Aix-Marseille Université
Faculté de Médecine
École Doctorale des Sciences De la Vie et de la Santé (EDSVS)

# THÈSE DE DOCTORAT/PhD

Présentée par

## Sandra APPELT

Née le 05 Décembre 1985 à Neuruppin, Allemagne

Pour obtenir le grade de docteur de L'UNIVERSITÉ AIX-MARSEILLE
SPÉCIALITÉ : **Pathologie Humaine, Maladie Infectieuse**

# Paléomicrobiologie des coprolithes

Présentée et publiquement soutenue devant
La Faculté de Médecine de Marseille
Le 09 Décembre 2013

COMPOSITION DU JURY

| | |
|---|---|
| M le Professeur Jean-Louis Mege | Président de Jury |
| Mme le Professeur Christine Keyser | Rapporteur |
| M le Professeur Antoine Andremont | Rapporteur |
| M le Professeur Michel Drancourt | Directeur de Thèse |

Assistance Publique
Hôpitaux de Marseille

## PREAMBULE

Le format de présentation de cette Thèse correspond à une recommandation de la spécialité Maladies Infectieuses et Microbiologie, à l'intérieur du Master des Sciences de la Vie et de la Santé qui dépend de l'Ecole Doctorale des Sciences de la Vie de Marseille. Le candidat est amené à respecter des règles qui lui sont imposées et qui comportent un format de thèse utilisé dans le Nord de l'Europe et qui permet un meilleur rangement que les thèses traditionnelles. Par ailleurs, la partie introduction et bibliographie est remplacée par une revue envoyée dans un journal afin de permettre une évaluation extérieure de la qualité de la revue et de permettre à l'étudiant de commencer le plus tôt possible une bibliographie exhaustive sur le domaine de cette thèse. Par ailleurs, la thèse présentée sur article publié, accepté ou soumis associé d'un bref commentaire donnant le sens général du travail. Cette forme de présentation a paru plus en adéquation avec les exigences de la compétition internationale et permet de se concentrer sur des travaux qui bénéficieront d'une diffusion internationale.

i

# SOMMAIRE

## ACRONYMES

| | |
|---|---|
| **aDNA** | ancient Deoxyribonucleic acid |
| **AD** | Anno Domini |
| **Amplicon** | product of artificial amplification |
| **BLAST** | Basic Local Alignment Tool |
| **bp** | base pair |
| **BP** | before present |
| **Contig** | Contiguous sequence |
| **dsDNA** | double stranded DNA |
| **NCBI** | National Center for Biotechnology Information |
| **ORF** | Open Reading Frame |
| **PCR** | Polymerase Chain Reaction |
| **Read** | A sequenced DNA fragment |
| **RNA** | Ribonucleic acid |
| **rRNA** | Ribosomal Ribonucleic Acid |
| **Shotgun sequencing** | Sequencing of random fragments |
| **ssDNA** | single stranded DNA |
| **Virome** | Viral metagenome |
| **WGS** | Whole Genome Shotgun sequencing |

Les coprolithes sont le produit de la dessiccation et de la minéralisation de matières fécales. En faisant le parallèle avec les selles modernes, les coprolithes peuvent être des échantillons appropriés à l'étude des habitudes alimentaires, de la flore intestinale et des maladies, des animaux et des hommes ayant vécu il y a plusieurs siècles. Leur analyse a d'abord reposé sur des observations macro et microscopiques, puis sur les détections d'ADN ou de molécules non-nucléotidiques en respectant les standards paléomicrobiologiques et en appliquant des méthodes couramment utilisées pour le diagnostic routine. Dans le travail de thèse ici présenté, un coprolithe datant des 14-15$^{ième}$ siècles et provenant de Namur en Belgique, a été étudié. Ce coprolithe a été retrouvé dans un tonneau, utilisé à l'époque comme latrine.

Dans un premier temps, nous avons caractérisé l'ensemble de la communauté microbienne associée au coprolithe et identifié certains pathogènes. Pour cela, des observations microscopiques, la culture et la métagénomique ont été combinées. La confirmation des résultats importants a reposé sur des PCR *ad hoc*. Les résultats ont montré qu'une partie du microbiote est similaire à l'environnement. De plus, des bactéries typiques de la flore intestinale, des parasites intestinaux et des pathogènes systémiques ont été trouvés. Pour en apprendre davantage, un second projet a visé à la purification de particules virales à partir du coprolithe de Namur et leur analyse par microscopie électronique et métagénomique virale. Des particules virales sphériques, ainsi que des bactériophages, ont ainsi été observés. Les virus associés au coprolithe correspondent à des virus d'eucaryotes, de procaryotes et d'archaea. La communauté virale était dominée par des bactériophages détectés dans le sol et les selles modernes. Bien que la composition phylogénétique diffère de celle des selles modernes, les fonctions métaboliques de ces communautés virales semblent être plus conservées. Parmi les fonctions métaboliques détectées, une correspond à des résistances aux antibiotiques.

Dans un troisième projet, des cultures et des identifications moléculaires ont été réalisées sur des kystes d'amibes précédemment observés dans le coprolithe de Namur. Les amibes ainsi isolées appartiennent au genre *Acanthamoeba* et pourraient avoir été conservées sous forme de kystes pendant des siècles dans le coprolithe de Namur. De plus, les co-cultures d'amibes permis l'isolement et l'identification d'une nouvelle bactérie bi-flagellée résistante aux amibes, proche des *Rickettsiales*.

S'agissant de la première étude globale portant sur un coprolithe retrouvé en Europe, notre travail contribue à la connaissance de la flore du Moyen-âge sur ce continent. De plus, notre approche démontre que l'application, sur des coprolithes, de méthodes utilisées habituellement pour le diagnostic de routine, permet de détecter des pathogènes dans des échantillons anciens.

**Mots-clés : paléomicrobiologie, coprolithes, flore intestinale, infections anciennes, virus, amibes**

# ABSTRACT

Coprolites result from the desiccation or mineralization of feces material. Accordingly, by drawing parallels to modern stools, coprolites can be suitable specimen to study diet habits, gut microbiota and diseases of animals and humans that have lived centuries ago. Their investigations relied first mainly on macroscopic and microscopic observations, then on nucleotidic and non-nucleotidic biomolecules detection through the application of paleomicrobiological standards and routine diagnostic methods. During this thesis work, a 14-15th century coprolite specimen from Namur, Belgium was analyzed. The Namur coprolite was found in a barrel – used at this period as pits or latrines.

At the initiation of this thesis work, it was aimed to characterize the entire microbial communities associated to the coprolite and to identify ancient pathogens. To accomplish this task microscopic observation, culture and metagenomics were combined. Confirmation of important results was targeted *in silico* or by *ad hoc* PCR. Results indicated that parts of the microbiota are similar to those coming from environment. Furthermore, typical gut microbiota inhabitants, intestinal parasites and systemic pathogens – still relevant nowadays – were found. To go further, in a second work, viral particles were purified from the Namur coprolite and analyzed by electron microscopic and viral metagenomic. Viral particles associated to spherical virions and bacteriophages were observable. Moreover, viruses infecting eukaryotes, bacteria and archaea were associated to the specimen. The viral community was dominated by bacteriophages commonly found in soil and in modern stools. Although the phylogenetic compositions differed from modern stool, the metabolic functions of the viral communities seem to have remained more conserved. Furthermore, antibiotic resistance was one of the metabolic functions detected.

In a third project, culture and molecular identification were performed on amoebal cysts previously observed within the Namur coprolite. The amoebas isolated belong to the genus *Acanthamoeba* and might have been conserved in form of cysts inside the Namur coprolite for centuries. Moreover, amoeba-co culturing leaded to the isolation and identification of a new bi-flagellar amoeba-resistant bacterium closely related to *Rickettsiales*.

Accordingly, represented is a broad analysis of a European coprolite that adds knowledge to ancient flora preserved within coprolites during medieval times. Furthermore it was showed that ancient pathogens can be detected – inside of coprolites – throughout the application used in routine diagnosis.

**Keywords: paleomicrobology, coprolites, gut microbiota, ancient infections, viruses, amoeba**

# INTRODUCTION ET OBJECTIFS

L'investigation des coprolithes donne des informations sur les conditions environnementales des temps passés et les habitudes alimentaires de nos ancêtres à travers l'analyse de leur flore intestinale et des maladies associées (Cleeland et al. 2013; Goncalves et al. 2003; Tito et al. 2012; Tito et al. 2008; Wood et al. 2013). Il est possible de déterminer par microscopie des résidus (pollen, os, œufs de helminthes, kystes de protozoaires) dans des coprolithes : la composition de la paléo-faune du Pléistocène Supérieur a ainsi pu être partiellement déterminée après la détection et l'identification de pollen (Yll 2006). Egalement, la détection de différents pollens en concentration élevée retrouvés dans des coprolithes humains témoignent de l'utilisation de plantes médicinales (Reinhard 2007; Reinhard & Bryant 2008). L'analyse de résidus alimentaires, des ossements animaux et humains, permet d'étayer la transmission de maladies par ingestion de nourritures contaminées (Chaves & Reinhard 2006; Chaves 2003; Dean 2006). L'étude des coprolithes fait appel à la détection d'ADN ancien et de molécules non-nucléotidiques, fournissant un support moeéculaire aux observations microscopiques (Leles et al. 2008; Loreille et al. 2001; Reinhard 2007; Reinhard & Bryant 2008; Ubaldi et al. 1998). Ces recherches peuvent être riches d'enseignement sur l'émergence, voire la ré-émergence, de maladies parasitaires et d'infections systémiques. Les avancées techniques dans le domaine moléculaire ont également grandement bénéficié à l'étude du microbiote intestinal humain : l'analyse par PCR est maintenant complétée par le séquençage à haut-débit (Cano et al. 2000; Luciani et al. 2006; Rollo et al. 2006; Santiago-Rodriguez et al. 2013; Tito et al. 2012; Tito et al. 2008; Ubaldi et al. 1998).

Nous avons revu l'ensemble des possibilités offertes par l'étude des coprolithes humains, ainsi que les techniques utilisées et les apports concernant la flore intestinale humaine à des époques passées, les parasites intestinaux ou encore les pathogènes systémiques, dans un travail de revue bibliographique présenté dans le **Chapitre I** de notre thèse. Ensuite, nous avons focalisé notre travail de thèse sur l'exploration paléomicrobiologique du microbiote associé à un coprolithe daté du Moyen Age. Le coprolithe étudié fut excavé en 1996 sur un site archéologique à Namur, en Belgique. Ce coprolithe a été découvert à l'intérieur d'un tonneau fermé, enterré à une profondeur de 3 mètres 80, tel que ceux utilisés comme latrines, daté aux alentours de 1400 – 1500 apr.

J.C. (après Jésus-Christ). Dans un premier temps (**Chapitre II**), le coprolithe de Namur a été analysé *via* une approche polyphasique, dans le but de caractériser le microbiote associé et de chercher des pathogènes. Suite à ce travail, la communauté virale associée au coprolithe a également été étudiée ce qui constitue la première analyse de ce type (**Chapitre III**). En supplément à ces analyses principalement moléculaires, le **Chapitre IV** présente une approche culturomique, ayant permit de recultiver des amibes enkystées précédemment observées et identifiées dans le **Chapitre II**. Le manuscrit se termine par la partie **Conclusions et Perspectives**, dans laquelle les résultats sont discutés.

# Chapitre I

REVUE DE LITTERATURE

## Des coprolithes humains comme source pour la paléomicrobiologie

Sandra Appelt[1], Matthieu Le Bailly[2] and Michel Drancourt[1*]

[1] Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, Inserm 1095, 13005 Marseille, France.[2] Franche-Comté University, CNRS UMR 6249 Chrono-Environment, 25 030 Besançon, France.

**Avant-propos**

Notre revue de littérature consacrée à la paléomicrobiologie des coprolithes permet à un lecteur non initié en paléomicrobiologie d'aborder facilement les différents concepts et méthodes décrits dans cette thèse. Elle permet également de situer notre travail expérimental dans un contexte global pour comprendre la contribution de nos travaux présentés dans les **Chapitres II**, **III** et **IV**. Au cours de fouilles archéologiques, des matières fécales fossilisées – appelées coprolithes – ont été découverts. Ces coprolithes sont de diverses origines, provenant d'espèces animales disparues ou existantes et des hommes. Vieux de plusieurs siècles voire de plusieurs millénaires, leurs analyses présentent un intérêt majeur en paléomicrobiologie, notamment pour la connaissance des conditions environnementales de l'époque, et pour l'étude de la flore intestinale humaine et ses pathogènes. Chaque étude est amorcée par les mêmes interrogations : *« Où trouver des coprolithes ? Comment les identifier et les stocker tout en les conservant d'éventuelles dégradations ? Quelles analyses en laboratoire peuvent être réalisées ? Qu'est-il possible d'en apprendre ? »*. La revue ici présentée est scindée en deux parties principales. La première est focalisée sur les méthodes utilisées pour manipuler les coprolithes sur le site archéologique puis dans le laboratoire. Les techniques d'analyse des coprolithes, comprenant la microscopie ainsi que la détection d'ADN et de molécules non-nucléotidiques, y sont également revues. En effet, des résidus microscopiques peuvent rester intacts pendant au moins 10 000 ans (Goncalves et al. 2003), tandis que des antigènes de protozoaires ont été détectés dans des coprolithes humains vieux de plus de 5 000 ans (Goncalves et al. 2002; Le Bailly et al. 2008). De plus, l'isolement d'ADN ancien à partir de coprolithes rend possible l'application de techniques de biologie moléculaire. La deuxième partie de la revue résume l'ensemble des informations obtenues par l'analyse de coprolithes humains au cours des 21 dernières années. Cette section est d'abord consacrée aux connaissances actuelles concernant la flore intestinale humaine ancienne, puis traite des pathogènes humains associés aux coprolithes, enfin des pathogènes systémiques qui peuvent être diagnostiqués dans des coprolithes.

# MINI REVIEW

## Human coprolites as a source for paleomicrobiology
**Running title:** Paleomicrobiology of coprolites

Sandra Appelt[1], Matthieu Le Bailly[2] and Michel Drancourt[1*]

[1] Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, Inserm 1095, 13005 Marseille, France. [2] Franche-Comté University, CNRS UMR 6249 Chrono-Environment, 25 030 Besançon, France.

[*]Corresponding author. Michel.Drancourt@medicine.univ-mrs.fr

**Abstract**

The paleomicrobiology of coprolites – fossilized fecal materials – yielded already data about different organisms including micro-eukaryotes, bacteria and archaea, expanding comprehension of ancient human dietary habits, gut microbiota, intestinal and systemic infections. This mini-review describes briefly previous works, including a summary of the main techniques used in handling coprolites and the findings obtained about ancient gut microbiota, past intestinal as well as systemic infections are outlined.

**INTRODUCTION**

Paleomicrobiology – the quest for ancient microbes – is based on the analyses of human bones, teeth and mummified soft tissues (Taylor et al. 1999; Zink et al. 2001; Tran et al. 2011). In addition, ancient fecal remains – preserved by mineralization or desiccation – in form of organic or permineralized coprolites, intestinal contents or latrines yielded data on environmental and gut microbiota of humans and animals that lived centuries to millennia ago (Reinhard and Bryant 1992; Araujo et al. 1998; Bouchet et al. 2003a; Dittmar 2009; Tito et al. 2012; Dentzien-Dias et al. 2013; Santiago-Rodriguez et al. 2013). The first ever described coprolite was discovered in the abdominal cavity of an 230-million-year old ichthyosaur and recently, a 270-million-year old shark coprolite was investigated (Reinhard and Bryant, 1992; Dentzien-Dias et al. 2013).

The oldest studied human coprolite was dated to approximately 12,450 before present (BP) (Jenkins et al. 2012) and human coprolites have been further recovered from 87 archeological sites (**Figure 1**). These coprolites were found in rock shelters, archeological layers, latrines, pits, and inside of mummified bodies (Bouchet et al. 2003b; Goncalves et al. 2003; Rollo et al. 2006; Reinhard and Bryant 2008; Tito et al. 2012). Human coprolites are usually fragmentized and flattered, preserved as amorphous masses of different sizes and textures (Reinhard and Bryant 2008). Consequently, they are often collected during the initial screening processes used to separate dirt from artifacts (Reinhard and Bryant 2008). Here, we review the techniques used to handle coprolites, and outline knowledge about ancient gut microbiota and intestinal as well as systemic infections which were obtained by the analyses of human coprolites.

**LOOKING FOR MICROORGANISMS IN COPROLITES: METHODOLOGIES**

During the excavation the use of gloves is recommended to minimize contaminations (Bouchet et al. 2003a), moreover when ancient DNA analyses are expected. Further, observations of the archeological site can be of major interest, providing information about the washing-off of microscopic remains and DNA during the fossilization process. For instance, the accumulation of coprolites or struvites (a soft mineral that precipitate in alkaline urine and form stones) can indicate such incidence that happen when coprolites did not have permanent protections from environmental conditions throughout centuries (Carrio et al. 2001; Bouchet et al. 2002; Bouchet et al. 2003a; Chaves and Reinhard 2006; Gilbert et al. 2008; Goldberg et al. 2009; Jenkins et al. 2012). After excavation the specimen can be stored into sterile forensic bags under adapted conditions to avoid contaminations, fungal growth and further degradations. Indeed, changes in the environmental conditions (light, temperature, oxygen concentration) from the place of specimen discovery can induce microbial growth or accelerate ancient DNA (aDNA) degradation (Bouchet et al. 2003a; Tito et al. 2008). The coprolite should be photographed, measured and weighted before aliquoting in the laboratory. Morphometric studies can inform about the host of the coprolite dependant on form, size, color, texture and inclusions (Reinhard et al. 2003). Nevertheless, for coprolites, host identification is more often based on the archeological context, host-specific food residues, host- specific parasites and bacteria found inside the coprolite and the analysis of host aDNA (Loreille et al. 2001; Poinar et al. 2001; Rocha and Serra-Freire 2009; Jenkins et al. 2012; Fugassa et al. 2013).

Before analyzing coprolites, due to subaerial exposition before recovery by sediment layers, it could be important to aseptically brush or remove the external layer (Iniguez et al. 2003a; Tito et al. 2012; Santiago-Rodriguez et al. 2013). A rehydration phase is needed and can be extended up to 10 days (Dufour & Le Bailly, 2013; Le Bailly et al. 2003; Le Bailly et al. 2005). Rehydration is done in aqueous 0.5% tri-sodium phosphate solution, sometimes completed with 5% glycerinated water or several drops of 10% formalin to avoid fungal or bacterial growth (Poinar 1998; Goncalves et al. 2003; Goncalves et al. 2004; Le Bailly et al. 2005). When the rehydration is performed in water only, sodium hydroxide, or Ethylene-diamine-tetra-acetic acid (EDTA), biogenic components can be destroyed and the coprolite can disintegrate (Goncalves et al. 2003). Accordingly, macromolecules including DNA, proteins and lipids can be stained using acridine orange, Fast Green FCF and Nile Red staining, respectively before analyses (Santiago-Rodriguez et al. 2013).

**MICROSCOPIC ANALYSES**

Microscopic observations can reveal helminths eggs and protozoa cysts (Goncalves et al. 2003; Ortega and Bonavia 2003). These microbial residues can stay uncollapsed for at least 11,000 years inside human coprolites (see **Figure 1**). Autofluorescence was described for the intestinal parasite *Cyclospora cayetanensis* that might be maintained for at least 3,000 years (Allison et al. 1999). The identification of intestinal parasites and protozoa is based on morphometric characterization of their eggs and cysts using light microscopy in first instance. For instance, length and wideness differentiate the human-infecting *Trichuris trichiura* eggs and the pig-infecting

17

*Trichuris suis* eggs (Fernandes et al. 2005; Rocha et al. 2006; Rocha and Serra-Freire 2009). Furthermore, *Metagonimus* spp. eggs are differentiated from those of *Clonorchisis* spp. only by scanning electron microscopic observation of differences on their surface structures (Shin et al. 2009). The observation and identification of protozoa cysts is rare, due to reduced resistance to natural decay and the rehydration protocol (Goncalves et al. 2002; Le Bailly et al. 2008).

**NON-NUCLEOTIDIC BIOMOLECULE DETECTION**

Cytochemical staining using FastGreen detected proteins in human coprolites dated to 100 BP – 600 Anno Domini (AD) (Santiago-Rodriguez et al. 2013). It was further possible to detect intestinal protozoa antigens by using either commercially available immunofluorescence assay (IFA) or enzyme-linked immunosorbent assay (ELISA) (Allison et al. 1999; Goncalves et al. 2002; Ortega and Bonavia 2003; Goncalves et al. 2004; Le Bailly & Bouchet 2006; Le Bailly et al. 2008; Mitchell et al. 2008). A monoclonal antibody IFA test detected *Cryptosporidium* and *Giardia* species in 3,000 BP to 1,000 AD coprolites (Allison et al. 1999; Ortega and Bonavia 2003; Le Bailly et al. 2008). The ELISA detected an *Entamoeba histolytica* specific adhesin as well as a *Giardia intestinalis* specific cyst wall protein 1 (CWP1) and a glycoprotein (GSA65) in samples dated from 5,300 BP to 1,900 AD (Goncalves et al. 2002; Goncalves et al. 2004; Le Bailly et al. 2008; Mitchell et al. 2008).

**MOLECULAR MICROBIOLOGY**

As for all ancient materials, coprolites needed to be manipulated according to standard protocols for aDNA. That implies to never use positive controls but instead to incorporate several negative controls. Positive results require confirmation by additional tests, either detection of another unrelated aDNA sequence;  detection of a specific protein;  or the independent replication of results in an appropriated second laboratory (Cooper and Poinar 2000; Hofreiter et al. 2001; Pääbo et al. 2004; Drancourt and Raoult 2005; Willerslev and Cooper 2005).

### aDNA extraction

Recently, acridine orange staining detected aDNA in human coprolites dated to 100 BP – 600 AD, thus indicating that molecular biological techniques can be applied to the specimens tested (Santiago-Rodriguez et al. 2013). Nevertheless, poorly characterized inhibitors may hamper the PCR-based detection of aDNA. Previous investigations showed that coprolites contained polysaccharides, lignin and Maillard reactions products (cross-links between reducing sugars and amino groups), all inhibitors of PCR (Poinar 1998). Those inhibitors have to be removed before applying molecular tests. N-phenacyl thiazolium bromide in sodium phosphate buffer solution (PTB) was described to be effective against Maillard reaction products (Poinar 1998; 2002). Incorporation of internal controls such as artificial non-targeted DNA templates allows monitoring for PCR inhibitors (Volossiouk et al. 1995; Rosenstraus et al. 1998; Honore-Bouakline et al. 2003). DNA extractions were performed either on eggs of intestinal parasites and bones pulled out from coprolites, or directly on the coprolites

(Poinar 1998; Loreille et al. 2001; Tito et al. 2011). Until now, six different protocols have been reported for the successful extraction of aDNA from coprolites. An initial protocol described in 1998 (Poinar 1998) was then slightly modified depending on the material (Poinar et al. 2001; Iniguez et al. 2003a; Iniguez et al. 2003b; Iniguez et al. 2006; Luciani et al. 2006). The MoBio PowerSoil and UltraClean Fecal DNA Kit, and a salting out extraction protocol were also used for coprolites dated from 7,315 BP to 1,450 AD. The latter is especially suitable for coprolites containing bone fragments (Tito et al. 2008; Tito et al. 2012; Cleeland et al. 2013; Santiago-Rodriguez et al. 2013). In 2012, another protocol normally used for ancient sediments, was used for human coprolites dated to 12,265 BP (Willerslev et al. 2003; Jenkins et al. 2012). Finally, an extraction protocol specifically for coprolites was described in 2012 (Kuch and Poinar 2012).

### PCR amplification and Sanger-sequencing

Nested-PCRs have been used to identify microorganisms in human coprolites, since nested-PCRs increased sensibility over standard PCRs (Kiatpathomchai et al. 2001). However, as for modern diagnosis, nested-PCR is plaqued by amplicon carryover during the second round of amplification and false positive results (Neumaier et al. 1998; Raoult et al. 2000; Kiatpathomchai et al. 2001; Zeaiter et al. 2003). Thus, a nested-PCR should not be longer use in coprolite investigations. PCR typically yielded products ranked in size from 98 to 918 base pairs (bp). Identification of intestinal parasites, fungi, bacteria and archaea were obtained by sequencing PCR amplified cytochrome b gene, 18S rDNA, 16S rDNA and 5S rDNA or nuclear ribosomal internal

transcribed spacer (ITS) regions (**Supplementary Table S2**). After amplification, the species identification was performed by blast annotations after sequencing or by terminal restricts fragments polymorphisms (Ubaldi et al. 1998; Cano et al. 2000; Loreille et al. 2001; Iniguez et al. 2003a; Iniguez et al. 2003b; Iniguez et al. 2006; Luciani et al. 2006; Rollo et al. 2006; Leles et al. 2008; Botella et al. 2010; Santiago-Rodriguez et al. 2013).

### *Next generation sequencing*

Large-scale sequencing yields a wide range of DNA sequences simultaneously, dependent or independent on targeted gene amplification (Kunin et al. 2008; Petrosino et al. 2009; Kuczynski et al. 2012). In a first strategy, PCR amplification of variable regions (V3, V6) of bacterial 16S rDNA using universal primers is followed by massive sequencing (Wang and Qian 2009). This approach yields data on the bacterial communities in coprolites (Tito et al. 2012). Moreover, additional source-tracking performed on 16S rDNA datasets permits to compare the mixture of bacterial taxa associated with coprolites to different modern sources (Tito et al. 2012).

The second strategy is the massive sequencing of the DNA without previous amplification (Petrosino et al. 2009; Kuczynski et al. 2012). Whole-genome shotgun sequencing (WGS) strategies can be used to gain access to the collection of genomes associated to specimens. These include sequences associated to the host, eukaryotes, bacteria and archaea as shown for two 1,300-year-old coprolites (Kunin et al. 2008; Tito et al. 2008). Moreover, WGS datasets can furthermore served to learn more about major

metabolic pathways associated to the collection of genomes associated to coprolites (Kunin et al. 2008; Tito et al. 2008).

**OUTPUTS**

*Resident gut microbiota*

The analyses of microbial communities preserved in coprolites, using massive parallel sequencing or PCR-amplifications, indicated that parts of the ancient gut microbiota are preserved (Ubaldi et al. 1998; Cano et al. 2000; Luciani et al. 2006; Rollo et al. 2006; Tito et al. 2008; Tito et al. 2012). The seven phyla mainly found in the modern human gut flora – i.e. *Firmicutes*, *Bacteroides*, *Actinobacteria*, *Fusobacteria*, *Proteobacteria*, *Verrumicrobia* and *Cyanobacteria* – have been all detected in coprolites (Ubaldi et al. 1998; Cano et al. 2000; Luciani et al. 2006; Rollo et al. 2006; Tito et al. 2008; Tito et al. 2012). Source-tracking performed on several pre-Columbian coprolites found a mixture of bacterial taxa similar to those coming from human children stool, primate gut or compost, respectively and in some cases no similarities to known sources were found (Tito et al. 2012). Moreover, recent analyses identified *Methanobrevibacter* spp., *Methanosphaera* spp. and *Sulfolobus* spp. in coprolites (Tito et al. 2008; Santiago-Rodriguez et al. 2013). Indeed, *Methanobrevibacter smithii* has been shown to be constant inhabitant of the human gut (Dridi et al. 2009).

The comparison of human coprolite samples from two different Indonesian cultures indicated that human intestinal microbiome was affected by diet related to cultural traditions (Santiago-Rodriguez et al. 2013). Furthermore, comparison to modern stool samples showed that coprolites exhibited more similarities to each other and to

22

stools from rural communities than to those coming from cosmopolitan communities. These findings suggest that the modern lifestyle may results in changes in the composition of the human gut flora (Tito et al. 2008; Tito et al. 2012).

*Pathogens*

Microbiological analyses of modern feces allowed the diagnosis of both intestinal and systemic infections (El Khechine et al. 2009; Jirku et al. 2012; Demeler et al. 2013; Keita et al. 2013). By analogy, coprolites are a material of choice for the retrospective diagnosis of diseases (**Figure 2**).

Intestinal tract pathogens – helminths and protozoa – have been diagnosed thanks to the microscopic observation of eggs and cysts; specific antigens and DNA detection. *Ascaris* spp. and *Trichuris* spp., causing ascariasis and trichuriasis respectively have been detected in 80% of human coprolites in European archeological sites and in 100% of those from medieval times (Bouchet et al. 2003b; Goncalves et al. 2003). Since *Ascaris* spp. are also known to infect pigs this observation questioned the initial source of transmission in connexion with the rising of agriculture the Neolithic times (Bouchet et al. 2003b; Goncalves et al. 2003; Leles et al. 2008).

*Dicrocoelium dendriticum*, responsible of bile duct infections, *Taenia* spp. causing cysticerosis and *Diphyllobothrium* spp. causing diphyllobothriasis have also been detected (**Figure 2**). *Diphyllobothrium* is known to infect fresh-water fishes and transmission to humans can occur through the consumption of raw or undercooked fish. *Diphyllobothrium* spp. eggs were observed in 75% of human coprolites from late

23

Neolithic times in Europe, questioning the importance of this parasitic infection at time (Le Bailly et al. 2005). Protozoa cysts belonging to *Entamoeba* spp., *Giardia* spp. and *Cyclospora* spp. were also identified in coprolites for Europe and America (Allison et al. 1999; Goncalves et al. 2003; Ortega and Bonavia, 2003).

Furthermore, unicellular eukaryotes and bacteria were also associated to coprolites. A dilated coprolite-rich colon of a Peruvian Inca mummy (1,400 AD) suggested that the Chagas disease, caused by *Trypanosoma* spp., might have circulated during time. This observation was further confirmed by PCR and by electron microscopic observations performed on soft tissues samples from the same mummy (Guhl et al. 1999). Considering the amplification products of the 16S rDNA gene with 74-100% similarities, bacteria from the genera *Acinetobacter* spp., *Clostridium* spp. and *Haemophilus* spp. were identified in human coprolites (Ubaldi et al. 1998; Luciani et al. 2006) (**Figure 2**). Sequences related to *Shigella* spp., responsible of shigellosis have also been detected (Araujo et al. 1998). Furthermore, metagenomic sequences of pathogenic microorganisms were found in WGS datasets from pre-Columbian coprolites including *Neisseria* spp. (*Neisseria gonorrhoeae*), *Yersinia* spp. (*Yersinia enterocolitica*), *Mycobacterium* spp. (*Mycobacterium tuberculosis*), *Plasmodium* spp., and *Shigella* spp. (Tito et al. 2008).

## CONCLUSION

Coprolite microbiology provides information regarding the human gut microbiota and diseases in past populations, after strict enforcement of standard practices in paleomicrobiology. These paleomicrobiological standards, developed over

24

the years, are essential precautions in order to avoid in-laboratory contamination and thus to ensure the interpretability of data. Currently, there are no established standardized procedures widely performed systematically to analyze coprolites but the techniques are the same than those used for routine diagnosis. Accordingly, investigations performed on coprolites already helped to estimate ancient human gut microbiota composition, and to expand knowledge about intestinal parasites that circulated in ancient populations. By drawing parallels to data obtained for modern stool samples, coprolites provided data about systemic infections. Yet, few data were obtained about bacteria and archaea, and even not at all about viruses, compared to those about intestinal parasites. Those disparities remain open fields for further investigations that will benefit from technical improvements and certainly provide useful data.

# REFERENCES

**Allison, M.J., Bergman, T., and Gerszten, E.** 1999. Further studies on fecal parasites in antiquity. Am J Clin Pathol **112**, 605-609.

**Araujo, A., Reinhard, K., Bastos, O.M., Costa, L.C., Pirmez, C., Iniguez, A., Vicente, A.C., Morel, C.M., and Ferreira, L.F.** 1998. Paleoparasitology: perspectives with new techniques. Rev Inst Med Trop Sao Paulo **40**, 371-376.

**Botella, H., Vargas, J., Arnay De La Rosa, M., Leles, D., Reimers, E., Vicente, A., and Iniguez, A.** 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 **105**, 1054-1056.

**Bouchet, F., Guidon, N., Dittmar, K., Harter, S., Ferreira, L.F., Chaves, S.M., Reinhard, K., and Araujo, A.** 2003a. Parasite remains in archaeological sites. Mem Inst Oswaldo Cruz **98** Suppl 1, 47-52.

**Bouchet, F., Harter, S., and Le Bailly, M.** 2003b. The state of the art of paleoparasitological research in the Old World. Mem Inst Oswaldo Cruz **98** Suppl 1, 95-101.

**Bouchet, F., Harter, S., Paicheler, J.C., Araujo, A., and Ferreira, L.F.** 2002. First recovery of *Schistosoma mansoni* eggs from a latrine in Europe (15-16th centuries). J Parasitol **88**, 404-405.

**Cano, R.J., Tiefenbrunner, F., Ubaldi, M., Del Cueto, C., Luciani, S., Cox, T., Orkand, P., Kunzel, K.H., and Rollo, F.** 2000. Sequence analysis of bacterial DNA in the colon and stomach of the Tyrolean Iceman. Am J Phys Anthropol **112**, 297-309.

**Carrio, J., Riquelme, J., Navarro, C., and Munuera, M.** 2001. Pollen in hyaena coprolites reflects late glacial landscape in southern Spain. Palaeogeography, Palaeoclimatology, Palaeoecology **176**, 193-205.

**Chaves, S.D.M., and Reinhard, K.** 2006. Critical analysis of coprolite evidence of medicinal plant use, Piauí, Brazil. Palaeogeography, Palaeoclimatology, Palaeoecology **237**, 110-118.

Cleeland, L.M., Reichard, M.V., Tito, R.Y., Reinhard, K.J., and Lewis, C.M., Jr. 2013. Clarifying Prehistoric Parasitism from a Complementary Morphological and Molecular Approach. J Archaeol Sci **40**, 3060-3066.

Cooper, A., and Poinar, H.N. 2000. Ancient DNA: do it right or not at all. Science **289**, 1139.

Demeler, J., Ramunke, S., Wolken, S., Ianiello, D., Rinaldi, L., Gahutu, J.B., Cringoli, G., Von Samson-Himmelstjerna, G., and Krucken, J. 2013. Discrimination of gastrointestinal nematode eggs from crude fecal egg preparations by inhibitor-resistant conventional and real-time PCR. PLoS One **8**, e61285.

Dentzien-Dias, P.C., Poinar, G., Jr., De Figueiredo, A.E., Pacheco, A.C., Horn, B.L., and Schultz, C.L. 2013. Tapeworm eggs in a 270 million-year-old shark coprolite. PLoS One **8**, e55007.

Dittmar, K. 2009. Old parasites for a new world: the future of paleoparasitological research. a review. *J* Parasitol **95**, 365-371.

Drancourt, M., and Raoult, D. 2005. Palaeomicrobiology: current issues and perspectives. *Nat Rev* Microbiol **3**, 23-35.

Dridi, B., Henry, M., El Khechine, A., Raoult, D., and Drancourt, M. 2009. High prevalence of Methanobrevibacter smithii and Methanosphaera stadtmanae detected in the human gut using an improved DNA detection protocol. PLoS One **4**, e7063.

Dufour, B., and Le Bailly, M. 2013. Testing new parasite egg extraction methods in paleoparasitology and an attempt at quantification. International Journal of Paleopathology (in press),

El Khechine, A., Henry, M., Raoult, D., and Drancourt, M. 2009. Detection of *Mycobacterium tuberculosis* complex organisms in the stools of patients with pulmonary tuberculosis. Microbiology **155**, 2384-2389.

Fernandes, A., Ferreira, L.F., Goncalves, M.L., Bouchet, F., Klein, C.H., Iguchi, T., Sianto, L., and Araujo, A. 2005. Intestinal parasite analysis in organic sediments collected from a 16th-century Belgian archeological site. Cad Saude Publica **21**, 329-332.

Fugassa, M., Gonzqlez Oliverez, E., and Petrigh, R. 2013. First paleoparasitological record of dioctophymatid egg in an archeological sample from Patagonia. Acta Tropica **128**, 175-177.

Gilbert, M.T., Jenkins, D.L., Gotherstrom, A., Naveran, N., Sanchez, J.J., Hofreiter, M., Thomsen, P.F., Binladen, J., Higham, T.F., Yohe, R.M., 2nd, Parr, R., Cummings, L.S., and Willerslev, E. 2008. DNA from pre-Clovis human coprolites in Oregon, North America. Science **320**, 786-789.

Goldberg, P., Berna, F., and Macphail, R.I. 2009. Comment on "DNA from pre-Clovis human coprolites in Oregon, North America". Science **325**, 148; author reply 148.

Goncalves, M.L., Araujo, A., Duarte, R., Da Silva, J.P., Reinhard, K., Bouchet, F., and Ferreira, L.F. 2002. Detection of *Giardia duodenalis* antigen in coprolites using a commercially available enzyme-linked immunosorbent assay. Trans R Soc Trop Med Hyg **96**, 640-643.

Goncalves, M.L., Araujo, A., and Ferreira, L.F. 2003. Human intestinal parasites in the past: new findings and a review. Mem Inst Oswaldo Cruz **98** Suppl 1, 103-118.

Goncalves, M.L., Da Silva, V.L., De Andrade, C.M., Reinhard, K., Da Rocha, G.C., Le Bailly, M., Bouchet, F., Ferreira, L.F., and Araujo, A. 2004. Amoebiasis distribution in the past: first steps using an immunoassay technique. Trans R Soc Trop Med Hyg **98**, 88-91.

Guhl, F., Jaramillo, C., Vallejo, G.A., Yockteng, R., Cardenas-Arroyo, F., Fornaciari, G., Arriaza, B., and Aufderheide, A.C. 1999. Isolation of *Trypanosoma cruzi* DNA in 4,000-year-old mummified human tissue from northern Chile. Am J Phys Anthropol **108**, 401-407.

Hofreiter, M., Serre, D., Poinar, H.N., Kuch, M., and Paabo, S. (2001. Ancient DNA. Nat Rev Genet **2**, 353-359.

Honore-Bouakline, S., Vincensini, J.P., Giacuzzo, V., Lagrange, P.H., and Herrmann, J.L. 2003. Rapid diagnosis of extrapulmonary tuberculosis by PCR: impact of sample preparation and DNA extraction. J Clin Microbiol **41**, 2323-2329.

Iniguez, A., Reinhard, K., Araujo, A., Ferreira, L., and Vicente, A. 2003a. *Enterobius vermicularis*: ancient DNA from North and South American human coprolites. Mem Inst Oswaldo Cruz **98** Suppl 1, 67-69.

Iniguez, A.M., Araujo, A., Ferreira, L.F., and Vicente, A.C. 2003b. Analysis of ancient DNA from coprolites: a perspective with random amplified polymorphic DNA-polymerase chain reaction approach. Mem Inst Oswaldo Cruz **98** Suppl 1, 63-65.

Iniguez, A.M., Reinhard, K., Carvalho Goncalves, M.L., Ferreira, L.F., Araujo, A., and Paulo Vicente, A.C. 2006. SL1 RNA gene recovery from *Enterobius vermicularis* ancient DNA in pre-Columbian human coprolites. Int J Parasitol **36**, 1419-1425.

Jenkins, D.L., Davis, L.G., Stafford, T.W., Jr., Campos, P.F., Hockett, B., Jones, G.T., Cummings, L.S., Yost, C., Connolly, T.J., Yohe, R.M., 2nd, Gibbons, S.C., Raghavan, M., Rasmussen, M., Paijmans, J.L., Hofreiter, M., Kemp, B.M., Barta, J.L., Monroe, C., Gilbert, M.T., and Willerslev, E. 2012. Clovis age Western Stemmed projectile points and human coprolites at the Paisley Caves. Science **337**, 223-228.

Jirku, M., Pomajbikova, K., Petrzelkova, K.J., Huzova, Z., Modry, D., and Lukes, J. 2012. Detection of *Plasmodium* spp. in human feces. Emerg Infect Dis **18**, 634-636.

Keita, A.K., Socolovschi, C., Ahuka-Mundeke, S., Ratmanov, P., Butel, C., Ayouba, A., Inogwabini, B.I., Muyembe-Tamfum, J.J., Mpoudi-Ngole, E., Delaporte, E., Peeters, M., Fenollar, F., and Raoult, D. 2013. Molecular evidence for the presence of *Rickettsia felis* in the feces of wild-living African apes. PLoS One **8**, e54679.

Kiatpathomchai, W., Boonsaeng, V., Tassanakajon, A., Wongteerasupaya, C., Jitrapakdee, S., and Panyim, S. 2001. A non-stop, single-tube, semi-nested PCR technique for grading the severity of white spot syndrome virus infections in Penaeus monodon. Dis Aquat Organ **47**, 235-239.

Kuch, M., and Poinar, H. 2012. Extraction of DNA from paleofeces. Methods Mol Biol **840**, 37-42.

**Kuczynski, J., Lauber, C.L., Walters, W.A., Parfrey, L.W., Clemente, J.C., Gevers, D., and Knight, R.** 2012. Experimental and analytical tools for studying the human microbiome. Nat Rev Genet **13**, 47-58.

**Kunin, V., Copeland, A., Lapidus, A., Mavromatis, K., and Hugenholtz, P.** 2008. A bioinformatician's guide to metagenomics. Microbiol Mol Biol Rev **72**, 557-578.

**Le Bailly, M., and Bouchet, F.** 2006. Paléoparasitologie et immunologie : l'exemple d'*Entamoeba histolytica*. Archéosciences, **30** : 129-135

**Le Bailly, M., Goncalves, M.L., Harter-Lailheugue, S., Prodeo, F., Araujo, A., and Bouchet, F.** 2008. New finding of *Giardia intestinalis* (Eukaryote, Metamonad) in Old World archaeological site using immunofluorescence and enzyme-linked immunosorbent assays. Mem Inst Oswaldo Cruz **103**, 298-300.

**Le Bailly, M., Leuzinger, U., and Bouchet, F.** 2003. *Dioctophymidae* eggs in coprolites from Neolithic site of Arbon-Bleiche 3 (Switzerland). J Parasitol **89**, 1073-1076. doi:10.1645/GE-3202RN

**Le Bailly, M., Leuzinger, U., Schlichtherle, H., and Bouchet, F.** 2005. *Diphyllobothrium*: Neolithic parasite? J Parasitol **91**, 957-959.

**Leles, D., Araujo, A., Ferreira, L.F., Vicente, A.C., and Iniguez, A.M.** 2008. Molecular paleoparasitological diagnosis of *Ascaris* sp. from coprolites: new scenery of ascariasis in pre-Colombian South America times. Mem Inst Oswaldo Cruz **103**, 106-108.

**Loreille, O., Roumat, E., Verneau, O., Bouchet, F., and Hanni, C.** 2001. Ancient DNA from *Ascaris*: extraction amplification and sequences from eggs collected in coprolites. Int J Parasitol **31**, 1101-1106.

