



以胶原蛋白为例浅述古蛋白质组学研究现状与未来*

李鑫 泮燕红**

南京大学, 地球科学与工程学院, 关键地球物质循环前沿科学中心, 南京 210023

提要 蛋白质作为参与构建生物体的重要生物大分子, 是生物功能与代谢的物质基础, 同时蛋白质的序列信息又源自生物的遗传编码信息, 因此是认识生物演化本质的重要研究对象。近年来, 随着质谱技术的发展, 获取生物化石中的古蛋白质序列信息不再遥不可及, 这在形态学与DNA序列信息缺乏的条件下对古生物的认识提供了一条新途径。胶原蛋白(Collagen)在动物骨骼中极为丰富, 又因为其特殊的结构, 易于在化石中保存, 故已成为古蛋白质组学研究的重要对象。本文将以胶原蛋白为例, 对古蛋白质组学现有的研究方法与已经取得的研究成果进行总结, 并对古蛋白质组学目前面临的挑战与困难及未来研究趋势进行讨论, 旨在展示古蛋白质组学的应用潜力, 并探讨其在生物演化研究中的意义。

关键词 古蛋白质组学 胶原蛋白 生物演化 ZooMS 鸟枪法

中文引用 李鑫, 泮燕红, 2023. 以胶原蛋白为例浅述古蛋白质组学研究现状与未来. 古生物学报, 62(2): 321–332. DOI: 10.19800/j.cnki.aps.2021065

英文引用 Li Xin, Pan Yan-hong, 2023. Review and prospects of paleoproteomics: using collagen as an example. Acta Palaeontologica Sinica, 62(2): 321–332. DOI: 10.19800/j.cnki.aps.2021065

Review and prospects of paleoproteomics: using collagen as an example

LI Xin, PAN Yan-hong

School of Earth Sciences and Engineering, Frontiers Science Center for Critical Earth Material Cycling, Nanjing University, Nanjing 210023, China

Abstract Proteins are macromolecules and they are the basis of many biological functions and metabolism. The sequence of the proteins is decided by genetic code. Therefore, proteins provide us unique information on evolution. Recent development in mass spectrometry technology makes it possible to obtain ancient protein sequence information from fossils, providing a new way to decipher the evolutionary secrets of animals from fossils when relevant morphological and DNA information is not available. Collagen is a widely distributed protein in bones. It has great preservation potential in fossils due to its special structure and has become a key research object in paleoproteomics. Using collagen as an example, this paper summarizes the research methods and research achievements of paleoproteomics and discusses the current challenges and future trends of this subject to show the application potential and research significance in paleontology.

Key words paleoproteomics, collagen protein, biological evolution, ZooMS, shotgun method

收稿日期: 2021-11-12; 改回日期: 2023-04-10; 录用日期: 2023-04-13

* 国家自然科学基金(41922011, 41872016)、中国科学院战略性先导科技专项(B类)(XDB26000000)和中央高校基本科研业务费专项资金(0206-14380137)联合资助。

** 通讯作者: 泮燕红, 教授, e-mail: panyanhong@nju.edu.cn

1 前 言

蛋白质是构成生物体结构的主要成分,是生物功能和代谢活动的物质基础,其序列包含着丰富的遗传信息。与DNA分子相比,蛋白质凭借其相对稳定的构象,在地质条件下具有更好的保存潜力。例如,目前被普遍承认的最早的古蛋白质材料距今约380万年,相较于目前已发现的最早的古DNA记录(约165万年)要早上许多(Demarchi *et al.*, 2016; van der Valk *et al.*, 2021)。因此,化石中遗存的蛋白质信息越来越受到关注,尤其在古DNA材料缺失的情况下,古蛋白质信息的作用与意义更加显著(Cappellini *et al.*, 2018)。学界对古蛋白质的研究与探索可以追溯到20世纪50年代,当时的研究人员发现化石中可能含有古氨基酸遗存(Abelson, 1954; Hare, 1967)。但直到20世纪70年代,研究人员才首次尝试从浮游有孔虫化石中获取其古蛋白序列信息,不过该实验结果一直被质疑可能受外源性污染影响(Schroeder and Bada, 1976)。在此之后的很长一段时间内,受限于分析技术,古蛋白质序列分析研究一直停滞不前。直到21世纪初,随着质谱技术的迅猛发展,才使得化石中非常稀少的古蛋白质的序列解析有了突破性进展。以Ostrom等(2000)获得晚更新世野牛化石中的骨钙素序列为标志,古蛋白质组学(Paleoproteomics)这一新兴研究方向正式诞生。古蛋白质组学以古蛋白质的肽段结构、氨基酸序列、单氨基酸多态性(Single amino acid polymorphism, SAPs, 即单一位点氨基酸的突变)以及翻译后修饰为主要研究对象,旨在恢复古代生物的遗传信息,并为进一步理解该生物的古生态、古生理与系统发育提供分子证据(Welker, 2018a; Hendy, 2021)。

由于结构的稳定性与在地质时间尺度下的良好保存潜力,胶原蛋白成为了当下古蛋白组学研究中的重要材料。Buckley等(2009)开创了动物考古质谱分析(Zooarchaeology by mass spectrometry screening, ZooMS)技术,并首次应用I型胶原蛋白对化石属种成功进行了分类。Rybczynski等(2013)分离出了约350万年的骆驼化石中I型胶原蛋白片段,并对其进行了肽质量指纹图谱(Peptide mass

fingerprinting, PMF)分析,展示了在地质历史尺度下I型胶原蛋白序列信息的保存潜力。之后,随着色谱-串联质谱(Liquid chromatography tandem-mass spectrometry, LC-MS/MS)技术的引入,鸟枪法(Shotgun method)使得胶原蛋白自下而上(bottom up)的测序成为了可能,一系列以胶原蛋白为材料的研究成果相继出炉,古蛋白质组学为古生物化石的研究开辟了新的视野(Cappellini *et al.*, 2018)。本文以胶原蛋白为例,从以下四个方面对古蛋白质组学研究展开综述: 1) 简要介绍胶原蛋白的结构、性质; 2) 列举古蛋白组学常用的分析方法; 3) 列举古蛋白质组学在古生物学与动物考古学中的应用; 4) 对古蛋白质组学研究的展望。

2 胶原蛋白概述

胶原蛋白是一类纤维状的蛋白质家族,是动物体内主要的结构蛋白,广泛参与到了皮肤、血管、软骨、骨骼等组织与器官的构建,是动物进行机械运动的重要基础。根据氨基酸及含糖量的不同,胶原蛋白被分为I型、II型、III型等数种。在许多动物体内,胶原蛋白是丰度最高的蛋白质类型。例如在哺乳动物体内,胶原蛋白的占比高达蛋白总量的20%–35%,尤其在骨骼组织中可达骨骼总有机质的80% (张达江、王亮, 2006; 王镜岩等, 2008)。在结构上,胶原蛋白拥有三条 α 肽链,其相互缠绕构成特殊的三螺旋结构,这赋予胶原蛋白良好的稳定性和抗溶解性,使其不溶于冷水和稀酸、稀碱溶液,并且能够耐受除胶原蛋白酶(Collagenase)之外的大多数蛋白酶的酶解作用。在动物骨骼中,最主要胶原蛋白类型为I型胶原蛋白,其含有大量甘氨酸(Glycine, Gly)和羟脯氨酸(Hydroxyproline, Hyp),高丰度的Gly和Hyp能够促进氢键和分子间形成交联,维持多肽链之间的紧密连接,有利于保持空间结构的稳定。并且在三螺旋结构中, Hyp的 γ -OH与水分子以及多肽链的酰胺基可产生架桥作用,因此高Hyp含量的胶原蛋白在高温下也能保持相对稳定的构象(Miles *et al.*, 1998)。此外,在骨组织的形成过程中, I型胶原蛋白还获得了一定程度的矿化,从而产生了紧密的层级结构,这使得它拥有比其他结缔组织中的胶原

