APPENDIX A. Statistical details for article examples.

Example 1 – statistical procedures. The nested linear model we employed in Example 1 takes the form $Y = X\beta + \epsilon$; where $Y$ is the matrix of observed values with dimensions $n$ (number of observations) $\times$ 2 (to account for $\delta^{13}C$ and $\delta^{15}N$ isotope ratios for each individual observation); $X$ is an $n \times k$ design matrix for $k$ groups (typically coded for an intercept and $k$-1 dummy variables); $\beta$ is a $2 \times k$ design matrix of group effects; and $\epsilon$ is an $n \times 2$ matrix of group residuals. Group effects are estimated as $\hat{\beta} = (X'X)^{-1}X'Y$, where $'$ is a matrix transpose and $^{-1}$ is a matrix inverse, and residuals are determined as $\epsilon = Y - X\hat{\beta}$. The model, $X$, can be reduced by removing the dummy variables for groups (in this case, producing a single vector of 1s to code for the overall centroid). Residuals from this reduced model, call it $X_r$, are similarly determined as $\epsilon_r = Y - X_r\hat{\beta}_r$. By randomly shuffling rows of the $\epsilon_r$ matrix, a new matrix is created, $\epsilon_r^*$, that has the simple property that sums of square and cross-products (SSCP) are the same as the observed, i.e., $\epsilon_r ' \epsilon_r = \epsilon_r^* ' \epsilon_r^*$ (see Collyer and Adams 2007), i.e., the log likelihood of the model is unchanged (Anderson and ter Braak 2003). Further, the equation $Y^* = X_r \beta_r + \epsilon_r^*$ produces randomly shuffled observed values, and is therefore, no different than shuffling the values themselves. A null distribution for hypothesis testing is created by repeating the randomization procedure many (e.g., 9999) times.

For comparisons of centroid locations and dispersion statistics $CD$, $NND$, and $E$, Euclidean distance is the same as the absolute value of the difference between, say, groups 1 and 2 for the metrics we tested. For example, $d_{c1,2} = \left[ (y_1 - y_2)(y_1 - y_2)^t \right]^{1/2}$, is the (bivariate) Euclidean distance between group centroids ($y_i$) and $d_{c1} = \left[ (E_1 - E_2)(E_1 - E_2)^t \right]^{1/2} = |E_1 - E_2|$ is the (univariate) Euclidean distance of eccentricities between groups 1 and 2. Repeating these calculations for many (i.e., 9,999) random permutations of residuals (as above) produces...
empirical null distributions of Euclidean distances for each metric. The rank percentile of observed Euclidean distance between groups is an empirical $P$-value for the test of the null hypothesis (e.g., the absolute value of among-group differences in metrics are equal to zero). Distances and null distributions are determined for all possible pairs of groups in the comparison.

*Examples 2 and 3 – statistical procedures.* To estimate guild $(g) \times$ site $(s)$ centroids, a two-factor linear model was used. The design matrix for this model contains parameters (dummy variables) for guild, sites, and guild$\times$site. Centroid values are determined as $\bar{Y} = \bar{X}_j \hat{\beta}_j$; where $\hat{\beta}_j$ is calculated as above; $\bar{X}_j$ is a $gs \times k$ design matrix for $g$ groups, $s$ sites, and $k$ parameters; and the subscript, $j$, indicates that the linear model is the “full” model.

If there are $g = 2$ guilds and $s = 2$ sites, then $\bar{Y}$ is a $4 \times 2$ matrix of group$\times$site centroids. This matrix is partitioned into two $2 \times 2$ matrices, $\bar{Y}_1$ and $\bar{Y}_2$ for each guild where the rows of these matrices are start and end points of a vector that describes the spatial difference in isotopic composition. When there are $g = 2$ guilds and $s = 3$ sites, then $\bar{Y}$ is a $6 \times 2$ matrix of group$\times$site centroids, and in this case, $\bar{Y}_1$ and $\bar{Y}_2$ each have three rows that describe a trajectory in the bivariate space, which is the spatial trend in isotopic composition. Analogously, any $gs \times 2$ $\bar{Y}$ can be portioned into $g$ matrices of $s \times 2$ dimensions that describes comparable spatial trajectories of isotopic composition. Test statistics are based on comparisons of attributes of guild trajectories.

