfaces in prebiotic environments." *Orig. Life Evol. Biosph.*, no. 33 (1):1-16.
- Friedberg E.C., Walker G.C. et Siede W., *DNA Repair and Mutagenesis*, 1995.
- Galan-Puchades M.T. et Fuentes M.V. (2013) "*Taenia asiatica*: the most neglected human *Taenia* and the possibility of cysticercosis." *Korean J. Parasitol.*, no. 51 (1):51-4.

- Galvin T.J. (1968) "Development of human and pig *Ascaris* in the pig and rabbit." *J. Parasitol.*, no. 54 (6):1085-91.
- Geigl E.M. "The domestication of cattle: Insights from a joint archaeozoological-palaeogenetical venture." In *Between sand and sea: The archaeology and human ecology of Southwest Asia*. ed. Tübingen Publications in Prehistory, 2011. p. 69-90.
- Gentilini M., *Médecine tropicale*. ed. Lavoisier, Paris, 2012.
- Gijon Botella H., Afonso Vargas J.A., Arnay de la Rosa M. *et al.* (2010) "Paleoparasitologic, paleogenetic and paleobotanic analysis of XVIII century coprolites from the church La Concepcion in Santa Cruz de Tenerife, Canary Islands, Spain." *Mem. Inst. Oswaldo Cruz*, no. 105 (8):1054-6.
- Ginolhac A., Rasmussen M., Gilbert M.T. *et al.* (2011) "mapDamage: testing for damage patterns in ancient DNA sequences." *Bioinformatics*, no. 27 (15):2153-5.
- Green R.E., Briggs A.W., Krause J. *et al.* (2009) "The Neandertal genome and ancient DNA authenticity." *EMBO J.*, no. 28 (17):2494-502.
- Grmek M.D. (1969) "Préliminaires d'une étude historique des maladies." *Annales. Économies, Sociétés, Civilisations*, no. 24 (6):1473-1483.
- Guhl F., Jaramillo C., Vallejo G.A. *et al.* (1999) "Isolation of *Trypanosoma cruzi* DNA in 4,000-year-old mummified human tissue from northern Chile." *Am. J. Phys. Anthropol.*, no. 108 (4):401-7.
- Guimaraes S., Pruvost M., Daligault J. *et al.* (soumis) "A new metabarcoding approach powerful enough to genotype ~50-70,000 year-old rodent remains from Northern Africa."
- Hancock K., Broughel D.E., Moura I.N. *et al.* (2001) "Sequence variation in the cytochrome oxidase I, internal transcribed spacer 1, and Ts14 diagnostic antigen sequences of *Taenia solium* isolates from South and Central America, India, and Asia." *Int. J. Parasitol.*, no. 31 (14):1601-7.
- Harkins K.M., Buikstra J.E., Campbell T. *et al.* (2015) "Screening ancient tuberculosis with qPCR: challenges and opportunities." *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, no. 370 (1660):20130622.
- Hauswirth W.W., Dickel C.D., Rowold D.J. *et al.* (1994) "Inter- and intrapopulation studies of ancient humans." *Experientia*, no. 50 (6):585-91.
- Herrmann B. (1988) "*Parasite Remains from Mediaeval Latrine Deposits: an Epidemiologic and Ecologic Approach. Note et monographie technique*", Paris.

- Higuchi R., Bowman B., Freiberger M. *et al.* (1984) "DNA sequences from the quagga, an extinct member of the horse family." *Nature*, no. 312 (5991):282-4.
- Hindiye M.Y. (2004) "*Integrated Guide to Sanitary Parasitology*", Organisation Mondiale de la Santé- Bureau régional de la Méditerranée de l'est.
- Hofreiter M., Jaenicke V., Serre D. *et al.* (2001) "DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA." *Nucleic Acids Res.*, no. 29 (23):4793-9.
- Hoss M., Jaruga P., Zastawny T.H. *et al.* (1996) "DNA damage and DNA sequence retrieval from ancient tissues." *Nucleic Acids Res.*, no. 24 (7):1304-7.
- Hsu S.J., Tseng P.H. et Chen P.J. (2005) "*Trichuris suis* therapy for ulcerative colitis: nonresponsive patients may need anti-helminth therapy." *Gastroenterology*, no. 129 (2):768-9; author reply 769.
- Hu M., Chilton N.B. et Gasser R.B. (2002) "The mitochondrial genomes of the human hookworms, *Ancylostoma duodenale* and *Necator americanus* (Nematoda: Secernentea)." *Int. J. Parasitol.*, no. 32 (2):145-58.
- Huang B.X., Kim H.Y. et Dass C. (2004) "Probing three-dimensional structure of bovine serum albumin by chemical cross-linking and mass spectrometry." *J. Am. Soc. Mass Spectrom.*, no. 15 (8):1237-47.
- Hurard S. (2012) "*Redécouverte du fort Saint-Sébastien, camp d'entraînement des troupes de Louis XIV dans la vallée de la Seine*", INRAP.
- Hurard S. (2013) "*Le Fort Saint-Sébastien de Saint-Germain-en-Laye. Camp d'entraînement des troupes de Louis XIV.*", JAIF.
- Hurard S. et Hanotaux M. Découverte du fort Saint-Sébastien de Louis XIV. Le Salon noir, France Culture, 20 juin 2012.
- Iniguez A.M., Araujo A., Ferreira L.F. *et al.* (2003a) "Analysis of ancient DNA from coprolites: a perspective with random amplified polymorphic DNA-polymerase chain reaction approach." *Mem. Inst. Oswaldo Cruz*, no. 98 Suppl 1:63-5.
- Iniguez A.M., Reinhard K., Carvalho Goncalves M.L. *et al.* (2006) "SL1 RNA gene recovery from *Enterobius vermicularis* ancient DNA in pre-Columbian human coprolites." *Int. J. Parasitol.*, no. 36 (13):1419-25.
- Iniguez A.M., Reinhard K.J., Araujo A. *et al.* (2003b) "*Enterobius vermicularis*: ancient DNA from North and South American human coprolites." *Mem. Inst. Oswaldo Cruz*, no. 98 Suppl 1:67-9.

