AUTHORS:
Palesa P. Madupe[1,2]
Fazeelah Munir[3]
Marc Dickinson[3]
Alberto J. Taurozzi[1]
Meaghan Mackie[1,4]
Mirriam Tawane[5]
Catherine Mollereau[6]
Nomawethu Hlazo[2,7]
Kirsty Penkman[3]
Lauren Schroeder[2,8]
Clément Zanolli[9]
Jesper V. Olsen[4]
Rebecca R. Ackermann[2,7]
Enrico Cappellini[1]

AFFILIATIONS:
[1]Globe Institute, University of Copenhagen, Copenhagen, Denmark
[2]Human Evolution Research Institute (HERI), University of Cape Town, Cape Town, South Africa
[3]Department of Chemistry, University of York, York, United Kingdom
[4]Proteomics Program, Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Copenhagen, Denmark
[5]National Heritage Council, Pretoria, South Africa
[6]Research Centre on Animal Cognition (CRCA), Centre of Integrative Biology (CBI), National Centre for Scientific Research (CNRS), University of Toulouse, Toulouse, France
[7]Department of Archaeology, University of Cape Town, Cape Town, South Africa
[8]Department of Anthropology, University of Toronto Mississauga, Mississauga, Canada
[9]From Prehistory to Today: Culture, Environment and Anthropology (PACEA), National Centre for Scientific Research (CNRS), University of Bordeaux, Bordeaux, France

CORRESPONDENCE TO:
Palesa Madupe

EMAIL:
palesa.madupe@palaeoproteomics.org

Results from an *Australopithecus africanus* dental enamel fragment confirm the potential of palaeoproteomics for South African Plio-Pleistocene fossil sites

The southern African Late Pliocene to Early Pleistocene hominin record is abundant and exhibits a high taxonomic diversity with three genera represented: *Australopithecus, Paranthropus* and *Homo*. Hominin fossil diversity and variation are often contextualised within other fossil assemblages or modern/extant counterparts. However, the incompleteness of the fossil record, sample selection bias and taphonomic condition of the specimens themselves constrain interpretations of diversity and variation within and between species. Thus, species identification and the nature of the observed variation are frequently debated. Palaeoproteomics can help improve our understanding of taxonomic variation, as demonstrated by the recently generated proteome of *Paranthropus* specimens from Swartkrans. Here, we demonstrate protein preservation for an *A. africanus* specimen from Sterkfontein Member 4, Sts 63, using minimally invasive analysis, and identify it as belonging to a male individual. We then discuss some of the current limitations of palaeoproteomics and how we can potentially overcome them. Although it is still in its infancy for Plio-Pleistocene hominin fossils, palaeoproteomics has the potential to help unravel the causes of observed morphological variation. Lastly, we strongly believe that the involvement of African researchers at all levels of this research, including leadership, is of great importance.

Significance:

We have successfully determined the biological sex of an *Australopithecus africanus* specimen (Sts 63) from Sterkfontein Member 4 with the age range of 3.5 to 2.01 Ma, with a high degree of confidence, and we have assessed the extent of protein preservation. These discoveries hold significant implications for our understanding of sexual dimorphism and intraspecies variation as observed in African Plio-Pleistocene hominins.

[Abstract in Setswana]

Introduction

Since the discovery of the Taung Child a century ago[1], South Africa has been a world leader in palaeoanthropology research. Much of this work has focused on understanding and interpreting the similarities and differences in the southern African Pleistocene fossil record and their relationships to hominins across the wider African continent. A key area of research is the study of early hominin taxonomy, phylogenetic relationships and variation, both among ancient taxa and between them and our species, *Homo sapiens*. Today, South Africa is well positioned to unpack these relationships, as it has a rich and taxonomically diverse hominin fossil record, particularly at sites located in the UNESCO World Heritage Site of the Cradle of Humankind. Important or iconic finds and specimens include individuals assigned to *Australopithecus africanus, A. sediba, A. prometheus, Paranthropus robustus, Homo erectus*, and *H. naledi*, with some of these species living contemporaneously[2].

A consequence of this rich and diverse fossil record is that there is considerable interindividual variation that can be attributed to numerous factors, including the potential sampling of morphological variation between species (i.e. taxonomic diversity), as well as variation within species (i.e. sexual dimorphism, inter-locality variation and microevolution/temporal depth variation). Teasing apart the presence of these different contributors to variation can be challenging.[3,4] As an example, *P. robustus*, a taxonomic group only found in South Africa, has been subject to varying hypotheses explaining the underlying causes of variation. Lockwood et al.[5] hypothesised that the variation in *P. robustus* is due to sexual dimorphism. However, the discovery of DNH 155, a purported male individual, and dental remains from the site of Drimolen attributed to *P. robustus* showing a less robust morphology than some of the material from the site of Swartkrans, led to the hypothesis that the observed variation is due to temporal depth variation[6,7] rather than a high degree of sexual dimorphism[5,8]. In particular, as DNH 115 is presumed male and more gracile relative to the presumed male individuals of Swartkrans and Komdraai B, the authors then hypothesised that the Drimolen collection is older than Swartkrans and Komdraai B.[6] Further analysis of dental remains of *P. robustus* from both Drimolen and Swartkrans suggested the variation may be due to different specimen compositions across localities.[9] This inter-locality variation hypothesis was corroborated by work on the differences in temporal bone shape and size in specimens from Drimolen, Swartkrans and Kromdraai B.[10] Nevertheless, it is noted that inter-locality variation and high sexual dimorphism hypothesis does not contradict the temporal depth hypothesis.[9,10]

Similarly, the *Australopithecus* assemblage from Sterkfontein, South Africa, is highly variable morphologically, and it is suggested that there are multiple species of *Australopithecus* (*A. prometheus* and *A. africanus*) found in Member 2 and Member 4 that are thought to have overlapped.[11,12] It is also hypothesised that there may be another

species in Member 4 besides *A. africanus*.[13] In addition, the variation observed has been proposed to be due to temporal depth[14,15], as Member 4 spans about 600 thousand years (ka)[16,17]. In both cases, for *P. robustus* as well as the Sterkfontein *Australopithecus* assemblage, controlling for sexual dimorphism by confidently attributing the sex of individuals would contribute significantly to our understanding of the underlying causes of variation.

