Does a Lack of Design and Repeatability Compromise Scientific Criticism? A Response to Smith et al. (2009).

Authors: Wunder, Michael B., Hobson, Keith A., Kelly, Jeff, Marra, Peter P., Wassenaar, Leonard I., et al.

Source: The Auk, 126(4) : 922-926

Published By: American Ornithological Society

URL: https://doi.org/10.1525/auk.2009.126.4.2

Letters

The Auk 126(4):922–926, 2009

Does a lack of design and repeatability compromise scientific criticism? A response to Smith et al. (2009).—In a recent paper published in *The Auk*, Smith et al. (2009) raised serious concerns over an apparent lack of reproducibility in their study of stable hydrogen isotope values (δD_f) in raptor feathers. The authors based their concerns on results obtained from different laboratories to which they submitted original and blind "repeats" over a multiyear period. A regression of the original sample δD versus "repeat" measurements showed an increase in the magnitude of residuals with increasing δD_f, especially for values greater than about –80‰ (Smith et al. 2009: fig. 2). Because of this, the authors "caution against the continued use of δD_f for predicting geographic origin, and for addressing important conservation questions" (p. 41) and conclude that "it is counterproductive to move forward [with hydrogen isotopes in avian studies] without first establishing full confidence in the technique that underlies

The Auk, Vol. 126, Number 4, pages 922–931. ISSN 0004-8038, electronic ISSN 1938-4254. DOI: 10.1525/auk.2009.126.4

such insights and conservation recommendations" (p. 45). We disagree with these sentiments.

The intent of Smith et al. (2009) was to report on the reliability of routine stable-hydrogen-isotope measurements from a client or user perspective, under the assumption that one lab is the same as the next. Although their aim is laudable, an interlaboratory comparison for the analysis of δD in organic material is, in fact, a rather complex endeavor that depends heavily on having an appropriate *a priori* study design and uniform isotopic analysis methodology in place (e.g., a ring test). Unfortunately, not all the isotope laboratories used were consulted about developing a proper ring-test study design, and additional design and methodology problems have led to inappropriately alarmist conclusions by Smith et al. (2009) that are largely indefensible and that do not accurately reflect the ongoing research and development of isotopic tracers in animal migration. Here, we highlight the reasons why we disagree with the recommendations in Smith et al. (2009) by (1) briefly discussing the important role of appropriate hydrogen isotopic standards in calibrating modern stable-isotope measurements, (2) pointing out a few of the confounding problems in Smith et al.'s (2009) design and analysis that help to explain some of their results, and (3) describing some general issues related to real progress in the use of stable hydrogen isotope ratios as geographic tracers for migratory birds and other animals.

Measurement standards.—Currently, there are no internationally accepted primary reference materials for δD in complex organic materials, such as feathers, that contain exchangeable hydrogen. All stable-isotope laboratories conform to strict quality-assurance and quality-control protocols that use primary isotopic reference materials regulated by the International Atomic Energy Agency (IAEA) in Vienna. Primary isotopic reference materials are used to calibrate daily working laboratory standards, with the general practice that working standards ideally should span the natural isotopic range and should be of the same type of material as the unknowns (principle of identical treatment). Unfortunately, because there are no IAEA primary reference materials for δD in complex organic tissues (feathers, other keratins, blood, etc.), the onus is on the researcher to (1) ensure that the laboratories of choice use appropriate or published calibration standards, (2) recognize that this is an evolving analytical field (e.g., not all labs employ identical analytical methods), and, therefore, (3) work closely with the selected laboratories toward a reasonable interpretation of their data, given the practices in place.

Because of the lack of formal primary standards for δD and the urgent need for reproducibility among labs, several key stable-isotope laboratories have independently and informally developed a "best practices" approach to organic hydrogen measurements through the creation of appropriate (and, in the case of feathers, keratin) working standards suitable for routine isotopic analyses (Wassenaar and Hobson 2003, 2006; Wassenaar 2008). A first attempt to achieve analytical compatibility among labs led to the dissemination of three provisional keratin working standards from the Environment Canada (EC) Lab in Saskatoon, Saskatchewan, that spanned a δD range between –200‰ and –100‰. It is worth recognizing that this isotopic range was largely arbitrary and reflected considerable effort to produce sufficient supplies of keratinous standard materials, given readily available material. Analytical linearity was anticipated within the bounds of the Vienna Standard Mean Ocean Water–Standard Light Antarctic Precipitation (VSMOW-SLAP) natural isotopic abundance range (approximately –400‰ to 0‰), an assumption supported by isotope laboratories and by our observations of good replicate agreement for high δD_f samples over the range –200‰ to 0‰ using online (pyrolysis, continuous flow–isotope ratio mass spectrometry [CF-IRMS]) and offline (zinc reduction, dual-inlet) techniques (see Kelly et al. 2009).

