Article:
Madupe, Palesa, Koenig, Claire, Patramanis, Ioannis et al. (35 more authors) (2025) Enamel Proteins Reveal Biological Sex and Genetic Variability within Southern African Paranthropus. Science. pp. 969-973. ISSN: 0036-8075

https://doi.org/10.1126/science.adt9539

Published in final edited form as:
Science. 2025 May 29; 388(6750): 969–973. doi:10.1126/science.adt9539.

Enamel proteins reveal biological sex and genetic variability in southern African *Paranthropus*

Palesa P. Madupe[#1,2], **Claire Koenig**[#3], **Ioannis Patramanis**[#1], **Patrick L. Rüther**[3], **Nomawethu Hlazo**[2,4], **Meaghan Mackie**[1,3], **Mirriam Tawane**[2,5], **Johanna Krueger**[6], **Alberto J. Taurozzi**[1], **Gaudry Troché**[1,3], **Job Kibii**[7], **Robyn Pickering**[2,8], **Marc R. Dickinson**[9], **Yonatan Sahle**[2,4], **Dipuo Kgotleng**[10], **Charles Musiba**[2,11], **Fredrick Manthi**[2,12], **Liam Bell**[13], **Michelle DuPlessis**[13], **Catherine Gilbert**[14], **Bernhard Zipfel**[15], **Lukas F.K. Kuderna**[6,‡], **Esther Lizano**[16], **Frido Welker**[1], **Pelagia Kyriakidou**[17], **Jürgen Cox**[17,18], **Catherine Mollereau**[19], **Caroline Tokarski**[14], **Jonathan Blackburn**[20], **Jazmín Ramos-Madrigal**[1], **Tomas Marques-**

[*]Corresponding authors. jesper.olsen@cpr.ku.dk; becky.ackermann@uct.ac.za; ecappellini@sund.ku.dk .
[‡]Lukas F. K. Kuderna is currently an employee of Illumina Inc.

Author contributions

P. P. M, Formal analysis, Investigation, Methodology, Software, Visualization, Writing - original draft and Writing -review & editing.
C.K, Formal analysis, Investigation, Methodology, Software, Visualization, Writing - original draft and Writing -review & editing.
I. P, Formal analysis, Methodology, Software, Visualization, Data curation, Writing - original draft and Writing -review & editing.
P. L. R, Formal analysis, Investigation, Methodology, Software and Writing -review & editing.
N. H, Investigation.
M. M., Investigation and Writing -review & editing.
M. T., Resources and Writing -review & editing.
J. K., Software, Visualization, and Writing -review & editing.
A. J. T., Investigation and Writing -review & editing.
G. T., Investigation.
J. K., Writing -original draft and Writing -review & editing.
R. P., Visualization, Writing -original draft and Writing -review & editing.
M. R. D., Investigation, Writing -original draft and Writing -review & editing.
Y. S., Writing -original draft and Writing -review & editing.
D. K., Writing -original draft and Writing -review & editing.
C. M, Writing -original draft and Writing -review & editing.
F. M., Writing -original draft and Writing -review & editing.
L. B., Validation.
M. D., Validation.
C. G., Software.
B. Z., Resources, Writing -review & editing.
L. F. K. K., Writing -review & editing.
E. L., Writing -review & editing.
F. W, Investigation and Writing -review & editing.
P. K., Software.
J. C., Supervision.
C. M., Writing -review & editing.
C. T., Supervision.
J. B., Supervision.
J. R-M., Data curation and Writing -review & editing.
T. M-B., Supervision and Writing -review & editing.
K. P., Supervision, Funding acquisition and Writing -review & editing.
C. Z., Data curation, Methodology, Software, Writing -original draft and Writing -review & editing.
L. S., Data curation, Methodology, Software, Writing -original draft and Writing -review & editing.
F. R., Supervision and Writing -review & editing.
J. V. O., Conceptualization, Funding acquisition, Supervision and Writing -review & editing. R. R. A., Conceptualization, Funding acquisition, Project administration, Supervision and Writing -review & editing.
E. C, Conceptualization, Methodology, Funding acquisition, Project administration, Supervision and Writing -review & editing.

Competing interests: Authors declare that they have no competing interests.

Europe PMC Funders Author Manuscripts

Bonet[6,16,21,22], **Kirsty Penkman**[9], **Clément Zanolli**[15,23], **Lauren Schroeder**[2,24], **Fernando Racimo**[1], **Jesper V. Olsen**[3,*], **Rebecca R. Ackermann**[2,4,*], **Enrico Cappellini**[1,*]

[1]Globe Institute, University of Copenhagen; Copenhagen K, 1172, Denmark.

[2]Human Evolution Research Institute (HERI), University of Cape Town; Cape Town, 7701, South Africa.

[3]Novo Nordisk Foundation Center for Protein Research, University of Copenhagen; Copenhagen N, 2200, Denmark.

[4]Department of Archaeology, University of Cape Town; Cape Town, 7701, South Africa.

[5]National Heritage Council South Africa; Pretoria, 0040, South Africa.

[6]Institute of Evolutionary Biology (UPF-CSIC), PRBB; Barcelona, 08003, Spain.

[7]Turkana Basin Institute; Nairobi, 00502, Kenya

[8]Department of Geological Sciences, University of Cape Town; Cape Town, 7701, South Africa.

[9]Department of Chemistry, University of York; York, YO10 5DD, United Kingdom.

[10]Palaeo-Research Institute, University of Johannesburg; Johannesburg, 2092, South Africa.

[11]Department of Evolutionary Anthropology, Duke University; Durham, NC 27708, United States of America.

[12]Palaeontology Section, National Museums of Kenya; Nairobi, 00100, Kenya.

[13]D-CYPHR, Centre for Proteomic and Genomic Research; Cape Town, 7925, South Africa

[14]CBMN, University of Bordeaux; Bordeaux, 33076, France.

[15]Evolutionary Studies Institute, University of the Witwatersrand; Johannesburg, 2017, South Africa.

[16]ICP-CERCA, Unidad Asociada al CSIC por el IBE UPF-CSIC; Barcelona, 08201, Spain

[17]Computational Systems Biochemistry, Max Planck Institute of Biochemistry; Martinsried, 82152, Germany.

[18]Department of Biological and Medical Psychology, University of Bergen; Bergen, 7807, Norway.

[19]Research Center on Animal Cognition (CRCA), University of Toulouse; Toulouse, 31400, France.

[20]Institute of Infectious Disease and Molecular Medicine, University of Cape Town; Cape Town, 7925, South Africa.

[21]Catalan Institution of Research and Advanced Studies (ICREA); Barcelona, 08010, Spain.

[22]CNAG, Centro Nacional de Analisis Genomico; Barcelona, 08028, Spain.

[23]PACEA, University of Bordeaux; Pessac, 33600, France.