**Luciani, S., Fornaciari, G., Rickards, O., Labarga, C.M., and Rollo, F.** 2006. Molecular characterization of a pre-Columbian mummy and in situ coprolite. Am J Phys Anthropol **129**, 620-629.

**Mitchell, P.D., Stern, E., and Tepper, Y.** 2008. Dysentery in the crusader kingdom of Jerusalem: an ELISA analysis of two medieval latrines in the City of Acre (Israel). *Journal of Archaeological* Science **35**, 1849-1853.

**Neumaier, M., Braun, A., and Wagener, C.** 1998. Fundamentals of quality assessment of molecular amplification methods in clinical diagnostics. International Federation of Clinical Chemistry Scientific Division Committee on Molecular Biology Techniques. Clin Chem **44**, 12-26.

**Ortega, Y.R., and Bonavia, D.** 2003. *Cryptosporidium*, *Giardia*, and *Cyclospora* in ancient Peruvians. *J* Parasitol **89**, 635-636.

**Pääbo, S., Poinar, H., Serre, D., Jaenicke-Despres, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L., and Hofreiter, M.** 2004. Genetic analyses from ancient DNA. Annu Rev Genet **38**, 645-679.

**P**etrosino, J.F., Highlander, S., Luna, R.A., Gibbs, R.A., and Versalovic, J. 2009. Metagenomic pyrosequencing and microbial identification. Clin Chem **55**, 856-866.

**Poinar, H.N.** 1998. Molecular Coproscopy: Dung and Diet of the Extinct Ground Sloth *Nothrotheriops* shastensis. Science **281**, 402-406.

**Poinar, H.N.** 2002. The genetic secrets some fossils hold. Acc Chem Res **35**, 676-684.

**Poinar, H.N., Kuch, M., Sobolik, K.D., Barnes, I., Stankiewicz, A.B., Kuder, T., Spaulding, W.G., Bryant, V.M., Cooper, A., and Paabo, S.** 2001. A molecular analysis of dietary diversity for three archaic Native Americans. Proc Natl Acad Sci U S A **98**, 4317-4322.

**Raoult, D., Aboudharam, G., Crubezy, E., Larrouy, G., Ludes, B., and Drancourt, M.** 2000. Molecular identification by "suicide PCR" of *Yersinia pestis* as the agent of medieval black death. Proc Natl Acad Sci U S A **97**, 12800-12803.

**Reinhard, K., and Bryant, V.J.** 2008. "Pathoecology and the Future of Coprolite Studies in Bioarchaeology," in Reanalysis and reinterpretation in southyestern bioarcheology, ed. A.W.M. Stodder. (Tempe, Arizona: Arizona State University Press).

**Reinhard, K., Fink, T.M., and Skiles, J.** 2003. A case of megacolon in Rio Grande valley as a possible case of Chagas disease. Mem Inst Oswaldo Cruz **98** Suppl 1, 165-172.

**Reinhard, K.J., and Bryant, V.M.J.** 1992. "Coprolite Analysis: A Biological Perspective on Archaeology," in Advances in Archaeological Method and Theory, ed. M. Shiffer. University of Arizona Press, Tucson, 245–288.

**Rocha, G.C., Harter-Lailheugue, S., Le Bailly, M., Araujo, A., Ferreira, L.F., Da Serra-Freire, N.M., and Bouchet, F.** 2006. Paleoparasitological remains revealed by seven historic contexts from "Place d'Armes", Namur, Belgium. Mem Inst Oswaldo Cruz **101** Suppl 2, 43-52.

**Rocha, G.C.D., and Serra-Freire, N.M.** 2009. Paleoparasitology at "Place d'Armes", Namur, Belgium: a biostatistics analysis of trichurid eggs between the Old and New World. Rev Bras Parasitol Vet **18**, 70-74.

**Rollo, F., Ermini, L., Luciani, S., Marota, I., and Olivieri, C.** 2006. Studies on the preservation of the intestinal microbiota's DNA in human mummies from cold environments. Med Secoli **18**, 725-740.

**Rosenstraus, M., Wang, Z., Chang, S.Y., Debonville, D., and Spadoro, J.P.** 1998. An internal control for routine diagnostic PCR: design, properties, and effect on clinical performance. *J Clin* Microbiol **36**, 191-197.

**Santiago-Rodriguez, T.M., Narganes-Storde, Y.M., Chanlatte, L., Crespo-Torres, E., Toranzos, G.A., Jimenez-Flores, R., Hamrick, A., and Cano, R.J.** 2013. Microbial communities in pre-columbian coprolites. PLoS One **8**, e65191.

**Shin, D.H., Lim, D.S., Choi, K.J., Oh, C.S., Kim, M.J., Lee, I.S., Kim, S.B., Shin, J.E., Bok, G.D., Chai, J.Y., and Seo, M.** 2009. Scanning electron microscope study of ancient parasite eggs recovered from Korean mummies of the Joseon Dynasty. J Parasitol **95**, 137-145.

**Taylor, G.M., Goyal, M., Legge, A.J., Shaw, R.J., and Young, D.** 1999. Genotypic analysis of *Mycobacterium tuberculosis* from medieval human remains. Microbiology **145** ( Pt 4), 899-904.

**Tito, R.Y., Belknap, S.L., 3rd, Sobolik, K.D., Ingraham, R.C., Cleeland, L.M., and Lewis, C.M., Jr.** 2011. Brief communication: DNA from early Holocene American dog. Am J Phys Anthropol **145**, 653-657.

Tito, R.Y., Knights, D., Metcalf, J., Obregon-Tito, A.J., Cleeland, L., Najar, F., Roe, B., Reinhard, K., Sobolik, K., Belknap, S., Foster, M., Spicer, P., Knight, R., and Lewis, C.M., Jr. 2012. Insights from characterizing extinct human gut microbiomes. PLoS One **7**, e51146.

Tito, R.Y., Macmil, S., Wiley, G., Najar, F., Cleeland, L., Qu, C., Wang, P., Romagne, F., Leonard, S., Ruiz, A.J., Reinhard, K., Roe, B.A., and Lewis, C.M., Jr. 2008. Phylotyping and functional analysis of two ancient human microbiomes. PLoS One **3**, e3703.

Tran, T.N., Aboudharam, G., Raoult, D., and Drancourt, M. 2011. Beyond ancient microbial DNA: nonnucleotidic biomolecules for paleomicrobiology. Biotechniques **50**, 370-380.

Ubaldi, M., Luciani, S., Marota, I., Fornaciari, G., Cano, R.J., and Rollo, F. 1998. Sequence analysis of bacterial DNA in the colon of an Andean mummy. Am J Phys Anthropol **107**, 285-295.

Volossiouk, T., Robb, E.J., and Nazar, R.N. 1995. Direct DNA extraction for PCR-mediated assays of soil organisms. Appl Environ Microbiol **61**, 3972-3976.

Wang, Y., and Qian, P.Y. 2009. Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *PLoS One* **4**, e7401.

Willerslev, E., and Cooper, A. 2005. Ancient DNA. Proc Biol Sci **272**, 3-16.

Willerslev, E., Hansen, A.J., Binladen, J., Brand, T.B., Gilbert, M.T., Shapiro, B., Bunce, M., Wiuf, C., Gilichinsky, D.A., and Cooper, A. 2003. Diverse plant and animal genetic records from Holocene and Pleistocene sediments. Science **300**, 791-795.

Zeaiter, Z., Fournier, P.E., Greub, G., and Raoult, D. 2003. Diagnosis of Bartonella endocarditis by a real-time nested PCR assay using serum. J Clin Microbiol **41**, 919-925.

Zink, A., Haas, C.J., Reischl, U., Szeimies, U., and Nerlich, A.G. 2001. Molecular analysis of skeletal tuberculosis in an ancient Egyptian population. J Med Microbiol **50**, 355-366.

**Legends of Figures**

**Figure 1. (A)** Bibliometry. **(B)** Repartition and dating **(B)** of human and mixed coprolites investigated during the last 21 years. The human coprolites reviewed herein were found in about 87 different archeological sites: 37 (42.5%) and 24 (27.6%) were located in South and North America respectively, whereas 18 (20.7%) were found in Europe and 3 - 5 (3.4 – 5.7%) in Africa and Asia. Analyses relying on microscopic analyses are labeled with yellow squares, non-nucleotidic biomolecules analyses are marked with blue triangles and molecular analyses are marked with orange circles. The corresponding references are listed in **Supplementary Table S1**.

**Figure 2.** Intestinal and systemic pathogens identified in human coprolites. The pathogens are grouped accordingly to their taxonomic classification into intestinal helminths, intestinal protozoa and unicellular eukaryotes and bacteria. The method of identification that yielded positive results is marked in green and negative tests are marked in grey. Microscopic observations were performed by light – or electron microscopy. Nonnucleotidic biomolecule detection included IFA and ELISA tests and molecular detections were performed using PCR-amplifications and next generation sequencing.

## A  Bibliometry

To focused border the literature of coprolite studies the search databases were limited to Pubmed and the Memórias do Instituto Oswaldo Cruz databases (journal homepage: http://www.scielo.br) and the Journal of Parasitology (journal homepage: http://www.journalofparasitology.org/).

Only publications since January 1992 in journals with an impact factor >1.4 were considered. "Paleo feces", "coprolites", "ancient DNA", "mummies", "human intestinal microbiota" and "human fossils" were used as keywords, independently and combined to search for papers published in English. References from previous reviews and references from selected papers are also reviewed.

Between January 1992 and September 2013, 124 papers related to coprolites studies were published. 76 (61.3%) reported non-DNA analysis: 50% were based on macroscopic/microscopic observations, and 11.3% were non-nucleotidic biomolecule analyses. 28 (22.6%) papers reported DNA analyses and 20 (16.1%) papers were reviews. 66 (53.2%) papers dealt with human or mixed coprolites. The bibliometry analysis indicates a trend in number of publications and an incipient shift of methods used to analyze coprolites from non-DNA to DNA analyses.

## B  Repartition of coprolites
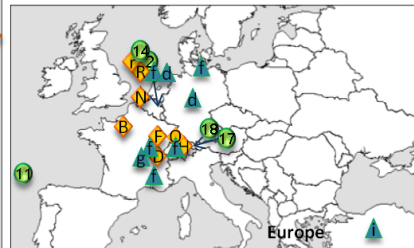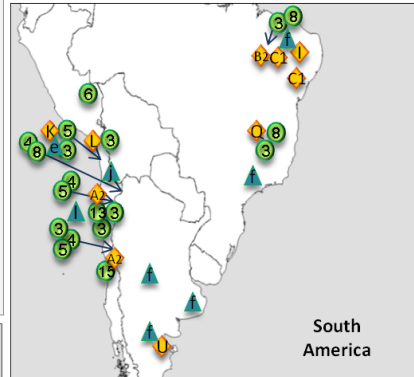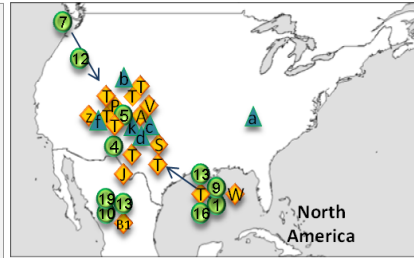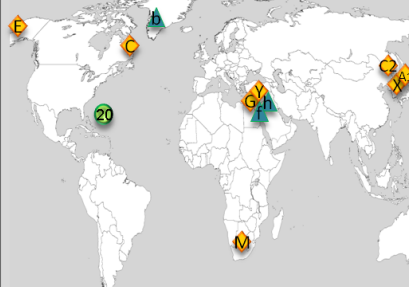


## C  Dating of coprolites



**Figure 1.**

37
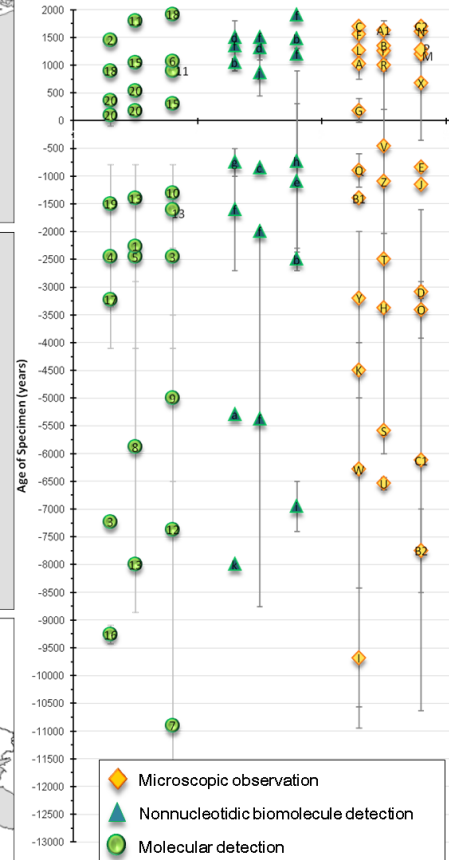
## Pathogens identified in coprolites



**Microorganisme** — **Methode of identification**

**Digestive tract pathogens**
**Intestinal helminths**
*Ancylostomid* spp.
*Ascaris* spp.*
*Capillaria* spp.
*Dicrocoelium* spp.
*Dioctophymidae* spp.
*Diphyllobothrium* spp.
*Dipylidium caninum*
*Enterobius* spp.
*Fasciola* spp.
*Gymnophalloides seoi*
*Metagonimus yokogawai*
*Physaloptera* spp.
*Schistosoma* spp.
*Taenia* spp.
*Trichostrongylus* spp.
*Trichuris* spp.*
**Intestinal protozoa**
*Cryptosporidium* spp.
*Cyclospora* spp.
*Entamoeba* spp.
*Giardia* spp.*
*Isospora belli*
*Sarcocystis hominis*
**Unicellular eukaryotes and bacteria**
**Protozoa**
*Leishmania* spp.
*Plasmodium* spp.
**Systemic bacterial pathogens**
*Clostridium* spp.
*Haemophilus* spp.
*Mycobacterium* spp.
*Neisseria* spp.
*Shigella* spp.
*Yersinia* spp.

Columns: Microscopy — Nonnucleotidic biomolecule detection — Molecular detection

* additionally detected in a potentially mixed animal and human coprolite

**Figure 2.**

39

**Human coprolites as a source for paleomicrobiology**

**Running title:** Paleomicrobiology of coprolites

Sandra Appelt[1], Matthieu Le Bailly[2] and Michel Drancourt[1*]

[1] Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, Inserm 1095, 13005 Marseille, France. [2] Franche-Comté University, CNRS UMR 6249 Chrono-Environment, 25 030 Besançon, France.

[*]Corresponding author. Michel.Drancourt@medicine.univ-mrs.fr

**SUPPLEMENTAL MATERIAL CONTENTS**

**SUPPLEMENTARY TABLES**

Supplementary Table S1-S2

**SUPPLEMENTARY REFERENCES**

**Supplementary Table S1. Scientific studies performed on human and potentially mixed animal and human coprolites.** Reported is the type of investigation, the archeological site were the studied coprolites were found, host nature and the corresponding references as the identifier corresponding to **Fig. 1**.

| Methode of Analysis | Sample Source | Archeological Site | Identifi | Reference |
|---|---|---|---|---|
| Microscopic observations | human | Canyon de Chelly (Arizona) | A | (Sutton and Reinhard, 1995) |
| Microscopic observations | human | Chungcheongnam (Korea) | A1 | (Shin et al. 2011) |
| Microscopic observations | human | Huaricanga,Caballete (Peru) | A2 | (Haas et al. 2013) |
| Microscopic observations | human | Paris (France) | B | (Bouchet, 1995) |
| Microscopic observations | human | Durango (Mexico) | B1 | (Jimenez et al. 2012) |
| Microscopic observations | human | Piaui (Brazil) | B2 | (Chaves and Reinhard, 2006) |
| Microscopic observations | mixed | Patagonia (Chile) | B3 | (Taglioretti et al. 2013) |
| Microscopic observations | mixed | Ferryland (Canada) | C | (Home and Tuck, 1996) |
| Microscopic observation | human | Yongin (Korea) | C1 | (Shin et al. 2009a) |
| Microscopic observation | human | São Raimundo Nonato, Piauí and | C2 | (Sianto et al. 2012) |
| Microscopic observations | human | Chalain Lake (France) | D | (Bouchet, 1997) |
| Microscopic observations | human | Aleutian Islands (Alaska) | E | (Bouchet et al. 2001) |
| Microscopic observations | human | Montbéliard (France) | F | (Bouchet et al. 2002) |
| Microscopic observations | human | Dakhleh Oasis (Egypt) | G | (Horne, 2002) |
| Microscopic observations | human | Arbon-Bleiche (Switzerland) | H | (Le Bailly et al. 2003) |
| Microscopic observations | mixed | Pernambuco (Brazil) | I | (Guerra et al. 2003) |
| Macroscopic observations | human | Chihuahuan Desert (Mexico) | J | (Reinhard and Urban, 2003) |
| Microscopic observations | human | Chinchorro (Chile) | K | (Reinhard et al. 2003) |
| Microscopic observations | human (suspected) | Lluta Valley (Chile) | L | (Santoro et al. 2003) |
| Microscopic observations | human | Greefswald (South Africa) | M | (Dittmar and Steyn, 2004) |
| Microscopic observations | human (suspected) | Raversjde (Belgiums) | N | (Fernandes et al. 2005) |
| Microscopic observations | human | Arbon Bleiche (Switzerland) | O | (Le Bailly et al. 2005) |
| Microscopic observations | human | Clen Canyon, Utha (USA) | P | (Moore et al. 1974) |

| | | | | |
|---|---|---|---|---|
| Microscopic observations | human | Minas Gerais (Brazil) | Q | (Sianto et al. 2005) |
| Microscopic observations | mixed | Namur (Belgium) | R | (Rocha et al. 2006) |
| Microscopic observations | mixed | Namur (Belgium) | r | (Rocha and Serra-Freire, 2009) |
| Microscopic observations | human | Hinds Cave (Texas) | S | (Dean, 2006) |
| Microscopic observations | human | Nevada, Texas, Utha (USA) | T | (Reinhard et al. 2007) |
| Microscopic observations | mixed | Patagonia (Agentinia) | U | (Fugassa et al. 2008) |
| Microscopic observations | human | Antelope cave (Arizona) | V | (Johnson et al. 2008) |
| Microscopic observations | mixed | Hinds Cave (Texas) | W | (Riley, 2008) |
| Microscopic observations | human | Yongin (Korea) | X | (Shin et al. 2009b) |
| Microscopic observations | human | El Deir, Oasis Kharga (Egypt) | Y | (Le Bailly et al. 2010) |
| Microscopic observations | mixed | Mojave Country (Arizona) | Z | (Fugassa et al. 2011) |
| Molecular Analysis | human | Hinds Cave (Texas) | 1 | (Poinar et al. 2001) |
| Molecular Analysis | mixed | Namur (Belgium) | 2 | (Loreille et al. 2001) |
| Molecular Analysis | human | Azapa, Antofagasta, Caserones (Chile) | 3 | (Iniguez et al. 2003a) |
| Molecular Analysis | human | Piaui, Unai (Brazil) | 3 | (Iniguez et al. 2003a) |
| Molecular Analysis | human | Antelope House, Canyon de Chelly | 4 | (Iniguez et al. 2003b) |
| Molecular Analysis | human | Tulan, Caserones, Tiliviche (Chile) | 4 | (Iniguez et al. 2003b) |
| Molecular Analysis | human | Antelope House (USA) | 5 | (Iniguez et al. 2006) |
| Molecular Analysis | human | Tulan, Caserones, Tiliviche (Chile) | 5 | (Iniguez et al. 2006) |
| Molecular Analysis | human | Cuzco (Peru) | 6 | (Luciani et al. 2006) |
| Molecular Analysis | human | Oregon (USA) | 7 | (Gilbert et al. 2008) |
| Molecular Analysis | human | Brazil, Peru, Chile | 8 | (Leles et al. 2008) |
| Molecular Analysis | human | Hinds Cave (Texas) | 9 | (Reinhard and Bryant, 2008) |
| Molecular Analysis | human | Rio Zape (Mexico) | 10 | (Tito et al. 2008) |
| Molecular Analysis | human | Tenerife (Spain) | 11 | (Botella et al. 2010) |
| Molecular Analysis | mixed | Oregon (USA) | 12 | (Jenkins et al. 2012) |
| Molecular Analysis | human | Caserones (Chile) | 13 | (Tito et al. 2012) |

| | | | | |
|---|---|---|---|---|
| Molecular Analysis | human | Hinds Cave (Texas) | 13 | (Tito et al. 2012) |
| Molecular Analysis | human | Rio Zape (Mexico) | 13 | (Tito et al. 2012) |
| Molecular Analysis | human | Andean | 15 | (Ubaldi et al. 1998) |
| Molecular Analysis | human | Hinds Cave (Texas) | 16 | (Tito et al. 2011) |
| Molecular Analysis | human | Alpes (Austria/Italy) | 17 | (Cano et al. 2000) |
| Molecular Analysis | human | Northern Italy | 18 | (Rollo et al. 2006) |
| Molecular Analysis | human | Rio Zape (Mexico) | 19 | (Cleeland et al. 2013) |
| Molecular Analysis | human | Sorcé and Guayanilla, Island of | 20 | (Santiago-Rodriguez et al. 2013) |
| nonnucleotidic biomolecule | human | Kentuckey (USA) | a | (Sobolik et al. 1996) |
| nonnucleotidic biomolecule | human | Westcoast of Greenland, Nevada | b | (Lin and Connor, 2001) |
| nonnucleotidic biomolecule | human | Cowboy Wash (USA) | c | (Marlar et al. 2000) |
| nonnucleotidic biomolecule | mixed | USA, Germany, Belgium | d | (Goncalves et al. 2002) |
| nonnucleotidic biomolecule | human | Los Gavilanes (Peru) | e | (Ortega and Bonavia, 2003) |
| nonnucleotidic biomolecule | human (suspected) | Argentina, Brazil, Chile, USA, | f | (Goncalves et al. 2004) |
| nonnucleotidic biomolecule | human | La Mothe (France) | g | (Le Bailly et al. 2008) |
| nonnucleotidic biomolecule | human (suspected) | City of Acre (Israel) | h | (Mitchell et al. 2008) |
| nonnucleotidic biomolecule | mixed | Konja (Turkey) | i | (Shillito et al. 2009) |
| nonnucleotidic biomolecule | human | Atacama Desert (Chile) | j | (Vinton et al. 2009) |
| nonnucleotidic biomolecule | human | Utha (USA) | k | (Hagey et al. 2010) |
| nonnucleotidic biomolecule analyses | human | Andean (South America) | l | (Allison et al. 1999) |

**Supplementary Table S2. PCR Systems used to amplify microbial aDNA out of coprolites.** Shown are the targeted microorganisms, the genomic regions, the length of amplification products (bp) as corresponding primer-sets, and the age of those coprolite specimens that yielded positive amplification results.

| Microorganism | Target | Method | Microbial aDNA (bp) | Primer | Sequence 5'→ 3' | Positive Tested Specimen-age | Reference |
|---|---|---|---|---|---|---|---|
| **Eukaryotes** | | | | | | | |
| *Ascaris* spp. | 18S rDNA* | PCR | 123 | Asc6 | CGAACGGCTCATTACAACAG | 1,392 -1,800 AD | (Loreille et al. 2001; Botella et al. 2010) |
| | | | | Asc7 | TCTAATAGATGCGCTCGTC | | |
| | | PCR | 99 | Asc8 | ATACATGCACCAAAGCTCCG | | |
| | | | | Asc9 | GCTATAGTTATTCAGAGTCACC | | |
| | | PCR | 147 | Asc10 | CCATGCATGTCTAAGTTCAA | | |
| | | | | Asc11 | CARAAAWTCGGAGCTTTGGT | | |
| | cyt b | PCR | 98-142 | Asc1 | GTTAGGTTACCGTCTAGTAAGG | 8,860 BP-1,905 AD | (Loreille et al. 2001; Leles et al. 2008) |
| | | | | Asc2 | CACTCAAAAAGGCCAAAGCACC | | |
| | cytochrome c oxidase subunit 1** | PCR | 248 | COX1F | GGATCTTGACTCTCGGGCTTA | 1,800 AD | (Botella et al. 2010) |
| | | | | As-Co1R | ACATAATGAAAATGACTAACAAC | | |
| | | PCR | 199 | As-Co1F COX1R | TTTTTTGGTCATCCTGAGGTTTAT GCCCGAGAGTCAAGATCCAT | | |
| | NADH dehydrogenase** | PCR | 152 | NAD1F | CTCCTCTGAATTCTTCGGAAA | 1,800 AD | (Botella et al. 2010) |
| | | | | NAD1R | CAGAAAACCCAATCAAACACA | | |
| *Enterobius* spp. | 5S rDNA° | Nested-PCR | 419 | Entf | CACTTGCTATACCAACAACAC | 4,110 BP-900 AD | (Iniguez et al. 2003b; Iniguez et al. 2006) |
| | | | | Entr | GCGCTACTAAACCATAGAG | | |
| | | | 198 | Eva | ACAACACTTGCACGTCTC | | |

| | | | | Evb | GAATTGCTCGTTTGC | | |
|---|---|---|---|---|---|---|---|
| diverse fungi | ITS region | PCR | 645 | ITS1F | CTTGGTCATTTAGAGGAAGTAA | 100 BP - 500 AD | (Santiago-Rodriguez et al. 2013) |
| | | | | ITS4B | TCCTCCGCTTATTGATATGC | | |
| **Bacteria** | | | | | | | |
| *Bacteroides* spp. | 16S rDNA | PCR | 541 | HF183F | ATCATGAGTTCACATGTCCG | 100 BP - 500 AD | (Santiago-Rodriguez et al. 2013) |
| | | | | Bac708R | CAATCGGAGTTCTTCGTG | | |
| *Bacteroides* spp. | 16S rDNA | PCR | 150 | BacCan | GGAGCGCAGACGGGTTTT | 100 BP - 500 AD | (Santiago-Rodriguez et al. 2013) |
| | | | | Bac708R | CAATCGGAGTTCTTCGTG | | |
| diverse bacteria (16S) | 16S rDNA | PCR | 523 | 8dF | AGAGTTTGTTCMTGGCTCAG | 100 BP - 500 AD | (Santiago-Rodriguez et al. 2013) |
| | | | | K2R | GTATTACCGCGGCTGCTGG | | |
| diverse bacteria (16S) | 16S rDNA | PCR | 98 | 29f | TGGCTCAGATTGAACGCTG | 980 - 1,170 AD | (Ubaldi et al. 1998; Luciani et al. 2006) |
| | | | | 98r | CCAGACATTACTCACCCGTCC | | |
| diverse bacteria (16S) | 16S rDNA | PCR | | 8F | AGCGTCAAACTTTTAAATTGAA | 3,350-3,100 BP | (Cano et al. 2000) |
| | | | 556 | 564R | CCTGCGTGCGCTTTACGCCC | | |
| | | | 576 | 584R | ACATCTGACTTAACAAACCG | | |
| | | | 797 | 805R | TCGACATCGTTTACGGCGTG | | |
| diverse bacteria (16S) | 16S rDNA V6 | PCR | 188 | 341F | CCTACGGGRSGCAGCAG | 1,400-8,000 BP | (Tito et al. 2012) |
| | | | | 529R | ACCGCGGCKGCTGGC | | |
| **Archeae** | | | | | | | |
| diverse archaea (16S) | 16S rDNA | PCR | 915 | Arch21F | TTCCGGTTGATCCYGCCGGA | 100 BP - 500 AD | (Santiago-Rodriguez et al. 2013) |
| | | | | Arch958R | GCGTTGAMTCCAATT | | |

*overlapping Primer-set amplifying totally 176 bp of the 18S rDNA gene; **Primer-set lead to unspecific amplification; ° no 419-bp amplification product was obtained

**SUPPLEMENTARY REFERENCES**

**Allison, M.J., Bergman, T., and Gerszten, E.** 1999. Further studies on fecal parasites in antiquity. Am J Clin Pathol **112**, 605-609.

**Botella, H., Vargas, J., Arnay De La Rosa, M., Leles, D., Reimers, E., Vicente, A., and Iniguez, A.** 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 **105,** 1054-1056.

**Bouchet, F.** 1995. Recovery of helminth eggs from archeological excavations of the Grand Louvre (Paris, France). J Parasitol **81,** 785-787.

**Bouchet, F.** 1997. Intestinal capillariasis in neolithic inhabitants of Chalain (Jura, France). Lancet **349,** 256.

**Bouchet, F., Harter, S., Paicheler, J.C., Araujo, A., and Ferreira, L.F.** 2002. First recovery of *Schistosoma mansoni* eggs from a latrine in Europe (15-16th centuries). J Parasitol **88,** 404-405.

**Bouchet, F., West, D., Lefevre, C., and Corbett, D.** 2001. Identification of parasitoses in a child burial from Adak Island (Central Aleutian Islands, Alaska). C R Acad Sci III **324,** 123-127.

**Cano, R.J., Tiefenbrunner, F., Ubaldi, M., Del Cueto, C., Luciani, S., Cox, T., Orkand, P., Kunzel, K.H., and Rollo, F.** 2000. Sequence analysis of bacterial DNA in the colon and stomach of the Tyrolean Iceman. Am J Phys Anthropol **112,** 297-309.

**Cleeland, L.M., Reichard, M.V., Tito, R.Y., Reinhard, K.J., and Lewis, C.M., Jr.** 2013. Clarifying Prehistoric Parasitism from a Complementary Morphological and Molecular Approach. J Archaeol Sci **40,** 3060-3066.

**Chaves, S.D.M., and Reinhard, K.** 2006. Critical analysis of coprolite evidence of medicinal plant use, Piauí, Brazil. Palaeogeography, Palaeoclimatology, Palaeoecology **237**, 110-118.

**Dean, G.** 2006. The science of coprolite analysis: The view from Hinds cave. Palaeogeography, Palaeoclimatology, Palaeoecology **237,** 67-79.

**Dittmar, K., and Steyn, M.** 2004. Paleoparasitological analysis of coprolites from K2, an Iron Age archaeological site in South Africa: the first finding of *Dicrocoelium sp.* eggs. J Parasitol **90,** 171-173.

**Fernandes, A., Ferreira, L.F., Goncalves, M.L., Bouchet, F., Klein, C.H., Iguchi, T., Sianto, L., and Araujo, A.** 2005. Intestinal parasite analysis in organic sediments collected from a 16th-century Belgian archeological site. Cad Saude Publica **21,** 329-332.

**Fugassa, M.H., Reinhard, K.J., Johnson, K.L., Gardner, S.L., Vieira, M., and Araujo, A.** 2011. Parasitism of prehistoric humans and companion animals from Antelope Cave, Mojave County, northwest Arizona. J Parasitol **97,** 862-867.

**Fugassa, M.H., Taglioretti, V., Goncalves, M.L., Araujo, A., Sardella, N.H., and Denegri, G.M.** 2008. Capillaria spp. eggs in Patagonian archaeological sites: statistical analysis of morphometric data. Mem Inst Oswaldo Cruz **103,** 104-105.

**Gilbert, M.T., Jenkins, D.L., Gotherstrom, A., Naveran, N., Sanchez, J.J., Hofreiter, M., Thomsen, P.F., Binladen, J., Higham, T.F., Yohe, R.M., 2nd, Parr, R., Cummings, L.S., and Willerslev, E.** 2008. DNA from pre-Clovis human coprolites in Oregon, North America. Science **320,** 786-789.

**Goncalves, M.L., Araujo, A., Duarte, R., Da Silva, J.P., Reinhard, K., Bouchet, F., and Ferreira, L.F.** 2002. Detection of *Giardia duodenalis* antigen in coprolites using a commercially available enzyme-linked immunosorbent assay. Trans R Soc Trop Med Hyg **96,** 640-643.

**Goncalves, M.L., Da Silva, V.L., De Andrade, C.M., Reinhard, K., Da Rocha, G.C., Le Bailly, M., Bouchet, F., Ferreira, L.F., and Araujo, A.** 2004. Amoebiasis distribution in the past: first steps using an immunoassay technique. Trans R Soc Trop Med Hyg **98,** 88-91.

**Guerra, R.M.S.M. C., Gazeta, G.S., Amorim, M., Duarte, A.N., and Serra-Freire, N.M.** 2003. Ecological analysis of acari recovered from coprolites from archaeological site of northeast Brazil. Mem Inst Oswaldo Cruz **98** Suppl 1**,** 181-190.

Haas, J., Creamer, W., Huaman Mesia, L., Goldstein, D., Reinhard, K., and Rodriguez, C.V. 2013. Evidence for maize (*Zea mays*) in the Late Archaic (3000-1800 B.C.) in the Norte Chico region of Peru. Proc Natl Acad Sci U S A **110,** 4945-4949.

Hagey, L.R., Vidal, N., Hofmann, A.F., and Krasowski, M.D. 2010. Evolutionary diversity of bile salts in reptiles and mammals, including analysis of ancient human and extinct giant ground sloth coprolites. BMC Evol Biol **10,** 133.

Home, P.D., and Tuck, J.A. 1996. Archaeoparasitology at a 17th century colonial site in Newfoundland. J Parasitol **82,** 512-515.

Horne, P.D. 2002. First evidence of enterobiasis in ancient Egypt. J Parasitol 88**,** 1019-1021.

Iniguez, A.M., Araujo, A., Ferreira, L.F., and Vicente, A.C. 2003a. Analysis of ancient DNA from coprolites: a perspective with random amplified polymorphic DNA-polymerase chain reaction approach. Mem Inst Oswaldo Cruz **98** Suppl 1**,** 63-65.

Iniguez, A.M., Reinhard, K., Carvalho Goncalves, M.L., Ferreira, L.F., Araujo, A., and Paulo Vicente, A.C. 2006. SL1 RNA gene recovery from *Enterobius vermicularis* ancient DNA in pre-Columbian human coprolites. Int J Parasitol **36,** 1419-1425.

Iniguez, A.M., Reinhard, K.J., Araujo, A., Ferreira, L.F., and Vicente, A.C. 2003b. *Enterobius vermicularis*: ancient DNA from North and South American human coprolites. Mem Inst Oswaldo Cruz **98** Suppl 1**,** 67-69.

Jenkins, D.L., Davis, L.G., Stafford, T.W., Jr., Campos, P.F., Hockett, B., Jones, G.T., Cummings, L.S., Yost, C., Connolly, T.J., Yohe, R.M., 2nd, Gibbons, S.C., Raghavan, M., Rasmussen, M., Paijmans, J.L., Hofreiter, M., Kemp, B.M., Barta, J.L., Monroe, C., Gilbert, M.T., and Willerslev, E. 2012. Clovis age Western Stemmed projectile points and human coprolites at the Paisley Caves. Science **337,** 223-228.

Jimenez, F.A., Gardner, S.L., Araujo, A., Fugassa, M., Brooks, R.H., Racz, E., and Reinhard, K.J. 2012. Zoonotic and human parasites of inhabitants of Cueva

de los Muertos Chiquitos, Rio Zape Valley, Durango, Mexico. J Parasitol **98,** 304-309.

**Johnson, K.L., Reinhardt, K.J., Sianto, L., Araujo, A., Gardner, S.L., and Janovy, J., Jr.** 2008. A tick from a prehistoric Arizona coprolite. J Parasitol **94,** 296-298. doi:10.1645/GE-1059.1

**Le Bailly, M., Goncalves, M.L., Harter-Lailheugue, S., Prodeo, F., Araujo, A., and Bouchet, F.** 2008. New finding of *Giardia intestinalis* (Eukaryote, Metamonad) in Old World archaeological site using immunofluorescence and enzyme-linked immunosorbent assays. Mem Inst Oswaldo Cruz **103,** 298-300.

**Le Bailly, M., Leuzinger, U., and Bouchet, F.** 2003. *Dioctophymidae* eggs in coprolites from Neolithic site of Arbon-Bleiche 3 (Switzerland). J Parasitol **89,** 1073-1076.

**Le Bailly, M., Leuzinger, U., Schlichtherle, H., and Bouchet, F.** 2005. *Diphyllobothrium*: Neolithic parasite? J Parasitol **91,** 957-959.

**Le Bailly, M., Mouze, S., Da Rocha, G.C., Heim, J.L., Lichtenberg, R., Dunand, F., and Bouchet, F.** 2010. Identification of *Taenia sp.* in a mummy from a Christian Necropolis in El-Deir, Oasis of Kharga, ancient Egypt. J Parasitol 96**,** 213-215.

**Leles, D., Araujo, A., Ferreira, L.F., Vicente, A.C., and Iniguez, A.M.** 2008. Molecular paleoparasitological diagnosis of *Ascaris sp.* from coprolites: new scenery of ascariasis in pre-Colombian South America times. Mem Inst Oswaldo Cruz **103,** 106-108.

**Lin, D.S., and Connor, W.E.** 2001. Fecal steroids of the coprolite of a Greenland Eskimo mummy, AD 1475: a clue to dietary sterol intake. Am J Clin Nutr 74**,** 44-49.

**Loreille, O., Roumat, E., Verneau, O., Bouchet, F., and Hanni, C.** 2001. Ancient DNA from *Ascaris*: extraction amplification and sequences from eggs collected in coprolites. Int J Parasitol **31,** 1101-1106.

**Luciani, S., Fornaciari, G., Rickards, O., Labarga, C.M., and Rollo, F.** 2006. Molecular characterization of a pre-Columbian mummy and in situ coprolite. Am J Phys Anthropol **129,** 620-629.

Marlar, R.A., Leonard, B.L., Billman, B.R., Lambert, P.M., and Marlar, J.E. 2000. Biochemical evidence of cannibalism at a prehistoric Puebloan site in southwestern Colorado. Nature **407,** 74-78.

Mitchell, P.D., Stern, E., and Tepper, Y. 2008. Dysentery in the crusader kingdom of Jerusalem: an ELISA analysis of two medieval latrines in the City of Acre (Israel). Journal of Archaeological Science **35,** 1849-1853.

Moore, J.G., Grundmann, A.W., Hall, H.J., and Fry, G.F. 1974. Human fluke infection in Glen Canyon at AD 1250. Am J Phys Anthropol **41,** 115-117.

Ortega, Y.R., and Bonavia, D. 2003. *Cryptosporidium, Giardia,* and *Cyclospora* in ancient Peruvians. J Parasitol **89,** 635-636.

Poinar, H.N., Kuch, M., Sobolik, K.D., Barnes, I., Stankiewicz, A.B., Kuder, T., Spaulding, W.G., Bryant, V.M., Cooper, A., and Paabo, S. 2001. A molecular analysis of dietary diversity for three archaic Native Americans. Proc Natl Acad Sci U S A **98,** 4317-4322.

Reinhard, K., and Bryant, V.J. 2008. "Pathoecology and the Future of Coprolite Studies in Bioarchaeology," in Reanalysis and reinterpretation in southwestern bioarcheology, ed. Stodder A.W.M.. Arizona State University Press: Tempe,199-216

Reinhard, K., Fink, T.M., and Skiles, J. 2003. A case of megacolon in Rio Grande valley as a possible case of Chagas disease. Mem Inst Oswaldo Cruz 98 Suppl 1, 165-172.

Reinhard, K., and Urban, O. 2003. Diagnosing ancient diphyllobothriasis from Chinchorro mummies. Mem Inst Oswaldo Cruz **98** Suppl 1, 191-193.

Reinhard, K.J., Ambler, J.R., and Szuter, C.R. 2007. Hunter-gatherer use of small animal food resources: coprolite evidence. International Journal of Osteoarchaeology **17,** 416-428.

Riley, T. 2008. Diet and seasonality in the Lower Pecos: evaluating coprolite data sets with cluster analysis. Journal of Archaeological Science **35,** 2726-2741.

Rocha, G.C., Harter-Lailheugue, S., Le Bailly, M., Araujo, A., Ferreira, L.F., Da Serra-Freire, N.M., and Bouchet, F. 2006. Paleoparasitological remains

revealed by seven historic contexts from "Place d'Armes", Namur, Belgium. Mem Inst Oswaldo Cruz **101** Suppl 2**,** 43-52.

Rocha, G.C.D., and Serra-Freire, N.M. 2009. Paleoparasitology at "Place d'Armes", Namur, Belgium: a biostatistics analysis of trichurid eggs between the Old and New World. Rev Bras Parasitol Vet **18,** 70-74.

**Rollo, F., Ermini, L., Luciani, S., Marota, I., and Olivieri, C.** 2006. Studies on the preservation of the intestinal microbiota's DNA in human mummies from cold environments. Med Secoli **18,** 725-740.

**Santiago-Rodriguez, T.M., Narganes-Storde, Y.M., Chanlatte, L., Crespo-Torres, E., Toranzos, G.A., Jimenez-Flores, R., Hamrick, A., and Cano, R.J.** 2013. Microbial communities in pre-columbian coprolites. PLoS One **8,** e65191.

**Santoro, C., Vinton, S.D., and Reinhard, K.J.** 2003. Inca expansion and parasitism in the lluta valley: preliminary data. Mem Inst Oswaldo Cruz **98** Suppl 1**,** 161-163.

**Shillito, L.M., Almond, M.J., Wicks, K., Marshall, L.J., and Matthews, W.** 2009. The use of FT-IR as a screening technique for organic residue analysis of archaeological samples. Spectrochim Acta A Mol Biomol Spectrosc **72,** 120-125.

**Shin, D.H., Chai, J.Y., Park, E.A., Lee, W., Lee, H., Lee, J.S., Choi, Y.M., Koh, B.J., Park, J.B., Oh, C.S., Bok, G.D., Kim, W.L., Lee, E., Lee, E.J., and Seo, M.** 2009a. Finding ancient parasite larvae in a sample from a male living in late 17th century Korea. J Parasitol **95,** 768-771.

**Shin, D.H., Lim, D.S., Choi, K.J., Oh, C.S., Kim, M.J., Lee, I.S., Kim, S.B., Shin, J.E., Bok, G.D., Chai, J.Y., and Seo, M.** 2009b. Scanning electron microscope study of ancient parasite eggs recovered from Korean mummies of the Joseon Dynasty. J Parasitol **95,** 137-145.

**Shin, D.H., Oh, C.S., Chai, J.Y., Lee, H.J., and Seo, M.** 2011. *Enterobius vermicularis* eggs discovered in coprolites from a medieval Korean mummy. Korean J Parasitol **49,** 323-326.

**Sianto, L., Reinhard, K.J., Chame, M., Chaves, S., Mendonca, S., Goncalves, M.L., Fernandes, A., Ferreira, L.F., and Araujo, A.** 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 **91,** 972-975.

Sianto, L., Teixeira-Santos, I., Chame, M., Chaves, S.M., Souza, S.M., Ferreira, L.F., Reinhard, K., and Araujo, A. 2012. Eating lizards: a millenary habit evidenced by Paleoparasitology. BMC Res Notes **5,** 586.

