DIVERSITY OF MALE IDENTITIES IN EARLY AND MIDDLE LA TÈNE PERIOD CEMETERIES IN CENTRAL EUROPE

Peter C. Ramsl

Archaeologica Slovaca Monographiae

STUDIA

Instituti Archaeologici Nitriensis
Academiae Scientiarum Slovacae

TOMUS XXXII

NITRAE MMXX

ARCHAEOLOGICKÝ ÚSTAV
SLOVENSKEJ AKADÉMIE VIED

DIVERSITY OF MALE IDENTITIES IN EARLY AND MIDDLE LA TÈNE PERIOD CEMETERIES IN CENTRAL EUROPE

Peter C. Ramsl

With contributions from
Friederike Novotny, Vienna (Anthropology),
and
Joscha Gretzinger – Stephan Schiffels, Jena (DNA)

This publication was funded by the Slovak Academy of Sciences, Marie Curie Actions
and FP 7 of the European Commission (SASPRO project 1340/03/03)
and by the Scientific Grant Agency of the Ministry of Education, Science, Research and Sport
of the Slovak Republic and the Slovak Academy of Sciences (VEGA project 1/0399/18).

NITRA 2020

Peter C. Ramsl

Diversity of Male Identities in Early and Middle La Tène Period Cemeteries in Central Europe

With contributions from

Friederike Novotny, Vienna (Anthropology),
and
Joscha Gretzinger – Stephan Schiffels, Jena (DNA)

Nitra 2020

Archaeologica Slovaca Monographiae
Studia, Tomus XXXII

Chief editor	doc. PhDr. Matej Ruttkay, CSc.
Reviewers	Dr. Marko Dizdar Dr. Mariana Egri Károly Tankó, PhD
Supervising editor	Lucia Benediková
Language editor	Philip Earl Steele
Layout	Beáta Jančíková
Cover	Peter C. Ramsl, Thomas Pertlwieser, Beáta Jančíková
Print	VEDA, vydavateľstvo SAV, Bratislava www.veda.sav.sk

Distribution	Archeologický ústav SAV Akademická 2 SK – 949 21 Nitra tel.: +421/37/6943 209 e-mail: nraukniz@savba.sk

Publisher

ISBN 978-80-8196-040-6

CONTENTS

REPORT ON GENETIC ANALYSES OF IRON AGE SAMPLES FROM SLOVAKIA

Joscha Gretzinger – Stephan Schiffels

INTRODUCTION

Genetic analyses of archaeological skeletal human remains have become routine thanks to next generation sequencing methodology, which allows us to analyze hundreds of millions of short DNA fragments in parallel. We obtained genetic data from 8 of 9 analyzed individuals from three archaeological sites in Slovakia (Palárikovo, Chotín, and Dubník), dating to the La Tène period. These analyses allowed us to confidently obtain the genetic sex of the 8 successful samples, as well as to identify preliminary ancestry information for two of the analyzed samples.

SAMPLE OVERVIEW

The following table shows a summary of the data, including individual CHT002 for which extremely few human DNA molecules could be found, and thus we did not analyze this individual further. The column 'Unique mapped reads' gives the number of sequenced DNA molecules that map to the human genome, i.e., which are likely of human origin. The endogenous DNA% indicates the fraction of human reads in the total DNA sample.

Table 1. Summary of the genetic analyses data.

Sample Name	Pandora ID	Sequenced reads	Unique mapped reads	Endogenous DNA (%)	AVG Cov NT	Damage first base 5'	Molecular Sex
I-2877 12/71	CHT001	6004618	282888	7.21	0.0046	0.18	male
I-2907 30/72	CHT002	4905954	874	0.04	0	0.15	unassigned
24/83	DUB001	6276704	677669	16.78	0.0106	0.16	male
73/72	PAL001	6368302	3173262	76.42	0.0562	0.14	male
53/71	PAL002	5080732	2203866	69.08	0.0369	0.18	female
5/70	PAL003	5105819	547189	16.87	0.008	0.15	male
70/72	PAL004	5394002	1382497	45.07	0.0195	0.17	male
58/72	PAL005	6167315	91183	2.41	0.0015	0.13	female
18/70	PAL006	7158739	372547	8.05	0.0055	0.17	male

Sample processing and sequencing

There are two established approaches to sampling the human petrous bone (pars petrosa) for DNA: Cutting the bone and drilling into the inside, or drilling into the bone from the outside. In this project, we have chosen the method of drilling from the outside, which is less invasive, but can sometimes lead to a lower yield of DNA or higher contamination. Figure 1 illustrates the two approaches before and after sampling.