[24]Department of Anthropology, University of Toronto Mississauga; Toronto, ON M5S 2S2, Canada.

[#] These authors contributed equally to this work.

Abstract

Paranthropus robustus is a morphologically well-documented Early Pleistocene hominin species from southern Africa with no genetic evidence reported so far. We describe the mass spectrometric sequencing of enamel peptides from four ca. 2 million year old dental specimens attributed morphologically to *P. robustus*, from the site of Swartkrans in South Africa. The identification of AMELY-specific peptides enabled us to assign two specimens to male individuals while semi-quantitative mass spectrometric data analysis attributed the other two to females. A single amino acid polymorphism and the enamel-dentine junction shape variation indicate potential subgroups present within southern African *Paranthropus*. This study demonstrates how palaeoproteomics can help distinguish sexual dimorphism from other sources of variation in African Early Pleistocene hominins.

While our understanding of the evolution of Middle to Late Pleistocene hominins is becoming increasingly clear, in large part due to ancient DNA (aDNA) sequencing data (1, 2), the biological and behavioral variation among earlier Plio-Pleistocene hominins remains poorly understood. The genus *Paranthropus* first appeared in the fossil record ca. 2.8 million years ago (Ma) and persisted until 1 Ma, coexisting in time and space with a number of other hominins, including *Australopithecus* species and members of the genus *Homo*. Most researchers consider *Paranthropus* to be monophyletic (3) however, morphological similarities between *P. robustus* and *A. africanus* (4, 5), and between *P. aethiopicus* and *A. afarensis* (6, 7), have raised the possibility of paraphyly or even admixture between species (8, 9). Furthermore, analyses of the enamel-dentine junction (EDJ) of southern African *Paranthropus* indicate significant variation, suggesting the possibility of detectable substructure within *P. robustus* (10), or even the presence of more than one species of this genus in southern Africa (11, 12). Other researchers have argued that the observed morphological differences stem from sexual dimorphism (13). Determining to what extent the variation within and between Plio-Pleistocene hominins is due to evolutionary diversification versus intraspecific variation, of which sexual dimorphism is likely a major contributor, is fundamental to interpreting their evolutionary history.

Although genetic data from the African continent have provided insights of unprecedented resolution into human demography and evolution (14, 15), aDNA has never been successfully recovered from African hominin material older than ~ 0.02 Ma (15). As phylogenetically informative ancient protein sequences have been retrieved beyond the limits of aDNA preservation in Eurasia (16–18), we attempted to recover them to help investigate the causes of Plio-Pleistocene hominin variation in Africa. We used liquid chromatography coupled to high-resolution tandem mass spectrometry to reconstruct dental enamel protein sequences from four southern African hominin specimens assigned to *P. robustus* (Supplementary Materials).

The four hominin fossils analyzed (SK 830, 835, 850, and 14132) originated from Swartkrans cave, located approximately 40 km northwest of Johannesburg, in South Africa's Cradle of Humankind World Heritage Site (Fig. 1). The teeth are from the oldest deposits at Swartkrans, Member 1 (MB1), which is dated to between 1.8 and 2.2 Ma (19, 20). Although Swartkrans has produced the largest collection of specimens attributed to *P. robustus*, the

relationships between this material and the *Paranthropus* fossils from other southern African sites has been the subject of various interpretations, summarized by Martin et al. (21).

Results

To maximize the breadth and depth of amino acid sequence coverage, manual off-line high-pH reversed-phase fractionation was carried out on StageTips. This strategy extends the dynamic range of the less complex fractions for subsequent MS analysis (22), increasing peptide identifications in all four *Paranthropus* samples (Fig. S1). The number of recovered amino acid positions increased up to 17%, and the number of peptide-spectrum matches increased up to 3-fold (Supplementary Materials). Further methodological development was achieved with the creation of an automated and open-source sequence assembly pipeline (Fig. S2). A site-based sequence reconstruction approach (23) was developed to generate consensus sequences directly from the MaxQuant output tables. This sequence assembly pipeline enables faster, more reproducible, and transparent data analysis processes. The generated outputs can be traced back to the original fragmentation spectra, thereby simplifying manual validation of ambiguous hits.

After successful proteomics analysis on faunal material (Fig. S3), dental enamel from four *Paranthropus* individuals was sampled. The combined analysis of the LC-MS/MS data obtained from fractionated and single-shot samples of each individual resulted in 4,600 to 8,500 PSMs covering 540 to 780 amino acid positions from 8 to 10 enamel-associated proteins (24–26), six of which, i.e. ALB, AMBN, AMELX, COL17A1, ENAM and MMP20, appeared in all analyzed specimens. A total of 425 amino acid positions were consistently identified in all four *Paranthropus* specimens, indicating that the majority of the covered positions are shared across all the samples (Fig. S4). For validation, the MS workflow was successfully replicated in a proteomics laboratory in Cape Town, South Africa (Fig. S5, (Supplementary Materials).

The authenticity of the recovered sequences is supported by multiple lines of evidence. First, the relationship between free and bound amino acids and their expected levels of racemization, as well as the extent of peptide bond hydrolysis, indicates that, in all the samples, the dental enamel is exhibiting behavior consistent with a closed system (Fig. S6, (Supplementary Materials)). Equally, across all four specimens, no, or negligible, exogenous contamination was supported by the high similarity of the amino acid composition profiles observed both within our sample set and in comparison to other ancient dental enamel specimens previously investigated (Figs. S6 and S7, (Supplementary Materials)). Second, all samples show advanced rates of diagenetically-induced amino acid modifications, such as glutamine and asparagine deamidation and arginine to ornithine conversion, compatible with the age and the geographic origin of the *Paranthropus* specimens (Fig. S4). Additionally, we observed extended oxidative modification of histidine, phenylalanine, tyrosine and tryptophan (Fig. S8, (Supplementary Materials)). Third, the peptide length distribution is skewed towards shorter amino acid chains compared to modern human dental enamel, as previously observed in other paleoanthropological material and in agreement with the high levels of peptide bond hydrolysis observed in the amino acid analysis (Fig. S4, (Supplementary Materials)). Altogether, these lines of evidence independently support the

Europe PMC Funders Author Manuscripts

authenticity of the ancient amino acid sequences we report. An attempt to detect protein-protein crosslinks did not lead to any confident identification (Supplementary Materials).

Specimens SK 850 and SK 835 were unambiguously identified as male *Paranthropus* individuals based on the observation of multiple overlapping AMELY-specific peptides (Fig. 2A, (Supplementary Materials)). No AMELY-specific peptide was detected in SK 830 and SK 14132. This absence alone, however, cannot necessarily lead to a female attribution (27, 28), for it is also consistent with these specimens belonging to male individuals whose signal for the AMELY-specific peptides would not be detectable by MS, or their acquisition would be stochastic. To exclude the latter scenario, we used a site-based semi-quantitative approach. Specifically, we defined an AMELX intensity threshold above which AMELY-specific peptides should be consistently detectable, if present in the sample. For validation, we apply this approach to 11 modern human enamel specimens, achieving sex attribution with 100% accuracy. Since the AMELX site intensities of both SK 830 and SK 14132 were measured above the defined intensity thresholds, and above the intensity of SK 835, we infer that both specimens originated from female individuals (Fig. 2 and fig. S9, (Supplementary Materials)).

