



# Sex determination of human remains from peptides in tooth enamel

Nicolas Andre Stewart<sup>a,1</sup>, Raquel Fernanda Gerlach<sup>b</sup>, Rebecca L. Gowland<sup>c</sup>, Kurt J. Gron<sup>c</sup>, and Janet Montgomery<sup>c</sup>

<sup>a</sup>School of Pharmacy and Biomolecular Sciences, University of Brighton, Brighton BN2 4AT, United Kingdom; <sup>b</sup>Department of Morphology, Physiology, and Basic Pathology, School of Dentistry of Ribeirão Preto, University of São Paulo, FORP/USP, 807 São Paulo, Brazil; and <sup>c</sup>Department of Archaeology, Durham University, Durham DH1 3LE, United Kingdom

Edited by Christopher Kuzawa, Northwestern University, Evanston, IL, and accepted by Editorial Board Member C. O. Lovejoy November 13, 2017 (received for review August 23, 2017)

The assignment of biological sex to archaeological human skeletons is a fundamental requirement for the reconstruction of the human past. It is conventionally and routinely performed on adults using metric analysis and morphological traits arising from postpubertal sexual dimorphism. A maximum accuracy of ~95% is possible if both the cranium and os coxae are present and intact, but this is seldom achievable for all skeletons. Furthermore, for infants and juveniles, there are no reliable morphological methods for sex determination without resorting to DNA analysis, which requires good DNA survival and is time-consuming. Consequently, sex determination of juvenile remains is rarely undertaken, and a dependable and expedient method that can correctly assign biological sex to human remains of any age is highly desirable. Here we present a method for sex determination of human remains by means of a minimally destructive surface acid etching of tooth enamel and subsequent identification of sex chromosome-linked isoforms of amelogenin, an enamel-forming protein, by nanoflow liquid chromatography mass spectrometry. Tooth enamel is the hardest tissue in the human body and survives burial exceptionally well, even when the rest of the skeleton or DNA in the organic fraction has decayed. Our method can reliably determine the biological sex of humans of any age using a body tissue that is difficult to cross-contaminate and is most likely to survive. The application of this method will make sex determination of adults and, for the first time, juveniles a reliable and routine activity in future bioarcheological and medico-legal science contexts.

sex determination | tooth enamel | amelogenin | human remains | mass spectrometry

Sex is a fundamental primary characteristic for the analysis of human skeletal remains in archaeological and medico-legal contexts. Techniques for determining other key identifying features, such as age at death and stature, are also sex-dependent (1). Accurate profiles of sex are crucial for reconstructing past societies in terms of demography, identity, and epidemiology, and are also essential in medico-legal contexts for identifying individuals (e.g., in mass disasters). In adults, sex can be determined with relative accuracy, usually estimated at 80–95% depending on such factors as skeletal preservation and degree of sexual dimorphism within the sample (2). One of the key limiting factors of osteological analyses to date has been an inability to reliably determine the sex of individuals from skeletal features before age ~18 y (3). Numerous methods have tried to do so, usually applying studies of the morphological and metrical characteristics of the infant and juvenile mandible, dentition, and ilium (e.g., refs. 4 and 5); however, no method has proven sufficiently reliable when tested on documented skeletal samples (6).

Unfortunately, ancient DNA analysis is not the answer to this problem due to issues of preservation, contamination, and expense. For example, several DNA studies attempted to determine the sex of infants from Romano-British sites with the aim of assessing whether preferential female infanticide was practiced (7, 8), but in all of those studies, viable results were obtained from only a small proportion of the overall number sampled. Owing to the destructive, costly, and inconclusive nature

of DNA sex determination, this method has rarely been attempted on any scale on human remains from archaeological sites.

Herein we present a method for secure biological sex determination of human remains by means of a minimally destructive surface acid etching of tooth enamel and subsequent identification of sex chromosome-linked isoforms of amelogenin, an enamel-forming protein, by nanoflow liquid chromatography mass spectrometry (nanoLC-MS). Tooth enamel is the hardest tissue in the human body and survives burial exceptionally well, even when the rest of the skeleton or DNA in the organic fraction has decayed. Therefore, this method holds promise for reliably determining the biological sex of humans of any age using a body tissue that is most likely to survive intact. It is minimally destructive, inexpensive, and reliable.

