UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Head Office: Università degli Studi di Padova

Department of Geosciences

Ph.D. COURSE IN GEOSCIENCES

SERIES XXXIV

ADVANCED ANALYTICAL DIAGNOSTICS APPLIED TO HUMAN OSTEOLOGICAL REMAINS

Thesis written with the financial contribution of Fondazione CARIPARO

Coordinator: Ch.ma Prof.ssa Claudia Agnini

Supervisor: Ch.mo Prof. Gilberto Artioli

Co-Supervisor: Prof. Luca Pagani, Dr. Luca Nodari

Ph.D. student: Cinzia Scaggion

ABSTRACT

Ancient bones, recovered from archaeological contexts and preserved in Museums, represent a valuable source of information on health, diet and mobility of ancient populations as well ason demographics and environmental conditions of the past. Because of the development of modern technologies of omic sciences, osteological finds are increasingly requested and this has led to an increase in the analysis of ancient DNA (aDNA). Sampling methods for ancient DNA extraction are predominantly destructive and may often compromise the osteological findings for future analysis or for studies in other research fields. Besides the invasive and destructive sampling, in the case of poor conservation conditions of the archaeological bone caused by taphonomic and diagenetic alterations, the sequencing of ancient DNA can be an extremely expensive operation.

Given the high costs of the aDNA sequencing procedure, in this research an analytical study based on infrared spectroscopy (FTIR) was conducted to develop a reliable, fast and inexpensive pre-screening method to determine presence/absence of genetic molecules in archaeological bonesamples. Infrared spectroscopy is a useful tool that is fast, minimally destructive, inexpensive and sensitive to changesin the structural properties of the organic (collagen) and inorganic (bioapatite nanocrystals) components that make up the bone. At the ultrastructural level, the organic and inorganic components ofbone may stabilize strong bounds with DNA, stabilizing it and determining its survival over time.

The sensitivity and efficiency of new IR parameters was tested on fresh bones and extremely altered archaeological samples, characterized by different chronology and origin. Bones that underwent diagenetic processes were characterized taking into account changes in climatic-environmental and burial conditions. The research was expanded by examining the changes induced by diagenesis on the secondary structure of preserved collagen, as well as by evaluating the changes on bioapatite crystals.

The obtained results demonstrate that the IR parameter used in this research, that describes the atomic order/disorder, is advantageous for monitoring minimal changes in the structure and chemical properties of bioapatite as well as indirectly of collagen. This method based on infrared spectra may improve the selection process of bone samples as well as their suitability for specific analyses, e.g. genetic, paleo-proteomicand stable isotope analysis. Here, I propose a functional predictive model based on infrared parameters, used to determine the most predictive parameter for the presence/absence of DNA and allowing to reduce the costs of genetic analyses. Theobtained results, shows that the quality of aDNA cannot be determined due to the influenceof local environmental factors.

SOMMARIO

Le ossa antiche, recuperate dai contesti archeologici e preservati all'interno dei Musei, rappresentano una preziosa fonte di informazioni sull'alimentazione, lo stato di salute, la mobilità delle popolazioni antiche nonché sulla demografia e condizioni ambientali del passato, utili a ricercatori e accademici. A seguito dello sviluppo di moderne tecnologie delle scienze omiche, i reperti osteologici sono sempre più richiesti e questo ha comportato un aumento dell'analisi del DNA antico (aDNA). I metodi di campionamento per l'estrazione del DNA antico sono prevalentemente distruttivi e spesso possono compromettere i reperti osteologici per ulteriori future analisi o per studi in altri campi di ricerca. Oltre al campionamento invasivo e distruttivo, in condizioni di scarsa conservazione dell'osso archeologico causata da alterazioni tafonomiche e diagenetiche, il sequenziamento del DNA antico può essere un'operazione estremamente costosa.

Dati gli elevati costi della procedura di sequenziamento dell'aDNA, in questo lavoro di ricerca è stato condotto uno studio analitico mediante spettroscopia a raggi infrarossi (FTIR) per sviluppare un metodo di pre-screening affidabile, veloce ed economico per determinare la presenza/assenza di molecole genetiche in un campione osseo archeologico.

La spettroscopia IR è uno strumento utile in quanto è rapida, minimamente distruttiva, economica e sensibile alle variazioni delle proprietà strutturali delle componenti organiche (collagene) e inorganiche (nano cristalli di bioapatite) che costituiscono l'osso. A livello ultrastrutturale, le componenti organiche e inorganiche possono stabilire forti legami con il DNA , stabilizzandolo e determinando la sua sopravvivenza nel tempo. Da campioni archeologici (di epoche e provenienze diverse) estremamente alterati a moderne ossa fresche, abbiamo valutato la sensibilità e l'efficacia di nuovi parametri IR per caratterizzare la diagenesi subita dalle ossa tenendo in considerazioni i cambiamenti delle condizioni climatico–ambientali e di seppellimento. Il lavoro è stato esteso per esaminare le modificazioni indotte dalla diagenesi sulla struttura secondaria del collagene conservato, valutandone gli effetti sui cristalli di bioapatite. I risultati ottenuti dimostrano che il parametro IR che descrive l'ordine/disordine atomico, utilizzato in questa ricerca, è vantaggioso per il monitoraggio di variazioni minime nella struttura e nelle proprietà chimiche della bioapatite nonché indirettamente nel collagene. Questo metodo potrebbe migliorare il processo di selezione dei campioni ossei nonché la loro idoneità per analisi specifiche, ad es. analisi genetiche, paleoproteomiche e degli isotopi stabili sulla base delle analisi spettrali. Viene qui proposto inoltre un modello predittivo funzionale con i parametri infrarossi utilizzati, al fine di determinare il parametro più predittivo per la prensenza/assenza di DNA, utile per ridurre i costi delle analisi genetiche. Dai dati ottenuti, la qualità/quantità di aDNA risulterebbe non essere determinabile a causa dell'influenza di fattori ambientali locali.

This essay is organized in five main chapters, treating different aspects of this research project.
Each chapter is formatted as a paper ready to be submitted to an international journal or already published.

TABLE OF CONTENTS

This essay is organized in five main chapters, treating different aspects of this research project. Each chapter is formatted as a paper ready to be submitted to an international journal or already published.

3. Results
3.1. Fourier Transform-IR spectroscopy data
3.2. Spectral analysis for bone characterization

4. Discussion

5. Conclusions

References

CHAPTER 4- Infrared spectroscopy as a human bone pre-screening method for molecular and isotope analysis

1. Introduction

2. Material and methods
2.1. Bone specimens
2.2. Micro-sampling for Infrared analysis
2.3. Sample preparation and IR spectra collected
2.3.1. Bone spectra analyses
2.3.2. Lyophilized bone collagen spectra analyses
2.4. Bone collagen extraction and isotopic analysis
2.5. Ancient DNA extraction

3. Results
3.1. Measurements of the infrared bone spectrum
3.2. FTIR measurement of bone collagen and quality control (QC)
3.3 Infrared data samples and DNA yields

4. Discussion

5. Conclusions

References

CHAPTER 5- (a case of study) **Life and death of Saint Theobald (1033-1066 AD) from chemical and forensic-anthropological analysis**

1. Introduction

2. Methods
2.1 Age determination
2.2 Sex determination
2.3 Palaeopathological methodology
2.4 Microsampling

3. Results and Discussion
3.1 Infrared Spectroscopy: sample preparation and spectrum analysis
3.2 Photogrammetric acquisition and digital modelling

4. Conclusions

References
Supplementary Information

This essay is organized in five main chapters, treating different aspects of this research project. Each chapter is formatted as a paper ready to be submitted to an international journal or already published.

CONCLUSIONS AND FUTURE PERSPECTIVES

REFERENCES

APPENDIX

Aneli, S., Saupe, T., Montinaro, F., Solnik, A., Molinaro, L., **Scaggion, C.**, Carrara, N., Raveane, A., Kivisild, T., Metspalu, M., Scheib, C.L., Pagani, L., 2022. The Genetic Origin of Daunians and the Pan- Mediterranean Southern Italian Iron Age Context. Mol. Biol. Evol. 39, 1–16. https://doi.org/10.1093/molbev/msac014

Saupe, T., Montinaro, F., **Scaggion, C.**, Carrara, N., Kivisild, T., D'Atanasio, E., Hui, R., Solnik, A., Lebrasseur, O., Larson, G., Alessandri, L., Arienzo, I., De Angelis, F., Rolfo, M.F., Skeates, R., Silvestri, L., Beckett, J., Talamo, S., Dolfini, A., Miari, M., Metspalu, M., Benazzi, S., Capelli, C., Pagani, L., Scheib, C.L., 2021. Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula. Curr. Biol. 31, 2576-2591.e12. https://doi.org/10.1016/j.cub.2021.04.022

INTRODUCTION

The animal and human skeletal tissues found in archaeological excavations play an important role in their eminent research function. They are often the only direct evidence of the past and hold valuable information useful to academics and researchers. Indeed, bones and teeth are useful biological archives (morphological, biochemical and histological) for archaeological and geological documentation providing information, regarding not only the individual whose remains are examined, but obtaining important information on paleopathology, paleoclimate conditions, migrations, evolution (Allende and Samplonius, 2022; Dupras and Schwarcz, 2001; Karakostis et al., 2021; Kieser et al., 2007; Kimmerle et al., 2008; Mays et al., 2001; Minozzi et al., 2012; Prowse et al., 2004; Raichlen et al., 2020).

These osteological finds of historical, archaeological, scientific and cultural interest are part of the so-called "biological assets", commonly preserved in museums, and are therefore important and protected in the context of the Conservation of Cultural Heritage (Licata et al., 2020).

Following the development of modern technologies of omic sciences, the investigation of osteological finds in museums is increasingly requested (Brown and Brown, 2011).

Among these, in the last decade, the advent of new technologies such as Next Generation Sequencing (NGS) (Behjati and Tarpey, 2013), has led to an increase in the analysis of ancient DNA (aDNA) (Der Sarkissian et al., 2015; Slatkin and Racimo, 2016) as it has given way to overcome some of the previous limitations concerning the study of ancient molecules such as: fragmentation, low number of copies, damage to the sequence and contamination. The improvement of the sequencing technique has made it possible to analyse genomes extracted from ancient bone tissues, providing a fundamental contribution to anthropological studies, reconstructing the diachronic and synchronic natural history of the organisms subjected to analysis (Slatkin and Racimo, 2016). This led to an inevitable destruction of bone samples as a consequence of often-invasive sampling. Sampling methods for ancient DNA extraction are predominantly destructive causing the loss of the entire bone portion through the drilling and cutting of the specimens, powdering or coring the bone findings (Mays et al., 2013; Orlando et al., 2021; Pálsdóttir et al., 2019).

To overcome this problem, several ancient DNA extraction protocols have recently been developed with the aim of safeguarding the most requested and sampled finds such as petrous bones and tooth roots, which are known to produce particular high concentration of good quality endogenous DNA (aDNA) (Higgins and Austin, 2013; Pinhasi et al., 2019). A forensic study by Misner et al., 2009 found that a significant factor influencing on the quality and quantity of extracted DNA is the type of bone. A very compact type of bone that is positively related to the preservation of bio-molecules over

time is the petrous bone (Gamba et al., 2014; Hansen et al., 2017; Parker et al., 2020; Pinhasi et al., 2019).

The petrous bone is a portion of the temporal bone of the skull, where the inner ear is located, surrounded by a particularly compact bone capsule.

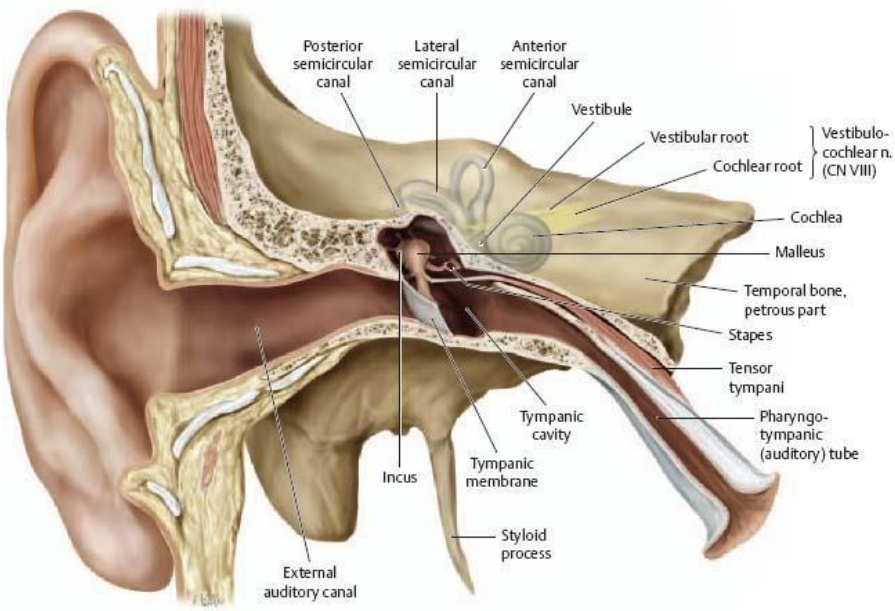

Figure 1. *Ear: Overview. Coronal section through right ear, anterior view. Imagine from* Gilroy et al., 2008.

It is divided into two portions: bone labyrinth and membranous labyrinth. The portion that survives and is the object of study is the bone labyrinth (*osseous labyrinth*) consisting of a complicated system of cavities dug into the thickness of the bone thunderstorm. It is composed of a central part called the vestibule, with three semicircular canals (upper or anterior, lateral or horizontal, and posterior, arranged on three floors perpendicular to each other) and from the spiral canal of the *cochlea* (Carlson, 2019; Wang et al., 2019).

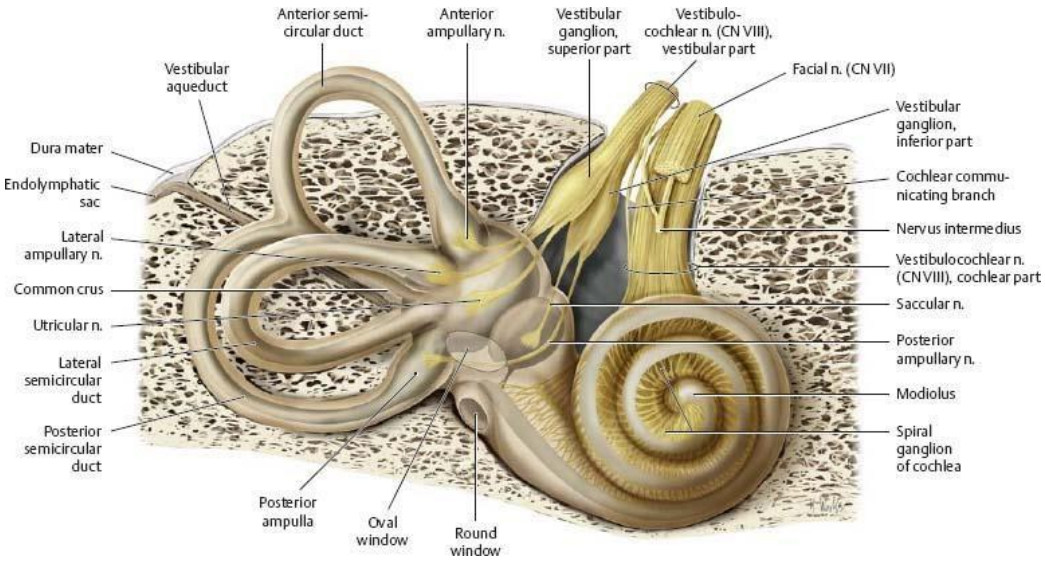

Figure 2. *Vestibular and auditory apparatus. The osseous labyrinth and its bony shell (cochlea). Imagine from* Gilroy et al., 2008

Gamba et al., 2014 found that the rates of endogenous DNA extracted from the petrous bones are 4 to 16-fold higher than those from the teeth, and up to 183-fold higher than those from other anatomic districts. The work of Gamba (2014), suggested that the densest part of the petrous bone reduces bacterial and chemically mediated *post-mortem* DNA decay.

Pinhasi et al., (2015) examined differences in the endogenous DNA yield within the petrous bone analizing three different parts of the bone: A) a spongy part, B) dense white bone surrounding the inner ear, and C) dense bone of the inner ear. High aDNA yields were obtained from the densest part of the petrous bone, that is the cochlea (part C), as well as samples from warm regions (either arid or humid) were less positive for the bonded conservation of the sample itself. In addition to petrous bones, also teeth are a useful skeletal source of DNA because of their unique chemical-physical composition and, being located in *alveoli* of mandible and maxilla, they are more protected from the environmental burial conditions that accelerate DNA decay (Garcìa et al., 1996). DNA extracted from teeth is often of higher quality (Cafiero et al., 2019) and is less prone to contamination than DNA extracted from bones (Gilbert et al., 2005). Human teeth are made up of the crown and the roots, which are encased in the mandibular and maxillary alveolar bone. Teeth roots are composed by cementum, dentine and pulp and these parts produce more aDNA yields than the crown, mainly made up of enamel (Gaytmenn and Sweet, 2003; Higgins et al., 2011). Enamel is far more mineralized than the other teeth structures, containing no organic components (collagen), and serves to protect the dentine and pulp. Furthermore, once formed, it is acellular and does not contain aDNA (Lacruz et al., 2017). Studies conducted by Gaytmenn and Sweet (Gaytmenn and Sweet, 2003) showed that the yield of DNA from the crown (containing layers of pulp and dentine with which it is intimately connected) is lower than the yields of DNA produced by the roots. This mineralyzed tissue provides a physical barrier protecting the cells within the tooth from external conditions such as microbes and environmental contaminants (Lacruz et al., 2017).

The pulp and dentine make up the majority of tooth and are rich in cells as odontoblasts, fibroblasts, plasma cells, nerve cells and undifferentiated mesenchymal cells (Lymperi et al., 2013). The high cell content in the pulp makes this tissue a very rich source of DNA in the teeth (Malaver and Yunis, 2003). However, this type of tissue is not always recoverable or in good condition as it may be absent in the teeth of old and/or sick individuals (Higgins and Austin, 2013). Dentine is a poor source of nuclear DNA compared to cementum (Higgins et al., 2013) as the latter is rich in cementocytes (Bosshardt and Selvig, 1997; Trammell and Kroman, 2013).

To reduce the destructive impact of DNA extraction sampling, numerous studies were performed to design a minimally invasive protocol for osteological findings while safeguarding their

morphological integrity useful for morphometric studies (Davies et al., 2019; Kolobova et al., 2020; Krenn et al., 2019; Kuzminsky et al., 2018; Müller et al., 1998).

Sirak et al., 2017 developed a method to sample complete crania without removing the *calvaria* (skullcap) or undermining its integrity in that the petrous is located inside the skull but is partially accessible of the external cranial base. The petrous pyramid (the spongiest portion of the petrous bone) is visible from the base of the skull while the *osseous labyrinth*, useful for genetic analysis, is arranged deeper and is not accessible without the removal of cranial elements. The method used to directly sample the *osseous labyrinth* is to lower a small bone crest between the carotid canal and the jugular foramen, thus creating a direct passage for drilling and sampling of *cochlea*. Another more recently developed and less invasive sampling method (also for incomplete crania with disarticulated petrous bones) concerns the sampling of the auricular ossicles (malleus, incus, and stapes), located inside the petrous bone and visible within the external auditory meatus (Pinhasi et al., 2019; Sirak et al., 2020). Ancient DNA sampling methods for teeth are destructive and often result in the loss of at least one entire root. Harney and colleagues developed a minimally destructive method for extracting ancient DNA from dental cementum present on the surface of tooth roots that does not require destructive procedures. The method consists of subjecting the tooth roots to surface disinfection using a 2% bleach solution and rinsing with 95% ethanol, followed by UV irradiation (Harney et al., 2021). Despite the numerous measures aimed at preserving osteological samples, the analysis of the aDNA remains a problem. Decoding ancient high-resolution DNA sequences is not simple: its survival is unpredictable and its extraction and analysis, as well as being expensive, is time-consuming and often fails.

When the individual dies, the genome is exposed to the effects of numerous processes that occur between the death and its extraction in the laboratory, as well as from the diagenetic processes that threaten its stability, leading to deterioration. These factors include intracellular nucleases which, being no longer inside the cell, can access and degrade DNA, as well as microorganisms that spread in decaying tissues (Dabney et al., 2013). The DNA molecule undergoes oxidative damage, depurination or hydrolytic damage.

Hydrolysis deprives the DNA of a nitrogenous base, creating an abasic site (AP site) that completely destabilizes the molecule until it breaks in a short time (Latham and Miller, 2019). This hydrolytic reaction can also affect the phosphodiester bond of the sugar-phosphate backbone causing the formation of a single-stranded nick in the double helix (that is a discontinuity in a double stranded DNA molecule) which destabilizes it, leading to its breaking (Lindahl, 1993) or can affects the amino groups of the bases leading to deamination of cytosine and guanine (Hofreiter et al., 2001; Poinar, 2002).

The presence of oxygen contributes to the diagenesis of DNA and causes its fragmentation. The oxidation of pyrimidines (such as cytosine and thymine) by free radicals (O_2, H_2O_2 and OH) can lead to the rupture of the double helix (Lindahl, 1993). Free radicals are also produced by microbial attacks by fungi and bacteria or ionizing radiation (Poinar, 2002). Temperature and pH play an important role in the deterioration of DNA. These factors influence oxidation and hydrolysis reactions and their impact on the nucleotide sequence. The extent of DNA fragmentation is strongly influenced by environmental conditions. More specifically, low temperatures inhibit enzymatic activity, thus preserving DNA better while high temperatures (in addition to promoting colonization by microorganisms) weaken the intermolecular forces (hydrogen bond and Van der Waals) damaging the integrity of the biomolecule (Dabney et al., 2013; Gilbert et al., 2007; Handt et al., 1994; Latham and Miller, 2019; Mitchell et al., 2005; Rollo et al., 2002). A basic pH increases the fragmentation of the polynucleotide chain (Lindahl, 1993) while an acid pH increases its de-amination (Mitchell et al., 2005).

Extensive research was done on osteological samples to determine where DNA is located and how it survives over time.

The structure of bone

Petrous bones and teeth are calcified tissue characterized by a complex hierarchical structure (divided into seven hierarchical levels) from the macro to the nano-scale and mainly consists of mineral and organic phases and water (Fratzl et al., 2004; Reiche and Gourrier, 2016; Reznikov et al., 2018; Weiner, 2010).

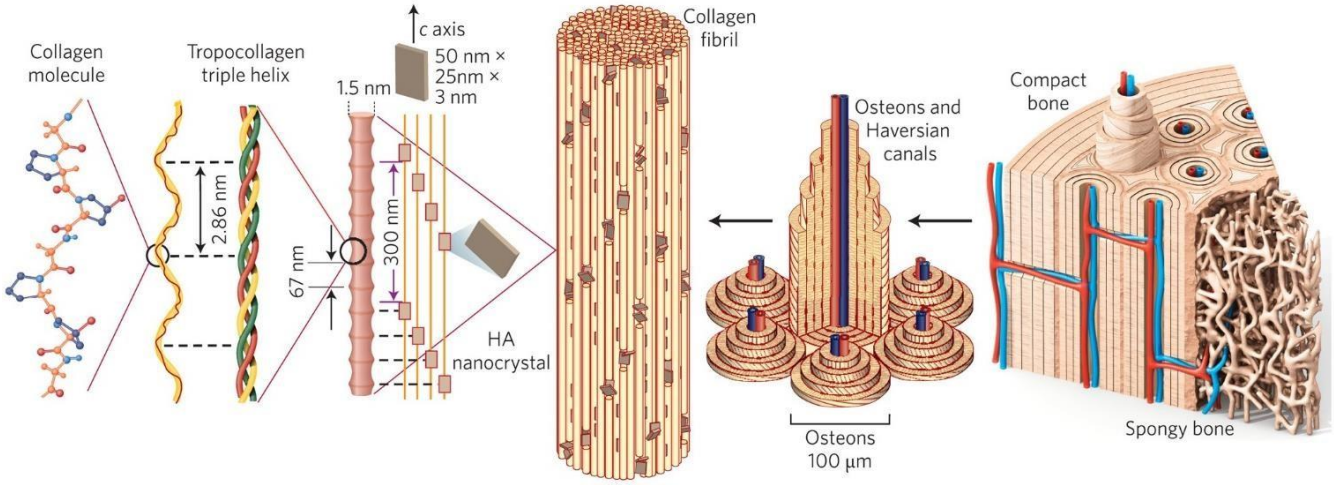

Figure 3. *Hierarchical structure of bone. Imagine from* Wegst et al., 2015.

The inorganic part of bone

The inorganic component, named bioapatite (BAp), is a nanocrystalline and defective carbonated-hydroxylapatite with a chemical composition significantly departing from that of hydroxylapatite

$Ca_5(PO_4)_3(OH)$, having hexagonal crystallographic symmetry with $P6_3/m$ space group, and constitutes about 60-70% of bone (Weiner, 2010). The bioapatite crystals appear as thin needles with extremely small dimensions (16-50 nm long, 16-50 nm wide and 2-5 nm thick) (Eppell et al., 2001) and are intimately associated with the structure of the organic component (organic matrix), which acts as a template for deposition and crystallization of minerals (Reiche and Gourrier, 2016; Weiner and Traub, 1992).

Biogenic apatite crystals are characterized by a disordered structure at the atomic scale due to many ionic substitutions, the most relevant of which is the tetrahedral phosphate ions substituted by planar carbonates. Moreover, because of the nano-size of crystals, bioapatite is characterized by a very high surface-to-bulk ratio (Asscher et al., 2012; Boskey and Pleshko Camacho, 2007; Weiner, 2010). The inherent disorder of the mature mineral phase induces a higher solubility of crystallites with respect to the geogenic hydroxylapatite. The bioapatite recrystallization feasibly occurs after the death of the organism by *Ostwald ripening* mechanisms (Ostwald, 1885). Hence, this suggest that with the time crystal size increase and increase its long-rage order, corresponding to a more stable phase (Berna et al., 2004; Weiner and Traub, 1992).

The organic part of bone

The organic phase is composed of bone tissue cells, such as osteoprogenitor cells, osteoblasts, osteocytes and osteoclasts that provide for the growth, production and reabsorption of bone tissue during the life of individual, and extracellular matrix made up of different glycoproteins, proteoglycans, various lipids and collagen (Bosman and Stamenkovic, 2003). The extracellular matrix acts as a scaffold for tissues and organs throughout the body, playing an essential role in their structural and functional integrity (Hynes, 2009). Its most abundant component is collagen of Type I, a fundamental fibrous protein, which constitutes 30% of all proteins present in the body (Veis, 2009).

Collagen type I protein has a conformation divided in *primary structure* and *secondary structure* (Ottani et al., 2002). Collagen building blocks are amino acids; small organic molecules that consist of an alpha carbon (C_α) atom linked to an amino group, a carboxyl group, a hydrogen atom and a side chain, that is a variable component (side chains) (Lawrence, 2012). Amino acids are linked together by peptide bonds, forming a repetitive structure of a tripeptide unit, (Gly-X-Pro)n; (Gly-X-Hypro)n; (Gly-Pro-Hypro)n, where X position is one of the twenty essential amino acids, while other positions are often occupied by proline and its post-translational modification in 4-hydroxyproline (Bella et al., 2006).

These amino acid sequences lead to the formation of precise chains, called α1(I), α2(I) (Garnero, 2015; Retief et al., 1985).

The constitutive fundamental unit of collagen is called tropocollagen (which determines its *secondary structure*) and is composed of three left-oriented polypeptide helices (α-chains of *primary structure* that form the backbone of the α-helix), bound together to form a right-handed ternary superhelix (triple helix, organized in two-dimensional structure) (Exposito et al., 2010; Weiner, 2010), while the side chains of amino acids exposed on the outside (Wess, 2005). Tropocollagen triple helices are about 300 nanometres long with a diameter of about 1.5 nanometres (Tzaphlidou and Berillis, 2005; Weiner, 2010) and are held together by cross-links and out of phase (¾ of its length) with a gap of 40 nanometres (Weiner, 2010). Hydrogen bonding between amino groups (N−H) and carboxyl groups (C=O) in neighbouring regions of the protein chain (up to distances of about 0.3 nm) causes certain patterns of folding known as a α-helices (alpha-helix) and β-sheet (beta-sheet) to occur (Greenfield, 2006). These stable folding patterns make up the *secondary structure* of a protein.

Each collagen micro-fibril constitutes a collagen fibril with a diameter of approximately 0.2 and 0.5 μm (Salz et al., 2006). In turn, these fibrils line up in even larger ordered matrices called fibers, which vary in size (Kadler et al., 1996).

In the structure of the bone system, collagen is intimately connected with the bioapatite crystals nucleated within the gap region, filling the space within and between the micro-fibrils during mineralization process (Weiner and Traub, 1992). The crystals grow with the c axis oriented parallel to the axis of the collagen fibril, within which they are located (Amornkitbamrung et al., 2022). In the hierarchical structure of the bone, the mineralized collagen fibrils are organized in arrays (~ 10-50 μm), where the protein matrix acts as a glue and the fibers are arranged with angular discontinuity (Colaço et al., 2019). Several layers of fibrils, each with its own orientation, assemble to form *lamellae* (~ 3-7 μm thick), long bones or *solid reticulum* in short, flat bones (Eliaz and Metoki, 2017). These *lamellae* are arranged concentrically around a vascular canal, also called Haversian canals, to form osteons (structural elements with a diameter of about 200-300 μm and a length of 1-2 cm), characteristic of the microstructure of the cortical bone (Kahla and Barkaoui, 2021). Transversely or obliquely to Havers' canals are Wolkmann's canals, containing small blood capillaries (Kim et al., 2015). The outer layer of each osteon is called a cement line and is continuously reformed thanks to the bone remodelling process, which occurs thanks to the metabolic activity operated by the blood flow of the vascular channels and *canaliculi* (gap junctions) (Kahla and Barkaoui, 2021; McGraw, 2018). In addition to the concentric organization, in the bone microstructure there are also interstitial *lamellae*, which fill the areas between multiple osteons (McGraw, 2018). At the macroscopic level, bone can be divided into cortical (compact bone) and cancellous (lamellar bone) (Ott, 2018). The different microstructure and density of these types of bone determines the different mechanical properties and consequently the different functions (Morgan et al., 2018). Both types of bone are

present in the human body, in different areas, depending on the mechanical properties required. Inside the *lacunae* in the bone matrix, cellular elements, such as osteocytes, remain trapped (Florencio-Silva et al., 2015). Osteocytes are irregularly shaped cells, with an evident nucleus and a cytoplasm that has several extensions (Mohamed, 2008). Through the *canaliculi*, the cytoplasmic extensions of different osteocytes make contact with each other, thus allowing metabolic exchanges between the osteocytes (Mohamed, 2008).

In teeth roots, cementocytes are connected by *canaliculi* that join periodontal ligament (Higgins and Austin, 2013). Cement, similar in physico-chemical composition to bone, has a different structure and functionality as it is avascular and it contains fewer inorganic salts (Bosshardt and Selvig, 1997; Higgins and Austin, 2013).

Alteration of bone

Archaeological bones undergo physico-chemical alterations due to diagenetic processes that can undermine their preservation, and that of aDNA too (Trueman et al., 2008).

One of the most important factors that can cause the diagenetic alterations of the bone tissue is the pH of the burial soil. This directly depend on the chemical/mineralogical composition of the soil and on the composition of circulating fluids, dissolved salts, acid ions H^+ or hydroxyls OH^-. A prevalence of hydrogen ions will make the soil acidic, while a prevalence of hydroxyls will result in a basic pH. Collagen at alkaline pH is more susceptible to alteration than if in contact with acidic soil (Collins et al., 2002, 1995; Rudakova and Zaikov, 1987). In acidic soils and in presence of circulating water, biogenic apatite can be solubilized more quickly (Nielsen-Marsh et al., 2007; Nielsen-Marsh and Hedges, 2000). Bioapatite is in fact more stable in soils with slightly alkaline pH and begins to dissolve at acidic pH (Berna et al., 2004). Bioapatite can recrystallize and incorporate exogenous ions from the circulating fluids (Berner, 1980).

In fact, the water leaches the mineral and its removal rate is related to the pH (Collins et al., 2002; Hedges, 2002; Kendall et al., 2018). Removal of the mineral causes the collagen to decay, exposing it to potential microbial attack, resulting in the destruction of bone tissue (Hedges, 2002; Klont et al., 1991; Nielsen-Marsh and Hedges, 2000; Reiche et al., 2003; Trueman et al., 2004). Draining soils involve a structural and chemical alteration of the bone as it will incorporate exogenous chemical elements within the crystalline structure as well as causing a dissolution and recrystallization of the bioapatite (Müller and Reiche, 2011; Reiche et al., 2002; Trueman et al., 2004). Other degradation agents are humic substances (such as humic acids, fulvic acids) that can form random, non-specific cross-links with bone collagen and causing the partial bone demineralization (Turner-Walker, 2007; van Klinken and Hedges, 1995). At certain condition of temperature and humidity the gelatinization of resulting in the interruption of the protein-mineral bond and leaving the bone tissues more

susceptible to microbial degradation (Kendall et al., 2018). Deterioration by microorganisms is associated with bioapatite dissolution, recrystallization and its re-configuration within the affected zones (Hedges, 2002; Keenan and Engel, 2017; Turner-Walker, 2007; Turner-Walker and Syversen, 2002). Collagen decays due to enzymatic activity, as these processes allow access to fungi and bacteria (Hedges, 2002). Organic acids and carbonate product by organic matter decomposition (Collins et al., 2002; Manning, 2000), produce H^+ ions which lead to collagen hydrolyses and swell, reducing apatite dissolution. The characteristics of the soil such as composition, physic-chemical properties, influenced by climatic conditions, play a key role in the survival of the osteological finding (Collins et al., 2002; Hedges, 2002; Hedges and Millard, 1995).

Chemical changes can also occur in acidic environments, transforming bioapatite into brushite (Berna et al., 2004) while the recrystallization and incorporation of exogenous ions into the crystal lattice from the burial environment (e.g. PO_4^{3-}, CO_3^{2-}, Ca_2^{+}, Mg_2^{+}) leads to a more stable and ordered crystalline structure characterized by larger dimensions of the nanocrystals (Keenan and Engel, 2017; Pollard et al., 2007; Reiche and Gourrier, 2016; Trueman, 2013).

Skeletal remains found in arid areas degrade much more rapidly than those deposited in temperate environments (Hedges, 2002; Maurer et al., 2014). Furthermore, sites with stagnant and anaerobic water show minimal microbial attack (Hedges, 2002). In summary, *post-mortem* alterations lead to the chemical degradation by hydrolysis of collagen with consequent increase in bone micro-porosity and increase in recrystallization of bioapatite due to extreme environmental temperatures and pH, humic acids and microbial attack that alter the collagen-apatite hierarchical structure (Collins et al., 2002; Hedges, 2002; Turner-Walker and Syversen, 2002).

The alteration of bone and its bearing on DNA survival

The main constituent components of the bone (organic and inorganic phases) are intimately connected not only in their organization but also concerning their survival during diagenetic alterations. The presence of collagen provides stability to the mineral matrix and its crystallinity increases following chemical loss of collagen (Collins et al., 2002; Hedges, 2002; Keenan and Engel, 2017; Kendall et al., 2018).

Post-mortem alterations are multifactorial complexes processes that act in the hierarchical structure of bone in a heterogeneous and non-linear way over time (Collins et al., 2002; Dal Sasso et al., 2014b, 2016; Hedges, 2002; Nielsen-Marsh et al., 2007; Reiche et al., 2002, 2003; Smith et al., 2007).

Ancient DNA in osteological samples strongly altered by exposure by degrading processes (Allentoft et al., 2012; Rollo et al., 2002) appears frayed, fragmented as well as contaminated by microbial genetic material (Smith et al., 2003; Tuross, 1994). Under the conditions of poor preservation, especially when the quantity of fragments found in an archaeological find is less than 1%, the

sequencing of ancient DNA can be an extremely expensive operation (Carpenter et al., 2013; Der Sarkissian et al., 2015; Nair, 2014). Several investigations were undertaken to understand how DNA is preserved and where it is preferentially located in the altered bone (Campos et al., 2012; Grunenwald et al., 2014; Okazaki et al., 2001).

At the ultrastructural level, the DNA can establish strong bonds with both the organic and the inorganic components, so that adsorption of DNA to bioapatite crystals surface as well as the linking with Type I collagen may stabilize it and may determine its survival in time (Campos et al., 2012; Collins et al., 1995; Sosa et al., 2013). The nucleotide chain can be strongly bound to both organic and inorganic bone phases, thus an extraction method to recover double helix molecules is applied to both fractions, with greater chance of success, employing EDTA (to demineralized the bioapatite) and proteinase K (to digest the osseous proteins) (Rohland and Hofreiter, 2007).

Some bones survive well, while others degrade rapidly; and since it is difficult to evaluate a *priori* the preservation of organic material within a bone, so far expensive and often destructive analyses have been the norm to assess the suitability of a given sample to be included in molecular studies (Orlando et al., 2021). Given the high costs and complexity of aDNA sequencing procedure, several analytical studies were carried out to develop reliable pre-screening methods to determine the presence/absence of suitable genetic molecules in an archaeological bone sample.

Different methods and different tools were employed (Dobberstein et al., 2008; Fernández et al., 2009; Fredericks et al., 2012; Gotherstrom et al., 2002; Kieser et al., 2007; Kontopoulos et al., 2020; Leskovar et al., 2020; Ottoni et al., 2009; Poinar et al., 1996; Poinar and Stankiewicz, 1999; Scorrano et al., 2015; Sosa et al., 2013).

Evaluation of novel parameters capable of identifying samples that might contain original sequences led to focus on proteins. In fact, organic macromolecules may be reliable indicators of the conservation status of a bone finding as they are as vulnerable as DNA to the same environmental conditions. Both collagen and DNA undergo hydrolysis processes (Iuliani et al., 2010; Poinar et al., 1996). In ancient samples showing a limited degree of collagen hydrolysis, greater peptide chains will be present with respect in poorly preserved tissues in which there is a more considerable fragmentation of the protein (León-López et al., 2019.

Poinar and Stankiewicz (1999) demonstrate that the relative quantity of pyrolysis products detected by gas chromatography – mass spectrometry (GC-MS) provides a good index of the quantity of peptide hydrolysis, which consequently provides information on the good preservation of ancient DNA.

