



不同城市居民头发中氧、氢稳定同位素检验和特征初探*

于子洋^{1,2)} 杨瑞琴¹⁾ 黄 阳^{1,2)} 秦 皓^{1,2)} 胡 灿²⁾ 朱 军^{2)**}王 萍²⁾ 郭洪玲²⁾ 梅宏成^{2)**}

(1) 中国人民公安大学侦查学院, 北京 100038; 2) 公安部物证鉴定中心, 北京 100038)

摘要 目的 人体组织的稳定同位素组成与其生长期间的个体饮食情况、所处环境及代谢状况有关。人头发一经长出便不再与身体进行物质交换, 化学性质稳定, 易于采集, 是研究人体组织稳定同位素组成的良好对象。构成人体的氧、氢元素主要来自于所摄入的水和食物, 其中氧、氢稳定同位素组成会通过角蛋白的形式被记录于头发当中。不同地区居民头发中氧、氢稳定同位素组成差异可被用于推断人的饮食情况、生活地域和活动轨迹信息, 在法庭科学等研究领域具有重要意义。

方法 本研究利用元素分析仪-稳定同位素质谱仪(EA-IRMS)对国内不同地区常住居民头发样本进行氧、氢稳定同位素比值检测和分析。

结果 部分城市间居民头发 $\delta^{18}\text{O}$ 和 $\delta^2\text{H}$ 存在显著性差异, 整体 $\delta^{18}\text{O}$ 和 $\delta^2\text{H}$ 存在显著正相关性。对所得稳定同位素数据进行判别分析推断头发的地域来源, 其交互验证整体判别准确率为63.9%, 结合碳、氮稳定同位素数据后, 其判别准确率大幅提升, 交互验证的整体判别准确率达到76.0%。随着判别分析中使用的稳定同位素种类的增加, 判别函数模型的判别能力明显增强。

结论 利用4种元素稳定同位素数据建立的多层感知器神经网络模型的整体判别准确率为82.8%, 径向基函数神经网络模型整体判别准确率为78.8%, 3种溯源推断数学模型中, 多层感知器神经网络模型的判别准确率最高。

关键词 人头发, 氧、氢稳定同位素, 溯源推断**中图分类号** R89**DOI:** 10.16476/j.pibb.2021.0317

稳定同位素分析具有示踪指示能力强、检测速度快、测量结果准确等特点, 被广泛应用于生态学和地球化学为代表的众多学科^[1-4]。大气中水蒸气的氧(O)、氢(H)稳定同位素比值受同位素分馏作用的影响, 表现出特定的空间变化规律^[5-6]。其以降水的形式成为地表水, 参与到动植物的生长和代谢活动, 最终反映在动植物体组织中, 由此形成的生物体内O、H稳定同位素比值分布规律可用于研究动物迁移轨迹、物证的真伪鉴定和溯源推断等^[7-12]。法庭科学研究人员为探究人体组织稳定同位素在个体溯源推断领域的应用潜力, 对头发、牙齿、指甲等人体组织进行了不同种类元素的稳定同位素检测和分析^[13-15]。人头发主要由氨基酸组成, 长出毛囊后便停止代谢, 不再与其他部分和环境发生交换, 化学性质稳定, 因此成为记录饮食、饮水中稳定同位素信息的良好载体^[16]。

目前研究人员已经针对人头发中O、H稳定同位素开展了部分研究。研究人员发现人头发中O、

H稳定同位素比值与日常饮用水中O、H稳定同位素比值存在正相关关系, 人头发中H元素大约有30%左右直接来源于饮水, 但当人体摄入大量含有不同 $\delta^2\text{H}$ 值的食物时, 其与当地饮水的相关关系会被极大弱化^[9, 17], 不同地区日常饮用水对居民头发中O、H稳定同位素的影响程度可能存在差异^[18]。在溯源推断领域, Bol等^[19]检测了英格兰当地居民和外来游客头发中O稳定同位素比值, 结合居民头发中碳(C)、氮(N)稳定同位素比值, 利用主成分分析法区分了当地居民和外来游客。

* 国家重点研发计划课题(2017YFC0803803), 公安部技术研究计划项目(2016JSYJB10), 公安部物证鉴定中心基科费发展支撑类项目(2018JB002)和现场物证溯源技术国家工程实验室开放课题(2019NELKFKT09)资助。

** 通讯联系人。

朱军 Tel: 010-83752661, E-mail: zhujun001cn@126.com

梅宏成 Tel: 010-83752662, E-mail: meihongcheng@163.com

收稿日期: 2021-10-19, 接受日期: 2022-03-21

Ehleringer 等^[20]和 Gautam 等^[21]通过大量样品测定, 制作了美国 and 韩国居民头发的 O、H 稳定同位素地图, 帮助实现个体活动轨迹的还原和来源地域的判定。中国地域辽阔, 汇集了大量不同的气候类型及不同饮食习惯, 中国居民头发稳定同位素组成必然存在特有的分布规律。但现有稳定同位素研究报道多来自美国、西欧等国家和地区, 针对中国居民头发中 O、H 等元素稳定同位素特征的研究相对较少^[22]。在溯源推断时, 相关研究多聚焦于单一种类稳定同位素数据的空间分布特征对比, 少有利用统计学手段进行溯源推断, 导致头发稳定同位素数据的溯源推断能力未被充分发掘。

本研究对华中 (郑州市、南阳市)、华东 (温州市)、华南 (贺州市)、西南 (昆明市) 和西北 (哈密市) 等 5 个地区 6 个城市常住居民的头发样本进行了 O、H 稳定同位素比值检验, 探讨中国不同城市居民头发 O、H 稳定同位素比值差异和影响因素, 考察利用不同统计方法对中国居民头发溯源推断的准确性和可靠性, 拓展对中国居民头发中 O、H 稳定同位素分布特征及其溯源能力的认识。

1 材料与方法

1.1 仪器装置与试剂

Flash EA 2000 型元素分析仪、253 Plus 型稳定

同位素比质谱仪、ConFlo IV 型连续流接口装置, 美国 Thermo Scientific 公司产品; KQ-250 DB 型数控超声波清洗器, 昆山市超声波仪器有限公司产品; 202-0AB 电热恒温干燥箱, 北京中兴伟业仪器有限公司产品; XPR 2 电子天平, Mettler Toledo 公司产品。CBS (鹿蹄, $\delta^{18}\text{O}_{\text{VSMOW}}=3.8\text{‰}$, $\delta^2\text{H}_{\text{VSMOW}}=-157\text{‰}$)、USGS42 (人头发, $\delta^{18}\text{O}_{\text{VSMOW}}=8.56\text{‰}$, $\delta^2\text{H}_{\text{VSMOW}}=-78.5\text{‰}$)、USGS43 (人头发, $\delta^{18}\text{O}_{\text{VSMOW}}=14.11\text{‰}$, $\delta^2\text{H}_{\text{VSMOW}}=-50.3\text{‰}$) 稳定同位素标准品, 购自美国 Reston 稳定同位素实验室; 甲醇、氯仿, 色谱纯, Thermo Fisher 公司产品; 实验用水, PURELAB Ultra 超纯水仪 (英国 ELGA 公司产品) 制备; 锡箔杯, 瑞士 Santis 公司产品。