Calibration range and standards development.—Recent studies and our own laboratory experiences have encountered δD_f values that extend well outside the keratin calibration range of –200‰ to –100‰ established by the EC lab (e.g., Wunder et al. 2005, Lott and Smith 2006, Powell and Hobson 2006, Langin et al. 2007, Kelly et al. 2008b), which underscores the need for keratin standards that span upwards to even +90‰. Such positive keratin standards do not currently exist or are not available in sufficient quantities to support wide distribution for comparison among labs. Moreover, all proposed keratin standards must conform to strict protocols of proper isotopic homogenization by cryogenic grinding, sieving, distribution, and extensive inter-lab comparison. For these reasons, the development of new keratin standards is costly in both time and money. Several labs are currently engaged in developing a wider isotopic range of keratin standards sampled from diverse sources that include Greater Kudu (*Tragelaphus strepsiceros*) horn from Ethiopia, feathers from desert-dwelling White-winged Doves (*Zenaida asiatica*) in Arizona, and Domestic Chickens (*Gallus gallus domesticus*) raised on deuterium-spiked water. Some studies, not surprisingly, have had to deal with "outside-the-calibration-range" data, and the stated concerns of Smith et al. (2009) are long known among isotopic research labs, as recently documented by Kelly et al. (2008a, 2009).

Although analytical linearity is often observed in δD_f analysis, the slope of the calibration line itself is not always unity. Laboratories that use fewer than two different working standards must rely on the false assumption of a strictly unit slope for the calibration line, an assumption that becomes more problematic as the measured sample values get farther away from the value of the standard. Smith et al. (2009) did not describe the names, values, or number of different organic keratin standards used for calibration purposes by each lab. Nor did they describe whether the calibration range bracketed by those standards (if the lab used >1 standard) was the same for all labs involved. It is therefore difficult to determine the extent to which this most basic complication contributed to the observed differences within and among lab results. For these reasons, researchers interested in using δD_f data must understand the importance of the values of the organic keratin standards used by the isotope lab to calibrate the measures against the VSMOW-SLAP scaling. More pragmatically, such researchers should guard against over-interpretation of δD_f values that fall well outside the laboratory calibration range defined by those organic keratin standards, and they should especially guard against over-interpretation of values that are derived from calibrations using fewer than two different organic keratin standards.

Design and analysis.—One of the major factors in our disagreement with the strong conclusions offered by Smith et al. (2009) is that their *ad hoc* study design did not consider important and well-known sources of variance, which in turn lead to an ambiguous interpretation of their results. There can be substantial systematic differences in δD_f along the length of a single feather,

an observation that has been documented in numerous species (Wassenaar and Hobson 2006), including the species of raptor considered by Smith et al. (2008). Smith et al. failed to acknowledge that this previous work (including their own) indicates that what they call "repeats" are not, in fact, repeats. Furthermore, because they also failed to classify feather measures as falling "in or out" of the specific laboratory calibration ranges discussed in the previous section, it is difficult to determine how these confounded factors each contributed to the observed differences in their measurements. Studying the relative influences of each of these two factors can be isolated by relatively simple study designs. For example, if the goal is to study measurement reproducibility outside the calibration range defined by the keratin standards, the influence of biologically derived differences among locations on the feather can be minimized by cryogenically grinding the feathers. Similarly, if the interest is in studying the potential biological factors contributing to the patterns observed along the length of a feather, then using feathers with δD_f values inside the calibration range of the standards will minimize the influence of measurement error from extrapolating beyond the calibration range. As described, the study design employed by Smith et al. (2009) controlled for neither of these, leaving us to wonder about the relative extent to which these two distinct and known variance-generating processes influenced their results.

