

## Discussion and Debate

### Mediterranean Bioarchaeology, Meta-Analysis and Migration

#### Towards Accurate Meta-Analyses in Mediterranean Bioarchaeology: A Critical Response to Leppard *et al.* (*JMA* 33, 2020)

Megan A. Perry<sup>1</sup>, Kristina Killgrove<sup>2</sup>, Lesley A. Gregoricka<sup>3</sup> and Tracy L. Prowse<sup>4</sup>

<sup>1</sup> Department of Anthropology, East Carolina University, Greenville, North Carolina 27858 USA  
E-mail: perrym@ecu.edu

<sup>2</sup> Department of Anthropology, University of North Carolina Chapel Hill, Chapel Hill, North Carolina 27599 USA  
E-mail: killgrove@unc.edu

<sup>3</sup> Department of Sociology, Anthropology and Social Work, University of South Alabama, Mobile, Alabama 36688 USA  
E-mail: lgregoricka@southalabama.edu

<sup>4</sup> Department of Anthropology, McMaster University, Hamilton Ontario, Canada L8S 4L9  
E-mail: prowset@mcmaster.ca

#### Introduction

We read with interest the article ‘The bioarchaeology of migration in the ancient Mediterranean: meta-analysis of radiogenic (<sup>87</sup>Sr/<sup>86</sup>Sr) isotope ratios’ by Thomas Leppard, Carmen Esposito and Massimiliano Esposito in *JMA* 33.2 (2020). As scholars also working with biogeochemistry in classical bioarchaeology, we understand the significant challenges inherent in meta-analysis of strontium isotope ratios. Unfortunately, Leppard and colleagues appear to misunderstand the nature or the structure of these data, and their study is inherently problematic. Here we explain the major problems we find with their work and conclude with suggestions for ways to move forward in tackling the question of mobility in antiquity.

In attempting to investigate ‘the inherent connectivity of the Mediterranean’ as ‘an assertion which has not been evaluated to an appropriate extent’ (p. 212), Leppard and co-authors identified 899 strontium isotope ratio data points from dozens of archaeological sites ‘spanning the period *ca.* 7500 BC to AD 700, ranging from Portugal to southwestern Jordan, and from burial contexts that cover the social spectrum’ (p. 221). Following a discussion of the sea-focused, connectivist mobility concepts of Braudel (1972), Broodbank (2013) and Horden and Purcell (2000), the authors note that, in order to better understand patterns and scale in human mobility around the ancient Mediterranean, they wish to investigate those concepts quantitatively. We are supportive of the use of archaeologically

derived data to answer questions of historical importance, but we have numerous questions about the data and methods that Leppard and co-authors chose for this analysis attempt.

## Data Issues

### *Inclusion and Exclusion Criteria*

Leppard and colleagues include in their paper 39 datasets from around the Mediterranean, representing the vast majority of published strontium isotope data. The authors note that published datasets containing fewer than five individuals were not included in the paper or the data analysis, but exclusion criteria beyond this were unclear, leading to probable dataset omissions (e.g., Cangemi 2016; Salesse *et al.* 2018). However, datasets containing between five and 19 samples were included in the paper, specifically in the tables and supplemental file, only to be later excluded in statistical analysis. While this appears to make sense from a mathematical perspective, logically this choice is problematic, for two reasons.

First, there is a tendency in academic publications to emphasize variation in data, not lack of variation. In the case of strontium isotope studies, those that show no evidence of immigration (and thus no isotopic variation) are less likely to be published than those which do show some evidence of migration and resulting variation in data. Relying on published data biases the meta-analysis toward heterogeneous datasets. Second, excluding small sample sizes, while logical for creating a statistically robust dataset, excludes publications of small samples with high frequencies of immigration (e.g., M. Candelaro, Grotta Scaloria, Fuente Hoz, Pharsala) and therefore introduces additional bias.

Second, the inclusion of  $n = 5-19$  datasets in 'original calculations' but exclusion from 'recomputed  $2\sigma$  from human enamel mean' leads to major errors. In fact, of the 39 datasets included for their original calculations, only 19 were subjected to reanalysis (48.7%). In order to under-

stand why this is problematic, take the Middle Bronze Age (MBA) to early Iron Age (IA) samples ( $n = 352$ ) from Table 1. The original publications identified (by comparison to baselines) 33 immigrants, for a frequency of 9.4%. Leppard and colleagues recalculated the expected number of immigrants based on their human mean  $\pm 2\sigma$  model, *after excluding small datasets* ( $n = 5-19$ ), and got 15 out of 258, or 5.8% immigrants. If we were to exclude those same datasets and recalculate *the original numbers*, however, it would sum to 12 nonlocals identified out of 258, or 4.7%, which is much closer to Leppard and colleagues' number (see Table 1). The same issue can be seen in their Neolithic data. They include in both the original calculation and their recalculations the dataset from M. Candelaro ( $n = 16$ ), even though they stated (p. 220) that datasets of this size would be excluded.

Consistently applying the rules that Leppard and colleagues set for themselves to both the original and the recalculations, the overall percentages as tabulated by them (pp. 222-23, table 1) need to be revised as shown in our Table 1. The lack of consistency results in a large discrepancy in the MBA-early IA time period due to their failure to include Kültepe ( $n = 25$ ) in their recalculation, as well as a discrepancy in the Neolithic in failing to exclude M. Candelaro ( $n = 16$ ). The authors provided no explanation for such decisions, in either the text or the supplemental file. As our revised Table 1 demonstrates, when the method the authors advocate is applied consistently across their collected dataset, the resulting differences in immigrant frequency are lessened, and a two-tailed Fisher's exact test of the two methods reveals no significant difference ( $p = 0.66$ ). In essence, Leppard and colleagues failed to compare apples to apples by selectively excluding more than half of the datasets in the  $n = 5-19$  range from their recalculation.

