PALAEOPROTEOMICS

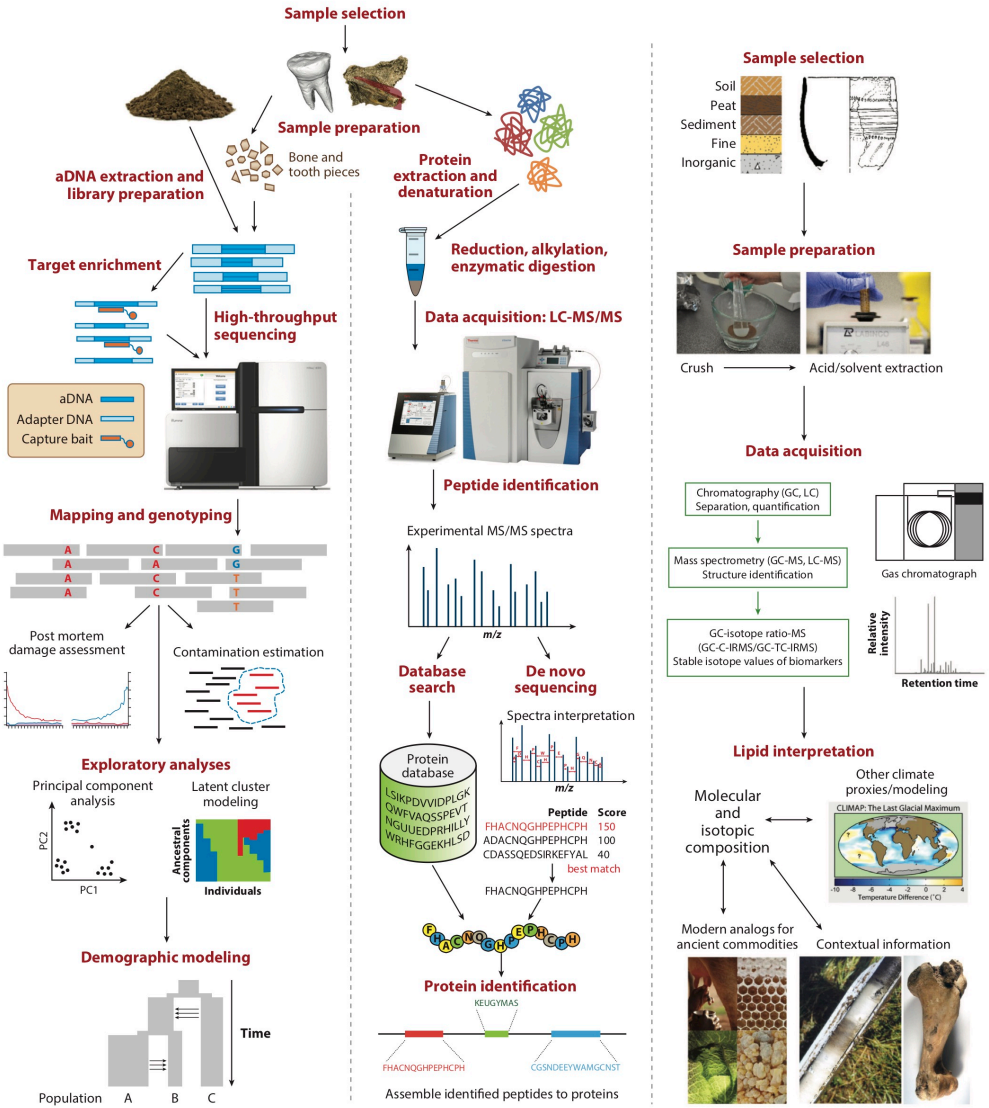
Peptide mass fingerprinting
or
Zooarchaeology by mass spectrometry (ZooMS)

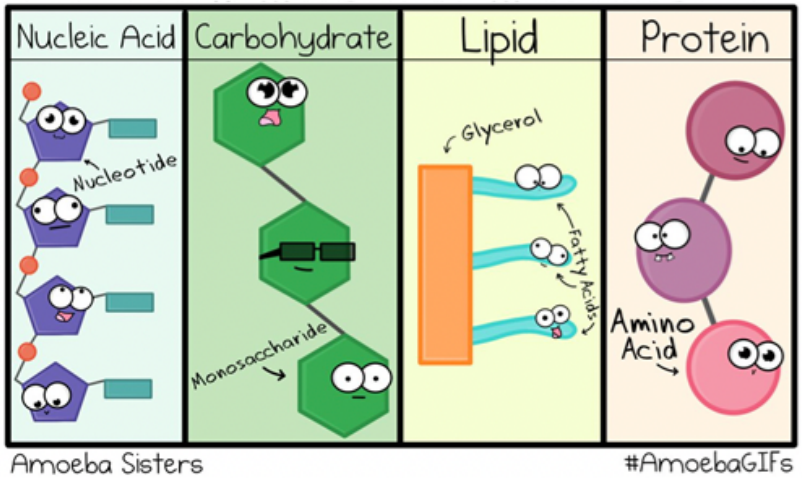
Katerina Douka

Group Leader, Max Planck Institute for the Science of Human History, Germany
Research Associate, University of Oxford, UK

Advances in ancient biomolecules research

Ancient DNA Ancient proteins Ancient lipids



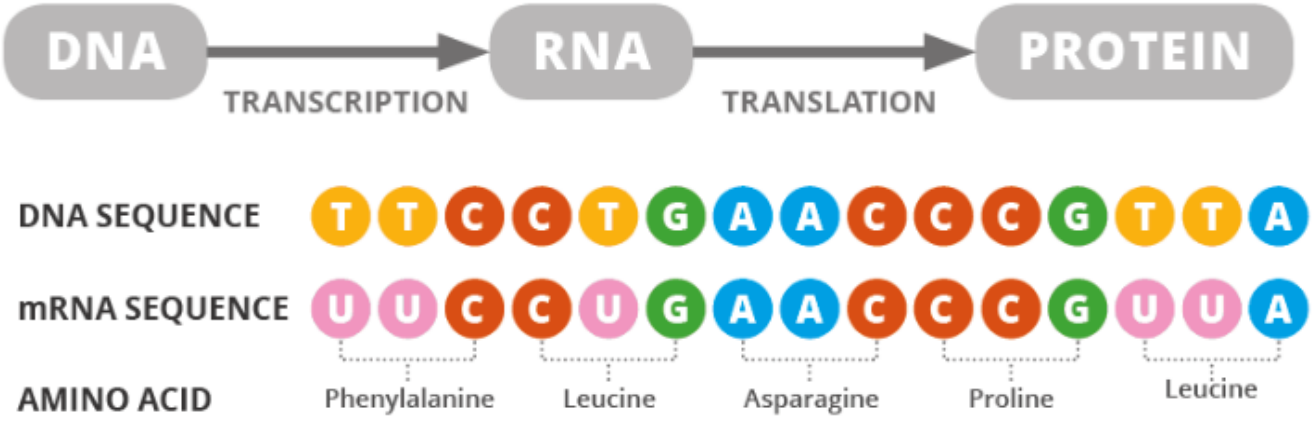


- ✓ technical innovations in instrumentation
- ✓ enhanced analytical protocols
- ✓ criteria for sample selection
- ✓ advanced computational approaches