## Results and Discussion

The acid-etching procedure to extract peptides from tooth enamel results in a complex base peak chromatogram when analyzed by nanoLC-MS/MS (Fig. 1). It has previously been shown that peptides can be identified using various acid-etching methods, and that these peptides originate from the major tooth enamel proteins: amelogenin, ameloblastin, and enamelin (9–12). During enamel maturation, the majority of these proteins

### Significance

The ability to assign biological sex to human skeletal remains is a fundamental requirement in archaeology, paleoanthropology, and medico-legal sciences. While DNA sequencing can be used, it is expensive, time-consuming, and often fails due to the poor quality of the remaining DNA. An easier, more reliable, and consistently applicable method is needed. We present a method for sex determination of human remains using peptides retrieved from tooth enamel. Amelogenin is an enamel-forming protein encoded for by both chromosomes X and Y, with slight differences in their amino acid sequences. Peptides with these differences were identified by nanoflow liquid chromatography mass spectrometry and found to correctly assign sex to archaeological human remains of various chronological ages, from hundreds to thousands of years old.

Author contributions: J.M. designed research; N.A.S., R.F.G., R.L.G., K.J.G., and J.M. performed research; N.A.S., R.F.G., and K.J.G. contributed new reagents/analytic tools; N.A.S. analyzed data; and N.A.S., R.F.G., R.L.G., K.J.G., and J.M. wrote the paper.

The authors declare no conflict of interest.

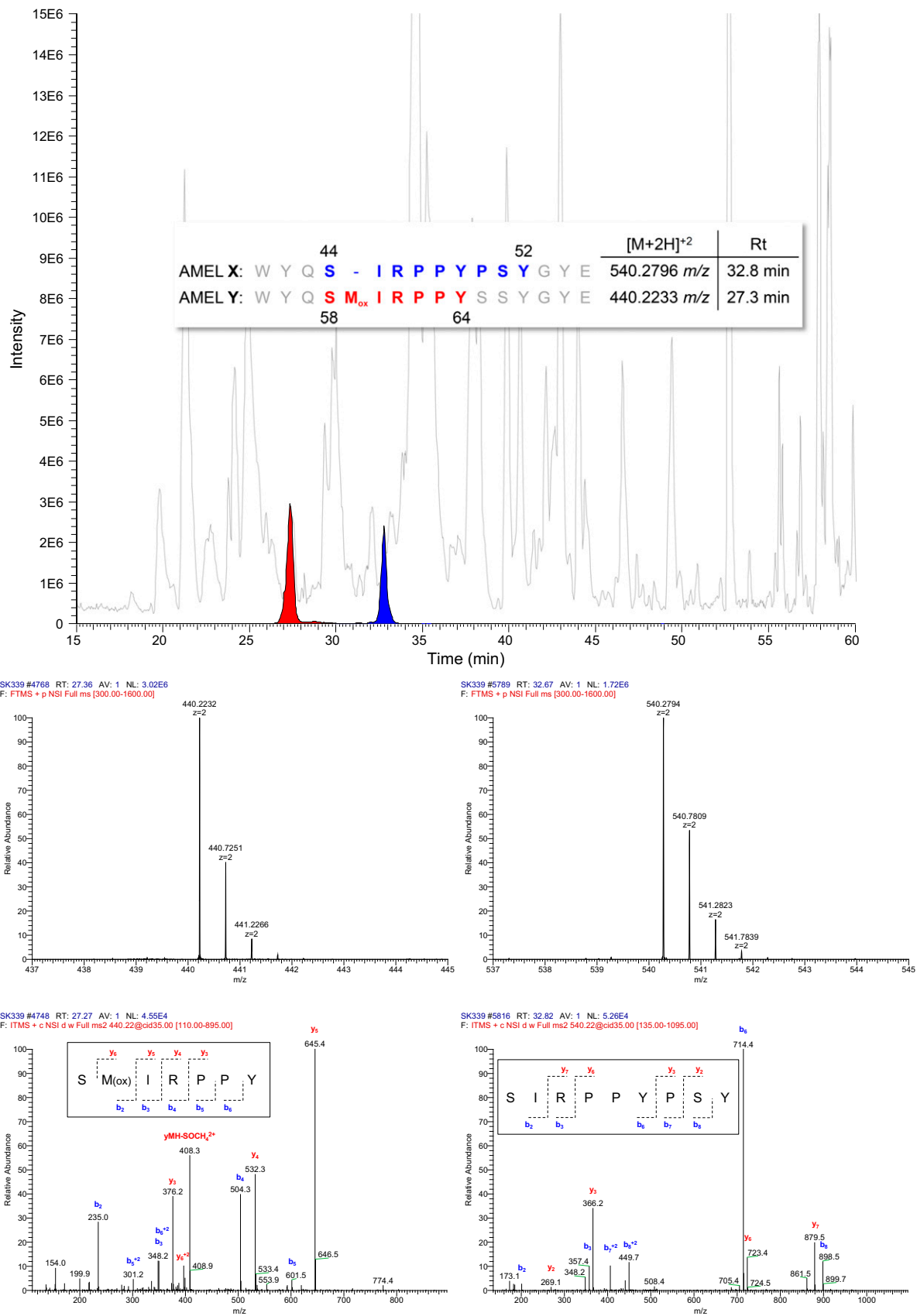
This article is a PNAS Direct Submission. C.K. is a guest editor invited by the Editorial Board.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