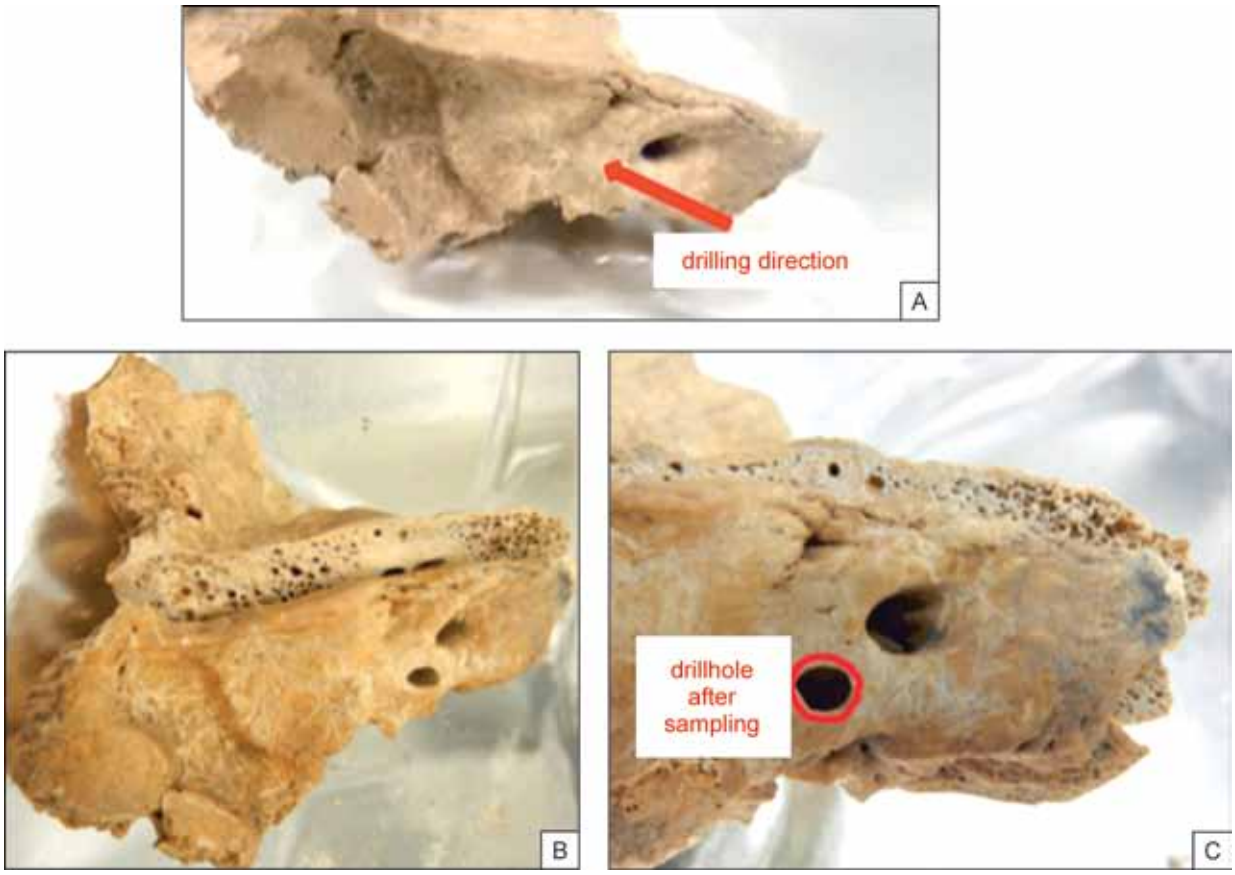

Fig. 1. Sampling the petrous bone for DNA. A – petrous bone before sampling; B – Petrous bone after sampling, and cut in half; C – Similar as in B but with marked drill hole. For illustrative purposes, images B and C show both the invasive and the non-invasive sampling procedures. The bones shown are not from the specimens analyzed here. Images kindly provided by Marie Himmel, MPI–SHH.