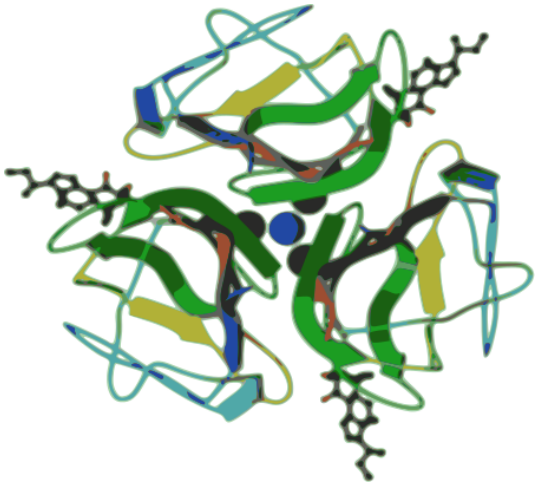
Cappellini et al. 2018

Proteins

- A chain of amino acids makes a protein



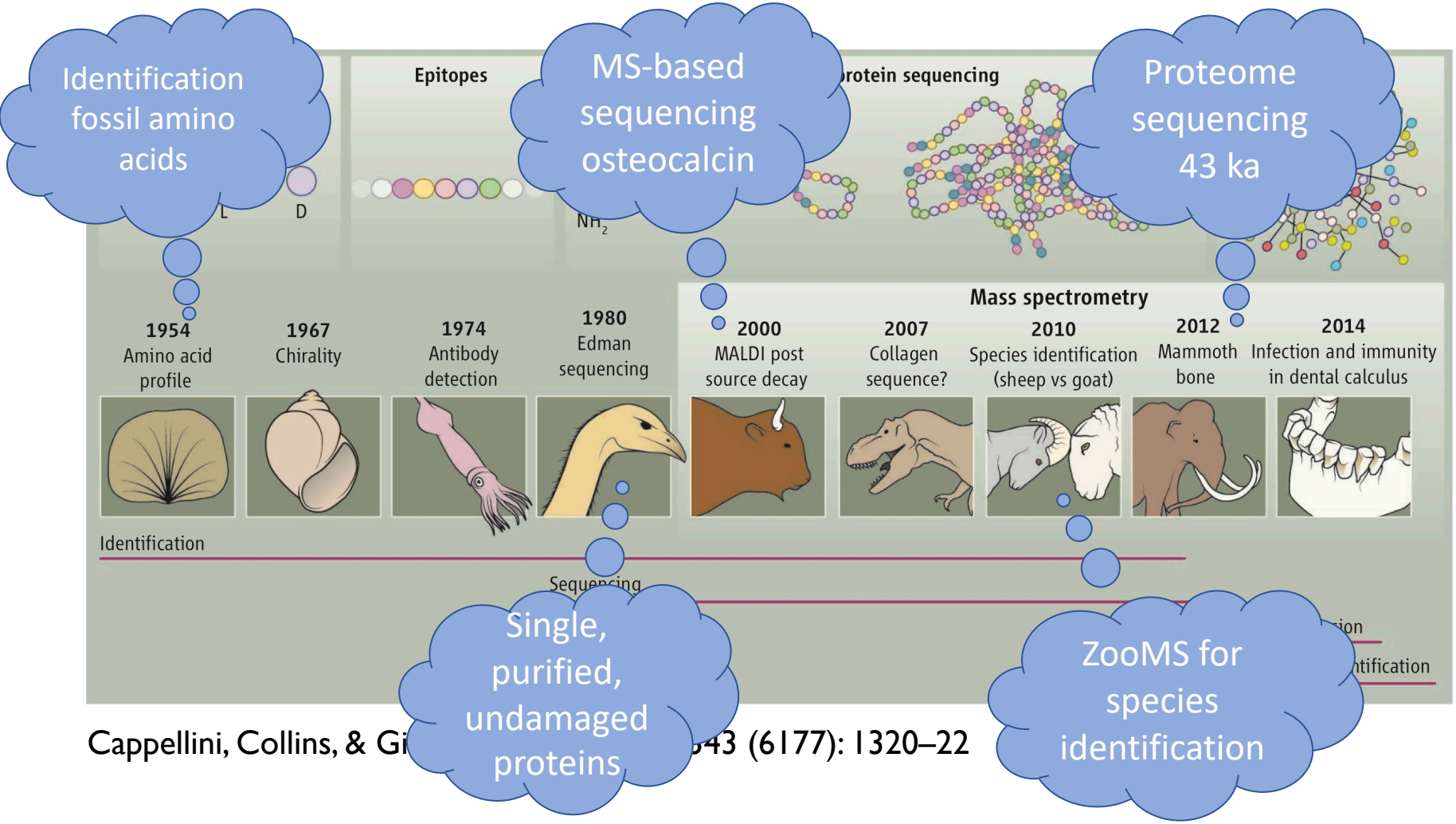
- The long chain folds around in on itself to create a shape
- This shape governs its function
- A group of proteins is called a proteome
- "Proteomics": mass spectrometric analysis of proteins



Source: PDB

Palaeoproteomics : Ancient Protein Analysis

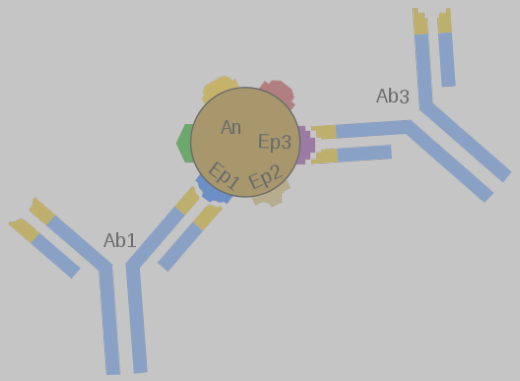
The study of ancient proteins (identification, quantification, modification)



Cappellini, Collins, & Gi 43 (6177): 1320–22

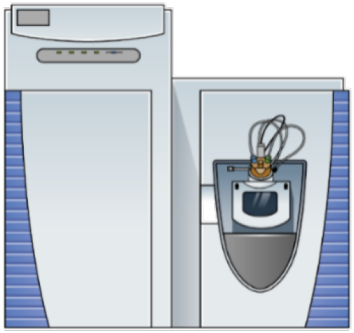
Protein Analysis Techniques & Technology

Antibody methods



- Based on epitope binding
- Targeting single proteins of interest
- Detects presence or absence
- Cheap

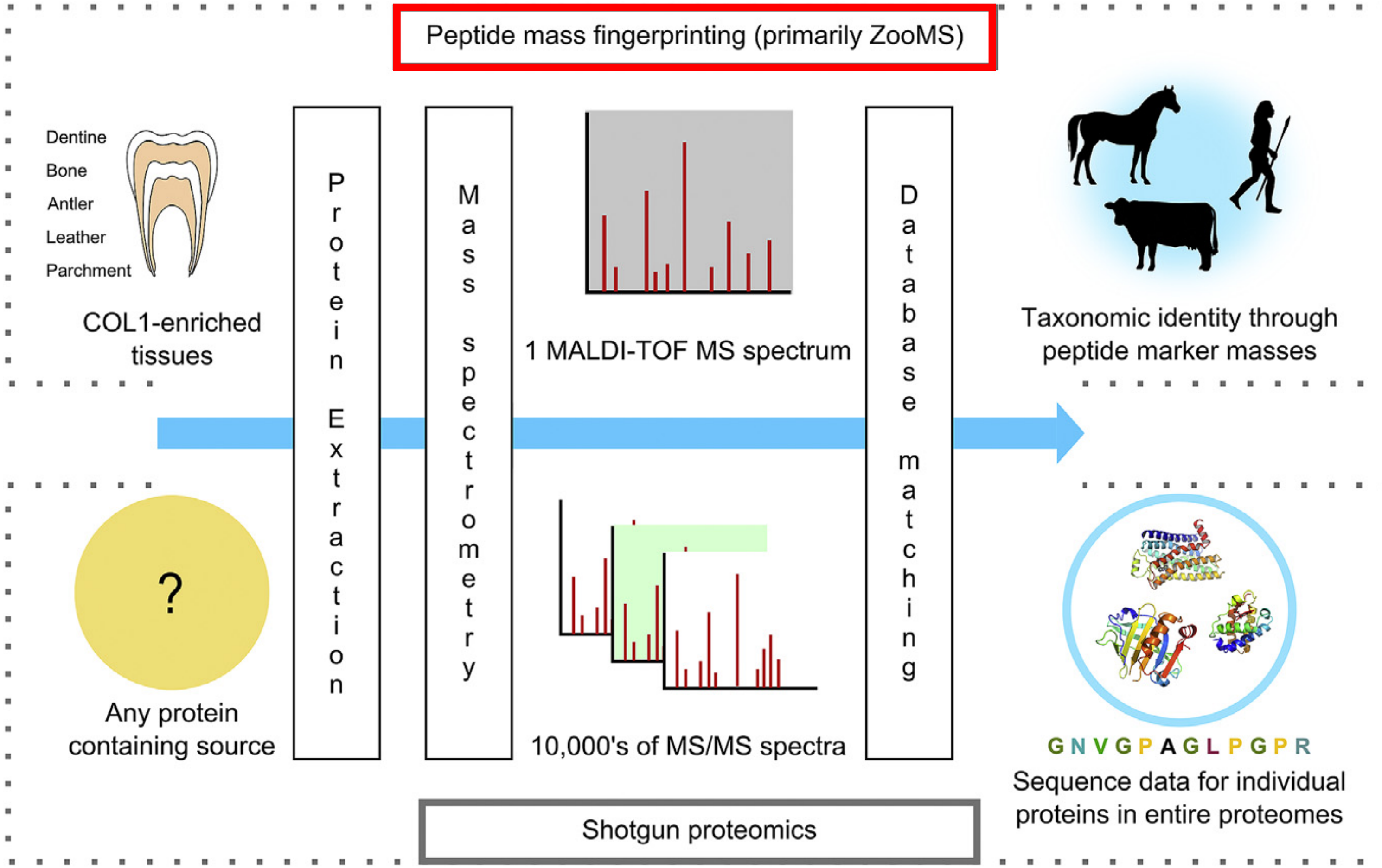
Mass Spectrometry



- Based on the detection of distinctive masses
- Generates peptide sequence information
- Expensive

Palaeoproteomics

Two approaches: **1. PMF**: MALDI TOF-MS **2. Shotgun proteomics**: LC- MS/MS



Welker et al. 2018

1. PMF => MALDI-TOF MS: total mass of individual peptides within an extract
2. Shotgun => LC-MS/MS: mass and exact amino acid sequence of the preserved peptides

Rely on presence and detection of SAPs (single amino acid polymorphisms) derived from nucleotide substitutions (SNPs, single nucleotide polymorphisms) at the genome level