### *Tissue Choice and the Definition of 'Immigrant'*

A second major problem in the paper is the lack of control for tissue type, and therefore the variably

**Table 1.** Number and percentage of nonlocals by time period, as presented by original researchers and by Leppard and colleagues.

Phase	Sites with $n \geq 19$	N. Samples	Original researchers' calculation of immigrants (cf. baseline data)	Leppard and colleagues' calculation of immigrants (cf. human mean $\pm 2\sigma$ )
Neolithic	Basta Alto d. Reinoso Bom Santo	76	10/76 (13.2%)	7/76 (9.2%)
EBA	El Robollosillo Marroquies	130	6/130 (4.6%)	7/130 (5.4%)
MBA-late IA	Los Berrocales Knossos Sant'Eurosia Kultepe Scalvinetto Tell Dothan Casinalbo Frattresina	258	12/258 (4.7%)	15/258 (5.8%)
Late IA	Monte Bibebe Castellaccio Europarco (Imperial) Casal Bertone Isola Sacra Velia Vagnari	222	18/222 (8.1%)	12/222 (5.4%)
<b>Total</b>		686	46/686 = 6.7%	41/686 = 6.0%

implied definition of 'immigrant'. Their abstract makes reference to 'post-juvenile migration', but the supplemental material demonstrates that the secondary data they utilize come from different teeth forming at different ages, likely chosen as a result of an opportunistic sampling strategy common in bioarchaeology. For example, in Roman Imperial Italy, the sample sets of Velia and Isola Sacra are mostly composed of second molars, whose crowns are complete around age eight (AlQahtani *et al.* 2010), while Vagnari and Casal Bertone are represented by first molars, whose crowns are completed by about four years old (AlQahtani *et al.* 2010). In these examples, depending on how the enamel was sampled, an immigrant in the first case could be defined as someone who changed geographical location after the age of eight, while an immigrant in the second case would be someone who moved at

age four or older. Other datasets mentioned by Leppard and colleagues employ canines (crown completion around age seven), premolars (age seven to eight) and third molars (which finish forming around age 15) (AlQahtani *et al.* 2010). Still other datasets do not reference the specific tooth (e.g., Bom Santo, Pharsala, Chloe), while one dataset includes deciduous dentition (Marroquies), and five more include petrous portions (El Robollosillo, Scalvinetto, Casinalbo, Frattresina) or unspecified skull data (Gòzquez).

There is no evidence that the authors excluded the deciduous data, in spite of the fact that this enamel is formed during a period in which an infant is typically breastfed exclusively and therefore likely to have its mother's strontium signature rather than its own; also, there is no evidence that they discussed issues involved in comparing petrous/skull data with enamel data

(e.g., Weber *et al.* 2020) or with using tissue that begins forming in utero (e.g., first molars and petrous portions).

We argue that we should not necessarily expect rates of migration to be similar among all non-adult age categories in antiquity (see Scaffidi and Knudson 2020 for an example from the Andes). A thorough meta-analysis needs to control for tissue type and to define ‘immigrant’ clearly, which their study does not do.

#### *Lack of Temporal and Chronological Controls*

A third issue with the authors’ choice of datasets is their temporal and spatial range. In order to attempt a meta-analysis, they have inappropriately collapsed these differences for the sake of using a single method to come up with a frequency of migration for each archaeological site. The cultural definition of ‘immigrant’ almost certainly differed across the ancient Mediterranean; subsequently, in many locations, the level of geological variation may mean that those individuals isotopically identified as ‘nonlocal’ would culturally not be considered ‘immigrants’ by other residents of the site (Perry 2018; Gregoricka 2021). Why should we assume that 11 migration data points from elite Grave Circle A in sixteenth-century BC Mycenae are at all comparable to 79 data points from non-elite Casal Bertone in second–fourth century AD Rome? We know from migration theory in anthropology (e.g., Brettell and Hollifield 2015) that different patterns of migration have been observed in different locations (e.g., towns, cities), in different circumstances (e.g., displacement following a war), and among different populations (e.g., slaves). Leppard and colleagues anticipate this objection (p. 221) but dismiss it, choosing instead to move ahead with their imposition of a statistical ‘baseline’ on 19 out of 39 datasets across 3000 km and 8000 years. How representative are these 700 individuals, and what information is lost when we use this minimal dataset to try to understand movement patterns across this vast expanse?

#### *Lack of Inter-Laboratory Data Normalization*

Fourth, the authors do not seem to be aware of how  $^{87}\text{Sr}/^{86}\text{Sr}$  values are measured and reported by laboratories, which is a key concern when engaging in an inter-site and inter-laboratory analysis. Strontium isotope data acquired via a thermal ionization mass spectrometer (TIMS) or multi-collector-inductively coupled plasma-mass spectrometer (MC-ICPMS) are reported relative to a NIST SRM/NBS 987 instrument standard (Goldstein *et al.* 2003), which can vary dramatically between laboratories. There is also a slight difference between  $^{87}\text{Sr}/^{86}\text{Sr}$  measured by TIMS versus MC-ICPMS (Wall *et al.* 2016: 16). Therefore,  $^{87}\text{Sr}/^{86}\text{Sr}$  values from different labs need to be normalized to a common value before their comparison or aggregation (Wall *et al.* 2016: 16). It is not clear that the authors did so, as no NBS standard or mention of normalization was reported in the article or any accompanying data.