蛋白更好的热稳定性(Trębacz and Wójtowicz, 2005)。鉴于含量的丰富度与性质的稳定性, 与其他蛋白相比, 胶原蛋白显示出更强的保存潜力, 因此胶原蛋白被视为分子古生物学研究中一类重要的材料(Buckley *et al.*, 2008b)。

3 古蛋白质组学常用分析方法

3.1 预筛选

由于化石材料本身的不可再生性与珍贵性, 在进行有损的质谱分析之前, 需要对可能保存了胶原蛋白的化石材料进行无损或低损的预筛选, 旨在挑选出胶原蛋白保存程度较好的材料进行后续的组学分析, 以此提高研究效率, 减少材料的损耗。常见的预筛选方法包括以下4种。

3.1.1 红外光谱法

红外光谱(Infrared Spectroscopy, IR)是一类吸收光谱, 也是蛋白质结构解析的基本方法。鉴于红外光谱检测具有高效、快速、对样品要求低或者样品检测环境要求低等优点, 非常适用于化石材料的在前期研究的大规模筛选与简单成分研究(Sponheimer *et al.*, 2019)。

衰减全反射-傅里叶变换红外光谱(Attenuated total reflection-Fourier transform infrared spectroscopy, ATR-FTIR)可以通过获取酰胺峰的强度直接反应样品中蛋白质丰度; 同时还可以获取碳酸盐-磷酸盐的比值(C/P)来评估化石材料中成岩矿化水平, 从侧面揭示蛋白质的保存状态(Presslee *et al.*, 2021)。Rao等(2020)利用ATR-FTIR对东亚更新世斑鬣狗(*Crocota crocuta*)的骨骼化石进行胶原蛋白的初期筛选, 并成功筛选得到了胶原蛋白保存状态良好的样品。近红外(near-infrared, NIR)光谱则兼具良好的灵敏度与检测效率, 还可用于野外环境的快速检测。如Sponheimer等(2019)就利用NIR对大量化石样品进行过快速筛选, 在几秒钟到几分钟的时间内, NIR便可以分离出胶原蛋白保存含量>1%的单个样品。这些研究都展示了红外光谱法在含胶原蛋白样品初期筛选中的应用潜力。

3.1.2 拉曼光谱法

拉曼光谱(Raman spectra)是一类散射光谱, 由于其可以识别红外不敏感的对称性振动引起的极化率变化, 故而可以识别红外光谱难以分辨的材料, 因此常用于红外光谱的互补研究。拉曼光谱分析具有微区、原位、无损的特点, 因此被广泛应用于生物材料与地球科学研究之中, 并且对于样品条件要求低, 无需复杂制备, 在样品含水或样品量极少的情况下也可以使用。与红外光谱相似, 拉曼光谱也可以识别出古胶原蛋白的酰胺峰, 例如Pestle等(2014)运用手持1030 nm激光激发拉曼光谱对骨骼样品进行检测, 成功识别出保存了胶原蛋白的样品, 并展示了拉曼光谱在野外环境下应用的潜力。

3.1.3 元素分析法(Elemental analysis)

骨骼中的氮元素主要由蛋白质提供, 其中胶原蛋白约占90% (Sillen and Parkinson, 1996)。因此, 化石骨骼中的氮元素百分比含量(%N)可以在一定程度上反映胶原蛋白保存水平。Brock等(2012)指出, 当以0.76%N作为临界值时, 有84%的概率成功筛选出胶原蛋白含量大于1%的样品。此外, 样品中碳、氮元素的莫尔比(C:N)也是重要的参考指标, 这一参数反映了外源含碳物质的污染情况, 一般来说, 当C:N<5时, 该样品保留胶原蛋白受到成岩作用与外界污染的影响较低, 检出胶原蛋白的概率较高(Tisnérat-Laborde, 2003)。元素分析法速度快、成本极低, 并且仅需要极少量的化石样品(1–10 mg)便可以进行检测, 因此是大批量样品进行早期快速筛查的重要手段。

3.1.4 氨基酸外消旋分析法(Amino acid racemization analysis)

在检出蛋白质后, 需要对蛋白质的年代进行简要评估。氨基酸外消旋分析是一类常用的蛋白质年代推测方法, 它通过高效液相色谱(High performance liquid chromatography, HPLC)将不同手性的氨基酸区分开来并测定比值, 以此判断蛋白质的保存情况。其理论基础是氨基酸在体外的外消旋作用。所谓外消旋作用, 即动物在生活状态下, 体

内的氨基酸为L型氨基酸,但在动物死亡后,这些L型氨基酸会逐渐自发地转化为D型氨基酸,直至两种手性的氨基酸数量持平,导致旋光性消失。因此,根据氨基酸外消旋程度,可以对蛋白质保存情况进行评估。例如, Bada等(1994)利用这一技术研究不同年代的虫珀中蛋白质的保存情况,并以此为基础构建了蛋白质在琥珀中降解的埋藏学模型。此外,由于同一时期的保存蛋白质在相同环境下呈现出近似的D/L比,采用这一技术还可以快速排除外源污染的蛋白,从而提高结果的可信度(Presslee *et al.*, 2021)。

3.2 古蛋白组学信息的获取

通过预筛选后,保存古蛋白的样品通过质谱法分析得到可能的蛋白序列,再用于后续属种鉴定、系统演化等研究。以下简要介绍古蛋白质组学研究中常用的3种质谱法。

3.2.1 动物考古质谱分析(ZooMS)法

ZooMS是一类主要依赖于基质辅助激光解吸电离飞行时间质谱技术的古胶原蛋白分析手段。通过识别生物材料中的物种特异性的蛋白标记来构建胶原肽质量指纹图谱(PMF), ZooMS可以在短时间内确定生物材料的物种来源(图1)。自开创以来, ZooMS凭借其高效、快速、准确的属种鉴定能力被广泛应用于骨骼碎片与化石材料的来源分析中,已成为当下古蛋白质研究的重要环节。

3.2.1.1 基质辅助激光解吸电离飞行时间质谱(MALDI-ToF MS)

基质辅助激光解吸电离(Matrix-Assisted Laser Desorption/Ionization, MALDI)技术是由Karas等(1985)发明的一类用于非挥发性样品的软电离方法。该方法的技术流程简述如下:首先,将适合的基质溶液与待分析样品的溶液混合,并将混合后的溶液点样于金属板之上。接着,迅速蒸发溶剂,使得基质与待分析样品一起形成晶体或半晶体状态。然后,用特定波长的激光照射金属板。此时板上的基质会将吸收到的激光的能量传递给待分析样品,使样品迅速气化并离子化。最终,离子化的样品会在金属板的电压驱动下进入检测器进行检