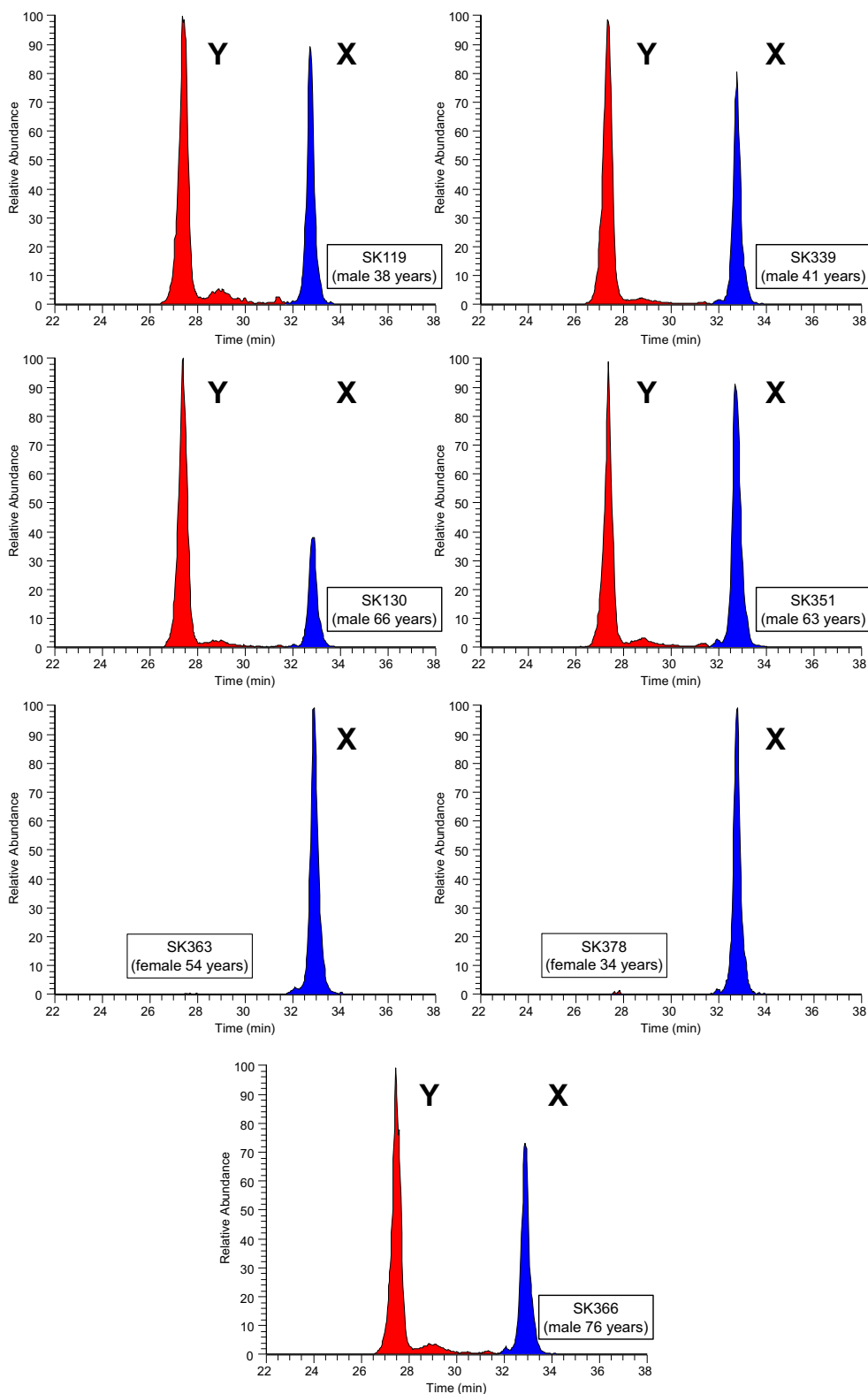
Data deposition: The mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD007856.

