GC-MS Analysis and PCA-Based Statistical Evaluation of Lipid Residues in Terracotta Pottery from the Buddhist Site of Badalpur, Taxila Valley, Pakistan

Abdul Basit[1] and Ghani-ur-Rahman[2]

https://doi.org/10.62345/jads.2025.14.1.119

Abstract

This study provides a detailed analysis of organic residues preserved in terracotta pottery from the Buddhist archaeological site of Badalpur, situated in the Taxila Valley, Pakistan. Gas Chromatography-Mass Spectrometry (GC-MS) was employed to identify and characterize lipid compounds in 63 pottery fragments, which were classified into five functional categories: oil lamps, cooking pots, storage vessels, bowls, and dish plates. The molecular analysis revealed a diverse range of saturated and unsaturated fatty acids, indicating the presence of both plant- and animal-derived substances. Principal Component Analysis (PCA) was applied to statistically evaluate compositional variations and to identify potential functional patterns among pottery types. The results indicate significant differentiation in fatty acid profiles, suggesting varied pottery use related to food preparation, storage, and ritual activities. This integrative analytical approach offers new insights into ancient dietary habits, resource utilization, and domestic practices at the Badalpur site.

Keywords: Badalpur, GC-MS, PCA, Fatty Acids, Pottery Residue, Archaeological Science.

Introduction

In recent decades, research has increasingly demonstrated that organic residues, particularly lipids, can be preserved within archaeological materials, most notably pottery (Briuer, 1976; Evershed et al., 1992; Loy, 1994; Pollard & Heron, 1996; Pollard et al., 2007). These residues, often absorbed into the porous, unglazed matrix of ceramic vessels or deposited as visible surface films, originate from a range of activities such as food processing, cooking, storage, and non-food uses like sealing (Heron & Evershed, 1993; Irto et al., 2022). Lipids, including fats and waxes, are particularly resilient and are frequently analyzed due to their ability to withstand microbial activity and environmental degradation over long periods (Evershed et al., 2008a, 2008b).

Among lipid compounds, free fatty acids are the most commonly encountered and extensively studied in archaeological ceramics (Evershed, 1993; Evershed et al., 2008b). However, their preservation and composition are affected by multiple factors, including burial environment, exposure to heat, and post-depositional chemical transformations such as hydrolysis, oxidation, and condensation (Gülacara et al., 1990; Hamman & Cramp, 2018; Rastogi et al., 2006).

Since the 1970s, the analysis of organic residues in archaeological finds has primarily utilized chromatographic techniques (Condamin et al., 1976; Passi et al., 1981), often combined with mass spectrometry (Heron et al., 1990; Mottram et al., 1999; Regert et al., 2003; Romanus et

[1]*PhD Scholar, Taxila Institute of Asian Civilizations, Quaid-i-Azam University, Islamabad, Pakistan.*
Email: abdulbasit@tiac.qau.edu.pk
[2]*Professor/Director, Taxila Institute of Asian Civilizations, Quaid-i-Azam University, Islamabad, Pakistan.*
Email: ghani@qau.edu.pk

Compliance with ethical standards: There are no conflicts of interest (financial or non-financial). This study did not receive any funding.

al., 2007). These methods are microinvasive, requiring the extraction of residual compounds from small amounts of ground potsherd in contact with the vessel's original contents (Gregg & Slater, 2010; Heron et al., 1990; Regert et al., 2003; Romanus et al., 2007). Gas Chromatography-Mass Spectrometry (GC-MS) is the principal analytical technique used to detect and identify lipid residues in archaeological ceramics. It facilitates the characterization of a diverse range of molecular biomarkers, including saturated and unsaturated fatty acids, sterols, terpenes, and benzoic acid derivatives (Bonaduce et al., 2017; Colombini et al., 2005; Colonese et al., 2015; Evershed et al., 2008a).

A common challenge in archaeology is organizing and interpreting multivariate datasets, where each record is linked to various quantitative and qualitative variables, such as provenance, size, typology, and composition for ceramics, or altitude, features, chronology, and material classes for sites (Cardarelli & Lapadula, 2022). Similarly, chemical data derived from archaeological pottery—such as the distribution of various fatty acids—are often high-dimensional and exhibit complex inter-correlations among variables. This complexity makes interpretation challenging when using traditional univariate or visual analytical methods. To address this, multivariate statistical techniques such as Principal Component Analysis (PCA) have been widely adopted in archaeometry (Baxter, 1994). PCA reduces the dimensionality of datasets containing multiple interrelated variables—such as the distribution of lipid compounds—by extracting the most significant components of variation. This approach helps to reveal underlying compositional patterns and grouping tendencies among archaeological samples.

In this context, the present study employs an integrative analytical approach, combining Gas Chromatography-Mass Spectrometry (GC-MS) and Principal Component Analysis (PCA), to investigate organic residues preserved in pottery from the Buddhist Monastic Complex of Badalpur, dating back to the 3rd-4th centuries. GC-MS enables the identification of lipid compounds absorbed into or associated with the ceramic matrices, while PCA allows for the visualization and interpretation of complex fatty acid profiles. This combined methodology not only facilitates the reconstruction of ancient culinary and storage practices but also offers broader insights into the dietary behaviours and subsistence strategies of the monastic community. Through this approach, the study aims to contribute meaningfully to the growing field of organic residue analysis in South Asian archaeology. The study addresses two main questions: first, whether GC-MS and PCA can be applied to Badalpur pottery, and second, what the identified compounds reveal about the dietary practices and food resources of the monastic inhabitants.

Study Site

The Buddhist Monastic Complex of Badalpur is situated in the culturally rich Taxila Valley of Pakistan. Strategically located at the crossroads of ancient trade and cultural exchange, the Taxila Valley has a history dating back to the 6th century BCE during the Achaemenid period (Marshall, 1960). From the 3rd century BCE onwards, numerous Buddhist structures, including monasteries and stupas, were constructed, with Badalpur being one of the significant sites. The Badalpur archaeological site is a significant Buddhist monastic complex in the Taxila Valley, located at 35°46'56" N and 72°52'09" E, at an elevation of 534 meters (see Figure 1**)**. Covering 2.9 acres near the Haro River, it lies 10 km northeast of the Taxila Museum and 2.5 km northwest of Julian. Excavations revealed eight cultural layers above natural soil, reflecting phases of construction, habitation, and resettlement. These layers span four major building periods from the 2nd to the 8th Century CE, with earlier traces possibly dating back to the 1st Century BCE. Artifacts such as pottery, coins, and burnt deposits, supported by carbon dating, indicate long-term occupation and provide valuable insights into the region’s cultural chronology (Khan et al., 2013, 2014). Archaeological excavations at

Badalpur have yielded a diverse and well-preserved assemblage of ceramics—including oil lamps, cooking pots, storage jars, bowls, plates, and other vessel forms— dating back to the 3rd-4th centuries CE (Aiyar, 1917; Arif et al., 2006, 2011; Khan et al., 2007, 2013).

Materials and Methods

Sample Collection and Classification

63 pottery samples from the Badalpur site underwent GC-MS analysis and were categorized into five groups: oil lamps, cooking pots, storage pots, bowls, and dish plates. Each sample was uniquely labeled. Surface-extracted samples were marked with "S" (e.g., ORA-22-S), while internal core samples had an "I" (e.g., ORA-22-I). When multiple samples were taken from the same pottery portion, they were labeled with a suffix, such as "A" for the first sample and "B" for the second (e.g., ORA-D-IA and ORA-D-IB). Samples were collected from various pottery parts, with the majority coming from the body and surface samples taken from residue areas (see Tables 2-6 and Appendix).

Sample Preparation and Extraction Protocol

Samples were prepared using a rotary tool (DREMEL Model 2050) with a disposable sandpaper head. Each pottery fragment underwent a two-stage process: the surface layer was first removed to create a primary fraction, followed by extraction of the underlying layer as a secondary fraction. Small amounts were then pulverized to a fine powder with a pestle and mortar. All work was done under a fume hood to maintain cleanliness and contain dust. A new sandpaper roller was used for each sample to avoid cross-contamination. The pestle and mortar were washed with water and dried, then cleaned twice each with methanol (MeOH), dichloromethane (DCM), and n-hexane between samples. Fume hood surfaces were cleaned with water after each sample and with DCM at the start and end of each day. All vials and tubes used were similarly washed with MeOH and DCM.

Up to 2 g of powdered sample per fraction was stored in metal foil. When possible, surface and core layers were analyzed separately; however, many samples with visible residues were combined and analyzed as single fractions. Before extraction, 100 µL of internal standard (5α-androstane, 50 ng/µL) was added. Each sample was treated with 7 mL of DCM: MeOH (2:1), sealed, and sonicated for 10 minutes. The solution was centrifuged at 1500 rpm for 5 minutes, and the extract was collected. This process was repeated twice, yielding ~20 mL of extract, which was reduced to ~2 mL using nitrogen gas at 25 °C. The extract was split into vials A and B. Both were dried, and vial B received 2 mL of 5% NaOH in MeOH, then heated at 70 °C for 60 minutes. After saponification, 550 µL of 6 M HCl was added to neutralize the solution. Vial B was extracted five times with 1 mL n-hexane, with each upper phase combined into vial A. Vial A was then dried under nitrogen.