We used available buccolingual and mesiodistal measurements of SK 830, SK 835, and SK 850 to compare our molecular-based sex attributions with those based on overall tooth size (Supplementary Materials). A sample of *A. africanus* is also included for comparison to better assess the extent of size variation in *Paranthropus* (29). SK 14132 is not included in the comparison, as its incompleteness precludes reliable measurement (Fig. S10). SK 830, which is assigned to a female individual based on molecular evidence, has mesiodistal and buccolingual measurements consistent with specimens previously regarded as female. SK 850, assigned to a male based on AMELY-specific peptides, has an mesiodistal measurement falling within the lower range of size variability seen among specimens previously considered to be males. SK 835, recently suggested to possibly belong to a female individual based on tooth crown dimensions (30), is here confidently identified through AMELY-specific peptides as originating from a male. Our results thus indicate that measurements of dental size are not necessarily accurate for correct sex estimation. The protein-based male attribution of SK 835 is consistent with its local strontium isotope signal previously suggesting philopatric male behavior (31).

After aligning the amino acid sequences assembled for each *Paranthropus* specimen, we identified a subset of 425 positions that are covered in all four individuals. Among these positions, we detect a variable site in enamelin (ENAM) position 137 (based on *Homo sapiens* canonical Ensembl transcript - ENST00000396073.4). Individuals SK 830 and SK 850 bear a fully deamidated glutamine (Q) in that position, while SK 835 has an arginine (R) (Fig. 3). Additionally, in SK 14132 the ENAM-137 site appears to be heterozygous, with the Q allele covered in 80% of the spectra (18 vs 4 PSMs). The confident identification of the two ENAM-137 alleles is further confirmed by the analysis of synthetic peptides (Fig. 3 and Fig. S11).

To assess how unexpected the within-sample variation would be if all four individuals belonged to a single species, we repeatedly sampled four randomly selected individuals

from a global sample of present-day humans (Supplementary Materials). We found it plausible that genetic variants segregating within a given species of equivalent diversity to modern humans could manifest as amino acid differences in a sample of the same size as the one we had for *Paranthropus*. However, we note that the effective population size of humans today most likely differs from that of the *Paranthropus* population we sampled (Supplementary Materials), leaving any taxonomic conclusion based on this genetic variation alone premature. When we compared the reconstructed protein sequences from the four *Paranthropus* individuals with their orthologs from a panel of extant and extinct hominids (great apes and humans), we detected a total of 16 species-informative single amino acid polymorphisms (Table S1). Based on these phylogenetically informative sites, all four *Paranthropus* protein sequences appear closer to those in the *Homo* clade than to any other primate (Figs. S12, S13-24). Thus, the placement of our *Paranthropus* samples agrees with the consensus view based on morphology (32, 33), further supporting the endogeneity of the protein sequences we retrieved. Given the limited total number of informative sites, conclusive phylogenetic results will require broader sequence coverage. Out of the 16 identified single amino acid polymorphisms, only two, i.e. COL17A1-636 and ENAM-137, showed an allelic state different from that of present-day humans, Neanderthals, and Denisovans. The mapping of these two variants on the primate phylogenetic tree (Fig. S12) indicates that, while COL17A1-636 is likely an ancestral variant compared to modern humans, the ENAM-137 Q is most likely a *Paranthropus* derived variant.

To integrate the evidence obtained with paleoproteomics and more established morphological approaches (11), geometric morphometric analyses of the enamel-dentine junction shape were carried out on the two best-preserved specimens (SK 835 and SK 830) (Fig. S10). *Paranthropus* M^3 and P_4 had a more asymmetric and taller enamel-dentine junction than those of early *Homo* and, to a lesser extent, than in *Australopithecus*. The results showed that both specimens belong to *Paranthropus* and differ from *Australopithecus* and early *Homo* (Supplementary Materials). Noticeably, the enamel-dentine junction of SK 835 M^3, bearing ENAM-137R, is statistically more similar to the *Paranthropus* specimens from the site of Drimolen, showing a wider occlusal basin, than to those from the Swartkrans and Kromdraai assemblages, the latter including the holotype of *P. robustus* (TM 1517). The enamel-dentine junction shape of the P_4 SK 830, in contrast, more closely resembles specimens from Swartkrans and Kromdraai and statistically differs from the Drimolen material (Fig. 4).

Discussion and broader implications

We report the recovery of Early Pleistocene hominin partial protein sequences from southern Africa. The four *Paranthropus* specimens we studied were recovered from cave sediments mostly composed of remobilized soil from outside the cave. Sedimentological evidence points to the fossil accumulations resulting from rapid, episodic flash floods, occurring during relatively arid climate conditions (34). This, coupled with extensive cementation of the fossil bearing sediments, explains the richness of fossils in the caves, and may have also favorably contributed to the preservation of the proteins within the fossil teeth. Whether or not protein preservation would be comparable in other early hominin-bearing deposits, including open-air sites such as those found elsewhere in Africa, is an open question. Future

work must give attention to the feasibility of biomolecular study, while minimizing damage to precious African heritage.

The application of manual off-line reversed-phase high-pH fractionation improved the dental enamel protein sequence coverage (Fig. S9), revealing the existence of diversity at the protein sequence level within southern African *Paranthropus* individuals. In addition, spectral prediction software and sequenced synthetic peptides helped validate spectra that provided mass spectrometric evidence to confirm this diversity and detect heterozygosity. Mass spectrometry has been previously utilized as a method of choice for the detection of both novel genetic variants and heterozygosity in modern human individuals (35–37). To our knowledge, however, this has never been previously applied in the context of paleoproteomics. Future studies should further explore this aspect.

The molecular identification of both male and female *Paranthropus* individuals demonstrates the limitations of sexing techniques based on tooth size (30). This capability has clear implications for our understanding and interpretation of morphological variation in the deep time hominin fossil record, as it enables us to exclude sexual dimorphism as one of the multiple variables affecting the range of anatomical variation. However, as the method we use is reliant on the identification of AMELY-specific amino acid sequences, male individuals with a deletion of the AMELY gene will not be detected. Deletions of this gene are uncommon, but have been recorded both in modern human populations (38, 39) and in one Neanderthal individual (40).