In sum, we have serious concerns about the number of errors in Leppard and colleagues’ data. We counted more than two dozen simple math errors in Table 1 alone, and these errors are replicated in the Supplemental Excel file. For example, they list the number of samples from Castellaccio Europarco (Republican) as six. However, they incorrectly calculate the percentage of locals (5/6, which they list as 80%) and nonlocals (1/6, which they list as 20%). Other numbers do not add up; we came up with different sample counts for Sant’Eurosia and Bologna from the Excel file than what is included in Table 1. Additionally, the authors appear to have chosen the number of decimal places more or less randomly, sometimes using one decimal place for percentages, sometimes two. While the occasional typo is understandable, the number of errors we noticed at first glance made us delve further into this paper, which is when we noticed additional statistical issues.

## Methodological Issues

### *Problems with Using Human Dental Enamel Mean $\pm 2\sigma$ to Establish the Local Range*

We take issue additionally with how the authors identified immigrants within a sample population. While we agree that identification of nonlocals is a general methodological issue in strontium isotope studies (e.g., Grimstead *et al.* 2017; Holt *et al.* 2021), the way the authors attempt to circumvent the issue with their analysis is flawed from both a mathematical and a logical perspective. After outlining correctly some of the methodological constraints in the sample populations (pp. 219-20) in terms of defining strontium baselines to which human data can be compared, the authors note that they ‘attempted to overcome’ the issue by

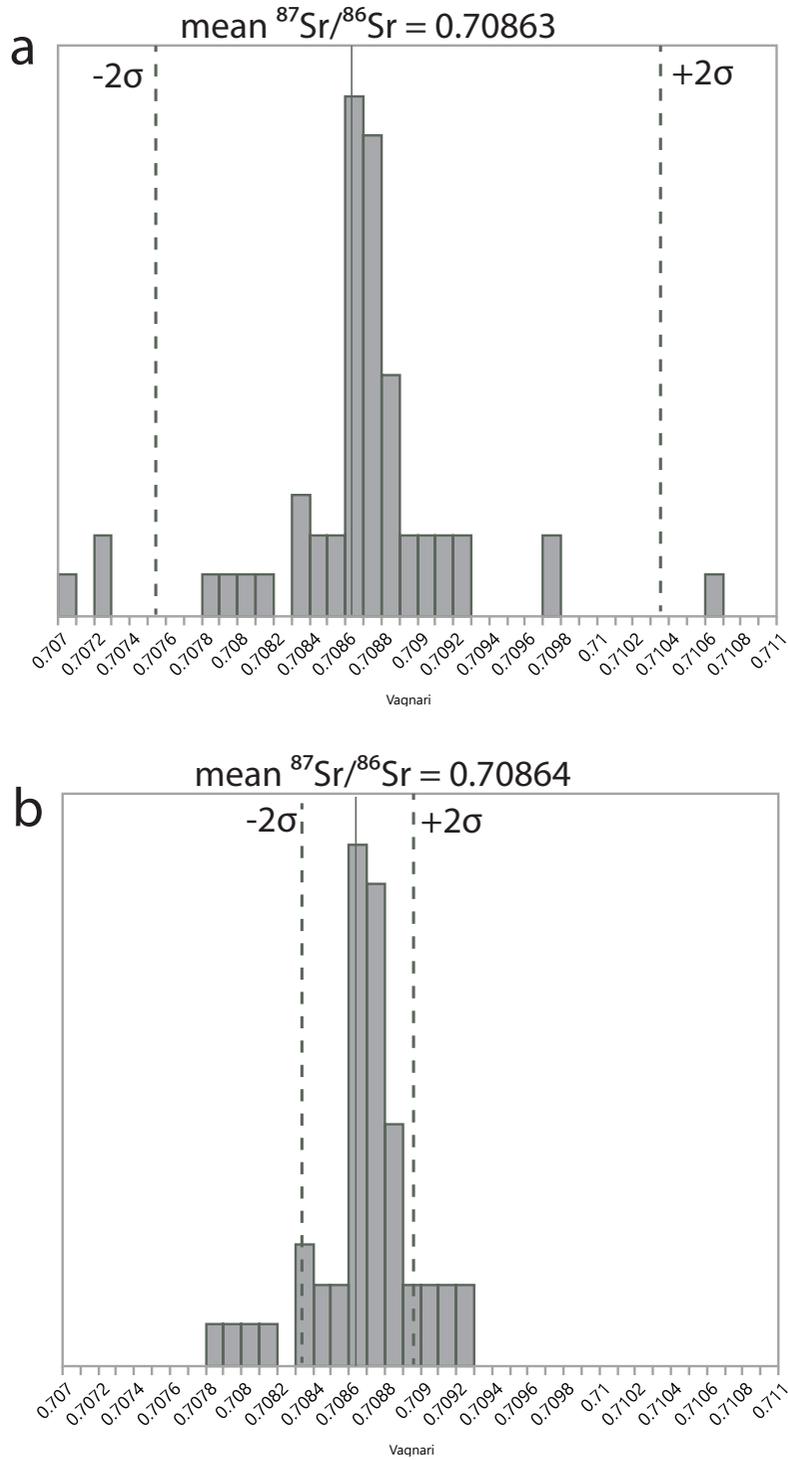
adopting a baseline computed from mean [human] tooth enamel values  $\pm 2\sigma$ , excluding baseline samples (flora, fauna, soil, water etc.) analyzed in the [original] papers, and [only] considering sites with a number of samples  $\geq 19$  (i.e., datasets large enough to compute robustly standard deviations from a mean).  
(p. 220)

Relying on identifying immigrants through outlying human dental enamel values was used in early archaeological isotope studies, but was quickly abandoned due to problems with identifying a cutoff threshold and relying on the central limit theorem for characterizing these data (e.g., discussion in Scaffidi and Knudson 2020). None of these studies, however, as far as we can tell, ever utilized a standard deviation of the sample under investigation to identify any nonlocals it might contain, instead relying on the inter-quartile range (IQR) *along with* environmental-based isotopic data (e.g., Wright 2005; Perry *et al.* 2009; 2017). In general, data points should not be identified as statistical outliers through calculating the mean and standard deviation, as the very presence of those outliers impacts the range of the standard deviation itself (Cousineau and Chartier 2010; Leys *et al.*

2013; see also Lightfoot and O’Connell 2016 in relevance to establishing a local  $\delta^{18}\text{O}$  range). Therefore, the presence of an extreme outlier may increase the standard deviation to the point that other possible outliers (i.e., nonlocals) with less divergent values would not fall outside of the standard deviation-established range.

