A Proteomic Analysis of Biological Sex and Health in Gurat, France

by

Cassidy G. Sellers

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Master of Arts
in
Public Issues Anthropology

Waterloo, Ontario, Canada, 2025

© Cassidy G. Sellers 2025

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Paleoproteomics, the study of ancient proteins, uses mass spectrometry to identify and characterize proteins by their amino acid sequences. This thesis explores the potential of paleoproteomics to inform bioarchaeological interpretations of biological sex and health in a small medieval population from Gurat, France. Ten milligram enamel samples from six individuals excavated from a rock-cut cave church were analyzed using liquid chromatography-tandem mass spectrometry (LC-MS/MS) to identify proteins in enamel and evaluate their interpretive potential. The primary aim of this project was to test a modified method to investigate its ability to successfully identify amelogenin (amelogenin-X and amelogenin-Y) and immune proteins (C-reactive protein and immunoglobulin-gamma). This method was intentionally modified to reduce analytical costs and resource requirements, while remaining applicable to very small quantities of dental enamel, thereby increasing its feasibility for archaeological and ethically sensitive sampling contexts.

The enamel-specific protein amelogenin was successfully identified in all six samples, allowing for proteomic estimations of biological sex to be made. In contrast, one non-enamel-specific immune protein (C-reactive protein) was identified in only one sample, reflecting uncertainty regarding the abundance and preservation of immune proteins in enamel, the immune histories of the Gurat individuals, and the sensitivity of the modified method for immune protein detection.

These results highlight both the strengths and limitations of paleoproteomics, offering avenues of exploration in future directions. Above all, this thesis finds that proteomic analyses can complement osteological analyses to offer valuable insight into archaeological populations.

Acknowledgements

This thesis would not have been possible without support from the following individuals. I first owe thanks to Dr. Alexis Dolphin, my supervisor and advisor, for her continual encouragement and advice, most of all during an unanticipated pivot in my research. I also owe deep thanks to Dr. Yiming Xiao, lab manager of the Mass Spectrometry Facility at the University of Waterloo, for not only allowing me access to his lab, but for his willingness to innovate and his unparalleled enthusiasm for this project. Additional thanks go to my committee members, Dr. Dolphin, Dr. Xiao, and Dr. Robert Stark for their appreciated feedback and insights.

To the entire Department of Anthropology, thank you for cultivating the invigorating academic environment that provided the initial spark for this investigation. Within this faculty, special thanks to Dr. Dolphin, Dr. Seçil Dağtas, and Dr. Adrienne Lo for their especially close observation of, and crucial suggestions regarding, my research. Thanks also go to the other members of my cohort, who formed a team of supportive, insightful, and encouraging voices throughout a process that was not without challenges.

I extend my deepest gratitude to the family of Iris Yuzdepski, the Social Sciences and Humanities Research Council of Canada, the Department of Anthropology, the Faculty of Arts, and the University of Waterloo, whose generous financial support facilitated the transformation of my research ideas into a reality.

I also owe thanks to Dr. Michael Gervers for his work in Gurat, and for granting me access to the collection.

I have been fortunate to lean on the support of my family and friends throughout the duration of this project. Thank you to my parents, brother, and grandparents for cheering me on from the shores of Nova Scotia. Additionally, special thanks go to Arjun Ravi, my partner and essential support; Maggie Lareau, my most trusted consult; and Macy Evans, Julie Brown, and Lauren Weir, whose presence throughout any undertaking is as indispensable to me as the shoes on my feet.

Lastly, I extend my gratitude towards the Gurat individuals themselves, who were the focus of my research. Thank you for allowing me to share parts of your stories.

Table of Contents

Author's Declaration ..... **ii**

Abstract ..... **iii**

Acknowledgements ..... **iv**

List of Figures ..... **ix**

List of Tables ..... **x**

List of Abbreviations ..... **xi**

Chapter 1: Introduction **1**

1.1 Research purpose and scope 1

1.2 Ethical and anthropological approaches to proteomic research 4

1.3 Ethics and responsibility in biomolecular research 6

Chapter 2: Background **11**

2.1 Applications of paleoproteomics 11

2.1.1 Applications of proteomics for biological sex 11

2.1.2 Applications of proteomics for health 13

2.2. Human enamel development: amelogenesis 19

2.3 Diagenesis 23

2.4 Social and archaeological context of Gurat, France 25

2.4.1 Introduction 25

2.4.2 The Gurat individuals.....30

2.4.2.1 Skeletal pathology of the Gurat individuals.....32

2.4.2.1.1 Paleopathology and observations of Gu 10.....34

2.4.2.1.2 Paleopathology and observations of the adult Gurat individuals.....35

Chapter 3: Materials and Methods 37

3.1 Sample selection 37

3.1.1 Diagenesis.....37

3.2 Sample collection and preparation 39

3.3 Mass spectrometry and data-dependent acquisition (DDA) 40

3.3.1 Overview of liquid chromatography-mass spectrometry-based proteomics.....40

3.3.2 Peptide ionization, MS1/MS2 acquisition, and identification.....41

3.3.3 Liquid chromatography-tandem mass spectrometry and acquisition parameters.....42

3.4 Peptide identification, analysis, and label-free quantification (LFQ) 44

Chapter 4: Results..... 47

4.2 C-reactive protein..... 54

4.3 Immunoglobulin-gamma 54

4.4 Albumin..... 55

Chapter 5: Discussion 58

5.1 Biological sex of the Gurat individuals 58

5.2 Immune proteins and health of the Gurat individuals 60

5.3 Limitations and future directions..... 63

Chapter 6: Conclusion **66**

References **68**

List of Figures

Figure 1. Location of Gurat in France. Image borrowed from Google Earth, 2025. 26

Figure 2. Location of St. Georges within Gurat. Image borrowed from Google Earth, 2025. 27

Figure 3. Exterior view of St. Georges from the North. Image borrowed from Navetat, 2019. .. 28

Figure 4. Exterior view of St. Georges from the South. Image borrowed from Navetat, 2019. .. 28

Figure 5. Interior view of St. Georges. Image borrowed from Navetat, 2019. 29

Figure 6. Diagram of Q Exactive Orbitrap used. Image borrowed from Thermo Fisher Scientific. 43

Figure 7. Peptides identified in Gu 2. Green peptides represent peptides specific to AMELY, where grey peptides represent the overlap with AMELX. Image provided by Dr. Yiming Xiao, Mass Spectrometry Facility, University of Waterloo (Waterloo, Ontario, Canada). 48

Figure 8. Peptide annotation for SMIRPPY (AMELY), provided by Dr. Yiming Xiao, Mass Spectrometry Facility, University of Waterloo (Waterloo, Ontario, Canada). Taller peaks represent higher intensities detected in fragmented ions. 50

Figure 9. Peptide annotation for MIRPPY (AMELY), demonstrating the oxidation of amino acid M, reflecting chemical alteration. Image provided by Dr. Yiming Xiao, Mass Spectrometry Facility, University of Waterloo (Waterloo, Ontario, Canada). Taller peaks represent higher intensities detected in fragmented ions. 51

Figure 10. Log10-transformed peptide intensities of identified AMELX and AMELY in each sample. 53

Figure 11. Log10-transformed peptide intensities of identified AMELX, AMELY, Albumin, CRP, and IgG in each sample 57

List of Tables

Table 1. List of individuals studied for this thesis, combined data from Meijer, 2018, and Seymour, 2019..... 32

Table 2. Instrument parameters provided by Dr. Yiming Xiao, Mass Spectrometry Facility, University of Waterloo (Waterloo, Ontario, Canada). 44

Table 3. FragPipe search parameters, provided by Dr. Yiming Xiao, Mass Spectrometry Facility, University of Waterloo (Waterloo, Ontario, Canada). 46

Table 4. Selected amelogenin peptide identification in Gurat individuals. 49

Table 5. Comparison of osteological and proteomic sex estimations for the Gurat individuals analyzed in this study. Log-transformed peptide intensities for amelogenin-X (AMELX) and amelogenin-Y (AMELY) are shown for each individual, alongside previously published osteological sex estimates (Meijer, 2018; Seymour, 2019) and proteomic sex estimations derived from the presence or absence of sex-specific amelogenin peptides. A value of 0 indicates that no peptides for that amelogenin group were detected. Proteomic sex estimation demonstrates clearer differentiation for males through the detection of AMELY peptides, while female estimations rely on the absence of AMELY and remain comparable in confidence to osteological assessments. . 52

Table 6. Log-transformed data of CRP protein present in the Gurat individuals. 54

Table 7. Selected serum albumin peptide sequences identified in Gurat individuals..... 56

Table 8. Log-transformed data for selected serum albumin peptide presence in Gurat individuals. 56

List of Abbreviations

ABC — Ammonium bicarbonate
ACN — Acetonitrile
AGC — Automatic gain control
AMELX — Amelogenin X
AMELY — Amelogenin Y
CE — Common Era
CI/mg — Counts intensity per milligram
CRP — C-reactive protein
DDA — Data-dependent acquisition
DEJ — Dentine–enamel junction
DNA — Deoxyribonucleic acid
DTT — Dithiothreitol
FA — Formic acid
FDR — False discovery rate
HCD — Higher-energy collisional dissociation
IAA — Indole-3-acetic acid
IgA — Immunoglobulin A
IgG — Immunoglobulin G
IgM — Immunoglobulin M
LC-MS — Liquid chromatography–mass spectrometry
LC-MS/MS — Liquid chromatography–tandem mass spectrometry
LFQ — Label-free quantification
m/z — Mass-to-charge ratio
MBR — Match between runs
mg — Milligram
MS — Mass spectrometry
MS1/MS2 — First- and second-stage mass spectrometry scans
MS/MS — Tandem mass spectrometry
µL — Microlitre
n.d. — No date
ppm — Parts per million
RNA — Ribonucleic acid
TIC — Total ion current
MaxLFQ — Maximal label-free quantification

Chapter 1: Introduction

1.1 Research purpose and scope

Proteomics, the study of proteins, uses mass spectrometry to identify and characterize proteins by their amino acid sequences (Hendy et al., 2019). Amino acids link together through peptide bonds to form peptides, which assemble and fold into complex, three-dimensional structures that constitute functional proteins (Hendy et al., 2018; Hendy et al., 2019; Warinner et al., 2022). Paleoproteomics extends these methods to the study of ancient proteins, applied to archaeological populations. While paleoproteomics is a relatively new discipline in anthropology, the survival and preservation of ancient proteins have been studied for decades, since the discovery of amino acids in fossils in 1954 (Abelson, 1954; Hendy et al., 2019).

The recent draw toward paleoproteomics in the anthropological discipline is attributed to proteins maintaining an exceptionally high level of preservation. Because proteins are generally more "robust" than DNA and RNA, owing to their greater molecular stability and frequent incorporation into mineralized tissues such as bone and dental enamel, they can survive in human remains long after nucleic acids have degraded, allowing for molecular analyses in contexts where ancient DNA recovery is not possible (Hendy et al., 2018; Hendy et al., 2019; Warinner et al., 2022). This resilience makes proteins an invaluable biomolecular resource for reconstructing aspects of the past, thereby expanding the scope of molecular anthropology beyond the limits of DNA-based research (Hendy et al., 2018; Hendy et al., 2019; Madupe et al., 2025; Warinner et al., 2022).

As the possibilities of what proteomics may share about ancient and past populations continue to be explored, it is critical that these methods are tested thoroughly. This thesis provides a proteomic analysis of teeth from an archaeological population in the hopes of assessing what proteomics may uncover in the case of a small group of individuals from Gurat, France.

Like proteins themselves, teeth and their tissues, such as enamel, exhibit remarkable preservation in archaeological contexts, providing an important means for anthropologists to better understand ancient populations and how they may have lived (Gil-Bona & Bidlack, 2020; Hillson, 2024). Research involving teeth has allowed us to better reconstruct aspects of past lives such as diet, age-at-death, health, or, more recently, biological sex (Bartlett, 2013; Castiblanco et al., 2015; Charlton et al., 2019; Gil-Bona & Bidlack, 2020; Hillson, 2024; Margolis et al., 2006). Through the application of proteomics using dental enamel, an emergent area of study within bioarchaeology, biological sex can be determined based on the presence of certain proteins that make up the structural foundations of enamel in as little as fifteen minutes (Ziganshin et al., 2020)

Buonasera et al., 2024, argue that proteomic sex estimations may represent only one application of enamel proteomics. In their study, they examine the presence and relative abundance of serum-derived immune proteins immunoglobulin-gamma (IgG) and C-reactive protein (CRP) in dental enamel to explore population-level differences in health and physiological stress. Analyzing enamel from Mission-Period ancestral Ohlone individuals and comparing these results with historic-period European-American and contemporary samples, they report elevated IgG levels among individuals interred at the mission site, as well as the

presence of CRP in a subset of archaeological samples (Buonasera et al., 2024). These patterns are cautiously interpreted as potentially reflecting higher burdens of disease and physiological stress within the Mission population, while emphasizing that such interpretations remain provisional (Buonasera et al., 2024). Importantly, their study demonstrates that immune-related proteins can be detected in archaeological enamel and suggests that enamel proteomics may provide insight into immune responses occurring during tooth formation (Buonasera et al., 2024). While further research is required to clarify the exact mechanisms of protein incorporation and preservation, this work establishes a foundation for investigating immune-related markers in enamel beyond sex estimation.

In this light, this thesis research adapts and slightly modifies existing proteomic methods previously used on bone and teeth to apply them to smaller samples of dental enamel and explore the boundaries of proteomic analysis. These modifications were implemented to account for differences in sample size and tissue type, while also reducing analytical costs, relying on readily available laboratory materials, and improving practical feasibility. Using mainly permanent teeth excavated from the early medieval church of St. Georges, Gurat, France, from individuals with previously produced osteobiographies, this project aims to identify proteins associated with biological sex and immune system activity during childhood. By observing specific protein presence with osteological indicators of biological sex and health, this research tests the interpretive opportunities of current proteomic methods. Proteins of particular interest include amelogenin, immunoglobulin-gamma, and C-reactive protein, which may shed light on the biological sex and childhood health of the Gurat individuals. Furthermore, this study simultaneously evaluates the potential of proteomic approaches for sex determination in

comparison to cranial and/or pelvic sex estimation strategies, offering valuable insights for both bioarchaeological and forensic applications.

1.2 Ethical and anthropological approaches to proteomic research

The importance of an anthropological approach to paleoproteomic research is that research remains people-focused. When investigating quantitative data as they relate to a population, biological anthropologists must put people at the forefront, meaning the individuals represented in the archaeological record and relevant descendant communities. It is critical to maintain that researchers are not simply working with data, but with the remains of real people, and that interpretations of molecular findings may have consequences for real lives. As such, anthropological research using molecular methods requires reflexivity, contextual awareness, and ethical responsibility. While the purpose of this thesis is primarily method development, it also hopes to contribute to the continued study of the ancient Gurat population in a manner consistent with the descendant community's expressed interest in seeing research carried out with these remains (A. Dolphin, personal communication, 2025).

A biocultural approach is often invoked in bioarchaeological research to integrate biological and cultural factors in the interpretation of human health and behaviour (Goodman & Leatherman, 1998). In this thesis, however, biocultural theory is not employed as a primary explanatory framework. Rather than using molecular data to generate broad claims about social organization or identity, biocultural perspectives are drawn on selectively to emphasize contextualization, interpretive restraint, and ethical limitation. This approach recognizes that human biology is not static or isolated, but shaped by social, political, economic, and

environmental conditions (Agarwal & Glencross, 2011; Goodman & Leatherman, 1998; Gregoricka, 2023; Schrader & Torres-Rouff, 2021; Zuckerman et al., 2023), while also acknowledging that proteomic data alone cannot meaningfully reconstruct these conditions.

Because proteomic data provide molecular-level information, such as the presence or absence of specific proteins, it can be tempting to link findings directly to broader cultural, social, or environmental explanations. Such interpretations are inherently speculative and risk exceeding what the data can support (Agarwal & Glencross, 2011; Gregoricka, 2023; Schrader & Torres-Rouff, 2021). Commenting beyond the scope of proteomic methods can jeopardize not only the integrity of the research, but also the populations represented and the credibility of the method itself. Large or absolute statements about identity, behaviour, or culture risk misrepresentation and may cause harm to descendant communities as well as to broader publics engaging with anthropological research (Agarwal & Glencross, 2011; Gregoricka, 2023; Martin et al., 2013; Schrader & Torres-Rouff, 2021; Zuckerman et al., 2023).