The degree of racemization of amino acids was thus examined as *proxy*. All amino acids, except glycine, can be found in the form of two optical isomers: D and L enantiomers (Abdulbagi et al.,

2021). The L-enantiomer is engaged *in vivo* for the biosynthesis of proteins and is found in greater proportion than the D form (MacGregor, 2001). Upon the death of the individuals, a balance occurs between the L and D forms (Wochna et al., 2018). The racemization rate depends on the same climatic and chemical-physical conditions of the soil that cause DNA depurination (Poinar et al., 1996). Aspartic amino acid is the most analysed amino acid since its rapid racemization within proteins allows determining the highest D/L ratios (D/L Asx value) in archaeological bones (Collins et al., 1999). Due to this property, this technique was widely used as a non-absolute dating method (Bada and Masters Helfman, 1975; Bravenec et al., 2018; Demarchi and Collins, 2014; Johnson and Miller, 1997). Its application in the archaeometric field was abandoned as it led to incorrect dating (Bravenec et al., 2018; Johnson and Miller, 1997).

Many researchers have investigated the potential of the amino acid racemization technique in that is a technique less invasive (less than 10 mg of bone) (Dobberstein et al., 2008; Fernández-Jalvo et al., 2016; Ottoni et al., 2009; Poinar et al., 1996; Poinar and Stankiewicz, 1999). Collins (2009) and Ottoni (2009) conclude that no correlation was found between DNA conservation and aspartic acid racemization, since the rate D/L Asx depends on biological and geochemical factors (Fernández et al., 2009). The total absence of collagen may still be indicative for excluding some samples for ancient DNA analysis, however the presence of the organic component does not necessarily indicate the presence of well-preserved DNA.

Fourier-transform infrared spectroscopy (FTIR) was recently applied to characterize changes in the properties of bone materials compromised by diagenesis (Al Sekhaneh et al., 2020; Chadefaux et al., 2009; Dal Sasso et al., 2016; Fleet, 2017; France et al., 2014; Kontopoulos et al., 2018; Lebon et al., 2014, 2010; Naito et al., 2020; Salesse et al., 2014; Stathopoulou et al., 2008). Infrared spectroscopy has been widely applied to bone characterization studies as an important tool for diagenesis research, as it offers many advantages in terms of low-cost application, low amount of required sample (~1 mg) and sensitivity to variations on the structural properties of both organic and inorganic components (Carden and Morris, 2000).

Therefore, this tool allows the rapid, minimally destructive and economic analysis of a large number of fossil bones. The crystallinity as well as the physico-chemical and microstructural properties of bioapatite are extremely meaningful indicators of archeological bone preservation (Weiner, 2010). Indeed, the characterization of archaeological bones using infrared spectroscopy provided valuable quantitative and qualitative information on bone preservation (Piga et al., 2016; Reiche et al., 2002; Shemesh, 1990; Smith et al., 2007; Thompson et al., 2011; Weiner and Bar-Yosef, 1990).

The potential of this spectroscopic technique was the basis available to design a pre-screening method of bone samples for genetic analysis. Diagenetic changes were examined in relation to the success or

failure of nuclear DNA amplification. Fredericks et al., (2012) compare infrared data with a range of DNA amplicons of different sizes on cremated and forensic bone samples.

IR data are widely used in the literature to describe the alteration state of apatite (often through the use of the IR splitting factor (IRSF), Weiner and Bar-Yosef, 1990) and the amide-to-phosphate ratio (mineral-matrix ratio), representative of the amount of collagen present in the bone as a proxy for collagen preservation (Boskey et al., 2003; Donnelly, 2011; Figueiredo et al., 2012; Paschalis et al., 2011). The results obtained from forensic dataset, presented by Fredericks and colleagues (2012); suggest that collagen degradation and DNA integrity are highly correlated. The breakdown of hydrogen bonds and covalent bonds in collagen and DNA occurs at a similar temperature, so that they are relatively stable over comparable amounts of time. The splitting factor (IRSF) of the mineral is also correlated with the amplification of the DNA but it seems not to be a very sensitive parameter since, the increase of IRSF, which indicates an alteration of bone conservation, is not constant and fluctuates over temperature range (Fredericks et al., 2012).

It is evident that artificially treated bones which have undergone only a single serious insult, such as the use of extremely high temperatures applied to fresh forensic bones, are not suitable for explaining more complex systems of alteration present in archaeological bones.

In 2013 Sosa used various analytical tools such as GC/MS, scanning electron microscopy (SEM) and energy dispersive X-ray analysis (EDX), X-ray diffraction (XRD) and FTIR, for the genetic pre-screening method. The aim of the work was to analyse medieval human bones by characterizing them by morphological, structural and chemical-biological aspects, and possibly relating them to the presence of DNA. A positive correlation was obtained with the presence of DNA from those examined samples which had a high collagen content (<12%), low racemization values of aspartic acid, leucine and glutamic acid, low IRSF value and a compact appearance of the bone (determined by SEM/EDX).

In this work, samples with the same chronology selected in three different archaeological sites from the same geographical area (Aragonese Pyrenees) were analysed. This protocol, besides not being really cheap, is not universally applicable to other archaeological samples, as it does not describe the particularities of different archaeological sites.

Scorrano and colleagues (2014) bound the presence of DNA to the conservation of archaeological bone by using various analytical tools: the racemization of amino acids, BJH (Barrett-Joyner-Halenda) pore size distribution method and infrared spectroscopy on archaeological bones with different origins and chronology. The data show that the IRSF is the parameter best associated with the presence of DNA. The analyzed samples all showed good conservation of the organic and inorganic component and the authors concluded that the spectroscopic parameter IRSF is the

parameter best associated with the presence of DNA and they recommend its use for simplicity and savings. This is proposed as an alternative to previous works that identify the organic component as the best proxy for the presence/absence of DNA.

In the 2002, Gotherstrom and colleagues studied experimentally degraded modern bones and teeth, in addition to ancient teeth and bones using XRD (which requires about 15 mg of material), and collagen analysis. They related the presence of collagen to the presence of mtDNA (mitochondrial DNA), more stable than nuclear DNA. The measured crystallinity and the quality of the extracted collagen were measured and their results showed that the presence of DNA was strongly related to both the crystallinity of the hydroxylapatite and the quantity of collagen. Their data suggest that hydroxylapatite plays a crucial role in DNA preservation and that the organic component likely indicates its presence.

X-ray diffractometry is undoubtedly one of the most important techniques for the analytical study of solids but both the qualitative analysis of the crystal phases and the full quantitative phase analysis (QDA) are less economical than FTIR analysis, and certainly not trivial to be carried out.

In the study by Wadsworth and colleagues (2017), in-depth proteomic analyses (LC-Orbitrap-MS/MS) on archaeological bovine dentin and bone from multiple European archaeological sites were performed, and the results compared with mtDNA and racemic amino acids data. Also in this work they suggest that DNA survival is more closely linked to the inorganic component, finding no correlation with the investigated proteins.

In 2020 Kontopoulos and colleagues publish a study based on spectroscopic techniques, such as Fourier-transform infrared spectroscopy in attenuated total reflectance mode (FTIR-ATR), besides proteomic and genetic analysis. They applied various infrared parameters widely used in literature for the evaluation of the conservation state of the finds such as IRSF, the carbonate/phosphate ratio as a measure of the relative structural carbonate content in the bioapatite crystals and the amide/phosphate ratio to evaluate the relative content of organic component preserved in the bone. This study proposes the use of FTIR-ATR for the identification of archaeological bone samples containing endogenous DNA and a good collagen content by weight, establishing arbitrary thresholds for the IR parameters used (IRSF, carbonate/phosphate, amide/phosphate). Similarly, Leskovar (2020) recent work explores the physico-chemical analysis of archaeological samples obtained from different origin and chronology, by means of FTIR- ATR spectroscopy, correlating the results with the DNA conservation status and developing a measurement protocol. The wide range of IR data collected were statistically analysed in order to validate a protocol for DNA conservation in ancient osteological remains.

According to Leskovar (2020), the state of collagen, phosphates and carbonates in the skeletal remains, determined by IR measurements, must all be taken into account to assess the conservation status of the endogenous DNA in the sample. Furthermore, they determined that in-depth analysis of the IR spectra could allow reliable separation of samples containing large amounts of extractable or free endogenous DNA. Not surprisingly, the data obtained in their work indicate that generally the age of the samples greatly affects preservation.

Although several researches have been conducted on the subject, it is not yet clear which parameter is more indicative of the presence/absence and conservation (%) of endogenous DNA in ancient bone samples.

A better understanding of bone diagenesis using FTIR spectroscopy and an improvement in the screening method would be useful to prevent unnecessary destruction of valuable biological samples and to minimize costs and resources.

One of the aims of this PhD project is to develop a valuable protocol applicable to several archeological contests, in order to preserve these collections from invasive analyses by allowing access to samples for scientific research. The evaluation of bone bioapatite degradation using the universal crystallinity curves developed by our research group (Dal Sasso et al., 2018) based on FTIR analysis, applied on a widely set of bone coming from different geographical areas and characterized by different chronology and diagenetic pathways due to different environmental conditions, could provide an adequate preliminary assessment of the presence of ancient DNA contained in the bone. In this work, we intend to examine the skeletal tissues, in particular the compact tissue of the petrous bones and the roots of the teeth as a pre-screening method to evaluate the bone suitability for specific analyzes, e.g. ancient DNA, paleoproteomic and stable isotope analysis.

The development of a molecular pre-screening method through the direct evaluation of *post-mortem* alteration of the degree of recrystallization and atomic order/disorder of bioapatite, correlating it with the quality/quantity of collagen, could allow a more careful selection of well-preserved bone samples, allowing to reduce the costs of genetic analysis. The micro-sampling of osteological findings, characterized by different chronology and origin, made it possible to investigate the variability of the sample in terms of conservation status. Comparing bone integrity with data obtained from genetic analysis is an essential step to provide an accurate assessment of the degree of conservation of the sample, aiming at the extraction of sequenceable endogenous aDNA, applicable to all osteoarchaeological findings. Furthermore, an in-depth study on the altered state of collagen will strengthen the knowledge on its conservation/alteration and on the possibility that it can be positively

correlated to the presence of ancient DNA. Since DNA survival is limited by multiple intercorrelations, the study is complemented by examining the microstructural changes induced by diagenesis on the secondary structure of conserved collagen (Chadefaux et al., 2009; Fredericks et al., 2012; Leskovar et al., 2020).

This thesis is divided in five chapters. The first chapter deals with the documentation of the archaeological specimens was carried out through X-ray computed micro-tomography (X-μCT) and close-range photogrammetry in order to preserved the morphological-textural information before sampling and eventually enabling detailed morphological studies. The three-dimensional surface models generated by the two techniques were compared with the aim of evaluating whether close-range photogrammetry is a valid alternative to construction of 3D models with μCT, which could alter future investigations.

The second chapter is addressed to the sampling strategy in order to develop a protocol, to be minimally invasive, and a standardized sample preparation method for Fourier transform infrared spectroscopy (FTIR) to produce accurate and reliable results.

The third chapter is focused on the physical-chemical characteristics of bone constituents retrieved from FTIR analysis were used to define the alteration state of bone material in relation to the local environmental conditions experienced by the bones during burial.

The fourth chapter shows the results obtained analyzing the alteration state of the inorganic and organic phases constituting the bone (mainly bioapatite nanocrystals and collagen) on a large and diverse set of samples. The yield of ancient DNA and bone spectral data were compared with those obtained from stable isotopes and spectra of lyophilized collagen in order to determine the parameters IR useful for the development of a pre-screening method.

The fifth chapter is a multidisciplinary case study that involved various professional figures in order to evaluate the state of conservation of the bones for radio-carbon analysis and to connect the skeletal remains to a historical figure through anthropological/paleopathological study. Finally, through forensic facial reconstruction techniques, his features were reconstructed.

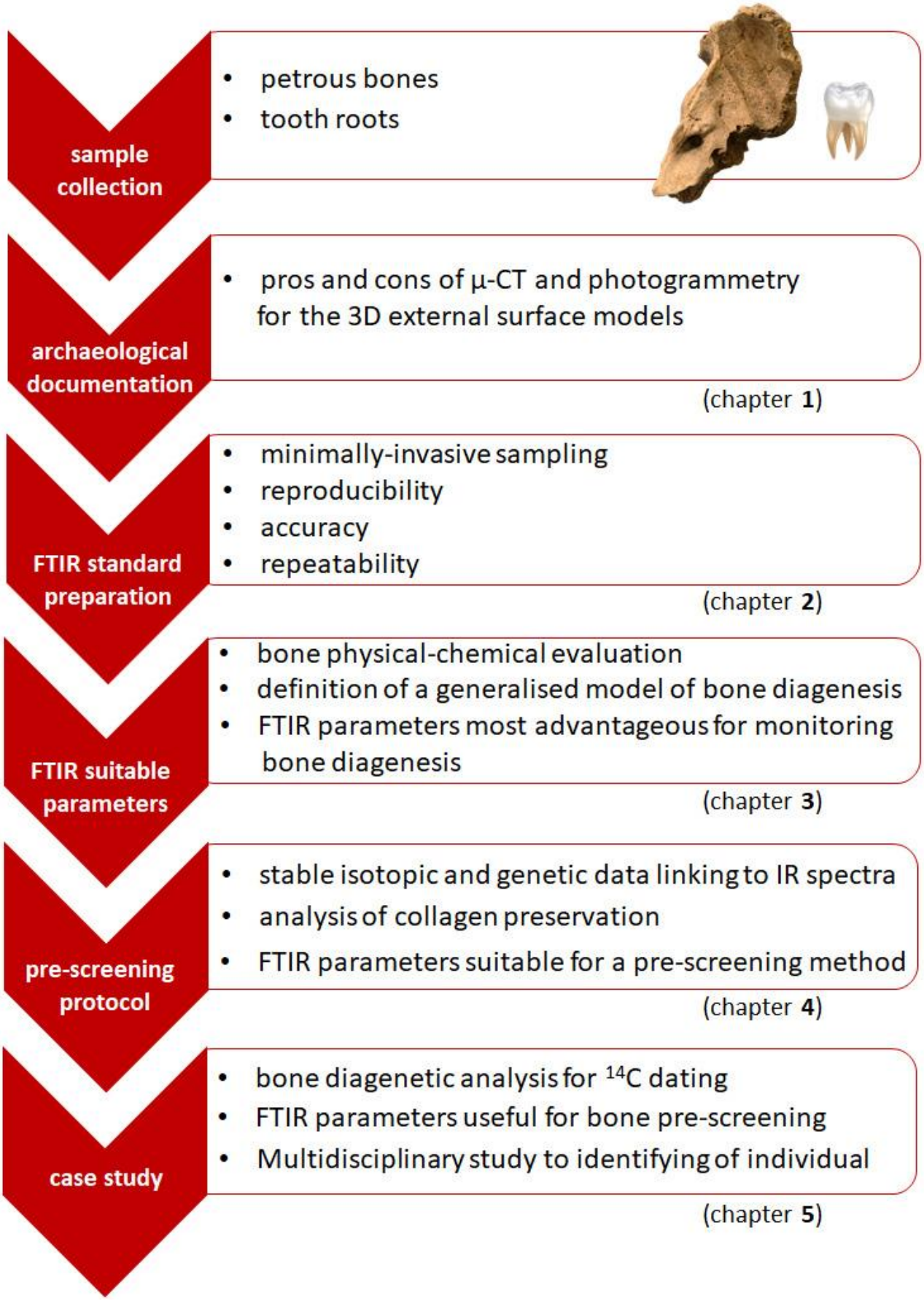

Figure 4. *Concept map of the thesis structure.*

MATERIALS AND METHODS

1. Archaeological context of bone specimens

The study samples selected for this research work included human bones of different origin, chronology and conservation status preserved in the Museum of Anthropology of the Padua University (Italy), in particular petrous bones and tooth roots.

The Museum of Anthropology holds a collection of 151 remains from archaeological excavations of cemeteries, castellieri (Bronze and Iron age fortified settlement, typical of Northeastern Italy) and caves. Some of the archaeological remains are attributed to the prehistoric era or pre-Roman and Etruscan era (13th century BC– 4th centuryAD), but most of them date from the late 19th and early 20th century (Alciati et al., 1996).

Recently, the Museum hold a very important osteological collection of about 200 human skeletons excavated in the archaeological site of Al-Khiday in central Sudan (an evolutionary area fundamental to better understanding the "Outside of Africa" with a wide chronological span ranging from about 7000 to the 4th-6th century a.C.).

All samples selected were sampled by means of a low-speed micro-drill (Dremel Micro Rotary) as detailed in **Chapter 2**, striving to reduce the impact of destructive sampling and retain as much information for future researchers as possible.

Below the archaeological contexts of the various provenance sites for the selected finds are reported, based on available information.

1.1. Tedeschi Collection

The Tedeschi Collection takes its name from Enrico Tedeschi (1860–1931).

Enrico Tedeschi was the first director of the Institute and Museum of Anthropology, from 1897 until his death. The craniological collection comes from cemetery exhumations, ossuaries or unclaimed bodies acquired from hospitals or other relief organizations and it includes demographic information. Most of the specimens come from one or more cemeteries in Bologna and from the Padua Hospital; they were probably exchanged between Padua and Bologna universities. Further individuals come from the Padua penal institute and from the Ferrara psychiatric Hospital, close to Padua. These individuals were collected between 1900 and 1910 and, in part, they preserve demographic data including age at death, date of death, sex, occupation, and cause of death. The remaining portion of the collection, represented by crania and mixed postcranial elements, does not have associated demographic information, but origin and/or chronological period are known.

More information is contained in (Carrara et al., 2018).

For this research work, 5 female skulls of the 19thc. ranging between the ages of 28 and 40 years old from the Tedeschi collection were selected, originally from the cemetery of Bologna and Rome and from the Hospital of Padua.

1.2. Al-Khiday (Central Sudan)

The samples come from an archaeological investigation carried out within the "El Salha Archaeological Project" and still in progress (Usai and Salvatori, 2005). Archaeological survey has identified around 200 sites along the White Nile, between Omdurman (Sudan) and the Jebel Aulia dam. The excavations were carried out near the village of Al Khiday (Omdurman, Khartoum, Sudan), on the western bank of the White Nile close to the confluence with the Blue Nile.

The sites examined revealed a Mesolithic occupation of the site and a multi-stratified cemetery (Dal Sasso et al., 2016; Usai et al., 2014, 2010; Usai and Salvatori, 2019) comprising at least three different burial phases (pre-Mesolithic, Neolithic and Classic/Late Meroitic). The pre-Mesolithic phase, the most ancient, is characterized by individuals buried in a prone and extended position, rarely documented elsewhere (Usai et al., 2017, 2010) without grave goods (with the exception of Grave 153 where an individual was found with an ivory arm-ring at the right wrist) (Salvatori et al., 2011). The attribution to the pre-Mesolithic period encountered many difficulties both for direct and indirect dating due to the poor conservation of the osteological remains; the bones are totally lacking of collagen useful for ^{14}C dating and for the absence of goods representative of known chronological phases. The determination of the chronological age is mainly constrained by the archaeological evidence of use of the site during the Mesolithic period, when pits with different functions were found to cut the graves (Dal Sasso et al., 2014a; Salvatori et al., 2011; Zerboni, 2011). Based on radiocarbon dating of charcoals and shells and typological studies performed on pottery fragments found within pits as filling material, the burials have been dated to approximately pre-6.700-6.300 BCE. The non-absolute dating was confirmed by 16 pits found that cut individuals in a prone position (Salvatori et al., 2011; Salvatori and Usai, 2019; Williams et al., 2015).

The site was later reused as a cemetery during the Neolithic period (4550-4250 cal. years BC) and later on during the Meroitic period (50 cal. years BC-250 cal. years AD). In the Neolithic period individuals were buried in flexed position with pottery vessels and personal adornments (Salvatori and Usai, 2019). The Meroitic grave structure is larger than the pre-Mesolithic one and consists of a rectangular pit that leads to a smaller chamber. The large dimensions of the Meroitic pits (3x2 m) strongly disturbed the most ancient sepulchral phases.

Within the Al-Khiday site, archaeological and geomorphological investigations were carried out, showing that the area was characterized by very different climatic environments along the Holocene. The per-Mesolithic phase experienced a paleo-marshy phase, while drier climatic conditions occurred during the Meroitic use of the site and later on (Usai, 2016; Williams et al., 2015).

In this research, bone samples belonging to different burial phases characterized by different paleo-

environmental conditions were studied.

1.3. Grottina Covoloni del Broion

The Grottina dei Covoloni del Broion is part of the oriental rocky wall complex of Colli Berici. It is close to the famous cave Grotta Broion in the province of Vicenza in Northern Italy, where among Paleolithic sediments, the Italian geologist Piero Leonardi of Padua University found significant traces of Neanderthal. This prehistoric shelter is located in Vallà di Lumignano in the municipality of Longare and was used as a funerary site. Through the analysis of the palaeo-surface, the site seems to have been used for a long time covering many generations belonging to the Chalcolithic, confirmed by the discovery of Eneolithic artifacts.

Grottina dei Covoloni del Broion was discovered in 1973 and systematic excavations of the different geological layers and of parts of the Grottina were conducted during four archeological campaigns between 1973 and 1977. The archaeological excavation of 1977 was entirely reserved for the study of funerary depositions. The inner part of the cave was intended for sepulchral use,as evidenced by the discovery of many human osteological findings incorporated in the calcareous sediment. The archaeologists had also found ceramic fragments and many flint artifacts: a flat blade of dagger, a significant amount of arrow cusps, blades, two flint cores, small discoid beads and a small calcite tube from a bracelet or necklace, and other personal ornaments found in the same layer. The human remains include bones and teeth under different states of conservation and distribution (Ligabue, 1973). The dating of the site (Bronze Age) was confirmed by radiocarbon analysis of the bone findings at 14CHRONO Center for Climate, the Environment, and Chronology in Belfast (UK).

1.4. San Giovanni Rotondo

The San Giovanni Rotondo samples come from the osteo-archaeological collection of the Museum of Anthropology of the University of Padova and are not associated to any record with the exception of a broad “Iron Age” archaeological label; they may be part of the samples brought to the Museum by Prof. Santo Tiné in the 1960s. The bone samples were dated at 14CHRONO Centre for Climate, the Environment, and Chronology in Belfast (UK), confirming the chronological phase of the Iron Age. Only one sample (SGR001_T) is Medieval (8^{th}c.AD).

1.5. Ordona

The necropolis of *Herdonia* (today’s Ordona, within the Apulian Foggia province in Italy) was studied during different archaeological campaigns in 1978 and 1981, interested in the Daunians, Roman as well as the

medieval settlements (Corrain, 1986). The human remains collected on such occasions were later cataloged and studied, revealingnew paleopathological evidence (Scaggion and Carrara, 2016). Firstly inhabited from the Neolithic period, Herdonia became an important Daunian center from the 6thc. BCE. Radiocarbon dating at 14CHRONO Centre of Belfast (UK) confirms this chronological period for these samples. Only one sample (ORD010_PB) is Medieval (11thc.AD).

1.6. Salapia

The necropolis of Salapia is an ancient town located 10 km from contemporary Cerignola, within the Apulian Foggia province in Italy. The osteological samples were brought by Prof. Santo Tiné to the Museum of Anthropology of the University of Padova with no further information on the archaeological context, and osteological studies were carried out by Cleto Corrain and colleagues in 1971 (Corrain et al., 1972) and dated to the ^{14}C at 14CHRONO Centre of Belfast (UK).

1.7. Oderzo

Oderzo is an Italian town in the province of Treviso, in the Veneto region, Northern Italy. Unfortunately, only partial anthropological data was produced and published in a short communication (Drusini et al., 1987). Although the ancient Opitergium (today's Oderzo) was from the pre-Roman/Roman period (2nd- 3nd c.AD), both the ritual of inhumation and of cremation are certified (Cipriano and Sandrini, 2015). In this study, only bones from inhumation burials were analyzed inasmuch they were not compromised by the high cremation temperatures, leading to collagen denaturation (Wagermaier and Fratzl, 2012). The Necropolis of Oderzo, found close to the western margin of the city during the archaeological dig occurred in 1986, was located at the river bank margin of the inactive Lia River (Cipriano and Sandrini, 2001), thus determining a low and humid ground, optimal for the conservation of skeletal remains.

1.8. Montecchio Maggiore

Montecchio Maggiore, a small town located at a dozen kilometers west from Vicenza (Veneto, Italy), seems to have been inhabited since the Eneolithic and Bronze Age, as attested by the discovery of artifacts, small flints, terracotta and human burials. The current center of Montecchio Maggiore is located at the base of plain corridor between the Lessini and the Berici where a stretch of the ancient Via Postumia passed, connected road between Vicenza and Verona and the great centers of Italy in the Po Valley. Archaeological excavations carried out in 1990 in an area behind the civil Hospital confirmed that the whole area, below a very small thickness of soil (0.30 - 0.50 m), was affected by ancient pre-existing structures (Rigoni and Bruttomesso, 2011). During the excavations

burials were identified, some of which, characterized by rich Lombard-type grave goods datati 7th-8thc. AD. The sampling of the osteological remains selected for this research was approved by the Soprintendenza Archeologia, Belle Arti e Paesaggio per l'Area metropolitana di Venezia e per le Province di Belluno, Padova e Treviso (Italy).

1.9. Sant'Anna Sopramonte

Sant'Anna Sopramonte (municipality of Trento, Northern Italy) was the site of a monastery with a mixed community of monks and nuns, dedicated to St. Augustine and the Dominican order, between the early 1200s AD and the end of 1400s AD. Historical sources describe the monastery as a political and religious center of the medieval age of absolute importance and exclusivity, in direct contact with the Vatican. Three popes (Gregorio IX, Urbano IV and Nicolò V) took care of the monastery, guaranteeing its economic subsistence, prosperity and the very important privilege of being able to bury the dead in the same place (Obermair and Brandstätter, Klaus. Curzel, 2006). On the site of Sant'Anna, a church dedicated to the patron saint and a large building known as the Casa del Preposto persist. Based on the important historical data of the sources, an archaeological campaign (currently in progress) has been carried out to verify the presence of structures of the famous medieval monastery in collaboration with the University of Padua, Appalachian State University (North Carolina, USA), the Sovrintendenza dei Beni Culturali of the Provincia Autonoma of Trento (Trento, Italy) and the A.S.U.C. (Amministrazione Separata dei Beni di Uso Civico) of Sopramonte. Two archaeological excavation campaigns were conducted in 2019 and 2021.

1.10. Biverone, San Stino di Livenza (VE)

Biverone is a village in the municipality of San Stino di Livenza, situated in the eastern part of the Venice province. The village is located close to the banks of the Livenza river. A section of the Via Annia – the consular Roman road that connected Adria to Aquileia, through Padua, Altino and Concordia – passed by the South of the village. In 1983, the Archaeological Superintendence of Veneto initiated a preliminary survey on a 300-sq.m. area that ended with the discovery of the necropolis (Croce da Villa, 1984). The stratigraphy indicated that the necropolis had been used for a long time, with an estimated presence of three generations at least. The first skeletons completely *in situ* and well-preserved were found at a depth of about 68 cm, surrounded by a blackish soil rich in organic matter, immediately under the humus and surrounded by brick structures slightly higher (50 cm) (Croce da Villa, 1984). Most of them are skeletons buried without funerary goods while South of the archaeological site was found a quadrangular pit used as a mass grave, as well as a large burnt area with numerous remains of pottery, glass and wood. The archaeological evidence

suggests the diffusion of the funeral meal for the celebration of the deceased. Supporting the thesis, within the anthropological study carried out at the Museum of Anthropology of the University of Padua, bones of animals (in particular of cattle and pigs), fragments of ollae, amphorae, plates and colored glass were found besides the skeletal remains (Gadioli et al., 2018). For this research, both burials deposited in single graves and skeletons found in mass graves were sampled.

1.11 Desenzano Del Garda

The church of San Lorenzo stands on a long ancient route, probably of Roman origin, coming from Desenzano del Garda in an agricultural landscape with significant evidence from the late antique and early medieval period. The strata relating to the construction of the church of San Lorenzo did not return findings capable of providing a chronology.

The absolute dating of some bone finds (7^{th} c.AD for the oldest) offer only a dating of the *ante quem*. The architectural features of the building suggest that the construction can be dated to the Late Antiquity (see Chavarria Arnau 2011), between the 6^{th} century AD - end of the 5^{th} century AD., compatible with the diffusion of the relics of San Lorenzo. Inside the church, a cemetery was identified including 24 graves. The graves are disposed uniformly without overlapping. The stratigraphic data enables to distinguish three phases of the cemetery. The oldest graves did not return skeletal remains or grave goods to perform an absolute dating. Graves from the later phases were dated by ^{14}C between the 7^{th} and 9^{th} c. AD (Canci et al., 2012).

1.12 Saint Theobald

Saint Theobald (Saint-Thibaut, 1033-1066 Provins, France) was an important medieval historical figure of Catholicism. The relevance of this Saint is attested by the statue dedicated to him (n.139) and located above Bernini's Colonnade in piazza San Pietro, Vatican City. Two hundred years before Saint Francis of Assisi (1181-1226), Theobald, coming from a noble family, was the first to embrace poverty and lived as a hermit and pilgrim. According to historical records, he died of leprosy (*Mycobacterium leprae*), a disease widespread in Europe at the time, at the age of 33. Some of his bones found their way from France as holy relics to the parish church of Saint Giovanni Battista in Badia Polesine, Italy. According to historical sources, Saint Theobald's body has never been buried in direct contact with the soil, preventing the *post-mortem* alterations typical of buried bones (Soffiantini, 2014). A multidisciplinary investigation was conducted to go beyond authentication and historical analysis to gain as much information as possible about his life, death and appearance, reported in **Chapter 5**.

2. The choice of methods

2.1. Close-Range Photogrammetry

In this research, short-range photogrammetry was applied to generate surface 3D models from a sequence of two-dimensional images (Luhmann et al., 2006). The obtained 3D models were compared with those generated by micro-CT with the aim of comparing the accuracy and precision of the models and measurements obtained using different methods and tools. Models were generated using images taken with oblique photography using a Nikon D810 Reflex Pro camera with 36.3Mp FF (1.0x) sensor, macro lenses 60 mm, f/2.8 aperture value, shutter speed 1/20 seconds and ISO 400, with definition of 3696×2448 pixels. A tripod was used to stabilize the process and photographs were taken rotating the turntable. Photographs were then processed to generate a 3D model through the dense point cloud for each sample with the photogrammetric reconstruction software 3DF Zephyr Free (3DFlow©). Refer to **Chapter 1** for more information on the methods used.

In **Chapter 5**, photogrammetry technique (Structure from Motion - Multiple-View Stereovision (SfM-MVS) was used to create a high-resolution, three-dimensional virtual model of the skull, which became the basis for the 3D reconstruction of the face of Saint Theobald. Precise metrical models were generated using several photo sets by means of a Nikon D800 Reflex (DSLR) camera, mounted on a tripod, with a CMOS sensor (36.3 Mp) and SIGMA 24 mm F 1.4 lens. The acquired photographs were processed to generate a 3D model of the two bone elements (cranium and mandible) with the openMVG101 photogrammetric reconstruction open-source software, within ArcheOS100 Linux distribution. The polygonal lattice that defines the virtual object (3D mesh) was processed by MeshLab102 software. The facial reconstruction/approximation of Saint Theobald was carried out completely in a digital environment using the ForensicOnBlender (for more details see **Chapter 5**).

2.2. Confocal laser scanning microscope (CLSM) and digital caliper

The precision of photogrammetry applied to small objects for optical dimensional metrology on the 3D model generated by photogrammetry was evaluated. Scaling of 3D models was carried out by acquiring micro-measurements by Confocal laser scanning microscope (CLSM) and as many by digital caliper and magnifying glass. Measurements of samples, as reported in **Chapter 1,** were taken by Confocal Laser Scanning Microscopy (CLSM), model Olympus LEXT OLS4000 at the Department of Geosciences of Padua University, with a magnification capacity of 5x/0.15 N.A., W.D. 23.50 mm, FOV (HxV) 2.630×2.630µm and a spatial resolution of 10 nm axially and 120 nm laterally. The same distance points were manually measured by digital caliper and magnifying glass. The goodness of the alignment was evaluated by the Final Mean residual reported by 3DF Zephyr (ICP function - Iterative Closest Point) by comparing the distances measured with the digital caliper and the confocal microscope. Using this function, 3D models with different scales were aligned.

2.3. X-ray computed microtomography

Samples (**Chapter 1**) were scanned at the Department of Geosciences of Padua University using a SkyScan® 1172 high-resolution computed micro-tomography (Salvo et al., 2003), equipped with a Cu X-ray tube and operating at 100kV and 100μA, with an Al filter 0.5 mm thick. No sample preparation was required. Radiographies were acquired through 360° rotation with an angular step of 0.30° and acquisition time of 1050ms, a resolution of 256Mp and pixel size of 8.96μm for sample 95T and of 8.00μm for sample 34T. Image stacks of the two scans were registered based on grey values and transformed into a 3D single file using CT-an® software (SkyScan® 1172) exported in .STL file format. For the comparison of 3D point clouds generated by X-μCT and photogrammetry was used an open-source software, CloudCompare 2.12.alpha (Anoia) version. See **Chapter 1**.

2.4. Fourier Transform Infrared Spectroscopy (FTIR)

The attenuated total reflectance Fourier transform infrared spectroscopy (ATR-FTIR) is a widely used technique in cultural heritage and its conservation as it requires very little preparation (Lebon et al., 2016; Margariti, 2019). Like Fourier Transform Infrared Spectroscopy (FTIR), ATR-FTIR is also minimally invasive/destructive and provides information on both inorganic and organic structures of bone by evaluating *post-mortem* changes of BAp nanocrystals and collagen (Kontopoulos et al., 2020). Compared to the transmission technique, ATR is sensitive to dispersion phenomena and consequently, the low intensity and high frequency bands (water vapor in the source region) can be difficult to observe creating a noisy signal. One of the limitations of the ATR technique is that the optical contact between the sample and the analyzer crystal must be guaranteed and, in consequence, the signal in transmission is generally more reproducible. In FTIR, much more signal is obtained than in attenuated mode over the entire spectrum and the analytical range is often wider (Bürgi, 2011).

Powder samples and lyophilized collagen (see **Chapters 2-3-4-5**) have been prepared, maintaining a ratio 1:100 mg of sample/KBr. Thus, 1 mg of sample was hand ground for 2 minutes in an agate mortar and KBr was added and homogenized. Powdered sample was then pressed, using a hydraulic press, under 11 tons/cm^2 pressure, and a 12 mm in diameter and 1.5 mm of thickness transparent pellet was obtained. Spectra were collected with a Nicolet 380 FTIR spectrometer equipped with a DTGS detector; 128 scans for each spectrum were acquired, in the range from 4000 to 400 cm^{-1}, with a spectral resolution of 4 cm^{-1}. Spectral analysis was performed using Omnic 9 software (Thermo Scientific). Statistical analysis was performed using the freely available Kernel Density Estimation (KDE) software developed by the Analytical Methods Committee of the Royal Society of Chemistry (Analytical Methods Committee, 2006). This function was applied to show a continuous substitution of the histograms obtaining an estimate of the underlying distribution of infrared parameters, modeled on the real measured data as reported in **Chapter 3**. A linear regression (function lm) in R free statistical software (4.1.2 version, available at www.r-project.org) was used (see **Chapter 4**). The infrared contribution parameters measured for each bone IR spectra were simultaneously combined

with endogenous aDNA yields to define a predictive model and highlighting the most predictive parameters.

2.5. Mass Spectrometry

The assessment of the degree of alteration of the organic component contained within the archaeological osteological samples was performed by IR spectroscopy (as reported in paragraph *2.4* and in **Chapter 4**) and by mass spectrometric analysis. 500 mg were collected from each selected sample (samples are reported in **Chapter 4**). The preparation of the samples was carried out in the Stable Isotope Mass Spectrometry laboratory of the Department of Geosciences (University of Padua) using the standard laboratory protocol based on Privat et al. (2002) and Richards and Hedges, (1999). Isotope analyses were performed using a Thermo Scientific Delta V Advantage Isotope Ratio Mass Spectrometer in continuous flow mode coupled to a Flash 2000 Elemental Analyzer and a ConFlo IV interface. The stable isotopic composition of carbon ($\delta^{13}C$) and nitrogen ($\delta^{15}N$) was calibrated respectively on the VPDB and AIR scales, using international standards (CH-6 and CH-7 for carbon, N-1 and N-2 for nitrogen and both UREA). The procedure performed and the standards used are reported in **Chapter 4**.

2.6. DNA sequencing

Laboratory extraction was performed in dedicated ancient DNA laboratories at the Estonian Biocentre, Institute of Genomics, University of Tartu, Tartu, Estonia. A minimally invasive sampling of ancient human remains was applied to extract the material useful for genetic analyses. The procedures performed are reported in detail in **Chapter 4**. Forty-one samples were selected (see **Table 1** of **Chapter 4**) and the data obtained from the molecular analysis were compared with the IR spectra from bone samples and lyophilized collagen in order to determine the most suitable IR parameters for the development of a pre-screening method.

2.7. Anthropological analysis

The anthropological and paleopathological analysis for the determination of sex, age and cause of death are reported in **Chapter 5** and were conducted to link the bone relics to the historical figure of Saint Theobald.

CHAPTER 1

Submitted to Digital Applications in Archaeology and Cultural Heritage

3D digital dental models' accuracy for anthropological study: comparing close-range Photogrammetry to µ-CT scanning

Cinzia Scaggion[1*], Stefano Castelli[1], Donatella Usai[2], Gilberto Artioli[1]

[1]Department of Geosciences, University of Padova, Padova, Italy
[2]Sudanese and Sub-Saharan Studies Center (CSSeS), Treviso, Italy

*corresponding author, e-mail address: cinzia.scaggion@phd.unipd.it (C. Scaggion)

Keywords: 3D models, X-µCT scan, ancient human teeth, photogrammetry, accuracy and precision

Abstract

X-rays micro-Computed Tomography (X-µCT) is a non-invasive and non-destructive technique primarily used for the 3D reconstruction of internal/external object structures. Despite the great potential, X-ray exposure for samples of archaeo-anthropological interest can compromise the integrity of the finds, altering the results for future investigations. The aim of this study is to evaluate whether close-range photogrammetry is a viable alternative in terms of non-invasiveness, low cost, efficiency and precision to the 3D external surface models with respect to X-µCT scanning. Photogrammetry was applied to small and highly diagenized archaeological human teeth characterized by irregular and fragmented surfaces. The accuracy of the 3D model and the metric information obtained from photogrammetry was evaluated comparing the micro-measurements obtained by digital caliper and Confocal laser scanning microscope (CLMS). The results obtained show the effectiveness of close-range photogrammetry to create a 3D model with metric information as accurate as those from micro-CT, with the advantage of reproducing textural information.

