RAPID COMMUNICATIONS IN MASS SPECTROMETRY
Rapid Commun. Mass Spectrom. 2009; **23**: 2316–2320
Published online in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/rcm.4150

RCM

A test of comparative equilibration for determining non-exchangeable stable hydrogen isotope values in complex organic materials

Jeffrey F. Kelly[1]*, Eli S. Bridge[1], Adam M. Fudickar[2] and Leonard I. Wassenaar[3]

[1]Oklahoma Biological Survey and Department of Zoology, 111 East Chesapeake St., University of Oklahoma, Norman, OK, USA
[2]Max Planck Institute for Ornithology, Department of Migration and Immuno-ecology, Vogelwarte Radolfzell, Schlossallee 2, D-78315 Radolfzell, Germany
[3]Environment Canada, 11 Innovation Boulevard, Saskatoon, Saskatchewan, S7N 3H5 Canada

Received 29 April 2009; Revised 3 June 2009; Accepted 10 June 2009

Comparative equilibration has been proposed as a methodological approach for determining the hydrogen isotopic composition (δD) of non-exchangeable hydrogen in complex organic materials, from feathers to blood and soils. This method depends on using homogenized standards that have been previously calibrated for their δD values of non-exchangeable H, that are compositionally similar to unknown samples, and that span an appropriate isotopic range. Currently no certified organic reference materials with exchangeable H exist, and so isotope laboratories have been required to develop provisional internal calibration standards, such as the keratin standards currently used in animal migration studies. Unfortunately, the isotope ratios of some samples fall outside the range of keratin standards currently used for comparative equilibration. Here we tested a set of five homogenized keratin powders as well as feathers from Painted Buntings and Dark-eyed Juncos to determine the effects of extrapolating comparative equilibration normalization equations outside the isotopic range of keratin standards. We found that (1) comparative equilibration gave precise results within the range of the calibration standards; (2) linear extrapolation of normalization equations produced accurate δD results to ~40‰ outside the range of the keratins standards used (−187 to −108); and (3) for both homogenized keratin powders and heterogeneous unknown samples there was no difference in variance between samples within and outside the range of keratin standards. This suggested that comparative equilibration is a robust and practical method for determining the δD of complex organic matrices, although caution is required for samples that fall far outside the calibration range. Copyright © 2009 John Wiley & Sons, Ltd.

Hydrogen stable isotope ratios (δD) of many complex organic materials (collagen, chitin, soil, plants, feathers, etc.) are controlled by the δD of precipitation at locations where they are grown and therefore are useful for inferring both origins and paleo-climatic trends.[1,2] Globally, δD values of precipitation are spatially and temporally predictable.[3] Most notably there are strong long-term latitudinal gradients in the δD values of precipitation, and this makes keratin (e.g. feathers) δD values useful for tracing the origins of long-distance avian migrants that move across large latitudinal δD gradients.[4–6] This utility has led to a huge increase in application of hydrogen stable isotopes to study movement of migrant species over the past decade.[7–9] A challenge for the future of H isotope applications in migratory ecology is to better understand the source and meaning of variance in δD isotope data obtained from various species.[10–12]

One of the first challenges encountered is that complex organic substrates like most biological tissues contain exchangeable H.[13] Left untreated, tissues dynamically exchange a proportion of their H atoms with ambient moisture, which leads to δD results that cannot be compared among laboratories. Some biological substrates (e.g. chitin) are amenable to nitration to remove exchangeable H, but most biological tissues are not. A corresponding challenge was the need for rapid and reproducible high sample throughput with comparably precise results among laboratories. The key innovation to address this challenge was the development of the 'Comparative Equilibration' method.[14] Comparative equilibration uses homogenized powdered complex organic standards that have previously been calibrated offline for their non-exchangeable δD. These standards are then analyzed with 'like' unknown samples and are used to correct measured δD values of bulk tissues

**Correspondence to*: J. F. Kelly, Oklahoma Biological Survey and Department of Zoology, 111 East Chesapeake St., University of Oklahoma, Norman, OK, USA.
E-mail: jeffrey.f.kelly@gmail.com
Contract/grant sponsor: US National Science Foundation; contract/grant number: 0541740.
Contract/grant sponsor: US Environmental Protection Agency; contract/grant number: R 833778.

