Research Proposal and Analytical Protocol

20 January 2019

Multi-proxy analyses of Iron Age palaeo-faecal specimens from Bushman Rock Shelter, Limpopo Province, South Africa

Riaan F. Rifkin [1,2]

[1] Center for Microbial Ecology and Genomics (CMEG), Department of Genetics, University of Pretoria, Hatfield, South Africa
[2] Human Origins and Palaeo-Environments Research Group (HOPE), Oxford Brookes University, Oxford, UK

Summary

This proposal forms part of a project which aims to cultivate the currently undeveloped southern African ancient DNA (aDNA) research niche and includes the establishment of a local aDNA analytical facility at the Department of Genetics, University of Pretoria. Currently, a manuscript following the metagenomic analyses of an ancient (c. AD 1420) human faecal specimen recovered from the Iron Age levels at Bushman Rock Shelter (BRS), Limpopo Province, is under review at Science Advances. The aim of this application is to apply the same multi-proxy analytical protocol (including shotgun metagenome sequencing, SEM analyses, intestinal parasitic analyses, 14C dating, and isotope analyses) to five additional ancient human faecal specimens.

Introduction

The archaeological incidence of human faecal material provides a rare opportunity to explore the ancestral human intestinal microbiome (IM). The human gastrointestinal tract (GI) harbours a dynamic population of bacteria, archaea, fungi, protozoa and viruses; the intestinal microbiota. This collection of microorganisms, comprising the human intestinal microbiome (IM) (1) performs critical functions in digestion, development, behaviour and immunity (2, 3). Modifications of the core IM composition (dysbiosis) have been associated with the pathogenesis of inflammatory diseases and infections (3, 4), including autoimmune and allergic diseases, obesity, inflammatory bowel disease and diabetes (5). Despite its clinical importance, the factors that contribute to changes in IM taxonomic composition and functionality, are not entirely understood (6, 7). This is attributed to the fact that most of what is known about the human IM, is based on contemporary industrialised and 'traditional' human societies (8-10). In evolutionary terms, our species have subsisted by hunting and gathering for >90% of our existence (11). Evidence derived from the analyses of the IMs of traditional societies, including the Tanzanian Hadza hunter-gatherers (8), the Venezuelan Yanomami Amerindians (5), the BaAka Pygmies in the Central African Republic (12) and the Arctic Inuit (13) are thus widely viewed as representing 'snapshots' of ancient human IM composition. However, as exposure to Western diets, medicines and microbes cannot be excluded, one must be cautious about the use of these ethnographic cohorts as proxies for prehistoric human IMs (14).

The transformation of the IMs of present-day humans to their current 'modernised' state commenced millennia ago, with the advent of the Neolithic, which, at *c*. 12,000 years ago (kya), resulted in the first major human dietary transition (15). But precisely how our IMs changed following the advent of the Neolithic, and the Industrial Revolution after *c*. AD 1760, remains ambiguous (16-18). In this regard, the analyses of ancient human IMs provide a unique view into the co-evolution of microbes and human hosts, host microbial ecology and changing human IM-related health states through time2. Indeed, over the past 15 million years, multiple lineages of intestinal bacterial taxa arose via co-speciation with African hominins and non-human primates, i.e., chimpanzees, bonobos and gorillas (19). The departure of behaviourally 'fully-modern' *Homo sapiens* from Africa *c*. 75 kya resulted in the global dispersal of our species (20). Significantly, various microbes accompanied these human dispersals 'out of Africa' (21, 22). Since the ancestral human IM is estimated to comprise a taxonomically and metabolically more diverse array of microbes than those found in contemporary societies (6,10), the IMs of pre-Clovis North Americas (23), pre-Columbian Puerto Rican

Amerindians (24) and pre-Columbian Andeans (25) represent more accurate indications of ancient human IM composition.