Six procedural blanks underwent the same process. For derivatization, 100 µL each of DCM and BSTFA were added to vial A, gently mixed, and transferred to pre-cleaned GC vials with 200 µL inserts. These were sealed and heated at 70 °C for 60 minutes, then left to incubate for 24 hours before GC-MS analysis.

GC-MS Analytical Conditions

GC-MS analysis was performed using an Agilent 7890A GC coupled with a 5975A MSD single-quadrupole mass spectrometer and autosampler, operated in full scan mode. A 1 µL sample was injected in splitless mode at 300 °C. The oven program started at 50 °C (5 min hold), ramped at 18 °C/min to 315 °C (20 min hold), followed by a 5 min post-run at 315 °C. The MSD source and quadrupole were set to 230 °C and 150 °C, respectively, with a scan range of m/z 50–550.

Chromatograms were analyzed using MSD ChemStation E.02.00.493. Compound identification was based on DRS_Demo, NIST05, and W8N08 libraries, with a match quality threshold of 90%. Peak areas were determined via manual integration in ChemStation.

Results

Total Ion Current (TIC) Chromatograms of the Samples

The TIC chromatograms obtained from the GC-MS analysis of Badalpur pottery samples display the separation of compounds based on their chemical properties. The x-axis represents retention time—how long each compound takes to elute from the column—while the y-axis reflects signal intensity, indicating compound abundance. These chromatograms provide a clear visual summary of the chemical composition, with each peak corresponding to a specific compound. As shown in Figure 2, distinct peaks appear between 11 and 23 minutes, with larger peaks representing higher concentrations. These patterns help in identifying and comparing organic residues preserved in the pottery.

Blank Samples

Blank samples play a crucial role in GC-MS analysis, helping to assess background contamination and ensure data accuracy. These blanks, which contain no analytes, were used to evaluate sample cleanliness and identify potential contaminants. Figure 3 displays a typical blank chromatogram, where the internal standard, androstane, is clearly visible. This peak aids in aligning retention times and normalizing signal intensity across chromatograms.

Two consistent peaks observed in all blanks (Table 1), at 18.5–18.7 min and 19.5–19.7 min, correspond to hexadecanoic acid trimethylsilyl ester and octadecanoic acid trimethylsilyl ester, respectively, based on spectral library matches. Because these compounds also appear in archaeological samples, their detection is statistically evaluated by normalizing peak areas to androstane and applying a significance threshold (2 standard deviations above the blank mean). Threshold values were calculated as follows: 0.04714 (hexanoic), 0.817163 (hexadecanoic), and 0.505557 (octadecanoic) esters. Statistical results (Tables 2–6) show that multiple pottery types—oil lamps, cooking pots, storage vessels, bowls, and dish plates—exhibited significantly higher levels of hexadecanoic and octadecanoic esters compared to blanks, in both surface and interior layers.

Fatty Acid Profiles across Pottery Types

In total, 38 distinct organic compounds, primarily fatty acids, were identified across all analyzed pottery samples from Badalpur, with each match demonstrating a spectral similarity of 90% or greater.

The TIC chromatograms of oil lamp samples (Figure 2 and 4) revealed a rich profile of fatty acids, including nonanoic (pelargonic), decanoic (capric), dodecanoic (lauric), tridecanoic (tridecylic), tetradecanoic (myristic), trans-9-hexadecenoic (palmitelaidic), hexadecanoic (palmitic), trans-9-octadecenoic (elaidic), octadecanoic (stearic), pentadecanoic (pentadecylic), heptadecanoic (margaric), and cis-9-octadecenoic (oleic) acids. Notably, samples ORA-28 and ORA-30 showed the highest concentrations of these compounds.

Cooking pot samples also exhibited a diverse array of fatty acids (Figure 5), both saturated and unsaturated. Saturated fatty acids detected include nonanoic, decanoic, dodecanoic, tridecanoic, tetradecanoic, pentadecanoic, hexadecanoic, heptadecanoic, octadecanoic, eicosanoic (arachidic), docosanoic (behenic), tricosanoic (tricosylic), and hexacosanoic acids. Unsaturated fatty acids identified include trans-9-hexadecenoic, trans-9-octadecenoic, cis-9-octadecenoic, and cis-6-octadecenoic acids. Additionally, compounds such as benzoic acid, benzophenone, and sterols like cholesterol and β-sitosterol were detected. Sample ORA-33

exhibited similar chemical profiles in both rim and body fragments, though the rim fragment contained a broader variety of compounds.

Storage pot samples were found to contain various saturated fatty acids (Figure 6), including octanoic (caprylic), nonanoic, decanoic, dodecanoic, pentadecanoic, heptadecanoic, hexadecanoic, octadecanoic, and docosanoic acids. Unsaturated acids present include trans-9-hexadecenoic, trans-9-octadecenoic, trans-12-octadecenoic, cis-9-octadecenoic, and cis-6-octadecenoic acids. Other notable compounds include benzoic acid and dehydroabietic acid.

The bowl samples exhibited a range of compounds categorized into three main groups (Figure 7). The first group, saturated fatty acids, included nonanoic, dodecanoic, tetradecanoic, hexadecanoic, octadecanoic, eicosanoic, pentadecanoic, heptadecanoic, and docosanoic acids. The second group, unsaturated fatty acids, included trans-9-octadecenoic (elaidic) and cis-9-octadecenoic (oleic) acids.

Finally, the TIC chromatograms of dish plate samples revealed a wide variety of compounds (Figure 8). Saturated fatty acids included nonanoic, decanoic, dodecanoic, tridecanoic, tetradecanoic, hexadecanoic, octadecanoic, eicosanoic, pentadecanoic, heptadecanoic, docosanoic, and tetracosanoic (lignoceric) acids. Unsaturated fatty acids included trans-9-hexadecenoic, trans-9-octadecenoic, cis-9-octadecenoic, 11-cis-octadecenoic (vaccenic), 12-hydroxy-cis-9-octadecanoic (ricinoleic), cis-6-octadecenoic, trans-12-octadecenoic, and 11-trans-octadecenoic acids. Other identified compounds included benzoic acid, acetic acid, cholesterol, and dehydroabietic acid, a diterpenoid compound.

Contaminants and Quality Control

During the identification of compounds using mass spectral libraries, some chromatograms showed peaks linked to inorganic contaminants such as phthalates from damaged vacuum seals and siloxanes from septum bleed. Additional contamination from hydrocarbon residues was also observed, affecting the chromatographic baseline (Figure 9). These compounds often produce mass spectra with evenly spaced m/z values, leading to multiple high-confidence but misleading library matches. Such issues can result from poor GC inlet maintenance or improper handling during ion source cleaning. Re-cleaning the ion source using lint-free nylon gloves is recommended. Nitriles, though absent in blanks, may have originated from nitrile gloves used during handling. Human metabolites like Androst-5,16-diene-3β-ol could also reflect contamination during excavation or lab procedures.

Principal Component Analysis (PCA)

Statistical data analysis on fatty acids distribution in pottery samples was carried out using multivariate techniques, such as principal component analysis (PCA). PCA has been widely used in archaeometry (see for example Baxter, 1994) and it is able to reduce the dimension of a dataset of samples (called observations) described by many inter-correlated dependent variables (i.e. the concentrations of chemical compounds). The aim of PCA is to extract the most significant information from the data and reduce the size of the dataset.

The dataset is structured in a matrix made of n row (n = number of samples) and m columns (m=number of chemical compounds) as schematized in Table 7, indicating the sample identification label (S) and the variable (V). Although not mandatory (see Chatfield, 1980), the dataset was standardized by subtracting the mean and dividing by the standard deviation. In this way, the dataset in centred and the compounds having higher concentration do not primarily contribute to the variance. PCA calculates new variables known as principal components (PC), derived as linear combinations of the original variables. The initial principal component is designed to possess the maximum variance possible, representing a significant portion of the information (inertia) contained in the data table. The second component is determined with the constraint of being orthogonal to the first component and

maximizing its own inertia. Subsequent components follow a similar computation process. The values assigned to these novel variables for each observation are termed factor scores. Geometrically, these factor scores can be interpreted as the projections of the observations onto the principal components. The correlation between a component and a variable reflects the information they have in common. Within the PCA context, this correlation is referred to as a loading. Loadings indicate the 'importance' of each variable in the construction of each principal component.