1.2 样品采集

在 6 个城市征集当地常住居民 (至少在当地连续生活 8 个月以上) 作为头发样本提供志愿者, 共采集头发样本 198 份, 具体样本信息如表 1 所示。所有用于 O、H 稳定同位素检验的头发样本均为志愿者贴近头皮 3 cm 左右的头发, 普通人每月头发生长速度约为 1 cm, 贴近头皮 3 cm 左右的头发样本可以保证其为摄入当地食物和水所长成。

Table 1 Information of samples and their origin

City	Longitude	Latitude	Altitude/m	Distance from the sea/km	Sampling time	Number of hair samples	Number of water samples
Hami, Xinjiang	93.54	42.83	692	2 074	2018	41	—
Zhengzhou, Henan	113.77	34.77	89	960	2018	30	3
Nanyang, Henan	112.55	33.00	106	650	2018	38	11
Wenzhou, Zhejiang	120.66	28.00	3	30	2018	33	24
Kunming, Yunnan	102.71	25.05	1 960	1 960	2018	28	27
Hezhou, Guangxi	111.52	24.43	121	121	2018	28	27

1.3 头发样品的前处理

参照 O'Connell 等^[23]的研究, 按照以下流程对头发样本进行清洗: 用甲醇-氯仿 (2 : 1, v/v) 超声清洗 1 h; 用去离子水超声清洗 2 次, 每次 15 min; 用甲醇-氯仿 (2 : 1, v/v) 超声清洗 0.5 h; 用去离子水超声清洗 3 次, 每次 15 min。将清洗后的头发在 85℃ 恒温箱中干燥 3.5 h。将干燥后的头发切割成小段, 称取约 300 μg, 用稳定同位素分析专用银杯包裹。分别称量约 300 μg 的 CBS、

USGS42、USGS43 标准物质, 用稳定同位素分析专用银杯包裹。包好的头发样本和标准物质用元素分析-稳定同位素比质谱仪进行检测。

1.4 水样的保存与处理

水样采集后于 4℃ 以下密封保存, 检测前用 0.22 μm 滤膜过滤。

1.5 稳定同位素比质谱仪的测试条件

元素分析仪: 氦气 (He) 吹扫流速 200 ml/min, 高温裂解炉温度 1 380℃, 色谱柱温度 80℃; 接口

装置ConFloIV: He稀释压力60 kPa。

1.6 O、H稳定同位素比值的检测

1.6.1 头发样本O、H稳定同位素比值测定

在1.5所设定仪器条件下对头发样本和标准物质同时检测,每个样本重复检测3次,采用Isodat3.0软件分析数据。

1.6.2 水样O、H稳定同位素比值测定

检测时,每6个样品插入一个标准样品,每个样品重复测定3次,每次进样2 μl ,利用标准样品校正得到检测结果。

1.6.3 样本O、H稳定同位素比值测定

O、H稳定同位素比值分别用 $\delta^{18}\text{O}$ 和 $\delta^2\text{H}$ 表示,两者相对标准为标准平均海水。计算公式如式(1)所示:

$$\delta = (R_{\text{样本}}/R_{\text{标准}} - 1) \times 1000\text{‰} \quad (1)$$

公式中 R 表示重同位素与轻同位素丰度比。

1.6.4 样本O、H稳定同位素比值的数据校正与分析

参照Qi等^[24]和刘昌景等^[25]制定的O、H稳定同位素比值数据校正方法,将CBS、USGS42和USGS43标准物质的 $\delta^{18}\text{O}$ 和 $\delta^2\text{H}$ 测定值与真实值进行线性拟合,得出线性校正方程。通过线性校正方程和头发样本 $\delta^{18}\text{O}$ 和 $\delta^2\text{H}$ 的测定值计算出其真实值。

1.7 数据处理与统计分析

利用SPSS 20.0软件对样本的O、H稳定同位素比值进行统计分析。

2 结果与讨论

2.1 不同城市居民头发O、H稳定同位素比值特征

检测所得6个城市居民头发 $\delta^{18}\text{O}$ 平均值是 $(9.3 \pm 1.0)\text{‰}$,范围为 $6.0\text{‰} \sim 11.7\text{‰}$ 。居民头发 $\delta^2\text{H}$ 平均值是 $(-71.2 \pm 12.0)\text{‰}$,范围为 $-105.2\text{‰} \sim -47.8\text{‰}$ 。对比不同城市居民头发O、H稳定同位素比值数据,不同城市存在显著性差异(ANOVA, $\delta^{18}\text{O}$: $F=72.04$, $P<0.01$; $\delta^2\text{H}$: $F=305.85$, $P<0.01$)。

对比不同城市居民头发 $\delta^{18}\text{O}$ 和 $\delta^2\text{H}$ 的标准偏差,二者存在显著相关性(PCCs(皮尔逊相关系数), $r=0.65$, $P<0.05$),表明随着城市中居民头发 $\delta^{18}\text{O}$ 离散程度的增加,居民头发 $\delta^2\text{H}$ 的离散程度增加。进一步对比居民头发 $\delta^{18}\text{O}$ 和 $\delta^2\text{H}$,二者存在显著正相关性(PCCs, $r=0.71$, $P<0.01$),表明二者具有相似变化趋势(图1)。居民头发 $\delta^2\text{H}$ 会随居民头发 $\delta^{18}\text{O}$ 增加而增加。

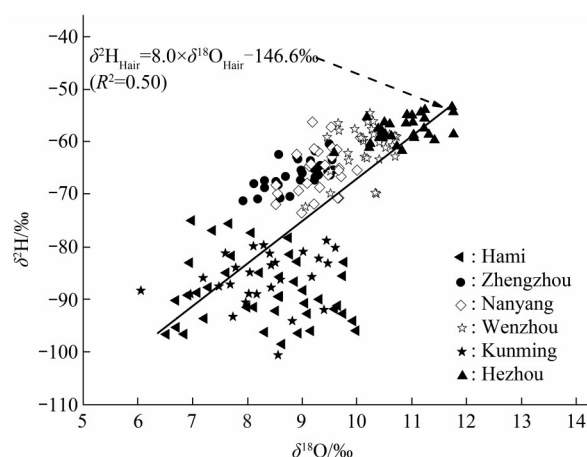


Fig. 1 $\delta^{18}\text{O}$ and $\delta^2\text{H}$ values of resident hair samples in different cities