<sup>1</sup>To whom correspondence should be addressed. Email: n.stewart@brighton.ac.uk.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1714926115/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1714926115/-DCSupplemental).



**Fig. 1.** A representative base peak chromatogram (300–1,600  $m/z$ ) produced from Fewston sample SK339. (Inset) Amino acid sequences of the two dimorphic peptides of amelogenin: AMELY-(58-64) peptide and AMELX-(44-52) peptide. The reconstructed ion chromatograms (to 4 ppm) for each of these are shown in red and blue, respectively, with full-scan MS and corresponding MS/MS below.

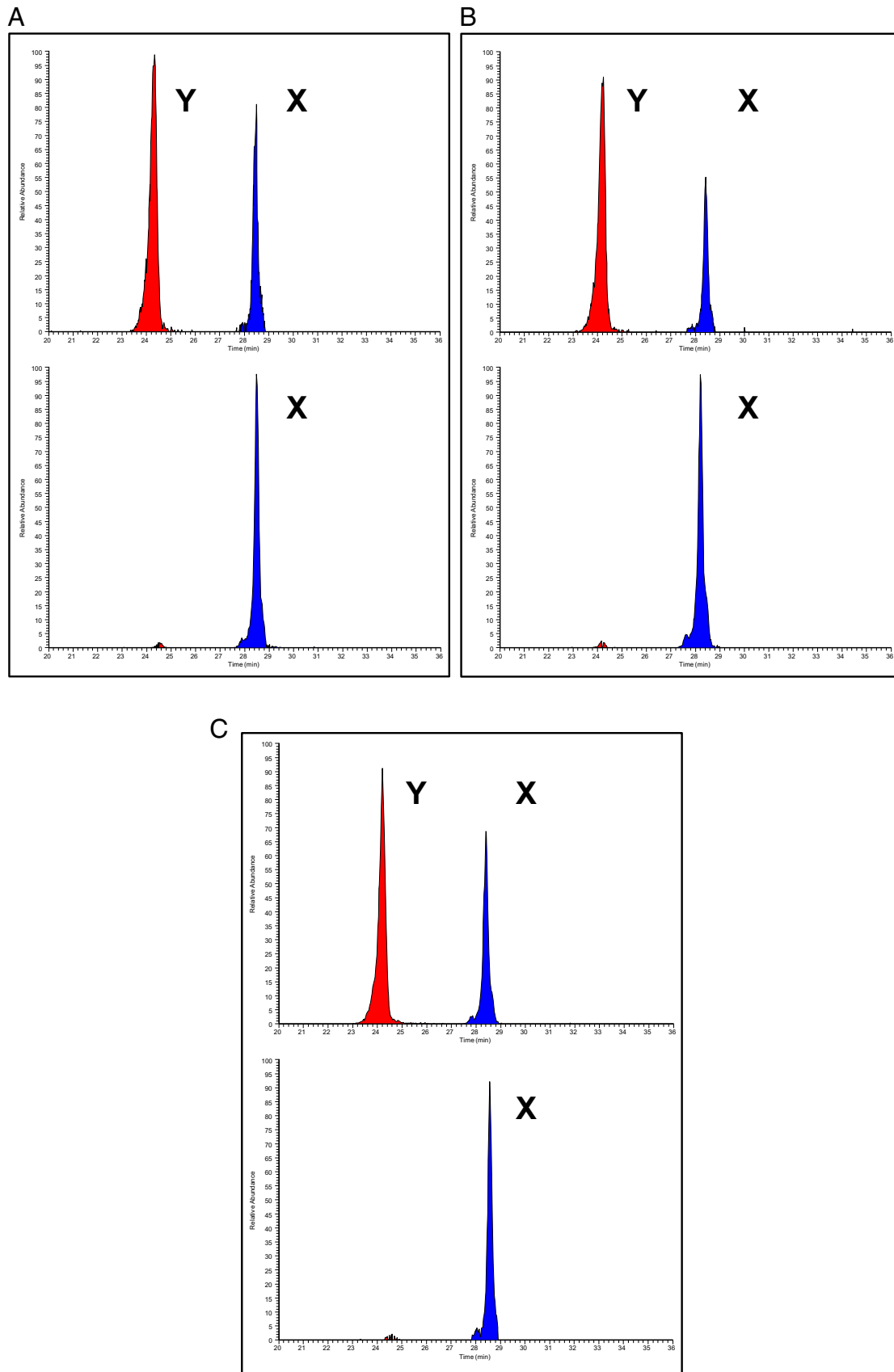


**Fig. 2.** Reconstructed ion chromatograms for the AMELY-(58-64) peptide (440.2233 *m/z*) and AMELX-(44-52) peptide (540.2796 *m/z*) (4 ppm mass tolerance) for the seven 19th century Fewston samples. Peaks corresponding to these are shown in red and blue, respectively. Known sex and age at death are indicated.

are processed by proteases, resulting in peptides of varying lengths, and remain in the mature enamel (9, 11, 12).

For amelogenin, peptides from the central portion of the protein are absent, but peptides from the N and C termini remain and

have been identified (PRIDE identifier PXD007856). The dimorphic differences between amelogenin X and Y are found in these regions, and one such peptide identified is the AMELY-(58-64) peptide, which possesses an additional methionine compared



**Fig. 3.** Reconstructed ion chromatograms for the AMELY-(58-64) peptide (440.2233  $m/z$ ) and AMELX-(44-52) peptide (540.2796  $m/z$ ) (4 ppm mass tolerance) for three male/female pairs of archaeological samples from St. Guthlac's Priory, 12th–16th century AD (A); Whitwell, ca. 5,700 BP (B); and Seaham, 7th–9th century AD (C) (previously published). Osteological age and sex determinations are indicated.

**Table 1. Details of samples used in this study**

Site location	Period	Type of burial	Skeleton no.	Age and sex*	Methods used to determine sex	Reference
Whitwell, Derbyshire, UK	Neolithic ca. 5,700 BP	Fragmentary, disarticulated cranium (SK485) and articulating mandible (SK219)	SK219	Adult female	Morphological traits of the mandible of SK219 and the articulating maxilla of SK485 (3, 14); marked sexual dimorphism	(15)
