Received: 7 August 2020 | Accepted: 24 December 2020

DOI: 10.1111/2041-210X.13556

RESEARCH ARTICLE

Methods in Ecology and Evolution | BRITISH ECOLOGICAL SOCIETY

Calibration chain transformation improves the comparability of organic hydrogen and oxygen stable isotope data

Sarah Magozzi[1,2] | Clement P. Bataille[3] | Keith A. Hobson[4] | Michael B. Wunder[5] | John D. Howa[6] | Andrea Contina[5] | Hannah B. Vander Zanden[7] | Gabriel J. Bowen[1]

[1]Department of Geology and Geophysics, University of Utah, Salt Lake City, UT, USA; [2]Department of Integrative Marine Ecology, Stazione Zoologica Anton Dohrn, Fano Marine Centre, Fano (PU), Italy; [3]Department of Earth and Environmental Sciences, University of Ottawa, Ottawa, ON, Canada; [4]Department of Biology and Environment and Climate Change Canada, Western University, London, ON, Canada; [5]Department of Integrative Biology, University of Colorado Denver, Denver, CO, USA; [6]Howa Analysis and Consulting LLC, Salt Lake City, UT, USA and [7]Department of Biology, University of Florida, Gainesville, FL, USA

Correspondence
Sarah Magozzi
Email: sarah.magozzi@szn.it

Funding information
National Science Foundation, Grant/Award Number: DBI-1565128 and DBI-1759730

Handling Editor: David Soto

Abstract

1. Stable hydrogen and oxygen isotopic compositions (δ^2H and $\delta^{18}O$, respectively) of animal tissues have been used to infer geographical origin or mobility based on the premise that the isotopic composition of tissue is systematically related to that of local water sources. Isotopic data for known-origin samples are required to quantify these tissue–environment relationships. Although many of such data have been published and could be reused by researchers, differences in the standards used for calibration and analytical procedures for different datasets limit the comparability of these data.
2. We develop an algorithm that uses results from comparative analysis of secondary standards to transform data among reference scales and estimate the uncertainty inherent in these transformations. We apply the algorithm to a compilation of known-origin keratin data published over the past ~20 years.
3. We show that transformation improves the comparability of data from different laboratories, and that the transformed data suggest ecophysiologically meaningful differences in keratin–water relationships among different animal groups and taxa.
4. The compiled data and algorithms are freely available in the ASSIGNR R-package to support geographical provenance research, and more generally offer a methodology overcoming several challenges in geochemical data integration and reuse.

KEYWORDS
calibration, comparative equilibration, database, isotopic exchange, keratin, known-origin, reference materials, standards

1 | INTRODUCTION

The stable hydrogen and oxygen isotope compositions (δ^2H and $\delta^{18}O$, respectively) of animal tissues have found widespread use as a geographical marker in wildlife, archaeological and forensic applications (reviewed by Carter & Chesson, 2017; Hobson & Wassenaar 2019; Meier-Augenstein, 2018). The underlying premise of these applications is that tissue isotopic composition is linked

to that of local water sources (i.e. precipitation) through relatively predictable relationships (e.g. Chamberlain et al., 1996; Ehleringer et al., 2008; Hobson & Wassenaar, 1997). The isotopic composition of precipitation varies predictably across space and time (Bowen & Revenaugh, 2003; Bowen, Wassenaar, et al., 2005; Craig, 1961), and thus tissue isotope ratios derived ultimately from precipitation-driven local food webs can be compared to environmental water isoscapes (predictive models of spatio-temporal isotope patterns) to infer tissue origin (Ma et al., 2020; Vander Zanden et al., 2014; Wunder, 2010). Keratinous tissues, such as feather, hair or nail, are metabolically inert once formed and so preserve an isotopic composition characteristic of the location of tissue growth (Hobson, 1999; Macko et al., 1999; West et al., 2004), and are the focus here and in many published studies.

Because tissue–environment relationships vary among taxa and regions (Magozzi et al., 2019 and references therein), samples of known origin are needed to quantify such relationships (e.g. Chamberlain et al., 1996; Ehleringer et al., 2008; Hobson et al., 2012; Hobson & Wassenaar, 1997). The collection and analysis of known-origin samples, however, is resource-intensive and in some cases prohibitive, reducing the efficiency and applicability of the approach. Although many known-origin datasets have already been published, and in theory could be reused to quantify tissue–environment relationships in new studies, different sample preparation, analytical and calibration practices used at different laboratories (or even at a single laboratory over time) have generated data that are not directly comparable (e.g. Bowen, Chesson, et al., 2005; Meier-Augenstein et al., 2013; Soto et al., 2017; Wassenaar & Hobson, 2003). As a result, responsible users of published data have thus far focused on measurements made in a single laboratory or obtained with identical protocols.

Different sample treatment and analysis methods are one source of inconsistency in published data. Exchange of H atoms from protein carboxyl and hydroxyl groups with ambient atmospheric water vapour molecules affects measured δ^2H values and must be corrected for (Bowen, Chesson, et al., 2005; Chesson et al., 2009; Schimmelmann, 1991; Wassenaar & Hobson, 2000, 2003), most commonly through comparative analysis against keratin standards for which non-exchangeable δ^2H values have been established (Kelly et al., 2009; Sauer et al., 2009; Wassenaar & Hobson, 2003). Water is also tightly adsorbed by keratin and adheres to sample capsules, contributing to the measured δ^2H and δ^{18}O values (Wortmann et al., 2001) unless samples are thoroughly dried prior to combustion (Bowen, Chesson, et al., 2005; Coplen & Qi, 2012; Soto et al., 2017). Analytical methods have themselves evolved since the advent of online thermal conversion/elemental analysis isotope ratio mass spectrometry (TC/EA-IRMS) for keratin H and O isotope analysis ~20 years ago. The use of different pyrolysis reactor fillings (e.g. chromium vs. glassy carbon; Gehre et al., 2015; Nair et al., 2015) and chromatographic conditions (Hunsinger et al., 2013; Qi et al., 2011) has been shown to affect measured δ^2H and δ^{18}O values, respectively.

Fortunately, isotope ratio analysis is performed as a comparative analysis, wherein the sample values are calibrated to the accepted values for co-analysed standards. Under ideal circumstances conforming to the principle of identical treatment (PIT; Werner & Brand, 2001), which requires the preparation and analysis of matrix-matched (chemically and physically equivalent) standards alongside the unknown samples, analytical biases should affect the samples and standards similarly and have little effect on sample values reported relative to the standard values (i.e. on a 'reference scale' defined by the standards used and their assigned values). However, ideal circumstances are not possible in many situations. Even where samples and standards are compositionally similar (i.e. keratin), differences in preparation (i.e. grinding/powdering) and biochemistry (i.e. amino acid profile) can affect H exchange and may impart bias to otherwise PIT-compliant comparisons (Alibardi, 2017; Bowen, Chesson, et al., 2005; Robbins, 2012). Additionally, the internationally accepted reference scale (here the Vienna Standard Mean Ocean Water, 'VSMOW', scale) for δ^2H and δ^{18}O data is defined by primary standards that are water, meaning that calibrating organic secondary standards to the VSMOW scale using PIT procedures is impossible. As methodologies have advanced, a set of 'optimal' procedures minimizing bias in these non-PIT comparisons has been developed, but in the interim many different secondary standard calibrations have been produced, leading to published known-origin data that are laboratory-specific and not robustly traceable to the VSMOW scale.

Here, we leverage cross-calibration studies, in which one set of secondary standards is analysed alongside and calibrated to a reference scale defined by a second set, to develop a method that transforms data between reference scales. Transformation from the original scale to a target scale proceeds along a chain of linked calibrations, ideally with each link consisting of a PIT-based cross-calibration or a non-PIT calibration using optimal methods (Qi et al., 2011; Soto et al., 2017; see also Coplen & Qi, 2016). The algorithm propagates uncertainty, and permits comparison of data reported on more than two dozen reference scales. We apply and test the method using a compilation of data for >4,000 keratin samples, showing that the method reduces, but does not always eliminate, discrepancies among data from different laboratories. The data and transformation algorithm are available in the ASSIGNR R-package (Ma et al., 2020) to support their open reuse.

2 | MATERIALS AND METHODS

2.1 | Secondary standard calibration history

We compile information on widely used keratin secondary standards. Our list is not exhaustive, but includes standards calibrated by five stable isotope laboratories (United States Geological Survey [USGS] Denver and Reston, the University of Ottawa, Environment Canada [Saskatoon] and the University of Utah) that have been used to generate a large fraction of the published keratin isotope data. Each reference scale is defined by two secondary standards and their assigned values (Dunn & Carter, 2018). In a typical laboratory application, a linear model relating instrument-reported values for the secondary standards to their assigned values is applied to the

measured values for unknowns to calibrate them to the reference scale.

Assigned values for secondary standards are the product of calibration (through co-analysis) relative to a reference scale defined by a different set of primary or secondary standards. These calibrations have used a wide range of methodologies, and for most keratin standards multiple calibrations have been generated using different methods. In general, these can be classified as calibrations based on (1) PIT methods, with full reporting of methods and uncertainty, (2) measurement against non-matrix-matched standards using optimal methods to minimize matrix effects or (3) measurements that are neither PIT-based nor optimal. For category 2, we consider optimal methods to include thorough sample drying (e.g. using a sealed, evacuated and heated carousel such as the Uni-Prep^{TM} device; Wassenaar et al., 2015; Soto et al., 2017), correction for exchangeable H via equilibration with multiple waters and use of a Cr-filled (vs. glassy carbon) pyrolysis reactor to avoid bias caused by HCN-producing reactions (Soto et al., 2017; see also Coplen & Qi, 2016). Optimal methods for $\delta^{18}O$ analyses include thorough drying and chromatographic methods allowing complete separation of N_2 and CO (Qi et al., 2011); for O standards, we are not aware of calibrations that combine sealed-carousel drying with optimal chromatography, and we accept somewhat less stringent (vacuum oven) drying procedures as acceptable. For categories 1 and 2, we accept that the assigned values (and their uncertainty based on replicate analyses) represent a strong and unbiased (or minimally biased) calibration of the newly calibrated standards to the reference scale. For category 3, the comparability of the comparison is questionable, and we generally do not consider a traceable link to be established. In this case, we treat the calibration as a 'floating' calibration, which cannot be linked to other reference scales through traceable cross-calibration. The floating calibration, however, can be linked to other calibrations for the same physical materials and their associated reference scales.