Generally, PCA provides a graph called bi-plot where scores are plotted in the new axis (component) together with the loading plot represented by arrows centred in the origin and pointing to their loading value of each component. In Figure 10, a hypothetical biplot obtained with 7 samples (S, objects) originally described by 5 variables (V) is reported. Figure 10 indicates that principal component 1 (PC1) that is generated by a linear combination of the 5 variables contains the 50% of the total variance, which means that about a half of the total information is contained in PC1. PC2 contains the 30% of the total variance. Therefore, Figure 10 represents with only 2 components the overall 75% of the total variance, which is generally accepted as an excellent result. Clearly, the 20% of the missing information is contained in the remaining 3 components. Figure 10 also show that V_1 and V_2 are correlated, while they are negatively correlated to V_4. This means that in samples where compounds V_1 and V_2 are abundant, V_4 is scarcely present. On the contrary, in samples where V_4 is abundant, V_1 and V_2 are scarcely present. It occurs for samples from S_3 to S_7. Indeed, while V_1 and V_2 are higher and V_4 is lower in samples from S_5 to S_7, V_1 and V_2 are lower and V_4 is higher in samples S_3 and S_4. The variables V_1, V_2 and V_4 are the main responsible of the distinction between samples S_3-S_4 from samples S_5-S_7.

Samples S_1 and S_2 are not characterized by V_1, V_2 and V_4. Nevertheless, they have lower values of V_3. Moreover, V_3 is orthogonal to V_1, V_2 and V_4, thus indicating that V_3 is not correlated with the other variables. V_3 is independent to the other variables, but it is informative of samples S_1 and S_2. Finally, PCA can also be used to remove redundant information. As evident from Figure 10, loadings of V_5 are small (short arrow) indicating that V_5 is not useful to describe this system, and it could be removed or not discussed.

Principal Component Analysis of Badalpur Pottery Collection

Principal component analysis was carried out in 49 archaeological samples (including 19 cooking pots, 8 storage pots, 6 bowls, 6 oil lamps and 11 dish plates) where fatty acids content was previously analysed through GC/MS. Although a total of 12 fatty acids were analysed, since Palmitelaidic, Elaidic and Pentadecyl acids were found only in a few samples, they were not considered. Therefore, statistical analysis was performed including: Pelargonic, Capric, Lauric, Myristic, Palmitic, Stearic, Margaric, Oleic and Petroselenic acids, thus representing a total of 9 variables.

Principal component analysis carried out with 49 samples and 9 variables (fatty acids) provided the biplot reported in Figure 11. As shown, two components (PC1 and PC2) are sufficient to represent the 52% of the total variance. It is also evident that some acids, such as Lauric, Oleic, Myristic, Pelargonic and Capric acids are correlated, as also demonstrated by the loadings reported in Table 8. The correlation between these acids may be an indicator of the same source of fatty acids. They were particularly abundant in dish plates and oil lamps, while storage and cooking pots have a scarce content of these acids.

Conversely, Palmitic, Petroselenic, Margaric and Stearic acids correlates but they are independent on the fatty acids that are significant in PC1. In particular, Petroselenic and Margaric acids correlate, likely indicating a similar source. PC2 mainly differentiate dish plates from oil lamps, where dish plates have a higher content in Petroselenic, Margaric, Stearic and Palmitic (or, oil lamps have a lower content of these fatty acids). Moreover, some

cooking pots seems to be enriched in Palmitic, Petroselenic and Margaric acids, although this is not a characteristic peculiar of cooking pots. Moreover, storage pot samples have a low variability as they are poorly dispersed.

Aiming to testing the possible differences between ceramic categories (i.e. cooking pot, storage pot, dish plates, bowls and oil lamp), a Hotelling T^2-test was performed. The Hotelling T^2-test is the 2D extension of the monodimensional Student t-test, generally used to test the significance of the differences between two populations. The results of the pairwise Hotelling T^2-test is reported in Table 9. As shown, oil lamp samples significantly differ from all the other categories because they have a higher content of oleic, capric, lauric, pelargonic and myristic acids, while they are depleted in margaric, palmitic and petroselenic acids.

Discussion

The GC-MS analysis of pottery samples from the Badalpur archaeological site led to the identification of 38 distinct organic compounds, primarily consisting of saturated and unsaturated fatty acids. These compounds were identified with high spectral similarity scores (≥90%), and their profiles are consistent with those commonly reported in archaeological ceramics (Evershed et al., 2008a, 2008b; Hammann et al., 2018; Kałużna-Czaplińska et al., 2016; Stear, 2008). Among the most frequently detected saturated fatty acids were dodecanoic acid (C12:0, lauric acid), tetradecanoic acid (C14:0, myristic acid), pentadecanoic acid (C15:0), hexadecanoic acid (C16:0, palmitic acid), heptadecanoic acid (C17:0, margaric acid), and octadecanoic acid (C18:0, stearic acid). Unsaturated fatty acids such as cis-9-hexadecenoic acid (C16:1, palmitoleic acid), cis-9-octadecenoic acid (C18:1, oleic acid), and cis-9, cis-12-octadecadienoic acid (C18:2, linoleic acid) were also identified, further supporting the presence of residues from both animal and plant sources (Kałużna-Czaplińska et al., 2016). The presence of these fatty acids provides valuable insights into past dietary habits, vessel function, and the use of organic materials in daily activities. Their detection in cooking pots, storage vessels, oil lamps, bowls, and dish plates suggests varied usage of these vessels, including food preparation, storage, and lighting. Notably, palmitic acid, stearic acid, and myristic acid—common in both animal fats and plant oils—were prominent across several vessel types, supporting interpretations of mixed resource use (Dunne et al., 2020; Irto et al., 2022). A particularly noteworthy finding is the presence of medium-chain fatty acids, including capric acid (C10:0), lauric acid (C12:0), and myristic acid (C14:0), which were especially prevalent in oil lamps and dish plates. These fatty acids are typically associated with substances such as coconut oil, palm oil, and certain dairy products (Irto et al., 2022; Sauter et al., 2003). Their significant concentration in oil lamps points to their likely use as fuel, indicating an informed selection of lighting materials by the inhabitants of the Badalpur site.

Principal Component Analysis (PCA) further clarified the functional distinctions between pottery types based on their lipid profiles. A notable cluster of fatty acids—lauric, oleic, myristic, pelargonic, and capric acids—was strongly associated with oil lamps and dish plates, suggesting a shared origin or function. In contrast, cooking and storage vessels showed higher concentrations of palmitic, stearic, margaric, and petroselenic acids, likely reflecting their roles in food preparation and preservation.

Collectively, the lipid residue evidence supports a nuanced understanding of vessel function in the Badalpur settlement. The results point to deliberate selection and use of animal and plant-based substances tailored to the needs of cooking, storage, and illumination, highlighting the complex culinary and domestic practices of the site's inhabitants.

Conclusion

In conclusion, this study has successfully demonstrated the effectiveness of the applied methodologies—particularly the use of GC-MS and PCA—in extracting, identifying, and characterizing organic residues from the Badalpur pottery collection. The GC-MS analysis revealed a wide range of fatty acids, some indicative of plant-based sources and others of animal-derived products, suggesting a diverse diet that likely included dairy. The PCA results showed that oil lamp samples were chemically distinct from other vessel types, highlighting variations in use and content. These findings provide meaningful insight into the dietary habits of the monastic community at Badalpur. However, due to the limited sample size, broader conclusions about the region's food practices cannot yet be drawn. Future research should expand the analysis to pottery from other archaeological sites in the Taxila Valley to gain a more complete understanding of historical foodways and cultural behaviors in the region.

