Aditya Singh

11 January 2026

Isotope Ratio Analysis for Determining the Country of Origin of Lactoferrin Samples

INTRODUCTION

Authentication of food and nutritional products has become increasingly important due to globalised supply chains and the high economic value associated with country-of-origin claims. Lactoferrin, a bioactive milk protein widely used in infant formula and functional foods, is particularly vulnerable to provenance ambiguity because raw materials may be sourced internationally while processing occurs locally.

Stable isotope ratio analysis (SIRA) is a well-established technique for determining the geographical origin of food products. Ratios of carbon ($\delta^{13}C$), nitrogen ($\delta^{15}N$), and sulfur ($\delta^{34}S$) reflect environmental, agricultural, and dietary conditions and therefore act as intrinsic fingerprints of biological origin.

This study applies multivariate statistical techniques to isotope ratio data from lactoferrin samples of known origin to:

1. Identify natural clustering by country of origin, and
2. Evaluate whether a pilot sample labelled as Australian is isotopically consistent with confirmed Australian samples.

DATA DESCRIPTION

The dataset consists of lactoferrin samples sourced from the following countries:

- Australia
- New Zealand
- United States
- China

For each sample, the following isotope ratios were measured:

- $\delta^{13}C$ (carbon isotope ratio)
- $\delta^{15}N$ (nitrogen isotope ratio)
- $\delta^{34}S$ (sulfur isotope ratio)

METHODOLOGY

Data Pre-processing

Prior to analysis:

- Only isotope variables ($\delta^{13}C$, $\delta^{15}N$, $\delta^{34}S$) were used.
- Variables were standardised (mean = 0, standard deviation = 1) to prevent dominance by variables with larger numeric ranges.

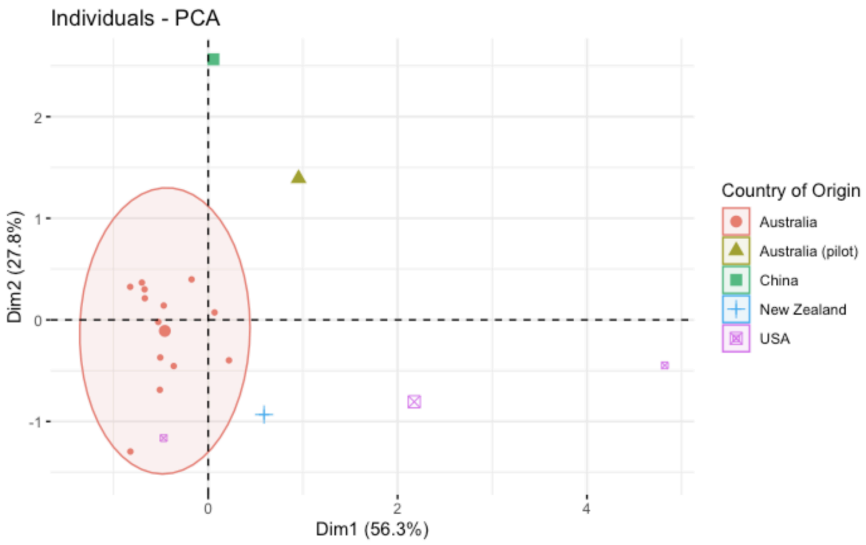
Principal Component Analysis (PCA)

PCA was applied to the scaled isotope ratios to:

- Reduce dimensionality,
- Visualise multivariate structure, and
- Identify natural groupings without using country labels during model fitting.

PCA is particularly appropriate for isotope fingerprinting because it preserves variance structure while allowing intuitive visual interpretation.

RESULTS



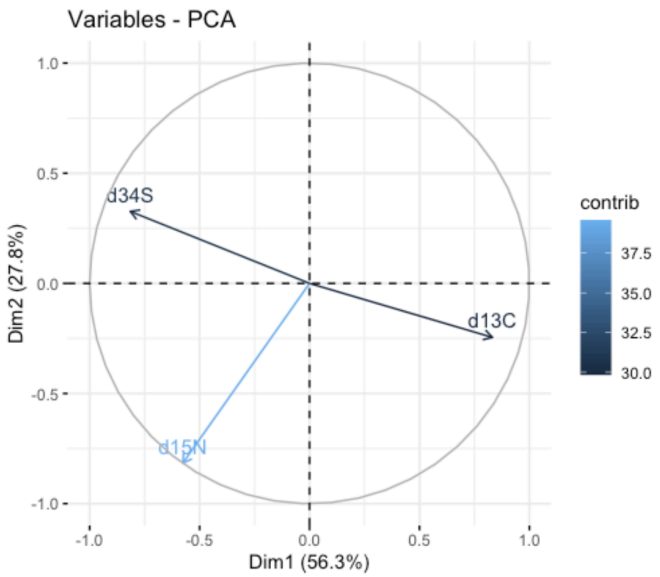
The figure above is the PCA score plot generated when PCA was performed on the data.