Sobolik, K.D., Gremillion, K.J., Whitten, P.L., and Watson, P.J. 1996. Sex determination of prehistoric human paleofeces. Am J Phys Anthropol **101,** 283-290.

Sutton, M.Q., and Reinhard, K.J. 1995. Cluster analysis of the coprolites from antelope house: Implications for anasazi diet and cuisine. Journal of Archaeological Science **22,** 741-750.

Taglioretti, V., Fugassa, M.H., Beltrame, M.O., and Sardella, N.H. 2013. Biometric identification of capillariid eggs from archaeological sites in Patagonia. J Helminthol**,** 1-7.

Tito, R.Y., Belknap, S.L., 3rd, Sobolik, K.D., Ingraham, R.C., Cleeland, L.M., and Lewis, C.M., Jr. 2011. Brief communication: DNA from early Holocene American dog. Am J Phys Anthropol **145,** 653-657.

Tito, R.Y., Knights, D., Metcalf, J., Obregon-Tito, A.J., Cleeland, L., Najar, F., Roe, B., Reinhard, K., Sobolik, K., Belknap, S., Foster, M., Spicer, P., Knight, R., and Lewis, C.M., Jr. 2012. Insights from characterizing extinct human gut microbiomes. PLoS One **7,** e51146.

Tito, R.Y., Macmil, S., Wiley, G., Najar, F., Cleeland, L., Qu, C., Wang, P., Romagne, F., Leonard, S., Ruiz, A.J., Reinhard, K., Roe, B.A., and Lewis, C.M., Jr. 2008. Phylotyping and functional analysis of two ancient human microbiomes. PLoS One **3,** e3703.

Ubaldi, M., Luciani, S., Marota, I., Fornaciari, G., Cano, R.J., and Rollo, F. 1998. Sequence analysis of bacterial DNA in the colon of an Andean mummy. Am J Phys Anthropol **107,** 285-295.

**Vinton, S.D., Perry, L., Reinhard, K.J., Santoro, C.M., and Teixeira-Santos, I.** 2009. Impact of empire expansion on household diet: the Inka in Northern Chile's Atacama Desert. PLoS One **4,** e8069.

# Chapitre II

ARTICLE

## Analyse polyphasique de la flore microbienne associée à un coprolithe du Moyen Âge, découvert en Belgique

Sandra Appelt[1], Fabrice Armougom[1], Matthieu Le Bailly[2], Catherine Robert[1], Michel Drancourt[1*]

[1] Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, Inserm 1095, 13005 Marseille, France. [2] Franche-Comté University, CNRS UMR 6249 Chrono-Environment, 25 030 Besançon, France.

[*]Corresponding author. Michel.Drancourt@medicine.univ-mrs.fr

**Avant-propos**

Les études paléomicrobiologiques ont permis d'approfondir la connaissance de la flore intestinale de populations anciennes, par détection de séquences amplifiées par PCR et séquençage à haut-débit sur des coprolithes et des échantillons de colon. Ces études ont porté sur 14 coprolithes humains et un échantillon de colon provenant de six sites archéologiques disséminés sur le continent Américain et deux échantillons de colon de deux sites archéologiques en Europe (Cano et al. 2000; Luciani et al. 2006; Rollo et al. 2006; Santiago-Rodriguez et al. 2013; Tito et al. 2012; Tito et al. 2008; Ubaldi et al. 1998). Ces études ont par exemple montré que la flore intestinale humaine ancienne présente un certain degré de similarité avec la flore intestinale des populations rurales modernes (Tito et al. 2012). Toutefois, les résultats sont encore peu nombreux tant ce domaine en est à ses balbutiements, et pour le moment peu comparables les uns aux autres en raison des différentes techniques employées (Cano et al. 2000; Luciani et al. 2006; Rollo et al. 2006; Santiago-Rodriguez et al. 2013; Tito et al. 2012; Tito et al. 2008; Ubaldi et al. 1998).

Afin d'enrichir la connaissance de la flore intestinale ancienne et des pathogènes associés, un coprolithe d'époque médiévale excavé à Namur a été étudié par une approche polyphasique combinant différentes techniques pour offrir une analyse détaillée en évitant toute contamination. Furent ainsi combinées des observations microscopiques, la culture bactérienne, la métagénomique et des amplifications par PCR, avec pour objectif l'obtention d'une vue synoptique de la flore associée au coprolithe découvert à Namur.

## Polyphasic analysis of a Middle Ages coprolite microbiota, Belgium
**Running title:** Middle Ages European coprolite flora

Sandra Appelt[1], Fabrice Armougom[1], Matthieu Le Bailly[2], Catherine Robert[1], Michel Drancourt[1*]

[1] Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, Inserm 1095, 13005 Marseille, France. [2] Franche-Comté University, CNRS UMR 6249 Chrono-Environment, 25 030 Besançon, France.

*Corresponding author. Michel.Drancourt@medicine.univ-mrs.fr

### Abstract

Paleomicrobiological investigations of a 14th century coprolite found inside a barrel in Namur, Belgium were done using microscopy, culture and metagenomics. Results were confirmed by *ad hoc* PCR – sequencing. Investigations yielded evidence for environmental flora preserved inside, indicated by microscopic observation of amoebal cysts, plant fibers, seeds, pollens and mold remains. Seventeen bacterial species were cultured, mixing environmental and gut microbiota organisms. Metagenomic analyses yielded 107,470 reads 31.9% of all reads were similar to known sequences stored in from public databases comprising 98.98% bacterial, 0.52% eukaryotic, 0.44% archaeal and 0.06% viral reads. Most abundant bacterial phyla were *Proteobacteria*, *Gemmatimonades*, *Actinobacteria* and *Bacteroides*. The 16S rDNA gene dataset yielded 132,000 trimmed reads and 673 Operational Taxonomic Units. Most abundant bacterial phyla observed in the 16S rDNA gene dataset belonged to *Proteobacteria*, *Firmicutes*, *Actinobacteria* and *Chlamydia*. The Namur coprolite yielded typical gut microbiota inhabitants, intestinal parasites *Trichuris* and *Ascaris* and systemic pathogens *Bartonella* and *Bordetella*. This study adds knowledge to gut microbiota in medieval times, and indicates that coprolites can be used as a source to detect pathogens responsible for past, systemic infections in addition to intestinal pathogens.

**Introduction**

Human paleomicrobiology, the quest for microbes in ancient specimens derived from humans, mainly relied on the investigations of old bone and dental pulp specimens [1]. Such investigations characterized past pathogens but did not provide data of ancient microbiota. In complement, investigating fossilized fecal material, i.e. coprolites, previously helped to gain knowledge on ancient human gut microbiota and intestinal parasites [2-6]. Currently, less than 20 coprolites and ancient colon content samples collected from six American and two European archeological sites have been investigated using large-scale sequencing and PCR-based analyses. These investigations yielded data about ancient gut microbiota, indicating that parts of the digestive flora were preserved in such specimens [2,4,5,7-10]. Moreover, these studies enabled to compare dietary habits of ancient populations and their impact on human gut flora composition [4,5,10].

In the present study, a further coprolite from a Middle Ages European site was investigated using a polyphasic approach in order to expand knowledge about gut microbiota in ancient Europe. Microscopic observations, culture and metagenomics (whole genome shotgun sequencing and 16S rDNA amplicon sequencing) were used to characterize the microbiota associated with the coprolite and to identify potential pathogens. Then, *ad hoc* suicide PCR amplifications were used for confirmation [11].

**Results**

**Microscopic observations**

In 1996, the exploration of an archeological Middle Ages site in Namur, Belgium yielded a closed barrel, such as those commonly used at that time as pits or latrines. The

barrel was located at a depth of 3.80 m beneath the modern soil level and contained a 121.4 g, dark-brown, well preserved coprolite specimen. The coprolite was further transferred in the laboratory for paleomicrobiology investigations (**Figure S1**). After aseptically peeling of its external portion, the inner portion of the coprolite was re-suspended in sterile Page's amoeba saline medium (PAS) and microscopic observations revealed the presence of several eggs. Thick-shelled and barrel-shaped eggs, with polar 'plugs' at the ends, 40-60 μm in length and 20-26 μm diameter, corresponded to the phenotypic description of *Trichuris* spp. eggs [12]. More precisely, broad eggs compatible with the pig-infecting *T. suis* species and thinner eggs compatible with human-infecting *T. trichiura* were observed (**Figure 1A** and **B**) [13]. Thick-shelled and brown eggs, corresponding to the description of *Ascaris* spp. eggs were also observed [12]. Among them, unfertilized elongated eggs (60 μm in length), fertilized round eggs (diameter between 40 and 50 μm), as well as eggs with embryos inside (**Figure 1C**) were found in the Namur coprolite. Microscopic observations also revealed the presence of suspected *Taenia* spp. eggs, plant fibers, pollens and mold remains (**Figure 1D** and **1E**). Cysts, plant fibers and seeds stained red using Congo red (**Figure 1E-G**). The cysts measured 4.1 to 13.5 μm and matched with the description of amoeba cysts [14].

**Culture**

After ten-day incubation at 30°C in the presence of negative controls, small colonies were visible in the aerobic and anaerobic layers of R2A and Schaedler brothes. Additionally, a tinny film was observed on the surface of the R2A solid medium. After 5-7-day subculture, matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry and 16S rDNA gene sequencing identified 10 different

bacterial species including *Paenibacillus macerans, Bacillus joetgali*, *Staphylococcus pasteuri, Staphylococcus epidermidis, Staphylococcus cohnii*, *Clostridium magnum*, *Micrococcus luteus*, *Bacillus horti*, *Pseudomonas geniculata* and *Stenotrophomonas maltophilia*. In addition, culturing the specimen in anaerobic and aerobic blood culture bottles in the presence of negative controls yielded one *Rhodanobacter* sp. organism, one *Paenibacillus* sp. organism, *Paenibacillus macerans*, *Paenibacillus thiaminolyticus*, *Paenibacillus ehimensis*, *Staphylococcus arlettae*, *Propionibacterium acnes* and *Enterobacter cloacae*.

**Metagenomics**

Acridine orange staining disclosed the presence of DNA in the coprolite (**Figure 2**), suggesting that molecular biological tools can be further applied to this specimen [10]. Accordingly, whole-genome shotgun sequencing (WGS) yielded a total of 37.5 millions base pairs and 107,470 reads, with an average sequence length of 375 bp and a GC content between 65 and 70% (MG-RAST accession number 4479942.3). Taxonomic assignment of the reads was performed using a BLASTX comparison with the National Center for Biotechnology Information (NCBI) database, with stringent parameters as previously described [15,16]. A significant similarity to known sequences was obtained for 31.9% of reads comprising 98.98% bacterial, 0.52% eukaryotic, 0.44% archaeal and 0.06% of viral reads. The most abundant bacterial phyla were *Proteobacteria* (58.12%), *Gemmatimonades* (15.18%), *Actinobacteria* (6.96%) and *Bacteroides* (5.10%) (**Figure 3**). More precisely, the WGS dataset yielded *Gemmatimonas*, *Rhizobium*, *Streptomyces* and *Burkholderia* known as environmental bacteria; as well as *Corynebacterium, Cytophaga, Enterobacter, Prevotella*, *Ruminococcus*, *Aeromonas*, *Escherichia*,

*Lactobacillus* and *Bacteroides* known as members of mammal gut microbiota (**Table S1**). In particular, contig reconstruction and annotation identified contigs belonging to human gut *Bacteroides* species, including *Bacteroides finegoldii*, *Bacteroides vulgatus*, *Bacteroides coprocola* along with *Bacteroides coprosuis* belonging to pig gut microbiota (**Table S2A**).

Some metagenomic reads were assigned to potential pathogenic bacteria (**Figure 4**, **Tables S1** and **S2B**). Reads and contigs of amoeba-resistant bacteria *Actinobacteria* spp., *Pseudomonas* spp., *Parachlamydia acanthamoebae*, *Legionella drancourtii* and *Legionella pneumophila* were found [17]. Moreover, contig annotations yielded *Burkholderia gladioli*, *Granulibacter bethesdensis*, *Leptospira borgpetersenii*, *Coxiella burnetii* and *Mycobacterium abscessus*. Contigs respectively encoding a hypothetical protein and an initiation factor 3 of *Brucella abortus* were also found, as well as reads and a contig matching with *Mycobacterium tuberculosis*. The latter contig encodes a possible toxin VapC46 protein (**Figure 4, Tables S1** and **S2B**). Two contigs assigned to *Clostridium botulinum* encode a perosamine synthetase and a methionyl-tRNA formyltransferase. Among *Bordetella* spp. sequences found in the metagenomic WGS dataset (**Figure 4**), contig reconstruction and annotation identified contigs encoding a hypothetical protein of *Bordetella parapertussis* and a putative hydrolase of *Bordetella bronchiseptica* (**Tables S1**, **S2**). For the hydrolase, phylogenetic analyses confirmed the BLAST based annotation (**Figure S2B**). The cultured *P. macerans, M. luteus, S. maltophilia* and *E. cloacae* were also found in the metagenomic WGS dataset.

In the WGS dataset, 0.09% of reads identified bacterial 16S rDNAs. For further analysis, the V6 region was amplified by PCR. This 16S rDNA gene dataset

yielded 132,000 trimmed reads. The Bayesian microbial source-tracking approach was performed to compare the 16S rDNA dataset associated with the Namur coprolite to those of modern stool samples, previously published coprolites, compost and soil as previously described [4]. The results indicated that the mixture of taxa associated to the Namur coprolite had no significant matches with any of the different sources used for comparison (**Figure S3**). A total of 673 Operational Taxonomic Units (OTUs) were assigned to the 16S rDNA dataset (MG-RAST accession number 4480463.3). Comparisons to the laboratory gut microbiota database including all 16S rDNA sequences generated by the 454FLX titanium platform showed that OTUs associated with the Namur coprolite did not clusterize with sequences previously amplified in the laboratory. The most abundant identified phyla were *Proteobacteria* (85.1%), *Firmicutes* (7.4%), *Actinobacteria* (2.8%) and *Chlamydia* (1.8%) (**Figure 3**). Furthermore, *Bartonella* species were detected with a 16S rDNA sequence identity of 98.7% (**Figure 4**); phylogenetic analyses indicated that these 16S rDNA amplicons were most closely related to *Bartonella henselae, Bartonella koehlerae* and *Bartonella quintana* (**Table S3** and **Figure S4**).

**Pathogen-specific PCR assays**

Mechanical lysis of the specimen reduced PCR inhibition as measured by PCR amplification of an internal nucleotide control. Further, microorganisms detected by microscopy or metagenomic were tentatively amplified by using species-specific PCR amplifications. Sequencing a 120-bp band generated using *Ascaris* spp.– specific primer pair [18] yielded 98% sequence identity with *Ascaris* sp. DHS-2010a cytochrome b gene sequence, similar to those previously reported in a human remain [19] (GenBank

Accession No. GU339224.1). The amplicon generated by an *Acanthamoeba* spp. – specific PCR showed a 99% sequence similarity with *Acanthamoeba castellanii* (GenBank Accession No. JF437606.1) [20]. *Bartonella* species, detected in the 16S rDNA dataset, were additionally amplified by PCR. The amplicon of the ribosomal RNA operon *rrsC* yielded 95% of sequence identity with *B. henselae* and 94% with *B. quintana* (GenBank Accession No. BX897699.1 and BX897700.1), respectively. *Bordetella* species, detected in the WGS dataset, were also amplified in the PCR assays (**Figure 4**). A 189-bp long fragment of the RNA polymerase C gene was generated, exhibiting 99% of sequence similarity with *B. bronchiseptica* (GenBank Accession No. HE965806.1, BX640437.1, HE965807.1), *B. parapertussis* (GenBank Accession No. HE965803.1, BX640423.1) and *Achromobacter xylosoxidans* (GenBank Accession No. CP002287.1). The obtained amplicon differed in one position (T → C transition) from the *B. parapertussis* and *B. bronchiseptica* reference strains. In two positions, the amplicon differed from *B. bronchiseptica* MO149 strain (T → C transition) as from the reference sequence of *A. xylosoxidans* (T → C transition and G → C mutation) (**Figure S5**) respectively.

**Synopsis of identified microorganisms**

The identified microorganisms were categorized into two groups. The first group comprises microorganisms for which the identification was confirmed by at least two independent methods among microscopy, culture, metagenomic and the PCR assay. This group includes *Ascaris* spp. and *Acanthamoeba* spp. identified by microscopy and PCR assay; *P. macerans, M. luteus, S. maltophilia* and *E. cloacae* found by culture and metagenomic; and *Bartonella* spp. – related to *B. henselae* and *B. quintana* – and

*Bordetella* spp. found in metagenomic WGS or 16S rDNA amplicon datasets and the PCR assay (**Figure 4**). The second group comprises of microorganisms identified only *in silico* by contig annotations (**Figure 4**). This group includes *B. abortus*, *C. botulinum, C. burnetii¸ G. bethesdensis*, *M. tuberculosis, M. absessus*, *B. gladioli*, *L. drancourtii*, *L. pneumophila*, *L. borgpetersenii*, *P. acanthamoebae*, *B. finegoldii, B. vulgatus*, *B. coprocola* and *B. coprosuis* (**Figure S2**).

**Discussion**

Paleomicrobiological investigations of a total of fourteen human coprolites and one colon sample have been reported from six different American archeological sites [2,4,5,7,10]. As for Europe, only two colon content specimens from two different sites, respectively dated to 3,350-3,100 BC and 1918 AD and no coprolite have been analyzed [8,9]. In the present study, a coprolite collected from a medieval site in Belgium was investigated using a polyphasic approach.

This coprolite was recovered from a sealed barrel which was still intact at the time of its discovery, and only the internal portion of the coprolite was investigated. Current recommendations for paleomicrobiology and paleoparasitology studies were strictly enforced in order to minimize in-laboratory contaminations [1,21-24]. No positive control was used and negative controls incorporated during all the experimentations remained negative. Nevertheless, the data herein reported – the presence of amoebal cysts, plant fibers, seeds, pollens and mold remains – indicate that the coprolite obviously contained environmental flora as previously reported in other investigations [4,5,10,25]. Accordingly, source tracking of the 16S rDNA amplicon dataset yielded no hits, as previously also observed for two pre-Columbian American coprolites [4]. More

coprolites need to be investigated before source tracking can be informative. Furthermore, the observation of pig- and human-infecting *Trichuris* spp. eggs in the coprolite indicates a possible common use of the barrel for animal and human feces. This point was supported by the results obtained in the metagenomic datasets, with bacterial sequences assigned to species found in modern gut microbiota of humans and pigs. Nevertheless, parts of the Namur coprolite correspond to human gut microbiota. Gut bacteria phyla – *Alpha-*, *Beta-* and *Gammaproteobacteria* and *Bacteroides* – were found herein as described for ancient coprolites and colon contents from American and European archeological sites [4,5,9,10]. Furthermore, *Corynebacterium*, *Enterobacter* and *Prevotella* are common inhabitants of the modern human digestive tract, further detected in the Namur coprolite [4,5,9,10,26]. Moreover, intestinal human parasites including *Trichuris* and *Ascaris* were also identified. The sequence obtained for the intestinal parasite *Ascaris* spp. supports the human origin of the Namur coprolite. Indeed, the same *Ascaris* spp. sequences were previously amplified from the remains of a human pelvic bone from a medieval Korean tomb [19].

On this basis, the Namur coprolite was further used as a source to detect past pathogens. Modern stool specimens diagnose systemic infections in humans and primates including malaria, rickettsiosis and tuberculosis [27-29]. Likewise, two pathogens *Haemophilus parainfluenzae* and *Clostridium botulinum* were already found in ancient fecal material [2,7] and *Neisseria, Yersinia, Shigella* and *Mycobacterium* sequences were identified in WGS datasets of pre-Columbian American coprolites [5]. To go further, it was herein attempted to ascertain the presence of potential pathogenic agents that might be associated to the analyzed specimen, *in silico* or by molecular tests. Bacterial pathogens

belonging *Bartonella* and *Bordetella* were detected here for the first time in a coprolite sample. *Bartonella* species are responsible for zoonoses [30] and *Bordetella* species cause respiratory tract infections [31].

**Conclusion**

Using a polyphasic approach including culture-dependant and culture-independent techniques, several microorganisms were consistently identified in one coprolite dating back to the European Middle Ages. These microorganisms included both common members of the gut microbiota and systemic pathogens. Coprolites are a source of knowledge regarding both microbiota and pathogens circulating in ancient populations and could be investigated using techniques routinely used for the modern diagnosis in clinical microbiology.

**Materials and Methods**

**Contamination prevention**

After excavation, the coprolite was stored in a sterile forensic specimen bag. In 2006, the coprolite was sent to our laboratory, where it was handled only in a positive pressure room with isolated ventilation under strict aseptic conditions. Workbenches were stringently disinfected using absolute ethanol and UV-irradiation for at least 30 min. Non-disposable instruments were autoclaved. Reagents and chemicals were from new stocks aliquoted into sterile, single-use tubes and immediately discarded after use. The external portion of the coprolite was aseptically removed; only the internal portion of the coprolite was used. DNA -extraction, PCR and post-PCR experiments were performed in separate rooms in isolated work areas. Positive controls were strictly avoided.

Negative controls for DNA -extraction and PCR were used at a 1:4 control: specimen ratio [1,21-24].

**Microscopy**

A 350-mg sample taken from the interior portion of the coprolite specimen was rehydrated for three days in 5 mL of phosphate buffered saline (PAS;Biotechnologie Appliquée, Taden, France) *prior* to observation under a microscope. Congo-red staining was applied to stain the cellulosic material of the coprolite. Then, 500 mg samples were taken from the interior of the coprolite specimen and rehydrated for two days in 5 mL of PAS. The diluted samples were fixed with absolute methanol on a glass slide, coated with 1 mg/mL Congo-red solution (Microm, Francheville, France) and incubated for 1 h at room temperature. The slide was carefully rinsed with 1 M sodium hydroxide (Aldrich-Chemie GmbH, Steinheim, Germany) and air-dried. Microscopic observations were performed using a Leica DM2500 microscope (Leica Microsystems, Nanterre, France) at 40x and 100x magnifications. Pictures were taken using a Nikon digital sight DS-U1 camera with Lucia G software (Nikon Instruments, Champigny sur Marne, France). Measurements were collected using the ImageJ software (http://rsbweb.nih.gov/ij/). Further, 500 mg of the coprolite were rehydrated into 1 mL PBS (Biotechnologie Appliquée), fixed with absolute ethanol on glass slides and stained using acridine orange solution as previously described [10].

**DNA recovery from the coprolite**

Two independent extraction protocols were used to obtain ancient DNA of the widest possible length range and to remove pigments, inhibitors of molecular detection methods. One gram of the coprolite was solubilized overnight at 4°C in 1mL TE buffer

(ethylenediamine tetraacetic acid; buffered solution, Tris HCl 10mM, EDTA 1mM, pH 8). A 500-μL aliquot of the solution was used for total DNA extraction as previously described [32], except that incubations into TE and digestion buffers were shortened to 1 day. TE buffer without coprolite was used as a negative control. For the DNA extraction using the PowerSoilR DNA Isolation Kit (MoBio Laboratories, Inc., Carlsbad, USA) [5], 500 μL of solubilized coprolite specimen were also used. Incubation was extended to 24h/56°C in PowerSoil[R] bead tubes containing sodium dodecyl sulfate (SDS) and digestion buffer C1, under contentiously rotation followed by shaking in a Bio 101 FastPrep instrument (Qbiogene) at level 6.5 (full speed) for 95s. DNA extraction was performed according to the manufacturer's instructions. Extraction batches contained also negative controls composed of PowerSoil[R] bead tubes without coprolite. Total DNA extracts from both protocols above were pooled together in a 1:1 ratio.

**Culture**

Four different culture conditions were used. Initially, Schaedler (Neogen, Lansingen, Michigan) and R2A (Neogen) brothes with an additive of 1.5% of agarose were used as culture media. To create an energetic potential sterile screw capped tubes (BIO-RAD, Marnes-la-Coquette, France) were filled with 15 mL of growth media and boiled for 30 min and then incubated at 47°C for 10 min. A 1-g inner portion of the coprolite was removed under anaerobic conditions and solubilized in 2 mL of sterile PBS (Biotechnologie Appliquée). Then 200 μL of this suspension were injected into the screw capped tubes containing Schaedler agar tubes under strict anaerobic atmosphere. Further, 1 mL of this suspension was incubated at 70°C for 20 min and then injected under strict anaerobiosis into the screw capped tubes containing R2A brothes. All tubes

were incubated at 30°C for daily inspection. When visible growth was observed, the tubes were sliced with a sterile glass cutter under strict anaerobiosis and subcultured on Schaedler or R2A agar plates under various atmospheres. The appropriated atmospheres were created in sterile incubation bags using gas generating pouch systems (BD Gas PakTM EZ, Maryland, USA). During the entire procedure, two negative controls (broth with and without PBS, (Biotechnologie Appliquée)) were carried out. Additionally, BD BACTEC, Lytic/10 Anaerobic and Aerobic bottles enriched with 7 mL of defibrinated sheep blood (bioMérieux, Marcy l'Etoile, France) were used for culturing. 1-g of the inner portion of the coprolite was solubilized into 1 mL of PBS (Biotechnologie Appliquée) and 300 µL of such suspension were injected into the culture bottle. After 2-day incubation at 37°C, 100 µL of the anaerobic and aerobic culture liquid were serially diluted (D1:D10$^{-10}$) and 10 µL of each dilution were plated onto COS culture plates (bioMérieux) and incubated at 37°C under strict anaerobe, microaerophile and aerobe atmosphere created by the use of gas generating pouch systems. Culture bottles with 300 µL of sterile PBS were run in parallel as negative controls. Likewise, COS plates (bioMérieux) inoculated with 10 µL culture liquid form the negative bottle culture and a COS plates (bioMérieux) that were opened under the Biosafety cabinet level 2 during the whole time of manipulation, were used as negative controls during subculture. Colonies were identification by MALDI-TOF mass spectrometry and 16S rRNA gene sequence analysis when MALDI-TOF identification failed.

**High-throughput pyrosequencing of the 16S rDNA V6 region**

The V6 region was amplified using 454-adapter primers [33]. PCR was performed in a final volume of 50 µL containing 1 x PCR buffer, 2 µL of 25 mM MgCl$_2$, 200 µM of

each d'NTP, 1 μL of 10 pM of each primer (10 pM), 31.15 μL ddH2O, 1 unit of HotStar Taq Polymerase (Invitrogen, Villebon sur Yvette, France) and 57-112 ng of DNA-extract. The amplification was performed by incubating at 95°C for 15 min, followed by 31 cycles of denaturation at 95°C for 45 sec, annealing at 58°C for 45 sec, elongation at 72°C for 90 sec, followed by an final elongation at 72°C for 10 min in an ABI Thermocycler (Applied Biosystems Gene Amp PCR System 2700, Villbon sur Yvette, France). PCR products were purified using Ampure beads (AgentcourtR AMPureR XP, Beckman Coulter, USA) and checked using a BioAnalyzer (Agilent 2,100 Bioanalyzer, Agilent Technology, Lithuania) LabChip with a DNA chip 7,500 at 548 bp (Agilent DNA 7,500 Reagent). Quantification of PCR products was performed using a Tecan GENios fluorometer was perfomed using a Quant-iTTM PicoGreenR ds DNA Assay Kit (Invitrogen) following the Amplicon Library Preparation Method in a manual provided by Roche (Roche, Mannheim, Germany). The concentration of the library was 27.7 ng/μL, corresponding to 8.84 E x 10 molecules/μL. Clonal amplification was performed using a GS Titanium LV emPCR Kit (Lib-A) v2 using only DNA Capture Beads A (Roche). Sequencing was performed with a GS FLX Titanium XLR70 Sequencing Kit (Roche).

### *OTU-based analysis*

The 16S pyrosequencing data were processed using Mothur package 1.5 [34]. No ambiguous bases 'N' and only one mismatch were allowed in the read and primer sequences. The quality read trimming used a moving window of 50 bp and required that the average quality score over the region did not drop below 35. The trimmed reads were dereplicated and aligned using the Sylva bacteria reference alignment provided by

the Mothur (http://www.mothur.org/). The multiple sequence alignment was filtered using a minimum read length of 200 bp. In addition, a pre-clustering step [35] was performed before chimera detection using the Uchime tool in Mothur. A distance matrix was built based on multiple sequence alignment, and OTUs clustering was performed at a 97% sequence identity. The taxonomic classification from phylum to genus level of each representative OUT sequence was performed using the RDP classifier tool and the RDP training set 9-database (http://www.mothur.org/). The relative abundance of reads per phyla was deduced from this classification. To exclude laboratory contaminations, we built an in-house gut microbiota database. This gut microbiota database contained 5,500,000 16S rDNA gene sequences using all of the data generated by the 454FLX titanium platform of the URMITE laboratory. The database includes data from obese (SRX118214), Senegalese (SRX118212, SRX118213), HIV (SRX209782), anorexic (SRX209240), post- antibiotic treatment (SRX189054, SRX189053) and drug resistance tuberculosis (SRX204218) stool specimens. The OTUs identified in the Namur coprolite were clustered at 97% sequence identity with sequences in our gut microbiota database. The OTUs and corresponding sequences that did not match with any sequence of the database were evaluated.

*Phylotype-based analysis*

To detect potential pathogens in the 16S pyrosequencing data, we used an alternative method that consisted of binning reads according to their taxonomic classification using BLAST searches against the Ribosomal Database Project (RDP) database. First, the OTU-based quality read trimming filter was applied. However, no multiple sequence alignment and no distance matrix were performed to increase read length and therefore,

improve the ability of assignment at the species level. Two databases were created using selected criteria from the Hierarchy Browser of the RDP 16S rRNA database, release 10 (http://rdp.cme.msu.edu/). A "Type" database was built with sequences labeled "Type strains", "Isolates" and size "length > 1,200 bp" with "good quality". A "Non-type" database was built using "Non-type strains", "Isolates", size "length > 1,200 bp" and "good quality". The two databases were formatted using Taxcollector [36]. The species level was defined with a minimum sequence identity of 98.7% [37] with the best BLAST hit from the "Type" database. The multiple best BLAST hit cases were checked for the most representative species (>50% of the multiple best BLAST hits). A second BLAST round was performed from the remaining reads with the same cutoffs but using the "Non-type" database. A total of 121 reads were assigned to the genus *Bartonella*, including 53 reads that were significantly assigned to *B. quintana* and 47 reads to *B. henselae*. For some 16S rDNA amplicons phylogenetic trees were constructed. Multiple sequence alignment was performed using MUSCLE [38] and curated by Gblocks [39]. The phylogenetic tree was built using the PhyML algorithm [40] with a bootstrap of 100 and the nucleotide substitution model HKY85 [41]. These tasks were all performed using the pipeline www.phylogenie.fr [42]. The phylogenetic trees were visualized using trees DrawTree [43].

*Source-tracking analysis*

To determine the potential sources of bacteria contained in the coprolite sample (defined as sink) we used the SourceTracker software package [44] as previously reported [45]. The Source Tracker sofware uses a Bayesian model that estimate the different source proportion found in a community sample (sink). The different sources examined (16S

rRNA datasets) correspond to 88 soil samples [46], 602 multiple human body sites (skin, nasal, tongues, urine) including 45 U.S. gut samples [47], 20 infant guts [48], 3 Polynesian guts from our laboratory, 60 mammal guts [49], eight different coprolites [4] and one compost [50]. If the community tested corresponds to a mixture of taxa that do not match with any of the source environments used, that portion of the community is classified as "unknown". The 16S rRNA reads were processed for all source and sink samples using the quantitative insights into microbial ecology 1.7.0 release (QIIME) [51]. High-quality read sequences (quality score >30, exact match to primer, and containing no ambiguous characters) were trimmed from the initial dataset. When compared datasets from non-overlapping amplicons the taxonomy binning with the closed-reference OTU picking strategy and using the Greengene gg_otus-12_10 release were performed. Then, all the taxonomy tables were converted and merged, for all the mapping files of the various source/sink datasets for SourceTracker analysis with the defaults parameters (iterations for Gibbs sampling =100, rarefaction depth =1000, alpha1=1e$^{-3}$, alpha2=1e$^{-1}$). The modification of the Source tracker parameters did not changed the results obtained for the analysis of the Namur coprolite. To control the analysis a soil sample [46] and a previously investigated coprolite ZA04 were used as positive controls [4](**Figure S3**). The coprolite specimen matched to the Sources of primate/mammalian gut and Burkina Faso children gut.

**High-throughput metagenomics**

The shotgun strategy was chosen for high-throughput pyrosequencing on a 454 Life Sciences Genome sequencer FLX instrument using titanium chemistry (Genome Sequencer RLX, Roche). Sequencing was performed using eight regions of the

PicoTiterPlate. The concentration of extracted DNA was measured with the QuAnt_IT Picogreen Kit (Invitrogen) on a Tecan Fluorometer (GENios) at a concentration of 28.2 ng /µL. A total of 500 ng of DNA was nebulized. The library was constructed according to the 454-titanium shotgun protocol and the manufacturer's instructions. DNA fragmentation was visualized using the BioAnalyzer 2,100 on a LabChip with high sensitivity and an optimal size of 872 bp. The DNA stock was measured on a TBS fluorometer at 8.776 E+08 molecules/ µL. The library was clonally amplified with 3cpb in 3 emPCR reactions using the GS Titanium SV emPCR Kit (Lib-L) version 2.The titration yield was 12.31%. In total 340,000 beads per project and per region were loaded onto the GS Titanium PicoTiterPlate Kit 70 x 75 which corresponded for this project with 383 µL of the clonal amplification and sequenced with the GS Titanium Sequencing Kit XLR70. The run was performed overnight, and then analyzed using the cluster. Using the Camera 2 [52] QC-Filter and the 454 duplicate clustering tool, reads that were low quality (average score < 19) or that were < 60 bp and identical duplicates (default sequence identity = 0.96) artificially produced by titanium technology, were deleted. Reads were blasted against the NCBI non-redundant protein database using a translated nucleotide query (BLASTX). The best BLAST-hits with $\geq$ 50% identity, $\geq$ 50 score and E-values < $1e^{-05}$ [15,16] were retained. Reads were assembled reads into contigs using a GS De Novo Assembler (Roche) with the following parameters: minimum overlap length of 35 bp, minimum identity of 98%. ORF searching was performed using Prodigal, and ORFs were blasted against the NCBI non-redundant database (BLASTP, E<$1e^{-05}$). When possible, phylogenetic trees of the ORFs of contigs, encoding proteins that were associated to potential bacterial pathogens, were built.

Regions that were homologous to the translated ORFs were searched using BLASTP against the non-redundant NCBI database. For multiple alignments of the sequences MUSCLE [38] was used and then curated by Gblocks [39]. The phylogenetic tree was built using the PhyML algorithm [40] with a bootstrap of 100 and the protein substitution model WAG. These tasks were all performed using the pipeline www.phylogenie.fr [42]. For the visualization of the phylogenetic trees DrawTree [43] was used. Principal coordinates analysis was performed using the correlated tool on the MG-RAST [54] metagenomic analysis server, to evaluate similarity between coprolite, soil and modern feces samples at the metabolic taxonomic level. Metabolic classifications were generated by a BLAST search against the SEED database using an E-value < 1e -05, a minimum identity cutoff of 80% and a minimum alignment length cutoff of 20 bp. The data were normalized to values between 0 and 1, distances were calculated using the Bray-Curtis method. The results of the analysis are represented in the **Figure S6**.

**References**

1. Drancourt M, Raoult D (2005) Palaeomicrobiology: current issues and perspectives. Nat Rev Microbiol 3: 23-35.

2. Ubaldi M, Luciani S, Marota I, Fornaciari G, Cano RJ, et al. (1998) Sequence analysis of bacterial DNA in the colon of an Andean mummy. Am J Phys Anthropol 107: 285-295.

3. Dittmar K (2009) Old parasites for a new world: the future of paleoparasitological research. a review. J Parasitol 95: 365-371.

4. Tito RY, Knights D, Metcalf J, Obregon-Tito AJ, Cleeland L, et al. (2012) Insights from characterizing extinct human gut microbiomes. PLoS One 7: e51146.

5. Tito RY, Macmil S, Wiley G, Najar F, Cleeland L, et al. (2008) Phylotyping and functional analysis of two ancient human microbiomes. PLoS One 3: e3703.

6. 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.

7. Luciani S, Fornaciari G, Rickards O, Labarga CM, Rollo F (2006) Molecular characterization of a pre-Columbian mummy and in situ coprolite. Am J Phys Anthropol 129: 620-629.

8. Rollo F, Ermini L, Luciani S, Marota I, Olivieri C (2006) Studies on the preservation of the intestinal microbiota's DNA in human mummies from cold environments. Med Secoli 18: 725-740.

9. Cano RJ, Tiefenbrunner F, Ubaldi M, Del Cueto C, Luciani S, et al. (2000) Sequence analysis of bacterial DNA in the colon and stomach of the Tyrolean Iceman. Am J Phys Anthropol 112: 297-309.

10. Santiago-Rodriguez TM, Narganes-Storde YM, Chanlatte L, Crespo-Torres E, Toranzos GA, et al. (2013) Microbial communities in pre-columbian coprolites. PLoS One 8: e65191.

11. Raoult D, Aboudharam G, Crubezy E, Larrouy G, Ludes B, et al. (2000) Molecular identification by "suicide PCR" of *Yersinia pestis* as the agent of medieval black death. Proc Natl Acad Sci U S A 97: 12800-12803.

12. Sheorey H, Biggs BA, Traynor P (2011) Nematodes. In: Versalovic J, Caroll KC, Funke G, Jorgensen JH, Landry ML, Warnock DW, Editors. Manual of Clinical Microbiology, 10<sup>th</sup> Edition: American Society of Microbiology, Washington DC. pp. 2122-2200.

13. Rocha GCd, Serra-Freire NM (2009) Paleoparasitology at "Place d'Armes", Namur, Belgium: a biostatistics analysis of trichurid eggs between the Old and New World. Rev Bras Parasitol Vet 18: 70-74.

14. Visvesvara GS, Moura H, Schuster FL (2007) Pathogenic and opportunistic free-living amoebae: *Acanthamoeba* spp., *Balamuthia mandrillaris*, *Naegleria fowleri*, and *Sappinia diploidea*. FEMS Immunol Med Microbiol 50: 1-26.

15. Kunin V, Copeland A, Lapidus A, Mavromatis K, Hugenholtz P (2008) A bioinformatician's guide to metagenomics. Microbiol Mol Biol Rev 72: 557-578.

16. Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, et al. (2009) A core gut microbiome in obese and lean twins. Nature 457: 480-484.

17. Greub G, Raoult D (2004) Microorganisms resistant to free-living amoebae. Clin Microbiol Rev 17: 413-433.

18. Loreille O, Roumat E, Verneau O, Bouchet F, Hanni C (2001) Ancient DNA from Ascaris: extraction amplification and sequences from eggs collected in coprolites. Int J Parasitol 31: 1101-1106.

19. Oh CS, Seo M, Lim NJ, Lee SJ, Lee EJ, et al. (2010) Paleoparasitological report on *Ascaris* aDNA from an ancient East Asian sample. Mem Inst Oswaldo Cruz 105: 225-228.

20. Schroeder JM, Booton GC, Hay J, Niszl IA, Seal DV, et al. (2001) Use of subgenic 18S ribosomal DNA PCR and sequencing for genus and genotype identification of *Acanthamoebae* from humans with keratitis and from sewage sludge. J Clin Microbiol 39: 1903-1911.

21. Cooper A, Poinar HN (2000) Ancient DNA: do it right or not at all. Science 289: 1139.

22. Hofreiter M, Serre D, Poinar HN, Kuch M, Paabo S (2001) Ancient DNA. Nat Rev Genet 2: 353-359.

23. Willerslev E, Cooper A (2005) Ancient DNA. Proc Biol Sci 272: 3-16.

24. Pääbo S, Poinar H, Serre D, Jaenicke-Despres V, Hebler J, et al. (2004) Genetic analyses from ancient DNA. Annu Rev Genet 38: 645-679.

25. Chaves SAdM, Reinhard KJ (2006) Critical analysis of coprolite evidence of medicinal plant use, Piauí, Brazil. Palaeogeography, Palaeoclimatology, Palaeoecology 237: 110-118.

26. Backhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JI (2005) Host-bacterial mutualism in the human intestine. Science 307: 1915-1920.

27. Keita AK, Socolovschi C, Ahuka-Mundeke S, Ratmanov P, Butel C, et al. (2013) Molecular evidence for the presence of *Rickettsia felis* in the feces of wild-living African apes. PLoS One 8: e54679.

28. Jirku M, Pomajbikova K, Petrzelkova KJ, Huzova Z, Modry D, et al. (2012) Detection of *Plasmodium* spp. in human feces. Emerg Infect Dis 18: 634-636.

29. El Khechine A, Henry M, Raoult D, Drancourt M (2009) Detection of *Mycobacterium tuberculosis* complex organisms in the stools of patients with pulmonary tuberculosis. Microbiology 155: 2384-2389.

30. Ricardog M, Kempf V, Chomel BB, Breitschyerdt E (2011) Bartonella. Versalovic J, Caroll KC, Funke G, Jorgensen JH, Landry ML, Warnock DW, Editors. Manual of Clinical Microbiology, 10th Edition: American Society of Microbiology, Washington DC. pp. 791-805.

31. Wirsing von König C, Riffelmann M, Coenye T (2011) Bordetella and Related Genera. In: Versalovic J, Caroll KC, Funke G, Jorgensen JH, Landry ML, Warnock DW, Editors. Manual of Clinical Microbiology, 10th Edition: American Society of Microbiology, Washington DC. pp 739-750.

32. Iniguez AM, Reinhard K, Carvalho Goncalves ML, Ferreira LF, Araujo A, et al. (2006) SL1 RNA gene recovery from *Enterobius vermicularis* ancient DNA in pre-Columbian human coprolites. Int J Parasitol 36: 1419-1425.

33. Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, et al. (2012) Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect 18: 1185-1193.

34. Schloss PD, Gevers D, Westcott SL (2011) Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. PLoS One 6: e27310.

35. Huse SM, Welch DM, Morrison HG, Sogin ML (2010) Ironing out the wrinkles in the rare biosphere through improved OTU clustering. Environ Microbiol 12: 1889-1898.

36. Giongo A, Davis-Richardson AG, Crabb DB, Triplett EW (2010) TaxCollector: Modifying Current 16S rRNA Databases for the Rapid Classification at Six Taxonomic Levels. Diversity 2: 1015-1025.

37. Schlaber R, Simmon KE, Fisher MA (2012) A systematic approach for discovering novel, clinically relevant bacteria. Emerg Infect Dis 18: 422-430.

38. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32: 1792-1797.

39. Talavera G, Castresana J (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol 56: 564-577.

40. Guindon S, Delsuc F, Dufayard JF, Gascuel O (2009) Estimating maximum likelihood phylogenies with PhyML. Methods Mol Biol 537: 113-137.

41. Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22: 160-174.

42. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, et al. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res 36: W465-469.

43. Chevenet F, Brun C, Banuls AL, Jacq B, Christen R (2006) TreeDyn: towards dynamic graphics and annotations for analyses of trees. BMC Bioinformatics 7: 439.

44. Knights D, Kuczynski J, Charlson ES, Zaneveld J, Mozer MC, et al. (2011) Bayesian community-wide culture-independent microbial source tracking. Nat Methods 8: 761-763.

45. Flores GE, Bates ST, Knights D, Lauber CL, Stombaugh J, et al. (2011) Microbial biogeography of public restroom surfaces. PLoS One 6: e28132.

46. Lauber CL, Hamady M, Knight R, Fierer N (2009) Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. Appl Environ Microbiol 75: 5111-5120.

47. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, et al. (2009) Bacterial community variation in human body habitats across space and time. Science 326: 1694-1697.

48. De Filippo C, Cavalieri D, Di Paola M, Ramazzotti M, Poullet JB, et al. (2010) Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. Proc Natl Acad Sci U S A 107: 14691-14696.

49. Muegge BD, Kuczynski J, Knights D, Clemente JC, Gonzalez A, et al. (2011) Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. Science 332: 970-974.

50. Watanabe K, Nagao N, Toda T, Kurosawa N (2008) Changes in bacterial communities accompanied by aggregation in a fed-batch composting reactor. Curr Microbiol 56: 458-467.

51. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, et al. (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7: 335-336.

52. Sun S, Chen J, Li W, Altintas I, Lin A, et al. (2011) Community cyberinfrastructure for Advanced Microbial Ecology Research and Analysis: the CAMERA resource. Nucleic Acids Res 39: D546-551.

53. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, et al. (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11: 119.

54. Glass EM, Wilkening J, Wilke A, Antonopoulos D, Meyer F (2010) Using the metagenomics RAST server (MG-RAST) for analyzing shotgun metagenomes. Cold Spring Harb Protoc 2010: pdb prot5368.

55. Greenberg DE, Ding L, Zelazny AM, Stock F, Wong A, et al. (2006) A novel bacterium associated with lymphadenitis in a patient with chronic granulomatous disease. PLoS Pathog 2: e28.

56. Greub G (2009) Parachlamydia acanthamoebae, an emerging agent of pneumonia. Clin Microbiol Infect 15: 18-28.

**Legends of Figures**

**Figure 1**. **Microscopic observation of unstained and stained Namur coprolite** (optical magnification: 100x). A) *Trichuris* spp. egg (ø 25.83 μm; length 43.65 μm); B) *Trichuris* spp. egg (ø 22.50 μm; length 41.10 μm); C) *Ascaris* spp. fertilized egg (ø 49.92 μm) and unfertilized egg (ø 43.16 μm; length 59.92 μm); D) pollen (ø 18.26 μm); E) suspected *Taenia* spp. egg (ø 14.65 μm); F) suspected *Acanthamoeba* spp. cyst; and G) seed remains. (The scale bar on the right indicates 20 μm).

**Figure 2. Nucleotide acid staining.** Nucleotide acids of the coprolite were stained using acridine orange (optical magnification: 40x, emission 525 nm).

**Figure 3. Composition of the Namur coprolite microbiota at the phylum level.** The dataset obtained from Whole-Genome Shotgun sequencing (WGS) and Operational Taxonomic Units (OTUs) assigned to the 16S rDNA amplicon were used to identify bacterial phyla. Only phyla comprising more than 0.3% of the datasets are shown.

**Figure 4. Pathogens associated with the Namur coprolite.** Only human and animal pathogens detected *in silico* or by at least two tests are shown. A positive test result is marked in green and negative tests are colored in red. Pathogenic microorganisms which were previously associated to ancient human coprolites or colon contents are marked by circles. * [55] ** [56]

**Figure 1**.

**Figure 2.**

Figure 3.

**Figure 4.**

**Polyphasic analysis of a Middle Ages coprolite microbiota, Belgium**

**Running title:** Middle Ages European coprolite flora

Sandra Appelt[1], Fabrice Armougom[1], Matthieu Le Bailly[2], Catherine Robert[1], Michel Drancourt[1*]

[1] Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, Inserm 1095, 13005 Marseille, France. [2] Franche-Comté University, CNRS UMR 6249 Chrono-Environment, 25 030 Besançon, France.

[*]Corresponding author. Michel.Drancourt@medicine.univ-mrs.fr

**SUPPLEMENTAL MATERIAL CONTENTS**

**SUPPLEMENTARY NOTE**

**SUPPLEMENTARY TABLES**

Supplementary Tables S1-S5

**SUPPLEMENTARY FIGURES**

Supplementary Figures S1-S6

**SUPPLEMENTARY REFERENCES**

**SUPPLEMENTARY NOTE**

**Specific PCR-amplifications and Sanger-sequencing**

Suicide PCR −amplification [1] was performed to reinforce the high-throughput pyrosequencing results and perform target-orientated searches for *Ascaris* sp., amoebae and viridae. For the molecular detection of *Ascaris* sp., a 145-bp region of the cytochrome b gene was amplified [2]. To detect *Acanthamoeba* spp., we targeted the *Acanthamoeba* spp. 18S rRNA gene [3]. In addition, seven primer pairs specific for the RNA-polymerase gene C of *Bordetella* sp., 16S rRNA gene of *Mycobacterium kumamotonense, Rickettsia* sp.*, Ehrlichia* sp.*, Brucella* sp., and *Bartonella* sp. were designed. The primers are summarized in Supplementary Table S5. Additionally an internal in-laboratory control was included, to check for PCR inhibitors in the DNA extract. PCR was performed in a 50-μL final volume containing 1x PCR buffer, 2 μL of 25 mM MgCl$_2$, 200 μM of each d'NTP, 1 μL of 10 pM of each primer, 31.15 μL of ddH$_2$O, 1 unit of HotStar Taq Polymerase (Invitrogen, Villebon sur Yvette, France) and 57-112 ng of DNA-extract. The steps of the PCR were an initial incubation at 95°C for 15 min; 36-40 cycles of denaturation for 95°C for 1 min, annealing for 45 sec at the corresponding primer annealing temperatures, and elongation at 72°C for 90 sec and a final elongation 72°C for 10 min; all of these steps were performed in a Gene Amp PCR System 2700 ABI Thermocycler (Applied Biosystems, Villbon sur Yvette, France). The PCR products were analyzed using a 2 % agarose gel (UltraPureTM agarose, Invitrogen, Villbon sur Yvette, France) and purified using the QIAquick PCR Purification Kit (QIAGEN, Courtaboeuf, France). For the F6-R6 and F8-R8 primer pairs, the PCR products were cloned into the pEGM-T-Easy Vector system I (Promega, Carbonnieres, France) and transferred *via* heat shock into competent JM109 *Escherichia coli* cells

according to the manufacturer's instructions. White colonies were screened on LB-Amp-XGal-IPTG agar plates using the following M13 PCR primers M13F 5'-GTAAAACGACGGCCAG-3' and M13R 5'-CAGGAAACAGCTATGAC-3'. All of the PCR and cloning products were sequenced in a final volume of 20 of μL (1 x sequencing buffer, 3.2 pM forward or reverse primer, 4 μL of BigDye Terminator V1.1 mix (Applied Biosystems), 7.4 μL of ddH$_2$O and 4 μL of PCR product) after purification using Sephadex Gel Filtration in the ABI PR ISM 3130xl genetic sequencer (Applied Biosystems, Villbon sur Yvette, France). The sequences were assembled using the ChromasPro software and compared with reference sequences of the GenBank database using NCBI BLAST searches.

**Quantitative real time PCR amplifications**

Quantitative real-time PCR (qPCR) assays targeting various genomic regions of twenty-two pathogens were performed to examine for their presence in the coprolite specimen. All of the targeted microorganisms and genes, including the primer pairs and probes, are listed in Supplementary Table S8. Serial ten-fold dilutions (from $10^{-1}$ to $10^{-3}$) of the DNA extract derived from the coprolite specimen were performed to dilute potential molecular detection inhibitors in the brown DNA supernatant. The primers and probes were used at final concentrations of 500 nM and 62.5 nM, respectively. For the qPCR mix, we used a commercial MasterMix of Eurogentec (Eurogentec, Angers, France) according to the manufacturer's instructions. The reactions were performed using a C1000TM Thermal cycler (CFX96TM Real-Time System, BIORAD, Marnes-la-Coquette, France). Negative controls consisting of the reaction mix without DNA were

added in a ratio of 1:3. To avoid contamination, the experiments were performed in working areas in which these systems had never been used. Overall, 30 different qPCR diagnostic systems were tested using various ten-fold dilutions of the coprolite DNA extract. Altogether, 30 different qPCR diagnostic systems were tested using various ten-fold dilutions of the coprolite DNA extract. All of the tested systems and all of the negative controls remained negative.

**Table S1. Typical gastro-intestinal, environmental and pathogenic bacteria assigned to the coprolite metagenome[#].**

| Gastro-intestinal Bacteria | | | Environmental Bacteria | | | Bacterial Pathogens | | | Opportunistic Pathogens | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Genus (Reads) | Species | Reads | Genus (Reads) | Species | Reads | Genus (Reads) | Species | Reads | Genus (Reads) | Species | Reads |
| Aeromonas (7) | *A. salmonicida* | 4 | Acidovorax (76) | *A. avenae* | 16 | Bartonella (9) | *B. henselae* | 1 | Achromobacter (114) | *A. xylosoxidans* | 80 |
|  | *A. veronii* | 3 |  | *A. citrulli* | 7 |  | *B. rochalimae* | 3 |  | *A. piechaudii* | 34 |
| Bacteroides (82) | *B. cellulosilyticus* | 6 |  | *A. delafieldii* | 28 |  | *Bartonella* sp. | 3 | Actinobacter (20) | *A. baumannii* | 4 |
|  | *B. fragilis* | 5 | Agrobacterium (199) | *A. radiobacter* | 93 | Bordetella (153) | *B. avium* | 15 | Aeromonas (9) | *A. veronii* | 3 |
|  | *B. intestinalis* | 5 |  | *A. tumefaciens* | 41 |  | *B. bronchiseptica* | 37 |  | *A. hydrphila* | 2 |
|  | *B. plebeius* | 5 |  | *A. vitis* | 38 |  | *B. parapertussis* | 16 | Bacillus (112) | *B. cereus* | 14 |
|  | *Bacteroides* sp. | 30 | Beutenbergia (10) | *B. cavernae* | 10 |  | *B. pertussis* | 14 |  | *B. thuringiensis* | 7 |
| Clostridium (45) | *C. botulinum* | 6 | Burkholderia (832) | *B. cenocepacia* | 36 | Brucella (164) | *B. abortus* | 9 | Burkholderia (832) | *B. gladioli* | 26 |
|  | *C. difficile* | 3 |  | *B. gladioli* | 26 |  | *B. melitensis* | 10 |  | *B. mallei* | 4 |
|  | *C. leptum* | 4 |  | *B. multivorans* | 34 |  | *B. neotomae* | 2 |  | *B. pseudomallei* | 41 |
| Collinsella (5) | *C. aerofaciens* | 4 |  | *B. pseudomallei* | 41 |  | *B. ovis* | 4 | Clostridium (45) | *C. botulinum* | 6 |
| Corynebacterium (25) | *C. aurimucosum* | 10 |  | *B. phytofirmans* | 34 |  | *B. suis* | 5 | Comamonas (20) | *C. testosteroni* | 19 |
| Eikenella (6) | *E. corrodens* | 3 |  | *B. ubonensis* | 35 | Chlamydia (7) | *C. pneumoniae* | 1 | Delftia (23) | *D. acidovorans* | 22 |
|  | *E. billigiae* | 3 |  | *B. graminis* | 46 |  | *C. trachomatis* | 3 | Escherichia (27) | *E. coli* | 21 |
| Enterobacter (15) | *E. cancerogenus* | 1 |  | *B. phymatum* | 59 |  | *C. psittaci* | 2 | Legionella (36) | *L. pneumophila* | 18 |
|  | *E. cloacae* | 9 |  | *B. xenovorans* | 72 | Coxiella (13) | *C. burnetii* | 13 |  | *L. drancourtii* | 7 |
|  | *Enterobacter* sp. | 5 | Chromobacterium (25) | *C. violaceum* | 25 | Granulibacter (45) | *G. bethesdensis* | 45 | Parachlamydia (31) | *P. acanthamoebae* | 31 |
| Escherichia (27) | *E. albertii* | 3 | Erythorbacter (170) | *E. litoralis* | 63 | Klebsiella (10) | *K. pneumoniae* | 4 | Pseudomonas (292) | *P. aeruginosa* | 47 |
|  | *E. coli* | 21 |  | *Erythorbacter* sp. | 107 | Leptospira (19) | *L. borgpetersenii* | 6 |  | *P. putida* | 29 |
|  | *E. fergusonii* | 3 | Gemmata (59) | *G. obscuriglobus* | 57 | Mycobacterium (113) | *M. intracellulare* | 9 |  | *P. stutzeri* | 37 |
| Eubacterium (6) | *E. limosum* | 4 | Gemmatimonas (5,370) | *G. aurantiaca* | 5,370 |  | *M. tuberculosis* | 9 | Rhodococcus (65) | *R. equi* | 21 |
| Flavobacterium (58) | *F. johnsoniae* | 28 | Laribacteria (16) | *L. hongkongensis* | 16 |  | *M. kansasii* | 9 |  | *R. erythropolis* | 8 |
| Lactobacillus (11) | *L. jensenii* | 3 | Legionella (36) | *L. drancourtii* | 7 | Salmonella (14) | *S. enterica* | 14 | Stenotrophomonas | *S. maltophilia* | 53 |
|  | *L. buchneri* | 2 |  | *L. pneumophila* | 18 | Vibrio (43) | *V. cholerae* | 7 | Micrococcus (16) | *M. luteus* | 16 |
|  | *L. fermentum* | 3 | Leptospirillum (10) | *L. ferrodiazotrophum* | 8 |  | *V. metschnikovii* | 5 |  |  |  |
| Parabacteroides ( 5) | *P. merdae* | 3 |  | *L. rubarum* | 7 | Yersinia (26) | *Y. pseudotuberculosis* | 3 |  |  |  |
| Prevotella (19) | *P. ruminicola* | 1 | Parachlamydia (31) | *P. acanthamoebae* | 31 |  | *Y. ruckeri* | 4 |  |  |  |
|  | *P. buccae* | 4 | Ralstonia (212) | *R. eutropha* | 110 |  | *Y. kirkensii* | 7 |  |  |  |
|  | *P. denticola* | 4 |  | *R. pickettii* | 19 |  |  |  |  |  |  |
|  | *P. disiens* | 4 |  | *R. solanacearum* | 69 |  |  |  |  |  |  |
| Roseburia (42) | *R. intestinalis* | 1 | Rhizobium (167) | *R. etli* | 144 |  |  |  |  |  |  |
|  | *Roseburia* sp. | 41 |  | *R. leguminosarum* | 123 |  |  |  |  |  |  |
| Ruminococcus (7) | *R. albus* | 2 | Streptomyces (449) | *S. bingchenggensis* | 26 |  |  |  |  |  |  |
|  | *Ruminococcus* sp. | 4 |  | *S. himastatinicus* | 37 |  |  |  |  |  |  |
| Serratia (14) | *S. odorifera* | 8 |  | *S. pristinaespiralis* | 19 |  |  |  |  |  |  |
|  | *S. proteamaculans* | 6 |  | *S. scabiei* | 27 |  |  |  |  |  |  |
| Streptococcus (8) | *S. agalactiae* | 3 |  | *S. violaceusniger* | 26 |  |  |  |  |  |  |

[#]Metagenomic reads were blasted against the NCBI protein database using a translated nucleotide query. Underrepresented taxonomic genera are not shown.

**Table S2. Number and size of contigs that were assigned to (A) *Bacteroides* spp. and (B) to bacterial pathogens associated to the coprolite sample#.**

| A)Conti ID | Contig length (bp) | Hit description | E-value | Hit accession ID | Percent ID |
|---|---|---|---|---|---|
| **contig00127°** | 555 | TonB-dependent receptor [***Bacteroides finegoldii***] | 3E-13 | WP_017142870.1 | 37% |
| **contig00570*** | 195 | ligand-gated channel protein [***Bacteroides coprosuis***] | 5 E-05 | WP_006744840.1 | 45% |
| **contig01043*** | 198 | GTPase HflX [***Bacteroides coprosuis***] | 1E-08 | WP_006744031.1 | 72% |
| **contig02059°** | 291 | collagen-binding protein [***Bacteroides vulgatus***] | 1E-17 | WP_005850005.1 | 65% |
| **contig01426°** | 576 | bacterial group 2 Ig-like protein [***Bacteroides coprocola CAG:162***] | 9E-17 | CDA71009.1 | 36% |

| B)Conti ID | Contig length (bp) | Hit description | E-value | Hit accession ID | Percent ID |
|---|---|---|---|---|---|
| **contig03375** | 504 | hypothetical protein BPP0674 [***Bordetella parapertussis 12822***] | 1e-32 | NP_883015.1 | 52% |
| **contig03491** | 933 | hydrolase [***Bordetella bronchiseptica 253***] | 2e-72 | YP_006966397.1 | 53% |
| **contig02527** | 503 | methyltransferase [***Coxiella burnetii RSA 493***] | 2e-65 | NP_819712.1 | 68% |
| **contig01914** | 1,568 | Possible toxin VapC46. Contains PIN domain [***Mycobacterium tuberculosis H37Rv***] | 8e-08 | NP_217901.1 | 42% |
| **contig03497** | 497 | putative transmembrane protein [***Mycobacterium abscessus***] | 2e-10 | WP_005102565.1 | 30% |
| **contig00341** | 529 | hypothetical protein [***Brucella abortus***] | 3e-62 | WP_006089712.1 | 80% |
| **contig00638** | 652 | Initiation factor 3 [***Brucella abortus S19***] and | 1e-24 | YP_001935939.1 | 79% |
| **contig01163** | 1,103 | hypothetical protein bgla_1g34580 [***Burkholderia gladioli BSR3***] | 5e-37 | YP_004362017.1 | 42% |
| **contig01000** | 1,019 | hypothetical protein BTI_2181 [***Burkholderia thailandensis MSMB121***] | 1e-78 | YP_007918661.1 | 70% |
| **contig02486** | 704 | paraquat-inducible protein B [***Granulibacter bethesdensis CGDNIH1***] | 3e-33 | YP_744011.1 | 49% |

| | | | | | |
|---|---|---|---|---|---|
| **contig01542** | 525 | cation efflux transporter [*Legionella pneumophila subsp. pneumophila*] | 6e-06 | YP_006506440.1 | 45% |
| **contig01760** | 833 | Fic/DOC family protein [*Leptospira borgpetersenii*] | 4e-75 | WP_002734514.1 | 52% |
| **contig02804** | 1,059 | perosamine synthetase [*Clostridium botulinum*] | 6e-22 | WP_003367053.1 | 70% |
| **contig03148** | 808 | methionyl-tRNA formyltransferase [*Clostridium botulinum* **A3 str. Loch Maree**] | 1e-34 | YP_001788030.1 | 36% |
| **contig01373** | 1,031 | 6-carboxy-5,6,7,8-tetrahydropterin synthase [*Vibrio cholerae*] | 2e-37 | WP_000980008.1 | 75% |
| **contig01064** | 537 | hypothetical protein [*Legionella drancourtii*] | 2e-27 | WP_006872680.1 | 40% |
| **contig00846** | 1,349 | CP4-6 prophage [*Yersinia kristensenii*] | 2e-124 | WP_004390432.1 | 63% |
| **contig01361** | 976 | DNA methylase N-4/N-6 domain protein [*Yersinia kristensenii*] | 3e-140 | WP_004390429.1 | 65% |
| **contig03337** | 847 | regulatory protein ada [*Parachlamydia acanthamoebae* **UV-7**] | 5e-68 | YP_004652478.1 | 64% |

[#]The contig identifier, its length (bp) and the annotation according to the best BLAST hit (BLASTX *versus* the non-redundant NCBI database, E-value<1e[-05]) are summarized. The E-value, the hit accession identifier and the percent of identity are also provided. Contigs of °human and *pig gut microbiota *Bacteroides* species.

**Table S3**. **Bacterial pathogens identified from the amplified 16S rRNA V6 region[#].**

| Genus | Species | Reads | Size (bp) | Reads > 230 bp |
|---|---|---|---|---|
| *Mycobacterium* | *M. kumamotonense* | 1 | 240 | 1 |
| *Bartonella* | *B. henselae* | 47 | 203-251 | 19 |
| | *B. quintana* | 53 | 270-375 | 53 |
| | *B. tribocorum* | 1 | 306 | 1 |

[#]The species level was defined with a minimum sequence identity of 98.7% using BLAST similarity searches against RDP databases.

**Table S4. Primers used to amplify DNA from intestinal parasites, bacterial pathogens and amoebae.**

| Desired specificity | Gene | Name | Sequence 5'→3' | Tm (C°) | Reference |
|---|---|---|---|---|---|
| *Ascaris* **sp.** | cytb | Asc1 | GTTAGGTTACCGTCTAGTAAGG | 53 | [2] |
|  |  | Asc2 | CACTCAAAAAGGCCAAAGCACC |  |  |
| *Acanthamoeba sp.* | 18S rRNA | IDP1 | TCTCACAAGCTGCTAGGGAGTCA | 60 | [3] |
|  |  | IDP2 | TCTCACAAGCTGCTAGGGAGTCA |  |  |
| *Bordetella* **sp.** | rpoC | F5 | ATGGCGAGGAAGTGCGCCAG | 60 | present study |
|  |  | R5 | CCGTCCGGCTTGGCCATCAG |  |  |
| *Bordetella* **sp.** | rpoC | F6 | GGCGAGCTCAAGGCCACCAG | 62 | present study |
|  |  | R6 | CAGGTCGGAGGTCGCGAAGC |  |  |
| *Bartonella* **sp.***; Brucella* **sp.** | rRNA_BAA | F7 | GCTGGCGCCCCTGCTTCAAA | 60 | present study |
|  |  | R7 | CCCCGCTGTCTCCAACGCAG |  |  |
| *Bartonella* **sp.***; Brucella* **sp.***;* | rRNA_rrsC | F8 | TGGCCTGCGATCTATGTTCT | 53 | present study |
|  |  | R8 | AGTCGTAACAAGGTAGCCGT |  |  |
| *Bartonella* **sp.***; Brucella* **sp.** *Rickettsia* **sp.***; Erlichia* **sp.** | BQ | F12 | ACAAAGGCTAGCGCCCCTGC | 60 | present study |
|  |  | R12 | TCCCCGTGAAGATGCGGGGT |  |  |
| *Bartonella henselae* | Spacer | S1 F | TTGCAAAGCAGGTGCTCTCC | 58 | diagnostical tool |
|  |  | S1 R | TAAGCGTGAGGTCGGAGGTT |  |  |
|  | Spacer | S3 F | CAATGGAGGCAACCGTTCTT | 49 | diagnostical tool |
|  |  | S3 R | GTGATATCGGGTACATTTTCAACTG |  |  |
|  | Spacer | S9F | CAACTTCACTGATTTCTGCGATAA | 47 | diagnostical tool |
|  |  | S9R | CGAGGAGTGGTTAATATGACAGCT |  |  |
| *Coxiella burnetii* | spacer | Cox 2F | CAACCCTGAATACCCAAGGA | 59 | diagnostical tool |
|  |  | Cox 2R | GAAGCTTCTGATAGGCGGGA |  |  |
|  | spacer | Cox 5F | CAGGAGCAAGCTTGAATGCG | 59 | diagnostical tool |
|  |  | Cox 5R | TGGTATGACAACCCGTCATG |  |  |

| *Legionella* sp. | rpoB | RL1 | GATGATATCGATCAYCTDGG | 55 | diagnostical tool |
|---|---|---|---|---|---|
| | | RL2 | TTCVGGCGTTTCAATNGGAC | | |
| *Mycobacterium* sp. | IS6110 | ISMtubF | CCTGCGAGCGTAGGCGTCGG | 68 | present study |
| | | ISMtubR | CTCGTCCAGCGCCGCTTCGG | | |
| | rpoB | MF | CGACCACTTCGGCAACCG | 60 | diagnostical tool |
| | | MR | TCGATCGGGCACATCCGG | | |
| *M. kumamotonense* | 16s rRNA | BF3 | GCGGGTTTTCTCGCAG | 50 | present study |
| | | BR3 | GCTCTTTACGCCCAGTAATTC | | |
| *M. tuberculosis* | MTC4 | MTC4F | ATGGGTTCGCCAGACGGCGAG | 60 | diagnostical tool |
| | | MTC4R | GATCAGCTACGGGTTGGCCG | | |
| *Rickettsia* sp. | dksA | dksA-F | TCCCATAGGTAATTTAGGTGTTTC | 54 | diagnostical tool |
| | | dksA-R | TACTACCGCATATCCAATTAAAAA | | |
| | mppA | mppA-F | GCAATTATCGGTCCGAATG | 54 | diagnostical tool |
| | | mppA-R | TTTCATTTATTTGTCTCAAAATTCA | | |
| | rpmE | rpmE-F | TTCCGGAAATGTAGTAAATCAATC | 54 | diagnostical tool |
| | | rpmE-R | TCAGGTTATGAGCCTGACGA | | |
| *Streptococcus* sp. | rpoB | Strepto_F | AARYTIGGMCCTGAAGAAAT | 46 | diagnostical tool |
| | | Strepto_R | TGIARTTTRTCATCAAACATGTG | | |
| *Vibrio cholerae* | ompW | ompW-tsense | CACCAAGAAGGTGACTTTATTGTG | 48 | [4] |
| | | ompW-antisense | GGTTTGTCGAATTAGCTTCACC | | |
| | cholera toxin | ctxA-s | TCAGACGGGATTTGTTAGGCACG | 50 | [5] |
| | | ctxA-a-s | TCTATCTCTGTAGCCCCTATTACG | | |
| | chromosom 1 | O1-F | GTTTCACTGAACAGATGGG | 50 | [5] |
| | | O1-R | GGTCATCTGTAAGTACAA | | |
| | SodB | SodB-F | AAGACCTCAACTGGCGGTA | 50 | [6] |
| | | SodB-R | GAAGTGTTAGTGATCGCCAGAGT | | |
| *Yersinia pestis* | glpD | glpD-F1 | GGCTAGCCGCCTCAACAAAAACAT | 58 | [7] |
| | | glpD-R2 | GGTGCCAGTTTCAGTAACAC | | |

**Table S5. The quantitative real-time PCR systems that were tested.**

| Microorganism | Gene | Name | SEQUENCES 5'→3' |
|---|---|---|---|
| *Brucella* sp. | IS711 | Brucellad | GCTCGGTTGCCAATATCAATG |
| | IS711 | Brucellar | GGGTAAAGCGTCGCCAGAAG |
| | | Brucellap | 6FAM-AAATCTTCCACCTTGCCCTTGCCATCA |
| | | | |
| *Brucella abortus* | IS711 | Abortusd | GCGGCTTTTCTATCACGGTATTC |
| | IS711 | Abortusr | CATGCGCTATGATCTGGTTACG |
| | | Abortusp | 6FAM-CGCTCATGCTCGCCAGACTTCAATG |
| | | | |
| *Brucella melitensis* | IS711 | Melitensisd | AACAAGCGGCACCCCTAAAA |
| | IS711 | Melitensisr | CATGCGCTATGATCTGGTTACG |
| | | Melitensisp | 6FAM-CAGGAGTGTTTCGGCTCAGAATAATCC |
| | | | |
| *Bartonella quintana* | yopP | B qui 11580F | TAAACCTCGGGGGAAGCAGA |
| | 1ère intention | B qui 11580R | TTTCGTCCTCAACCCCATCA |
| | | B qui 11580P | 6FAM- CGTTGCCGACAAGACGTCCTTGC -TAMRA |
| | | | |
| | fabF3 | B qui 05300F | GCTGGCCTTGCTCTTGATGA |
| | fabF3 | B qui 05300R | GCTACTCTGCGTGCCTTGGA |
| | | B qui 05300P | 6FAM- TGCAGCAGGTGGAGGAGAACGTG -TAMRA |
| | | | |
| *Bartonella grahamii* | badA2 | B_gra_3_F | AGATGGAAAAATCCGCTCCA |
| | 1ere intention | B_gra_3_R | AGGCAAGGGCAAAGAGCATA |
| | | B_gra_3_P | 6FAM- TCCGCAACGAGTTCTGGTGGTCA -TAMRA |
| | | | |
| *Bordetella pertussis* | Toxine | Toxpert2d | CCTACCAGAGCGAATATCTGGCA |
| | Toxine | Toxpert2r | GCGTTACCCTGCGGATGTTTT |
| | Toxine | Toxpert2p | 6FAM-ACCGGCGCATTCCGCCC |

111

| | | | |
|---|---|---|---|
| *Chlamydia pneumoniae* | omp2 | ChPnMGBd | GATTCGTCGCTAGTGCGGA |
| | omp2 | ChPnMGBr | GTCTAACCTTCTTCGCTGTCA |
| | omp2 | ChPnMGB | 6FAM-ACAAAGCCAGCACCTGTTCCT -Mgb |
| | | | |
| *Chlamydia trachomatis* | unknown | Chlam-tracho_1_F | AGCTCCCAAAGCAACCAGAR |
| | unknown | Chlam-tracho_1_R | BTGTCGCTGCGTTGGTTTTA |
| | unknown | Chlam-tracho_1_P | 6FAM-CAACAGCACCACCAGCAGCTGC |
| | | | |
| *Coxiella burnetii* | IS1111A | IS 1111 0706 F | CAAGAAACGTATCGCTGTGGC |
| | IS1111A | IS 1111 0706 R | CACAGAGCCACCGTATGAATC |
| | IS1111A | IS1111 07-06 P | 6FAM-  CCGAGTTCGAAACAATGAGGGCTG  -TAMRA |
| | | | |
| *Escherichia coli* | ompG | ECOmpGMGBAluId | GCTGCGCGTGCAAATGCG |
| | ompG | ECOmpGMGBAluIr | CATGGTCATCGCTTCGGTCT |
| | ompG | ECOmpGMGB | 6FAM-CATCAGAAACTGAACACCAC -Mgb |
| | | | |
| **Klebsiella pneumoniae** | β-lactamase | KPCall-2F | CGCCGTGCAATACAGTGATA |
| **carbapenemase (KPC)** | β-lactamase | KPCall-2R | GCAGAGCCCAGTGTCAGCTT |
| | | KPCprobeall-2 | 6FAM- CTCTATCGGCGATACCACGT |
| | | | |
| **New Delhi metallo-** | metallo-β-lactamase | NDM1-F | GCGCAACACAGCCTGACTTT |
| **beta-lactamase-1** | metallo-β-lactamase | NDM1-R | CAGCCACCAAAAGCGATGTC |
| **(NDM-1)** | | | |
| | | NDM1 | 6FAM- CAACCGCGCCCAACTTTGGC |
| | | | |
| *Legionella pneumophila* | gyrB | Lpneumo_gyrB_MBF | TGGAACCGGTTTGCATCATA |
| | gyrB | Lpneumo_gyrB_MB R | CGACAGGAATACCACGRCCA |
| | | Lpneumo_gyrB_MBP | FAM- TGAATCCCTGGCAGGTTATTGCAAGG |

112

| | | | |
|---|---|---|---|
| *Leptospira interrogans* | 16S | Lep16SMGBd | GCGGCGAACGGGTGAGTAA |
| | 16S | Lep16SMGBr | GGAAAGTTATCCAGACTC |
| | | Lep16SMGB | 6FAM-ACGTGGGTAATCTT -Mgb |
| | hsp | Lint_hsp_MBF | TTCTCGCTCCCAGAGTAGACA |
| | hsp | Lint_hsp_MBR | TTTTCCAATTGAACTTGAACGTC |
| | | Lint_hsp_MBP | FAM- ACTGGCCGACCTTCCAGGTGTAGA |
| *Listeria monocytogenes* | hlyQ | hlyQf | CATGGCACCACCAGCATC |
| | hlyQ | hlyQr | ATCCGCGTGTTTCTTTTCGA |
| | hlyQ | hlyQp | 6FAM-CGCCTGCAAGTCCTAAGACGCCA |
| *Mycobacterium tuberculosis* | ITS | ITSd | GGGTGGGGTGTGGTGTTTGA |
| | ITS | ITSr | CAAGGCATCCACCATGCGC |
| | ITS | sonde tub | 6FAM-GCTAGCCGGCAGCGTATCCAT |
| *Rickettsia* **sp.** | unknown | 1029-F1 | GAMAAATGAATTATATACGCCGCAAA |
| | DIAGNOSTIC | 1029-R1 | ATTATTKCCAAATATTCGTCCTGTAC |
| | DIAGNOSTIC | Rick1029_MBP | 6FAM- CGGCAGGTAAGKATGCTACTCAAGATAA |
| | gltA (CS) | RKND03_F | GTGAATGAAAGATTACACTATTTAT |
| | EPIDEMIO | RKND03_R | GTATCTTAGCAATCATTCTAATAGC |
| | EPIDEMIO | RKND03 P | 6FAM- CTATTATGCTTGCGGCTGTCGGTTC  -TAMRA |
| *Rickettsia prowazekii* | ompB | Rpr_ompB_F | AATGCTCTTGCAGCTGGTTCT |
| | ompB | Rpr_ompB_R | TCGAGTGCTAATATTTTTGAAGCA |
| | ompB | Rpr_ompB_P | 6FAM- CGGTGGTGTTAATGCTGCGTTACAACA  -TAMRA |
| *Rickettsia felis* | bioB | R_fel0527_F | ATGTTCGGGCTTCCGGTATG |

| | | | |
|---|---|---|---|
| | bioB | R_fel0527_R | CCGATTCAGCAGGTTCTTCAA |
| | bioB | R_fel0527_P | 6FAM- GCTGCGGCGGTATTTTAGGAATGGG -TAMRA |
| | orfb | orfb-f | CCCTTTTCGTAACGCTTTGCT |
| | orfb | orfb-r | GGGCTAAACCAGGGAAACCT |
| | orfb | orfb- p (sonde tamra) | 6-FAM-TGTTCCGGTTTTAACGGCAGATACCCA-TAMRA |
| **Shiga toxine de type 2** | stx2 | stx2F | GTGGCATTAATACTGAATTGTCATCA |
| | stx2 | stx2R | GCGTAATCCCACGGACTCTTC |
| | stx2 | stx2P | 6FAM- CGGACCTCTGTATCTGCCTGAAGCGTAAGGGTCCG |
| *Streptococcus pneumoniae* | systeme poc | plyNd | GCGATAGCTTTCTCCAAGTGG |
| | systeme poc | plyNR | TTAGCCAACAAATCGTTTACCG |
| | systeme poc | plyN_P | 6FAM-CCCAGCAATTCAAGTGTTCGCCGA |
| | lytA | Pneumo_lytA_F | CCTGTAGCCATTTCGCCTGA |
| | lytA | Pneumo_lytA_R | GACCGCTGGAGGAAGCACA |
| | lytA | LytA_P | 6FAM-AGACGGCAACTGGTACTGGTTCGACAA |

**Legends of Supplementary Figures**

**Figure S1. Working overview of the polyphasic approach used to analyze the Namur coprolite.**

**Figure S2. Phylogenetic tree of a hydrolase.** A phylogenetic tree was generated from the translated open reading frame of a contig encoding a hydrolase close to *Bordetella* species. The tree was constructed using the PhyML algorithm with a bootstrap of 100. The bootstrap support is reported for each branch.

**Figure S3. Bayesian source-tracking results.** (**A**) The mixture of taxa associated to the 16S rDNA gene amplicon dataset of the coprolite specimen was compared to known dataset of various environments. To control the workflow used to perform the analyses, two known samples (**B**) one coprolite previously investigated [9] and (**C**) a soil sample [10] were positively tested.

**Figure S4. Phylogenetic tree of 16S rDNA amplicons matching to *Bartonella sp.*.** The tree was constructed using the PhyML algorithm with a bootstrap of 100. The bootstraps are reported for each branch. Phylogenetic tree of 16S rDNA amplicons closely related to (**A**) *B. henselae*, *B. koehlerae* and (**B**) *B. quintana*.

**Figure S5. Alignment and of the amplicon matching to *Bordetella* and *Achromobacter*.** The multiple sequence alignment of the amplicon with *Bordetella* species and *Achromobacter xylosoxidans* was performed using CLUSTALW multiple alignment tool [11].

**Figure S6. Metabolic comparison of modern metagenomes to the coprolite metagenome.** The Principal coordinates analysis was based on read classification according to BLASTX searches against the SEED Database. For each metagenome included the MG-RAST accession number is given. Compared metagenomes are from soil (yellow cluster), healthy mammalian and human feces (blue cluster); and the

coprolite (red). The coprolite metagenome does not group with either the modern gut or soil microbiota.
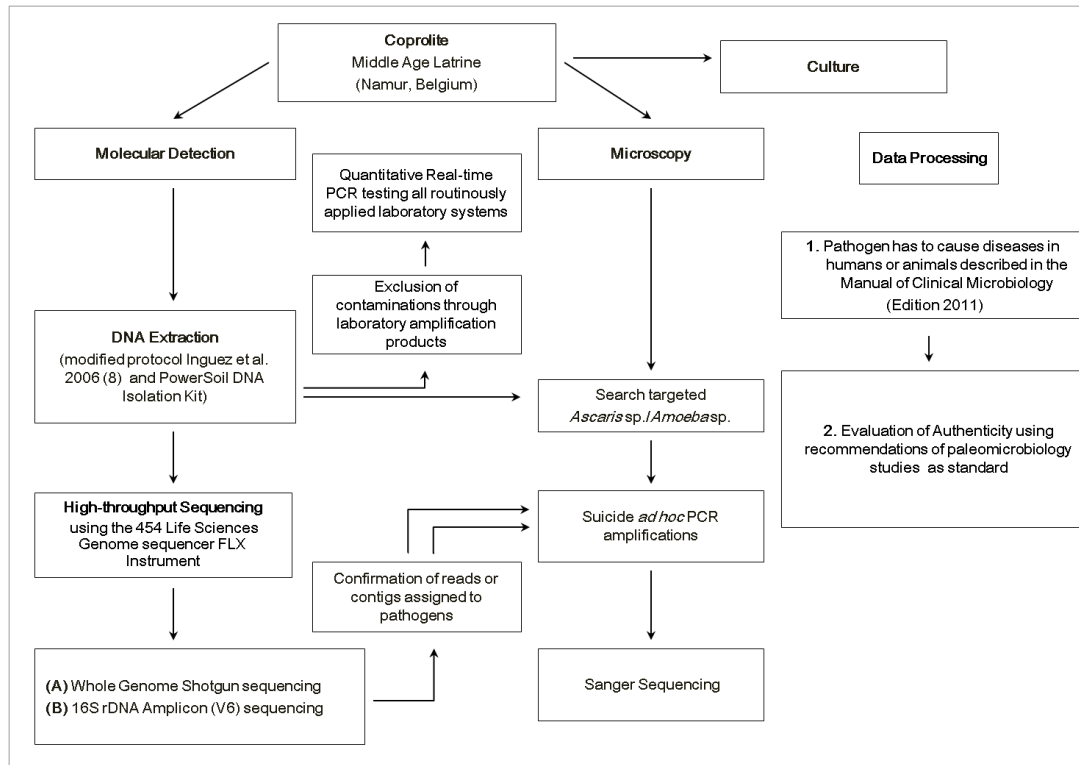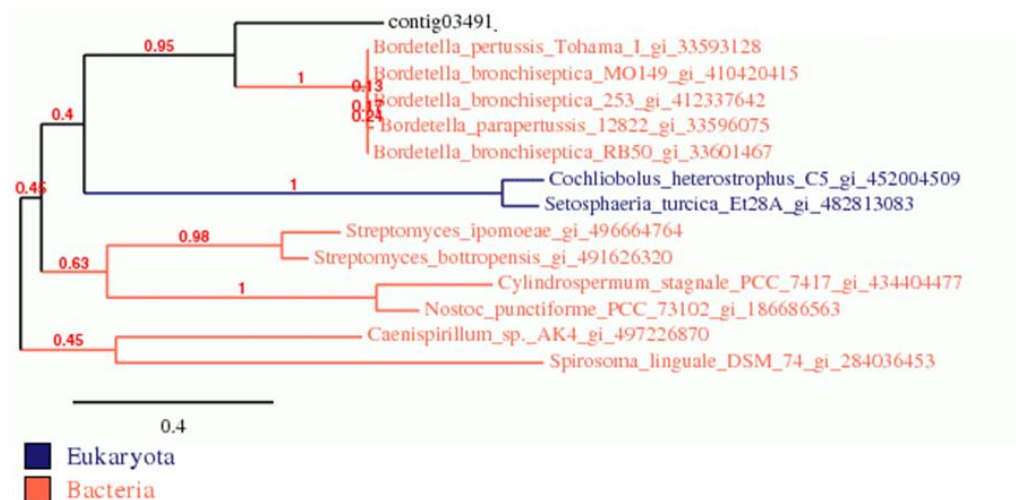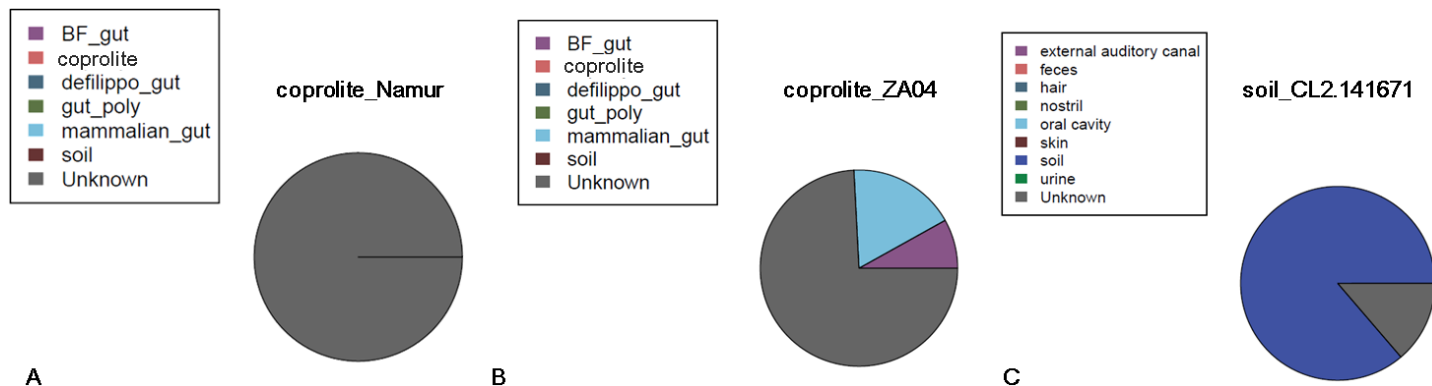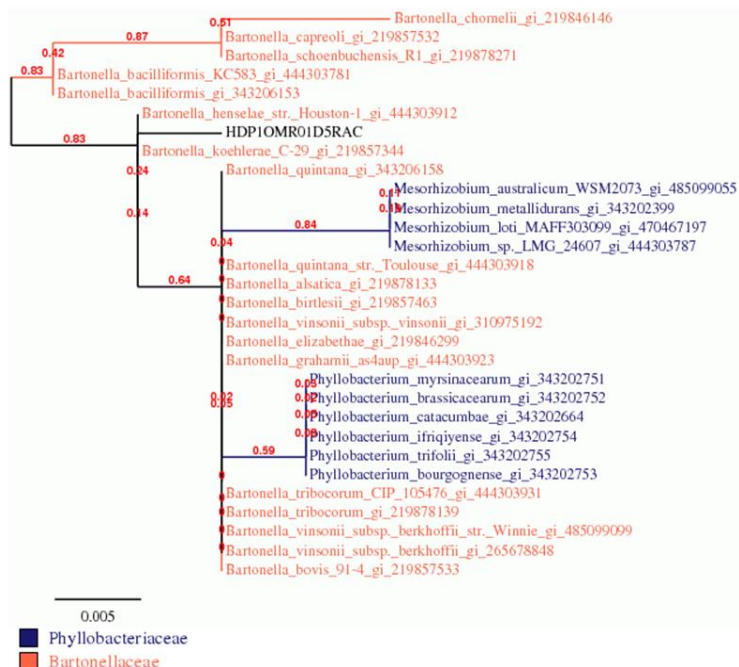
**Figure S1.**

Figure S2.