Palaeoproteomics

Well-preserved remains

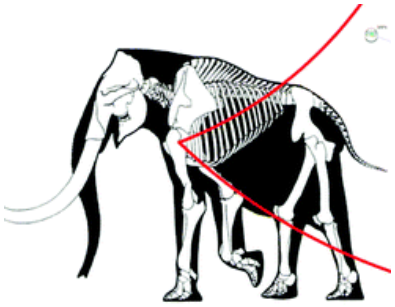
Grape seeds
Cappellini et al. 2010

Cheese remains
Yang et al. 2014

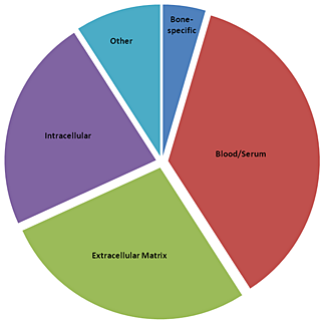
Sourdough bread
Shevchenko et al. 2014

Dairy products
Xie et al. 2016

Bone biology and taphonomy

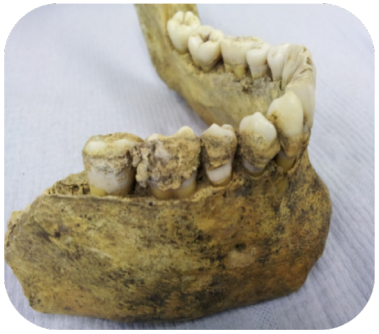


Mammoth Bone
Cappellini et al. 2012



Bone Proteomes
Buckley et al. 2016

Dental Calculus



Oral Microbiome
Warinner et al. 2014

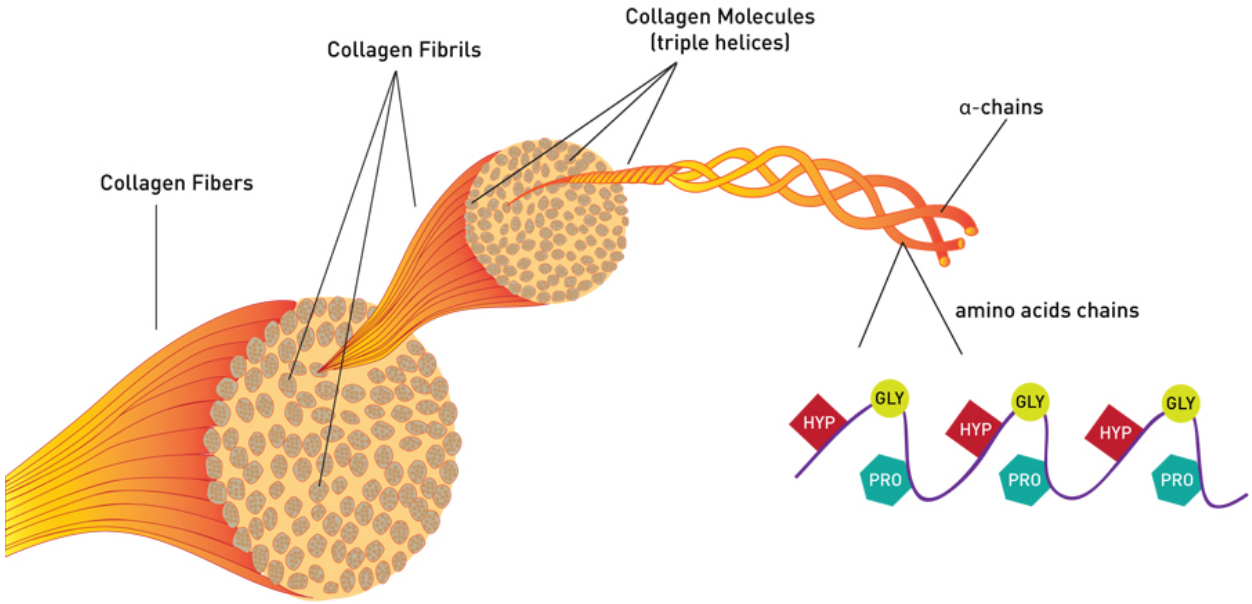
Milk Proteins
Warinner et al. 2014

Palaeoproteomics : Advantages

Proteins are more abundant than other biomolecules

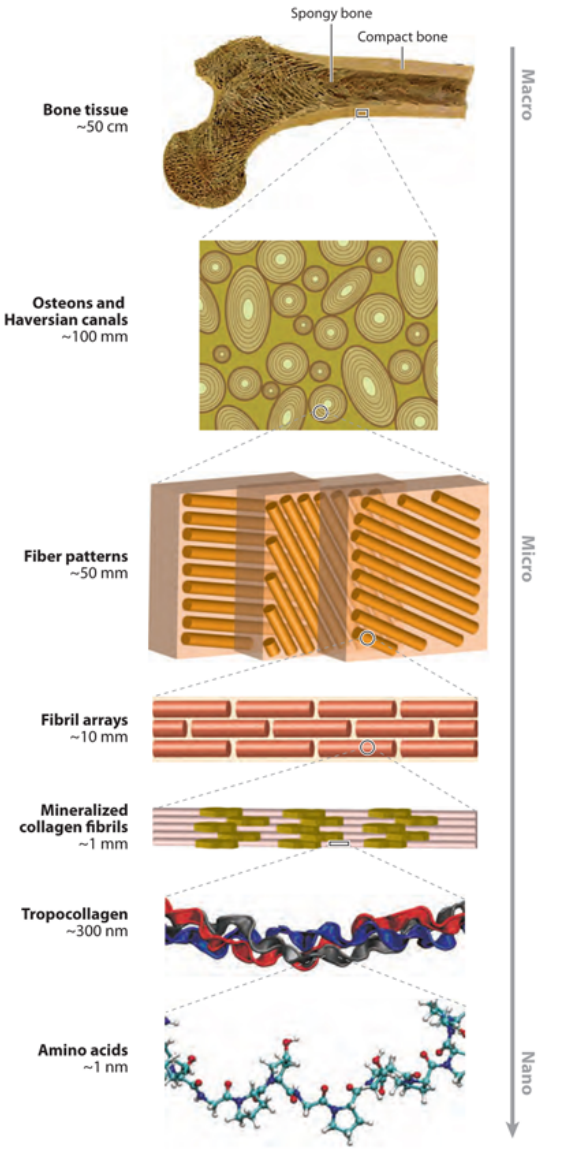
&

Collagen is the most abundant protein in the vertebrate kingdom



>28 types of collagen

Human body 80-90% type I collagen



Hierarchical structure of bone

[from Launey et al. (2010) Annual Review of Materials Research]

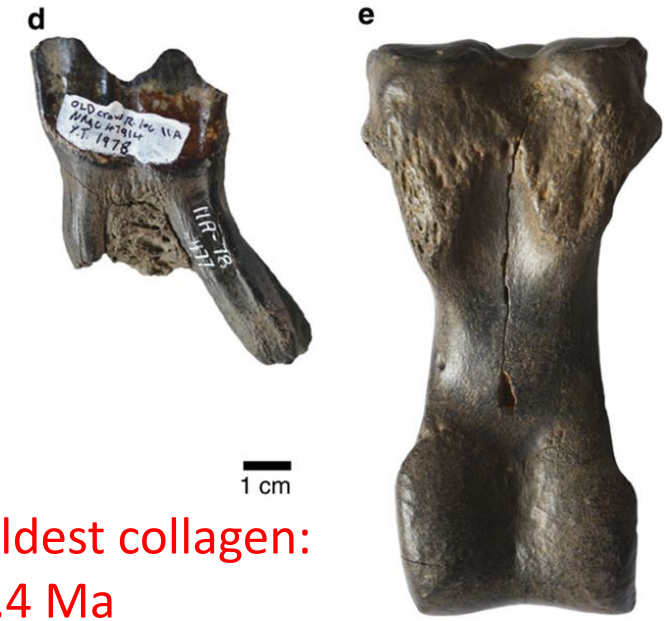
Palaeoproteomics : Advantages

Proteins are more stable and can survive longer than other biomolecules

Orlando et al. 2013

Yao-Chang Lee et al. 2017

Rybczynski et al. 2013



Oldest collagen:
3.4 Ma

Palaeoproteomics : Advantages

Protein analyses are fast, high-throughput and at low(er) cost

Peptide mass fingerprinting

Anal. Chem. **2002,** *74,* 5960−5968

Identification and Quantification of Feathers, Down, and Hair of Avian and Mammalian Origin Using Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry

Klaus Hollemeyer,*,† Wolfgang Altmeyer,‡ and Elmar Heinzle†

Available online at www.sciencedirect.com
ScienceDirect
Analytical Biochemistry 374 (2008) 325–334
ANALYTICAL BIOCHEMISTRY
www.elsevier.com/locate/yabio

A method of isolating the collagen (I) α2 chain carboxytelopeptide for species identification in bone fragments

Michael Buckley [a,*], Matthew Collins [b], Jane Thomas-Oates [c]

[a] *Department of Biology, University of York, York YO10 5YW, UK*
[b] *Department of Archaeology, University of York, York YO1 7EP, UK*
[c] *Department of Chemistry, University of York, York YO10 5DD, UK*

Received 20 September 2007
Available online 27 December 2007

Abstract

We present a novel method for the isolation and analysis of the bone collagen (I) α2 chain carboxytelopeptide as a species biomarker. Conventional methods for the analysis and sequencing of mixtures of proteins and peptides commonly involve using the protease trypsin to cleave proteins present in the sample. However, in the study of collagen, these methods result in very complex mixtures of peptides that are difficult to analyze and the acquired results are not reproducible. Here we use bacterial collagenase (from *Clostridium histolyticum*) for its ability to cleave the highly unusual Gly-Xaa-Yaa repeating sequence of collagen, where Xaa usually is Pro and Yaa often is Hyp. Followed by a simple isolation step using a reverse phase solid phase extraction cartridge, the α2 (I) chain carboxytelopeptide can be readily analyzed by matrix-assisted laser desorption/ionization–mass spectrometry (MALDI–MS) and the results can be used to distinguish between different species of origin.