2.2 日常饮用水中O、H稳定同位素比值特征

本研究检测中国6个城市居民饮用水的O、H稳定同位素比值(暂无哈密饮用水数据),不同城市居民饮用水 $\delta^{18}\text{O}$ 平均值是 $(-7.7 \pm 1.6)\text{‰}$,范围为 $-12.1\text{‰} \sim -4.7\text{‰}$ 。居民饮用水 $\delta^2\text{H}$ 平均值是 $(-53.4 \pm 16.1)\text{‰}$,范围为 $-88.9\text{‰} \sim -27.5\text{‰}$ (图2)。不同城市居民饮用水 $\delta^{18}\text{O}$ 和 $\delta^2\text{H}$,二者存在显著相关性(PCCs, $r=0.99$, $P<0.01$)。Chen等^[26]研究认为,大气降水中的O、H同位素比值受空间变化的影响,存在明显的地域差异,其与纬度和海拔存在明显负相关关系。本研究并未发现饮用水 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 与经度、纬度存在明显相关关系,但饮用水中 $\delta^{18}\text{O}$ 和 $\delta^2\text{H}$ 与海拔存在明显负相关性(PCCs, $\delta^{18}\text{O}$: $r=-0.84$, $P=0.03$; $\delta^2\text{H}$: $r=-0.84$, $P=0.01$)。这表明居民居住地大气降水O、H同位素比值与饮用水可能不完全相同,不同来源水的自然或人工混合会影响居民饮用水中O、H同位素的组成。

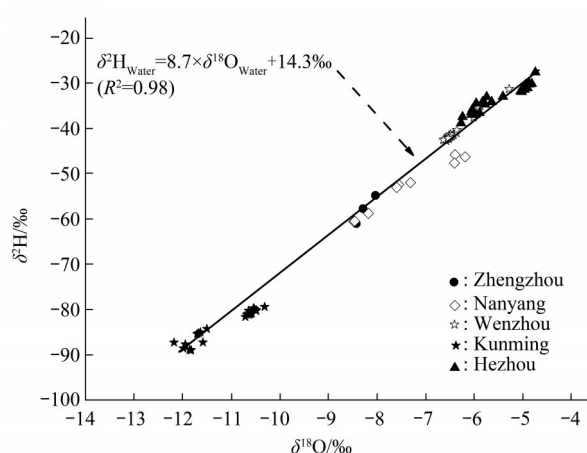


Fig. 2 Distribution of $\delta^{18}\text{O}$ and $\delta^2\text{H}$ values of drinking water samples in different cities

2.3 居民头发O、H稳定同位素比值与饮用水的关系

不同城市居民头发和饮用水中 $\delta^{18}\text{O}$ 和 $\delta^2\text{H}$ 关系如图3所示。整体而言,居民头发中O、H稳定同位素与水中相比完全不同。居民头发中 ^{18}O 相较饮用水中 ^{18}O 是富集的,居民头发中 ^2H 相较饮用水 ^2H 是贫化的,表明饮用水中O、H稳定同位素在被人体组织同化的过程中出现了明显了同位素分馏^[17]。不同城市居民头发 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 和日常饮用水中的 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 存在一定对应关系。居民饮用水 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 较高的贺州、温州居民头发中 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 同样表现出高值,居民饮用水 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 最低的昆明,其居民头发中 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 较低,这表明尽管在人体内存在同位素分馏现象,饮用水中O、H稳定同位素的空间分布特征还是在很大程度上保留在了居民头发当中。与饮用水中 $\delta^{18}\text{O}$ 和 $\delta^2\text{H}$ 相似,居民头发中 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 与空间因素存在一定相关关系,虽与经

度、纬度无明显相关关系,但与海拔存在明显负相关关系(PCCs, $\delta^{18}\text{O}$: $r=-0.67$, $P=0.05$; $\delta^2\text{H}$: $r=-0.75$, $P=0.01$),且相关性与饮用水相比明显减弱。这可能是因为空间因素是经由饮用水影响头发中稳定同位素组成,在此过程中空间因素的影响能力逐渐减弱。

对不同城市居民头发和饮用水 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 平均值进行回归分析,二者存在明显的正相关关系。对于居民头发和饮用水 $\delta^{18}\text{O}$ 平均值,其线性回归方程为: $\delta^{18}\text{O}_{\text{头发}}=0.4\times\delta^{18}\text{O}_{\text{水}}+12.7$ ($R^2=0.91$)。对于H稳定同位素,其线性回归方程为: $\delta^2\text{H}_{\text{头发}}=0.5\times\delta^2\text{H}_{\text{水}}-39.6$ ($R^2=0.90$)。根据不同城市居民头发和饮用水 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 平均值的回归方程,不同城市居民头发中O、H稳定同位素分别有40%和50%来自于日常饮用水。何亚等^[22]同样发现中国居民头发中O、H稳定同位素比值与水中的O、H稳定同位素比值显著相关。

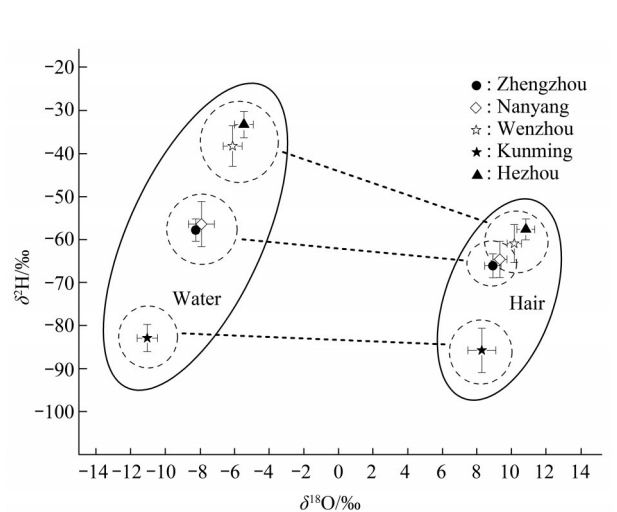


Fig. 3 Distribution of mean $\delta^{18}\text{O}$ and mean $\delta^2\text{H}$ values of resident hair and drinking water samples in different cities

不同国家居民头发和饮用水中 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 如表2所示,居民头发 $\delta^{18}\text{O}$ 范围约为4‰~20‰, $\delta^2\text{H}$ 为-140‰~-48‰; 饮用水 $\delta^{18}\text{O}$ 范围约为-18‰~4‰, $\delta^2\text{H}$ 为-140‰~-8‰。不同国家居民头发 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 均存在明显差异,饮用水中 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 的变化是导致这种差异的重要原因。目前关于头发中稳定同位素的研究尚处于起步阶段,不同研究中样本量和采样地区的选择不同会导致所得数据存在差异(例如本研究 and Thompson等^[18]的研究)。另外,居民饮食数据收集受限,不同饮食习惯与居民组织中稳定同位素组成的关系尚待深入研究,导致不同地区间数据的直接比较相对困难。分析头发和饮用水中O、H稳定同位素的关联性,不同地区中二者的关联程度存在差异,美国居民头发O、H稳定同位素