(e.g. feathers) to determine the δD value of the non-exchangeable fraction of the feather H.[14,15] This method allows for rapid sample throughput and makes it possible to compare measurements within and among different sampling periods within and among laboratories.[2]

Comparative equilibration is typically implemented by measuring unknown complex organic samples and chemically similar calibration standards within the same analysis runs (e.g. keratin standards with keratinous feathers, hair, etc.). A regression equation (corrected for instrumental drift) is then fitted using the measured values of standards as the independent variable and accepted values of the non-exchangeable fraction of the H isotopes in the standards as the dependent variable. This best-fit regression equation is then used to estimate δD values of the non-exchangeable fraction of H of the unknown samples.[14] This method requires an appropriate isotopic range of homogenous powdered organic matrices that have accepted δD values for the non-exchangeable fraction of H. These homogenous complex organic standard powders are time consuming and costly to develop and to calibrate; although recent efforts have reduced these difficulties somewhat.[16] At present, for keratins used in animal studies, the most widely used provisional δD standards are chicken feathers (CFS), cow hooves (CHS) and bowhead whale baleen (BWB), which span about 80‰ (−187‰ to −108‰).[14]

Many animals grow keratins that fall within the range of these standards, but a significant proportion of these keratins and other tissues do not. When unknown samples fall outside the range of the standards it is often assumed that the comparative equilibration regression equation can be extrapolated to these values; that is, analytical linearity is assumed. This assumption is not unique to measurement of H isotope ratios in organic materials and has a long history in stable isotope applications.[17] Nonetheless, we are not aware of any specific attempts to test the validity of comparative equilibration when extrapolating beyond the range of the standards, especially with keratin. Recent studies have criticized this practice, despite a weak empirical basis for doing so.[18] We tested the assumption of linearity and its effect on the mean and variation in powders and heterogeneous samples measured outside the calibration range of known standards.

We used three keratin standards (BWB, CFS, CHS) and two additional homogenous keratin powders that had more positive δD values than the accepted keratin standards to test the assumption of linearity and to compare the variance among raw values and normalized values within and outside the range of standards. We then examined the variance in two sets of non-homogenized feather samples to examine effects of extrapolation.

EXPERIMENTAL

We analyzed samples in batch sequences of 49 samples and references that we refer to as autoruns. Each autorun contained three replicates of five homogenized keratin powders. Of these keratin powders, three were supplied by one of the authors (LIW) and were documented by Wassenaar and Hobson;[14,15] chicken feathers (CFS), cow hooves (CHS) and bowhead whale baleen (BWB). The remaining two were prepared at the University of Oklahoma; human hair from one of the authors – JFK (HOJ), and Brown headed cowbird feathers (BHCO; Table 1). Powders prepared at the University of Oklahoma were cut into small pieces and cryoground in liquid nitrogen (Spex Certiprep 6750 freezer mill; Metuchen, NJ, USA). Each subsample was cryoground for three cycles of 3 min with 1 min between cycles. The resulting powder was sieved to remove particles larger than 63 μm and then blended to ensure isotopic homogeneity.

Table 1. Measured and equilibrated H stable isotope ratios of five keratin powders and feathers from two species of birds. For each equilibrated value the standards used to generate the equilibration equations are provided. Equilibrated values were used to test for effects on linearity and variance. All values are means of 13 autoruns reported as δD ± SD per mil VSMOW except where noted

Sample (accepted)	Raw	Equilibrated	Standards used	Test
CHS (−187)	−175.8 (3.6)			None
CFS (−147.4)	−145.8 (3.5)	−146.8 (3.5)	CHS, BWB	Variance
BWB (−108)	−116.9 (4.2)	−109.7[a] (7.6)	CHS, CFS	Linearity
HOJ	−90.5 (4.2)	−73.0 (5.0)	CHS, BWB	Variance
BHCO	−76.2 (5.0)	−54.0 (4.3)	CHS, BWB	Variance
Juncos	−148.6 (11.8)	−151.8 (14.0)	CHS, CFS, BWB	Variance
Buntings	−78.4 (12.9)	−62.0 (16.1)	CHS, CFS, BWB	Variance

[a] n = 39.