Keywords: Bacterial collagenase; Collagen; Telopeptide; Species variability; Meat and bone meal; Archaeology; Gelatin; Mass spectrometry

Buckley et al. 2008

RAPID COMMUNICATIONS IN MASS SPECTROMETRY
Rapid Commun. Mass Spectrom. 2009; **23**: 3843–3854
Published online in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/rcm.4316
RCM

Species identification by analysis of bone collagen using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry

Michael Buckley[1]*, Matthew Collins[2], Jane Thomas-Oates[3] and Julie C. Wilson[4]

[1] Department of Biology, University of York, Heslington, York YO10 5YW, UK
[2] Department of Archaeology, University of York, Heslington, York YO10 5YW, UK
[3] Centre of Excellence in Mass Spectrometry, Department of Chemistry, University of York, Heslington, York, YO10 5YW, UK
[4] York Centre for Complex Systems Analysis, Departments of Mathematics and Chemistry, University of York, Heslington, York, YO10 5YW, UK

Received 2 August 2009; Revised 23 September 2009; Accepted 29 September 2009

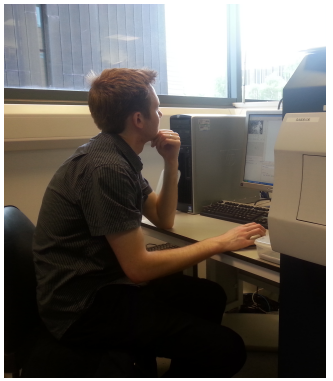
Species identification of fragmentary bone, such as in rendered meat and bone meal or from archaeological sites, is often difficult in the absence of clear morphological markers. Here we present a robust method of analysing genus-specific collagen peptides by mass spectrometry simply by using solid-phase extraction (a C18 ZipTip®) for peptide purification, rather than liquid chromatography/mass spectrometry (LC/MS). Analysis of the collagen from 32 different mammal species identified a total of 92 peptide markers that could be used for species identification, for example, in processed food and animal feed. A set of ancient (>100 ka@10°C) bone samples was also analysed to show that the proposed method has applications to archaeological bone identification. Copyright © 2009 John Wiley & Sons, Ltd.

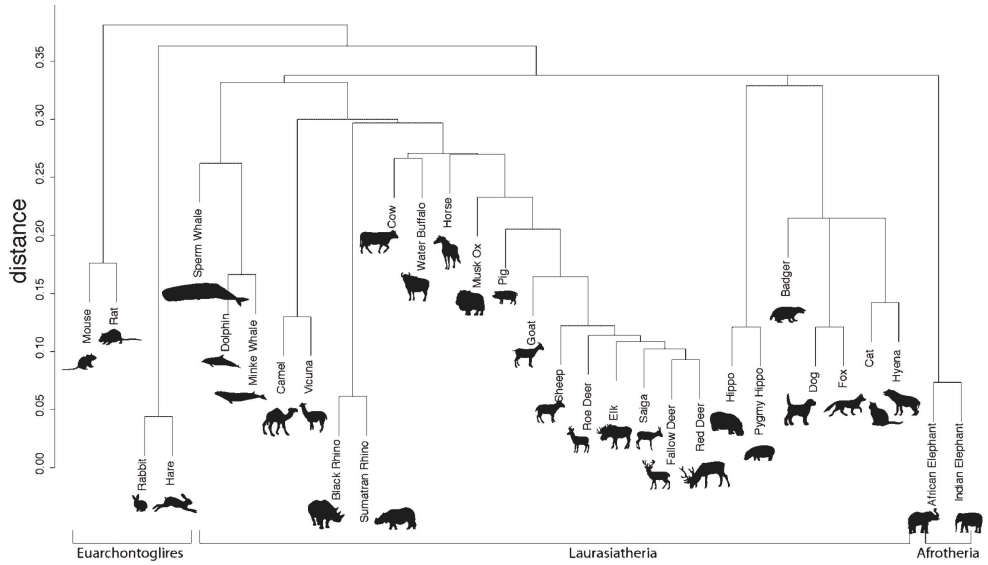
Buckley et al. 2009

Peptide mass fingerprinting

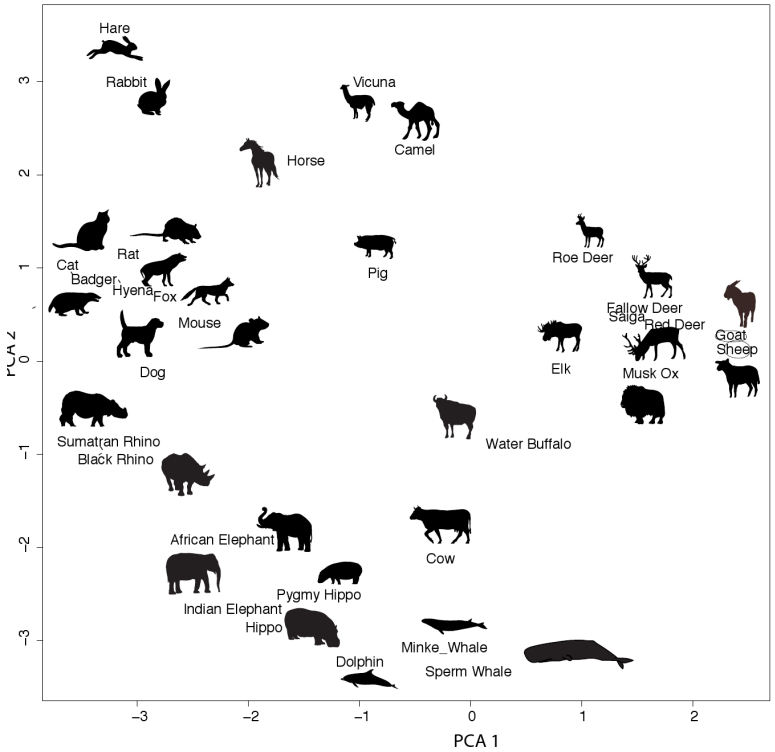
ZooMS (Zooarchaeology by Mass Spectrometry)

Taxonomic identification of collagen type I (COL1)





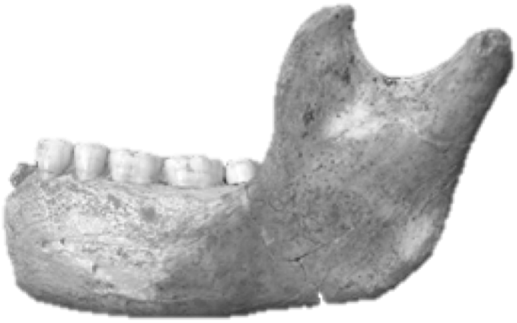
Buckley et al. 2009



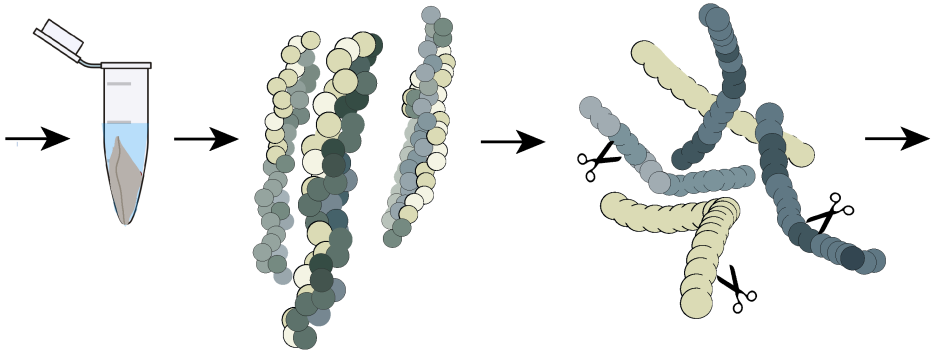
Buckley et al. 2010

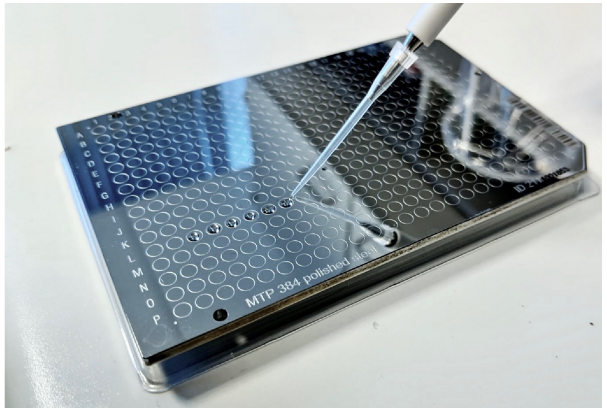
Peptide mass fingerprinting

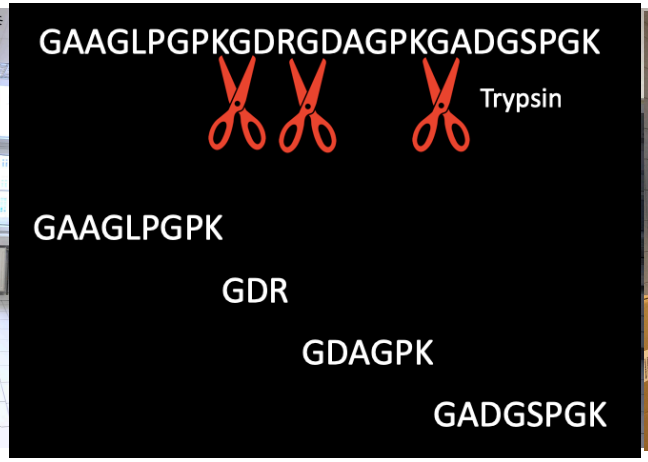
ZooMS (Zooarchaeology by Mass Spectrometry)