Table 2 $\delta^{18}\text{O}$ and mean $\delta^2\text{H}$ values of resident hair and drinking water samples between the Chinese and people of other countries

Country	Hair samples		Water samples		Coefficients k of the regression equation for hair and water ($y=kx+b$)		References
	$\delta^{18}\text{O}/\text{‰}$	$\delta^2\text{H}/\text{‰}$	$\delta^{18}\text{O}/\text{‰}$	$\delta^2\text{H}/\text{‰}$	$\delta^{18}\text{O}$	$\delta^2\text{H}$	
China	6.0–11.7	-105.2–-47.8	-12.1–-4.7	-88.9–-27.5	0.40	0.50	This study
Korea	9.1–12.2	-88.7–58.4	-17.8–0.0	-131.9–7.7	0.27	0.36	Gautam <i>et al.</i> ^[21]
China	10.4–15.3	-122–71	-10.8–4	-89–30	0.39	0.42	Thompson <i>et al.</i> ^[18]
India	10.9–20	-109–63	-7.1–3.5	-50–4	–	–	Thompson <i>et al.</i> ^[18]
Mongolia	4.7–13.5	-140–89	-15.1–-0.2	-114–33	–	–	Thompson <i>et al.</i> ^[18]
Pakistan	11.4–15.1	-121–78	-9.6–3.2	-63–18	–	–	Thompson <i>et al.</i> ^[18]
United States	7.4–15.9	-125–77	-18–1.3	-140–4	0.35	0.27	Ehleringer <i>et al.</i> ^[20]

分别有35%和27%来自于日常饮用水^[20]，韩国居民头发比例分别为27%和36%^[21]，回归方程中较高的斜率与居民更多使用当地食物和水有关^[27]，本研究Thompson等^[18]的研究中，中国居民均存在较大的斜率，意味着中国居民可能相对更多的直接或间接食用当地食物和水，这为其在溯源推断领域的应用提供了便利。

2.4 利用头发中稳定同位素信息进行居民居住地的溯源推断

在法庭科学领域，针对头发的检测目前主要集中于表观形态学特征和DNA分析，对其中稳定同位素组成的分析和应用相对较少。结合人头发中多种元素稳定同位素组成进行综合分析，研究人员可以推断人的饮食习惯，进而推断居住地域乃至活动轨迹，辅助人员身份鉴定，为司法活动提供帮助。目前利用多种元素的稳定同位素进行统计学分析从而对人的居住地进行溯源推断的研究较少。本研究利用典型判别和神经网络将测得的6个城市居民头发中的O、H稳定同位素数据结合已有的碳、氮稳定同位素数据进行统计学分析，建立相应数学模型，检验模型的判别准确率，探讨这些统计方法对居民居住地推断的有效性。

2.4.1 居民头发O、H稳定同位素比值的典型判别分析

典型判别分析是在样品类别明确的前提下，建立最优化的线性组合模型来概括类间差异的统计分析方法，是一种应用广泛的多元统计技术。对所构建典型判别函数有效性的验证方法主要包括自身验证和交互验证等。典型判别函数的有效性由判别准确率表示，其数值为判别准确的样本数与总样本数的比值。

本研究对6个城市198份居民头发 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 进行判别分析，以考察居民头发O、H稳定同位素对居民居住城市的溯源推断能力。所得标准化的判别函数式如下：判别函数式1为 $Y_1=0.35X_{\text{O}}+0.90X_{\text{H}}$ ，判别函数式2为 $Y_2=0.93X_{\text{O}}-0.43X_{\text{H}}$ 。这组判别函数

特征值如表3所示，判别函数式1的方差贡献率为96.5%，携带了绝大部分的原始数据信息，判别函数式2的方差贡献率为3.5%，二者合计携带100%的可用于判别分析的原始数据信息。两个函数式 P 值均小于0.01，表明两个函数式在进行判别分析时均存在统计学意义，两个函数式均应当保留作为判别函数模型。

由以上两个判别函数式所构建判别函数模型对6个城市居民头发 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 进行判别分析可获得各类别的联合分布图（图4）。两条坐标轴分别由函数判别式1和判别函数式2构成。

由图4可见，各城市居民头发在判别函数式1上的区别效果较判别函数式2更为明显。哈密、昆明与其他城市明显分开，贺州分布区域与其他城市亦有较为明显的区别。南阳、郑州、温州等3个城市分布集中，区分效果相对较差。由上判别函数所得分类结果如表4所示。6个城市居民头发样本自身验证和交叉验证的整体判别准确率分别为66.1%和63.9%，不同城市居民头发的验证判别准确率为55.2%~84.6%。

不同城市居民头发样本的自身验证判别准确率的分布趋势与交叉验证基本一致，整体略高于交叉验证。其中，哈密、温州、昆明、贺州等城市判别效果相对较好，其判别准确率均在60%以上。郑州和南阳两城市判别效果相对较差，错判主要发生在二者之间。造成错判的可能原因在于郑州和南阳距离相对较近，因此容易相互错判，而这两个城市均位于中国中部地区，与位于中国东部的温州、南部的贺州和西南部的昆明和西北部的哈密相距很远，O、H稳定同位素比值差异较大（表4），所以分类效果相对较好。值得注意的是，地理位置对头发O、H稳定同位素比值的影响并不是唯一的，哈密和昆明虽然距离远，但是其居民头发和饮用水中O、H稳定同位素比值相差不大，这是导致哈密和昆明两地居民头发错判的主要原因。

Table 3 Abstract of Fisher discriminant function of $\delta^{18}\text{O}$ and $\delta^2\text{H}$ values of resident hair

Function	$\delta^{18}\text{O}$ Standardization factor	$\delta^2\text{H}$ Standardization factor	Eigenvalue	Variance contribution/%	Cumulative variance contribution/%	P
1	0.35	0.93	9.42	96.5	96.5	< 0.01
2	0.90	- 0.43	0.34	3.5	100	< 0.01

