Natural mummification of the human gut preserves bacteriophage DNA

Tasha M. Santiago-Rodriguez[1,2,*], Gino Fornaciari[3,4], Stefania Luciani[5], Scot E. Dowd[6], Gary A. Toranzos[7], Isolina Marota[5] and Raul J. Cano[1,2]

[1]Center for Applications in Biotechnology, California Polytechnic State University, San Luis Obispo, CA, USA; [2] Department of Biology, California Polytechnic State University, San Luis Obispo, CA, USA; [3]Department of Translational Research on New Technologies in Medicine and Surgery, Division of Paleopathology, University of Pisa, Pisa, Italy; [4]Center for Anthropological, Paleopathological and Historical Studies of the Sardinian and Mediterranean Populations, Department of Biomedical Sciences, University of Sassari, Sassari, Italy; [5]Laboratory of Molecular Archaeo-Anthropology/ancient DNA, School of Biosciences and Veterinary Medicine, University of Camerino, Camerino, Italy; [6] Molecular Research LP (MR DNA), Shallowater, Texas, USA; [7]Department of Biology, University of Puerto Rico, San Juan, PR

Running title: Gut phageome of mummified human remains

***Corresponding author:** Tasha M. Santiago-Rodriguez

Mailing Address: Center for Applications in Biotechnology, California Polytechnic State University, San Luis Obispo, CA, USA 93407

Email: tasantia@calpoly.edu

This is a pre-copyedited, author produced PDF of an article accepted for publication in FEMS Microbiology Letters following peer review. The version of record is available online at http://dx.doi.org/10.1093/femsle/fnv219

ABSTRACT

The natural mummification process of the human gut represents a unique opportunity to study the resulting microbial community structure and composition. While results are providing insights into the preservation of bacteria, fungi, pathogenic eukaryotes and eukaryotic viruses, no studies have demonstrated that the process of natural mummification also results in the preservation of bacteriophage DNA. We characterized the gut microbiome of three pre-Columbian Andean mummies, namely FI3, FI9 and FI12, and found sequences homologous to viruses. From the sequences attributable to viruses, 50.4% (mummy FI3), 1.0% (mummy FI9) and 84.4% (mummy FI12) were homologous to bacteriophages. Sequences corresponding to the *Siphoviridae, Myoviridae, Podoviridae* and *Microviridae* families were identified. Predicted putative bacterial hosts corresponded mainly to the Firmicutes and Proteobacteria, and included *Bacillus, Staphylococcus, Clostridium, Escherichia, Vibrio, Klebsiella, Pseudomonas* and *Yersinia*. Predicted functional categories associated with bacteriophages showed a representation of structural, replication, integration, and entry and lysis genes. The present study suggests that the natural mummification of the human gut results in the preservation of bacteriophage DNA, representing an opportunity to elucidate the ancient phageome and to hypothesize possible mechanisms of preservation.

Keywords: ancient microbiomes; bacteriophages; microbiome; mummy; phageome; virome

INTRODUCTION

The human gut microbiome is home to diverse communities comprised of bacteria, archaea and eukaryotes (Yatsunenko *et al.*, 2012, Hoffmann *et al.*, 2013); yet, an increasing number of studies have demonstrated that the human gut is also inhabited by diverse viral communities, many of which are bacteriophages (Minot *et al.*, 2011, Minot *et al.*, 2013). Bacteriophages play important roles in biogeochemical cycles (Fuhrman, 1999) and in the evolution of their bacterial hosts (Ai *et al.*, 2000, Bollback & Huelsenbeck, 2001, Coberly *et al.*, 2009, Minot *et al.*, 2013, Cvirkaite-Krupovic *et al.*, 2015); however, we are just beginning to understand the role of bacteriophages as part of the human microbiome (Sun & Relman, 2013, Abeles & Pride, 2014). Previous studies have demonstrated that bacteriophages are part of the human oral (Pride *et al.*, 2012, Edlund *et al.*, 2015), skin (Robles-Sikisaka *et al.*, 2013, Denesvre *et al.*, 2015), genitourinary tract (Santiago-Rodriguez *et al.*, 2015a) and gut microbiomes (Breitbart *et al.*, 2003, Minot *et al.*, 2011, Hofer, 2013, Cadwell, 2015, Ray, 2015). Bacteriophages also have major impacts in human health and disease (Willner *et al.*, 2009, Willner *et al.*, 2011, Ly *et al.*, 2014, Landini *et al.*, 2015, Norman *et al.*, 2015, Santiago-Rodriguez *et al.*, 2015b, Wang *et al.*, 2015). In diseases such as periodontitis, the relative abundance of bacteriophages belonging to the *Myoviridae* family is higher in subjects with the disease compared to subjects with good periodontal health. While results may be influenced by the representation of myoviruses in databases, these bacteriophages are believed to shape oral bacterial communities by lysing their hosts, thus, are believed to promote periodontal disease (Ly *et al.*, 2014, Santiago-Rodriguez *et al.*, 2015b). Other more serious diseases, including inflammatory bowel disease (IBD), have also associated bacteriophages with a dysbiosis of the gut bacterial communities, probably resulting in the disease (Norman *et al.*, 2015).

Human microbiomes dating to hundreds and thousands of years are just beginning to be characterized, and have also been associated with dietary shifts, dietary habits, and periodontal health and disease (Adler *et al.*, 2013, Cano *et al.*, 2014, Warinner *et al.*, 2015, Weyrich *et al.*, 2015). Yet, very few studies have focused on the viral composition of ancient human samples. Previous studies have focused on viruses in ancient human specimens including retroviruses and those from the *Flaviridae*, *Rhabdoviridae*, *Parvoviridae* families (Emerman & Malik, 2010, Patel *et al.*, 2011, Aswad & Katzourakis, 2012, Katzourakis, 2013, Lavialle *et al.*, 2013, Rivera-Perez *et al.*, 2015). It is feasible to hypothesize that ancient microbiomes are also home to a community of bacteriophages homologous to those present in modern human microbiomes. A previous study characterizing the virome of fossilized fecal material from the14th century found that bacteriophages comprised a good proportion of the viral communities (Appelt *et al.*, 2014).