Due to a single SAP (ENAM–137), our reconstructed trees (Fig. S13-S24) suggest that one of the *Paranthropus* individuals (SK 835) might be more distantly-related to the other three individuals. Although potentially the result of incomplete lineage sorting, it is also possible that this individual may have belonged to a distinct *Paranthropus* group, which separated from the other individuals relatively recently, a hypothesis compatible with both the paleoproteomics and enamel dentine junction morphology, but microevolution of a single taxon over time (21, 41, 42) or a large effective population size, could also explain the observed genetic and morphological variability. The southern African *Paranthropus* assemblage exhibits considerable size variation, most of which has previously been attributed to sexual dimorphism, possibly reflecting a gorilla-like pattern of extended growth for males, i.e. bimaturation (13). However, recent studies have suggested that these morphological differences might indicate either different taxa (11, 21), or site-related diversity within a single species over time, i.e. micro-evolutionary changes, following a morphocline (41, 43, 44). The recent description of P. capensis, a gracile species of *Paranthropus*, indicates that taxonomic diversity within this genus is currently underestimated and needs to be investigated further (45). Regardless, as the small-crowned SK 835 is now confidently assigned to a male, sexual dimorphism is an insufficient explanation for the observed variation within our *Paranthropus* sample set.

The four specimens we analyzed, along with other fossils from Swartkrans Member 1, were not accurately mapped when they were collected, thus they come from sediments accumulating within a 500 ka time window. Nevertheless, while Member 1 sediments potentially cover a long time interval, the fossil-bearing deposits likely accumulated rapidly

Europe PMC Funders Author Manuscripts

(34). A rapid accumulation would imply that the *Paranthropus* teeth analyzed here are penecontemporaneous and, consequently, that the observed variation may more likely result from taxon diversity than changes of a single taxon over time. Further combined paleoproteomic and morphometric sampling of *Paranthropus* individuals, from Swartkrans and Drimolen, the latter of which hosts samples that are morphologically closer to SK 835, could inform the source of this variation. Given that ENAM-137 was covered with high confidence in all four of our samples, differences in allele frequencies of this SAP between specimens from various paleoanthropological sites would consolidate site-specific *Paranthropus* variation, while more precise and direct dating techniques would help distinguish between microevolution of a single taxon and intra-taxonomic diversity.

We show that the analysis of multiple individuals, in conjunction with morphological evidence, can better explore their genetic history and illuminate variation potentially indicative of inter- or intra-taxon diversity. Successful protein extraction should be achievable for hominins recovered in other southern African cave sites of similar age and geology, making biological sex identification and intra-species analysis possible. This study, as well as preliminary results recently published (46) also raise the possibility of extracting similar data from other African early hominin material, such as *A. africanus* and *A. afarensis* that are represented by a plethora of isolated and fragmentary dental remains. The Cradle of Humankind has yielded an exceptionally large number of hominin fossils, yet the greatest diversity of hominin taxa is currently known from eastern African sites, mainly in the rift valley regions of Ethiopia, Kenya and Tanzania. Whether and how much of this phyletic diversity is real, and not the result of methodological limitations and/or research(er) bias, remains a debated topic. The coherent results obtained from this study combining molecular and morphological data have implications for addressing such long-standing controversies surrounding the nature and extent of Plio-Pleistocene hominin diversity (47–49).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

All the *Paranthropus* fossil specimens analyzed for this study are part of the permanent collection of the Ditsong National Museum of Natural History, in Pretoria (South Africa). All non-hominin faunal fossil specimens analyzed for this study are part of the permanent collection of the University of the Witwatersrand, in Johannesburg (South Africa). Permits for temporary export and sampling were issued by the South African Heritage Resources Agency (SAHRA permit IDs: 2946, 3079 and 3132). Access to these specimens can be requested through Ditsong National Museum and the University of the Witwatersrand.

- Contact person for Ditsong National Museum of Natural History is Dr. Lazarus Kgasi the junior curator of Plio-Pleistocene Paleontology, lkgasi@ditsong.org.za phone number: +27 12 492 5807
  - Specimen Ids: SK 830, SK 835, SK 850 and SK 14132
- Contact person at the University of the Witwatersrand is Dr. Bernhard Zipfel the curator of fossils and rock collections, Bernhard.Zipfel@wits.ac.za, phone +27 11 717 6683
  - Specimen Ids: SKX 37041, SKX 3730, SKX 4996 a, CD. 5410

We acknowledge Ricardo Fong-Zazueta for creating and providing the protein translations of the ‘independent’ reference dataset. This research contributes towards the output of the Biogeochemistry Research Infrastructure Platform (BIOGRIP), supported by the Department of Science and Innovation, South Africa.

Europe PMC Funders Author Manuscripts

Funding

European Research Council (ERC) under the European Union's Horizon 2020 grant agreements No. 861389 "PUSHH" (PPM, CK, IP, JK, CG, PK, JC, KP, FR, TMB, JVO and EC)

European Research Council (ERC) under the European Union's Horizon 2020 grant agreements No. 722606 No. 722606 "TEMPERA" (PLR, EC, and JVO)

European Research Council (ERC) under the European Union's Horizon 2020 grant agreements No. 101021361 "BACKWARD" (EC, PPM, IP, CZ, LS, FR, JVO, RRA, AJT, MM and GT)

VILLUM FONDEN grant agreement No. 17649 (EC, FW, JRM)

Danish National Research Foundation "PROTEIOS" (MM)

Danish National Research Foundation "DNRF128" (MM)

The Novo Nordisk Foundation grant agreement No. NNF14CC0001 (CK, PLR, JVO)

National Research Foundation of South Africa grant agreement No. 117670 (RRA, NH)

National Research Foundation of South Africa grant agreement No. 136512 (RRA, NH)

Natural Sciences and Engineering Research Council of Canada Discovery Grant grant agreement No. RGPIN-2020-04159 (LS)

VILLUM FONDEN grant agreement No. 00025300 (FR)

Novo Nordisk Foundation Data Science Ascending Investigator Award grant agreement No. NNF22OC0076816 (FR)

European Research Council (ERC) under the European Union's Horizon 2020 grant agreements No. 101077592 (FR)

European Research Council (ERC) under the European Union's Horizon 2020 grant agreements No. 951385 (FR)

European Research Council (ERC) under the European Union's Horizon 2020 grant agreement No. 864203 (TMB)

Natural Environment Research Council (NERC) grant agreement No. NE/S010211/1 (MD, KP)

Department of Science and Innovation in South Africa "DIPLOMICS" (LB, MLP)

French National Research Agency "GenoMorph" grant agreement No. ANR-20-CE12-0018 (CZ)

IdEx "Investments for the Future" program / GPR "Human Past" (CZ)

Data and materials availability

Data availability

- The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository (50) with the dataset identifier **PXD040221**.
- Reference Datasets, XML files and phylogenetic results files are available on Zenodo: https://zenodo.org/records/10843737

Code availability

- Custom R-code for sequence assembly is available on GitHub at: https://github.com/ClaireKoenig/ProteinSequenceAssembly

- Genetic Variation analysis code is available on GitHub at: https://github.com/johnpatramanis/Code_for_Genetic_Diversity_Sampling
- Commands for the generation of the phylogenetic workflow is available in the Supplementary Materials.