Enamel palaeoproteomics studies of fossil hominins

Palaeoproteomics is the study of proteins from fossilised material, and it exists at the intersection of multiple disciplines: chemistry, molecular biology, archaeology, palaeontology, palaeoanthropology, palaeoecology, computational biology and history.[18,19] Mineral-bound proteins have recently been shown to survive deeper in time and in warmer regions[20] relative to DNA[21]. In 2009, Nielsen-Marsh et al.[22] demonstrated the feasibility of extracting enamel peptides from late Pleistocene Neanderthal specimens using a trypsin-aided digestion process coupled with matrix-assisted laser desorption/ionisation (MALDI) sequencing. Their analysis successfully identified sex chromosome linked amelogenin-specific[23] peptides, highlighting the potential of this technique in ancient protein studies.

A few years ago, a study employed a digestion-free peptide extraction protocol[24] and liquid chromatography coupled to tandem mass spectrometry, which generated the dental enamel proteome of the extinct *Gigantopithecus blacki* dated to 1.9 Ma from Chifeng Cave, China[25]. Using the same technique, Welker et al.[26] recovered proteins from *Homo antecessor* and *H. erectus*, providing the oldest genetic information for the genus *Homo*. For the *H. antecessor* specimen, they recovered amelogenin-Y (AMELY), demonstrating that it belonged to a male individual, while there was no detection of an AMELY signal for the *H. erectus* specimen. These studies demonstrated the feasibility of using the enamel proteome to understand fossil variation from temperate and subtropical regions of the world. In these contexts, the enamel proteome can provide both tentative phylogenetic signals and confident biological sex identification of ancient male individuals.

One of the major questions that arose from the above studies is whether it is feasible to recover ancient proteomes from an African context, given the differences in ancient climate, geology and taphonomy. Subsequently, the southern African *Paranthropus* dental enamel proteome demonstrated the feasibility of palaeoproteomics[27] and the potential of investigating within-species variation within the African context. Here, we aim to demonstrate the potential of using a minimally invasive extraction protocol, expand the sample set of southern African hominins being analysed via palaeoproteomics and further explore fossil variation through a genetic lens. In this work, and in recognition of the centenary of the announcement of the species *A. africanus*, we report a palaeoproteomic profile of specimen Sts 63 (Figure 1), morphologically identified as *A. africanus*, from Sterkfontein Member 4. We then provide additional examples from a recent palaeoproteomic investigation of *Paranthropus*[27] and further discuss the current limitations of palaeoproteomics. Finally, we discuss the challenges of transformation, focusing on how it can be achieved through meaningful and impactful collaborative efforts that build capacity in Africa.

Methods

Permission for temporary export and sampling (permit IDs 3026 and 3079) was granted by the South African Heritage Resource Agency for palaeoproteomic analysis of Sts 63 (Figure 1), an *A. africanus* molar fragment, with no significant morphology preserved, from Sterkfontein Member 4.

Biomolecular preservation

Chiral amino acid analysis was undertaken on enamel (± 5 mg) from Sts 63 following the protocols of Dickinson et al.[28] After bleaching, the specimen was divided into two fractions: one for determining free amino acids (FAA) and one for quanitfication of the total hydrolysable amino acids (THAA). The concentration of the intra-crystalline

STS 63

1 cm

Figure 1: The analysed Sts 63 molar fragment (no orientation could be identified).

amino acids and their extent of racemisation (D/L value) were then quantified using RP-HPLC (Agilent 1100 series HPLC fitted with HyperSil C18 base deactivated silica column [5 μm, 250 x 3 mm] and fluorescence detector) following a modified method of Kaufman and Manley[29]. To provide estimation of intra-crystalline protein degradation, D and L ratios were measured for the following amino acids: aspartic acid and asparagine (Asx), glutamic acid and glutamine (Glx), serine (Ser), alanine (Ala), valine (Val), phenylalanine (Phe), isoleucine (Ile), leucine (Leu), threonine (Thr), arginine (Arg), tyrosine (Tyr) and glycine (Gly).

Etching extraction

Sts 63 peptides were extracted using a minimally destructive extraction method, specifically acid etching, as first demonstrated by Stewart et al.[30] Briefly, the sample surface was first cleaned using molecular biology grade water, the varnish coating was gently scraped off, and then the surface was wiped with low dust laboratory tissue (Kimtech) to remove debris. To further clean the surface, a volume of 130 mL of 10% v/v trifluoroacetic acid (TFA) was placed into the cap of a 0.5 mL Eppendorf tube and the sample was manually held so that the surface of the sample was in contact with the acid solution for an initial 15 s. The tube and acidic solution were then discarded. The acid cleaning step was done twice. The sample was reintroduced to the new 10% TFA in the new tube, and contact was maintained for a total of 10–15 min, with visual inspection every 3–5 min. The acidic solution (sample extract) was removed from the 0.5 mL tube cap and placed into a fresh Protein LoBind Eppendorf tube and the cap was washed with 100 mL 10% TFA and combined with the sample extract. C18 StageTip[31] peptide concentration/clean-up was performed as described by Cappellini et al.[24] and Taurozzi et al.[32] An extraction blank was prepared simultaneously with the sample.

Liquid chromatography with tandem mass spectrometry analysis

The peptides were eluted with 30 μL of 40% acetonitrile (ACN) and 0.1% formic acid (FA) into a 96-well mass spectrometry (MS) plate from the C18 StageTip.[31] They were then resuspended in 4 μL of 5% of ACN 0.1% TFA. The solution containing the peptides was analysed through liquid chromatography with tandem mass spectrometry (LC MS/MS) following protocols published for palaeoproteomics samples.[24,33] Peptide separation took place on a 15-cm column (75 μm inner diameter), in-house laser-pulled and packed with 1.9 μm C18 beads (Dr Maisch, Germany), on an EASY-nLC 1200 (Proxeon, Odense, Denmark) connected to an Exploris 480 mass spectrometer (Thermo Scientific, Bremen, Germany), on a 77-min gradient with wash-blanks in between the injections of samples to hinder cross-contamination.