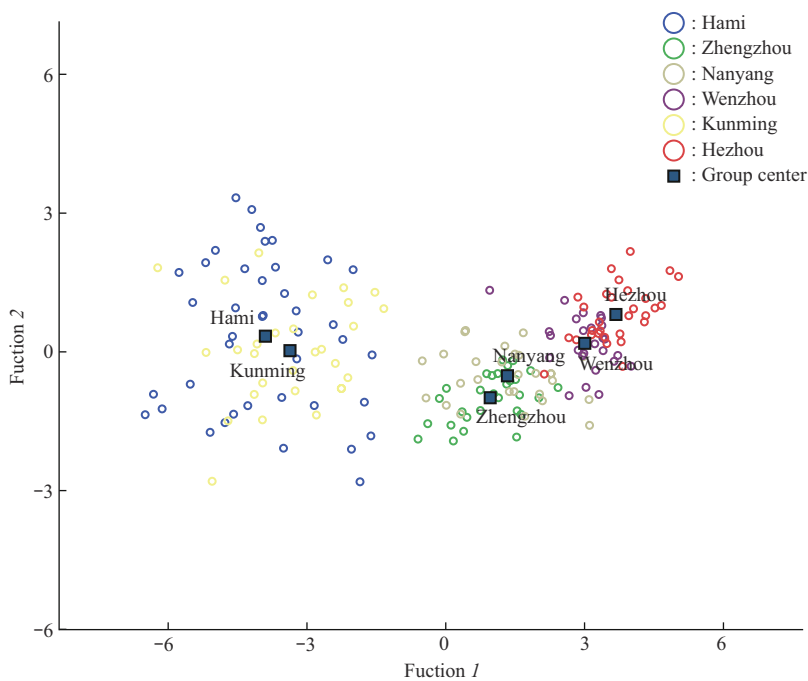


Fig. 4 Distribution of hair samples in different cities on function 1 and function 2 obtained from $\delta^{18}\text{O}$ and $\delta^2\text{H}$

Table 4 Classification results of hair samples

City	Discriminant accuracy (self-validation/cross-validation)/%						Total
	Hami	Zhengzhou	Nanyang	Wenzhou	Kunming	Hezhou	
Hami	63.4/56.1	0/0	0/0	0/0	36.6/41.5	0/0	100.0
Zhengzhou	0/0	55.2/55.2	41.4/41.4	3.4/3.4	0/0	0/0	100.0
Nanyang	0/0	28.6/28.6	60.7/60.7	10.7/10.7	0/0	0/0	100.0
Wenzhou	0/0	0/0	3.8/3.8	84.6/84.6	0/0	11.5/11.5	100.0
Kunming	31.0/34.5	0/0	0/0	0/0	69.0/65.5	0/0	100.0
Hezhou	0/0	0/0	3.3/3.3	30.0/30.0	0/0	66.7/66.7	100.0

2.4.2 居民头发C、N、O、H稳定同位素比值的典型判别分析

结合已有的6个城市居民头发C、N稳定同位素比值(数据未发表),本研究对6个城市居民头发 $\delta^{13}\text{C}$ 、 $\delta^{15}\text{N}$ 、 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 等4种稳定同位素数据进行典型判别分析。典型判别分析所得标准化的判别函数式如表5所示,判别函数式1的方差贡献率为74.8%,判别函数式2的方差贡献率为23.4%,判

别函数式3的方差贡献率为1.7%,三者合计携带99.9%的原始数据信息。3个判别函数式的 P 值均小于0.01,表明3个函数式在进行判别分析时均存在统计学意义,3个函数式均应当保留作为判别函数模型。判别函数式4的方差贡献率仅为0.1%, P 值为0.51,不具有统计学意义,因此排除判别函数式4。

Table 5 Abstract of Fisher discriminant function of $\delta^{13}\text{C}$, $\delta^{15}\text{N}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$ values of resident hair

Function	$\delta^{13}\text{C}$ Standardization factor	$\delta^{15}\text{N}$ Standardization factor	$\delta^{18}\text{O}$ Standardization factor	$\delta^2\text{H}$ Standardization factor	Eigenvalue	Variance contribution /%	Cumulative variance contribution/%	P
1	-0.136	0.375	0.44	0.866	9.23	74.8	74.8	< 0.01
2	-0.366	1.051	0.258	-0.314	2.88	23.4	98.1	< 0.01
3	0.871	-0.021	-0.469	0.158	0.21	1.70	99.9	< 0.01
4	0.684	-0.399	0.736	-0.386	0.01	0.10	100	0.51

以 $\delta^{13}\text{C}$ 、 $\delta^{15}\text{N}$ 、 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 等4种稳定同位素数据构建的典型判别函数模型对6个城市居民头发样本进行判别分析得到联合分布图(图5)。分别以函数判别式1和判别函数式2作为分布图的 x 轴和 y 轴。由图5可见,各组质心在 x 轴和 y 轴均可得到

较为明显的区分。昆明、哈密两城市样本可以与其他城市完全分开。各城市样品组质心之间区分程度要明显好于图4。这表明C、N、O、H等4种稳定同位素数据构建的判别函数模型区分效果好于以O、H两种稳定同位素数据构建的判别函数模型。

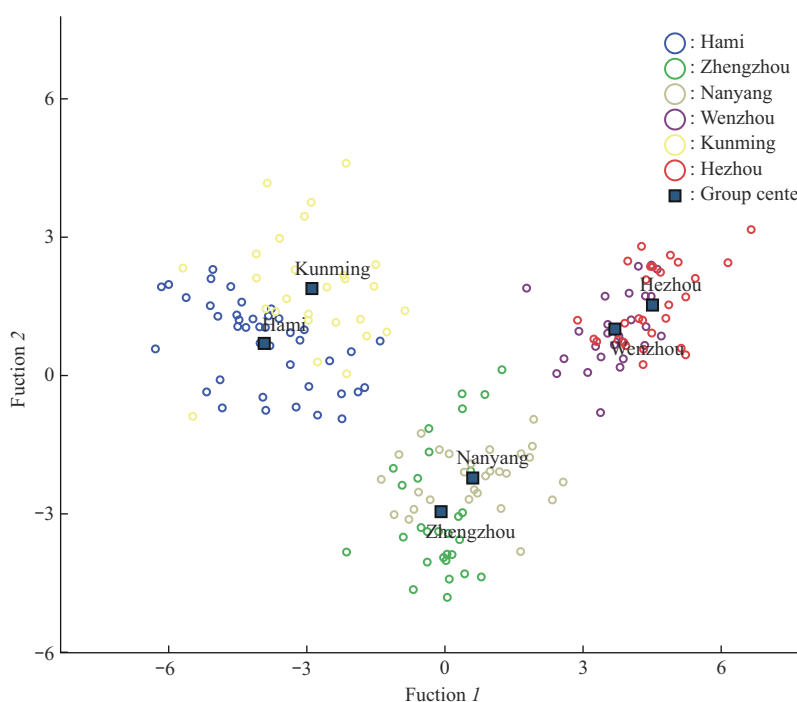


Fig. 5 Distribution of hair sample on function 1 and function 2 obtained from $\delta^{13}\text{C}$, $\delta^{15}\text{N}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$