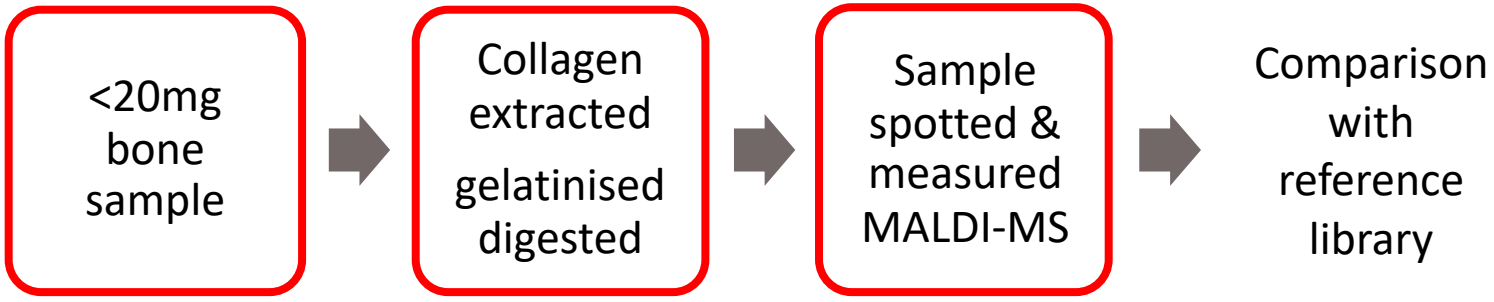


Peptide mass fingerprinting





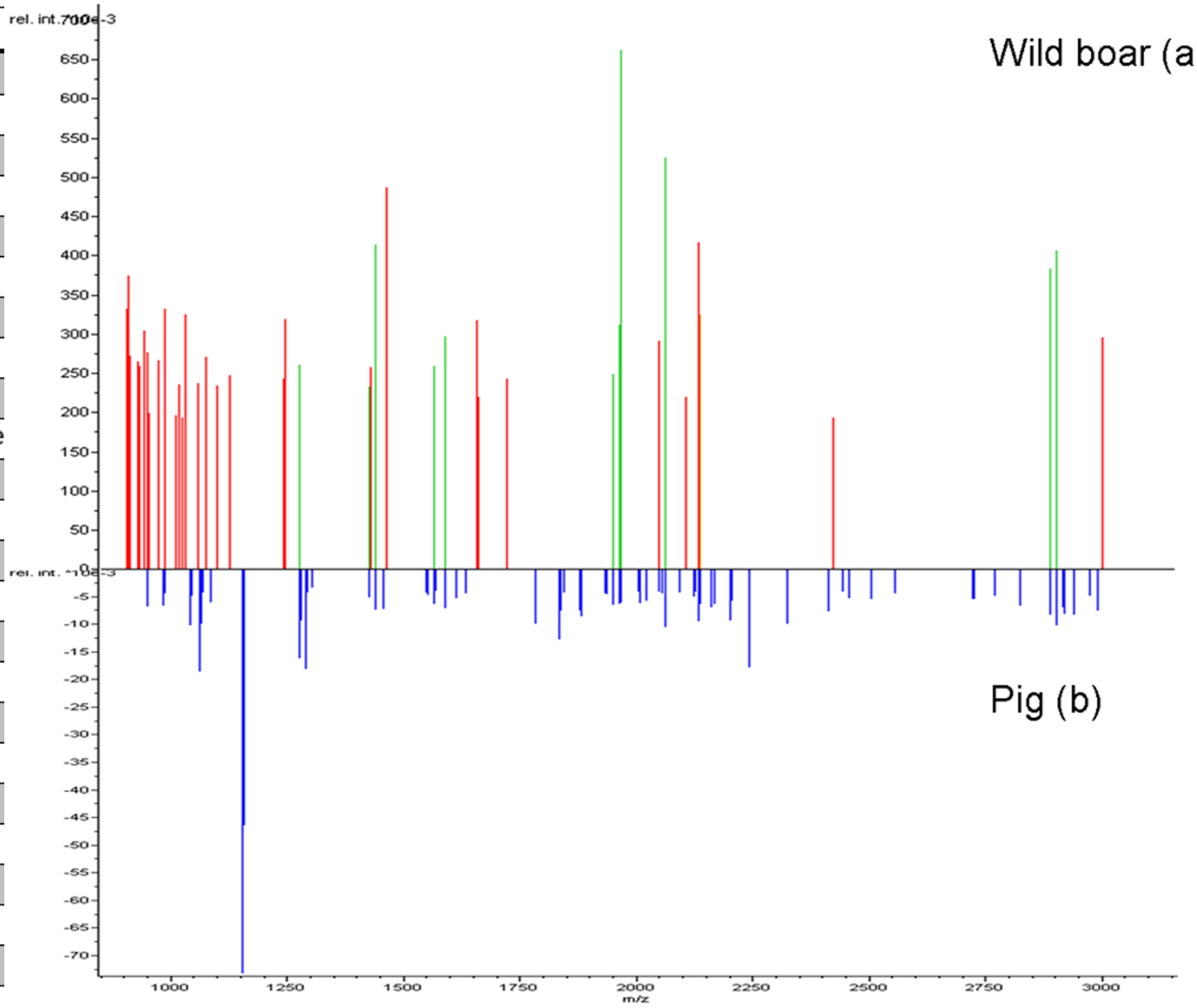


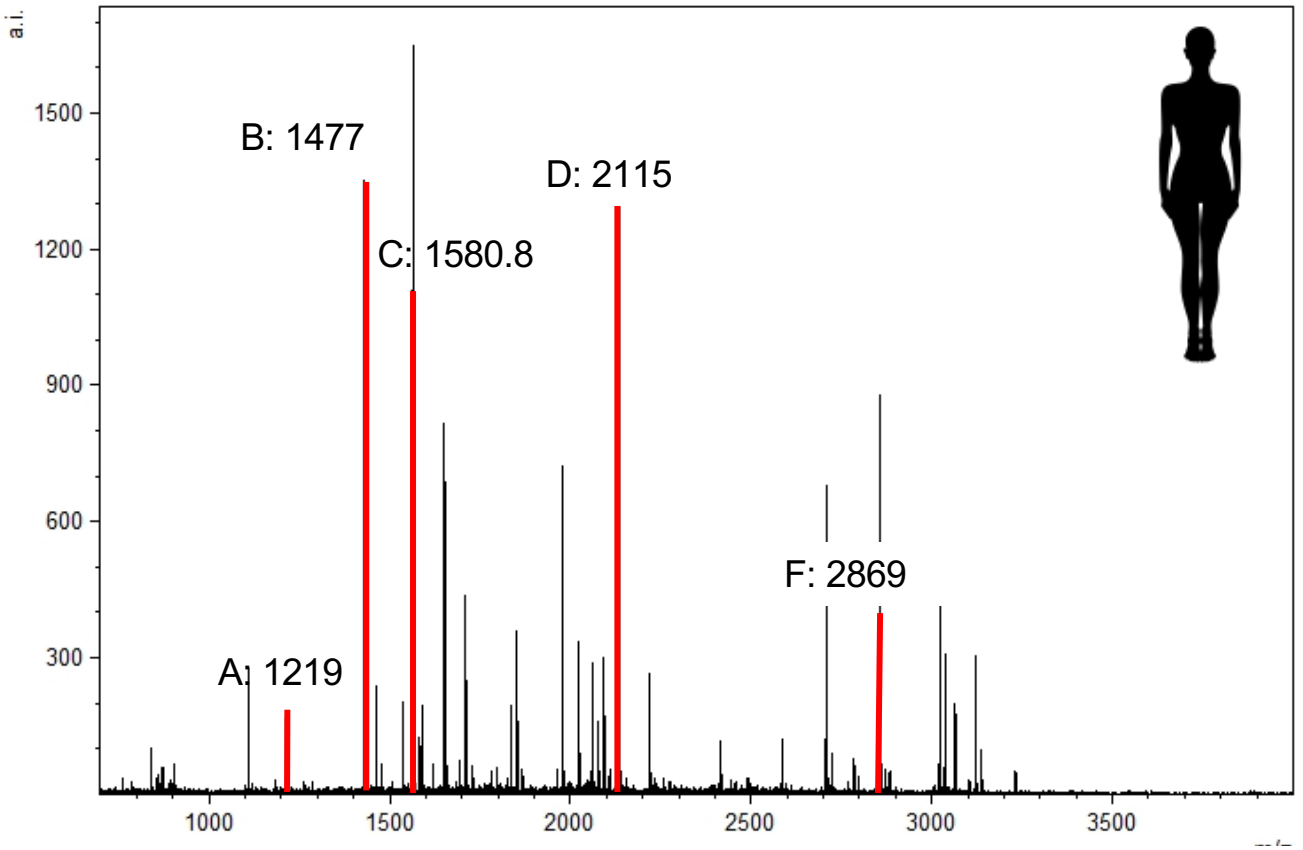


Extraction

Peptide mass fingerprinting

Species	G2
Sheep	3033.5
Goat	3093.5
Cattle	3033.5
Red fox	2999.4
Dog	2999.5
Wolf	2999.5
Black bear	2973.5
Polar bear	2973.5
Brown bear	2973.5
Caribou/Reindee	x
Roe deer	3059.4
Red deer	3033.5
Fallow deer	3033.4
Sperm whale	3039.5
Pilot whale	3023
Narwhal	3051
Mouse	2947.5
Rat	3003.5
Olive baboon	2973.5
Human	2973.5
Pig	3033.5
Common zebra	2999.4
Horse	2999.5





<20mg bone sample → Collagen extracted and digested → Sample measured MALDI-TOF MS → Comparison with reference library

ZooMS : applications

Domesticated taxa, sheep (*O. aries*) vs goat (*C. hircus*)