1. INTRODUCTION

Digital imaging is an innovative tool employed in cultural heritage diagnostics, scientific analysis and valorization, including reproduction. As far small concerned, two techniques are mainly used: X-ray micro-Computed Tomography (X-µCT) and optical photogrammetry, though laser scanning is also used. X-ray micro-Computed Tomography is a noninvasive and non-destructive technique primarily used for the 3d reconstruction of internal structures. Micro-CT reconstructs digital cross sections (slices) of an object that can be stacked to create 3D volumes. Virtual 3D volumes of specimens can be manipulated, sectioned, prepared, dissected and measured, producing an accurate reproduction of both the internal and external morphology. This allows access to the morphological information contained in fragile, valuable or small findings, including extinct fossils as comparative materials.

Moreover, different phases constituting the sample can be easily identified and separately analyzed, based on their X-ray absorption coefficient contrast.

Structure from Motion (SfM) photogrammetry is a technique used to generate external 3D model reconstructions starting from sequence of bidimensional images (Luhmann et al., 2006). It applies a mathematical model, fundamentally based on trigonometry (Szeliski, 2009), so that the 3-dimensional coordinates of the object surface are estimated by identifying the homologous features in the sequence of images taken from different perspectives (Luhmann et al., 2006). Photogrammetry is relatively low-cost, fast, non-invasive and provides information related to the external surface of an object. It is a reproducible method that allows for the generation of precise 3D models with high texture rendering and color (Luhmann et al., 2006).

Furthermore, photogrammetry allows high precision measurements (Ackermann et al., 2008; Hallert, 1963) by using reference points in photographs providing information on the size, shape, location, volume and areas of an object (Gillies, 2015; Sholarin and Awange, 2015).

A disadvantage of CT scanning technologies, besides the high cost of equipment and long processing time, is the generation of a huge amount of individual data points - having a set coordinate on the X, Y, and Z axis - in a 3-dimensional plane (called point cloud). This point cloud cannot be easily handled due to the huge density unlike the dense point cloud generated by photogrammetry, which can be easily calculated and controlled by photogrammetry software (James et al., 2015).

How the X-ray irradiation may affect the retrieval of pristine information from tooth samples has long been debated, in particular in the case of studies involving highly degradable molecules i.e. the case of ancient DNA (aDNA) or compounds sensitive to radiation doses which are then impossible to investigate with dosebades techniques, such as dating techniques based on electron-trap accumulation (i.e. for example ESR dating). The scientific community debated whether X-ray radiation from micro-CT scanning may damage ancient DNA (Grieshaber et al., 2008; Hall et al., 2016; Immel et al., 2016). On the one hand, medical devices are tuned to lower radiation levels, which may favour the preservation of aDNA as low radiation doses (below 200 Gy) does not damage it to an appreciable degree (Immel et al., 2016). On the other hand, examination of fossil samples is often performed using industrial or research micro-CT devices that exert higher radiation doses (Hall et al., 2016), which can reduce the amount of amplifiable aDNA. This is particularly evident when using high-brilliance synchrotron radiation causing significant degradation of aDNA, as radiation doses could exceed 2000 Gy (Immel et al., 2016). Moreover, photogrammetric modelling scanning is more cost-effective and less time-consuming than μ-CT scanning and in paleoanthropological contexts eliminates the possibility of exposing fossils to radiation doses (Duval and Martín-Francés, 2017).

The issue is especially important in Electron Spin Resonance (ESR) is a dating that relies upon the number of unpaired electrons in a crystalline material produced (e.g. tooth enamel) by the exposure to natural radiation. Knowing the rate of natural radiation, the ESR signal can be used to estimate the

chronology of the palaeo-anthropological findings. To determine the chronology of materials beyond the sensitivity of radiocarbon dating (older than c.a. 50.000 years), the ESR method has become more and more essential. Fossils subject to high-energy artificial radiation suffer a significant overestimation of the natural radiation dose. Given the sensitivity of ESR, even low radiation doses can significantly affect the age estimation of fossils (Hershkovitz et al., 2018).

3D digitization has become increasingly common and widely used in different fields of research (Aramendi et al., 2017; Bates et al., 2010; Buzi et al., 2018; Fau et al., 2016; Gaboutchian et al., 2020, 2019; Giacomini et al., 2019; Knyaz and Gaboutchian, 2016; Lussu and Marini, 2020; Medina et al., 2020; Moraes et al., 2014) including forensic sciences (Donato et al., 2020; Santoro et al., 2017; Thali et al., 2000; Villa et al., 2018). Nowadays photogrammetric 3D surface models are widely used for research, exhibition, and museum archiving. Their importance is also well known in medicine thanks to radiation-free image acquisition (Meulstee et al., 2017).

In this research, dental models will be compared to corresponding µ-CT models using point-to-point surface deviation analysis, which measures the distance between analogous points to each nearest point in the reference 3D point cloud, using CloudCompare software. Comparing different point clouds, it will be possible to calculate the metric differences between the clouds. The results of this study, therefore, may indicate the pros and cons of close-range photogrammetry, that may provide an affordable alternative to micro-CT scanning, although the latter offers high-resolution internal (as well as external) reconstruction of objects. The goal is to generate 3D surface models of isolated teeth, color and surface texture data with the accuracy and precision of the models and measurements.

Human teeth were chosen inasmuch they are currently highly studied in anthropology and human evolutionary studies (Clement et al., 2012; Hillson et al., 2010; Margherita et al., 2017). With advancements in imaging techniques, morphometric studies exploring teeth trait and development are emerging (Davies et al., 2019; Krenn et al., 2019). In this view, 3D model calibration becomes essential to achieve the precision of the measurement task. The small dimensions, the very irregular areas and surfaces characterized by micro cracks on enamel and cementum, have allowed us to evaluate the potential as well as the limits of the two techniques.

2. MATERIAL AND METHODS

2.1. Samples

Tooth samples were made available come from the multiple-phase cemetery (used since c.a. 8000 BC until 150 AD) of Al-Khiday (Khartoum, Central Sudan), located along the western bank of the White Nile (Usai et al., 2017, 2014; Usai and Salvatori, 2019). Two fragmented tooth samples of the Neolithic period (Grave95; Grave34) (**Figure.1**) belonging to two female individuals aged between

36-45 and 25-35 years, were selected (Iacumin et al., 2016). The upper molar sample 34T (Grave34) has an age-related worn complete crown, the sample 95T (Grave95) lower molar has no crown and retains both complete roots. Both specimens were selected as case studies to compare rough and heterogeneous surfaces.

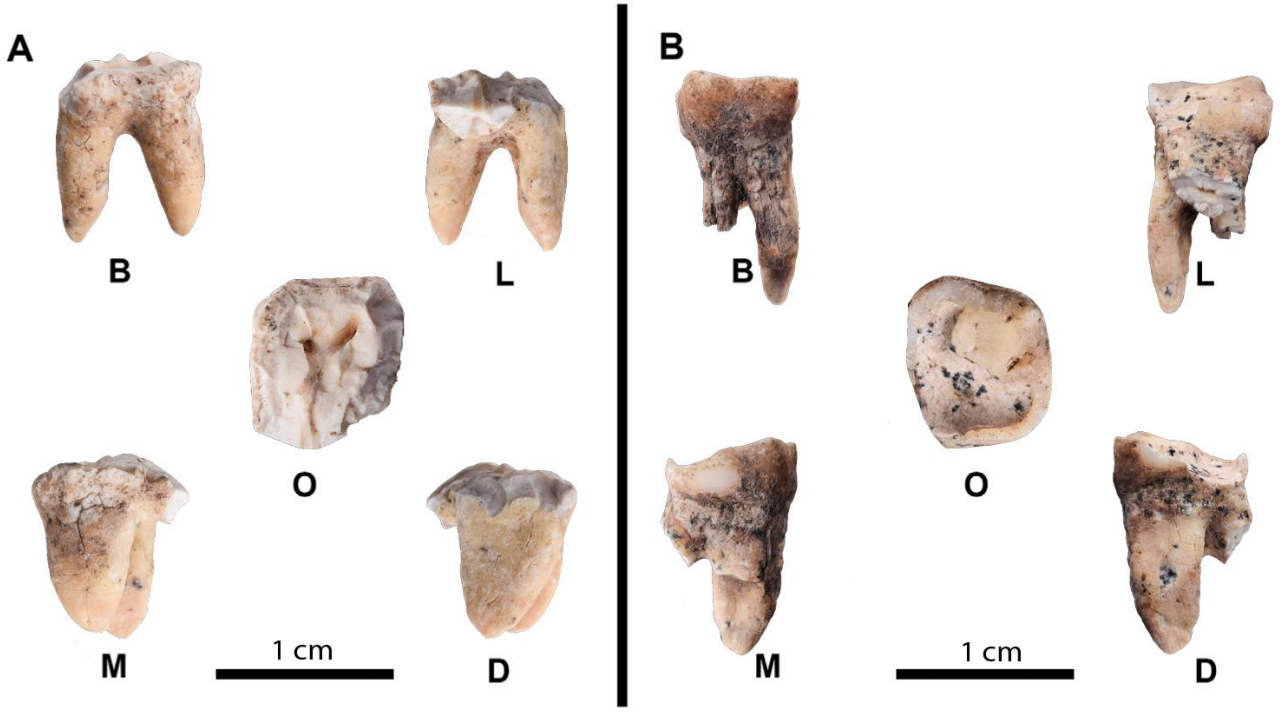

Figure 1. *Side views: buccal **B**; distal **D**; lingual **L**; mesial **M**; occlusal **O** of 95T (**A**) and 34T (**B**).*

2.2. Acquisition and elaboration micro-CT scanning

Samples were scanned at the Department of Geosciences of Padua University using a SkyScan® 1172 high-resolution computed micro-tomography (Salvo et al., 2003), equipped with a Cu X-ray tube and operating at 100 kV and 100 nA, with an Al filter 0.5 mm thick. No sample preparation was required. Radiographies to 360° rotations were acquired with an angular step of 0.30° and acquisition time of 1050 ms, a resolution of 256 Mp and pixel size of 8.96 µm for the 95T scan and of 8.00 µm for the 34T scan. Image stacks of the two scans were registered based on grey values and transformed to a 3D single file using CT-an® software (SkyScan® 1172) exported in .STL file format. Finally, in order to ensure comparability, point clouds of both resulting 3D surface models were reduced to 1.250.6000 in CloudCompare sub-sampling the clouds with *random* method and they were saved as .PLY file format. The two micro-CT reconstructed point clouds (95T and 34T) surfaces are constituted by 119.189 and 170.601 points, respectively.

2.3. Micro-photogrammetry and virtual reconstruction

The micro-photogrammetric technique Structure from Motion and Multiple-View Stereovision (SfM-MVS) was used to create high-resolution three-dimensional virtual models (Orekhova et al., 2019) of the teeth. Precise metrical models were generated using images taken with oblique photography using a Nikon D810 Reflex Pro camera with 36.3Mp FF (1.0x) sensor, macro lenses 60 mm, f/2.8 aperture value, shutter speed 1/20 seconds and ISO 400, with definition of 3696×2448 pixels (**Table 1**).

Table 1. *Technical specifications of the photographic sensor with macro-lens utilized.*

Nikon D810 Reflex Pro	
Type	Md-size SLR
Sensor	CMOS
Image size	Full frame (35.9x24mm)
Total pixels	36.3Mp
Focal length	60 mm
Focused distance to object	120-150 mm
Max resolution	7360x4912
Effective pixel	36 megapixels

Photos were acquired using a graduated turntable, light box, tripod, fixed light lamp. 96 and 116 photos were acquired for samples 34T and 95T, respectively. The number of photos varies depending on the morphology of the object: to ensure optimal 3D reconstruction, a larger sequence of photos was acquired for 95T due to its smaller dimensions. This has led to more shooting from the same point of grip with different focus plans, in order to focus on all the sample areas. Virtual reconstructions were performed from photographs that were taken following the norms of convergent close-range Photogrammetry (Gillies, 2015) from horizontal and vertical planes (**Figure 2**), using the guidelines for 3D reconstruction of small sized objects (Gallo et al., 2014).

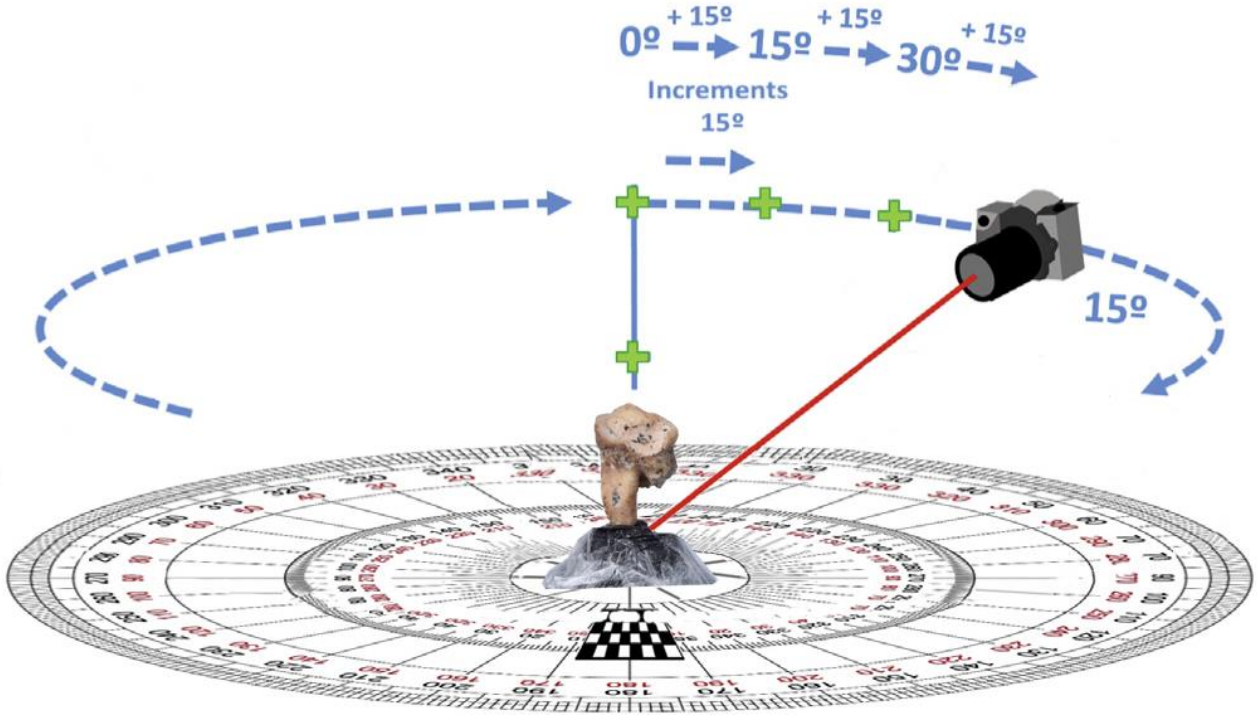

Figure 2. *Protocol for image capture utilized to acquired teeth samples by the Photogrammetric Method (modified from* Aramendi et al., 2017).

An almost complete overlap between photographs is necessary. Photos were collected maintaining the camera with an angle of 15° in the horizontal and the vertical axis while the light source remained constant during the image data acquisition. The photographs were taken in two planes perpendicular and parallel to occlusal plane. A tripod was used to stabilize the process and photographs were taken rotating the turntable. A manual exposure was performed with constant color calibration, aperture and sensor sensitivity (ISO) value, to achieve the same exposure times for all images to reduce variables during the photo shoot. Color temperature was measured by white manual balance (WB) as it is more accurate. Manual balance is used for color-critical photographs and to define with greater accuracy the original color rendering. Photographs were then processed to generate a 3D model through the dense point cloud for each sample with the photogrammetric reconstruction software 3DF Zephyr Free (3DFlow©): complete and closed photogrammetry system. Through mathematical methods, photogrammetry provides good information concerning the size, shape, and location of an object from one or more photographs (Gillies, 2015). The bundle triangulation algorithm is the primary algorithm for 3D reconstruction used in photogrammetry. It is a robust algorithm capable of simultaneously solving the coordinates of the perspective center, and determining the principal distance and principal point of images. In other words, when the positions of the camera are known, algorithm looks for common points in a set of photos to localize the 3D object in space (Luhmann et al., 2006). The reconstruction with Structure from Motion (SfM) photogrammetric process takes place to an unknown scale factor. To scale the 3D artefact metrically and make measurements, it is

necessary to introduce at least one distance constraint or control point constraints. A 3D model is generated as the point cloud; it has no real size, giving the user the ability to resize the object by scaling it (Moraes and Santandrea, 2020).

2.4. Digital and manual micro-scale measurements

As explained in the Introduction, photogrammetry is a technique allowing to obtain precise measurements of 3D models from photos. The precision of photogrammetry applied to small objects for optical dimensional metrology on the 3D model generated by photogrammetry was evaluated. Scaling of these 3D models was carried out by acquiring micro-measurements by Confocal laser scanning microscope (CLSM) and as many by digital caliper and magnifying glass, on the surface of both the dental samples. The small dimensions, the very irregular areas and surfaces characterized by micro cracks on enamel and cementum, have allowed us to evaluate the potential as well as the limits of the two tools (CLSM and caliper) compared with the measures acquired by micro-CT. Subsequently, the measurements were compared on 3DF Zephyr photogrammetric software, giving way to determine the precision of the taken measures.

Some known metric distance constraints were introduced to enforce scaling and orientation of the model. Six distances were taken for both specimens by a more optic accurate and versatile tools, such as Confocal Laser Scanning Microscopy (CLSM), model Olympus LEXT OLS4000 at the Department of Geosciences of Padua University, with a magnification capacity of 5x/0.15 N.A., W.D. 23.50 mm, FOV (HxV) 2.630×2.630μm and a spatial resolution of 10 nm axially and 120 nm laterally. The same distance points taken earlier were measured manually by digital caliper and magnifying glass. Both measurements were compared between them and with scale defined by μ-CT scan.

The goodness of the alignment was evaluated by the Final Mean residual reported by 3DF Zephyr by comparing the distances measured with the digital caliper and the confocal microscope. The similarity of the dimensions of the models generated are compared by ICP function (Iterative Closest Point, ICP) of 3DF Zephyr which made it possible to determine the precision of the model generated. ICP function is an algorithm employed to minimize the difference between two point clouds of points and considering translation, rotation and scale (3DF Zephyr User Manual-version 6.0, 2013). Using the function, it was possible to align two point clouds with different scale.

2.5. Comparison and distance calculation between two point clouds.

The alignment process of the 3D models (**Figure 3**) was performed using an open-source software, CloudCompare 2.12.alpha (Anoia) version. The micro-CT models were used as reference model in the entire process. CloudCompare software was chosen because it provides excellent registration and visualization features, as well as statistical tools for the evaluation.

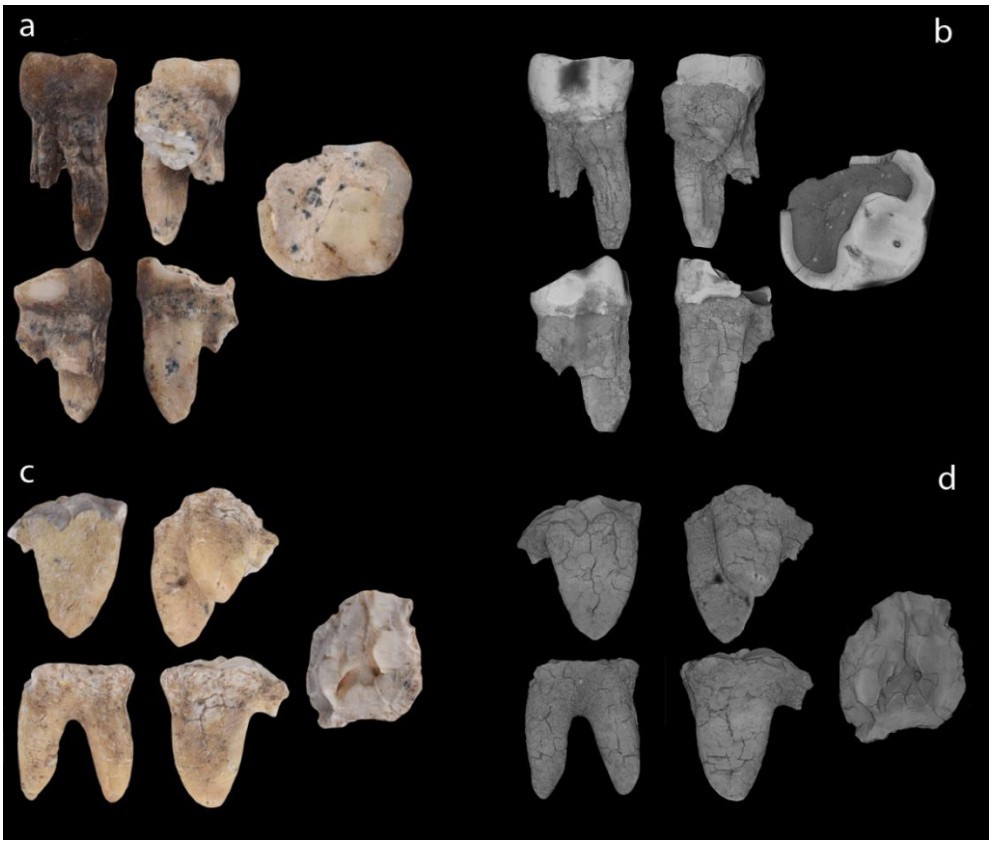

Figure 3. *On the left,* ***a)*** *photogrammetric three-dimensional digital model of LM1 (upper left first molar, 34T);* ***b)*** *three-dimensional reconstruction by X-µCT scanning of 34T;* ***c)*** *model of RM1 (lower right first root of molar, 95T);* ***d)*** *µCT model of 95T sample.*

One of the most common and utilized cloud distance computation is the Cloud-to-Cloud method (C2C method). C2C method is the computation of distances between two clouds or between a point cloud and a mesh (that is the lattice that defines an object in space, made up of vertices, edges and faces (Siddi F., 2014). In this study, the Cloud-to-Cloud method, aims to determine the distance difference between two cloud points acquired with micro-CT and macro-lens Photogrammetry (**Figure 4**).

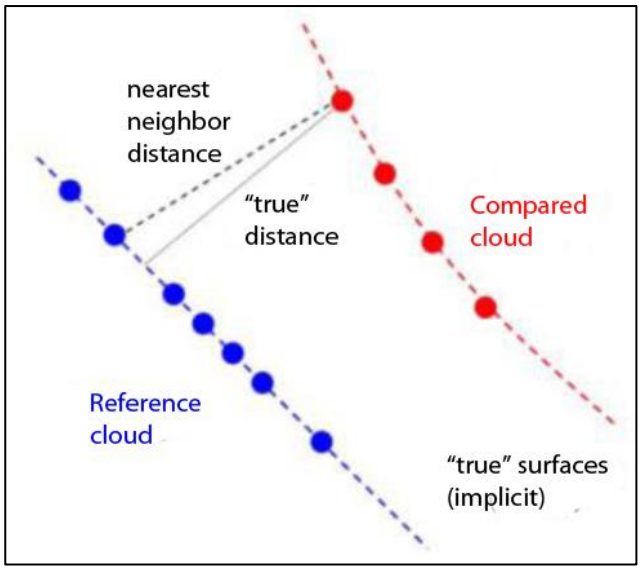

Figure 4.*The figure shows the basic concept of C2C computation method (reported by Ahmad Fuad et al., 2018).*

Micro-CT models were aligned in the same coordinate system with the corresponding photogrammetric model, modifying their "*boxes dimensions"*. At first, a rough alignment was performed by selecting four points pairs for both samples maintaining a final RMS below 1 mm. During the fine alignment phase (Iterative Closest Point, ICP) 20 iterations were used with a random sampling limit of 50.000 points with rotation and translation, but scaling adjustment was not allowed during alignment. Fine alignment was performed through the function "cloud registration" (ICP rms). To achieve precise alignment, the pairs of points were selected with different heights and distances along the surface. The quality of clouds alignment was monitored by the RMS (Root Mean Square), a statistical parameter computed in CloudCompare on the set of distances between a point cloud and another one, giving information about the distances distribution. The theoretical overlap between each 3D models was set to 100% for both clouds. Subsequently this preliminary alignment procedure, photogrammetric 3D surfaces comparison was performed with that generated by micro-CT by the "*compute cloud/cloud distance*" tool.

C2C distances computation method is based on *k*-nearest neighbours (kNN) searching algorithm. The distance between two points is defined using the Euclidean metric, widely used for finding *k*-nearest neighbours for each point in a reference cloud model with compared cloud dataset, removing the noise and computing the surface curvature (Ahmad Fuad et al., 2018; Shen et al., 2017). CloudCompare software provides three local modeling methods: Least Square Plane, 2D1/2 Triangulation and Quadric. The effectiveness of the local surface model is statistically dependent on how is appropriate the local surface approximation. The Quadric function can represent smooth and curvy surfaces but as being the most versatile local model option, it is the recommended function (Girardeau-Montaut, 2015). In this research, six nearby points (kNN) and seven octree level were used. Octree is an important CloudCompare function structure designed to speed up the processing of spatial data by decomposing the space into cubes. It has been chosen 7^{th} octree level dividing the structure in 2^{21} cubes storing sub-cubes that contain at least one cloud point (Girardeau-Montaut, 2015). Surfaces generated with photogrammetry were compared with those generated using micro-CT.

3. RESULTS

Known the pixels size of the micro-CT model, precise size along Z axis of the scanned objects were obtained. The dimensions calculated from micro-CT were 19.592 mm and 10.447 mm for 34T and 95T, respectively, which were taken as references. From measurements acquired by CLSM and by digital caliper, paralleled dimensions were obtained for 34T and 95T samples, measuring 19.2 mm and 10.0 mm (CLSM), respectively, while 20.0 mm and 9.90 mm were measured by caliper (**Table 2**).

Table 2. Comparison of results between model measurements.

Comparison	µ-CT (mm)	Digital caliper (mm)	CLSM (mm)
34T	19.592	20.0	19.2
95T	10.447	9.90	10.0

The sets of measurements acquired with the Confocal microscope and with the caliper were loaded on 3DF Zephyr to scale the models and the generated models reported different errors on model scale (Final Mean residual) on the points alignment (ICP). The Final Mean residual is a parameter of 3DF Zephyr, which shows the quality of the projection and alignment calculating the mean deviation between the position in the cloud scattered of each control point and its coordinates. When this value is lower than 1 then the orientation can be considered valid (3DF Zephyr User Manual-version 6.0, 2013). Final Mean residuals of 0.012 mm and 0.217 mm were recorded with the caliper and 0.010 mm and 0.015 mm with the CLSM for 95T and 34T, respectively.

All values are reported in **Table 3**.

Table 3. Comparison between 95T and 34T point clouds calibration (caliper vs. CLSM).

Comparison 3DF Zephyr	Digital caliper 34T (mm)	Digital caliper 95T (mm)	CLSM 34T (mm)	CLSM 95T (mm)	Caliper vs. CLSM
Finale Mean residual	0.217	0.012	0.015	0.010	
ICP error (34T)					0.085
ICP error (95T)					0.001

From the results obtained by ICP function of 3DF Zephyr, a chart was plotted that graphically explains the global residuals as histograms. From the comparison of the two cloud points (**Figure 5**) the ICP error calculated for 34T (**a**) and 95T (**b**) is 0.085 mm and 0.001 mm respectively (**Table 3**).

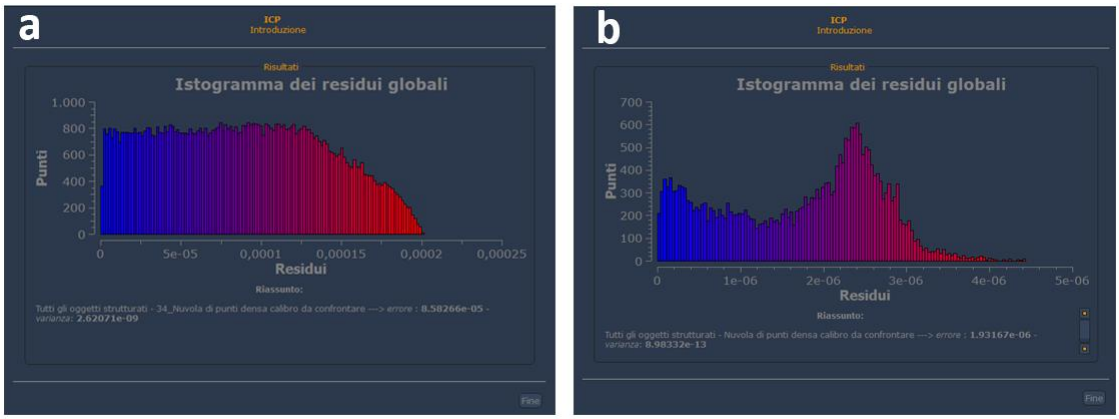

*Figure 5. Global residual histograms show the errors between resizing of 34T (**a**) and 95T (**b**).*

For both comparisons, cloud points resized by confocal microscope were used as references as it allows high precision measurements, both on smooth to very rough surface.

Considering the high spatial resolution of micro-CT (pixel size of c.a. 8 µm), the estimated error between the manually reported measurement and the micro-CT is 0.5 mm for both samples, which can be considered minimal. Quantitative and qualitative comparisons were conducted between two couple of 3D models at a time in order to compare the efficiency of the different acquisition 3D methods and surface model generation. In the comparison of surfaces generated with different methods, we have registered a fine overlap with final RMS of 0.24 mm for 34T and 0.08 mm for 95T. Absolute distances of C2C results shows cloud differences are represented in a color scale to show difference value (**Figure 6**).

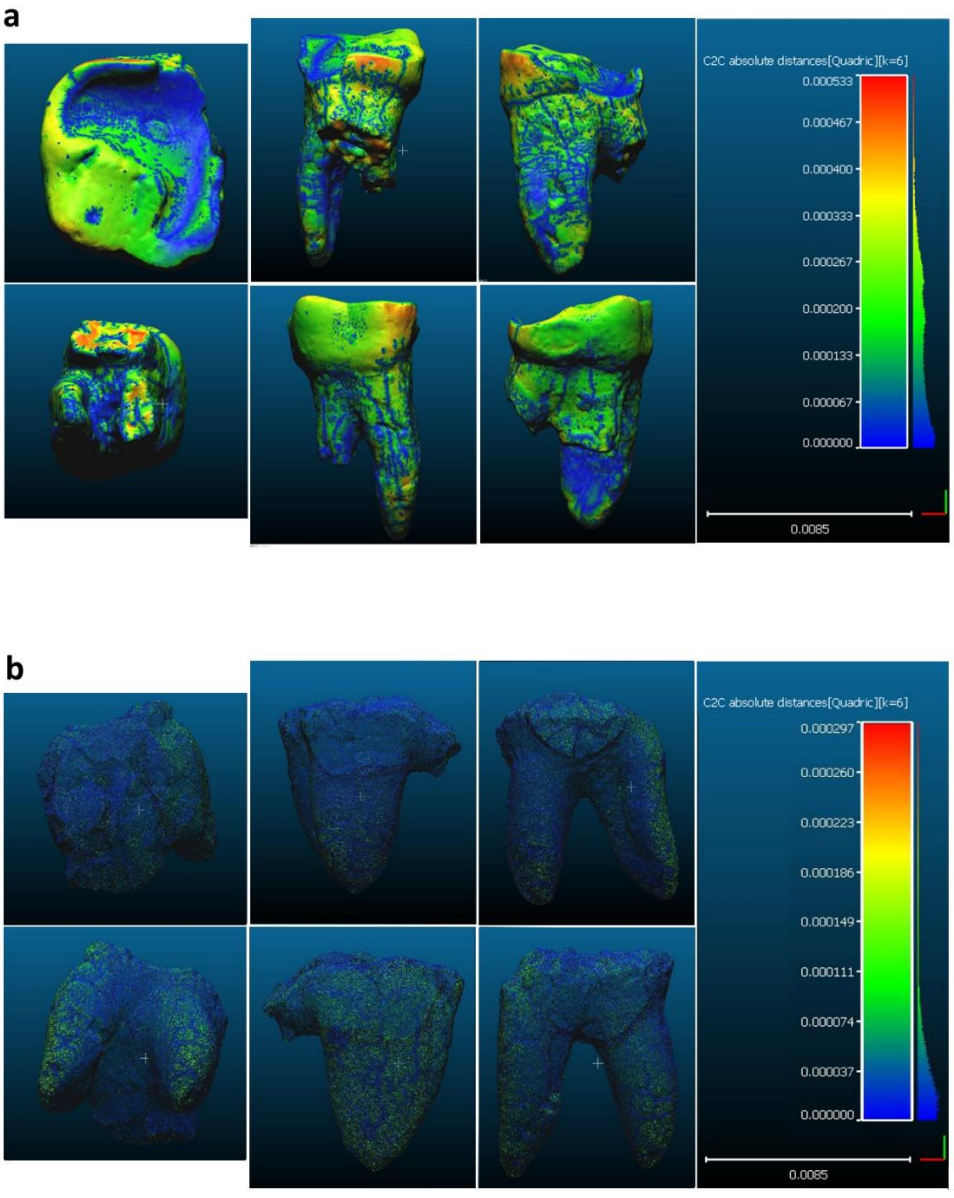

Figure 6.*Color map of 3D surface differences between micro-CT and photogrammetry models of 34T (**a**) and 95T (**b**) generated by CloudCompare.*

Areas with high accuracy are shown with positive value around zero (in blue), while in green color are the areas with less perfect overlap between two surfaces. The absolute values deviation expressed in red and yellow colors, determines that points in photogrammetry models lie above the reference micro-CT surface. The same color scale was applied for all comparisons. **Figure 6a** shows higher complete consistency with cloud references than 95T sample (**Figure 6b**). 34T model has a small area with large deviation illustrated in red color in correspondence with the root tip. After computing C2C absolute distances (errors), histogram of the resulted errors for both 3D objects were generated (shown in **Figure 7**) where in the Y axis is reported the points number (scalar value) per class (or bin) while the X axis spans between the min and max range of point distances. Final RMS error, mean distance and standard deviation are computed for C2C differences as reported in **Table 4** for each comparison. As shown in **Figure 7**, the histograms of C2C differences between two comparisons indicate that the distance distribution is different where the average C2C distances are always close to zero. The standard deviation is the statistical indicator that delivers information on C2C computed distance: narrower shape is always better as it means closeness to the mean value with low residual values, as visible in **Figure 7b**. In the comparison of surfaces 95T micro-CT model deviated with a mean distance from that photogrammetric by 0.34 mm on average and by 0.14 mm for 34T, where 95% of point clouds have errors less than ±0.08 mm for 95T and ±0.35 mm for 34T. Standard deviation calculated for both samples has lower value in 95T with a value of 0.02 mm than 34T with a value of 0.11 mm, illustrated in **Table 4**.

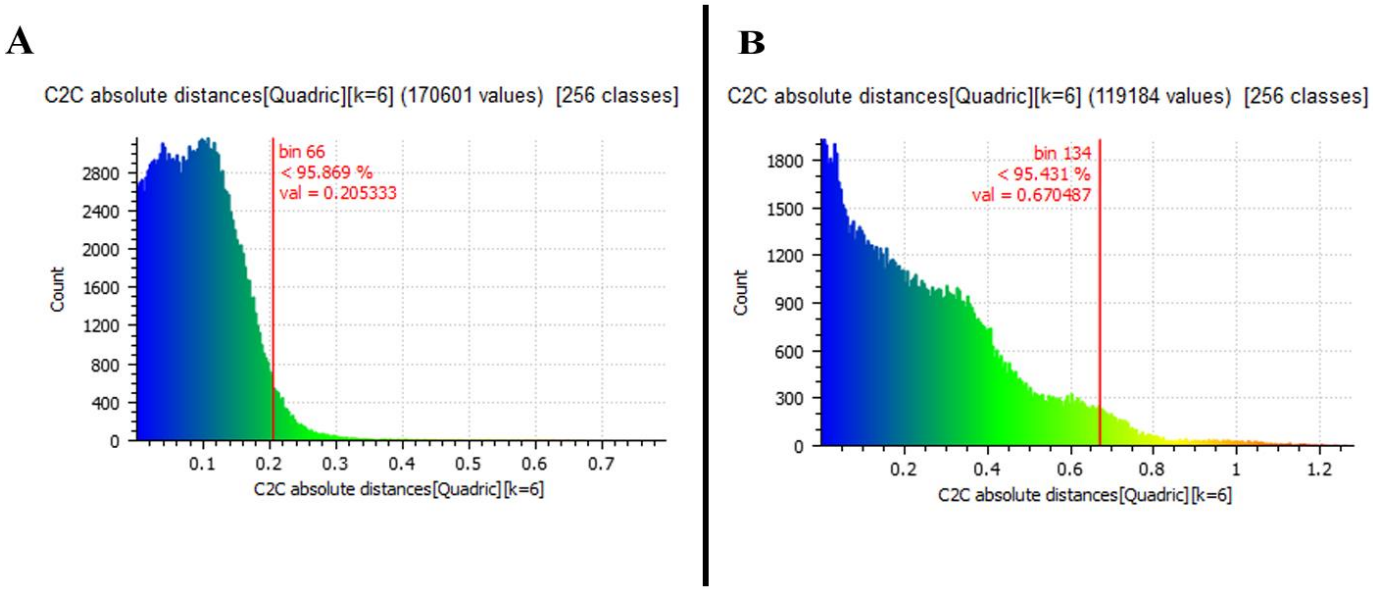

Figure 7. *Histograms of C2C differences for the 3D surface comparison test:* ***a****) 95T,* ***b****) 34T.*

Table 4. Descriptive statistics of C2C differences between two tested techniques.