测。除了传递能量外,基质还有一定保护作用,使待分析样品能避免一定程度的激光损伤(Cleland and Schroeter, 2018)。这一技术的最大优势在于能够电离复杂的生物大分子,如蛋白质,且不产生过多碎片干扰检测。此外, MALDI技术对于样品的处理要求也不高,一些原始样品甚至可以直接进行检测,极大地减少了纯化过程中样品的丢失与污染。

MALDI技术与飞行时间(Time of Flight, ToF)质量分析器耦合使用,就构成了MALDI-ToF MS。ToF的原理为:当样品在激光照射电离化后,会在加速电场的作用下加速到一定的动能水平,随后样品会经过一段自由飞行后并抵达接收器。此时,离子的飞行时间与离子质量相关,质量越大,飞抵接收器所需要的时间就越长,通过记录飞行时间,就可以得到不同离子的质荷比(Mass-to-charge ratio),从而实现样品的分离。由于MALDI-ToF MS检测速度快,所需的骨骼样品量低(15–30 mg),适用于生物大分子,所以被视为研究古蛋白质的理想手段(Buckley *et al.*, 2009; Presslee *et al.*, 2021)。

3.2.1.2 胶原肽质量指纹图谱(Collagen peptide mass fingerprinting, PMF)

PMF是ZooMS法进行蛋白质种类鉴定的关键步骤,其构建过程大致如下:首先,在采集到古蛋白质样品后,会使用特定的蛋白酶进行消化,这一步目的在于获得具有已知特定位点切口的肽片段,便于后续比对。酶解完成后,再对得到的肽片段进行MALDI-ToF MS分析,获得对应的质荷比值。由于同源胶原蛋白在演化中的突变,导致胶原蛋白在不同的属种之间存在一定程度的氨基酸序列差异,这种差异能反映为质谱结果中质荷比的差异。因此,当获得某一化石材料中胶原蛋白的质荷比后,便可以通过比对样品与数据库中其他动物胶原蛋白的质荷比来判断样品所属的生物属种。PMF用于属种鉴定的优势在于其速度快,可信度较高,因此多被用于破碎骨骼或组织的鉴定,尤其是新生代哺乳动物以及古人类的碎骨(Brown *et al.*, 2016)。此外,由于能够解析出蛋白质氨基酸序列信息, PMF在一定程度上还能揭示演化关系,为构建演化序列提供参考(Buckley *et al.*, 2019b)。

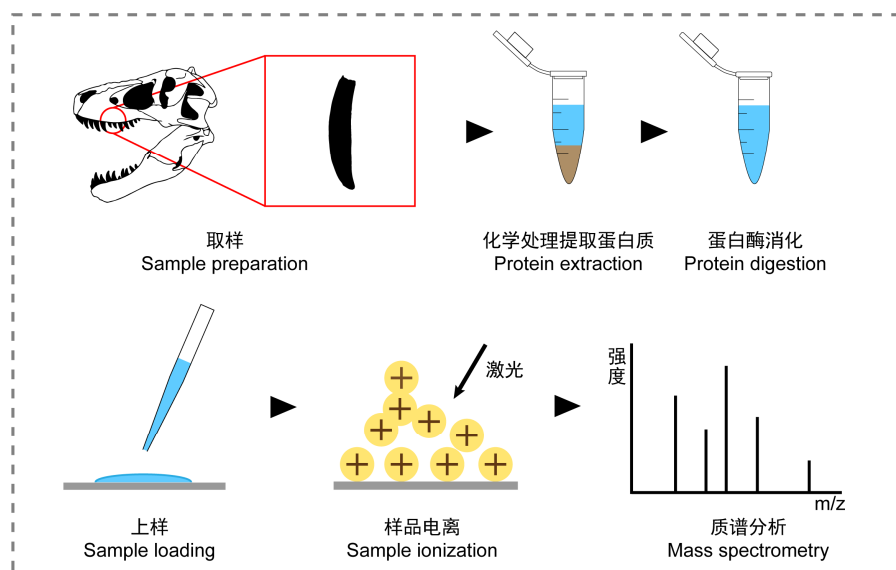


图1 ZooMS 流程示意图

Fig. 1 Flow chart of ZooMS

3.2.2 鸟枪法

鸟枪法(Shotgun method)是利用酶切将生物大分子剪切为小片段, 再将这些短片段单独进行分析, 最后再拼接起来进行序列信息读取的方法。因为这一过程类似于用霰弹枪打鸟, 故而得名。不同于ZooMS, 鸟枪法一次性可以解析成百上千种的蛋白质, 还能识别蛋白质序列翻译后修饰(Welker, 2018a)。

色谱-串联质谱(LC-MS/MS)是目前用于蛋白质鸟枪法测序的主要技术, 因为实验过程是从读取蛋白质碎片信息再到恢复蛋白质整体信息, 故而又被称作自下而上(bottom up)的蛋白质测序方法(图2)。LC-MS/MS分为两个部分: 前段的LC用于材料的分离与纯化, 而后段的MS/MS主要用于纯化后样品的质荷比分析。

LC-MS/MS的技术流程大致如下: 首先, 样品溶液需要经过物理化学方法或酶解法处理, 以获得蛋白质断裂后的肽段; 之后将肽段溶液上样于LC反相色谱柱中进行分离; 最后通过电喷雾电离(Electrospray ionization, ESI)使得材料有顺序地进入质谱仪内进行分析。LC目的就是不同物质进行分离, 因此LC-MS/MS不像ZooMS一样需要胶原蛋白的纯化过程, 可以直接对生物的整体组织进行分析。在进入MS/MS后, 电离样品便可以

进行质荷比分析, 但不同于普通的MS分析, MS/MS加入了解离过程, 在完成初次质荷比分析后, 电离后的肽段需要通过解离过程被破碎(常用解离方法有碰撞诱导解离、电子转移解离、电子捕获解离等), 进而获得更小的肽段碎片。通过一次或多次的碎裂, 样品最后呈现为一系列单位大小不同的肽片段组合, 而每一次碰撞结果都会用于质荷比分析, 从而得到一系列的蛋白质序列信息。通过这些序列信息的相互比对、拼接以及蛋白质序列的数据库检索, 研究人员便可以快速鉴定出混合样本中蛋白质的种类, 并且可以获得完整的蛋白序列用于后续研究(Welker, 2018a)。相较于ZooMS法, 鸟枪法适用范围更广, 分析能力更强, 目的更倾向于下游的蛋白结构解析与生物演化研究(图3)。

3.2.3 自上而下(top-down)的蛋白质测序方法

因为自下而上的蛋白质测序方法(LC-MS/MS)会在酶解过程中破坏蛋白质的空间结构, 同时也在一定程度上带来了人为修饰的风险, 如氨基酸的脱酰胺作用(Li *et al.*, 2008)。因此, 自上而下(top-down)的蛋白质测序方法有望被引入古蛋白质研究中。自上而下的蛋白质测序方法跳过了酶解步骤, 直接将提取到的蛋白质进行电离化处理,

