

This is a repository copy of *Synthesis of a Vocal Sound from the 3,000 year old Mummy, Nesyamun 'True of Voice'*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/159500/>

Version: Published Version

Article:

Howard, David, Schofield, John orcid.org/0000-0001-6903-7395, Fletcher, Joann et al. (3 more authors) (2020) *Synthesis of a Vocal Sound from the 3,000 year old Mummy, Nesyamun 'True of Voice'*. Scientific Reports. 45000. ISSN 2045-2322

<https://doi.org/10.1038/s41598-019-56316-y>

Reuse


This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:
<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

OPEN

Synthesis of a Vocal Sound from the 3,000 year old Mummy, Nesyamun 'True of Voice'

D. M. Howard^{1*}, J. Schofield^{2*}, J. Fletcher², K. Baxter³, G. R. Iball⁴ & S. A. Buckley^{2,5} 

The sound of a 3,000 year old mummified individual has been accurately reproduced as a vowel-like sound based on measurements of the precise dimensions of his extant vocal tract following Computed Tomography (CT) scanning, enabling the creation of a 3-D printed vocal tract. By using the Vocal Tract Organ, which provides a user-controllable artificial larynx sound source, a vowel sound is synthesised which compares favourably with vowels of modern individuals.

The sound of a vocal tract from the past has been synthesised to be heard again in the present, allowing people to engage with the past in completely new and innovative ways. The precise dimensions of an individual's vocal tract produce a sound unique to them¹. If the tract dimensions can be scientifically established, vocal sounds can be synthesised by using an electronic larynx sound source² and a 3-D printed vocal tract³. Since the restoration of an exact vocal sound requires the perfect preservation of the soft tissues, this is impossible for individuals whose remains are only skeletal. Even where soft tissue does survive, for example in mummified remains⁴, the vocal tract can either be missing or distorted⁵. The process is only feasible when the relevant soft tissue is reasonably intact, as in the case of the 3,000 year-old mummified body of the Egyptian priest Nesyamun⁶, whose 'in death' vocal tract acoustic output has been scientifically synthesised. This acoustic output is for the single sound for the extant vocal tract shape; it does not provide a basis for synthesising running speech. To do so would require knowledge of the relevant vocal tract articulations, phonetics and timing patterns of his language. The synthesised vowel sound based on the precise dimensions of his unique vocal tract is here compared to modern vowels as proof of method and to demonstrate future research potential.

Having established the scientific recreation of a 3-D printed vocal tract unique to a living individual, the 'Voices from the Past' Project was set up to investigate this possibility for those long dead in cases where their remains are sufficiently well preserved. With the need for optimum preservation of the vocal tract an essential requirement, combined with the practical necessity for precise CT-imaging in close proximity to the individual selected, the mummified body of Nesyamun was a highly appropriate choice. This was also true for archaeological reasons.

The Egyptian Nesyamun (Fig. 1) lived during the politically volatile reign of pharaoh Ramses XI (c.1099–1069 BC) over 3000 years ago, working as a scribe and priest at the state temple of Karnak in Thebes (modern Luxor). His voice was an essential part of his ritual duties which involved spoken as well as sung elements⁷.

With his mummified remains now displayed in Leeds City Museum, the current project is only the most recent to examine Nesyamun, whose remains have been at the forefront of mummy studies for almost two centuries. Following the unwrapping of his body in 1824, it was examined by members of the Leeds Philosophical and Literary Society including three surgeons and a chemist whose multidisciplinary scientific investigation published in 1828⁸ was the first of its kind. Following the development of X-rays, the body underwent radiological examination in 1931 at the University of Leeds' School of Medicine, in 1964 by the University of Sheffield School of Dentistry, and in 1990 at the University of Manchester by a team using endoscopy, histology, X-ray and early CT scanning techniques^{9,10}. These combined studies revealed that Nesyamun had died in his mid-50s⁹ and had suffered from gum disease and severely worn teeth, yet nonetheless "had a strong well-developed mandible", which, like the maxilla, was 'prognathic', and "clearly Nubian blood had once coursed through his veins"¹¹.

¹Department of Electronic Engineering, Royal Holloway, University of London, Egham, Surrey, United Kingdom.