The five keratin powders were analyzed in sequence positions 1–5, 22–26, and 45–49 in each autorun, as is typically done in laboratories. Samples 5–21 and 27–44 were the feathers of unknown δD value. Many of the unknown feather samples in these autoruns were from Painted Buntings (*Passerina ciris*, n = 80) captured at Fort Sill, OK, USA, in the summer of 2007 and 2008 or Dark-eyed Juncos (*Junco hyemalis*, n = 72) captured in Norman, OK, USA, in the winter of 2009. Because neither of these species molt at the sampling location we made no assumption that these unknown samples represent a homogeneous grouping. We also had no *a priori* expectation that each feather would be isotopically homogenous.

All feather sample materials were cleaned with dilute detergent and then 2:1 chloroform/methanol following the method of Parrite and Kelly.[19] We packed 140 to 160 μg of each sample into a 3.5 mm × 5 mm silver capsule. A tight allowable range of ±10 μg was required to avoid variance in δD values due to variable H yields on the mass spectrometer.[2] All isotope ratio data were collected at the University of Oklahoma with a ThermoFinnigan Delta V isotope ratio mass spectrometer connected to a high-temperature pyrolysis elemental analyzer (TC/EA, ThermoFinnigan, Bremen, Germany) through an open split valve (Conflo III, ThermoFinnigan). The trap and box currents of the isotope ratio mass spectrometer were 0.8 mA and 0.7 mA, respectively. The TC/EA reactor was operated at 1450°C, and contained glassy C, quartz wool and silver wool according to ThermoFinnigan specifications.[20] The crucible in the reactor was changed every 100 samples and the reactor

DOI: 10.1002/rcm

contents were re-packed every <200 samples. The He flow rate was 100 mL/min maintained at a pressure of 21 psi, and the gas chromatography (GC) column was held at 100°C. H^{3+} corrections were carried out when the column was repacked and the H^{3+} signal ranged from 4.48 to 4.53 ppm/nA.

For each sample run, two reference pulses of ultra-high purity H_2 gas (99.999%, Air Gas) were injected into the ion source; the first at 40 s and the second at 90 s after the start of acquisition. Each pulse lasted 24 s with an intensity of 3000 mV at mass 2. The second reference peak was used in calculation of the sample δD value. Based on the δD values of standard materials the approximate δD value of the reference gas was −320‰. We used this value of the reference gas to calculate the raw δD values of the samples. Samples were automatically dropped from a 50-position zero-blank autosampler with a 1 s drop time at 110 s after the beginning of acquisition. The H_2 gas from the sample was chromatographically separated from N_2 and CO using a 5 Å packed molecular sieve. The sample peak was detected at about 135 s after the beginning of acquisition. The total analysis time of 300 s per sample allowed complete elution of CO and N_2 prior to the next sample acquisition.

For each autorun we corrected all measurements for instrumental drift between the first and last sample. Instrumental drift corrections were based on the slopes of best-fit lines for δD values regressed against analysis time of references within each autorun. A slope was calculated for the five powders in the run and these five slopes were averaged to achieve the drift correction coefficient. All data are reported in per mil notation (‰) relative to VSMOW (Vienna Standard Mean Ocean Water).

We used results from the three known keratin standards (BWB, CFS and CHS) to determine if it was reasonable to assume analytical linearity outside the range of known standards. For this analysis we treated CHS and CFS standards as known samples and treated the BWB standard as an unknown. For each of the 13 autoruns we calculated a comparative equilibration equation based on CHS and CFS samples and then used this equation to correct the measurements of BWB. We use a one-sample t-test to determine if the estimated value of BWB was different from the accepted provisional value (−108). Because unknowns are typically analyzed as single samples, we did not average our BWB replicates in each run for this analysis (i.e., n = 39 for BWB).