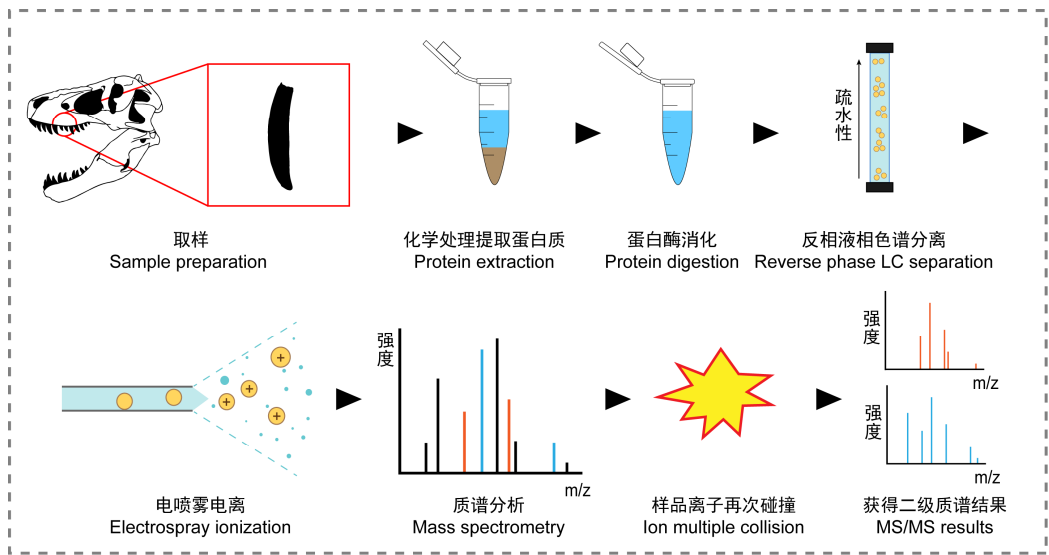


图 2 LC-MS/MS 流程示意图

Fig. 2 Flow chart of LC-MS/MS

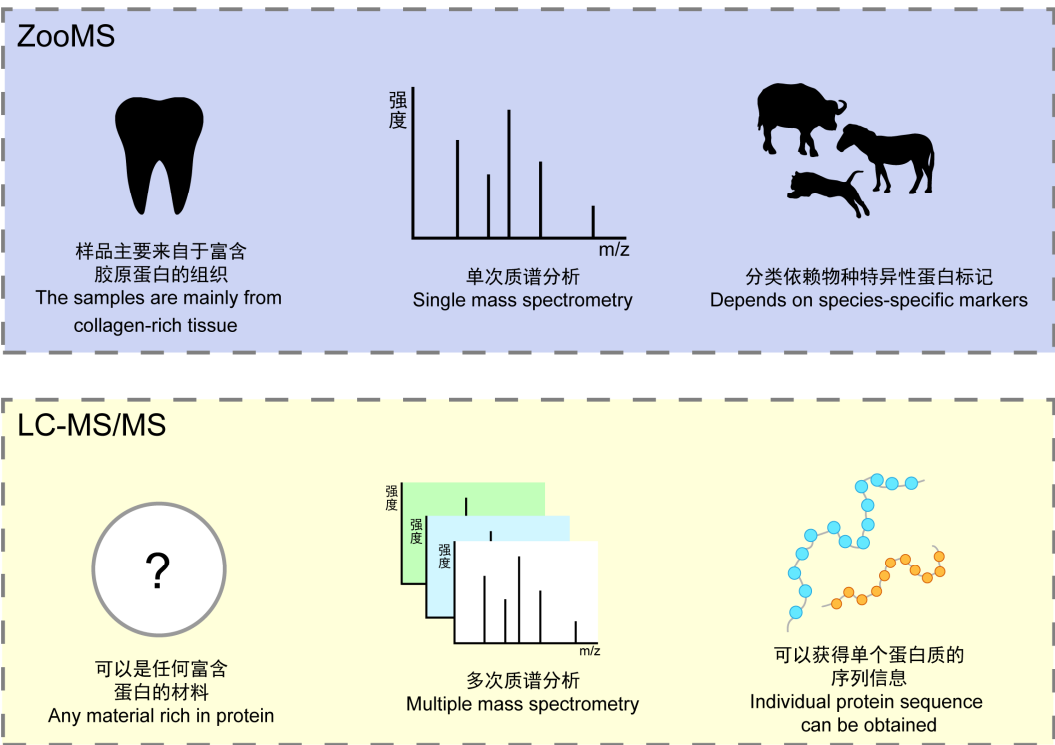


图 3 ZooMS 与 LC-MS/MS 的对比

Fig. 3 Comparison between ZooMS and LC-MS/MS

然后导入质谱进行检测，以期直接得到整个蛋白质的序列信息。因为未引入酶解步骤，故而蛋白质的空间结构信息以及修饰信息在质谱分析中得以保存。并且步骤的减少也节省了整个分析所需要的时间，所以这一技术的效率也更高(Toby *et al.*,

2016)。不过目前自上而下的测序方法对蛋白质的分子量有一定限定，通常适用于分子量小于 30 kDa 的蛋白，当分子量超过 30 kDa 后鉴定效率会有下降(Fornelli *et al.*, 2017)。并且成本较高，不适用于大规模的筛选性研究。

4 古蛋白质组学在古生物学与动物考古学中的应用

4.1 属种鉴定

以胶原蛋白为例,不同属种生物的胶原蛋白序列具有特征性差异,因此通过对比其序列信息可以进行属种鉴定。Buckley等(2008a)尝试利用I型胶原蛋白的序列差异对不同物种骨骼进行分类学研究,他们的实验证实了其可行性。随后他们在此基础上建立了ZooMS技术,成功区分了新石器考古遗址中的山羊(*Capra hircus*)与绵羊(*Ovis aries*)的碎骨(Buckley *et al.*, 2010)。借助古蛋白组学技术的发展,许多在早期无法进行形态学研究或者古DNA分析的化石材料得以发挥价值,如Buckley等(2011)用LC-MS/MS鉴定出两块未知的哺乳动物化石属于草原猛犸象(*Mammuthus trogontherii*); van der Sluis等(2014)对2006年发掘得到的巨兔股骨进行ZooMS分析,成功的鉴定出了其属种。在古人类学研究中,大量碎骨以及骨制品样品也在古蛋白组学技术的帮助下获得了更好的研究。Brown等(2016)结合ZooMS和古线粒体DNA分析,在2315块碎骨中识别出了属于人属的骨骼,并在其中找到了隶属于尼安德特人的线粒体基因;同年, Welker等(2016)也采用了ZooMS快速筛选出了人属碎骨,并结合LC-MS/MS及线粒体DNA技术判断其属于尼安德特人。Chen等(2019)借助ZooMS与鸟枪法蛋白质测序提取并分析了夏河人下颌骨的胶原蛋白,在未能提取得到古DNA的情况下,研究人员仅仅利用了一个SAP的差异便实现了亚种一级的分类,鉴定出夏河人实际上属于丹尼索瓦人,为青藏高原古人类与古生态研究开启了新的纪元。

4.2 系统发育研究

蛋白质序列信息可以帮助研究人员刻画物种的系统演化史。古蛋白质组学研究的目标之一就是通过获取古蛋白质的序列信息,重建远古生物的系统发育过程。尤其在古DNA材料缺失或信息不足的情况下,古蛋白质的序列信息就可能成为唯一的分子遗传信息参考。例如, Welker等(2015)借