2.1.1 | Hydrogen secondary standards

We compiled H calibration information for most commonly used keratin standards (Table 1), and here introduce different calibrations associated with each set of physical materials. In 2003, the Environment Canada (now Environment and Climate Change Canada) stable isotope laboratory in Saskatoon, Canada, introduced the CHS, CFS and BWB (cow hoof, chicken feathers and bowhead whale baleen) secondary standards, and reported calibrations based on dual-inlet analysis of steam-equilibrated aliquots of these materials with offline H_2 derived from zinc reduction techniques versus online combustion (calibrations OldEC.1_H_1 and OldEC.2_H_1, respectively; Wassenaar & Hobson, 2003). In 2011, the Environment Canada laboratory introduced the KHS and CBS (kudu horn and caribou hoof) standards to replace CHS, CFS and BWB, and analysed these with online combustion continuous flow isotope ratio mass spectrometry techniques; between 2011 and 2015, nine calibrations for these secondary standards were reported (EC_H_1-9; Qi & Coplen, 2011; Soto et al., 2017; Wassenaar et al., 2015; L.I. Wassenaar, pers. comm.). EC_H_7 and 8 offered uncertainty-quantified, PIT-based calibrations to the OldEC.1_H_1 and OldEC.2_H_1 reference scales, respectively (Soto et al., 2017; L.I. Wassenaar, pers. comm.). EC_H_9 represented certified values for the KHS and CBS secondary standards based on calibration to VSMOW using optimal methods (https://isotopes.usgs.gov/lab/referencematerials/KHS.pdf; https://isotopes.usgs.gov/lab/referencematerials/CBS.pdf; Soto et al., 2017).

The USGS Denver laboratory produced two hair standards designated AK and LA (Alaska and Louisiana bear hair). These were initially calibrated through offline zinc-reduction and dual-inlet analyses against non-matrix-matched secondary standards (DEN_H_1; C. Stricker, pers. comm.). A round-robin experiment in 2005 produced a PIT-based calibration to the OldEC.1_H_1 reference scale (DEN_H_2; C. Stricker, pers. comm.).

In 2005, the SIRFER laboratory (University of Utah) prepared the FH and UH (Florida and Utah horse hair) standards. Several non-PIT calibrations to other organic secondary standards have been published (OldUT_H_1, 3 and 4; Bowen, Chesson, et al., 2005; L.A. Chesson & J.D. Howa, pers. comm.; Table S3). A PIT calibration to the US_H_1 reference scale (see below) was published in 2011 (OldUT_H_2; Coplen & Qi, 2012; Qi & Coplen, 2011). In 2012, IsoForensics Inc., introduced the DS and ORX (Dall sheep and oryx antelope horns) standards to replace FH and UH. Five calibrations for these are available (UT_H_1-5; L.A. Chesson & J.D. Howa, pers. comm.; J.D. Howa, pers. comm.; Table S3). Each of these is either non-PIT or lacking uncertainty, however, and therefore considered a floating calibration.

In 2011–2012, the USGS Reston isotope laboratory introduced the USGS42 and USGS43 (Tibetan and Indian human hair) standards. Seven calibrations exist, including a series of calibrations to primary standards using non-optimal methods (US_H_1-4 and 6; Coplen & Qi, 2012, 2016; Qi & Coplen, 2011; Wassenaar et al., 2015; summarized by Soto et al., 2017) and one using optimal methods (US_H_7; Soto et al., 2017). The certified values for these materials comprise the US_H_6 scale (https://isotopes.usgs.gov/lab/referencematerials/USGS42.pdf; https://isotopes.usgs.gov/lab/referencematerials/USGS43.pdf; Coplen & Qi, 2016), and, although this calibration used non-optimal drying methods, the values obtained are indistinguishable from US_H_7 (Soto et al., 2017) and thus adopted here as the authoritative calibration to VSMOW. USGS42 and USGS43 were also calibrated to the UT_H_2 reference scale (US_H_5; Table S3).

Two human hair standards (AND and CAL-CAN) were prepared at the University of Ottawa G.G. Hatch (now Jan Veizer) stable isotope laboratory in 2011. Four non-PIT calibrations have been presented (CAN_H_1-3 and 5; Coplen & Qi, 2012; Meier-Augenstein et al., 2011). Although CAN_H_5 represents a keratin-to-keratin calibration, the two sets of secondary standards were not uniformly powdered prior to analysis (T.B. Coplen, pers. comm.); therefore, this calibration did not meet PIT requirements. A single PIT-based calibration to the US_H_1 reference scale was reported by Meier-Augenstein et al. (2011; CAN_H_4).

TABLE 1 Calibrated δ^2H values for most commonly used secondary and primary standards for H isotope analysis of keratins

Calibration	Standard	δ^2H (mean)	δ^2H (*SD*)	*n*	Reference scale	Citation	Methods
DEN_H_1	LA	−63			DEN_H_1	C. Stricker (pers. comm.)	Equilibration with waters; Zn-reduction and dual inlet-IRMS; calibration to non-keratin organic standards (PEF-1)
	AK	−164					
DEN_H_2	LA	−78.1	3	29	OldEC.1_H_1	C. Stricker (pers. comm.)	Co-analysis with OldEC.1_H_1; TCEA-IRMS; calibration to OldEC.1_H_1
	AK	−171.5	3.3	30			
OldEC.1_H_1	BWB	−108			OldEC.1_H_1	Wassenaar and Hobson (2003)	Equilibration with waters at high temperature; Zn-reduction and dual inlet-IRMS method described by Wassenaar and Hobson (2000); calibration to water standards
	CHS	−187					
OldEC.2_H_1	BWB	−108			OldEC.2_H_1	Wassenaar and Hobson (2003)	Equilibration with waters at high temperature; Zn-reduction and dual inlet-IRMS method described by Wassenaar and Hobson (2000); calibration to water standards
	CFS	−138					
EC_H_1	KHS	−37.1			EC_H_1	Soto et al. (2017)	Equilibration with waters at room temperature for 6 days, dried using different methods, assumed $\varepsilon = 0‰$; TCEA-IRMS with glassy C-filled reactor; calibration to water standards
	CBS	−166.3					
EC_H_2	KHS	−41			EC_H_2	Soto et al. (2017)	Equilibration with waters at 120°C, evacuated using different methods, assumed $\varepsilon = 0‰$ per mil; TCEA-IRMS with glassy C-filled reactor; calibration to water standards
	CBS	−177.9					
EC_H_3	KHS	−35.5			EC_H_3	Soto et al. (2017)	Equilibration with waters at 70°C for 2 hr, evacuated and flushed with He online, assumed $\varepsilon = 80‰$; TCEA-IRMS with glassy C-filled reactor; calibration to water standards
	CBS	−153.3					
EC_H_4	KHS	−47.5			EC_H_4	Soto et al. (2017)	Equilibration with waters at 105°C for 2 hr, evacuated and flushed with He online, assumed $\varepsilon = 80‰$; TCEA-IRMS with glassy C-filled reactor; calibration to water standards
	CBS	−178.8					
EC_H_5	KHS	−54.1			EC_H_5	Soto et al. (2017)	Co-analysis with OldEC.1_H_1; TCEA-IRMS; calibration to OldEC.1_H_1; no uncertainty reported
	CBS	−197					
EC_H_6	KHS	−54.1			EC_H_6	Soto et al. (2017)	Co-analysis with OldEC.2_H_1; TCEA-IRMS; calibration to OldEC.2_H_1; no uncertainty reported
	CBS	−197					
EC_H_7	KHS	−54.7	3.12	711	OldEC.1_H_1	L.I. Wassenaar (pers. comm.)	Co-analysis with OldEC.1_H_1; TCEA-IRMS; calibration to OldEC.1_H_1; data until 2012
	CBS	−197.2	4.34	689			
EC_H_8	KHS	−54.7	3.12	711	OldEC.2_H_1	L.I. Wassenaar (pers. comm.)	Co-analysis with OldEC.2_H_1; TCEA-IRMS; calibration to OldEC.2_H_1; data until 2012
	CBS	−197.2	4.34	689			
EC_H_9*	KHS	−35.3	1.1	5	VSMOW_H	Soto et al. (2017)	Equilibration with waters over a range of temperatures (25–105°C) and drying procedures, assumed $\varepsilon = 0‰$; TCEA-IRMS with Cr-filled reactor; calibration to water standards (optimal method)
	CBS	−157	0.9	5			
OldUT_H_1	FH	−76.3			OldUT_H_1	Bowen, Chesson, et al. (2005), L.A. Chesson and J.D. Howa (pers. comm.)	Equilibration with waters at room temperature; TCEA-IRMS; calibration to non-keratin organic standards (cellulose and PEF-1)
	UH	−141.6					
OldUT_H_2	FH	−52.9	0.6	6	US_H_1	Qi and Coplen (2011), Coplen and Qi (2012)	Co-analysis with US_H_1; TCEA-IRMS; calibration to US_H_1
	UH	−117	1.7	6			
OldUT_H_3	FH	−58.8			OldUT_H_3	L.A. Chesson and J.D. Howa (pers. comm.)	Equilibration with waters at room temperature; TCEA-IRMS; calibration to non-keratin organic standards (dual inlet calibrated *n*-C24 and *n*-C28)
	UH	−128.7					

(Continues)

TABLE 1 (Continued)