Comparison C2C	Mean distance (mm)	St.deviation (mm)	95% val. (mm)	RMS error (mm)
34T	0.14	0.11	±0.35	0.24
95T	0.34	0.02	±0.08	0.08

4. DISCUSSION

Accuracy, precision and pros and cons in 3D surface generation by different tools are discussed. With accuracy in 3D modeling surface we refer to the highest possible degree of similarity between the micro-CT data and photogrammetric 3D model. One factor that greatly affects the accuracy of the virtual model is undoubtedly the image resolution. The higher the polygon count in the respective area of the surface, the more accurately the minute anatomical features are resolved. The total number of points influences resolution realistically. The data was collected with the highest possible quality using the best available resolution by the photographic equipment used. The amount of points in a cloud affects resolution, which is essential for a digital model. Due to the excessive size of original files, their digital manipulation was difficult in terms of file loading and model measurements. Regardless of point cloud quantity, model surfaces out of manual photogrammetric software Zephyr®, provide enough detail for morphological analysis. Even with 119.189 (95T) and 170.601 (34T) points on the photogrammetric 3D surface, morphological details were clear and resolved. The initial quality of images is not the only factor influencing the accuracy of the 3D model surface of the original specimen. In fact, another parameter that affects the quality of the model is the resizing of the object. The incorrect scaling, above all in small objects, can lead to invalid morphometric measurements altering the data. Only micro-CT scanning provides automatic scaling in that information of pixel sizes are provided in the log file. The automatic scale of the model does not happen with manual photogrammetry and for this reason an external measuring cup is inserted during the capture of images or are reported on the software with a caliper. Usually this manual operation involves a certain degree of error, leading to incorrect scaling of the object. Evaluating the mean distance between two surfaces for both objects, the reproducibility is good at around 0.08 mm at 95% of the surface. For all samples, the mean distance and the standard deviation between two techniques is substantially equivalent with a high degree of precision below millimeter. In our case, the greatest distances recorded on the 34T surface occurred on the crown mainly, whereas a better overlap could be found in areas not covered by enamel or with translucent surfaces. This is mainly because the enamel and dentine texture are more reflecting. This particular property of teeth dentine/enamel leads to affect the alignment results and therefore dense point cloud has noisy points in this particular area.

In the comparison of surfaces, RMS value is lower in the 95T sample despite its dimensions are more reduced and the rendering is less detailed with 119.184 points rather than 170.601points of 34T. Despite the small dimensions, the standard deviation is lower resulting in a higher quality in the definition of details.

The dense point clouds generated by photogrammetry are accurate and differ little (a few hundred microns) from the corresponding tomography.

Another potential of photogrammetry is the possibility of reuniting detached and fragmented parts, virtually restoring its original integrity. **Figure 8** shows the restoration result of the 95T sample, as the roots were detached from the crown.

Figure 8. *Virtual restoration of 95T sample.*

To match the root fragment to the tooth, virtual reconstruction was performed by matching two separate photogrammetric scans of the two divided objects. Once the 3D models were acquired, they were imported into a single modeling software (3DF Zephyr) to view them simultaneously, approaching them in virtual space and aligning their scan centers. Subsequently, by analyzing the original images, six pairs of points uniformly distributed on the respective contact surfaces of the two fragments were selected. The two 3D models, scaled and correctly oriented, were collimated by recording the control points (or contact points) and clouds of rototranslating points, until the spatial positions of the previously identified contact points coincide.

The latest innovation on photogrammetric field led to the fusion of the exterior information and interior scans processed by a single software. Significant samples can be observed in a complete way,

becoming a useful tool for preliminary investigation for researchers. Currently, a software on the market available is Astromaterials 3D software, developed by NASA Johnson Space Center, joining micro-CT slices and high-resolution precision photos with a resolution as fine as 30 to 60μm, creating interactive 3D models. This technology could be extended the study of objects insert in a Museum scientific database. Besides it can be extremely important to use photogrammetry to create databases of 3D virtual models in that they represent the possibility of explore the volumes, otherwise inaccessible, useful for explain odontometric analysis. In the field of paleoanthropology and paleontology, teeth play a fundamental role as elements for taxonomic discrimination and evolution studies (de Bonis and Viriot, 2002; Irish and Usai, 2021; Richard Scott and Irish, 2017; Smith et al., 2005). Over time, dental morphometry has developed with a view to increasing the information potential and repeatability factors of the measurements performed, to create more and more opportunities for comparing the results obtained by the various research groups. The obtained results showed that the inexpensive photogrammetric technique is sufficiently accurate and we recommend its use as an alternative tool to X-ray imaging to provide surface 3D information with the additional advantage of providing texture rendering.

ACKNOWLEDGMENTS

We thank Dr. Luca Penasa Ph.D., researcher at Department of Geosciences of the Padua University and Dr. Gregorio Dal Sasso Ph.D., researcher at the Institute CNR-IGG for their valuable scientific advices.

CONFLICT OF INTEREST

The authors do not have a conflict of interest to declare.

AUTHOR CONTRIBUTIONS

Cinzia Scaggion: conceptualization; data curation; formal analysis; investigation; writing-original draft.

Stefano Castelli: data curation; formal analysis; investigation; supervision; writing-review & editing.

Donatella Usai: resources; writing-review & editing.

Gilberto Artioli: project administration; supervision; writing-review & editing.

REFERENCES

Ackermann, S., Menna, F., Scamardella, A., Troisi, S., 2008. Digital photogrammetry for high precision 3D measurements in shipbuilding field. 6th CIRP Int. COnf. ICME 1, 6–11.

Ahmad Fuad, N., Yusoff, A.R., Ismail, Z., Majid, Z., 2018. Comparing the performance of point cloud registration methods for landslide monitoring using mobile laser scanning data. Int.

Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch. 42, 11–21. https://doi.org/10.5194/isprs-archives-XLII-4-W9-11-2018

Aramendi, J., Maté-González, M.A., Yravedra, J., Ortega, M.C., Arriaza, M.C., González-Aguilera, D., Baquedano, E., Domínguez-Rodrigo, M., 2017. Discerning carnivore agency through the three-dimensional study of tooth pits: Revisiting crocodile feeding behaviour at FLK- Zinj and FLK NN3 (Olduvai Gorge, Tanzania). Palaeogeogr. Palaeoclimatol. Palaeoecol. 488, 93–102. https://doi.org/10.1016/j.palaeo.2017.05.021

Bates, K., Falkingham, P., Rarity, F., 2010. Application of high-resolution laser scanning and photogrammetric techniques to data acquisition, analysis and interpretation in palaeontology. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XXXVIII, 68–73.

Buzi, C., Micarelli, I., Profico, A., Conti, J., Grassetti, R., Cristiano, W., Di Vincenzo, F., Tafuri, M.A., Manzi, G., 2018. Measuring the shape: Performance evaluation of a photogrammetry improvement applied to the Neanderthal skull Saccopastore 1. Acta IMEKO 7, 79–85. https://doi.org/10.21014/acta_imeko.v7i3.597

Clement, A.F., Hillson, S.W., Aiello, L.C., 2012. Tooth wear, Neanderthal facial morphology and the anterior dental loading hypothesis. J. Hum. Evol. 62, 367–376. https://doi.org/10.1016/j.jhevol.2011.11.014

Davies, T.W., Delezene, L.K., Gunz, P., Hublin, J.J., Skinner, M.M., 2019. Endostructural morphology in hominoid mandibular third premolars: Geometric morphometric analysis of dentine crown shape. J. Hum. Evol. 133, 198–213. https://doi.org/10.1016/j.jhevol.2019.06.004

de Bonis, L., Viriot, L., 2002. Teeth and paleoanthropology. Connect. Tissue Res. 43, 87–93. https://doi.org/10.1080/03008200290001032

Donato, L., Cecchi, R., Goldoni, M., Ubelaker, D.H., 2020. Photogrammetry vs CT Scan: Evaluation of Accuracy of a Low-Cost Three-Dimensional Acquisition Method for Forensic Facial Approximation. J. Forensic Sci. 65, 1260–1265. https://doi.org/10.1111/1556-4029.14319

Duval, M., Martín-Francés, L., 2017. Quantifying the impact of µCT-scanning of human fossil teeth on ESR age results. Am. J. Phys. Anthropol. 163, 205–212. https://doi.org/10.1002/ajpa.23180

Fau, M., Cornette, R., Houssaye, A., 2016. Apport de la photogrammétrie à la numérisation 3D d'os de spécimens montés : potentiel et limites. Comptes Rendus - Palevol 15, 968–977. https://doi.org/10.1016/j.crpv.2016.08.003

Gaboutchian, A., Knyaz, V., Vasilyev, S., Maksimov, A., 2020. 3D data analysis for teeth morphology study. CEUR Workshop Proc. 2744, 1–10. https://doi.org/10.51130/graphicon-

2020-2-3-38

Gaboutchian, A. V., Knyaz, V.A., Leybova, N.A., Petrosyan, G., Simonyan, H., Vasilyev, S. V., 2019. Application of photogrammetric techniques in palaeoodontological studies trough automated digital shape analysis of human teeth. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch. 42, 75–80. https://doi.org/10.5194/isprs-archives-XLII-2-W12-75-2019

Gallo, A., Muzzupappa, M., Bruno, F., 2014. 3D reconstruction of small sized objects from a sequence of multi-focused images. J. Cult. Herit. 15, 173–182. https://doi.org/10.1016/j.culher.2013.04.009

Giacomini, G., Scaravelli, D., Herrel, A., Veneziano, A., Russo, D., Brown, R.P., Meloro, C., 2019. 3D Photogrammetry of Bat Skulls: Perspectives for Macro-evolutionary Analyses. Evol. Biol. 46, 249–259. https://doi.org/10.1007/s11692-019-09478-6

Gillies, D., 2015. Close Range Photogrammetry, Photogrammetric Record. https://doi.org/10.1111/phor.12114

Girardeau-Montaut, D., 2015. CloudCompare - User manual. Webpage: http://www.cloudcompare.org.

Grieshaber, B.M., Osborne, D.L., Doubleday, A.F., Kaestle, F.A., 2008. A pilot study into the effects of X-ray and computed tomography exposure on the amplification of DNA from bone. J. Archaeol. Sci. 35, 681–687. https://doi.org/10.1016/j.jas.2007.06.001

Hall, A.C., Sherlock, E., Sykes, D., 2016. Does Micro-CT scanning damage DNA in museum specimens ? Authors Date Submitted Does Micro-CT scanning damage DNA in museum specimens ?

Hallert, B.P., 1963. Accuracy and Precision in Photogrammetry. Photogramm. Eng. 29, 108–112.

Hershkovitz, I., Duval, M., Grün, R., Mercier, N., Valladas, H., Ayalon, A., Bar-Matthews, M., Weber, G.W., Quam, R., Zaidner, Y., Weinstein-Evron, M., 2018. Response to Comment on “The earliest modern humans outside Africa.” Science (80-.). 362. https://doi.org/10.1126/science.aat8964

Hillson, S.W., Parfitt, S.A., Bello, S.M., Roberts, M.B., Stringer, C.B., 2010. Two hominin incisor teeth from the middle Pleistocene site of Boxgrove, Sussex, England. J. Hum. Evol. 59, 493–503. https://doi.org/10.1016/j.jhevol.2010.06.004

Iacumin, P., Di Matteo, A., Usai, D., Salvatori, S., Venturelli, G., 2016. Stable isotope study on ancient populations of central sudan: Insights on their diet and environment. Am. J. Phys. Anthropol. 160, 498–518. https://doi.org/10.1002/ajpa.22987

Immel, A., Le Cabec, A., Bonazzi, M., Herbig, A., Temming, H., Schuenemann, V.J., Bos, K.I., Langbein, F., Harvati, K., Bridault, A., Pion, G., Julien, M.A., Krotova, O., Conard, N.J.,

Münzel, S.C., Drucker, D.G., Viola, B., Hublin, J.J., Tafforeau, P., Krause, J., 2016. Effect of X-ray irradiation on ancient DNA in sub-fossil bones - Guidelines for safe X-ray imaging. Sci. Rep. 6, 1–14. https://doi.org/10.1038/srep32969

Irish, J.D., Usai, D., 2021. The transition from hunting-gathering to agriculture in Nubia: Dental evidence for and against selection, population continuity and discontinuity. Proc. R. Soc. B Biol. Sci. 288. https://doi.org/10.1098/rspb.2021.0969

James, D.W., Eckermann, J., Belblidia, F., Sienz, J., 2015. Point cloud data from Photogrammetry techniques to generate 3D Geometry. Proc. 23rd UK Conf. Assoc. Comput. Mech. Eng. 8 5.

Knyaz, V.A., Gaboutchian, A. V., 2016. Photogrammetry-based automated measurements for tooth shape and occlusion analysis. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch. 41, 849–855. https://doi.org/10.5194/isprsarchives-XLI-B5-849-2016

Krenn, V.A., Fornai, C., Wurm, L., Bookstein, F.L., Haeusler, M., Weber, G.W., 2019. Variation of 3D outer and inner crown morphology in modern human mandibular premolars. Am. J. Phys. Anthropol. 169, 646–663. https://doi.org/10.1002/ajpa.23858

Luhmann, T., Robson, S., Kyle, S., Harley, I., 2006. Introducción, Close Range Photogrammetry: Principles, techniques and applications.

Lussu, P., Marini, E., 2020. Ultra close-range digital photogrammetry in skeletal anthropology: A systematic review. PLoS One 15, 1–29. https://doi.org/10.1371/journal.pone.0230948

Margherita, C., Oxilia, G., Barbi, V., Panetta, D., Hublin, J.J., Lordkipanidze, D., Meshveliani, T., Jakeli, N., Matskevich, Z., Bar-Yosef, O., Belfer-Cohen, A., Pinhasi, R., Benazzi, S., 2017. Morphological description and morphometric analyses of the Upper Palaeolithic human remains from Dzudzuana and Satsurblia caves, western Georgia. J. Hum. Evol. 113, 83–90. https://doi.org/10.1016/j.jhevol.2017.07.011

Medina, J.J., Maley, J.M., Sannapareddy, S., Medina, N.N., Gilman, C.M., McCormack, J.E., 2020. A rapid and cost-effective pipeline for digitization of museum specimens with 3D photogrammetry. PLoS One 15, 1–14. https://doi.org/10.1371/journal.pone.0236417

Meulstee, J.W., Verhamme, L.M., Borstlap, W.A., Van der Heijden, F., De Jong, G.A., Xi, T., Bergé, S.J., Delye, H., Maal, T.J.J., 2017. A new method for three-dimensional evaluation of the cranial shape and the automatic identification of craniosynostosis using 3D stereophotogrammetry. Int. J. Oral Maxillofac. Surg. 46, 819–826. https://doi.org/10.1016/j.ijom.2017.03.017

Moraes, C., Santandrea, G., 2020. Protocollo di tomografia per la pianificazione della chirurgia ortognatica. https://doi.org/10.6084/m9.figshare.12923726

Moraes, C.A. da C., Dias, P.E.M., Melani, R.F.H., 2014. Demonstration of protocol for computer-

aided forensic facial reconstruction with free software and photogrammetry. J. Res. Dent. 2, 77. https://doi.org/10.19177/jrd.v2e1201477-90

Orekhova, E.A., Romashkin, T. V., Ustinova, N.G., 2019. Synergetic Efficiency of the Economic System in the Conditions of Digital Economy Development: Conceptual Approach. https://doi.org/10.2991/iscfec-18.2019.45

Richard Scott, G., Irish, J.D., 2017. Human Tooth Crown and Root Morphology: The Arizona State University Dental Anthropology System.

Salvo, L., Cloetens, P., Maire, E., Zabler, S., Blandin, J.J., Buffière, J.Y., Ludwig, W., Boller, E., Bellet, D., Josserond, C., 2003. X-ray micro-tomography an attractive characterisation technique in materials science. Nucl. Instruments Methods Phys. Res. Sect. B Beam Interact. with Mater. Atoms 200, 273–286. https://doi.org/10.1016/S0168-583X(02)01689-0

Santoro, V., Lubelli, S., De Donno, A., Inchingolo, A., Lavecchia, F., Introna, F., 2017. Photogrammetric 3D skull/photo superimposition: A pilot study. Forensic Sci. Int. 273, 168–174. https://doi.org/10.1016/j.forsciint.2017.02.006

Shen, Y., Lindenbergh, R., Wang, J., 2017. Change analysis in structural laser scanning point clouds: The baseline method. Sensors (Switzerland) 17, 1–25. https://doi.org/10.3390/s17010026

Sholarin, E.A., Awange, J.L., 2015. Photogrammetry. Environ. Sci. Eng. (Subseries Environ. Sci. 213–230. https://doi.org/10.1007/978-3-319-27651-9_10

Smith, J.B., Vann, D.R., Dodson, P., 2005. Dental morphology and variation in theropod dinosaurs: Implications for the taxonomic identification of isolated teeth. Anat. Rec. - Part A Discov. Mol. Cell. Evol. Biol. 285, 699–736. https://doi.org/10.1002/ar.a.20206

Szeliski, R., 2009. Computer Vision : Algorithms and Applications Chapter 3 Image processing. Comput. Vis. Algorithms Appl. 110.

Thali, M.J., Braun, M., Brüschweiler, W., Dirnhofer, R., 2000. Matching tire tracks on the head using forensic photogrammetry. Forensic Sci. Int. 113, 281–287. https://doi.org/10.1016/S0379-0738(00)00234-6

Usai, D., Maritan, L., Dal Sasso, G., Artioli, G., Salvatori, S., Jakob, T., Salviato, T., 2017. Late pleistocene/early holocene evidence of prostatic stones at al khiday cemetery, Central Sudan. PLoS One 12, 1–12. https://doi.org/10.1371/journal.pone.0169524

Usai, D., Salvatori, S., 2019. The Mesolithic period on the White Nile region and the al-Khiday sites. Azania 54, 445–467. https://doi.org/10.1080/0067270X.2019.1691846

Usai, D., Salvatori, S., Jakob, T., David, R., 2014. The Al Khiday Cemetery in Central Sudan and its “Classic/Late Meroitic” Period Graves. J. African Archaeol. 12, 183–204.

https://doi.org/10.3213/2191-5784-10254

Villa, C., Flies, M.J., Jacobsen, C., 2018. Forensic 3D documentation of bodies: Simple and fast procedure for combining CT scanning with external photogrammetry data (Journal of Forensic Radiology and Imaging (2018) 12 (e1)(S2212478017300904)(10.1016/j.jofri.2017.11.002)). J. Forensic Radiol. Imaging 12, e2–e7. https://doi.org/10.1016/j.jofri.2017.11.003

CHAPTER 2

Evaluation of the influence of sample preparation for Infrared spectroscopy measurements of bone specimens for a minimally invasive method

Cinzia Scaggion[1*], Luca Nodari[2], Luca Pagani[3], Nicola Carrara[4], Gilberto Artioli[1].

[1] Department of Geosciences, University of Padova, Padova, 35131, Italy
[2] Italian National Research Council-CNR, ICMATE, Padova, 35127, Italy
[3] Department of Biology, University of Padova, Padova, 35122, Italy
[4] Museum of Anthropology, University of Padova, Padova, 35121, Italy

*corresponding author, e-mail address: cinzia.scaggion@phd.unipd.it (C. Scaggion)

Keywords: FTIR spectroscopy; pre-screening method; sampling preparation;

Abstract

Following the development of modern technologies, the investigation of museum osteological finds is increasingly required. One of the goals of this project is to develop viable protocols to preserve these collections from exceedingly invasive analysis, allowing access to the specimens for scientific research. The measure of the degradation of bone bioapatite, using crystallinity and atomic order/disorder parameters based on FTIR analysis, could provide an adequate preliminary evaluation of the bone conservation. Basically, it was necessary to seek new methodologies of extraction of a minimal quantity of material and to avoid contaminations. As a first step, all experimental sampling strategies, including reproducibility and minimal invasiveness, have been carefully calibrated in order to obtain reliable and safe procedures. The samples have been characterized by Fourier transform infrared spectroscopy (FTIR) techniques and subsequently tested for ancient DNA (aDNA) extraction.

1. INTRODUCTION

The evaluation of the altered state of the bones was found to be crucial for the understanding of the involved diagenetic processes.

Diagenesis is a complex and multifactorial phenomenon causing various changes in the physicochemical properties of the bone, and mechanical and histological alterations that happen at different extent. During burial, numerous alterations can occur such as absorption and exchange of ions, leaching, microbiological attacks with varying degrees of intensity depending on the local geochemical conditions of the burial site (Hedges, 2002; Hedges and Millard, 1995; Jans, 2008; Reiche et al., 2003; Trueman, 2013; Trueman et al., 2004). The diagenesis of archaeological bones is a widely discussed topic in the literature (Collins et al., 2002; Hedges, 2002; Keenan, 2016; Kendall et al., 2018; Weiner, 2010), and several studies were conducted to determine how and in which way diagenetic processes affect the conservation (Dal Sasso et al., 2014; Hedges, 2002; Lebon et al., 2010; Lee-Thorp and Sealy, 2008; Nielsen-Marsh and Hedges, 2000; Smith et al., 2007).

The geochemical aspects of bone diagenesis were studied through a variety of analytical techniques (Jans et al., 2004; Stathopoulou et al., 2008; Trueman and Tuross, 2019). Lately, Infrared

spectroscopy (FTIR) is one of the most used techniques for this purpose, as it is extremely sensitive, inexpensive and it requires minimal amounts of bone material to be analysed (~1 mg). Furthermore, IR spectroscopy simultaneously provides chemical-physical information related to the organic and inorganic components of the bone.

Bone is a heterogeneous composite material consisting of an intimate association between bioapatite crystallites (carbonate hydroxylapatite), collagen fibrils and water (Traub et al., 1992; Weiner, 2010). Several spectroscopic parameters were proposed to describe the diagenetic changes affecting the conservation state of bone from the macro to the nanoscale (Dal Sasso et al., 2018, 2014; Kontopoulos et al., 2020, 2019; Le Meillour et al., 2018; Lebon et al., 2016, 2010; Leskovar et al., 2020; Weiner and Traub, 1992). Among these, the most used parameter is related to the indirect monitoring of the bioapatite crystallinity. Crystallinity provides information about the state of conservation of the bone mineral as it undergoes dissolution/recrystallization processes during diagenesis (Berna et al., 2004; Weiner, 2010; Wopenka and Pasteris, 2005). Bioapatite is a nanocrystalline phase departing from the stoichiometric geological hydroxylapatite $Ca_{10}(PO_4)_6(OH)$, allowing a variety of structural ion substitutions and vacancies (Cacciotti, 2016; Foley et al., 2020; LeGeros, 1981). Among these, the most common substitution consists of about 5-8% weight of carbonate replacing phosphate groups (Type B) and to a lower extent hydroxyl ions (Type A) (Pan and Darvell, 2010; Pasteris et al., 2004; Wopenka and Pasteris, 2005). The occurrence of such substitutions in the bioapatite structure introduces a structural disorder in the crystallites resulting in changes in the physical-chemical properties, such as solubility and crystallite size (LeGeros, 1981; Wopenka and Pasteris, 2005). The loss of carbonate species within the bone apatite structure causes the formation of a thermodynamically more stable and more ordered crystalline structure, leading to the increase of the crystallites size and lower solubility (Berna et al., 2004; Greiner et al., 2019; Hedges, 2002; Nielsen-Marsh and Hedges, 2000).

The most used index to quantitatively monitor and describe the changes in bone crystallinity using infrared spectroscopy is the splitting factor (**IRSF**). The splitting factor quantifies the separation between two peaks calculated by the sum of two peak intensities of ν_4 (PO_4^{3-}) divided by the intensity of the valley between them (Weiner and Bar-Yosef, 1990). Another recently developed parameter (Dal Sasso et al., 2018) is the infrared peak width at 85% intensity of the 604 cm^{-1} (**FW85%**) that can be a useful tool to reliably estimate the minimal changes in the chemical composition and structure of bioapatite (**Figure.1**).

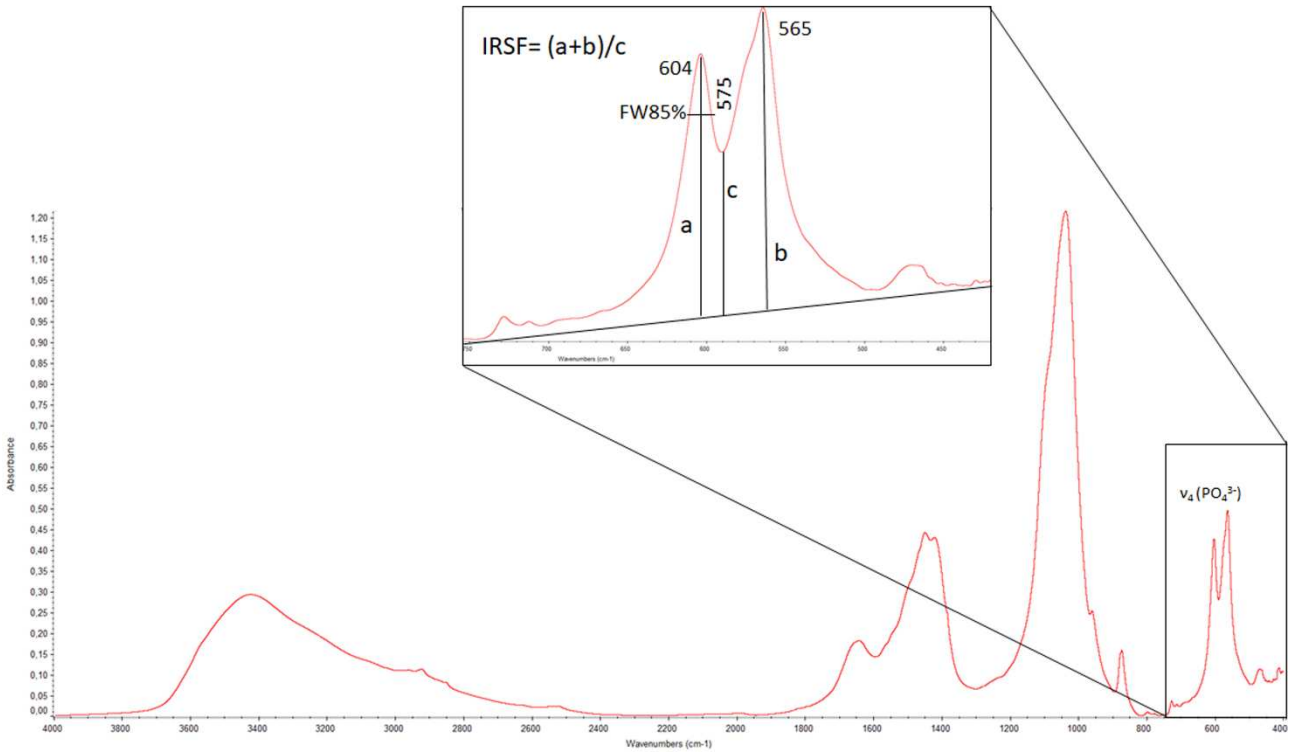

Figure 1. *Diagenetic parameters (IRSF and FW85%) calculated from FTIR spectra.*

Therefore, peak width could be the most promising parameter to indicate the degree of bone alteration, as it is closely related to the atomic order/disorder in the crystal lattice.

The archaeological skeletal tissues of animals and humans often contain valuable information for different research fields (Allende and Samplonius, 2022; Dupras and Schwarcz, 2001; Karakostis et al., 2021; Kieser et al., 2007; Kimmerle et al., 2008; Mays et al., 2001; Minozzi et al., 2012; Prowse et al., 2004; Raichlen et al., 2020).

The osteological finds of historical, archaeological, scientific and cultural interest are part of the so-called "biological assets", commonly preserved in museums, and are therefore important and protected in the context of the Conservation of Cultural Heritage (Licata et al., 2020). Bones, being real ”biobanks”, are increasingly required for scientific purposes (Brown and Brown, 2011). Sampling, which is often damaging and destructive, can be a serious problem when only limited bone samples are available. Furthermore, besides the evident aesthetical damage, sampling may compromise the morphological studies, which provide important information on the estimate of the age of individuals, but also dietary, cultural and evolutionary information (Balzeau, 2015; Brace et al., 1991; Demes and Creel, 1988; Doden and Halves, 1984; Oxilia et al., 2018; Rathmann and Reyes-Centeno, 2020).

The objectives of this preliminary study are: 1) the development of a micro-invasive and effective sampling method on bone finds of museum interest; 2) the application of a standard preparation of

bone powder for FTIR analysis, which significantly improves the accuracy, the reproducibility as well as the reliability of the data, allowing a systematic evaluation of bone in the archaeological and bio-anthropological fields; 3) the direct evaluation of the degree of *post-mortem* recrystallization of the bone bioapatite. For this purpose, separated petrous bones were selected inasmuch they are the most requested anatomical elements for genetic analysis and often are not easily accessible in complete *crania*.

2. MATERIALS AND METHODS

2.1. Bone specimen

Two human petrous bones (IND.A and IND.D) were selected from the graves of Oderzo (Southern Italy), preserved in the Anthropology Museum of Padua University (Italy), (**Table.1**).

SAMPLE	CHRONOLOGY	ARCHAEOLOGICAL SITE	SKELETAL ELEMENT
IND.A	2^{nd}- 3^{rd} c.AD	Orderzo, Italy	petrous bone
IND.D	2^{nd}- 3^{rd} c.AD	Oderzo, Italy	petrous bone

Table 1. *Information on the included samples: label, skeletal anatomic element, chronology, origin.*

Unfortunately, anthropological data have been only partially published (Drusini et al., 1987). In the ancient *Opitergium* (today's Oderzo), dated to the pre-Roman and Roman period, both the rituals of inhumation and of cremation are certified (Cipriano and Sandrini, 2015). In this study, only bones from inhumated burials with were analysed since they were not compromised by the high temperatures leading to collagen denaturation during cremation (Wagermaier and Fratzl, 2012). The Necropolis of Oderzo, found close to the western border of the city during the archaeological excavation in 1986, was located on the river bank of the extinct Lia river (Cipriano and Sandrini, 2001). The burial is typical of riverine conditions.

2.2. micro-CT investigation for a minimally invasive bone sampling

The sample INDIV.D was selected for its good state of conservation.

Through the preliminary MicroCT analysis, using a cone-beam X-ray geometry, a stack of sequential two-dimensional projections of the specimen was acquired and virtually reconstructed to generate a high-resolution three-dimensional image of the object (Elliott and Dover, 1982). Image analysis software (CTAn v.1.18 Micro-CT software) was used to determine the standard morphometric indices of bone establishing the most effective point to sample without compromising the integrity of the same sample.

These parameters were used to verify the area with the highest bone density of the petrous bone, aiming at the precise extraction of the powder for FTIR analysis (or micro-core), without damaging the *otic capsule*, preserving very precious material for genetic analysis (**Figure.2**).

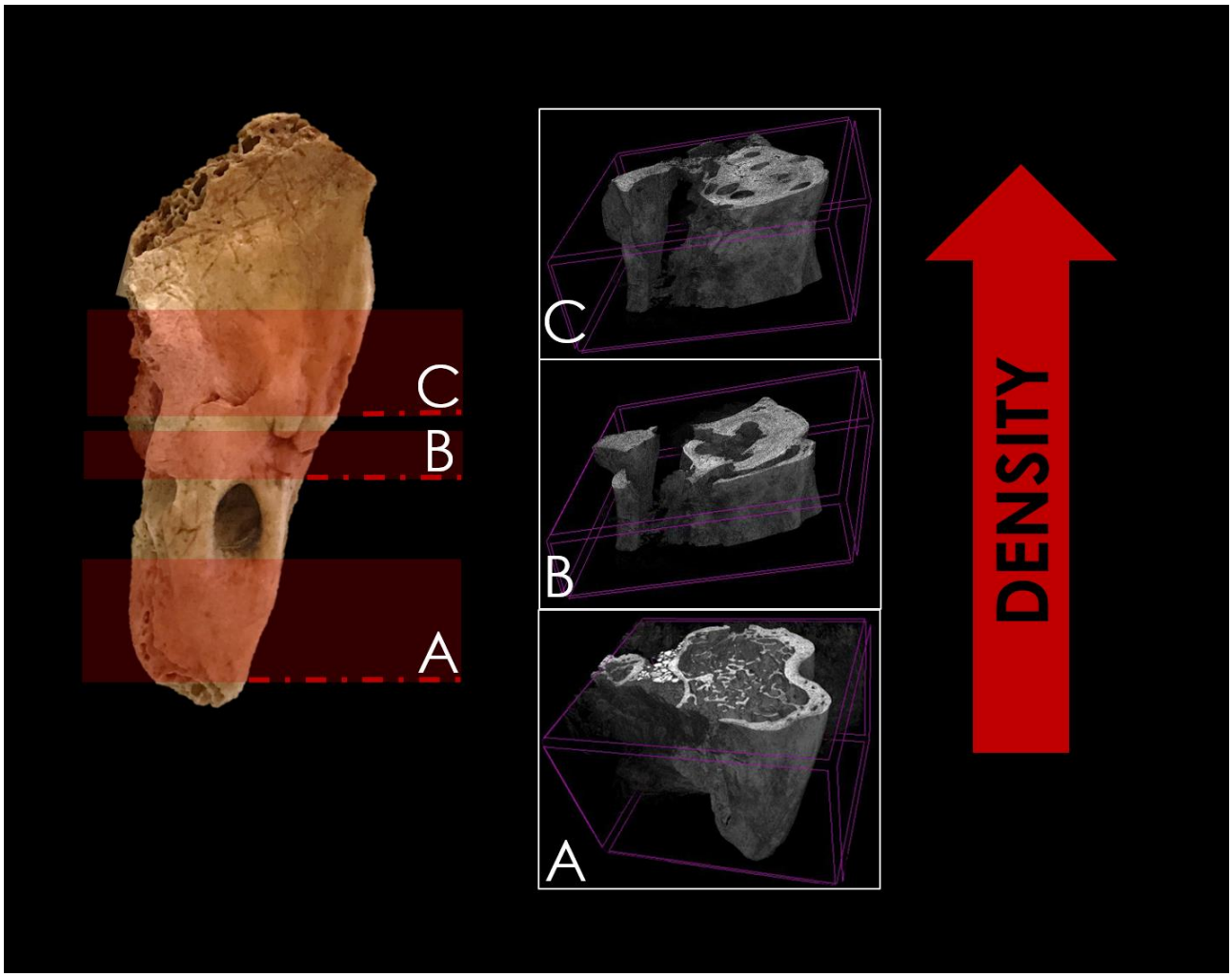

Figure 2. *Petrous bone at the X-µCT examination in cross section before micro-sampling and identification of three areas characterized by different densities.*

On these bases, a procedure for powder extraction was set and applied on the same archaeological find by Dremel Micro Rotary drilling with two different micro-tools. At first, the external bone surface, encrusted by sediments, was mechanically removed by means of low-speed micro-drill, equipped with an abrasive round point of 2.4 mm. Another helical micro-tool (2.4 mm) at moderate speed was used to get the necessary quantity of bone powder. Particular attention was paid to avoid the overheating of the samples during the mechanical cleaning process.

The operating procedure for standard sampling of bone of the ICMP (ICMP, 2015) was followed to disinfect all drill bits to avoid external contamination.

2.3 Infrared spectroscopy analysis

IND.A sample was used to extract enough powder (~1 mg for each sampling) to execute a series of FTIR analyses. Infrared spectroscopy provides a sensitive probe for specific functional groups in organic and inorganic compounds; it is widely used not only to obtain qualitative information (Rubinson and Rubinson, 2002) but also semi-quantitative ones (Weiner, 2010) in order to investigate the nature as well as the atomic order/disorder of mineral and organic components in the analysed samples. Therefore, FTIR spectroscopy can provide valuable information on the structural and chemical properties of archaeological bones and, consequently on their preservation (Kontopoulos et al., 2019; Poralan et al., 2015; Weiner and Bar-Yosef, 1990).

We sought a protocol for sample preparation to guarantee the accuracy, and replicability of the FTIR spectra. Powdered sample diluted in 100mg of KBr was then pressed, using a hydraulic press, under 11 tons/cm^2 pressure, and a 12 mm in diameter and 1.5 mm of thickness transparent pellet was obtained. Spectra were collected with a Nicolet 380 FTIR spectrometer equipped with a DTGS detector; 128 scans for each spectrum were acquired, in the range from 4000 to 400 cm^{-1}, with a spectral resolution of 4 cm^{-1}. Spectral analysis was performed using Omnic 9 software (Thermo Scientific).

Although previous studies have reported the changes in crystallinity caused by the grinding intensity (Asscher et al., 2011; Lebon et al., 2010), in this work the effect of sample preparation was evaluated by investigating the percentage of dilution (bone/KBr ratio) and time/extent of grinding. IRSF, as defined by Weiner and Bar-Yosef (1990), and FW85% defined by Dal Sasso et al. (2018) of the 604 cm^{-1} peak were calculated for each spectrum. The baseline was defined by two points calculated as the local minimum in the region 850–620 cm^{-1} and 510–470 cm^{-1} (**Figure.1**).

2.3.1 Experiment 1: grinding time

Bone powder from IND.A was taken from the petrous bones by first removing the external sediments to avoid exogenous contamination coming from the soil of deposition as reported in paragraph 2.2. Four pellets were prepared keeping the sample/KBr ratio at 1:100 mg. Each pellet was prepared by hand-grinding the sample with KBr in an agate mortar at different times (30 seconds; 2 minutes; 5 minutes and 10 minutes) in order to determine the optimal grinding time to guarantee reproducible peak shape and to minimize the effects of particle size and homogenisation (Asscher et al., 2017). The pellets were pressed for 1 minute and analysed with the infrared spectroscopy. One pellet was prepared with ratio 1:100 mg of KBr and, after the first measurement, the pellet was reground to create a new pellet and probed again with different grinding times for 30 seconds, 5 and 10 minutes.

2.3.2 Experiment 2: dilution

The external bone surface, encrusted by sediments, was mechanically removed by means of a low-speed micro-drill equipped with an abrasive point (paragraph *2.2*). Attention was paid to avoid the overheating of the samples during the mechanical cleaning process. One microgram of sample was hand ground in an agate mortar while the potassium bromide (KBr) was added and homogenised with it. Thus, three pellets were prepared by increasing the amount of KBr (bone powder/KBr ratio = 1:100 mg; 1:300 mg; 1:700 mg) and grinding constantly for 2 minutes and pressing for about 1 minute.

2.3.3 Experiment 3: sampling point and depth

Through the same procedures, six different sampling points were chosen on the same petrous bone (IND.A) at different depths: 2 mm; 4 mm; 15 mm (**Figure.3**). The extracted bone powder was measured through FTIR spectroscopy preparing pellets keeping the sample/KBr ratio at 1:100 mg and hand grinding for 2 minutes, as described before.