由上判别函数所得分类结果如表6所示。6个城市居民头发样本自身验证的整体判别准确率为86.6%,不同城市居民头发样本自身验证判别准确率为75.0%~92.7%。交互验证的整体判别准确率为76.0%,不同城市居民头发样本交互验证判别准确

率为64.3%~87.8%,判别效果相对较好,对比表5,各个城市的判别准确率相较O、H两种稳定同位素数据构建的判别函数模型均有不同程度的上升,其中以哈密和昆明上升最为明显,这与图5分布相一致。

Table 6 Discrimination accuracy of the living cities based on $\delta^{13}\text{C}$, $\delta^{15}\text{N}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$

City	Discriminant accuracy (self-validation/cross-validation)/%						Total
	Hami	Zhengzhou	Nanyang	Wenzhou	Kunming	Hezhou	
Hami	92.7/87.8	0/2.4	0/0	0/0	7.3/9.8	0/0	100.0
Zhengzhou	0/0	75.9/65.5	24.1/31.0	0/3.4	0/0	0/0	100.0
Nanyang	0/0	25.0/35.7	75.0/64.3	0/0	0/0	0/0	100.0
Wenzhou	0/0	0/0	0/0	76.9/73.1	0/0	23.1/26.9	100.0
Kunming	11.1/14.8	0/0	0/0	0/0	88.9/85.2	0/0	100.0
Hezhou	0/0	0/0	0/0	25.0/25.0	0/0	75.0/75.0	100.0

利用C、N、O、H不同稳定同位素数据分别构成的判别函数模型其判别准确率如表7所示。 $\delta^{13}\text{C}$ 、 $\delta^{15}\text{N}$ 、 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 各自单独构建判别函数时,整体交互验证判别准确率分别为30.3%、52.0%、44.0%、50.0%。 $\delta^{15}\text{N}$ 和 $\delta^2\text{H}$ 单独构建的判别函数模型的判别效果要好于 $\delta^{18}\text{O}$ 和 $\delta^{13}\text{C}$,表明不同种类元素稳定同位素的溯源推断能力存在差异,H稳定同位素在4种元素中具有最优的溯源推断潜力。利用 $\delta^{13}\text{C}$ 、 $\delta^{15}\text{N}$ 和 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 两种稳定同位素组成的判别

函数模型,其判别准确率分别为59.1%和66.1%,其判别效果要好于单一元素。利用4种稳定同位素构建判别函数的判别准确率为76.0%,相较两种元素构成的判别函数模型,其判别准确率提高了20%,优于何亚等^[22]利用O、H稳定同位素数据进行判别分析的判别准确率,这表明随着判别分析中使用的稳定同位素种类的增加,判别函数模型的判别能力逐渐增强,这与Thompson等^[18]利用多种稳定同位素进行地域区分时所得到的结论相似。

Table 7 Discrimination accuracy of the living cities based on $\delta^{13}\text{C}$, $\delta^{15}\text{N}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$ or their combination

City	Discriminant accuracy (self-validation/cross-validation)/%						
	$\delta^{13}\text{C}$	$\delta^{15}\text{N}$	$\delta^{18}\text{O}$	$\delta^2\text{H}$	$\delta^{13}\text{C}+\delta^{15}\text{N}$	$\delta^{18}\text{O}+\delta^2\text{H}$	$\delta^{13}\text{C}+\delta^{15}\text{N}+\delta^{18}\text{O}+\delta^2\text{H}$
Hami	68.3/68.3	82.9/82.9	12.2/12.2	63.4/63.4	90.2/82.9	63.4/56.1	92.7/87.8
Zhengzhou	10.0/10.0	66.7/66.7	24.1/24.1	58.6/58.6	63.3/63.3	55.2/55.2	75.9/65.5
Nanyang	13.2/13.2	52.6/52.6	67.9/64.3	27.0/24.3	60.5/60.5	60.7/60.7	75.0/64.3
Wenzhou	54.5/54.5/	15.2/15.2	69.2/69.2	38.5/38.5	48.5/42.4	84.6/84.6	76.9/73.1
Kunming	17.9/17.9	32.1/32.1	53.3/46.7	62.1/58.6	21.2/17.9	69.0/65.5	88.9/85.2
Hezhou	3.6/3.6	53.6/53.6	63.3/66.3	60.0/56.7	57.1/57.1	66.7/66.7	75.0/75.0
Total	30.3/30.3	52.0/52.0	45.7/44.0	51.6/50.0	59.1/56.1	66.1/63.9	81.6/76.0

2.4.3 居民头发C、N、O、H稳定同位素比值的神

经网络分析

神经网络分析是借鉴神经元之间的联系,利用计算机的计算能力模拟神经网络的信息传递过程,构成一个大规模、非线性、自适应的动力系统,具有自组织、自适应、自学习的能力。其优点在于对变量没有假设要求,有能力自动且精确的给出复杂联系的近似表达式,具有一定容错性,为处理模糊的、数据不完全的、不精确的模式识别提供了一个全新途径,处理非线性问题的能力优于传统统计分析方法。本研究尝试利用目前应用较多的多层感知

器神经网络(MLP)和径向基函数神经网络(RBF)对6个城市居民头发中C、N、O、H稳定同位素数据进行溯源推断,探究两种神经网络模型对居民头发的溯源推断能力。

利用SPSS软件的MLP对6个城市居民头发的 $\delta^{13}\text{C}$ 、 $\delta^{15}\text{N}$ 、 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 进行分类建模,所得模型摘要如表8所示。将各城市居民头发C、N、O、H稳定同位素数据按照训练集70%、验证集30%的比例构建MLP模型,其输入层为 $\delta^{13}\text{C}$ 、 $\delta^{15}\text{N}$ 、 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 4种稳定同位素数据,输出层为6个城市类别。

Table 8 Abstract of MLP model

Training/validation sample ratio%	Number of input layers	Input layer covariance	Number of hidden layers/units	Hidden layer startup function	Number of output layers	Output level dependent variable	Output layer startup function	Type of training
70/30	4	$\delta^{13}\text{C}$, $\delta^{15}\text{N}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$	1/5	Hyperbolic tangent	6	City	Softmax	Batch process

上述MLP模型对不同城市居民头发的分类情况如表9所示。训练集中各城市居民头发的分类准确率为50.0%~100.0%,整体分类准确率为78.5%。验证集中各城市居民头发的分类准确率为50%~100%,整体分类准确率为82.8%,哈密、郑州、昆明、贺州4个城市的分类准确率相对较高,准确率

为83.3%~100%。这与2.4.1中所构建判别函数模型的判别结果相似。

利用SPSS软件RBF对6个城市居民头发的 $\delta^{13}\text{C}$ 、 $\delta^{15}\text{N}$ 、 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 进行分类建模,所得模型摘要如表10所示。

Table 9 Discrimination accuracy of the living cities based on $\delta^{13}\text{C}$, $\delta^{15}\text{N}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$ in MLP model