A related factor is that the mean  $\pm 2\sigma$  method assumes a normal distribution, which is generally not the case for strontium isotope data. As noted by Scaffidi and Knudson (2020), the attritional nature of isotopes in dental enamel bulk samples and the vagaries of human diet by age may negate any expectations of centrality in these data. In addition, in roughly normal distributions, using a  $2\sigma$  cutoff will always identify nonlocal individuals because, by mathematical definition, 5% of the data points are  $2\sigma$  from the mean. These are in fact not outliers, but simply the tails of the distribution. In this case,  $2\sigma$  provides an arbitrary cutoff with no explanatory value.

We can demonstrate both of these issues with the dataset from Vagnari ( $n = 56$ ) that Leppard and co-authors cite. Figure 1a below presents a histogram of the dataset with the human dental enamel-derived  $\pm 2\sigma$  range identified via the vertical dotted lines ( $^{87}\text{Sr}/^{86}\text{Sr} = 0.7075293 - 0.7097349$ ). In this case, three individuals fall below and one individual falls above the range, although there appear to be at least two possible additional outliers with values between 0.7097 and 0.7098. However, if we were to exclude these outliers and recalculate the  $\pm 2\sigma$  range, as shown in Figure 1b, some individuals initially identified as ‘locals’ now are considered to be ‘immigrants’, even though this is theoretically the same sample, just one that is completely locally born. This happens because, whatever the distribution, and whether it is normal or not, there will always be observations falling outside of the  $\pm 2\sigma$  range, and thus, immigrants will always be identified in every sample analyzed. Thus, across multiple sites (El Robollosillo, Knossos, Tell Dothan, Frattesina, Velia) where no immigrants were identified by the



**Figure 1.** Comparing distributions of Vagnari strontium isotope data with (a) and without (b) outliers (the dashed lines represent the  $\pm 2\sigma$  from the mean).

original authors, the method used by Leppard and colleagues forces ‘migrants’ to be generated.

Figure 1b also demonstrates that the  $\pm 2\sigma$  range may not present a logical threshold in the data. Why in this case is the upper boundary of 0.70893 a ‘magic number’, separating those with fractionally different diets into immigrants and locals? This threshold has little to do with the actual data or, more importantly, with its context. Indeed, the data’s explanatory value (or the ability to identify an immigrant and their possible place of origin) is generated when human strontium data are compared to other datasets. This is the reason that independent baseline data from local geology, water, animals and/or plants were created separately by every researcher whose study was included by the authors. And this is the reason that the authors’ study is so problematic: they ‘recalculate’ geographically and temporally specific human data that have been calibrated to a local baseline and, in almost all cases, reduce the site-specific variation to a mathematically general *ca.* 5% (p. 226, fig. 3; compare with our Figure 2 and Table 1). Furthermore, their assertion that datasets over 19 individuals are more statistically sound also appears to relate to assumptions about the normal distribution; in a sample size of 20, application of the human enamel mean  $\pm 2\sigma$  method is likely to identify one whole individual as an immigrant.

These problems with the standard deviation method used by Leppard and colleagues can also be demonstrated in an ancient population with a high frequency of immigration at a particular time. For example, in a study not included in the meta-analysis, analysis of human burials from Late Byzantine Aila in Jordan identified 19 out of 23 nonlocal individuals based on a local  $^{87}\text{Sr}/^{86}\text{Sr}$  range established through a number of environmental and cultural variables (Perry *et al.* 2017). The four definitively locally born individuals (and even the additional 10 possible locals) all cluster at the low end of the distribution of  $^{87}\text{Sr}/^{86}\text{Sr}$  from the site, with the nonlocals having higher values. Using Leppard and

colleagues’ human enamel mean  $\pm 2\sigma$  method, the locals at the lowest end of the range and the nonlocals with the highest values would all have been considered as immigrants, which completely disregards the geological and environmental context of the site.

By employing the simplistic mathematical technique of mean  $\pm 2\sigma$ , the authors have naively critiqued research studies that are far more complex than they understood them to be. The result is that they appear to be begging the question, in the logical sense of the phrase: they imply that the frequency of migration in the Mediterranean has been exaggerated and set out to prove it by choosing a mathematical method with no explanatory value and by ignoring half of their datasets in order to deflate artificially the frequency of migration found by the original researchers. We understand the desire to perform a meta-analysis to answer questions about ancient migration—but the reason that the Leppard and colleagues’ article is ‘the first meta-analysis of Mediterranean strontium isotope datasets derived from human funerary populations’ (p. 212) is not because they were the first to consider or even to attempt this. It is because decades of work on archaeological strontium studies have shown that meta-analysis of strontium isotope ratios is incredibly difficult to do correctly.