The natural mummification process is also known to preserve ancient microbial DNA due to cold temperatures and low oxygen levels (Cano *et al.*, 2000, Zink *et al.*, 2000, Tito *et al.*, 2012). Our previous study characterizing the gut microbiome of a pre-Columbian Andean mummy identified sequences associated with bacteria, archaea, fungi, pathogenic eukaryotes and eukaryotic viruses (Santiago-Rodriguez *et al.* 2015c); yet, no studies have demonstrated that the process of natural mummification also results in the preservation of bacteriophage DNA. Bacteriophage communities are usually characterized using viral metagenomics, which consists in the enrichment of viruses by CsCl gradient ultracentrifugation (Rosario *et al.*, 2009, Walker, 2010, Ly *et al.*, 2014, Santiago-Rodriguez *et al.*, 2015a). A previous study characterized the viral

communities of a 14th century coprolite using viral metagenomics, but the method needs to be tested in mummified human specimens (Appelt *et al.,* 2014). Shotgun metagenomics has also shown to provide information on microbial communities in ancient human samples (Adler *et al.*, 2013). While shotgun metagenomics is not selective for bacteriophage DNA, it is useful in characterizing phage communities in modern samples (Belda-Ferre *et al.*, 2012, Santiago-Rodriguez *et al.,* 2015b). Therefore, by using metagenomics, we aim to: (i) determine the percentage of sequences associated with bacteriophages, (ii) identify bacteriophages sharing sequence homology to modern bacteriophages, and (iii) determine predicted functional categories associated with bacteriophages in the gut of naturally-preserved human mummies.

MATERIALS AND METHODS

Description of mummified human remains

The specimens studied are presently stored at the Museum of Anthropology and Ethnology of the University of Florence, Italy. Autopsies were performed by paleopathologists G. Fornaciari and colleagues, and specimens were collected from internal organs by cutting the skin and the ribs. The first mummy, FI3, was an adult male dating to the 14-15th century which showed a good preservation of the skin with the adnexa and a massive presence of fungi and ectoparasites. The presence of microscopic, non-pathological fungi, including the genus *Aspergillus* (easily identifiable with Periodic Acid Schiff staining), is a very common finding in mummies as a post-mortem invasion phenomenon. DNA was extracted from abdominal viscera. The second mummy, FI9, was a female of estimated 18-23 years of age, dating to the 11th century A.D. DNA was extracted from the descending colon, but the ascending and transverse

colon, as well as paleofeces were previously characterized by our group (Santiago-Rodriguez *et al.*, 2015c). The third mummy, FI12, was an adult female, estimated age 20-25 years, and autopsy showed that she was afflicted by bronchopneumonia. An exact date for the mummy is unknown, but it is evident that she belonged to the Inca culture from the fetal position found in the burial basket. DNA was extracted from the transverse colon.

Avoidance of contamination

We employed the standard precautions for ancient DNA work including the use of sterile gloves, pretreatment of mortars, pestles, and homogenizers with HCl, use of UV-irradiated safety cabinets, dedicated gel trays, tanks and reagents. The autopsy was performed by paleopathologists wearing sterile surgical coats, sterile latex gloves, sterile masks, headdresses and overshoes. The outermost portions of the specimens were not used to eliminate the risk of surface contamination, and one replicate per sample was obtained for further analyses. The mummified specimens were immediately kept and sealed in sterile containers, reducing the possibility of subsequent contamination. The samples were stored aseptically in hermetic plastic containers in a dry environment with silica gel at 18-20 °C. DNA extraction and further precautions were performed as described previously (Santiago-Rodriguez *et al.*, 2015c) and are detailed in **Supplementary methods.**

Metagenome analyses for viruses

DNA library preparation for metagenome sequencing was performed at the Next-Generation sequencing provider Molecular Research Laboratory (MRDNA)

(www.mrdnalab.com; Shallowater, TX, USA) under strict procedures to eliminate cross-contamination with modern DNA as described previously (Santiago-Rodriguez *et al.*, 2015c). Libraries were sequenced using Illumina MiSeq following Truseq DNA library preparation protocol, and sequence files were processed as described previously (Santiago-Rodriguez *et al.*, 2015c). Data were then uploaded and annotated using the MG-RAST pipeline and taxonomic assignments were determined using the SEED database with a minimum e-value of 80% (Meyer *et al.*, 2008). To determine the percentage of sequences associated with viruses and the predicted putative hosts at the phylum level, data were acquired from the Virus category. Sequences were also mapped to a virus database that included both prokaryotic and eukaryotic viruses (www.phantome.org; ftp://ftp.ncbi.nih.gov/genomes/Viruses/). Mapping was performed using CLC Genomics Workbench with the following parameters: no masking, mismatch cost=2, insertion cost=3, deletion cost=3, with an 80% identity over a minimum of 50% of the read length. Mapped reads were also retrieved from CLC Genomics Workbench as a SAM file and processed using mapDamage for further ancient DNA authentication as described previously (Ginolhac *et al.*, 2011). Predicted functional categories associated with bacteriophages were analyzed using MG-RAST with a minimum e-value of 80%.

16S rRNA gene analyses

16S rRNA gene data from these mummies were used to associate the phages predicted putative hosts with the bacterial taxonomy at the phylum level. SourceTracker analyses were also performed to identify possible sources of contamination (Knights *et al.*, 2011, Santiago-

Rodriguez *et al.*, 2015c, Santiago-Rodriguez *et al.*, 2015d). 16S rRNA gene methods and analyses are described in **Supplementary Methods**.

RESULTS

Metagenome and 16S rRNA gene high-throughput sequencing data

For the metagenome analyses, a total of 16,805,260 (mummy FI3), 146,081,692 (mummy FI9) and 16,537,474 (mummy FI12) sequences, with an average length of 100 bp were analyzed. For the 16S rRNA gene analyses, a total of 79,752 (mummy FI3), 8,731 (mummy FI9) and 57,979 (mummy FI12) sequences with an average length of 270 bp were acquired, but data was rarefied to 8,000 sequences to minimize the effect of disparate sequence number in the results (Santiago-Rodriguez *et al.*, 2015c). We utilized SourceTracker using 135 human (45 oral, 45 skin, and 45 gut), and 45 soil microbiomes to identify possible sources of contamination in the mummified gut tissues (**Supplementary Figure 1**) (Santiago-Rodriguez *et al.*, 2015c). Given that the samples were obtained from the mummies' colon, it is expected some of the ancient sequences to match modern gut microbiomes, as in the case of mummy FI3. However, given that the sequences did not match the most likely sources of contamination (shown as unknown) in any of the mummies, namely skin and soil microbiomes, suggests that no external sources of contamination contributed to the findings reported in the present study (Santiago-Rodriguez *et al.*, 2015c, Santiago-Rodriguez *et al.*, 2015d). We also performed mapDamage analyses for mummy FI3 mapped viral reads, but did not note the typical DNA damage pattern at the 5' or 3' ends as described for eukaryotic genomes (Knapp *et al.*, 2012, Der Sarkissian *et al.*, 2014) (**Supplementary Figure 2**).