Calibration	Standard	δ^2H (mean)	δ^2H (*SD*)	*n*	Reference scale	Citation	Methods
OldUT_H_4	FH UH	−60.2 −130.1			OldUT_H_4	Table S3	Equilibration with waters at room temperature; TCEA-IRMS; calibration to non-keratin organic standards (continuous flow calibrated *n*-C24 and *n*-C28)
UT_H_1	ORX DS	−34 −172.7			UT_H_1	L.A. Chesson and J.D. Howa (pers. comm.)	Equilibration with waters at room temperature; TCEA-IRMS; calibration to non-keratin organic standards (dual inlet calibrated *n*-C24 and *n*-C28)
UT_H_2	ORX DS	−35.4 −174.1			UT_H_2	Table S3	Equilibration with waters at room temperature; TCEA-IRMS; calibration to non-keratin organic standards (continuous flow calibrated *n*-C24 and *n*-C28); calibrated at IsoForensics Inc.
UT_H_3	ORX DS	−51.2 −200			UT_H_3	J.D. Howa (pers. comm.)	Equilibration with waters at room temperature; TCEA-IRMS; calibration to non-keratin organic standards (continuous flow calibrated *n*-C24 and *n*-C28); calibrated at Cornell University stable isotope laboratory
UT_H_4	ORX DS	−47.2 −195.9			UT_H_4	J.D. Howa (pers. comm.)	Co-analysis with EC_H_5; TCEA-IRMS; calibration to EC_H_5; calibrated at UC Davis stable isotope facility; no uncertainty reported
UT_H_5	ORX DS	−43.1 −191.7			UT_H_5	J.D. Howa (pers. comm.)	Equilibration with waters at room temperature; TCEA-IRMS; calibration to non-keratin organic standards (continuous flow calibrated *n*-C24 and *n*-C28); calibrated at Texas A&M University stable isotope laboratory
CAN_H_1	AND CAL-CAN	−71.6 −106.8			CAN_H_1	Meier-Augenstein et al. (2011)	Equilibration with waters at room temperature; TCEA-IRMS; calibration to non-keratin organic standards (PEF-1 and KGa-1); calibrated at University of Ottawa Jan Veizer stable isotope laboratory; mean values obtained with different desiccators
CAN_H_2	AND CAL-CAN	−70.6 −103.2			CAN_H_2	Meier-Augenstein et al. (2011)	Equilibration with waters at room temperature; TCEA-IRMS; calibration to non-keratin organic standards (PEF-1 and KGa-1); calibrated at University of Ottawa Jan Veizer stable isotope laboratory; values obtained with plastic desiccator with lubricant
CAN_H_3	AND CAL-CAN	−72.9 −105.9			CAN_H_3	Meier-Augenstein et al. (2011)	Equilibration with waters at room temperature; TCEA-IRMS; calibration to non-keratin organic standards (PEF-1 and coumarin); calibrated at James Hutton Institute stable isotope unit in Dundee (UK); values obtained with glass desiccator evacuated and filled with Sicapent
CAN_H_4	AND CAL-CAN	−70.4 −109.8	1.1 0.8	10 10	US_H_1	Meier-Augenstein et al. (2011)	Equilibrated with waters at high temperature; TCEA-IRMS; calibration to US_H_1; calibrated at USGS Reston stable isotope laboratory
CAN_H_5	AND CAL-CAN	−67.9 −103			CAN_H_5	Qi and Coplen (2011), Coplen and Qi (2012)	Co-analysis with US_H_1; TCEA-IRMS; calibration to US_H_1; PIT not met
US_H_1	USGS43 USGS42	−50.3 −78.5			US_H_1	Qi and Coplen (2011), Coplen and Qi (2012), Soto et al. (2017)	Equilibration with waters at room temperature for 6 days, evacuated using different methods, assumed $\varepsilon = 0‰$; TCEA-IRMS with glassy C-filled reactor; calibration to water standards
US_H_2	USGS43 USGS42	−58.5 −88.3			US_H_2	Qi and Coplen (2011), Soto et al. (2017)	Equilibration with waters at 120°C, evacuated using different methods, assumed $\varepsilon = 0‰$; TCEA-IRMS with glassy C-filled reactor; calibration to water standards
US_H_3	USGS43 USGS42	−48.7 −75.7			US_H_3	Wassenaar et al. (2015), Soto et al. (2017)	Equilibration with waters at 70°C for 2 hr, evacuated and flushed with He, assumed $\varepsilon = 80‰$; TCEA-IRMS with glassy C-filled reactor; calibration to water standards

(Continues)

TABLE 1 (Continued)

Calibration	Standard	δ^2H (mean)	δ^2H (*SD*)	*n*	Reference scale	Citation	Methods
US_H_4	USGS43 USGS42	−60.3 −89.6			US_H_4	Wassenaar et al. (2015), Soto et al. (2017)	Equilibration with waters at 105°C for 2 hr, evacuated and flushed with He, assumed $\varepsilon = 80‰$; TCEA-IRMS with glassy C-filled reactor; calibration to water standards
US_H_5	USGS43 USGS42	−57.5 −88.7	2 1.7	87 104	UT_H_2	Table S3	Co-analysis with UT_H_2; TCEA-IRMS; calibration to UT_H_2
US_H_6*	USGS43 USGS42	−44.4 −72.9			VSMOW_H	Coplen and Qi (2016)	Equilibration with waters at room temperature for 6 days, dried in vacuum oven at 60°C or in glass desiccators with Sicapent, assumed $\varepsilon = 0‰$; TCEA-IRMS with Cr-filled reactor; calibration to water standards
US_H_7	USGS43 USGS42	−44.2 −72.2	0.9 1	5 5	US_H_7	Soto et al. (2017)	Equilibration with waters over a range of temperatures (25–105°C) and drying procedures, assumed $\varepsilon = 0‰$; TCEA-IRMS with Cr-filled reactor; calibration to water standards (optimal method)
VSMOW_H	VSMOW SLAP	0 −428			VSMOW_H	https://nucleus.iaea.org/rpst/documents/VSMOW_SLAP.pdf	

Note: Certified reference material (CRM) calibrations are noted by *.

2.1.2 | Oxygen secondary standards

A smaller number of secondary standards have been widely used for keratin O isotope analysis (Table 2). The International Atomic Energy Agency (IAEA) prepared two benzoic acid (IAEA-601 and IAEA-602) standards for use in organic O analysis. Values based on a series of calibrations to water primary standards and inorganic secondary standards at numerous laboratories have been published (IAEA_O_1 and 3–8; Brand et al., 2009; Schimmelmann, 2002), and a consensus value has been adopted based on an inter-laboratory average (IAEA_O_2; Brand et al., 2009, 2014). Although we include the benzoic acids here, we note that these are not matrix-matched standards for keratins and are not calibrated to VSMOW following the criteria that we have accepted as optimal. Thus, we discourage their use in keratin analysis and suggest caution in the interpretation of transformations involving these standards.

Several calibrations for the Environment Canada standard materials have been reported based on non-PIT, non-optimal methods (OldEC.3_O_1 and EC_O_9 and 10). These were also calibrated to VSMOW using optimal methods (Qi et al., 2011). Certified values for KHS and CBS (EC_O_9; https://isotopes.usgs.gov/lab/referencematerials/KHS.pdf; https://isotopes.usgs.gov/lab/referencematerials/CBS.pdf; Wassenaar & Hobson, 2010) differed significantly from values for the same standards determined with optimal methods (EC_O_11), and here we adopt the latter as the most robust calibration to VSMOW. Laboratories have used different combinations of these standards in sample analysis, including CHS and CFS (OldEC.3_O_2), BWB and CFS (OldEC.2_O_2) and KHS and CBS (EC_O_11).

The FH and UH standards were initially calibrated using non-optimal, non-PIT methods (OldUT_O_1; Bowen, Chesson, et al., 2005; L.A. Chesson & J.D. Howa, pers. comm.). They were later PIT calibrated to the US_O_1 reference scale (see below; OldUT_O_2; Coplen & Qi, 2012), and a calibration to the IAEA_O_1 scale was also conducted (OldUT_O_4; Table S4). The ORX and DS standards have been calibrated to the IAEA_O_1 scale at two laboratories (UT_O_2 and 4; J.D. Howa, pers. comm.), and to the EC_O_11 scale at Texas A&M University (UT_O_5; J.D. Howa, pers. comm.). Because the UT_O_4 and UT_O_5 calibrations lack uncertainty estimates, we treat them as floating calibrations. Although they reflect neither optimal protocols nor PIT conditions, we accept the OldUT_O_4 and UT_O_2 calibrations as the best available calibrations between the IAEA benzoic acids and keratin secondary standards, and use them for transformations involving the benzoic acid standards.

The AND and CAL-SAL standards were calibrated to waters using non-optimal (CAN_O_6) and optimal (CAN_O_7) analytical methods (Qi et al., 2011). They were also PIT calibrated to the US_O_1 reference scale (CAN_O_5; Coplen & Qi, 2012).

The USGS42 and USGS43 standards were calibrated to waters using non-PIT, non-optimal methods (US_O_8) as well as using optimal methods (US_O_1; Coplen & Qi, 2012; Qi et al., 2011), giving values adopted in the USGS-issued certificates for these materials (https://isotopes.usgs.gov/lab/referencematerials/USGS42.pdf; https://isotopes.usgs.gov/lab/referencematerials/USGS43.pdf). USGS42 and USGS43 were also calibrated to the UT_O_2 reference scale (US_O_5; Table S4).

TABLE 2 Calibrated $\delta^{18}O$ values for most commonly used secondary and primary standards for keratin O isotope analysis

Calibration	Standards	$\delta^{18}O$ (mean)	$\delta^{18}O$ (*SD*)	*n*	Reference scale	Citation	Methods
OldEC.2_O_2	BWB	12.82	0.13	3	VSMOW_O	Qi et al. (2011)	HTC-IRMS with new GC column; calibration to water standards sealed in silver tubes (optimal method)
	CFS	5.24	0.02	3			
OldEC.3_O_1	CHS	5.59			OldEC.3_O_1	Qi et al. (2011)	HTC-IRMS with new GC column, peak detected on mass 28; calibration to water standards sealed in silver tubes
	CFS	5.38					
OldEC.3_O_2	CHS	5.31	0.19	3	VSMOW_O	Qi et al. (2011)	HTC-IRMS with new GC column; calibration to water standards sealed in silver tubes (optimal method)
	CFS	5.24	0.02	3			
EC_O_9*	KHS	20.3			EC_O_9	Wassenaar and Hobson (2010)	HTC-IRMS with new GC column; calibration to IAEA_O_1; no drying method reported
	CBS	3.8					
EC_O_10	KHS	21.46			EC_O_10	Qi et al. (2011)	HTC-IRMS with new GC column, peak detected on mass 28; calibration to water standards sealed in silver tubes
	CBS	2.5					
EC_O_11	KHS	21.21	0.17	4	VSMOW_O	Qi et al. (2011)	HTC-IRMS with new GC column; calibration to water standards sealed in silver tubes (optimal method)
	CBS	2.39	0.13	4			
OldUT_O_1	FH	14.85			OldUT_O_1	Bowen, Chesson, et al. (2005), L.A. Chesson and J.D. Howa (pers. comm.)	TCEA-IRMS; calibration to non-keratin organic standards (cellulose and PEF-1)
	UH	5.67					
OldUT_O_2	FH	13.22	0.21	6	VSMOW_O	Coplen and Qi (2012)	TCEA-IRMS; calibration to US_O_1
	UH	3.49	0.16	6			
OldUT_O_4	FH	14.67	0.27	48	IAEA_O_1	Table S4	TCEA-IRMS; calibration to IAEA_O_1
	UH	5.35	0.18	48			
UT_O_2	ORX	25.09	0.3	49	IAEA_O_1	Table S4	TCEA-IRMS; calibration to IAEA_O_1; calibrated at IsoForensics Inc.
	DS	6.02	0.24	50			
UT_O_4	ORX	22.82			UT_O_4	J.D. Howa (pers. comm.)	TCEA-IRMS; calibration to IAEA_O_1; calibrated at UC Davis stable isotope facility
	DS	4.66					
UT_O_5	ORX	23.45			UT_O_5	J.D. Howa (pers. comm.)	TCEA-IRMS; calibration to EC_O_7; no uncertainty reported
	DS	4.93					
CAN_O_6	AND	12.26			CAN_O_6	Qi et al. (2011)	HTC-IRMS with new GC column, peak detected on mass 28; calibration to water standards sealed in silver tubes
	CAL-SAL	5.95					
CAN_O_7	AND	12.22	0.12	3	VSMOW_O	Qi et al. (2011)	HTC-IRMS with new GC column; calibration to water standards sealed in silver tubes (optimal method)
	CAL-SAL	5.97	0.08	3			
CAN_O_5	AND	23.45			VSMOW_O	Coplen and Qi (2012)	TCEA-IRMS; calibration to US_O_1
	CAL-SAL	5.49					
US_O_1*	USGS43	14.11	0.10	18	VSMOW_O	Qi et al. (2011), Coplen and Qi (2012)	TCEA-IRMS with new GC column; calibration to water standards sealed in silver tubes (optimal method)
	USGS42	8.56	0.10	18			
US_O_5	USGS43	15.96	0.41	44	IAEA_O_1	Table S4	TCEA-IRMS; calibrated to UT_O_2
	USGS42	9.94	0.36	56			
US_O_8	USGS43	14.4			US_O_8	Qi et al. (2011)	HTC-IRMS with new GC column, peak detected on mass 28; calibration to water standards sealed in silver tubes
	USGS42	8.85					
IAEA_O_1	IAEA601	23.3			IAEA_O_1	Schimmelmann (2002), Brand et al. (2009)	Online (Max Planck Institute for Biogeochemistry) and offline (Indiana University stable isotope research facility) methods; average values
	IAEA602	71.4					