Replication is important.—Smith et al.'s (2009) study did not meet key standards for the publication of experimental and comparative results because sufficient analytical information was not provided for any of the labs in the study. As such, this research cannot currently be replicated by others. Smith et al. (2009) did not provide enough details concerning the names or locations of the labs that were used, the distribution of sample varieties among lab trials (e.g., batches from which species, which range of feather values, and which pre-analysis treatment types went to which labs), or the laboratory procedures followed by each lab (CF-IRMS instrumentation, reference gas, carrier flow rates, standards and normalization procedures, pyrolysis temperatures, etc.). When using methodologies from a single lab, referencing analytical techniques from previously published work is sufficient. However, when comparing among multiple labs, information on how these facilities differ in their subtle approaches is important and must be provided in any published work.

We have learned that the editors of *The Auk* told the authors to remove lab names from their manuscript in an effort to protect the reputations of the labs, given the negative nature of the article (J. Jones pers. comm.). Despite the fact that we represent some of the labs used in the study, we would much rather have had our lab names along with all others listed explicitly in the manuscript. Furthermore, we believe that failing to inform the labs that the data they generated were to be used in such a publication represents a missed opportunity to make real constructive progress toward improving the way ornithologists work with stable-isotope laboratories. We stand firmly in support of our lab practices and would have preferred to be contacted and named along with all other labs used in the study; this would have promoted a more constructive discussion among the labs involved toward understanding factors that contributed to apparent discrepancies, whether attributable to methods used by the labs or to the methods of the investigators. We cannot understand why *The Auk* would ask the authors to censor this information and are greatly disappointed by that decision. If Smith et al.'s (2009) study was truly designed for the purpose of examining laboratory-derived causative factors in the repeatability of measures and was done in a rigorous manner, there should be every reason to alert laboratories that were producing questionable data. Altogether, these oversights seriously detract from the study's scientific rigor and stand in stark contrast to the overly strong conclusions about the importance of analytical reproducibility.

Additional issues of progress.—Smith et al. (2009) warned that their concerns are not restricted to raptors and, in so doing, implied that the community of researchers using the hydrogen isotope approach for passerines and other taxonomic groups have been unaware, thus far, of the issues of within- and between-individual isotopic variation. This is clearly not the case. Several interlaboratory comparisons of homogenized keratin materials from passerines were designed to look at this very issue; those properly designed studies show agreement among at least four labs and even for nonhomogenized feathers from 18 different samples (Wassenaar 2008; see also Wassenaar and Hobson 2006).

Using the logic of Smith et al. (2009), the raptor δD_f basemap produced by Lott and Smith (2006) might be considered suspect because vast regions of that "raptor isoscape" involve birds with δD_f values outside the range of available keratin-calibration standards. We examined this further by regressing the raptor feather values reported by Lott and Smith (2006) against predicted amount-weighted growing-season average precipitation δD_p from Bowen et al. (2005). That regression (Fig. 1) shows good agreement through the entire range of δD_f values (–175‰ to 0‰), which suggests that >60% of the variance in raptor feathers is explained by predicted long-term patterns of hydrogen isotopes in rainfall, as has been observed in numerous other species (Hobson 2008). Thus, while we also are concerned about the need to develop much wider coverage in the range of keratin standards and discourage single-point or mineral calibration approaches, patterns such as those shown in Figure 1 provide an excellent basis for productive approaches to geographic assignment of individuals.

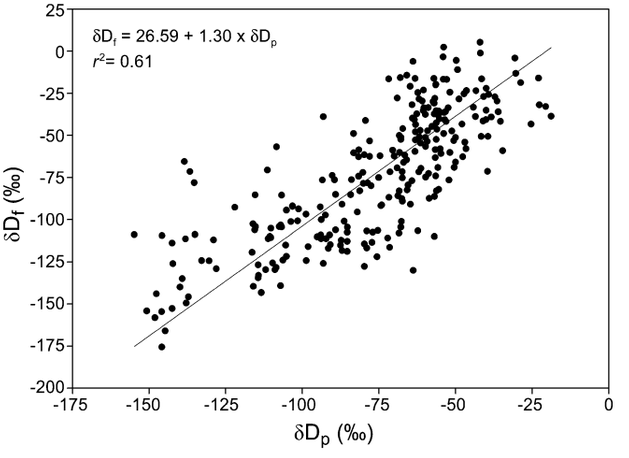

FIG. 1. Reduced-major-axis regression model showing the relationship between δD_f in known-source raptor feathers (n = 264) from Lott and Smith (2006), and expected growing-season δD in precipitation (from Bowen et al. 2005).