Accordingly, proteomic evidence in this thesis is treated as one line of evidence among many, interpreted cautiously and in conversation with osteological, archaeological, and historical data when relevant and possible. The goal is not to produce sweeping narratives from molecular findings, but to define their limits, contextual relevance, and interpretive implications. This chapter therefore adopts an ethical anthropological framework that prioritizes transparency, restraint, and respect for the individuals and communities represented by the data, rather than the production of biocultural narratives that extend beyond what proteomic methods can support. In this thesis, these ethical and anthropological principles are operationalized through cautious

interpretation, method modification to minimize unnecessary destruction, and a deliberate avoidance of claims that extend beyond what proteomic data can reasonably support.

1.3 Ethics and responsibility in biomolecular research

Working with human remains requires a heightened level of reflexivity and ethical consideration (Agarwal & Glencross, 2011). In recent years, proteomic research has emerged as a transformative method in biological anthropology, yet the pace of scientific advancement has, at times, outstripped sustained reflection on why this work is undertaken and for whom. At present, there are no formal ethical guidelines specific to proteomic research, nor widely recognized standards governing sampling, analysis, or interpretation (Hendy et al., 2018; Hendy et al., 2019). This gap should be addressed by appropriate professional associations. In the absence of such guidance, it remains essential to remember that the human remains analyzed in paleoproteomic research represent individuals and communities documented in the archaeological record. As researchers, we have a responsibility to uphold respect, transparency, and care, ensuring that ethical considerations are central to both research design and interpretation. These ethical considerations directly informed the methodological decisions made in this project, including sample selection, method modification, the language used to describe destructiveness, and the interpretive boundaries applied to the results.

One of the primary ethical challenges of molecular research, including proteomics, is the risk of reducing human beings to datasets, thereby obscuring the lived realities of the individuals represented. This form of reductionism can emerge subtly through the framing of research questions, the categorization of samples, or the presentation of results. Ethical responsibility in proteomic research therefore requires careful attention to how molecular findings are

contextualized and communicated, and to the potential consequences such interpretations may carry for descendant communities and broader publics.

Ethical practice in anthropological research further involves the adoption of visible, sustainable, and community-centered research habits. These include clear justification for sampling and analysis, thorough pre-sampling documentation, evaluation and minimization of destructive methods, and transparent reporting of results where appropriate (Gregoricka, 2023; Martin et al., 2013). In this thesis, such considerations informed the development of both methods and results. To preserve the integrity of the skeletal collection, measurements and photographs of teeth were taken prior to destructive sampling and retained for future research, and each tooth was selected based on existing osteological analyses to minimize damage to the overall collection.

Proteomic research, while promising, continues to face significant ethical and methodological challenges. Particular care must be taken in the language used to describe analytical techniques. In proteomic literature, methods are frequently described as “minimally invasive” (Brůžek et al., 2024; Fabrizi et al., 2024; Fiddyment et al., 2015; Hansen et al., 2024; Multari et al., 2022; Stewart et al., 2016). However, this terminology is not used here. All research involving human remains is, by nature, invasive. It is therefore more accurate and ethically transparent to describe methods in terms of their destructiveness. In this thesis, “minimally destructive” refers to methods that preserve visible morphology, limit chemical or physical alteration, and retain potential for future analysis, while “destructive” refers to methods that partially or fully destroy samples and preclude their reuse. Destructive methods should

always be evaluated with caution, particularly when working with irreplaceable archaeological materials.

In this project, minimally destructive methods were not feasible due to time and resource constraints. As a result, destructive methods were carefully evaluated prior to use. Although the overall morphology of the teeth was preserved, nearly all enamel was removed through drilling. The resulting enamel powder was collected for proteomic analysis and retained for future research, rather than discarded.

Method development was also undertaken to reduce unnecessary destruction. While enamel samples of approximately 20mg are commonly used in proteomic studies, preliminary experiments conducted for this thesis demonstrated that smaller samples of 15mg and 10mg yielded comparable proteomic results. Samples weighing 6mg and 8mg were also tested but did not produce meaningful data. These experiments are critical when working with archaeological remains, as unnecessary destructive sampling poses a greater ethical concern than destruction itself. Although larger samples often yield richer datasets (Hansen et al., 2024; Multari et al., 2022), they do so at the cost of future analytical potential. Where destructive methods are required, efforts to minimize sample size while maintaining data quality are therefore essential. This project also maintains a clear distinction between biological sex and gender. Biological sex refers here to chromosomal traits identified through molecular or osteological analysis, whereas gender refers to social, cultural, and personal identities (Agarwal & Wesp, 2017; Bruckett & Law, 2018). Determining biological sex is not equivalent to determining gender, and conflating these terms risks reinforcing biologically deterministic interpretations of identity. Such misrepresentations can contribute to exclusionary or harmful narratives and are ethically

inappropriate in proteomic research. Accordingly, this thesis does not compare biological sex with gender and avoids commentary that extends beyond what molecular data can support.

This research contributes to ongoing discussions in biological anthropology by investigating the potential of dental enamel to serve as a bioindicator of health in comparison to existing paleopathological evidence within an early medieval population from Gurat, France. Through the application of proteomics, this project aims to assess whether ancient enamel retains a lasting memory of the body's immune responses during the time of its development throughout childhood (Buonasera et al., 2024). While only one targeted immune protein was detected in this study using the developed method, this result holds important implications for understanding both the preservation limits of enamel and the interpretive boundaries of paleoproteomic methods. The absence of detectable immune proteins, especially when paleopathological indicators are present, suggests not only that there is more to be understood when it comes to protein integration in the enamel matrix, but also that this method requires further optimization and has several limitations to work through. These findings also support the possibility that antibodies may not survive the same way that structural proteins do over prolonged periods of time (Kendall et al., 2016). In any case, interpretations of immune proteins must be made with significant caution, and technological and scientific advances are critical to the continuation of this field. The findings in this thesis highlight the existing limitations and boundaries in immune-related paleoproteomics.

In addition to its focus on health-related markers, this project also assesses the utility of enamel proteomics for sex determination in comparison to osteological methods. By simultaneously incorporating both proteomic and traditional osteological methods, this thesis

maintains that these methods may work best when used together, rather than alone. In the case of subadult remains where osteological methods were not applicable, this thesis could not reliably conduct sex estimation based on the proteomic evidence alone. This also highlights the need for research specifically with subadult remains, particularly those of known biological sex.

In summary, this thesis aims to contribute to the development of proteomic research by testing its boundaries, refining its applications, and contributing to broader discussions about the interpretations of molecular information in anthropological contexts.

Chapter 2: Background

2.1 Applications of paleoproteomics

2.1.1 Applications of proteomics for biological sex

Of the limited proteomic studies published to date, most focus on the structural proteins of parts of human remains, such as enamel-specific proteins amelogenin, ameloblastin, and enamelin, or collagen in dentine and bone (Bartlett, 2013; Castiblanco et al., 2015; Gil-Bona & Bidlack, 2020; Hillson, 2024; Margolis et al., 2006). In particular, amelogenin proteins amelogenin-X and amelogenin-Y, which represent the X and Y chromosomes respectively, have been of much focus to provide proteomic estimates of biological sex (Bartlett, 2013; Castiblanco et al., 2015; Gil-Bona & Bidlack, 2020; Hillson, 2024; Margolis et al., 2006).

Amelogenin is an enamel-specific structural protein critical to the process of enamel formation and mineralization as the most abundant structural protein in enamel (Bartlett, 2013; Brookes et al., 2017; Castiblanco et al., 2015; Duverger & Lee, 2024; Gil-Bona & Bidlack, 2020; Hillson, 2024; Margolis et al., 2006). Though amelogenin-X is present in both males and females, amelogenin-Y is male-sex-specific, indicating that its presence in teeth is representative of male biological sex (Castiblanco et al., 2015; Gil-Bona & Bidlack, 2020; Greco et al., 2023; Jágr et al., 2019; Parker et al., 2019; Procopio et al., 2018; Sharma et al., 2020; Shaw et al., 2023).

The distinction between amelogenin-X (AMELX) and amelogenin-Y (AMELY) arises from key differences in their amino acid sequences, making them distinguishable sex-specific peptide markers detectable by mass spectrometry (Hendy et al., 2018; Hendy et al., 2019; Hendy,

2021). For example, short amino acid sequences such as SMIRPPY and MIRPPY have proven diagnostic for the identification of Y-chromosome-linked amelogenin in proteomic datasets (Parker et al., 2019). The successful identification of these proteins in ancient and contemporary enamel allows the inferences of biological sex directly from protein evidence, independent of morphological or genetic data (Gil-Bona & Bidlack, 2020; Parker et al., 2019). This can be very useful in sex estimation in skeletal remains, especially where the pelvis or cranial bones are not present, or in the case of subadults, where sex estimation is not possible (Castiblanco et al., 2015; Gil-Bona & Bidlack, 2020; Greco et al., 2023; Jágr et al., 2019; Parker et al., 2019; Procopio et al., 2018; Sharma et al., 2020; Shaw et al., 2023). As the field of proteomics continues to grow, so do the opportunities for more routine use of proteomics for sex estimation in forensic cases.

Proteomic sex estimation is especially valuable in contexts where DNA preservation is poor or contamination risk is high (Hendy et al., 2018; Hendy et al., 2019; Hendy, 2021; Warinner et al., 2022). Amelogenin proteins are far more stable than DNA over archaeological timescales, as they are shielded with the highly mineralized structure of enamel, which resists both degradation and environmental alteration (Gil-Bona & Bidlack, 2020; Hendy et al., 2018; Hendy et al., 2019; Hendy, 2021; Warinner et al., 2022). Consequently, proteomic sex estimation provides an alternative for fragmentary skeletal remains, where traditional osteological methods are not always feasible (Castiblanco et al., 2015; Gil-Bona & Bidlack, 2020; Greco et al., 2023; Jágr et al., 2019; Parker et al., 2019; Procopio et al., 2018; Sharma et al., 2020; Shaw et al., 2023).

In recent years, several methodological refinements have improved both the sensitivity and the reproducibility of proteomics (Hendy, 2021). Advances in liquid chromatography tandem

mass spectrometry (LC-MS/MS) and the development of targeted peptide databases have enhanced the ability to distinguish AMELX and AMELY peptides with high confidence (Hendy, 2021; Parker et al., 2019). Additionally, there has been increased interest in minimally destructive methods for biological sex estimation and enamel proteomics (Hansen et al., 2024; Multari et al., 2022), expanding the applicability of proteomic methods to different archaeological populations, including early hominins (Madupe et al., 2025).

Despite its strengths, several challenges remain in proteomic estimations of biological sex. Female sex remains difficult to determine with any confidence, with the absence of AMELY being its only evidence. This is unreliable due to the numerous potential reasons that AMELY could be obscured from results, such as degradation, oxidation, or fragmentation (Hendy, 2021). Statistical analysis has also not yet been refined to determine what kinds of quantities may be indicative of female sex, and this is not something that will be achieved until the needed evolution of proteomic methods—especially those that destroy less enamel. Proteomic sex estimations for females require much further research before they can work reliably on their own, with multiple peptide confirmations and replicated analyses to minimize the risk of false negatives or positives. While proteomic analyses have been increasingly helpful in estimating biological sex, particularly in the cases of males, it is important not to limit their uses to only these questions.

2.1.2 Applications of proteomics for health

Though it is important to study what must be present, expanding into other proteins without structural functions may tell us more about the people we study. Few other studies have examined non-structural proteins in human remains, which is a significant gap in existing

literature (Buonasera et al., 2024; Kendall et al., 2016; Oliveira et al., 2023). Of these studies, discovering the presence of antibodies and immune proteins has been of increasing interest (Buonasera et al., 2024; Kendall et al., 2016; Oliveira et al., 2023). In finding these proteins, one may be able to comment on an individual's health, and, down the line, possibly even the presence of specific diseases (Ansar, 2020; Buonasera et al., 2024; Chiang et al., 2019; Cundiff et al., 2024; Kendall et al., 2016; Oliveira et al., 2023). As it applies to tooth enamel specifically, though this has not been thoroughly confirmed, these diseases would have most likely been present during the process of enamel formation during early childhood.

As enamel matures, it becomes very in low protein content, making the identification of proteins all the more difficult (Duverger & Lee, 2024; Hillson, 2024; Moradian-Oldak & George, 2021). Though it is not known exactly how serum proteins like those related to the immune system become incorporated into human tooth enamel (Buonasera et al., 2024; Gil-Bona & Bidlack, 2020), both C-reactive protein (CRP) and immunoglobulin-gamma (IgG) have been identified (Ansar, 2020; Buonasera et al., 2024; Chiang et al., 2019; Cundiff et al., 2024; Kendall et al., 2016; Oliveira et al., 2023). The presence of these proteins indicates that immune system activity has occurred, but the question of when remains, especially due to the unknown nature of how these proteins have become incorporated (Buonasera et al., 2024; Gil-Bona & Bidlack, 2020). For the sake of this thesis, results are discussed primarily based on the assumption that serum immune proteins become secreted in enamel during its development in childhood through blood, rather than through saliva during an individual's lifetime. While the secretion of proteins into saliva across an individual's lifetime is possible, this process remains largely unexplored and is unlikely to produce concentrations sufficient for consistent detection. In this study, protein

abundance refers to the relative quantity of a given protein detected by mass spectrometry, inferred from peptide ion intensities rather than absolute concentration. Because no universally accepted threshold currently exists to define high- versus low-abundance proteins, abundance is interpreted comparatively across samples within the population examined. At present, successful proteomic detection relies heavily on sufficient protein abundance, which is more likely associated with the persistence of structural proteins within the enamel matrix. As proteomic methods continue to advance, it may become possible to detect less abundant immune proteins, including those potentially introduced through salivary pathways. Further research is therefore required to clarify both the mechanisms and timing of serum protein incorporation into dental enamel.

While paleopathology remains an extremely important subdiscipline in bioanthropological research involving human remains, it is also constrained by both diagnostic and interpretive limitations (Appleby et al., 2015; Brickley & Morgan, 2023; Wei & Cooper, 2023). Morphological indicators of sickness and disease are very uniform, as skeletons can only change in so many different ways (Wei & Cooper, 2023). Thus, finding the presence of specific diseases is very difficult to do, as morphological indicators of one disease often overlap with others (Appleby et al., 2015; Brickley & Morgan, 2023). Furthermore, it is known that the vast majority of illnesses and diseases do not remodel our bones— such as the flu, or a common cold. This limitation is central to the osteological paradox, which highlights that skeletal evidence reflects patterns of survival rather than a direct record of disease prevalence, complicating interpretations of health in past populations (Wood et al., 1992). In this context, immune-related proteins have the potential to provide insight into a broader range of illness experiences that may

not be captured osteologically, particularly if such proteins are incorporated into enamel during childhood development (Wei & Cooper, 2023). Proteomic approaches may therefore complement paleopathological analyses as research continues to clarify when and how immune proteins are incorporated into dental tissues, and how low-abundance proteins can be reliably detected. However, at present, proteomic research remains limited in scope. Current analyses largely focus on C-reactive protein and immunoglobulin-gamma, both of which are only loosely associated with immune response and lack specificity with respect to particular diseases or health experiences (Ansar, 2020; Chiang et al., 2019; Cundiff et al., 2024; VanDuijn et al., 2010). Although typically abundant, their presence provides only coarse insight into immune activity and cannot yet be used to reconstruct detailed or temporally precise health profiles.

It is without question that current literature focused on proteomic analyses of health has been vague as this research continues to develop (Buonasera et al., 2024; Kendall et al., 2016). Even though presence or absence of immune proteins tell us about whether or not inflammation was experienced, their presence or absence does not tell us how, or why, or how long the inflammation was experienced for, nor do we have an idea of what quantities may or may not be significant. Because of this, further studies involving specific immune proteins are needed. As these studies currently stand, CRP and IgG have offered the following insights on people and their health.