References

- Aiyar, N. (1917). *Frontier Circle for 1916-17: Annual Report of the Archaeological Survey of India*. Peshawar.
- Arif, M., Khan, A., & Ibrahim, M. (2011). Excavation at the Buddhist site of Badalpur, Taxila Valley (campaign 2005). *Gandharan Studies*, 5, 27–42.
- Arif, M., Khattak, M. H., & Khan, M. H. (2006). Fascinating discoveries from Buddhist Sanctuary of Badapur, District Haripur, Taxila Valley. *Ancient Pakistan*, *17*, 119-126.
- Baxter, M. J. (1994). *Exploratory multivariate analysis in archaeology*. ISD.
- Bonaduce, I., Ribechini, E., Modugno, F., & Colombini, M. P. (2017). Analytical approaches based on gas chromatography mass spectrometry (GC/MS) to study organic materials in artworks and archaeological objects. In R. Mazzeo (Ed.), *Analytical Chemistry for Cultural Heritage*. Topics in Current Chemistry Collections. Springer, Cham.
- Briuer, F. I. (1976). New clue to stone tool function: plant and animal residues. *American Antiquity*, *41*(4), 478-484.
- Cardarelli, L., & Lapadula, A. (2022). Dimensionality reduction for data visualization and exploratory analysis of ceramic assemblages. *Archeologia e Calcolatori*, *33*(2).
- Chatfield, C. (1980). *Introduction to multivariate analysis*. 1st ed. Routledge.
- Colombini, M. P., Modugno, F., & Ribechini, E. (2005). Organic mass spectrometry in archaeology: Evidence for Brassicaceae seed oil in Egyptian ceramic lamps. *Journal of Mass Spectrometry*, *40*(7), 890-898.
- Colonese, A. C., Farrell, T., Lucquin, A., Firth, D., Charlton, S., Robson, H. K., Alexander, M., & Craig, O. E. (2015). Archaeological bone lipids as palaeodietary markers. *Rapid Communications in Mass Spectrometry*, *29*(7), 611-618.
- Condamin, J., Formenti, F., Metais, M. O., Michel, M., & Blond, P. (1976). The application of gas chromatography to the tracing of oil in ancient amphorae. *Archaeometry*, *18*(2), 195-201.
- Evershed, R. P. (1993). Biomolecular archaeology and lipids. *World Archaeology*, *25*(1), 74-93.
- Evershed, R. P. (2008a). Organic residue analysis in archaeology: The archaeological biomarker revolution. *Archaeometry*, *50*(6), 895-924.
- Evershed, R. P. (2008b). Experimental approaches to the interpretation of absorbed organic residues in archaeological ceramics. *World Archaeology*, *40*(1), 26-47
- Evershed, R. P., Heron, C., & Goad, L. J. (1990). Analysis of organic residues of archaeological origin by high-temperature gas chromatography and gas chromatography-mass spectrometry. *Analyst*, *115*(10), 1339-1342.
- Evershed, R. P., Heron, C., & Goad, L. J. (1992). The survival of food residues: New methods of analysis, interpretation and application. In A. M. Pollard (Ed.), *New Developments in Archaeological Science*, 429-436.
- Gregg, M. W., & Slater, G. F. (2010). A new method for extraction, isolation and transesterification of free fatty acids from archaeological pottery. *Archaeometry*, *52*(5), 833-854.

- Gülacar, F. O., Susini, A., & Klohn, M. (1990). Preservation and post-mortem transformations of lipids in samples from a 4000-year-old Nubian mummy. *Journal of Archaeological Science*, *17*(6), 691-705.
- Gurr, M. I., Harwood, J. L., & Frayn, K. N. (2002). *Biochemistry: An introduction*. Blackwell Science Ltd.
- Hammann, S. & Cramp, L. J. (2018). Towards the detection of dietary cereal processing through absorbed lipid biomarkers in archaeological pottery. *Journal of Archaeological Science*, *93*, 74-81.
- Hammann, S., Cramp, L. J. E., Whittle, M., & Evershed, R. P. (2018). Cholesterol degradation in archaeological pottery mediated by fired clay and fatty acid pro-oxidants. *Tetrahedron Letters, 59*(50), 4401–4404.
- Heron, C., & Evershed, R. P. (1993). The analysis of organic residues and the study of pottery use. *Journal of Archaeological Method and Theory*, *5*, 247-284.
- Irto, A., Micalizzi, G., Bretti, C., Chiaia, V., Mondello, L., & Cardiano, P. (2022). Lipids in archaeological pottery: A review on their sampling and extraction techniques. *Molecules*, *27*(11), 3451.
- Kałużna-Czaplińska, J., Rosiak, A., Kwapińska, M., & Kwapiński, W. (2016). Different analytical procedures for the study of organic residues in archaeological ceramic samples with the use of gas chromatography-mass spectrometry. *Critical Reviews in Analytical Chemistry*, *46*(1), 67–81.
- Khan, M. A. et al. (2007). A Preliminary Report of the Third Season (2007-08), Excavation on the Buddhist site of Badalpur, Taxila Valley. *Museum Journal*, *National Museum of Pakistan*, 39-52.
- Khan, M. A., Rehman, A. U., Khan, M. M., Khan, M. T., & Khan, I. (2013). Excavation at Badalpur Monastery, District Haripur (Khyber Pakhtunkhwa), Pakistan: A preliminary report of season 2013. *Journal of Asian Civilizations, 36*(2), 65–80.
- Khan, M. A., Arif, S., Butt, A., Pervaiz, A., & Arif, M. (2014). Excavation at Badalpur Monastery, District Haripur (Khyber Pakhtunkhwa), Pakistan: A preliminary report of season 2014. *Journal of Asian Civilizations, 37*(1), 85.
- Marshall, J. (1960). *A guide to Taxila*. Department of Archaeology in Pakistan: Karachi.
- Mottram, H. R., Dudd, S. N., Lawrence, G. J., Stott, A. W., & Evershed, R. P. (1999). New chromatographic, mass spectrometric and stable isotope approaches to the classification of degraded animal fats preserved in archaeological pottery. *Journal of Chromatography A*, *833*(2), 209-221.
- Passi, S., Rothschild-Boros, M. C., Fasella, P., Nazzaro-Porro, M., & Whitehouse, D. (1981). An application of high performance liquid chromatography to analysis of lipids in archaeological samples. *Journal of Lipid research*, *22*(5), 778-784.
- Pollard, A. M., & Heron, C. (1996). *Archaeological chemistry*. Washington DC: Royal Society of Chemistry.
- Pollard, A. M., Batt, C. M., Stern, B., & Young, S. M. M. (2007). *Analytical chemistry in archaeology*. Cambridge: Cambridge University Press.
- Rastogi, P., Mathur, B., Rastogi, S., Gupta, V. P., & Gupta, R. (2006). Fatty acid oxidation and other biochemical changes induced by cooking in commonly used Indian fats and oils. *Nutrition & Food Science*, *36*(6), 407-413.
- Regert, M., Garnier, N., Decavallas, O., Cren-Olivé, C., & Rolando, C. (2003). Structural characterization of lipid constituents from natural substances preservedin archaeological environments. *Measurement science and technology*, *14*(9), 1620.
- Romanus, K., Poblome, J., Verbeke, K., Luypaerts, A., Jacobs, P., De Vos, D., & Waelkens, M. (2007). An evaluation of analytical and interpretative methodologies for the extraction and identification of lipids associated with pottery sherds from the site of Sagalassos, Turkey. *Archaeometry*, *49*(4), 729-747.
- Stear, N. A. (2008). *Changing patterns of animal exploitation in the prehistoric Eurasian steppe: An integrated molecular, stable isotopic and archaeological approach* (Doctoral dissertation, University of Bristol).

Appendix

Figure 1: Map showing location of Badalpur in Pakistan

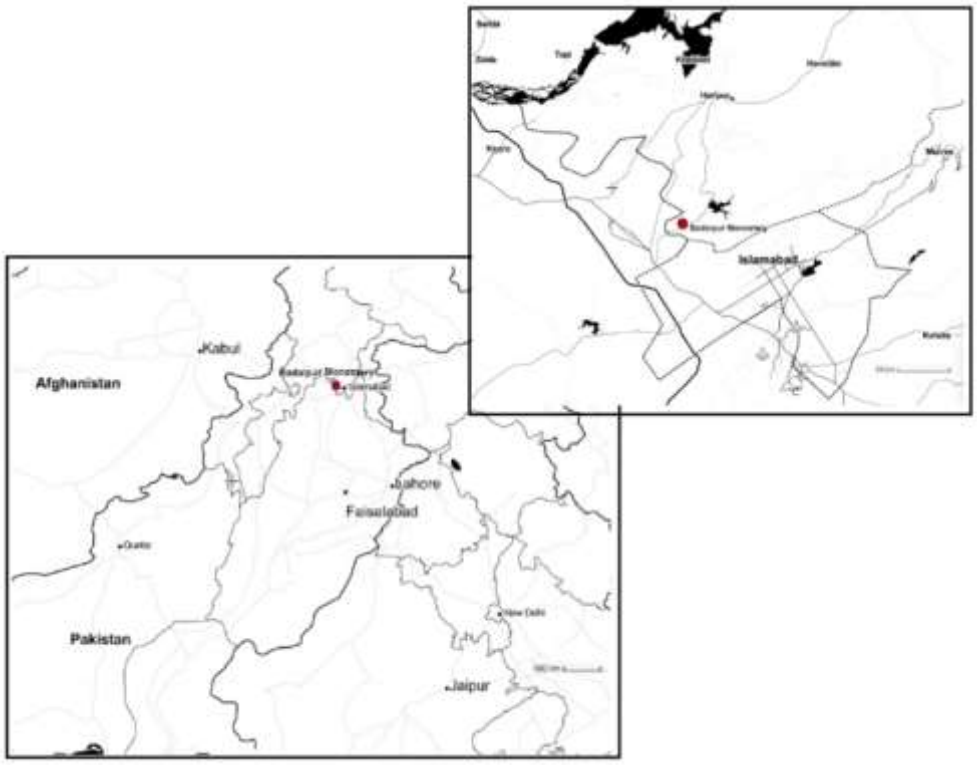


Figure 2: Total Ion current (TIC) chromatogram of an oil lamp sample ORA-C surface residue fraction with their library match