Natural mummification preserves bacteriophage DNA

Metagenome analyses showed that the mummified guts included in the present study had sequence homology to viral genomes, with a proportion corresponding to bacteriophages. A total of 2,198 (mummy FI3), 74,052 (mummy FI9), and 275 (mummy FI12) sequences were homologous to viruses. Approximately 50.4%, 4.0% and 45.6% of the viral sequences in mummy FI3 were homologous to bacteriophages, eukaryotic viruses and unclassified viruses, respectively. Mummy FI9 had the majority of the viral sequences (93.7%) not matching bacteriophages or eukaryotic viruses (unclassified). Approximately 84.4%, 6.9% and 8.7% of the viral sequences in mummy FI12 were homologous to bacteriophages, eukaryotic viruses and unclassified viruses, respectively (**Figure 1A**). Analysis of the phage families showed that 42.9% (mummy FI3), 0.29% (mummy FI9) and 79.9% (mummy FI12) of the sequences were homologous to siphoviruses. Approximately 4.6% (mummy FI3), 0.04% (mummy FI9) and 3.6% (mummy FI12) of the sequences were homologous to myoviruses. Podoviruses represented 0.2% and 0.03% of the viral sequences in mummies FI3 and FI9, respectively. No podovirus-homologous sequences were present in mummy FI12. Microviruses contributed 0.2%, 0.05% and 1.8% of the viral sequences in mummies FI3, FI9 and FI12, respectively. The remaining sequences associated with bacteriophages could not be classified (**Figure 1B**).

Analyses of predicted putative hosts at the phylum level showed that the majority (80.4%) of the sequences associated with bacteriophages in mummies FI3 and FI12 were homologous to those having Firmicutes as the bacterial hosts. Other putative hosts at the phylum

level included the Proteobacteria, Actinobacteria or Cyanobacteria. Notably, mummy FI9 had the majority of the bacteriophage sequences sharing homology to those infecting Proteobacteria (93.6%) (**Figure 2**). 16S rRNA gene analyses showed that the Firmicutes were the most represented bacterial group in mummies FI3 (99.9%), FI9 (98.5%) and FI12 (99.4%) (**Figure 2**).

Mapping results also demonstrated that reads in the mummified guts corresponded to phage homologs. **Table 1** shows examples of the mapping results to presumptive bacteriophages with the highest number of reads and coverage. Examples include *Staphylococcus*, *Cronobacter* and *Brochothrix* phages in mummy FI3, *Lactobacillus* and *Staphylococcus* phages in mummy FI9, and *Bacillus* and *Cronobacter* phages in mummy FI12. Reads also mapped across bacteriophage genomes, although not broadly, in mummies FI3 (**Figure 3A**), FI9 (**Figure 3B**) and FI12 (**Figure 3C**). Examples shown include *Staphylococcus* bacteriophage StB20 (**Figure3A**), *Lactobacillus* phage AQ113 (**Figure 3B**), and Enterobacteria phage phiX174 sensu lato (**Figure 3C**). Regions that mapped to modern Enterobacteria phage phiX174 in mummies FI3 (**Panel A**), FI9 (**Panel B**) and FI12 (**Panel C**) are shown in **Supplementary Figure 3**. Nucleotide differences between modern and ancient sequences are shown in red, and indicate that ancient sequences do not correspond to the standard spike-in control used in Illumina sequencing.

Predicted functional categories associated with bacteriophage genes were divided into structure (head and tail), entry and lysis, integrases, replication, packaging, antirepressors, repressors, virulence, introns and hypothetical proteins. The majority of the bacteriophage

categories in mummy FI3 corresponded to packaging (35.3%), and entry and lysis (31.2%) (**Figure 4A**). Mummy FI9 had the majority of the bacteriophage categories (98.7%) corresponding to integrases (**Figure 4B**). Most of the bacteriophage categories in mummy FI12 corresponded to the entry and lysis (65.1%) (**Figure 4C**).

DISCUSSION

Our data is consistent with authentic ancient DNA, as shown with the high level of fragmentation (Ubaldi *et al.*, 1998) and the SourceTracker analyses (Santiago-Rodriguez *et al*., 2015c). We also performed mapDamage analyses to assess patterns of DNA damage that could be consistent with ancient DNA (Der Sarkissian *et al.*, 2014), but did not note these patterns with the phageomes tested. While mapDamage has proven to be useful in determining DNA damage in eukaryotic genomes (Knapp *et al.*, 2012), the method still needs to be further tested in microbiomes. Damage artifacts are often used to authenticate ancient DNA, but these may represent a challenge when characterizing ancient microbiomes. DNA damage analyses may not always provide reliable information of ancient microbiomes as nucleotide differences could also represent a novel microorganism (Warinner *et al.*, 2015). In addition, ancient DNA originating from microorganisms or eukaryotes would possibly need to be analyzed differently as different degrees of damage may be associated with specific taphonomic conditions.

Little is known about bacteriophage DNA preservation in ancient human specimens (Appelt *et al.*, 2014). Our study adds to the knowledge of ancient viruses by showing that the natural mummification process of the human gut results in the preservation of bacteriophage

DNA. While the recovering and sequencing methods differed, the relative abundances of siphoviruses, myoviruses, podoviruses and microviruses were virtually similar to previously reported viromes in extant human guts and coprolites (Breitbart *et al.*, 2003, Appelt *et al.*, 2014). Of interest was also the presence of sequences that were not homologous to known bacteriophages, consistent with their rapid evolution in modern human guts (Minot *et al.*, 2013). These results are also consistent with previous studies showing that a proportion of viral sequences in the human gut usually cannot be assigned to existing reference genomes (Breitbart *et al.*, 2003, Minot *et al.*, 2011, Muniesa & Jofre, 2014).

While we have demonstrated that phage DNA is preserved in mummified gut tissue, the preservation processes remain to be elucidated. A possible explanation for bacteriophage DNA preservation in mummified gut tissue may include the replication cycles. A proportion of the sequences associated with bacteriophages in the mummies guts corresponded to integrases, antirepressors and repressors. While these genes are known to be markers of lysogeny and may suggest the presence of prophages, it is difficult to demonstrate with our current data that temperate bacteriophages were in a prophage state in the mummified gut tissues. However, it is reasonable to hypothesize that prophages were preserved along with their bacterial host genomes. Of interest is also the possibility that the process of natural mummification resulted in the induction of prophages due to desiccation, which is known to trigger the lytic cycle (Brovko, 2007).