**Figure S3**.

**Figure S4.**

```
                                          10        20        30        40        50        60        70        80        90       100
                                           |         |         |         |         |         |         |         |         |         |
SequenceCoprolite                CAGGTCGGAGGTCGCGAAGCGGCCGCCGTCCAGCGGCACCAGCGGACGCAGGTCCGGCGGCAGCACGGGCAGCACTTCCATGACCATCCACTCGGCCTTG
HE965806.1 B. bronchiseptica 253 ..........................................................................................T.........
HE965803.1 B. parapertussis Bpp5 ..........................................................................................T.........
BX640437.1 B. bronchiseptica RB50 ..........................................................................................T.........
BX640423.1 B. parapertussis 12822 ..........................................................................................T.........
HE965807.1 B. bronchiseptica MO149 ..........................................................................................T.........
CP002287.1 A. xylosoxidans A8    ....................................................................................................
                                 *********************************************************************************** *********

                                          110       120       130       140       150       160       170       180
                                           |         |         |         |         |         |         |         |
SequenceCoprolite                ATGCCGGACTTCTGGAAGCCTTCCAGCACCTTCAGGCGCTTGGAGATCTTCTTGATCTTGGCTTCCGAGCTGGTGGCCTTGAGCTCGCC
HE965806.1 B. bronchiseptica 253 ......................................................................................
HE965803.1 B. parapertussis Bpp5 ......................................................................................
BX640437.1 B. bronchiseptica RB50 ......................................................................................
BX640423.1 B. parapertussis 12822 ......................................................................................
HE965807.1 B. bronchiseptica MO149 ...............................................................................T.....
CP002287.1 A. xylosoxidans A8    .............................................................................C..T.....
                                 ******************************************************************************* ** *****
```

**Figure S5.**

125

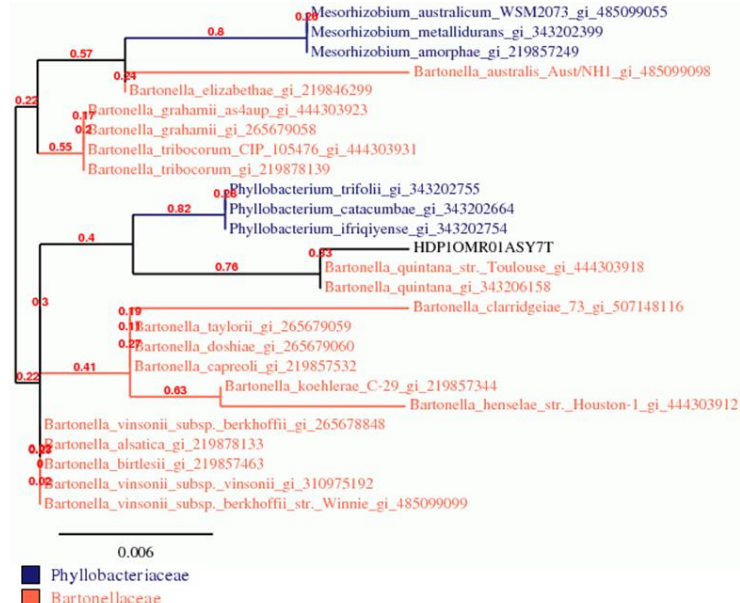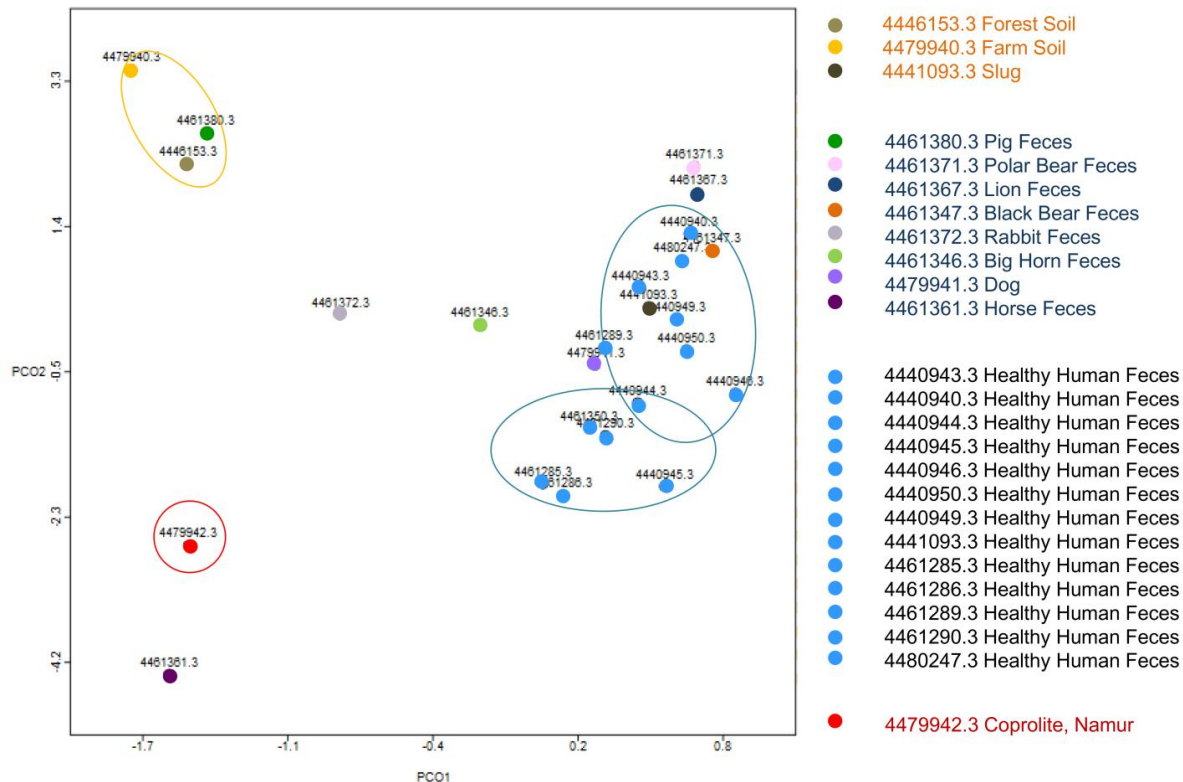**Figure S6.**

# Chapitre III

ARTICLE

## Des virus dans un coprolithe du 14ième siècle

Sandra Appelt[1,*], Laura Fancello[1,*], Matthieu Le Bailly[2], Didier Raoult[1], Michel Drancourt[1], Christelle Desnues[†,1]

[1] Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, Inserm 1095, 13385 Marseille, France.[2] Franche-Comté University, CNRS UMR 6249 Chrono-Environment, 25 030 Besançon, France.

[*] These authors have contributed equally to this work

† Corresponding author: christelle.desnues@univ-amu.fr

## Avant-propos

Très peu de travaux paléomicrobiologiques sont consacrés aux virus. Certains travaux ont rapporté l'observation de particules virales dans des fossiles : des virions du virus de la variole ont ainsi été observés dans des restes humains datés de 400 ans (Fornaciari & Marchetti 1986; Marennikova et al. 1990). Ces observations obtenues par microscopique électronique on été récemment confirmées par la détection de séquences spécifique des Poxvirus de la variole, dans des échantillons de 300 ans (Biagini et al. 2012). Des séquences d'un *Anelloviridae* – pouvant provoquer des infections virales chroniques – et d'un virus lymphotrophique des cellules T ont été respectivement identifiées dans la pulpe dentaire d'un soldat de l'armée napoléonienne (Kalingrad, 1812), et dans les restes d'une momie des Andes datés de 1 500 ans (Bedarida et al. 2011; Li et al. 1999; Sonoda et al. 2000). Plus proche de nous la detection de virus de la grippe espagnole (Taubenberger 1997). Les coprolithes pourraient également servir de base à des investigations sur les infections virales, mais aucun travail de paléovirologie des coprolithes n'a été publié. Ici, nous avons caractérisé par microscopie électronique à transmission et par métagénomique virale, la communauté virale associée au coprolithe de Namur.

JOURNAL ARTICLE

# Viruses in a 14th-century coprolite

**Running title:** Viruses in a 14th-century coprolite

Sandra Appelt[1,*], Laura Fancello[1,*], Matthieu Le Bailly[2], Didier Raoult[1], Michel Drancourt[1], Christelle Desnues[†,1]

[1] Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, Inserm 1095, 13385 Marseille, France. [2] Franche-Comté University, CNRS UMR 6249 Chrono-Environment, 25 030 Besançon, France.

[*] These authors have contributed equally to this work
† Corresponding author: christelle.desnues@univ-amu.fr

## Abstract

Coprolites are fossilized fecal material that can reveal information about ancient intestinal and environmental microbiota. Viral metagenomics has allowed systematic characterization of viral diversity in environmental and human-associated specimens, but little is known about the viral diversity in fossil remains. Here, we analyzed the viral community of a 14th-century coprolite from a closed barrel in a Middle Age site in Belgium using electron microscopy and metagenomics. Viruses that infect eukaryotes, bacteria and archaea were detected, and we confirmed the presence of some of them by *ad hoc* suicide PCR. The coprolite DNA viral community was dominated by bacteriophages commonly found in modern stools and soil. Although their phylogenetic compositions differed, the metabolic functions of the viral communities have remained conserved across centuries. Antibiotic resistance was one of the reconstructed metabolic functions detected.

**Introduction**

Viral metagenomics is a sequencing-based analysis of all of viral genomes isolated from a sample. It has promoted the characterization of viral community diversity. Viral metagenomics has already been successfully applied to the exploration of modern environmental specimens sampled from marine water, freshwater, stromatolites and thrombolites and soil (1-4) and to modern human-associated specimens collected from the liver, blood, nasopharyngeal aspirates and stool (5-9). The DNA viromes generated from modern stools have been demonstrated to be dominated by bacteriophages (10, 11) and to be less diverse than environmental samples (8, 12).

Viral metagenomics does not require culturing viruses or *a priori* knowledge of the sequences that will be targeted, which allows for the identification of new, unknown or unexpected viruses and for the global assessment of the virome. Viral metagenomics is thus particularly suitable for paleomicrobiological studies, as little is known about which viruses are characteristic of ancient specimens. Indeed, the majority of ancient DNA (aDNA) studies are based on the analysis of human and bacterial aDNA (13-15), and viral persistence and its detectability in ancient specimens remains unclear. Electron microscopy has previously revealed that viral particles can persist for over 400 years, but their viability was lost (16). Moreover, PCR amplifications yielded positive results for viral aDNA in ancient specimens such as mummified soft tissues, bones and teeth. The amplification products varied between 100 and 570 bp in size, which indicated that viral aDNA can be detected for at least 1,500 years (17-20).

Here, we used electron microscopy and, for the first time, viral metagenomics to characterize the viral community of an ancient stool specimen. A viral DNA metagenome was generated from a 14th-century coprolite sample that was recovered from a Middle Age site in Namur (Belgium).

**Material and Methods**

**Virus-like particle isolation, transmission electron microscopy (TEM) and DNA extraction.** First, 5.8 grams of the interior of the coprolite were aseptically removed and solubilized overnight at 4°C under continuous rotation in 40 mL of phosphate saline buffer (PBS), pH 7.4 (bioMérieux, Marcy-l'Etoile, France), which had previously been passed through a 0.02-μm filter. The coprolite solution was centrifuged for 10 min at 500 g; then, the upper layer was removed and filtered in stages using sterile Whatman filters (pore sizes: 0.8 μm, 0.45 μm, and 0.22 μm, (Whatman Part of GE Healthcare, Dassel, Germany)). Twenty-five milliliters of the coprolite filtrate were used to precipitate and purify viral particles onto a cesium chloride density gradient using ultracentrifugation, and DNAse treatment was then performed (21). A 40-μl aliquot of the purified viral particles was stained with 1.5% ammonium molybdate (Euromedex) and observed by transmission electron microscopy using a Philips Morgagni 268D electron microscope (FEI Co., Eindhoven, Netherlands). To isolate the nucleic acids from the purified viral particles, the formamide procedure previously described by Thurber et al. (21) was used. A standard 18S rDNA PCR was performed to verify the absence of human DNA contamination.

**Viral metagenomic library preparation and sequencing.** Nucleic acids were amplified using the Illustra GenomiPhi V2 DNA Amplification Kit (GE Healthcare Life Sciences, Freiburg, Germany). Phi29 DNA polymerase preferentially amplifies circular and single-stranded DNA (4, 12). To minimize this bias, duplicate reactions were performed as previously suggested (21). Amplification products were pooled and ethanol purified.

A shotgun strategy was chosen for high-throughput pyrosequencing on a 454 Life Sciences Genome FLX sequencer using titanium chemistry (Genome Sequencer RLX, Roche). Sequencing was performed using 1/16 of a picotiter plate.

**Preprocessing of sequencing data.** The reads were screened for quality using mothur (22). Only reads longer than 50 bp and with an average quality score greater than 21 were kept. Reads with more than two ambiguous base calls and/or reads with homopolymers longer than ten bases were eliminated. Identical sequences artificially generated by the pyrosequencing technology were also excluded using the "unique.seqs" mothur command. The preprocessed viral metagenome is publicly available on the Metavir server (www.metavir-meb.univ-bpclermont.fr) under the identifier "Namur_viral."

**Annotation of Reads.** A BLASTN search against the non-redundant NCBI database (E-value$<1e^{-05}$) was performed. Reads with no significant similarity to sequences stored in the NCBI database were classified as "unknown reads." The relative abundances of the viral genotypes were estimated using GAAS (23), which is based on a BLASTX search

against the RefSeq Viral Genomes database (E-value<1e$^{-05}$) and normalizes the number of reads matching each viral genotype by the length of the genome.

Functional annotation was performed on the MG-RAST server (24) using the non-redundant SEED database (E-value<1e$^{-05}$). A stringent search of virulence factors was also performed using BLASTX on the Virulence Factor Database (25), with 60% as the minimum identity and a cutoff E-value<1e$^{-05}$.

**Assembly and contig annotation.** The reads were assembled into contigs using the Newbler de novo assembler (Roche) with at least 98% identity and 35 bp of overlap. Only contigs longer than 400 bp were used in subsequent analyses.

Known and unknown contigs were identified on the basis of the BLASTN search against the non-redundant NCBI database (E-value<1e$^{-05}$). The taxonomic and functional contig classification was based on a BLASTX search against the non-redundant NCBI database (E-value<1$^{-05}$). A specific search for contigs encoding antibiotic resistance genes was also performed using BLAST on the ARDB (Antibiotic Resistance Genes Database) with an E-value<1e$^{-05}$ (26). Significant hits were manually verified.

**Phylogenetic trees.** When possible, phylogenetic trees of the contigs encoding antibiotic resistance genes were built. The program Prodigal was used to search for open reading frames (ORFs) in these contigs (27). Homologs to the translated ORFs were searched against the non-redundant NCBI database using BLASTP. A multiple alignment was constructed using MUSCLE (28) and curated using Gblocks (29). The phylogenetic tree was then built using the PhyML algorithm (30) with a bootstrap value

of 100. These tasks were all performed using the pipeline freely available on www.phylogenie.fr (31). The trees were visualized using MEGA v.4 (32).

**Evidence of temperate bacteriophages.** Contigs generated from the assembly were analyzed to search for indicators of temperate bacteriophages, as previously described (33). We searched for three indicators: 1) nucleotide identity to bacterial genomes (BLASTN, E-value<1e$^{-05}$, 90% minimum identity, 90% minimum query coverage), 2) presence of integrase-encoding genes using annotations from the COG and PFAM databases (E-value<1e$^{-05}$) and 3) significant similarity to prophage proteins available on the ACLAME database (BLASTX on the ACLAME prophages database, E-value<1e$^{-05}$). Data were graphically represented using the R package "VennDiagram."

**Comparative metagenomics.** The coprolite-associated virome was taxonomically and functionally compared to 21 published viromes of modern stools from healthy adult humans (12, 33). All viromes were taxonomically (Genbank database, E-value<1e$^{-05}$) and functionally annotated (SEED database, E-value<1e$^{-05}$) using MG-RAST. The taxonomic and functional virome profiles were compared using principal component analysis on the MG-RAST server (normalized data, Bray-curtis measure of distance). Species richness estimations were obtained from the MG-RAST server. Functional diversity (measured by the Shannon-Wiener index) was calculated using the "estimateDiversity" function of the ShotGunFunctionalizeR package on the SEED-based functional metagenome annotations (E<1e$^{-05}$) (34).

**Specific PCR amplifications and sequencing.** Suicide PCR amplifications were performed to confirm high-throughput pyrosequencing results. To perform suicide PCR,

the primer pairs were used only once in working areas, and no positive controls were used (35). For giant virus detection, primer pairs targeting the nonfunctional B-family DNA polymerase were used. Additional primer pairs were designed to specifically target ORFs identified in some viral contigs assembled *de novo* from the virome and matching *Cyanophages*, *Mycobacterium* phages, *Bacillus* phages, *Burkholderia* phages, *Celeribacter* phages and *Clostridium* phages (Supplementary Table S1 and Supplementary Material, Section 4).

**Results**

The specimen was excavated in 1996 and collected from the interior of a closed barrel, which was commonly used during this period as a pit or latrine (36). The barrel was buried at a depth of 3.80 m. The 121.4-g coprolite specimen was dark brown and well preserved under anaerobic taphonomic conditions. Extensive precautions were undertaken to avoid contaminating the coprolite specimen in our laboratory environment; no positive control was used (15) and suicide PCR protocols were applied (35). All negative controls, used in a 1:4 control:specimen ratio, were consistent with current recommendations for paleomicrobiological and paleoparasitological studies (13, 15, 37, 38) and remained negative. Virus-like particles (VLPs) purified from the internal region of the coprolite, after the external layer was removed, were morphologically diverse and varied in size and shape. Oval particles of different lengths (up to 200 nm) and diameters (up to 100 nm), as well as rod-shaped structures (up to 250 nm in length), were observed (Fig. 1A). We identified a VLP with a dense core and a diameter of approximately 150 nm, apparently surrounded by an envelope-like structure (Fig. 1B).

Viral particles exhibiting characteristics typical of the *Siphoviridae* bacteriophage family (icosahedral head, long tail) were also observed (Figs. 1C-1E).

High-throughput sequencing generated 30,654 reads corresponding to approximately 10.8 million bp. After quality trimming and duplicate removal, 29,811 reads remained (Supplementary Table S2). The preprocessed read lengths ranged between 77 bp and 574 bp and had an average GC content of 47% (Supplementary Fig. S2). Finally, 41.93% of the reads were assembled into 1,464 contigs that ranged from 421 to 12,500 bp (Supplementary Tables S2). In total, 22.15% of all reads and 17.28% of all contigs were significantly similar to known sequences from public databases (Fig. 2A and Supplementary Table S2). The known fraction of the viral DNA community was dominated by double-stranded DNA viruses (85.21%) and single-stranded DNA viruses only represented 0.81% of the community (Fig. 2B). The most abundant double-stranded DNA viral families were *Siphoviridae* (58.89%), *Myoviridae* (8.79%) and *Podoviridae* (5.95%). We found viral families that can infect eukaryotes (*Ascoviridae*, *Poxviridae*, *Iridoviridae*, *Adenoviridae*, *Mimiviridae*, *Herpesviridae*, *Baculoviridae*, *Polydnaviridae* and *Phycodnaviridae*), archaea (*Lipothrixiviridae*, *Tectiviridae* and *Bicaudaviridae*) and bacteria (*Siphoviridae*, *Myoviridae* and *Podoviridae*) (Fig. 2B).

Eukaryotic viruses were present at a low abundance, with *Phycodnaviridae* being the most abundant family (0.81%) (Fig. 2B). We also identified a contig encoding a hypothetical protein of invertebrate iridescent virus 3 (IIV-3). IIV-3 is a member of the *Iridoviridae* family, genus *Chloriridovirus*, with a large particle size (180 nm) that infects mosquitoes (Supplementary Table S3). Metagenomic results were confirmed by *ad hoc* suicide PCR (35). In the presence of negative controls, a 167-bp fragment of a

141

*Mimiviridae*-like nonfunctional B-family DNA polymerase was amplified and sequenced, revealing 84% identity to that of the Moumouvirus of the *Mimiviridae* family (GenBank Accession No. GU265560.1).

Viral families infecting archaea were also identified at low abundances. These families corresponded to *Lipothrixiviridae* (0.04%), *Tectiviridae* (0.11%) and *Bicaudaviridae* (0.02%) (Fig. 2B). One contig was identified as an environmental *Halophage eHP-6*, an unclassified bacteriophage that infects *Haloarchaea* (Supplementary Table S3).

In contrast, the majority of the identifiable sequences were related to viruses infecting bacteria (bacteriophages), especially those of the genus *Bacillus* (14.08%). The identified bacteriophages can infect as many as 37 different bacterial genera, including bacterial genera commonly associated with the human gut, such as *Enterobacteria* phages (11.54%), *Lactobacillus* phages (2.23%) and *Lactococcus* phages (2.14%) (Fig. 3). Other findings included bacteriophages that infect typical soil-dwelling bacteria: *Geobacillus* phages (7.53%), *Streptomyces* phages (3.98%) and *Delftia* phages (0.11%). Several reads were identified from bacteriophages whose bacterial hosts belong to genera that also include human pathogens, such as *Mycobacterium* phages (7.89%), *Vibrio* phages (0.29%), *Pseudomonas* phages (4.01%), *Streptococcus* phages (5.06%), *Staphylococcus* phages (5.07%), *Listeria* phages (3.48%), *Burkholderia* phages (3.38%) and *Clostridium* phages (3.83%) (Fig. 3). The presence of some of these bacteriophages (*Bacillus*, *Clostridium*, *Mycobacterium* and *Burkholderia* phages) was further supported by contig reconstruction, *ad hoc* PCR amplification and sequencing (Supplementary Tables S1 and S3). Moreover, contigs were also identified from bacteriophages that

likely infect hosts known to live in aquatic environments (*Cyanophage* S-TIM5, *Synechococcus* phage S-CBS3, *Celeribacter* phage P12053L, a prophage of *Planctomyces limnophilus* DSM 3776 and one uncultured phage identified in a viral metagenomic study of water from the Mediterranean Sea). An ORF encoding a putative phage tail fiber protein of *Celeribacter* phage P12053L, which was identified on one of these contigs, was further amplified by specific PCR, and the 280-bp amplicon was verified by Sanger sequencing. In addition, a 1,939-bp contig matched an unidentified phage previously described in a viral metagenomic study performed on modern human stools (33) (Supplementary Table S3). Only a scaffold is available for the unidentified phage, and the matched protein is annotated as a hypothetical protein; however, this hypothetical protein is predicted to contain a conserved domain corresponding to an N-acetylmuramoyl-L-alanine amidase. This domain is characteristic of autolysins that degrade peptidoglycans and is typically observed in bacteriophage, prophage and bacterial genomes.

Evidence of a temperate life cycle for the detected bacteriophage sequences was observed using three indicators (33): 1) nucleotide identity to bacterial genomes (an indicator of prophage formation), 2) the presence of integrase-encoding genes (markers of temperate bacteriophages); 3) similarity to prophage sequences available in the ACLAME database (see Materials and Methods). We observed that 329 contigs (22.47%) significantly matched prophage proteins in the ACLAME database; 52 contigs (3.55%) had significant nucleotide identity to bacterial genomes (especially genomes of *Escherichia coli* strains) and 32 contigs (2.18%) harbored integrase-encoding genes (Supplementary Figure S2). This strategy provides only a minimal estimate of the

number of temperate bacteriophages, as stated by Minot et al. (33). Overall, 375 contigs (25.61%) presented at least one of the three indicators and could be tentatively classified as temperate bacteriophages (Supplementary Fig. S2).

The coprolite-associated DNA virome was compared to the viromes of 21 modern human stool specimens (Fig. 4). At the taxonomic level, the coprolite virome did not group with modern stool viromes, whereas it was functionally more similar to some of the modern stool samples (Fig. 4). Overall, the coprolite virome displayed higher species richness (315.279) and seemed to be more functionally diverse (average Shannon-Wiener index of 4.8693) than modern stool viromes (average species richness of 77.824 and average Shannon-Wiener index of 4.1264) (Supplementary Table S4). A more extensive functional analysis of the assembled contigs revealed that most of the identifiable ORFs harbored by these contigs coded for genes involved in DNA metabolism (n=80), as is typical of viromes, and virulence genes (n=87). The most abundant virulence genes were those involved in resistance to antibiotics and toxic compounds. In particular, a contig encoding a chloramphenicol O-acetyltransferase gene that mediates chloramphenicol resistance was observed. This gene was found to belong to *Chryseobacterium* sp., and the BLAST-based annotation was further confirmed by a phylogenetic tree constructed from the ORF of this contig (Supplementary Fig. S3). To further investigate the presence of virulence genes, virome reads were examined using the Virulence Factor Database, which includes both conventional factors directly involved in the pathogenesis and factors important to establishing infection. A stringent search allowed the identification of 166 reads encoding virulence factors. In particular, virulence factors of the bacterial genera *Escherichia* (n=42), *Salmonella* (n=39) and

*Shigella* (n=34) were observed (Supplementary Table S5). A pathway-centric analysis based on COG annotation revealed that virulence (defense mechanisms) was overrepresented in the coprolite compared to modern stools. Other differences included an overrepresentation of lipid transport and metabolism, fatty acid biosynthesis and amino acid transport and metabolism (Supplementary Table S6). Indeed, 12 ORFs on annotated contigs were found to encode genes involved in lipid metabolism, in particular fatty acid biosynthesis (n=3), glycerolipid and glycerophospholipid metabolism (n=3), isoprenoid metabolism (n=3) and polyhydroxybutyrate metabolism (n=3). Annotated contigs also contained 36 ORFs encoding functions related to the metabolism of amino acids, especially lysine, threonine, methionine and cysteine (n=11) and arginine, urea and polyamines (n=10).

**Discussion**

We report the first metagenomic analysis of an ancient human DNA virome. The use of viral metagenomics allowed us to perform a systematic research of known and unknown viruses without *a priori* targeting of expected viruses.

Because minimizing contamination is vital in paleomicrobiology, extensive precautions established by previously published recommended protocols were implemented to avoid contamination of the coprolite specimen (13, 15, 37, 38). The coprolite studied here was recovered from a sealed barrel that was still intact at the time it was found, suggesting that the coprolite was protected from contamination by environmental material for centuries. Only the internal region of the coprolite was used in our experiments. We ascertained the presence of viruses by three independent approaches, *i.e.*, electron microscopy, metagenomics and suicide PCR. The PCR

amplification product sequences were original, *i.e.*, they had not been previously observed in our laboratory, and all negative controls remained negative.

The viral metagenome was generated using a multiple-displacement amplification of viral genomes *via* the phi29 polymerase. This method is known to preferentially amplify circular and single-stranded DNA (12). To minimize this bias, a duplicate amplification reaction was performed as previously suggested (21). Indeed, the taxonomic composition of the generated virome shows low abundance of single-stranded DNA viruses (0.81% according to the GAAS estimates). Therefore, amplification bias does not seem to be significant, although it should be taken into account when estimating viral abundance. Additionally, it should be considered that these estimations are based on the identifiable viral sequences (known sequences) of the DNA virome.

The majority of the generated metagenomic sequences were of unknown origin. The known sequences corresponded to DNA viruses that infect eukaryotes, bacteria and archaea. Eukaryotic and archaeal viral sequences were detected only at low abundances, and their presence was supported by contig recovery or confirmed by *ad hoc* suicide PCR. The majority of the identifiable sequences recovered from the coprolite corresponded to bacteriophages, especially *Siphoviridae*. Apparently, the bacteriophages originated from two sources: the environment and digestive tract microbiota. Some of the identified bacteriophages infect bacteria belonging to genera that include mammalian pathogens. These findings are consistent with those obtained for the previously generated bacterial metagenome (39). In addition to bacteriophages that infect bacteria already described in the digestive tract, bacteriophages infecting

bacterial genera typically found in the soil environment, such as *Geobacillus* phages, were also detected. As the fossilization process is long, coprolites can also incorporate environmental microorganisms between the moment after defecation and the end of desiccation. Thus, the study of coprolites offers the opportunity to explore ancient intestinal and environmental microbiota (40, 41). We detected bacterial viruses associated with hosts that can be found in aquatic and marine environments, such as *Cyanophages* phages and *Synechococcus* phages. It could be speculated that this may be related to dietary habits. The consumption of fish has already been described for Namur's citizens during the Middle Ages on the basis of archeological remains and the location of the town at the intersection of two rivers (42, 43).

Comparative analyses to previously published viromes show that modern human stool viromes do not group with the coprolite virome at the taxonomic level. All previous works on viral communities associated with the stool of healthy individuals showed a high prevalence of bacteriophages, in particular double-stranded DNA bacteriophages of the *Siphoviridae* family (8, 11, 33) or single-stranded DNA bacteriophages of the *Microviridae* family (10, 12), with high inter-individual variability. Accordingly, the coprolite virome shows a high prevalence of *Siphoviridae* (8, 11, 33). Moreover, as in modern stools, we found evidence for temperate bacteriophages (10, 33). However, we did not observe significant abundance of single-stranded DNA viruses (10, 12) or the same most abundant prophages identified in other modern stool viromes (33), and we observed a high abundance of bacteriophages infecting bacteria typically found in soil. At the functional level, no clear separation can be observed between the coprolite virome and modern stool viromes, and functions

might be conserved between the coprolite and some modern stool samples. This finding is consistent with those of a recent study that demonstrated that despite inter-individual taxonomic variability, the metabolic profile was significantly conserved within viromes from the same ecological niche (44). This persistence of metabolic functionalities across centuries may reinforce the crucial role of the viral community in the human gastrointestinal tract.

Finally, the coprolite virome is more functionally diverse and rich in virulence genes than modern stool samples viromes. One contig encoding a gene for chloramphenicol resistance (the chloramphenicol O-acetyltransferase), a broad-spectrum antibiotic that inhibits bacterial protein synthesis, was identified. The presence of antibiotic resistance genes in viral metagenomes has been reported in modern human stools (33). Indeed, bacteriophages constitute a reservoir of resistance genes (45-47) and bacteriophage transduction represents one important mode of lateral transfer of resistance genes between bacterial species. Phylogenetic studies have demonstrated that the evolution and dissemination of resistance genes started well before the use of antibiotics (48-50). Accordingly, direct evidence for the presence of antibiotic resistance genes in pre-antibiotic era specimens was provided by *ad hoc* PCR amplifications using DNA extracted from 30,000-year-old permafrost sediments in Canada (51). Here, we demonstrate that bacteriophages are an ancient reservoir of resistance genes associated with human samples that date back as far as the Middle Ages. Moreover, we provide evidence for the lysogenic lifestyle of these bacteriophages, which may support their role in the mobilization and lateral transfer of genes in bacterial communities.

Overall, this study furthers our understanding of past viral diversity and distribution and promotes the further exploration of ancient viral communities using coprolite specimens.

**Acknowledgments**

# References

1.	**Angly FE, Felts B, Breitbart M, Salamon P, Edwards RA, Carlson C, Chan AM, Haynes M, Kelley S, Liu H, Mahaffy JM, Mueller JE, Nulton J, Olson R, Parsons R, Rayhawk S, Suttle CA, Rohwer F.** 2006. The marine viromes of four oceanic regions. PLoS Biol **4:** e368

2.	**Lopez-Bueno A, Tamames J, Velazquez D, Moya A, Quesada A, Alcami A.** 2009. High diversity of the viral community from an Antarctic lake. Science **326:** 858-61

3.	**Desnues C, Rodriguez-Brito B, Rayhawk S, Kelley S, Tran T, Haynes M, Liu H, Furlan M, Wegley L, Chau B, Ruan Y, Hall D, Angly FE, Edwards RA, Li L, Thurber RV, Reid RP, Siefert J, Souza V, Valentine DL, Swan BK, Breitbart M, Rohwer F.** 2008. Biodiversity and biogeography of phages in modern stromatolites and thrombolites. Nature **452:** 340-3

4.	**Kim KH, Chang HW, Nam YD, Roh SW, Kim MS, Sung Y, Jeon CO, Oh HM, Bae JW.** 2008. Amplification of uncultured single-stranded DNA viruses from rice paddy soil. Appl Environ Microbiol **74:** 5975-85

5.	**Briese T, Paweska JT, McMullan LK, Hutchison SK, Street C, Palacios G, Khristova ML, Weyer J, Swanepoel R, Egholm M, Nichol ST, Lipkin WI.** 2009. Genetic detection and characterization of Lujo virus, a new hemorrhagic fever-associated arenavirus from southern Africa. PLoS Pathog **5:** e1000455

6.	**Anderson NG, Gerin JL, Anderson NL.** 2003. Global screening for human viral pathogens. Emerg Infect Dis **9:** 768-74

7.	**Nakamura S, Yang CS, Sakon N, Ueda M, Tougan T, Yamashita A, Goto N, Takahashi K, Yasunaga T, Ikuta K, Mizutani T, Okamoto Y, Tagami M, Morita R, Maeda N, Kawai J, Hayashizaki Y, Nagai Y, Horii T, Iida T, Nakaya T.** 2009. Direct metagenomic detection of viral pathogens in nasal and fecal specimens using an unbiased high-throughput sequencing approach. PLoS One **4:** e4219

8.      **Breitbart M, Hewson I, Felts B, Mahaffy JM, Nulton J, Salamon P, Rohwer F.** 2003. Metagenomic analyses of an uncultured viral community from human feces. J Bacteriol **185:** 6220-3

9.      **Breitbart M, Rohwer F.** 2005. Method for discovering novel DNA viruses in blood using viral particle selection and shotgun sequencing. Biotechniques **39:** 729-36

10.     **Reyes A, Haynes M, Hanson N, Angly FE, Heath AC, Rohwer F, Gordon JI.** 2010. Viruses in the faecal microbiota of monozygotic twins and their mothers. Nature **466:** 334-8

11.     **Breitbart M, Haynes M, Kelley S, Angly F, Edwards RA, Felts B, Mahaffy JM, Mueller J, Nulton J, Rayhawk S, Rodriguez-Brito B, Salamon P, Rohwer F.** 2008. Viral diversity and dynamics in an infant gut. Res Microbiol **159:** 367-73

12.     **Kim MS, Park EJ, Roh SW, Bae JW.** 2011. Diversity and abundance of single-stranded DNA viruses in human feces. Appl Environ Microbiol **77:** 8062-70

13.     **Pääbo S, Poinar H, Serre D, Jaenicke-Despres V, Hebler J, Rohland N, Kuch M, Krause J, Vigilant L, Hofreiter M.** 2004. Genetic analyses from ancient DNA. Annu Rev Genet **38:** 645-79

14.     **Willerslev E, Cooper A.** 2005. Ancient DNA. Proc Biol Sci **272:** 3-16

15.     **Drancourt M, Raoult D.** 2005. Palaeomicrobiology: current issues and perspectives. Nat Rev Microbiol **3:** 23-35

16.     **Marennikova SS, Shelukhina EM, Zhukova OA, Yanova NN, Loparev VN.** 1990. Smallpox diagnosed 400 years later: results of skin lesions examination of 16th century Italian mummy. J Hyg Epidemiol Microbiol Immunol **34:** 227-31

17.     **Bedarida S, Dutour O, Buzhilova AP, de Micco P, Biagini P.** 2011. Identification of viral DNA (Anelloviridae) in a 200-year-old dental pulp sample (Napoleon's Great Army, Kaliningrad, 1812). Infect Genet Evol **11:** 358-62

18. **Biagini P, Theves C, Balaresque P, Geraut A, Cannet C, Keyser C, Nikolaeva D, Gerard P, Duchesne S, Orlando L, Willerslev E, Alekseev AN, de Micco P, Ludes B, Crubezy E.** 2012. Variola virus in a 300-year-old Siberian mummy. N Engl J Med **367:** 2057-9

19. **Li HC, Fujiyoshi T, Lou H, Yashiki S, Sonoda S, Cartier L, Nunez L, Munoz I, Horai S, Tajima K.** 1999. The presence of ancient human T-cell lymphotropic virus type I provirus DNA in an Andean mummy. Nat Med **5:** 1428-32

20. **Sonoda S, Li HC, Cartier L, Nunez L, Tajima K.** 2000. Ancient HTLV type 1 provirus DNA of Andean mummy. AIDS Res Hum Retroviruses **16:** 1753-6

21. **Thurber RV, Haynes M, Breitbart M, Wegley L, Rohwer F.** 2009. Laboratory procedures to generate viral metagenomes. Nat Protoc **4:** 470-83

22. **Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF.** 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol **75:** 7537-41

23. **Angly FE, Willner D, Prieto-Davo A, Edwards RA, Schmieder R, Vega-Thurber R, Antonopoulos DA, Barott K, Cottrell MT, Desnues C, Dinsdale EA, Furlan M, Haynes M, Henn MR, Hu Y, Kirchman DL, McDole T, McPherson JD, Meyer F, Miller RM, Mundt E, Naviaux RK, Rodriguez-Mueller B, Stevens R, Wegley L, Zhang L, Zhu B, Rohwer F.** 2009. The GAAS metagenomic tool and its estimations of viral and microbial average genome size in four major biomes. PLoS Comput Biol **5:** e1000593

24. **Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M, Paczian T, Rodriguez A, Stevens R, Wilke A, Wilkening J, Edwards RA.** 2008. The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinformatics **9:** 386

25. **Chen L, Xiong Z, Sun L, Yang J, Jin Q.** 2012. VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. Nucleic Acids Res **40:** D641-5

26. **Liu B, Pop M.** 2009. ARDB--Antibiotic Resistance Genes Database. Nucleic Acids Res **37:** D443-7

27. **Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ.** 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics **11:** 119

28. **Edgar RC.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32:** 1792-7

29. **Talavera G, Castresana J.** 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol **56:** 564-77

30. **Guindon S, Delsuc F, Dufayard JF, Gascuel O.** 2009. Estimating maximum likelihood phylogenies with PhyML. Methods Mol Biol **537:** 113-37

31. **Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O.** 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res **36:** W465-9

32. T**amura K, Dudley J, Nei M, Kumar S.** 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol **24:** 1596-9

33. **Minot S, Sinha R, Chen J, Li H, Keilbaugh SA, Wu GD, Lewis JD, Bushman FD.** 2011. The human gut virome: inter-individual variation and dynamic response to diet. Genome Res **21:** 1616-25

34. **Kristiansson E, Hugenholtz P, Dalevi D.** 2009. ShotgunFunctionalizeR: an R-package for functional comparison of metagenomes. Bioinformatics **25:** 2737-8

35. **Raoult D, Aboudharam G, Crubezy E, Larrouy G, Ludes B, Drancourt M.** 2000. Molecular identification by "suicide PCR" of Yersinia pestis as the agent of medieval black death. Proc Natl Acad Sci U S A **97:** 12800-3

36.    **Rocha GCd.** 2003. Praça das Armas, Namur, Bélgica. Contribuição de um estudo paleoparasitológico. Escola Nacional de Saúde Pública-Fiocruz, Rio de Janeiro, Rio de Janeiro. 142 pp.

37.    **Cooper A, Poinar HN.** 2000. Ancient DNA: do it right or not at all. Science 289: 1139

38.    **Hofreiter M, Serre D, Poinar HN, Kuch M, Paabo S.** 2001. Ancient DNA. Nat Rev Genet **2:** 353-9

39.    **Appelt S, Armougom F, Le Bailly M, Robert C, Drancourt M.** 2013. Polyphasic analysis of a Middle Ages coprolite microbiota, Belgium. unpublished data

40.    **Yll R, Carrión JS, Marra AC, Bonfiglio L.** 2006. Vegetation reconstruction on the basis of pollen in Late Pleistocene hyena coprolites from San Teodoro Cave (Sicily, Italy). Palaeogeography, Palaeoclimatology, Palaeoecology **237:** 32-9

41.    **Wood JR, Wilmshurst JM, Wagstaff SJ, Worthy TH, Rawlence NJ, Cooper A.** 2012. High-resolution coproecology: using coprolites to reconstruct the habits and habitats of New Zealand's extinct upland moa (Megalapteryx didinus). PLoS One **7:** e40025

42.    **Boone I, De Cupere B, Van Neer W.** 2002. Social statuts as reflected in the food refuse form late medieval sites in Namur (Belgium). Revue belge de philologie et d'histoire **80:** 1391-4

43.    **Rocha GC, Harter-Lailheugue S, Le Bailly M, Araujo A, Ferreira LF, da Serra-Freire NM, Bouchet F.** 2006. Paleoparasitological remains revealed by seven historic contexts from "Place d'Armes", Namur, Belgium. Mem Inst Oswaldo Cruz 101 Suppl **2:** 43-52

44.    **Willner D, Furlan M, Haynes M, Schmieder R, Angly FE, Silva J, Tammadoni S, Nosrat B, Conrad D, Rohwer F.** 2009. Metagenomic analysis of respiratory tract DNA viral communities in cystic fibrosis and non-cystic fibrosis individuals. PLoS One **4:** e7370

45.     **Muniesa M, Garcia A, Miro E, Mirelis B, Prats G, Jofre J, Navarro F.** 2004. Bacteriophages and diffusion of beta-lactamase genes. Emerg Infect Dis **10:** 1134-7

46.     **Colomer-Lluch M, Jofre J, Muniesa M.** 2011. Antibiotic resistance genes in the bacteriophage DNA fraction of environmental samples. PLoS One 6: e17549

47.     **Mazaheri Nezhad Fard R, Barton MD, Heuzenroeder MW.** 2011. Bacteriophage-mediated transduction of antibiotic resistance in enterococci. Lett Appl Microbiol **52:** 559-64

48.     **Hall BG, Barlow M.** 2004. Evolution of the serine beta-lactamases: past, present and future. Drug Resist Updat **7:** 111-23

49.     **Garau G, Di Guilmi AM, Hall BG.** 2005. Structure-based phylogeny of the metallo-beta-lactamases. Antimicrob Agents Chemother **49:** 2778-84

50.     **Aminov RI, Mackie RI.** 2007. Evolution and ecology of antibiotic resistance genes. FEMS Microbiol Lett **271:** 147-61

51.     **D'Costa VM, King CE, Kalan L, Morar M, Sung WW, Schwarz C, Froese D, Zazula G, Calmels F, Debruyne R, Golding GB, Poinar HN, Wright GD.** 2011. Antibiotic resistance is ancient. Nature **477:** 457-61

**Legends of Figures**

**Figure 1. Transmission electron microscopy of negative stained viral particles.** (**A**) Overview of stained viral particles, which vary in size and shape, isolated from the Middle Age coprolite. (**B**) A representative virion and (**C-E**) viral-like particles with icosahedral nucleocapsids and a long filament tail characteristic of *Siphoviridae* bacteriophages.

