Pathogen Species Identification from Metagenomes in Ancient Remains: the Challenge of Identifying Human Pathogenic Species of Trypanosomatidae Via Bioinformatic Tools

D Sereno[1*], F Dorkeld[2], M Akhoundi[3], P Perrin[4*]

1 IRD, Montpellier University, InterTryp, Montpellier, France
2 INRA-UMR 1062 CBGP (INRA, IRD, CIRAD), Montpellier SupAgro, Montferrier-sur-Lez, Languedoc-Roussillon 34988, France
3 Parasitology-Mycology Department, Avicenne Hospital, AP-HP, Bobigny, France
4 Montpellier University, IRD, CNRS, MIVEGEC, Montpellier, France

Abstract

Proper species identification from ancient DNA samples is a difficult task that sheds light on the evolutionary history of pathogenic microorganisms. The field of palaeomicrobiology has undoubtedly benefited from the advent of untargeted metagenomic approaches that use next-generation sequencing methodologies. Nevertheless, assigning ancient DNA at the species level is a challenging process. Recently, the gut microbiome analysis of three pre-Columbian Andean mummies [1](Santiago-Rodriguez et al. 2016) has called into question the identification of Leishmania in South America. Here, the metagenomic data filed in MG-RAST (Metagenomics RAST server) were used for a further attempt to identify members of the Trypanosomatidae family infecting these ancient remains. For this purpose, we used two metagenomic analysis tools. In the first step, data were analysed using the ultrafast metagenomic sequence classifier, based on exact alignment of k-mers (Kraken). In the second step, we used Bowtie2, an ultrafast and memory-efficient tool for aligning sequencing reads to long reference sequences. We then compared the output results. These approaches highlight some interesting findings on potential infections by human pathogenic trypanosomatids in these three pre-Columbian mummies.

Keywords
Trypanosomatidae; Kraken taxonomic assignment tool; Bowtie2 fast short reads aligner; Ancient DNA; Parasitome; Co-infection

Corresponding authors: P Perrin Pascale.perrin@umontpellier.fr; D Sereno denis.sereno@ird.fr

1

Introduction

Santiago-Rodriguez et al. [1] reported for the first time some evidence on the occurrence of *Leishmania* DNA in the guts of Andean mummies dating to pre-Columbian times, and they proposed assignment to *Leishmania donovani*. The circulation of such *Leishmania* species in South America is not currently known and has never before been documented in human remains. Therefore, the proper identification of the *Leishmania* species that would have infected mummies before Iberian colonization remains a major concern and would bring new elements to the puzzle of the possible evolutionary scenarios [2]. *Leishmania donovani* and *Leishmania infantum* are, by far, the most common *Leishmania* species responsible for the visceral form (visceral leishmaniasis, VL) of the disease in both the Old World and the New World. So far, all cases of leishmaniasis described on pre-Columbian mummies are reminiscent of cutaneous (CL) or muco-cutaneous (MCL) lesions: CL was observed on a mummy dating to a cultural group from 700-800 AD and found in a cemetery in Peru [3] and MCL on 4 samples in the archaeological cemetery of Coyo Oriente (skulls approximately 1,000 years old) located in the desert of San Pedro de Atacama, Northern Chile [4]. These sites come from a time period predating European contacts. In these remains [4], confirmed a *Leishmania* infection using a PCR approach (amplification of fragments of the IMP dehydrogenase gene, the kinetoplast minicircle, the amino acid permease AAP13LD and the adenylate kinase gene). Further, they mention that the amplified sequences differ from those of *L. donovani*.