²Department of Archaeology, University of York, The King's Manor, York, United Kingdom. ³Leeds Museums and Galleries, Leeds, United Kingdom. ⁴Medical Physics Department, Old Medical School, Leeds General Infirmary, Leeds, United Kingdom. ⁵Institute for Prehistory, Early History and Medieval Archeology, University of Tübingen, Tübingen, Germany. *email: david.howard@rhul.ac.uk; john.schofield@york.ac.uk



Figure 1. The mummified body of Nesyamun laid on the couch to be CT scanned at Leeds General Infirmary. © Leeds Teaching Hospitals/Leeds Museums and Galleries.



Figure 2. Nesyamun's name in hieroglyphs as shown in his coffin inscriptions.

His coffin inscriptions give the name Nesyamun (Fig. 2)⁷, but as one of the first ancient Egyptian names to be translated following the decipherment of hieroglyphs in 1822, this was initially read 'Natsif-Amon'⁸ with at least nine later variants^{9,12} until eventually corrected to Nesyamun⁹. This was a vital clarification within ancient Egyptian culture in which the name was regarded as essential to an individual as their physical (mummified) body and their soul (ka) and spirit (ba). It was also a fundamental belief that 'to speak the name of the dead is to make them live again' (alternatively translated: 'a man is revived when his name is pronounced'¹³), both by living relatives and by the deceased themselves when appearing before the gods of judgement. Only those able to verbally confirm that they had led a virtuous life were granted entry into eternity and awarded the epithet 'maat kheru', 'true of voice'¹⁴, as applied to Nesyamun himself throughout his coffin inscriptions. In these texts, Nesyamun asks that his soul receives eternal sustenance, is able to move around freely and to see and address the gods⁹ as he had in his working life. Therefore his documented wish to be able to speak after his death, combined with the excellent state of his mummified body, made Nesyamun the ideal subject for the 'Voices from the Past' project for which his body was re-examined using state-of-the-art CT scanning equipment.

Since human remains have unique status not as 'objects' but as the remains of once-living people (see SI), it was also necessary to consider the ethical issues raised by the research and its possible heritage outcomes^{15,16}. The team concluded that the potential benefits outweighed the concerns, particularly because Nesyamun's own words express his desire to 'speak again' and that the scientific techniques used were non-destructive.

The CT images confirmed that a significant part of the structure of Nesyamun's larynx and throat remains *in situ* as a result of the elaborate mummification process, thus enabling the vocal tract shape to be measured. The tongue, however, has lost its muscular bulk over time and the soft palate is not present as illustrated in Fig. 3. The dimensions of Nesyamun's tract are 81.4 mm between the external front of the upper lip and the hard/soft palate boundary and 68.4 mm between the thyroid notch and the hard/soft palate boundary. Comparable measurements for two living adult males are 103.6/111.0 mm and 80.0/86.0 mm respectively. Nesyamun's tract therefore appears notably smaller than those of contemporary adult males.

Following the scans, a 3-D printed tract was created for Nesyamun and designed to be used with the Vocal Tract Organ¹⁷ which provides an appropriate acoustic larynx source as a time domain waveform synthesis of the Liljencrants-Fant (LF) larynx source which is commonly employed in speech synthesis¹⁸. The fundamental frequency, loudness and vibrato rate and depth can be individually controlled. The tract incorporates a coupler at its larynx end that is designed to fit snugly over the output end of an Adastra model 952-210 (16 ohm, 60 Watt) loudspeaker drive unit.

Figure 4 shows long-term average spectra for (1) the source signal from the Vocal Tract Organ via the Adastra loudspeaker alone (dotted line) and (2) the output from the 3-D printed vocal tract for Nesyamun placed atop the Adastra loudspeaker (solid line). The joystick controlled version of the Vocal Tract Organ was used to create a larynx input to the 3-D printed vocal tract of Nesyamun and the joysticks were not altered during the sound. Four