References and Notes

1. Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MH-Y, Hansen NF, et al. A draft sequence of the Neandertal genome. Science. 2010; 328: 710–722. DOI: 10.1126/science.1188021 [PubMed: 20448178]
2. Reich D, Green RE, Kircher M, Krause J, Patterson N, Durand EY, Viola B, Briggs AW, Stenzel U, Johnson PLF, Maricic T, et al. Genetic history of an archaic hominin group from Denisova Cave in Siberia. Nature. 2010; 468: 1053–1060. DOI: 10.1038/nature09710 [PubMed: 21179161]
3. Strait DS, Grine FE. Inferring hominoid and early hominid phylogeny using craniodental characters: the role of fossil taxa. J Hum Evol. 2004; 47: 399–452. [PubMed: 15566946]
4. Harrison T. Evolutionary history of the "robust" australopithecines. Int J Primatol. 1991; 12: 85–88.
5. Wood, B, Schroer, K. Vertebrate Paleobiology and Paleoanthropology. Springer International Publishing; Cham: 2017. 95–107.
6. Kimbel WH, White TD. A revised reconstruction of the adult skull of *Australopithecus afarensis*. J Hum Evol. 1988; 17: 545–550.
7. Suwa G. Evolution of the "robust" australopithecines in the Omo succession: evidence from mandibular premolar morphology. Evolutionary history of the "robust" australopithecines. 1988.
8. Schroeder L, Ackermann RR. Moving beyond the adaptationist paradigm for human evolution, and why it matters. J Hum Evol. 2023; 174 103296 [PubMed: 36527977]
9. Hall R. Paleobiology and systematics of canids and hominids. J Hum Evol. 1977; 6: 519–533.
10. Braga J, Zimmer V, Dumoncel J, Samir C, de Beer F, Zanolli C, Pinto D, Rohlf FJ, Grine FE. Efficacy of diffeomorphic surface matching and 3D geometric morphometrics for taxonomic discrimination of Early Pleistocene hominin mandibular molars. J Hum Evol. 2019; 130: 21–35. [PubMed: 31010541]
11. Zanolli C, Davies TW, Joannes-Boyau R, Beaudet A, Bruxelles L, de Beer F, Hoffman J, Hublin J-J, Jakata K, Kgasi L, Kullmer O, et al. Dental data challenge the ubiquitous presence of *Homo* in the Cradle of Humankind. Proc Natl Acad Sci USA. 2022; 119 e2111212119 doi: 10.1073/pnas.2111212119 [PubMed: 35787044]
12. Skinner MM, Bailey SE, Gunz P, Kimbel WH, Alemseged Z, Delezene LK, Menter C, Moggi-Cecchi J, Kupczik K. Below the crown: examining interspecies variation in postcanine enamel thickness, EDJ, and root form in the *Paranthropus* clade. 2018.
13. Lockwood CA, Menter CG, Moggi-Cecchi J, Keyser AW. Extended male growth in a fossil hominin species. Science. 2007; 318: 1443–1446. [PubMed: 18048687]
14. Skoglund P, Thompson JC, Prendergast ME, Mittnik A, Sirak K, Hajdinjak M, Salie T, Rohland N, Mallick S, Peltzer A, Heinze A, et al. Reconstructing prehistoric African population structure. Cell. 2017; 171: 59–71. e21 doi: 10.1016/j.cell.2017.08.049 [PubMed: 28938123]
15. Lipson M, Sawchuk EA, Thompson JC, Oppenheimer J, Tryon CA, Ranhorn KL, de Luna KM, Sirak KA, Olalde I, Ambrose SH, Arthur JW, et al. Ancient DNA and deep population structure in sub-Saharan African foragers. Nature. 2022; 603: 290–296. DOI: 10.1038/s41586-022-04430-9 [PubMed: 35197631]
16. Cappellini E, Welker F, Pandolfi L, Ramos-Madrigal J, Samodova D, Rüther PL, Fotakis AK, Lyon D, Moreno-Mayar JV, Bukhsianidze M, Rakownikow Jersie-Christensen R, et al. Early Pleistocene enamel proteome from Dmanisi resolves *Stephanorhinus* phylogeny. Nature. 2019; 574 doi: 10.1038/s41586-019-1555-y [PubMed: 31511700]
17. Welker F, Ramos-Madrigal J, Kuhlwilm M, Liao W, Gutenbrunner P, de Manuel M, Samodova D, Mackie M, Allentoft ME, Bacon A-M, Collins MJ, et al. Enamel proteome

shows that *Gigantopithecus* was an early diverging pongine. Nature. 2019; 576 doi: 10.1038/s41586-019-1728-8 [PubMed: 31723270]