The rise of NGS (next-generation sequencing) technologies has opened a new field of systematic investigations in metagenomics. Until recently, BLAST (Basic Local Alignment Search Tool) alignments, which rely on finding the best alignment to a panel of genomic sequences, were the traditional approach to assign a taxonomic label to an unknown sequence. However, unambiguous assignment at the species level is very hard and this tool is very expensive in CPU time for NGS data analysis. To shed light on the causative agent of leishmaniasis infection (and more broadly on parasites belonging to the Trypanosomatidae family) in these ancient remains, we decided to use new software dedicated to metagenomic data analysis. Kraken, a bioinformatic program [5], presents numerous advantages over other programs, including its speed of performing analysis on metagenomes. The identification at the species level by Kraken is based on the use of exact-match database queries of k-mers, rather than on alignment similarity. We applied this new approach to MG-RAST pre-processed metagenomic data of Santiago-Rodriguez et al.'s study [1] and compared the output alignments by using Bowtie2 [6].

Materials and methods

Data

We used the metagenomic data available in MG-RAST (Metagenomics RAST server http://blog.mg-rast.org/) concerning the gut microbiomes of three pre-Columbian Andean mummies, F3, F9 and F12 [1].

Methods

We analysed the data with Kraken (version 1.0), a system for ultrafast metagenomic sequence classification using exact alignment [5] (http://ccb.jhu.edu/software/kraken/). It relies on the development of "a database that contains records consisting of a k-mer and the LCA (lowest common ancestor) of all organisms whose genomes contain that k-mer" [5]. We built a reference database with default parameters of Kraken (k=31). Our custom-built database is composed of all the complete Trypanosomatidae genomes collected from the NCBI

(http://www.ncbi.nlm.nih.gov/genome/) and complemented by genomes from TriptrypDB release 37 (25 April 2018) (Kinetoplastid Genomics Resource http://tritrypdb.org/tritrypdb/) [7]. The total number of Trypanosomatidae genomes in the database is 79 (Table S1). Filtered reads were then compared with our non-redundant, custom-built Trypanosomatidae database, and the results were visualized with a metagenomic visualization tool, Krona [8]. To complete the analysis, the same set of metagenome data was analysed with Bowtie2, a program for rapid alignment of gapped reads [6] (Figure 1).

Results & Discussion

In a first attempt and to compare the respective limits of the two tools in species assignment, we focused on microorganisms belonging to the Trypanosomatidae family. These were chosen for two reasons, first, because of the presence of lesions observed in mummies by [1] and second, because of the endemicity of Chagas disease (WHO www.who.int/chagas/resources/en/) and leishmaniasis in this South American region (WHO www.who.int/leishmaniasis/resources/en/).

Trypanosomes and Chagas disease. The genome representativeness of our database is 55% (11 genomes out of approximately 20 Trypanosoma species currently described) (Figure 2A). The reference database includes trypanosomes responsible for American trypanosomiasis (*T. cruzi*) and trypanosomes responsible for human African trypanosomiasis and animal trypanosomiasis.

The analysis performed with Kraken reveals that a non-negligible proportion of reads (5%, 9% and 10% for FI3, FI9 and F12) is attributed to the Trypanosomatidae family (Figure 3A). The genus *Trypanosoma* was clearly identified in ancient DNA from the guts of the three mummies FI3, FI9 and FI12, with 63%, 69% and 52%, respectively, of the detected Trypanosomatidae community (Figure 3B). As expected, almost all reads belonging to the genus Trypanosoma can be attributed to *T. cruzi* (Figure 4A). These results corroborate the conclusion of Santiago-Rodriguez et al.[1]. Currently, *T. cruzi* is split into six genetic lineages or discrete typing units (DTUs) named TcI, TcIV, TcII, TcIII, TcV, and TcVI, respectively, and a seventh one called TcBat. For both FI3 and FI9, the highest number of reads matches the *T. cruzi* strain Tula cl2 (Table S2), which belongs to DTU I. This DTU is widely represented in the genomic database we gathered (5 of the 11 genomes currently available). For FI12, reads that match Tula cl2 are scarce (Table S2). Nevertheless, because of the relatively low proportion (<10%) of reads that match with a genome filed in our database, it is probable that the DTU of the infecting *T. cruzi* strain is not yet represented.