DNA from 9 samples from Chotín, Palárikovo, and Dubník was extracted in the ancient DNA facilities of the Max Planck Institute for the Science of Human History (MPI-SHH) in Jena, Germany. For each specimen, ~50 mg of dentine powder was used for an extraction procedure specifically designed for ancient DNA retrieval (*Dabney et al. 2013*). Extraction buffer containing 0.45 M EDTA, pH 8.0 (Life Technologies) and 0.25 mg/ml Proteinase K (Sigma-Aldrich) was added to the bone powder and incubated at 37 °C with rotation overnight. The supernatant was separated from the pellet of bone powder by centrifugation (14,000 rpm). A binding buffer consisting of 5 M GuHCL (Sigma Aldrich) and 40% Isopropanol (Merck), together with 400 µl of 1 M sodium acetate (pH 5.5) was added to the supernatant, and the solution purified by spinning it through a purification column attached to a High Pure Extender Assembly funnel (8 min in 1500 rpm, with slow acceleration). The column was then spun into a collection tube (1 min 14,000 rpm) 1–2 times to maximize the yield. This was followed by two subsequent washing steps of 450 µl of wash buffer (High Pure Viral Nucleic Acid Large Volume Kit) and two dry spin steps of 1 min centrifugation at 14,000 rpm. The final total volume of 100 µl eluate was reached by two separate elution rounds of 50 µl of TET (10 mMTris-HCL, 1 mM EDTA pH 8.0, 0.1% Tween20), each spun for 1 min at 14,000 rpm into a fresh Eppendorf 1.5 ml tube. Negative controls (buffer instead of sample) were processed in parallel at a ratio of 1 control per 7 samples.

Of the 100 µl extract, 20 µl was used to immortalize the sample DNA as a double-stranded library. The procedure included a blunt-end repair, adapter ligation, and adapter fill-in steps, as described by Meyer and Kircher (*Meyer/Kircher 2010*). During the blunt-end repair step, a mixture of 0.4 U/µl T4 PNK (polynucleotide kinase) and 0.024 U/µl T4 DNA polymerase, 1× NEB buffer 2 (NEB), 100 µM dNTP mix (Thermo Scientific), 1 mM ATP (NEB) and 0.8 mg/ml BSA (NEB) was added to the template DNA, followed by incubation in a thermocycler (15 min 15 °C, 15 min 25 °C) and purification with a MinElute kit (QIAGEN). The

product was eluted in 18 µl TET buffer. The adapter ligation step included a mixture of 1× Quick Ligase Buffer (NEB), 250 nM Illumina Adapters (Sigma-Aldrich) and 0.125 U/µl Quick Ligase (NEB), added to the 18 µl eluate, followed by a 20 min incubation, and second purification step with MinElute columns, this time in 20 µl eluate. For the fill-in step, a mixture of 0.4 U/µlBst-polymerase and 125 µM dNTP mix was added and the mixture then incubated in a thermocycler (30 min 37 °C, 10 min 80 °C). To introduce the UDG-half treatment, an initial stage was included in the library preparation, in which 250 U USER enzyme (NEB) was added into the 20 µl of extract, followed by an incubation at 37 °C for 30 min, and then 12 °C for 1 min. This again was followed by the addition of 200 U UGI (Uracil Glycosylase inhibitor, by NEB) and another identical incubation to stop the enzymatic excision of deaminated sites, as described in (*Rohland et al. 2015*). For each library, a unique pair of eight-bp-long indexes was incorporated using a Pfu Turbo CxHotstart DNA Polymerase and a thermocycling program with the temperature profile as follows: initial denaturation (98 °C for 30 sec), cycle of denaturation/annealing/ elongation (98 °C for 10 sec/ 60 °C for 20 sec/ 72 °C for 20 sec) and final extension at 72 °C for 10 min (*Kircher/Sawyer/Meyer 2012*). Bone powder from a cave bear was processed in parallel serving as a positive control. Negative controls for both extraction and library preparation stages were kept alongside the samples throughout the entire workflow.

Experiment efficiency was ensured by quantifying the concentration of the libraries on qPCR (Roche) using aliquots from libraries before and after indexing. The molecular copy number in pre-indexed libraries ranged from ~ 10E8 to ~ 10E9 copies/µl, indicating a successful library preparation, whereas the indexed libraries ranged from ~ 10E10 to ~ 10E12 copies/µl, stating an admissible indexing efficiency. The negative controls showed 4–5 orders of magnitude lower concentration than the samples, indicating low contamination levels from the laboratory processing stages.