Using δD_f for making geographic assignments of migratory species is not inherently problematic. It is entirely reasonable to move forward despite uncertainty over the process of measuring δD_f by using modeling approaches to transparently incorporate estimable variance (Wunder and Norris 2008). One should not view isotopic assignment as a tool that provides unambiguous geographic coordinates for each individual, because no single location is expected to be isotopically homogenous or unique in space and time, and we expect variation in isotope data for many reasons. We can make probabilistic (rather than absolute) statements about a location as a potential origin, given known and measurable uncertainties derived from biological and analytical processes. Wunder (2007, 2010) and Hobson et al. (2009) illustrate how these approaches model the current state of the science, however advanced it may (or may not) be; inferences from such assignment models will be only as strong as the data are reliable. Because of this, we agree with Smith et al. (2009) that data reliability is worth considering directly, but we respectfully disagree with how they tried to do that.

Uncertainty does not warrant inaction.—We believe that to "caution against continued use of δD_f for predicting geographic origin" or to "caution . . . against addressing important conservation questions" (Smith et al. 2009:41) is to miss an important and pragmatic point. Rarely are conservation efforts based on only a single bit of information. More typically, conservation decisions are complex optimizations that consider a wide range of scientific, economic, social, and political factors. Conservation decision-makers use all the information at their disposal and, even still, are most often faced with the need to make decisions in the face of incomplete information. It is true that incorrect scientific information can be worse than no information, but ignoring scientific information is far worse. Rather than withholding such information, as suggested by Smith et al. (2009), we believe that so long as interpretative ambiguities are transparently described or modeled, conservation planners will benefit from isotope-based inferences.

In short, the conclusion of Smith et al. (2009:45) that "it is counterproductive to move forward [with hydrogen isotopic studies] without first establishing full confidence in the technique that underlies such insights and conservation recommendations" is itself counterproductive. We are disappointed that Smith et al. (2009) chose to make such strong negative recommendations on the basis of results from a poorly designed investigation. The use of stable isotopes in assigning geographic origins is a new, evolving, valuable, and innovative scientific field of study that has moved rapidly within a decade and will continue to be refined and improved only through continued research and publication of appropriately reviewed work.

Acknowledgments.—We thank Smith et al. for the opportunity to discuss these underappreciated points about the use of stable hydrogen isotopes in applied ornithology. We thank the editors of *The Auk* for the opportunity to contribute this letter. Two anonymous reviewers and J. Jones provided comments that helped to improve the clarity of our manuscript.—Michael B. Wunder, *University of Colorado Denver, Department of Biology, Campus Box 171, Denver, Colorado 80217, USA;* Keith A. Hobson, *Environment Canada, 11 Innovation Boulevard, Saskatoon, Saskatchewan S7N 3H5, Canada (address correspondence to this author. E-mail: keith.hobson@ec.gc.ca);* Jeff Kelly, *Oklahoma Biological Survey and Department of Zoology, 111 E. Chesapeake Street University of Oklahoma, Norman, Oklahoma 73072, USA;* Peter P. Marra, *Smithsonian Migratory Bird Center, National Zoological Park, 3001 Connecticut Avenue, Washington, D.C. 20008, USA;* Leonard I. Wassenaar, *Environment Canada, 11 Innovation Boulevard, Saskatoon, Saskatchewan S7N 3H5, Canada;* Craig A. Stricker, *U.S. Geological Survey, Denver Federal Center, MS963, Denver, Colorado 80225, USA;* and Richard R. Doucett, *Colorado Plateau Stable Isotope Laboratory, Northern Arizona University, Box 5640, Flagstaff, Arizona 86011, USA.*

Literature Cited

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

Hobson, K. A. 2008. Applying isotopic methods to tracking animal movements. Pages 45–78 *in* Tracking Animal Migration with Stable Isotopes (K. A. Hobson and L. I. Wassenaar, Eds.). Academic Press, London.