In our next analysis, we focused on variability that may result from comparative equilibration for samples within and outside the range of the known standards. For this analysis we used the same data described above from three keratin standards powders, but in this analysis we treated the CFS samples as unknown along with two additional keratin powders: human hair (HOJ) and Brown-headed Cowbird feathers (BHCO). We calculated comparative equilibration normalization equations for each of the 13 autoruns from CHS and BWB samples. We then used a Levene's test to check for homogeneity of variances among the raw measurements of the five powders and the normalized values of these powders. We note that the raw values of individual samples are not meaningful, but the variation among replicate measurements is informative. Because all the keratin powders were homogenized, we expected them to produce equally variable raw measurements. Therefore, if a Levene's test on the corrected values indicated significant heterogeneity in variance among the powders, we attributed this variation to the comparative equilibration method. Because it has the most positive δD value, we expected variance to be greatest in the corrected BHCO samples followed by HOJ with CFS having the least variation.

In a similar manner we used raw data values and comparative equilibration normalized values of Dark-eyed Junco and Painted Bunting feathers to infer the likely impact of comparative equilibration on sample variation. Because the Junco samples tended to be within the range of the standards we expected the variation in this sample to be unaffected by comparative equilibration. In contrast, the δD values of Painted Buntings tend to be more positive than the range of keratin standards. Therefore, we expected to see an increase un the variance among these samples. We again used Levene's test to compare the variation between Junco and Bunting feathers both before and after equilibration.

RESULTS

We failed to reject the null hypothesis of linearity outside the range of standard values. We calculated a grand mean of −109.7‰ for BWB based on the CFS and CHS samples (Table 1, Fig. 1). This value does not significantly differ from the accepted one of −108‰ ($t_{1,38} = 1.35$, $p = 0.19$). More importantly, the total effect size of 1.7‰ across a spread of 40‰ was small relative to other sources of error, and is acceptable for δD measurements. Generally, ±2‰ is considered to be an acceptable level of precision for δD measurements.

When using CFS, BWB and CHS as standards, our 13 calibration slopes ranged from 1.24 to 1.47 with intercepts that ranged from 32.8 to 70.0. Using these corrections we

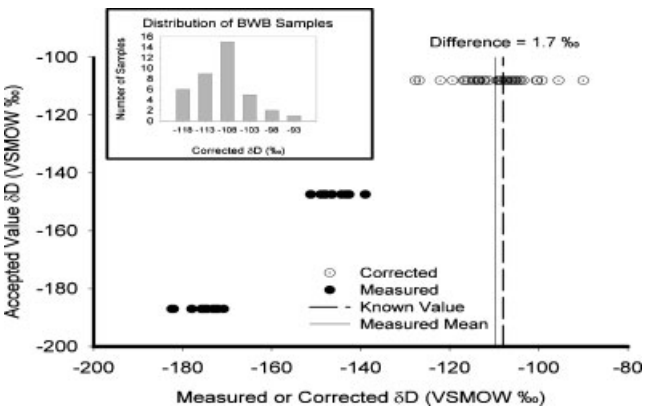

Figure 1. Solid circles are measured mean values (n = 13 autoruns) for CHS and CFS plotted against their accepted values (CHS = −187; CFS = −147.4); comparative equilibrium normalizations were based on these data. Open circles are corrected values of BWB measurements based on the data from the solid circles (n = 39, x axis) plotted against the accepted value for BWB (−108). A one-sample t-test indicates no difference between the normalized value for BWB and the accepted value. The insert is a histogram of BWB values.