The libraries were amplified with PCR, for the amount of cycles corresponding to the concentrations of the indexed libraries, using AccuPrimePfx polymerase (5 µl of library template, 2 U AccuPrimePfx DNA polymerase by Invitrogen, 1 U of readymade 10× PCR mastermix, and 0.3 µM of primers IS5 and IS6, for each 100 µl reaction) with thermal profile of 2 min denaturation at 95 °C, 3–9 cycles consisting of 15 sec denaturation at 95 °C, 30 sec annealing at 60 °C, 2 min elongation at 68 °C and 5 min elongation at 68 °C. The amplified libraries were purified using MinElute spin columns with the standard protocol provided by the manufacturer (Qiagen), and quantified for sequencing using an Agilent 2100 Bioanalyzer DNA 1000 chip.

We used EAGER (*Peltzer et al. 2016*; version 1.92.50) to process the sequenced reads, using default parameters (see below) for human-originated, UDG-half treated, single-end sequencing data. Specifically, AdapterRemoval was used to trim the sequencing adapters from our reads, with a minimum overlap of 1 bp, and using a minimum base quality of 20 and minimum sequence length of 30 bp. BWA aln (version 0.7.12-r1039, https://sourceforge.net/projects/bio-bwa/files). *Li/Durbin* (*2009*) was used to map the reads to the GRCh37 hg19 human reference sequence, with a seed length (-l) of 32, max number of differences (-n) of 0.01 while applying a quality filter (-q) of 30. Duplicate removal was carried out using DeDup v0.12.1. Terminal base deamination damage calculation was done using mapDamage, specifying a length (-l) of 100 bp.

For downstream analyses, we used bamutils (version 1.0.13, https://github.com/ statgen/bamUtil.git) TrimBam to trim two bases at the start and end of all reads. This procedure eliminates the positions that are affected by deamination, thus removing genotyping errors that could arise due to ancient DNA damage.

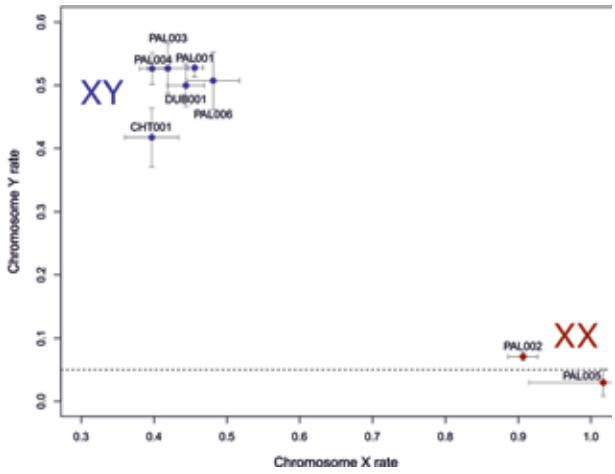

Fig. 2. Illustration of coverage on the X and Y chromosome. Shown are relative genomic coverages of the sex chromosomes relative to the autosomes. In males, we expect both the X and the Y relative coverage at 1/2 of the autosomes. In females, we expect Y coverage to be close to zero, and the X coverage to be around the same as that for the autosomes.

GENETIC SEX DETERMINATION

We used the short-read alignments to the human genome to measure mean coverage on the X and Y chromosomes, relative to the autosomes. Since male individuals have one Y and