	Discrimination accuracy/%						Total
	Hami	Zhengzhou	Nanyang	Wenzhou	Kunming	Hezhou	
Training	100	50.0	82.4	66.7	80.0	86.4	78.5
Validation	94.1	88.9	72.7	50.0	100.0	83.3	82.8

Table 10 Abstract of RBF model

Training/validation sample ratio%	Number of input layers	Input layer covariance	Number of hidden layers/units	Hidden layer startup function	Number of output layers	Output level dependent variable	Output layer startup function
70/30	4	$\delta^{13}\text{C}$, $\delta^{15}\text{N}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$	1/6	Standardized radial basis functions	6	City	Softmax

上述径向基函数神经网络模型对不同城市居民头发的分类情况如表 11 所示。训练集中各城市居民头发的分类准确率为 63.6%~93.1%，整体分类准确率为 79.5%。验证集中各城市居民头发的分类准

确率为 69.2%~100.0%，整体分类准确率为 78.8%。训练集中，哈密的分类准确率最高，准确率达到 93.1%。除郑州外，其他城市的分类准确率均在 70% 以上。测试集中，哈密、南阳、昆明、温州的分类准确率较高。

Table 11 Discrimination accuracy of the living cities based on $\delta^{13}\text{C}$, $\delta^{15}\text{N}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$ in RBF model

	Discrimination accuracy/%						Total
	Hami	Zhengzhou	Nanyang	Wenzhou	Kunming	Hezhou	
Training	93.1	63.6	76.2	75.0	90.0	73.3	79.5
Validation	75.0	71.4	85.7	83.3	100.0	69.2	78.8

对比 MLP 和 RBF 两种神经网络模型利用不同城市居民头发 C、N、O、H 稳定同位素数据的溯源推断能力，MLP 验证集的整体分类准确率 (82.8%) 略优于 RBF (78.8%)。需要注意的是，虽然 MLP 模型在训练集中对郑州居民头发的分类准确率达到 88.9%，但在验证集中仅为 50.0%，构建的两种神经网络模型对郑州居民头发的区分能力均较差，反映出神经网络模型在进行头发分类时的局限性。

本文利用不同城市居民头发 C、N、O、H 稳定同位素数据所构建的典型判别函数、MLP 模型和 RBF 模型 3 种溯源推断模型，其整体判别准确率依次为 76.0% (交互验证)、82.8% (测试集)、78.8% (测试集)，3 种溯源模型分类准确率差别不大，其中 MLP 模型整体分类准确率最高，其次为 RBF 函数，典型判别函数相对较差。对比不同模型所得的分类结果，即便是同一城市的头发样本，不同模型之间构建逻辑的差异会导致部分样本错判到不同的城市。虽然 MLP 模型达到了最高的分类准确率，但相对典型判别函数，其出现了对某些城市样本分类准确率较低的情况，这可能与本研究数据量相对较小有关，在应用时应加以注意。

3 结 论

本文对中国 6 个城市常住居民头发进行了 O、H 稳定同位素比值分析，结果表明，部分城市间居民头发 $\delta^{18}\text{O}$ 、 $\delta^2\text{H}$ 存在显著差异，居民头发整体 $\delta^{18}\text{O}$ 和 $\delta^2\text{H}$ 存在显著正相关性。居民头发 O、H 稳定同位素比值与日常饮用水 O、H 稳定同位素比值呈显著正相关，中国城市居民头发中 O、H 稳定同位素分别有 35% 和 50% 来自于日常饮用水。在溯源推断中，利用居民头发中 O、H 稳定同位素数据进行典型判别分析，其交互验证的整体判别准确率为 63.9%，结合 C、N 稳定同位素数据建立判别模型，其判别准确率大幅提升，交互验证的整体判别准确率达到 76.0%，且随着判别分析中使用的稳定同位素种类的增加，判别函数模型的判别能力显著增强。本文建立的多层感知器神经网络模型的整体分类准确率为 82.8%，径向基函数神经网络模型整体分类准确率为 78.8%，3 种溯源推断数学模型中，多层感知器神经网络模型的判别准确率最高，但需注意，相对典型判别函数，神经网络模型出现了对某些城市样本分类准确率较低的情况，在应用时应加以注意。