Data analysis

The files generated by the mass spectrometer in the '.raw' file format were then processed using MaxQuant version 2.1.0.3[34], to confidently match the spectra against peptides from a custom-made reference database of amelogenin proteins of extant *Homo sapiens*, publicly available ancient hominins, and members of *Pan, Gorilla* and *Pongo* downloaded from Uniprot and NCBI and translated in-house[35]. The peptide identification was performed, setting the digestion parameter to unspecific, and the minimum length for unspecific peptides was set to seven amino acids. In the main search, the peptide mass tolerance was left at 4.5 parts per million (ppm), also leaving the setting of the fragment mass tolerance at 20 ppm.

The Andromeda threshold score for both unmodified and modified proteins was set to 40, to filter out peptide spectral matches (PSM) with a low-quality score. No fixed post-translational modifications were set. Glutamine and asparagine deamidation, oxidation of methionine, oxidation of proline, oxidation of tyrosine, phosphorylation of serine/threonine/tyrosine, ornithine conversion from asparagine, and N-terminal pyroglutamic acid from glutamic and aspartic acids were all included as possible variable modifications.

Proteins included in the database of common contaminants provided by MaxQuant, for example, proteinaceous laboratory reagents and human skin keratins, as well as reverse sequences, were manually removed and not considered any further. Similarly, proteins detected in the laboratory blank were not considered further.

Results

A total of 142 amino acids was recovered for both AMELX and AMELY, with 118 peptides, 4 unique to AMELX and 3 unique to AMELY (Table 1). Thus, we were able to identify Sts 63 as belonging to a male individual, with the confident detection of three specific AMELY peptides (Figure 2). Subsequently, we observed similar diagenetic markers as seen in Cappellini et al.[24], Welker et al.[25,26] and Madupe et al.[27], i.e. the peptide length distribution and rate of deamidation, albeit at higher amounts (Figure 3A and 3B). Moreover, we observed higher levels of intra-crystalline protein decomposition in Sts 63 relative to *Paranthropus* specimens from Swartkrans, including higher levels of racemisation (conversion of the L-amino acids to their D-form). The high intra-crystalline protein decomposition patterns in the enamel are consistent with a closed system behaviour, thus indicating that the recovered proteins are endogenous to the enamel matrix (Figure 3C). The higher levels of intra-crystalline protein decomposition are consistent with radiometric dating that indicates Sts 63 (Sterkfontein Member 4) is older than the *Paranthropus* specimens studied in Madupe et al.[27] (Swartkrans Member 1).

Discussion

A preliminary protein profile of A. africanus

Studies carried out on ancient hominin specimens allow us to start to unpack whether hominin morphological variation is due to sexual dimorphism, taxonomic differences or potentially other forms of variation. However, these studies are still in their infancy. Madupe et al.[27] reported the recovery of the enamel proteome from four *Paranthropus* teeth dated to ca 2 Ma[36] from Swartkrans, South Africa. The most abundant enamel proteins, namely enamelin, amelogenin and ameoloblastin, were recovered as part of the suite of proteins sequenced via tandem mass spectrometry. The identification of AMELY-specific peptides and semi-quantitative mass spectrometry data analysis enabled confident identification of the biological sex of all the specimens. Intraspecies amino acid sequence variation was also observed among the four *Paranthropus* specimens, corroborating independent observations made on morphology.[37] The recovered molecular data also confirmed the taxonomic placement of *Paranthropus* within the hominin clade, which formed the outgroup of the clade, including *H. sapiens*, Neanderthals and Denisovans.

In contrast, the analysis carried out here on Sts 63 is via a minimally invasive extraction protocol, which generated a minimal proteome (Table 1). Excitingly, this allowed us to confirm the presence of ancient proteins

Table 1: Summary statistics of the number of peptides, protein sequence coverage and the total amino acids recovered in each protein. The amelogenin lengths refer to the human versions: ensemble transcript ENSG00000125363 for AMELX and ENSG00000099721 for AMEY.

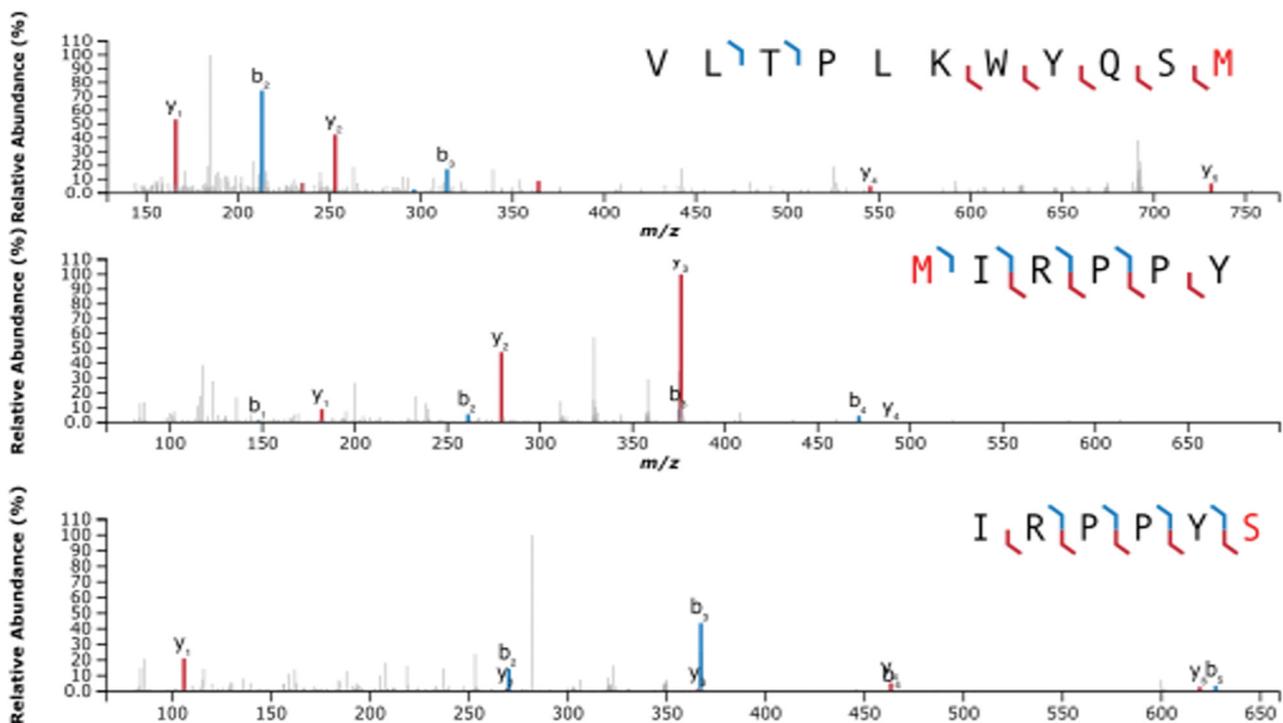| Protein ID	Total number of peptides	Unique peptides	Percentage coverage (%)	Sequence length	Total amnio acids recovered
Amelogenin X	67	4	41.9	205	86
Amelogenin Y	51	3	27.1	206	56