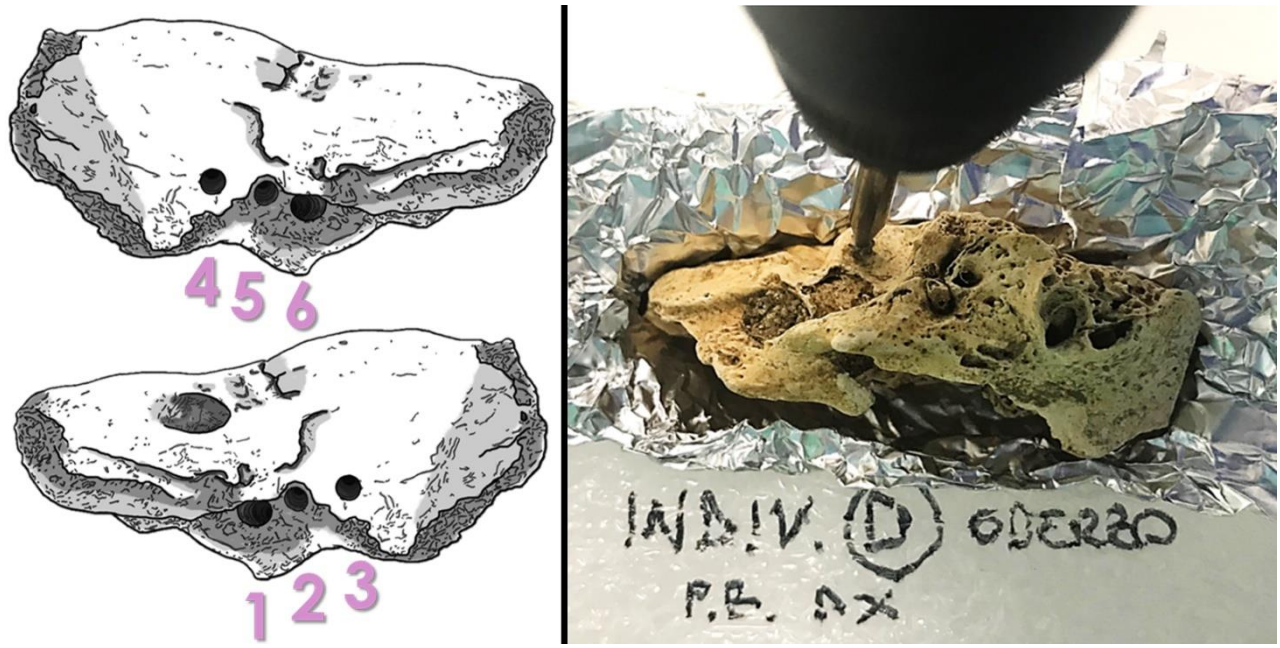

Figure 3. *The figure shows the various sampling points performed on the sample. Each sampling point was sampled twice by increasing the depth by 2 mm. It is suggested to see* ***Figure.3*** *e* ***Table.3*** *(Experiment 3) to connect sampling points and the different depths.*

3. RESULTS AND DISCUSSION

3.1 Infrared spectral analysis

In the spectral analysis of the examined sample (INDIV.A), a scarcity of organic content is detected as the spectrum does not present the characteristic bands of collagen.

Indeed, the spectrum shows a wide band within the range 3500-2500 cm^{-1}, generated by the stretching modes ν_1 and ν_3 of water molecules (H–O–H bonds) while the wide resonance doublet of the

hydrogen bond at ~ 3310-3270 cm^{-1} and ~ 3100-3030 cm^{-1}, related to amide A and amide B (N–H stretching) proteins that are insensitive to the conformation of the polypeptide backbone (Barth, 2007), is scarcely evident. Moving the focus in the fingerprint region (1800-400 cm-1), specific collagen absorption band was identified and related to C=O stretching vibrations, N–H bonding and to the coupling of the C–N stretch to the N–H bend, corresponding to Amide I (1660 cm^{-1}), while the characteristic vibrational absorption bands of Amide II (1550 cm^{-1}) and Amide III (at 1340-1240 cm^{-1}) are very weak. The shift to lower wavenumbers of the Amide I band is due to the superimposed contribution of ν_2 (O–H) at 1645 cm^{-1} bending mode of water (Socrates, 2004). This confirms the scarcity of organic material within the bone sample. The IR spectra show the ν_3 (CO_3^{2-}) asymmetric stretching peak at 1452-1425 cm^{-1} and the ν_2 (CO_3^{2-}) out of plane bend vibration, at 875 cm^{-1} (Fleet, 2009). For all spectra, the contribution of the ν_3 (PO_4^{3-}) band at 1090 cm^{-1}, due to the asymmetrical vibration of the phosphate ions, appears as a weak shoulder, indicating a less ordered bioapatite structure (Dal Sasso et al., 2018; Pleshko et al., 1992). The absorption bands at 1035, 604 and 565 cm^{-1} correspond to the asymmetric stretching and bending (ν_3, ν_4) of the phosphates, respectively, while a weak band at 960 cm^{-1} is attributed to the symmetrical stretching mode ν_1 (PO_4^{3-}). Peak at 712 cm^{-1} of ν_4 (CO_3^{2-}) in-plane bending, related to the precipitated exogenous calcite, was detected. Characteristic vibrational bands were also detected at 809 cm^{-1} and 472 cm^{-1} corresponding to asymmetric and symmetric stretching vibration of the Si–O–Si bond of the silica as reported in literature (Saikia et al., 2008).

3.2 Results of the experiments performed

After repeated grindings of the same pellet (*Experiment 1*), the phosphate peak width at 85% (FW85%) should decrease and the splitting factor (IRSF) should increase until significant value changes are noted (Asscher et al., 2011) (**Table.2**).

EXPERIMENT N.1 (same pellet)	**FW85%**	**IRSF**
IND.A (30 sec.)	10.86	3.31
IND.A (5 min.)	10.02	3.46
IND.A (10 min.)	9.61	3.59
EXPERIMENT N.1 (diff.pellet)		
IND.A (30 sec.)	10.40	3.53
IND.A (2 min.)	10.11	3.58
IND.A (5 min.)	9.35	3.73
IND.A (5 min.)	9.41	3.74
IND.A (10 min.)	10.40	3.53

EXPERIMENT N.2 (ratio /KBr mg)	**FW85%**	**IRSF**
IND.A (1:100)	10.20	3.45
IND.A (1:300)	9.08	3.63
IND.A (1.700)	10.34	3.18

EXPERIMENT N.3	FW85%	IRSF
IND.A_point 1 (2 mm-sup)	9.94	3.72
IND.A_point 1 (2 mm-deep)	9.86	3.56
IND.A_point 2 (2 mm-sup)	10.28	3.60
IND.A_point 2 (2 mm-deep)	10.11	3.35
IND.A_point 3 (2 mm-sup)	10.09	3.47
IND.A_point 3 (2 mm-deep)	10.45	3.37
IND.A_point 4 (2 mm-sup)	9.97	3.48
IND.A_point 4 (2 mm-deep)	9.95	3.47
IND.A_point 5 (15 mm)	9.93	3.70
IND.A_point 6 (15 mm)	10.03	3.60

Table 2. *The table shows the data collected from the Infrared analysis after the grinding and dilution experiments. Experiment n.3 indicates the various sampling points. Same points of petrous bone were sampled at different depths and it is suggested to look at* ***Figure.3*** *as the number of points assigned correspond to the sampling site.*

Indeed, in *Experiment 1* a steady decrease in FW85% values and an increase in IRSF were observed, already after the second grinding step. The FW85% and IRSF values after the second grinding, with a time of 5 minutes, already show a significant reduction in the peak width from 10.86 to 10.11 and a consequent increase in IRSF from 3.31 to 3.46, respectively. Increasing the grinding time (10 minutes) these variations are increased with FW85% value of 9.61 and IRSF of 3.59 (**Table.2**).

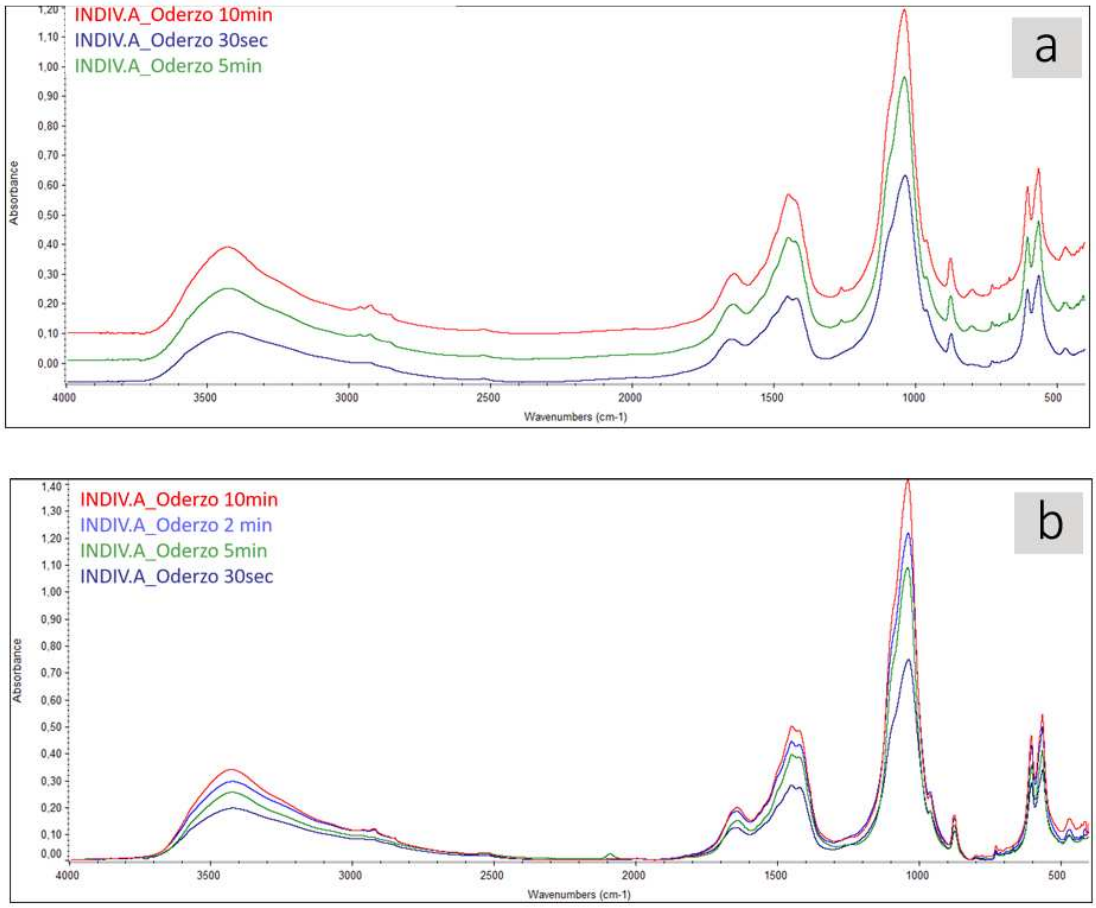

Figure 4. *The figure* ***4a*** *shows the spectra acquired after grinding the same pellet in several steps (Experiment 1in* ***Table3****) while* ***4b*** *shows the spectra acquired after the grinding of different pellets (Experiment 1in* ***Table3****) prepared with the same protocol as reported in Materials and Methods.*

The growth effect of the IRSF, due to different sample properties including particle size because of grinding(Asscher et al., 2011), is expressed in the IR spectra (**Figure.4a**) but there are no substantial differences immediately after the second grinding step. By analysing the IR spectra, acquired from pellets ground at different times, significant variations in the values of IRSF and FW85% were found. After the second grinding (2 minutes) the values of FW85% decrease from 10.40 to 10.11 while no significant variations of IRFS were recorded with values ranging from 3.53 to 3.58. Any relevant difference was found between the samples ground at 5 and 10 minutes. Significant differences were not found between the IR spectra shown in **Figure 4b** and grinding at 2-5 and 10 minutes.

In *Experiment 2* the effect of dilution was investigated. Previous studies (Lebon et al., 2010) have shown that dilution heavily affects the parameters measurement from FTIR spectra, both in terms of absolute value and trend. Data, collected from IR measurements, show a fluctuation in the values of the splitting factor and the phosphate peak width, as reported in **Table.2**.
The dilution seems to considerably affect the IRSF determination: this may be caused by the increase of signal-to-noise ratio (especially for much diluted samples) and probably by the decrease of the instrument sensitivity due to dilution effects. For these reasons, to get a good quality of spectra, it was considered to maintain the ratio of sample/KBr to 1:100.

In *Experiment 3* powders extracted from INDIV.D were used. Six different sampling points (external and internal side of petrous bone, **Figure.2**) at depth of 2 mm, 4 mm and 15 mm per each sampling point were prepared as reported by the protocol used in paragraph *2.3*.
This experiment was necessary to verify whether there is some detectable differences between samples extracted at different depths and sampling points of the same specimen.
Data collected for FW85% and IRSF parameters from ten spectra (**Table.2**) show that these parameters are not correlated by depth and sampling position. Considering the FW85% and IRSF values obtained from IR data collected on several sampling position, the variability of sample preparation is higher than that derived by the choice of the depth and sampling position: differences are due to statistical fluctuations caused by noise-limited sensitivity of the tool (**Figure.5**).

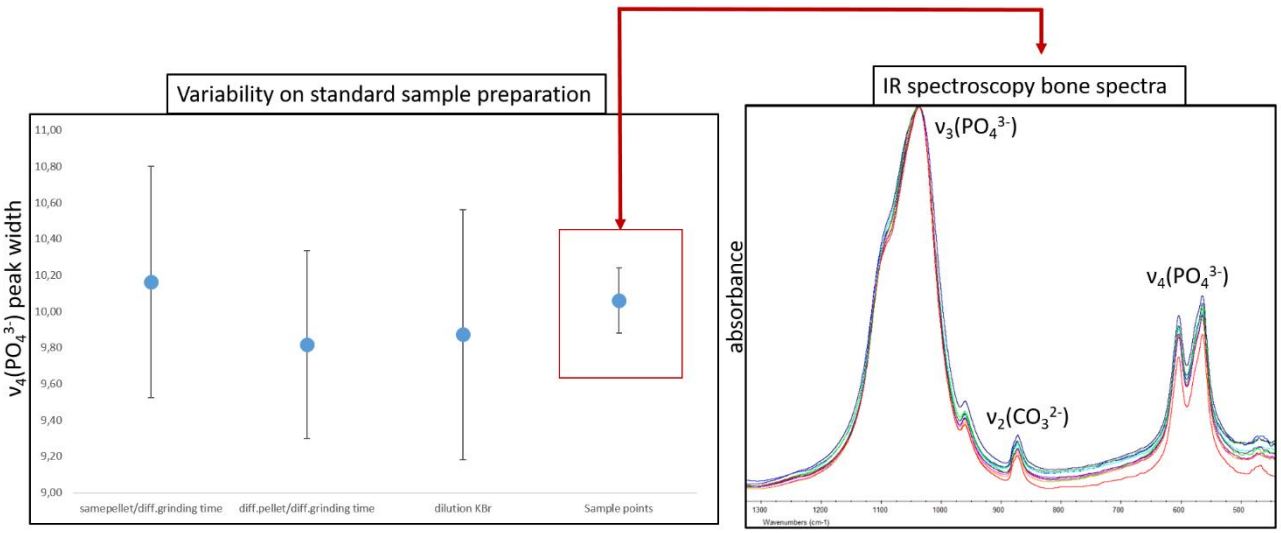

Figure 5.*The preparation has a strong impact on the analysis of the spectra unlike the sampling point. This result indicates that diagenesis acts uniformly despite varying degrees of bone density.*

This result highlights that in the case of a relatively homogeneous and well preserved petrous bone diagenesis uniformly acts on the same sample and there are no significant differences on the different sub-samples of the bone tissue analysed.

By means of analysis software (CTAn v.1.18 Micro-CT software), it was possible to determine the standard morphometric indices of bone: bone surface, trabecular thickness and trabecular separation. From the images reconstructed in micro-CT, a greater accumulation of residues of the soil has been detected in the apex of the petrous bone (part A show **Figure.2**), much more cancellous and mainly consisting of bone *trabeculae* than the *cochlea*.

The presence of exogenous material can represent a problem for the determination of the sample preservation state as signals may overlap, possibly invalidating the analysis.

Area C, identified through images acquired with micro-CT (**Figure.2**), allows extracting sufficient material for the spectroscopic analysis with a minimum invasiveness. Furthermore, being a denser and more resistant area than area A, it allows sampling of the specimen avoiding the risk of shattering the apex of the petrous bone. Although area B is located near the *subarcuate fossa* of the petrous bone (consisting of compact bone tissue), the incomplete apical region and the middle region of the *cochlea* reside inside (**Figure.2**), from which a little bone material is extracted. The densest area with the greatest chance of obtaining enough powders at a depth of no more than 2 mm was found in the upper proximity of the internal opening of the carotid canal near the *otic capsule*.

4. CONCLUSIONS

Based on these preliminary results, significant improvement in the quality of the spectra, allowing a better direct evaluation of the degree of *post-mortem* bioapatite recrystallization, was obtained after grinding for two minutes, while no particular changes were recorded with longer grinding times. All

experimental sampling parameters were carefully calibrated in terms of reproducibility and minimal invasiveness in order to obtain reliable and safe procedures.

The sampling protocol, developed for archaeological bone finds, consists in the mechanical removal of the material (~ 1 mg) near to the internal opening of the carotid canal (at a depth of 2 mm) by means of a low-speed micro-drill, equipped with an abrasive round point, to remove sediments on the surface. Subsequently, the required amount of bone powder was extracted using a helical micro-tool at moderate speed. Each micro-tool was disinfected according to the ICMP standard procedures. Furthermore, the application of a standard bone powder preparation for FTIR analysis has significantly improved the accuracy, repeatability and reproducibility, as well as the reliability of the data, allowing a systematic evaluation of bones samples suitable for archaeological and anthropological-molecular studies. The standard protocol, delineated in this research, has been used to elaborate a pre-screening method for molecular analysis and, for this purpose, it has been applied to all bone samples selected and reported in **Chapters 3**, **4** and **5**.

AUTHOR CONTRIBUTION

Cinzia Scaggion: conceptualization, data curation, formal analysis, investigation, writing - original draft; **Luca Nodari:** conceptualization, validation, supervision, writing - review & editing; **Luca Pagani:** supervision, writing - review & editing; **Nicola Carrara:** resources, writing - review & editing; **Gilberto Artioli:** project administration, conceptualization, supervision, writing - review & editing.

ACKNOWLEDGEMENTS

We thank Dr. Christophe Snoeck Ph.D., researcher at Vrije Universiteit Brussel and the head of the Brussels Bioarchaeology Lab (BB-LAB) and Dr. Ina Reiche Ph.D., research director of French National Centre for Scientific Research-CNRS for their valuable scientific advices.

REFERENCES

Allende, M.K.-G., Samplonius, A., 2022. Dental anthropological report: Exploring plant-based treatments through the analysis of dental calculus and sediment of dental caries in a woman from the Late Preceramic period, Peru. Ann. Anat. - Anat. Anzeiger 240, 151849. https://doi.org/10.1016/j.aanat.2021.151849

Asscher, Y., Dal Sasso, G., Nodari, L., Angelini, I., Boffa Ballaran, T., Artioli, G., 2017. Differentiating between long and short range disorder in infra-red spectra: On the meaning of “crystallinity” in silica. Phys. Chem. Chem. Phys. 19, 21783–21790. https://doi.org/10.1039/c7cp03446f

Asscher, Y., Regev, L., Weiner, S., Boaretto, E., 2011. Atomic Disorder in Fossil Tooth and Bone

Mineral: An FTIR Study Using the Grinding Curve Method. ArchéoSciences 135–141. https://doi.org/10.4000/archeosciences.3062

Balzeau, A., 2015. Comparative aspects of temporal bone pneumatization in some African fossil hominins. BMSAP 27, 135–141. https://doi.org/10.1007/s13219-015-0126-5

Barth, A., 2007. Infrared spectroscopy of proteins. Biochim. Biophys. Acta - Bioenerg. 1767, 1073–1101. https://doi.org/10.1016/j.bbabio.2007.06.004

Berna, F., Matthews, A., Weiner, S., 2004. Solubilities of bone mineral from archaeological sites: The recrystallization window. J. Archaeol. Sci. 31, 867–882. https://doi.org/10.1016/j.jas.2003.12.003

Brace, C.L., Smith, S.L., Hunt, K.D., 1991. What big teeth you had, Grandma! Human tooth size, past and present, in: M.K. Kelley and C.S. Larsen (Ed.), Advances in Dental Anthropology. Wiley-Liss, New York, pp. 33–57.

Brown, T., Brown, K., 2011. Biomolecular Archaeology. Wiley-Blackwell, Oxford, UK. https://doi.org/10.1002/9781444392449

Cacciotti, I., 2016. Cationic and Anionic Substitutions in Hydroxyapatite, in: Handbook of Bioceramics and Biocomposites. Springer International Publishing, Cham, pp. 145–211. https://doi.org/10.1007/978-3-319-12460-5_7

Cipriano, S., Sandrini, G.., 2001. La banchina fluviale di Opitergium, in: EUT Edizioni Università di Trieste (Ed.), Antichità Altoadriatiche XLVI (2001), Strutture Portuali e Rotte Marittime Nell'Adriatico Di Età Romana. Trieste, pp. 454–455.

Cipriano, S., Sandrini, G.M., 2015. ...Redditur enim terrae corpus (Cic., De.Leg.,2,56) Necropoli di Oderzo tra media e tarda età Imperiale., in: Federica Rinaldi e Alberto Vigoni (Ed.), Le Necropoli Della Media e Tardaetà Imperiale (Iii-Iv Secolo d.c.) a Iulia Concordia e Nell'arco Altoadriatico.Organizzaizone Spaziale, Aspetti Monumentali e Sociali. Grafiche Turato Edizioni, Rubano (PD), pp. 225–244.

Collins, M.J., Nielsen-Marsh, C.M., Hiller, J., Smith, C.I., Roberts, J.P., Prigodich, R. V., Wess, T.J., Csapò, J., Millard, A.R., Turner-Walker, G., 2002. The survival of organic matter in bone: A review. Archaeometry 44, 383–394. https://doi.org/10.1111/1475-4754.t01-1-00071

Dal Sasso, G., Asscher, Y., Angelini, I., Nodari, L., Artioli, G., 2018. A universal curve of apatite crystallinity for the assessment of bone integrity and preservation. Sci. Rep. 8, 1–13. https://doi.org/10.1038/s41598-018-30642-z

Dal Sasso, G., Maritan, L., Usai, D., Angelini, I., Artioli, G., 2014. Bone diagenesis at the micro-scale: Bone alteration patterns during multiple burial phases at Al Khiday (Khartoum, Sudan) between the Early Holocene and the II century AD. Palaeogeogr. Palaeoclimatol. Palaeoecol.

416, 30–42. https://doi.org/10.1016/j.palaeo.2014.06.034

Demes, B., Creel, N., 1988. Bite force, diet, and cranial morphology of fossil hominids. J. Hum. Evol. 17, 657–670. https://doi.org/10.1016/0047-2484(88)90023-1

Doden, E., Halves, R., 1984. On the functional morphology of the human petrous bone. Am. J. Anat. 169, 451–462. https://doi.org/10.1002/aja.1001690407

Drusini, A., Rippa Bonati, M., Furlan, C., 1987. Oderzo: necropoli in via Spinè in: Riviste e Annali Archeologia. Quad. di Archeol. del Veneto 81.

Dupras, T.L., Schwarcz, H.P., 2001. Strangers in a Strange Land: Stable Isotope Evidence for Human Migration in the Dakhleh Oasis, Egypt. J. Archaeol. Sci. 28, 1199–1208. https://doi.org/10.1006/jasc.2001.0640

Elliott, J.C., Dover, S.D., 1982. X-ray microtomography. J. Microsc. 126, 211–213. https://doi.org/10.1111/j.1365-2818.1982.tb00376.x

Fleet, M.E., 2009. Infrared spectra of carbonate apatites: ν2-Region bands. Biomaterials 30, 1473–1481. https://doi.org/10.1016/j.biomaterials.2008.12.007

Foley, B., Greiner, M., McGlynn, G., Schmahl, W.W., 2020. Anatomical Variation of Human Bone Bioapatite Crystallography. Crystals 10, 859. https://doi.org/10.3390/cryst10100859

Greiner, M., Rodríguez-Navarro, A., Heinig, M.F., Mayer, K., Kocsis, B., Göhring, A., Toncala, A., Grupe, G., Schmahl, W.W., 2019. Bone incineration: An experimental study on mineral structure, colour and crystalline state. J. Archaeol. Sci. Reports 25, 507–518. https://doi.org/10.1016/j.jasrep.2019.05.009

Hedges, R.E.M., 2002. Bone diagenesis: an overview of processes. Archaeometry 44, 319–328. https://doi.org/10.1111/1475-4754.00064

Hedges, R.E.M., Millard, A.R., 1995. Bones and Groundwater: Towards the Modelling of Diagenetic Processes. J. Archaeol. Sci. 22, 155–164. https://doi.org/10.1006/jasc.1995.0017

ICMP, 2015. Standard Operating Procedure for Sampling Bone and Tooth Specimens From Human Remains 1–19.

Jans, M.M.E., 2008. Microbial bioerosion of bone – a review, in: Current Developments in Bioerosion. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 397–413. https://doi.org/10.1007/978-3-540-77598-0_20

Jans, M.M.E., Nielsen-Marsh, C.M., Smith, C.I., Collins, M.J., Kars, H., 2004. Characterisation of microbial attack on archaeological bone. J. Archaeol. Sci. 31, 87–95. https://doi.org/10.1016/j.jas.2003.07.007

Karakostis, F.A., Haeufle, D., Anastopoulou, I., Moraitis, K., Hotz, G., Tourloukis, V., Harvati, K., 2021. Biomechanics of the human thumb and the evolution of dexterity. Curr. Biol. 31, 1317-

1325.e8. https://doi.org/10.1016/j.cub.2020.12.041

Keenan, S.W., 2016. From bone to fossil: A review of the diagenesis of bioapatite. Am. Mineral. 101, 1943–1951. https://doi.org/10.2138/am-2016-5737

Kendall, C., Eriksen, A.M.H., Kontopoulos, I., Collins, M.J., Turner-Walker, G., 2018. Diagenesis of archaeological bone and tooth. Palaeogeogr. Palaeoclimatol. Palaeoecol. 491, 21–37. https://doi.org/10.1016/j.palaeo.2017.11.041

Kieser, J.A., Bernal, V., Neil Waddell, J., Raju, S., 2007. The Uniqueness of the Human Anterior Dentition: A Geometric Morphometric Analysis. J. Forensic Sci. 52, 671–677. https://doi.org/10.1111/j.1556-4029.2007.00403.x

Kimmerle, E.H., Ross, A., Slice, D., 2008. Sexual Dimorphism in America: Geometric Morphometric Analysis of the Craniofacial Region. J. Forensic Sci. 53, 54–57. https://doi.org/10.1111/j.1556-4029.2007.00627.x

Kontopoulos, I., Penkman, K., McAllister, G.D., Lynnerup, N., Damgaard, P.B., Hansen, H.B., Allentoft, M.E., Collins, M.J., 2019. Petrous bone diagenesis: a multi-analytical approach. Palaeogeogr. Palaeoclimatol. Palaeoecol. 518, 143–154. https://doi.org/10.1016/j.palaeo.2019.01.005

Kontopoulos, I., Penkman, K., Mullin, V.E., Winkelbach, L., Unterländer, M., Scheu, A., Kreutzer, S., Hansen, H.B., Margaryan, A., Teasdale, M.D., Gehlen, B., Street, M., Lynnerup, N., Liritzis, I., Sampson, A., Papageorgopoulou, C., Allentoft, M.E., Burger, J., Bradley, D.G., Collins, M.J., 2020. Screening archaeological bone for palaeogenetic and palaeoproteomic studies. PLoS One 15, 1–17. https://doi.org/10.1371/journal.pone.0235146

Le Meillour, L., Zazzo, A., Lesur, J., Cersoy, S., Marie, A., Lebon, M., Pleurdeau, D., Zirah, S., 2018. Identification of degraded bone and tooth splinters from arid environments using palaeoproteomics. Palaeogeogr. Palaeoclimatol. Palaeoecol. 511, 472–482. https://doi.org/10.1016/j.palaeo.2018.09.013

Lebon, M., Reiche, I., Bahain, J.J., Chadefaux, C., Moigne, A.M., Fröhlich, F., Sémah, F., Schwarcz, H.P., Falguères, C., 2010. New parameters for the characterization of diagenetic alterations and heat-induced changes of fossil bone mineral using Fourier transform infrared spectrometry. J. Archaeol. Sci. 37, 2265–2276. https://doi.org/10.1016/j.jas.2010.03.024

Lebon, M., Reiche, I., Gallet, X., Bellot-Gurlet, L., Zazzo, A., 2016. Rapid quantification of bone collagen content by ATR-FTIR spectroscopy. Radiocarbon 58, 131–145. https://doi.org/10.1017/RDC.2015.11

Lee-Thorp, J., Sealy, J., 2008. Beyond documenting diagenesis: The fifth international bone diagenesis workshop. Palaeogeogr. Palaeoclimatol. Palaeoecol. 266, 129–133.

https://doi.org/10.1016/j.palaeo.2008.03.025

LeGeros, R.Z., 1981. Apatites in Biological Systems. Prog. Cryst. Growth Charact. 4, 1–45.

Leskovar, T., Zupanič Pajnič, I., Geršak, Ž.M., Jerman, I., Črešnar, M., 2020. ATR-FTIR spectroscopy combined with data manipulation as a pre-screening method to assess DNA preservation in skeletal remains. Forensic Sci. Int. Genet. 44, 102196. https://doi.org/10.1016/j.fsigen.2019.102196

Licata, M., Bonsignore, A., Boano, R., Monza, F., Fulcheri, E., Ciliberti, R., 2020. Study, conservation and exhibition of human remains: The need of a bioethical perspective. Acta Biomed. 91, 1–10. https://doi.org/10.23750/abm.v91i4.9674

Mays, S., Taylor, G.M., Legge, A.J., Young, D.B., Turner-Walker, G., 2001. Paleopathological and biomolecular study of tuberculosis in a medieval skeletal collection from England. Am. J. Phys. Anthropol. 114, 298–311. https://doi.org/10.1002/ajpa.1042

Minozzi, S., Catalano, P., Caldarini, C., Fornaciari, G., 2012. Palaeopathology of Human Remains from the Roman Imperial Age. Pathobiology 79, 268–283. https://doi.org/10.1159/000338097

Nielsen-Marsh, C.M., Hedges, R.E.M., 2000. Patterns of diagenesis in bone I: The effects of site environments. J. Archaeol. Sci. 27, 1139–1150. https://doi.org/10.1006/jasc.1999.0537

Oxilia, G., Bortolini, E., Martini, S., Papini, A., Boggioni, M., Buti, L., Figus, C., Sorrentino, R., Townsend, G., Kaidonis, J., Fiorenza, L., Cristiani, E., Kullmer, O., Moggi-Cecchi, J., Benazzi, S., 2018. The physiological linkage between molar inclination and dental macrowear pattern. Am. J. Phys. Anthropol. 166, 941–951. https://doi.org/10.1002/ajpa.23476

Pan, H., Darvell, B.W., 2010. Effect of Carbonate on Hydroxyapatite Solubility. Cryst. Growth Des. 10, 845–850. https://doi.org/10.1021/cg901199h

Pasteris, J.D., Wopenka, B., Freeman, J.J., Rogers, K., Valsami-Jones, E., van der Houwen, J.A.M., Silva, M.J., 2004. Lack of OH in nanocrystalline apatite as a function of degree of atomic order: implications for bone and biomaterials. Biomaterials 25, 229–238. https://doi.org/10.1016/S0142-9612(03)00487-3

Pleshko, N.L., Boskey, A.L., Mendelsohn, R., 1992. An FT-IR microscopic investigation of the effects of tissue preservation on bone. Calcif. Tissue Int. 51, 72–77. https://doi.org/10.1007/BF00296221

Poralan, G.M., Gambe, J.E., Alcantara, E.M., Vequizo, R.M., 2015. X-ray diffraction and infrared spectroscopy analyses on the crystallinity of engineered biological hydroxyapatite for medical application. IOP Conf. Ser. Mater. Sci. Eng. 79. https://doi.org/10.1088/1757-899X/79/1/012028

Prowse, T., Schwarcz, H.P., Saunders, S., Macchiarelli, R., Bondioli, L., 2004. Isotopic paleodiet

studies of skeletons from the Imperial Roman-age cemetery of Isola Sacra, Rome, Italy. J. Archaeol. Sci. 31, 259–272. https://doi.org/10.1016/j.jas.2003.08.008

Raichlen, D.A., Pontzer, H., Zderic, T.W., Harris, J.A., Mabulla, A.Z.P., Hamilton, M.T., Wood, B.M., 2020. Sitting, squatting, and the evolutionary biology of human inactivity. Proc. Natl. Acad. Sci. 117, 7115–7121. https://doi.org/10.1073/pnas.1911868117

Rathmann, H., Reyes-Centeno, H., 2020. Testing the utility of dental morphological trait combinations for inferring human neutral genetic variation. Proc. Natl. Acad. Sci. 117, 10769–10777. https://doi.org/10.1073/pnas.1914330117

Reiche, I., Favre-Quattropani, L., Vignaud, C., Bocherens, H., Charlet, L., Menu, M., 2003. A multi-analytical study of bone diagenesis: the Neolithic site of Bercy (Paris, France). Meas. Sci. Technol. 14, 1608–1619. https://doi.org/10.1088/0957-0233/14/9/312

Rubinson, K.A., Rubinson, J.F., 2002. Chimica analitica strumentale. Zanichelli.

Saikia, B.J., Parthasarathy, G., Sarmah, N.C., 2008. Fourier transform infrared spectroscopic estimation of crystallinity in SiO2 based rocks. Bull. Mater. Sci. 31, 775–779. https://doi.org/10.1007/s12034-008-0123-0

Smith, C.I., Nielsen-Marsh, C.M., Jans, M.M.E., Collins, M.J., 2007. Bone diagenesis in the European Holocene I: patterns and mechanisms. J. Archaeol. Sci. 34, 1485–1493. https://doi.org/10.1016/j.jas.2006.11.006

Socrates, G., 2004. Infrared and Raman Characteristic Group Frequencies: Tables and Charts, 3° edition. ed.

Stathopoulou, E.T., Psycharis, V., Chryssikos, G.D., Gionis, V., Theodorou, G., 2008. Bone diagenesis: New data from infrared spectroscopy and X-ray diffraction. Palaeogeogr. Palaeoclimatol. Palaeoecol. 266, 168–174. https://doi.org/10.1016/j.palaeo.2008.03.022

Traub, W., Arad, T., Weiner, S., 1992. Origin of Mineral Crystal Growth in Collagen Fibrils. Matrix 12, 251–255. https://doi.org/10.1016/S0934-8832(11)80076-4

Trueman, C.N., 2013. CHEMICAL TAPHONOMY OF BIOMINERALIZED TISSUES 56, 475–486. https://doi.org/10.1111/pala.12041

Trueman, C.N., Tuross, N., 2019. Trace elements in recent and fossil bone apatite. Phosphates Geochemical, Geobiol. Mater. Importance 48, 489–522. https://doi.org/10.2138/rmg.2002.48.13

Trueman, C.N.G., Behrensmeyer, A.K., Tuross, N., Weiner, S., 2004. Mineralogical and compositional changes in bones exposed on soil surfaces in Amboseli National Park, Kenya: Diagenetic mechanisms and the role of sediment pore fluids. J. Archaeol. Sci. 31, 721–739. https://doi.org/10.1016/j.jas.2003.11.003

Wagermaier, W., Fratzl, P., 2012. Collagen, in: Matyjaszewski, K., Möller, M. (Eds.), Polymer Science: A Comprehensive Reference. Elsevier Science, pp. 35–55.

Weiner, S., 2010. Microarchaeology. Beyond the visible archaeological record, first edit. ed. Cambridge University Press, UK.

Weiner, S., Bar-Yosef, O., 1990. States of preservation of bones from prehistoric sites in the Near East: A survey. J. Archaeol. Sci. 17, 187–196. https://doi.org/10.1016/0305-4403(90)90058-D

Weiner, S., Traub, W., 1992. Bone structure: from angstroms to microns. FASEB J. 6, 879–85. https://doi.org/10.1096/fasebj.6.3.1740237

Wopenka, B., Pasteris, J.D., 2005. A mineralogical perspective on the apatite in bone. Mater. Sci. Eng. C 25, 131–143. https://doi.org/10.1016/j.msec.2005.01.008

CHAPTER 3

Diagenetic alterations model of heterogeneous archaeo-osteological tissues

Cinzia Scaggion[1*], Gregorio Dal Sasso[2], Luca Nodari[3], Luca Pagani[4], Nicola Carrara[5], Donatella Usai[6], Giulia Gadioli[5], Gilberto Artioli[1]

[1] Department of Geosciences, University of Padova, Padova, 35131, Italy
[2] Institute of Geosciences and Earth resources, Italian National Research Council-CNR, Padova, 35131, Italy
[3] Italian National Research Council-CNR, ICMATE, Padova, 35127, Italy
[4] Department of Biology, University of Padova, Padova, 35122, Italy
[5] Museum of Anthropology, University of Padova, Padova, 35121, Italy
[6] Sudanese and Sub-Saharan Studies Center (CSSeS), Treviso, Italy

*corresponding author, e-mail address: cinzia.scaggion@phd.unipd.it (C. Scaggion)

Keywords: diagenesis; FTIR; bone quality; collagen preservation; atomic disorder; infrared peak width

Abstract

Here we intend to propose a broad model of diagenetic alteration, comparing the behavior of the organic and inorganic phase on a large and diversified set of bone samples characterized by different origins and chronological phases. From extremely altered archaeological specimens to modern, fresh bones, we evaluated the sensitivity and efficacy of the Infrared phosphate peak width including it as the most predictive parameter on the degree of alteration of the bone system. The work was extended to examine the modifications induced by diagenesis on the secondary structure of conserved collagen, evaluating their effects on bioapatite crystals. The results obtained prove that the width of the phosphate peak is advantageous for monitoring minimal changes in the structure and chemical properties of bioapatite as well as indirectly in collagen. This method could improve the selection process of bone samples for different fields of interest. Therefore, we will intend to proceed with this research idea by developing a minimally invasive and low-cost pre-screening method for molecular and isotope analysis.