Objective

Despite the fact that African populations are not underrepresented in studies concerning 'traditional' human IMs (8, 12, 26), there is, currently, no information concerning the taxonomy and metabolic capacity of ancestral (i.e., archaeologically-derived) African IM composition. To gain insight into the ancient African human IM, and the prehistoric incidence of intestinal parasites, pathogenic microbes and antibiotic resistance genes, we performed shotgun metagenomic sequencing of a prehistoric (pre-colonial) faecal specimen recovered from an Iron Age (*c*. AD 1420) context at Bushman Rock Shelter (BRS), Limpopo Province, South Africa. Comparison with ancient (Ötzi), traditional (Hadza and Malawian) and contemporary 'Western' (Italian) IM datasets indicate that the IM of the BRS individual represents a unique taxonomic and metabolic configuration observed in neither contemporary African, nor European, populations. Specifically, the results of this study indicate that the IM of the Iron Age (*c*. AD 1420) Bantu-speaking individual exhibits IM features characteristic of a transitional forager-agro-pastoralist diet. Comparison with the Tyrolean Iceman, Hadza hunter-gatherers, Malawian agro-pastoralists and contemporary Italians, reveals that the BRS IM precedes IM adaptation to 'Western' diets, including the consumption of coffee, tea, chocolate, citrus and soy, and the use of antibiotics, analgesics and exposure to toxic environmental pollutants. Our analyses therefore elucidates the ways in which human IMs responded to recent dietary changes, prescription medications and environmental pollutants, providing insight into human IM evolution since the advent of the Neolithic *c*. 12,000 years ago.

However, these results are derived from only a single ancient faecal sample representing a single human individual. Five additional palaeo-faecal specimens have since been recovered from BRS, and the aim of this application is to apply the recently-devised multi-proxy analytical protocol (including shotgun metagenome sequencing, SEM analyses, intestinal parasitic analyses, 14C dating, and isotope analyses) to these ancient samples. The five palaeo-faecal specimens were recovered *in situ* from an exposed stratigraphic section (labelled 'Angel') at BRS (Table 1). The occupation level from which the specimens derive comprises the two upper layers of the rock-shelter (*i.e.*, Layers 1 and 2) and relates to the arrival of Bantu-speaking Iron Age agro-pastoralist communities in the region after *c.* 1,800 years ago (ya). This occupation reflects the advent of the Neolithic in South Africa, which entailed the introduction of domesticated taxa such as sorghum (*Sorghum bicolor*), cattle (*Bos taurus*) and various other Iron Age-related species and cultural practices (*i.e*., ceramic and iron-smelting technologies) into the region. All the preceding archaeological layers at BRS are representative of occupations by Holocene (*i.e*., the Oakhurst and Robberg techno-complexes ~10 kya) and Pleistocene (the Pietersburg techno-complex ~70 kya) hunter-gatherers (27, 28).

Table 1. The five palaeo-faecal specimens to be analysed during the proposed study.

Sample	Square	Unit (layer)	Plot number (comments)	SAHRIS ObjectID	ObjectAutoID
BRS S1 2019	N3D	Angel 5	23	[nid: 520602]	62542
BRS S2 2019	M3A	Angel 5	'coprolite (human)'	[nid: 520603]	62543
BRS S3 2019	M3A	Angel 2	'coprolite'	[nid: 520604]	62544
BRS S4 2019	M3D	Angel 2	'copro (human)'	[nid: 520605]	62545
BRS S5 2019	O3A	Angel 6	36	[nid: 520606]	62546

Outcomes

Previously, we performed a comprehensive analysis of an ancient palaeo-faecal specimen derived from an early-15th century Iron Age (Neolithic) South African Bantu-speaking hunter-pastoralist. Although representative of the IM composition, metabolic capacity and ARG configuration of the distal (*i.e.*, the colon including the cecum, rectum and anal canal) IM of a single human individual, the characterisation of an authenticated ancient African Bantu-speaker IM is an important step towards understanding the ancestral (*i.e.*, pre-colonial African) state of the human IM. The large number of unassigned sequence reads in the analysed palaeo-faecal specimen (90.98%) is suggestive of substantial unknown IM taxonomic diversity and metabolic functionality. In the future, the identification of these and other novel taxa and metabolic capacities might have significant implications for identifying health risks specific to the sub-Saharan African Bantu-speaker population which has increased in prevalence with the adoption of Western diets, medical treatments and exposure to modern pollutants. Given that sub-Saharan Africans living outside Africa exhibits a high prevalence of complex diseases, the comparison of ancient African IM data to those of modern Africans might facilitate not only retrospective disease diagnosis, but also the identification of IM-related risk factors that contribute to the onset of certain diseases.