Journal of Archaeological Science 37 (2010) 13–20

Contents lists available at ScienceDirect

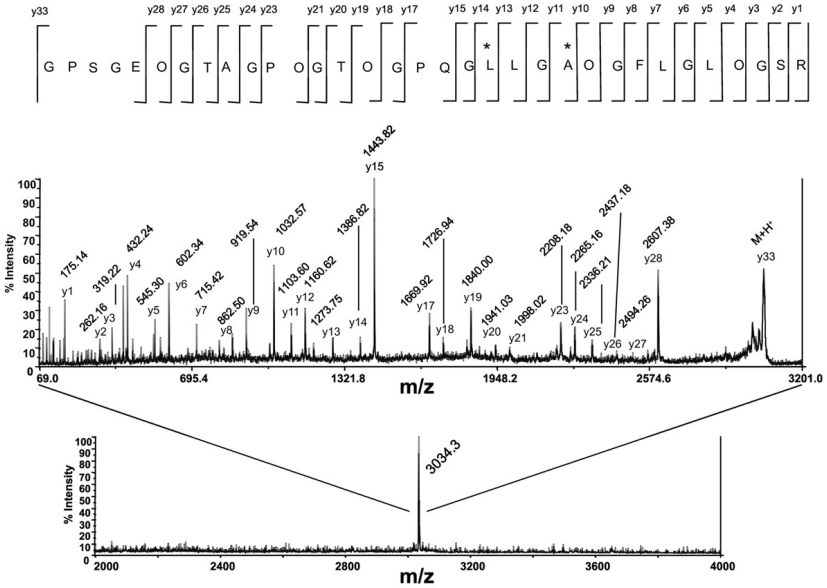
Journal of Archaeological Science

journal homepage: http://www.elsevier.com/locate/jas

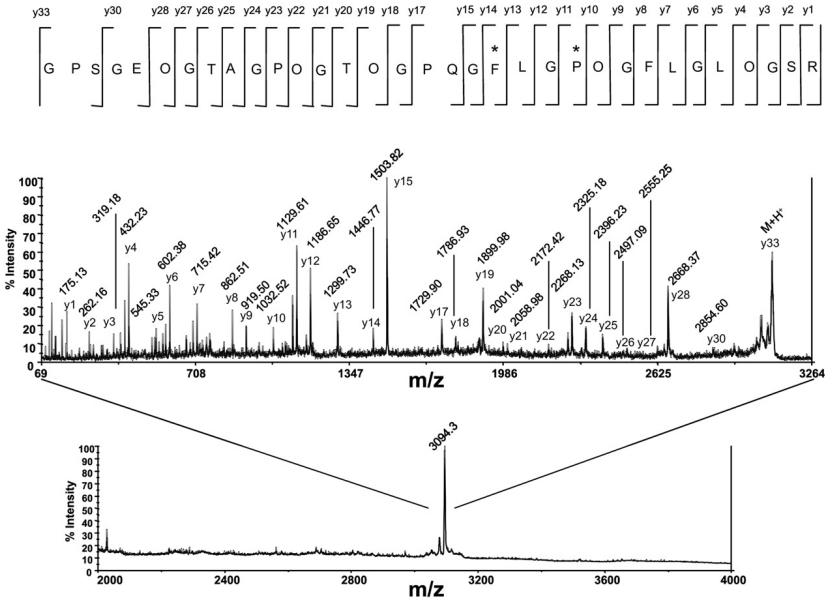
Distinguishing between archaeological sheep and goat bones using a single collagen peptide

Mike Buckley [a,*], Sarah Whitcher Kansa [b], Sarah Howard [c], Stuart Campbell [c], Jane Thomas-Oates [d], Matthew Collins [e]

Buckley et al. 2010



Sheep



Goat

ZooMS : applications

Ötzi the Iceman

Research Article

Rapid Communications in Mass Spectrometry

Received: 12 March 2012 Revised: 11 May 2012 Accepted: 11 May 2012 Published online in Wiley Online Library

Rapid Commun. Mass Spectrom. **2012**, *26*, 1735–1745
(wileyonlinelibrary.com) DOI: 10.1002/rcm.6277

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry combined with multidimensional scaling, binary hierarchical cluster tree and selected diagnostic masses improves species identification of Neolithic keratin sequences from furs of the Tyrolean Iceman Oetzi

Klaus Hollemeyer[1]*, Wolfgang Altmeyer[2], Elmar Heinzle[1] and Christian Pitra[3]

~5500 years ago (3300-3100 BC)

Dog/red fox

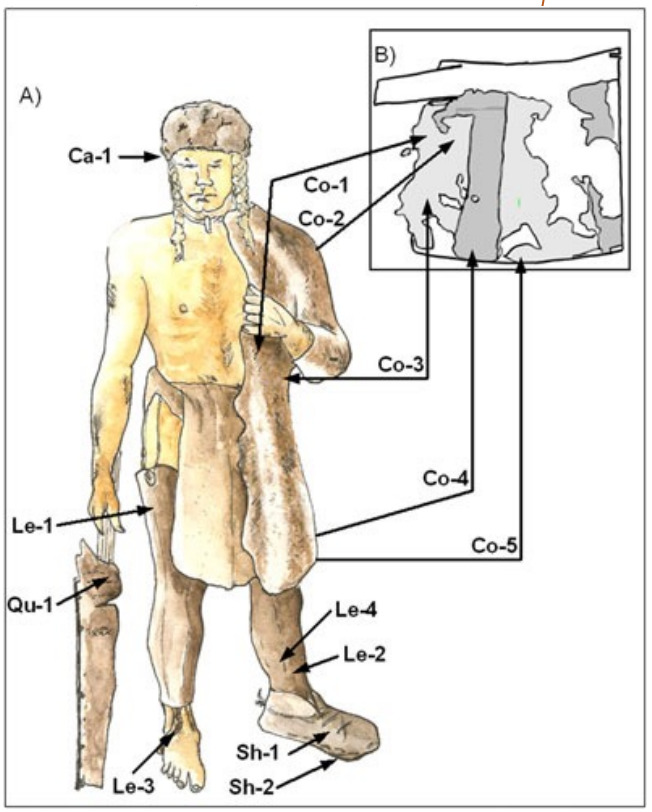
Goat

Red deer

Cattle

Brown bear/canid

Sheep



Hollemeyer et al. 2012

80 pieces of leather/ 13 identified/ 6 species (wild & domesticated)

ZooMS : applications

Prehistoric bone tools from Morocco

Journal of Archaeological Science 98 (2018) 149–157

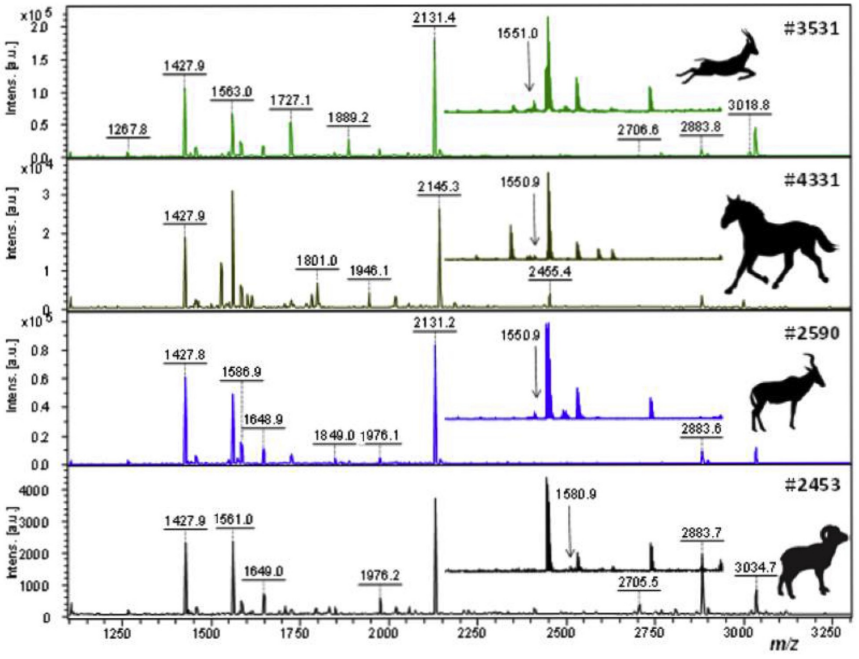
Contents lists available at ScienceDirect

Journal of Archaeological Science

journal homepage: www.elsevier.com/locate/jas

ZooMS identification of bone tools from the North African Later Stone Age

Abigail Desmond[a,*], Nick Barton[a], Abdeljalil Bouzouggar[b], Katerina Douka[c,d], Philippe Fernandez[e], Louise Humphrey[f], Jacob Morales[g], Elaine Turner[h], Michael Buckley[i]