1. INTRODUCTION

Hard tissues such as bones and teeth are often the only direct proof remains of animals and humans of the past and hence represent valuable archives for paleogenomic, palaeoecology, palaeo-environment and palaeopathology studies. For these reasons, the osteological remains of historical, archaeological, medical-forensic and paleoanthropological interest are of fundamental importance due to their eminent research function. *Post-mortem* alterations of archaeological bone depend upon a number of various processes, collectively included in the concept of bone diagenesis, that can lead to degrade it rapidly due to extrinsic/intrinsic factors (i.e. burial environment, bone biochemistry and microstructure), which may make the specimen less suitable for the planned studies aiming at recovering pristine information on bone material. The bone tissue consists of an inorganic component, formed by carbonate-hydroxylapatite nanocrystals known as bioapatite, and organic components as collagen, which accounts for approximately 90% of the organic fraction of bone, protein-mucopolysaccharide complexes, and glycoproteins (Kendall et al., 2018).

In a live organism the chemical composition, solubility, mechanical function, crystal size and structure of biogenic apatite are regulated by metabolic processes (Clarke, 2008; Rey et al., 2009). After the death of individual, bioapatite nanocrystals are thermodynamically unstable and prone to recrystallization, so that the removal of the collagen leaves bioapatite crystals unprotected giving rise to a more brittle biomaterial (Wescott, 2019) undergoing the dissolution/reprecipitation of apatite nanocrystals. The bioapatite recrystallization feasibly occurs by *Ostwald ripening* mechanisms, for which small crystallites dissolve favoring the growth of larger crystal (Berna et al., 2004). This mechanism, as well as incorporation of exogenous ions incorporations in the crystal lattice from the burial environment (e.g., PO_4^{3-}, CO_3^{2-}, Ca^{2+}, Mg^{2+}), leads to a more stable and ordered crystal structure characterized by larger dimensions of nanocrystals with decreased specific surface area (Pollard et al., 2007; Trueman, 2013; Trueman et al., 2004). Bioapatite can undergo chemical/physical transformations or mechanical alterations as a result of the combined action of the soil chemistry, microorganisms or plants (Hedges, 2002; Trueman and Martill, 2002; Weiner and Addadi, 2011).

Collagen is one of the most important components for bioarchaeological studies (radiocarbon dating, stable isotopes analysis etc.) and accounts for 10-25% of the total bone weight, most of which is Type I collagen (France et al., 2014). The loss of collagen is caused by the change from insoluble collagen to soluble phase (gelatin) with an impact upon the physical properties of the bone. Collagen loss occurs due to both biological and chemical degradation for hydrolysis (Collins et al., 1995; Keenan, 2021). The transition of the collagen structure into gelatin is accelerated by high temperature, extreme pH levels and presence of humic acids while it is delayed by the preservation of bioapatite mineral and by close packing of the collagen fibrils (Collins et al., 2002). Being diagenesis site-specific and strongly influenced by local environmental and burial conditions, and considering the high number of variables affecting the alteration processes, a predictive procedure to evaluate the preservation state of archaeological bones and the survival of the organic fraction is not straightforward.

Infrared spectroscopy has been widely applied to bone characterization studies as an important tool for diagenesis research, as it offers many advantages in terms of low-cost application, low amount of required sample (~1 mg) and sensitivity to variations on the structural properties of both organic and inorganic components (Carden and Morris, 2000).

Therefore, this tool allows the rapid, minimally destructive and economic analysis of a large number of fossil bones. The crystallinity as well as the physico-chemical and microstructural properties of bioapatite are extremely meaningful indicators of archeological bone preservation (Weiner, 2010). Indeed, the characterization of archaeological bones using infrared spectroscopy provided valuable quantitative and qualitative information on bone preservation (Piga et al., 2016; Querido et al., 2018; Reiche et al., 2002; Shemesh, 1990; Smith et al., 2007; Thompson et al., 2011; Weiner and Bar-

Yosef, 1990). Here we propose an analytical approach previously developed by Dal Sasso et al. (2018), which uses FTIR spectroscopy as a potential diagnostic tool to indirectly describe or quantify changes in bioapatite crystallinity by monitoring its variations between different human bone samples.

In the previous work a detailed study on the relationship between ν_4 phosphate peak width at FW85%, IRSF and crystallinity (determined by X-ray powder diffraction - XRD) has been carried out by investigating fresh and fossil bones. It emerged that infrared peak width at 604 cm^{-1} can be a novel way useful tool to reliably assess minimal changes in bioapatite chemical composition and structure, acquiring significant indirect physical information on biomaterials constituting bone system. Thus, the broadening of the peak could be the most promising parameter on the degree of bone alteration, as it is closely related to the atomic order/disorder in the crystal lattice.

The main aim of this study is to provide a reliable method to retrieve indirect significant physical-chemical information on inorganic and organic bone constituents. A set of parameters measured by FTIR are evaluated and discussed to determine the conservation status of an archaeological bone. A comprehensive model of diagenetic alteration will be proposed by comparing the behavior of the organic and inorganic phase on a large and diverse set of samples. The general model here proposed enables a comprehensive description of diagenetic alterations of a wide set of human bones with different origins, chronology and diagenetic history.

2. MATERIALS AND METHODS

2.1. Bone samples

The study samples included human bones from different origin, chronology and state of conservation from the Anthropology Museum of Padua University, in particular petrous bones (n.57) and tooth roots (n.59). Analyzed bones were available from 2^{nd}- 3^{rd} c.AD graves of Oderzo (IT), 7^{th}-8^{th} c.AD graves of Montecchio Maggiore (IT), 4^{th}-6^{th} c.AD graves of Biverone, Bronze Age burials of Grottina dei Covoloni del Broion (IT), the high medieval 11^{th}-12^{th} c.AD cemetery of Sant'Anna Sopramonte (IT), pre-Mesolithic, Neolithic and Meroitic graves of Al-Khiday (Central Sudan, for an overview: Usai et al., 2014; Usai and Salvatori, 2019) and other samples from the three Southern Italian Iron Age archaeological contexts of Salapia, Ordona and San Giovanni Rotondo. The sample selection comprises the high medieval Holy Relic of St. Theobald in Badia Polesine (Rovigo, Italy) that has never been buried. Ancient samples are compared with the human modern bones of the Tedeschi Collection, a precious skeletal collection dating to the late 19^{th} and early 20^{th} century (Carrara et al., 2018), fresh human tooth (Asscher et al., 2012; Dal Sasso et al., 2014; Iacumin et al., 2016), fresh animal bones and specimens of fossil femurs found in the Al-Khiday (Sudan) site used in a previous

work Dal Sasso et al. (2018), the full sample list is reported in **Table.1** (reported at the end of the Chapter).

2.2. Micro-sampling

Inspired by the work of Sirak et al. (2017) and Pinhasi et al. (2015), we applied an alternative and minimally destructive method for the extraction of powders or micro-cores from the petrous bone for FTIR analysis, since in some cases it was not possible to extract the bone materials because of complete crania. Moreover, safeguarding complete osteological remains from invasive analysis and sampling is fundamental to preserve as much physical-anthropological information as possible.

In complete crania, cochlea is not easily accessible. Based on Sirak et al. (2017), a practical method to sample the osseous labyrinth consists in entering with a Dremel micro-tool at the site of the bony ridge, separating the jugular foramen from the carotid canal on the basal surface of the cranium. Once reduced the bony ridge, the external bone surface – encrusted by soil sediments – was mechanically removed by means of low-speed micro-drill, equipped with an abrasive round point of 2.4 mm. To avoid contamination issues, overturning the cranium to collect the powder is not practicable, as sediments contained within the endocranium would also be collected. Therefore, the inner part of osseous labyrinth was extracted by a micro-core with an internal diameter of 1.8 mm. The sample fragment was thus extracted without being contaminated and particular attention was paid to use low speed in order to avoid the overheating of the samples during mechanical cleaning and sampling.

For incomplete crania (where the temporal bone is easily accessible), a similar procedure was adopted: once removed the soil sediment encrusted on the surface with a round abrasive point of 2.4 mm, the necessary amount of bone powder was obtained by using another helical micro-tool (2.4 mm) at moderate speed.

The operating procedure for standard sampling of bone of the ICMP (International Commission on Missing Persons - ICMP 2015) was adopted, disinfecting and sterilizing all drill bits to avoid external contamination. The powdered material from the roots of the molars was extracted with a round engraver bit (2.4 mm diameter) after removal of the external deposits.

The cementum/dentine powder from the roots of the molars was extracted with a round engraver bit (2.4 mm diameter) after removal of the external deposits following the same cleaning procedures of tools from ICMP.

2.3. Fourier Transform-IR spectroscopy: sample preparation and measurements

FTIR spectroscopy provides a sensitive probe for specific functional groups in organic and inorganic compounds; it is widely used not only to obtain qualitative information (Rubinson et al., 2002) but also semi-quantitative ones (Weiner, 2010) in order to investigate the nature and structure of mineral

and organic matter in the analyzed samples. Therefore, FTIR spectroscopy can provide valuable information on the preservation of the structural and the chemical properties of archaeological bones (Kontopoulos et al., 2019; Poralan et al., 2015; Weiner and Bar-Yosef, 1990).

We emplaced a protocol for sample preparation to guarantee the accuracy and replicability of the FTIR spectra. Samples have been prepared maintaining a ratio 1:100 mg of sample/KBr. Thus, 1 mg of sample was hand ground in an agate mortar and KBr was added and homogenized. The powdered sample was then pressed, using a hydraulic press, under 11 tons/cm^2 pressure, and a 12 mm in diameter and 1.5 mm of thickness transparent pellet was obtained. Spectra were collected with a Nicolet 380 FTIR spectrometer equipped with a DTGS detector; 128 scans for each spectrum were acquired, in the range from 4000 to 400 cm^{-1}, with a spectral resolution of 4 cm^{-1}. Spectral analysis was performed using Omnic 9 software (Thermo Scientific). To semi-quantifying of carbonate and amide presents in the spectrum, the ratios have been calculated dividing the intensity of absorption by the intensity of ν_3 phosphate peak at 1035 cm^{-1} (**Figure.1b**).

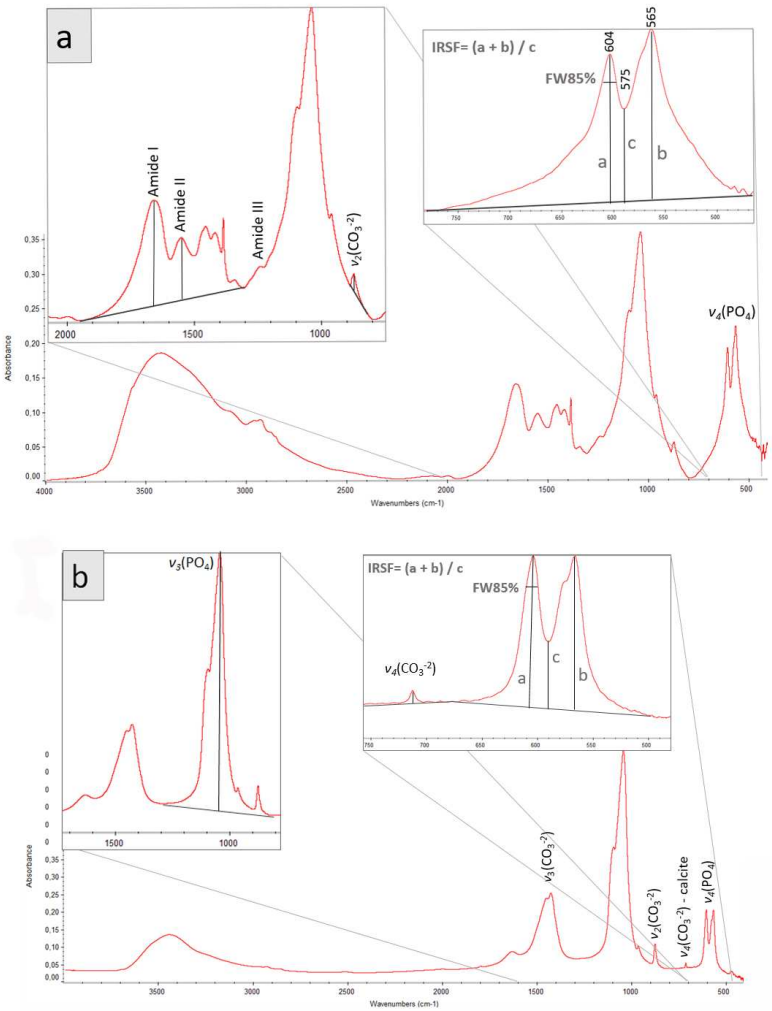

Figure 1. *FTIR spectra of archaeological human bones a comparison: an archaeological sample well preserved (**a**) and another one diagenetically altered (**b**). Bands are attributed to Amide I-II, phosphate and carbonate vibrational modes. Peak intensities and baseline used to calculate AmI/PO_4^{3-}, AmII/PO_4^{3-}, CO_3^{2-}/PO_4^{3-}, CaI/PO_4^{3-}, IRSF (splitting factor) and the full width at half maximum (FWHM) of the phosphate peak at 604 cm^{-1}(FW85%) parameters are shown.*

Peaks height is calculated from the baselines, which are defined by two points calculated as the local minimum ranging in selected regions of the spectrum, reported in **Table.2**.

Table.2. *Wavenumbers and baselines adopted to calculate peaks intensity and width of tooth and bone sample properties by FTIR spectroscopy (adapted from references Dal Sasso et al., 2018; Paschalis et al., 2011).*

Vibrational mode and functional group	Wavenumber (cm^{-1})	Baseline (cm^{-1})
AmideI ν (C=O) stretch	1660	2000–1800/1400–1200
AmideII ν (C–N) stretch and σ (N–H) bend	1550	2000–1800/1400–1200
$\nu_3(PO_4^{3-})$ antisymmetric stretching	1035	1400–1200/900–750
$\nu_2(CO_3^{2-})$ out of plane bend vibration	875	891–816
$\nu_4(CO_3^{2-})$ in plane bend vibration	712	718–730/709–700
$\nu_4(PO_4^{3-})$ bend	604	850–620/510–470
$\nu_4(PO_4^{3-})$ bend	565	850–620/510–470

The infrared splitting factor (IRSF) is a function that quantifies the splitting extent of the anti-symmetric bending frequency of the phosphate ν_4 vibrational mode calculated from Infrared (FTIR) spectra of bone samples; it is one of the most used parameters to assess the crystallinity as index of preservation (Weiner and Bar-Yosef, 1990).

The ν_4 (PO_4^{3-}) vibrational band shows two separate peaks at 604 and 565 cm^{-1} related to the atomic disorder and/or the crystalline size of the bioapatite (Dal Sasso et al., 2016) (**Figure.1a**). **IRSF**, calculated for each spectrum, is a parameter that quantifies the extent of splitting of the two peaks calculated with the sum of two peak intensities of ν_4 (PO_4^{3-}) divided by the intensity of the valley between them (Weiner and Bar-Yosef, 1990). The splitting factor is used to monitor the degree of recrystallization of the bioapatite as higher values of IRSF show a more ordered crystal structure with an increase in crystal size and vice versa. This correlation between crystal dimension and IR parameter was determined by comparing other analytical techniques with IRSF values (Psycharis et al., 2001; Stathopoulou et al., 2008).

The height percentage used to calculate the width of peak at 604 cm^{-1} at the 85% (**FW85%**) was considered for each spectrum in order to monitor the physical-chemical variations of bioapatite (Dal Sasso et al., 2018b). Mineral-to-matrix ratios were calculated by amide I (1660 cm^{-1}) divided by main phosphate peak intensity at 1035 cm^{-1} (ν_3 (PO_4^{3-}) (**AmI/PO_4^{3-}**). Amide I (1660 cm^{-1}) absorption bands divided by the intensity of amide II (1550 cm^{-1}) were used to monitor the collagen denaturation through the changing of bands intensity ratios (**AmI/AmII**) (Derrick MR, Stulik D, Landry JM, 1999). It is widely representative of the amount of collagen to the amount of mineral present in a bone to value collagen preservation (Paschalis et al., 2011; Figueiredo et al., 2012; Boskey, 1992, 2003; Donnelly, 2010). The carbonate content of bioapatite it was monitored by dividing the intensity of the band at 875 cm^{-1} ν_2 (CO_3^{2-}) vibrational mode of B-type carbonate on the ν_3 phosphate peak intensity (**CO_3^{2-}/PO_4^{3-}**). The amount of secondary calcite precipitated in bone structure during

diagenetic processes was calculated by dividing the characteristic vibrational band ν_4 (CO_3^{2-}) of secondary calcite at 712 cm^{-1} by the intensity peak at 1035 cm^{-1} of phosphate (**Cal/PO_4^{3-}**). For a more detailed description, see **Figure.1**, **Table.2**, **Table.3**.

Table 3. *List of parameters used.*

Parameters	**Wavenumber (cm^{-1})**	**Characterization**
IRSF	(604+565)/intensity of the valley	crystallinity index
FW85%	width at 85% of the height of the 604 cm^{-1} peak	atomic order/disorder
AmII/PO_4^{3-}	1550 cm^{-1}/1035 cm^{-1}	relative amount of collagen
AmI/PO_4^{3-}	1660 cm^{-1}/1035 cm^{-1}	amount of amide on phosphate
CO_3^{2-}/PO_4^{3-}	875 cm^{-1}/1035 cm^{-1}	amount of carbonate on phosphate
Cal/ PO_4^{3-}	712 cm^{-1}/1035 cm^{-1}	amount of calcite on phosphate

2.4. Statistical processing

Statistical analysis of mineral/matrix ratio (AmI/PO_4^{3-}) and FW85% was performed using the freely available Kernel Density Estimation (KDE) software developed by the Analytical Methods Committee of the Royal Society of Chemistry (Analytical Methods Committee, 2006). Kernel density estimation is a non-parametric method of estimating the probability density function (PDF) of a continuous random variable

$$f(x) = \frac{1}{nh} + \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right)$$

(KDE formula)

where a normal Gaussian as Kernel function (K) was chose, while smoothing parameter h (bandwidth) was statistically determined (Silverman, 1986), n is the sample size, x random variable while x_i data points locations. This function was applied to show a continuous substitution of the histograms obtaining an estimate of the underlying distribution of AmI/PO_4^{3-} and FW85% parameters, modeled on the real measured data. The dataset was divided in eight Necropolis where Ordona, Salapia and San Giovanni Rotondo have been merged into a single population, as it is possible to consider them as a single group by origin and chronology, while Sant'Anna Sopramonte was excluded as the small number of samples makes the analysis of the samples unrepresentative.

3. RESULTS

3.1. Fourier Transform-IR spectroscopy data

The spectra analyzed show signals that can be attributable to the amide, phosphates and carbonate groups. In details, peak at ~3080, 1660, 1550 cm^{-1}, correspond to Amide B, Amide I and Amide II

vibrational bands, respectively, peaks at 1340 and 1240 cm^{-1} correspond to the Amide III bands, peaks at 1450, 1420, 875, 712 cm^{-1} to ν_3, ν_2, ν_4 of carbonate absorption bands while peak at 1035, 604 and 565 cm^{-1} the stretching and bending of phosphates ν_3, ν_4 (PO_4^{3-}) vibrational bands, respectively (Derrick et al., 1999; Singh et al., 1993; Stani et al., 2020; Weiner, 2010). The data analyzed focus on the region between 3400 cm^{-1} and 480 cm^{-1}, which corresponds to most of the organic and inorganic compounds in the skeletal remains, covering Amide I-II, ν_2, ν_4 (CO_3^{2-}) and ν_3, ν_4 (PO_4^{3-}) peaks (**Figure.1**). Bioapatite crystallinity indicators as IRSF, as defined by Weiner and Bar-Yosef (1990), and the width of the 604 cm^{-1} phosphate peak (defined as full width at 85% of maximum height - FW85% (Dal Sasso et al., 2018b) was chosen and calculated for each spectrum. The carbonate ion is known to occupy two different positions in the hydroxylapatite bone structure in the hydroxyl lattice position and in that of phosphate (Sønju Clasen and Ruyter, 1997). The B-type carbonate (substituting the phosphate) is much more abundant than A-type (substituting the hydroxyl) and therefore it is assumed to produce reliable results in monitoring the carbonate content of bioapatite (Wopenka and Pasteris, 2005). Measuring the ν_3 carbonate doublet at 1450-1423 cm^{-1} is not straightforward because of the partial overlap of both A and B types and of other absorption bands of organic functional groups (Fleet, 2009; Rey et al., 1989). The ν_2 875 cm^{-1} and is attributed to carbonate B-type and is considered more useful for our study as it better reflects the real portions of carbonate principal components without overlapping to organic components. Secondary calcite may precipitate inside the micro-porosity of the archaeological bones during diagenesis (Dal Sasso et al., 2018a), for which the contribution to the IR vibrational bands overlaps to that of the structural carbonate, thus altering the CO_3^{2-}/PO_4^{3-} parameter (Gueta et al., 2007) (**Figure.2**). The amount of precipitated exogenous calcite can be monitored through the parameter Cal/PO_4^{3-} calculated by dividing the intensity of the peak at 712 cm^{-1} (ν_4 CO_3^{2-} in-plane bending characteristic vibrational mode of calcite) with the 1035 cm^{-1} intensity of the phosphate band (Dal Sasso et al., 2016). To monitor the state of sample conservation several comparisons and ratios were considered to provide a semi-quantitative measure of bone integrity based on the phosphate, carbonate and amide absorption bands.

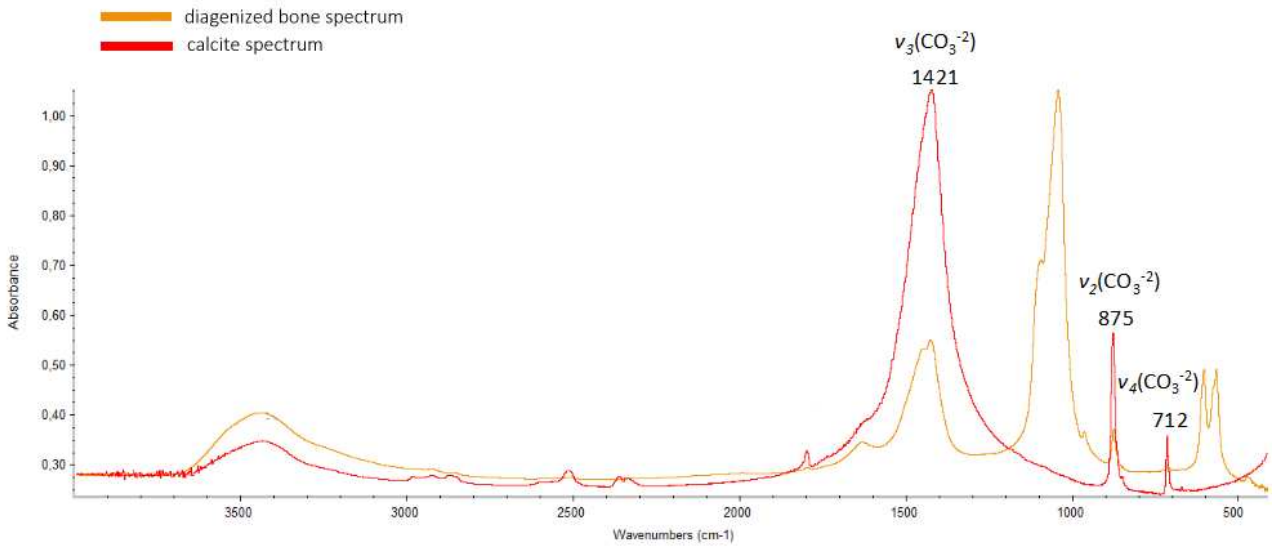

Figure 2. *FTIR spectra of calcite (red) and archaeological human bone sample (yellow, Al-Khiday 128gr5)*

3.2. Spectral analysis for bone characterization

Results (**Table.4**) show a few variations of FW85% and IRSF values for the less crystalline bioapatite (fresh and modern bone), ranging from 17.60 to 11.50 cm^{-1} and the corresponding a minimal variation of IRSF values, ranging from 2.77 to 3.24. Archaeological bones from Italy and Sudan show higher variations of IRSF and FW85%: ranges from 8.47 to 12.25 cm^{-1} and from 6.72 to 9.34 cm^{-1} for FW85%, whereas IRSF ranges from 5.13 to 2.99 and from 6.33 to 3.35, respectively.

Table 4. *Diagenetic parameters calculated from FTIR spectrum of each sample and referring to: bone crystallinity (IRSF); atomic order/disorder FW85%; secondary calcite content (Cal/PO_4^{3-}); structural carbonate content in bioapatite (CO_3^{2-}/PO_4^{3-}); collagen preservation ($Amide\ I/PO_4^{3-}$), ($Amide\ II/PO_4^{3-}$).*

Archaeological site	Bone type	Label	IRSF	FW85%	Cal/PO_4	CO_3/PO_4	$Amide\ I/PO_4$	$Amide\ II/\ PO_4$
	petrous bone	Tb.1/1	4,00	9,50	0,000	0,05	0,09	0,08
	petrous bone	Tb.2	3,71	9,85	0,000	0,06	0,09	0,09
	petrous bone	Tb.6	3,80	9,54	0,000	0,06	0,10	0,09
Montecchio	petrous bone	Tb.9	4,42	10,82	0,000	0,05	0,09	0,07
Maggiore	petrous bone	Tb.12	4,44	10,52	0,000	0,05	0,13	0,09
	tooth root	Tb.1/1	3,57	9,85	0,001	0,06	0,12	0,09
	tooth root	Tb.2	3,33	8,94	0,000	0,07	0,32	0,20
	tooth root	Tb.6	3,30	9,48	0,000	0,08	0,29	0,19
	tooth root	Tb.9	3,32	10,95	0,000	0,06	0,19	0,13
	tooth root	Tb.12	3,87	9,69	0,000	0,05	0,15	0,10
Badia Polesine	petrous bone	STheo	3,39	11,19	0,000	0,06	0,35	0,21
	petrous bone	BR-03	3,68	9,57	0,019	0,12	0,12	0,11
	petrous bone	BR-04	3,41	10,36	0,002	0,07	0,14	0,12
	petrous bone	BR-01	3,66	9,54	0,001	0,06	0,11	0,10
	petrous bone	BR-05	3,41	10,04	0,005	0,08	0,16	0,12
	tooth root	BR-25	3,18	11,20	0,002	0,09	0,25	0,18
	tooth root	BR-06	3,56	9,67	0,001	0,08	0,31	0,20
	tooth root	BR-07	3,00	12,02	0,003	0,11	0,35	0,24
	tooth root	BR-09	3,53	9,65	0,012	0,09	0,13	0,11
Grottina	tooth root	BR-16	3,10	11,54	0,000	0,08	0,31	0,21
Covoloni del	tooth root	BR-22	3,64	9,39	0,010	0,08	0,11	0,10

Broion	tooth root	BR-24	3,09	11,97	0,003	0,12	0,40	0,27
	tooth root	BR-10	3,70	9,41	0,005	0,07	0,32	0,19
	tooth root	BR-27	2,99	12,24	0,007	0,14	0,38	0,27
	tooth root	BR-13	2,99	12,25	0,001	0,09	0,37	0,24
	tooth root	BR-17	3,29	9,98	0,011	0,09	0,17	0,14
	tooth root	BR-21	3,76	9,40	0,012	0,10	0,24	0,18
	tooth root	BR-19	3,37	10,12	0,005	0,09	0,17	0,14
	tooth root	BR-12	3,11	11,80	0,005	0,11	0,38	0,25
	tooth root	BR-14	3,05	11,74	0,012	0,13	0,36	0,25
	tooth root	BR-18	3,13	11,28	0,006	0,09	0,35	0,22
	petrous bone	Tb.O.198	3,29	11,16	0,000	0,08	0,12	0,11
	petrous bone	Indiv.A	3,58	10,97	0,004	0,10	0,13	0,13
	petrous bone	Indiv.B	3,10	10,34	0,000	0,08	0,15	0,14
	petrous bone	Indiv.C	3,23	10,86	0,000	0,08	0,19	0,15
Oderzo	petrous bone	Indiv.D	3,72	11,07	0,002	0,07	0,11	0,11
	tooth root	Tb.17	3,36	9,89	0,000	0,09	0,27	0,20
	tooth root	Tb.O.1986	3,31	10,71	0,000	0,07	0,33	0,21
	tooth root	senzanome3	3,21	9,94	0,000	0,08	0,24	0,17
	petrous bone	Tb.BFC008	3,23	10,78	0,001	0,07	0,16	0,13
	petrous bone	Tb.BFC10	4,11	9,60	0,001	0,05	0,10	0,09
	petrous bone	Tb.BFC13	3,15	11,16	0,004	0,09	0,22	0,16
	petrous bone	Tb.10	3,28	10,76	0,000	0,07	0,15	0,13
	petrous bone	Tb.13	4,06	9,62	0,004	0,08	0,12	0,11
	petrous bone	Tb.19.1	3,30	11,30	0,001	0,07	0,15	0,12
	petrous bone	Tb.40	3,27	10,61	0,002	0,08	0,17	0,14
Biverone,	petrous bone	Tb.B6.1	3,73	10,01	0,003	0,07	0,11	0,10
San Stino di	tooth root	Tb.10	3,07	11,83	0,000	0,07	0,34	0,21
Livenza	tooth root	Tb.13	3,35	10,99	0,000	0,06	0,16	0,12
	tooth root	Tb.19.1	3,03	11,48	0,000	0,07	0,24	0,17
	tooth root	Tb.40	3,14	11,09	0,000	0,07	0,22	0,15
	tooth root	Tb.B6.1	3,10	10,86	0,001	0,06	0,21	0,14
	tooth root	Tb.B-111	3,25	10,86	0,000	0,08	0,29	0,19
Modern bone	petrous bone	0S1242	2,87	13,25	0,004	0,12	0,44	0,29
	petrous bone	0S419	3,03	12,09	0,000	0,07	0,50	0,27
	petrous bone	0S426	3,11	13,45	0,004	0,08	0,39	0,23
	petrous bone	0S428	3,21	11,50	0,009	0,08	0,31	0,23
	petrous bone	0S431	3,06	13,22	0,006	0,08	0,35	0,23
	petrous bone	0S1453	3,05	13,22	0,004	0,09	0,38	0,25
	tooth root	0S419	3,24	11,57	0,000	0,05	0,40	0,22
	tooth root	0S426	3,17	11,97	0,000	0,05	0,39	0,22
	tooth root	0S428	3,09	12,35	0,002	0,06	0,37	0,21
	tooth root	0S431	3,23	11,76	0,000	0,06	0,38	0,23
	tooth root	0S1453	3,05	13,67	0,002	0,06	0,35	0,21
	femur	3F-gr2	4,71	8,45	0,069	0,19	0,12	0,08
	femur	31F-gr3	5,64	7,64	0,029	0,11	0,05	0,06
	femur	50F-gr5	6,17	7,12	0,000	0,02	0,04	0,02
	femur	55-gr2	3,74	8,59	0,012	0,09	0,08	0,08
	femur	103-gr3	4,78	7,74	0,011	0,06	0,06	0,03
	femur	115F-gr4	6,24	6,72	0,000	0,02	0,04	0,02
	femur	128-gr5	4,60	8,61	0,023	0,10	0,06	0,06
	femur	136-gr5	6,33	6,77	0,002	0,03	0,04	0,02
	femur	156-gr5	4,00	8,55	0,008	0,07	0,06	0,04
	femur	158-gr4	3,90	8,90	0,007	0,07	0,06	0,06
	femur	159F-gr3	4,65	8,27	0,000	0,03	0,11	0,05
	femur	163-gr5	6,07	6,82	0,003	0,03	0,03	0,02
Al-Khiday	femur	170-gr5	4,68	7,91	0,001	0,04	0,03	0,04
	femur	177-gr4	3,35	8,48	0,003	0,06	0,06	0,06
	tooth root	Al-K_12T	3,83	8,91	0,025	0,13	0,05	0,08
	tooth root	Al-K_34T	5,65	7,53	0,003	0,03	0,02	0,02
	tooth root	Al-K_35T	3,85	8,65	0,019	0,10	0,05	0,07
	tooth root	Al-K_58T	5,46	8,21	0,000	0,02	0,03	0,03

	tooth root	Al-K_95T	5,76	7,53	0,005	0,03	0,02	0,02
	tooth root	Al-K_104T	4,90	7,93	0,015	0,06	0,02	0,04
	tooth root	Al-K_163T	5,47	7,66	0,012	0,05	0,03	0,03
	tooth root	Al-K_164T	5,35	7,47	0,009	0,04	0,03	0,03
	tooth root	Al-K185T	3,47	9,13	0,022	0,13	0,05	0,10
	tooth root	Al-K_186T	5,28	8,36	0,000	0,02	0,04	0,03
	tooth root	Al-K_187T	5,77	7,46	0,000	0,02	0,03	0,02
	petrous bone	Al-K_12PB	3,89	9,34	0,009	0,08	0,04	0,06
	petrous bone	Al-K_34PB	5,53	8,11	0,005	0,03	0,06	0,06
	petrous bone	Al-K_35PB	3,62	9,27	0,007	0,08	0,06	0,09
	petrous bone	Al-K_58PB	5,75	8,45	0,001	0,02	0,02	0,02
	petrous bone	Al-K_95PB	5,79	7,62	0,004	0,03	0,02	0,02
	petrous bone	Al-K104PB	4,33	8,08	0,005	0,07	0,07	0,07
	petrous bone	Al-K163PB	5,28	8,03	0,002	0,03	0,03	0,04
	petrous bone	Al-K164PB	5,87	7,53	0,003	0,03	0,02	0,02
	petrous bone	Al-K185PB	3,57	9,27	0,004	0,07	0,04	0,07
	petrous bone	Al-K186PB	5,79	8,44	0,000	0,02	0,03	0,03
	petrous bone	Al-K187PB	5,12	8,19	0,000	0,03	0,03	0,03
	tooth root	Tb.1	4,06	8,47	0,002	0,03	0,24	0,11
Sant'Anna	tooth root	Tb.3	3,19	11,18	0,014	0,05	0,30	0,19
Sopramonte	petrous bone	US48	3,35	9,94	0,007	0,09	0,11	0,11
	tooth root	US48	3,00	11,60	0,006	0,08	0,32	0,22
	petrous bone	Tb.1	4,20	9,15	0,000	0,09	0,11	0,08
	tooth root	SAL001	3,73	9,58	0,002	0,06	0,18	0,12
	petrous bone	SAL003	4,11	9,94	0,000	0,04	0,12	0,08
	petrous bone	SAL005	4,45	9,59	0,001	0,05	0,11	0,07
Salapia	petrous bone	SAL007	5,13	8,67	0,002	0,03	0,11	0,05
	tooth root	SAL008	3,67	9,99	0,001	0,06	0,23	0,14
	petrous bone	SAL010	5,12	8,96	0,018	0,11	0,04	0,07
	tooth root	SAL011	3,22	11,44	0,003	0,09	0,40	0,24
	tooth root	SAL012	4,58	9,12	0,006	0,05	0,07	0,04
	petrous bone	ORD001	4,07	9,09	0,001	0,04	0,08	0,06
	petrous bone	ORD004	3,87	10,39	0,003	0,05	0,10	0,08
	petrous bone	ORD006	4,39	9,43	0,003	0,04	0,05	0,05
	petrous bone	ORD009	4,61	9,75	0,001	0,04	0,05	0,04
Ordona	petrous bone	ORD010	4,16	9,59	0,051	0,17	0,14	0,13
	petrous bone	ORD011	3,51	10,78	0,023	0,17	0,19	0,17
	petrous bone	ORD012	4,36	9,16	0,000	0,04	0,05	0,04
	petrous bone	ORD014	4,37	9,04	0,027	0,14	0,11	0,10
	petrous bone	ORD018	4,13	9,23	0,001	0,04	0,06	0,05
	petrous bone	ORD019	4,29	9,70	0,000	0,05	0,06	0,05
	tooth root	SGR001	3,47	10,41	0,002	0,05	0,18	0,12
San Giovanni	tooth root	SGR002	3,81	9,28	0,000	0,04	0,10	0,07
Rotondo	tooth root	SGR003	3,27	11,44	0,003	0,05	0,27	0,16
	femur	Dog08	12,93	3,04	0,001	0,09	0,44	0,29
	femur	Horse10	16,09	2,82	0,002	0,11	0,62	0,41
	femur	O.sheep11	13,13	2,99	0,001	0,09	0,39	0,26
Fresh bone	femur	Pig10	17,60	2,85	0,003	0,13	0,64	0,44
	tooth root	Y.sheep06	14,34	2,85	0,002	0,07	0,39	0,23
	tooth root	O.sheep05	12,23	3,05	0,001	0,06	0,32	0,20
	tooth root	O.sheep06	12,40	3,04	0,001	0,06	0,32	0,20
	tooth root	Horse07	14,33	2,77	0,000	0,08	0,36	0,23
	femur	Oxbone	12,51	3,01	0,001	0,06	0,39	0,24
	tooth root	Human25	2,99	2,99	0,000	0,08	0,37	0,23

Sudanese archaeological bones show significantly higher value in IRSF and lower value of FW85% than modern bones while the archaeological bones from Italian sites fall within this range of values (view **Figure.3a**).