CRP has been found in both dentine and enamel, and may provide insight into any number of causes (Ansar, 2020; Buckley et al., 2009; Chiang et al., 2019; Cundiff et al., 2024; Graham-Engeland et al., 2018; Köhler-Forsberg et al., 2017; Marsland et al., 2017). CRP has been most frequently used in living individuals as biomarkers for tissue

injury/infection/inflammation, trauma, and even as a potential predictor for cardiovascular disease (Ansar, 2020; Buckley et al., 2009; Graham-Engeland et al., 2018). It is known that CRP purifies serum to promote healing, as has been associated with individuals suffering from tuberculosis and malaria (Ansar, 2020). Though somewhat controversial, it has also been suggested that their heightened levels can be caused by any kind of significant stress, even outside of physical and bodily health issues (Buonasera et al., 2024; Chiang et al., 2019; Cundiff et al., 2024; Graham-Engeland et al., 2018; Köhler-Forsberg et al., 2017; Marsland et al., 2017; Steptoe et al., 2007; Uchino et al., 2016). Considering this, CRP is reflective of experiences of infection, inflammation, and healing, and maybe even to psychosocial stressors (Ansar, 2020; Cundiff et al., 2024; Graham-Engeland et al., 2018; Köhler-Forsberg et al., 2017; Marsland et al., 2017; Steptoe et al., 2007; Uchino et al., 2016), making it extremely meaningful to understanding health in a bioarchaeological context. However, it must be recognized that CRP is a relatively vague biomarker overall. While its presence may indicate infection, inflammation, healing, and/or stress, this has not been thoroughly tested (Ansar, 2020; Chiang et al., 2019; Cundiff et al., 2024). Levels of CRP in living people are highly variable, and this may also be true of their presence in dental tissues/enamel (Ansar, 2020). While documenting the presence of CRP in proteomic analyses is a relevant result, there is much to be discovered in terms of what their presence may indicate about the lives of those whose teeth have been sampled. It is also important to note that CRP is typically found at much lower concentrations than comparative serum proteins (Buonasera et al., 2024).

Similarly, IgG may be considered just as intriguing. IgG is the most abundant immune protein, and is the second most abundant serum protein after albumin (Buonasera et al., 2024;

Kendall et al., 2016; Oliveira et al., 2023; Shaw et al., 2023). Its exact role and function are not entirely known, but it has been linked with inflammation and immune processes (VanDuijn et al, 2010). Like CRP, IgG indicates that inflammation and healing have occurred, even though there is no exact timeline or etiology known (Ansar, 2020; Buonasera et al., 2024; Chiang et al., 2019; Cundiff et al., 2024; Kendall et al., 2016; Oliveira et al., 2023). IgG is currently the only immunoglobulin emphasized in enamel proteomic literature, despite the existence of several other immunoglobulins that are biologically plausible candidates for incorporation into dental tissues (Buonasera et al., 2024; Kendall et al., 2016; Oliveira et al., 2023). Other immunoglobulins, such as immunoglobulin-A (IgA) or immunoglobulin-M (IgM), play more specific and well-characterized roles in immune system processes, including mucosal immunity and early-stage immune response, which could, in theory, offer more temporally or functionally specific insights into health and disease (VanDuijn et al., 2010). Their absence from current paleoproteomic research is likely not due to biological impossibility, but rather to methodological limitations, including their lower abundance relative to IgG. As a result, IgG remains the primary focus of immunological inquiry in enamel proteomics, not because it is the most informative immune marker, but because it may be the most detectable given existing analytical constraints. As proteomic techniques continue to improve, the identification of additional immunoglobulins may become feasible, opening new avenues for more refined interpretations of immune activity in past populations.

While many questions remain, the investigation of CRP and IgG provides exciting new avenues of research regarding the proteomic analyses of health. As method testing continues to evolve, so should the targeted proteins, expanding into those with known roles and functions in

immune processes, which may offer more specific information. As protein abundance becomes less of a requirement in proteomic analyses, other immune proteins will be important to explore.

2.2. Human enamel development: amelogenesis

The process of amelogenesis, which refers to the formation of enamel, is critical to the understanding of this thesis. During amelogenesis, different proteins are secreted to form the enamel layer of each tooth (Bartlett, 2013; Brookes et al., 2017; Duverger & Lee, 2024; Hillson, 2024; Moradian-Oldak & George, 2021). Most of these are structural proteins, meaning proteins that function specifically to the creation of the structure of enamel (Bartlett, 2013; Brookes et al., 2017; Duverger & Lee, 2024; Hillson, 2024; Moradian-Oldak & George, 2021). Amelogenin is one of these structural proteins, and is a protein of central focus in this research project. The multi-stage process of amelogenesis is outlined in this section to demonstrate the known incorporation of proteins into the enamel matrix, and to emphasize the timing of when these incorporations occur.

During the presecretory stage, the first stage of amelogenesis, odontoblasts begin releasing predentin proteins at the site of the future dentine-enamel junction (DEJ) (Bartlett, 2013; Duverger & Lee, 2024; Hillson, 2024). These proteins are composed primarily of collagen, which later becomes the most dominant structural protein within dentine (Bartlett, 2013; Hillson, 2024). Mineralization of the predentin begins at the future DEJ, first at the cusps and progressing to the cervical regions of the developing crown (Bartlett, 2013; Duverger & Lee, 2024). After this initial mineralization of dentine at the DEJ, ameloblasts (part of the enamel organ) remove and destroy the basal lamina (a basement membrane separating the enamel organ from the odontoblast layer) and start secreting enamel matrix proteins, rapidly initiating mineralization of

the enamel matrix (Bartlett, 2013; Hillson, 2024). This marks the earliest appearance of enamel-specific proteins and the formation of the enamel matrix (Bartlett, 2013; Hillson, 2024).

During the secretory stage of amelogenesis, ameloblasts become fully differentiated and begin secreting large quantities of enamel matrix proteins, or proteins specific to enamel (Bartlett, 2013; Brookes et al., 2017; Duverger & Lee, 2024; Hillson, 2024; Moradian-Oldak & George, 2021). This thickens the enamel layer, and largely involves two amelogenin proteins: amelogenin-X and amelogenin-Y (Bartlett, 2013; Brookes et al., 2017; Duverger & Lee, 2024; Hillson, 2024; Moradian-Oldak & George, 2021). Amelogenin-X exists in greater abundance than amelogenin-Y, due to amelogenin-Y being male-sex-specific (Bartlett, 2013; Brookes et al., 2017; Duverger & Lee, 2024; Hillson, 2024; Moradian-Oldak & George, 2021). At the onset of the secretory stage, ameloblasts elongate and develop Tomes' processes, which project toward the forming of the enamel surface (Bartlett, 2013; Brookes et al., 2017; Duverger & Lee, 2024; Moradian-Oldak & George, 2021). Following the development of the Tomes' processes, ameloblasts begin secreting larger amounts of enamel matrix proteins as they move away from the dentine surface to thicken the enamel layer (Bartlett, 2013; Brookes et al., 2017; Duverger & Lee, 2024; Hillson, 2024; Moradian-Oldak & George, 2021). The main protein secreted in this process is amelogenin (80—90%), though small amounts of ameloblastin, enamelin are also released (Bartlett, 2013; Brookes et al., 2017; Duverger & Lee, 2024; Fincham et al., 1999; Hillson, 2024; Margolis et al., 2006; Moradian-Oldak & George, 2021). Both ameloblastin and enamelin are essential for the formation of enamel crystals (Bartlett, 2013). Conversely, in rare cases where amelogenin is not present, enamel crystals can still abnormally form (Bartlett, 2013). The enamel matrix produced during the secretory stage is protein rich and relatively soft,

being only partially mineralized (Bartlett, 2013; Brookes et al., 2017; Duverger & Lee, 2024; Hillson, 2024; Moradian-Oldak & George, 2021). Though hardening does not take place until the maturation stage, the secretory stage is where the thickness and volume of enamel are established (Bartlett, 2013; Brookes et al., 2017; Duverger & Lee, 2024; Moradian-Oldak & George, 2021; Smith, 1998).

The transition stage marks the brief but critical shift between enamel matrix secretion and full mineral maturation, where ameloblasts undergo extensive structural and functional reorganization (Bartlett, 2013; Duverger & Lee, 2024; Smith, 1998). Secretory ameloblasts shorten and lose their Tomes' processes, signaling the end of the enamel matrix secretion (Bartlett, 2013; Duverger & Lee, 2024). These cells transform into shorter, protein-reabsorbing ameloblasts, which function less as active protein distributors, and more as barriers and transporters that prepare the enamel for its final mineralization phase (Bartlett, 2013; Smith, 1998). Finally, roughly 25% of ameloblasts die during this stage, forcing the surviving ones to shift their functions and express different sets of proteins (Bartlett, 2013; Brookes et al., 2017; Smith, 1998). This allows for the maturation stage to take place.

The maturation stage is the final and longest phase of enamel development, characterized by the removal of residual organic material (proteins), and the completion of enamel mineralization (Bartlett, 2013; Brookes et al., 2017; Duverger & Lee, 2024; Fincham et al., 1999; Hillson, 2024; Margolis et al., 2006; Moradian-Oldak & George, 2021; Smith, 1998). During this stage, nearly all enamel matrix proteins are degraded and reabsorbed, while crystallites expand and fuse to form the highly mineralized, hardened structure of mature enamel (Bartlett, 2013; Brookes et al., 2017; Duverger & Lee, 2024; Fincham et al., 1999; Hillson, 2024; Margolis et al.,

2006; Moradian-Oldak & George, 2021; Smith, 1998). As mineralization progresses, approximately 75% of the remaining ameloblasts undergo programmed cell death, dropping the percentage of proteins by weight significantly, by upwards of 50% (Fincham et al., 1999; Smith, 1998). By the end of this stage, only trace amounts of organic material remain, including remnants of the three primary structural proteins: amelogenin, ameloblastin, and enamelin (Bartlett, 2013; Brookes et al., 2017; Duverger & Lee, 2024; Fincham et al., 1999; Hillson, 2024; Margolis et al., 2006; Smith, 1998). The maturation stage is also the most time-consuming phase of amelogenesis, lasting approximately 960—1187 days (Smith, 1998). This duration reflects the asynchronous nature of enamel maturation, where different teeth mature at varying rates (Hillson, 2024).

Overall, the timing of amelogenesis or enamel formation is somewhat predictable in nature among individuals. In general, enamel first begins to form around the 14th to 16th week of intrauterine life, starting with deciduous incisors (Hillson, 2024; Smith, 1998). For deciduous teeth, amelogenesis typically begins between 14—18 weeks *in utero*, and is completed by approximately 11 months after birth (Hillson, 2024). For permanent teeth, enamel formation begins around the time of birth for the first permanent molars, and continues sequentially for the rest of dentition (Hillson, 2024). First permanent molars begin around birth, and complete between 2.5—3 years of age; incisors and canines begin between 3—5 months and complete between 4—7 years of age; premolars begin between 1.5—2.5 years of age and complete around 5—7 years of age; second molars begin between 2.5—3 years of age and complete around 7—8 years of age; and third molars begin around 8—10 years of age and continue up to 12—16 years

of age or later (Hillson, 2024). This means that enamel development of all teeth spans from roughly 14 weeks *in utero* to late adolescence.

For the sake of this specific project, the timing of enamel formation of the permanent and deciduous canines are most important to take note of, as these are the types of teeth used. The enamel of the permanent canines develops between 3—5 months and 4—7 years of age, while the deciduous canines develop between 14—18 weeks *in utero* and 11 months after birth (Hillson, 2024). These specific timings are crucial to the interpretation of this method's results.

2.3 Diagenesis

Diagenesis, the post-mortem chemical and physical alteration of remains frequently seen in burial conditions (Lambert et al., 1985; Kendall et al., 2018; Kohn et al., 1999; Madupe et al., 2025), is always important to consider when analyzing human remains from an archaeological context. Diagenesis directly affects protein degradation, oxidation, and fragmentation, and thus is critical to results and their interpretations (Hendy, 2021; Lambert et al., 1985; Kendall et al., 2018; Kohn et al., 1999; Madupe et al., 2025). Because it was evaluated in this study as a precautionary measure, it is explained in detail below.

Diagenesis is caused by the interactions between porous organic components and elements associated with deposition, as frequently seen in long-term burials (Hillson, 2024; Kendall et al., 2018; Lambert et al., 1985). Where teeth are generally more durable than bone in archaeological contexts, some tissues within them are more vulnerable to the process of diagenesis than others due to their porosity and high organic contents. Dentine, for example, holds high amounts of collagen proteins similar to bone, making their organic to inorganic ratios relatively comparable (Hillson, 2024; Kendall et al., 2018; Lambert et al., 1985). Because of this abundance of collagen,

dentine is very porous, making it susceptible to chemical alterations (Hillson, 2024; Kendall et al., 2018; Kohn et al., 1998; Lambert et al., 1985). In contrast, enamel is much more resistant to diagenesis due to its low porosity and low organic content, though not immune to diagenesis altogether (Kendall et al., 2018; Lambert et al., 1985). Although enamel does tend to preserve better than dentine and other dental tissues, it may still be chemically altered, especially in certain environments for long periods of time (Kendall et al., 2018; Kohn et al., 1998; Lambert et al., 1985). For this reason, dental tissues that have undergone diagenetic alteration are less desirable for bioarchaeological analysis, since their physical and chemical integrity cannot be reliably predicted (Kendall et al., 2018; Kohn et al., 1998; Lambert et al., 1985).

Environmental conditions play a decisive role in the rate and severity of diagenesis. Wet, oxygen-rich, or acidic soils accelerate the process by facilitating the infiltration of water, ions, sediments, and microorganisms, which in turn promote erosion, dissolution, and molecular degradation (Kendall et al., 2018; Lambert et al., 1985). Conversely, dry and stable burial environments generally promote better long-term preservation (Kohn et al., 1999). Structural damage to the tooth, such as wear or cracking, can further increase permeability and accelerate diagenetic alteration (Kendall et al., 2018; Lambert et al., 1985).

The consequences of diagenesis extend beyond the visible damage of remains, sometimes compromising the molecular history of the remains through deterioration, contamination, and introduction of new proteins and elements embedded in remains (Kendall et al., 2018; Kohn et al., 1998). As a result, studies that rely on diagenetically altered human remains face significant challenges in yielding reliable and significant data (Kohn et al., 1998). This is why careful evaluation of diagenetic alteration is essential in bioarchaeological studies, especially those that

rely on molecular evidence. This evaluation helps to ensure the distinction of authentic biological data from diagenetic noise.

2.4 Social and archaeological context of Gurat, France

2.4.1 Introduction

Gurat is a small village located in the region of Nouvelle-Aquitaine, France, and is home to a rock-cut cave church called St. Georges that was excavated during the late 1960s and early 1970s under the supervision of Dr. Michael Gervers and Veronika Gervers (Franklin & Gervers, 1978; Gervers & Gervers, 1975; Navetat, 2021). Figure 1 situates Gurat within southwestern France, while Figure 2 shows the location of the St. Georges cave church within the village itself. Eighteen individuals were excavated from the cave church, which is believed to have originated as a hermitage and later evolved into a monastery (Gervers, 1967). Of the eighteen individuals excavated, men, women, and children were identified (Clements & Gruspier, n.d.; Meijer, 2018; Seymour, 2019).



Figure 1. Location of Gurat in France. Image borrowed from Google Earth, 2025.



Figure 2. Location of St. Georges within Gurat. Image borrowed from Google Earth, 2025.

The Romanesque cave church was built in the thirteenth century (approximately 1237—1286 CE), and remained active into the seventeenth century (approximately 1655 CE) (Franklin & Gervers, 1978; Gervers & Gervers, 1975; Navetat, 2021). This cave church is one of many rock-cut structures in southwestern France, and one of the eleven known rock-cut churches that exist within a one-hundred-kilometer radius of Gurat (Gervers, 1967). Two pilgrimage routes passed near the village of Gurat and these cave churches, suggesting possible high mobility in the area (Meijer, 2018; Meijer et al., 2019). Figures 3 and 4 illustrate the exterior setting of St. Georges, and Figure 5 depicts the interior of the church.



Figure 3. Exterior view of St. Georges from the North. Image borrowed from Navetat, 2019.



Figure 4. Exterior view of St. Georges from the South. Image borrowed from Navetat, 2019.



Figure 5. Interior view of St. Georges. Image borrowed from Navetat, 2019.

Of the excavated eighteen individuals, thirteen were complete enough for osteobiographies (Seymour, 2019). The individuals have been tentatively dated to approximately 1390 CE based on the recovery of a dated coin within the burials, as well as radiocarbon dating on excavated charcoal and pottery fragments (Franklin & Gervers, 1978; Gervers, 1967). It remains unknown if these individuals represent one population at one time, or if they were part of different populations that occupied the site over the span of a few centuries due to the distribution of the graves and the possibility of secondary burials (Franklin & Gervers, 1978; Gervers, 1967). Though St. Georges has been tentatively dated to the thirteenth century, the cave church may have originated as early as the fourth century, as only speculated in an unpublished report (Clements & Gruspier, n.d.).