The detection of strictly virulent bacteriophage DNA is intriguing as it may suggest that there may be other mechanisms supporting bacteriophage preservation during the process of natural mummification of the human gut. It is known that lytic bacteriophages are persistent members of the human microbiome and can be detected up to 60 days in the oral cavity, suggesting that they may attach to mucous layers (Barr *et al.*, 2013, Abeles & Pride, 2014, Edlund *et al.*, 2015). Although similar studies have not been carried out in ancient samples, it is feasible that mucous layer(s) in mummified human specimens may also act as a reservoir of lytic bacteriophages (Fornaciari, 1993, Corthals *et al.*, 2012). Intact capsids may aid in the preservation of bacteriophage DNA, but this still needs to be demonstrated with mummified gut tissue using electron microscopy (Appelt *et al.*, 2014). Another possible explanation for the detection of bacteriophages in the mummified guts is their seemingly high proportions in the human gut, where concentrations may range between 10^7 to 10^{10} per gram of feces (Muniesa & Jofre, 2014). This relatively high initial concentration may aid in the detection of bacteriophage DNA even if some inactivation has occurred during natural mummification.

Given that lysogenic bacteriophages co-evolve with their bacterial hosts, culture-independent methods using sequence homology have shown to be relatively accurate in predicting putative bacterial hosts up to the genus level in modern human viromes (Minot *et al.*, 2011, Ly *et al.*, 2014). While the same techniques have shown to not possess this same specificity with strictly lytic bacteriophages, they are still useful in providing insights into predicted putative bacterial hosts (Ly *et al.*, 2014). Previous studies have also associated predicted phage putative hosts with 16S rRNA gene data (Pride *et al.*, 2012, Ly *et al.*, 2014, Abeles *et al.*, 2015, Santiago-Rodriguez *et al.*, 2015a, Santiago-Rodriguez *et al.*, 2015e), but

phageome relative abundances do not always mirror those of their bacterial hosts (Edlund *et al.*, 2015). This may be due to different dynamic relationships present for different host/phage pairs (Abeles *et al.*, 2014, Ly *et al.*, 2014, Abeles *et al.*, 2015, Santiago-Rodriguez *et al.*, 2015a). While associations between the phageome and 16S rRNA gene data exhibit limitations, results may still provide insights into phage-host interactions.

Other possible reasons for specific bacteriophages being preserved in ancient gut phageomes may include differences in gender (Abeles *et al.*, 2014), dietary habits (Minot *et al.*, 2011), and health status (Cadwell, 2015), which are known to affect the phageome in extant human guts. We can only speculate that differences in the mummies gender, dietary habits, culture and health status may influence their phageomes (Santiago-Rodriguez *et al.*, 2015c). This may have been the case for mummy FI9, where, although the metagenome analyses generated > 140 million sequences (compared to >16 million sequences for mummies FI3 and FI12), we can only hypothesized that differences in her phageome may be associated to the mentioned factors. Ideally, virome metagenomics would better capture these trends when compared to shotgun metagenomics. While we are in the process of developing a method to study bacteriophages and other viruses in ancient specimens using viral metagenomics, from the data it is evident that shotgun metagenomics provide insights on bacteriophage sequences in ancient human gut microbiomes. Results also provide insights into bacteriophage community structure and composition in the gut of naturally-preserved mummies.

References

Abeles SR & Pride DT (2014) Molecular bases and role of viruses in the human microbiome. *J Mol Biol* **426**: 3892-3906.

Abeles SR, Ly M, Santiago-Rodriguez TM & Pride DT (2015) Effects of Long Term Antibiotic Therapy on Human Oral and Fecal Viromes. *PLoS One* **10**: e0134941.

Abeles SR, Robles-Sikisaka R, Ly M, Lum AG, Salzman J, Boehm TK & Pride DT (2014) Human oral viruses are personal, persistent and gender-consistent. *ISME J* **8**: 1753-1767.

Adler CJ, Dobney K, Weyrich LS, *et al.* (2013) Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. *Nat Genet* **45**: 450-455, 455e451.

Ai Y, Meng F & Zeng Y (2000) [The evolution of pathogen-host interactions mediated by bacteriophages]. *Wei Sheng Wu Xue Bao* **40**: 657-660.

Appelt S, Fancello L, Le Bailly M, Raoult D, Drancourt M & Desnues C (2014) Viruses in a 14th-century coprolite. *Appl Environ Microbiol* **80**: 2648-2655.

Aswad A & Katzourakis A (2012) Paleovirology and virally derived immunity. *Trends Ecol Evol* **27**: 627-636.

Barr JJ, Auro R, Furlan M, *et al.* (2013) Bacteriophage adhering to mucus provide a non-host-derived immunity. *Proc Natl Acad Sci U S A* **110**: 10771-10776.

Belda-Ferre P, Alcaraz LD, Cabrera-Rubio R, Romero H, Simon-Soro A, Pignatelli M & Mira A (2012) The oral metagenome in health and disease. *ISME J* **6**: 46-56.

Bollback JP & Huelsenbeck JP (2001) Phylogeny, genome evolution, and host specificity of single-stranded RNA bacteriophage (family Leviviridae). *J Mol Evol* **52**: 117-128.

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**: 6220-6223.

Brovko L (2007) Identification of Microorganisms and Specific Detection of Bacteria using Bioluminescence. *Bioluminescence for Food and Environmental Microbiology Safety,*(Brovko L, ed.) p.^pp. 43-50. SPIE Press, Bellingham, Washington, USA.

Cadwell K (2015) The virome in host health and disease. *Immunity* **42**: 805-813.

Cadwell K (2015) Expanding the role of the virome: commensalism in the gut. *J Virol* **89**: 1951-1953.

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

Cano RJ, Rivera-Perez J, Toranzos GA, Santiago-Rodriguez TM, Narganes-Storde YM, Chanlatte-Baik L, Garcia-Roldan E, Bunkley-Williams L & Massey SE (2014) Paleomicrobiology: revealing fecal microbiomes of ancient indigenous cultures. *PLoS One* **9**: e106833.

Coberly LC, Wei W, Sampson KY, Millstein J, Wichman HA & Krone SM (2009) Space, time, and host evolution facilitate coexistence of competing bacteriophages: theory and experiment. *Am Nat* **173**: E121-138.

Corthals A, Koller A, Martin DW, Rieger R, Chen EI, Bernaski M, Recagno G & Davalos LM (2012) Detecting the immune system response of a 500 year-old Inca mummy. *PLoS One* **7**: e41244.