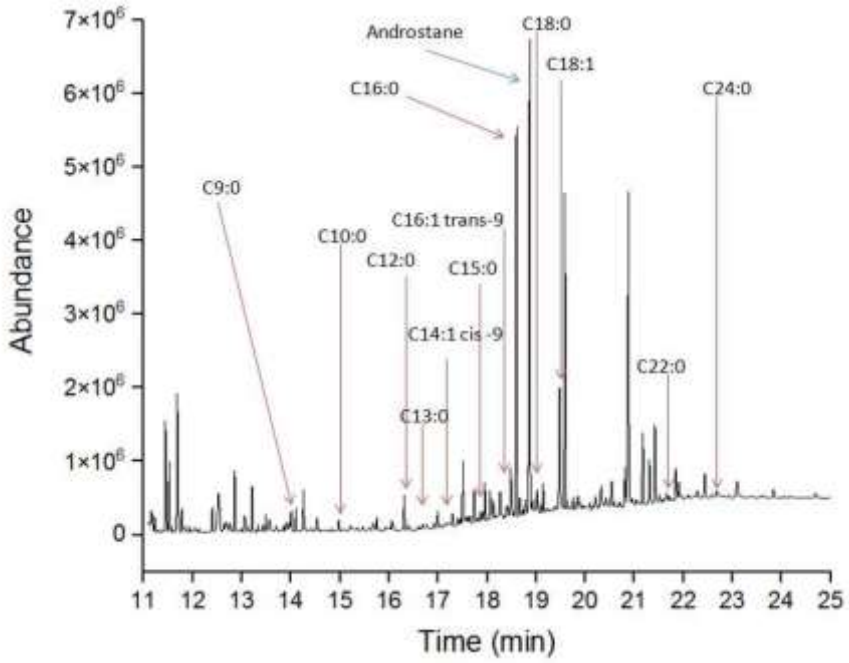


Figure 3: a Total Ion Current (TIC) chromatogram of the blank sample 3.