Of the research pertaining to the Gurat individuals, much has been related to their identification, where osteobiographies were produced and pathologies were studied (Clements & Gruspier, n.d.; Gretenhart, n.d.; Meijer, 2018; Seymour, 2019). There are also a number of conflicting unpublished reports, all of which begin skeletal analysis but remain uncompleted (Clements & Gruspier, n.d.; Gretenhart, n.d.; Webber & Gretenhart, n.d.). Other studies include chemical analysis of the bones (Hancock et al., 1987), blood typing (Lengyel, 1975), and a bone density study (Webber & Gretenhart, n.d.), though all of these studies focus on method testing and development rather than the contextual information of Gurat (Seymour, 2019). In terms of results, none of these three analyses provided new information regarding the Gurat individuals specifically, other than Lengyel's (1975) article providing age-at-death and biological sex estimates, most of which conflicted from morphological evidence checked by several different authors (Clements & Gruspier, n.d.; Meijer, 2018; Seymour, 2019). In some of the most recent research (Meijer, 2018; Meijer et al., 2019), strontium isotope analysis was conducted, and it was determined that many of the Gurat individuals were likely migrants to the village, demonstrating physical evidence of mobility in an area with historical evidence for that practice. It was also found that most individuals correlated to highland or lowland France, or in the case of one individual, outside of France entirely (Meijer, 2018; Meijer et al., 2019). Even with this research on the Gurat individuals, there is still much to learn about the population, their practices, and their identities.

2.4.2 The Gurat individuals

Individuals were selected for this thesis based on the visible available surface area of their enamel. Many of the teeth within the Gurat collection show signs of extreme wear, so

ensuring that enamel was visibly present for sample collection was imperative. On this basis, six teeth from six different Gurat individuals were analyzed. Some of their prior established information combined from previous studies can be seen in Table 1 below. It should be noted that osteological sex estimates were not conducted as a part of this thesis, and that the presented data is combined from research by Meijer, 2018, and Seymour, 2019. The use of "possible" and "probable" designations of biological sex reflects varying levels of confidence in osteological sex estimation, as recommended by Buikstra and Ubelaker (1994), based on trait expression, skeletal completeness, and internal consistency of indicators. As per Buikstra and Ubelaker (1994), individuals classified as "probable" male or female display multiple, consistent sexually dimorphic features with moderate to strong expression, allowing for greater confidence in the assessment despite the absence of complete or unequivocal skeletal indicators. In contrast, biological sex estimations classified as "possible" male or female reflect more ambiguous and indeterminate traits.

Grave	**Individual**	**Tooth Type**	**Osteological Sex Estimation**	**Age**
Ha	Gu 1	Lower left permanent canine	Probable female	40—50+
Hb	Gu 2	Lower left permanent canine	Probable male	35—50
Dc	Gu 7	Lower right permanent canine	Possible male	35—50
Dc	Gu 8	Upper left permanent canine	Probable male	35—50
Dc	Gu 9	Lower left permanent canine	Possible male	35—50

Dd	Gu 10	Upper left deciduous canine	Indeterminate	3 +/- 12 months

Table 1. List of individuals studied for this thesis, combined data from Meijer, 2018, and Seymour, 2019.

2.4.2.1 Skeletal pathology of the Gurat individuals

Because dental enamel forms primarily *in utero* and early childhood, skeletal pathology observed in adulthood will not reflect physiological conditions present during enamel formation of the canines. This thesis assumes that the organic component preserved within enamel primarily captures developmental history, as it remains uncertain to what extent serum-derived proteins—present in bodily fluids such as saliva or blood—may be incorporated into enamel during later life, particularly under pathological conditions such as carious lesions or extreme dental wear. As the samples analyzed in this thesis were collected from internal enamel surfaces of teeth without visible caries or extensive wear, post-developmental protein infiltration is not expected. Accordingly, interpretations of health in this study focus on physiological processes active during amelogenesis, when enamel proteins are secreted and incorporated into the developing tissue.

It is important to note that this thesis did not conduct original paleopathological assessments. All observations of skeletal pathology discussed here are drawn from previously completed osteological analyses presented by Seymour, 2019. These data are used comparatively to contextualize proteomic findings, rather than as independent lines of observation generated within this project, consistent with the approach taken for osteological sex estimation. By relying on established contextual and osteological data, this research prioritizes methodological development,

reinforcing that proteomic results must be interpreted within existing biological and archaeological frameworks rather than treated as standalone evidence.

Indicators of poor health during childhood that are temporally relevant to enamel formation include linear enamel hypoplasia (Hillson, 2024), Harris lines indicative of growth disruption (Harris, 1931), and skeletal lesions that form during early development, such as cribra orbitalia in subadults (Anderson et al., 2025; Walker et al., 2009). In contrast, most skeletal pathologies observed in adult remains cannot be confidently linked to specific periods of childhood stress due to bone remodeling and the cumulative nature of skeletal change (Wei & Cooper, 2023). As such, enamel provides a uniquely stable archive of early-life physiological conditions, making it particularly well suited for investigating childhood health through proteomic analysis.

On this basis, pathology is only thoroughly and individually discussed for a single individual (Gu 10), whose age-at-death is estimated at approximately 3 years of age. Full osteobiographies, including lists of observed pathologies, can be found in Seymour, 2019, though they are briefly summarized for the other individuals below. Because the observed paleopathology in Gu 10 may have occurred during the narrow window of enamel development, it is reasonable to presume that immune proteins may have been incorporated into their enamel matrix. While the extent to which immune-related proteins may be incorporated into enamel under such conditions remains unknown, the skeletal pathology observed in Gu 10 provides relevant contextual information for interpreting proteomic signals associated with early-life health. Gu 10's skeletal pathology is therefore discussed below.

2.4.2.1.1 Paleopathology and observations of Gu 10

As documented in previous research, Gu 10 exhibits severe cribra orbitalia of the orbital roofs, visible within the eye orbits (Seymour, 2019). Cribra orbitalia is characterized by porous lesions on the superior surfaces of the orbits, and is generally associated with physiological stress during early childhood, most commonly linked to anemia, nutritional deficiency, infectious disease, or chronic inflammation (Anderson et al., 2025; Walker et al., 2009). The presence of severe cribra orbitalia is widely interpreted as evidence of a systemic stress response occurring during periods of rapid growth and development, rather than as a condition acquired later in life (Anderson et al., 2025; Walker et al., 2009).

Because the age-at-death of Gu 10 falls closely within the period of dental enamel formation for the sampled dentition, the observed skeletal pathology is temporally relevant to the proteomic data generated in this study. Any physiological stress or immune response responsible for the development of cribra orbitalia would have occurred during or near the window of amelogenesis, when enamel proteins are actively secreted and incorporated into the enamel matrix. While the presence of cribra orbitalia cannot be directly linked to specific protein expression patterns within enamel, it provides important contextual evidence suggesting that Gu 10 experienced a significant immune or metabolic challenge during early life. This contextual information is therefore considered alongside the proteomic results, particularly in the interpretation of immune-related proteins detected in this individual.

2.4.2.1.2 Paleopathology and observations of the adult Gurat individuals

The adult individuals included in this study (Gu 1, Gu 2, Gu 7, Gu 8, and Gu 9) display extensive skeletal and dental pathology consistent with physically demanding lives, repeated trauma, and persistent biomechanical stress (Seymour, 2019). Common features across the Gurat individuals include severe dental wear and tooth loss, periodontal disease, degenerative joint disease, vertebral compression, Schmorl's nodes, and widespread enthesophyte formation (Seymour, 2019). These patterns are most consistent with cumulative lifetime stress and habitual activity rather than discrete episodes of acute illness (Seymour, 2019).

Among the adult individuals, Gu 2 was found to exhibit the most extensive and complex skeletal pathology (Seymour, 2019). This individual presents with multiple healed fractures, including severe trauma to the cervical spine involving the first three cervical vertebrae, as well as a healed fracture of the sixth thoracic vertebra (Seymour, 2019). Additional evidence of trauma includes a healed coronal fracture of the left talus and localized periosteal reactions on both tibiae and fibulae, suggesting injury or infection affecting the lower limbs (Seymour, 2019). Despite the severity of these injuries, the degree of healing indicates long-term survival following trauma (Seymour, 2019). Gu 2 also displays substantial dental pathology, including multiple abscesses, extensive calculus buildup, and an expanded incisive foramen consistent with systemic infection (Seymour, 2019).

Markers suggestive of childhood physiological stress are present in several adult individuals. Harris lines were observed in Gu 2, Gu 7, Gu 8, and Gu 9, indicating episodes of growth disruption during childhood followed by recovery (Seymour, 2019). While Harris lines reflect nonspecific

stress rather than discrete disease processes, they represent interruptions to growth that necessarily occurred during skeletal development (Harris, 1931; Seymour, 2019). However, Harris lines cannot be precisely correlated with the short timing of enamel formation in the sampled canines.

In contrast, the majority of skeletal pathologies observed in the adult individuals—such as degenerative joint disease, vertebral compression, healed fractures, and extensive enthesophyte formation—are cumulative in nature and most likely developed or progressed during adulthood (Seymour, 2019). These conditions therefore fall outside the developmental window captured by dental enamel. As a result, adult paleopathology in this study is only used to contextualize the Gurat individuals.

Chapter 3: Materials and Methods

3.1 Sample selection

When selecting samples for this thesis, not only was the visible surface area of enamel in each tooth considered, but also the previous studies that had already been done with the collection (Meijer, 2018; Seymour, 2019). To minimize excess destruction of samples, only teeth that had been previously sliced (from cusp to dental root tip) and embedded in resin were considered for this study (n=10) (Meijer, 2018). To ensure a broader representation of individuals, only one tooth from each grave/individual was selected (n=7). Finally, the dimensions of the enamel surface visible on both halves of each tooth were examined through a microscope. One tooth was excluded from analysis due to a very small amount of visible enamel, and for the remaining teeth, the half containing the greatest visible enamel surface was selected for drilling during sample collection (n = 6). This left one tooth from each of the following six individuals: Gu 1, Gu 2, Gu 7, Gu 8, Gu 9, and Gu 10. While most proteomic studies typically use molars due to their size and availability, the teeth used in this study are predominantly canines, as tooth type has not shown to impact results (Buonasera et al., 2024; Castiblanco et al., 2015; Gil-Bona & Bidlack, 2020; Greco et al., 2023; Hillson, 2024; Jágr et al., 2019; Parker et al., 2019; Procopio et al., 2018; Sharma et al., 2020; Shaw et al., 2023). All sampling and analytical decisions in this study were guided by the principle of minimizing destruction while maximizing interpretive value, with enamel powder quantities reduced wherever possible and selected only from teeth already documented in previous osteological research.

3.1.1 Diagenesis

Diagenesis, the post-mortem chemical and physical alteration of remains frequently seen

in burial conditions, was also considered during sample selection (Lambert et al., 1985; Kendall et al., 2018; Kohn et al., 1999; Madupe et al., 2025). Because diagenesis occurs most frequently in the reactive, porous tissues of bone (made predominantly of collagen), the Gurat teeth were closely examined where any dentine exposure was present, given its highly collagenous components (Hillson, 2024; Parker et al., 2019). While enamel, the tissue studied in this thesis, has been shown to be extremely resistant to the processes of diagenesis and more chemically stable than reactive, dentine contamination is possible in sample collection and must be considered (Buonasera et al., 2020; Gil-Bona & Bidlack, 2020; Hendy et al., 2018; Hendy, 2021; Hillson, 2024; Kendall et al., 2016; Kendall et al., 2018; Koenig et al., 2024; Kotli et al., 2025; Madupe et al., 2025; Parker et al., 2019; Stewart et al., 2016). Evaluation of diagenesis is crucial in studies with ancient populations, where remains have been in a variety of environmental conditions over long periods of time (Lambert et al., 1985; Kohn et al., 1999).

From the entire Gurat collection, the teeth that were already embedded in resin were teeth that showed the least amount of wear, caries, cracks, and discolouration (Meijer, 2018). Even still, each tooth selected for this thesis was carefully examined with a microscope for any visual indicators of diagenesis, such as discolouration, pitting, or cracks, to limit the potential of contaminant proteins. Because these teeth had already undergone extensive diagenetic examination in the past when they were cut (Meijer, 2018), and because the enamel samples taken were located within the inside halves of the teeth, limiting their potential of exposure to the elements, this was only precautionary. It is not believed that any of the six samples used in this thesis have undergone diagenesis.

3.2 Sample collection and preparation

Each tooth's surface was gently wiped with acetone and air dried prior to being drilled using a Micromill for sample collection. To limit contamination as much as possible, all surfaces of the drill and scales were also wiped with acetone and Kimtech Kimwipes before and between the drilling of each tooth, and gloves were worn for the duration of sample collection. Lo Bind Eppendorf (EP) tubes were used to collect and weigh all enamel samples, with each sample weighing approximately 10mg.

Sample preparation procedures were modified and combined from Schroeter et al., 2016 and Parker et al., 2019, and took place over the course of several days at the University of Waterloo's Mass Spectrometry Facility (Waterloo, Ontario, Canada). First, each sample was soaked in 1mL of 10% formic acid, stored at room temperature for 68 hours for demineralization (FA, TCI America, Oregon, USA). Samples were then centrifuged, and their supernatants collected and dried. The leftover powder materials from the teeth were labelled and kept in a freezer, where they are available for use in further proteomic analyses. Next, the pH of each sample was neutralized using small amounts of 50mM ammonium bicarbonate (ABC) buffer (Sigma-Aldrich, Missouri, USA), along with 27% ammonium hydroxide (Sigma-Aldrich, Missouri, USA) whenever necessary. Once the pH of each sample reached 7, their total volumes were recorded and used to calculate volumes of materials used for reduction and alkylation. After this step, a blank sample of myoglobin (1µg per 1µL) underwent the sample preparation process at the same time as the Gurat samples to be used as a reference point during mass spectrometry analysis. For the first reduction, 10% of each sample's volume of 50mM dithiothreitol (DTT, Sigma-Aldrich, Missouri, USA) was mixed into each sample before they were transferred to a

dry bath to incubate for 30 minutes at 56℃. After 30 minutes, samples were cooled to room temperature, and the same amount of 100mM indole-3-acetic acid (IAA, Sigma-Aldrich, Missouri, USA) was added to each sample in the dark, where they soaked for 30 minutes at room temperature. After 30 minutes, the same volume of 50mM DTT was added again to reduce each sample, and pH values were verified to ensure successful reduction and alkylation processes (pH = 6). Digestion was carried out overnight at 37℃ using 4µL of MS-grade trypsin (Promega, Wisconsin, USA) per sample, and 0.01% protease enhancer (Promega, Wisconsin, USA), equal to 1% of the samples' total volume. After 18 hours, samples were centrifuged, their supernatants collected and dried down, and then lyophilized and reconstituted in 0.1% FA. Finally, samples were desalted using Pierce™ Peptide Desalting Spin Columns (Thermo Fisher Scientific, Massachusetts, USA). Peptides were eluted with 80% acetonitrile (ACN) and 1% formic acid (FA). Following the desalting process, samples were dried down once more before undergoing mass spectrometry analysis.

3.3 Mass spectrometry and data-dependent acquisition (DDA)

3.3.1 Overview of liquid chromatography-mass spectrometry-based proteomics

Mass spectrometry (MS) is an analytical technique used to identify and quantify biomolecules by measuring the mass-to-charge ratios of ionized molecules (Aebersold & Mann, 2003; Aebersold & Mann, 2016; Watson & Sparkman, 2007). In MS-based proteomics, proteins are typically enzymatically digested into peptides, most commonly using trypsin, because intact protein analysis is less robust in highly complex samples with a wide dynamic range of protein abundances (Aebsersold & Mann, 2016; Cox & Mann, 2008). Peptides are more consistently

ionized, fragmented, and detected by mass spectrometry, enabling more reliable identification and quantification (Aebersold & Mann, 2016; Cox & Mann, 2008). These peptides are introduced into the mass spectrometer through a liquid chromatography-mass spectrometry (LC-MS) workflow, in which liquid chromatography separates peptides over time based on their physicochemical properties, reducing sample complexity and improving analytical sensitivity (Watson & Sparkman, 2007). As peptides elute from the chromatographic column, they are converted into charged ions via electrospray ionization, a soft ionization technique that introduces peptides into the gas phase with minimal fragmentation, in contrast to harsher ionization methods that can break molecules apart and reduce the information available for accurate identification (Fenn et al., 1989; Watson & Sparkman, 2007). This LC-MS workflow enables reproducible detection of peptides in complex biological samples, including degraded archaeological materials where protein concentrations may be low (Hendy et al., 2018; Warinner et al., 2022).