#### *Inter-Quartile Range (IQR) as a Better Approach when Disregarding Baseline Data*

This article initially came to our attention because, to our knowledge, only one successful meta-analysis of archaeological strontium isotopes has been published (Scaffidi and Knudson 2020). In that article, Scaffidi and Knudson carefully lay out a complex model for comparing  $^{87}\text{Sr}/^{86}\text{Sr}$  values in the prehistoric Andes. The dataset they generated for their meta-analysis included time periods, biological sex and tissue type, as well as the developmental age for the tissue (particularly for adult dental enamel; deciduous teeth were excluded). Descriptive statistics helped them understand rough proxies of population diversity,

and a slew of statistical tests (ANOVA, Bartlett, CI/Bonferroni,  $\chi^2$ ) were used to investigate statistical outliers between sites. Importantly, they also included geospatial analysis using ArcMap, both to investigate altitudinal differences in strontium signatures and to identify outlier  $^{87}\text{Sr}/^{86}\text{Sr}$  values in tissues and plants. Finally, they generated a predictive model for the strontium isoscape using Universal Kriging and cross-validated it with ArcMap. In doing this important mathematical background work, Scaffidi and Knudson additionally found that, for the populations they were working with, the inter-quartile range (IQR) of the human samples was a better measure of variance than  $\pm 2\sigma$  for their meta-analysis, as it is less affected by outliers skewing the measure.

Across the prehistoric Andes, this meta-analysis found the following: (1) there were greater fre-

quencies of immigrants in later time periods, consistent with past research; (2) enamel strontium values from early childhood were more variable, suggesting family mobility patterns; and (3) males were more mobile than females, particularly as adults, suggesting more limited social roles for women. They concluded that their meta-analysis is useful but with limitations: in particular, a single-isotope model suffers from issues of equifinality or overlapping of ranges. Scaffidi and Knudson (2020: 15), however, suggest that the creation of baseline food chain models and strontium isotope models using spatially-explicit methods is important because ‘these models, combined with multifaceted and deeply-contextualized studies of prehistoric subsistence, trade, and mobility patterns’, will ‘help to define expected baseline ranges within isotopic

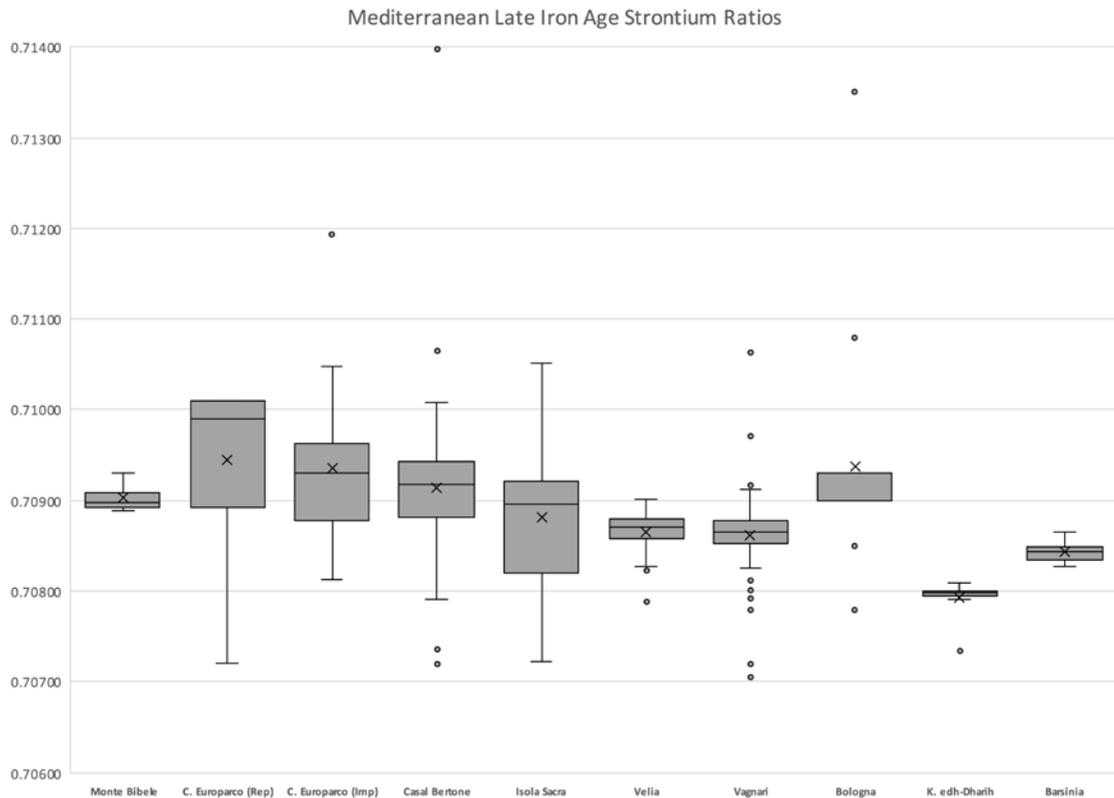
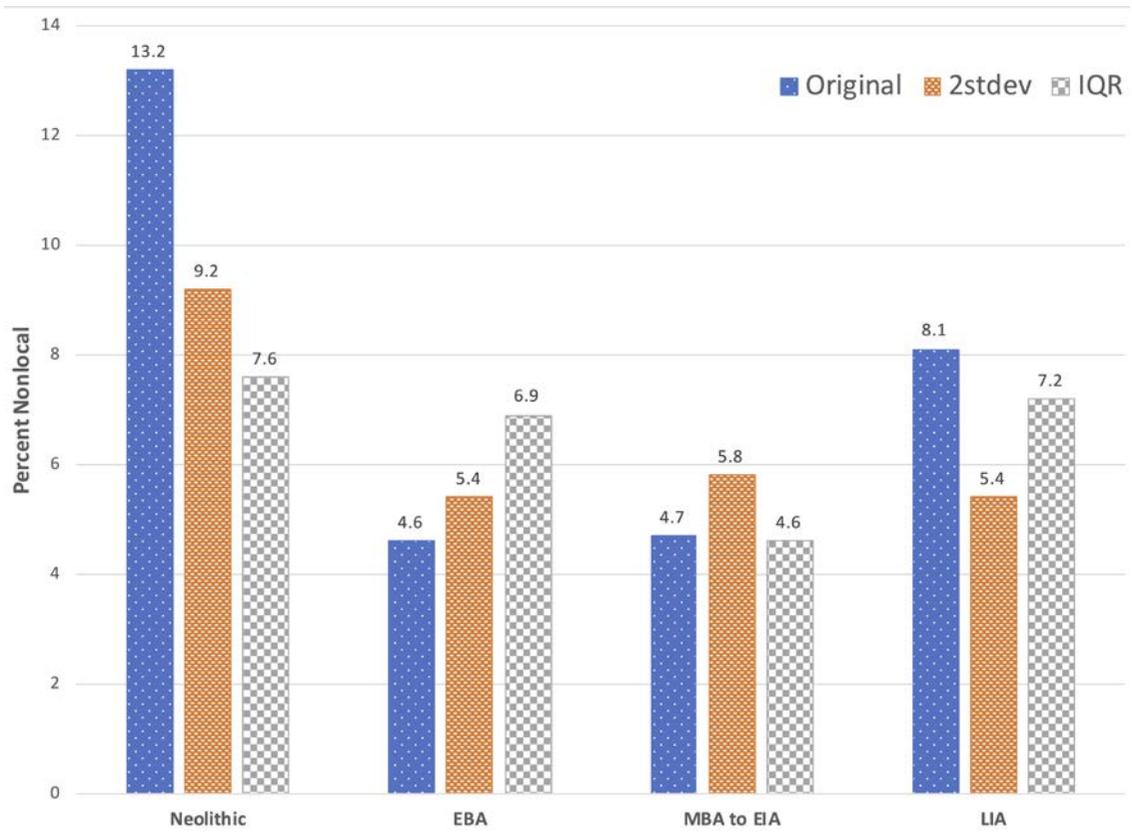