Figure 2: The top frame is the human AMELX (position 47 to 66) and AMELY (position 57 to 67) aligned, and highlighted in red are the two different amino acids in the alignment; with the insertion of methionine (**M**) in position 59 and a serine (**S**) at position 66 instead of a proline (P) in the AMELX corresponding position. Below that are three peptide-spectrum matches from *Australopithecus africanus* Sts 63 for human AMELY with **M** and **S** highlighted in red. Note peptide spectrum graphs plot mass-to-charge (m/z) values of ions on the *x*-axis, as measured during the mass spectrometric analysis of peptide fragments, and the relative peak abundance (%) on the *y*-axis. The red peaks represent y-ions, which are generated from fragmentation at the C-terminal side of peptide bonds, and they correspond to red bars between amino acids in the peptide sequence. The blue peaks represent b-ions, which result from fragmentation at the N-terminal side of peptide bonds, and they correspond to blue bars between amino acids in the peptide sequence. These peaks are matched to the theoretical spectrum of the peptide, aiding in the identification of the peptide sequence. This figure was generated using the publicly free site www.proteomicsdb.org/use/ by inputting MS2 mass to charge ratios.

in the sample, and to identify the sex of Sts 63 as male, confirming the potential for these studies using material from Sterkfontein. However, we are currently limited in the scope of a comparative analysis with *P. robustus*, which will require a larger *A. africanus* enamel proteome generated from extracting higher amounts of enamel from additional individuals. Moreover, in South Africa, species of *Australopithecus*, *Paranthropus* and early *Homo* were contemporaneous, and, ideally, phylogenetic comparisons based on genetic data should include information from all these taxa, allowing us to answer important questions about sexual dimorphism and taxonomic relationships. This is an exciting avenue for future research.

Consequently, this raises the question of whether hominins outside the South African Cradle cave systems will also have sufficiently good preservation for ancient protein recovery. The fossil *A. africanus* tooth studied here, as well as the *Paranthropus* specimens studied by Madupe et al.[27], were all recovered from cave sediments composed of re-mobilised soil outside the cave[38]. It is possible that favourable protein preservation is due to factors specific to these systems, including rapid fossil accumulation and relative aridity.[27,39] However, there are currently no published hominin protein preservation data on fossils recovered from other sites in South Africa and no data from open-air fossil sites, particularly in eastern Africa, where the very different depositional environments are known to have detrimental diagenetic consequences for enamel preservation.[40] Expanding palaeoproteomic studies of enamel outside the Cradle cave system context will therefore require considerable

exploration of preservation – an important future avenue for research into African Plio-Pleistocene hominins.

The current limitations – and future opportunities of palaeoproteomics

Palaeoproteomics provides deep-time genetic data that were previously inaccessible. The ancient genetic data allow us to draw tentative hominin phylogenies and study interspecies and intraspecies variation, sexual dimorphism, and temporal depth variation. These are all incredibly exciting breakthroughs that are poised to revolutionise our understanding of human evolution. However, there are still some limitations and pitfalls to overcome.