3.3.2 Peptide ionization, MS1/MS2 acquisition, and identification

Once ionized, peptides are analyzed in the mass spectrometer using a two-stage scanning approach. In the first stage (MS1), the instrument measures the mass-to-charge ratios and relative intensities of intact peptide ions, generating the quantitative signal used for label-free protein abundance estimation in this thesis (Aebersold & Mann, 2016). In the second stage (MS2), selected peptide ions, generally prioritized based on their abundance in the MS1 scan, are isolated and fragmented, producing tandem mass spectra that contain sequence-specific information used for peptide identification by comparison with theoretical spectra derived from protein sequence databases (Aebersold & Mann, 2016). Because database searching involves large numbers of

spectral comparisons, false discovery rate (FDR) control is applied to limit incorrect peptide-spectrum matches, with a 1% FDR threshold commonly used at both the peptide and protein levels to ensure confidence in identifications (Benjamini & Hochberg, 1995; Nishimura & Tojo, 2014). For this project, MS2 spectra were used exclusively for peptide and protein identification, while MS1 ion intensities were used for label-free quantification of enamel proteins, allowing relative comparisons of protein abundance across samples while maintaining stringent identification criteria.

3.3.3 Liquid chromatography-tandem mass spectrometry and acquisition parameters

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis was conducted by lab manager Dr. Yiming Xiao at the Mass Spectrometry Facility at the University of Waterloo (Waterloo, Ontario, Canada). Desalted and dried samples were resuspended in 20µL of buffer (0.1% FA), and 10µL of each sample was injected for LC-MS/MS analysis on a Q Exactive Orbitrap mass spectrometer (Figure 6) (Thermo Fisher Scientific, Massachusetts, USA) coupled to a Thermo Ultimate 3000 high-performance liquid chromatography (HPLC) system (Y. Xiao, personal communication, 2025). Chromatographic separation was performed on an ACQUITY UPLC CSH C18 column (2.1 mm × 150 mm, Waters, Massachusetts, USA) maintained at 40°C and operated at a constant flow rate of 200 µL/min (Y. Xiao, personal communication, 2025). Peptides were separated on a column maintained at 40℃ using mobile phases of 0.1% FA in water (A) and 0.1% FA in acetonitrile (B) with the following 82-minute gradient: 5% B (0—2 min), 5—35% B (2—62 min), 35—95% B (62—65 min), 95% B (65—77 min) and 5% B (77—81 min) for re-equilibration (Y. Xiao, personal communication, 2025).

Mass spectrometry data were acquired using a Q Exactive Orbitrap (Thermo Fisher Scientific, Waltham, Massachusetts, USA), operating in data-dependent acquisition (DDA) mode using the top-ten method, selecting the ten most abundant ions within the specified range for fragmentation and MS2 analysis (Tsou et al., 2015; Y. Xiao, personal communication, 2025). In data-dependent acquisition (DDA), precursor ions are selected for fragmentation based on their signal intensity in the MS1 scan, prioritizing the most abundant ions for MS2 analysis (Aebersold & Mann, 2016). This method enables broad proteomic coverage while maintaining accuracy (Tsou et al., 2015). A detailed summary of instrument parameters is provided in Table 2.

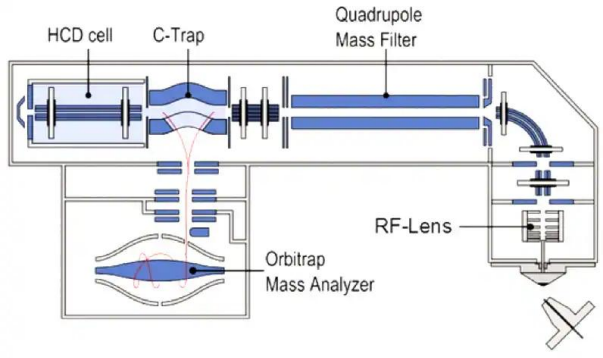


Figure 6. Diagram of Q Exactive Orbitrap used. Image borrowed from Thermo Fisher Scientific.

Parameter	Setting
MS1 Resolution	70,000 (at m/z 400)
MS1 Scan Range	300—1,650 m/z
MS1 AGC Target	3 x 10^6
MS1 Maximum Injection Time	100 ms
RF Lens	50%

MS2 Resolution	17,500
MS2 AGC Target	1×10^5
MS2 Maximum Injection Time	50 ms
Isolation Window	2 m/z
Fragmentation Method	Higher-energy collisional dissociation (HCD)
Normalized Collision Energy (NCE)	30%
Dynamic Exclusion	15 s
Charge States Excluded	7+, 8+, and >8+
Microscans per Spectrum	1

Table 2. Instrument parameters provided by Dr. Yiming Xiao, Mass Spectrometry Facility, University of Waterloo (Waterloo, Ontario, Canada).

3.4 Peptide identification, analysis, and label-free quantification (LFQ)

Raw MS files were processed using FragPipe (v23.1) developed by the Nesvizhskii Lab (University of Michigan, Ann Arbor, MI, USA). Each sample was categorized into a separate experimental group under the Basic-Search workflow (Y. Xiao, personal communication, 2025). Database searches were conducted against a custom protein database containing Amelogenin-X (AMELX), Amelogenin Y (AMELY), Immunoglobulin-Gamma (IgG) heavy chain, and C-reactive protein (CRP), with contaminant and decoy sequences appended automatically by FragPipe (Y. Xiao, personal communication, 2025). This workflow is slightly modified from Parker et al., 2019.

Peptide-spectrum matches (PSMs) were filtered to a 1% false discovery rate (FDR) at both the peptide and protein levels (Y. Xiao, personal communication, 2025). All FragPipe

search parameters are summarized in Table 3. It is important to note that search parameters included a setting to detect oxidation (+15.99 Da) on methionine, and deamidation (+0.98 Da) on asparagine and glutamine. Including oxidation and deamidation as searchable modifications ensures that chemically altered peptides are not overlooked during protein identification. Because these modifications accumulate through post-mortem processes, such as diagenesis, their detection provides important context for evaluating protein preservation and interpreting the presence of diagenetically sensitive proteins such as albumin, a common lab contaminant. While these modification patterns cannot independently confirm protein antiquity, they contribute to a broader assessment of peptide authenticity when considered alongside consistency across samples.

Parameter	**Setting**
Database	Custom (AMELX, AMELY, IgG heavy chain, CRP) + contaminants + decoys
Precursor Mass Tolerance	20 ppm
Fragment Mass Tolerance	0.05 Da
Digestion Type	Non-specific
Peptide Length Range	4—20 amino acids
Fixed Modifications	Carbamidomethylation (+ 57.02 Da) on cysteine
Variable Modifications	Oxidation (+15.99 Da) on methionine, deamidation (+0.98 Da) on asparagine/glutamine
Maximum Variable Modifications per Peptide	4
FDR Control	1% (peptide and protein level)

Table 3. FragPipe search parameters, provided by Dr. Yiming Xiao, Mass Spectrometry Facility, University of Waterloo (Waterloo, Ontario, Canada).

Finally, protein quantification was performed using the IonQuant module (Nesvizhskii Lab, University of Michigan) implementing the MaxLFQ algorithm (Max Planck Institute of Biochemistry, Martinsried, Germany) (Y. Xiao, personal communication, 2025). Match-between-runs (MBR) was enabled to improve identification consistency across samples, allowing peptide matches within a 1-minute retention time window. Peptide intensities were normalized using total ion current (TIC) normalization to correct for technical variability in overall signal intensity between runs, and these normalized peptide intensities were subsequently used for label-free quantification (Y. Xiao, personal communication, 2025). This approach allows LFQ values to be interpreted as relative measures of protein abundance while reducing the influence of technical variation (Merrick, 2025).

Chapter 4: Results

A total of 178 peptides were identified across the six samples. Each protein previously identified as a protein of specific interest to this project is discussed below.

4.1 Amelogenin

Forty-nine peptides associated with amelogenin were identified across the six samples. In order to conduct sex estimation of the Gurat individuals, ambiguous peptides (peptides that may be associated with both AMELX and AMELY) were ruled out from analysis (n=23; 46.93%). This left 20 peptides specific to AMELX (40.82%), and 6 peptides specific to AMELY (12.22%). To further narrow this analysis, the remaining peptides were analyzed by the following criteria: the number of times they occurred across the six samples (1—2+), the match type at which they were identified (MS/MS prioritized over MBR), the spectral counts associated with their identifications (generally 2+), the length of the peptides (between 6—25 amino acids), and their MaxLFQ intensities (>0). The application of these conservative filtering criteria substantially reduced the number of peptides retained for downstream interpretation. Consequently, the relative abundance of AMELX observed in this study is lower than that reported in some previous amelogenin-based sex estimation studies. This pattern reflects methodological stringency rather than biological absence, as AMELX peptides are more likely to be excluded due to sequence overlap with AMELY and lower confidence identifications. Following this multi-criteria assessment, the most accurate, reliable peptides (n=9) were identified as ensuring the most accurate analysis of the biological sex of each sample. Table 4 lists these peptide sequences, the associated protein with that sequence, and which samples positively identified these sequences. Figure 7 demonstrates an example of peptide identification in the case of Gu 2.

Sequences highlighted in green represent peptides specific to AMELY, and peptides highlighted in grey represent the overlap with AMELX, or ambiguous peptides. Figures 8 and 9 illustrate examples of the annotated spectra for the peptide sequences SMIRPPY and MIRPPY, respectively, both of which are specific to AMELY. These spectra provide visual confirmation of the peptide identification generated through database searching, showing the fragmentation patterns and ion matches that support confident peptide assignment. Figure 9 further highlights the oxidation of amino acid M, or methionine residue within the MIRPPY sequence. The identification of this modification demonstrates the sensitivity of proteomic analysis in detecting subtle molecular changes.

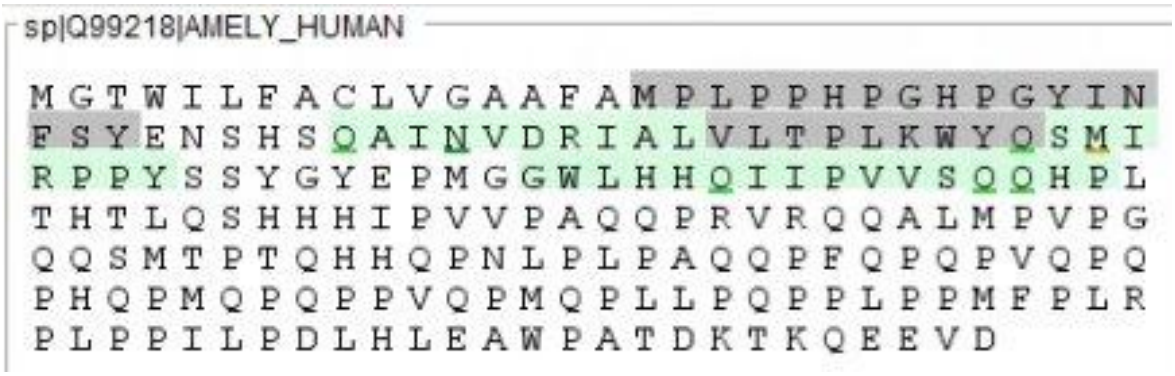


Figure 7. Peptides identified in Gu 2. Green peptides represent peptides specific to AMELY, where grey peptides represent the overlap with AMELX. Image provided by Dr. Yiming Xiao, Mass Spectrometry Facility, University of Waterloo (Waterloo, Ontario, Canada).

Peptide Sequence	Protein (AMELX or AMELY)	Samples Identified
GHPGYINFSYEVLTPK	AMELX	Gu 2, Gu 8, Gu 9
SIRPPYSPY	AMELX	Gu 2, Gu 8, Gu 9, Gu 10
SIRPPYPSYG	AMELX	Gu 2, Gu 7, Gu 8

SYEVLTPLK	AMELX	Gu 2, Gu 7, Gu 8, Gu 9
WYQSIRPPYPSYGYEPMGGW	AMELX	Gu 1, Gu 2, Gu 7, Gu 9
YEVLTPLK	AMELX	Gu 2, Gu 7, Gu 8, Gu 9
IALVLTPK	AMELY	Gu 2, Gu 7, Gu 8
MIRPPY	AMELY	Gu 2, Gu 7, Gu 8, Gu 9
SMIRPPY	AMELY	Gu 2, Gu 8

Table 4. Selected amelogenin peptide identification in Gurat individuals.

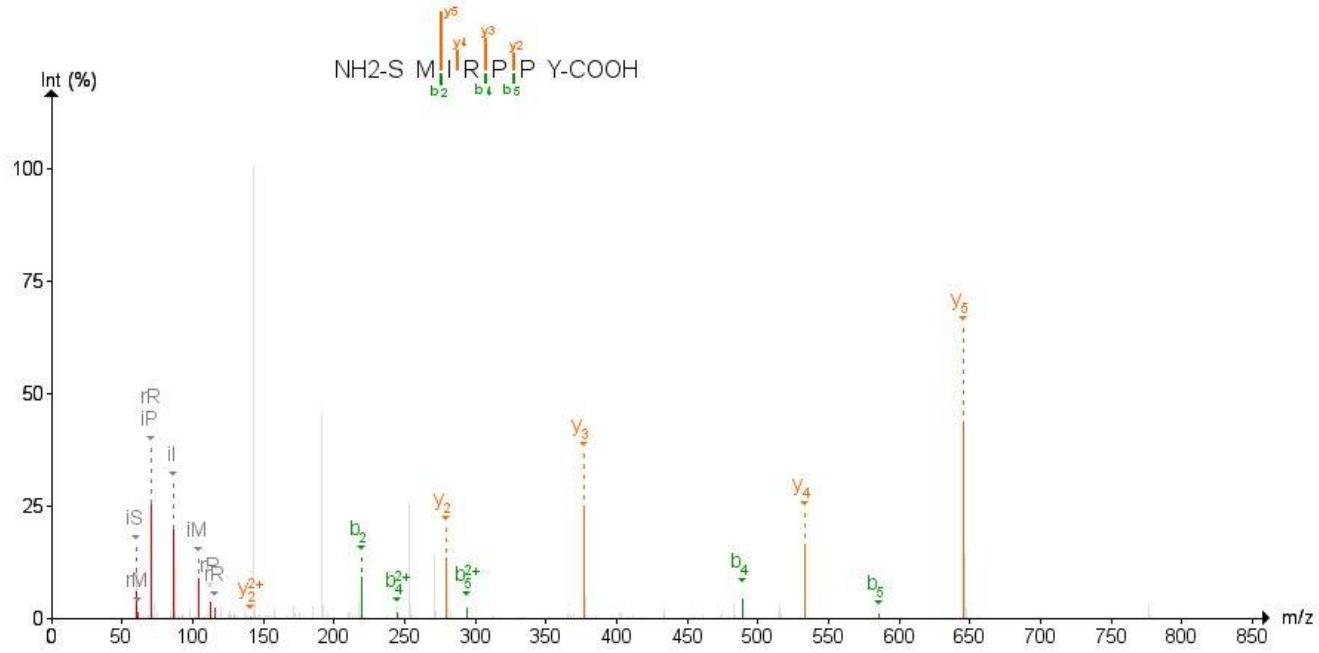


Figure 8. Peptide annotation for SMIRPPY (AMELY), provided by Dr. Yiming Xiao, Mass Spectrometry Facility, University of Waterloo (Waterloo, Ontario, Canada). Taller peaks represent higher intensities detected in fragmented ions.