Analysis with Bowtie2 clearly confirmed a *T. cruzi* infection (53%, 49% and 57% for FI3, FI9 and FI12, respectively) in these mummies (Figure 3C). A majority of the reads matched the *T. cruzi* strains Y and Tula cl2 (Table S3) rather than other genomes filed in the database. Concerning the alignment with *T. cruzi* Y, the reads matched 178 and 203 contigs in FI3 and FI9, respectively, out of 9821 (Table S4). A large majority of reads (80% in FI3 and 90% in FI9) matched 52 identical contigs for both mummies. For *T. cruzi* Tula cl2, reads of mummy FI3 and mummy FI9 matched 42 and 39 contigs, respectively, out of 5300, and almost 100% of the reads for both mummies matched 28 identical contigs (Table S4). Overall, this set of results indicates that the infecting *T. cruzi* strain is probably not in our reference database.

A large majority of the genome of *Trypanosoma cruzi* strains is lacking: currently, out of more than 1,902 strains inventoried to date (Brenière et al. 2016), complete genomes are deposited in the NCBI and/or Tritryp databases for only 11. Likewise, the database we used does not contain representative isolates belonging to DTUs III, IV or VII (TcBat). Interestingly, ancient *T. cruzi* DNA was also identified in human mummies dating from the same period (Chinchorro culture) and in the same geographical region, Southern Peru [9]. In Bolivia and

Peru, strains belonging to the DTU I and DTU V clades are the main *T. cruzi* strains isolated, followed by strains of DTUs IV, III and VI [10]. The knowledge of the complex evolutionary history of *T. cruzi*, which involved genetic exchanges [11], and the existence of hybrid DTUs will certainly benefit from the identification of the infecting strains at the DTU level from this ancient DNA material. Insight into the circulating strains in these mummies would generate important elements for the calibration of such a dynamically evolving scenario. Nevertheless, our analysis highlights that the ancient *T. cruzi* DNA present in the three mummies cannot be assigned to *T. cruzi* CL Brener (DTU VI).

Leishmania and leishmaniasis. All the Leishmania genomes currently available in NCBI or in TRITRYPdb are included in the reference database. The database genome representativeness is 38%, with 21 genomes out of 54 *Leishmania* species currently identified [12]. The representativeness of the database appears to be far better for microorganisms belonging to the subgenera *Leishmania* and *Viannia* (Figure 2B). Concerning human pathogenic *Leishmania* species, the representativity is far better than 70% (14 genomes out of 20), with the subgenera *Leishmania* and *Viannia* well represented (80 and 83%, respectively). However, no genome available concerning the subgenus *Porsicia* (*Paraleishmania* section) is currently available (Table 1 below). This subgenus includes some human pathogenic *Leishmania sp.*, such as *L. colombiensis*.

Subgenus	*Leishmania*	*Viannia*	*Sauroleishmania*	*Mundinia*	*L. australiensis*	*Porsicia*	*Endotrypanum*
Number of species pathogenic for humans	10	6	0	1	1	2	0
Number of genomes available	8	5	0	1	0	0	0
% of genomes available	**80**	**83**	**0**	**100**	**0**	**0**	**0**

Table 1. Representativity of sequenced genomes available in relation to the number of known *Leishmania* species pathogenic for humans, according the updated classification of Akroundi et al. 2017)[12].

The NGS reads analysed with the metagenomic sequence classifier Kraken highlight a probable co-infection with Leishmania parasite for the 3 mummies FI3, FI9 and FI12 (Figure 3B), with 3%, 3% and 10% of the reads attributed to Trypanosomatidae that belong to the genus *Leishmania*, respectively. From figure 4, it is clear that Kraken assigns reads to a large number of *Leishmania* species, each with a low percentage. Clearly, no predominant *Leishmania* species is detected, and *Leishmania donovani* does not appear as a probable infecting *Leishmania* species in all these mummies. Therefore, these results prompt us to question the *Leishmania* species infecting the mummies. Because DNA originates from internal tissues, we have to look for *Leishmania sp.* currently known to affect mainly internal organs and cause VL, namely, *L. donovani, L. infantum, L. tropica, L. martiniquensis, L. colombiensis*. At this point, we can exclude an infection by the first four L. species cited above. The sequence of the *L. colombiensis* genome (*Paraleishmania* section) is currently not available and may be the Leishmania agent we are looking for. Currently, in Peru, the country that mummies originate, Leishmania pathogenic for humans include *L. peruviana*, *L. guyanensis, L. amazonensis*, *L. lainsoni* and *L. braziliensis* [12]. Some *Leishmania* species that usually cause cutaneous forms have, under some circumstances, the capacity to disseminate into internal organs [13,14]. Such unusual clinical presentation is frequently associated with an immunocompromised state or