**Figure 2. (A) The proportion of known and unknown reads (in percent).** Reads were defined as "unknown" if they lacked homology to the non-redundant NCBI database according to a BLASTN search (E-value<1e$^{-05}$) and as "known" otherwise. (**B**) **The relative abundance of viral families.** The relative abundance of identified viral families was estimated using the GAAS software.

**Figure 3. Relative bacteriophage abundance.** The relative bacteriophage abundances were estimated using the GAAS software. The hosts of the bacteriophages that were also identified in a previous study on the bacterial community associated with this specimen (unpublished data (39) are marked with a red point.

**Figure 4. Comparison between the modern human stool viromes and the coprolite virome.** Principal component analysis was used to compare the viral metagenomes associated with the coprolite (highlighted in red) to those associated with modern human stool samples (S1-S21) at the taxonomic (**A**) and functional (**B**) levels.
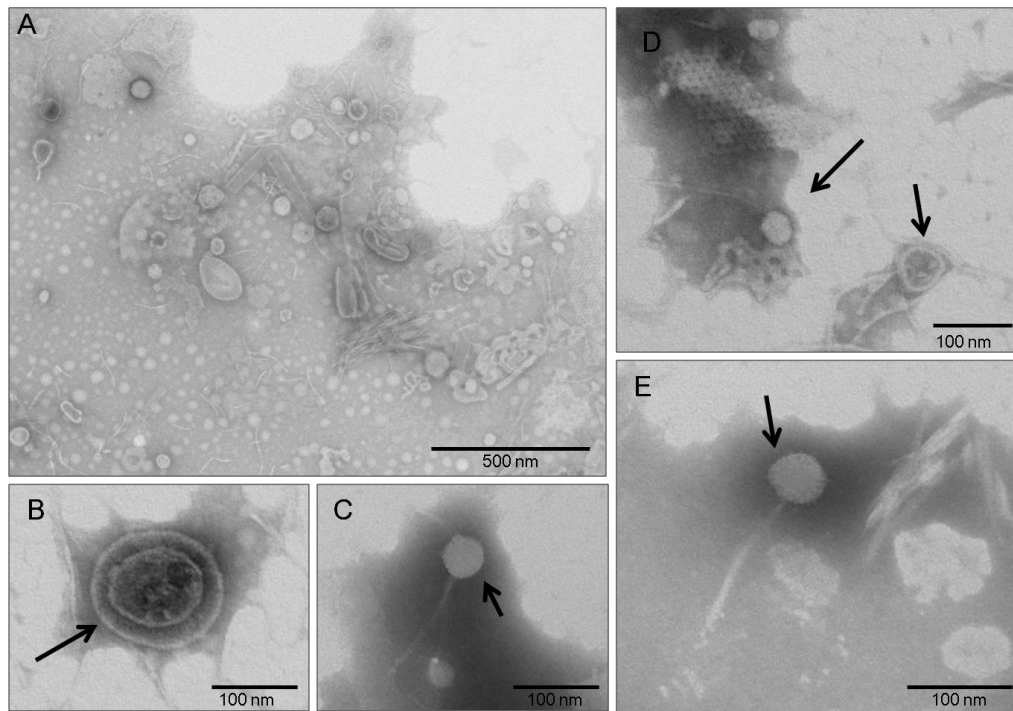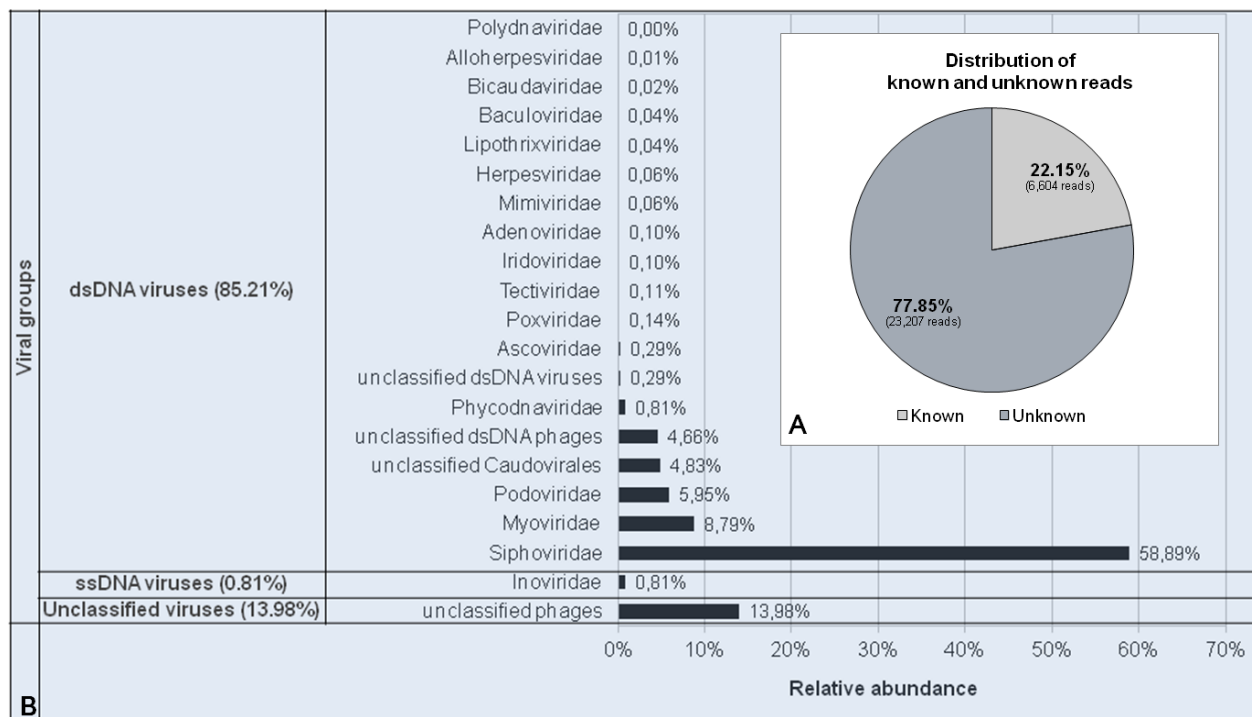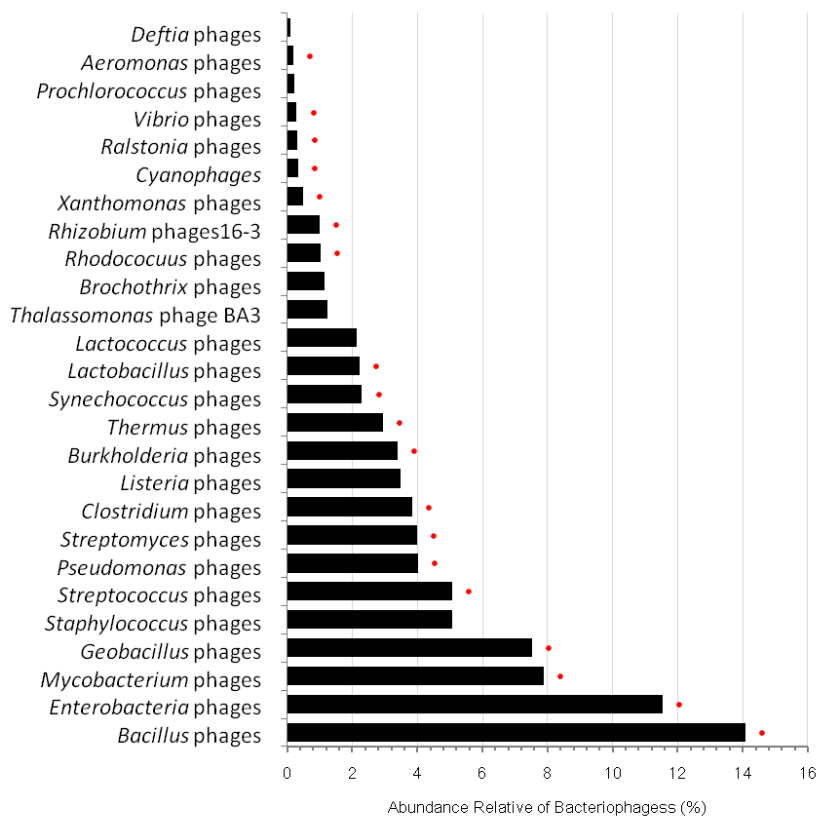
**Figure 1.**
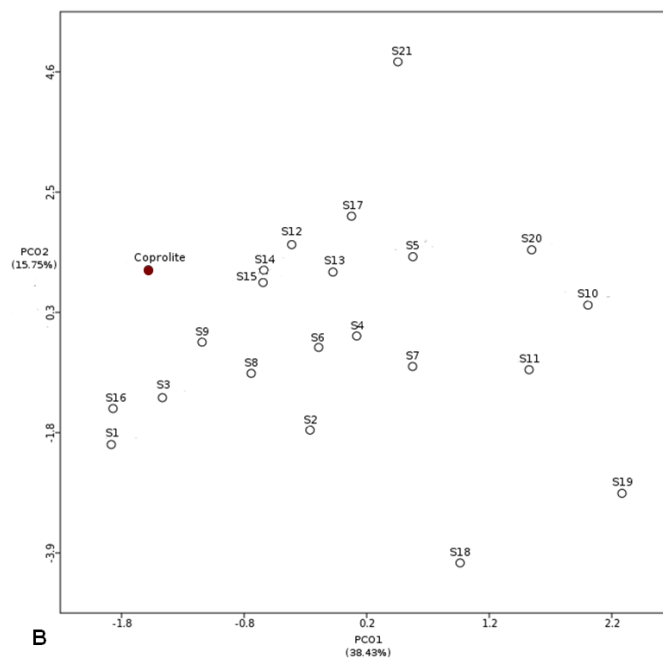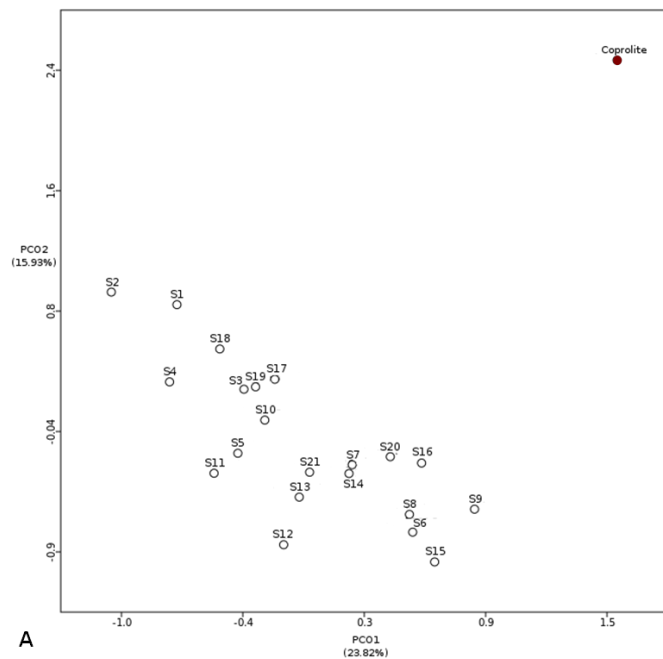
**Figure 2.**

Figure 3.

163

**Figure 4.**

# Viruses in a 14<sup>th</sup>-century coprolite

**Running title:** Viruses in a 14<sup>th</sup>-century coprolite

Sandra Appelt[1,*], Laura Fancello[1,*], Matthieu Le Bailly[2], Didier Raoult[1], Michel Drancourt[1], Christelle Desnues[†,1]

[1] Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, Inserm 1095, 13385 Marseille, France. [2] Franche-Comté University, CNRS UMR 6249 Chrono-Environment, 25 030 Besançon, France.

[*] These authors have contributed equally to this work
† Corresponding author: christelle.desnues@univ-amu.fr

---

**SUPPLEMENTAL MATERIAL CONTENTS**

**SUPPLEMENTARY NOTE**

**SUPPLEMENTARY TABLES**

Supplementary Tables S1-S6

**SUPPLEMENTARY FIGURES**

Supplementary Figures S1-S3

**SUPPLEMENTARY REFERENCES**

**SUPPLEMENTARY NOTE**

**Supplementary Material Section 1.** Prevention of contamination.

After excavation, the coprolite was stored in a sterile forensic specimen bag. In 2006, the coprolite was sent to our laboratory, where it was handled only in a positive pressure room with isolated ventilation under strict aseptic conditions. Workbenches were stringently disinfected using absolute ethanol and were UV irradiated for at least 30 min. Non-disposable instruments were autoclaved. Reagents and chemicals were aliquoted from new stocks into sterile, single-use tubes and immediately discarded after use. The external portion of the coprolite was aseptically removed, and only the internal portion was used in this study. DNA extraction, PCR and post-PCR experiments were performed in separate rooms in isolated work areas. Positive controls were strictly avoided. Negative controls for DNA extraction and PCR were used in a 1:4 control:specimen ratio (1-5).

**Supplementary Material Section 2.** 454 sequencing.

The concentration of the extracted DNA (20.5 ng/μL) was assessed using the Quant-iT PicoGreen kit (Invitrogen) and a Tecan GENios fluorometer. The DNA (500 ng) was nebulized, and a library was constructed according to the 454 Titanium shotgun protocol and the manufacturer's instructions. DNA fragmentation was visualized using the Agilent 2100 Bioanalyzer and a high-sensitivity labchip at an optimal size of 872 bp. The DNA stock concentration was measured using a TBS fluorometer at $2.47e^{+09}$ molecules/μL, and the DNA was stored at -20°C. The library was clonally amplified with 2 cpb in an emPCR reaction using the GS Titanium SV emPCR Kit (Lib-L) v2. The

titration yield was 14.72%. A total of 125,000 beads per project and per region were loaded onto the GS Titanium PicoTiterPlate kit 70x75. The reaction was sequenced using the GS Titanium XLR70 Sequencing kit. The run was performed overnight and then analyzed on the ES Titanium computing cluster.

**Supplementary Material Section 3.** Recovery of total DNA from the coprolite.

Two different DNA extraction protocols were used, including the DNA extraction method recommended by Iniguez et al. (6) and the PowerSoilR DNA Isolation Kit (MoBio Laboratories, Inc., Carlsbad, SA) (7).One gram of the coprolite was solubilized overnight at 4°C in 1 mL of TE buffer (ethylenediamine tetraacetic acid; buffered solution, Tris HCl 10 mM, EDTA 1 mM, pH 8). A 500-μL aliquot of the solution was used for total DNA extraction, as previously described (Iniguez et al. 2006), but the incubations in TE and digestion buffers were shortened to 1 day. TE buffer without coprolite was used as a negative control. For DNA extraction using the PowerSoilR DNA Isolation Kit (MoBio Laboratories, Inc., Carlsbad, USA), 500 μL of solubilized coprolite specimen was also used. Incubation was extended to 24 h at 56°C in PowerSoilR bead tubes containing sodium dodecyl sulfate (SDS) and digestion buffer C1 under contentiously rotation, followed by shaking in a Bio 101 FastPrep instrument (Qbiogene) at level 6.5 (full speed) for 95 s. DNA extraction was performed according to the manufacturer's instructions. Extraction batches also contained negative controls composed of PowerSoilR bead tubes without coprolite. Total DNA extracts from both protocols above were pooled in a 1:1 ratio.

**Supplementary Material Section 4.** Specific PCR amplifications and sequencing.

The PCRs were performed in a 50-µL final volume that included 1x PCR buffer, 2 µL of 25 mM $MgCl_2$, 200 µM of each dNTP, 1 µL of each 10 pM primer, 31.15 µL of $ddH_2O$, 1 unit of HotStar Taq Polymerase (Invitrogen, Villebon sur Yvette, France) and 57-112 ng of DNA extract. The DNA recovered from the coprolite, as described above, served as the template in the reaction targeting giant viruses. Due to DNA template limitations, the GenomiPhi-amplified DNA served as the template in all other reactions. The PCR steps comprised an initial incubation at 95°C for 15 min, 36-40 denaturation cycles at 95°C for 1 min, annealing at the corresponding primer annealing temperatures for 30 sec and elongation at 72°C for 90 sec and a final elongation at 72°C for 10 min; all of these steps were performed in a Gene Amp PCR System 2700 ABI Thermocycler (Applied Biosystems, Villbon sur Yvette, France). The PCR products were analyzed using a 2% agarose gel (UltraPureTM agarose, Invitrogen, Villbon sur Yvette, France) and purified using the QIAquick PCR Purification Kit (QIAGEN, Courtaboeuf, France). All of the PCR products were sequenced in a final volume of 20 µL (1x sequencing buffer, 3.2 pM forward or reverse primer, 4 µL of BigDye Terminator V1.1 mix (Applied Biosystems), 7.4 µL of $ddH_2O$ and 4 µL of PCR product) after purification using Sephadex Gel Filtration in the ABI PR ISM 3130xl genetic sequencer (Applied Biosystems, Villbon sur Yvette, France). The sequences were assembled using Chromaspro software and compared with reference sequences from GenBank using NCBI BLAST searches.

**Supplementary Table S1. Primer systems used for *ad hoc* suicide PCR amplifications[#].**

| Targeted virus or contig ORF | Annotation | Primer | Primer sequence (5' → 3') | Amplicon size (bp) |
|---|---|---|---|---|
| Giant virus | nonfunctional B-family DNA polymerase (giant viruses) | A10F01 A10R01 | AAGGGGACAAGGAGTTAAAATAT TAGATATACGTTTGGTTTTGGAGTGA | 160 |
| ORF1 contig01108 | tail fiber protein (*Cyanophage* S-TIM5) | 1108F 1108R | AATCGGCTGAGGTGAACCAG TTGCCGGTCAGGGTTGTATC | 328 |
| contig00021 range 9-257 | gp160 (*Mycobacterium* phage Optimus) | 21F 21R | GTCTACCCTCGAATTCGCCA GAGAACGGGGTGAGGGAATG | 292 |
| ORF1 contig00261 | gp218 (*Bacillus* phage G) | 261F 261R | TCGGAGGATATGCAGAACGC GCGCATCAGGTAAGGAGGTT | 309 |
| ORF2 contig00101 | putative prophage repressor protein (*Burkholderia* phage BcepC6B) | 101F 101R | TGCTCAGAGGCACGAAACAT AATTATCCGGGCCACTTCCG | 350 |
| ORF1 contig00184 | putative phage tail fiber protein (*Celeribacter* phage P1253L) | 184F 184R | GCCAAGGCTTCCTCATCCAT CATGTACGTCGCATTGTCCG | 282 |
| ORF1 contig00702 | hypothetical protein (*Clostridium* phage D-1873) | 702F 702R | CTCAATTGAACGACAAGGAAGCA ATGTCAGCAGACCAAGCGTC | 262 |

[#]The primers were designed on the basis of ORFs identified on contigs, except for primers targeting giant viruses.

**Supplementary Table S2. Reads before and after preprocessing and the read assembly process output.**

| | |
|---|---|
| **Total number of raw reads** | **30,654** |
| Total number of preprocessed reads | 29,811 |
| Average read length (bp) | 358 |
| Minimal read length (bp) | 77 |
| Maximal read length (bp) | 574 |
| Average GC content | 47% |
| **Total number of contigs** | **1,464** |
| Total number of large contigs | 104 |
| Length of the longest contig (bp) | 12,500 |
| Number of assembled reads | 13,682 |
| Number of partially assembled reads | 6,282 |
| Number of singletons | 9,481 |
| Percentage of known contigs | 17.28% (n=253) |
| Percentage of unknown contigs | 82.78% (n=1,211) |

**Supplementary Table S3. Viral contigs[#].**

| Contig ID | Contig length (bp) | Hit description | E-value | Hit accession ID | Percent ID |
|-----------|--------------------|-----------------|---------|------------------|------------|
| **contig00065** | 1,048 | hypothetical protein MIV033L [Invertebrate iridescent virus 3] | 2e-07 | YP_654605 | 40.79 |
| **contig00021** | 543 | gp160 [*Mycobacterium* phage Optimus] | 5e-13 | AEJ92216 | 45.78 |
| **contig00068** | 1,155 | unnamed protein product [*Synechococcus* phage S-CBS3] | 5e-55 | YP_004421762 | 34.48 |
| **contig00101** | 1,005 | putative prophage repressor protein [*Burkholderia* phage BcepC6B] | 7e-08 | YP_024964 | 48.39 |
| **contig00157** | 1,939 | hypothetical protein 2200_scaffold2278_00035 [unidentified phage] stool sample from healthy person. Minot et al.. 2012 | 9e-25 | AFB75852 | 52.44 |
| **contig00184** | 445 | putative phage tail fiber protein [*Celeribacte*r phage P12053L]gb\|AFM54660.1\| | 2e-13 | YP_006560940 | 66.15 |
| **contig00261** | 926 | gp218 [*Bacillus* phage G] | 3e-26 | AEO93477 | 39.56 |
| **contig00396** | 689 | hypothetical protein [uncultured phage MedDCM-OCT-S04-C136] | 3e-06 | ADD94244 | 35.71 |
| **contig00474** | 1,719 | gp27 [*Bacillus* phage G] | 7e-07 | AEO93298 | 28.98 |
| **contig00611** | 488 | phage terminase. large subunit. PBSX family [*Thermoanaerobacterium* phage THSA-485A] | 2e-26 | YP_006546303 | 49.64 |
| **contig00629** | 463 | putative nuclease SbcCD D subunit [*Bacillus* phage BCP78]gb\|AEW47679.1 | 6e-07 | AEW47191 | 35.45 |
| **contig00656** | 402 | gp128 [*Bacillus* phage G] | 1e-22 | AEO93390 | 68.42 |
| **contig00664** | 477 | hypothetical protein OSG_eHP6_00230 [Environmental *Halophage* eHP-6] | 1e-07 | AFH21679 | 56.00 |
| **contig00702** | 554 | conserved hypothetical protein [*Clostridium* phage D-1873]gb\|EES90342.1\| | 2e-19 | ZP_04863741 | 46.94 |
| **contig00926** | 473 | hypothetical protein [*Riemerella* phage RAP44] | 2e-15 | AEB71650 | 44.19 |
| **contig01108** | 731 | tail fiber protein [*Cyanophage* S-TIM5] | 4e-07 | AEZ65603 | 36.27 |
| **contig01456** | 734 | DNA polymerase [Bacteriophage APSE-4] | 6e-28 | ACJ10154 | 34.80 |
| **contig00075** | 3,946 | phage tape measure protein [*Planctomyces limnophilus* DSM 3776] | 9e-62 | YP_003632269 | 58.99 |
| **contig00078** | 8,733 | phage major capsid protein. HK97 family [*Clostridium cellulovorans* 743B] | 4e-100 | YP_003842316 | 46.20 |
| **contig00073** | 436 | phage-related protein-like protein [*Acetivibrio cellulolyticus* CD2] | 1e-07 | ZP_09464888 | 41.79 |

| | | | | | |
|---|---|---|---|---|---|
| **contig00076** | 2,218 | phage tail tape measure protein. TP901 family [*Thermoanaerobacter italicus* Ab9] | 2e-54 | YP_003477199 | 64.49 |
| **contig01464** | 1,248 | putative DNA-binding phage protein [*Burkholderia cenocepacia* J2315] | 2e-18 | YP_002234532 | 39.44 |
| **contig00055** | 639 | phage-like protein [*Thermoanaerobacterium saccharolyticum* JW/SL-YS485] | 1e-06 | YP_006393008 | 41.54 |
| **contig00048** | 6,115 | phage Gp37Gp68 [*Mycobacterium massiliense* 1S-151-0930] | 3e-75 | EIU63155 | 47.46 |
| **contig00230** | 7,394 | TP901 family phage tail tape measure protein [*Thermosipho melanesiensis* BI429] | 7e-10 | YP_001306681 | 53.33 |
| **contig00143** | 1,495 | Prophage LambdaBa04. site-specific recombinase. phage integrase [*Bacillus cereus* AH621] | 3e-22 | ZP_04293086 | 39.16 |
| **contig00181** | 3,804 | phage-like protein [*Bacillus licheniformis* WX-02] | 1e-66 | EID48485 | 63.45 |

[#]The contig identifier, length (bp) and annotation according to the best BLAST hit (BLASTX *versus* the non-redundant NCBI database, E-value<1e[-05]) are summarized. The E-value, hit accession identifier and percent identity are also provided.

**Supplementary Table S4. Viral metagenomes used for comparison, illustrated in Fig. 4.**

| Identifier | Sample name | Sample type | Species richness (effective number of species) | Functional diversity (Shannon-Wiener index) | Reference |
|---|---|---|---|---|---|
| S1 | F-A | Modern stool | 138.760 | 4.6275 | [8] |
| S2 | F-B | Modern stool | 141.912 | 4.1130 | [8] |
| S3 | F-C | Modern stool | 127.776 | 4.7134 | [8] |
| S4 | F-D | Modern stool | 20.232 | 3.9289 | [8] |
| S5 | F-E | Modern stool | 42.160 | 4.1415 | [8] |
| S6 | X-1 | Modern stool | 112.061 | 4.4603 | [9] |
| S7 | L1-1 | Modern stool | 67.221 | 4.1169 | [9] |
| S8 | L1-8 | Modern stool | 99.298 | 4.4523 | [9] |
| S9 | L2-2 | Modern stool | 121.056 | 4.6202 | [9] |
| S10 | L2-7 | Modern stool | 65.893 | 3.4547 | [9] |
| S11 | L2-8 | Modern stool | 40.181 | 3.2639 | [9] |
| S12 | H1-1 | Modern stool | 130.359 | 4.0687 | [9] |
| S13 | H1-2 | Modern stool | 118.034 | 4.1585 | [9] |
| S14 | H1-7 | Modern stool | 60.617 | 4.4489 | [9] |
| S15 | H1-8 | Modern stool | 105.572 | 4.5599 | [9] |
| S16 | H2-1 | Modern stool | 56.515 | 4.7019 | [9] |
| S17 | H2-8 | Modern stool | 76.075 | 4.0324 | [9] |
| S18 | L3-1 | Modern stool | 32.353 | 4.2648 | [9] |
| S19 | L3-2 | Modern stool | 36.772 | 3.1021 | [9] |
| S20 | L3-7 | Modern stool | 21.855 | 3.6276 | [9] |
| S21 | L3-8 | Modern stool | 19.611 | 3.7963 | [9] |
| Coprolite | Coprolite | Middle Age stool | 315.279 | 4.8693 | Present study |

**Supplementary Table S5. Significant matches to virulence factors associated to pathogenic bacteria** (BLASTX of metagenomic reads against the Virulence Factor Database). The bacterial genera and the number of matching reads are reported.

| Genera | Number of hits |
|:---:|:---:|
| *Escherichia* | 42 |
| *Salmonella* | 39 |
| *Shigella* | 34 |
| *Yersinia* | 13 |
| *Pseudomonas* | 11 |
| *Listeria* | 8 |
| *Legionella* | 7 |
| *Mycobacterium* | 5 |
| *Vibrio* | 3 |
| *Bordetella* | 2 |
| *Neisseria* | 1 |
| *Haemophilus* | 1 |
| **Total** | **166** |

**Supplementary Table S6. Statistically significant differences in COG pathways between the coprolite virome and modern stool viromes.** The first 15 significantly different COG pathways as well as the associated p-values (adjusted for multiple comparisons) are reported. The COG pathways, which are overrepresented in the coprolite virome, are indicated in bold.

| COG pathway | Adjusted p-value |
| --- | --- |
| General function prediction only | 1.76853150434351e-85 |
| Thymidylate biosynthesis | 6.48788404441294e-69 |
| **Energy production and conversion** | **2.8107332246192e-40** |
| **Lipid transport and metabolism** | **1.04271960007454e-39** |
| **Amino acid transport and metabolism** | **2.26843518010469e-35** |
| Basal replication machinery | 7.15265582902521e-30 |
| **TCA cycle** | **1.26195281556932e-26** |
| Nucleotide transport and metabolism | 3.14695805939879e-25 |
| Function unknown | 4.69221917501543e-25 |
| **Fatty acid biosyntheis** | **7.80050284274998e-25** |
| **Defense mechanisms** | **3.68599997778464e-18** |
| **Translation, ribosomal structure and biogenesis** | **2.17547655548738e-17** |
| **Heme biosynthesis** | **1.05366947514799e-14** |
| **Glyoxylate bypass** | **1.01093235970868e-13** |
| Cell cycle control, cell division | 1.17938484034507e-12 |

**Legends of Supplementary Figures**

**Supplementary Figure S1. GC content and read length distributions.** (**A**) The GC content distribution of preprocessed metagenomic reads. The y-axis indicates the number of reads, whereas the x-axis indicates (**B**) the read length distribution. The y-axis indicates the number of reads, and the x-axis indicates the length (in base pairs).

**Supplementary Figure S2. Analysis of temperate bacteriophages.** The Venn diagram represents the frequency of indicators associated with a temperate lifecycle in the assembled viral contigs (9): presence of integrases-encoding genes ("integrases," blue) and prophage proteins ("prophage proteins," red) and similarity to bacterial genomes ("bacterial genomes," green). The number of contigs with each indicator is reported, as is the total number of contigs.

**Supplementary Figure S3. Phylogenetic tree of a chloramphenicol O-acetyltransferase.** A phylogenetic tree was generated from the translated open reading frame of a contig encoding a chloramphenicol O-acetyltransferase. The tree was constructed using the PhyML algorithm with a bootstrap of 100. The bootstrap support is reported for each branch.
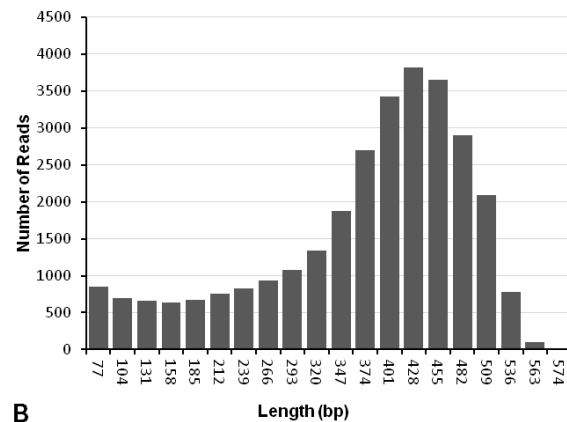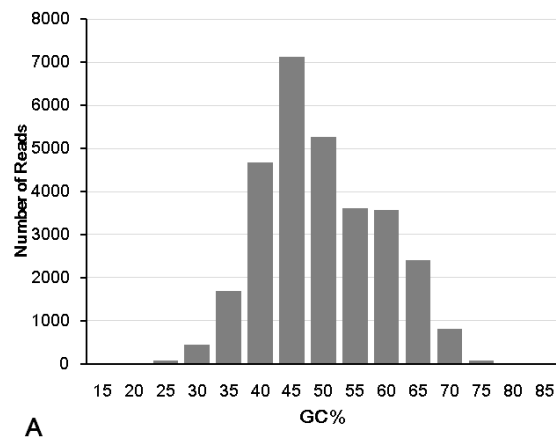
**Supplementary Figure S1.**
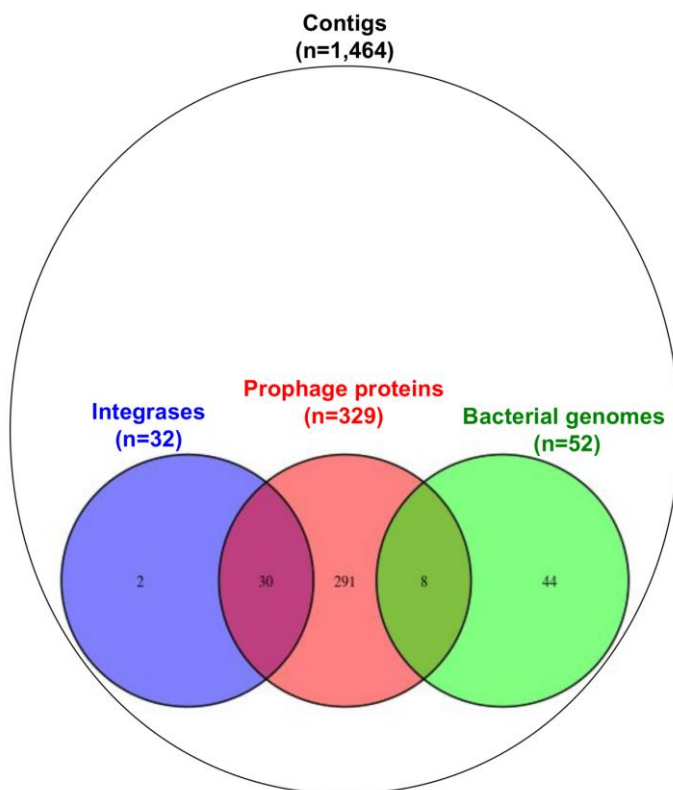
**Supplementary Figure S2.**
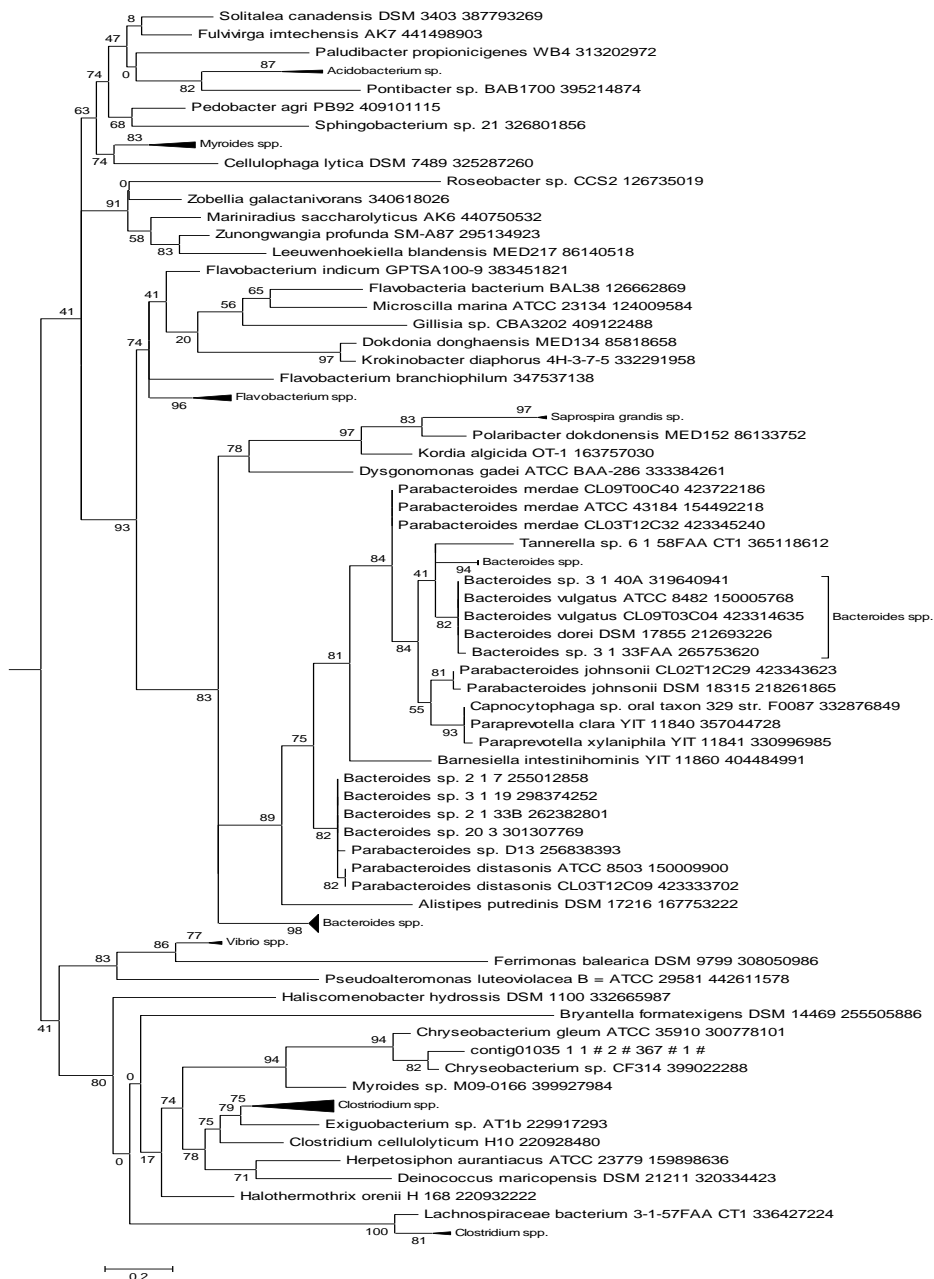
**Supplementary Figure S3.**

**SUPPLEMENTARY REFERENCES**

1.     **Cooper A, Poinar HN.** 2000. Ancient DNA: do it right or not at all. Science 289: 1139

2.     **Willerslev E, Cooper A. 2005.** Ancient DNA. Proc Biol Sci **272:** 3-16

3.     **Hofreiter M, Serre D, Poinar HN, Kuch M, Paabo S.** 2001. Ancient DNA. Nat Rev Genet **2:** 353-9

4.     **Pääbo S, Poinar H, Serre D, Jaenicke-Despres V, Hebler J, Rohland N, Kuch M, Krause J, Vigilant L, Hofreiter M.** 2004. Genetic analyses from ancient DNA. Annu Rev Genet **38:** 645-79

5.     **Drancourt M, Raoult D.** 2005. Palaeomicrobiology: current issues and perspectives. Nat Rev Microbiol **3:** 23-35

6.     **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:** 1419-25

7.     **Tito RY, Macmil S, Wiley G, Najar F, Cleeland L, Qu C, Wang P, Romagne F, Leonard S, Ruiz AJ, Reinhard K, Roe BA, Lewis CM, Jr.** 2008. Phylotyping and functional analysis of two ancient human microbiomes. PLoS One **3:** e3703

8.     **Kim MS, Park EJ, Roh SW, Bae JW.** 2011. Diversity and abundance of single-stranded DNA viruses in human feces. Appl Environ Microbiol **77:** 8062-70

9.     **Minot S, Sinha R, Chen J, Li H, Keilbaugh SA, Wu GD, Lewis JD, Bushman FD.** 2011. The human gut virome: inter-individual variation and dynamic response to diet. Genome Res **21:** 1616-25

# Chapitre IV

ARTICLE

## Culture d'*Acanthamoebae* et d'un endosymbionte associés à un coprolithe de Moyen Âge

Sandra Appelt[1], Morgan Gaia[1], Matthieu Le Bailly[2], Bernard La Scola[1]*, Michel Drancourt[1]*

[1] Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, Inserm 1095, 13005 Marseille, France. [2] Franche-Comté University, CNRS UMR 6249 Chrono-Environment, 25 030 Besançon, France.

*Corresponding author. Michel.Drancourt@medicine.univ-mrs.fr

Article en préparation.

## Avant-propos

Comme décrit dans l'étude **Chapitre II**, des kystes d'amibes ont été observés par microscopie dans le coprolithe de Namur. Ces observations ont été confirmées par la détection dans cet échantillon d'une séquence ADN spécifique de l'espèce *Acanthamoeba* aprés microdissection laser des kystes d'amibes. Des études antérieures controversées ont montré qu'il était possible de cultiver des micro-organismes à partir d'échantillons vieux de plusieurs siècles (Christine Helms et al. 2004; Kennedy 1994; Khanaeva et al. 2013; La Farge et al. 2013; Lambert et al. 1998; Redmond et al. 1998; Rhodes et al. 1998; Vreeland et al. 2000; Willerslev et al. 2004; Yashina et al. 2012). Même si certains de ces travaux sont sujets à controverses de part la suspicion de contamination de l'échantillon par l'environnement du laboratoire (Satterfield et al. 2005; Willerslev et al. 2005), néanmoins certains micro-organismes sont capables de survivre pendant plusieurs décades en adoptant des formes persistantes. Il a ainsi été montré que des graines de plantes (*Silene stenophylla,* des espèces de Bryophytes) avaient conservé leur capacité germinative après une période de dormance de 30 000 ans (La et al. 2013; Yashina et al. 2012); certains protozoaires (*Acanthamoeba,* des espèces de levure) forment des kystes pouvant survivre plus d'une vingtaine d'années sans perte de leur viabilité (Kennedy 1994; Mazur et al. 1995; Sriram et al. 2008) et certaines bactéries sporulées (*Bacillus*, *Clostridium*) survirent plusieurs décennies, voire des siècles (Christine Helms et al. 2004; Khanaeva et al. 2013; La et al. 2013; Redmond et al. 1998; Rhodes et al. 1998). En particulier, des travaux récents ont montré que des bactéries sporulantes et non-sporulantes associées à des échantillons du permafrost vieux de 500 mille ans, possédaient des mécanismes de réparation de l'ADN encore actifs pouvant être corrélés à leur persistance (Johnson et al. 2007; Lewis et al. 2008; Price et al. 2004). Ceci témoigne des limites actuelles des connaissances des mécanismes impliqués dans la survie à long-terme des micro-organismes. Sur ces bases, nous avons cultivé des amibes à partir de kystes du coprolithe de Namur, permettant leur transition en état trophozoïte, ce qui a offert l'opportunité de découvrir une nouvelle

espèce bactérienne, endosymbionte de l'amibe ; et de séquencer le génome de cette amibe.