Figure 2. IQR box plots of the late IA datasets included in the article by Leppard and colleagues.



**Figure 3.** Identification of outliers using original (baseline) method, the mean  $\pm 2\sigma$  method and the IQR method, for every site with 19 or more individuals.

catchment areas, and help us move toward probabilistic determination of sample origins within these complex biogeographical networks’.

Because Leppard and colleagues have provided a (mostly correct) dataset in their Supplemental File, it was easy to create a series of box plots to illustrate what IQRs for the individual sites look like. For example, taking the late IA sites (Figure 2), the series of box plots demonstrates the mean and range of the strontium isotope data, and also identifies outliers as dots.

If we again reduce the data to exclude sites under 19 individuals, we are left with the Italian sites Monte Bibebe (0 immigrants), Castellaccio E. Imperial (1), Casal Bertone (4), Isola Sacra (0), Velia (2) and Vagnari (9), for a total of 16 statistical outliers who may be immigrants. This

number is closer to the original researchers’ calculation (18) than it is to Leppard and colleagues’ standard deviation calculation (12). While we calculated IQRs for every site, we are not including all of the box plot graphs here. Instead, again using the criteria the authors cite for inclusion of datasets, we produced an alternative to their Figure 3, showing the original researchers’ outlier findings compared to baseline data, their standard deviation method and the IQR method (Figure 3). If we exclude all the datasets from the original calculations that they do from their recalculation, it would look something more like Figure 3.

The frequency of nonlocals per historical period is clearly still not consistent across the three methods of baseline, standard deviation

and IQR. However, only the standard deviation method produced values consistently surrounding a statistically expected 5%. Therefore, the human dental enamel mean  $\pm 2\sigma$  model employed by Leppard and colleagues is not a good approximation of the actual frequency of either outliers or immigrants in the ancient Mediterranean. IQR has its own issues as a statistical model, and as such is typically only used in an exploratory manner or when there are no known baseline data. Comparing ancient human strontium isotope ratios to a separately constructed baseline made up of nonhuman samples remains the standard for both identifying and interpreting these ratios (Price *et al.* 2002; Bentley 2006), regardless of the statistical methods used to establish the range (standard deviation, IQR, etc.). That said, an exploratory statistical model for human migration in antiquity that cannot take into account differences in chronology or geography would be better if based on the IQR, as it is robust to outliers. Any future meta-analysis of the ancient Mediterranean should deal more fully with statistical tests of range and variance in addition to the geography of strontium distribution and the biocultural understandings of immigrants.

## Conclusion

We value the idea of a meta-analysis of immigration in the ancient Mediterranean and have written numerous publications on this topic using not only strontium but also oxygen and lead isotopes (e.g., Prowse *et al.* 2007; Perry *et al.* 2009; Montgomery *et al.* 2010; Killgrove and Montgomery 2016; Gregoricka and Sheridan 2017). We also find it interesting that overall rates of migration appear to have decreased over time from the Neolithic through the Roman Imperial period, which is a result that holds in the original research calculations, in Leppard and colleagues' human dental enamel mean  $\pm 2\sigma$  simplifications and in our IQR models. A well-done meta-analysis of human

strontium isotopes, however, has the potential to show us where there were unexpectedly high frequencies of immigrants and unexpectedly low frequencies, in addition to variation across time and space, as well as patterning of migration in terms of sex and age.