Analyses

The extraction and analyses of aDNA from African prehistoric samples is challenging, primarily because no suitable aDNA extraction facilities currently exist in Africa. Globally, there are 62 laboratories dedicated to aDNA research (Europe - 25, North America - 22, Australasia - 6, Asia - 5, Central America - 3 and South America - 1). Given the fact that sub-Saharan Africa should form the very focus of aDNA research it is curious that, excluding Antarctica, Africa is the only continent on which no aDNA dedicated facilities exist. Although the Centre for Microbial Ecology and Genomics (CMEG) has mastered the extraction of environmental DNA from soils derived from the Namib Desert and Antarctica, archaeological sample processing must take place at a suitable 'clean laboratory' facility. Accordingly, collaborative analytical arrangements have been made with the Centre for GeoGenetics (CGG) (Genetic Identification and Discovery Group (GID)) at the Natural History Museum in Copenhagen, Denmark. The five palaeo-faecal specimens will therefore be exported to Denmark and analysed in the 'clean' laboratory fascilities at the CGG.

All pre-PCR amplification steps will be performed applying established aDNA protocols. Extractions for shotgun metagenome sequencing will be carried out using a phenol-chloroform- and kit-based extraction protocol optimized for ancient sedimentary and faecal samples. This will entailed dissolving a total of ~16 g of palaeo-faecal matter in 40 ml digestion buffer. Specimen libraries and two negative control (*i.e.*, an extraction and library preparation control referred to as 'E-LPCs') libraries will be constructed using the NEBNext DNA Library Prep Master Mix for 454 (E6070) and sequenced on an Illumina HiSeq 2500 platform at the Danish National High-Throughput DNA Sequencing Centre. Taxonomic profiling will be preceded by several data pre-processing steps. First, raw sequence reads will be processed to remove all Illumina PhiX spikes, human reads and all exact duplicate reads present in the extraction (n = 1) and library preparation (n = 1) negative controls (E-LPCs) using BBDuk. Second, barcodes, adapters, reads shorter than 25 base-pairs (bp) and 'quality score' <25 will be removed from the dataset using AdapterRemoval V2. Taxonomic binning will then performed by BLASTn and BLASTx comparisons against the National Center for Biotechnology Information non-redundant nucleotide sequence databases (NCBI-nr and NCBI-nt). Taxa will be identified using MEGAN6 Community Edition (CE) v6.8.19 by using the weighted lowest common ancestor (wlca) option (80%), with parameter values set as follows: min. bit score: 50, expect value (e-value): 1.0e-10, top 10% hits, top percent: 10, min. support: 10 and min complexity: 0.45. Species identifications were based on significant hits (bit score>40) and on MEGAN6 CE v6.8.19 parameters established at identities: 100%, positives: 100% and no (0%) gaps. Comparisons of BRS IM sequence reads with those derived from other (comparative) IMs will be performed by the subsampling in of the reads to the lowest number of reads present in any library. Molecular damage following death is a standard feature of all

aDNA molecules. The accumulation of deaminated cytosine (uracil) within the overhanging ends of aDNA templates results in increasing cytosine (C) to thymine (T) misincorporation rates toward read starts, with matching guanine (G) to adenine (A) misincorporations increasing toward read ends in double-stranded library preparations. MapDamage is widely used to determine the incidence of cytosine (C) to thymine (T) and guanine (G) to adenine (A) substitution rates at the 5'-ends and 3'-ends of strands (29, 30).

Following aDNA extraction and sequencing in Copenhagen, accelerator mass spectrometry (AMS) dating will be performed on two sub-samples derived from the interior regions of each specimen at the iThemba LABS AMS facility in Johannesburg, South Africa. To investigate the dietary composition of the specimens, one sub-sample derived from the interior regions of the specimens will be subjected to isotopic analyses at the Mammal Research Institute, University of Pretoria, South Africa. In addition to genomic taxonomic profiling, we will also perform microscopic analysis to determine the incidence of intestinal parasitic helminths and protozoa. The extraction protocols applied in palaeo-parasitology used to extract parasitic markers (*i.e.*, eggs or oocysts) typically entails rehydration, homogenisation and micro-sieving. Using 0.5 g sub-samples from each specimen, this will be performed at the University of Bourgogne France-Comte, Besancon, France. For SEM analyses, we will immobilise 0.5 g palaeo-faecal material on double-sided carbon tape (SPI supplies), and coat this with gold with an Emitech K450X sputter coater (Quorum Technologies, UK). SEM images will be acquired at the University of Pretoria on a Zeiss Ultra Plus Field Emission Scanning Electron microscopes (Carl Zeiss, Oberkochen, Germany), at an accelerating voltage of 1kV.