As amelogenin-X is expressed in both female and male individuals, it is challenging to identify female individuals unambiguously; absence of evidence (e.g. no detection of AMELY) does not always mean evidence of absence. Currently, there is no experimental way of positively identifying male individuals with a deletion of the amelogenin-Y gene – a condition that has been documented in modern humans[41,42] and in a Neanderthal individual[43]. Additionally, in this current iteration of the palaeoproteomic workflow, male individuals whose amelogenin-Y protein has been degraded below instrumental detection limits due to diagenesis will also be misidentified as female. Several recent attempts to identify females through semi-quantitative analyses have proven fruitful.[27,44-46] However, these methods rely on having at least one positively identified
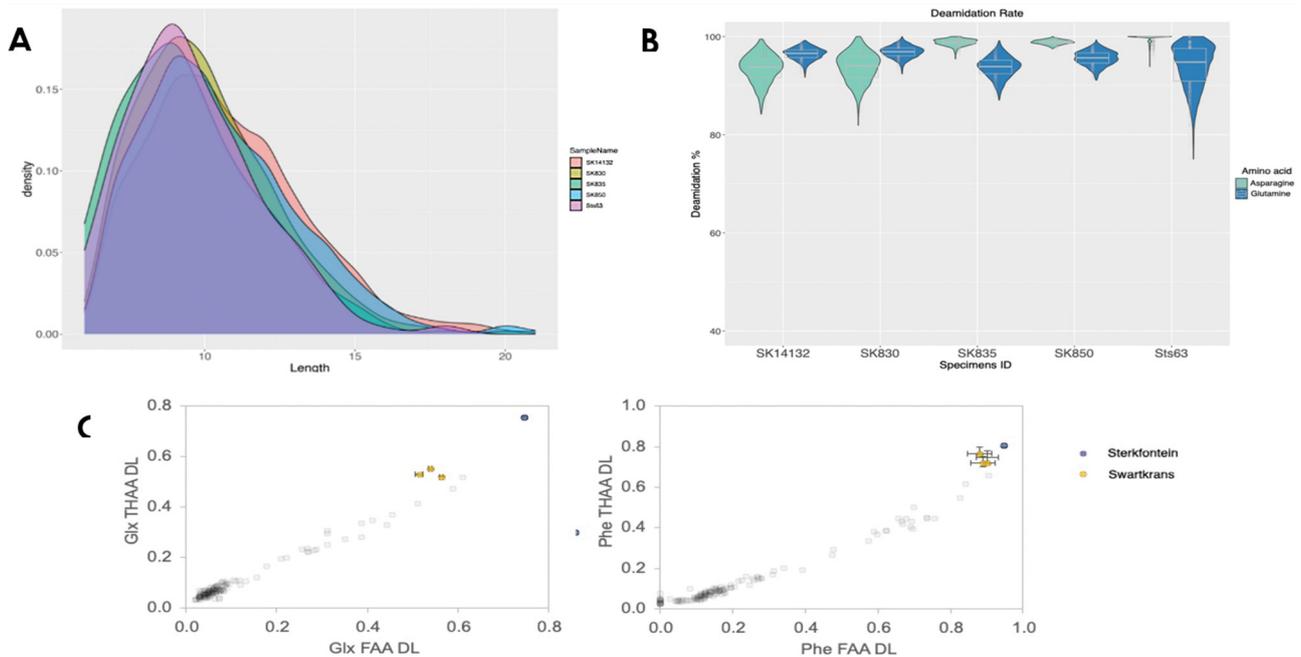
Figure 3: Diagenetic modifications of the Sterkfontein *Australopithecus africanus* Sts 63 relative to the Swartkrans *Paranthropus* specimens from Madupe et al.[27] (A) Peptide length distribution is skewed toward shorter fragments due to spontaneous terminal hydrolysis, with the *x*-axis indicating the peptide length and the *y*-axis indicating the density distribution of peptide lengths. (B) Asparagine and glutamine deamidation levels, with the *x*-axis indicating the specimens and the *y*-axis showing the percentage of asparagine and glutamine deamidations in each specimen. (C) Free amino acid (FAA) vs total hydrolysable amino acid (THAA) racemisation for glutamine/glutamic acid (Glx) and phenylalanine (Phe), with the Sts 63 specimen in blue being higher than the Swartkrans specimens in yellow. A reference data set of previously analysed enamel is shown in grey to indicate the expected correlation between FAA and THAA racemisation for closed-system enamel.

male individual in the sample to establish a probabilistic framework. Therefore, currently, confident semi-quantitative female detection is sample-set dependent.

Another important consideration is the small amount of genetic information currently retrieved by enamel palaeoproteomic analysis. Proteins only represent the expression of the exonic part of the genome, and the ancient enamel proteome is not particularly rich, counting only about 12 proteins.[24-26] Furthermore, enamel proteins are hydrolysed by proteases in the final phase of amelogenesis during tooth maturation. Specifically, matrix metalloproteinase-20 (MMP20) and kallikrein-related peptidase 4 (KLK4) break down enamelins, amelogenins, ameloblastins and amelotins[47,48], leaving in mature dental enamel only a limited subset of the protein sequences initially synthesised. Furthermore, phylogenetic incongruence, in which evolutionary trees constructed from individual genes differ from each other and from the expected species trees, affects the accuracy of the phylogenies we generate from enamel proteomes.[49] For this reason, phylogenies built with this approach are based on amino acid sequences only a few hundred amino acids long and cannot be considered very informative.[50]

In Madupe et al.[27], the authors emphasise that the observed phylogenetic placement of *P. robustus* is tentative due to the size of the recovered proteome, and here we did not include a reconstructed phylogeny as the minimally invasive peptide extraction protocol resulted in a very small proteome, making any phylogeny even less reliable. For comparison, in the initial phases of a DNA analysis, researchers relied on short DNA sequences, such as portions of the mitochondrial DNA (mtDNA). Although these studies provided preliminary insights, they had limitations. They did not detect gene flow between Neanderthals and modern humans because they focused on uniparentally inherited markers.[51] This approach masked the complexities of interbreeding and gene exchange among different hominin groups. Only later, with the introduction of high-throughput next generation DNA sequencing[52,53], did more comprehensive genomic analyses reveal these phenomena[54-56]. For proteins, we are already seeing glimpses of technology improving modern proteomic modes of data acquisition with single-molecule protein sequencing[57], the merits of which are discussed by Paterson et al.[50]

Final thoughts

The relatively new field of palaeoproteomics has the potential to revolutionise our understanding of Plio-Pleistocene hominin diversity in southern Africa, and possibly in Africa more broadly. Recent and ongoing studies have demonstrated its application in interpreting morphological variation. Madupe et al.[27] observed the presence of substantial molecular variation within *Paranthropus*, in addition to identifying biological sex. Additionally, here we have presented the protein preservation of a specimen identified morphologically as *A. africanus*. This is the first step to attempt the recovery of the enamel proteome for this specimen. In addition to the preliminary palaeoproteomic characterisation, we also identified the sex of the specimen and validated the endogeneity of the recovered enamel proteins. The studies of *Paranthropus* proteomes, combined with the initial palaeoproteomic analysis of the *A. africanus* specimen presented here, demonstrate the feasibility and utility of palaeoproteomic studies in South Africa. Even though palaeoproteomics is still in its infancy and caution should be used in interpreting the results, it is still poised to be able to answer some of palaeoanthropology's most fundamental questions about sexual dimorphism, variation and phylogeny.

Future studies should focus on improving protein recovery and on increasing the breadth and depth of amino acid sequence coverage, as well as on the number of studied samples and taxa. Moreover, less destructive protein extraction methods need to be explored. Currently, the most common approach is to extract proteins by destructively sampling approximately 100 mg of dental enamel. In the future, alternative methods, such as the minimally destructive method used in this study, would make the application of palaeoproteomics more broadly applicable.