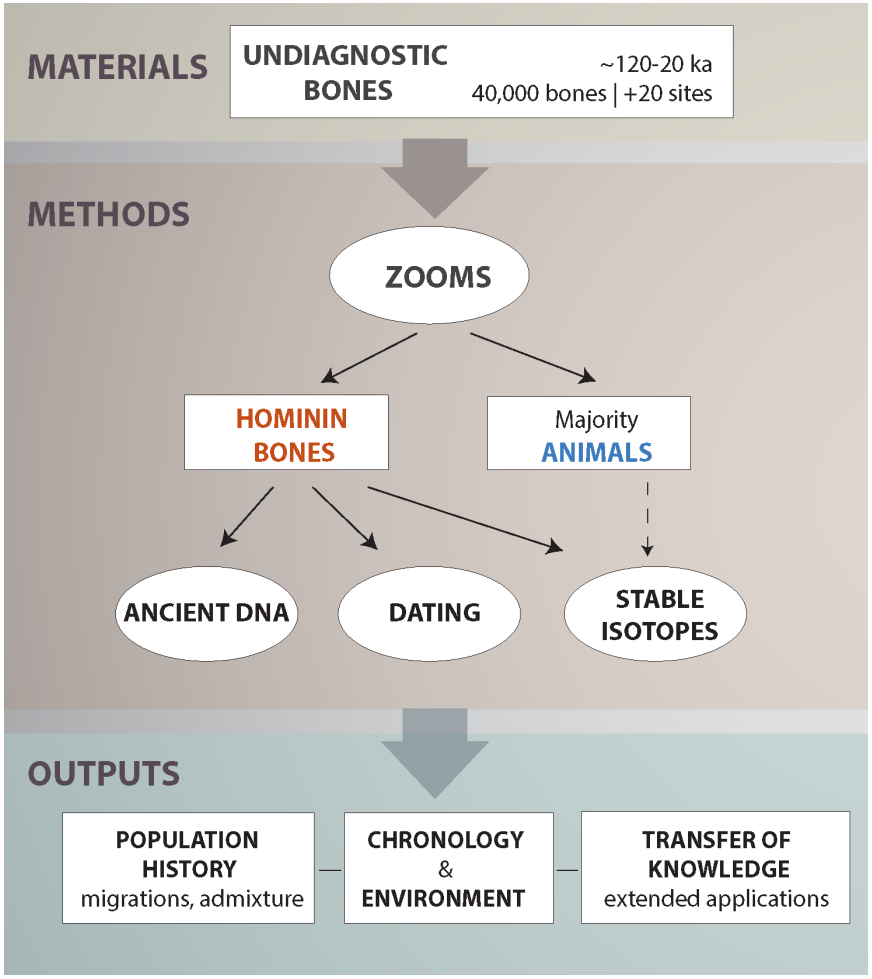
Desmond et al. 2018

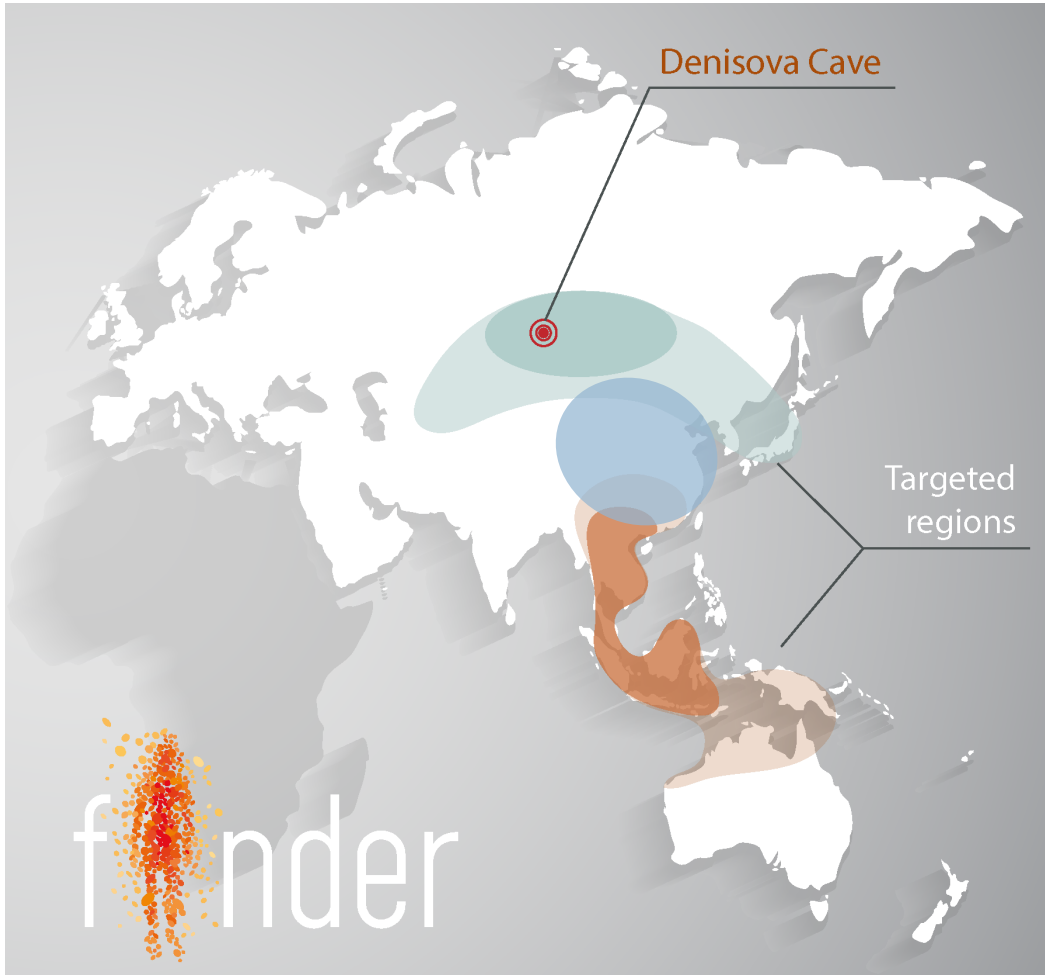
Detecting humans in the prehistoric record?

Material culture = human biology ?

Unidentified bone fragments

FINDER Project 2017-2022





20,000+ bone samples, >20 Pleistocene sites, 6 countries

Scientific methodologies

AMS
radiocarbon dating

Peptide mass
fingerprinting
(ZooMS)

Mass spectrometry-
Stable isotopes

Next generation
sequencing (aDNA)

FINDER

D-N

Neanderthal

Homo sp.

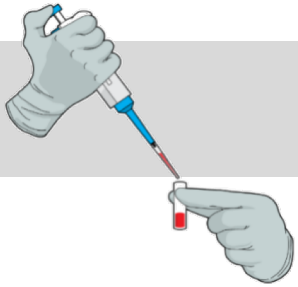
Conclusions

- Ancient proteomics is a fast developing field
- ZooMS is a rapid and reliable method of identifying fragmented bone to taxon;
- FINDER aims to reveal the spatial distribution of archaic hominins in the late Pleistocene, particularly in eastern Eurasia; we have identified 8 hominin bone fragments from Denisova using this method, based on screening of >9000 samples;
- Many archaeological assemblages waiting to be analysed.

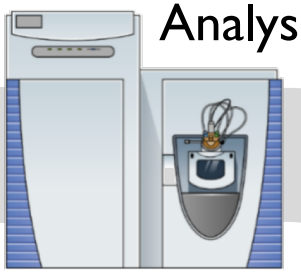
‘Shotgun’ Proteomics

Sampling

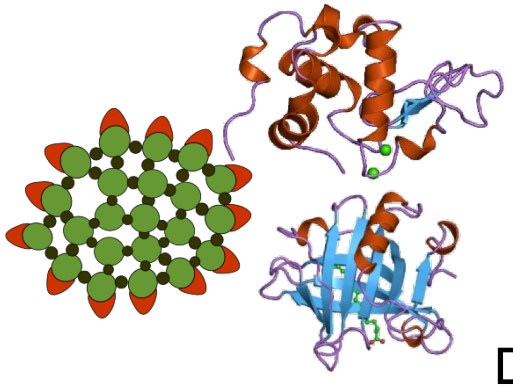
Protein Extraction



LC-MS/MS Analysis



Bioinformatics (Mascot)