Instead, for the reasons discussed above, the paper by Leppard and colleagues is not a valid meta-analysis of ancient human strontium isotope ratios. And while we agree with them that 'meta-analysis of diachronic human migration in the Mediterranean' is worthy of research, we disagree that it is 'a matter of urgent interest' (p. 231). Given the currently published database of well under 1000 individual human  $^{87}\text{Sr}/^{86}\text{Sr}$  isotope ratios across dozens of sites, across hundreds of kilometers, and over 8000 years (see also Salese *et al.* 2018 on IsoArch), we do not find the strontium isotope data robust enough to go beyond site-level interpretations right now. That said, we do agree that 'synthetic studies of human residence mobility are now clearly in order, as are more nuanced studies of geographic variation and broader attempts to knit together radiogenic isotope, artifactual, and ancient genomic data' (p. 231); indeed, such studies have been published for several years (e.g., Killgrove and Montgomery 2016; Killgrove 2017).

While we appreciate the effort of researchers to engage in *post hoc* analyses of published isotope data and welcome differing opinions on our own work, we strongly advise these scholars to ensure they fully understand both the creation and structure of the isotope data they purport to use. Misunderstandings of isotope data have been published in classical archaeology journals in the past (e.g., Bruun 2010 on Prowse *et al.* 2007, to which Killgrove 2010 is a response), and it is unfortunate that Leppard and colleagues similarly failed to understand the complexity of the data they gathered. Each of us has long advocated a bioarchaeology of the Roman world that represents a fully multidisciplinary approach, where experts in osteology, biochemistry, history, statistics, archaeology and material

culture come together to understand the ancient past better. Leppard and colleagues certainly raise several interesting theoretical points about the classical conception of an interconnected Mediterranean in their *JMA* article, but the bulk of their contribution is a mathematically flawed attempt to harness complex biochemical data. As such, we caution scholars not to take the conclusions of the article at face value and, instead, suggest the interested reader return to the original publications, which appropriately contextualize the archaeological sites and bioarchaeological samples.

### About the Authors

Megan Perry is a bioarchaeologist and Professor of Anthropology at East Carolina University whose research focuses on paleopathology and biogeochemistry of communities in first-century BC to seventh-century AD Jordan. She currently directs the Petra North Ridge Project in Jordan.

Kristina Killgrove is a bioarchaeologist whose research focuses on understanding the lives of non-elites in Italy during the Imperial period. Her latest projects are a DNA analysis of human remains from Gabii and a full osteological report of the individuals who died at Oplontis in AD 79.

Lesley Gregoricka is a bioarchaeologist and Associate Professor of Anthropology at the University of South Alabama. Her research focuses on prehistoric mortuary practices and the biogeochemistry of ancient human skeletal remains to examine changing patterns of mobility and social complexity across Arabia and the Levant.

Tracy Prowse is an Associate Professor in the Department of Anthropology at McMaster University. She is the director of excavations at the Roman site of Vagnari, southern Italy. She has published on diet, mobility and health in

Roman Italy through the use of stable isotopes and paleopathological analyses.

### References

- AlQahtani, S., M. Hector and H. Liversidge  
2010 Brief communication: the London atlas of human tooth development and eruption. *American Journal of Physical Anthropology* 142: 481-90. <https://doi.org/10.1002/ajpa.21258>
- Bentley, R.A.  
2006 Strontium isotopes from the earth to the archaeological skeleton: a review. *Journal of Archaeological Method and Theory* 13: 135-87. <https://doi.org/10.1007/s10816-006-9009-x>
- Braudel, F.  
1972 *The Mediterranean and the Mediterranean World in the Age of Philip II*. 2 vols. New York: Harper and Row.
- Brettell, C., and J.F. Hollifield  
2015 *Migration Theory: Talking across Disciplines*. Boca Raton, Florida: Taylor and Francis.
- Broodbank, C.  
2013 *The Making of the Middle Sea: A History of the Mediterranean from the Beginning to the Emergence of the Classical World*. London: Thames and Hudson.
- Bruun, C.  
2010 Water, oxygen isotopes, and immigration to Ostia-Portus. *Journal of Roman Archaeology* 23: 109-32. <https://doi.org/10.1017/S1047759400002324>
- Cangemi, I.  
2016 A Scale-Free, Relational Approach to Social Development in Late-Prehistoric Tyrrhenian Central Italy. PhD dissertation, University of Michigan.
- Cousineau, D., and S. Chartier  
2010 Outliers detection and treatment: a review. *International Journal of Psychological Research* 3: 58-67. <https://doi.org/10.21500/20112084.844>
- Goldstein, S.L., P. Deines, E.H. Oelkers, R.L. Rudnick and L.M. Walter  
2003 Standards for publication of isotope ratio and chemical data in *Chemical Geology*. *Chemical Geology* 202: 1-4. <https://doi.org/10.1016/j.chemgeo.2003.08.003>
- Gregoricka, L.  
2021 Moving forward: a bioarchaeology of mobility