18. Welker F, Ramos-Madrigal J, Gutenbrunner P, Mackie M, Tiwary S, Rakownikow Jersie-Christensen R, Chiva C, Dickinson MR, Kuhlwilm M, de Manuel M, Gelabert P, et al. The dental proteome of *Homo antecessor*. Nature. 2020; 580 doi: 10.1038/s41586-020-2153-8 [PubMed: 32269345]
19. 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.
20. Kuman K, Granger DE, Gibbon RJ, Pickering TR, Caruana MV, Bruxelles L, Clarke RJ, Heaton JL, Stratford D, Brain CK. A new absolute date from Swartkrans Cave for the oldest occurrences of *Paranthropus robustus* and Oldowan stone tools in South Africa. J Hum Evol. 2021; 156 103000 [PubMed: 34020297]
21. Martin JM, Leece AB, Baker SE, Herries AIR, Strait DS. A lineage perspective on hominin taxonomy and evolution. Evol Anthropol. 2024; 33 e22018 [PubMed: 38217397]
22. Batth TS, Francavilla C, Olsen JV. Off-line high-pH reversed-phase fractionation for in-depth phosphoproteomics. J Proteome Res. 2014; 13: 6176–6186. [PubMed: 25338131]
23. Rüther PL, Husic IM, Bangsgaard P, Gregersen KM, Pantmann P, Carvalho M, Godinho RM, Friedl L, Cascalheira J, Taurozzi AJ, Jørkov MLS, et al. SPIN enables high throughput species identification of archaeological bone by proteomics. Nat Commun. 2022; 13 2458 doi: 10.1038/s41467-022-30097-x [PubMed: 35513387]
24. Simmer J, Hu J. Expression, structure, and function of enamel proteinases. Connect Tissue Res. 2002; 43: 441–449. [PubMed: 12489196]
25. Fincham AG, Simmer JP. Amelogenin proteins of developing dental enamel. Ciba Found Symp. 1997; 205: 118–30. [PubMed: 9189621]
26. Simmer JP, Hu JC-C, Hu Y, Zhang S, Liang T, Wang S-K, Kim J-W, Yamakoshi Y, Chun YH, Bartlett JD, Smith CE. A genetic model for the secretory stage of dental enamel formation. J Struct Biol. 2021; 213 107805 doi: 10.1016/j.jsb.2021.107805 [PubMed: 34715329]
27. Parker GJ, Yip JM, Eerkens JW, Salemi M, Durbin-Johnson B, Kiesow C, Haas R, Buikstra JE, Klaus H, Regan LA, Rocke DM, et al. Sex estimation using sexually dimorphic amelogenin protein fragments in human enamel. J Archaeol Sci. 2019; 101: 169–180.
28. Cleland TP, McGuire SA, Beatrice JS, Moran KS, France CAM. SPEED-E: A modified version of the sample preparation by Easy extraction and Digestion(-free) protocol for enamel-based sex estimation in archaeological remains. J Archaeol Sci. 2024; 168 106006
29. Wood, B. Hominid Cranial Remains. Vol. 4. Oxford University Press; USA: 1991.
30. Dean C, Zanolli C, Le Cabec A, Tawane M, Garrevoet J, Mazurier A, Macchiarelli R. Growth and development of the third permanent molar in *Paranthropus robustus* from Swartkrans, South Africa. Sci Rep. 2020; 10 19053 doi: 10.1038/s41598-020-76032-2 [PubMed: 33149180]
31. Copeland SR, Sponheimer M, de Ruiter DJ, Lee-Thorp JA, Codron D, le Roux PJ, Grimes V, Richards MP. Strontium isotope evidence for landscape use by early hominins. Nature. 2011; 474: 76–78. [PubMed: 21637256]
32. Dembo M, Radovčic D, Garvin HM, Laird MF, Schroeder L, Scott JE, Brophy J, Ackermann RR, Musiba CM, de Ruiter DJ, Mooers AØ, et al. The evolutionary relationships and age of *Homo naledi*: An assessment using dated Bayesian phylogenetic methods. J Hum Evol. 2016; 97: 17–26. [PubMed: 27457542]
33. Mongle CS, Strait DS, Grine FE. An updated analysis of hominin phylogeny with an emphasis on re-evaluating the phylogenetic relationships of *Australopithecus sediba*. J Hum Evol. 2023; 175 103311 [PubMed: 36706599]
34. Pickering R, Herries AIR, Woodhead JD, Hellstrom JC, Green HE, Paul B, Ritzman T, Strait DS, Schoville BJ, Hancox PJ. U-Pb-dated flowstones restrict South African early hominin record to dry climate phases. Nature. 2019; 565: 226–229. [PubMed: 30464348]
35. Daniel YA, Turner C, Haynes RM, Hunt BJ, Dalton RN. Rapid and specific detection of clinically significant haemoglobinopathies using electrospray mass spectrometry-mass spectrometry. Br J Haematol. 2005; 130: 635–643. [PubMed: 16098080]

36. Mason KE, Anex D, Grey T, Hart B, Parker G. Protein-based forensic identification using genetically variant peptides in human bone. Forensic Sci Int. 2018; 288: 89–96. [PubMed: 29738994]
37. Wu P-W, Mason KE, Durbin-Johnson BP, Salemi M, Phinney BS, Rocke DM, Parker GJ, Rice RH. Proteomic analysis of hair shafts from monozygotic twins: Expression profiles and genetically variant peptides. Proteomics. 2017; 17 doi: 10.1002/pmic.201600462 [PubMed: 28544375]
38. Pang Q, Lin Q, Wang D, Sun Z, Wang J. Molecular characterization of the Yp11.2 region deletion in the Chinese Han population. Int J Legal Med. 2021; 135: 1351–1358. DOI: 10.1007/s00414-021-02596-x [PubMed: 33903958]
39. 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. [PubMed: 16684695]
40. Skov L, Peyrégne S, Popli D, Iasi LNM, Devièse T, Slon V, Zavala EI, Hajdinjak M, Sümer P, Grote S, Bossoms Mesa, et al. Genetic insights into the social organization of Neanderthals. Nature. 2022; 610: 519–525. DOI: 10.1038/s41586-022-05283-y [PubMed: 36261548]
41. Martin JM, Leece AB, Neubauer S, Baker SE, Mongle CS, Boschian G, Schwartz GT, Smith AL, Ledogar JA, Strait DS, Herries AIR. Drimolen cranium DNH 155 documents microevolution in an early hominin species. Nat Ecol Evol. 2021; 5: 38–45. [PubMed: 33168991]
42. 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. [PubMed: 37812213]
43. Rak Y, Kimbel WH, Moggi-Cecchi J, Lockwood CA, Menter C. The DNH 7 skull of *Australopithecus robustus* from Drimolen (Main Quarry), South Africa. J Hum Evol. 2021; 151 102913 [PubMed: 33388495]
44. 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 13956 doi: 10.1038/s41598-022-18223-7 [PubMed: 35977986]
45. Zanolli C, Hublin J-J, Kullmer O, Schrenk F, Kgasi L, Tawane M, Xing S. Taxonomic revision of the SK 15 mandible based on bone and tooth structural organization. J Hum Evol. 2025; 200 103634 [PubMed: 39752989]
46. Madupe PP, Munir F, Dickinson M, Taurozzi AJ, Mackie M, Tawane M, Mollereau C, Hlazo N, Penkman K, Schroeder L, Zanolli C, et al. Results from an *Australopithecus africanus* dental enamel fragment confirm the potential of palaeoproteomics for South African Plio-Pleistocene fossil sites. South African Journal of Science. 2025; 121 18571
47. Haile-Selassie Y, Gibert L, Melillo SM, Ryan TM, Alene M, Deino A, Levin NE, Scott G, Saylor BZ. New species from Ethiopia further expands Middle Pliocene hominin diversity. Nature. 2015; 521: 483–488. [PubMed: 26017448]
48. Spoor F, Leakey MG, O'Higgins P. Middle Pliocene hominin diversity: *Australopithecus deyiremeda* and *Kenyanthropus platyops*. Philos Trans R Soc Lond B Biol Sci. 2016; 371 20150231 doi: 10.1098/rstb.2015.0231 [PubMed: 27298462]
49. White T. Paleoanthropology: five's a crowd in our family tree. Curr Biol. 2013; 23: R112–5. [PubMed: 23391385]

Europe PMC Funders Author Manuscripts

One-sentence abstract

Palaeoproteomics enabled the biological sex attribution of, and the detection of an amino acid polymorphism in four dental specimens from South Africa attributed to *Paranthropus robustus*.