Cvirkaite-Krupovic V, Carballido-Lopez R & Tavares P (2015) Virus evolution toward limited dependence on nonessential functions of the host: the case of bacteriophage SPP1. *J Virol* **89**: 2875-2883.

Denesvre C, Dumarest M, Remy S, Gourichon D & Eloit M (2015) Chicken skin virome analyzed by high-throughput sequencing shows a composition highly different from human skin. *Virus Genes* **51**: 209-216.

Der Sarkissian C, Brotherton P, Balanovsky O, Templeton JE, Llamas B, Soubrier J, Moiseyev V, Khartanovich V, Cooper A & Haak W (2014) Mitochondrial genome sequencing in Mesolithic North East Europe Unearths a new sub-clade within the broadly distributed human haplogroup C1. *PloS one* **9**.

Der Sarkissian C, Brotherton P, Balanovsky O, *et al.* (2014) Mitochondrial genome sequencing in Mesolithic North East Europe Unearths a new sub-clade within the broadly distributed human haplogroup C1. *PLoS One* **9**: e87612.

Edlund A, Santiago-Rodriguez TM, Boehm TK & Pride DT (2015) Bacteriophage and their potential roles in the human oral cavity. *J Oral Microbiol* **7**: 27423.

Emerman M & Malik HS (2010) Paleovirology--modern consequences of ancient viruses. *PLoS Biol* **8**: e1000301.

Fornaciari G (1993) Adenocarcinoma in the mummy of Ferrante I of Aragon, King of Naples (1431-1494). *Paleopathol Newsl* 5-8.

Fuhrman JA (1999) Marine viruses and their biogeochemical and ecological effects. *Nature* **399**: 541-548.

Ginolhac A, Rasmussen M, Gilbert MT, Willerslev E & Orlando L (2011) mapDamage: testing for damage patterns in ancient DNA sequences. *Bioinformatics* **27**: 2153-2155.

Hofer U (2013) Viral evolution: Variation in the gut virome. *Nat Rev Microbiol* **11**: 596.

Hoffmann C, Dollive S, Grunberg S, Chen J, Li H, Wu GD, Lewis JD & Bushman FD (2013) Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents. *PLoS One* **8**: e66019.

Katzourakis A (2013) Paleovirology: inferring viral evolution from host genome sequence data. *Philos Trans R Soc Lond B Biol Sci* **368**: 20120493.

Knapp M, Horsburgh KA, Prost S, Stanton JA, Buckley HR, Walter RK & Matisoo-Smith EA (2012) Complete mitochondrial DNA genome sequences from the first New Zealanders. *Proc Natl Acad Sci U S A* **109**: 18350-18354.

Knights D, Kuczynski J, Charlson ES, Zaneveld J, Mozer MC, Collman RG, Bushman FD, Knight R & Kelley ST (2011) Bayesian community-wide culture-independent microbial source tracking. *Nat Methods* **8**: 761-763.

Landini MM, Borgogna C, Peretti A, *et al.* (2015) Identification of the skin virome in a boy with widespread human papillomavirus-2-positive warts that completely regressed after administration of tetravalent human papillomavirus vaccine. *Br J Dermatol* **173**: 597-600.

Lavialle C, Cornelis G, Dupressoir A, Esnault C, Heidmann O, Vernochet C & Heidmann T (2013) Paleovirology of 'syncytins', retroviral env genes exapted for a role in placentation. *Philos Trans R Soc Lond B Biol Sci* **368**: 20120507.

Ly M, Abeles SR, Boehm TK, Robles-Sikisaka R, Naidu M, Santiago-Rodriguez T & Pride DT (2014) Altered oral viral ecology in association with periodontal disease. *MBio* **5**: e01133-01114.

Meyer F, Paarmann D, D'Souza M, *et al.* (2008) The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **9**: 386.

Minot S, Bryson A, Chehoud C, Wu GD, Lewis JD & Bushman FD (2013) Rapid evolution of the human gut virome. *Proc Natl Acad Sci U S A* **110**: 12450-12455.

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**: 1616-1625.

Muniesa M & Jofre J (2014) Identifying and analyzing bacteriophages in human fecal samples: what could we discover? *Future Microbiol* **9**: 879-886.

Norman JM, Handley SA, Baldridge MT, Droit L, Liu CY, Keller BC, Kambal A, Monaco CL, Zhao G & Fleshner P (2015) Disease-specific alterations in the enteric virome in inflammatory bowel disease. *Cell* **160**: 447-460.

Patel MR, Emerman M & Malik HS (2011) Paleovirology - ghosts and gifts of viruses past. *Curr Opin Virol* **1**: 304-309.

Pride DT, Salzman J, Haynes M, Rohwer F, Davis-Long C, White RA, 3rd, Loomer P, Armitage GC & Relman DA (2012) Evidence of a robust resident bacteriophage population revealed through analysis of the human salivary virome. *ISME J* **6**: 915-926.

Ray K (2015) Gut microbiota: The gut virome and bacterial microbiome-the early years. *Nat Rev Gastroenterol Hepatol*.

Rivera-Perez JI, Cano RJ, Narganes-Storde YM, Chanlate-Baik L & Toranzos GA (2015) Retroviral DNA sequences as a means for determining ancient diets. *PLosOne* **(Accepted)**.

Robles-Sikisaka R, Ly M, Boehm T, Naidu M, Salzman J & Pride DT (2013) Association between living environment and human oral viral ecology. *ISME J* **7**: 1710-1724.

Rosario K, Nilsson C, Lim YW, Ruan Y & Breitbart M (2009) Metagenomic analysis of viruses in reclaimed water. *Environ Microbiol* **11**: 2806-2820.

Santiago-Rodriguez TM, Ly M, Bonilla N & Pride DT (2015a) The human urine virome in association with urinary tract infections. *Front Microbiol* **6**: 14.

Santiago-Rodriguez TM, Naidu M, Abeles SR, Boehm TK, Ly M & Pride DT (2015b) Transcriptome analysis of bacteriophage communities in periodontal health and disease. *BMC Genomics* **16**: 549.

Santiago-Rodriguez TM, Fornaciari G, Luciani S, Dowd SE, Toranzos GA, Marota I & Cano RJ (2015c) Gut Microbiome of an 11th Century A.D. Pre-Columbian Andean Mummy. *PLoS One* **10**: e0138135.

Santiago-Rodriguez TM, Fornaciari G, Luciani S, Dowd SE, Toranzos GA, Marota I & Cano RJ (2015d) Taxonomic and predicted functional profiles of the gut microbiome of naturally-preserved pre-Columbian Andean mummies. *mBio* **(Under review)**.