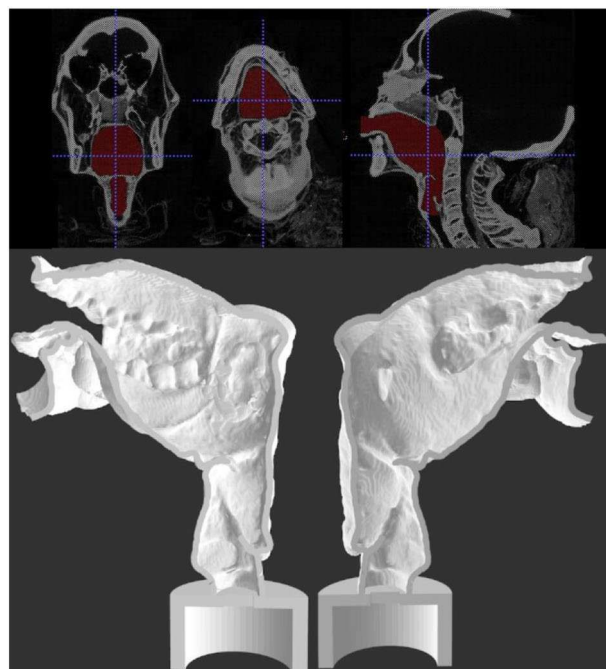


Figure 3. Final segmentation view (upper) and sagittal section of the two halves of 3-D printed Nesyamun's vocal tract (lower). The lack of tongue muscular bulk and soft palate is clear.

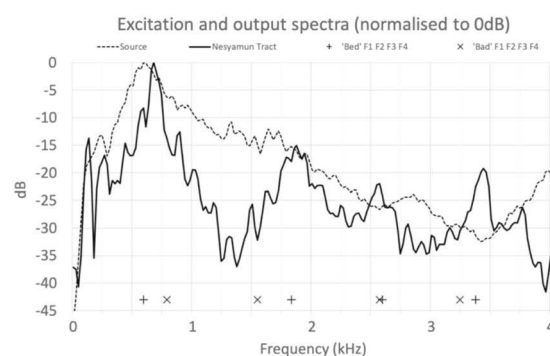


Figure 4. Long-term average spectra for (a) the Vocal Tract Organ output larynx source (dashed), and (b) the output from the 3-D printed vocal tract for Nesyamun acoustically excited by the Vocal Tract Organ output (solid). The first four formant positions averaged from six male English speakers are shown for the words 'bad' (+) and 'bed' (x).

formant peaks are evident in the output spectrum for Nesyamun's 3-D printed vocal tract which are indicative of four resonances in Nesyamun's vocal tract within this frequency range.

Of vital perceptual importance in recreating a natural vocal sound is the application of some form of fundamental frequency variation. Nesyamun's duties included speaking as well as chanting or singing the daily liturgy⁷, so the vocal tract organ was used to provide a falling intonation in the male speech fundamental frequency range. Narrowband (30 ms window) and wide-band (5 ms window) spectrograms for a few vibrato cycles of the output are shown in Fig. 5 and formant frequency values averaged throughout the sound measured using Praat^{19,20} are shown in Table 1.

The measured lower three formant values for the vowel of Nesyamun fall between the vowel in 'bed' and the vowel in 'bad' within the formant data quoted in the classic 1952 work by Peterson and Barney²¹, based on the closeness in frequency of the second and first formants respectively. By way of a modern comparison, the first four formant frequencies averaged for six adult male English speakers for the vowels in 'bed' and 'bad', measured using Praat, are given in Table 1 and plotted in Fig. 4. The match for F1 and F3 is remarkably good but there are differences for F2 and F4. There will not be an exact match because: (1) no two vocal tracts are exactly the same so there will always be formant frequency differences between speakers, (2) any acoustic similarity between a modern English pronunciation of 'bed' and 'bad' and the language of Nesyamun cannot be assumed, and (3)

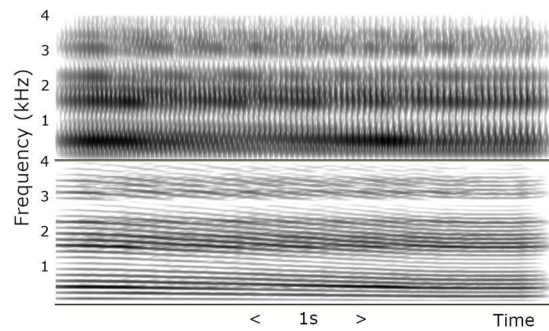


Figure 5. Wide band 30 ms window (upper) and narrow band 5 ms window (lower) spectrograms for a quasi-spoken falling intonation generated using a joystick controlled Vocal Tract Organ driving the 3-D printed vocal tract for Nesyamun.