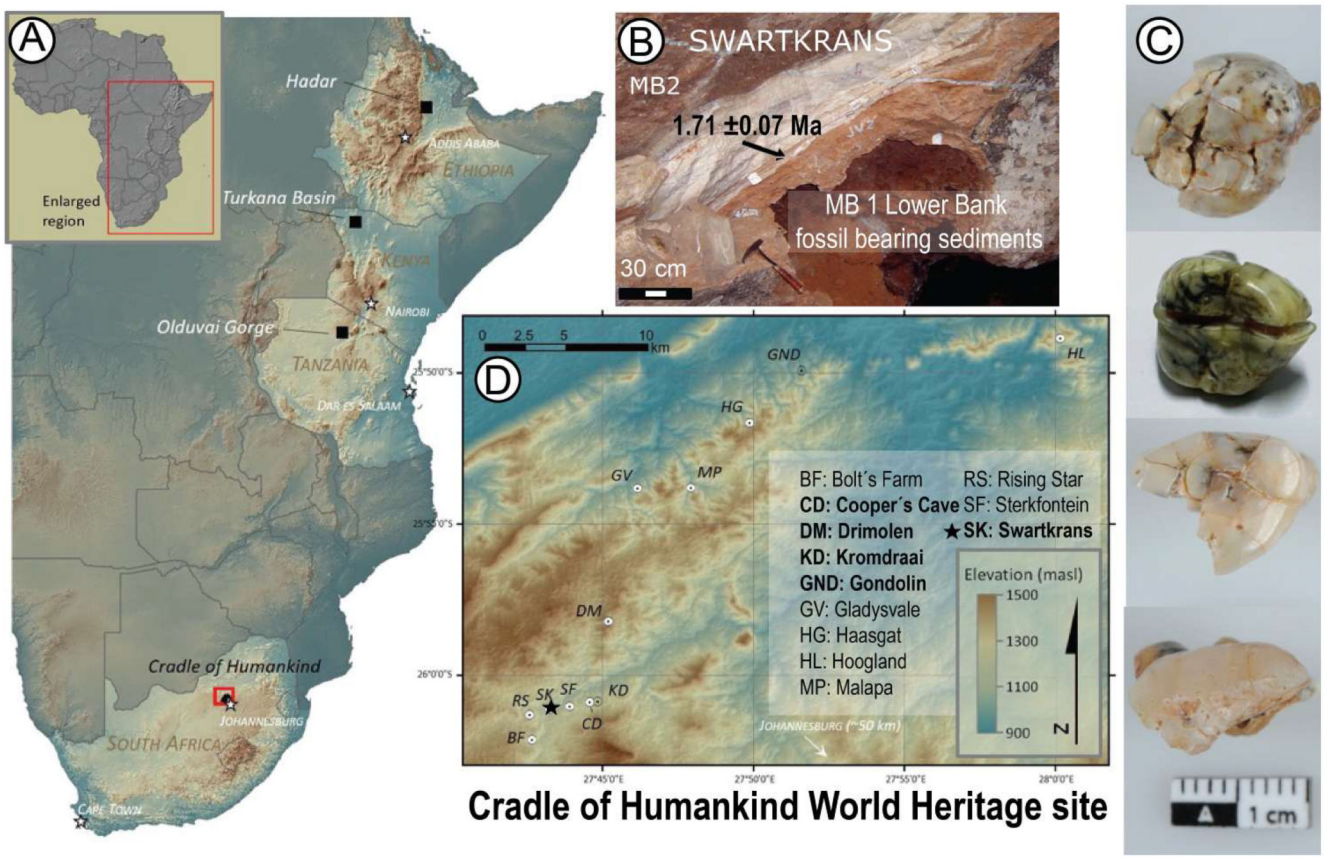

Fig. 1. Location and cave structure of the site of Swartkrans, South Africa.

A- Topographical map of the African continent (inset) showing the major early hominin fossil bearing regions. B- Photograph of the *Paranthropus* bearing paleocave Swartkrans, showing the Member 1 fossil bearing sediments and dated flowstone (19). C- The occlusal view of *Paranthropus* teeth analyzed, from top to bottom: SK 830, a left P_4; SK 835, a left M^3; SK 850, a right P_3; SK 14132, a fragment of an isolated postcanine tooth (Supplementary Materials). D- Enlarged view of the Cradle of Humankind World Heritage Site in South Africa (shown in A), with *Paranthropus* fossil locality names in bold. Swartkrans is marked with a star.

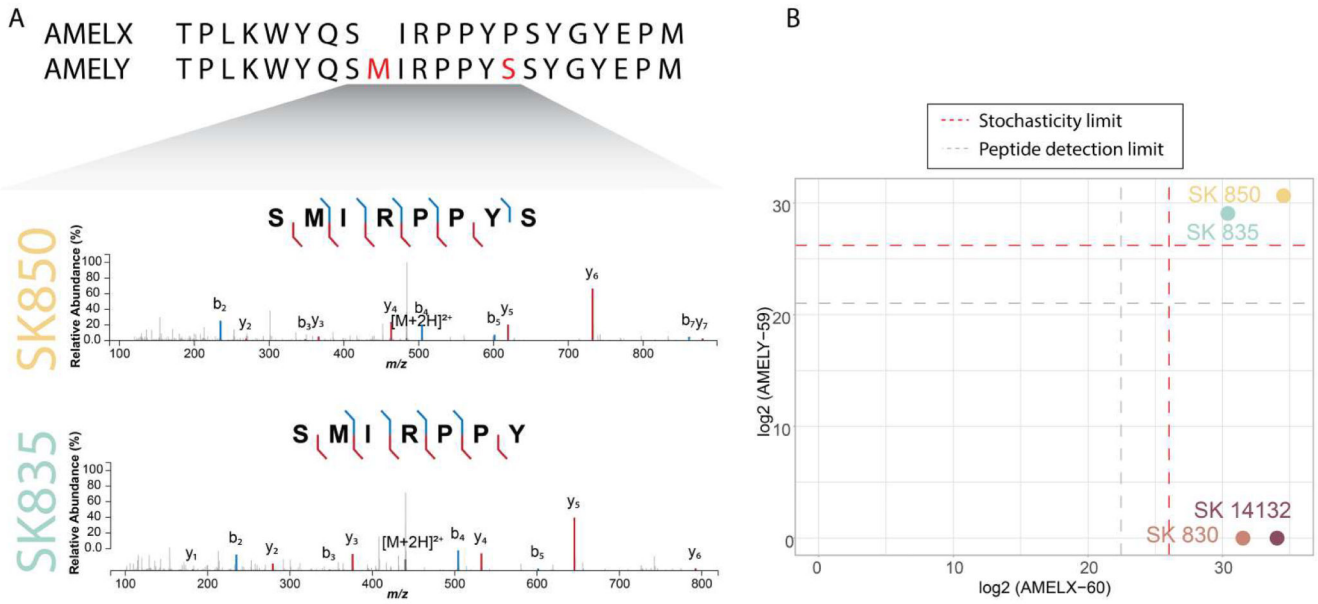

Fig. 2. Sex identification of the four *Paranthropus* specimens.