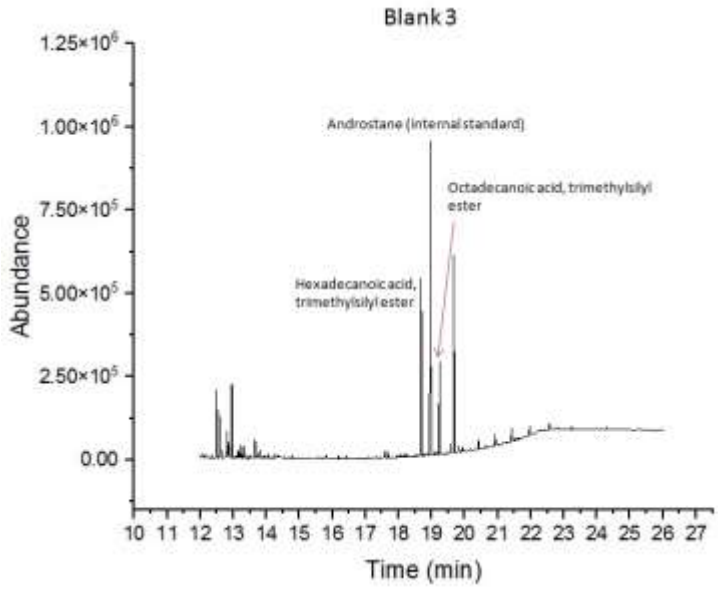


Figure 4: Fatty Acids Identified in Oil Lamp Samples

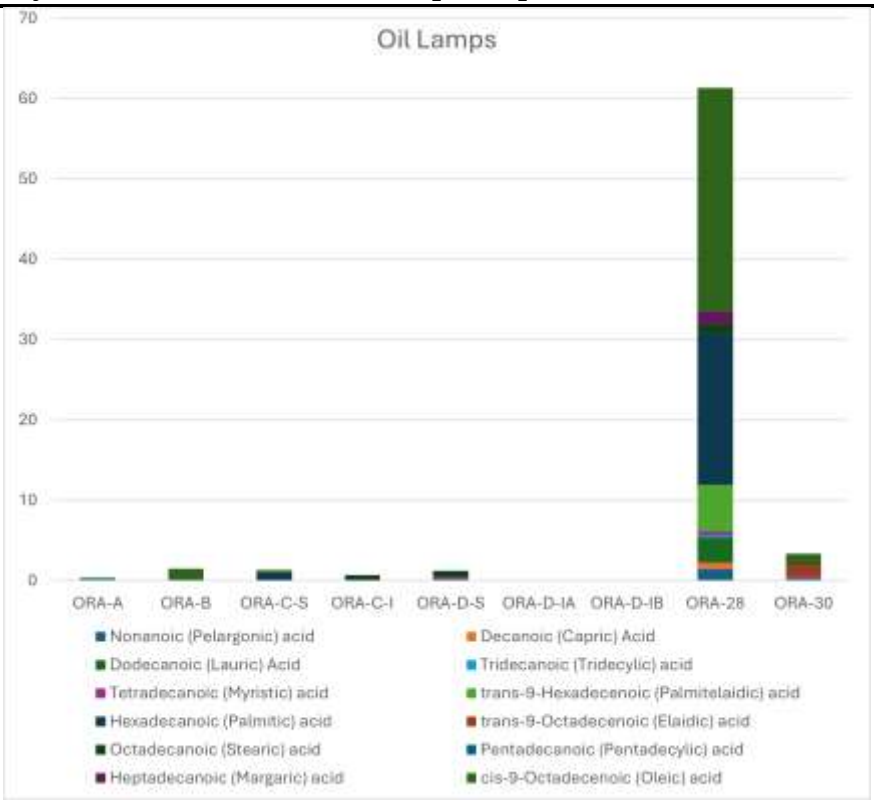


Figure 5: Fatty Acids Identified in Cooking Pot Samples

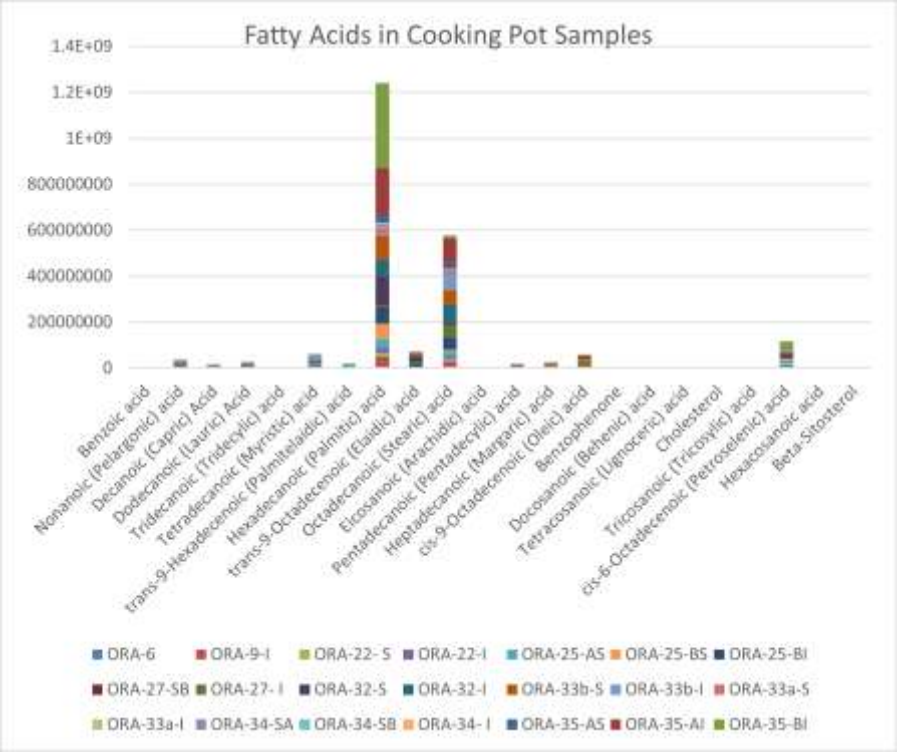


Figure 6: Fatty Acids Identified in Storage Pot Samples

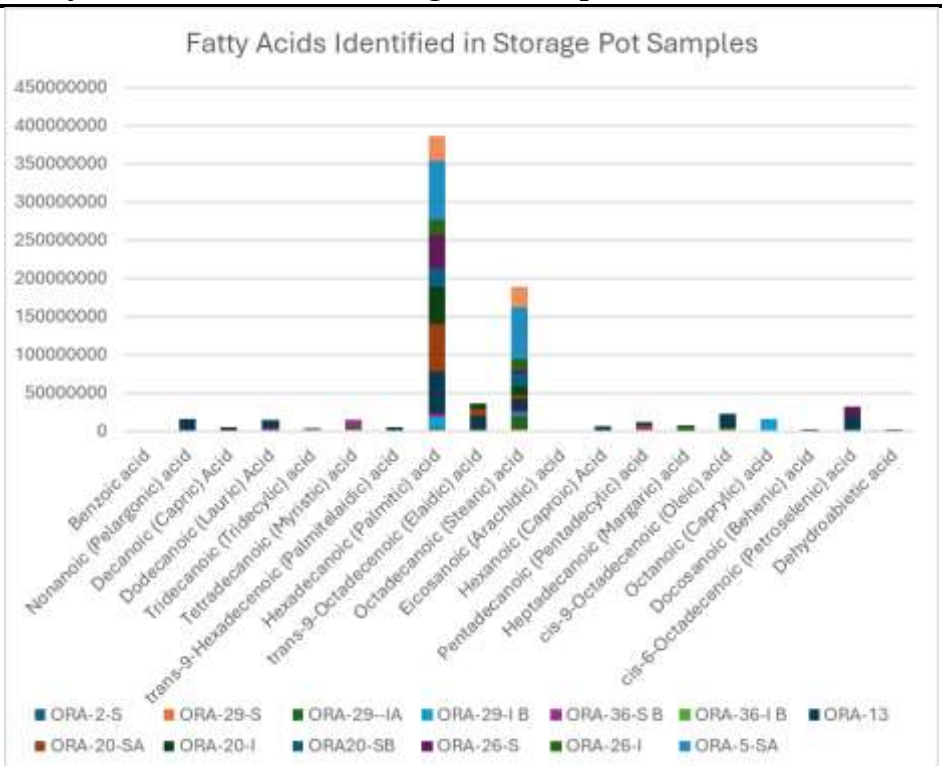


Figure 7: Fatty Acids Identified in Bowl Samples

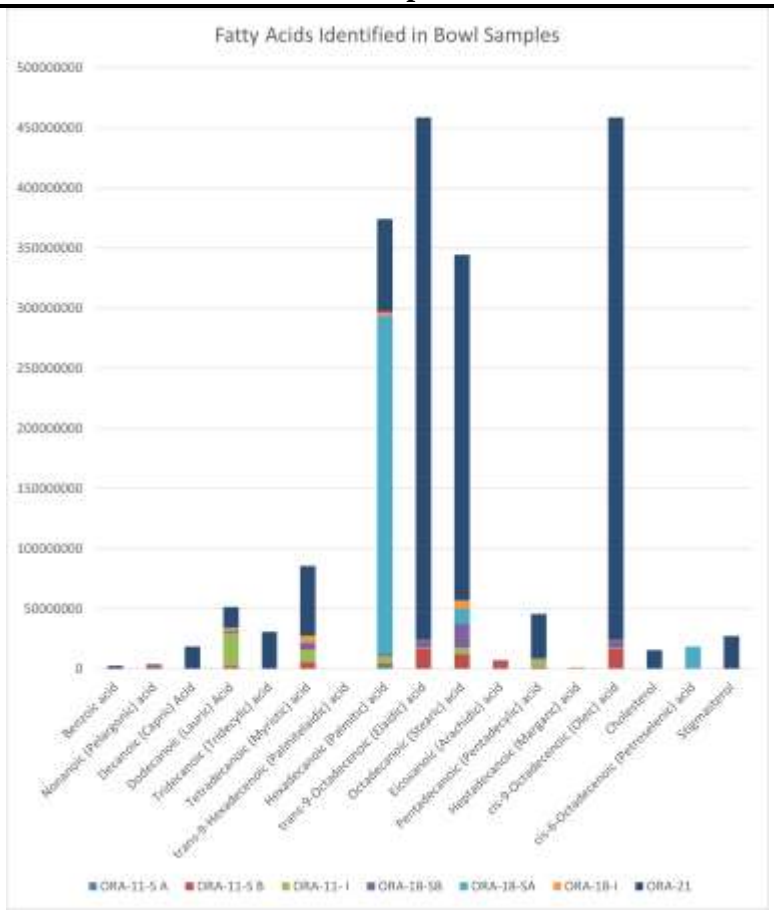


Figure 8: Fatty Acids Identified in Dish Plate Samples

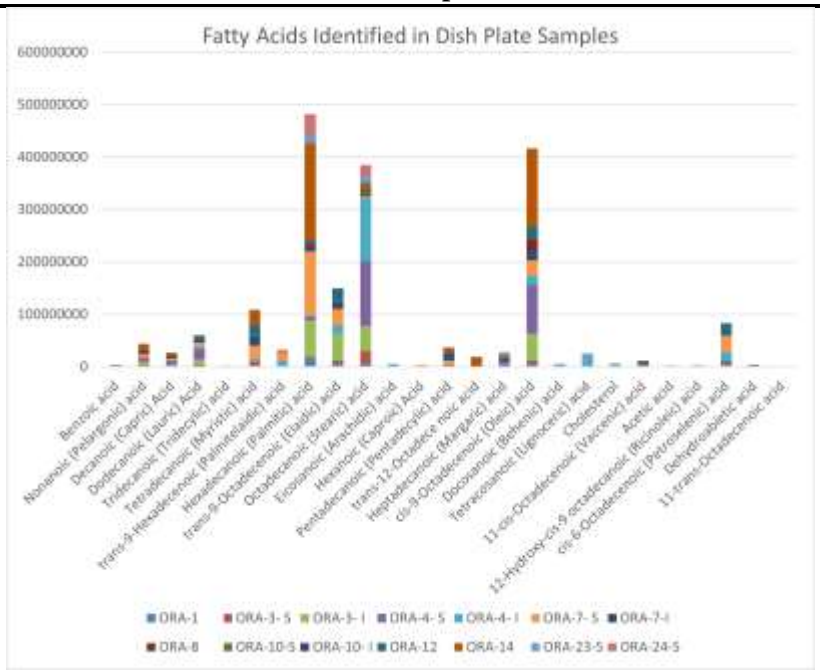


Figure 9: Total Ion Current (TIC) Chromatogram of ORA-1 with Hydrocarbon Peaks

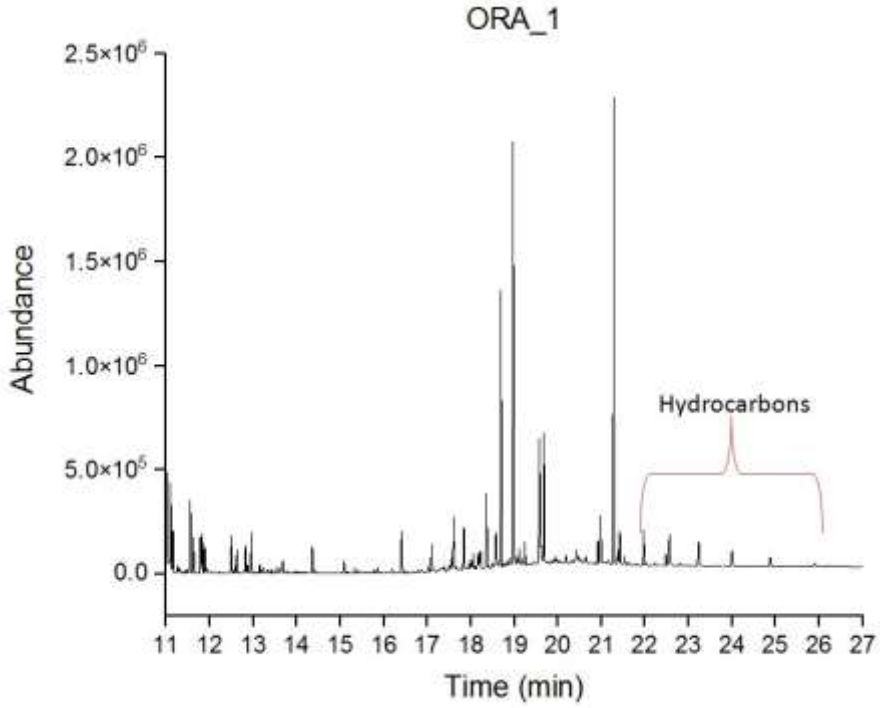


Figure 10: Schematic representation of a biplot from an original dataset containing 7 objects (S) and 5 variables (V)