Palaeoproteomic research is a new and burgeoning field that has the potential to increase our understanding of the deep past. We see huge potential for the application of palaeoproteomics in understanding Plio-Pleistocene hominin diversity. As we have explored here, a lot of work still needs to be done, and this provides a unique and exciting opportunity for this field to be developed collaboratively, together with African researchers at the forefront. In this current special issue,

Lee-Thorp and Sponheimer[58] provide historical examples of how meaningful collaboration can be achieved and how it led to the field of fossil biogeochemistry expanding via the investment in scientific infrastructure and capacity building in South Africa. To actuate this, we are currently working on establishing an ancient biomolecules laboratory in South Africa, so that at least the first part of the palaeoproteomics workflow can be carried out in Africa, in collaboration with international labs for the sequencing part of the workflow. This would mean that fossils could be sampled locally with no need for them to leave the continent. This would represent a big step in ensuring both capacity building and the safety and safekeeping of African heritage.

In conclusion, palaeoproteomics research is at the cusp of remarkable discoveries, making this an ideal time to develop new ways in which research could be done. We also want to emphasise that palaeoproteomics should not be another field of study in which marginalised communities are left out, or in which parachute/helicopter[59] science takes place. We can take the initiative in this nascent research field to halt colonial science[60-62] and to realise that research is greatly improved by meaningful co-creation and collaboration[63]. Knowledge comes in different forms and diversity improves the quality of research.[64] We are excited by what lies ahead.

Funding

Data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository with the data set identifier PXD054431.

Declarations

We have no competing interests to declare. We have no AI or LLM use to declare.

Authors' contributions

P.P.M.: Conceptualisation, data collection, sample analysis, data analysis, writing – the initial draft, writing – revisions, project leadership, project management. F.M.: Data collection, sample analysis, data analysis. M.D.: Data collection, sample analysis, data analysis, writing – revisions. A.J.T.: Data collection, sample analysis, writing – revisions. M.M.: Data collection, sample analysis, writing – revisions. M.T.: Provided the specimens. C.M.: Writing – revisions. N.H.: Writing – revisions. K.P.: Writing – revisions. L.S.: Writing – revisions, student supervision. C.Z.: Writing – revisions. J.V.O.: Funding acquisition. R.R.A.: Conceptualisation, writing – revisions, student supervision, project leadership, project management, funding acquisition. E.C.: Conceptualisation, writing – revisions, student supervision, project leadership, project management, funding acquisition. All authors read and approved the final manuscript.

References

1. Dart RA. The Taungs skull. Nature. 1925;116:462. https://doi.org/10.1038/116462a0

2. Herries AIR, Martin JM, Leece AB, Adams JW, Boschian G, Joannes-Boyau R, et al. Contemporaneity of *Australopithecus*, *Paranthropus*, and early *Homo erectus* in South Africa. Science. 2020;368(6486), eaaw7293. https://doi.org/10.1126/science.aaw7293

3. Frayer DW, Wolpoff MH. Sexual dimorphism. Annu Rev Anthropol. 1985;14:429–473. https://doi.org/10.1146/annurev.an.14.100185.002241

4. Thackeray JF, Dykes S. Morphometric analyses of hominoid crania, probabilities of conspecificity and an approximation of a biological species constant. Homo. 2016;67:1–10. https://doi.org/10.1016/j.jchb.2015.09.003

5. Lockwood CA, Menter CG, Moggi-Cecchi J, Keyser AW. Extended male growth in a fossil hominin species. Science. 2007;318:1443–1446. https://doi.org/10.1126/science.1149211

6. Martin JM, Leece B, Neubauer S, Baker SE, Mongle CS, Boschian G, et al. Drimolen cranium DNH 155 documents microevolution in an early hominin species. Nat Ecol Evol. 2021;5:38–45. https://doi.org/10.1038/s41559-020-01319-6

7. Leece AB, Martin JM, Baker S, Wilson C, Strait DS, Schwartz GT, et al. New hominin dental remains from the ~2.04–1.95 Ma Drimolen Main Quarry, South Africa. Ann Hum Biol. 2023;50:407–427. https://doi.org/10.1080/03014460.2023.2261849

8. Keyser AW. The Drimolen skull: The most complete australopithecine cranium and mandible to date. S Afr J Sci. 2000;96:189–193.

9. Leece AB, Martin JM, Herries AIR, Riga A, Menter CG, Moggi-Cecchi J. New hominin dental remains from the Drimolen Main Quarry, South Africa (1999–2008). Am J Biol Anthropol. 2022;179:240–260. https://doi.org/10.1002/ajpa.24570

10. Braga J, Chinamatira G, Zipfel B, Zimmer V. New fossils from Kromdraai and Drimolen, South Africa, and their distinctiveness among *Paranthropus robustus*. Sci Rep. 2022;12, Art. #13956. https://doi.org/10.1038/s41598-022-18223-7

11. Clarke RJ. Latest information on Sterkfontein's *Australopithecus* skeleton and a new look at *Australopithecus*. S Afr J Sci. 2008;104:443–449. https://doi.org/10.1590/S0038-23532008000600015

12. Clarke RJ, Kuman K. The skull of StW 573, a 3.67 Ma *Australopithecus prometheus* skeleton from Sterkfontein Caves, South Africa. J Hum Evol. 2019;134, Art. #102634. https://doi.org/10.1016/j.jhevol.2019.06.005

13. Lockwood CA, Tobias PV. Morphology and affinities of new hominin cranial remains from Member 4 of the Sterkfontein Formation, Gauteng Province, South Africa. J Hum Evol. 2002;42:389–450. https://doi.org/10.1006/jhev.2001.0532

14. Moggi-Cecchi J, Grine FE, Tobias PV. Early hominid dental remains from Members 4 and 5 of the Sterkfontein Formation (1966–1996 excavations): Catalogue, individual associations, morphological descriptions and initial metrical analysis. J Hum Evol. 2006;50:239–328. https://doi.org/10.1016/j.jhevol.2005.08.012

15. Beaudet A. The *Australopithecus* assemblage from Sterkfontein Member 4 (South Africa) and the concept of variation in palaeontology. Evol Anthropol. 2023;32:154–168. https://doi.org/10.1002/evan.21972

16. Pickering R, Kramers JD. Re-appraisal of the stratigraphy and determination of new U-Pb dates for the Sterkfontein hominin site, South Africa. J Hum Evol. 2010;59:70–86. https://doi.org/10.1016/j.jhevol.2010.03.014

17. Granger DE, Stratford D, Bruxelles L, Gibbon RJ, Clarke RJ, Kuman K. Cosmogenic nuclide dating of *Australopithecus* at Sterkfontein, South Africa. Proc Natl Acad Sci USA. 2022;119, e2123516119. https://doi.org/10.1073/pnas.2123516119

18. Warinner C, Korzow Richter K, Collins MJ. Paleoproteomics. Chem Rev. 2022;122:13401–13446. https://doi.org/10.1021/acs.chemrev.1c00703

19. Cappellini E, Prohaska A, Racimo F, Welker F, Pedersen MW, Allentoft ME, et al. Ancient biomolecules and evolutionary inference. Annu Rev Biochem. 2018;87:1029–1060. https://doi.org/10.1146/annurev-biochem-062917-012002

20. Demarchi B, Hall S, Roncal-Herrero T, Freeman CL, Woolley J, Crisp MK, et al. Protein sequences bound to mineral surfaces persist into deep time. eLife. 2016;5, e17092.