Whitwell, Derbyshire, UK	Neolithic ca. 5,700 BP	Fragmentary, disarticulated mandible	SK534	Adult male	Morphological traits of the mandible only (3, 14); marked sexual dimorphism	(15)
Seaham, County Durham, UK	7th–9th century AD	Inhumation cemetery	FF5 SK15	Female 26–45 y	Morphological traits of the pelvis and skull (3, 14, 16)	(17)
Seaham, County Durham, UK	7th–9th century AD	Inhumation cemetery	FF5 SK3	Male 36+ y	Morphological traits of the pelvis and skull (3, 14, 16)	(17)
St Guthlac's Priory, Hereford, UK	12th–16th century AD	Inhumation cemetery	SK 9503	Female old adult	Morphological traits of the pelvis and skull (3, 14, 16)	(18)
St Guthlac's Priory, Hereford, UK	12th–16th century AD	Inhumation cemetery	SK 9515	Male old adult	Morphological traits of the pelvis and skull (3, 14, 16); significant grave goods: chalice	(18)
Fewston, North Yorkshire, UK	19th century AD	Inhumation cemetery	SK363	Female 54 y	Documented age and sex: coffin plate; morphological traits of the pelvis and skull (2, 3, 16)	(13)
Fewston, North Yorkshire, UK	19th century AD	Inhumation cemetery	SK378	Female 34 y	Documented age and sex: coffin plate; morphological traits of the pelvis and skull (2, 3, 16)	(13)
Fewston, North Yorkshire, UK	19th century AD	Inhumation cemetery	SK366	Male 76 y	Documented age and sex: coffin plate; morphological traits of the pelvis and skull (2, 3, 16)	(13)
Fewston, North Yorkshire, UK	19th century AD	Inhumation cemetery	SK119	Male 38 y	Documented age and sex: coffin plate; morphological traits of the pelvis and skull (2, 3, 16)	(13)
Fewston, North Yorkshire, UK	19th century AD	Inhumation cemetery	SK339	Male 41 y	Documented age and sex: coffin plate; morphological traits of the pelvis and skull (2, 3, 16)	(13)
Fewston, North Yorkshire, UK	19th century AD	Inhumation cemetery	SK130	Male 66 y	Documented age and sex: coffin plate; morphological traits of the pelvis and cranium (2, 3, 16)	(13)
Fewston, North Yorkshire, UK	19th century AD	Inhumation cemetery	SK351	Male 63 y	Documented age and sex: grave stone; morphological traits of the pelvis and skull (2, 3, 16)	(13)