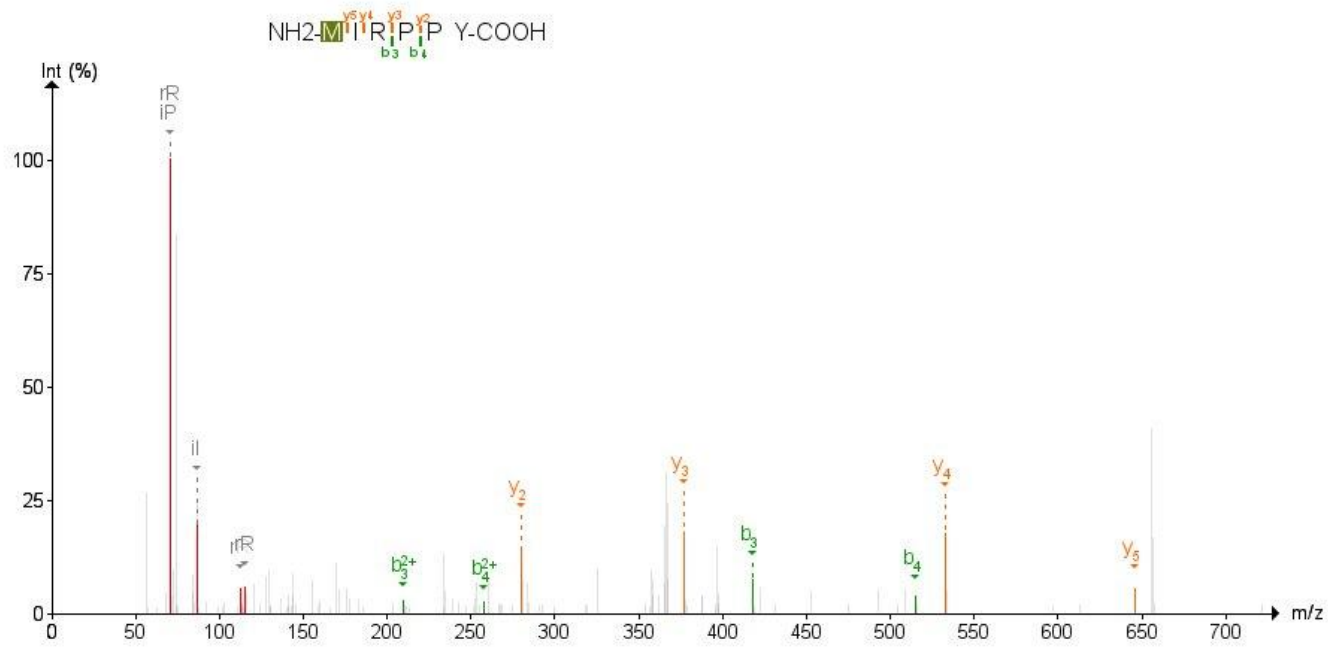


Figure 9. Peptide annotation for MIRPPY (AMELY), demonstrating the oxidation of amino acid M, reflecting chemical alteration. Image provided by Dr. Yiming Xiao, Mass Spectrometry Facility, University of Waterloo (Waterloo, Ontario, Canada). Taller peaks represent higher intensities detected in fragmented ions.

Some samples (ex: Gu 2) identified several more amelogenin peptides than other samples (ex: Gu 1, Gu 10), and were generally more protein abundant. Following Parker et al. (2019), TIC (total ion current)–normalized peptide intensities were log10-transformed prior to analysis of biological sex in order to facilitate comparison of peptide signal distributions and sample performance under the tested method. AMELX_HUMAN and AMELY_HUMAN were subsequently quantified using combined MaxLFQ ion intensity measurements for each sample (Parker et al., 2019). Results were not adjusted based on the total powder each sample originated with, due to their very close similarity in weight. A value of 1 was added to each combined value

before the data were log-transformed (Parker et al., 2019). Based on these final values, male or female sex of each Gurat individual was predicted, assuming that the presence of any AMELY (>0) indicated male biological sex. With such a small sample size, and under the assumption that any presence of AMELY peptides would be indicative of male biological sex, statistical analysis and logistical regression were not performed. Table 5 represents the log-transformed values, along with osteological and proteomic sex estimations for the Gurat individuals, and Figure 10 shows these log-transformed values across each sample.

Sample ID	Log AMELX	Log AMELY	Osteological Sex Estimation	Proteomic Sex Estimation
Gu 1	6.118530282	0	Probable female	Possible female
Gu 2	8.61680402	7.57440834	Probable male	Male
Gu 7	7.449649252	6.755856436	Possible male	Male
Gu 8	9.002567981	7.27243044	Probable male	Male
Gu 9	7.706878431	6.717287997	Possible male	Male
Gu 10	5.87335592	0	Indeterminate	Possible female

Table 5. Comparison of osteological and proteomic sex estimations for the Gurat individuals analyzed in this study. Log-transformed peptide intensities for amelogenin-X (AMELX) and amelogenin-Y (AMELY) are shown for each individual, alongside previously published osteological sex estimates (Meijer, 2018; Seymour, 2019) and proteomic sex estimations derived from the presence or absence of sex-specific amelogenin peptides. A value of 0 indicates that no peptides for that amelogenin group were detected. Proteomic sex estimation demonstrates clearer differentiation for males through the detection of AMELY peptides, while female estimations rely on the absence of AMELY and remain comparable in confidence to osteological assessments.

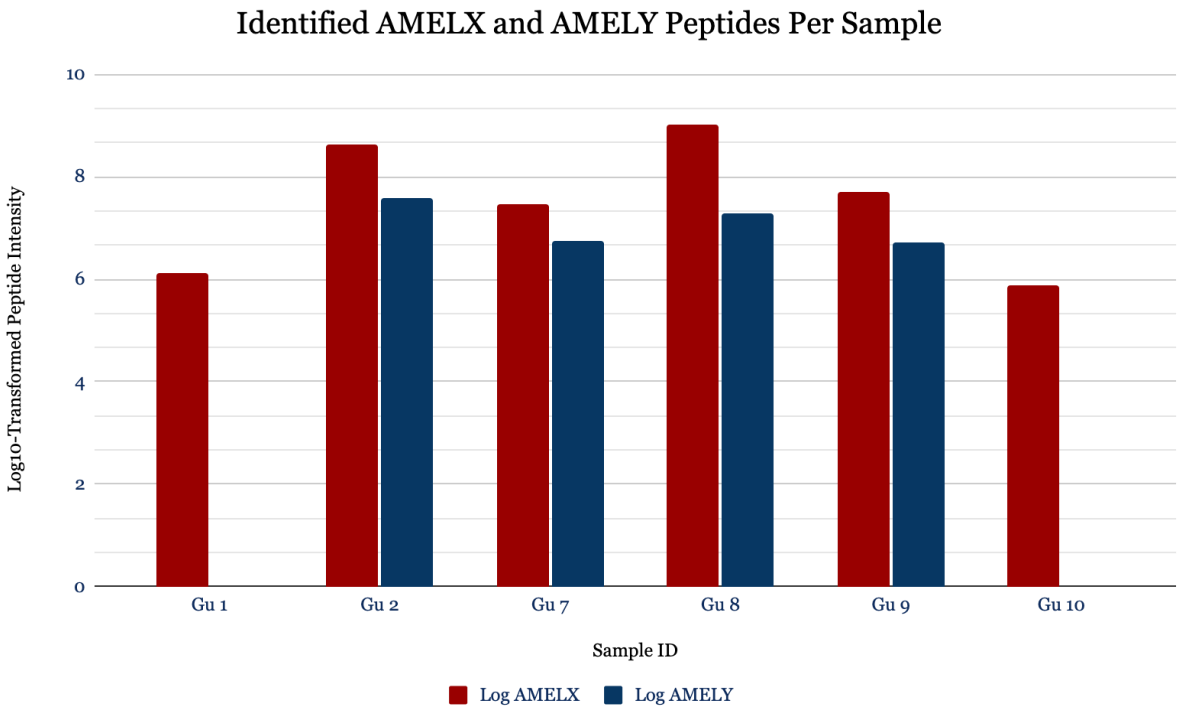


Figure 10. Log10-transformed peptide intensities of identified AMELX and AMELY in each sample.

It was not possible to definitively assign female sex to Gu 1 and Gu 10 due to a general lack of evidence. Not only were these samples comparatively not protein abundant overall but their AMELX signals were relatively low in comparison to those of the other individuals included in this analysis. Furthermore, it is important to note that an absence of evidence does not equal an evidence of absence, and that it may be possible AMELY peptides were present, but not identified. On proteomic evidence alone, the biological sexes of Gu 1 and Gu 10 can only be estimated as possible females.

4.2 C-reactive protein

Only one peptide specific to CRP was identified amongst the six samples (KAFVFPK for Gu 2). Based on the criteria applied to the amelogenin peptides (i.e. the number of times a peptide occurred across the six samples, the match type at which they were identified, the spectral counts associated with their identifications, the length of the peptides, and their MaxLFQ intensities), this peptide does fit most of these criteria, with its isolated detection making it only somewhat preferable for selection. As with the amelogenin peptides, the data for CRP in Gu 2 were log-transformed in order to compare them to the other proteins identified in this thesis. These data are reflected in Table 6.

Sample ID	Log CRP
Gu 1	0
Gu 2	6.757034472
Gu 7	0
Gu 8	0
Gu 9	0
Gu 10	0

Table 6. Log-transformed data of CRP protein present in the Gurat individuals.

4.3 Immunoglobulin-gamma

IgG peptides were not reliably identified in any of the six Gurat individuals. Any peptides that were identified did not fit the above-utilized criteria for identifying significant results.

4.4 Albumin

Though it was not first determined as a protein of interest in this study, the most abundant protein found in serum, albumin, was also identified. Albumin has been used in other proteomic studies of immune proteins to measure degradation, as it is soluble and globular like IgG and CRP (Buonasera et al., 2024). Because of this, albumin peptides were also assessed in hopes of comparing them to the CRP and IgG peptides found in the Gurat individuals. Selected peptides are detailed in Table 7. While the Gurat individuals did not have any IgG and very little CRP proteins identified in their teeth, serum albumin levels were comparatively high. Because albumin is a known and common laboratory contaminant, its identification was evaluated carefully, using mass spectrometer parameters and cross-sample patterns to assess whether detected albumin peptides were more consistent with ancient preservation or modern contamination. Table 8 demonstrates these values taken from seven selected peptides, and Figure 11 reflects the normalized values of all discussed peptides found in this study across each sample. The comparatively high quantities of albumin indicate that the low levels of CRP and IgG that were found may not be due to protein degradation over time, and that CRP and IgG may not have been abundantly present to begin with (Buonasera et al., 2024). Conversely, it can be argued that this abundance of albumin and lack of IgG and CRP suggests that antibodies degrade differently than other structural proteins in biological materials of human remains, and that immune proteins lose their immunoreactivity, as has been suggested in other work attempting to extract IgG peptides (Kendall et al., 2016).

Peptide Sequence	Samples Identified
SLHTLFGDKLCTVATLR	Gu 1, Gu 2, Gu 9
RHPDYSVVLLLR	Gu 1, Gu 2, Gu 7, Gu 8, Gu 10
QTALVELVK	Gu 2, Gu 7, Gu 8, Gu 9
PDYSVVLLLR	Gu 1, Gu 2, Gu 7, Gu 9
LVAASQAALGR	Gu 1, Gu 7, Gu 8, Gu 10
KVPQVSTPTLVEVSR	Gu 1, Gu 2, Gu 7, Gu 8, Gu 9, Gu 10
KLVAASQAALGL	Gu 1, Gu 2, Gu 7, Gu 9, Gu 10

Table 7. Selected serum albumin peptide sequences identified in Gurat individuals.

Sample ID	Log ALBUMIN
Gu 1	7.891094349
Gu 2	7.75363686
Gu 7	7.506398531
Gu 8	7.479440567
Gu 9	8.052562199
Gu 10	7.656522577

Table 8. Log-transformed data for selected serum albumin peptide presence in Gurat individuals.

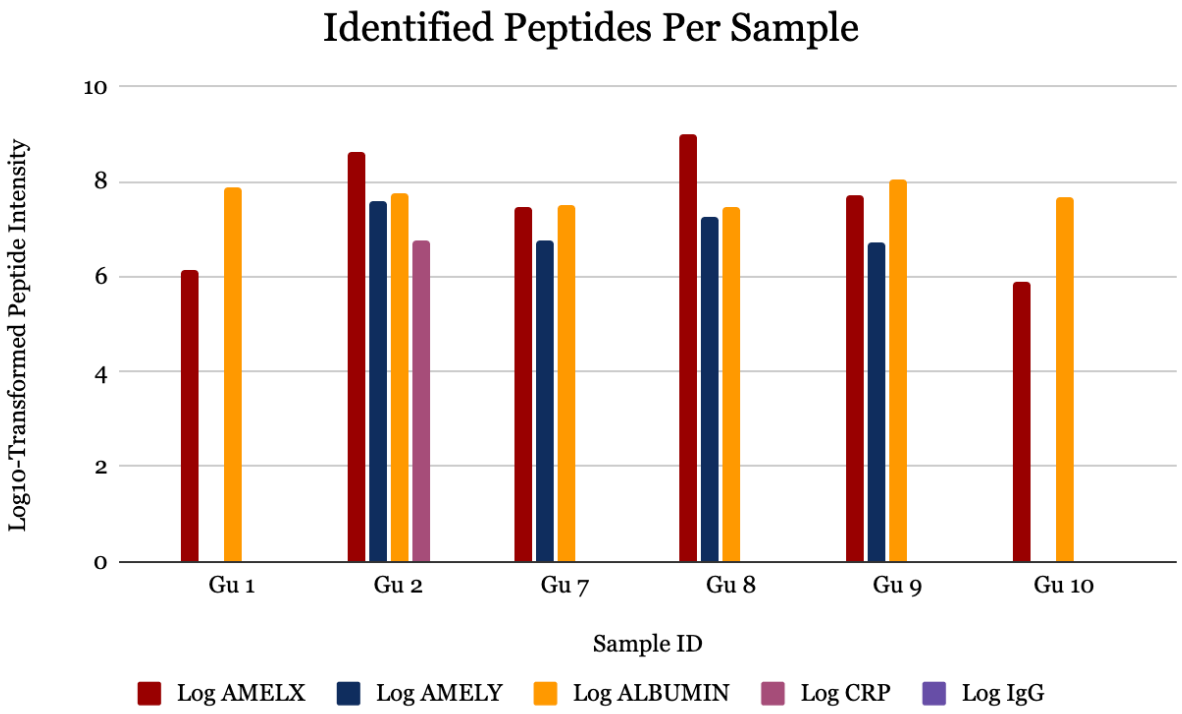


Figure 11. Log10-transformed peptide intensities of identified AMELX, AMELY, Albumin, CRP, and IgG in each sample

Chapter 5: Discussion

5.1 Biological sex of the Gurat individuals

This study demonstrates that amelogenin peptides (AMELX and AMELY) can be successfully identified using the modified analytical methods employed here, and highlights the potential contributions of paleoproteomics to biological sex estimation in archaeological contexts. Where the biological sex of Gu 2, Gu 7, Gu 8, and Gu 9 had previously been estimated as possible or probable males based on osteological analyses (Meijer, 2018; Seymour, 2019), the proteomic results presented here confirm these individuals as male with greater confidence through the identification of AMELY peptides. In contrast, the proteomic sex estimation for Gu 1 and Gu 10 remains uncertain. Both individuals yielded comparatively low overall protein abundance, and no non-ambiguous AMELY peptides were identified. Each sample produced only a single reliable AMELX-specific peptide, limiting interpretive confidence. Osteological analyses have previously identified Gu 1 as a probable female based on pelvic morphology (Meijer, 2018; Seymour, 2019), while Gu 10's biological sex was indeterminate due to their age. In both cases, proteomic evidence can only suggest a possible female classification. Importantly, this uncertainty reflects an intentional ethical position adopted throughout this thesis: where evidence is limited or ambiguous, interpretations are constrained rather than overstated. This approach is particularly critical for juvenile individuals such as Gu 10, where developmental biology, preservation variability, and methodological limitations intersect most strongly.

Comparative studies have demonstrated that proteomic sex estimation can be as accurate as, or more accurate than, osteological and genomic approaches (Adair et al., 2025; Buonasera et al., 2020; Buonasera et al., 2024; Koenig et al., 2024; Parker et al., 2019). However, this study

deliberately avoids equating the absence of AMELY peptides with definitive female biological sex. While some proteomic frameworks would classify Gu 1 and Gu 10 as female based solely on the absence of AMELY, this thesis maintains that absence of evidence does not constitute evidence of absence, particularly in samples characterized by low overall protein yield.