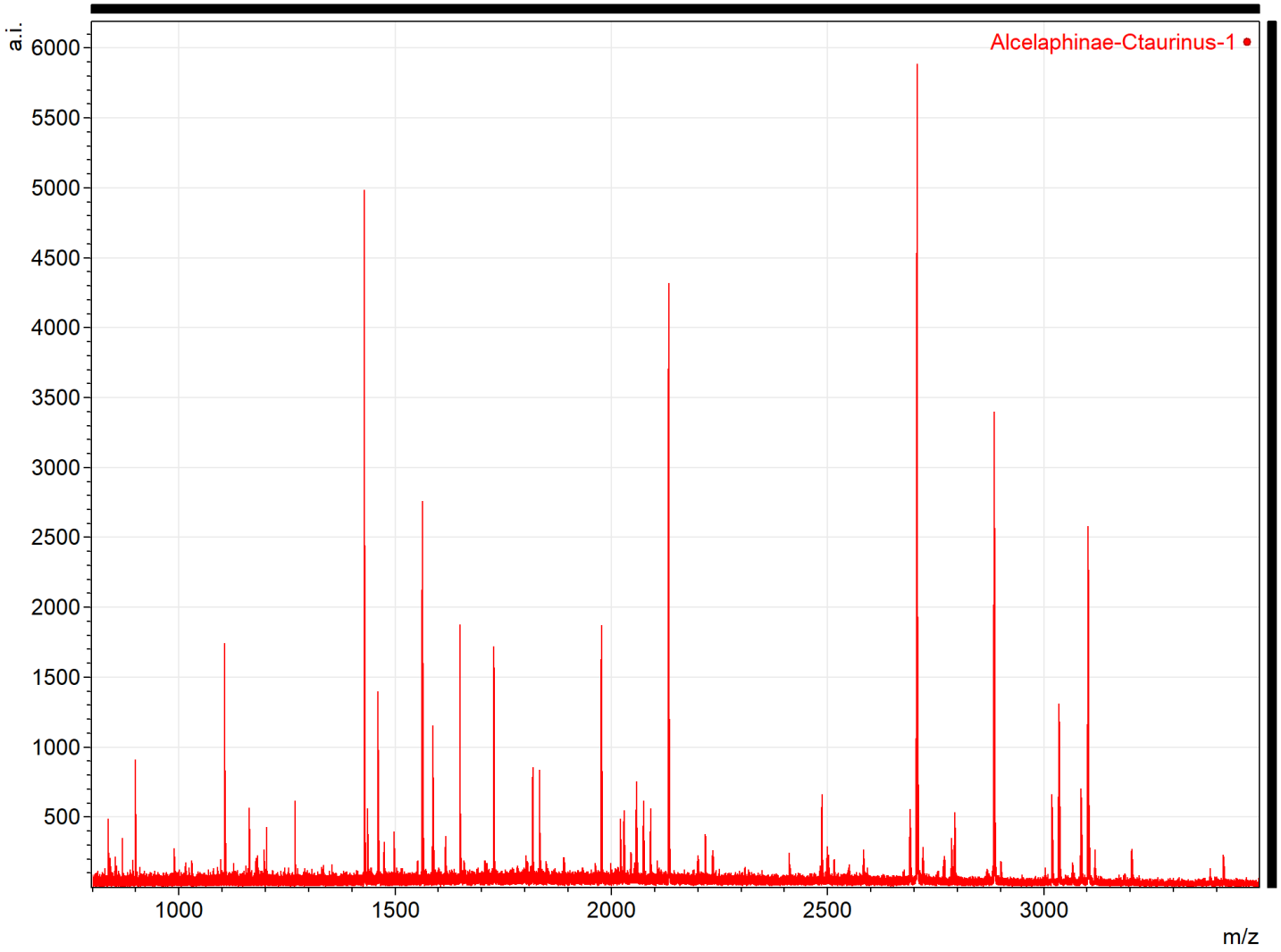
Goat
TPEVDKEALEKFD

Sheep
TPEVDNEALEKFD

Data Analysis and Interpretation

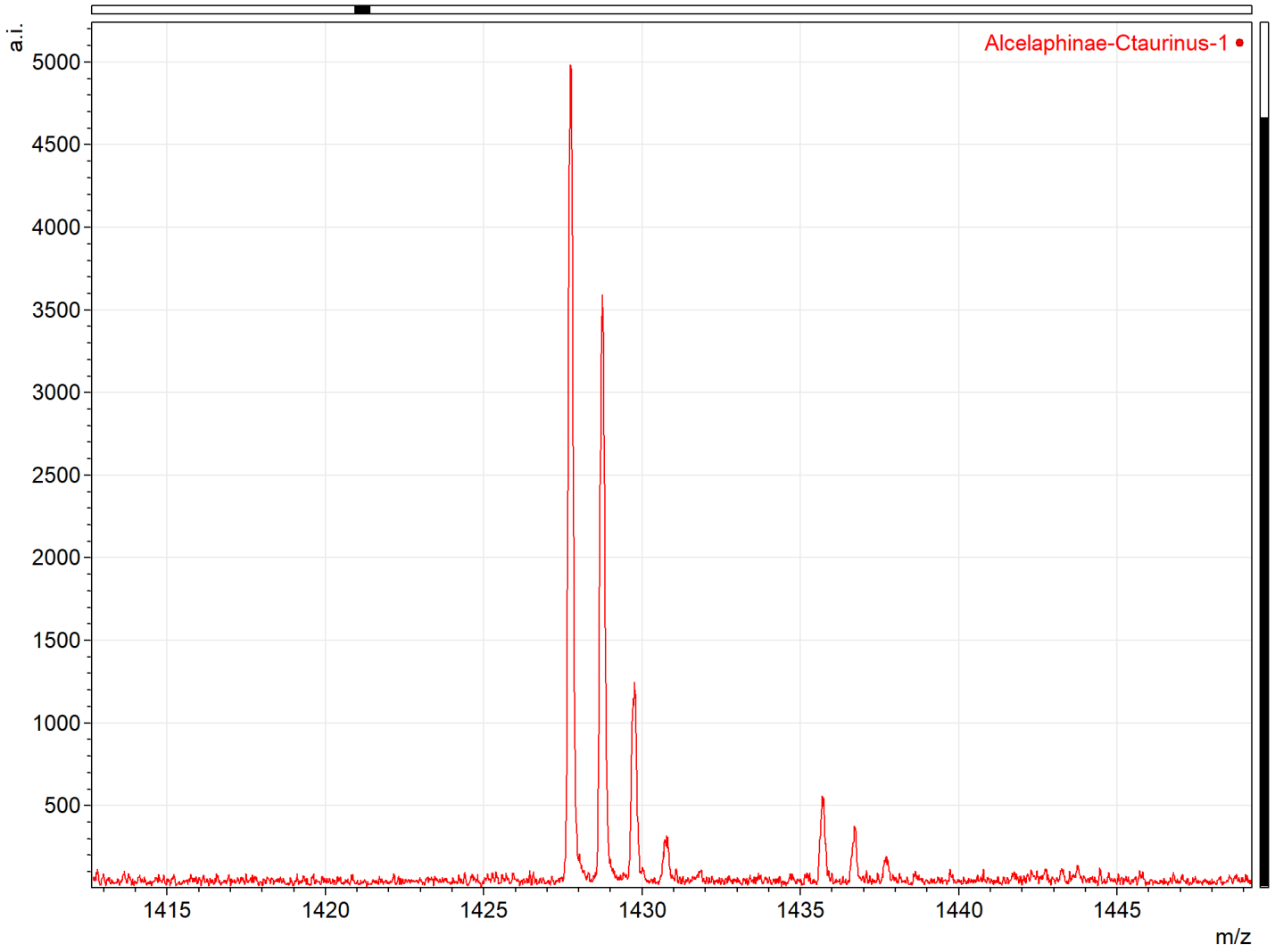
MALDI MS1

MALDI

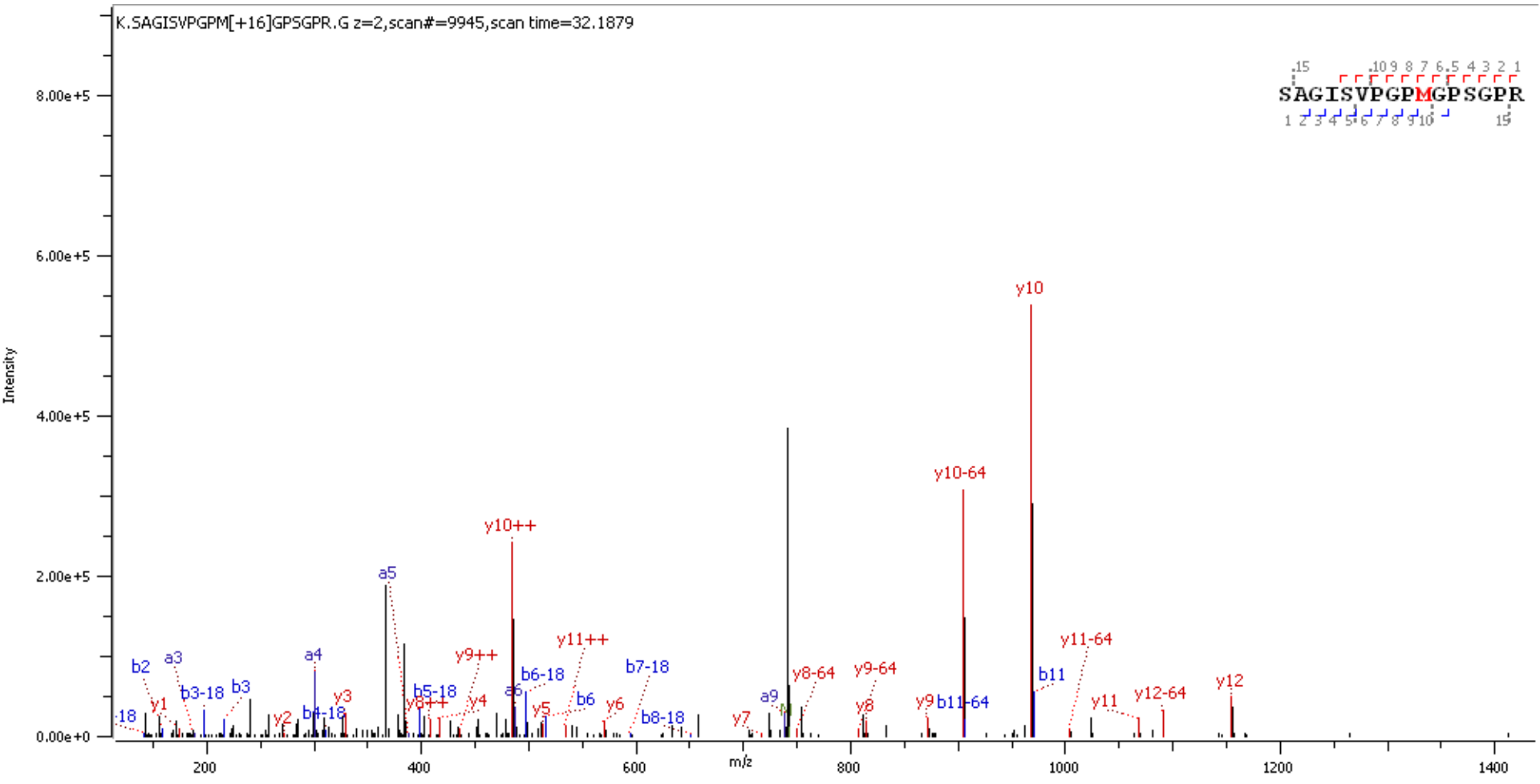


MALDI MS1

MALDI



MALDI MS2


K.SAGISVPGPM[+16]GPSGPR.G z=2,scan#=9945,scan time=32.1879
SAGISVPGPMGPSGPR
Intensity
8.00e+5
6.00e+5
4.00e+5
2.00e+5
0.00e+0
m/z
200
400
600
800
1000
1200
1400
y10
y10-64
y10++
a5
y11++
y9++
b2
a3
a4
b3-18
b3
b5-18
b6-18
b7-18
y1
y3
y4
y5
b6
y6
b8-18
y7
a9
y8-64
y9-64
y8
y9
b11-64
b11
y11-64
y11
y12-64
y12

MALDI MS1 - MS2

MS1

MS2