Santiago-Rodriguez, T.M., Ly, M., Daigneault, M.C., Brown, I.H.L., McDonald, J.A.K., Bonilla, N., Allen-Vercoe, E. and Pride, D.T. (2015e). Chemostat culture systems support diverse bacteriophage communities from human feces. *Microbiome* **3**: 58.

Sun CL & Relman DA (2013) Microbiota's 'little helpers': bacteriophages and antibiotic-associated responses in the gut microbiome. *Genome Biol* **14**: 127.

Tito RY, Knights D, Metcalf J, *et al.* (2012) Insights from characterizing extinct human gut microbiomes. *PLoS One* **7**: e51146.

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

Walker A (2010) Gut metagenomics goes viral. *Nat Rev Microbiol* **8**: 841.

Wang J, Gao Y & Zhao F (2015) Phage-bacteria interaction network in human oral microbiome. *Environ Microbiol*.

Warinner C, Speller C & Collins MJ (2015) A new era in palaeomicrobiology: prospects for ancient dental calculus as a long-term record of the human oral microbiome. *Philos T R Soc B* **370**.

Warinner C, Speller C & Collins MJ (2015) A new era in palaeomicrobiology: prospects for ancient dental calculus as a long-term record of the human oral microbiome. *Philos Trans R Soc Lond B Biol Sci* **370**: 20130376.

Weyrich LS, Dobney K & Cooper A (2015) Ancient DNA analysis of dental calculus. *J Hum Evol* **79**: 119-124.

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.

Willner D, Furlan M, Schmieder R, *et al.* (2011) Metagenomic detection of phage-encoded platelet-binding factors in the human oral cavity. *Proc Natl Acad Sci U S A* **108 Suppl 1**: 4547-4553.

Yatsunenko T, Rey FE, Manary MJ, *et al.* (2012) Human gut microbiome viewed across age and geography. *Nature* **486**: 222-227.

Zink A, Reischl U, Wolf H & Nerlich AG (2000) Molecular evidence of bacteremia by gastrointestinal pathogenic bacteria in an infant mummy from ancient Egypt. *Arch Pathol Lab Med* **124**: 1614-1618.

Table 1. Examples of bacteriophage sequences mapping known bacteriophage genomes in mummies FI3, FI9 and FI12. Bacterial taxonomic classification at the family level was determined using 16S rRNA gene analyses. Family and relative abundance percentages are shown.

Presumptive bacteriophage	Phage family	Number of reads	Average coverage	Accession number	Reference length (bp)
Mummy FI3					
Acinetobacter phage 133	*Myoviridae*	48	7.51E-03	NC_015250.1	159,801
Aeromonas phage 65	*Myoviridae*	93	9.62E-03	NC_015251.1	235,229
Bacillus phage 0305phi8-36	*Myoviridae*	110	0.01	NC_009760.1	218,948
Brochothrix phage A9	*Myoviridae*	295	0.1	NC_015253.1	127,065
Campylobacter phage CP21	*Myoviridae*	101	0.02	NC_019507.1	182,833
Cellulophaga phage phiST	*Siphoviridae*	67	0.03	NC_020842.1	79,114
Clostridium phage CDMH1	*Myoviridae*	60	0.05	NC_024144.1	54,279
Cronobacter phage vB_CsaM_GAP32	*Myoviridae*	411	0.05	NC_019401.1	358,663
Cyanophage P-RSM6	*Myoviridae*	54	6.28E-03	NC_020855.1	192,497
Enterobacteria phage phiX174 sensu lato	*Microviridae*	21	0.26	NC_001422.1	5,386
Erwinia phage Ea35-70	*Myoviridae*	69	5.41E-03	NC_023557.1	271,084
Klebsiella phage K64-1 DNA	*Myoviridae*	54	4.77E-03	NC_027399.1	346,602
Mycobacterium phage Myrna	*Myoviridae*	135	0.02	NC_011273.1	164,602
Pelagibacter phage HTVC008M	*Myoviridae*	77	0.01	NC_020484.1	147,284
Prochlorococcus phage P-HM2	*Myoviridae*	98	0.01	NC_015284.1	183,806
Pseudomonas phage 201phi2-1	*Myoviridae*	93	6.85E-03	NC_010821.1	316,674
Pseudomonas phage PaBG	*Myoviridae*	176	0.03	NC_022096.1	258,139
Sphingomonas phage PAU	*Myoviridae*	90	0.01	NC_019521.1	219,372
Staphylococcus phage StB20	*Siphoviridae*	1127	2.5	NC_019915.1	40,917
Staphylococcus phage Twort	*Myoviridae*	172	0.06	NC_007021.1	130,706
Synechococcus phage ACG-2014f	*Myoviridae*	219	0.04	NC_026927.1	228,143
Synechococcus phage S-SKS1	*Siphoviridae*	149	0.02	NC_020851.1	208,007
Vibrio phage KVP40	*Myoviridae*	65	5.93E-03	NC_005083.2	244,834
Yersinia phage phiR1-37	*Myoviridae*	106	0.01	NC_016163.1	262,391
Mummy FI9					
Enterobacteria phage phiX174 sensu lato	*Microviridae*	416	1.86	NC_001422	5,386
Lactobacillus phage AQ113	*Myoviridae*	4215	0.05	NC_019782	36,566
Staphylococcus phage StB20	*Siphoviridae*	7695	0.04	NC_019915	40,917
Mummy FI12					
Bacillus phage G	*Myoviridae*	263	0.41	NC_023719.1	497,513
Cronobacter phage vB_CsaM_GAP32	*Myoviridae*	168	0.07	NC_019401.1	358663
Enterobacteria phage phiX174 sensu lato	*Microviridae*	47	1.86	NC_001422.1	5,386
Pseudomonas phage phiKZ	*Myoviridae*	51	0.06	NC_004629.1	280,334

*Not detected.