| | F1 (Hz) | F2 (Hz) | F3 (Hz) | F4 (Hz) |
|------------------------|---------|---------|---------|---------|
| Nesyamun tract | 690 | 1870 | 2560 | 3420 |
| Mean for 6 males 'bed' | 594 | 1821 | 2578 | 3350 |
| Mean for 6 males 'bad' | 783 | 1542 | 2559 | 3227 |

Table 1. Measured average frequencies for formants one to four (F1–F4) throughout the sound created with the 3-D printed vocal tract for Nesyamun excited by the Vocal Tract Organ. Comparison data of averaged first four formant frequencies for six adult male English speakers for the vowels in 'bed' and 'bad' for comparison.

Nesyamun's vocal tract posture is not set for speaking any specific vowel; rather it is set appropriate for his burial position. In addition, his tongue has lost much of its muscle bulk and his soft palate is missing.

Aside from the archaeological possibilities based on the relatively rare survival of intact vocal tract soft tissue as in the case of Nesyamun, it is possible to estimate vocal tract shapes purely from skeletal information²² via, for example, ellipses of dispersion for skull bony landmarks²³. Making use of 3-D printed vocal tracts is one way to recreate the sound output; it is also feasible to calculate the acoustic output based on digital waveguides²⁴ and this offers a future possibility of producing a dynamic acoustic output of spoken or sung utterances.

Previous attempts to recreate the voice of an ancient individual have employed software techniques to animate a facial reconstruction image to give an approximation of the original voice^{25,26}. Recordings do exist of individuals with extraordinary voices who died soon after the introduction of sound recording, such as the last castrato (Alessandro Moreschi - recorded in 1902 and 1904)²⁷, but here is offered a vocal recreation that is based on an extant vocal tract preserved over three millennia. This innovation has implications for the way in which the past is presented to the public, either through conventional heritage interpretation displays or via digital interventions.

As a rare witness to a cataclysmic period in Egypt's ancient history, Nesyamun also has a pre-eminent place in the history of Egyptology. His body and coffin have been on permanent display in Leeds Museum for almost two centuries, and although few visitors can read his coffin's hieroglyphic texts for themselves, the possibility of transmitting their vocalisation would not only fulfil Nesyamun's own wishes as he himself expressed, but make them accessible to all²⁸. Having considered and accommodated all ethical implications, the transmission of sound resulting from his actual vocal tract after a three millennia silence would mean that those who come to see him would also be able to *hear* a sound from his vocal tract as an initial step, emphasising his humanity with the potential to excite and inspire.

Similarly, the well-preserved temple of Karnak in which Nesyamun undertook his duties is the destination for over a million visitors each year, providing further exciting possibilities for heritage interpretation within Egypt's tourist economy.

This innovative interdisciplinary collaboration has produced the unique opportunity to hear the vocal tract output of someone long dead by virtue of their soft tissue preservation and new developments in technology, digital scanning and 3-D printing. While this approach has wide implications for heritage management/museum display, its relevance conforms exactly to the ancient Egyptians' fundamental belief that 'to speak the name of the dead is to make them live again'. Given Nesyamun's stated desire to have his voice heard in the afterlife in order to live forever, the fulfilment of his beliefs through the synthesis of his vocal function allows us to make direct contact with ancient Egypt by listening to a sound from a vocal tract that has not been heard for over 3000 years, preserved through mummification and now restored through this new technique.