助第四纪晚期*Toxodon*与*Macrauchenia*两个属的化石中保留的胶原蛋白序列信息建立了南美洲有蹄类的系统演化模型,解决了有关南美洲大型有蹄类的演化起源疑难; Buckley等(2015)同时应用ZooMS与LC-MS/MS建立了已经灭绝的巨树懒与现存树懒的亲缘关系,为研究树懒的起源与演化提供了新的思路。类似的研究还应用于骆驼、野牛、河狸、犀牛、嵌齿象、岛鼯、鬣狗、巨猿等哺乳动物化石的分子水平系统演化研究中,足以看出古蛋白质组学在系统发育研究中的潜力(Rybczynski *et al.*, 2013; Hill *et al.*, 2015; Cleland *et al.*, 2016; Welker *et al.*, 2017, 2019; Buckley *et al.*, 2019a, 2020; Rao *et al.*, 2020)。

4.3 古生态学重建

假设在同一埋藏条件下的多种生物的胶原蛋白的保存潜力相似,那么就可以根据胶原蛋白的组学信息识别出各种曾经共同生活的生物类群,并可以由此勾勒出该地层对应时期内的古生态群落组成以及各个种群之间的关系。因为ZooMS在大批量快速的动物属种鉴定中具有相当的优势,故而借助这一技术来复原生态学关系也是古蛋白组学研究的重要方向。例如, Buckley等(2018)对晚更新世的洞穴啮齿类化石进行研究时,便借助PMF发现了田鼠与旅鼠在不同层位数量的变化,表明两者在生态位上的竞争关系; Sinet-Mathiot等(2019)对意大利晚更新世的人类遗迹进行研究,在食物碎屑中发现了野牛的骨骼,推测早期人类就有对野牛的捕食行为;同年, Buckley与Herman(2019)对洞穴沉积的蝙蝠骨骼进行了PMF分析,显示当时蝙蝠在洞穴中聚群方式的生活状态。此外,对于古人类学来说,古蛋白质组学还应用在一些骨制品或者食物垃圾中,从而揭示古人类的生活模式,为农业的起源提供证据(Fernandes *et al.*, 2015; Le Meillour *et al.*, 2020; Hendy, 2021)。

4.4 古生理学探索

通过古蛋白质序列的研究,使研究人员有机会直接复原已灭绝动物的蛋白结构,研究其功能位点与空间构型,进而得到古生物的生理学信息。例如, Buckley等(2011)利用LC-MS/MS技术获得了

猛犸象与乳齿象完整的胶原蛋白序列;同年,Cappellini等(2011)发现了猛犸象特有的氨基酸取代位点,并发现其与之前研究中根据古DNA数据得到的突变位点相吻合,从而印证了古蛋白质序列信息的可靠性,使古生物基因突变的研究更进一步(Miller *et al.*, 2008)。此外,Cappellini等(2011)还从猛犸象化石中提取出了球蛋白、角蛋白等一百多种蛋白质,借助这些古蛋白质,研究者得以认识猛犸象体内的不同蛋白质之间的相互作用,从而复原其生化代谢途径。之后,Hill等(2015)对中更新世野牛骨骼化石提取出的胶原蛋白进行分析,发现了与骨骼钙化过程有关的胶原蛋白羟赖氨酸糖基化修饰;同年,Cleland等(2015)在恐鸟科(Dinornithidae)化石中提取到的胶原蛋白片段中发现了甲基化、二甲基化、烷基化、羟基化、岩藻糖基化等翻译后修饰,使得研究人员能直接观察到灭绝物种与现生物种在分子水平的异同。以上这些发现说明:在未来,人们有希望借助古蛋白质的结构与修饰信息,重建古生物的新陈代谢与生理活动,从而更深层次地理解生物演化的规律。

5 古蛋白组学的未来展望

当下古蛋白质组学研究已经从早期的探索阶段逐渐转向标准化技术流程的构建。标准化技术体系的构建,使得不同地区的科学家得以在统一标准下考量对比研究成果,这既便于研究人员检查结果的可靠性,也利于相互之间的交流与共享。这要求研究者设计一套包含样本提取、预处理与分析、质谱分析、误差分析、数据处理与共享的研究构架(Hendy *et al.*, 2018)。就当前的发展趋势来看,未来古蛋白质组学研究将围绕着局部技术的优化与研究范围的拓展两个方面展开。

对于局部技术来说,比较重要的是样本提取技术的优化。在目前的古蛋白质研究中,蛋白材料的可信度一直是备受争议的话题。尤其在中生代非鸟恐龙的胶原蛋白研究中,许多研究者都对其胶原蛋白序列结果表示怀疑(Buckley *et al.*, 2017; Schweitzer *et al.*, 2019)。为提高可信度,就必须强化提取与鉴别技术,以表现出更强的说服力。由于

古生物材料中蛋白质含量较低,这就需要引入多种预筛选方法来提高蛋白获得率(Presslee *et al.*, 2021)。除此之外,在提取方法上,可以在尽可能减少化石材料破坏的同时,采用磁珠法等其他技术手段来替代传统的超滤法也可以提高蛋白质的捕获效率(Palmer *et al.*, 2021)。在获得样品分子之后,还可以使用多种蛋白酶进行处理,以取代之前的单一胰蛋白酶处理,这就可以获得更多样的已知断裂位点,便于后续进行序列分析(Lanigan *et al.*, 2020)。

值得注意的是,利用质谱法进行序列分析时,由于蛋白质的脱酰胺作用,会导致一些氨基酸被转化为其他氨基酸(如天冬酰胺脱酰胺后变为天冬氨酸),从而获得与生物原始状态不同的氨基酸序列信息(Cleland *et al.*, 2015; Welker *et al.*, 2017; Hendy *et al.*, 2018)。如果发生脱酰胺的位点是判断物种类别的关键位点,则可能导致完全错误的物种分析结果。这一问题可能无法仅凭改进质谱技术解决,故而是未来亟待解决的一个难题。

在获得蛋白信息后,结合其他生物大分子进行综合分析也是未来的重要研究方向。在古人类研究中,古DNA与古蛋白质组的协同研究已经取得了一系列的成果,展示了多种生物大分子协同应用的强大效力(Cappellini *et al.*, 2011; Brown *et al.*, 2016)。而不同蛋白质的协同作用也可能成为未来的重要研究方向,例如角蛋白、乳球蛋白、骨钙素、 α -2-HS-糖蛋白、纤维蛋白原 α 链等的结合(Ostrom *et al.*, 2000; Wadsworth and Buckley, 2014; Hendy *et al.*, 2018; Welker, 2018b)。这些在古生物化石中被认为有较好保存潜力的蛋白质的整合研究,能够提供更为细节的蛋白质信息,更全面展示分子演化水平,为深时系统发育构建提供更详细的综合证据(Froment *et al.*, 2021)。

除局部技术的更新外,古蛋白质组学的研究范围也在进一步拓展。目前,古蛋白材料大多数来自于新生代,且主要为更新世的哺乳动物。因此,更久远的化石材料以及其他类群的动物材料将会是未来较为关注的重点。目前可靠的胶原蛋白保存上限约为350万年,而能否将这一上限继续推进,则需要进一步的研究(Buckley *et al.*, 2019a)。不过通过热带地区的巨猿牙齿化石牙釉质蛋白的研究,

指示出古蛋白质对于较高的温度也有一定的耐受能力,而不仅仅局限于高纬度的寒冷地区(Welker *et al.*, 2019)。除了哺乳动物外,其他类群的研究仅有零星报道。如鸟类化石中有报道残余氨基酸的发现,但关于鸟类古蛋白组学的研究还在起步阶段(Huq *et al.*, 1990; McCoy *et al.*, 2019)。而最近也有研究人员展开了两栖类与鱼类的胶原蛋白研究,拓宽了古胶原蛋白的研究视野(Buckley *et al.*, 2020)。而一些对于无脊椎动物的古蛋白质序列研究也在进行当中(Drake *et al.*, 2020)。这些不同类群的蛋白质信息将会帮助研究人员重建与恢复远古时期的生态面貌,为理解与研究古生物的生命活动、生态关系以及演化起到重要的作用。

致谢 评审专家提出宝贵意见与建议,特此感谢!