- Inrap. "Un fort éphémère pour préparer la prise de Maastricht en 1673 ". Communiqué de presse du 12 juin 2012.
- Inrap. 2015. Communication internet sur le site de Passel.
- Jaeger L.H. et Iniguez A.M. (2014) "Molecular paleoparasitological hybridization approach as effective tool for diagnosing human intestinal parasites from scarce archaeological remains." *PLoS ONE*, no. 9 (8):e105910.
- Jawertz E., Melnick J.L. et Adelberg E.A., *Microbiologie médicale*. ed. Les Presses de l'Université Laval, Québec, 1973.
- Jeon H.-K., Kim K.-H. et Eom K.S. (2007) "Complete sequence of the mitochondrial genome of *Taenia saginata*: Comparison with *T. solium* and *T. asiatica*." *Parasitology International*, no. 56 (3):243 - 246.
- Jeon H.K., Lee K.H., Kim K.H. *et al.* (2005) "Complete sequence and structure of the mitochondrial genome of the human tapeworm, *Taenia asiatica* (Platyhelminthes; Cestoda)." *Parasitology*, no. 130 (Pt 6):717-26.
- Jex A.R., Liu S., Li B. *et al.* (2011) "*Ascaris suum* draft genome." *Nature*, no. 479 (7374):529-33.
- Jex A.R., Nejsum P., Schwarz E.M. *et al.* (2014) "Genome and transcriptome of the porcine whipworm *Trichuris suis*." *Nat. Genet.*, no. 46 (7):701-6.
- Jex A.R., Waeschenbach A., Hu M. *et al.* (2009) "The mitochondrial genomes of *Ancylostoma caninum* and *Bunostomum phlebotomum*--two hookworms of animal health and zoonotic importance." *BMC Genomics*, no. 10:79.
- Jia T.W., Melville S., Utzinger J. *et al.* (2012) "Soil-transmitted helminth reinfection after drug treatment: a systematic review and meta-analysis." *PLoS Negl. Trop. Dis.*, no. 6 (5):e1621.
- Kacki S. et Romon T. (2013) "From burials to population identity: archaeological appraisal of the status of a Lesser Antilles colonial cemetery (Baillif, Guadeloupe)." *Antiquity*, no. 87 (337):829-839.
- Kang S., Sultana T., Eom K.S. *et al.* (2009) "The mitochondrial genome sequence of *Enterobius vermicularis* (Nematoda: Oxyurida)--an idiosyncratic gene order and phylogenetic information for chromadorean nematodes." *Gene*, no. 429 (1-2):87-97.
- Kanz C., Aldebert P., Althorpe N. *et al.* (2005) "The EMBL Nucleotide Sequence Database." *Nucleic Acids Res.*, no. 33 (Database issue):D29-33.

- Khairat R., Ball M., Chang C.C. *et al.* (2013) "First insights into the metagenome of Egyptian mummies using next-generation sequencing." *J. Appl. Genet.*, no. 54 (3):309-25.
- Khan G., Kangro H.O., Coates P.J. *et al.* (1991) "Inhibitory effects of urine on the polymerase chain reaction for *Cytomegalovirus* DNA." *J. Clin. Pathol.*, no. 44 (5):360-5.
- Koboldt D.C., Zhang Q., Larson D.E. *et al.* (2012) "VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing." *Genome Res.*, no. 22 (3):568-76.
- Koskivaara M., Valtonen E.T. et Prost M. (1991) "Dactylogyrids on the gills of roach in central Finland: features of infection and species composition." *Int. J. Parasitol.*, no. 21 (5):565-572.
- Krause J., Dear P.H., Pollack J.L. *et al.* (2006) "Multiplex amplification of the mammoth mitochondrial genome and the evolution of Elephantidae." *Nature*, no. 439 (7077):724-7.
- Kreder C.A. (1996) "Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein." *Appl. Environ. Microbiol.*, no. 62 (3):1102-6.
- Lamarck M., *Histoire naturelle des animaux sans vertèbres*. ed. Déterville, Paris, 1815-1822, vol. 1.
- Lavrov D.V. et Brown W.M. (2001) "*Trichinella spiralis* mtDNA: a nematode mitochondrial genome that encodes a putative ATP8 and normally structured tRNAs and has a gene arrangement relatable to those of coelomate metazoans." *Genetics*, no. 157 (2):621-37.
- Le Bailly M. et Bouchet F. (2010) "Ancient dicrocoeliosis: Occurrence, distribution and migration." *Acta Trop*, no. 115 (3):175-80.
- Le Bailly M. et Bouchet F. (2013) "*Diphyllobothrium* in the past: Review and new records." *Int. J. of Paleopathology*, no. 3 (3):182-187.
- Le Bailly M., Goncalves M.L., Lefevre C. *et al.* (2006) "Parasitism in Kansas in the 1800s: a glimpse to the past through the analysis of grave sediments from Meadowlark cemetery." *Mem. Inst. Oswaldo Cruz*, no. 101 Suppl 2:53-6.
- Le Bailly M., Landolt M. et Bouchet F. (2012) "First World War German soldier intestinal worms: an original study of a trench latrine in France." *J. Parasitol.*, no. 98 (6):1273-5.
- Le Bailly M., Landolt M., Mauchamp L. *et al.* (2014a) "Intestinal parasites in First World War German soldiers from "Kilianstollen", Carspach, France." *PLoS ONE*, no. 9 (10):e109543.