(Continues)

TABLE 2 (Continued)

Calibration	Standards	$\delta^{18}O$ (mean)	$\delta^{18}O$ (*SD*)	*n*	Reference scale	Citation	Methods
IAEA_O_2	IAEA601 IAEA602	23.14 71.28			IAEA_O_2	Brand et al. (2009, 2014)	HTC-IRMS methods; calibration to water and barium sulphates standards
IAEA_O_3	IAEA601 IAEA602	23.02 71.26			IAEA_O_3	Brand et al. (2009, 2014)	TCEA-IRMS; calibration to water and inorganic standards; calibrated at USGS Reston stable isotope laboratory
IAEA_O_4	IAEA601 IAEA602	23.14 71.24			IAEA_O_4	Brand et al. (2009, 2014)	HTC-IRMS; calibration to water and inorganic standards; calibrated at Eidgenoessische Technische Hochschule Zurich stable isotope laboratory
IAEA_O_5	IAEA601 IAEA602	23.03 71.17	1.1 0.8	10 10	IAEA_O_5	Brand et al. (2009, 2014)	HTC-IRMS; calibration to water and inorganic standards; calibrated at University of Groningen Centre for Isotope Research
IAEA_O_6	IAEA601 IAEA602	22.95 71.01			IAEA_O_6	Brand et al. (2009, 2014)	HTC-IRMS; calibration to water and inorganic standards; calibrated at Max Planck Institute for Biogeochemistry
IAEA_O_7	IAEA601 IAEA602	23.43 71.01			IAEA_O_7	Brand et al. (2009, 2014)	HTC-IRMS; calibration to water and inorganic standards; calibrated at Helmholtz Centre for Environmental Research isotope laboratory
IAEA_O_8	IAEA601 IAEA602	23.48 71.31			IAEA_O_8	Brand et al. (2009, 2014)	HTC-IRMS; calibration to water and inorganic standards; calibrated at Australian National University stable isotope laboratory
VSMOW_O	VSMOW SLAP	0 −55.5			VSMOW_O	https://nucleus.iaea.org/rpst/documents/VSMOW_SLAP.pdf	

Note: CRM calibrations are noted by *.

2.2 | Secondary standard database

We compiled summary data and methodological information for each secondary standard calibration (Tables 1 and 2), a description of all unpublished calibration data (Table S1; Table 2), and, when available, the raw unpublished calibration data themselves (Tables S3 and S4).

A summary of the calibrations was added as a new object *stds* within the ASSIGNR R-package (Ma et al., 2020). This list object contains two data frames (*hstds* and *ostds*) that record assigned secondary standard values, uncertainty and calibration methods for each calibration. A *Ref_scale* field records the reference scale to which the assigned values are calibrated (either another scale, or, for a floating calibration, the calibration itself). Two adjacency matrices (*ham* and *oam*) record which scales are linked, either because they share a common reference scale or because they represent different assigned values for the same physical materials.

2.3 | Known-origin database

We updated ASSIGNR's database of known-origin tissue samples by adding additional keratin data and information on sample preparation and analysis. The new *knownOrig* database consists of three objects.

knownOrig$sources includes attribution and methodological information, where available, for the compiled datasets (see Table S5). Documentation includes (a) sample type (e.g. feather); (b) sampling method (e.g. vane vs. rachis); (c) whether the samples were powdered (Y/N); (d) whether lipids (i.e. surface oils) were extracted (Y/N) and (e) lipid extraction method (e.g. 2:1 chloroform:methanol); (f) whether H exchange was corrected for (Y/N) and (g) the H exchange correction method (multiple waters/comparative equilibration); (h) equilibration temperature (high/ambient); (i) the reference scale originally used for data calibration (e.g. OldEC.1_H_1); (j) whether the standards used for calibration were powdered (Y/N); (k) whether procedures that limit the contribution of adsorbed water to the analyses, either through dedicated sample preparation devices (Soto et al., 2017; Wassenaar et al., 2015) or careful drying and rapid handing of dried samples (e.g. Bowen, Chesson, et al., 2005), were used (Y/N); (l) the analysis method (TCEA/H device, Cr vs. glassy carbon reduction) and (m) the analysis type (H/HO).

knownOrig$sites is a spatial object that records location information for all sample collection sites, including geographical coordinates, site name, state and/or country.

knownOrig$samples is a data frame containing information about taxonomy, age class, material type (e.g. hair) and matrix (e.g. keratin),

and δ^2H and/or δ^{18}O values and uncertainty for each known-origin sample. This object can be linked to the *knownOrig$sources* object through the *Dataset_ID* field and to the *knownOrig$sites* object through the *Site_ID* field. If available, analytical uncertainty is calculated from replicate analyses of the sample itself (we do not distinguish between replicates from a single or multiple analytical runs, as this is often not reported). In most cases, information on sample replicates was not available and we record a generic estimate of reproducibility based on replicate analyses of a quality control material.

2.4 | Data analysis

2.4.1 | Recalibration functions

A new ASSIGNR function *refTrans* identifies and conducts transformations among reference scales based on the above-compiled calibration data. Briefly, a breadth-first algorithm is applied to the H or O adjacency matrix to identify the shortest chain linking the starting and target reference scales. This chain is composed of a sequence of calibrations, each linked to the next through a shared reference scale (i.e. two sets of secondary standards, one calibrated to the other) or a shared set of standards (i.e. two calibrations for the same physical materials). Links involving a single set of standards imply a reassignment of values and addition of uncertainty associated with the new calibration. For each of such links, the algorithm calculates a linear scale transformation relating δ values on the new and old reference scales (Equations 1–3):

$$\delta_{\text{new}} = \delta_{\text{old}} \times m_{\text{new-old}} + b_{\text{new-old}}, \quad (1)$$

$$m_{\text{new-old}} = \frac{SH_{\text{new}} - SL_{\text{new}}}{SH_{\text{old}} - SL_{\text{old}}}, \text{ and} \quad (2)$$

$$b_{\text{new-old}} = SH_{\text{new}} - SH_{\text{old}} \times m_{\text{new-old}}, \quad (3)$$

where m and b refer to the slope and intercept of the transformation and *SH* and *SL* new and old are the high and low values assigned to the standards in the new and old calibrations, respectively. Uncertainty is propagated by randomly sampling (default = 5,000 times) from normal distributions defined by the means and standard errors of (a) the reported sample values; (b) the standard values for the original reference scale, if they have associated uncertainty and (c) each calibrated standard value in the chain with associated uncertainty; and refitting the transformation equations. Because floating calibrations are self-referenced, they have no associated uncertainty; calibrations that are linked to a second reference scale add uncertainty associated with the underlying cross-calibration.

These functions allow transformations between most reference scales compiled here. We used calibration chains (Figures S3 and S4) to transform all compiled datasets to the VSMOW reference scale for analysis.

2.4.2 | Tissue–water relationships

We compared within-species site-average keratin values for known-origin samples with local precipitation amount-weighted annual average δ^2H and δ^{18}O values extracted from precipitation isoscapes (http://www.waterisotopes.org; Bowen & Revenaugh, 2003; Bowen, Wassenaar, et al., 2005). These values offer a standardized, first-order estimate of spatial variation in local environmental water isotope ratios that can be used to characterize tissue–water relationships across different groups of taxonomically and/or ecologically related species. We derived such relationships using both original and recalibrated keratin data.

2.4.3 | Validation

We validated the quality of the calibration chain transformations using several datasets in which data for the same or related samples were originally reported on different reference scales. These include modern human hair δ^2H data reported on the OldUT_H_1 (Ehleringer et al., 2008; Thompson et al., 2010) and CAN_H_1 scales (Bataille et al., 2020; C.P. Bataille, pers. comm.), a collection of Lesser scaup (*Aythya affinis*) feathers calibrated to the OldEC.1_H_1 (Hobson et al., 2009) and OldUT_H_1 (G.J. Bowen, pers. comm.) scales, and data from taxonomically and ecologically similar birds reported on several different scales. Our expectation was that values for the scaup samples would be more similar after transformation to a common scale, and that the tissue–water relationship for other ecologically related samples would be more uniform following transformation. For two sample groups (modern humans and ground-foraging non-passerine birds), enough data were available to allow statistical testing. Levene's test was used to assess whether residual variance from tissue–water regressions was reduced using different regressions for data originally calibrated to different reference scales; the test was repeated for pre- and post-transformation data.