参考文献 (References)

- 王镜岩, 朱圣庚, 徐长法, 2008. 生物化学教程. 北京: 高等教育出版社. 62–63.
- 张达江, 王亮, 2006. I型胶原蛋白的结构、功能及其应用研究的现状与前景. 生物技术通讯, 17: 265–269.
- Abelson P H, 1954. Paleobiochemistry: organic constituents of fossils. *In*: Carnegie Institution of Washington (ed.), Year Book, 53. Washington: Carnegie Institution of Washington. 97–101.
- Bada J L, Wang X S, Poinar H N, Pääbo S, Poinar G O, 1994. Amino acid racemization in amber-entombed insects: implications for DNA preservation. *Geochimica et Cosmochimica Acta*, 58: 3131–3135. DOI: 10.1016/0016-7037(94)90185-6
- Brock F, Wood R, Higham T F G, Ditchfield P, Bayliss A, Ramsey C B, 2012. Reliability of nitrogen content (%N) and carbon: nitrogen atomic ratios (C: N) as indicators of collagen preservation suitable for radiocarbon dating. *Radiocarbon*, 54: 879–886. DOI: 10.1017/s0033822200047524
- Brown S, Higham T, Slon V, Pääbo S, Meyer M, Douka K, Brock F, Comeskey D, Procopio N, Shunkov M, Derevianko A, Buckley M, 2016. Identification of a new hominin bone from Denisova Cave, Siberia using collagen fingerprinting and mitochondrial DNA analysis. *Scientific Reports*, 6: 23559. DOI: 10.1038/srep23559
- Buckley M, Anderung C, Penkman K, Raney B J, Götherström A, Thomas-Oates J, Collins M J, 2008a. Comparing the survival of osteocalcin and mtDNA in archaeological bone from four European sites. *Journal of Archaeological Science*, 35: 1756–1764. DOI: 10.1016/j.jas.2007.11.022
- Buckley M, Cheylan M, 2020. Collagen fingerprinting for the species identification of archaeological amphibian remains. *Boreas*, 49: 709–717. DOI: 10.1111/bor.12443
- Buckley M, Collins M, Thomas-Oates J, 2008b. A method of isolating the collagen (I) $\alpha 2$ chain carboxyteleopeptide for species identification in bone fragments. *Analytical Biochemistry*, 374: 325–334.
- Buckley M, Collins M, Thomas-Oates J, Wilson J C, 2009. Species identification by analysis of bone collagen using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, 23: 3843–3854.
- Buckley M, Fariña R A, Lawless C, Tambusso P S, Varela L, Carlini A A, Powell J E, Martinez J G, 2015. Collagen sequence analysis of the extinct giant ground sloths *Lestodon* and *Megatherium*. *PLoS One*, 10: 105–109.
- Buckley M, Gu Mu-xin, Herman J, Junno J, Denys C, Chamberlain A T, 2018. Species identification of voles and lemmings from Late Pleistocene deposits in Pin Hole Cave (Creswell Crags, UK) using collagen fingerprinting. *Quaternary International*, 483: 83–89.
- Buckley M, Harvey V L, Orihuela J, Mychajliw A M, Keating J N, Almonte Milan J N, Lawless C, Chamberlain A T, Egerton V M, Manning P L, Teeling E, 2020. Collagen sequence analysis reveals evolutionary history of extinct West Indies *Nesophontes* (Island-Shrews). *Molecular Biology and Evolution*, 37: 2931–2943. DOI: 10.1093/molbev/msaa137
- Buckley M, Herman J, 2019. Species identification of Late Pleistocene bat bones using collagen fingerprinting. *International Journal of Osteoarchaeology*, 29: 1051–1059. DOI: 10.1002/oa.2818.
- Buckley M, Kansa S, Howard S, Campbell S, Thomas-Oates J, Collins M, 2010. Distinguishing between archaeological sheep and goat bones using a single collagen peptide. *Journal of Archaeological Science*, 37: 13–20. DOI: 10.1016/j.jas.2009.08.020
- Buckley M, Larkin N, Collins M, 2011. Mammoth and Mastodon collagen sequences; survival and utility. *Geochimica et Cosmochimica Acta*, 75: 2007–2016.
- Buckley M, Lawless C, Rybczynski N, 2019a. Collagen sequence analysis of fossil camels, *Camelops* and c.f. *Paracamelus*, from the Arctic and sub-Arctic of Plio-Pleistocene North America. *Journal of Proteomics*, 194: 218–225. DOI: 10.1016/j.jprot.2018.11.014
- Buckley M, Recabarren O P, Lawless C, García N, Pino M, 2019b. A molecular phylogeny of the extinct South American gomphothere through collagen sequence analysis. *Quaternary Science Reviews*, 224: 105882. DOI: 10.1016/j.quascirev.2019.105882
- Buckley M, Warwood S, van Dongen B, Kitchener A C, Manning P L, 2017. A fossil protein chimera; difficulties in discriminating dinosaur peptide sequences from modern cross-contamination. *Proceedings of the Royal Society B: Biological Sciences*, 284:

20170544. DOI: 10.1098/rspb.2017.0544
- Cappellini E, Jensen L J, Szklarczyk D, Ginolhac A, da Fonseca R A R, Stafford T W, Holen S R, Collins M J, Orlando L, Willerslev E, Gilbert M T P, Olsen J V, 2012. Proteomic analysis of a Pleistocene Mammoth femur reveals more than one hundred ancient bone proteins. *Journal of Proteome Research*, 11: 917–926. DOI: 10.1021/pr200721u
- Cappellini E, Prohaska A, Racimo F, Welker F, Pedersen M W, Al-lentoft M E, de Barros Damgaard P, Gutenbrunner P, Dunne J, Hammann S, Roffet-Salque M, Ilardo M, Moreno-Mayar J V, Wang Yu-cheng, Sikora M, Vinner L, Cox J, Evershed R P, Willerslev E, 2018. Ancient biomolecules and evolutionary inference. *Annual Review of Biochemistry*, 87: 1029–1060. DOI: 10.1146/annurev-biochem-062917-012002
- Chen Fa-hu, Welker F, Shen Chuan-chou, Bailey S E, Bergmann I, Davis S, Xia Huan, Wang Hui, Fischer R, Freidline S E, Yu Tsai-luen, Skinner M M, Stelzer S, Dong Guang-rong, Fu Qi-ao-mei, Dong Guang-hui, Wang Jian, Zhang Dong-ju, Hublin J J, 2019. A late Middle Pleistocene Denisovan mandible from the Tibetan Plateau. *Nature*, 569: 409–412. DOI: 10.1038/s41586-019-1139-x
- Cleland T P, Schroeter E R, 2018. A comparison of common mass spectrometry approaches for paleoproteomics. *Journal of Proteome Research*, 17: 936–945. DOI: 10.1021/acs.jproteome.7b00703
- Cleland T P, Schroeter E R, Feranec R S, Vashishth D, 2016. Peptide sequences from the first *Castoroides ohioensis* skull and the utility of old museum collections for palaeoproteomics. *Proceedings of the Royal Society B: Biological Sciences*, 283: 20160593. DOI: 10.1098/rspb.2016.0593
- Cleland T P, Schroeter E R, Schweitzer M H, 2015. Biologically and diagenetically derived peptide modifications in moa collagens. *Proceedings of the Royal Society B: Biological Sciences*, 282: 20150015. DOI: 10.1098/rspb.2015.0015
- Demarchi B, Hall S, Roncal-Herrero T, Freeman C L, Woolley J, Crisp M K, Wilson J, Fotakis A, Fischer R, Kessler B M, Rakownikow Jersie-Christensen R, Olsen J V, Haile J, Thomas J, Marean C W, Parkington J, Presslee S, Lee-Thorp J, Ditchfield P, Hamilton J F, Ward M W, Wang C M, Shaw M D, Harrison T, Domínguez-Rodrigo M, DE MacPhee R, Kwekason A, Ecker M, Kolska Horwitz L, Chazan M, Kröger R, Thomas-Oates J, Harding J H, Cappellini E, Penkman K, Collins M J, 2016. Protein sequences bound to mineral surfaces persist into deep time. *eLife*, 5: 17092. DOI: 10.7554/elife.17092
- Drake J L, Whitelegge J P, Jacobs D K, 2020. First sequencing of ancient coral skeletal proteins. *Scientific Reports*, 10: 19407. DOI: 10.1038/s41598-020-75846-4
- Fernandes R, Grootes P, Nadeau M J, Nehlich O, 2015. Quantitative diet reconstruction of a Neolithic population using a Bayesian mixing model (FRUITS): the case study of Ostorf (Germany). *American Journal of Physical Anthropology*, 158: 325–340. DOI: 10.1002/ajpa.22788
- Fornelli L, Durbin K R, Fellers R T, Early B P, Greer J B, LeDuc R D, Compton P D, Kelleher N L, 2017. Advancing top-down analysis of the human proteome using a benchtop quadrupole-orbitrap mass spectrometer. *Journal of Proteome Research*, 16: 609–618. DOI: 10.1021/acs.jproteome.6b00698
- Froment C, Zanolli C, Hourset M, Mouton-Barbosa E, Moreira A, Burlet-Schiltz O, Mollereau C, 2021. Protein sequence comparison of human and non-human primate tooth proteomes. *Journal of Proteomics*, 231: 104045. DOI: 10.1016/j.jprot.2020
- Hare P E, 1967. Nonprotein amino acids in fossils. *Carnegie Inst Year Book*, 65: 362–364.
- Hendy J, 2021. Ancient protein analysis in archaeology. *Science Advances*, 7: eabb9314. DOI: 10.1126/sciadv.abb9314
- Hendy J, Welker F, Demarchi B, Speller C, Warinner C, Collins M J, 2018. A guide to ancient protein studies. *Nature Ecology & Evolution*, 2: 791–799. DOI: 10.1038/s41559-018-0510-x
- Hill R C, Wither M J, Nemkov T, Barrett A, D'Alessandro A, Dzieciatkowska M, Hansen K C, 2015. Preserved proteins from extinct *Bison latifrons* identified by tandem mass spectrometry; hydroxylysine glycosides are a common feature of ancient collagen. *Molecular & Cellular Proteomics: MCP*, 14: 1946–1958. DOI: 10.1074/mcp.M114.047787
- Huq N L, Tseng A, Chapman G E, 1990. Partial amino acid sequence of osteocalcin from an extinct species of ratite bird. *Biochemistry International*, 21: 491–496.
- Karas M, Bachmann D, Hillenkamp F, 1985. Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. *Analytical Chemistry*, 57: 2935–2939. DOI: 10.1021/ac00291a042
- Lanigan L T, Mackie M, Feine S, Hublin J J, Schmitz R W, Wilcke A, Collins M J, Cappellini E, Olsen J V, Taurozzi A J, Welker F, 2020. Multi-protease analysis of Pleistocene bone proteomes. *Journal of Proteomics*, 228: 103889. DOI: 10.1016/j.jprot.2020.103889
- Le Meillour L, Zirah S, Zazzo A, Cersoy S, Détroit F, Imalwa E, Lebon M, Nankela A, Tombret O, Pleurdeau D, Lesur J, 2020. Palaeoproteomics gives new insight into early southern African pastoralism. *Scientific Reports*, 10: 14427. DOI: 10.1038/s41598-020-71374-3
- Li Xiao-juan, Cournoyer J J, Cheng Lin, O'Connor P B, 2008. Use of ¹⁸O labels to monitor deamidation during protein and peptide sample processing. *Journal of the American Society for Mass Spectrometry*, 19: 855–864. DOI: 10.1016/j.jasms.2008.02.011
- McCoy V E, Gabbott S E, Penkman K, Collins M J, Presslee S, Holt J, Grossman H, Wang Bo, Solórzano Kraemer M M, Delclòs X, Peñalver E, 2019. Ancient amino acids from fossil feathers in amber. *Scientific Reports*, 9: 6420. DOI: 10.1038/s41598-019-