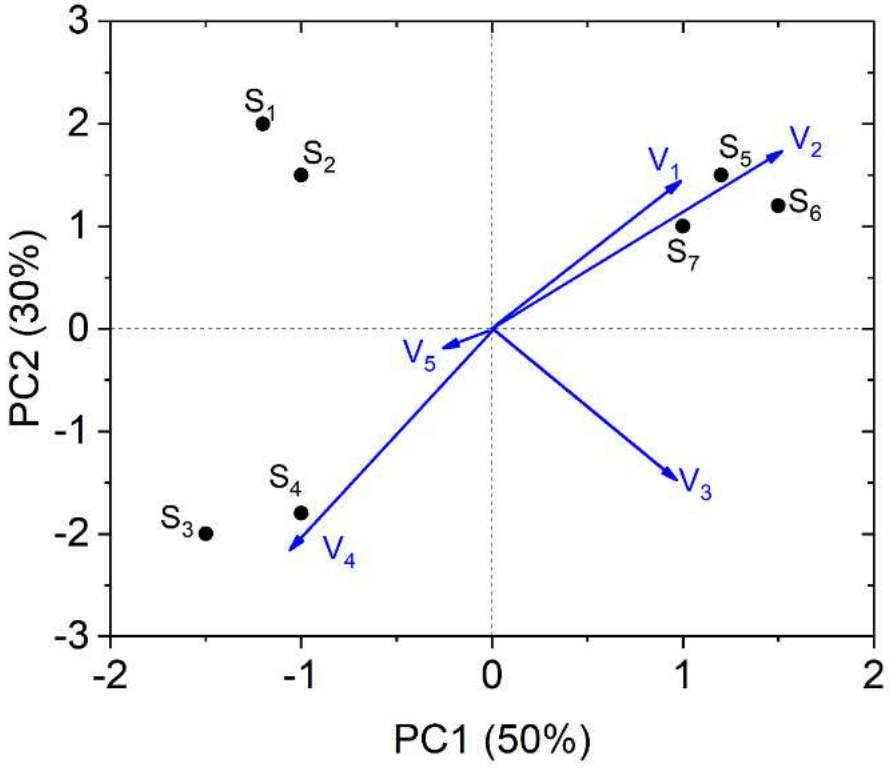
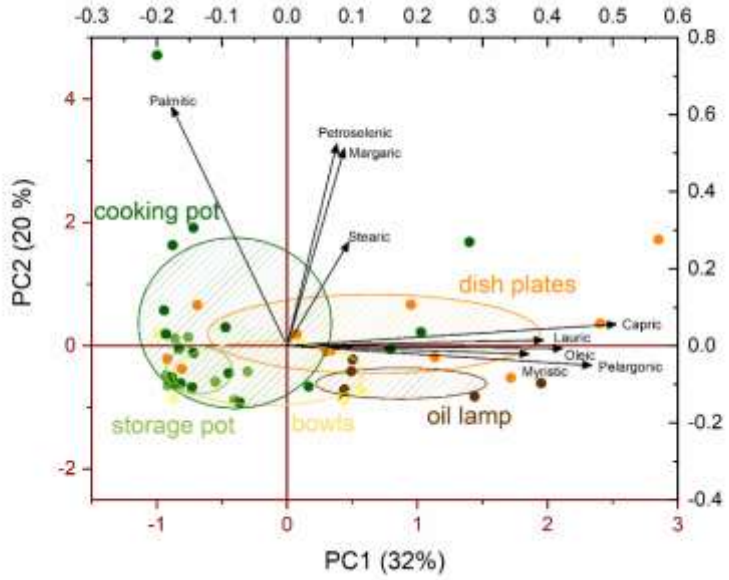


Figure 11: Biplot of the PCA Performed in Archaeological Pottery Samples of Badalpur.



Tables

Table 1: Blank Sample 3

Blank 3					
Peak	Ret. Time	Width	Area	Compound	% Match
1	18.962	0.022	12194173	Androstane	
2	18.685	0.017	5427489	Hexadecanoic acid, trimethylsilyl ester	94
3	19.675	0.017	5885926	Octadecanoic acid, trimethylsilyl ester	98

Table 2: Threshold Values of Compounds in Blank Samples versus Normalized Peak Ratios (Compound:Androstane) in Archaeological Samples. Note: Compounds with peak ratios below the threshold or below the instrument's Limit of Detection were not considered.

Oil Lamps								
Compound/ Sample	**ORA-A**	**ORA-B**	**ORA-C-S**	**ORA-D-S**	**ORA-D-IA**	**ORA-D-IB**	**ORA-28**	**ORA-30**
Androstane	1	1	1	1	1	1	1	1
Hexadecanoic (Palmitic) acid			0.85				18.9	
Octadecanoic (Stearic) acid				0.73			1.09	
Nonanoic (Pelargonic) acid				0.13			1.41	0.11
Decanoic (Capric) Acid				0.04			0.80	
Dodecanoic (Lauric) Acid				0.09			3.12	0.16
Tridecanoic (Tridecylic) acid							0.29	0.04
Tetradecanoic (Myristic) acid				0.19			0.44	0.35
Pentadecanoic (Pentadecylic) acid	0.14	0.10	0.06					
trans-9-Hexadecenoic (Palmitelaidic) acid							5.81	
Heptadecanoic (Margaric) acid							1.62	0.15
trans-9-Octadecenoic (Elaidic) acid								1.24
cis-9-Octadecenoic (Oleic) acid		1.38	0.32	0.18			27.7	1.24

Table 3: Threshold Values of Compounds in Blank Samples versus Normalized Peak Ratios (Compound:Androstane) in Archaeological Samples. Compounds with peak ratios above threshold values are indicated in green.

Cooking Pots								
Compound/ Sample	**ORA -6**	**ORA -9**	**ORA -22- S**	**ORA -22-I**	**ORA-25-AS**	**ORA-25-BS**	**ORA -25-BI**	**ORA-27-SB**
Androstane	1	1	1	1	1	1	1	1
Hexadecanoic (Palmitic) acid						1.02		
Nonanoic (Pelargonic) acid	0.22							
Decanoic (Capric) Acid	0.10							
Dodecanoic (Lauric) Acid	0.17							
Tetradecanoic (Myristic) acid	0.35							
Pentadecanoic (Pentadecylic) acid	0.10							
cis-9-Octadecenoic (Oleic) acid			0.21					
Compound/ Sample	**ORA -27- I**	**ORA -32-S**	**ORA -32-I**	**ORA -33b-S**	**ORA-33b-I**	**ORA-33a-S**	**ORA -33a-I**	**ORA-34-SA**
Androstane	1	1	1	1	1	1	1	1
Hexadecanoic (Palmitic) acid		1.32	2.50	1.40				
Octadecanoic (Stearic) acid	1.17		2.52	0.86	0.52			
Nonanoic (Pelargonic) acid				0.13				
Decanoic (Capric) Acid				0.05				
Dodecanoic (Lauric) Acid	0.09							
Tetradecanoic (Myristic) acid	0.23							
Pentadecanoic (Pentadecylic) acid				0.09				
Heptadecanoic (Margaric) acid	0.03			0.06				
trans-9-Octadecenoic (Elaidic) acid	0.25	0.25	0.17	0.14				
cis-9-Octadecenoic (Oleic) acid	0.25			0.14				
Compound/ Sample	**ORA -34-SB**	**ORA -34-I**	**ORA -35-AS**	**ORA -35-AI**	**ORA-35-BI**			
Androstane	1	1	1	1	1			
Hexadecanoic (Palmitic) acid				2.3	4.56			
Octadecanoic (Stearic) acid				1.17				
Nonanoic (Pelargonic) acid	0.11							
Tridecanoic (Tridecylic) acid	0.07							
Pentadecanoic (Pentadecylic) acid	0.07							
Heptadecanoic (Margaric) acid				0.06	0.09			
cis-9-Octadecenoic (Oleic) acid				0.19				

Table 3: Threshold Values of Compounds in Blank Samples versus Normalized Peak Ratios (Compound:Androstane) in Archaeological Samples. Compounds with peak ratios above threshold values are indicated in green.

Storage Pots							
Compound/ Sample	**ORA-2-S**	**ORA-29-S**	**ORA-29—IA**	**ORA-29-I B**	**ORA-36-S B**	**ORA-36-I B**	**ORA-13**
Androstane	1	1	1	1	1	1	NOT FOUND
Hexadecanoic (Palmitic) acid				0.95			
Tridecanoic (Tridecylic) acid		0.09					
Pentadecanoic (Pentadecylic) acid		0.09			0.05		
cis-9-Octadecenoic (Oleic) acid		0.15					

Compound/ Sample	ORA-20-SA	ORA-20-I	ORA20-SB	ORA-26-S	ORA-26-I	ORA-5-SA	ORA-5-SB
Androstane	1	1	1	1	1	1	1
Octadecanoic (Stearic) acid						0.55	
Heptadecanoic (Margaric) acid					0.05		

Table 4: Threshold Values of Compounds in Blank Samples versus Normalized Peak Ratios (Compound:Androstane) in Archaeological Samples. Compounds with peak ratios above threshold values are indicated in green.

Bowls							
Compound/ Sample	**ORA-11-S A**	**ORA-11-S B**	**ORA-11- I**	**ORA-18-SB**	**ORA-18-SA**	**ORA-18-I**	**ORA-21**
Androstane	1	1	1	1	1	1	NOT FOUND
Hexadecanoic (Palmitic) acid					1.31		
Octadecanoic (Stearic) acid		0.6		0.57			
Nonanoic (Pelargonic) acid		0.08					
Dodecanoic (Lauric) Acid		0.11	0.22				
Tetradecanoic (Myristic) acid		0.27				0.19	
Pentadecanoic (Pentadecylic) acid		0.08					
Heptadecanoic (Margaric) acid						0.05	
trans-9-Octadecenoic (Elaidic) acid		0.82		0.19			
cis-9-Octadecenoic (Oleic) acid		0.82		0.19			

Table 5: Threshold Values of Compounds in Blank Samples versus Normalized Peak Ratios (Compound:Androstane) in Archaeological Samples. Compounds with peak ratios above threshold values are indicated in green.