dysfunction of the T-helper-mediated immune response [2]. Our study clearly shows that a low proportion of reads are attributed to *L. peruviana*, *L. amazonensis*, *L. braziliensis* and *L. guyanensis* (species causing cutaneous leishmaniasis (CL)). Such low frequency in read assignment calls into question the validity of the identification (Figure 4) (Table S2). Nevertheless, the occurrence of an infection by multiple *Leishmania sp.* cannot be ruled out. From these results, it is clear that to be accurate in its species assignment, the reference database used by Kraken needs to be highly representative of the described species and strains. The more exhaustive the database, the better the assignment accuracy will be.

The alignment program Bowtie2 detects the presence of Leishmania in the three mummies (FI3, FI9 and FI12) (Figure 3C). But unlike the mummy FI12, in the mummies FI3 and FI9, the reads match a large panel of *Leishmania sp.,* especially in the mummy FI3 (Table S3). This differs from the results performed with Kraken. However, it is not possible to unequivocally assign DNA to a specific *Leishmania* species and *L. donovani* in particular.

Our analysis reveals that the genome of the infecting *Leishmania* species, whose DNA is detected in these ancient remains, is probably not present in our reference database. We may consider the occurrence of an ancient Leishmania species. Nevertheless, since our reference database is not exhaustive, we cannot rule out the presence of *Leishmania sp.* causing cutaneous or muco-cutaneous lesions or the co-infection of multiple *Leishmania species*. Still, the detection of *Leishmania* DNA is clearly not the result of an artefact or a contamination.

Other Leishmaniinae (Crithidiatae). *Trypanosoma* and *Leishmania* are dixenous parasites, meaning that their life cycle includes an invertebrate as a first host and a vertebrate as a second host. Several genomes of monoxenous (one-host) Trypanosomatidae are available and belong to the genera *Angomonas*, *Crithidia*, *Leptomonas*, *Lotmaria* and *Strigomonas* (Table S1). The database genome representativeness for *Leptomonas* is 5%, with only 2 genomes available out of 39 *Leptomonas* species currently described (Figure 2C). The analysis performed on the three mummies revealed that a relatively high proportion of reads attributed to the family Trypanosomatidae are reminiscent of the genus *Leptomonas* (13%, 18%, and 13% for FI3, FI9 and FI12, respectively) (Figure 3B). A large majority of these reads are assigned to *Leptomonas seymouri* and not *Leptomonas pyrrochoris* after an analysis with Kraken (Figure 4) (Table S2). Nevertheless, because of the lack of available genomes for 37 out of 39 *Leptomonas sp.*, the assignment to *L. seymouri* has to be taken with caution, and an assignment at the strain level cannot be made.

Bowtie2 also confirms the recurrent occurrence of *Leptomonas seymouri* in all mummies (Figure 3C), with a high number of reads matching the *Leptomonas seymouri* strain BHU-1095 (Table S3) in mummies FI3 and F9 (56245 and 125577, respectively). In both mummies, the reads match with 8 and 9 scaffolds (out of 1216), respectively, and a large majority of reads match on only 5 scaffolds (Table S4).

Leptomonas species are usually found in the gut of insects, but they have the potential to infect mammals as an opportunistic parasite. Nevertheless, their infective capacity in mammals seems to be limited to immunocompromised hosts [15]. Interestingly, *Leptomonas seymouri* has been repeatedly isolated from VL patients infected by *Leishmania donovani* in India [16].