References

1. J. Lederberg, A. T. McCray, 'Ome sweet' omics: A genealogical treasury of words. Scientist 15, 8 (2010).
2. C. Warinner, C. Speller, M. J. Collins, C. M. Lewis, Ancient human microbiomes. J Hum Evol. 79, 125-136 (2015).
3. E. Thursby, N. Juge, Introduction to the human gut microbiota. Biochem J. 474, 1823-1836 (2017).
4. L. W. van den Elsen, H. C. Poyntz, L. S. Weyrich, W. Young, E. E. Forbes-Blom, Embracing the gut microbiota: The new frontier for inflammatory and infectious diseases. Clin Transl Immunology 6, doi: 10.1038/cti.2016.91 (2017).
5. J. C. Clemente. E. C. Pehrsson, M. J. Blaser, K. Sandhu, Z. Gao, B. Wang, M. Magris, G. Hidalgo, M Contreras, Ó. Noya-Alarcón, O. Lander, J. McDonald, M. Cox, J. Walter, P. L. Oh, J. F. Ruiz, S. Rodriguez, N. Shen, S. J. Song, J. Metcalf, R. Knight, G. Dantas, M. G. Dominguez-Bello, The microbiome of uncontacted Amerindians. Sci Adv. 1, doi: 10.1126/sciadv.1500183 (2015).
6. E. R. Davenport, J. G. Sanders, S. J. Song, K. R. Amato, A. G. Clark, R. Knight, The human microbiome in evolution. BMC Biol. 15, doi: 10.1186/s12915-017-0454-7 (2017).
7. M. J. Blaser, The past and future biology of the human microbiome in an age of extinctions. Cell 172, 1173-1177 (2018).
8. S. L. Schnorr, M. Candela, S. Rampelli, M. Centanni, C. Consolandi, G. Basaglia, S. Turroni, E. Biagi, C. Peano, M. Severgnini, J. Fiori, R. Gotti, G. De Bellis, D. Luiselli, P. Brigidi, A. Mabulla, F. Marlowe, A. G. Henry, A. N. Crittenden, Gut microbiome of the Hadza hunter-gatherers. Nat Commun. 5, doi: 10.1038/ncomms4654 (2014).
9. J. Obregon-Tito, R. Y. Tito, J. Metcalf, K. Sankaranarayanan, J. C. Clemente, L. K. Ursell, Z. Zech Xu, W. Van Treuren, R. Knight, P. M. Gaffney, P. Spicer, P. Lawson, L. Marin-Reyes, O. Trujillo-Villarroel, M. Foster, E. Guija-Poma, L. Troncoso-Corzo, C. Warinner, A. T. Ozga, C. M. Lewis, Subsistence strategies in traditional societies distinguish gut microbiomes. Nat Commun. 6505, doi: 10.1038/ncomms7505 (2015).
10. Warinner, C. M. Lewis, Microbiome and health in past and present human populations. Am Anthrop. 117, 740-741 (2015).