- Le Bailly M., Leuzinger U., Schlichtherle H. *et al.* (2007) " « Crise économique » au Néolithique à la transition Pfyn-Horgen (3400 BC) : contribution de la paléoparasitologie." *Anthropozoologica*, no. 42 (2):175-185.
- Le Bailly M., Romon T. et Kacki S. (2014b) "New evidence of *Entamoeba histolytica* infections in pre-Columbian and colonial cemeteries in the Caribbean." *J. Parasitol.*, no. 100 (5):684-6.
- Le T.H., Pearson M.S., Blair D. *et al.* (2002) "Complete mitochondrial genomes confirm the distinctiveness of the horse-dog and sheep-dog strains of *Echinococcus granulosus*." *Parasitology*, no. 124 (Pt 1):97-112.
- Lee S.U. et Huh S. (2004) "Variation of nuclear and mitochondrial DNAs in Korean and Chinese isolates of *Clonorchis sinensis*." *Korean J. Parasitol.*, no. 42 (3):145-8.
- Leles D., Araujo A., Ferreira L.F. *et al.* (2008) "Molecular paleoparasitological diagnosis of *Ascaris* sp. from coprolites: new scenery of ascariasis in pre-Colombian South America times." *Mem. Inst. Oswaldo Cruz*, no. 103 (1):106-8.
- Leles D., Araujo A., Vicente A.C. *et al.* (2009) "Molecular diagnosis of ascariasis from human feces and description of a new *Ascaris* sp. genotype in Brazil." *Vet. Parasitol.*, no. 163 (1-2):167-70.
- Leles D., Araujo A., Vicente A.C. *et al.* (2010) "ITS1 intra-individual variability of *Ascaris* isolates from Brazil." *Parasitol. Int.*, no. 59 (1):93-6.
- Leles D., Cascardo P., Freire Ados S. *et al.* (2014) "Insights about echinostomiasis by paleomolecular diagnosis." *Parasitol. Int.*, no. 63 (4):646-9.
- Leles D., Gardner S.L., Reinhard K. *et al.* (2012) "Are *Ascaris lumbricoides* and *Ascaris suum* a single species?" *Parasit. Vectors*, no. 5:42.
- Leonard J.A., Shanks O., Hofreiter M. *et al.* (2007) "Animal DNA in PCR reagents plagues ancient DNA research." *J. of Archaeol. Sci.*, no. 34:1361-1366.
- Levy-Booth D.J., Campbell R.G., Rgulden R.H. *et al.* (2007) "Cycling of extracellular DNA in the soil environment." *Soil Biology and Biochemistry*, no. 39:2977-2991.
- Li H., Handsaker B., Wysoker A. *et al.* (2009) "The Sequence Alignment/Map format and SAMtools." *Bioinformatics*, no. 25 (16):2078-9.
- Liao Y., Smyth G.K. et Shi W. (2014) "featureCounts: an efficient general purpose program for assigning sequence reads to genomic features." *Bioinformatics*, no. 30 (7):923-30.
- Lima V.S., Iniguez A.M., Otsuki K. *et al.* (2008) "Chagas disease in ancient hunter-gatherer population, Brazil." *Emerg. Infect. Dis.*, no. 14 (6):1001-2.