To our surprise, the analysis performed with Kraken reveals that an equivalent proportion of reads (2%) is assigned to *Lotmaria passim* in the three mummies (Figure 3B and 4). Via Bowtie2 analysis, a very low number of reads match this genus in the mummies (Figure 3C and Table S3). *Lotmaria passim* is the founding member of this genus (Figure 2). This member of Crithidiatae is described as a common parasite of the honey bee *Apis mellifera* [17], which calls into question its presence in a human gut microbiome.

Conclusion

The strength of our approach is to work also with the most exhaustive reference database of Trypanosomatidae genomic sequences possible and two programs with complementary approaches in metagenomic analysis (k-mer searching vs local alignment) to gain insight into the identity of the infecting Trypanosomatidae. Analysis using the sequence classifier (Kraken) unambiguously confirmed *Trypanosoma cruzi* infection and undoubtedly Leishmania infection in the three Andean mummies. In addition, our analysis provides new information on co-infection by at least two human pathogenic trypanosomatids, Leishmania species and *T. cruzi*, in all mummies with available metagenomes. This type of co-infection is known to occur in human individuals as well in some wild mammals [18][19]. Unfortunately, it is not possible to go further in the assignment to the species level for *Leishmania* and DTU level for *T. cruzi*. This may be due to the lack of a number of *Leishmania* genomes, particularly those of the *Paraleishmania* section, and of members of *T. cruzi* belonging to some DTUs. It also highlights a pattern of poly-infection coupled with an opportunistic trypanosomatid, i.e., *Leptomonas seymouri* [20]. Therefore, future studies with an exhaustive reference database are necessary to better understand the interrelationships that shape the microbial community and play a role in the evolution of the parasitome.

Supplementary materials

Supplementary File 1, Table S1. List of Trypanosomatidae genome sequences available and included in the reference database.
Supplementary File 2, Table S2. A) Results of Kraken analysis on pre-processed metagenomic data for the mummy FI3 (mgm_46.29033.3); B) results of Kraken analysis on pre-processed metagenomic data for the mummy FI9 (mgm_46.30170.3); C) results of Kraken analysis on pre-processed metagenomic data for the mummy FI12 (mgm_46.26489.3).
Supplementary File 3, Table S3. Results of Bowtie2 analysis concerning the three mummies FI3, FI9 and FI12.
Supplementary File 4, Table S4. Distribution of reads matching a) *Leptomonas seymouri*, b) *Trypanosoma cruzi* Y and c) *Trypanosoma cruzi* Tula cl2 on the contigs for the three mummies

Author contributions

PP and DS drafted the manuscript; FD conducted all the bioinformatic analyses. All authors finalized the manuscript. All authors read and approved the final version of the manuscript.

Funding

Conflicts of interest

The authors declare no conflict of interest.

References

1. Santiago-Rodriguez T.M.; Fornaciari G.; Luciani S.; Dowd S.E.; Toranzos G.A.; Marota I.; Cano R.J. Taxonomic and predicted metabolic profiles of the human gut microbiome in pre-Columbian mummies. *FEMS Microbiol Ecol.* **2016**, 92. pii: fiw182. doi: 10.1093/femsec/fiw182.

2. Sereno, D.; Akroundi, M.; Dorkeld, F.; Oury, B.; Momen, H.; Perrin, P. What pre-Columbian mummies could teach us about South American leishmaniases ? *Pathog. Dis.* **2017**, 75, ftx019. doi: 10.1093/femspd/ftx019.

3. Gerszten E.; Allison M.J.; Maguire B. Paleopathology in South American mummies: a review and new findings. *Pathobiology.* **2012**, 79, 247-256. doi: 10.1159/000334087.

4. Costa, M.A.; Matheson, C.; Lachetta, L.; Llagostera, A.; Appenzeller, O. Ancient Leishmaniasis in a highland desert of Northern Chile. *PLoS One* **2009** , 4 , e6983, doi: 10.1371/journal.pone.0006983.