11. R. Lee, R. H. Daly, Cambridge encyclopaedia of hunters and gatherers. Cambridge, Cambridge University Press ISBN 9780521609197 (1999).
12. Gomez, K. J. Petrzelkova, M. B. Burns, C. J. Yeoman, A. R. Amato, K. Vlckova, D. Modry, A. Todd, C. A. Jost Robinson, M. J. Remis, M. G. Torralba, E. Morton, J. D. Umaña, F. Carbonero, H. R. Gaskins, K. E. Nelson, B. A. Wilson, R. M. Stumpf, B. A. White, S. R. Leigh, Gut microbiome of coexisting BaAka Pygmies and Bantu reflects gradients of traditional subsistence patterns. Cell Rep. 14, 2142-2153 (2016).
13. Girard, N. Tromas, M. Amyot, B. J. Shapiro, Gut microbiome of the Canadian Arctic Inuit. mSphere 2, doi: 10.1128/mSphere.00297-16 (2017).
14. Y. Voigt, P. I. Costea, J. R. Kultima, S. S. Li, G. Zeller, S. Sunagawa, P. Bork, Temporal and technical variability of human gut metagenomes. Genome Biol. 16, doi: 10.1186/s13059-015-0639-8 (2015).
15. J. Walter, R. Ley, The human gut microbiome: Ecology and recent evolutionary changes. Annu Rev Microbiol. 65, 411-429 (2011).
16. M. J. Blaser, S. Falkow, What are the consequences of the disappearing human microbiota? Nat Rev Microbiol. 7, 887-894 (2009).
17. J. Adler, K. Dobney, L. S. Weyrich, J. Kaidonis, A. W. Walker, W. Haak, C. J. Bradshaw, G. Townsend, A. Soltysiak, K. W. Alt, J. Parkhill, A. Cooper, Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. Nat Genet. 45, 450-455 (2013).
18. S. Quercia, M. Candela, C. Giuliani, S. Turroni, D. Luiselli, S. Rampelli, P. Brigidi, C. Franceschi, M. G. Bacalini, P. Garagnani, C. Pirazzini, From lifetime to evolution: Timescales of human gut microbiota adaptation. Front Microbiol. 5, doi: 10.3389/fmicb.2014.00587 (2014).
19. H. Moeller, A. Caro-Quintero, D. Mjungu, A. V. Georgiev, E. V. Lonsdorf, M. N. Muller, A. E. Pusey, M. Peeters, B. H. Hahn, H. Ochman, Cospeciation of gut microbiota with hominids. Science 353, 380-382 (2016).
20. H. Reyes-Centeno, K. Harvati, G. Jäger, Tracking modern human population history from linguistic and cranial phenotype. Sci Rep. 6, doi: 10.1038/srep36645 (2016).
21. Houldcroft, J. B. Ramond, R. F. Rifkin, S. J. Underdown, Migrating microbes: What pathogens can tell us about population movements and human evolution. Ann Hum Biol. 44, 397-407 (2017).
22. B. Linz, F. Balloux, Y. Moodley, A. Manica, H. Liu, P. Roumagnac, D. Falush, C. Stamer, F. Prugnolle, S. W. van der Merwe, Y. Yamaoka, D. Y. Graham, E. Perez-Trallero, T. Wadstrom, S. Suerbaum, M. Achtman, An African origin for the intimate association between humans and Helicobacter pylori. Nature 445, 915-918 (2007).
23. M. T. Gilbert, D. L. Jenkins, A. Götherstrom, N. Naveran, J. J. Sanchez, M. Hofreiter, P. F. Thomsen, J. Binladen, T. F. Higham, R. M. Yohe, R. Parr, L. S. Cummings, E. Willerslev, DNA from pre-Clovis human coprolites in Oregon, North America. Science 320, 786-789 (2008).
24. R. J. Cano, J. Rivera-Perez, G. A. Toranzos, T. M. Santiago-Rodriguez, Y. M. Narganes-Storde, L. Chanlatte-Baik, E. García-Roldán, L. Bunkley-Williams, S. E. Massey, Paleomicrobiology: Revealing fecal microbiomes of ancient indigenous cultures. PLoS One 9, doi: 10.1371/journal.pone.0106833 (2014).
25. T. M. Santiago-Rodriguez, G. Fornaciari, S. Luciani, S. E. Dowd, G. A. Toranzos, I. Marota, R. J. Cano, Gut microbiome of an 11th century A.D. Pre-Columbian Andean mummy. PLoS One 10, doi: 10.1371/journal.pone.0138135 (2015).
26. S. Rampelli, S. L. Schnorr, C. Consolandi, S. Turroni, M. Severgnini, C. Peano, P. Brigidi, A. N. Crittenden, A. G. Henry, M. Candela, Metagenome sequencing of the Hadza hunter-gatherer gut microbiota. Curr Biol. 25, 1682-1693 (2015).
27. G. Porraz, A. Val, L. Dayet, P. de la Pena, K. Douze, C. E. Miller, M. Murungi, C. Tribolo, V. Schmid, C. Sievers, Bushman Rock Shelter (Limpopo, South Africa): A perspective from the edge of the Highveld. S Afr Archaeol Bull. 70, 166-179 (2015).

28. G. Porraz, A. Val, C. Tribolo, N. Mercier, P. de la Peña, M. M. Haaland, M. Igreja, C. E. Miller, V. Schmid, The MIS5 Pietersburg at '28' Bushman Rock Shelter, Limpopo Province, South Africa. PLoS One 13, doi: https://doi.org/10.1371/journal.pone.0202853 (2018).
29. Ginolhac, M. Rasmussen, M. T. P. Gilbert, E. Willerslev, L. Orlando, mapDamage: Testing for damage patterns in ancient DNA sequences. Bioinformatics 27, 2153-2155 (2011).
30. H. Jónsson, A. Ginolhac, M. Schubert, P. L. Johnson, L. Orlando, mapDamage 2.0: Fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics 29, 1682-1684 (2013).