- Lindahl T. (1993) "Instability and decay of the primary structure of DNA." *Nature*, no. 362 (6422):709-15.
- Lindahl T. (1996) "The Croonian Lecture, 1996: Endogenous damage to DNA." *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, no. 351 (1347):1529-38.
- Liu G.-H., Yan H.-B., Otranto D. *et al.* (2014a) "*Dicrocoelium chinensis* and *Dicrocoelium dendriticum* (Trematoda: Digenea) are distinct lancet fluke species based on mitochondrial and nuclear ribosomal DNA sequences." *Molecular Phylogenetics and Evolution*, no. 79 (0):325 - 331.
- Liu G.H., Gasser R.B., Su A. *et al.* (2012a) "Clear genetic distinctiveness between human- and pig-derived *Trichuris* based on analyses of mitochondrial datasets." *PLoS Negl Trop Dis*, no. 6 (2):e1539.
- Liu G.H., Wu C.Y., Song H.Q. *et al.* (2012b) "Comparative analyses of the complete mitochondrial genomes of *Ascaris lumbricoides* and *Ascaris suum* from humans and pigs." *Gene*, no. 492 (1):110-6.
- Liu G.H., Zhou W., Nisbet A.J. *et al.* (2014b) "Characterization of *Trichuris trichiura* from humans and *T. suis* from pigs in China using internal transcribed spacers of nuclear ribosomal DNA." *J. Helminthol.*, no. 88 (1):64-8.
- Liu W.-Q., Liu J., Zhang J.-H. *et al.* (2007) "Comparison of ancient and modern *Clonorchis sinensis* based on ITS1 and ITS2 sequences." *Acta tropica*, no. 101 (2):91-94.
- Loreille O. et Bouchet F. (2003) "Evolution of ascariasis in humans and pigs: a multi-disciplinary approach." *Mem. Inst. Oswaldo Cruz*, no. 98:39-46.
- Loreille O., Roumat E., Verneau O. *et al.* (2001) "Ancient DNA from *Ascaris*: extraction amplification and sequences from eggs collected in coprolites." *Int. J. Parasitol.*, no. 31 (10):1101-6.
- Ludwig A., Pruvost M., Reissmann M. *et al.* (2009) "Coat color variation at the beginning of horse domestication." *Science*, no. 324 (5926):485.
- Lysek H., Malinsky J. et Janisch R. (1985) "Ultrastructure of eggs of *Ascaris lumbricoides* Linnaeus, 1758. I. Egg-shells." *Folia Parasitol. (Praha)*, no. 32 (4):381-4.
- Madden W., Salo W.L., Streitz J. *et al.* (2001) "Hybridization screening of very short PCR products for paleoepidemiological studies of Chagas' disease." *Biotechniques*, no. 30:102-109.

- Malhotra K., Foltz L., Mahoney W.C. *et al.* (1998) "Interaction and effect of annealing temperature on primers used in differential display RT-PCR." *Nucleic Acids Res.*, no. 26 (3):854-6.
- Mayer M.P. (2013) "Hsp70 chaperone dynamics and molecular mechanism." *Trends Biochem. Sci.*, no. 38 (10):507-14.
- Mayr G.W., Windhorst S. et Hillemeier K. (2005) "Antiproliferative plant and synthetic polyphenolics are specific inhibitors of vertebrate inositol-1,4,5-trisphosphate 3-kinases and inositol polyphosphate multikinase." *J. Biol. Chem.*, no. 280 (14):13229-40.
- Mégnin J.-P., *Les parasites articulés chez l'homme et les animaux utiles (maladies qu'ils occasionnent)*. ed. Masson, Paris, 1895.
- Mercier J. (2013) "*À l'ombre des remparts : 1 600 ans d'évolution urbaine à Auxerre. La fouille de la place des Véens*", Bulletin du centre d'études médiévales d'Auxerre, Disponible à <http://cem.revues.org/13131>.
- Merriman B., Ion Torrent R., Team D. *et al.* (2012) "Progress in ion torrent semiconductor chip based sequencing." *Electrophoresis*, no. 33 (23):3397-417.
- Meyer M., Fu Q., Aximu-Petri A. *et al.* (2014) "A mitochondrial genome sequence of a hominin from Sima de los Huesos." *Nature*, no. 505 (7483):403-6.
- Miller R.L., Ikram S., Armelagos G.J. *et al.* (1994) "Diagnosis of *Plasmodium falciparum* infections in mummies using the rapid manual ParaSight-F test." *Trans. R. Soc. Trop. Med. Hyg.*, no. 88 (1):31-2.
- Moriceau T. Rapport temporaire sur le site de la Turquerie. communication e-mail, mai 2015.
- Tricentenaire du Service de santé des armées, *Le service de santé des armées de l'évolution du concept hospitalier en France*, Ministère de la Défense, 2008, vol. 35.
- Mullis K., Faloona F., Scharf S. *et al.* (1986) "Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction." *Cold Spring Harb. Symp. Quant. Biol.*, no. 51 Pt 1:263-73.
- Myskova E., Ditrich O., Sak B. *et al.* (2014) "Detection of ancient DNA of *Encephalitozoon intestinalis* (Microsporidia) in archaeological material." *J. Parasitol.*, no. 100 (3):356-9.
- Nakao M., Abmed D., Yamasaki H. *et al.* (2007) "Mitochondrial genomes of the human broad tapeworms *Diphyllobothrium latum* and *Diphyllobothrium nihonkaiense* (Cestoda: Diphyllobothriidae)." *Parasitol. Res.*, no. 101 (1):233-236.