- and migration. *Journal of Archaeological Research* 29: 581-635. <https://doi.org/10.1007/s10814-020-09155-9>
- Gregoricka, L., and S.G. Sheridan  
2017 Continuity or conquest? A multi-isotope approach to investigating identity in the Early Iron Age of the southern Levant. *American Journal of Physical Anthropology* 162: 73-89. <https://doi.org/10.1002/ajpa.23086>
- Grimstead, D.N., S. Nugent and J. Whipple  
2017 Why a standardization of strontium isotope baseline environmental data is needed and recommendations for methodology. *Advances in Archaeological Practice* 5: 184-95. <https://doi.org/10.1017/aap.2017.6>
- Holt, E., J.A. Evans and R. Madgwick  
2021 Strontium ( $^{87}\text{Sr}/^{86}\text{Sr}$ ) mapping: a critical review of methods and approaches. *Earth-Science Reviews* 216: 103593. <https://doi.org/10.1016/j.earscirev.2021.103593>
- Horden, P., and N. Purcell  
2000 *The Corrupting Sea: A Study in Mediterranean History*. Oxford: Blackwell.
- Killgrove, K.  
2010 A response to C. Bruun, 'Water, oxygen isotopes and immigration to Ostia-Portus'. *Journal of Roman Archaeology* 23: 133-36. <https://doi.org/10.1017/S1047759400002336>
- 2017 Imperialism and physiological stress in Rome and its environs (1st–3rd centuries AD). In M.S. Murphy and H.D. Klaus (eds.), *Colonized Bodies, Worlds Transformed: Toward a Global Bioarchaeology of Contact and Colonialism*, 247-77. Gainesville, Florida: University Press of Florida. <https://doi.org/10.2307/j.ctvx0725r.14>
- Killgrove, K., and J. Montgomery  
2016 All roads lead to Rome: exploring human migration to the eternal city through biochemistry of skeletons from two Imperial-era cemeteries (1st–3rd c AD). *PLoS One* 11: e0147585. <https://doi.org/10.1371/journal.pone.0147585>
- Leys, C., C. Ley, O. Klein, P. Bernard and L. Licata  
2013 Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology* 49: 764-66. <https://doi.org/10.1016/j.jesp.2013.03.013>
- Lightfoot, E., and T.C. O'Connell  
2016 On the use of biomineral oxygen isotope data to identify human migrants in the archaeological record: intra-sample variation, statistical methods and geographical considerations. *PLoS ONE* 11: e0153850. <https://doi.org/10.1371/journal.pone.0153850>
- Montgomery, J., J.A. Evans, S.R. Chenery, V. Pashley and K. Killgrove  
2010 'Gleaming, white and deadly': using lead to track human exposure and geographic origins in the Roman period in Britain. In H. Eckardt (ed.), *Roman Diasporas: Archaeological Approaches to Mobility and Diversity in the Roman Empire*. Journal of Roman Archaeology Supplementary Series 78: 199-226. Portsmouth, Rhode Island: Journal of Roman Archaeology
- Perry, M.A.  
2018 Anatomy of restlessness: strontium isotopes and human migration in the Graeco-Roman Near East. In J. Yoo, A. Zerbini and C. Barron (eds.), *Migration and Migrant Identities in the Near East from Antiquity to Middle Ages*, 187-213. London: Routledge. <https://doi.org/10.4324/9781351254762-9>
- Perry, M.A., D.S. Coleman, D.L. Dettman and A.H. al-Shiyab  
2009 An isotopic perspective on the transport of Byzantine mining camp laborers into southwestern Jordan. *American Journal of Physical Anthropology* 140: 429-44. <https://doi.org/10.1002/ajpa.21085>
- Perry, M.A., C. Jennings and D.S. Coleman  
2017 Strontium isotope evidence for long-distance immigration into the Byzantine port city of Aila, modern Aqaba, Jordan. *Archaeological and Anthropological Sciences* 9: 943-64. <https://doi.org/10.1007/s12520-016-0314-3>
- Price, T.D., J.H. Burton and R.A. Bentley  
2002 The characterization of biologically available strontium isotope ratios for the study of prehistoric migration. *Archaeometry* 44: 117-35. <https://doi.org/10.1111/1475-4754.00047>
- Prowse, T.L., H.P. Schwarcz, P. Garnsey, M. Knyf, R. Macchiarelli and L. Bondioli  
2007 Isotopic evidence for age-related immigration to imperial Rome. *American Journal of Physical Anthropology* 132: 510-19. <https://doi.org/10.1002/ajpa.20541>
- Salesse, K., R. Fernandes, X. de Rochefort, J. Brůžek, D. Castex and E. Dufour  
2018 IsoArcH.eu: an open-access and collaborative

- isotope database for bioarchaeological samples from the Graeco-Roman world and its margins. *Journal of Archaeological Science, Reports* 19: 1050-55. <https://doi.org/10.1016/j.jasrep.2017.07.030>
- Scaffidi, B.K., and K.J. Knudson  
2020 An archaeological strontium isoscape for the prehistoric Andes: understanding population mobility through a geostatistical meta-analysis of archaeological  $^{87}\text{Sr}/^{86}\text{Sr}$  values from humans, animals, and artifacts. *Journal of Archaeological Science* 117: 105121. <https://doi.org/10.1016/j.jas.2020.105121>
- Wall, A.J., R.C. Capo, B.W. Stewart, T.T. Phan, J.C. Jain, J.A. Hakala and G.D. Guthrie  
2016 *High-Throughput Method for Strontium Isotope Analysis by Multi-Collector Inductively Coupled Plasma Mass Spectrometer* NETL Technical Reports Series 12-2016. Pittsburg, Pennsylvania: National Energy Technology Laboratory. <https://doi.org/10.2172/1361494>
- Weber, W., T. Tacail, F. Lugli ... R. Mertz-Kraus and T. Tüken  
2020 Strontium uptake and intra-population  $^{87}\text{Sr}/^{86}\text{Sr}$  variability of bones and teeth – controlled feeding experiments with rodents (*Rattus norvegicus*, *Cavia porcellus*). *Frontiers in Ecology and Evolution* 8: 569940. <https://doi.org/10.3389/fevo.2020.569940>
- Wright, L.E.  
2005 Identifying immigrants to Tikal, Guatemala: defining local variability in strontium isotope ratios of human tooth enamel. *Journal of Archaeological Science* 32: 555-566. <https://doi.org/10.1016/j.jas.2004.11.011>
- 

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]