# JOURNAL ARTICLE

## Growing *Acanthamoeba* out of medieval coprolite reveals a new endosymbiont
**Running title:** *Acanthamoeba* and Endosymbiont out of coprolite

Sandra Appelt[1], Morgan Gaia[1], Matthieu Le Bailly[2], Bernard La Scola[1]*, Michel Drancourt[1]*

[1] Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, Inserm 1095, 13005 Marseille, France. [2] Franche-Comté University, CNRS UMR 6249 Chrono-Environment, 25 030 Besançon, France.

*Corresponding author. Michel.Drancourt@medicine.univ-mrs.fr

**Keywords: coprolite, Amoeba, *Acanthamoeba* endosymbiont**
Article in preparation.

---

## Abstract

*Acanthamoeba* have been grown from soil, water, cooling towers, human and animal speciemens, but have never been investigated in archeological specimens. Microscopic examination of a Middle Ages coprolite excavated from an archeological site in Namur, Belgium, yielded amoebal cysts. After micros-dissection, cysts were grown and identified through 18S rDNA sequencing as *Acanthamoeba*. The amoeba also revealed the presence of an amoebal-resistant endosymbiont possibly associated to *Rickettsiales*, yet exhibiting bipolar immobile flagella. This is the first report of successful cultures of protozoa and associated endosymbiont isolated from a five-century-old specimen.

**INTRODUCTION**

The lifespan greatly varies among living organisms with some organisms persisting in form of seeds, spores and cysts (La Farge et al. 2013; Lewis et al. 2008; Mazur et al. 1995). Such persisting stages typically show limited metabolic activity, preserving living organisms for decades when living conditions are not favorable. Likewise, dormant bacteria trapped inside permafrozen samples had surprisingly low metabolic activities indicative of DNA repair (Johnson et al. 2007; Krisko and Radman 2010; Lewis et al. 2008; Price and Sowers 2004). Indeed, plant seeds were grown after more than 30,000 years stay in permafrozen samples and bacteria might have survived inside ancient samples for decades and even for centuries (Christine Helms et al. 2004; Khanaeva et al. 2013; La Farge et al. 2013; Redmond et al. 1998; Rhodes et al. 1998; Yashina et al. 2012).

Cysts such as those of amoeba are also known to be highly resistant to extreme temperatures, desiccation and disinfections (Coulon et al. 2010; Sriram et al. 2008). Free-living encysted amoeba can stay persistent for at least twenty years (Mazur et al. 1995; Sriram et al. 2008). In particular several studies showed that *Acanthamoeba* cysts survived a period of twenty to twenty-four years stored at 4℃ in water or in a desiccated stage without losing pathogenicity (Mazur et al. 1995; Sriram et al. 2008). Adverse environmental conditions induce changes in its lifecycle form vegetative (trophozoites, 12-35 μm) to resistant forms (cysts, 10-20 μm) (WHO 2003). Today little is known about the maximal viability of amoeba cysts and associated viruses or bacteria. Amoeba have a potential of pathogenicity, mainly due to their association with amoeba-resistant

bacteria, such as *Mycobacterium* spp. or *Legionnella* spp. (Greub 2009; La Scola et al. 2000; Thomas et al. 2008), responsible of respiratory tract infections.

Coprolites are fossilized fecal material of human or animal origin that lived centuries ago (Araujo et al. 1998). As the fossilization process is long, coprolites can also incorporate external microorganisms after defecation to the end of desiccation and thus can reflect ancient environments (Wood et al. 2012; Yll et al. 2006). During investigations performed on a 14th century coprolite specimen from Namur, Belgium, amoebal cysts were observed. The observations were completed by culture and molecular test. Herein, the excystation of this ancient amoeba is reported, as well as the isolation and identification of an associated new amoeba-resistant bacterium.

## RESULTS

### Microscopic observation and Culture of amoeba

Cellulosic components of the Namur coprolite material were stained using Congo-red and observed by light-microscopy (**Figure 1**). Among stained materials, cysts were observed with sizes varying from 4.1 μm to 13.5 μm, leading to the presumption of possible amoebal cysts (Khan 2007). The amplicons generated by *Acanthamoeba* spp.-specific PCR (Schroeder et al. 2001) performed on the coprolite DNA extract in the presence of negative controls, futher yielded a sequence with 99% of similarity to *Acanthamoeba castellanii* sequence (GenBank Accession No. JF437606.1) (**Supplementary Table S1**). Additionally, on the bias of laser-microdissection, 25 of these cysts were collected and DNA was extracted. In the presence of negative controls, *ad hoc* suicide PCR-amplifications (Raoult et al. 2000) using *Acanthamoeba* spp.

specific primer yielded positive results. The amplicons yielded 95% to 100% identity to *Acanthamoeba* spp. (*A. castellanii* GenBank Accession No. JF437606 and *Acanthamoeba polyphaga* (GenBank Accession No. AF260724). Accordingly, throughout the attempt to culture the amoeba – using different conditions (**Supplementary Figure S1**) – three different *Acanthamoeba* spp. were cultured. Amoebal vegetative cells were observed onto gelose plates after seven days of incubation. All negative culture controls incorporated during experimentation remained sterile.

**Identification of amoeba**

Molecular identification by *ad hoc* PCR amplifications and sequencing of the DNA extracted from the cultured amoebas indicated that they belong to the genus *Acanthamoeba*. Amplicons generated for cultured amoeba A and B yielded 100% identity and coverage with *A. polyphaga* and *A. castellanii* species (GenBank accession No., JX983592.1; KC164234.1; KC164230.1; KC164226.1), respectively. The amplicon generated from cultured amoeba BF yielded 99% identity and 98% coverage with *A. castellanii* (GenBank accession No., U07413.1; KC164234.1) (**Supplementary Table S1**). Genotyping further indicated that the three cultures belong to the genotype clade T4 of *Acanthamoeba*. Furthermore, the sequences were compared to those obtained for *Acanthamoeba* spp. detected and handled in the laboratory, including *A. polyphaga* AP1 (Lab strain 1), *A. castellanii* Neff (Lab strain 2), *Acanthamoeba* sp. lentille 1 (Lab strain 3), *Acanthamoeba* sp. sputum (Lab strain 4) and *Acanthamoeba* sp. lentille 2 (Lab strain 5). The comparison of the sequences, and the phylogenetic construction based on the

18S rDNA gene confirmed that the three amoeba cultured from the Namur coprolite did not match any of the five *Acanthamoeba* spp. isolates handled in the laboratory excluding in-laboratory contaminations (**Supplementary Figure S2**). To go further, additional protein spectra of the three cultured amoeba and of the laboratory strains were created by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS). Unique profiles were obtained for each *Acanthamoeba* (see **Supplementary Figures S3** and **S4**). All negative controls remained negative and the strain-specific profiles were reproducible over 3 times. The MALDI-TOF MS profile of each cultured amoeba showed minor derivations. The dendrogram created on the basis of these individual protein spectra showed that the cultured amoeba strain A could be associated to laboratory strains, yet different (**Figure 2**).

**Endosymbiont**

The attempt to cultivate amoeba-resistant bacteria out of the cultured amoeba B yielded a positive result. After five-day incubation, growth of an *Acanthamoebae* endosymbiont was observed. Transmission electron microscopy showed that this endosymbiont was bacillary with bipolar flagella (**Figure 3**). However, no movement of the flagella was observed by light-microscopy. Molecular identification based on 16S rDNA sequencing yielded 100% of coverage and 99% of identity to an uncultured endosymbiont of *Acanthamoeba* spp. (GenBank Accession No. AF239294.1). With respect to the source material – a coprolite – and the place of culture – Marseille –, the endosymbiont was named *Coprolita marseillensis* (GenBank Accession No. JQ409353.1). Phylogenetic investigations indicated that the *C. marseillensis* clustered as a sister group closely related to *Rickettsiale*s bacteria (**Figure 4**).

204

**DISCUSSION**

Here, three different amoeba of the genus *Acanthamoeba* were grown from a medieval coprolite excavated from an archeological site in Belgium. These amoeba were certainly present in the coprolite itself and do not result from laboratory contamination. Indeed, cysts compatible with amoebal cysts were microscopically observed in the coprolite, an were directly excited for further analyses by microdissection. Indeed, the re-cultivations of these microorganisms might give access to living organisms from ancient samples. Here, the first isolation and presumptive characterization of amoebas successfully cultured from a coprolite is reported. Furthermore, a new amoeba-resistant bacterium was isolated and identified by amoeba-co culturing out of one of these amoebas. Dating the amoeba is indirect, relying on the direct observation of cysts in a 5 centuries-old coprolite. Thus, extensive precautions were required in order to avoid any contamination of the coprolite through the laboratory environment. Accordingly, all used negative controls have been in agreement with current recommendations for paleomicrobiology and paleoparasitology studies, and remained negative in both DNA – and culture-based experiments (Cooper and Poinar 2000; Drancourt and Raoult 2005; Hebsgaard et al. 2005; Hofreiter et al. 2001; Pääbo et al. 2004; Willerslev and Cooper 2005) and only the internal region of the Namur coprolite was used for experimentation. Furthermore, the Namur coprolite was recovered found the inside of a sealed barrel buried at a depth of 3.80 m, suggesting an absence of contamination by surrounding environment for centuries and the anaerobic conditions were favorable for preservation of the material. Yet, it is difficult to estimate the part of material exchanged with environment before sealing the barrel, but all evidences indicate that the described results are authentic.

Moreover, the presence of amoebas in the coprolite was confirmed by three independent tests microscopic observations, molecular detection, and culture. The sequences analyzed were original and never been detected before in the laboratory. Accordingly, the micro-organisms identified are not consequence of laboratory contaminations but were indeed present in the specimen at the time it was excavated. Moreover, analysis of sequences derived from the amoeba, after culturing and axenisation, yielded 100% of identity to *A. polyphaga* and *A. castellanii* species that were previously isolated in another laboratory from compost of composting facilities. The composition of compost might be quiet similar to those of the archeological Namur coprolite sample, because both are composed of a mixture of organic matter including feces, and soil, and this gives credit to the results. Identification of the isolated amoebas indicated that they belong to the genotypic clade T4, commonly associated to amoeba infections such as keratitis cases (Booton et al. 2004). Nevertheless, the implication of these amoebas in ancient populations is difficult to determine, as amoebas of the *Acanthamoeba* genus are ubiquitously found in various environments (Lorenzo-Morales et al. 2007; Magnet et al. 2012; Schuster and Visvesvara 2004; Siddiqui and Khan 2012; Visvesvara et al. 2007; Walochnik et al. 1999; Wildschutte and Lawrence 2007). The pathogenic impact of *Acanthamoeba* spp. on human populations is often related to associated microorganisms. When environmental conditions are not favorable for the amoeba growth, they can adopt a dormant state in forms of cysts until the conditions are changing (WHO 2003). Microorganisms such as *Legionella* spp.; *Mycobacteria* spp. or viruses which are able survive to phagocytosis of amoebas can thus benefits from this resistance of amoeba cysts (Greub 2009; La Scola et al. 2000; Thomas et al. 2008). Several studies have

shown that *Acanthamoeba* cysts are able to survive for 24-years of storage at 4°C in water without losing pathogenicity and even desiccation for up to 21 years (Mazur et al. 1995; Sriram et al. 2008). Until now, amoeba resistant microorganisms have not been investigated in cysts conserved for a long time. The possibility of survival of endosymbionts inside an ancient cyst has thus to be considered. Cultivation of one of the found amoebas allowed observation and then isolation of an *Acanthamoeba* endosymbiont corresponding to a bi-flagellated bacillary bacterium. Phylogenetic investigations indicated that it might be a member of the *alpha-Proteobacteria* class that formed an independent lineage related to *Rickettsiales*. Sequences of its 16S rDNA gene were also previously detected in clinical *Acanthamoeba* spp. isolates (Fritsche et al. 1999). The *Rickettsiales* were thought to be without flagella, until *Candidatus Midichloria mitochondrii* was identified. *Candidatus Midichloria mitochondrii* possess and express some genes that are implicated in the formation of flagella; even if no flagella can be observed (Mariconti et al. 2012; Sassera et al. 2008). Interestingly the herein cultivated endosymbiont owns bipolar flagella apparently immotile, and thus could represent an evolutionary intermediate.

**CONCLUSION**

In conclusion the investigations were focused on cultivation of microorganisms potentially conserved inside fossils. This study is the first report of amoebas and associated resistant bacterium cultivation coming from an ancient latrine sample. The specimen was dated to medieval times and definitive confirmation of the isolated protozoan age is complicated to estimate. However, all evidences provided suggest an advanced age of the cultivated amoebas. Thus, they could be the oldest *Acanthamoebae*

alive. Moreover, by amoebal co-culture an amoebal-resistant endosymbiont was identified and isolated by amoebal co-culture. The endosymbiont is associated to *Rickettsiales*, yet exhibiting bipolar immobile flagella. This is the first report of successful protozoa culture and associated endosymbiont isolation from a centuries old sample.

## MATERIAL AND METHODS

### Nature of the coprolite.

The 14th century coprolite specimen was collected from an ancient latrine – i.e. a mix of feces of different individuals – in 1996 from a Middle Ages site in Namur, Belgium. The Namur coprolite was found in the interior of a closed barrel stacked at a depth of 3.80 m, such as those commonly used in this period as pits or latrines. The 121.4 g coprolite specimen was dark brown, well preserved under anaerobic conditions and was described as a mix of soil and organic matter. During previous investigations cellulosic components of the Namur coprolite were stained using Congo-red and observed by light-microscopy, revealing the presence of amoeba cysts (Appelt et al. 2013).

### Prevention of contaminations.

The coprolite was handled only in a positive pressure room with isolated ventilation under strict aseptic conditions. Workbenches were stringently disinfected using absolute ethanol and UV-irradiation for at least 30 min. Non-disposable instruments were autoclaved. Reagents and chemicals were from new stocks aliquoted into sterile, single-use tubes and immediately discarded after use. The external portion of the coprolite was aseptically removed, only the internal portion was used in this

study. DNA -extraction, PCR and post-PCR experiments were performed in separate rooms in isolated work areas. Positive controls were strictly avoided. Negative controls were used in a 1:4 control: specimen ratio (Cooper and Poinar 2000; Drancourt and Raoult 2005; Hebsgaard et al. 2005; Hofreiter et al. 2001; Pääbo et al. 2004).

**Microscopy of stained coprolite material and laser dissection.** The cellulosic material (cysts, plant fibers, seeds) of the coprolite was stained using Congo-red as previously discribed (Appelt et al. 2013). Laser dissection of cysts was then performed using silicone membrane-coated slides (Molecular Machines & Industries, Aartselaar, Belgium), a Nikon ECLIPS TE2000-U microscope (Nikon Instruments, Champigny sur Marne, France) and a mmi CellcutR laser and its controlling program mmi cell tool (Molecular Machines & Industries, Aartselaar, Belgium).

**Culturing of *Acanthamoeba*.**

Two times a 500 mg portion of the interior region of the Namur coprolite was solubilized in 1mL of sterile PAS (Page's Amoeba Saline medium). To avoid fungal growth, 50 µL of amphotericin B (10 µg/mL) were added to one tube of 1 mL suspension.

For culturing, non-nutritive agar plates (Biotechnologie Appliquée) were over-coated with a thin film of living *Enterobacter aerogenes* cells. Then, 50 µL of the coprolite suspension were immediately plated onto a plate for the culture A, without amphotericin B. For the culture B (without amphotericin B) and BF (with amphotericin B) the coprolite solutions were incubated for 4 days at 4°C before being plated (**Supplementary Fig. S5**). Negative controls consisted of PAS without coprolite

material, added in a ratio 1:1. The cultures were incubated at 28 °C and daily observed by light-microscopy. To axenisate the cultivated *Acanthamoeba* spp., three subcultures were carried-out using each time a single clone, first on non-nutritive agar plates with living *E. aerogenes* cells, then on 3 successive non-nutritive agar plates with UV-irradiated (2 hours) dead *E. aerogenes* cells and finally in sterile liquid proteose-yeast-glucose (PYG) growth media (Biotechnologie Appliquée). When a visible growth was observed, the PYG-growth medium was changed. After 1-2.5 month of procedure, amoebas were axenised and used for molecular identification and MALDI-TOF-MS.

**Culturing and Identification of *Acanthamoeba* endosymbiont.**

To culture the amoeba endosymbiont, 7 g of coprolite material were solubilized in 25 mL of ddH$_2$O and incubated over night at 4 °C. The supernatant was

filtered with a KNF Neuberger Labor ramp filter (Neuberger Laboport, Freiburg-Munzingen, Germany) using 0.1 µm Millipore Isopore$^{TM}$ membrane filters (Millipore, Molsheim, France) and used for amoeba-co culturing as previously described (Pagnier et al. 2008). The supernatant of the culture was then used for transmission electron microscopy observations after negative staining using 1.5 % molybdate solution or after inclusion in acryl-resin. The supernatant of the culture was furthermore used for DNA-extraction.

**DNA-Extraction.**

DNA of the coprolite was extracted as described previously (Iniguez et al. 2006; Tito et al. 2008). As for the isolation of DNA out of cysts, collected via microdissection, a standard Phenol-Chloroform DNA extraction protocol was used. DNA extractions out of cultured *Acanthamoeba* spp. including the cultured endosymbiont, and all ever handled

210

or used *Acanthamoeba* spp. of the laboratory, were performed using the QIAampR DNA Extraction Mini Kit (QIAGEN, Courtaboeuf, France), following the supplier's instructions. Each extraction included several extraction-blanks consisting of sterile water.

**Molecular Analysis.**

*Ad hoc* PCRs protocols were applied to amplify the 18S rRNA gene region of *Acanthamoeba spp.* (Schroeder et al. 2001). Amplicons were generated from the DNA extract from the coprolite using suicide PCR amplification (Raoult et al. 2000) and then from the cultured *Acanthamoeba* spp. and all *Acanthamoeba* spp. of the laboratory. The PCRs were performed in a final volume of 50-µL containing 1x PCR buffer, 2 µL of 25 mM MgCl2, 200 µM of each d'NTP, 1 µL of 10 pM of each primer, 31.15 µL of ddH2O, 1 unit of HotStar Taq Polymerase (Invitrogen, Villebon sur Yvette, France) and 57-112 ng of DNA-extract. The steps of the PCR were an initial incubation at 95 °C for 15 min; 38 cycles of denaturation for 95 °C for 1 min, annealing for 45 sec at the corresponding primer annealing temperatures, and elongation at 72 °C for 90 sec and a final elongation 72 °C for 10 min; all of these steps were performed in a Gene Amp PCR System 2700 ABI Thermocycler (Applied Biosystems, Villbon sur Yvette, France). The PCR products were analyzed using a 2 % agarose gel (UltraPureTM agarose, Invitrogen, Villbon sur Yvette, France) and purified using the QIAquick PCR Purification Kit (QIAGEN, Courtaboeuf, France). The sequences were assembled using the ChromasPro software and compared with reference sequences of the GenBank database using NCBI BLAST searches. To generate the genotype depending on *Rns* profiles the sequences

211

were aligned and compared to the *Acanthamoeba* genotype names stored in the NCBI Gen Bank database. The same amplicons were also used for phylogenetic trees constructions. Multiple sequence alignment was performed using MUSCLE (Edgar 2004) and gaps were removed. The phylogenetic tree was built using the PhyML algorithm (Guindon et al. 2009) with a bootstrap of 100 and the nucleotide substitution model HKY85 (Hasegawa et al. 1985). These tasks were all performed using the pipeline www.phylogenie.fr (Dereeper et al. 2008). The phylogenetic trees were visualized using MEGA v.5 (Tamura et al. 2011). To identify the amoeba-endosymbiont, the supernatant of the culture was used for DNA extraction as described above, and the 16S rRNA gene was sequenced. For phylogenetic analysis, multiple sequence alignment was performed using MUSCLE (Edgar 2004) and curated by Gblocks (Talavera and Castresana 2007). The phylogenetic tree was built according to the protocol described above for amoeba's identification.

**Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry.**

1 mL of *Acanthamoeba* culture suspension was washed twice with distilled water. Then, 1.5 µL of the suspensions were spotted onto a 96 MSP target plate (Bruker Daltonik) in quadruplicates. 1.5 µL of matrix solution (saturated solution of α-cyano-4-hydroxycinnamic acid (Sigma), 50% acetonitrile, 2.5% tri-fluoracetic-acid and and high-performance liquid chromatography (HPLC)-grade water) was directly spotted on each sample. Cocrystallization of the target plate was performed by air-drying at room temperature for 5 min (Seng et al. 2009). Measurements were performed with a

microflex LT mass spectrometer (Bruker Daltonik) equipped with a 337-nm nitrogen laser. Spectra were recorded in the positive linear mode (delay, 170 ns; ion source 1 voltage, 20 kV; ion source 2 voltage, 18.5 kV; lens voltage, 7 kV; mass range, 2–20 kDa). Each spectrum was obtained after 240 shots by an automatic mode at variable laser power with an acquisition time ranged of 30 sec per spot. Data were automatically acquired using AutoXecute acquisition flexcontrol software. The 4 raw spectra obtained for each isolate were imported into BioTyper 2.0 software (Bruker Daltonik) and analyzed by standard pattern matching (with default parameter settings) against the reference amoeba spectra in the BioTyper database. The method of identification included the m/z from 3 to 15 kDa. For each spectrum, no more than 100 peaks were taken into account and compared with peaks in the database. To validate the analysis of a whole MSP96 target, bacterial test standard (protein extract of *Escherichia coli* DH5alpha, Bruker Daltonics) and matrix solution were added to the analysis batch as positive and negative controls, respectively. To validate the analysis, two deposits of 1.5 µL of bacterial test standard covered with 1.5 µL of alpha-cyano-4-hydroxycinnamic acid Maldi matrix were used as a standard for the Bruker Biotyper analysis method and as a validation standard for the analysis. To validate the analysis using the Maldi Biotyper software, the bacterial identification score must be higher than 1.9 and the identification result must be *E. coli*. Both deposits of 1.5 µL of matrix solution must give a score lower than 1.7.

To compare the obtained protein mass profiles of the *Acanthamoeba* spp., the average spectral profiles obtained from the four spots for each specimen were created,

analyzed and compared using the ClinProTools, version 2.2, program (Bruker Daltonics). A hierarchical clustering of the mass spectra of all *Acanthamoeba* spp. was performed using the mean spectrum projection (MSP) dendrogram function of MALDI Biotyper, version 3.0. The objective was to determine whether the cultured *Acanthamoeba* spp. clusterize among the laboratory *Acanthamoeba* spp. strains.

## Acknowledgment

# REFERENCES

Appelt S, Armougom F, Le Bailly M, Robert C, and Drancourt M. 2013. Polyphasic analysis of a Middle Ages coprolite microbiota, Belgium. PloS one under revision.

Araujo A, Reinhard K, Bastos OM, Costa LC, Pirmez C, Iniguez A, Vicente AC, Morel CM, and Ferreira LF. 1998. Paleoparasitology: perspectives with new techniques. Rev Inst Med Trop Sao Paulo 40(6):371-376.

Booton GC, Rogerson A, Bonilla TD, Seal DV, Kelly DJ, Beattie TK, Tomlinson A, Lares-Villa F, Fuerst PA, and Byers TJ. 2004. Molecular and physiological evaluation of subtropical environmental isolates of *Acanthamoeba* spp., causal agent of Acanthamoeba keratitis. The Journal of eukaryotic microbiology 51(2):192-200.

Christine Helms A, Camillo Martiny A, Hofman-Bang J, K. Ahring B, and Kilstrup M. 2004. Identification of bacterial cultures from archaeological wood using molecular biological techniques. International Biodeterioration & Biodegradation 53(2):79-88.

Cooper A, and Poinar HN. 2000. Ancient DNA: do it right or not at all. Science 289(5482):1139.

Coulon C, Collignon A, McDonnell G, and Thomas V. 2010. Resistance of Acanthamoeba cysts to disinfection treatments used in health care settings. J Clin Microbiol 48(8):2689-2697.

Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M et al. . 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic acids research 36(Web Server issue):W465-469.

Drancourt M, and Raoult D. 2005. Palaeomicrobiology: current issues and perspectives. Nat Rev Microbiol 3(1):23-35.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research 32(5):1792-1797.

Fritsche TR, Horn M, Seyedirashti S, Gautom RK, Schleifer KH, and Wagner M. 1999. In situ detection of novel bacterial endosymbionts of *Acanthamoeba* spp. phylogenetically related to members of the order Rickettsiales. Applied and environmental microbiology 65(1):206-212.

Greub G. 2009. *Parachlamydia acanthamoebae*, an emerging agent of pneumonia. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases 15(1):18-28.

Guindon S, Delsuc F, Dufayard JF, and Gascuel O. 2009. Estimating maximum likelihood phylogenies with PhyML. Methods Mol Biol 537:113-137.

Hasegawa M, Kishino H, and Yano T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of molecular evolution 22(2):160-174.

Hebsgaard MB, Phillips MJ, and Willerslev E. 2005. Geologically ancient DNA: fact or artefact? Trends in microbiology 13(5):212-220.

Hofreiter M, Serre D, Poinar HN, Kuch M, and Paabo S. 2001. Ancient DNA. Nat Rev Genet 2(5):353-359.

Iniguez AM, Reinhard K, Carvalho Goncalves ML, Ferreira LF, Araujo A, and Paulo Vicente AC. 2006. SL1 RNA gene recovery from *Enterobius vermicularis* ancient DNA in pre-Columbian human coprolites. Int J Parasitol 36(13):1419-1425.

Johnson SS, Hebsgaard MB, Christensen TR, Mastepanov M, Nielsen R, Munch K, Brand T, Gilbert MT, Zuber MT, Bunce M. 2007. Ancient bacteria show evidence of DNA repair. Proc Natl Acad Sci U S A 104(36):14401-14405.

Khan NA. 2007. *Acanthamoeba* invasion of the central nervous system. Int J Parasitol 37(2):131-138.

Khanaeva TA, Suslova MY, Zemskaya TI, Molodin VI, Pilipenko AS, and Parzinger G. 2013. Microbial diversity in the samples from archeological complexes of the Pazyryk culture (IV–III centuries BC) in northwestern Mongolia. Microbiology 82(1):43-51.

Krisko A, and Radman M. 2010. Protein damage and death by radiation in *Escherichia coli* and *Deinococcus radiodurans*. Proc Natl Acad Sci U S A 107(32):14373-14377.

La Farge C, Williams KH, and England JH. 2013. Regeneration of Little Ice Age bryophytes emerging from a polar glacier with implications of totipotency in extreme environments. Proc Natl Acad Sci U S A 110(24):9839-9844.

La Scola B, Barrassi L, and Raoult D. 2000. Isolation of new fastidious alpha Proteobacteria and *Afipia felis* from hospital water supplies by direct plating and amoebal co-culture procedures. FEMS microbiology ecology 34(2):129-137.

Lewis K, Epstein S, Godoy VG, and Hong SH. 2008. Intact DNA in ancient permafrost. Trends in microbiology 16(3):92-94.

Lorenzo-Morales J, Lopez-Darias M, Martinez-Carretero E, and Valladares B. 2007. Isolation of potentially pathogenic strains of *Acanthamoeba* in wild squirrels from the Canary Islands and Morocco. Experimental parasitology 117(1):74-79.

Magnet A, Galvan AL, Fenoy S, Izquierdo F, Rueda C, Fernandez Vadillo C, Perez-Irezabal J, Bandyopadhyay K, Visvesvara GS, da Silva AJ et al. . 2012. Molecular characterization of *Acanthamoeba* isolated in water treatment plants and comparison with clinical isolates. Parasitology research 111(1):383-392.

Mariconti M, Epis S, Sacchi L, Biggiogera M, Sassera D, Genchi M, Alberti E, Montagna M, Bandi C, and Bazzocchi C. 2012. A study on the presence of flagella in the order Rickettsiales: the case of 'Candidatus Midichloria mitochondrii'. Microbiology 158(Pt 7):1677-1683.

Mazur T, Hadas E, and Iwanicka I. 1995. The duration of the cyst stage and the viability and virulence of *Acanthamoeba* isolates. Trop Med Parasitol 46(2):106-108.

Pääbo S, Poinar H, Serre D, Jaenicke-Despres V, Hebler J, Rohland N, Kuch M, Krause J, Vigilant L, and Hofreiter M. 2004. Genetic analyses from ancient DNA. Annu Rev Genet 38:645-679.

Pagnier I, Raoult D, and La Scola B. 2008. Isolation and identification of amoeba-resisting bacteria from water in human environment by using an *Acanthamoeba polyphaga* co-culture procedure. Environmental microbiology 10(5):1135-1144.

Price PB, and Sowers T. 2004. Temperature dependence of metabolic rates for microbial growth, maintenance, and survival. Proc Natl Acad Sci U S A 101(13):4631-4636.

Raoult D, Aboudharam G, Crubezy E, Larrouy G, Ludes B, and Drancourt M. 2000. Molecular identification by "suicide PCR" of *Yersinia pestis* as the agent of medieval black death. Proc Natl Acad Sci U S A 97(23):12800-12803.

Redmond C, Pearce MJ, Manchee RJ, and Berdal BP. 1998. Deadly relic of the Great War. Nature 393(6687):747-748.

Rhodes AN, Urbance JW, Youga H, Corlew-Newman H, Reddy CA, Klug MJ, Tiedje JM, and Fisher DC. 1998. Identification of bacterial isolates obtained from intestinal contents associated with 12,000-year-old mastodon remains. Applied and environmental microbiology 64(2):651-658.

Sassera D, Lo N, Bouman EA, Epis S, Mortarino M, and Bandi C. 2008. "*Candidatus Midichloria*" endosymbionts bloom after the blood meal of the host, the hard tick Ixodes ricinus. Applied and environmental microbiology 74(19):6138-6140.

Schroeder JM, Booton GC, Hay J, Niszl IA, Seal DV, Markus MB, Fuerst PA, and Byers TJ. 2001. Use of subgenic 18S ribosomal DNA PCR and sequencing for genus and genotype identification of *Acanthamoebae* from humans with keratitis and from sewage sludge. J Clin Microbiol 39(5):1903-1911.

Schuster FL, and Visvesvara GS. 2004. Free-living amoebae as opportunistic and non-opportunistic pathogens of humans and animals. Int J Parasitol 34(9):1001-1027.

Seng P, Drancourt M, Gouriet F, La Scola B, Fournier PE, Rolain JM, and Raoult D. 2009. Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry.

Clinical infectious diseases : an official publication of the Infectious Diseases Society of America 49(4):543-551.

Siddiqui R, and Khan NA. 2012. Biology and pathogenesis of *Acanthamoeba*. Parasites & vectors 5:6.

Sriram R, Shoff M, Booton G, Fuerst P, and Visvesvara GS. 2008. Survival of *Acanthamoeba* cysts after desiccation for more than 20 years. J Clin Microbiol 46(12):4045-4048.

Talavera G, and Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Systematic biology 56(4):564-577.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular biology and evolution 28(10):2731-2739.

Thomas V, Loret JF, Jousset M, and Greub G. 2008. Biodiversity of amoebae and amoebae-resisting bacteria in a drinking water treatment plant. Environmental microbiology 10(10):2728-2745.

Tito RY, Macmil S, Wiley G, Najar F, Cleeland L, Qu C, Wang P, Romagne F, Leonard S, Ruiz AJ et al. . 2008. Phylotyping and functional analysis of two ancient human microbiomes. PLoS One 3(11):e3703.

Visvesvara GS, Booton GC, Kelley DJ, Fuerst P, Sriram R, Finkelstein A, and Garner MM. 2007. In vitro culture, serologic and molecular analysis of *Acanthamoeba* isolated from the liver of a keel-billed toucan (Ramphastos sulfuratus). Veterinary parasitology 143(1):74-78.

Walochnik J, Hassl A, Simon K, Benyr G, and Aspock H. 1999. Isolation and identification by partial sequencing of the 18S ribosomal gene of free-living amoebae from necrotic tissue of *Basilliscus plumifrons* (Sauria: Iguanidae). Parasitology research 85(7):601-603.

WHO. 2003. Guidelines for Safe Recreational Water Environments: Coastal and Fresh Waters. Free-living microorganisms. Geneva, Switzerland: World Health Organization. p 102-117.

Wildschutte H, and Lawrence JG. 2007. Differential *Salmonella* survival against communities of intestinal amoebae. Microbiology 153(Pt 6):1781-1789.

Willerslev E, and Cooper A. 2005. Ancient DNA. Proceedings Biological sciences / The Royal Society 272(1558):3-16.

Wood JR, Wilmshurst JM, Wagstaff SJ, Worthy TH, Rawlence NJ, and Cooper A. 2012. High-resolution coproecology: using coprolites to reconstruct the habits and habitats of New Zealand's extinct upland moa (*Megalapteryx didinus*). PLoS One 7(6):e40025.

Yashina S, Gubin S, Maksimovich S, Yashina A, Gakhova E, and Gilichinsky D. 2012. Regeneration of whole fertile plants from 30,000-y-old fruit tissue buried in Siberian permafrost. Proc Natl Acad Sci U S A 109(10):4008-4013.

Yll R, Carrión JS, Marra AC, and Bonfiglio L. 2006. Vegetation reconstruction on the basis of pollen in Late Pleistocene hyena coprolites from San Teodoro Cave (Sicily, Italy). Palaeogeography, Palaeoclimatology, Palaeoecology 237(1):32-39.

**Legends of Figures**

**Figure 1. Microscopic observation of stained coprolite material.** Cellulosic components were stained using Congo-Red. **A-E**) cysts **F**) plant fiber (Camera: Nikon digital sight DS-U1 camera, optical magnification: x 100; scale bare on the right 10 μm).

**Figure 2. Dendrogram of *Acanthamoeba* species.** The dendrogram was created on the basis of unique protein profiles obtained by MALDI-TOF. Illustrated are the cluster arrangements of *Acanthamoeba* spp. cultured herein and those that are used or handled in the laboratory. LS means Lab strain, LS 1 corresponds to *A. polyphaga* AP1, LS 2 to *A. castellanii* Neff, LS 3 *Acanthamoeba* sp. to lentille 1, LS 4 to *Acanthamoeba* sp. crachat-Bernadet and LS 5 to *Acanthamoeba* sp. lentille 2.

**Figure 3. Transmission electron microscopy observations of the *Acanthamoeba* endosymbiont. A-B**) Observations of the endosymbiont by TEM after negative staining. **C-D**) Observations of thin layers of resin-embedded inclusion of the endosymbiont inside its host *Acanthamoeba* Culture B, by TEM.

**Figure S5. Phylogenetic tree of the 16S rDNA gene.** A phylogenetic tree was generated from the amplicons encoding the 16S rDNA gene region of the cultivated *Acanthamoeba* endosymbiont (red point). The tree was constructed using the PhyML algorithm with a bootstrap of 100. The bootstrap support (>40) is reported for each branch. Unreported GenBank Accession Numbers are listed corresponding to their identifier (triangle or kite) in **Supplementary Table S2**.
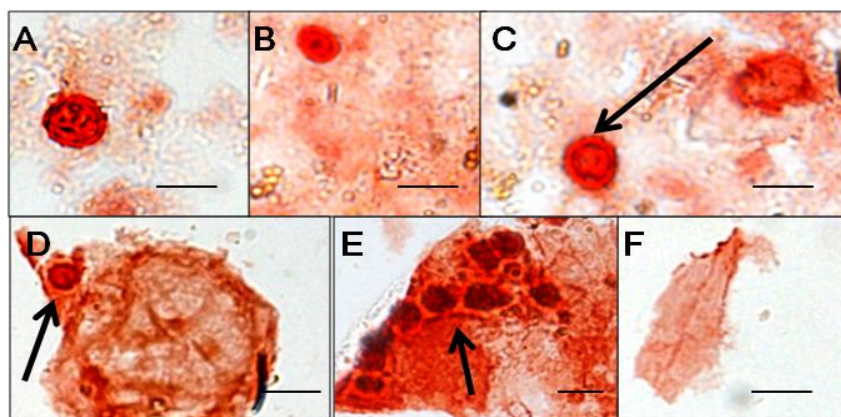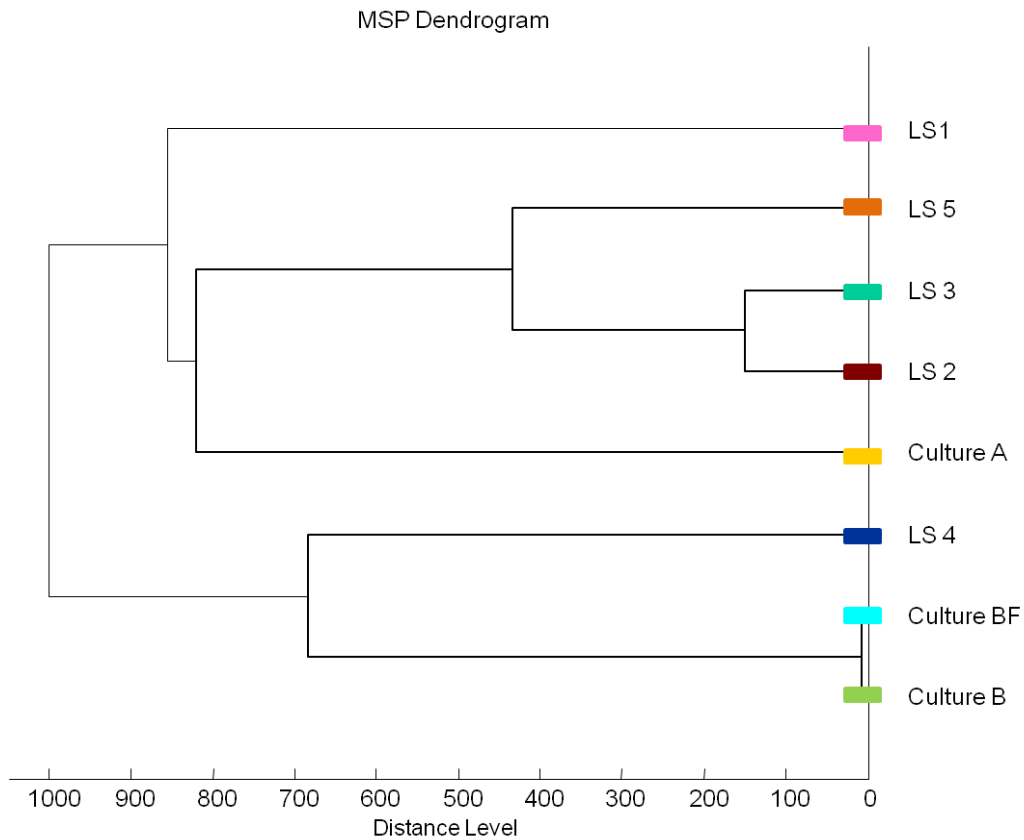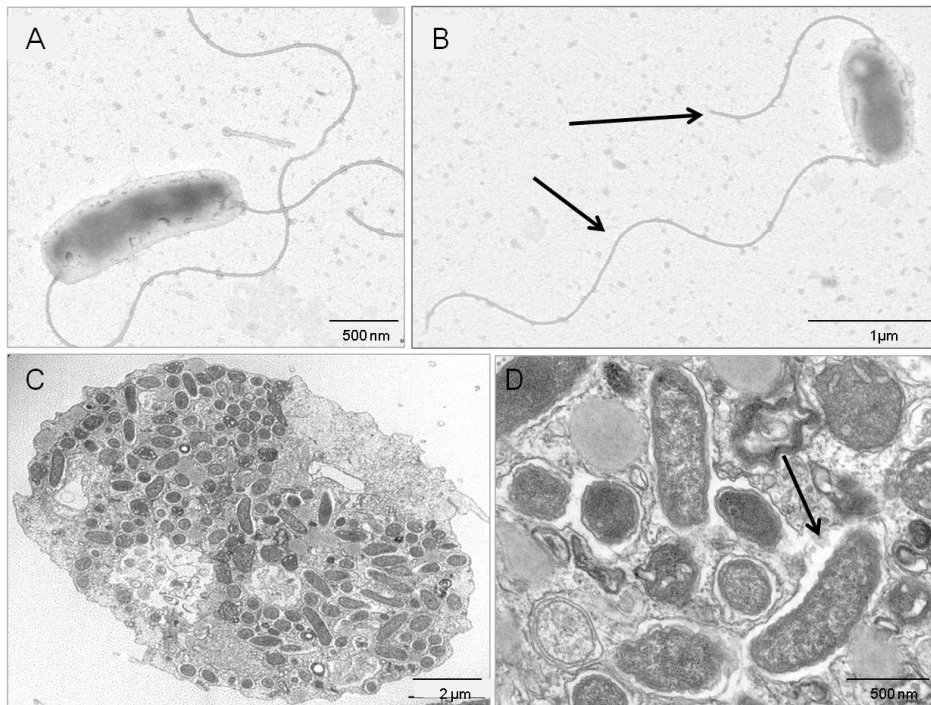
**Figure 1.**

MSP Dendrogram

LS1
LS 5
LS 3
LS 2
Culture A
LS 4
Culture BF
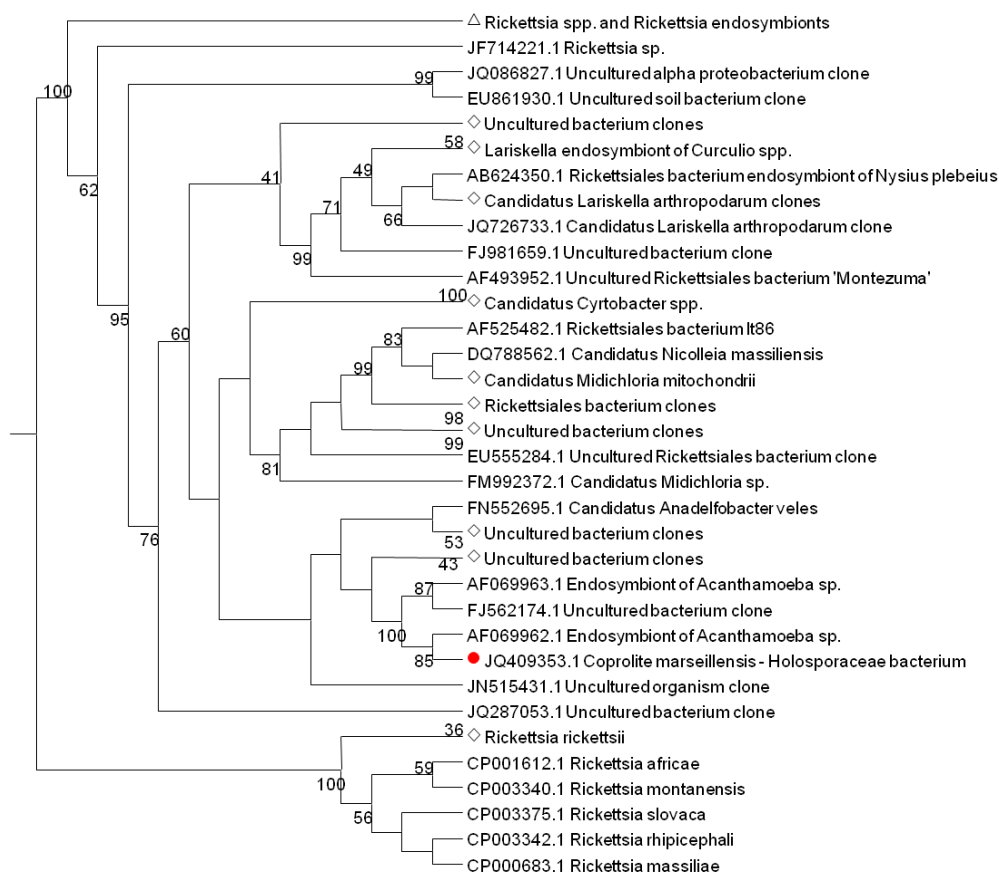Culture B

Distance Level

**Figure 2.**

**Figure 3.**

**Figure 4.**

## Growing *Acanthamoeba* out of medieval coprolite reveals a new endosymbiont

**Running title:** *Acanthamoeba* and Endosymbiont out of coprolite

Sandra Appelt[1], Morgan Gaia[1], Matthieu Le Bailly[2], Bernard La Scola[1]*, Michel Drancourt[1]*

[1] Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, Inserm 1095, 13005 Marseille, France. [2] Franche-Comté University, CNRS UMR 6249 Chrono-Environment, 25 030 Besançon, France.