A- MS2 spectra of AMELY-specific peptides detected in the two male individuals SK 850 and SK 835. In both spectra, the detection of the methionine in position 59, characteristic of males, is well supported by the fragment ion series. B- Site intensities of AMELY-59 as a function of the site intensities of AMELX-60. Site intensities are calculated as the sum of all precursor intensities covering the given amino acid position. The red dashed lines represent the stochasticity limits inherent to the MS acquisition strategy and setup (Supplementary Materials). Below the lines, we would expect the acquisition to be stochastic. The gray dashed lines represent the peptide detection limits. They have been assessed as the minimum intensity of a precursor covering the given amino acid position in a male individual for the studied data set. Above the line, precursors can technically and consistently be measured in the MS. Each point represents a different *Paranthropus* specimen.

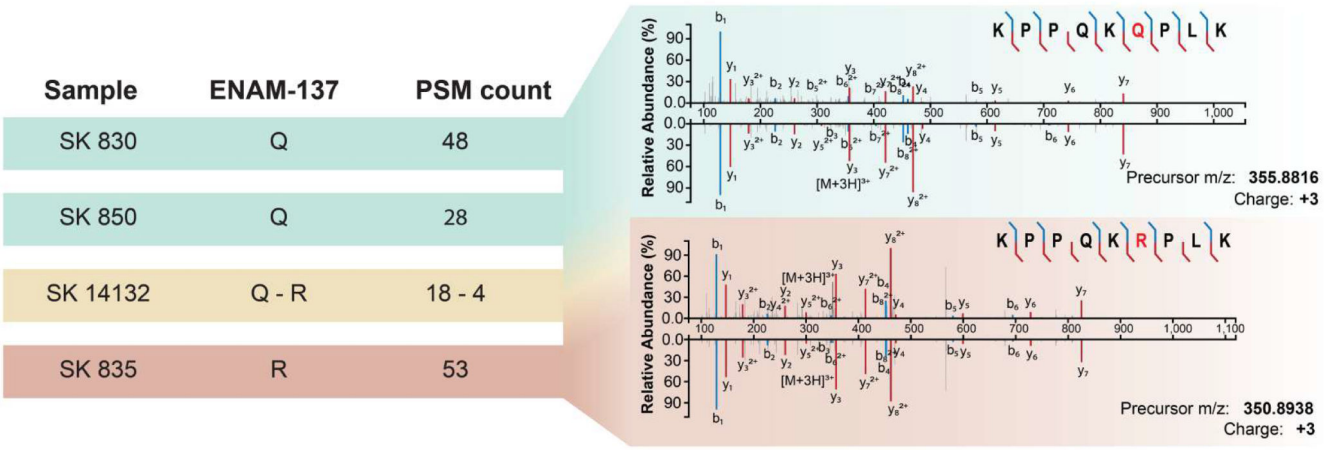

Fig. 3. Sequence variation within the *Paranthropus* individuals.

Amino acid sequence variation at ENAM-137 in the four *Paranthropus* samples and number of peptide spectrum matches (PSMs) supporting their detection. The detection of glutamine (Q) and arginine (R) at ENAM-137 was validated with the analysis of synthetic peptides. On the right, the mirror plots represent the MS2 spectra covering glutamine, in peptide KPPQKQPLK, and arginine, in peptide KPPQKRPLK, in the *Paranthropus* samples (top) compared to the MS2 spectra of the corresponding synthetic peptides (bottom).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

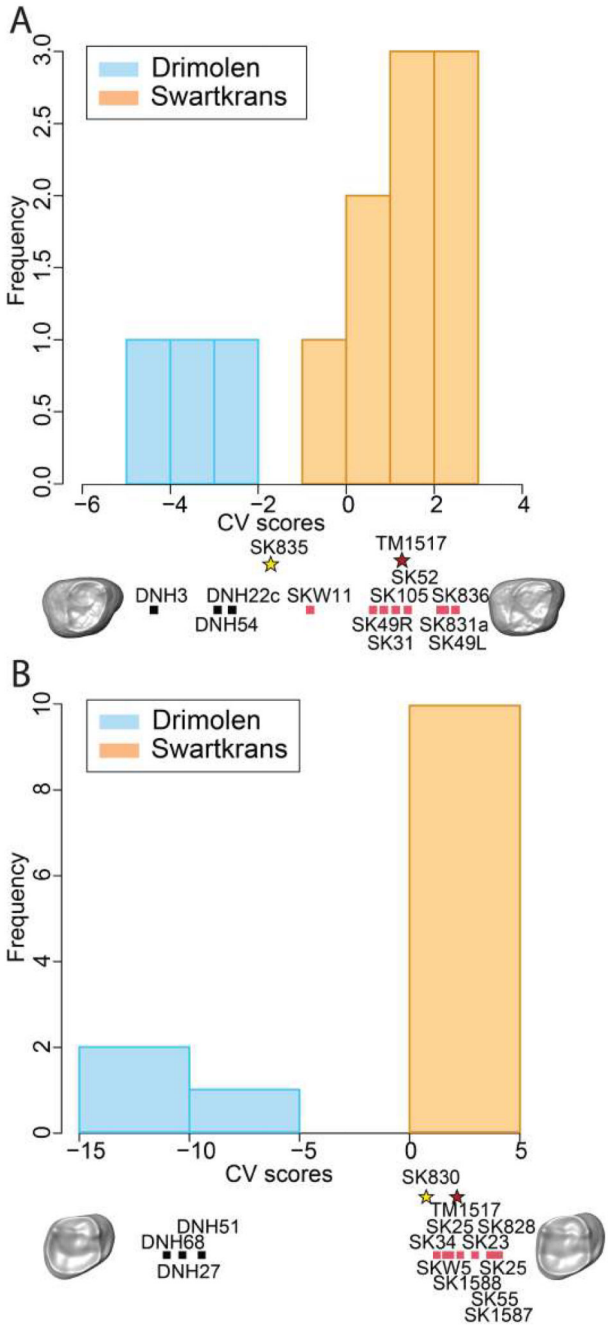

Fig. 4. Geometric morphometric analysis of the enamel-dentine junction.

A- Frequency and distribution of canonical variate scores of the Swartkrans (and Kromdraai), and Drimolen M^3s enamel-dentine junction shape analysis. This analysis shows SK 835 as statistically closer to specimens from Drimolen compared to other specimens from Swartkrans and the holotype of *P. robustus* TM 1517 from Kromdraai. B- Same as in A but using P_4s, showing that SK 830 is closer to Swartkrans and Kromdraai specimens compared to other specimens from Drimolen.