- Nakao M., Yokoyama N., Sako Y. *et al.* (2002) "The complete mitochondrial DNA } sequence of the cestode *Echinococcus multilocularis* (Cyclophyllidea: Taeniidae)." *Mitochondrion*, no. 1 (6):497 - 509.
- Nerlich A.G., Schraut B., Dittrich S. *et al.* (2008) "*Plasmodium falciparum* in ancient Egypt." *Emerg. Infect. Dis.*, no. 14 (8):1317-9.
- Nezamabadi M., Aali A., Stöllner T. *et al.* (2013a) "Paleoparasitological analysis of samples from the Chehrabad salt mine (Northwestern Iran)." *Int. J. of Paleopathology*, no. 3:229-233.
- Nezamabadi M., Mashkour M., Aali A. *et al.* (2013b) "Identification of *Taenia* sp. in a natural human mummy (third century BC) from the Chehrabad salt mine in Iran." *J. Parasitol.*, no. 99 (3):570-2.
- Oh C.S., Seo M., Chai J.Y. *et al.* (2010a) "Amplification and sequencing of *Trichuris trichiura* ancient DNA extracted from archaeological sediments." *J. of Archaeol. Sci.*, no. 37 (6):1269-1273.
- Oh C.S., Seo M., Lim N.J. *et al.* (2010b) "Paleoparasitological report on *Ascaris* aDNA from an ancient East Asian sample." *Mem. Inst. Oswaldo Cruz*, no. 105 (2):225-228.
- Okimoto R., Macfarlane J.L., Clary D.O. *et al.* (1992) "The mitochondrial genomes of two nematodes, *Caenorhabditis elegans* and *Ascaris suum*." *Genetics*, no. 130 (3):471-98.
- OMS. 2002. The partnership for parasite control. Papier lu à la 54 ème assemblée de l'OMS, Rome.
- OMS. (2013) "*Assessing the efficacy of anthelmintic drugs against schistosomiasis and soil-transmitted helminthiasis*", Disponible à <http://www.who.int/iris/handle/10665/79019>.
- OMS et UNICEF. (2013) "*Progress on sanitation and drinking-water - 2013 update*", Disponible à http://www.who.int/water_sanitation_health/publications/.
- Orlando L., Ginolhac A., Zhang G. *et al.* (2013) "Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse." *Nature*, no. 499 (7456):74-8.
- Ortner D.J. et Putschar W.G.J., *Identification of pathological conditions in human skeletal remains*. ed. Smithsonian Institution Press, Washington, 1981, Smithsonian Contributions to Anthropology, vol. 28.
- Paabo S. (1989) "Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification." *Proc. Natl. Acad. Sci. U.S.A.*, no. 86 (6):1939-43.
- Parducci L., Matetovici I., Fontana S.L. *et al.* (2013) "Molecular- and pollen-based vegetation analysis in lake sediments from central Scandinavia." *Mol. Ecol.*, no. 22 (13):3511-24.

- Park Y.C., Kim W. et Park J.K. (2011) "The complete mitochondrial genome of human parasitic roundworm, *Ascaris lumbricoides*." *Mitochondrial DNA*, no. 22 (4):91-3.
- Pastor-Valle J., Gonzalez L.M., Martin-Clemente J.P. *et al.* (2014) "Molecular diagnosis of diphyllbothriasis in Spain, most presumably acquired via imported fish, or sojourn abroad." *New Microbes New Infect.*, no. 2 (1):1-6.
- Patel P., Garson J.A., Tettmar K.I. *et al.* (2012) "Development of an ethidium monoazide-enhanced internally controlled universal 16S rDNA real-time polymerase chain reaction assay for detection of bacterial contamination in platelet concentrates." *Transfusion*, no. 52 (7):1423-32.
- Peng W., Yuan K., Hu M. *et al.* (2005) "Mutation scanning-coupled analysis of haplotypic variability in mitochondrial DNA regions reveals low gene flow between human and porcine *Ascaris* in endemic regions of China." *Electrophoresis*, no. 26 (22):4317-26.
- Peng W., Yuan K., Zhou X. *et al.* (2003) "Molecular epidemiological investigation of *Ascaris* genotypes in China based on single-strand conformation polymorphism analysis of ribosomal DNA." *Electrophoresis*, no. 24 (14):2308-15.
- Petersen L.W., Moldrup P., Jacobsen O.H. *et al.* (1996) "Relations between specific surface area and soil physical and chemical properties " *Soil Science*, no. 161:9-21.
- Poinar G., Duarte D. et Santos M.J. (2010) "*Halomonhystera parasitica* n. sp. (Nematoda: Monhysteridae), a parasite of *Talorchestia brito* (Crustacea: Talitridae) in Portugal." *Syst. Parasitol.*, no. 75 (1):53-8.
- Poinar G., Jr. et Boucot A.J. (2006) "Evidence of intestinal parasites of dinosaurs." *Parasitology*, no. 133 (Pt 2):245-9.
- Polz M.F. et Cavanaugh C.M. (1998) "Bias in template-to-product ratios in multitemplate PCR." *Appl. Environ. Microbiol.*, no. 64 (10):3724-30.
- Pruvost M., Schwarz R., Correia V.B. *et al.* (2007) "Freshly excavated fossil bones are best for amplification of ancient DNA." *Proc. Natl. Acad. Sci. U.S.A.*, no. 104 (3):739-44.
- Qu W., Shen Z. et Zhang C. PriDimerCheck. Logiciel disponible à <http://biocompute.bmi.ac.cn/MPprimer/>.
- Rabino Massa E., Cerutti N. et Savoia D. (2000) "Malaria in ancient Egypt: paleoimmunological investigations in predynastic mummified remains." *Chungara*, no. 32:7-9.
- Ramaroli V., Hamilton J., Ditchfield P. *et al.* (2010) "The Chehr Abad "Salt men" and the isotopic ecology of humans in ancient Iran." *Am. J. Phys. Anthropol.*, no. 143 (3):343-54.