5. Wood, D.E.; Salzberg, S. L. Kraken: ultrafast metagenomic sequence classification using exact alignment. *Genome Biology.* **2014**, 15, R46. doi: 10.1186/gb-2014-15-3-r46

6. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie2. *Nat. Methods.* **2012,** 9, 357-359. doi. 10.1038/nmeth.1923.

7. Aslett, M.; Aurrecoechea, C.; Berriman, M.; Brestelli, J.; Brunk, B.P.; Carrington, M.; Depledge, D.P.; Fischer, F.; Gajria, B.; Gao, X.; Gardner, M.J.; Gingle, A.; Grant, G.; Harb, O.S.; Heiges, M.; Hertz-Fowler, C.; Houston, R.; Innamorato, F.; Lodice, J.; Kissinger, J.C.; Kraemer, E.; Li, W.; Logan, F.J.; Miller, J.A.; Mitra, S.; Myler, P.J.; Nayak, V.; Pennington, C.; Phan, I.; Pinney, D.F.; Ramasamy, G.; Rogers, M.B.; Roos, D.S.; Ross, C.; Sivam, D.; Smith, D.F.; Srinivasamoorthy, G.; Stoeckert, C.J.; Subramanian, Jr S.; Thibodeau, R.; Tivey, A.; Treatman, C.; Velarde, G.; Wang, H. TriTrypDB: a functional genomic resource for the Trypanosomatidae. *Nucl. Acids Res.* **2010,** 38 (Database issue), D457-D462; doi: 10.1093/nar/gkp851.

8. Ondov, B.D.; Bergman N.H.; Phillippy, A.M. Interactive metagenomic visualization in a Web browser. *BMC Bioinformatics.* **2011,** 12, 385 https://doi.org/10.1186/1471-2105-12-385.

9. Aufderheide, A.C.; Salo, W.; Madden, M.; Streitz, J.; Buikstra, J.; Guhl, F.; Arriaza, B.; Renier, C.; Wittmers, L.E.; Jr., Fornaciari, G.; Allison, M. A 9,000-year record of Chagas' disease. *Proc. Natl. Acad. Sci. USA* **2004,** 101, 2034–2039. .doi: 10.1073/pnas.0307312101.

10. Brenière, S.F.; Waleckx, E.; Barnabé, C. Over six thousand *Trypanosoma cruzi* strains classified into discrete typing units (DTUs) : attempt at an inventory. *PLoS Negl. Trop. Dis.* **2016,** 10, e0004792. doi: 10.1371/journal.pntd.0004792.

11. Tomasini N. ; Diosque P. Evolution of *Trypanosoma cruzi* : clarifying hybridisations, mitochondrial introgressions and phylogenetic relationships between majour lineages. *Mem. Inst. Oswaldo Cruz.* **2015,** 110, 403-413. doi: 10.1590/0074-02760140401.

12. Akhoundi, M.; Downing, T.; Votypka, J.; Kuhls, K.; Lukes, J.; Cannet, A.; Ravel, C.; Marty, P.; Delaunay, P.; Kasbari, M.; Granouillac, B.; Gradoni, L.; Sereno, D. Leishmania infections: molecular targets and diagnosis. *Mol. Aspects Med.* **2017,** 57, 1-29. doi: 10.1016/j.mam.2016.11.012.

13. Barral, A.; Badaro, R.; Barral-Netto, M.; Jr. Grimaldi, G.; Momen, H.; Carvalho, E.M. Isolation of *Leishmania mexicana amazonensis* from the bone marrow in a case of American

visceral leishmaniasis. *Am. J. Trop. Med. Hyg.* **1986,** 35, 732-734. doi: 10.4269/ajtmh.1986.35.732.

14. Barral, A.; Pedral-Sampaio, D.; Jr Grimaldi, G.; Momen, H.; MacMahon-Pratt, D.; Ribeiro de Jesus, A.; Almeida, R.; Badaro, R.; Barral-Netto, M.; Carvalho, E.M. et al. Leishmaniasis in Bahia, Brazil: evidence that *Leishmania amazonensis* produces a wide spectrum of clinical diseases. *Am. J. Trop. Med. Hyg.* **1991,** 44, 536-546. PMID: 2063957.