DOI: 10.1002/rcm

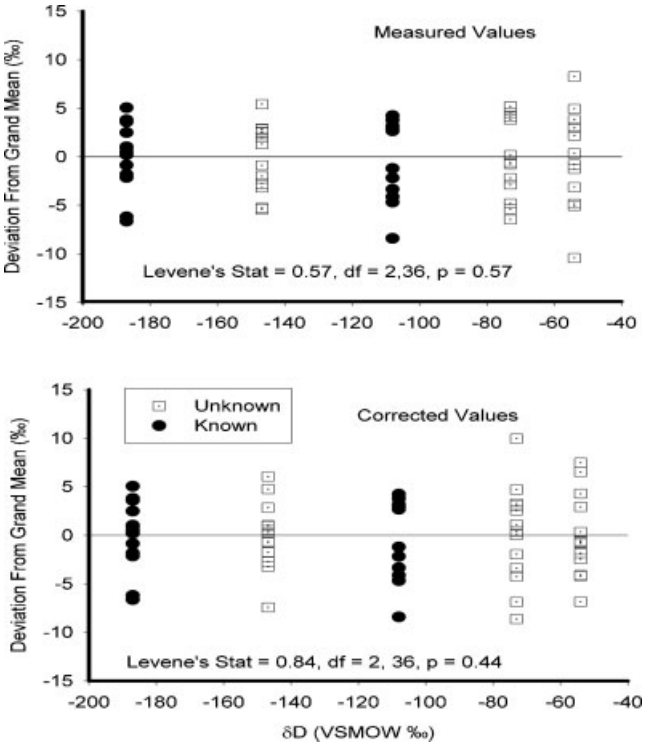

Figure 2. Deviations from the grand mean for measured (top) and equilibrated (bottom) values for five keratin powders. Solid symbols indicate powders that were treated as lab standards and used to generate comparative equilibration normalization equations. Open symbols were corrected with these equations. Levene's tests indicate that the measured and corrected values of the powders were equally variable.

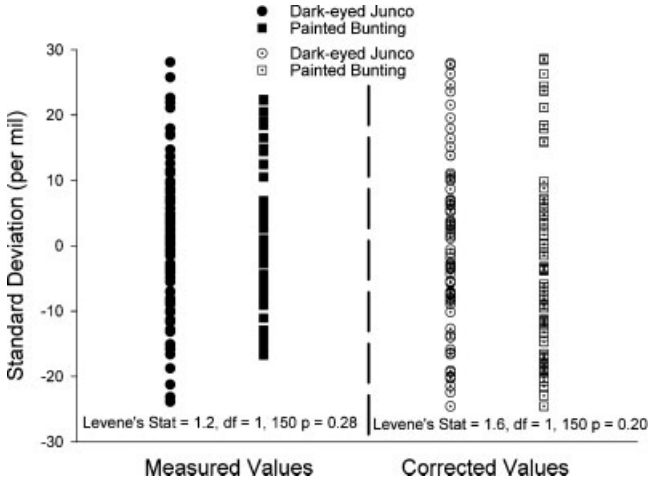

Figure 3. Deviations from the grand mean for samples of non-homogenized 80 Dark-eyed Junco and 70 Painted Bunting feathers. There was no difference in the variation between measured and corrected values of juncos or buntings.

failed to reject the hypothesis of equal variances across the raw measured values and the corrected values of the keratin powders (Fig. 2) suggesting that extrapolation to values outside the range of standards does not result in an inflation of the variance in measurements. All the Dark-eyed Junco feathers that we measured were within the range of our three keratin standards, while all the Painted Buntings were outside this range (Table 1). Comparison of the raw measured values of these birds showed that there was no significant difference in the variance between Junco and Bunting samples prior to or after comparative equilibration normalization corrections (Fig. 3).

DISCUSSION

Overall when comparative equilibration was used as designed, that is samples are within the range of the standards, the method was highly effective at correcting measured values to the accepted value (CFS data, Table 1). Further, our results supported the assumption that equilibrium H exchange is a linear process, at least over the range from −187 to −108‰. The variation in slopes and intercepts of our calibration lines suggested that routine measurement of multiple standards with samples was necessary for adequate correction. It is unlikely that single-point or offset calibration approaches would produce similar results. There was minimal difference between the grand mean of our comparatively equilibrated samples of BWB (−109.7) and the accepted value derived through offline steam equilibration (−108) by Wassenaar and Hobson.[14] This finding suggested that it was possible to infer the values of samples outside the range of known standards and it supported the validity of this approach. Further testing is needed to validate this method with samples that greatly exceed the range of the standards. Nonetheless, our data indicate that comparative equilibration provides a viable approach for arriving at precise values for new standards without having to resort to offline steam equilibration. Wunder *et al*. also employed this approach effectively to validate a working standard at −57‰ that was used in their comparative equilibrium corrections.[21]

Our results also indicated no detectable inflation in variation in δD values as the distance between the standards and unknowns increased. The absence of this pattern was surprising, and probably indicated that within our range of samples comparative equilibration was a minimal source of error. Even for our unknown feather samples the standard deviations of the population of samples were not significantly inflated by comparative equilibration (Table 1).