Dish Plates							
Compound/Sample	**ORA-1**	**ORA-3-S**	**ORA-3-I**	**ORA-4-S**	**ORA-4- I**	**ORA-7-S**	**ORA-7-I**
Androstane	1	1	1	1	1	1	1
Hexadecanoic (Palmitic) acid			0.83			2.5	
Octadecanoic (Stearic) acid			0.57	1.08	2.8		
Nonanoic (Pelargonic) acid						0.14	
Decanoic (Capric) Acid				0.06		0.06	
Dodecanoic (Lauric) Acid	0.09		0.1	0.17	0.1	0.16	
Tetradecanoic (Myristic) acid						0.51	0.22
Pentadecanoic (Pentadecylic) acid						0.15	0.08
Heptadecanoic (Margaric) acid				0.04	0.05		0.05
trans-9-Octadecenoic (Elaidic) acid	0.23		0.61		0.39	0.62	0.19
cis-9-Octadecenoic (Oleic) acid	0.23		0.61	0.81	0.39	0.62	0.19
Compound/Sample	**ORA-8**	**ORA-10-S**	**ORA-10- I**	**ORA-12**	**ORA-14**	**ORA-23-S**	**ORA-24-S**
Androstane	1	1	1	1	1	1	1

Hexadecanoic (Palmitic) acid					11		0.86
Octadecanoic (Stearic) acid		0.84			0.63		
Nonanoic (Pelargonic) acid	0.11				0.63		
Decanoic (Capric) Acid	0.08			0.05	0.2		
Dodecanoic (Lauric) Acid				0.14			
Tridecanoic (Tridecylic) acid					0.09		
Tetradecanoic (Myristic) acid				0.4	1.74		
Pentadecanoic (Pentadecylic) acid			0.076	0.07	0.49		
Heptadecanoic (Margaric) acid			0.03	0.05	0.18	0.09	
trans-9-Octadecenoic (Elaidic) acid				0.47			
cis-9-Octadecenoic (Oleic) acid	0.51			0.47	8.77		

Table 6: Scheme of the Dataset for PCA

	Variables						
Observations	V_1	V_2	V_3	V_4	…	…	V_m
S_1	$x_{1,1}$	$x_{1,2}$	$x_{1,3}$	…	…	…	$x_{1,m}$
S_2	$x_{2,1}$	$x_{2,2}$	…	…	…	…	…
…	…						
S_n	$x_{n,1}$						

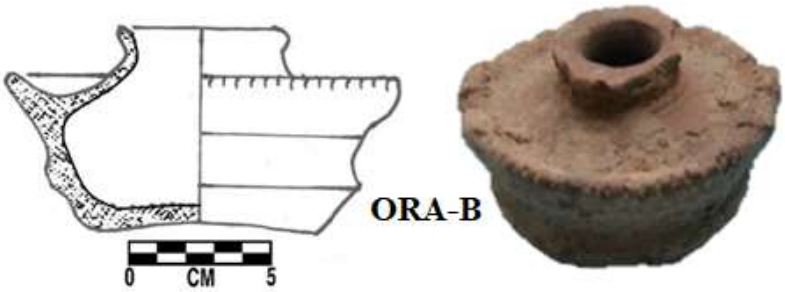
Table 7: Loadings values of the principal components

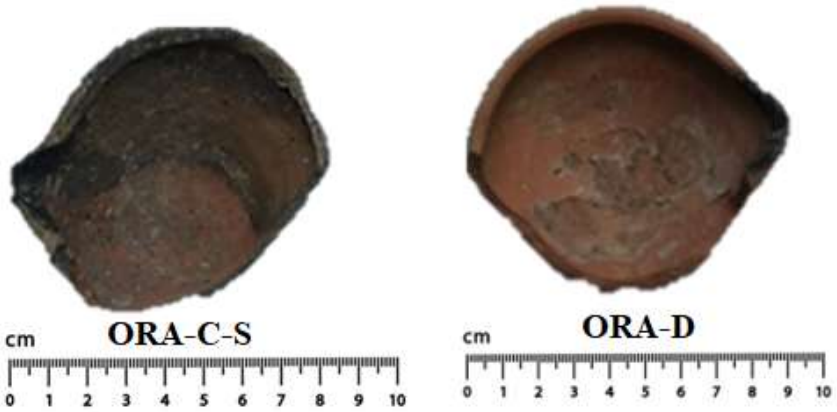
	PC1 coefficient	PC2 coefficient
Pelargonic	**0.46789**	-0.05097
Capric	**0.50532**	0.0559
Lauric	**0.42301**	-0.00698
Myristic	**0.37158**	-0.02151
Palmitic	-0.17705	**0.61955**
Stearic	0.09448	0.2689
Margaric	0.08776	**0.51242**
Oleic	**0.39355**	0.01453
Petroselenic	0.07658	**0.52424**

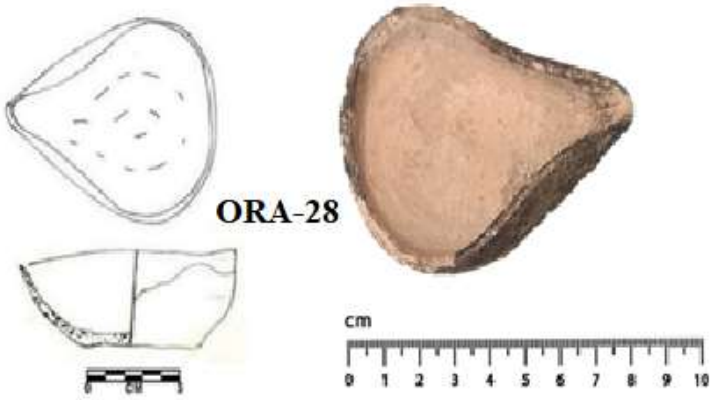
Table 8: Results of the pairwise Hotelling T2-test. The symbol = indicate that there's not any statistical differences. *, ** and * indicate a significance level of a = 0.1, 0.05 and 0.01, respectively)**

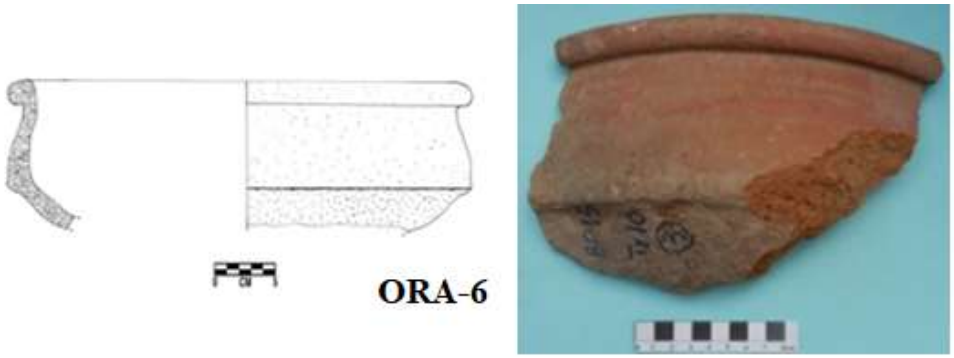
	Cooking Pot	Dish Plate	Oil lamp	Storage Pot
Bowls	=	=	*	=
Cooking Pot		**	***	=
Dish plates			**	**
Oil Lamp				***

Appendix: Some Representative Pottery Samples from the Badalpur Site



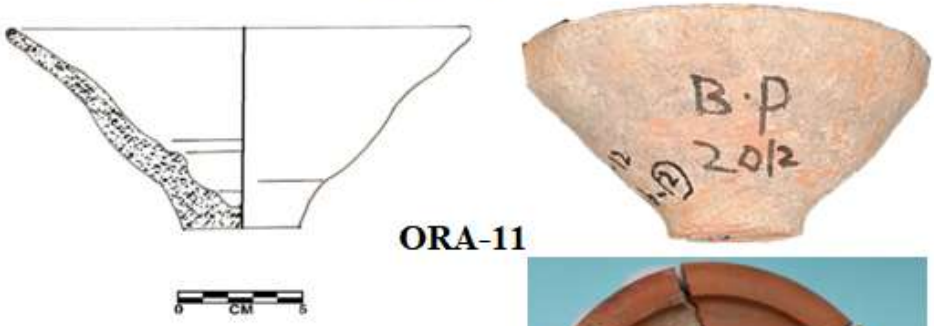





ORA-6


B·P
2011
ORA-27


B·P 2012
ORA-13


B·P
2012
ORA-11


ORA-1


ORA-4