\*As previously determined by osteological, epigraphic, and grave goods.

with the aligned sequence of AMELX (Fig. 1, *Inset*). In one of the samples (SK130), the AMELX peptide seemed to be relatively lower in abundance compared with the AMELY peptide. This most likely reflects a higher relative amount of the AMELY peptide, as this peptide contains a methionine, and it may be oxidized in greater amount in this sample. Therefore, the oxidized version of the AMELY peptide was chosen for Y sex confirmation, as it is expected to predominate in old samples, as opposed to the unoxidized peptide, which is anticipated to be of low abundance or absent.

A peptide from amelogenin X, AMELX-(44-52) peptide, with similar ion intensity to the AMELY-(58-64) peptide, was used to clearly assign sex based on the presence or absence of AMELY-

(58-64) from the seven Fewston samples (Fig. 2) and the three sets of sex-paired samples (Fig. 3).

The sexual dimorphic protein isoforms of amelogenin, along with the fact that peptides from the dimorphic regions of these isoforms serendipitously remain in tooth enamel, provide an excellent means for unequivocally establishing the sex of human remains by nanoLC-MS. The unique nature of tooth enamel as the body's densest and hardest tissue makes it a perfect repository for these peptides. The two dimorphic peptides chosen, AMELX-(44-52) peptide and AMELY-(58-64) peptide, clearly differentiate between male and female, with both the AMELY and AMELX peptides found in male samples and only the AMELX peptide found in female samples. In all cases, our

results agree with the assignment of sex by either coffin plates or standard osteological methods.

It should be noted that this method identifies the presence of peptides originating from sex chromosome-linked isoforms of amelogenin and currently cannot identify polymorphisms of multiple copies of these chromosomes (e.g., aneuploid 47, XXY or 47, XYY). Quantitation of these peptides may allow for this in these rare cases and warrants further investigation.

The ability to determine the sex of infant and juvenile remains completely revolutionizes studies of growth, child care, epidemiology, and demography in the past. For the first time, it will allow osteologists to examine sex-specific cultural treatment and differentiate between the health of boys and girls, as well as sex-specific growth trajectories and past developmental milestones, such as age of puberty and subsequent repercussions for fertility. Sites with poor preservation are common in archaeological contexts, and at such sites teeth generally survive better than bone, and thus sex can be established for adults as well as juvenile skeletons in the absence of key skeletal identifiers. In addition, the dimorphic peptide sequence is identical in apes (Fig. S1) and so should be present in all hominins. Finally, this technique will also have a transformative effect on human identification in medico-legal contexts, such as mass disasters and war graves, allowing sex to be established both reliably and cost-effectively.

## Materials and Methods

Fewston is a small village located in the Washburn Valley, near Harrogate in North Yorkshire, United Kingdom. The skeletal assemblage was excavated from the parish churchyard in advance of building work in 2009–2010 and was reburied in September 2016. Twenty-one of the excavated individuals were confidently identified based on coffin plates and grave monuments. All of these identified individuals date to the late 19th century. Etches were performed on teeth from seven adult individuals of known identity and sex. Sex was assigned by coffin plates and confirmed by osteological analysis (13) (Table 1). All samples were anonymized after removal of peptides and analyzed blind.