Each trajectory has three attributes of interest: size, orientation, and shape. Trajectory size measures the amount of change in location in the bivariate plot among sites. Different size measures can be used, but path distance – the cumulative sum of vector lengths between adjacent sites – is perhaps the most informative, as it expresses the amount of change in location most precisely. To calculate trajectory (path) size, first the $i^{th}$ trajectory is described as the expected values (i.e., means) for sites within guilds as $Z_i = X_i \hat{\beta}_i$. Each row of $Z_i$ is a site mean. For $s$ sites there are $s-1$ vectors which describe the difference in site means, i.e., if every row of $Z_i$ is represented as $z$, then $\Delta z_j = z_{j+1} - z_j$ for the $j$ levels of $s-1$. Thus, $d_{path} = \sum_{j=1}^{s-1} \sqrt{\Delta z_j \Delta z'_j} = \sum_{j=1}^{s-1} \sqrt{(Z_{i,j+1} - Z_{i,j})(Z_{i,j+1} - Z_{i,j})}$. The test statistic describing pairwise
differences in $d_{path}$ is then calculated as $\Delta d_{path} = |d_{path 1} - d_{path 2}|$ for every combination of two trajectory sizes. Trajectory direction is a vector in the bivariate space that describes the principal direction from start point to end point (the site values at the beginning and end of the trajectory). This vector can be estimated through principal components analysis (PCA) using site means (the vector is the first component – PC1). If only two site means exist, then PC1 is the same as the vector of differences between means. For any pairwise comparison of trajectories, the vector correlation between PCs is the inner product of the PCs, standardized by their Euclidean distances. The angle between vectors is the arccosine of this value: 

$$\theta_{12} = \cos^{-1}\left(\frac{\mathbf{p}_1 \cdot \mathbf{p}_2}{\sqrt{\mathbf{p}_1 \cdot \mathbf{p}_1} \sqrt{\mathbf{p}_2 \cdot \mathbf{p}_2}}\right),$$

where $\mathbf{p}$ is the vector representation of a principal component. It is essential to correct for arbitrarily described PC directions as a result of rotation (as exact PC vectors can be arbitrarily described to point in opposite directions; see Adams and Collyer 2009).

Trajectory shape is more challenging to define. In geometric morphometrics studies, the shape of a trajectory (or configuration) of 2- or 3-dimensional landmarks is the description of the trajectory after it is rendered invariant to other trajectories in terms of its size, position, and orientation (Rohlf and Slice 1990). The difference in shape between two trajectories is based on the differences in location of their corresponding points, after a generalized Procrustes analysis (GPA) (Rohlf and Slice 1990). GPA centers configurations scaled to unit size, and rotates them via generalized least squares iterations such that variance among points is minimized (Rohlf and Slice 1990). In morphological studies, a shape difference is the square root of the summed squared differences in anatomical landmarks described by Cartesian coordinates of “aligned” configurations, a metric called “Procrustes distance” (Bookstein 1991). Likewise, a difference in shape between two isotope trajectories is the Procrustes distance between them after they are scaled to the same size, centered, and rotated to minimize variation among site means.

Two descriptions of GPA are provided by Rohlf and Slice (1990) and Adams and Collyer (2009) and only briefly summarized here. First, each trajectory is mean-centered by subtracting its trajectory centroid to produce $\mathbf{Z}_c$. Next, each trajectory is scaled to unit size by dividing the trajectory by its centroid size, $CS$, the square root of the summed squared distances between the site means and the centroid (Bookstein 1991), producing $\mathbf{Z}_c'$. (Note that the order of the first two steps is not important.) These scale-free, centered trajectories are invariant in size and
position, and only differ in orientation and shape. The final step is to use a GLS superimposition procedure which rotates $Z'_c$ such that variance among corresponding site means is minimized without changing the relative locations of site means within trajectories (i.e., trajectories are rigid; see Adams and Collyer 2009), thus making invariant directional differences. The subsequent (aligned) trajectories, $Z''_c$, vary only in shape. Difference in shape between pairwise comparisons of trajectories are estimated as the Procrustes distance between them is calculated as $D_{shape12} = \sqrt{\left((Z''_{c1} - Z''_{c1})^T (Z''_{c1} - Z''_{c1})^T\right)}$, using $D$ instead of $d$ as a convention not to cause confusion with trajectory size. As with trajectory size, the test statistic describing pairwise differences in $D_{shape}$ is then calculated as $\Delta D_{shape} = |D_{shape1} - D_{shape2}|$ for every combination of two trajectory shapes.

**LITERATURE CITED**