- Raoult D., Reed D.L., Dittmar K. *et al.* (2008) "Molecular identification of lice from pre-Columbian mummies." *J. Infect. Dis.*, no. 197 (4):535-43.
- Renault-Miskovsky J. "Contribution of carpology and palynology to the knowledge of past vegetal environments, diet, diseases and pharmacopoeia of prehistoric human populations." In *Des sources du savoir aux médicaments du futur*. ed. La Société française d'ethnopharmacologie, 2002.
- Romon T. (2010) "*Liaison Baillif-Rivière des Pères*", Inrap.
- Rosseel T., Pardon B., De Clercq K. *et al.* (2014) "False-positive results in metagenomic virus discovery: a strong case for follow-up diagnosis." *Transbound Emerg. Dis.*, no. 61 (4):293-9.
- Rueckert A. et Morgan H.W. (2007) "Removal of contaminating DNA from polymerase chain reaction using ethidium monoazide." *J. Microbiol. Methods*, no. 68 (3):596-600.
- Ruffer M.-A. (1913) "On pathological lesions found on coptic bodies (400-500 AD)." *Journal of Pathol. Bact.*, no. 18:149-162.
- Ruffer M.-A., *Studies in the palaeopathology of Egypt*. ed. The University of Chicago Press, Chicago, 1921.
- Santos R.V., Nunes Jda S., Camargo J.A. *et al.* (2013) "High occurrence of *Entamoeba histolytica* in the municipalities of Ariquemes and Monte Negro, State of Rondonia, Western Amazonia, Brazil." *Rev. Inst. Med. Trop. Sao Paulo*, no. 55 (3).
- Sarkar G. et Sommer S. (1990) "More light on PCR contamination." *Nature*, no. 347 (6291):340-1.
- Shin D.H., Oh C.S., Lee H.J. *et al.* (2013) "Ancient DNA analysis on *Clonorchis sinensis* eggs remained in samples from medieval Korean mummy." *J. of Archaeol. Sci.*, no. 40:211-216.
- Shu F.F., Lv R.Q., Zhang Y.F. *et al.* (2012) "Characterization of *Fasciola* samples by ITS of rDNA sequences revealed the existence of *Fasciola hepatica* and *Fasciola gigantica* in Yunnan Province, China." *J. Parasitol.*, no. 98 (4):889-90.
- Sianto L., Reinhard K.J., Chame M. *et al.* (2005) "The finding of *Echinostoma* (Trematoda: Digenea) and hookworm eggs in coprolites collected from a Brazilian mummified body dated 600-1,200 years before present." *J. Parasitol.*, no. 91 (4):972-5.
- Slater L., Ntarlagiannis D. et Wishart D. (2006) "On the relationship between induced polarization and surface area in metal-sand and clay-sand mixtures." *Geophysics*, no. 71:1-5.

- Soe M.J., Nejsum P., Fredensborg B.L. *et al.* (2015) "DNA typing of ancient parasite eggs from environmental samples identifies human and animal worm infections in Viking-age settlement." *J. Parasitol.*, no. 101 (1):57-63.
- Spangenberg J.E., Jacomet S. et Schibler J. (2006) "Chemical analyses of organic residues in archaeological pottery from Arbon Bleiche 3, Switzerland and evidence for dairying in the late Neolithic." *J. of Archaeol. Sci.*, no. 33:1-13.
- Spigelman M. et Lemma E. (1993) "The use of Polymerase Chain Reaction to detect *Mycobacterium tuberculosis* in ancient skeletons." *Int. J. Osteoarcheol.*, no. 3:137-143.
- Stiller M., Knapp M., Stenzel U. *et al.* (2009) "Direct multiplex sequencing (DMPS)--a novel method for targeted high-throughput sequencing of ancient and highly degraded DNA." *Genome Res.*, no. 19 (10):1843-8.
- Summers R.W., Elliott D.E., Urban J.F., Jr. *et al.* (2005) "*Trichuris suis* therapy for active ulcerative colitis: a randomized controlled trial." *Gastroenterology*, no. 128 (4):825-32.
- Tadayon S., Sharifiyazdi H., Moazeni M. *et al.* (2015) "Molecular Differentiation of *Fasciola* Species and Characterization of Genetic Diversity of *F. gigantica* Using NADH Dehydrogenase I (ND1) Gene in the Endemic Areas of Iran." *Iran J. Parasitol.*, no. 10 (1):9-18.
- Tardy C. 2015. Communication par email sur le site de Mas de Vignoles.
- Tarrus J., Chinchilla J. et Bosch À. (1994) "*La Draga (Banyoles) : un site lacustre du Néolithique ancien cardial en Catalogne*", Bulletin de la Société préhistorique française, vol. 91, Disponible à:
http://www.persee.fr/web/revues/home/prescript/article/bspf_0249-7638_1994_num_91_6_9800.
- Tateno Y., Imanishi T., Miyazaki S. *et al.* (2002) "DNA Data Bank of Japan (DDBJ) for genome scale research in life science." *Nucleic Acids Res.*, no. 30 (1):27-30.
- Thompson R.C. et McManus D.P. (2002) "Towards a taxonomic revision of the genus *Echinococcus*." *Trends Parasitol.*, no. 18 (10):452-457.
- Torgerson P.R., Pilkington J., Gulland F.M.D. *et al.* (1995) "Further evidence for the long distance dispersal of taeniid eggs." *Int. J. Parasitol.*, no. 25 (2):265-267.
- Tsai I.J., Zarowiecki M., Holroyd N. *et al.* (2013) "The genomes of four tapeworm species reveal adaptations to parasitism." *Nature*, no. 496 (7443):57-63.