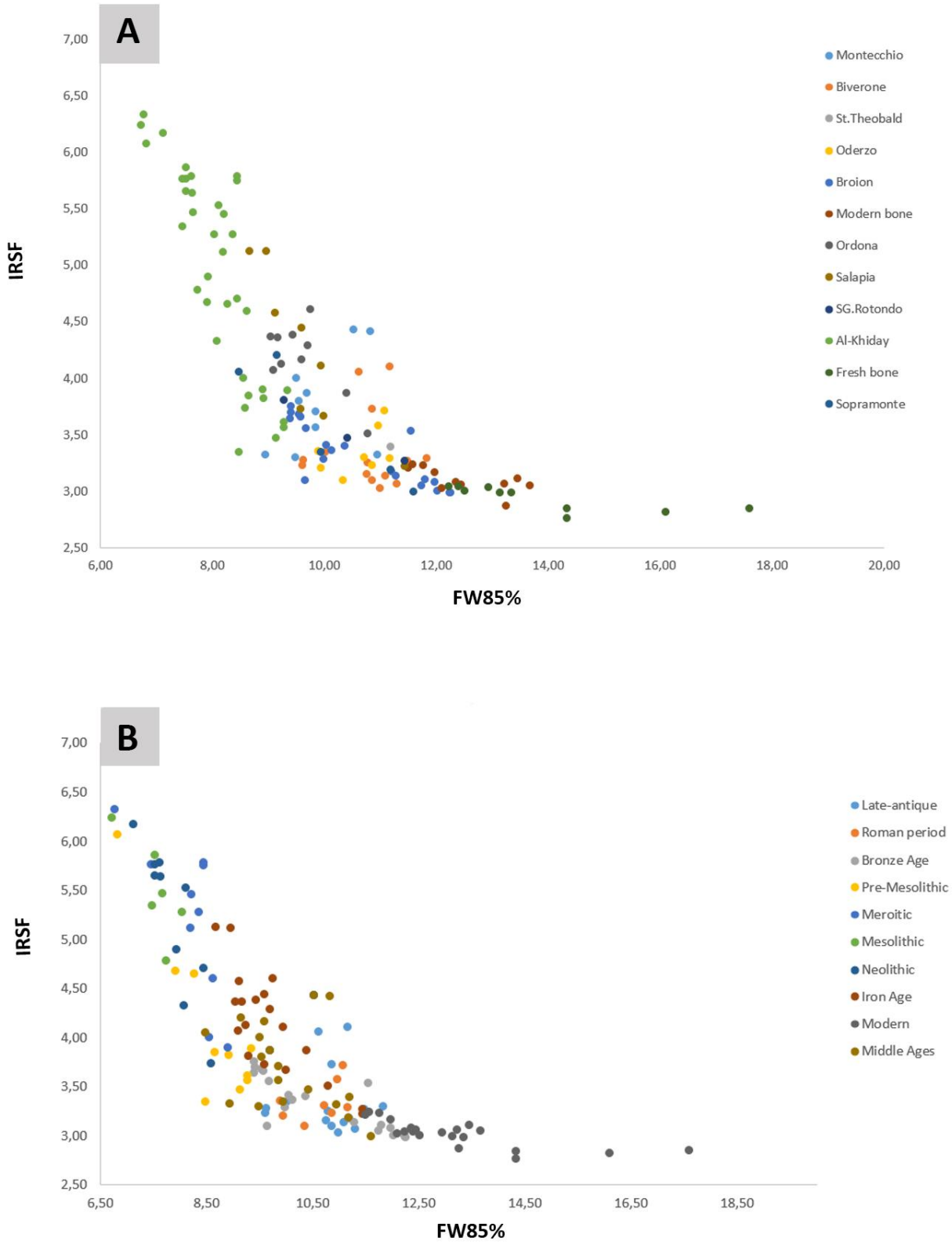

Figure 3. *Correlation between the IRSF and FW85% among bioapatite samples. The bone samples were divided by origin (**a**) chronological phases (**b**).*

Figure.3b and **Figure.4** highlights distribution of spectral data dividing them by chronological phase and anatomical elements. **Figure.4** shows the distribution of modern and archaeological context of

teeth and petrous bones and the IRSF-FW85% plot describes slight differences between the sets of samples divided by anatomical elements, as the roots of the teeth are more clustered in the final end of the curve indicating a better preservation within values between 7.53 to 12.45 of FW85% and 5.77 to 3.06 of IRSF, with a median value 10.67 for peak width and 3.32 for crystallinity index.

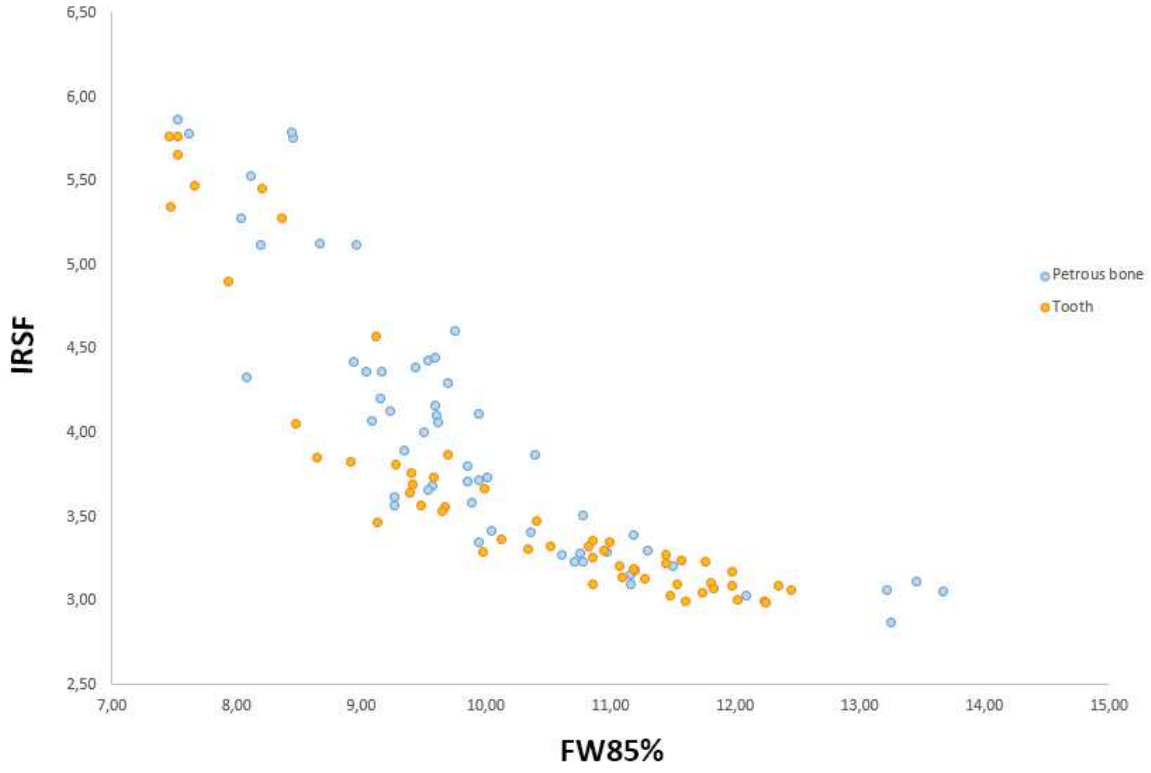

Figure 4. *Correlation between the IRSF and FW85% division by anatomical elements.*

The CO_3^{2-}/PO_4^{3-} ratio, calculated for all samples and reported in **Table.4**, ranges from 0.19 to 0.02 for Sudanese samples, from 0.17 to 0.03 for the Italian ones and from 0.12 to 0.05 for fresh and modern bones. Secondary calcite ($CaCO_3$) deriving from dissolved carbonate ions and free calcium ions present in the burial environment (Fernández-Jalvo et al., 2016) was identified in 73 samples identified by the presence of the peak at 712 cm^{-1} (Gueta et al., 2007). The Cal/PO_4^{3-} ratio displays the amount of exogenous calcite precipitated in the bone micro-structure with a content between 0.069 and 0.001 (**Table.4**). Post-yield mechanical properties of bone are mostly determined by collagen (Currey, 2008a, 2008b; Garnero, 2015) and it is essential to define the spectroscopic biomarkers to investigate the quantity of collagen. The changes in these ratios with denaturation may reflect alterations in the collagen secondary structure, specifically a transition from ordered to less-ordered structure. A variation of levels of preserved collagen content is recorded by the measured values of

AmI/PO_4^{3-} (from 0.38 to 0.04, from 0.12 to 0.02 for Italian and Al-Khiday samples, respectively and from 0.64 to 0.15 for fresh and modern bones) (**Table.4**). The AmII/ PO_4^{3-} were obtained values are from 0.27 to 0.03, 0.10 to 0.02 for Italian and Sudanese samples, respectively and ranging from 0.44 to 0.23, 0.29 to 0.21 for fresh and modern bones.

4. DISCUSSION

The set of osteoarchaeological human samples here analyzed is large and diverse for chronology and origin with a high variability of in terms of chemical composition and structure due to different diagenetic processes experienced during burial. The same bone elements, such as teeth roots and petrous bones, were examined for each selected necropolis.

The width at 85% of the height of the ν_4 (PO_4^{3-}) vibrational mode at 604^{-1} (FW85%) - chosen as it produces a well-defined signal being less influenced by other overlapping vibrational modes - describes the variation of bioapatite structural properties in IR spectra (Dal Sasso et al., 2018b).

IRSF is less sensitive to describe the significant variation of well-preserved and fresh bones, in that from results, this parameter does not show high difference of splitting factor's values while the FW85% parameter values show higher variations in terms of physical-chemical properties. Conversely, more diagenized archaeological bones reflect high variations of the splitting factor and smaller change on peak broadening. Hence, IRSF variations may misleadingly reflect the structural changes that occur in the nanostructure, whereas chemical/structural information on the phosphate environment in bioapatite is provided by the infrared peak width (FW85%). These characteristics determined the choice of the peak width used to monitor bioapatite crystallinity.

Phosphate peak width has been shown to effectively describe the chemical environment of bioapatite crystallites by monitoring minimal changes in the chemical composition and structure of crystal lattice. Bioapatite in bone system is chemically complex, in that it is structural by defect-rich and prone to substitutions (i.e. carbonate ions, LeGeros, 1981; Rey et al., 2009) with respect to crystalline hydroxylapatite of geologic origin. The occurrence of ion substitutions and vacancies in the bioapatite crystal lattice directly influence the IR vibrational modes of functional groups, e.g. phosphate, which can be detected through the measurement of the peak width (Wopenka and Pasteris, 2005). Moreover, the occurrence of structural defects plays a major role in confining the crystallites size at the nanoscale (Dal Sasso et al., 2018b). This correlation is expressed well comparing FW85% values with those of carbonate (CO_3^{2-}/PO_4^{3-} - FW85%), measured by the contribution of ν_2 (CO_3^{2-}) vibrational band divided by main ν_3 (PO_4^{3-}) mode (CO_3^{2-}/PO_4^{3-}). Although it may be extremely useful and indicative of the atomic order/disorder of the crystal lattice, this tool is of little use for the study of archaeological bones.

The state of conservation among the different bone finds is highly represented by the IR data, which translate into clear and distinct spectra showing variability of the bioapatite properties, and diagenetic recrystallization. By comparing the new data with of a previously published dataset (Dal Sasso et al., 2018) which used different types of bone tissue from human archaeological remains and fresh animal (indicated in the **Figure.5** as "Series2") it is observed that FTIR spectra reflect the degree of conservation of the bone samples utilizing FW85%-IRSF parameters.

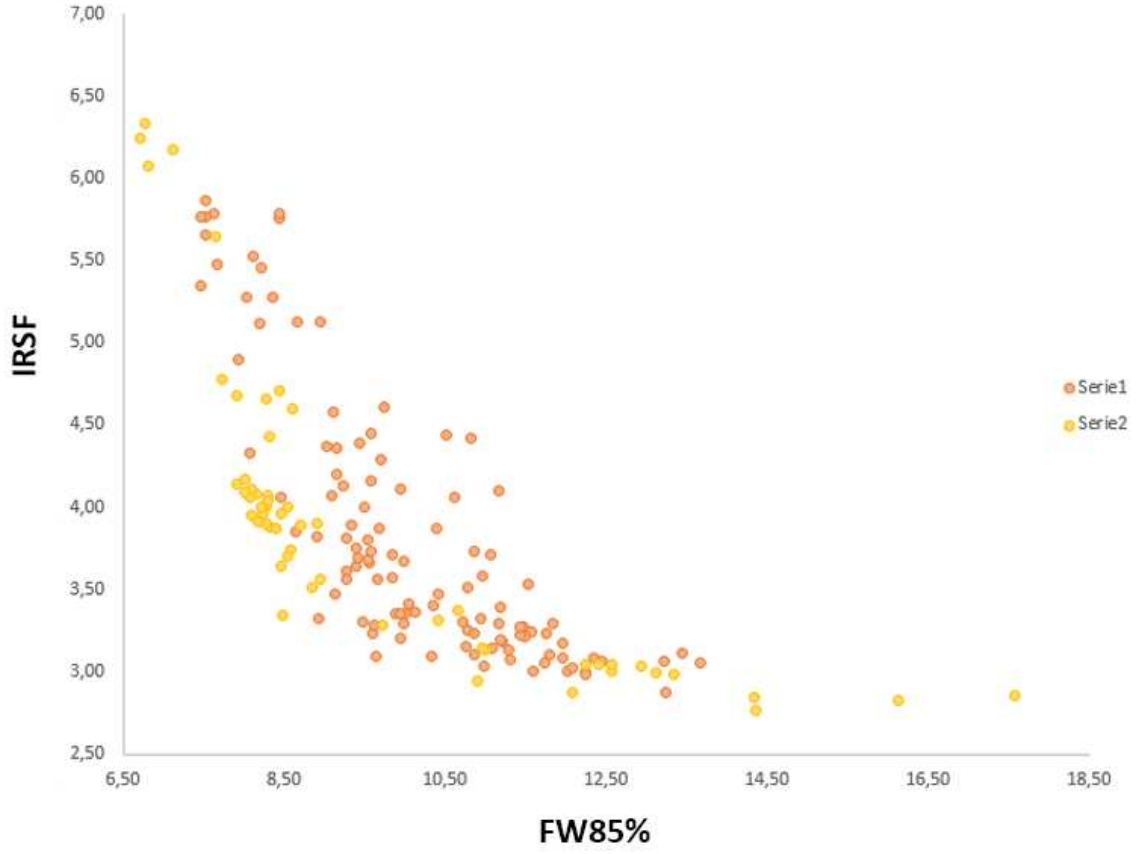

Figure 5. *Comparison between new data with of a previously published dataset.*

Comparing the two datasets (Series1-Series 2), the biogenic alteration from different diagenetic pathways would appear to follow the same trend. It is known that diagenesis impairs histological integrity increasing bone porosity and mineral crystallinity, resulting in the loss collagen (Hedges and Millard, 1995). The extent of these alterations depends on burial conditions, external features and time, the latter especially when this concerns paleontological bones. Therefore, the degree of alteration can vary but not the mechanisms of how diagenesis. As a general trend, the increase of the phosphate peak width is followed by a correspondent decrease of IRSF, indicating chemical/structural changes in the crystallites.

Figure.6a and **Figure.6b** show that the contribution of secondary calcite heavily affects the assessment of the carbonate content of the bioapatite mineral phase. The secondary calcite signal,

frequently detected in archaeological bones and precipitated within the bone micro-pores during diagenesis, overlaps with the vibrational modes ν_2 (CO_3^{2-}) and ν_3 (CO_3^{2-}) characteristic of the carbonate ions of bioapatite. The peak at 712 cm^{-1} is characteristic of calcite and reveals its presence when it is present in conspicuous quantities (over ~ 2.5 wt.%) (Dal Sasso et al., 2016).

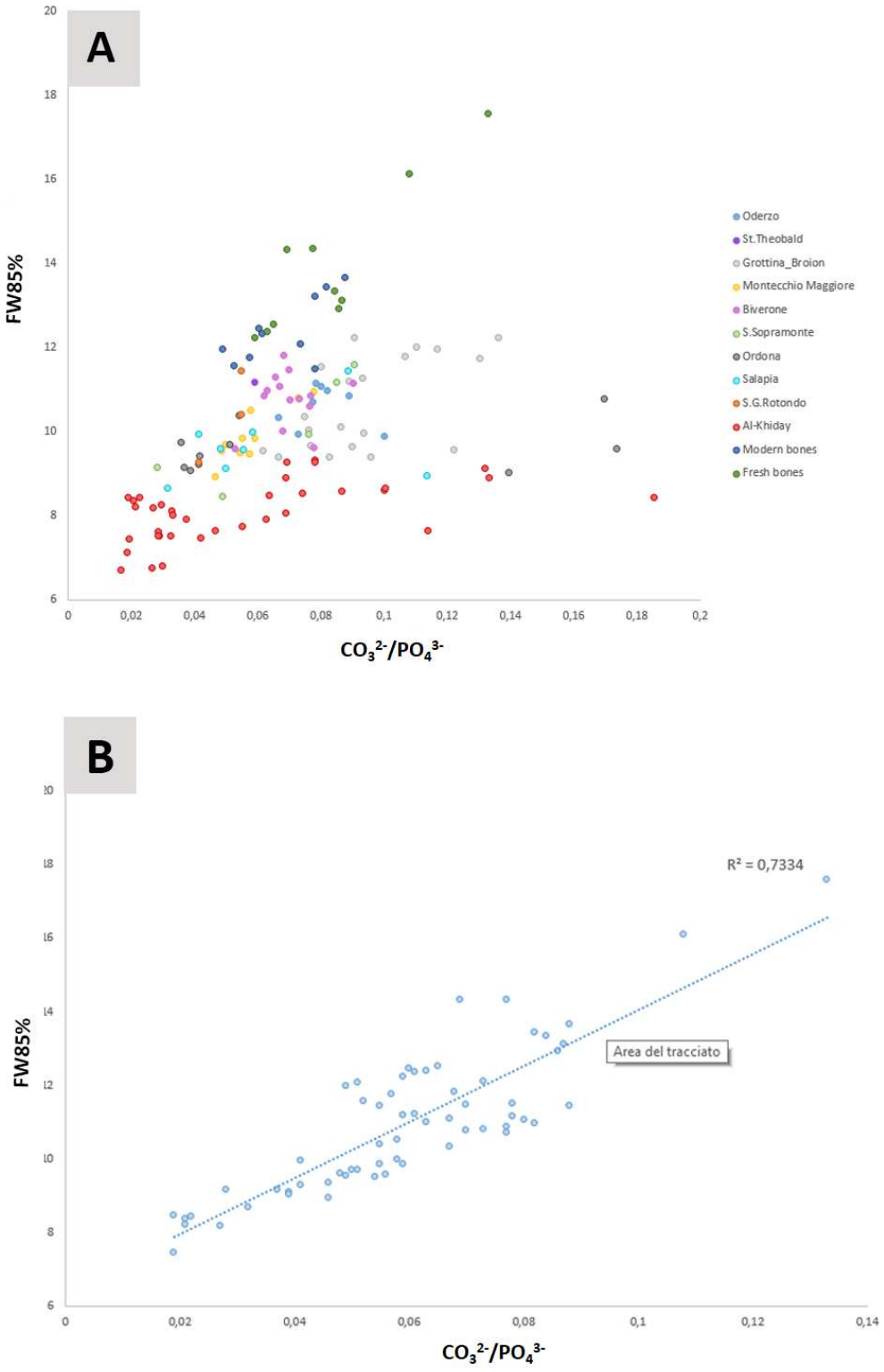

Figure 6. *The atomic order/disorder (FW85%), morphology and size of BAp crystals are highly correlated and dependent on carbonate concentrations (CO_3^{2-}/PO_4^{3-}). The limit of this parameter is the overlap with secondary calcite that influence the results.*

This correlation is improved when 73 samples, showing the secondary calcite peak at 712 cm^{-1}, were removed (R^2 = 0.733. **Figure.6b**), proving that the exposure to the conditions of the burial environment significantly influence the amount of calcite in the bone matrix and consequently the measurements on the carbonate, while the FW85% parameter remains the most reliable indicator.

Infrared spectroscopy is a very powerful tool to investigate the degree of alteration of bone structure in that allows deducing quality as well presence/absence of organic matrix, collagen of type I.

The absorption bands of Amide I and Amide II at 1660 cm^{-1} and 1550 cm^{-1} respectively, were selected for this study. The comparison between the AmI/PO_4^{3-} and AmII/PO_4^{3-} absorption bands, characteristic of the peptide bond, was used to monitor the presence/absence of the organic matter distinguishing the altered samples that present only structural or weakly bound water. In fact, the absorption band of water overlaps that of the amides altering its determination (Fellows et al., 2020). These parameters discriminate discreetly well the samples without organic matter from the richer ones as evidenced in **Figure.7**.

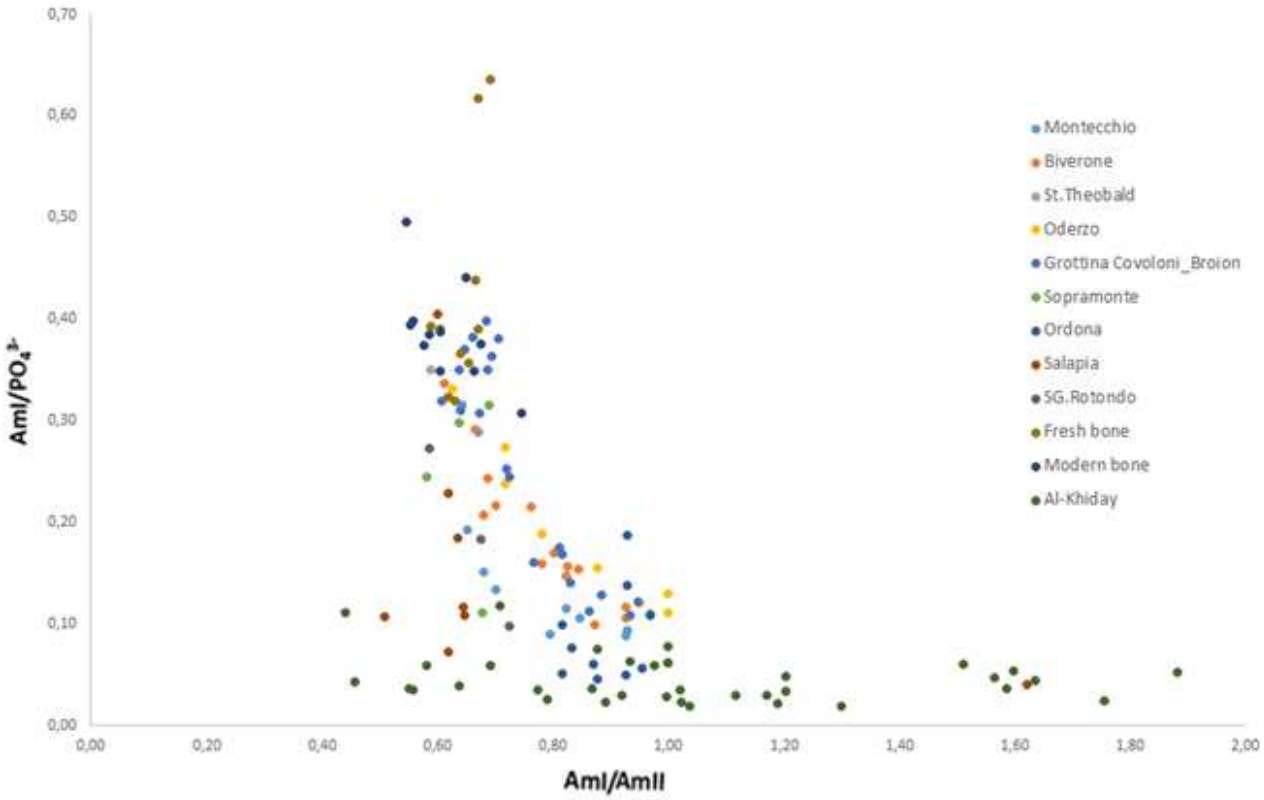

Figure 7. *Plot shows the correlation between the AmI/PO_4^{3-} and AmI/AmII to determine the presence of organic matter preserved in archaeological bones.*

Comparing AmI/PO_4^{3-} - AmI/AmII, the graph discriminates the well-preserved samples from the poorly ones. Taking as a reference the "fresh bones" (containing excellent quality and large quantity of collagen, i.e. AmI/PO_4^{3-} > 0.1), the best preserved samples in terms of quality/quantity of the organic component show a high value of AmI/PO_4^{3-} and a constant ratio between the two intensities

of the amides (i.e. AmI/AmII = 0.55-0.75). Samples that contain collagen, but in small quantities, show a constant value of AmI/AmII and a decreasing value of AmI/PO_4^{3-}, while the more degraded samples show low values of mineral to matrix ratio (AmI/PO_4^{3-}), mostly referring to absorbed water, and a lack of correlation between AmI and AmII.

The statistical analysis of AmI/PO_4^{3-} values distribution through the Kernel Density Estimation (KDE) (**Figure.8**) of eight necropolis shows significant variations between different groups of samples in terms of frequency distribution of collagen content. KDE shows that fresh and modern bones can be easily distinguished from altered ones by dividing them into three distinct macro groups. Within fresh and modern bone samples, minimal differences in terms of the preservation of the organic component can be identified.

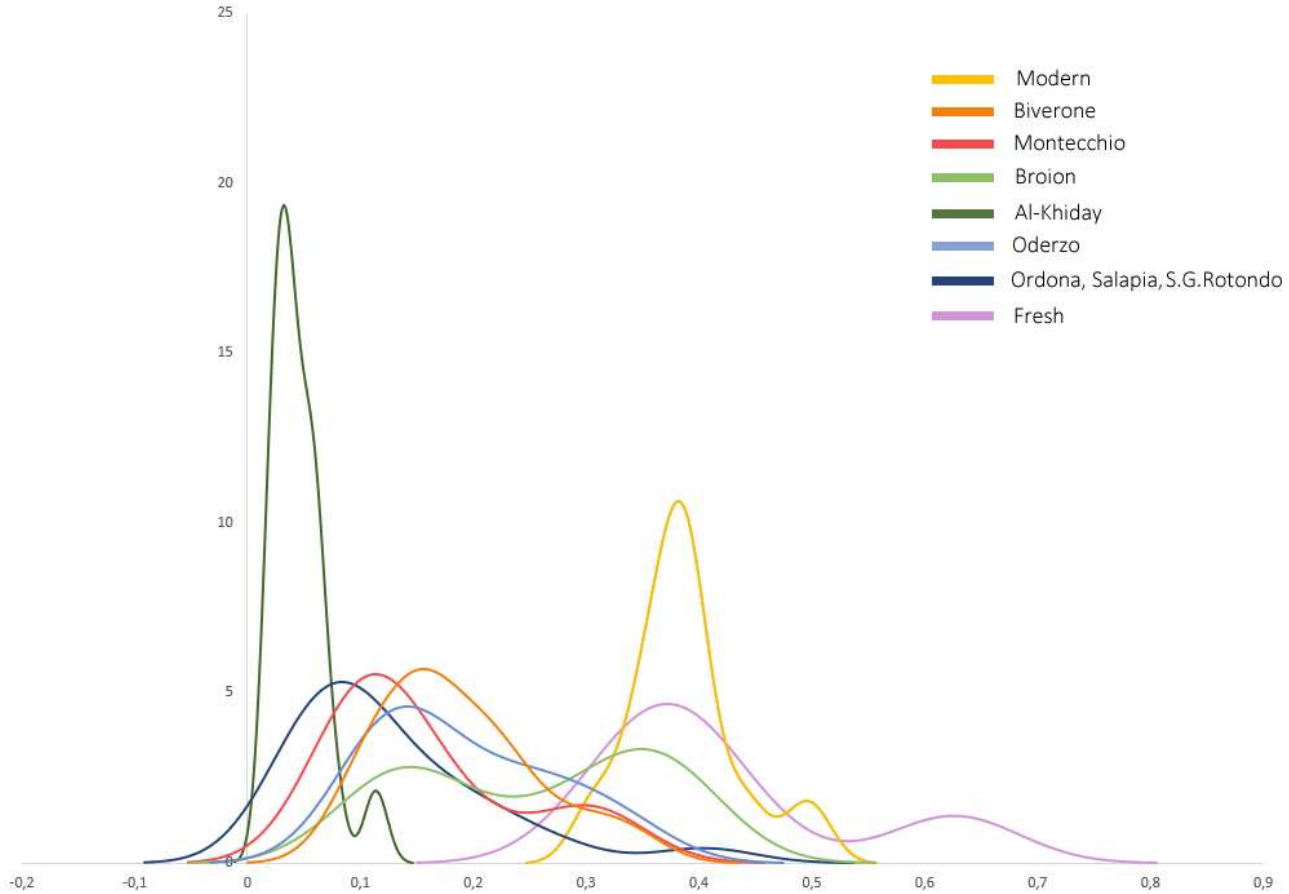

Figure 8. *Kernel Density Estimation (KDE) of the AmI/PO_4^{3-} measured for all samples. Graph shows significant variations between different groups of samples in terms of frequency distribution of collagen content. The y-axis in a density plot is the Kernel density function for the kernel density estimation while in x-axis are reported FTIR values of parameter used.*

The archaeological bones from the Al-Khiday site (Central Sudan), which have undergone extreme diagenetic processes, show a highly reduced range of variation in terms of samples variability. Its bimodal distribution could be an artificial behavior since the AmI/PO_4^{3-} - AmI/AmII graph shows that there is virtually no bone collagen. For the archaeological bones of the "Covoloni del Broion" cave

site, having a conspicuous quantity of collagen, this parameter distinguishes a variety of organic conservation within a set of samples from the same site.

The results show that AmI/PO_4^{3-} is a very useful tool for quantifying the presence of collagen preserved in the bone system but is not very reliable for the highly degraded samples that are free of organic matter and contain mainly structural water, due to environmental and burial conditions. Archaeological bones can incorporate water within the bone structure, the signal of which overlaps the amide band. In this regard, AmI/PO_4^{3-} - AmI/AmII could be a useful and more sensitive plot for discriminating well samples with collagen from those without.

Teeth are often the most sought-after anatomical elements for test involving biological molecules (i.e. aDNA, istopic analysis) in that the root substrate is anatomically protected and better preserved because of their alveolar location, the enamel cover, and the lower porosity with respect to the bony tissues (Adler et al., 2011).

Like teeth, also dense bone parts such as petrous bones are particularly suitable for biomolecular research, providing relevant results (Gonzalez et al., 2020; Pinhasi et al., 2015). This suggests that the inner, denser part is protected from external factors. However, this assumption may be questioned by the results showed in the **Figure.4**, where close similarities in diagenetic processes due to tissue organization and bone chemical composition can be observed. Overall, it can be deduced that the teeth are preserved better, but not always, and this mainly depends on the conditions of burial and it should not be used as an assumption.

This behavior is intimately correlated with collagen loss due to environmental factors and microorganisms that alters the inorganic portion of bone and corrodes the bone microstructures creating tunnelling (Dal Sasso et al., 2014; Hedges, 2002). The presence of collagen ensures stability of the mineral component while its alteration or its loss causes its recrystallization. Therefore, there is an intimate chemical-physical association between the collagen fibrils and crystallites, and this chemical affinity determines a structural rearrangement of bioapatite (Mendes et al., 2012). **Figure.9** shows that the width of phosphate peak is linearly correlated to the amount of organic matrix, the measurement of which is provided by the AmI/PO_4^{3-} parameter. The width of the phosphate peak (FW85%) correlates to the collagen content (AmI/PO_4^{3-}) showing the dependence between the crystallite size and the loss of collagenous tissue, since the presence of collagenous material in the hierarchical structure of the bone plays an important role in controlling the crystallite size in bioapatite. As reported in the literature (Collins et al., 1995; Trueman and Tuross, 2019), recrystallization proceeds together with hydrolysis of collagen, so that inter-crystalline space left by collagen loss is replaced by growth of reprecipitated apatite and of secondary calcium carbonate, when allowed by environmental conditions. Regardless of the chronology, origin and diagenetic

pathways, the figure shows that samples containing different quantities of collagenous material follow a single trend that can be reasonably approximated by a linear regression model.

Figure.3a shows that chronology does not affect conservation: the Meroitic samples from Al-Khiday (150 BCE - 250 CE) are poorly-preserved compared to the pre-Mesolithic ones (before 7000 BCE). The bone finds from the same archaeological site have a different state of conservation regardless of the chronological phase. For example, looking at Salapia (1100 BCE - 600 BCE, southeastern Italy) and Al-Khiday (150 BCE - 250 CE, central Sudan) the first group of samples is more preserved than the last. Conversely, the local burial/environmental conditions are the main factors affecting the preservation (**Figure.3b**). Similarly, the roughly coeval samples found in karst environment such as "Covolini del Broion" cave (Bronze Age, northern Italy) and in arid environment such as the Meroitic samples of Al-Khiday show significant differences in terms of crystallinity and collagen content.

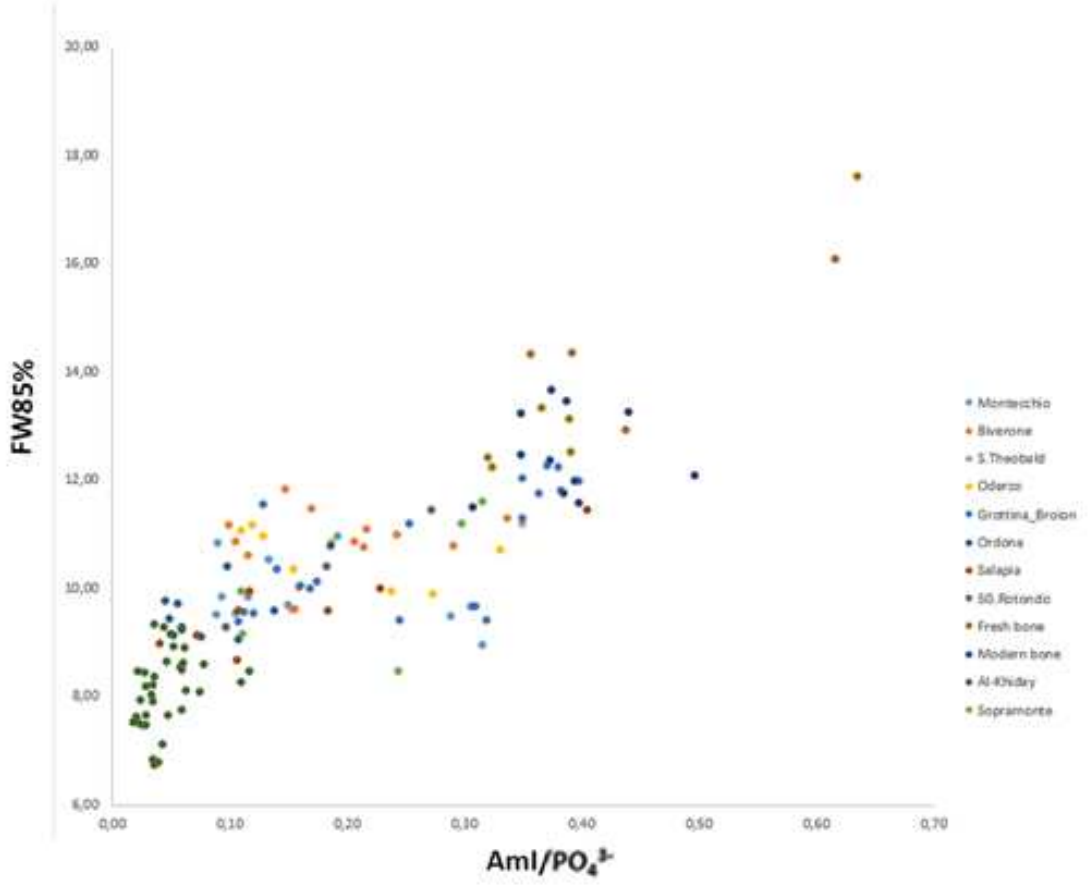

Figure 9. *Correlation between FW85% and AmI/PO_4^{3-}. The AmI/PO_4^{3-} data correlates with the structural order/disorder indicating that a greater recrystallization determines a loss of organic component.*

The KDE of (**Figure.10**) of the phosphate peak width distribution (FW85%) shows significant variations among different necropolises. This parameter is extremely significant to subdivide highly altered bones (Al-Khiday) from well-preserved ones, distinguishing effectively more altered from poorly diagenized samples.

Observing the subsets of bone samples, minimal significant differences are identified on the degree of alteration/recrystallization among samples of the same group. The trend of FW85% distribution and AmI/PO_4^{3-} establishes that FW85% is not only indirectly indicative of the quantity of carbonate in the crystal lattice of bioapatite but is sensitive to the presence of collagen left during bioapatite recrystallization. The high sensitivity of the phosphate peak width could therefore be used to select less diagenized and less altered samples useful for isotopic analysis or ancient DNA extraction. FW85% provides a reliable evaluation of material properties even in subsamples pertaining to the same necropolis with minimal variations: bones with low crystallinity are characterized by very high and diversified values of FW85% assuming different biomechanical and chemical properties.

The work discussed here shows the global action of diagenesis and how it affects bone tissue.

Although bone samples may undergo different levels of alteration (e.g. due to particular environmental and burial conditions), based on the analyzed data we can see that diagenesis acts according to a generalized model, progressively losing collagen, and structural carbonate with consequent strong recrystallization of apatite characterized by a higher structural order.

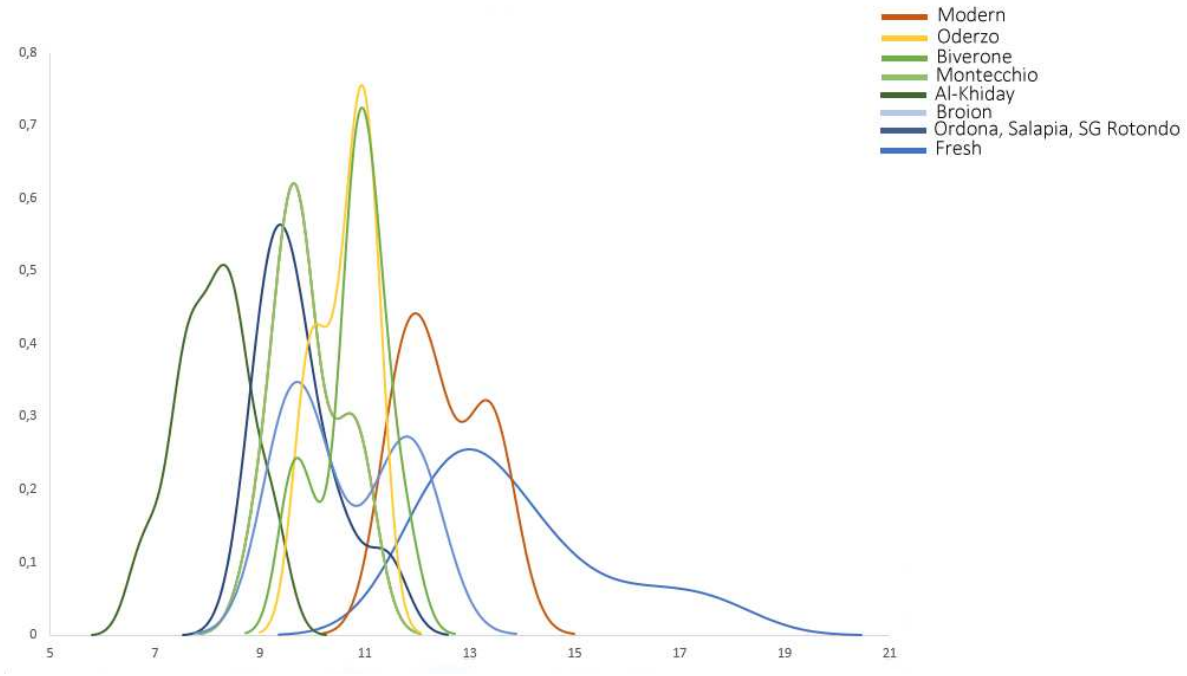

Figure 10. *Kernel Density Estimation (KDE) of the FW85%. In y-axis Kernel density function, in x-axis are reported FTIR values of parameter used. Graph shows significant variations among different necropolises using FW85% parameter. From the statistical analysis it is possible to see the distribution of the more or less narrow diagenetic alterations which indicates that the diagenetic processes have acted uniformly. While there are more or less wide distributions in the sites where diagenesis has acted differently.*

5. CONCLUSIONS

The results obtained with a low cost, highly sensitive and minimally invasive analytical technique such as IR spectroscopy, demonstrate the correlation between bioapatite structural order and

crystallinity and the presence of organic material. The wide set of osteological specimens characterized by different diagenetic factors ranging from extremely altered specimens to modern and fresh bones has allowed to attest and confirm the validity of phosphate peak width at FW85% as a reliable parameter to obtain relevant chemical/structural information on bone material both for the organic and inorganic phases. This parameter is advantageous for monitoring minimal changes in the structure and on chemical properties of bioapatite as well as indirectly in collagen. This results show that diagenesis affect the preservation of bones in different extent according to a generalized model showing a progressive loss of collagen together with the recrystallization of bioapatite and the loss of structural carbonate. This minimally invasive spectroscopic technique can in fact be applied as a pre-screening method for the preservation of bone collagen as it adequately contributes to describing the *post-mortem* processes undergone by the samples. This method could improve the specimens selection process for isotopic or genetic analysis and radiocarbon dating in archaeological, molecular anthropological and forensic fields.