21. Lipson M, Sawchuk EA, Thompson JC, Oppenheimer J. Ancient DNA and deep population structure in sub-Saharan African foragers. Nature. 2022;603:290–296. https://doi.org/10.1038/s41586-022-04430-9

22. Nielsen-Marsh CM, Stegemann C, Hoffmann R, Smith T, Feeney R, Toussaint M, et al. Extraction and sequencing of human and Neanderthal mature enamel proteins using MALDI-TOF/TOF MS. J Archaeol Sci. 2009;36:1758–1763. https://doi.org/10.1016/j.jas.2009.04.004

23. Fincham AG, Bessem CC, Lau EC, Pavlova Z, Shuler C, Slavkin HC, et al. Human developing enamel proteins exhibit a sex-linked dimorphism. Calcif Tissue Int. 1991;48:288–290. https://doi.org/10.1007/BF02556382

24. Cappellini E, Welker F, Pandolfi L, Ramos-Madrigal J, Samodova D, Rüther PL, et al. Early Pleistocene enamel proteome from Dmanisi resolves *Stephanorhinus* phylogeny. Nature. 2019;574:103–107. https://doi.org/10.1038/s41586-019-1555-y

25. Welker F, Ramos-Madrigal J, Kuhlwilm M, Liao W, Gutenbrunner P, de Manuel M, et al. Enamel proteome shows that *Gigantopithecus* was an early diverging pongine. Nature. 2019;576:262–265. https://doi.org/10.1038/s41586-019-1728-8

26. Welker F, Ramos-Madrigal J, Gutenbrunner P, Mackie M, Tiwary S, Jersie-Christensen RR, et al. The dental proteome of *Homo* antecessor. Nature. 2020;580:235–238. https://doi.org/10.1038/s41586-020-2153-8

27. Madupe PP, Koenig C, Patramanis I, Rüther PL, Hlazo N, Mackie M, et al. Enamel proteins reveal biological sex and genetic variability within southern African *Paranthropus* [preprint]. bioRxiv 2023.07.03.547326. https://doi.org/10.1101/2023.07.03.547326

28. Dickinson MR, Lister AM, Penkman KEH. A new method for enamel amino acid racemization dating: A closed system approach. Quat Geochronol. 2019;50:29–46. https://doi.org/10.1016/j.quageo.2018.11.005

29. Kaufman DS, Manley WF. A new procedure for determining dl amino acid ratios in fossils using reverse phase liquid chromatography. Quat Sci Rev. 1998;17:987–1000. https://doi.org/10.1016/S0277-3791(97)00086-3

30. Stewart NA, Molina GF, Issa JPM, Yates NA, Sosovicka M, Vieira AR, et al. The identification of peptides by nanoLC-MS/MS from human surface tooth enamel following a simple acid etch extraction. RSC Adv. 2016;6:61673–61679. https://doi.org/10.1039/C6RA05120K

31. Rappsilber J, Ishihama Y, Mann M. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. Anal Chem. 2003;75:663–670. https://doi.org/10.1021/ac026117i

32. Taurozzi AJ, Rüther PL, Patramanis I, Koenig C, Paterson RS, Madupe PP, et al. Deep-time phylogenetic inference by paleoproteomic analysis of dental enamel. Nat Protoc. 2024;19:2085–2116. https://doi.org/10.1038/s41596-024-00975-3

33. Mackie M, Rüther P, Samodova D, Di Gianvincenzo F, Granzotto C, Lyon D, et al. Palaeoproteomic profiling of conservation layers on a 14th century Italian wall painting. Angew Chem Int Ed. 2018;57:7369–7374. https://doi.org/10.1002/anie.201713020

34. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol. 2008;26:1367–1372. https://doi.org/10.1038/nbt.1511

35. Patramanis I, Ramos-Madrigal J, Cappellini E, Racimo F. PaleoProPhyler: A reproducible pipeline for phylogenetic inference using ancient proteins. Peer Commun J. 2023;3, e112. https://doi.org/10.24072/pcjournal.344

36. Pickering R, Kramers JD, Hancox PJ, de Ruiter DJ, Woodhead JD. Contemporary flowstone development links early hominin bearing cave deposits in South Africa. Earth Planet Sci Lett. 2011;306:23–32. https://doi.org/10.1016/j.epsl.2011.03.019

37. Zanolli C, Davies TW, Joannes-Boyau R, Beaudet A, Bruxelles L, de Beer F, et al. Dental data challenge the ubiquitous presence of *Homo* in the Cradle of Humankind. Proc Natl Acad Sci USA. 2022;119, e2111212119. https://doi.org/10.1073/pnas.2111212119

38. Pickering R, Hancox PJ, Lee-Thorp JA, Grün R, Mortimer GE, McCulloch M, et al. Stratigraphy, U-Th chronology, and paleoenvironments at Gladysvale Cave: Insights into the climatic control of South African hominin-bearing cave deposits. J Hum Evol. 2007;53:602–619. https://doi.org/10.1016/j.jhevol.2007.02.005

39. Pickering R, Herries AIR, Woodhead JD, Hellstrom JC, Green HE, Paul B, et al. U-Pb-dated flowstones restrict South African early hominin record to dry climate phases. Nature. 2018;565:226–229. https://doi.org/10.1038/s41586-018-0711-0