3 | RESULTS

3.1 | Recalibration

3.1.1 | Hydrogen

The compiled known-origin dataset includes 935 human hair and 3,075 bird feather samples analysed for δ^2H values. Hair data were originally reported on the OldUT_H_1 or CAN_H_1 reference scales. The majority of the feather data were referenced to the OldEC.1_H_1 scale, although five other scales were represented (Figure 1). None of the data were reported on scales that were directly traceable to the VSMOW reference scale based on the criteria used here (e.g. EC_H_9 and US_H_6), so all values were shifted during transformation. Most sample values are higher after transformation, but the magnitude of

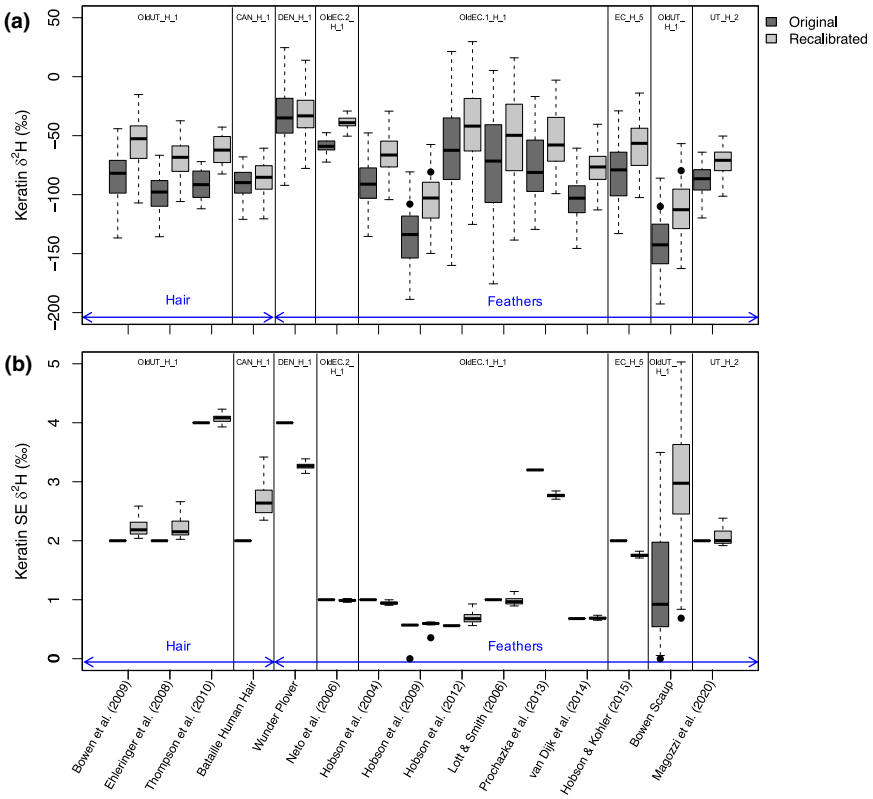

FIGURE 1 Boxplots of original and recalibrated δ^2H values (a) and corresponding standard errors (b). Vertical grey lines separate data based on the original reference scale (annotation). Black dots represent values for a quality control keratin sample (BWB) analysed in two different laboratories

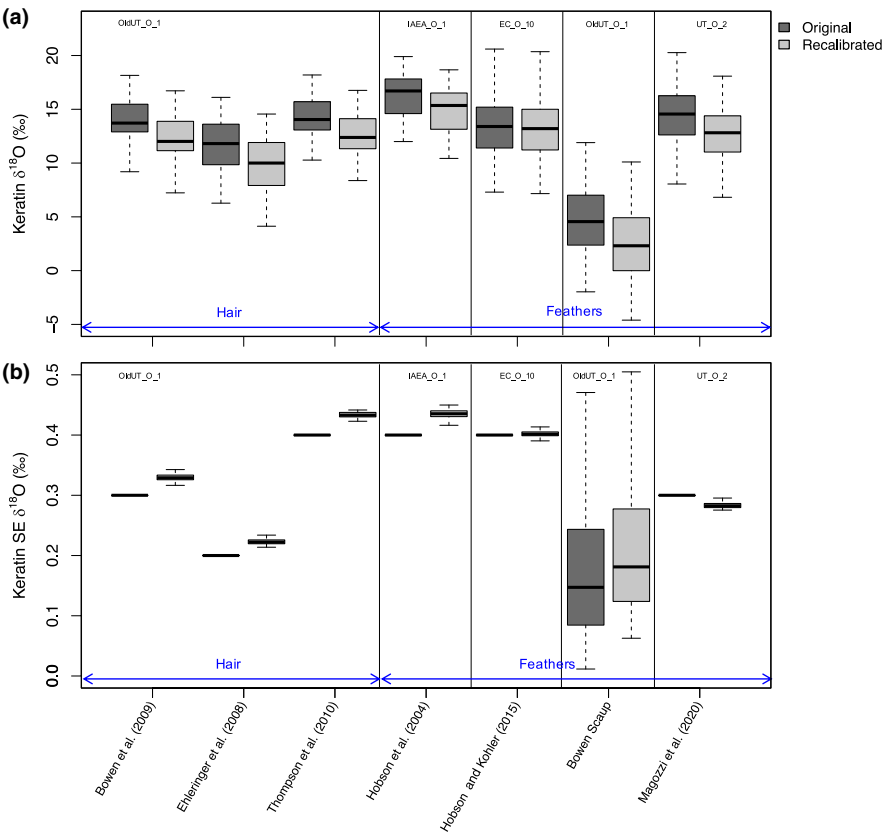

FIGURE 2 Boxplots of original and recalibrated δ^{18}O values (a) and corresponding standard errors (b). Vertical grey lines separate data based on the original reference scale (annotation)

change varies greatly among scales (Figure 1a). For most datasets, the range of recalibrated values is compressed compared to that of the original data, with the exception of data calibrated to the CAN_H_1 scale. The largest contributors to scale compression are generally the calibrations of the keratins to the VSMOW scale based on optimal methods. The estimated uncertainty of transformed sample values

reflects a combination of the originally reported analytical uncertainty, contraction (or expansion) of the δ^2H scale during transformation and uncertainty added within the calibration chain. Reported uncertainties varied widely among datasets, and uncertainties converge somewhat following transformation: δ^2H scale compression drives a reduction in estimated uncertainty for many samples with high reported uncertainty, and the addition of uncertainty from transformation dominates for those with low reported uncertainty (Figure 1b).

3.2 | Oxygen

The compilation contains 358 human hair and 337 bird feather samples analysed for δ^{18}O values. All hair δ^{18}O data used the OldUT_O_1 reference scale. Feather δ^{18}O data were referenced to four different scales, including one dataset using the benzoic-acid-based IAEA_O_1 scale (Figure 2). As for δ^2H values, all δ^{18}O datasets required transformation to be linked to the VSMOW reference scale. Transformed values are somewhat lower than values calibrated to the original scales, but scale compression/expansion during the transformations is more modest and less uniform than for δ^2H values (Figure 2a). Estimated uncertainty is slightly increased for most recalibrated datasets, with the exception of those originally reported on the IAEA_O_1 and UT_O_2 scales where scale compression drives small reductions in uncertainty (Figure 2b).

4 | DISCUSSION

4.1 | Validation

Calibration chain transformation leverages cross-calibration of keratin standards to develop scale transformations that are intended to improve data comparability. Examples in which known or presumed relationships exist between data originally calibrated to different scales allow us to test for improved comparability of transformed data and evaluate that improvement relative to other sources of variability among datasets.

We used the method to transform two sets of modern human hair δ^2H data originally calibrated to the OldUT_H_1 (Ehleringer et al., 2008; Thompson et al., 2010) and CAN_H_1 (Bataille et al., 2020; C.P. Bataille, pers. comm.) reference scales. Although we expect some regional variation in hair isotope ratios due to dietary differences (Bowen et al., 2009), hair δ^2H values are known to correlate strongly with local environmental water values (Ehleringer et al., 2008). Before recalibration, values in the CAN_H_1 dataset, consisting of samples from Canadian residents, were ~8‰ higher, on average, than those calibrated to OldUT_H_1, which included samples from the United States and east Asia (Figure 1a). This pattern is opposite to that expected based on water δ^2H values for these regions (Bowen, Wassenaar, et al., 2005), and regression relationships between the two groups of data and local precipitation δ^2H values were statistically distinct (Levene's test p value << 0.05; Figure 3a,b). After transformation to the VSMOW reference scale, values in the Canadian dataset are lower than those for USA/Asia, as expected (Figure 1a), and the hair–water relationships are no longer distinct (Levene's test p value = 0.96; Figure 3a,b). This strongly suggests that the originally calibrated data were not comparable, and that the calibration chain transformation eliminates (or greatly reduces) the disparity between these datasets.

We also evaluated transformation effects on multi-species data for two broadly defined ecological guilds of birds. The vast majority of passerine δ^2H values in the database were originally calibrated to the OldEC.1_H_1 and EC_H_5 reference scales. However, the compilation contains a dataset for a population of spotted towhees (*Pipilo maculatus*) from a single site in Utah (Magozzi et al., 2020) that was calibrated to UT_H_2. Both pre- and post-transformation, the towhee site-average value clusters well with other passerine data from environments with similar water δ^2H values, but the post-transformation value falls closer to the mean tissue–water relationship for the composite dataset (Figure 3c,d). The same is true for pre- and post-transformation δ^{18}O data for passerine feathers originally calibrated to IAEA_O_1, UT_O_1 and EC_O_10 scales (Figure 4c,d). These are relatively weak tests in that we lack a firm basis for predicting expected differences between species, but the results are consistent with the idea that calibration chain transformation increases comparability among datasets.

Values of δ^2H for ground-foraging non-passerine birds originally calibrated to DEN_H_1 (Wunder et al., 2005; M.B. Wunder, pers. comm.) and OldEC.1_H_1 (Hobson et al., 2004) defined two discrete clusters when plotted against local water values (Figure 3e). Following transformation, the offset between these groups is reduced but not eliminated (Levene's test p value pre-transformation = $5e^{-4}$; post-transformation = $4e^{-3}$; Figure 3f). In this case, the residual offset may represent real, ecologically driven differences among taxa, rather than an analytical artefact. The DEN_H_1-calibrated data represent a single species (*Charadrius montanus*) that occupies dry grasslands with sparse vegetation cover, in which evaporative isotope effects might lead to higher food web δ^2H values (e.g. Magozzi et al., 2019) than for the other species represented in the database.