Methods

In September 2016 Nesyamun's mummified body was transferred from Leeds City Museum to the nearby Computed Tomography (CT) Scanning Department at Leeds General Infirmary. Once within the scanning room it was removed from its coffin and transferred onto the couch of a Siemens Definition (Erlangen, Germany) multi-detector CT scanner. Positioned on the couch in a head-first supine orientation (Fig. 1), a high resolution helical CT scan was performed from which contiguous axial images of 0.6 mm slice thickness were

reconstructed covering the range from cranial vertex to hallux. Images were acquired at a tube voltage of 120kVp, tube current-time product of 180mAs, detector coverage of 64×0.6 mm and helical pitch factor of 1.0 and were reconstructed using a soft tissue convolution kernel and a 450 mm reconstructed field of view. In order to improve visualisation of the vocal tract, a further set of axial, coronal and sagittal images covering cranial vertex to lung apices were reconstructed using a 0.6 mm slice thickness and 220 mm reconstructed field of view. All image sets were exported in uncompressed DICOM format for further manipulation and processing.

ITK-SNAP²⁹, which allows a three-dimensional structural representation of human tissues to be observed, was used to view the airway between the larynx and lips which is itself isolated as a solid shape to enable the 3-D printing process. On-screen, air is generally represented in black, with volumes of soft tissue and bone being represented in grey to white. The process of creating the vocal tract model itself involves semi-automatic growing of user-defined starting spheres within the black (air) volumes outwards to stop at soft tissue/bone boundaries which are denoted by a change in contrast. This process involves trial and error alongside close observation of the determined boundaries. Making changes to the starting positions of the user-defined spheres and repeating the process as necessary is a core element of this process. Post-processing hand-editing of the final airway with direct reference to the original CT data enables final minor changes to be made as appropriate. In particular, the lack of a soft palate (see Fig. 3) meant that its position had to be estimated prior to printing.

The resulting airway volume represents the inside of Nesyamun's vocal tract as it is preserved. Whilst the vocal tract soft tissue is essentially intact and the oral and pharyngeal cavities are well represented (see Fig. 3), the tongue is desiccated therefore losing the majority of its bulk. In addition, the soft palate is missing and the tract boundary it normally forms has therefore been estimated. A virtual sheath is created around the airway to which a loudspeaker coupler is added. The resulting vocal tract model is 3-D printed (Stratysys Connex 260 machine – 200 micron maximum linear printing error at this scale).

Data availability

The data supporting the findings are fully available without restriction. Relevant data are available from the corresponding author upon request.

Received: 4 June 2019; Accepted: 29 November 2019;