Hobson, K. A., H. Lormée, S. L. Van Wilgenburg, L. I. Wassenaar, and J. M. Boutin. 2009. Stable isotopes (δD) delineate the origins and migratory connectivity of harvested animals: The case of European woodpigeons. Journal of Applied Ecology 46:572–581.

Kelly, J. F., S. Bearhop, G. J. Bowen, K. A. Hobson, D. R. Norris, L. I. Wassenaar, J. B. West, and M. B. Wunder. 2008a. Future directions and challenges for using stable isotopes in advancing terrestrial animal migration research. Pages 129–139 *in* Tracking Animal Migration with Stable Isotopes (K. A. Hobson and L. I. Wassenaar, Eds.). Academic Press, London.

Kelly, J. F., E. S. Bridge, A. M. Fudickar, and L. I. Wassenaar. 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.

Kelly, J. F., M. J. Johnson, S. Langridge, and M. Whitfield. 2008b. Efficacy of stable isotope ratios in assigning endangered migrants to breeding and wintering sites. Ecological Applications 18:568–576.

Langin, K. M., M. W. Reudink, P. P. Marra, D. R. Norris, T. K. Kyser, and L. M. Ratcliffe. 2007. Hydrogen isotopic variation in migratory bird tissues of known origin: Implications for geographic assignment. Oecologia 152:449–457.

Lott, C. A., and J. P. Smith. 2006. A geographic-information-system approach to estimating the origin of migratory raptors in North America using stable hydrogen isotope ratios in feathers. Auk 123:822–835.

Powell, L. A., and K. A. Hobson. 2006. Enriched feather hydrogen isotope values for Wood Thrushes sampled in Georgia, USA, during the breeding season: Implications for quantifying dispersal. Canadian Journal of Zoology 84:1331–1338.

Smith, A. D., K. Donohue, and A. M. Dufty, Jr. 2008. Intrafeather and intraindividual variation in the stable-hydrogen isotope (δD) content of raptor feathers. Condor 110:500–506.

Smith, A. D., C. A. Lott, J. P. Smith, K. C. Donohue, S. Wittenberg, K. G. Smith, and L. Goodrich. 2009. Deuterium measurements of raptor feathers: Does a lack of reproducibility compromise geographic assignment? Auk 126:41–46.

WASSENAAR, L. I. 2008. An introduction to light stable isotopes for use in terrestrial animal migration studies. Pages 21–44 *in* Tracking Animal Migration with Stable Isotopes (K. A. Hobson and L. I. Wassenaar, Eds.). Academic Press, London.

WASSENAAR, L. I., AND K. A. HOBSON. 2003. Comparative equilibration and online technique for determination of non-exchangeable hydrogen of keratins for use in animal migration studies. Isotopes in Environmental Health Studies 39:211–217.

WASSENAAR, L. I., AND K. A. HOBSON. 2006. Stable-hydrogen isotope heterogeneity in keratinous materials: Mass spectrometry and migratory wildlife tissue subsampling strategies. Rapid Communications in Mass Spectrometry 20:2505–2510.

WUNDER, M. B. 2007. Geographic structure and dynamics in Mountain Plover. Ph.D. dissertation. Colorado State University, Fort Collins.

WUNDER, M. B. 2010. Using isoscapes to model probability surfaces for determining geographic origins. *In* Isoscapes: Understanding Movement, Pattern, and Process on Earth through Isotope Mapping (J. B. West, G. J. Bowen, T. E. Dawson, and K. P. Tu, Eds.). Springer, New York. In press.

WUNDER, M. B., C. L. KESTER, F. L. KNOPF, AND R. O. RYE. 2005. A test of geographic assignment using isotope tracers in feathers of known origin. Oecologia 144:607–617.

WUNDER, M. B., AND D. R. NORRIS. 2008. Analysis and design for isotope-based studies of migratory animals. Pages 107–128 *in* Tracking Animal Migration with Stable Isotopes (K. A. Hobson and L. I. Wassenaar, Eds.). Academic Press, London.

Received 13 March 2009, accepted 10 August 2009
Associate Editor: J. Jones