These results contribute to ongoing discussions surrounding confidence thresholds in proteomic sex estimation. Statistical classification was not applied in this study due to both the small sample size (n = 6) and the absence of consensus regarding quantitative thresholds for female sex estimation. Parker et al. (2019), for example, propose that the absence of AMELY combined with AMELX intensities exceeding 5.31×10^8 CI/mg corresponds to a greater than 90% probability of female biological sex. Neither Gu 1 nor Gu 10 meets this threshold. Moreover, Parker's criteria cannot be directly applied here due to differences in sample mass (10mg vs. 20mg), extraction protocols, and analytical workflows. Applying externally derived thresholds under these conditions would risk overstating interpretive certainty.

Though the presence of AMELY alone should be considered sufficient evidence of male sex, determining female sex is not yet something that can be confidently done with proteomic evidence alone. The approach of absence of AMELY equating to the estimation of female biological sex does not account for the potential of environmental or diagenetic factors that cannot be seen with the naked eye (Hendy, 2021). Additionally, differences in organic composition between permanent and deciduous teeth—and how these tissues preserve over time—remain insufficiently understood. As such, it is not possible to determine whether the limited AMELX representation observed in Gu 1 and Gu 10 reflects biological variation, developmental timing, or taphonomic processes. In this context, restrained interpretation in cases of ambiguity should be

viewed not as a limitation of proteomic analysis, but as an ethical safeguard in bioarchaeological research.

5.2 Immune proteins and health of the Gurat individuals

The lack of immune proteins identified in this study could be for any number of reasons, which are explored throughout this chapter. In terms of interpreting health among the Gurat individuals, this study does not seek to replace or replicate osteological assessments, but rather to evaluate whether enamel proteomics can provide complementary insight into physiological processes active during tooth formation (Meijer, 2018; Seymour, 2019). While Harris lines can be observed on Gu 2, Gu 7, Gu 8, and Gu 9 (Seymour, 2019), it cannot be confirmed that these Harris lines coincided with the narrow window of enamel development for all four individuals. This temporal uncertainty may help explain why immune proteins were not identified for Gu 7, Gu 8, and Gu 9. It is also possible that immune proteins were not secreted in high abundance, leaving them undetected by the instrumentation used.

Further, though CRP was identified in Gu 2, it can only be speculated that this individual experienced a period of immune response during the development of their enamel—from approximately 3—5 months until approximately 4—7 years of age (Hillson, 2024). Due to the unknown exact functions of CRP and lack of meaningful paleopathological evidence, the cause of this immune response cannot be determined. However, Gu 2's evidence of Harris lines (Seymour, 2019) provides limited but temporally relevant context for interpreting the presence of CRP, even though a direct causal link cannot be established.

In the case of Gu 10, where there is morphological evidence of immune activity potentially during their years of enamel development (Seymour, 2019), no targeted immune proteins were

detected. This absence may reflect limitations of the modified method and instrumentation, reduced preservation of immune-related proteins, or a biological scenario in which Gu 10's immune response did not result in the incorporation of immune proteins into the enamel matrix at detectable levels. It is also possible that Gu 10's cribra orbitalia did not develop until after enamel formation.

The findings of this research support the hypothesis of Kendall et al., 2016, who argue that IgG may be an unsuitable target for detecting disease-associated antigens. The inconsistency across samples in this study and in others does not suggest that immune proteins may reliably inform interpretations of health in ancient populations using the methods that have been employed thus far (Buonasera et al., 2024; Kendall et al., 2016). As suggested by Kendall et al., antibodies (a term also used for IgG) may preserve poorly, and differently than other organic components of teeth and bone (Kendall et al., 2016). The same authors also found IgG difficult to extract after testing three different methods, which all differ from the one employed in this thesis (Kendall et al., 2016). In considering these results, along with the ones provided by Buonasera et al. (2024), immune proteins and their quantification have a long way to go before any specific interpretations can be made, due to the vague nature of immune proteins and their functions. Their identification requires more sensitive methods than the identification of other proteins, such as amelogenin. Though the instrument used in this study is not optimized for these kinds of identifications, the discovery of even one immune protein demonstrates the potential immune proteomics holds as these methods continue to advance. Though these results support the hypothesis of Kendall et al., (2016), the last decade of proteomic research demonstrates further avenues of research worth exploring.

Buonasera et al., 2024, anticipate that serum albumin should react similarly to IgG and CRP, due to their forms being similar in nature. They argue that the level of protein degradation should apply with similar kinetics amongst those three proteins, and that their relative intensities could provide a record of stress/disease during tooth formation (Buonasera et al., 2024). This pattern is not supported by the results of the present study. While it is entirely possible that ancient IgG cannot be identified with the method used here, very consistent quantities of albumin were found in all six samples. It could be possible that albumin preserves better than immune proteins, though much further research in this avenue is required. Even still, the vast difference in protein abundance between IgG and CRP and albumin and amelogenin calls for further exploration, especially in consideration of Gu 10.

Further, the differences in timing between immune activity responses and enamel formation may also explain the absence of immune proteins in most of the individuals included in this study. Though they are not completely understood, antibodies like IgG are believed to circulate in response to short-term infections, whereas enamel formation is generally a more gradual process (Ansar, 2020; Buonasera et al., 2024; Chiang et al., 2019; Cundiff et al., 2024; Hillson, 2024; Kendall et al., 2016; Oliveira et al., 2023). Under this assumption, immune activity must coincide precisely with the enamel formation process for immune proteins to successfully be detected in enamel (Buonasera et al., 2024). With an organic component so low, capturing immune proteins could be difficult unless they existed in abundance. In this light, the absence of immune proteins found in the enamel of the Gurat individuals would not necessarily be interpreted as an absence of inflammation or healing, but as a reflection of the timing of the incorporation of proteins/organic content in dentition.

In keeping with the ethical framework outlined at the beginning of this thesis, these results are interpreted with particular caution and restraint. The Gurat individuals are not simply datasets, but people whose biological remains reflect complex life histories that cannot be reduced to isolated molecular signals. As such, the absence or presence of immune-related proteins is not treated as definitive evidence of health or illness, but rather as one line of contextual information considered alongside osteological data, archaeological context, and known methodological limitations. By avoiding overinterpretation and explicitly acknowledging uncertainty, this study aims to balance scientific inquiry with respect for the individuals and communities represented, emphasizing transparency, proportionality, and care in the interpretation of paleoproteomic data.

5.3 Limitations and future directions

The largest obstacle in this thesis research was the time constraint. With more time and further resources, current workflow and instrumentation set up could have been modified further to improve results. As it currently stands, this method can successfully detect amelogenin, but cannot reliably detect immune proteins. Modifications made to employ the method used in this project were made in the interest of cost, resources, and time, and these slight modifications could be the reason for these results. Outside of the sample preparation process, the use of newer, more sensitive mass spectrometry equipment could have potentially provided different results. Further, the use of data processing methods outside of what was achieved here may also have yielded better results.

Beyond time and resource limitations, several methodological constraints deserve discussion, such as sample size and representativeness. Only six individuals could be included for this study, which clearly limits the ability to provide any population-level interpretations.

Additionally, multiple teeth from the same individuals were not run as replicates, which could have better spoken to the success and reliability of this method. Exploring these avenues could have allowed for better assessment of individual variability in protein preservation.

Another limitation involves the possibility of contamination. Though this study did not identify collagenous proteins associated with dentine, there is, of course, the possibility that dentine was included in small amounts throughout the samples. There is also the possibility that modern contaminants could have been included in the samples, despite the database of contaminants that was run during mass spectrometry analysis. Though many precautionary measures were taken to limit contamination, such as the cleaning of teeth surfaces, tool surfaces, and all instrumentation used, the heavy-handling of these samples over time may have introduced modern contaminants. In this event, it could be possible that low-abundance ancient proteins were overshadowed by the contaminants.

As this field continues to grow, there is very much left to explore. Though this study used the smallest enamel powder sample that has been published to date, less destructive enamel sampling techniques applicable to paleoproteomics should continue to be developed. In the case of the smaller enamel sample size used here, preliminary research found that equivalent results were yielded in comparison to samples double the size. The preservation of human remains is paramount, and this should be kept in mind during any proteomic study that is destructive in nature, particularly when methodological gains are incremental rather than transformative. Also left to be explored is the specific preservation of ancient antibodies. Though various studies have confirmed that enamel and its structural proteins preserve remarkably well, the preservation of proteins such as serum proteins should continue to be investigated (Bartlett, 2013; Buonasera et al., 2024;

Castiblanco et al., 2015; Gil-Bona & Bidlack, 2020; Hillson, 2024; Kendall et al., 2016). Experiments with this focus may shed light on whether the inconsistent amounts of immune proteins present in archaeological samples reflects individual biology or taphonomic conditions (Buonasera et al., 2024; Kendall et al., 2016). And finally, and perhaps most importantly, future research should emphasize comparisons between contemporary populations and ancient ones when it comes to proteomic method development. Having definitive and known information would greatly assist in testing accuracy and applicability of proteomic methods to bioarchaeological populations where little is known. This is critical to the foundation of this research. Taken together, these findings emphasize the importance of ethical proportionality in paleoproteomic research, where the pursuit of new molecular insights must be balanced against the limits of current methods and the responsibility to represent past individuals with accuracy, care, and restraint.

Chapter 6: Conclusion

This thesis investigated the potential of enamel proteins using a modified method to contribute to bioarchaeological interpretations of biological sex and health in six individuals from a monastery in Gurat, France. The results confirmed that amelogenin peptides could be successfully identified from archaeological enamel samples and could serve as reliable markers for biological sex in most individuals. These findings reinforce the reliability of amelogenin-based sex determinations, while emphasizing the need for caution in interpreting the results from samples with low protein abundance. In contrast, the results suggest that immune proteins may not serve as reliable bioindicators of ancient health using the modified method employed here. Though another serum protein, albumin, was consistently present across all samples, immune proteins could generally not be identified with confidence. For this sample then, enamel proteomics does not yet offer a dependable means of reconstructing health or immune activity in archaeological contexts through the use of this method.

While this study's results were modest in scope, they hope to provide meaningful contributions to the growing field of paleoproteomics. The methods used here are sufficient for amelogenin protein identification and biological sex estimation. While the limited success in detecting immune proteins was initially disappointing, this highlights the current methodological needs for future research. This study also demonstrated that even with limitations in resources, a modified extraction protocol with smaller samples can yield meaningful proteomic data. With continual technological refinement and interdisciplinary collaboration, paleoproteomic studies using tooth enamel hold immense potential for reconstructing the unknown histories of archaeological populations. Taken together, this research emphasizes the importance of

methodological transparency, interpretive restraint, and proportionality in paleoproteomic analyses, ensuring that molecular insights are balanced with respect for the individuals and communities represented in the archaeological record.

References

Abelson, P. H. (1954). Amino acids in fossils. Science 119, 576

Adair, L. R., Lewis, M. E., Collins, M. J., & Cramer, R. (2025). LAP-MALDI MS analysis of amelogenin from teeth for biological sex estimation. *Journal of Pharmaceutical and Biomedical Analysis*, *255*, 116599-

Aebersold, R., & Mann, M. (2003). Mass spectrometry-based proteomics. *Nature (London)*, *422*(6928), 198–207.

Aebersold, R., & Mann, M. (2016). Mass-spectrometric exploration of proteome structure and function. *Nature (London)*, *537*(7620), 347–355.

Agarwal, S. C., & Glencross, B. A. (2011). *Social Bioarchaeology* (1. Aufl.). Wiley-Blackwell.

Agarwal, S. C., & Wesp, J. K. (Eds.). (2017). *Exploring Sex and Gender in Bioarchaeology*. University of New Mexico Press.

Anderson, A. S., Blackwell, A., Sutherland, M. L., Kraft, T., Sutherland, J., Beheim, B., Cummings, D., Ghafoor, S., Hooper, P. L., Rodriguez, D. E., Irimia, A., Gatz, M., Mack, W., Rowan, C., Miyamoto, M., Buetow, K., Finch, C., Wann, L. S., Allam, A., … Gurven, M. D. (2025). Childhood skeletal lesions common in prehistory are present in living forager-farmers and predict adult markers of immune function. *Science Advances*, *11*(29), eadw3697

Ansar, W. (2020). Multiple Faces of C-Reactive Protein: Structure–Function Relationships. In: Ansar, W., Ghosh, S. (eds) Clinical Significance of C-reactive Protein. Springer, Singapore.

Appleby, J., Thomas, R., & Buikstra, J. (2015). Increasing confidence in paleopathological diagnosis – Application of the Istanbul terminological framework. *International Journal of Paleopathology*, *8*, 19–21

Bartlett, J. D. (2013). Dental Enamel Development: Proteinases and Their Enamel Matrix Substrates. *ISRN Dentistry*, *2013*, 684607–684624.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B, Methodological*, *57*(1), 289–300.

Brickley, M. B., & Morgan, B. (2023). Assessing diagnostic certainty for scurvy and rickets in human skeletal remains. *American Journal of Biological Anthropology*, *181*(4), 637–645.

Brookes, S. J., Barron, M. J., Dixon, M. J., & Kirkham, J. (2017). The Unfolded Protein Response in Amelogenesis and Enamel Pathologies. *Frontiers in Physiology*, *8*, 653.

Bruckett, C. & Law, T. (2018). An Intersectional Lens on Gendered Violence. *Women and Gendered Violence in Canada* (19-41). Toronto: University of Toronto Press.

Brůžek, J., Mikšík, I., Pilmann Kotěrová, A., Morvan, M., Drtikolová Kaupová, S., Santos, F., Danielisová, A., Zazvonilová, E., Maureille, B., & Velemínský, P. (2024). Undertaking the biological sex assessment of human remains: The applicability of minimally-invasive methods for proteomic sex estimation from enamel peptides. *Journal of Cultural Heritage*, *66*, 204–214.

Buckley, D. I., Fu, R., Freeman, M., Rogers, K., & Helfand, M. (2009). C-reactive protein as a risk factor for coronary heart disease: A systematic review and meta-analyses for the U.S. Preventive Services Task Force. Annals of Internal Medicine, 151(7), 483–495.

Buikstra, Jane E., and Douglas H. Ubelaker. 1994. "Standards for data collection from human skeletal remains: proceedings of a seminar at the Field Museum of Natural History." Arkansas Archeological Survey Research Series No. 44.

Buonasera, T., Eerkens, J., de Flamingh, A., Engbring, L., Yip, J., Li, H., Haas, R., DiGiuseppe, D., Grant, D., Salemi, M., Nijmeh, C., Arellano, M., Leventhal, A., Phinney, B., Byrd, B. F., Malhi, R. S., & Parker, G. (2020). A comparison of proteomic, genomic, and osteological methods of archaeological sex estimation. *Scientific Reports*, *10*(1), 11897-.

Buonasera, T., Eerkens, J., Malarchik, D., Panich, L. M., Canzonieri, C., Zimmer, C., Clough, C., Ostrander, T., Sutton, A., Salemi, M., & Parker, G. (2024). Immune proteins recovered in tooth enamel as a biochemical record of health in past populations: Paleoproteomic analysis of Mission Period Native Californians. *Journal of Archaeological Science*, *171*.

Castiblanco, G. A., Rutishauser, D., Ilag, L. L., Martignon, S., Castellanos, J. E., & Mejía, W. (2015). Identification of proteins from human permanent erupted enamel. *European Journal of Oral Sciences*, *123*(6), 390–395.

Charlton, S., Ramsøe, A., Collins, M., Craig, O. E., Fischer, R., Alexander, M., & Speller, C. F. (2019). New insights into Neolithic milk consumption through proteomic analysis of dental calculus. *Archaeological and Anthropological Sciences*, *11*(11), 6183–6196.

Chiang, J. J., Park, H., Almeida, D. M., Bower, J. E., Cole, S. W., Irwin, M. R., McCreath, H., Seeman, T. E., Fuligni, A. J., & Freedland, K. E. (2019). Psychosocial Stress and C-Reactive Protein From Mid-Adolescence to Young Adulthood. Health Psychology, 38(3), 259–267.

Clements, V., & Gruspier, K. (n.d). An analysis of skeletal material from the rock-cut church at Gurat. Unpublished report

Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, *26*(12), 1367–1372.