Figure legends

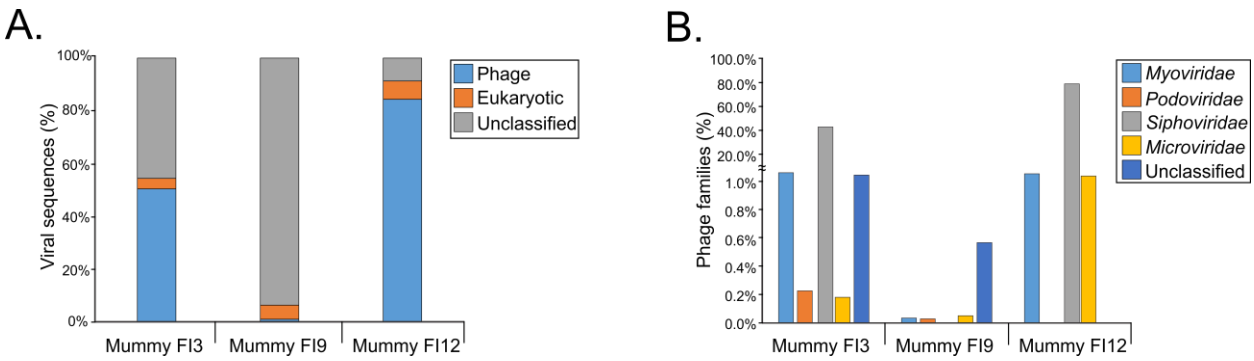

Figure 1. Panel A shows the percentage of sequences homologous to phages, eukaryotic viruses and unclassified viruses. Percentage was calculated based on the total number of sequences corresponding to viruses. Panel B shows the percentage of sequences corresponding to phage families. Families included the *Siphoviridae*, *Myoviridae*, *Podoviridae*, *Microviridae* and unclassified, and were determined based on sequence homology to known phages.

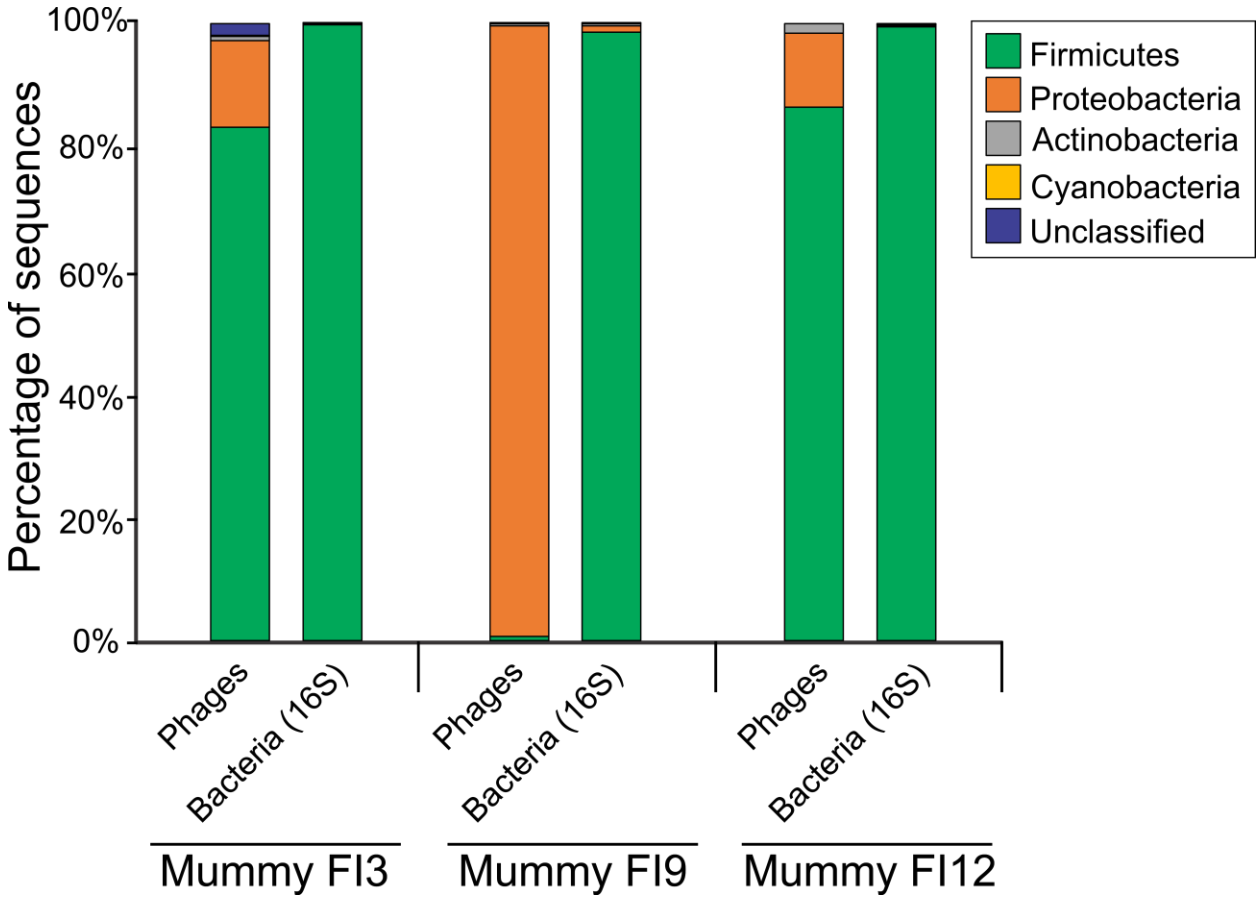

Figure 2. Bacteriophage predicted putative bacterial host at the phylum level was determined based on sequence homology. Predicted putative hosts included the Firmicutes, Proteobacteria, Actinobacteria, Cyanobacteria and unclassified. Figure also shows the bacterial phylum based on analysis of the 16S rRNA gene variable region V4.