The database includes only one case in which the same samples were calibrated to two different scales. Lesser scaup feathers analysed and calibrated to OldEC.1_H_1 (Hobson et al., 2009) were subsequently reanalysed at Purdue University relative to OldUT_H_1 (G.J. Bowen, pers. comm.); the data show a small but consistent (mean = 6.7‰) offset. In this case, the transformation results in almost no relative shift in values (Figure 1a). The BWB secondary standard was analysed alongside the feathers at Purdue, and both the originally calibrated and transformed data show good agreement with its value on the OldEC.1_H_1 scale. This suggests that the original reference scales are themselves closely comparable, and that the small offset in the scaup sample data might result from other methodological effects. One possibility may be inaccurate correction for H exchange due to differences in the physical condition (powdered vs. cut) of the samples and standards (Coplen & Qi, 2012) or difference in their amino acid composition. Thus, these data highlight the potential importance of

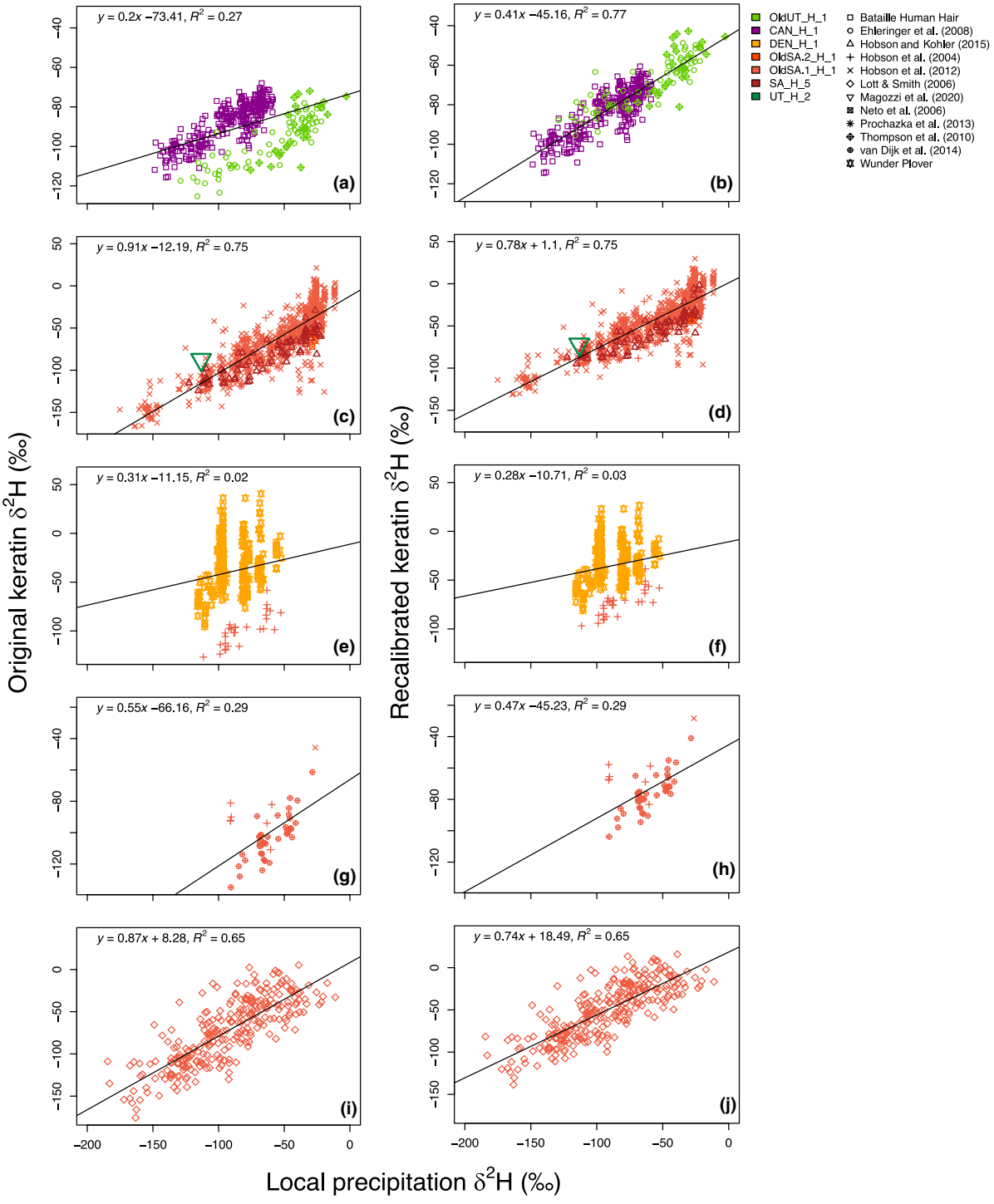

FIGURE 3 Comparison of original (left panels) and VSMOW-recalibrated (right panels) site-average keratin and local precipitation δ^2H values for taxonomically and/or ecologically related animals: modern humans (USA, Canada and Asia; a and b), passerines (c and d), ground-foraging non-passerine birds (e and f), waterbirds (g and h) and raptors (i and j). Local precipitation δ^2H values are extracted from the precipitation amount-weighted annual average δ^2H isoscape on http://www.waterisotopes.org (Bowen & Revenaugh, 2003; Bowen, Wassenaar, et al., 2005). Colours represent different reference scales used in calibration of the original data, symbols represent different datasets

minor deviations from PIT, which remain common in analytical work with complex organic materials, as an unresolved source of uncertainty in data compilations. Standardization of protocols across laboratories will be needed to reduce or eliminate this uncertainty.

4.2 | Tissue–water relationships

Differences in tissue–water isotope relationships may reflect environmentally or biologically controlled isotope effects during

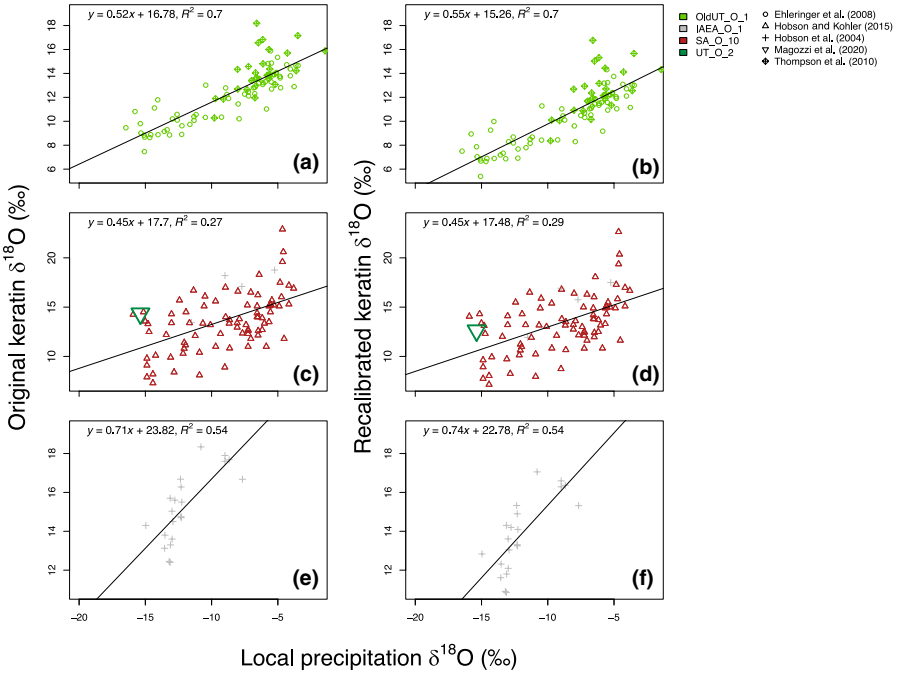

FIGURE 4 Comparison of original (left panels) and VSMOW-recalibrated (right panels) site-average keratin and local precipitation $\delta^{18}O$ values for taxonomically and/or ecologically related animals: modern humans (USA; a and b), passerines (c and d) and ground-foraging non-passerine birds (e and f). Symbology and precipitation estimation as in Figure 3

the routing of elements from precipitation through hydrological and ecological systems to consumer tissues (Hobson et al., 2012; Magozzi et al., 2019; Meehan et al., 2003). Questions about the comparability of data from different studies have previously limited the extent to which cross-group comparisons could be made. Here we explore such comparisons based on the VSMOW-transformed keratin data. We caution that the compiled data are a non-random sample, and comparisons may be complicated by differences in geographical range, body size and other ecophysiological characteristics. Regardless, several patterns consistent with first-order expectations emerge, supporting the idea that differences among groups may be meaningful and, to some degree, predictable.

Slopes of the tissue–water relationships reflect the degree to which geographical differences in water isotope ratios are transferred to local consumers as opposed to being damped by H or O from sources not tied to the local water signature (e.g. atmospheric O_2 and stored or transported non-local resources). For δ^2H values, the slopes for all avian groups except ground-foraging non-passerine birds (for which weak correlation lowers the OLS regression slope) are similar, and slopes for all bird groups are substantially higher than that for humans (Figure 3). This is consistent with the idea that widespread consumption of non-local food resources dampens the isotopic variability that would otherwise be expected in human hair due to geography (Bowen et al., 2009; Ehleringer et al., 2008). Slope comparisons for $\delta^{18}O$ values are complicated by the smaller sample sizes and weaker tissue–water correlations for avian groups, but at minimum suggest that human–avian differences are less apparent than for δ^2H values (Figure 4). The smaller contribution of food-derived O (relative to H) to keratin (Ehleringer et al., 2008) is consistent with this result.

The data also show differences in keratin isotope ratios between different groups across a range of water isotopic compositions (Figures 3 and 4). Values of δ^2H for waterbirds are lowest, and for raptors highest, across most of the sampled range (Figure 3). The close association of waterbirds with aquatic systems may reduce evaporation-driven ^2H-enrichment in the food webs on which they depend (Hobson et al., 2012 and references therein). In contrast, the longer food chain of raptors involves additional potential for evaporation and dilution of ^2H-depleted organic H by relatively ^2H-enriched environmental water (Magozzi et al., 2019). Enrichment of ^2H in feather keratin in ground-foraging relative to shrub- and canopy-feeding birds, reported by Hobson et al. (2012), is apparent here for mountain plovers but not for other ground-foraging non-passerine species (Figure 3). This may reflect the dry, high elevation habitat occupied by the plovers, characterized by high levels of evapotranspiration. Human hair δ^2H values are low to intermediate, and $\delta^{18}O$ values quite low, relative to all avian groups (Figures 3 and 4). Dietary effects (dominantly non-aquatic food resources, low to intermediate trophic level) might be expected to give intermediate values for human hair δ^2H, whereas the large body size and relatively low rate of evaporative water loss relative to other body water fluxes, plus high drinking water consumption, could contribute to reducing both δ^2H and $\delta^{18}O$ values of human keratin samples (Kohn, 1996; Magozzi et al., 2019).