In addition, male and female pairs from three archaeological sites ranging in date from the early Neolithic (ca. 5,700 BP) to the Medieval period were tested to determine the survival and recovery of sufficient proteins over archaeological timescales and a variety of burial contexts. For each pair, sex had been previously assigned using standard osteological methods, as described in Table 1. These pairs were not analyzed blind.

Samples from Whitwell, Seaham, and St. Guthlac's were prepared and analyzed as described previously (9). All reagents used were of analytical grade, and solvents used for nanoLC-MS/MS were MS grade. The Fewston samples were prepared using a modified streamlined version of the protocol as follows. The tooth surface was abraded using a dental burr to remove obvious surface contaminants. The enamel was washed with 3% H<sub>2</sub>O<sub>2</sub> for 30 s and then rinsed with ultrapure water (Elga Purelab Ultra, 18.2 MΩ-cm). Approximately 60 μL of

5% (vol/vol) HCl was placed in the cap of a 0.2-mL Eppendorf tube, leaving a convex meniscus protruding above the lip. An initial etch was performed by lowering the tooth onto the HCl and maintaining contact for 2 min. This first etch was discarded. A second 2-min etch was then performed and retained as the etch solution. In an adjustable 0.5- to 10-μL pipette set to 10 μL, a C18 resin-loaded ZipTip (ZTC18S096; EMD Millipore) was conditioned three times with 100% acetonitrile, then three times with 0.1% (vol/vol) formic acid, with each draw discarded. The etch solution was bound to the ZipTip by pipetting the solution up and down 10 times, with the last draw discarded. The ZipTip was washed six times with 0.1% (vol/vol) formic acid, with each wash discarded. The adjustable pipette was set to 4 μL, and the resin-bound peptides were eluted into a 4-μL 60% acetonitrile/0.1% formic acid elution buffer and lyophilized.

Samples were dissolved in 12 μL of 0.1% trifluoroacetic acid (TFA) in water, centrifuged on a benchtop centrifuge for 5 min to remove any particulate matter, and transferred (10 μL) to glass autosampler vials. Then 5 μL of sample was injected for analysis by reversed-phase nanoLC-MS with a liquid chromatograph (nanoRS U3000; Thermo Fisher Scientific) coupled to a hybrid linear ion trap mass spectrometer (Orbitrap XL; Thermo Fisher Scientific). Peptides were first loaded onto a C18 trapping cartridge (Pepmap100 C18; Thermo Fisher Scientific; 0.3 × 5 mm i.d.; 5 μm particle size) for 10 min at a flow rate of 5 μL/min with 0.1% TFA. Separation was achieved at a flow rate of 300 nL/min on an analytical column (PepMap100 C18; 25 cm × 75 μm; 5 μm particle size) with a gradient starting at 1%, increased to 13.3% solvent B over 20 min, then to 25.6% over 15 min, 45% over 10 min, and 99% over 15 min, held constant at 99% for 5 min, returned to 1%, and equilibrated for 20 min. Nano-electrospray ionization was performed with a 10-μm uncoated silica tip emitter (FS360-20-10-N-20; New Objective). The MS was operated in data-dependent MS/MS mode in which each full MS scan was collected in the Orbitrap (300–1,600 m/z; R = 60,000 at 400 m/z), followed by up to nine MS/MS scan events performed in the linear ion trap, with the most abundant peptide molecular ions selected for collision-induced dissociation, using a normalized collision energy of 35%. Total MS acquisition time was 64 min. Data were searched against the human proteome (UniprotKB, 10/15) with MaxQuant v 1.5.1.2 using default search settings with methionine oxidation as a variable modification, unspecific digestion mode, and a minimum peptide length of 6. The MS data have been deposited to the ProteomeXchange Consortium via the PRIDE (19) partner repository with the dataset identifier PXD007856.