15. Kraeva, N.; Butenko, A.; Hlavacova, J.; Kostygov, A.; Myskova, J.; Grybchuk, D.; Lestinova, T.; Votypka, J.; Volf, P.; Opperdoes, F.; Flegontov, P.; Lukes, J.; Yurchenko, V. *Leptomonas seymouri*: Adaptations to the Dixenous Life Cycle Analyzed by Genome Sequencing, Transcriptome Profiling and Co-infection with *Leishmania donovani*. *PLoS Pathogens.* **2015**, 11, e1005127.

16. Ghosh S.; Banerjee P.; Sarkar A. ; Datta S. ; Chatterjee M. ; Coinfection of *Leptomonas seymouri* and *Leishmania donovani* in Indian leishmaniasis. *J. Clin. Microbiol.* **2012** 50:2774-8. doi: 10.1128/JCM.00966-12.

17. Tritschler, M.; Retschnig, G.; Yañez, O.; Williams, G.R.; Neumann, P. Host sharing by the honey bee parasites *Lotmaria passim* and *Nosema ceranae. Ecol. Evol.* **2017**, 7, 1850-1857. doi: 10.1002/ece3.2796.

18. da Costa A.P. ; Costa F.B. ; Soares H.S. ; Ramirez D.G.; Mesquita E.T. ; Gennari S.M., ; Marcili A. *Trypanosoma cruzi* and *Leishmania infantum chagasi* Infection in Wild Mammals from Maranhão State, Brazil. *Vector Borne Zoonotic Dis.* **2015,** 15, 656-66, doi: 10.1089/vbz.2015.1771.

19. Martínez D.Y.; Verdonck K.; Kaye P.M.; Adaui V.; Polman K.; Llanos-Cuentas A.; Dujardin J.C.; Boelaert M. Tegumentary leishmaniasis and coinfections other than HIV. *PLoS Negl Trop Dis.* **2018,** 12, e0006125, doi:10.1371/journal.pntd.0006125.

20. Dedet, J.P.; Pratlong, F. *Leishmania*, *Trypanosoma* and monoxenous trypanosomatids as emerging opportunistic agents. *J. Eukaryot. Microbiol.* **2000,** 47, 37-39. doi: 10.1111/j.1550-7408.2000.tb00008.x

Figure Legends

Figure 1. Schematic comparison of the Bowtie aligner with the Kraken taxonomic classifier.

Figure 2. Proportion of available genomes (in blue) on the number of known species (in red) in various genera of Trypanosomatidae. (A) in *Trypanosoma*, (B) in the genus *Leishmania*, (C) in other genera of Trypanosomatidae.

Figure 3. Composition of gut microbiota for the three Andean mummies studied by Santiago-Rodriguez et al. [1]: FI3, FI9 and FI12, respectively. (A) The proportion of reads assigned to the Trypanosomatidae family for each mummy by the Kraken analysis. Reads not matching Trypanosomatidae sequences appear in grey. Reads matching Trypanosomatidae sequences are coloured in orange. (B) In Trypanosomatidae: numbers and percentages of reads assigned to the genera *Leishmania* (blue), *Leptomonas* (red), *Lotmaria* (green), and *Trypanosoma* (purple). Reads from mummies FI3, FI9, and FI12 and matching any other Trypanosomatidae are shown in orange. (C) Results obtained by the Bowtie2 alignment sequence for mummies FI3, FI9 and FI12, respectively. We used the same colour reference for a direct comparison.

Figure 4. Profiles of colon microbiomes from three Andean mummies [1]: FI3, FI9 and FI12, for Trypanosomatidae. Circles represent taxonomic classification in ascending order up to the species level (outermost circle) with their relative abundance. These graphs were generated using the program Krona. Less-abundant taxa are listed outside the charts.