Variance Explained

The first two principal components explained approximately **84% of total variance**:

- **PC1:** ~56%
- **PC2:** ~28%

This indicates that a two-dimensional PCA plot provides a robust and representative summary of the isotopic variability in the dataset.



Correlation Plot Interpretation

The correlation circle helps explain which isotope ratios are driving the separation between countries. PC1 is mainly influenced by $\delta^{13}C$ and $\delta^{34}S$, which contribute in opposite directions. This indicates that carbon and sulfur isotope ratios vary inversely across the samples and are key drivers of geographic differentiation. PC2 is primarily influenced by $\delta^{15}N$, meaning that nitrogen

isotope variation contributes independently to separating countries. The relatively long vectors and their close proximity to the unit circle indicate that all three isotope systems are well represented in the first two principal components. This supports the use of these isotope ratios as effective markers for country-of-origin discrimination.

PCA Score Plot Interpretation

The PCA score plot revealed clear separation between countries of origin:

Australian samples

- Form a **tight, compact cluster**.
- Exhibit low internal variability.
- Indicate a consistent isotopic fingerprint.

The red ellipse represents ~95% confidence region for Australian samples.

New Zealand samples

- Separated from Australian samples primarily along PC2.
- Reflect distinct agricultural and environmental nitrogen and sulfur signatures.

United States samples

- Strongly separated along PC1.
- Show the most distinct isotopic profile among all groups.

Chinese samples

- Separated primarily along PC2.
- Do not overlap with the Australian confidence region.

Evaluation of the Australian Pilot Sample

The pilot sample labelled as Australian:

- Lies outside the Australian confidence ellipse.
- Does not cluster with confirmed Australian samples.
- Occupies a region of PCA space closer to non-Australian samples.

This indicates that the isotopic signature of the pilot sample is inconsistent with the established Australian reference group in this dataset.

DISCUSSION

Interpretation of Isotopic Separation

The PCA results demonstrate clear isotopic differentiation between lactoferrin samples from different countries of origin. The strong separation observed in the score plot, combined with the high proportion of variance explained by the first two principal components (84%), indicates that $\delta^{13}C$, $\delta^{15}N$, and $\delta^{34}S$ collectively capture meaningful geographic structure within the dataset.

PC1 is primarily driven by opposing contributions from $\delta^{13}C$ and $\delta^{34}S$, suggesting that carbon and sulfur isotope systems are the dominant factors distinguishing certain countries. PC2 is largely influenced by $\delta^{15}N$, indicating that nitrogen isotopic variation provides an additional and independent dimension of separation. Together, these isotope systems form a consistent multivariate fingerprint for each country group.

Confirmed Australian samples form a tight and well-defined cluster with limited internal dispersion, indicating a stable and consistent isotopic signature within the reference group. In contrast, other countries occupy distinct regions of PCA space, reflecting differences in environmental conditions, agricultural practices, feed composition, and regional sulfur cycling.

Interpretation of the Pilot Sample

The pilot sample labelled as Australian lies outside the Australian confidence ellipse and does not cluster with the confirmed Australian reference group. Its position in PCA space suggests isotopic characteristics more consistent with non-Australian samples.

Based on the three isotope systems analysed, this result indicates that the pilot sample does not share the established isotopic fingerprint of confirmed Australian lactoferrin in this dataset. While PCA is an exploratory technique and does not provide probabilistic classification, the observed separation is strong and visually distinct, supporting the conclusion of isotopic inconsistency.

Possible explanations for this discrepancy include:

- Use of imported raw milk or whey during pilot production,
- Differences in animal feed sources,
- Sourcing from a region not represented in the Australian reference dataset,
- Blended or mixed-origin inputs.

Further targeted isotope analysis or supervised classification modelling (e.g., discriminant analysis) could strengthen evidentiary confidence.

CONCLUSION

- Stable isotope ratios ($\delta^{13}C$, $\delta^{15}N$, $\delta^{34}S$) successfully discriminate lactoferrin samples by country of origin.
- Confirmed Australian samples exhibit a consistent and well-defined isotopic signature.
- The pilot sample labelled as Australian does **not** align with this signature.
- This indicates isotopic inconsistency with Australian reference material and suggests a different biological origin.