参 考 文 献

- [1] Smith K F, Sharp Z D, Brown J H. Isotopic composition of carbon and oxygen in desert fauna: investigations into the effects of diet, physiology, and seasonality. *J Arid Environ*, 2002, **52**(4): 419-430
- [2] 李明财,黎贞发,易现峰,等. 青藏高原东部高寒草甸植物 $\delta^{13}\text{C}$ 年间变化及其环境分析. *生态环境*, 2007, **16**(4): 1205-1210
Li M, Li Z, Yi X, *et al.* *Ecology and Environmental Sciences*, 2007, **16**(4): 1205-1210
- [3] Welp L R, Keeling R F, Meijer H A, *et al.* Interannual variability in the oxygen isotopes of atmospheric CO_2 driven by El Niño. *Nature*, 2011, **477**(7366): 579-582
- [4] Bowen G J, Ehleringer J R, Chesson L A, *et al.* Stable isotope ratios of tap water in the contiguous united states. *Water Resour Res*, 2007, **43**(3): W03419
- [5] Bowen G J, Wilkinson B. Spatial distribution of $\delta^{18}\text{O}$ in meteoric precipitation. *Geology*, 2002, **30**(4): 315-318
- [6] Zhao S, Hu H, Tian F, *et al.* Divergence of stable isotopes in tap water across China. *Sci Rep*, 2017, **7**(1): 43653
- [7] Bowen G J, Wassenaar L I, Hobson K A. Global application of stable hydrogen and oxygen isotopes to wildlife forensics. *Oecologia*, 2005, **143**(3): 337-348
- [8] Gentile N, Siegwolf R T, Esseiva P, *et al.* Isotope ratio mass spectrometry as a tool for source inference in forensic science: a critical review. *Forensic Sci Int*, 2015, **251**(1): 139-158
- [9] O'brien D M, Wooller M J. Tracking human travel using stable oxygen and hydrogen isotope analyses of hair and urine. *Rapid Commun Mass Spectrom*, 2007, **21**(15): 2422-2430
- [10] Kennedy C D, Bowen G J, Ehleringer J R. Temporal variation of oxygen isotope ratios ($\delta^{18}\text{O}$) in drinking water: implications for specifying location of origin with human scalp hair. *Forensic Sci Int*, 2011, **208**(1): 156-166
- [11] Zhao Y, Yang S, Wang D. Stable carbon and nitrogen isotopes as a potential tool to differentiate pork from organic and conventional systems. *J Sci Food Agric*, 2016, **96**(11): 3950-3955
- [12] Ehleringer J R, Casale J F, Lott M J, *et al.* Tracing the geographical origin of cocaine. *Nature*, 2000, **408**(6810): 311-312
- [13] Chesson L A, Meier-Augenstein W, Berg G E, *et al.* Basic principles of stable isotope analysis in humanitarian forensic science//Parra R C, Zapico S C, Ubelaker D H. *Forensic Science and Humanitarian Action: Interacting With the Dead the Living*. New York: Wiley, 2020: 285-310
- [14] Ammer S T, Bartelink E J, Vollner J M, *et al.* Spatial distributions of oxygen stable isotope ratios in tap water from Mexico for region of origin predictions of unidentified border crossers. *J Forensic Sci*, 2020, **65**(4): 1049-1055
- [15] Bartelink E J, Chesson L A. Recent applications of isotope analysis to forensic anthropology. *Forensic Sci Res*, 2019, **4**(1): 29-44
- [16] Petzke K J, Boeing H, Metges C C. Choice of dietary protein of vegetarians and omnivores is reflected in their hair protein ^{13}C and ^{15}N abundance. *Rapid Commun Mass Spectrom*, 2005, **19**(11): 1392-1400
- [17] Sharp Z D, Atudorei V, Panarello H O, *et al.* Hydrogen isotope systematics of hair: archeological and forensic applications. *J Archaeol Sci*, 2003, **30**(12): 1709-1716
- [18] Thompson A H, Chesson L A, Podlesak D W, *et al.* Stable isotope analysis of modern human hair collected from Asia (China, India, Mongolia, and Pakistan). *Amer J Phys Anthropol*, 2010, **141**(3): 440-451
- [19] Bol R, Marsh J, Heaton T H. Multiple stable isotope (^{18}O , ^{13}C , ^{15}N and ^{34}S) analysis of human hair to identify the recent migrants in a rural community in SW England. *Rapid Commun Mass Spectrom*, 2007, **21**(18): 2951-2954
- [20] Ehleringer J R, Bowen G J, Chesson L A, *et al.* Hydrogen and oxygen isotope ratios in human hair are related to geography. *Proc Natl Acad Sci USA*, 2008, **105**(8): 2788-2793
- [21] Gautam M K, Song B Y, Shin W J, *et al.* Spatial variations in oxygen and hydrogen isotopes in waters and human hair across south Korea. *Sci Total Environ*, 2020, **726**(2020): 138365
- [22] 何亚,刘昌景,梅宏成,等. 阿克苏、连云港和淮安城市居民头发中氢、氧稳定同位素比值差异研究. *质谱学报*, 2019, **40**(6): 575-583
He Y, Liu C, Mei H, *et al.* *J Chin Mass Spectr Soc*, 2019, **40**(6): 575-583
- [23] O'Connell T C, Hedges R E M. Isotopic comparison of hair and bone: archaeological analyses. *J Archaeol Sci*, 1999, **26**(6): 661-665
- [24] Coplen T B, Qi H. A revision in hydrogen isotopic composition of USGS42 and USGS43 human-hair stable isotopic reference materials for forensic science. *Forensic Sci Int*, 2016, **266**: 222-225
- [25] 刘昌景,梅宏成,朱军,等. 头发氢稳定同位素比值检测的标准品选择研究. *质谱学报*, 2019, **40**(4): 342-348
Liu C, Mei H, Zhu J, *et al.* *J Chin Mass Spectr Soc*, 2019, **40**(4): 342-348
- [26] Chen Z, Cheng J, Guo P, *et al.* Distribution characters and its control factors of stable isotope in precipitation over China. *Trans Atmos Sci*, 2010, **33**: 667-679
- [27] Fraser I, Meier-Augenstein W. Stable ^2H isotope analysis of modern-day human hair and nails can aid forensic human identification. *Rapid Commun Mass Spectrom*, 2007, **21**(20): 3279-3285

Examination and Analysis on Oxygen and Hydrogen Stable Isotope Ratios of Resident Hair*

YU Zi-Yang^{1,2)}, YANG Rui-Qin¹⁾, HUANG Yang^{1,2)}, QIN Hao^{1,2)}, HU Can²⁾, ZHU Jun^{2)**},
WANG Ping²⁾, GUO Hong-Ling²⁾, MEI Hong-Cheng^{2)**}

(¹⁾School of Criminal Investigation, Peoples' Public Security University of China, Beijing 100038, China;

²⁾Institute of Forensic Science, Ministry of Public Security, Beijing 100038, China)

Abstract Objective The stable isotopic composition of human tissue is related to the individual's diet, environment and metabolic status. Once grown, human hair is no longer exchanged with the body and is chemically stable and easy to collect, making it a good subject for studying the stable isotope composition of human tissue. The oxygen and hydrogen elements that make up the human body are mainly derived from the water and food, of which the oxygen and hydrogen stable isotope composition is recorded in hair keratin. Differences in the stable isotope composition of oxygen and hydrogen in the human hair of different regions can be used to infer information about people's diets, living areas and activity trajectories, which is of great importance in forensic science and other research fields. **Methods** In this study, oxygen and hydrogen stable isotope ratios of hair samples from permanent residents in different regions of China were detected and analyzed using an elemental analyzer-isotope ratio mass spectrometer (EA-IRMS). **Results** The results showed that there were significant differences in $\delta^{18}\text{O}$ and $\delta^2\text{H}$ between some cities, and a significant positive correlation between overall $\delta^{18}\text{O}$ and $\delta^2\text{H}$. The overall discrimination accuracy of the stable isotope data was 63.9% for the cross-validation, and increased significantly when combined with the carbon and nitrogen stable isotope data, reaching an overall discrimination accuracy of 76.0% for the cross-validation. As the type of stable isotopes used in the discriminant analysis increased, the model's discriminant ability was significantly enhanced. **Conclusion** The overall discriminative accuracy of the multi-layer perceptron neural network model built by using the stable isotope data of the 4 elements was 82.8%, and the overall discriminative accuracy of the radial basis function neural network model was 78.8%. The multilayer perceptron neural network model had the highest discriminative accuracy among the three traceability mathematical models.

Key words human hair, oxygen and hydrogen stable isotope, traceability

DOI: 10.16476/j.pibb.2021.0317

* This work was supported by grants from National Key R&D Program of China (2017YFC0803803), Technology Research Program of Ministry of Public Security of China (2016JSYJB10), Basal Research Fund of Institute of Forensic Science, Ministry of Public Security (2018JB002), and Open Research Grants Program of National Engineering Laboratory for Forensic Science (2019NELKFKT09).

** Corresponding author.

ZHU Jun. Tel: 86-10-83752661, E-mail: zhujun001cn@126.com

MEI Hong-Cheng. Tel: 86-10-83752662, E-mail: meihongcheng@163.com

Received: October 19, 2021 Accepted: March 21, 2022