40. Kohn MJ, Schoeninger MJ, Barker WW. Altered states: Effects of diagenesis on fossil tooth chemistry. Geochim Cosmochim Acta. 1999;63:2737–2747. https://doi.org/10.1016/S0016-7037(99)00208-2

41. Cheng JB, Liu Q, Long F, Huang DX, Yi SH. Analysis of the Yp11.2 deletion region of phenotypically normal males with an AMELY-Null allele in the Chinese Han population. Genet Test Mol Biomark. 2019;23:359–362. https://doi.org/10.1089/gtmb.2018.0231

42. Mitchell RJ, Kreskas M, Baxter E, Buffalino L, Van Oorschot RAH. An investigation of sequence deletions of amelogenin (AMELY), a Y-chromosome locus commonly used for gender determination. Ann Hum Biol. 2006;33:227–240. https://doi.org/10.1080/03014460600594620

43. Skov L, Peyrégne S, Popli D, Iasi LNM, Devièse T, Slon V, et al. Genetic insights into the social organization of Neanderthals. Nature. 2022;610:519–525. https://doi.org/10.1038/s41586-022-05283-y

44. Parker GJ, Yip JM, Eerkens JW, Salemi M, Durbin-Johnson B, Kiesow C, et al. Sex estimation using sexually dimorphic amelogenin protein fragments in human enamel. J Archaeol Sci. 2019;101:169–180. https://doi.org/10.1016/j.jas.2018.08.011

45. Froment C, Hourset M, Sáenz-Oyhéréguy N, Mouton-Barbosa E, Willmann C, Zanolli C, et al. Analysis of 5000 year-old human teeth using optimized large-scale and targeted proteomics approaches for detection of sex-specific peptides. J Proteomics. 2020;211, Art. #103548. https://doi.org/10.1016/j.jprot.2019.103548

46. Koenig C, Bortel P, Paterson RS, Rendl B, Madupe PP, Troché GB, et al. Automated high-throughput biological sex identification from archaeological human dental enamel using targeted proteomics [preprint]. bioRxiv 2024.02.20.581140. https://doi.org/10.1101/2024.02.20.581140

47. Lu Y, Papagerakis P, Yamakoshi Y, Hu JCC, Bartlett JD, Simmer JP. Functions of KLK4 and MMP-20 in dental enamel formation. Biol Chem. 2008;389:695–700. https://doi.org/10.1515/BC.2008.080

48. Yamakoshi Y, Hu JCC, Fukae M, Yamakoshi F, Simmer JP. How do enamelysin and kallikrein 4 process the 32-kDa enamelin? Eur J Oral Sci. 2006;114(Suppl 1):45–51. Discussion 93-5, 379-80. https://doi.org/10.1111/j.1600-0722.2006.00281.x

49. Degnan JH, Rosenberg NA. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol Evol. 2009;24:332–340. https://doi.org/10.1016/j.tree.2009.01.009

50. Paterson RS, Madupe PP, Cappellini E. Paleoproteomics sheds light on million-year-old fossils. Nat Rev Mol Cell Biol. 2024;26:1–2. https://doi.org/10.1038/s41580-024-00803-2

51. Paijmans JLA, Gilbert MTP, Hofreiter M. Mitogenomic analyses from ancient DNA. Mol Phylogenet Evol. 2013;69:404–416. https://doi.org/10.1016/j.ympev.2012.06.002

52. Orlando L, Allaby R, Skoglund P, Der Sarkissian C, Stockhammer PW, Ávila-Arcos MC, et al. Ancient DNA analysis. Nat Rev Methods Primers. 2021;1:14. https://doi.org/10.1038/s43586-020-00011-0

53. Knapp M, Hofreiter M. Next Generation Sequencing of Ancient DNA: Requirements, Strategies and Perspectives. Genes. 2010;1:227–243. https://doi.org/10.3390/genes1020227

54. Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, et al. A draft sequence of the Neandertal genome. Science. 2010;328:710–722. https://doi.org/10.1126/science.1188021

55. Meyer M, Kircher M, Gansauge MT, Li H, Racimo F, Mallick S, et al. A high-coverage genome sequence from an archaic Denisovan individual. Science. 2012;338:222–226. https://doi.org/10.1126/science.1224344

56. Reich D, Green RE, Kircher M, Krause J, Patterson N, Durand EY, et al. Genetic history of an archaic hominin group from Denisova cave in Siberia. Nature. 2010;468:1053–1060. https://doi.org/10.1038/nature09710

57. Floyd BM, Marcotte EM. Protein sequencing, one molecule at a time. Annu Rev Biophys. 2022;51:181–200. https://doi.org/10.1146/annurev-biophys-102121-103615

58. Lee-Thorp J, Sponheimer M. The development of chemical approaches to fossil hominin ecology in South Africa. S Afr J Sci. 2025;121(1/2), Art. #18529. https://doi.org/10.17159/sajs.2025/18529

59. Adame F. Meaningful collaborations can end 'helicopter research'. Nature Career Column. 29 June 2021. https://doi.org/10.1038/d41586-021-01795-1

60. Schroeder L, Ackermann RR. Moving beyond the adaptationist paradigm for human evolution, and why it matters. J Hum Evol. 2023;174, Art. #103296. https://doi.org/10.1016/j.jhevol.2022.103296

61. Elbein A. Decolonizing the hunt for dinosaurs and other fossils. The New York Times. 22 March 2021. Available from: https://www.nytimes.com/2021/03/22/science/dinosaurs-fossils-colonialism.html

62. Athreya S, Ackermann R. Colonialism and narratives of human origins in Asia and Africa. In: Porr M, Matthews J. Interrogating human origins. London: Routledge; 2019. p. 72–95. https://doi.org/10.4324/9780203731659-4

63. Stevens KR, Masters KS Imoukhuede PI, Haynes KA, Setton LA, Cosgriff-Hernandez E, et al. Fund Black scientists. Cell. 2021;184:561–565. https://doi.org/10.1016/j.cell.2021.01.011

64. Gomez LE, Bernet P. Diversity improves performance and outcomes. J Natl Med Assoc. 2019;111:383–392. https://doi.org/10.1016/j.jnma.2019.01.006