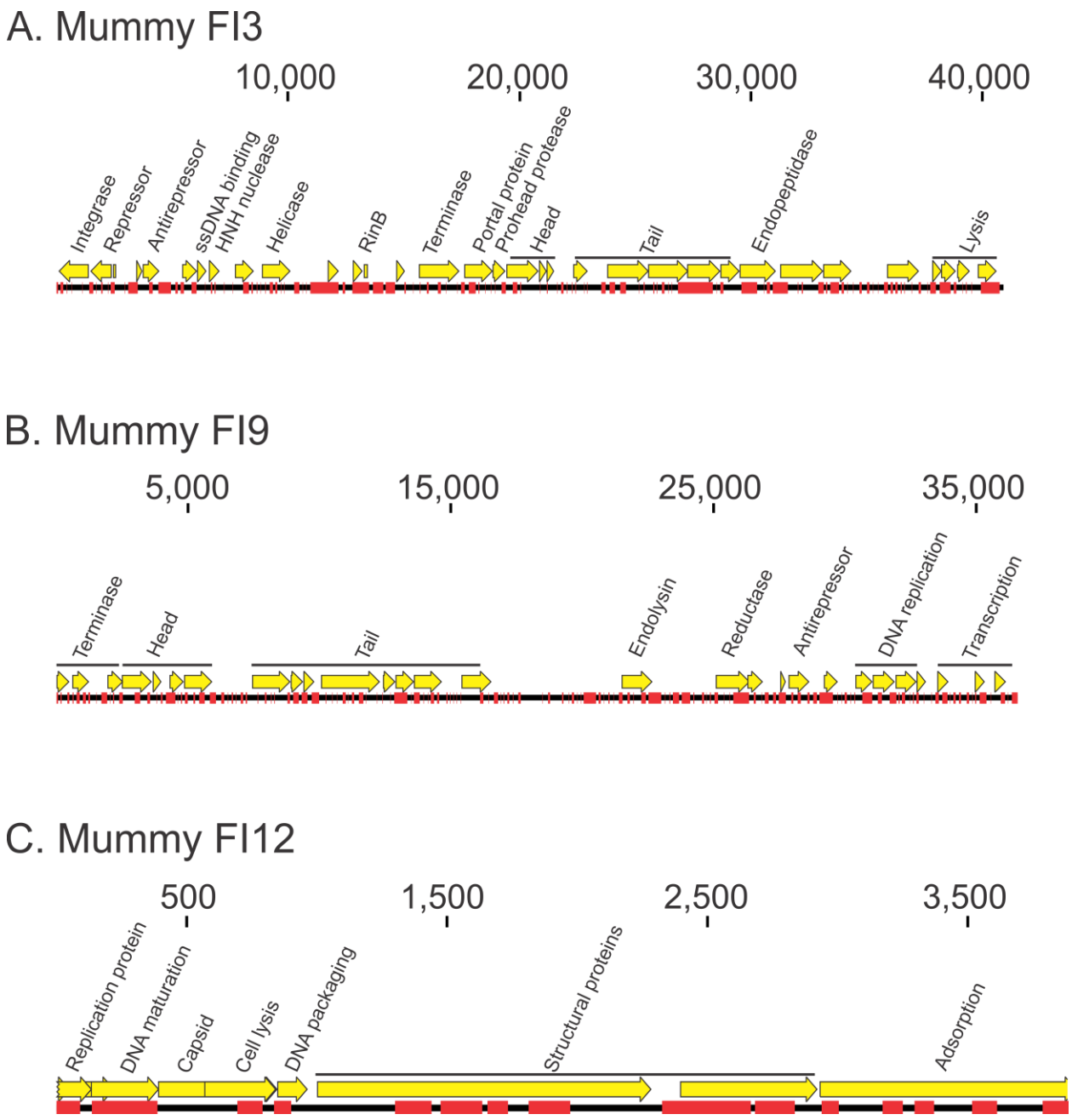

Figure 3. Examples of mapping results in mummies FI3 (Panel A), FI9 (Panel B), and FI12 (Panel C). Examples included presumptive *Staphylococcus* phage StB20 (mummy FI3), *Lactobacillus* phage AQ113 (mummy FI9) and Enterobacteria phage phiX174 sensu lato (mummy FI12). Reads mapping to the phage genomes are shown in red.

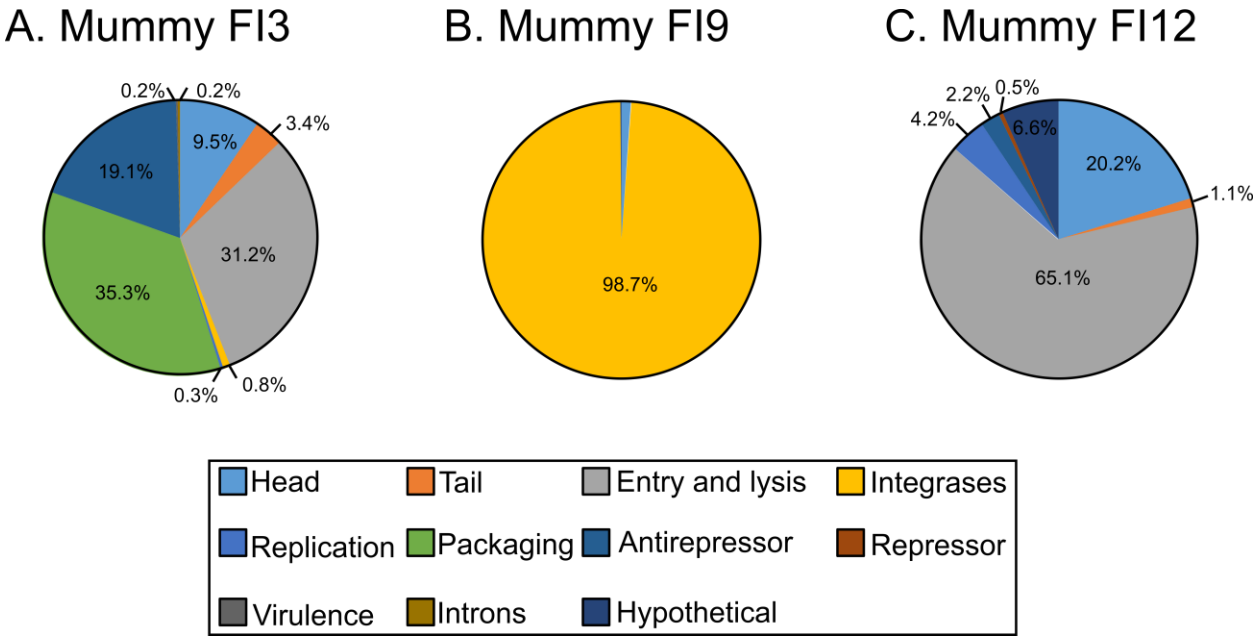

Figure 4. Functional categories attributed to bacteriophages. Categories included structure (head and tail), entry and lysis, integrases, replication, packaging, antirepressors, repressors, virulence genes, introns and hypothetical proteins.

Supplementary Figure 1. Bayesian-SourceTracker analyses for the determination of possible sources of contamination. Gut sequences from mummies FI3 (Panel A), FI9 (Panel B) and FI12 (Panel C) were compared to extant gut (purple), oral (blue), skin (green) and soil (pink) microbiomes. Sequences not matching any of the human and /or soil sources (unknown) are shown in gray.

Supplementary Figure 2. mapDamage results for mummy FI3 viral reads at the first and last 25 bases.

Supplementary Figure 3. Selected regions that mapped to modern Enterobacteria phage phiX174 in mummies FI3 (Panel A), FI9 (Panel B) and FI12 (Panel C). Differences between modern and ancient Enterobacteria phage phiX174 are shown in red.