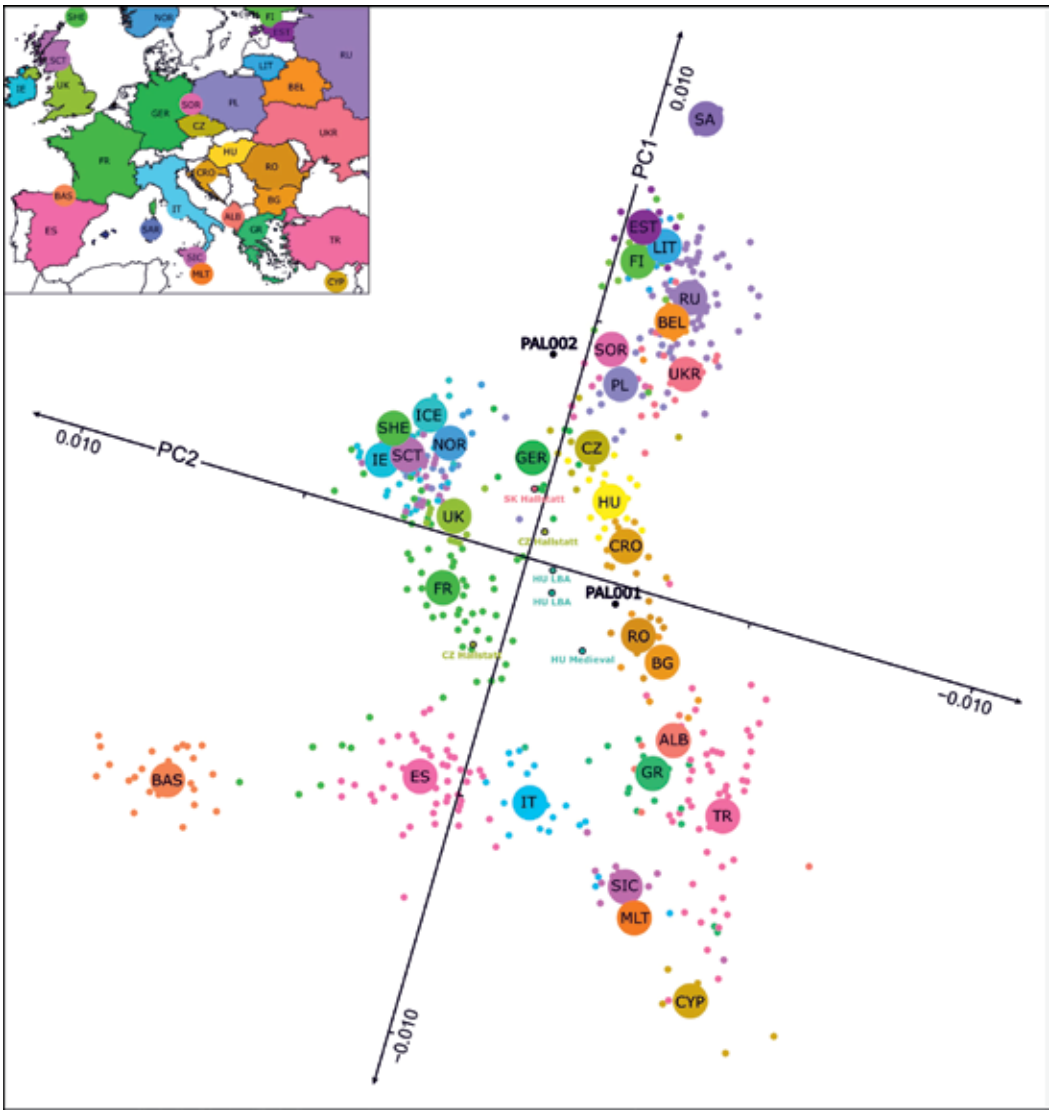

Fig. 3. Principal components analysis with 623 published modern European samples (*Lazaridis et al. 2014*) and several other published ancient individuals (*Damgaard et al. 2018*; *Gamba et al. 2014*; *Mathieson et al. 2015*).

one X chromosome, but two of each autosome, we expect both X and Y to have relative genomic coverage of ½. In contrast, females have two copies of the X chromosome, and no Y chromosome, so their relative coverage on X should be 1 (relative to autosomes) and 0 on the Y (Fig. 2).

With the exception of CHT002, which failed and produced not enough human DNA, we indeed find two clusters of samples, with only two female samples (PAL002 and PAL005) and all other samples being male.

PRELIMINARY ANCESTRY ANALYSIS

We used Principal Components Analysis to analyze the ancestry of the two best preserved individuals. We first computed principal components of 623 modern European samples from 37 populations (*Lazaridis et al. 2014*), and then projected the genetic data from PAL001 and PAL002 onto those principal components (Fig. 3). The analysis reveals some heterogeneity of the two individuals, with PAL001 falling more towards the south/southeast, clustering together with present-day Romanians and Croatians, whereas PAL002 clusters with present-day Sorbian and Polish people.

OUTLOOK AND CONCLUSION

DNA Preservation in the three analyzed sites was remarkably good. We could successfully retrieve human DNA, reaching up to 60% in PAL001, from all but one individual. This proved sufficient for accurate determination of the genetic sex of the successful eight sample. At this stage, only small fractions of the human genomes of these individuals were reconstructed, and deeper shotgun sequencing, or application of Capture Enrichment technology (*Haak et al. 2015*; *Mathieson et al. 2015*), to retrieve the complete genomes of these individuals, which reveal fine-scale population ancestry of all samples. At this point, we could however already obtain ancestry estimates for two of the eight successful samples, which revealed an overall Central European ancestry, with some difference between the two, spanning from Polish/Sorb genetic ancestry towards more southern Croatian or Romanian ancestry. Deeper sequencing will reveal whether this difference reflects a single outlier, or a general pattern of heterogeneity in these three Iron Age populations.