We urge researchers to continue to apply and test the comparative equilibration method when measuring stable H isotope ratios of feathers and other biological substrates. While we are hesitant to dismiss the possibility of inflated variation well outside the range of lab standards based on our data alone, we think that within the range of our measured values this problem is not a primary concern. However, researchers should be aware of this possibility (as well as potential intra-sample isotopic heterogeneity) when unknown samples are far more enriched (or depleted) relative to the standards used to develop comparative equilibration correction equations. Finally, development of large batches of complex organic substrates that span the range of normally encountered keratins and other biological substrates is a pressing need for laboratory standardization in this field.

DOI: 10.1002/rcm

Acknowledgements

We thank R. Maynard and M. Engel for assistance with analysis. This research was supported by the US National Science Foundation (IOS: 0541740) and the US Environmental Protection Agency (R 833778). Although the research described in this article has been funded wholly or partially by the United States Environmental Protection Agency through grant/cooperative agreement R 833778, it has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred.

REFERENCES

1. Bowen GJ, Wassenaar LI, Hobson KA. *Oecologia* 2005; **143**: 337.
2. Wassenaar LI. In *Tracking Animal Migration with Stable Isotopes*. Academic Press: London, 2008; 21.
3. Bowen GJ, Revenaugh J. *Water Resources Res.* 2003; **39**: 1299.
4. Rubenstein DR, Chamberlain CP, Holmes RT, Ayres MP, Waldbauer JR, Graves GR, Tuross NC. *Science* 2002; **295**: 1062.
5. Kelly JF, Atudorei V, Sharp ZD, Finch DM. *Oecologia* 2002; **130**: 216.
6. Paxton KL, Van Riper C III, Theimer TC, Paxton EH. *Auk* 2007; **124**: 162.
7. Hobson KA, Norris RD. In *Tracking Animal Migration with Stable Isotopes*. Academic Press: London, 2008; 1.
8. Hobson KA, Wassenaar LI. *Oecologia* 1997; **109**: 142.
9. Chamberlain CP, Blum JD, Holmes RT, Feng X, Sherry TW. *Oecologia* 1997; **109**: 132.
10. Wassenaar LI, Hobson KA. *Rapid Commun. Mass Spectrom.* 2006; **20**: 1.
11. Wunder MB, Norris DR. In *Tracking Animal Migration with Stable Isotopes*. Academic Press: London, 2008; 107.
12. Wunder MB. In *Isoscapes: Isotope Mapping and its Applications*. Springer-Verlag: New York, 2009; in press.
13. Chesson LA, Podlesak DW, Cerling TE, Ehleringer JR. *Rapid Commun. Mass Spectrom.* 2009; **23**: 1275.
14. Wassenaar LI, Hobson KA. *Isot. Environ. Health Stud.* 2003; **39**: 1.
15. Wassenaar LI, Hobson KA. *Environ. Sci. Technol.* 2000; **34**: 2354.
16. Sauer PE, Schimmelmann A, Sessions AL, Topalov K. *Rapid Commun. Mass Spectrom.* 2009; **23**: 949.
17. Coplen TB. *Chem. Geol.* 1988; **72**: 293.
18. Smith AD, Lott CA, Smith JP, Donohue KC, Wittenberg S, Smith KG, Goodrich L. *Auk* 2009; **126**: 41.
19. Paritte JM, Kelly JF. *Auk* 2009; **126**: 165.
20. ThermoFinnigan. 2003. *TC/EA Operating Manual – High Temperature Conversion Elemental Analyzer*, Issue **9/2003** Ident No. 1127601.
21. Wunder MB, Kester CL, Knopf FL, Rye RO. *Oecologia* 2005; **144**: 607.

DOI: 10.1002/rcm