- 42938-9
- Miles C A, Knott L, Sumner I G, Bailey A J, 1998. Differences between the thermal stabilities of the three triple-helical domains of type IX collagen. *Journal of Molecular Biology*, 277: 135–144. DOI: 10.1006/jmbi.1997.1603
- Miller W, Drautz D I, Ratan A, Pusey B, Qi Ji, Lesk A M, Tomsho L P, Packard M D, Zhao Fang-qing, Sher A, Tikhonov A, Raney B, Patterson N, Lindblad-Toh K, Lander E S, Knight J R, Irzyk G P, Fredrikson K M, Harkins T T, Sheridan S, Pringle T, Schuster S C, 2008. Sequencing the nuclear genome of the extinct woolly mammoth. *Nature*, 456: 387–390. DOI: 10.1038/nature07446
- Ostrom P H, Schall M, Gandhi H, Shen Tun-li, Hauschka P V, Strahler J R, Gage D A, 2000. New strategies for characterizing ancient proteins using matrix-assisted laser desorption ionization mass spectrometry. *Geochimica et Cosmochimica Acta*, 64: 1043–1050. DOI: 10.1016/S0016-7037(99)00381-6
- Palmer K S, Makarewicz C A, Tishkin A A, Tur S S, Chunag A, Diimajav E, Jamsranjav B, Buckley M, 2021. Comparing the use of magnetic beads with ultrafiltration for ancient dental Calculus proteomics. *Journal of Proteome Research*, 20: 1689–1704. DOI: 10.1021/acs.jproteome.0c00862
- Pestle W J, Ahmad F, Vesper B J, Cordell G A, Colvard M D, 2014. Ancient bone collagen assessment by hand-held vibrational spectroscopy. *Journal of Archaeological Science*, 42: 381–389. DOI: 10.1016/j.jas.2013.11.014
- Presslee S, Penkman K, Fischer R, Richards-Slidel E, Southon J, Hospitaleche C A, Collins M, MacPhee R, 2021. Assessment of different screening methods for selecting palaeontological bone samples for peptide sequencing. *Journal of Proteomics*, 230: 103986. DOI: 10.1016/j.jprot.2020.103986
- Rao Hui-yun, Yang Yi-min, Liu Jin-yi, Westbury M V, Zhang Chi, Shao Qing-feng, 2020. Palaeoproteomic analysis of Pleistocene cave hyenas from east Asia. *Scientific Reports*, 10: 16674. DOI: 10.1038/s41598-020-73542-x
- Rybczynski N, Gosse J C, Richard Harington C, Wogelius R A, Hidy A J, Buckley M, 2013. Mid-Pliocene warm-period deposits in the High Arctic yield insight into camel evolution. *Nature Communications*, 4: 1550. DOI: 10.1038/ncomms2516
- Schroeder R A, Bada J L, 1976. A review of the geochemical applications of the amino acid racemization reaction. *Earth-Science Reviews*, 12: 347–391. DOI: 10.1016/0012-8252(76)90011-8
- Schweitzer M H, Schroeter E R, Cleland T P, Zheng Wen-xia, 2019. Paleoproteomics of Mesozoic dinosaurs and other Mesozoic fossils. *Proteomics*, 19: e1800251. DOI: 10.1002/pmic.201800251
- Sillen A, Parkington J, 1996. Diagenesis of bones from Eland's Bay Cave. *Journal of Archaeological Science*, 23: 535–542. DOI: 10.1006/jasc.1996.0050
- Sinet-Mathiot V, Smith G M, Romandini M, Wilcke A, Peresani M, Hublin J J, Welker F, 2019. Combining ZooMS and zooarchaeology to study Late Pleistocene hominin behaviour at Fumane (Italy). *Scientific Reports*, 9: 12350. DOI: 10.1038/s41598-019-48706-z
- Sponheimer M, Ryder C M, Fewlass H, Smith E K, Pestle W J, Talamo S, 2019. Saving Old Bones: a non-destructive method for bone collagen prescreening. *Scientific Reports*, 9: 13928. DOI: 10.1038/s41598-019-50443-2
- Tisnérat-Laborde N, Valladas H, Kaltnecker E, Arnold M, 2003. AMS radiocarbon dating of bones at LSCE. *Radiocarbon*, 45: 409–419. DOI: 10.1017/S003382220003277x
- Toby T K, Fornelli L, Kelleher N L, 2016. Progress in top-down proteomics and the analysis of proteoforms. *Annual Review of Analytical Chemistry: Palo Alto, Calif*, 9: 499–519. DOI: 10.1146/annurev-anchem-071015-041550
- Trębacz H, Wójtowicz K, 2005. Thermal stabilization of collagen molecules in bone tissue. *International Journal of Biological Macromolecules*, 37: 257–262. DOI: 10.1016/j.ijbiomac.2005.04.007
- van der Sluis L G, Hollund H I, Buckley M, De Louw P G B, Rijdsdijk K F, Kars H, 2014. Combining histology, stable isotope analysis and ZooMS collagen fingerprinting to investigate the taphonomic history and dietary behaviour of extinct giant tortoises from the Mare aux Songes deposit on Mauritius. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 416: 80–91. DOI: 10.1016/j.palaeo.2014.06.003
- van der Valk T, Pečnerová P, Diez-del-Molino D, Bergström A, Oppenheimer J, Hartmann S, Xenikoudakis G, Thomas J A, Dehasque M, Sağlıcan E, Fidan F R, Barnes I, Liu Shan-lin, Somel M, Heintzman P D, Nikolskiy P, Shapiro B, Skoglund P, Hofreiter M, Lister A M, Götherström A, Dalén L, 2021. Million-year-old DNA sheds light on the genomic history of mammoths. *Nature*, 591: 265–269. DOI: 10.1038/s41586-021-03224-9
- Wadsworth C, Buckley M, 2014. Proteome degradation in fossils: investigating the longevity of protein survival in ancient bone. *Rapid Communications in Mass Spectrometry*, 28: 605–615. DOI: 10.1002/rcm.6821
- Wang Jing-yan, Zhu Sheng-geng, Xu Chang-fa, 2008. *Essential Biochemistry*. Beijing: Higher Education Press. 62–63 (in Chinese).
- Welker F, 2018a. Palaeoproteomics for human evolution studies. *Quaternary Science Reviews*, 190: 137–147. DOI: 10.1016/j.quascirev.2018.04.033
- Welker F, 2018b. Elucidation of cross-species proteomic effects in human and hominin bone proteome identification through a bioinformatics experiment. *BMC Evolutionary Biology*, 18: 123. DOI: 10.1186/s12862-018-1141-1
- Welker F, Collins M J, Thomas J A, Wadsley M, Brace S, Cappellini E, Turvey S T, Reguero M, Gelfo J N, Kramarz A, Burger J, Thomas-Oates J, Ashford D A, Ashton P D, Rowsell K, Porter D M, Kessler B, Fischer R, Baessmann C, Kaspar S, Olsen J V,

- Kiley P, Elliott J A, Kelstrup C D, Mullin V, Hofreiter M, Willerslev E, Hublin J J, Orlando L, Barnes I, MacPhee R D E, 2015. Ancient proteins resolve the evolutionary history of Darwin's South American ungulates. *Nature*, 522: 81–84. DOI: 10.1038/nature 14249
- Welker F, Hajdinjak M, Talamo S, Jaouen K, Dannemann M, David F, Julien M, Meyer M, Kelso J, Barnes I, Brace S, Kamminga P, Fischer R, Kessler B M, Stewart J R, Pääbo S, Collins M J, Hublin J J, 2016. Palaeoproteomic evidence identifies archaic hominins associated with the Châtelperronian at the Grotte du Renne. *Proceedings of the National Academy of Sciences of the United States of America*, 113: 11162–11167. DOI: 10.1073/pnas.1605834113
- Welker F, Ramos-Madriral J, Kuhlwiilm M, Liao Wei, Gutenbrunner P, de Manuel M, Samodova D, Mackie M, Allentoft M E, Bacon A M, Collins M J, Cox J, Lalueza-Fox C, Olsen J V, Demeter F, Wang Wei, Marques-Bonet T, Cappellini E, 2019. Enamel proteome shows that *Gigantopithecus* was an early diverging pongine. *Nature*, 576: 262–265. DOI: 10.1038/s41586-019-1728-8
- Welker F, Smith G M, Hutson J M, Kindler L, Garcia-Moreno A, Villaluenga A, Turner E, Gaudzinski-Windheuser S, 2017. Middle Pleistocene protein sequences from the rhinoceros genus *Stephanorhinus* and the phylogeny of extant and extinct Middle/Late Pleistocene Rhinocerotidae. *PeerJ*, 5: e3033. DOI: 10.7717/peerj.3033
- Zhang Da-jiang, Wang Liang, 2006. Structure and Function of Type I Collagen and the Situation and Prospect of Exploratory Development. *Letters in Biotechnology*, 17: 265–269 (in Chinese with English abstract).

(责任编辑: 邓 涛)