5 | CONCLUSIONS

Although keratin δ^2H and $\delta^{18}O$ data from known-origin biological samples are important in movement ecology research, they are difficult to compare among studies and reuse due to heterogeneity in analytical methods. The calibration chain method introduced here attempts to resolve one major source of heterogeneity by reducing or eliminating differences related to the use of different

secondary standards in data calibration. Comparisons of pre- and post-transformation data show that the method improves comparability, and suggests systematic isotopic differences between groups of organisms that likely reflect differences in isotope routing and fractionation within food webs. The method is implemented in the ASSIGNR R-package, and provides a basis for improved reuse of the database of >4,000 samples included therein and for transformation of user-generated data. The approach developed here could be extended to other matrices and isotope systems in future work.

We emphasize, however, that post-hoc correction such as that introduced here is not ideal, and encourage continued efforts within the research community to increase the availability of suitable standards and adoption of optimal analytical methods. Standard development has thus far been conducted largely by self-organized groups, leading to the diversity of heterogeneity of materials evidenced here. In addition, other methodological factors continue to contribute uncertainty to comparisons among studies. Some have been more thoroughly discussed in other papers (e.g. Soto et al., 2017), and are documented in the ASSIGNR database but not accounted for in calibration chain transformation. To leverage the method developed here and continue to support improvements in the standardization and comparability of organic H and O isotope data, we suggest the following three priorities for the community:

1. Organize and support coordinated efforts to develop, characterize (including regular round-robin comparisons among laboratories) and distribute large amounts of standard materials for keratin and other commonly studied biological materials (e.g. chitin).
2. Conduct analyses in compliance with PIT principles to the maximum extent possible, and, where identical treatment is not possible, adopt technologies and methodologies that are demonstrated to eliminate analytical effects known to impart bias in non-PIT analyses (e.g. Hunsinger & Stern, 2012; Wassenaar et al., 2015).
3. Ensure that all new data reports include essential quality control and methodological information, such as measured weight % H and O, the identity and accepted values of the standards used, and details of sample preparation, handling and drying, needed to assess and conduct post-hoc re-evaluation of the results.

ACKNOWLEDGEMENTS

We profoundly thank C Stricker, LI Wassenaar and DX Soto for providing mean δ^2H values and associated standard errors for DEN_H_2 and EC_H_7 and 8 calibrations. Their hard work on producing data for these secondary standards over time has provided the foundation for this valuable study comparing calibrations and recalibrating available sample data to comparable reference scales. We also thank LI Wassenar and an anonymous reviewer for providing constructive feedback on our manuscript.

AUTHORS' CONTRIBUTIONS

S.M. and G.J.B. conceived the ideas and designed the methodology; S.M. compiled and the data; S.M. and G.J.B. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/2041-210X.13556.

DATA AVAILABILITY STATEMENT

The ASSIGNR package version 2.0.0 is available on the CRAN repository: https://cran.r-project.org/web/packages/assignR/index.html. The release (Bowen & Ma, 2021) is archived on Zenodo: https://doi.org/10.5281/zenodo.4430639. Previously unpublished non-aggregated δ^2H and δ^{18}O data (Tables S3 and S4) are archived on Figshare: https://doi.org/10.6084/m9.figshare.12769928.

ORCID

Sarah Magozzi https://orcid.org/0000-0003-4070-9105
Clement P. Bataille https://orcid.org/0000-0001-8625-4658
Keith A. Hobson https://orcid.org/0000-0002-2525-1178
Michael B. Wunder https://orcid.org/0000-0002-8063-2408
John D. Howa https://orcid.org/0000-0002-9855-126X
Andrea Contina https://orcid.org/0000-0002-0484-6711
Hannah B. Vander Zanden https://orcid.org/0000-0003-3366-5116
Gabriel J. Bowen https://orcid.org/0000-0002-6928-3104

REFERENCES

Alibardi, L. (2017). Review: Cornification, morphogenesis and evolution of feathers. *Protoplasma*, *254*, 1259–1281.

Bataille, C. P., Chartrand, M. M. G., Raposo, F., & St-Jean, G. (2020). Assessing geographic controls of hair isotopic variability in human populations: A case-study in Canada. *PLoS ONE*, *15*, e0237105.

Bowen, G. J., Chesson, L., Nielson, K., Cerling, T. E., & Ehleringer, J. R. (2005). Treatment methods for the determination of δ^2H and δ^{18}O of hair keratin by continuous-flow isotope-ratio mass spectrometry. *Rapid Communications in Mass Spectrometry*, *19*, 2371–2378.

Bowen, G. J., Ehleringer, J. R., Chesson, L. A., Thompson, A. H., Podlesak, D. W., & Cerling, T. E. (2009). Dietary and physiological controls on the hydrogen and oxygen isotope ratios of hair from mid-20th century indigenous populations. *American Journal of Physical Anthropology*, *139*, 494–504.

Bowen, G. J., & Ma, C. (2021). SPATIAL-Lab/assignR: Redstart (Version 2.0.0). *Zenodo*, https://doi.org/10.5281/zenodo.4430639

Bowen, G. J., & Revenaugh, J. (2003). Interpolating the isotopic composition of modern meteoric precipitation. *Water Resources Research*, *39*, 1299.

Bowen, G. J., Wassenaar, L. I., & Hobson, K. A. (2005). Global application of stable hydrogen and oxygen isotopes to wildlife forensics. *Oecologia*, *143*, 337–348.

Brand, W. A., Coplen, T. B., Aerts-Bijma, A. T., Bohlke, J. K., Gehre, M., Geilmann, H., Groning, M., Jansen, H. G., Meijer, H. A. J., Mroczowski, S. J., Qi, H., Soergel, K., Stuart-Williams, H., Weise, S. M., & Werner, R. A. (2009). Comprehensive inter-laboratory calibration of reference materials for δ^{18}O versus VSMOW using various on-line high-temperature conversion techniques. *Rapid Communications in Mass Spectrometry*, *23*, 999–1019.

Brand, W. A., Coplen, T. B., Vogl, J., Rosner, M., & Prohaska, T. (2014). Assessment of international reference materials for isotope-ratio analysis (IUPAC Technical Report). *Pure Applied Chemistry*, *86*, 425–467.

Carter, J. F., & Chesson, L. A. (Eds.). (2017). *Food forensics: Stable isotope as a guide to authenticity and origin*. CRC Press.

Chamberlain, C. P., Blum, J. D., Holmes, R. T., Feng, X., Sherry, T. W., & Graves, G. R. (1996). The use of isotope tracers for identifying populations of migratory birds. *Oecologia*, *109*, 132–141. https://doi.org/10.1007/s004420050067

Chesson, L. A., Podlesak, D. W., Cerling, T. E., & Ehleringer, J. R. (2009). Evaluating uncertainty in the calculation of non-exchangeable hydrogen fractions within organic materials. *Rapid Communications in Mass Spectrometry*, *23*, 1275–1280. https://doi.org/10.1002/rcm.4000

Coplen, T. B., & Qi, H. (2012). USGS42 and USGS43: Human hair stable hydrogen and oxygen isotopic reference materials and analytical methods for forensic science and implications for published measurement results. *Forensic Science International*, *214*, 135–141. https://doi.org/10.1016/j.forsciint.2011.07.035

Coplen, T. B., & Qi, H. (2016). A revision in hydrogen isotopic composition of USGS42 and USGS43 human-hair stable isotopic reference materials for forensic science. *Forensic Science International*, *266*, 222–225. https://doi.org/10.1016/j.forsciint.2016.05.029

Craig, H. (1961). Isotopic variations in meteoric waters. *Science*, *133*, 1702. https://doi.org/10.1126/science.133.3465.1702

Dunn, P. J. H., & Carter, J. F. (Eds.). (2018). *Good practice guide for isotope ratio mass spectrometry* (2nd ed.). FIRMS. ISBN: 978-0-948926-33-4.

Ehleringer, J. R., Bowen, G. J., Chesson, L. A., West, A. G., Podlesak, D. W., & Cerling, T. E. (2008). Hydrogen and oxygen isotope ratios in human hair are related to geography. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 2788–2793. https://doi.org/10.1073/pnas.0712228105

Gehre, M., Renpenning, J., Gilevska, T., Qi, H., Coplen, T. B., Meijer, H. A. J., Brand, W. A., & Schimmelmann, A. (2015). On-line hydrogen-isotope measurement of organic samples using elemental chromium: An extension of high temperature elemental-analyzer techniques. *Analytical Chemistry*, *87*, 5198–5205.

Hobson, K. A. (1999). Tracing origins and migration of wildlife using stable isotopes: A review. *Oecologia*, *120*, 314–326. https://doi.org/10.1007/s004420050865

Hobson, K. A., Bowen, G. J., Wassenaar, L. I., Ferrand, Y., & Lormee, H. (2004). Using stable hydrogen and oxygen isotope measurements of feathers to infer geographical origins of migrating European birds. *Oecologia*, *141*, 477–488. https://doi.org/10.1007/s00442-004-1671-7

Hobson, K. A., & Kohler, G. (2015). On the use of stable oxygen isotope ($\delta^{18}O$) measurements for tracking avian movements in North America. *Ecology and Evolution*, *5*, 799–806.

Hobson, K. A., Van Wilgenburg, S. L., Wassenaar, L. I., & Larson, K. (2012). Linking hydrogen ($\delta^{2}H$) isotopes in feathers and precipitation: Sources of variance and consequences for assignment to isoscapes. *PLoS ONE*, *7*, e35137. https://doi.org/10.1371/journal.pone.0035137

Hobson, K. A., & Wassenaar, L. I. (1997). Linking brooding and wintering grounds of neotropical migrant songbirds using stable hydrogen isotopic analysis of feathers. *Oecologia*, *109*, 142–148.

Hobson, K. A., & Wassenaar, L. I. (Eds.). (2019). *Tracking animal migration with stable isotopes*. Elsevier.