LITERATURE

Dabney et al. 2013 — J. Dabney/M. Knapp/I. Glocke/M.-T. Gansauge/A. Weihmann/B. Nickel/C. Valdiosera, et al.: "Complete Mitochondrial Genome Sequence of a Middle Pleistocene Cave Bear Reconstructed from Ultrashort DNA Fragments." *Proceedings of the National Academy of Sciences of the United States of America* 110 (39), 2013, 15758–63.

Damgaard et al. 2018 — P. de Barros Damgaard/N. Marchi/S. Rasmussen/M. Peyrot/G. Renaud/Th. Korneliussen/J. V. Moreno-Mayar et al.: "137 Ancient Human Genomes from across the Eurasian Steppes." *Nature* 557 (7705), 2018, 369–374.

Gamba et al. 2014 — C. Gamba/E. R. Jones/M. D. Teasdale/R. L. McLaughlin/G. Gonzalez-Fortes/V. Mattiangeli/L. Domboróczki et al.: "Genome Flux and Stasis in a Five Millennium Transect of European Prehistory." *Nature Communications* 5 (October) 2014, 5257.

Haak et al. 2015 — W. Haak/I. Lazaridis/N. Patterson/N. Rohland/S. Mallick/B. Llamas/G. Brandt et al.: "Massive Migration from the Steppe Was a Source for Indo-European Languages in Europe." *Nature* 522 (7555), 2015, 207–211.

Kircher/Sawyer/Meyer 2012 — M. Kircher/S. Sawyer/M. Meyer: "Double Indexing Overcomes Inaccuracies in Multiplex Sequencing on the Illumina Platform." *Nucleic Acids Research* 40 (1), 2012, e3.

Lazaridis et al. 2014 — I. Lazaridis/N. Patterson/A. Mittnik/G. Renaud/S. Mallick/K. Kirsanow/P. H. Sudmant et al.: "Ancient Human Genomes Suggest Three Ancestral Populations for Present-Day Europeans." *Nature* 513 (7518), 2014, 409–413.

Li/Durbin 2009 — H. Li/R. Durbin: "Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform." *Bioinformatics* 25 (14), 2009, 1754–1760.

Mathieson et al. 2015 — I. Mathieson/I. Lazaridis/N. Rohland/S. Mallick/N. Patterson/S. Alpaslan Roodenberg/E. Harney et al.: "Genome-Wide Patterns of Selection in 230 Ancient Eurasians." *Nature*, November 2015. http://www.nature.com/doifinder/10.1038/nature16152.

Meyer/Kircher 2010 — M. Meyer/M. Kircher: "Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing." *Cold Spring Harbor Protocols* 2010 (6), 2010, db. prot 5448.

Peltzer et al. 2016 — A. Peltzer/G. Jäger/A. Herbig/A. Seitz/Ch. Kniep/J. Krause/K. Nieselt: "EAGER: Efficient Ancient Genome Reconstruction." *Genome Biology* 17 (1), 2016, 416.

Rohland et al. 2015 — N. Rohland/E. Harney/S. Mallick/S. Nordenfelt/D. Reich: "Partial Uracil-DNA-Glycosylase Treatment for Screening of Ancient DNA." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 370 (1660), 2015, 20130624.

AUTHORS' ADDRESSES

Privatdozent Mag. Dr. Peter C. Ramsl
Institut für Urgeschichte und Historische Archäologie
Universität Wien
Franz-Klein-Gasse 1
AT – 1190 Wien
peter.ramsl@univie.ac.at

and

Archeologický ústav SAV
Akademická 2
SK – 949 21 Nitra

Mag. Friederike Novotny
Anthropologische Abteilung
Naturhistorisches Museum Wien
Burgring 7
AT – 1010 Wien
friederike.novotny@nhm-wien.ac.at

Joscha Gretzinger
Max-Planck-Institut für Menschheitsgeschichte
Abteilung Archäogenetik
Kahlaische Str. 10
D – 077 45 Jena
gretzinger@shh.mpg.de

Stephan Schiffels
Max-Planck-Institut für Menschheitsgeschichte
Abteilung Archäogenetik
Kahlaische Str. 10
D – 077 45 Jena
schiffels@shh.mpg.de