Published online: 23 January 2020

References

1. Story, B. H., Titze, I. R. & Hoffman, E. A. The relationship of vocal tract shape to three voice qualities. *J. Acoust. Soc. Amer.* **109**(4), 1651–1667 (2001).
2. Sleeth, L. E. & Doyle P. C. In *Clinical Care and Rehabilitation in Head and Neck Cancer* (ed. Doyle, P. C.) 231–246 (Springer, Basel, 2019).
3. Delvaux, B. & Howard, D. A. New Method to Explore the Spectral Impact of the Piriform Fossae on the Singing Voice: Benchmarking Using MRI-Based 3D-Printed Vocal Tracts. *PLoS ONE* **9**(7) (2014).
4. Aufderheide, A. C. *The Scientific Study of Mummies* (Cambridge University Press, Cambridge, 2003).
5. Watson, E. & Myers, M. The mummy of Baket-en-her-nakht in the Hancock. *Museum: A radiological update. J. Egypt. Archaeol.* **79**, 179–187 (1993).
6. Tapp, E. & Wildsmith, K. In *The Mummy's Tale* (eds. David, A. & Tapp, E.) 132–153 (Michael O'Mara, London, 1992).
7. Moret, A. Le rituel du culte divin journalier en Égypte, d'après les papyrus de Berlin et les textes du temple de Sêti Ier, à Abydos, (Leroux, Paris 1902).
8. Osburn, W. *An account of an Egyptian mummy, presented to the museum of the Leeds Philosophical and Literary Society, by the late John Blayds, Esq.* (Robinson & Hernaman, Leeds, 1828).
9. Wassell, B. *The Coffin of Nesyamun the 'Leeds Mummy': the first complete translation of the hieroglyphic texts* (Leeds Philosophical and Literary Society Ltd., Leeds, 2008).
10. David, A. & Tapp, E. (eds) *The Mummy's Tale* (Michael O'Mara, London, 1992).
11. Prag, J. & Neave, R. Making Faces: using forensic and archaeological evidence (British Museum Press, London, 1997).
12. Porter, B. & Moss, R. *Topographical Bibliography of Ancient Egyptian Hieroglyphic Texts, Statues, Reliefs and Paintings, Volume I. Part 2: The Theban Necropolis, Royal Tombs and Smaller Cemeteries* (Oxford University Press, Oxford, 1989).
13. Lichtheim, M. *Ancient Egyptian Literature Volume III: the Late Period* (University of California Press, Berkeley, 1980).
14. Allen, J. P. *Middle Egyptian: an Introduction to the Language and Culture of Hieroglyphs* (Cambridge University Press, Cambridge, 2000).
15. DCMS Guidance for the Care of Human Remains in Museums, www.britishmuseum.org/pdf/DCMS%20Guide.pdf (2005).
16. Leeds Museums and Galleries Human Remains Policy, www.leeds.gov.uk/museumsandgalleries/about-us/policies-and-practice (2015).
17. Howard, D. M. The Vocal Tract Organ: A New Musical Instrument Using 3-D Printed Vocal Tracts. *J. Voice* **32**, 660–667 (2018).
18. Fant, G., Liljencrants, J. & Lin, Q. G. A four parameter model of glottal flow. *STL-QPSR* **2**, 119–156 (1985).
19. Boersma, P. Praat, a system for doing phonetics by computer. *Glott International* **5**, 341–345 (2001).
20. Boersma, P. & Weenink, D. Praat: doing phonetics by computer [Computer program] Version 6.0.14, retrieved 13 May 2017 from, <http://www.fon.hum.uva.nl/praat/>.
21. Peterson, G. E. & Barney, H. L. Control methods used in a study of the vowels. *J. Acoust. Soc. Am.* **24**, 175–184 (1952).
22. Budil, I. Functional reconstruction of the supralaryngeal vocal tract of fossil human. *Human Evolution* **9**(1), 35–52 (1994).
23. Boë, L. J., Heim, J. L., Autesserre, D. & Badin, P. Prediction of geometrical vocal tract limits from bony landmarks: Modern humans and Neandertallians in *Speech Production: Models, Phonetic Processes and Techniques* (eds. Harrington, J. & Tabain, M.) 103–119. (Psychology Press, New York, 2005).
24. Speed, M., Murphy, D. T. & Howard, D. M. Three-dimensional digital waveguide mesh simulation of cylindrical vocal tract analogs. *IEEE Trans. Audio, Speech, Language Process.* **21**, 449–455 (2013).
25. Ogden, A. In *The Facial Reconstruction, in Gristhorpe Man: a Life and Death in the Bronze Age* (eds. Melton, N., Montgomery, J. & Knüsel, C.) 142–145 (Oxbow Books, Oxford, 2013).
26. Avanzini, F., Cosi, P., Füstös, R. & Sandi, A. When Fantasy meets science: An attempt to recreate the voice of Ötzi the "Iceman". *Associazione Italiana Scienze della Voce (AISV)* **3**, 425–431 (2017).
27. Howard, D. M. Acoustic and physiological aspects of the castrato voice, In: *Moreschi and the voice of the castrato*, Clapton, N., London: Haus Books, 227–258 (2008).

28. University College London *Audio examples of different languages spoken in Egypt*, <http://www.ucl.ac.uk/museums-static/digitalegypt/sound/index.html> (2002)
29. Yushkevich, P. A. *et al.* User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage* **31**, 1116–28 [Computer program] Version 3.6.0, retrieved 15 July 2017 from www.itksnap.org (2006).

Acknowledgements

We thank Steve Alty for technical support with tract measurements. Support for this research has been provided by the Centre for Digital Heritage, University of York and Pharos Research. Thanks to Mengli Feng and Alex Clarke for help with the 3-D printing process.

Author contributions

Original concept: J.S. Project design: J.S. and D.M.H. Historical and archaeological background: J.F. and S.A.B. Heritage and ethical considerations: J.S., J.F. and K.B. Performed the experiments: D.M.H. and G.I. Analyzed the data: D.M.H. and G.I. Contributed to the writing of the manuscript: D.M.H., J.S., J.F., K.B., G.I. and S.A.B.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-019-56316-y>.

Correspondence and requests for materials should be addressed to D.M.H. or J.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020