- Tsai Y.L. et Olson B.H. (1992) "Rapid method for separation of bacterial DNA from humic substances in sediments for polymerase chain reaction." *Appl. Environ. Microbiol.*, no. 58 (7):2292-5.
- Urbina J.A., Lazard K., Marchan E. *et al.* (1993) "Mevinolin (lovastatin) potentiates the antiproliferative effects of ketoconazole and terbinafine against *Trypanosoma (Schizotrypanum) cruzi*: in vitro and in vivo studies." *Antimicrob. Agents Chemother.*, no. 37 (3):580-91.
- Van Beneden P.-J., *Commensaux et Parasites*. ed. Bibliothèque Internationale, Paris, 1875.
- Van Kruiningen H.J. et West A.B. (2005) "Potential danger in the medical use of *Trichuris suis* for the treatment of inflammatory bowel disease." *Inflamm. Bowel Dis.*, no. 11 (5):515.
- VanBuskirk K.M., O'Neill M.T., De La Vega P. *et al.* (2009) "Preerythrocytic, live-attenuated *Plasmodium falciparum* vaccine candidates by design." *Proc. Natl. Acad. Sci. U.S.A.*, no. 106 (31):13004-9.
- Vasan S., Zhang X., Zhang X. *et al.* (1996) "An agent cleaving glucose-derived protein crosslinks in vitro and in vivo." *Nature*, no. 382 (6588):275-8.
- Vidal. *Vidal: la base de donnée en ligne des prescripteurs libéraux* 2015. Disponible à <http://www.vidal.fr/>.
- Wagner A., Blackstone N., Cartwright P. *et al.* (1994) "Surveys of Gene Families Using Polymerase Chain Reaction: PCR Selection and PCR Drift." *Systematic biology*, no. 42 (2):250-261.
- Wang X.Y., Zhao G.H., Liu G.H. *et al.* (2013) "Characterization of *Dicrocoelium chinensis* from domestic yaks in Gansu and Sichuan provinces, China, using genetic markers in two mitochondrial genes." *Mitochondrial DNA*, no. 24 (3):263-6.
- Wannasan A., Khositharattanakool P., Chaiwong P. *et al.* (2014) "Identification of *Fasciola* species based on mitochondrial and nuclear DNA reveals the co-existence of intermediate *Fasciola* and *Fasciola gigantica* in Thailand." *Exp. Parasitol.*, no. 146:64-70.
- Wei T., Lu G. et Clover G. (2008) "Novel approaches to mitigate primer interaction and eliminate inhibitors in multiplex PCR, demonstrated using an assay for detection of three strawberry viruses." *J. Virol. Methods*, no. 151 (1):132-9.
- Wéry M., *Protozoologie médicale*. ed. De Boeck Université, Bruxelles, 1995.
- Wicht B., de Marval F., Gottstein B. *et al.* (2008) "Imported diphyllbothriasis in Switzerland: molecular evidence of *Diphyllbothrium dendriticum* (Nitsch, 1824)." *Parasitol. Res.*, no. 102 (2):201-4.

- Wicht B., Yanagida T., Scholz T. *et al.* (2010) "Multiplex PCR for differential identification of broad tapeworms (Cestoda: Diphylobothrium) infecting humans." *J. Clin. Microbiol.*, no. 48 (9):3111-6.
- Willerslev E., Cappellini E., Boomsma W. *et al.* (2007) "Ancient biomolecules from deep ice cores reveal a forested southern Greenland." *Science*, no. 317 (5834):111-4.
- Willerslev E., Davison J., Moora M. *et al.* (2014) "Fifty thousand years of Arctic vegetation and megafaunal diet." *Nature*, no. 506 (7486):47-51.
- Wood J.R., Wilmshurst J.M., Rawlence N.J. *et al.* (2013) "A Megafauna's Microfauna: Gastrointestinal Parasites of New Zealand's Extinct Moa (Aves: Dinornithiformes)." *PLoS ONE*, no. 8 (2):e57315.
- Wood J.R., Wilmshurst J.M., Wagstaff S.J. *et al.* (2012) "High-resolution coproecology: using coprolites to reconstruct the habits and habitats of New Zealand's extinct upland moa (*Megalapteryx didinus*)." *PLoS ONE*, no. 7 (6):e40025.
- Zammit J. (2005) "Les conséquences écologiques de la néolithisation dans l'histoire humaine." *Bulletin de la Société préhistorique française*, no. 102 (2):371-379.
- Zheng H., Jia H., Shankar A. *et al.* (2011) "Detection of murine leukemia virus or mouse DNA in commercial RT-PCR reagents and human DNAs." *PLoS ONE*, no. 6 (12):e29050.
- Zheng H., Zhang W., Zhang L. *et al.* (2013) "The genome of the hydatid tapeworm *Echinococcus granulosus*." *Nat. Genet.*, no. 45 (10):1168-1175.
- Zhou C., Li M., Yuan K. *et al.* (2012) "Pig *Ascaris*: an important source of human ascariasis in China." *Infect. Genet. Evol.*, no. 12 (6):1172-7.
- Zink A.R., Spigelman M., Schraut B. *et al.* (2006) "Leishmaniasis in ancient Egypt and Upper nubia." *Emerg. Infect. Dis.*, no. 12 (10):1616-7.