**ACKNOWLEDGMENTS.** We thank Anwen Caffell, Malin Holst (York Osteoarchaeology), John Buglass, The Heritage Lottery Fund, and the Washburn Heritage Centre for access to and information on the Fewston skeletal assemblage; Andrew Chamberlain and Charlotte Roberts for access to the Whitwell and St. Guthlac's skeletal collections and their unpublished sex determinations of the individuals included herein; Niall Hammond (Archaeo-Environment) and Northern Archaeological Associates for access to the Seaham skeletal collection excavated in 1999; Tessa Loeffelmann for sex determination of these individuals in the course of her 2014 MSc dissertation "Of Doomed and Transient Men: The Anglo-Saxon Skeletons of Seaham in Their Context" at Durham University; and Dr. Steve M. Sweet (University of Sussex) for access to the nanoLC-Orbitrap MS instrument. K.J.G. thanks the Leverhulme Trust for funding. R.F.G. acknowledges support from the State of Sao Paulo Research Foundation and the Brazilian Research Council for studies on enamel peptides and trace elements.

- Gowland R, Thompson TJU (2013) *Human Identity and Identification* (Cambridge Univ Press, Cambridge, UK).
- Cox M, Mays S (2000) *Human Osteology in Archaeology and Forensic Science* (Greenwich Medical Media, London), pp 117–120.
- Buikstra JE, Ubelaker DH (1994) *Standards for Data Collection from Human Skeletal Remains: Proceedings of a Seminar at the Field Museum of Natural History* (Arkansas Archeological Survey, Fayetteville, AR).
- Schutzkowski H (1993) Sex determination of infant and juvenile skeletons, I: Morphognostic features. *Am J Phys Anthropol* 90:199–205.
- Vlak D, Roksandic M, Schillaci MA (2008) Greater sciatic notch as a sex indicator in juveniles. *Am J Phys Anthropol* 137:309–315.
- Lewis M (2007) *The Bioarchaeology of Children: Perspectives from Biological and Forensic Anthropology* (Cambridge Univ Press, Cambridge, UK).
- Abu-Mandil Hassan N, Brown KA, Evers J, Brown TA, Mays S (2014) Ancient DNA study of the remains of putative infanticide victims from the Yewden Roman villa site at Hambleden, England. *J Archaeol Sci* 43:192–197.
- Mays S, Faerman M (2001) Sex identification in some putative infanticide victims from Roman Britain using ancient DNA. *J Archaeol Sci* 28:555–559.
- Stewart NA, et al. (2016) The identification of peptides by nanoLC-MS/MS from human surface tooth enamel following a simple acid etch extraction. *RSC Adv* 6: 61673–61679.
- Castiblanco GA, et al. (2015) Identification of proteins from human permanent erupted enamel. *Eur J Oral Sci* 123:390–395.
- Porto IM, Laure HJ, de Sousa FB, Rosa JC, Gerlach RF (2011) New techniques for the recovery of small amounts of mature enamel proteins. *J Archaeol Sci* 38:3596–3604.
- Porto IM, et al. (2011) Recovery and identification of mature enamel proteins in ancient teeth. *Eur J Oral Sci* 119:83–87.
- Caffell A, Holst M (2010) *Osteological analysis, the Church of Saint Michael and Saint Lawrence, Fewston, UK* (York Osteoarchaeology Ltd., York, UK).
- Bass WM (1987) *Human Osteology: A Laboratory and Field Manual* (Missouri Archaeological Society, Columbia, MO), 3rd Ed.
- Chamberlain AT, Witkin AV (2011) A Neolithic cairn at Whitwell, Derbyshire. *Derbyshire Archaeol J* 131:1–131.
- Brickley M, McKinley JI (2004) Guidelines to the standards for recording human remains. Institute of Field Archaeologists paper no. 7. Available at www.archaeologists.net/sites/default/files/ifa\_paper\_7.pdf. Accessed November 20, 2017.
- Loeffelmann T (2014) Of doomed and transient men: The Anglo-Saxon skeletons of Seaham in their context. MSc dissertation (Durham University, Durham, UK).
- Roberts C (2005) The human skeletal report. Appendix to Crooks K. H. Excavations at Hereford County Hospital 1998-2003 (Archaeological Investigations Ltd., Hereford, UK), Report SMR 31923.
- Vizcaino JA, et al. (2016) 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* 44:11033.