Hobson, K. A., Wunder, M. B., Van Wilgenburg, S. L., Clark, R. G., & Wassenaar, L. I. (2009). A method for investigating population declines of migratory birds using stable isotopes: Origins of harvested Lesser Scaup in North America. *PLoS ONE*, *4*, e7915. https://doi.org/10.1371/journal.pone.0007915

Hunsinger, G. B., & Stern, L. A. (2012). Improved accuracy in high-temperature conversion elemental analyzer $\delta^{18}O$ measurements of nitrogen-rich organics. *Rapid Communications in Mass Spectrometry*, *26*, 554–562. https://doi.org/10.1002/rcm.6132

Hunsinger, G. B., Tipple, C. A., & Stern, L. A. (2013). Gaseous byproducts from high-temperature thermal conversion elemental analysis of nitrogen- and sulfur-bearing compounds with considerations for $\delta^{2}H$ and $\delta^{18}O$ analyses. *Rapid Communications in Mass Spectrometry*, *27*, 1649–1659.

Kelly, J. F., Bridge, E. S., Fudickar, A. M., & Wassenaar, L. I. (2009). A test of comparative equilibration for determining non-exchangeable stable hydrogen isotope values in complex organic materials. *Rapid Communications in Mass Spectrometry*, *23*, 2316–2320. https://doi.org/10.1002/rcm.4150

Kohn, M. J. (1996). Predicting animal $\delta^{18}O$ accounting for diet and physiological adaptation. *Geochimica Cosmochimica Acta*, *60*, 4811–4829. https://doi.org/10.1016/S0016-7037(96)00240-2

Lott, C. A., & Smith, J. P. (2006). A Geographic-Information-System approach to estimating the origin of migratory raptors in North America using stable hydrogen isotope ratios in feathers. *The Auk*, *123*, 822–835. https://doi.org/10.1093/auk/123.3.822

Ma, C., Vander Zanden, H. B., Wunder, M. B., & Bowen, G. J. (2020). An R package for isotope-based geographic assignment. *Methods in Ecology and Evolution*, *11*, 996–1001.

Macko, S. A., Engel, M. H., Andrusevich, V., Lubec, G., O'Connel, T. C., & Hedges, E. M. (1999). Documenting the diet in ancient human populations through stable isotope analysis of hair. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *354*, 65–67. https://doi.org/10.1098/rstb.1999.0360

Magozzi, S., Vander Zanden, H. B., Wunder, M. B., & Bowen, G. J. (2019). Mechanistic model predicts tissue-environment relationships and trophic shifts in animal hydrogen and oxygen isotope ratios. *Oecologia*, *191*, 777–789. https://doi.org/10.1007/s00442-019-04532-8

Magozzi, S., Vander Zanden, H. B., Wunder, M. B., Trueman, C. N., Pinney, K., Peers, D., Dennison, P. E., Horns, J. J., Sekercioglu, C. H., & Bowen, G. J. (2020). Combining models of environment, behavior, and physiology to predict tissue hydrogen and oxygen isotope variance among individual terrestrial animals. *Frontiers in Ecology and Evolution*, *8*, 536109. https://doi.org/10.3389/fevo.2020.536109

Meehan, T. D., Rosenfield, R. N., Atudorei, V. N., Bielefeldt, J., Rosenfield, L. J., Stewart, A. C., Stout, W. E., & Bozek, M. A. (2003). Variation in hydrogen stable-isotope ratios between adult and nestling Cooper's Hawks. *The Condor*, *105*, 567–572. https://doi.org/10.1093/condor/105.3.567

Meier-Augenstein, W. (Ed.). (2018). *Stable isotope forensics: An introduction to the forensic application of stable isotope analysis* (2nd ed.). John Wiley & Sons Ltd.

Meier-Augenstein, W., Chartrand, M. M. G., Kemp, H. F., & St-Jean, G. (2011). An inter-laboratory comparative study into sample preparation for both reproducible and repeatable forensic ^{2}H isotope analysis of human hair by continuous flow isotope ratio mass spectrometry. *Rapid Communications in Mass Spectrometry*, *25*, 3331–3338.

Meier-Augenstein, W., Hobson, K. A., & Wassenaar, L. I. (2013). Critique: Measuring hydrogen stable isotope abundance of proteins to infer origins of wildlife, food and people. *Bioanalysis*, *5*, 751–767. https://doi.org/10.4155/bio.13.36

Nair, S., Geilmann, H., Coplen, T. B., Qi, H., Gehre, M., Schimmelmann, A., & Brand, W. A. (2015). Isotopic disproportionation during hydrogen isotopic analysis of nitrogen-bearing organic compounds. *Rapid Communications in Mass Spectrometry*, *29*, 878–884. https://doi.org/10.1002/rcm.7174

Neto, J. M., Newton, J., Gosler, A. G., & Perrins, C. M. (2006). Using stable isotope analysis to determine the winter moult extent in migratory birds: The complex moult of Savi's warblers *Locustella luscinoides*. *Journal of Avian Biology*, *37*, 117–124.

Prochazka, P., Van Wilgenburg, S. L., Neto, J. M., Yosef, R., & Hobson, K. A. (2013). Using stable hydrogen isotopes ($\delta^{2}H$) and ring recoveries to trace natal origins in a Eurasian passerine with a migratory divide. *Journal of Avian Biology*, *44*, 541–550.

Qi, H., & Coplen, T. B. (2011). Investigation of preparation techniques for $\delta^{2}H$ analysis of keratin materials and a proposed analytical protocol. *Rapid Communications in Mass Spectometry*, *25*, 2209–2222.

Qi, H., Coplen, T. B., & Wassenaar, L. I. (2011). Improved online $\delta^{18}O$ measurements of nitrogen- and sulfur-bearing organic materials and a proposed analytical protocol. *Rapid Communications in Mass Spectrometry*, *25*, 2049–2058. https://doi.org/10.1002/rcm.5088

Robbins, C. R. (2012). *Chemical and physical behavior of human hair*. Springer-Verlag Berlin Heidelberg.

Sauer, P. E., Schimmelmann, A., Sessions, A. L., & Topalov, K. (2009). Simplified batch equilibration for D/H determination of non-exchangeable hydrogen in solid organic material. *Rapid Communications in Mass Spectrometry*, *23*, 949–956. https://doi.org/10.1002/rcm.3954

Schimmelmann, A. (1991). Determination of the concentration and stable isotopic composition of nonexchangeable hydrogen in organic matter. *Analytical Chemistry*, *63*, 2456–2459. https://doi.org/10.1021/ac00021a013

Schimmelmann, A. (2002). Final report to the International Atomic Energy Agency (IAEA). Contract No. 11792/R0 2002.

Soto, D. X., Kohler, G., Wassenaar, L. I., & Hobson, K. A. (2017). Re-evaluation of the hydrogen stable isotopic compositions of keratin calibration standards for wildlife and forensic science applications. *Rapid Communications in Mass Spectrometry*, *31*, 119–1203.

Thompson, A. H., Chesson, L. A., Podlesak, D. W., Bowen, G. J., Cerling, T. E., & Ehleringer, J. R. (2010). Stable isotope analysis of modern human hair collected from Asia (China, India, Mongolia, and Pakistan). *American Journal of Physiological Anthropology*, *141*, 440–451.

van Dijk, J. G. B., Meissner, W., & Klaassen, M. (2014). Improving provenance studies in migratory birds when using feather hydrogen stable isotopes. *Journal of Avian Biology*, *45*, 103–108. https://doi.org/10.1111/j.1600-048X.2013.00232.x

Vander Zanden, H. B., Wunder, M. B., Hobson, K. A., Van Wilgenburg, S. L., Wassenaar, L. I., Welker, J. M., & Bowen, G. J. (2014). Contrasting assignment of migratory organisms to geographic origins to using long-term versus year-specific precipitation isotope maps. *Methods in Ecology and Evolution*, *9*, 891–900.

Wassenaar, L. I., & Hobson, K. A. (2000). Improved method for determining the stable-hydrogen isotopic composition (δD) of complex organic materials of environmental interest. *Environmental Science and Technology*, *34*, 2354–2360. https://doi.org/10.1021/es990804i

Wassenaar, L. I., & Hobson, K. A. (2003). Comparative equilibration and online technique for determination of non-exchangeable hydrogen of keratins for use in animal migration studies. *Isotopes in Environmental and Health Studies*, *39*, 21Springer1–217. https://doi.org/10.1080/1025601031000096781

Wassenaar, L. I., & Hobson, K. A. (2010). Two new keratin standards (δ^2H, $\delta^{18}O$) for daily laboratory use in wildlife and forensic isotopic studies. 7th IsoEcol, August 9th-13th 2010.

Wassenaar, L. I., Hobson, K. A., & Sisti, L. (2015). An online temperature-controlled vacuum-equilibration preparation system for the measurement of δ^2H values of non-exchangeable-H and of $\delta^{18}O$ values in organic materials by isotope-ratio mass spectrometry. *Rapid Communications in Mass Spectrometry*, *29*, 397–407.

Werner, R. A., & Brand, W. A. (2001). Referencing strategies and techniques in stable isotope ratio analysis. *Rapid Communications in Mass Spectrometry*, *15*, 501–519. https://doi.org/10.1002/rcm.258

West, A. G., Ayliffe, L. K., Cerling, T. E., Robinson, T. F., Karren, B., Dearing, M. D., & Ehleringer, J. R. (2004). Short-term diet changes revealed using stable carbon isotopes in horse tail-hair. *Functional Ecology*, *18*, 616–624. https://doi.org/10.1111/j.0269-8463.2004.00862.x

Wortmann, F.-J., Augustin, P., & Popescu, C. (2001). Temperature dependence of the water-sorption isotherms of wool. *Journal of Applied Polymer Science*, *79*, 1054–1061. https://doi.org/10.1002/1097-4628(20010207)79:6<1054:AID-APP80>3.0.CO;2-X

Wunder, M. B. (2010). Using isoscapes to model probability surfaces for determining geographic origins. In J. B. West, G. J. Bowen, T. E. Dawson, & K. P. Tu (Eds.), *Isoscapes: Understanding movement, pattern, and process on Earth through isotope mapping* (pp. 251–270). Springer.

Wunder, M. B., Kester, C. L., Knopf, F. L., & Rye, R. O. (2005). A test of geographic assignment using isotope tracers in feathers of known origin. *Oecologia*, *144*, 607–617. https://doi.org/10.1007/s00442-005-0071-y

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Magozzi S, Bataille CP, Hobson KA, et al. Calibration chain transformation improves the comparability of organic hydrogen and oxygen stable isotope data. *Methods Ecol Evol*. 2021;12:732–747. https://doi.org/10.1111/2041-210X.13556