Résumé

APPORTS DE LA PALÉOGÉNÉTIQUE À L'ÉTUDE DES HELMINTHES GASTRO-INTESTINAUX ANCIENS

La paléoparasitologie est l'étude des restes de parasites préservés dans des échantillons archéologiques et permet de mieux comprendre l'état de santé des populations anciennes et d'obtenir des informations d'ordre anthropologique ou ethnologique, sur les régimes alimentaires ou les conditions d'hygiène au quotidien. Les restes de parasites peuvent être retrouvés sous forme de macro-restes (vers ou larves), d'antigènes, d'ADN ou d'œufs. Ces derniers peuvent être particulièrement bien préservés au cours du temps car ils sont composés en partie de chitine, les rendant résistants aux processus de dégradation. L'observation microscopique de leurs caractéristiques morphologiques et micrométriques permet d'identifier les taxons au niveau du genre ou de la famille. Dans le cadre de cette thèse, plusieurs helminthes gastro-intestinaux, dont les œufs sont fréquemment retrouvés dans des échantillons archéologiques, ont été ciblés par une approche génétique. Il s'agit des vers plats *Taenia saginata*, *T. solium*, *T. asiatica*, *Echinococcus granulosus*, *E. multilocularis*, *Diphyllobothrium latum*, *D. dendriticum* et *D. nihonkaiense*, des nématodes *Trichuris trichiura*, *Enterobius vermicularis* et *Ascaris* sp. et des douves *Fasciola hepatica*, *F. gigantica*, *Dicrocoelium dendriticum* et *D. chinensis*.

La méthode « *aMPlex Torrent* » permet de détecter, dans un grand nombre d'échantillons archéologiques, une faible quantité d'ADN de parasites. Cette approche combine la spécificité et la sensibilité de la PCR au haut-débit du séquençage de nouvelle génération. Plusieurs vestiges, provenant de périodes et de régions géographiques diverses, ont été analysés. Des résultats génétiques ont été obtenus pour des échantillons aussi anciens que 7200 BP. Nous avons par ailleurs obtenus les premières séquences anciennes de *Taenia* sp., *Diphyllobothrium* sp., *Echinococcus* sp., et les premières séquences européennes d'*Enterobius vermicularis*. Au vu de ces résultats, notre approche apparaît comme étant complémentaire à la microscopie.

Mots clés : paléogénétique, paléoparasitologie, helminthes gastro-intestinaux, génotypage, séquençage de nouvelle génération, PCR multiplex, ADN, œufs

Abstract

PALAEOGENETICS TO STUDY ANCIENT GASTROINTESTINAL HELMINTHS

Palaeoparasitology, the study of parasite remains from archaeological samples, is a discipline that can highlight questions about the health status of the ancient populations. It can give important anthropological or ethnological information such as the diet and the hygiene conditions of past societies. The remains can be preserved as macroremains (worms or larvae), antigens, DNA or eggs. Because they are partially made of chitin, eggs of gastrointestinal helminths resist well over time to the taphonomic degradation process. It is possible to distinguish between different families or genera of parasites by looking at the morphological features of eggs. However, since several taxa share common features, the determination is rarely possible at the species level. For this thesis, several parasite species for which eggs are commonly observed in archeological samples have been studied by a genetic approach. We studied the tapeworms *Taenia saginata*, *T. solium*, *T. asiatica*, *Echinococcus granulosus*, *E. multilocularis*, *Diphyllobothrium latum*, *D. dendriticum*, and *D. nihonkaiense*; the nematodes *Trichuris trichiura*, *Enterobius vermicularis*, and *Ascaris* sp.; and the flukes *Fasciola hepatica*, *F. gigantica*, *Dicrocoelium dendriticum*, and *D. chinensis*.

The “aMPlex Torrent” approach has been set up to detect minute amounts of DNA from parasites in multiple archaeological samples. This approach combines the specificity and sensitivity of PCR to the throughput of Next-Generation sequencing. Several samples have been analyzed by this approach. We obtained genetic results for samples as old as 7200 BP and from various geographical and archeological contexts. We obtained the first ancient DNA sequences for *Taenia* sp., *Diphyllobothrium* sp., *Echinococcus* sp. and the first European sequences for *Enterobius vermicularis*. Genetic analyses and microscopic observations appear to be complementary. Indeed, at least one taxon per sample was detected by one of the two approaches.

Keywords: palaeogenetics, palaeoparasitology, gastrointestinal helminths, genotyping, Next-Generation Sequencing, multiplex PCR, DNA, eggs