Cundiff, J. M., Bennett, A., Williams, A., Cushman, M., Howard, V. J., & Ruiz, J. M. (2024). Association Between Psychosocial Factors and C-Reactive Protein Across Income, Race, and Sex. Health Psychology, 43(1), 7–18.

Duverger, O., & Lee, J. S. (2024). The intricacies of tooth enamel: Embryonic origin, development and human genetics. *Journal of Structural Biology*, *216*(4), Article 108135.

Fabrizi, I., Flament, S., Delhon, C., Gourichon, L., Vuillien, M., Oueslati, T., Auguste, P., Rolando, C., & Bray, F. (2024). Low-Invasive Sampling Method with Tape-Disc Sampling for the Taxonomic Identification of Archeological and Paleontological Bones by Proteomics. *Journal of Proteome Research*, *23*(8), 3404–3417.

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., & Whitehouse, C. M. (1989). Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science (American Association for the Advancement of Science)*, *246*(4926), 64–71.

Fiddyment, S., Holsinger, B., Ruzzier, C., Devine, A., Binois, A., Albarella, U., Fischer, R., Nichols, E., Curtis, A., Cheese, E., Teasdale, M. D., Checkley-Scott, C., Milner, S. J., Rudy, K. M., Johnson, E. J., Vnouček, J., Garrison, M., McGrory, S., Bradley, D. G., & Collins, M. J. (2015). Animal origin of 13th-century uterine vellum revealed using noninvasive

peptide fingerprinting. *Proceedings of the National Academy of Sciences - PNAS*, *112*(49), 15066–15071

Fincham, A. G., Moradian-Oldak, J., & Simmer, J. P. (1999). The Structural Biology of the Developing Dental Enamel Matrix. *Journal of Structural Biology*, *126*(3), 270–299.

Franklin, U. M., & Gervers, M. (1978). The technical analysis of coinage from the medieval rock-cut church at Gurat, France. Journal of the International Institute for Conservation, 3(2), 2–6

Gervers, M. (1967). The cave church at Gurat (Charente): Preliminary report. Gesta, 6, 10–20.

Gervers, M., & Gervers, V. (1975). Cliffs, caves and hermits: The rock-cut church at Gurat. Rotunda VIII: 5–13.

Gervers, M., & Lengyel, I. (1976). *Église rupestre de Gurat : compte rendu de fouille, 1974.* Sociétế archéologique et historique de la Charente.

Gil-Bona, A., & Bidlack, F. B. (2020). Tooth Enamel and its Dynamic Protein Matrix. *International Journal of Molecular Sciences*, *21*(12), 4458-.

Goodman, A. H., & Leatherman, T. L. (1998). Traversing the chasm between biology and culture: An introduction. In A. L. Goodman & T. L. Leatherman (Eds.), Building a new biocultural synthesis: Political-economic perspectives on human biology (pp. 3–41). The University of Michigan Press

Google. (2). *Google Earth* [Satellite imagery]. https://earth.google.com/

Graham-Engeland, J. E., Sin, N. L., Smyth, J. M., Jones, D. R., Knight, E. L., Sliwinski, M. J., Almeida, D. M., Katz, M. J., Lipton, R. B., & Engeland, C. G. (2018). Negative and

positive affect as predictors of inflammation: Timing matters. *Brain, Behavior, and Immunity*, *74*, 222–230.

Greco, E., Gennaro, A. M., Piombino-Mascali, D., Costanzo, D., Accardo, S., Licen, S., Barbieri, P., Fornasaro, S., Semeraro, S., Marin, E., Signoretti, S., Gabriele, C., & Gaspari, M. (2023). Dental proteomic analyses and Raman spectroscopy for the estimation of the biological sex and age of human remains from the Greek cemetery of San Giorgio Extra, Reggio Calabria (Italy). Microchemical Journal, 195, 109472-.

Gregoricka, L. A. (2023). The ethics of excavating: bioarchaeology and the case for rehabilitating legacy human skeletal collections in the Near East. *Levant (London)*, *55*(3), 294–303.

Gretenhart, C.E. n.d. Skeletal Remains From a Medieval Cemetery, Gurat, Charente. Draft.

Hancock, R. et.al. 1987. Are Archaeological Bones Similar to Modern Bones? An INAA Assessment. Journal of Radioanalytical Analysis and Nuclear Chemistry. 110.. 283-291

Hansen, J., Dekker, J., Troché, G., Fagernäs, Z., Olsen, J. V., Saña Seguí, M., & Welker, F. (2024). A comparative study of commercially available, minimally invasive, sampling methods on Early Neolithic humeri analysed via palaeoproteomics. *Journal of Archaeological Science*, *167*, 106002

Harris, H.A. 1931. Lines of arrested growth in the long bones in childhood: the correlation of histological and radiographic appearances in clinical and experimental conditions. The British Journal of Radiology. 18. 622-40.

Hendy, J., Welker, F., Demarchi, B., Speller, C., Warinner, C., & Collins, M. J. (2018). A guide to ancient protein studies. *Nature Ecology & Evolution*, *2*(5), 791–799.

Hendy, J., van Doorn, N., & Collins, M. (2019). Proteomics. In *Archaeological Science* (pp. 35–69)

Hendy, J. (2021). Ancient protein analysis in archaeology. *Science Advances*, *7*(3).

Hillson, S. (2024). *Dental anthropology* (Second edition.). Cambridge University Press.

Jágr, M., Ergang, P., Pataridis, S., Kolrosová, M., Bartoš, M., & Mikšík, I. (2019). Proteomic analysis of dentin–enamel junction and adjacent protein-containing enamel matrix layer of healthy human molar teeth. *European Journal of Oral Sciences*, *127*(2), 112–121.

Kendall, R., Hendy, J., Collins, M. J., Millard, A. R., & Gowland, R. L. (2016). Poor preservation of antibodies in archaeological human bone and dentine. *Science and Technology of Archaeological Research*, *2*(1), 15–24.

Kendall, C., Eriksen, A. M. H., Kontopoulos, I., Collins, M. J., & Turner-Walker, G. (2018). Diagenesis of archaeological bone and tooth. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *491*, 21–37

Koenig, C., Bortel, P., Paterson, R. S., Rendl, B., Madupe, P. P., Troché, G. B., Hermann, N. V., Martínez de Pinillos, M., Martinón-Torres, M., Mularczyk, S., Schjellerup Jørkov, M. L., Gerner, C., Kanz, F., Martinez-Val, A., Cappellini, E., & Olsen, J. V. (2024). Automated High-Throughput Biological Sex Identification from Archeological Human Dental Enamel Using Targeted Proteomics. *Journal of Proteome Research*, *23*(11), 5107–5121.

Köhler-Forsberg, O., Buttenschøn, H. N., Tansey, K. E., Maier, W., Hauser, J., Dernovsek, M. Z., Henigsberg, N., Souery, D., Farmer, A., Rietschel, M., McGuffin, P., Aitchison, K. J., Uher, R., & Mors, O. (2017). Association between C-reactive protein (CRP) with depression

symptom severity and specific depressive symptoms in major depression. Brain, Behavior, and Immunity, 62, 344–350.

Kohn, M. J., Schoeninger, M. J., & Barker, W. W. (1999). Altered states: effects of diagenesis on fossil tooth chemistry. *Geochimica et Cosmochimica Acta*, *63*(18), 2737–2747.

Kotli, P., Morgenstern, D., Nagar, Y., Katina, C., 'Adawi, Z., Arbiv, K., & Boaretto, E. (2025). Sexing remains of a Byzantine ascetic burial using enamel proteomics. *Journal of Archaeological Science, Reports*, *62*, 104972-.

Lambert, Joseph B., Sharon Vlasak, Carole B. Szpunar, and Jane E. Buikstra. 1985. "Bone diagenesis and dietary analysis." Journal of Human Evolution 14(5): 477- 482.

Lengyel, I. 1975. Paleoserology. Blood Typing with the fluorescent antibody method. Akademiai Kiado, Budapest.

Madupe, P. P., Munir, F., Dickinson, M., Taurozzi, A. J., Mackie, M., Tawane, M., Mollereau, C., Hlazo, N., Penkman, K., Schroeder, L., Zanolli, C., Olsen, J. V., Ackermann, R. R., & Cappellini, E. (2025). Results from an Australopithecus africanus dental enamel fragment confirm the potential of palaeoproteomics for South African Plio-Pleistocene fossil sites. *South African Journal of Science*, *121*(1–2), 66–73.

Margolis, H. C., Beniash, E., & Fowler, C. E. (2006). Role of Macromolecular Assembly of Enamel Matrix Proteins in Enamel Formation. *Journal of Dental Research*, *85*(9), 775–793.

Marsland, A. L., Walsh, C., Lockwood, K., & John-Henderson, N. A. (2017). The effects of acute psychological stress on circulating and stimulated inflammatory markers: A systematic review and meta-analysis. Brain, Behavior, and Immunity, 64, 208–219.

Martin, D. L., Harrod, R. P., & Pérez, V. R. (2013). The Practice of Bioarchaeology. In *Bioarchaeology* (pp. 1–21). Springer New York.

Meijer, J. 2018. Reconstructing the Origins of the Medieval Monastic Inhabitants of a Cave Church in Gurat, France, using Strontium Isotope Analysis. UW Space.

Meijer, J., Dolphin, A. E., Yakymchuk, C., & Gervers, M. (2019). Interpreting medieval mobility from burials at the rock-hewn church of St. Georges, Gurat (France): Insights from strontium isotope analysis of bones and teeth. *International Journal of Osteoarchaeology*, *29*(4), 574–583.

Merrick, T. (2025). Data normalization techniques in metabolomics: Best practices. *Analytical Techniques,* Metabolome Express.

Moradian-Oldak, J., & George, A. (2021). Biomineralization of Enamel and Dentin Mediated by Matrix Proteins. *Journal of Dental Research*, *100*(10), 1020–1029.

Multari, D. H., Ravishankar, P., Sullivan, G. J., Power, R. K., Lord, C., Fraser, J. A., & Haynes, P. A. (2022). Development of a novel minimally invasive sampling and analysis technique using skin sampling tape strips for bioarchaeological proteomics. *Journal of Archaeological Science*, *139*, 105548-.

Navetat, M. 2019. Église monolithe Gurat, Charente, Nouvelle-Aquitaine, Hadès, 2019

Navetat, M. 2021. Église monolithe Saint-Georges, Gurat, Charente, Nouvelle-Aquitaine, Hadès, 2021

Nishimura, T., & Tojo, H. (2014). Mass Spectrometry-Based Protein Sequencing Platforms. In G. Marko-Varga (Ed.), *Genomics and Proteomics for Clinical Discovery and Development* (Vol. 6, pp. 69–99). Springer Netherlands.

Oliveira, B. P., Buzalaf, M. A. R., Silva, N. C., Ventura, T. M. O., Toniolo, J., & Rodrigues, J. A. (2023). Proteomic profile of the acquired enamel pellicle of children with early childhood caries and caries-free children. *European Journal of Oral Sciences*, *131*(4),

Parker, G. J., Yip, J. M., Eerkens, J. W., Salemi, M., Durbin-Johnson, B., Kiesow, C., Haas, R., Buikstra, J. E., Klaus, H., Regan, L. A., Rocke, D. M., & Phinney, B. S. (2019). Sex estimation using sexually dimorphic amelogenin protein fragments in human enamel. Journal of Archaeological Science, 101, 169–180.

Procopio, N., Chamberlain, A. T., & Buckley, M. (2018). Exploring Biological and Geological Age-related Changes through Variations in Intra- and Intertooth Proteomes of Ancient Dentine. *Journal of Proteome Research*, *17*(3), 1000–1013.

Schrader, Sarah A. and Christina Torres-Rouff. 2021. "Embodying Bioarchaeology: Theory and Practice." In Theoretical Approaches in Bioarchaeology, edited by Colleen M. Cheverko, Julia R. Prince-Buitenhuys, and Mark Hubbe: 15-27. New York: Routledge.

Schroeter, E. R., DeHart, C. J., Schweitzer, M. H., Thomas, P. M., & Kelleher, N. L. (2016). Bone protein "extractomics": comparing the efficiency of bone protein extractions of Gallus gallus in tandem mass spectrometry, with an eye towards paleoproteomics. *PeerJ (San Francisco, CA)*, *4*, Article e2603.

Seymour, G. 2019. Reconstructing the Life Histories of the Individuals Buried in the Rock-cut Cave Church of St. Georges, in Gurat, France. UW Space.

Sharma, V., Rastogi, S., Kumar Bhati, K., Srinivasan, A., Roychoudhury, A., Nikolajeff, F., & Kumar, S. (2020). Mapping the Inorganic and Proteomic Differences among Different

Types of Human Teeth: A Preliminary Compositional Insight. *Biomolecules (Basel, Switzerland)*, *10*(11), 1540.

Shaw, B., McDonnell, T., Radley, E., Thomas, B., Smith, L., Davenport, C. A. L., Gonzalez, S., Rahman, A., & Layfield, R. (2023). Preservation of whole antibodies within ancient teeth. *iScience*, *26*(9), 107575.

Smith, C. E. (1998). Cellular and chemical events during enamel maturation. *Critical Reviews in Oral Biology and Medicine*, *9*(2), 128–161.

Steptoe, A., Hamer, M., & Chida, Y. (2007). The effects of acute psychological stress on circulating inflammatory factors in humans: A review and meta-analysis. Brain, Behavior, and Immunity, 21(7), 901–912.

Stewart, N. A., Molina, G. F., Mardegan Issa, J. P., Yates, N. A., Sosovicka, M., Vieira, A. R., Line, S. R. P., Montgomery, J., & Gerlach, R. F. (2016). The identification of peptides by nanoLC-MS/MS from human surface tooth enamel following a simple acid etch extraction. *RSC Advances*, *6*(66), 61673–61679

Tsou, C.-C., Avtonomov, D., Larsen, B., Tucholska, M., Choi, H., Gingras, A.-C., & Nesvizhskii, A. I. (2015). DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics. *Nature Methods*, *12*(3), 258–264.

Uchino, B. N., Ruiz, J. M., Smith, T. W., Smyth, J. M., Taylor, D. J., Allison, M., & Ahn, C. (2016). Ethnic/racial differences in the association between social support and levels of C-reactive proteins in the North Texas Heart Study. Psychophysiology, 53(1), 64–70

VanDuijn, M. M., Dekker, L. J. M., Zeneyedpour, L., Smitt, P. A. E. S., & Luider, T. M. (2010). Immune Responses Are Characterized by Specific Shared Immunoglobulin Peptides That

Can Be Detected by Proteomic Techniques[S]. *The Journal of Biological Chemistry*, *285*(38), 29247–29253.

Warinner, C., Korzow Richter, K., & Collins, M. J. (2022). Paleoproteomics. Chemical Reviews, 122(16), 13401–13446.

Walker, P. L., Bathurst, R. R., Richman, R., Gjerdrum, T., & Andrushko, V. A. (2009). The causes of porotic hyperostosis and cribra orbitalia: A reappraisal of the iron-deficiency-anemia hypothesis. *American Journal of Physical Anthropology*, *139*(2), 109–125.

Watson, J. T., & Sparkman, O. D. (2007). *Introduction to mass spectrometry : instrumentation, applications and strategies for data interpretation* (4th ed.). John Wiley & Sons.

Webber, C and Gretenhart, C.E. n.d. Unpublished preliminary report on bone density. McMaster University.

Wei, X., & Cooper, D. M. L. (2023). The various meanings and uses of bone "remodeling" in biological anthropology: A review. *American Journal of Biological Anthropology*, *182*(2), 318–329.

Wood, J. W., Milner, G. R., Harpending, H. C., & Weiss, K. M. (1992). *The osteological paradox: Problems of inferring prehistoric health from skeletal samples.* Current Anthropology, 33(4), 343–370.

Ziganshin, R. H., Berezina, N. Y., Alexandrov, P. L., Ryabinin, V. V., & Buzhilova, A. P. (2020). Optimization of Method for Human Sex Determination Using Peptidome Analysis of Teeth Enamel from Teeth of Different Biological Generation, Archeological Age, and Degrees of Taphonomic Preservation. *Biochemistry. Biokhimiia*, *85*(5), 614–622.

Zuckerman, M. K., Tribble, A. G., Austin, R. M., DeGaglia, C. M. S., & Emery, T. (2023). Biocultural perspectives on bioarchaeological and paleopathological evidence of past pandemics. American Journal of Biological Anthropology, 182(4), 557–582.