*Corresponding author. Michel.Drancourt@medicine.univ-mrs.fr

**SUPPLEMENTAL MATERIAL CONTENTS**

**SUPPLEMENTARY TABLES**

Supplementary Tables S1-S2

**SUPPLEMENTARY FIGURES**

Supplementary Figures S1-S4

**SUPPLEMENTARY REFERENCES**

**Supplementary Table S1 :** Amplicons generated by *ad hoc* PCR amplifications[#].

| ID | Amplification length (bp) | Hit description | E-value | Hit Accession ID | Percentage Coverage | Percentage ID | Environment of Hit Isolation | Reference |
|---|---|---|---|---|---|---|---|---|
| Culture A | 326 | *Acanthamoeba castellanii* | 4e-169 | JX983592.1 | 100% | 100% | compost of composting facilities | Conza et al. 2013 |
| | | *Acanthamoeba castellanii clone CF1-119b* | 4e-169 | KC164234.1 | 100% | 100% | compost of composting facilities | Conza et al. 2013 |
| | | *Acanthamoeba castellanii clone CF3-107* | 4e-169 | KC164230.1 | 100% | 100% | compost of composting facilities | Conza et al. 2013 |
| | | *Acanthamoeba polyphaga clone CF3-77b* | 4e-169 | KC164226.1 | 100% | 100% | compost of composting facilities | Conza et al. 2013 |
| Culture B | 325 | *Acanthamoeba castellanii* | 9e-161 | JX983592.1 | 100% | 98% | compost of composting facilities | Conza et al. 2013 |
| | | *Acanthamoeba castellanii clone CF1-119b* | 9e-161 | KC164234.1 | 100% | 98% | compost of composting facilities | Conza et al. 2013 |
| | | *Acanthamoeba castellanii clone CF3-107* | 9e-161 | KC164230.1 | 100% | 98% | compost of composting facilities | Conza et al. 2013 |
| | | *Acanthamoeba polyphaga clone CF3-77b* | 9e-161 | KC164226.1 | 100% | 98% | compost of composting facilities | Conza et al. 2013 |
| Culture BF | 416 | *Acanthamoeba castellanii clone CF1-119b* | 0 | KC164234.1 | 99% | 98% | compost of composting facilities | Conza et al. 2013 |
| | | *Acanthamoeba castellanii ATCC 50374* | 0 | U07413.1 | 99% | 98% | ---- | Gast et al. 1997 |
| PCR Coprolite* | 323 | *Acanthamoeba sp. Had_008* | 4e-164 | FJ042636.1 | 100% | 99% | cornea (host:human/Israel) | unpublished° |
| | | *Acanthamoeba polyphaga isolate A8/SB2* | 4e-162 | GU596994.1 | 99% | 99% | environmental biofilm | Hsu et al. 2011 |
| PCR cysts * | 405 | *Acanthamoeba polyphaga Page-23* | 0 | AF019061.1 | 100% | 100% | ---- | Gast et al. 1997 |

[#] The amplicon identifier, its length (bp) and the annotation according to the best BLAST hits (BLASTN *versus* the non-redundant NCBI database) are su[...] E-value, the hit accession identifier, the percent of identity, the environment of described isolation and the reference are also provided. * Suicide PCR amplifications were [...] primer pairs were used only ones in working areas and no positive controls were incorporated (Raoult et al. 2000).

**Supplementary Table S2:** GenBank Accession Numbers corresponding to the phylogenetic construct shown in **Supplementary Figure S5**.

| Phylogenetic tree Identifier | GenBank Accession No | Species |
|---|---|---|
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | DQ365809.1 | *Rickettsia raoultii* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | AF120026.1 | *Rickettsia sp. RpA4 strain* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | JX432017.1 | *Uncultured Rickettsia sp.* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | EU071486.1 | *Uncultured Rickettsia sp.* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | AF120024.1 | *Rickettsia sp. DnS28 strain* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | EU036982.1 | *Rickettsia raoultii* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | AB604665.1 | *Rickettsia secondary endosymbiont of Curculio camelliae* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | NR 074472.1 | *Rickettsia montanensis* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | CP003319.1 | *Rickettsia massiliae* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | L36106.1 | *Rickettsia massiliae* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | L36102.1 | *Rickettsia sp.* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | NR 074486.1 | *Rickettsia massiliae* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | NR 025919.1 | *Rickettsia massiliae* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | JQ339355.1 | Uncultured *Rickettsia* sp. |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | JQ412124.1 | Uncultured *Rickettsia* sp. |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | AB604667.1 | *Rickettsia secondary endosymbiont Curculio hilgendorfi* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | NR 074480.1 | *Rickettsia conorii* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | NR 074462.1 | *Rickettsia slovaca* |

| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | NR 074474.1 | *Rickettsia slovaca* |
| --- | --- | --- |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | NR 025920.1 | *Rickettsia montanensis* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | U11016.1 | *Rickettsia montana* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | NR 074527.1 | *Rickettsia africae* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | L36098.1 | *Rickettsia africae* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | NR 074483.1 | *Rickettsia felis* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | CP003338.1 | *Rickettsia australis* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | NR 074496.1 | *Rickettsia australis* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | NR 102941.1 | *Rickettsia rickettsii* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | NR 025921.1 | *Rickettsia rhipicephali* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | CP000053.1 | *Rickettsia felis* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | AB746407.1 | *Rickettsia* endosymbiont of *Curculio hilgendorfi* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | AB604666.1 | *Rickettsia* secondary endosymbiont of *Curculio lateritius* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | L28944.1 | *Rickettsia* sp. |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | JQ726775.1 | *Rickettsia* endosymbiont of *Kleidocerys resedae* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | AF060705.2 | *Rickettsia honei* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | NR 074473.1 | *Rickettsia rhipicephali* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | AB604668.1 | *Rickettsia* secondary endosymbiont *Curculio hilgendorfi* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | NR 074484.1 | *Rickettsia bellii* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | NR 026042.1 | *Rickettsia aeschlimannii* |
| Δ *Rickettsia* spp. and *Rickettsia* endosymbionts | U11019.1 | *Rickettsia rhipicephali* |

◊ *Candidatus Midichloria mitochondrii*  CP002130.1  *Candidatus Midichloria mitochondrii*

◊ *Candidatus Midichloria mitochondrii*  AJ566640.2  *Candidatus Midichloria mitochondrii*

◊ *Candidatus Midichloria mitochondrii*  NR 074492.1  *Candidatus Midichloria mitochondrii*

◊ Rickettsiales bacterium clones  DQ379965.1  *Rickettsiales* bacterium clone

◊ Rickettsiales bacterium clones  DQ379964.1  *Rickettsiales* bacterium clone

◊ *Lariskella* endosymbiont of *Curculio* spp.  AB746417.1  *Lariskella endosymbiont of Curculio okumai*

◊ *Lariskella* endosymbiont of *Curculio* spp  AB746412.1  *Lariskella endosymbiont of Curculio morimotoi*

◊ *Lariskella* endosymbiont of *Curculio* spp  AB746414.1  *Lariskella endosymbiont of Curculio morimotoi*

◊ *Lariskella* endosymbiont of *Curculio* spp  AB746413.1  *Lariskella endosymbiont of Curculio morimotoi*

◊ *Lariskella* endosymbiont of *Curculio* spp  AB746416.1  *Lariskella endosymbiont of Curculio okumai*

◊ *Candidatus Lariskella arthropodarum* clones  JQ726735.1  *Candidatus Lariskella arthropodarum* clone

◊ *Candidatus Lariskella arthropodarum* clones  JQ726715.1  *Candidatus Lariskella arthropodarum* clone

◊ *Candidatus Lariskella arthropodarum* clones  JQ726727.1  *Candidatus Lariskella arthropodarum* clone

◊ *Candidatus Lariskella arthropodarum* clones  JQ726723.1  *Candidatus Lariskella arthropodarum* clone

◊ *Candidatus Lariskella arthropodarum* clones  JQ726733.1  *Candidatus Lariskella arthropodarum* clone

◊ *Candidatus Cyrtobacter* spp.  FN552696.1  *Candidatus Cyrtobacter comes*

◊ *Candidatus Cyrtobacter* spp.  FN552698.1  *Candidatus Cyrtobacter comes*

◊ *Candidatus Cyrtobacter* spp.  HE978250.1  *Candidatus Cyrtobacter zanobii*

◊ *Candidatus Cyrtobacter* spp.  FN552697.1  *Candidatus Cyrtobacter comes*

◊ uncultured bacteria clones  AY942762.1  Uncultured bacterium clone

◊ uncultured bacteria clones  EF667892.1  Uncultured *Rickettsiales* bacterium clone

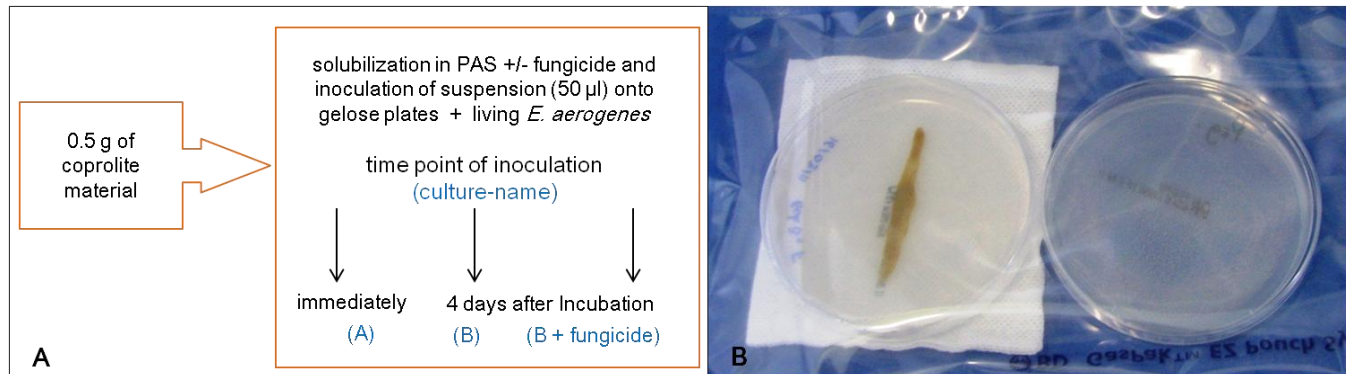| | | |
|---|---|---|
| ◊ uncultured bacteria clones | KC682789.1 | Uncultured bacterium clone |
| ◊ uncultured bacteria clones | GU189017.1 | Uncultured bacterium clone |
| ◊ uncultured bacteria clones | JQ086827.1 | Uncultured *alpha proteobacterium* clone |
| ◊ uncultured bacteria clones | JQ337869.1 | Uncultured bacterium clone |
| ◊ uncultured bacteria clones | EU861930.1 | Uncultured soil bacterium clone |
| ◊ uncultured bacteria clones | JX170254.1 | Uncultured bacterium clone |
| ◊ uncultured bacteria clones | GU118498.1 | Uncultured bacterium clone |
| ◊ uncultured bacteria clones | KF037677.1 | Uncultured bacterium clone |
| ◊ uncultured bacteria clones | EU236349.1 | Uncultured bacterium clone |
| ◊ uncultured bacteria clones | EU789894.1 | Uncultured bacterium clone |
| ◊ uncultured bacteria clones | GQ302530.1 | Uncultured *alpha proteobacterium* clone |
| ◊ uncultured bacteria clones | HE999308.1 | Uncultured bacterium |
| ◊ uncultured bacteria clones | GQ480088.1 | Uncultured bacterium clone |
| ◊ uncultured bacteria clones | AF497583.1 | Uncultured bacterium clone |
| ◊ uncultured bacteria clones | AF497582.1 | Uncultured bacterium clone |

**Legends of Supplementary Figures**

**Supplementary Figure S1.** *Acanthamoeba* **Culturing.** (**A**) Procedure of cultivation and sampling. (**B**) Coprolite suspension plated onto a non-nutritive agar plate over-coated with a thin film of living *Enterobacter aerogenes* cells (on the left) and the introduced negative control consisting of non-nutritive agar plate coated with a thin film of living *E. aerogenes* cells without coprolite suspension (on the right).
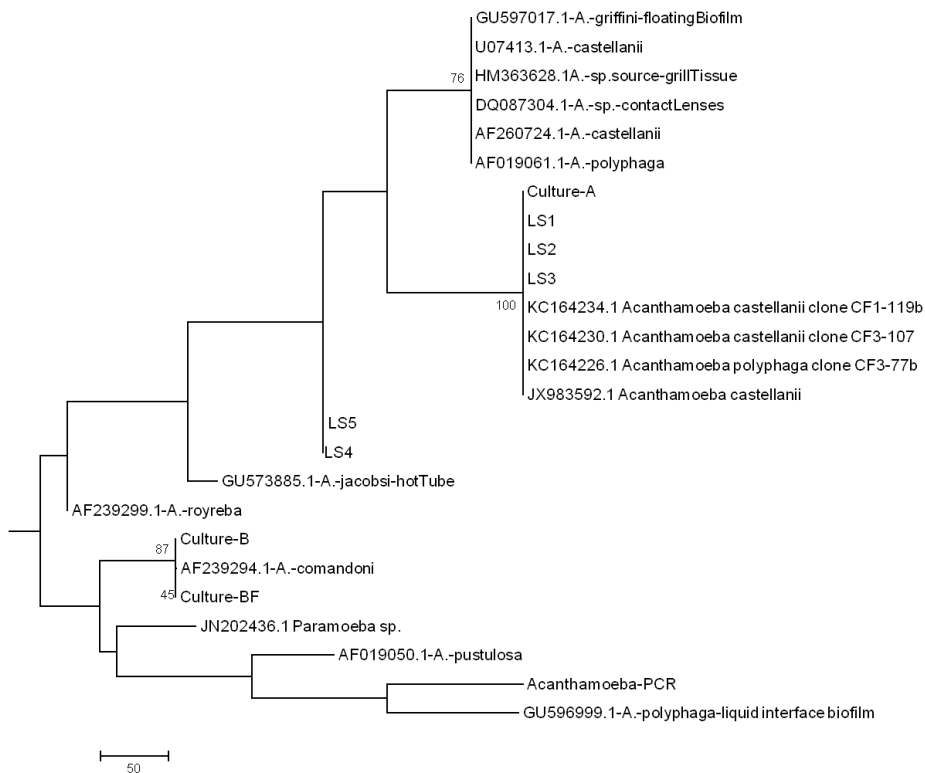
**Supplementary Figure S2: Phylogenetic tree of the 18S rDNA gene.** A phylogenetic tree was generated from the amplicons encoding the 18S rDNA gene region of *Acanthamoeba*. The tree was constructed using the PhyML algorithm with a bootstrap of 100. The bootstrap support (>40) is reported for each branch. LS mean laboratory strain.

**Supplementary Figure S3. MALDI-TOF profile of** *Acanthamoeba*. **Supplementary Figure S3. MALDI-TOF profile of** *Acanthamoeba* **species.** Reported are the identifier and the strain specific protein profile of *Acanthamoeba* spp. cultured herein and those that are used or handled in the laboratory. The strain-specific profiles were reproducible over 3 replicates. LS mean laboratory strain.
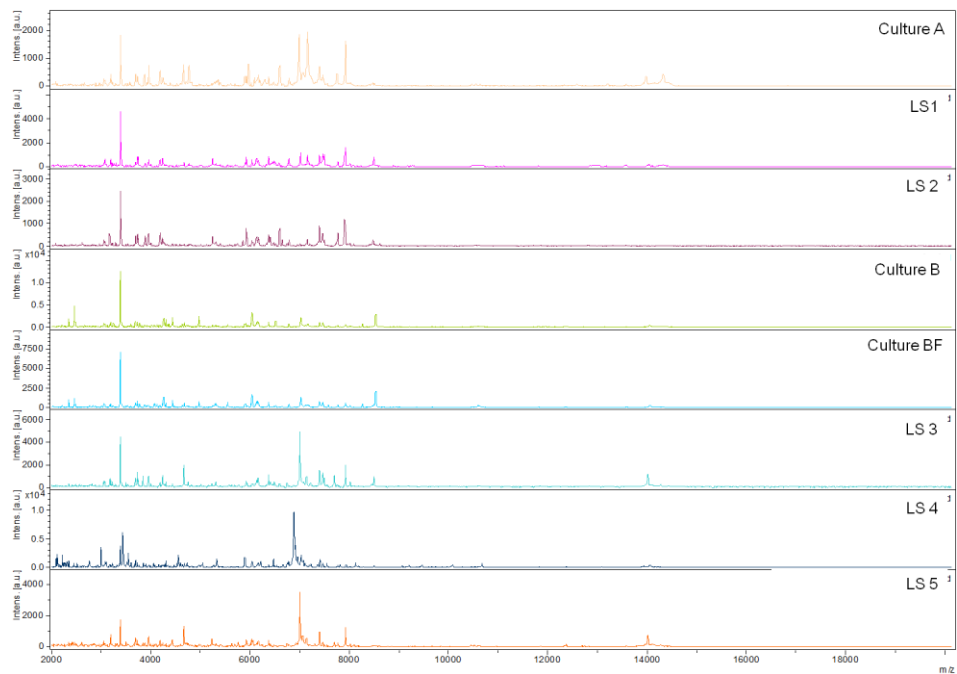
**Supplemental Figure S4. Inter-species comparison of** *Acanthamoeba* **MALDI-TOF profiles.** The peptide profiles of cultured *Acanthamoeba* were compared among them and to those of laboratory strains. Strain specific peptide peaks (red), peaks that are similarly (yellow) and identical peaks (green) are shown. **A)** cultured strain B compared to cultured strain BF, **B)** strain BF compared to lab strain 4, **C)** strain B compared to lab strain 4, **D)** cultured strain A compared to lab strain 1, **E)** strain A compared to lab strain 2.
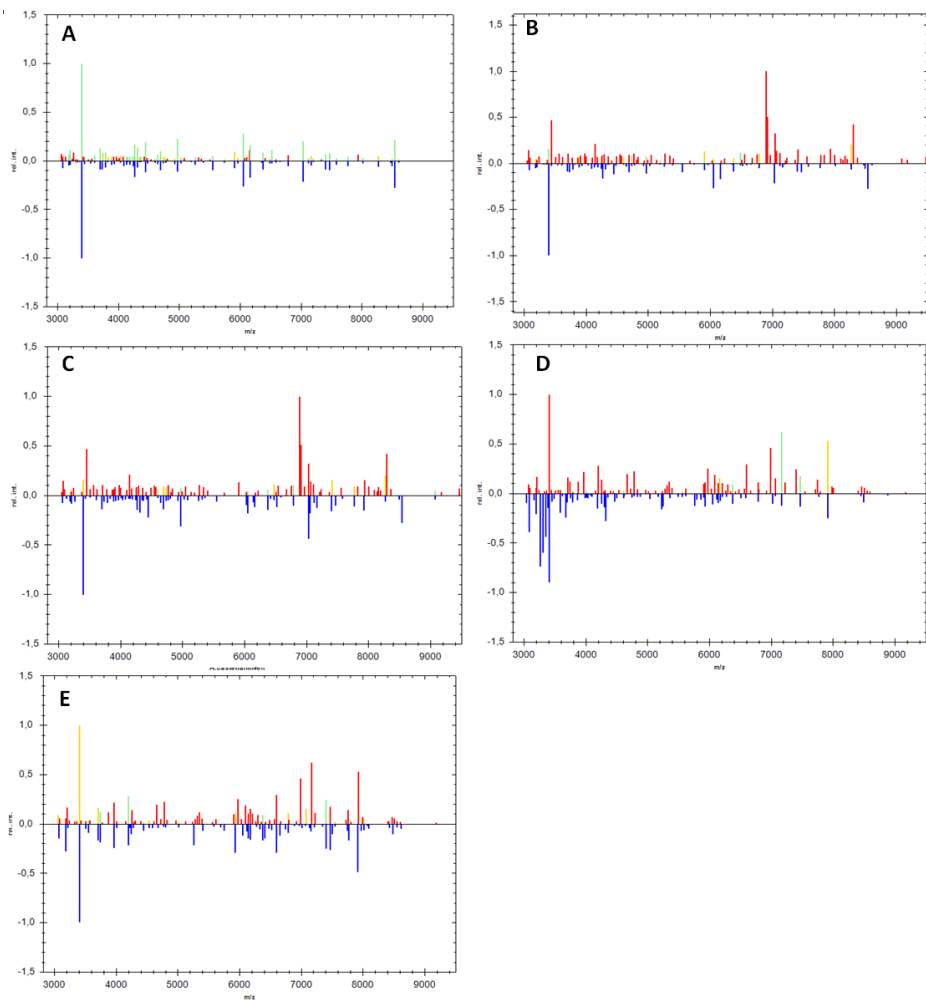
**Supplementary Figure S1.**

GU597017.1-A.-griffini-floatingBiofilm
U07413.1-A.-castellanii
76  HM363628.1A.-sp.source-grillTissue
DQ087304.1-A.-sp.-contactLenses
AF260724.1-A.-castellanii
AF019061.1-A.-polyphaga
Culture-A
LS1
LS2
LS3
100  KC164234.1 Acanthamoeba castellanii clone CF1-119b
KC164230.1 Acanthamoeba castellanii clone CF3-107
KC164226.1 Acanthamoeba polyphaga clone CF3-77b
JX983592.1 Acanthamoeba castellanii
LS5
LS4
GU573885.1-A.-jacobsi-hotTube
AF239299.1-A.-royreba
87  Culture-B
AF239294.1-A.-comandoni
45  Culture-BF
JN202436.1 Paramoeba sp.
AF019050.1-A.-pustulosa
Acanthamoeba-PCR
GU596999.1-A.-polyphaga-liquid interface biofilm

50

**Supplementary Figure S2.**

243

**Supplementary Figure S3.**

**Supplemental Figure S4.**

# SPPLEMENTARY REFERENCES

Conza L, Pagani SC, and Gaia V. 2013. Presence of legionella and free-living amoebae in composts and bioaerosols from composting facilities. PLoS One 8(7):e68244.

Gast RJ, Ledee DR, Fuerst PA, and Byers TJ. 1996. Subgenus systematics of Acanthamoeba: four nuclear 18S rDNA sequence types. The Journal of eukaryotic microbiology 43(6):498-504.

Hsu BM, Huang CC, Chen JS, Chen NH, Huang JT. 2011. Comparison of potentially pathogenic free-living amoeba hosts by *Legionella* spp. in substrate-associated biofilms and floating biofilms from spring environments. Water Research 45: 5171-5181.

Raoult D, Aboudharam G, Crubezy E, Larrouy G, Ludes B, and Drancourt M. 2000. Molecular identification by "suicide PCR" of Yersinia pestis as the agent of medieval black death. Proc Natl Acad Sci U S A 97(23):12800-12803.

# CONCLUSIONS ET PERSPECTIVES

Dans notre travail de thèse, un coprolithe provenant d'une latrine des 14-15 ème siècles découvert à Namur en Belgique, a été étudié. Afin d'accéder au microbiote préservé à l'intérieur du coprolithe, différentes techniques fournissant des informations complémentaires ont été utilisées, au nombre desquelles la microscopie, la culture (**Chapitre II**, **Chapitre IV**) ou des analyses par biologie moléculaire (**Chapitre II**, **Chapitre III**). Ce faisant, en respectant les standards paléomicrobiologiques (Cooper & Poinar 2000; Drancourt & Raoult 2005; Hofreiter et al. 2001; Pääbo et al. 2004; Willerslev & Cooper 2005), nous avons trouvé des résultats se complétant les uns les autres, renforçant de fait leur authenticité. Les micro-organismes détectés associés au coprolithe de Namur proviennent de deux origines distinctes, d'une part l'environnement et de l'autre la flore intestinale. L'observation par microscopie a permis de révéler la présence de fibres végétales, de restes de moisissures, de kystes d'amibes, d'helminthes intestinaux, de bactériophages et de particules ressemblant à des virus. Ces résultats ont été renforcés par la culture, qui a mis en évidence des bactéries ayant pour origine l'environnement et la flore intestinale, et des amibes avec un endosymbionte associé. De même, des eucaryotes, des procaryotes et des virus ont été identifiés par biologie moléculaire. Les résultats obtenus après analyse des données métagénomiques – issues du séquençage à haut-débit de l'ADN total (**Chapitre II**) et de l'ADN viral (**Chapitre III**) extraits du coprolithe – correspondent à des bactéries et leurs bactériophages. Parmi ces derniers, des bactériophages infectant des bactéries du sol et de l'intestin, comme les *Enterobacteriaceae* et les *Lactobacillus*, ont été trouvés (Kim et al. 2013; Yatsunenko et al. 2012). Le virome du coprolithe a en outre été caractérisé par une forte abondance de facteurs de virulence associés à des bactéries pathogenes, parmi lesquelles certaines – *Legionella*, *Mycobacteria* et *Bordetella* notamment – avaient été identifiées dans le premier projet.

Pour approfondir ce point, nous avons cherché, par BLAST contre la banque de données ARDB (Antibiotic Resistance Genes Database), des gènes de résistance aux antibiotiques qui pourraient être associés au virome. Ainsi, un gène codant pour une chloramphenicol O-acétyltransferase, médiatrice de la résistance au chloramphenicol, a été détecté. Sa présence a été par la suite confirmée par une analyse phylogénétique. Au vu de ces résultats, nous avons cherché, en collaboration avec une autre équipe de notre laboratoire (Jean-Marc Rolain, Seydina Diene), des gènes de résistance dans la banque

de données WGS par BLAST contre la banque de données de gènes de résistance établie au laboratoire. Les premiers résultats ont idiqué la présence de gènes asociés aux *Klebsiella* codant pour une β-lactamase, confirmée par les reconstructions phylogénétiques (données non publiées). Ces résultats concernant la présence de gènes impliqués dans des résistances aux antibiotiques ne sont pour le moment que préliminaires, suggerant de futures études pour obtenir plus d'informations sur le résistome associé au coprolithe de Namur.

Pour résumer, cette thèse s'est intéressée à divers objectifs : la caractérisation de la communauté microbienne associée au coprolithe de Namur, et l'identification d'agents pathogènes associés. Nous avons identifié des eucaryotes, des procaryotes ainsi que des virus, augmentant la liste des micro-organismes découverts dans des coprolithes et des spécimens archéologiques. Au vu des résultats, nos études soulignent une nouvelle fois l'intérêt d'analyser les coprolithes afin d'élargir les connaissances des écosystèmes anciens et des conditions de vie des populations anciennes. Nos études ont permis d'en apprendre davantage sur la flore intestinale et les pathogènes anciens (*Bordetella* et *Bartonella*) de l'Europe du Moyen-âge. Il a également été montré que la métagénomique virale est un outil applicable aux coprolithes pour l'analyse des communautés virales anciennes. La méthode peut donc être appliquée au domaine de la paléomicrobiologie, en apportant de nouvelles informations sur les particules virales et leurs acides nucléiques, en complément des résultats obtenus par PCR.

Des amibes séculaires et leur endosymbionte associé ont également été cultivées. En considérant les résultats déjà obtenus et le fait que ces micro-organismes sont des témoins des flores anciennes, il serait intéressant d'approfondir leurs analyses, et notamment de reconstruire et d'annoter leur génome afin de mieux comprendre leurs possibles relations phylogénétiques et les processus d'évolution. Plus largement, le travail réalisé pendant cette thèse démontre les possibilités qu'offrent les coprolithes comme source de connaisance sur les environnements anciens à partir d'un unique échantillon. La latrine moyenâgeuse étudiée ici a permis des analyses biomoléculaires et des cultures, techniques fournissant des informations indépendantes sur les environnements et populations anciens. La quantité de données recueillies via ces approches à partir d'environ 100 g d'un unique échantillon archéologique, laisse présager des possibilités offertes par les futures analyses des coprolithes.

# REFERENCES

**Bedarida, S., Dutour, O., Buzhilova, AP., De Micco, P., Biagini P.** 2011. Identification of viral DNA (*Anelloviridae*) in a 200-year-old dental pulp sample (Napoleon's Great Army, Kaliningrad, 1812). *Infection, Genetics and Evololution* 11: 358-62.

**Cano, RJ., Tiefenbrunner, F., Ubaldi, M., Del Cueto, C., Luciani, S., Cox, T., Orkand, P., Kunzel, KH., and Rollo, F.** 2000. Sequence analysis of bacterial DNA in the colon and stomach of the Tyrolean Iceman. *American Journal of Physical Anthropology* 112.

**Chaves, SDM., and Reinhard, KJ.** 2006. Critical analysis of coprolite evidence of medicinal plant use, Piauí, Brazil. *Palaeogeography, Palaeoclimatology, Palaeoecology* 237, 110-118.

**Christine Helms A, Camillo Martiny, A., Hofman-Bang, JK., Ahring, B., and Kilstrup, M.** 2004. Identification of bacterial cultures from archaeological wood using molecular biological techniques. *International Biodeterioration & Biodegradation* 53(2):79-88.

**Cleeland, LM., Reichard, MV., Tito, RY., Reinhard, KJ., and Lewis, CM. Jr.** 2013. Clarifying Prehistoric Parasitism from a Complementary Morphological and Molecular Approach. *Journal of Archaeological Science* 40.

**Cooper, A. and Poinar, HN.** 2000. Ancient DNA: do it right or not at all. *Science* 289.

**Dean, GW.** 2006. The science of coprolite analysis: The view from Hinds cave. *Palaeogeography, Palaeoclimatology, Palaeoecology* 237, 67-79.

**Drancourt, M. and Raoult, D.** 2005. Palaeomicrobiology: current issues and perspectives. *Nature Review Microbiology* 3.

**Fornaciari, G and Marchetti, A..** 1986. Intact smallpox virus particles in an Italian mummy of sixteenth century. *Lancet* 2.

Goncalves, ML., Araujo, A., Duarte, R., Da Silva, JP., Reinhard, K., Bouchet, F., and Ferreira, LF. 2002. Detection of *Giardia duodenalis* antigen in coprolites using a commercially available enzyme-linked immunosorbent assay. *Transaction of the Royal Society of Tropical Medicine and Hygiene* 96.

Goncalves, ML., Araujo, A., and Ferreira, LF. 2003. Human intestinal parasites in the past: new findings and a review. *Memórias do Instituto Oswaldo Cruz* 98 Suppl 1.

Hofreiter, M., Serre, D., Poinar, H. N., Kuch, M., and Paabo, S. 2001. Ancient DNA. *Nature Reviews Genetics* 2.

Johnson, SS., Hebsgaard, MB., Christensen, TR., Mastepanov, M., Nielsen, R., Munch, K., Brand, T., Gilbert, MT., Zuber, MT., Bunce, M. 2007. Ancient bacteria show evidence of DNA repair. *Proceedings of the National Academy of Sciences of the United States of America* 104(36):14401-14405.

Kennedy, MJ., Reader, SL., Swierczynski, LM. 1994. Preservation records of micro-organisms: evidence of the tenacity of life. *Microbiology* 140, 2513-2529.

Khanaeva TA, Suslova MY, Zemskaya TI, Molodin VI, Pilipenko AS, and Parzinger G. 2013. Microbial diversity in the samples from archeological complexes of the Pazyryk culture (IV–III centuries BC) in northwestern Mongolia. *Microbiology* 82(1):43-51.

Kim, MS., Park, EJ., Roh, SW., Bae, JW. 2011. Diversity and abundance of single-stranded DNA viruses in human feces. *Applied and Environmental Microbiology* 77: 8062-70.

La Farge, C., Williams, KH., and England, JH. 2013. Regeneration of Little Ice Age bryophytes emerging from a polar glacier with implications of totipotency in extreme environments. *Proceedings of the National Academy of Sciences of the United States of America* 110(24):9839-9844.

Lambert, LH. , Cox, T., Mitchell, K., Rossello-Mora, RA., Del Cueto, C., Dodge, DE., Orkand, P. Cano, RJ. 1998. *Staphylococcus succinus* sp. nov., isolated from Dominican amber. *International Journal of Systematic Bacteriology* 48, 511-518.

Le Bailly, M., Goncalves, M. L., Harter-Lailheugue, S., Prodeo, F., Araujo, A., and Bouchet, F. 2008. New finding of *Giardia intestinalis* (Eukaryote, Metamonad) in Old World archaeological site using immunofluorescence and

enzyme-linked immunosorbent assays. *Memórias do Instituto Oswaldo Cruz* 103.

**Leles, D., Araujo, A., Ferreira, LF., Vicente, A. C., and Iniguez, AM.** 2008. Molecular paleoparasitological diagnosis of *Ascaris* sp. from coprolites: new scenery of ascariasis in pre-Colombian South America times. *Memórias do Instituto Oswaldo Cruz* 103.

**Lewis, K., Epstein, S., Godoy, VG., and Hong, SH.** 2008. Intact DNA in ancient permafrost. *Trends in microbiology* 16(3):92-94.

**Li, HC., Fujiyoshi, T., Lou, H., Yashiki, S., Sonoda, S., Cartier, L., Nunez, L., Munoz, I., Horai, S., Tajima, K.** 1999. The presence of ancient human T-cell lymphotropic virus type I provirus DNA in an Andean mummy. *Nature Medicine* 5: 1428-32.

**Loreille, O., Roumat, E., Verneau, O., Bouchet, F., and Hanni, C.** 2001. Ancient DNA from *Ascaris*: extraction amplification and sequences from eggs collected in coprolites. *International Journal of Parasitology* 31.

**Luciani, S., Fornaciari, G., Rickards, O., Labarga, C. M., and Rollo, F.** 2006. Molecular characterization of a pre-Columbian mummy and in situ coprolite. *American Journal of Physical Anthropology* 129.

**Marennikova, SS., Shelukhina, EM., Zhukova, OA., Yanova, NN., Loparev, VN.** 1990. Smallpox diagnosed 400 years later: results of skin lesions examination of 16th century Italian mummy. *Journal of hygiene, epidemiology, microbiology, and immunology* 34: 227-31.

**Mazur, T, Hadas, E, and Iwanicka, I.** 1995. The duration of the cyst stage and the viability and virulence of *Acanthamoeba* isolates. *Tropical Medicine and Parasitology* 46(2):106-108.

**Pääbo, S., Poinar, H., Serre, D., Jaenicke-Despres, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L., and Hofreiter, M.** 2004. Genetic analyses from ancient DNA. *Annual Review of Genetics* 38.

**Price, PB. and Sowers, T.** 2004. Temperature dependence of metabolic rates for microbial growth, maintenance, and survival. *Proceedings of the National Academy of Sciences Proceedings of the National Academy of Sciences* 101, 4631-4636.

**Redmond, C., Pearce, MJ., Manchee, RJ., and Berdal, BP.** 1998. Deadly relic of the Great War. *Nature* 393(6687):747-748.

**Reinhard, KJ.** 2007. Hunter-gatherer use of small animal food resources : coprolite evidence. *International journal of osteoarchaeology* 17, 416-428.

**Reinhard, KJ. and Bryant, VMJ.** 2008. Pathoecology and the Future of Coprolite Studies in Bioarchaeology. In (ed. Ann W.M.Stodder). Tempe, Arizona: Arizona State University Press.

**Rhodes, AN., Urbance, JW., Youga, H., Corlew-Newman, H., Reddy, CA., Klug, MJ., Tiedje, JM., and Fisher, DC.** 1998. Identification of bacterial isolates obtained from intestinal contents associated with 12,000-year-old mastodon remains. *Applied and Environmental Microbiology* 64(2):651-658.

**Rollo, F., Ermini, L., Luciani, S., Marota, I., and Olivieri, C.** 2006. Studies on the preservation of the intestinal microbiota's DNA in human mummies from cold environments. *Medicina nei secoli* 18.

**Santiago-Rodriguez, TM., Narganes-Storde, YM., Chanlatte, L., Crespo-Torres, E., Toranzos, G. A., Jimenez-Flores, R., Hamrick, A., and Cano, RJ.** 2013. Microbial communities in pre-columbian coprolites. *PLoS One* 8.

**Satterfield, CL., Lowenstein, TK., Vreeland, RH., Rosenzeig, WD. and Powers, DW.** 2005. New evidence for 250 Ma age of halotolerant bacterium from a Permian salt crystal: Comment and reply: REPLY. *Geology* 33, e93-e94.

**Sonoda, S., Li, HC., Cartier, L., Nunez, L., and Tajima, K.** 2000. Ancient HTLV type 1 provirus DNA of Andean mummy. *AIDS Research and Human Retroviruses* **16:** 1753-6.

**Sriram, R., Shoff, M., Booton, G., Fuerst, P., and Visvesvara, GS.** 2008. Survival of *Acanthamoeba* cysts after desiccation for more than 20 years. *Journal of Clinical Microbiology* 46(12):4045-4048.

**Taubenberger, JK.** 1997. Initial genetic characterization of the 1918 'Spanish' influenza virus. *Science* 275, 1793-1796.

**Tito, RY., Knights, D., Metcalf, J., Obregon-Tito, AJ., Cleeland, L., Najar, F., Roe, B., Reinhard, K., Sobolik, K., Belknap, S., Foster, M., Spicer, P., Knight,**

**R., and Lewis, CM. Jr.** 2012. Insights from characterizing extinct human gut microbiomes. *PLoS One* 7.

**Tito, RY., Macmil, S., Wiley, G., Najar, F., Cleeland, L., Qu, C., Wang, P., Romagne, F., Leonard, S., Ruiz, AJ., Reinhard, K., Roe, BA., and Lewis, CM., Jr.** 2008. Phylotyping and functional analysis of two ancient human microbiomes. *PLoS One* 3.

**Ubaldi, M., Luciani, S., Marota, I., Fornaciari, G., Cano, RJ., and Rollo, F.** 1998. Sequence analysis of bacterial DNA in the colon of an Andean mummy. *American Journal of Physical Anthropology* 107.

**Vreeland, RH., Rosenzweig, WD., and Powers, DW.** 2000. Isolation of a 250 million-year-old halotolerant bacterium from a primary salt crystal. *Nature* 407(6806):897-900.

**Willerslev, E. and Cooper, A.** 2005. Ancient DNA. *Proceedings Biological sciences / The Royal Society* 272.

**Willerslev, E. and Hebsgraad, MB.** 2005. New evidence for 250 Ma age of halotolerant bacterium from a Permian salt crystal: Comment and reply: COMMENT. *Geology* 33, e93.

**Willerslev, E., Hansen, AJ. and Poinar, HN.** 2004. Isolation of nucleic acids and cultures from fossil ice and permafrost. *Trends in Ecology & Evolution* 19, 141-147.

**Wood, JR., Wilmshurst, JM., Rawlence, NJ., Bonner, KI., Worthy, TH., Kinsella, JM., and Cooper, A.** 2013. A megafauna's microfauna: gastrointestinal parasites of New Zealand's extinct moa (Aves: *Dinornithiformes*). *PLoS One* 8, e57315.

**Yashina, S., Gubin, S., Maksimovich, S., Yashina, A., Gakhova, E., and Gilichinsky, D.** 2012. Regeneration of whole fertile plants from 30,000-y-old fruit tissue buried in Siberian permafrost. *Proceedings of the National Academy of Sciences of the United States of America* 109(10):4008-4013.

**Yatsunenko, T., Rey, FE., Manray, MJ., Trehan, I., Dominguez-Bello, MG., Contreras, M., Magris, M., Hildalgo, G., Baldassano, RN., Anokhin, AP., Heath, AC., Warner, B., Reeder, J., Kuczynski, J., Caporaso, JG., Lozupone, CA., Lauber, C., Clemente, JC., Knights, D., Knights, R.,**

**Gorden, JI.** 2012. Human gut microbiome viewed across age and geography. *Nature* 486, 222-227.

**Yll, R.** 2006. Vegetation reconstruction on the basis of pollen in Late Pleistocene hyena coprolites from San Teodoro Cave (Sicily, Italy). *Palaeogeography, Palaeoclimatology, Palaeoecology* 237, 32-39.

# REMERCIEMENTS

*Je tiens à adresser en premier lieu mon profond respect ainsi que mes remerciements à mon directeur de thèse le Pr Michel Drancourt. Il n'a pas simplement accepté de diriger ma thèse; il m'a également permis de confirmer ma passion pour la recherche en biologie. Même si cela n'a pas toujours été facile, il n'a eu de cesse de m'encourager et m'a permis d'apprendre énormément à ses côtés. J'en profite pour lui exprimer ici ma plus profonde gratitude.*

*Je remercie le Pr Didier Raoult pour l'accueil qu'il m'a réservé au sein de son laboratoire.*

*Je tiens également à remercier le membre du jury qui ses sont rendus disponible pour juger la qualité de mes travaux. Merci à Pr Christine Keyser et au Pr Antoine Andremont pour avoir accepter de faire partie de mon jury en qualité de rapporteurs. C'est également un très grand honneur d'avoir comme président du jury le Pr Jean-Louis Mège, dont l'écoute n'a d'égal que ses compétences en pathologie humaine.*

*J'aimerai remercier Christelle Desnues pour son regard critique sur mon travail mais aussi ses conseils d'orientation tout au long de ma thèse. Elle a su être là durant certains moments difficiles et je n'aurai sûrement pas pu mener à bien certains travaux sans son aide.*

*Je tiens aussi à remercier toutes les personnes qui ont participé de près ou de loin à mon travail : l'ensemble des techniciens tous plus serviables les uns que les autres ainsi que tous mes collègues de l'URMITE. Aussi un grand merci à mes amies ici et en Allemagne pour leur soutien et leur amitié.*