ETICHS STATEMENT

Fresh human bone sample was provided by Dr Sunita Ho, University of California, San Francisco, with the necessary approvals, as reported in Asscher et al. (2011) while Sudanese fossil bones were provided by Sudanese and Sub-Saharan Studies Center (CSSeS) as reported in Dal Sasso et al. (2018). The animal fresh bones used do not come from animals threatened or endangered species and were not sacrificed for this research purposes. Fresh samples were taken from a slaughterhouse while samples of dogs and horses were obtained from a veterinary hospital (Dal Sasso et al., 2018).

AUTHOR CONTRIBUTION

Cinzia Scaggion: conceptualization, data curation, formal analysis, investigation, writing - original draft; **Gregorio Dal Sasso:** conceptualization, data curation, formal analysis, writing - review & editing; **Luca Nodari:** validation, supervision, writing - review & editing; **Luca Pagani:** supervision, writing - review & editing; **Nicola Carrara:** resources, writing - review & editing; **Donatella Usai:** resources, writing - review & editing; **Giulia Gadioli:** resources, writing - review & editing; **Gilberto Artioli:** project administration, conceptualization, supervision, writing - review & editing.

ACKNOWLEDGEMENTS

We thank Dr. Christophe Snoeck Ph.D., researcher at Vrije Universiteit Brussel and the head of the Brussels Bioarchaeology Lab (BB-LAB) and Dr. Ina Reiche Ph.D., research director of French National Centre for Scientific Research-CNRS for their valuable scientific advices.

DECLARATION OF COMPETING INTEREST

The authors do not have a conflict of interest to declare and the manuscript has not been published or presented to another journal.

Table 1. *Information on the included samples: sample number, skeletal anatomic element, chronology, origin.*

SAMPLE	CHRONOLOGY	ARCHAEOLOGICAL SITE	SKELETAL ELEMENT
Tb.1/1_MonPbsx	7th-8thc.AD	Montecchio Maggiore, Italy	petrous bone
Tb.2_MonPbsx	7th-8thc.AD	Montecchio Maggiore, Italy	petrous bone
Tb.6_MonPbsx	7th-8thc.AD	Montecchio Maggiore, Italy	petrous bone
Tb.9_MonPbsx	7th-8thc.AD	Montecchio Maggiore, Italy	petrous bone
Tb.12_MonPbdx	Neolithic (?)	Montecchio Maggiore, Italy	petrous bone
Tb.1/1_MonT	7th-8thc.AD	Montecchio Maggiore, Italy	tooth
Tb.2_MonT	7th-8thc.AD	Montecchio Maggiore, Italy	tooth
Tb.6_MonT	7th-8thc.AD	Montecchio Maggiore, Italy	tooth
Tb.9_MonT	7th-8thc.AD	Montecchio Maggiore, Italy	tooth
Tb.12_MonT	Neolithic	Montecchio Maggiore, Italy	tooth
Tb.BFC008_BMGPbdx	4th-6thc.AD	Biverone-San Stino di Livenza, Italy	petrous bone
Tb.BFC10_BMG Pbdx	4th-6thc.AD	Biverone-San Stino di Livenza, Italy	petrous bone
Tb.BFC13_BMG Pbdx	4th-6thc.AD	Biverone-San Stino di Livenza, Italy	petrous bone
Tb.10_BPbsx	4th-6thc.AD	Biverone-San Stino di Livenza, Italy	petrous bone
Tb.13_BPbdx	4th-6thc.AD	Biverone-San Stino di Livenza, Italy	petrous bone
Tb.19.1_BPbdx	4th-6thc.AD	Biverone-San Stino di Livenza, Italy	petrous bone
Tb.40_BPbsx	4th-6thc.AD	Biverone-San Stino di Livenza, Italy	petrous bone
Tb.B6.1_BPbsx	4th-6thc.AD	Biverone-San Stino di Livenza, Italy	petrous bone
Tb.10_BM2lowsx	4th-6thc.AD	Biverone-San Stino di Livenza, Italy	tooth
Tb.13_BM2uppdx	4th-6thc.AD	Biverone-San Stino di Livenza, Italy	tooth

Tb.19.1_BM2uppdx	4th-6thc.AD	Biverone-San Stino di Livenza, Italy	tooth
Tb.40_BM3uppdx	4th-6thc.AD	Biverone-San Stino di Livenza, Italy	tooth
Tb.B6.1_BM2uppdx	4th-6thc.AD	Biverone-San Stino di Livenza, Italy	tooth
Tb.B-111_BMG	4th-6thc.AD	Biverone-San Stino di Livenza, Italy	tooth
BR-03_ PbGCB O-Ce	Bronze Age	Grottina Covoloni del Broion, Italy	petrous bone
BR-04_PbGCB O-Ce	Bronze Age	Grottina Covoloni del Broion, Italy	petrous bone
BR-01_Pb GCB O:38	Bronze Age	Grottina Covoloni del Broion, Italy	petrous bone
BR-05_ PbRIMANEGGIATO	Bronze Age	Grottina Covoloni del Broion, Italy	petrous bone
BR-25_GCB	Bronze Age	Grottina Covoloni del Broion, Italy	tooth
BR-06_GCB	Bronze Age	Grottina Covoloni del Broion, Italy	tooth
BR-07_GCB	Bronze Age	Grottina Covoloni del Broion, Italy	tooth
BR-09_GCB	Bronze Age	Grottina Covoloni del Broion, Italy	tooth
BR-16_ GCB 4B+4S 5S	Bronze Age	Grottina Covoloni del Broion, Italy	tooth
BR-22_ GCB 7-353	Bronze Age	Grottina Covoloni del Broion, Italy	tooth
BR-24_ GCB 53	Bronze Age	Grottina Covoloni del Broion, Italy	tooth
BR-10_GCB 223	Bronze Age	Grottina Covoloni del Broion, Italy	tooth
BR-27_GCB 275	Bronze Age	Grottina Covoloni del Broion, Italy	tooth
BR-13_ GCB274	Bronze Age	Grottina Covoloni del Broion, Italy	tooth
BR-17_GCB 629	Bronze Age	Grottina Covoloni del Broion, Italy	tooth
BR-21_GCB 807	Bronze Age	Grottina Covoloni del Broion, Italy	tooth
BR-19_ GCB 909	Bronze Age	Grottina Covoloni del Broion, Italy	tooth
BR-12_ GCB O:156	Bronze Age	Grottina Covoloni del Broion, Italy	tooth
BR-14_ GCB SeH6 Str248	Bronze Age	Grottina Covoloni del Broion, Italy	tooth
BR-18_GCB 221	Bronze Age	Grottina Covoloni del Broion, Italy	tooth
STheo _Pbdx	11th c.AD	Badia Polesine, Italy	petrous bone
Tb.sn1A_OPbsx	2nd- 3rd c.AD	Oderzo, Italy	petrous bone

Tb.sn1B_OPbdx	2nd- 3rd c.AD	Oderzo, Italy	petrous bone
Tb.sn1C_OPbdx	2nd- 3rd c.AD	Oderzo, Italy	petrous bone
Tb.sn1D_OPbdx	2nd- 3rd c.AD	Oderzo, Italy	petrous bone
O.1986_Pbdx	2nd- 3rd c.AD	Oderzo, Italy	petrous bone
O.1986_M2lowdx	2nd- 3rd c.AD	Oderzo, Italy	tooth
Tb.17_OM1lowdx	2nd- 3rd c.AD	Oderzo, Italy	tooth
Tb.sn3_OM2lowsx	2nd- 3rd c.AD	Oderzo, Italy	tooth
Al-K_12PB	pre-Mesolithic	Al-Khiday, Sudan	petrous bone
Al-K_12T	pre-Mesolithic	Al-Khiday, Sudan	tooth
Al-K_34PB	Neolithic	Al-Khiday, Sudan	petrous bone
Al-K_34T	Neolithic	Al-Khiday, Sudan	tooth
Al-K_35PB	pre-Mesolithic	Al-Khiday, Sudan	petrous bone
Al-K_35T	pre-Mesolithic	Al-Khiday, Sudan	tooth
Al-K_58PB	Meroitic	Al-Khiday, Sudan	petrous bone
Al-K_58T	Meroitic	Al-Khiday, Sudan	tooth
Al-K_95PB	Neolithic	Al-Khiday, Sudan	petrous bone
Al-K_95T	Neolithic	Al-Khiday, Sudan	tooth
Al-K_104PB	Neolithic	Al-Khiday, Sudan	petrous bone
Al-K_104T	Neolithic	Al-Khiday, Sudan	tooth
Al-K_163PB	Mesolithic (?)	Al-Khiday, Sudan	petrous bone
Al-K_163T	Mesolithic (?)	Al-Khiday, Sudan	tooth
Al-K_164PB	Mesolithic (?)	Al-Khiday, Sudan	petrous bone
Al-K_164T	Mesolithic (?)	Al-Khiday, Sudan	tooth
Al-K_185PB	pre-Mesolithic	Al-Khiday, Sudan	petrous bone
Al-K_185T	pre-Mesolithic	Al-Khiday, Sudan	tooth
Al-K_186PB	Meroitic	Al-Khiday, Sudan	petrous bone
Al-K_186T	Meroitic	Al-Khiday, Sudan	tooth
Al-K_187PB	Meroitic	Al-Khiday, Sudan	petrous bone
Al-K_187T	Meroitic	Al-Khiday, Sudan	tooth
3F-gr2	Neolithic	Al-Khiday, Sudan	femur

31F-gr3	Mesolithic (?)	Al-Khiday, Sudan	femur
50Fgr5	Meroitic	Al-Khiday, Sudan	femur
55gr2	pre-Mesolithic	Al-Khiday, Sudan	femur
103gr3	Neolithic	Al-Khiday, Sudan	femur
115F-gr4	Meroitic	Al-Khiday, Sudan	femur
128gr5	Meroitic	Al-Khiday, Sudan	femur
136gr5	Meroitic	Al-Khiday, Sudan	femur
156gr5	Meroitic	Al-Khiday, Sudan	femur
158gr4	Meroitic	Al-Khiday, Sudan	femur
159F-gr3	pre-Mesolithic	Al-Khiday, Sudan	femur
163gr5a	pre-Mesolithic	Al-Khiday, Sudan	femur
170gr5	pre-Mesolithic	Al-Khiday, Sudan	femur
177gr4	pre-Mesolithic	Al-Khiday, Sudan	femur
SPRM_Tb.1T	11th-12thc.AD	Sant'Anna Sopramonte, Trento, Italy	tooth
SPRM_Tb.3T	11th-12thc.AD	Sant'Anna Sopramonte, Trento, Italy	tooth
SPRM_US48PB	11th-12thc.AD	Sant'Anna Sopramonte, Trento, Italy	petrous bone
SPRM_US48T	11th-12thc.AD	Sant'Anna Sopramonte, Trento, Italy	petrous bone
SPRM_Tb.1T	11th-12thc.AD	Sant'Anna Sopramonte, Trento, Italy	petrous bone
TEDESCHI_OS1242PB	19th c.AD	Modern	petrous bone
TEDESCHI_OS419T	19th c.AD	Modern	tooth
TEDESCHI_OS426PB	19th c.AD	Modern	petrous bone
TEDESCHI_OS428PB	19th c.AD	Modern	petrous bone
TEDESCHI_OS428T	19th c.AD	Modern	tooth
TEDESCHI_OS431T	19th c.AD	Modern	tooth
TEDESCHI_OS1453T	19th c.AD	Modern	tooth
TEDESCHI_OS426T	19th c.AD	Modern	tooth
TEDESCHI_O419PB	19th c.AD	Modern	petrous bone
TEDESCHI_O431PB	19th c.AD	Modern	petrous bone
TEDESCHI_OS1453PB	19th c.AD	Modern	petrous bone
ORD001_PB	Iron Age	Ordona, Italy	petrous bone
ORD004_PB	Iron Age	Ordona, Italy	petrous bone
ORD006_PB	Iron Age	Ordona, Italy	petrous bone
ORD009_PB	Iron Age	Ordona, Italy	petrous bone
ORD010_PB	11thc.AD	Ordona, Italy	petrous bone
ORD011_PB	Iron Age	Ordona, Italy	petrous bone
ORD012_PB	Iron Age	Ordona, Italy	petrous bone
ORD014_PB	Iron Age	Ordona, Italy	petrous bone
ORD018_PB	Iron Age	Ordona, Italy	petrous bone
ORD019_PB	Iron Age	Ordona, Italy	petrous bone
SAL001_T	Iron Age	Salapia, Italy	tooth
SAL003_PB	Iron Age	Salapia, Italy	petrous bone

SAL005_PB	Iron Age	Salapia, Italy	petrous bone
SAL007_PB	Iron Age	Salapia, Italy	petrous bone
SAL008_T	Iron Age	Salapia, Italy	tooth
SAL010_PB	Iron Age	Salapia, Italy	petrous bone
SAL011_T	Iron Age	Salapia, Italy	tooth
SAL012_T	Iron Age	Salapia, Italy	tooth
SGR001_T	8thc.AD	San Giovanni Rotondo, Italy	tooth
SGR002_T	Iron Age	San Giovanni Rotondo, Italy	tooth
SGR003_T	Iron Age	San Giovanni Rotondo, Italy	tooth
M3 Human male 25yrs Dentine	21st c.AD	Fresh bone	tooth
Bone Dog 08	21st c.AD	Fresh bone	femur
Bone Horse 10	21st c.AD	Fresh bone	femur
Bone O.Sheep 11	21st c.AD	Fresh bone	femur
Bone Pig 10	21st c.AD	Fresh bone	femur
Dentin Young Sheep 06	21st c.AD	Fresh bone	tooth
M1 Dentine Old Sheep 05	21st c.AD	Fresh bone	tooth
M2 Dentine Old Sheep 06	21st c.AD	Fresh bone	tooth
M2 Modern Horse Cementum	21st c.AD	Fresh bone	tooth
oxbone.SPA	21st c.AD	Fresh bone	femur

REFERENCES

Adler, C.J., Haak, W., Donlon, D., Cooper, A., 2011. Survival and recovery of DNA from ancient teeth and bones. J. Archaeol. Sci. 38, 956–964. https://doi.org/10.1016/j.jas.2010.11.010

Analytical Methods Committee, 2006. Representing data distributions with kernel density estimates. AMC Tech. Br. 4, 2.

Asscher, Y., Regev, L., Weiner, S., Boaretto, E., 2012. Atomic disorder in Fossil TooThand bone mineral: An FTIR study using the grinding curve method. ArcheoSciences 35, 135–141. https://doi.org/10.4000/archeosciences.3062

Berna, F., Matthews, A., Weiner, S., 2004. Solubilities of bone mineral from archaeological sites: The recrystallization window. J. Archaeol. Sci. 31, 867–882. https://doi.org/10.1016/j.jas.2003.12.003

Carden, A., Morris, M.D., 2000. Application of vibrational spectroscopy to the study of mineralized tissues (review). J. Biomed. Opt. 5, 259. https://doi.org/10.1117/1.429994

Carrara, N., Scaggion, C., Holland, E., 2018. The Tedeschi collection: A collection of documented and undocumented human skeletal remains at the Museum of Anthropology, Padua University (Italy). Am. J. Phys. Anthropol. 166. https://doi.org/10.1002/ajpa.23471

Clarke, B., 2008. Normal bone anatomy and physiology. Clin. J. Am. Soc. Nephrol. 3 Suppl 3, 131–139. https://doi.org/10.2215/CJN.04151206

Collins, M.J., Nielsen-Marsh, C.M., Hiller, J., Smith, C.I., Roberts, J.P., Prigodich, R. V., Wess, T.J., Csapò, J., Millard, A.R., Turner-Walker, G., 2002. The survival of organic matter in bone: A review. Archaeometry 44, 383–394. https://doi.org/10.1111/1475-4754.t01-1-00071

Collins, M.J., Riley, M.S., Child, A.M., Turner-Walker, G., 1995. A Basic Mathematical Simulation of the Chemical Degradation of Ancient Collagen. J. Archaeol. Sci. 22, 175–183. https://doi.org/10.1006/jasc.1995.0019

Currey, J., 2008a. The structure and mechanical properties of bone, Bioceramics and their Clinical Applications. Woodhead Publishing Limited. https://doi.org/10.1533/9781845694227.1.3

Currey, J., 2008b. Collagen and the mechanical properties of bone and calcified cartilage. Collagen Struct. Mech. 397–420. https://doi.org/10.1007/978-0-387-73906-9_14

Dal Sasso, G., Angelini, I., Maritan, L., Artioli, G., 2018a. Raman hyperspectral imaging as an effective and highly informative tool to study the diagenetic alteration of fossil bones. Talanta 179, 167–176. https://doi.org/10.1016/j.talanta.2017.10.059

Dal Sasso, G., Asscher, Y., Angelini, I., Nodari, L., Artioli, G., 2018b. A universal curve of apatite crystallinity for the assessment of bone integrity and preservation. Sci. Rep. 8. https://doi.org/10.1038/s41598-018-30642-z

Dal Sasso, G., Lebon, M., Angelini, I., Maritan, L., Usai, D., Artioli, G., 2016. Bone diagenesis variability among multiple burial phases at Al Khiday (Sudan) investigated by ATR-FTIR spectroscopy. Palaeogeogr. Palaeoclimatol. Palaeoecol. 463, 168–179. https://doi.org/10.1016/j.palaeo.2016.10.005

Dal Sasso, G., Maritan, L., Usai, D., Angelini, I., Artioli, G., 2014. Bone diagenesis at the micro-scale: Bone alteration patterns during multiple burial phases at Al Khiday (Khartoum, Sudan) between the Early Holocene and the II century AD. Palaeogeogr. Palaeoclimatol. Palaeoecol. 416, 30–42. https://doi.org/10.1016/j.palaeo.2014.06.034

Derrick, M., Stulik, D., Landry, J., 1999. Infrared Spectroscopy in Conservation Science. Scientific Tools for Conservation. Los Angeles.

Fellows, A.P., Casford, M.T.L., Davies, P.B., 2020. Spectral Analysis and Deconvolution of the Amide I Band of Proteins Presenting with High-Frequency Noise and Baseline Shifts. Appl. Spectrosc. 74, 597–615. https://doi.org/10.1177/0003702819898536

Fernández-Jalvo, Y., Pesquero, M.D., Tormo, L., 2016. Now a bone, then calcite. Palaeogeogr. Palaeoclimatol. Palaeoecol. 444, 60–70. https://doi.org/10.1016/j.palaeo.2015.12.002

Fleet, M.E., 2009. Infrared spectra of carbonate apatites: ν2-Region bands. Biomaterials 30, 1473–1481. https://doi.org/10.1016/j.biomaterials.2008.12.007

France, C.A.M., Thomas, D.B., Doney, C.R., Madden, O., 2014. FT-Raman spectroscopy as a method for screening collagen diagenesis in bone. J. Archaeol. Sci. 42, 346–355. https://doi.org/10.1016/j.jas.2013.11.020

Garnero, P., 2015. The Role of Collagen Organization on the Properties of Bone. Calcif. Tissue Int. 97, 229–240. https://doi.org/10.1007/s00223-015-9996-2

Gonzalez, A., Cannet, C., Zvénigorosky, V., Geraut, A., Koch, G., Delabarde, T., Ludes, B., Raul, J.S., Keyser, C., 2020. The petrous bone: Ideal substrate in legal medicine? Forensic Sci. Int. Genet. 47, 102305. https://doi.org/10.1016/j.fsigen.2020.102305

Gueta, R., Natan, A., Addadi, L., Weiner, S., Refson, K., Kronik, L., 2007. Local atomic order and infrared spectra of biogenic calcite. Angew. Chemie - Int. Ed. 46, 291–294. https://doi.org/10.1002/anie.200603327

Hedges, R.E.M., 2002. Bone diagenesis: an overview of processes. Archaeometry 44, 319–328. https://doi.org/10.1111/1475-4754.00064

Hedges, R.E.M., Millard, A.R., 1995. Bones and Groundwater: Towards the Modelling of Diagenetic Processes. J. Archaeol. Sci. 22, 155–164. https://doi.org/10.1006/jasc.1995.0017

Iacumin, P., Di Matteo, A., Usai, D., Salvatori, S., Venturelli, G., 2016. Stable isotope study on ancient populations of central sudan: Insights on their diet and environment. Am. J. Phys.

Anthropol. 160, 498–518. https://doi.org/10.1002/ajpa.22987

Keenan, S.W., 2021. Fossilization, in: Alderton, D., Scott, A.E. (Eds.), Encyclopedia of Geology. Academic press, pp. 1–11.

Kendall, C., Eriksen, A.M.H., Kontopoulos, I., Collins, M.J., Turner-Walker, G., 2018. Diagenesis of archaeological bone and tooth. Palaeogeogr. Palaeoclimatol. Palaeoecol. 491, 21–37. https://doi.org/10.1016/j.palaeo.2017.11.041

Kontopoulos, I., Penkman, K., McAllister, G.D., Lynnerup, N., Damgaard, P.B., Hansen, H.B., Allentoft, M.E., Collins, M.J., 2019. Petrous bone diagenesis: a multi-analytical approach. Palaeogeogr. Palaeoclimatol. Palaeoecol. 518, 143–154. https://doi.org/10.1016/j.palaeo.2019.01.005

LeGeros, R.Z., 1981. Apatites in Biological Systems. Prog. Cryst. Growth Charact. 4, 1–45.

Mendes, L.C., Ribeiro, G.L., Marques, R.C., 2012. In Situ Hydroxyapatite Synthesis: Influence of Collagen on Its Structural and Morphological Characteristic. Mater. Sci. Appl. 03, 580–586. https://doi.org/10.4236/msa.2012.38083

Piga, G., Gonçalves, D., Thompson, T.J.U., Brunetti, A., Malgosa, A., Enzo, S., 2016. Understanding the Crystallinity Indices Behavior of Burned Bones and Teeth by ATR-IR and XRD in the Presence of Bioapatite Mixed with Other Phosphate and Carbonate Phases. Int. J. Spectrosc. 2016, 1–9. https://doi.org/10.1155/2016/4810149

Pinhasi, R., Fernandes, D., Sirak, K., Novak, M., Connell, S., Alpaslan-Roodenberg, S., Gerritsen, F., Moiseyev, V., Gromov, A., Raczky, P., Anders, A., Pietrusewsky, M., Rollefson, G., Jovanovic, M., Trinhhoang, H., Bar-Oz, G., Oxenham, M., Matsumura, H., Hofreiter, M., 2015. Optimal ancient DNA yields from the inner ear part of the human petrous bone. PLoS One 10, 1–13. https://doi.org/10.1371/journal.pone.0129102

Pollard, A.M., Batt, C.M., Stern, B., Young, S.M.., 2007. Analytical Chemistry in Archaeology. Cambridge University Press, Cambridge.

Poralan, G.M., Gambe, J.E., Alcantara, E.M., Vequizo, R.M., 2015. X-ray diffraction and infrared spectroscopy analyses on the crystallinity of engineered biological hydroxyapatite for medical application. IOP Conf. Ser. Mater. Sci. Eng. 79. https://doi.org/10.1088/1757-899X/79/1/012028

Psycharis, V., Kalamakis, N., Boukos, N., Trapalis, C., Bourlinos, A., Karakasides, M., 2001. Chemical and X-ray diffraction peak broadening analysis, electron microscopy and IR studies of biological apatites. Mater. Sci. Forum 378–381, 759–764. https://doi.org/10.4028/www.scientific.net/msf.378-381.759

Querido, W., Ailavajhala, R., Padalkar, M., Pleshko, N., 2018. Validated Approaches for

Quantification of Bone Mineral Crystallinity Using Transmission Fourier Transform Infrared (FT-IR), Attenuated Total Reflection (ATR) FT-IR, and Raman Spectroscopy. Appl. Spectrosc. 72, 1581–1593. https://doi.org/10.1177/0003702818789165

Reiche, I., Vignaud, C., Menu, M., 2002. The crystallinity of ancient bone and dentine: New insights by transmission electron microscopy. Archaeometry 44, 447–459. https://doi.org/10.1111/1475-4754.00077

Rey, C., Collins, B., Goehl, T., Dickson, I.R., Glimcher, M.J., 1989. The carbonate environment in bone mineral: A resolution-enhanced fourier transform infrared spectroscopy study. Calcif. Tissue Int. 45, 157–164. https://doi.org/10.1007/BF02556059

Rey, C., Combes, C., Drouet, C., Glimcher, M.J., 2009. Bone mineral: Update on chemical composition and structure. Osteoporos. Int. 20, 1013–1021. https://doi.org/10.1007/s00198-009-0860-y

Rubinson, K.A., Rubinson, J.F., Taddia M. (curated by), 2002. Chimica analitica strumentale. Zanichelli.

Shemesh, A., 1990. Crystallinity and diagenesis of sedimentary apatites. Geochim. Cosmochim. Acta 54, 2433–2438. https://doi.org/10.1016/0016-7037(90)90230-I

Silverman, B.W., 1986. Density Estimation for Statistics and Data Analysis. Monographs on Statistics and Applied Probability. Chapman and Hall, London, Great Britain.

Singh, B.R., DeOliveira, D.B., Fu, F.-N., Fuller, M.P., 1993. Fourier transform infrared analysis of amide III bands of proteins for the secondary sctructure estimation. Biomol. Spectrosc. III 1890, 47–55. https://doi.org/10.1117/12.145242

Sirak, K.A., Fernandes, D.M., Cheronet, O., Novak, M., Gamarra, B., Balassa, T., Bernert, Z., Cséki, A., Dani, J., Gallina, J.Z., Kocsis-Buruzs, G., Kővári, I., László, O., Pap, I., Patay, R., Petkes, Z., Szenthe, G., Szeniczey, T., Hajdu, T., Pinhasi, R., 2017. A minimally-invasive method for sampling human petrous bones from the cranial base for ancient DNA analysis. Biotechniques 62, 283–289. https://doi.org/10.2144/000114558

Smith, C.I., Nielsen-Marsh, C.M., Jans, M.M.E., Collins, M.J., 2007. Bone diagenesis in the European Holocene I: patterns and mechanisms. J. Archaeol. Sci. 34, 1485–1493. https://doi.org/10.1016/j.jas.2006.11.006

Sønju Clasen, A.B., Ruyter, I.E., 1997. Quantitative determination of type A and type B carbonate in human deciduous and permanent enamel by means of Fourier transform infrared spectrometry. Adv. Dent. Res. 11, 523–527. https://doi.org/10.1177/08959374970110042101

Stani, C., Vaccari, L., Mitri, E., Birarda, G., 2020. FTIR investigation of the secondary structure of type I collagen: New insight into the amide III band. Spectrochim. Acta - Part A Mol. Biomol.

Spectrosc. 229, 118006. https://doi.org/10.1016/j.saa.2019.118006

Stathopoulou, E.T., Psycharis, V., Chryssikos, G.D., Gionis, V., Theodorou, G., 2008. Bone diagenesis: New data from infrared spectroscopy and X-ray diffraction. Palaeogeogr. Palaeoclimatol. Palaeoecol. 266, 168–174. https://doi.org/10.1016/j.palaeo.2008.03.022

Thompson, T.J.U., Islam, M., Piduru, K., Marcel, A., 2011. An investigation into the internal and external variables acting on crystallinity index using Fourier Transform Infrared Spectroscopy on unaltered and burned bone. Palaeogeogr. Palaeoclimatol. Palaeoecol. 299, 168–174. https://doi.org/10.1016/j.palaeo.2010.10.044

Trueman, C.N., 2013. CHEMICAL TAPHONOMY OF BIOMINERALIZED TISSUES 56, 475–486. https://doi.org/10.1111/pala.12041

Trueman, C.N., Tuross, N., 2019. Trace elements in recent and fossil bone apatite. Phosphates Geochemical, Geobiol. Mater. Importance 48, 489–522. https://doi.org/10.2138/rmg.2002.48.13

Trueman, C.N.G., Behrensmeyer, A.K., Tuross, N., Weiner, S., 2004. Mineralogical and compositional changes in bones exposed on soil surfaces in Amboseli National Park, Kenya: Diagenetic mechanisms and the role of sediment pore fluids. J. Archaeol. Sci. 31, 721–739. https://doi.org/10.1016/j.jas.2003.11.003

Usai, D., Salvatori, S., 2019. The Mesolithic period on the White Nile region and the al-Khiday sites. Azania 54, 445–467. https://doi.org/10.1080/0067270X.2019.1691846

Usai, D., Salvatori, S., Jakob, T., David, R., 2014. The Al Khiday cemetery in central Sudan and its “classic/late Meroitic” period graves. J. African Archaeol. 12, 183–204. https://doi.org/10.3213/2191-5784-10254

Weiner, S., 2010. Microarchaeology. Beyond the visible archaeological record., first edit. ed. Cambridge University Press, UK.

Weiner, S., Bar-Yosef, O., 1990. States of preservation of bones from prehistoric sites in the Near East: A survey. J. Archaeol. Sci. 17, 187–196. https://doi.org/10.1016/0305-4403(90)90058-D

Wescott, D.J., 2019. Postmortem change in bone biomechanical properties: Loss of plasticity. Forensic Sci. Int. 300, 164–169. https://doi.org/10.1016/j.forsciint.2019.04.017

Wopenka, B., Pasteris, J.D., 2005. A mineralogical perspective on the apatite in bone. Mater. Sci. Eng. C 25, 131–143. https://doi.org/10.1016/j.msec.2005.01.008

CHAPTER 4

Infrared spectroscopy as a human bone pre-screening method for molecular and isotope analysis

Cinzia Scaggion[1*], Luca Nodari[2], Maurizio Marinato[3], Gregorio Dal Sasso[4], Manuel Rigo[1,4], Tina Saupe[5], Serena Aneli[6,7], Luca Pagani[6], Christiana L. Scheib[5], Nicola Carrara[8], Gilberto Artioli[1].

[1] Department of Geosciences, University of Padova, Padova, 35131, Italy
[2] Italian National Research Council-CNR, ICMATE, Padova, 35127, Italy
[3] Department of Cultural Heritage, University of Padua, 35139, Italy
[4] Institute of Geosciences and Earth resources, Italian National Research Council-CNR, Padova, 35131, Italy
[5] Estonian Biocentre, Institute of Genomics, University of Tartu, Tartu, 51010, Estonia
[6] Department of Biology, University of Padova, Padova, 35122, Italy
[7] Department of Public Health Sciences and Pediatrics, University of Turin, Turin, 10126, Italy
[8] Museum of Anthropology, University of Padova, Padova, 35121, Italy

*corresponding author, e-mail address: cinzia.scaggion@phd.unipd.it (C. Scaggion)

Keywords: bone preservation; FTIR spectroscopy; ancient DNA; pre-screening method; isotopic analysis

Abstract

Following the development of modern technologies, investigation of museum osteological finds is increasingly required. By the project we intend to develop viable protocols in order to preserve these collections from exceedingly invasive analyzes, allowing access to the specimens for scientific research. The main aim of this work is to survey skeletal tissues, specifically petrous bones and root of teeth using Infrared Spectroscopy as a pre-screening method to assess the bone quality for molecular analyses. This approach could overcome a major problem of identifying useful genetic material in archaeological bone collections without resorting to demanding and expensive laboratory studies. A minimally invasive sampling of archaeological bone finds was developed and bone structural and compositional changes were examined, linking isotopic and genetic data to Infrared spectra. The predictive model proposed here is functional with the infrared parameters used but the quality/quantity of aDNA cannot be determined because of the influence of environmental local factors.

1. INTRODUCTION

Often osteoarchaeological remains are physical evidence of human or animal presence, containing valuable biological information about the past. In particular, the dense bone parts of the petrous bone and the dental cementum of tooth roots can provide high concentration of endogenous DNA (aDNA) (Gamba et al., 2014; Hansen et al., 2017; Parker et al., 2020) and, because of their unique informational value, advantageous for forensic/medical sciences, exhibition purposes, isotopic and evolutionary studies, are the most widely used sample source. Bones and teeth are calcified tissue characterized by a complex hierarchical structure from the macro to the nano-scale and consist mainly of mineral and organic phases (Fratzl et al., 2004; Reznikov et al., 2018). The inorganic component bioapatite (BAp) is a nanocrystalline and defective carbonated-hydroxylapatite with a chemical composition significantly departing from the stoichiometric hydroxylapatite $Ca_5(PO_4)_3(OH)$ and constitutes about 60% of bone. The 30% of bone material is constituted by organic components as

collagen, protein-mucopolysaccharide complexes and glycoproteins and collagen, whereas the remaining 10% is bone structural water (Dey, 2020). Collagen fibrils (the majority is Type I collagen which accounts ~ 90% of the organic fraction of a bone) is the major fibrous protein which gives elasticity and strength to the bone system (Shoulders and Raines, 2009) and it is the primary component for osteological analysis (radiocarbon, stable isotopes, genetics etc.). Deoxyribonucleic acid (DNA) in living bone is located within several cells type as preosteoblast, osteoblast, osteoclast and osteocyte (Florencio-Silva et al., 2015). Osteocytes, typical of mature bone, represent 90% of all bone cells and are enclosed in a niche carved into the intracellular bone substance, called *lacunae* (Florencio-Silva et al., 2015). Tooth cementum is rich in cementocytes, which are DNA-containing cells that remain encased in the mineral structure of the tooth after death (Bosshardt and Selvig, 1997). *Post-mortem* alterations lead to the chemical degradation for hydrolysis of the collagen with a consequent increase of bone micro-porosity and bioapatite recrystallization due to extreme temperature and environmental pH, humic acids, and microbial attack that alter collagen-apatite hierarchical structure (Collins et al., 2002; Hedges, 2002; Turner-Walker and Syversen, 2002). Diagenetic changes alter the bone microscopic structure provoking chemical alterations, favorable for the DNA survival (Allentoft et al., 2012; Rollo et al., 2002) while progressively reducing the content of organic components in ancient samples (> 1%), making the analysis on a genomic scale impossible and/or very expensive (Der Sarkissian et al., 2015). Hence, at ultrastructural level, the DNA can establish strong bonds with both the organic and the inorganic components, so that adsorption of DNA to bioapatite crystals surface as well as the linking with Type I collagen may stabilize it and may determine its outliving over time (Campos et al., 2012; Collins et al., 1995; Sosa et al., 2013). Genetic molecules can strongly bound to both organic and inorganic bone phases, thus an extraction method to recover double helix molecules is applied to both fractions, with greater chance of success, employing EDTA (to demineralized the BAp) and Proteinase K (to digest the osseous proteins) (Rohland and Hofreiter, 2007). It is assumed that bones found in regions characterized by high temperatures yield less aDNA than bones from temperate climate regions, whereas higher yields are retrieved from finds buried in caves. However, this assumption is not always true: the endogenous DNA quality/quantity recovered from apparently similar sites are often different, as highlighted in genetic research (Aneli et al., 2022; Saupe et al., 2021).

Sampling methods for ancient DNA extraction are predominantly destructive causing the loss of the entire root, the drilling and cutting of the *cochlea* (located in the petrous portion of the temporal bone), for powdering or coring the bone findings (Mays et al., 2013). Some bones survive well, while others degrade rapidly; and since it is difficult to evaluate a *priori* the preservation of organic material within a bone, so far expensive and often destructive analyses have been the norm to assess the

suitability of a given sample to be included in molecular studies (Orlando et al., 2021). In this paper, we intend to analyse skeletal tissues, specifically the compact part of petrous bones and tooth roots as a pre-screening method to assess the bone suitability for specific analyses, i.e. ancient DNA, palaeoproteomic, and stable isotope analysis.

This approach can overcome the major problem of identifying suitable materials in archaeological bone collections. Tooth and petrous bone sampling can be destructive, in particular when only one or a few of them are preserved. Moreover, invasive sampling may cause the decrease of the exhibitional value and may hamper morphological studies that provide important information about the age estimate of an individual for forensic and anthropologic cases, but also the diet, cultural habits and evolution (Balzeau, 2015; Brace et al., 1991; Demes and Creel, 1988; Doden and Halves, 1984; Oxilia et al., 2018; Rathmann and Reyes-Centeno, 2020). Thus, before irreversibly damaging of precious materials it is important that the sample choice is based on solid evidence about the positive outcome of a planned analyses, so that the pros and cons can be evaluated by interacting with museum curators, conservators and archaeologists. Recent methodological advances in several research fields have increased the request for sampling from museum collections, and raised ethical concerns over the destruction of human remains (Mays et al., 2013; Squires et al., 2019). Currently, several strategies have been adopted to protect the long-term integrity of collections devising minimally invasive sampling protocols (Brown et al., 2016; Harney et al., 2021; Sirak et al., 2020). The development of a molecular pre-screening method through the direct evaluation of the degree of *post-mortem* recrystallization of the bioapatite and correlating it with the quality/quantity of collagen could allow a more careful selection of well-preserved bone samples, enabling to reduce the costs of genetic analysis. The micro-sampling of osteological findings, characterized by different chronology and origin allowed us to investigate the sample variability in terms of preservation state. Comparing the bone integrity with the data obtained from the genetic analysis is an essential step to provide a precise evaluation of the degree of sample conservation, aiming at the extraction of sequentiable endogenous aDNA, applicable to all osteoarchaeological findings. Various methods have been applied to the pre-screening of bone/collagen preservation, but these are often destructive, expensive, and/or require extraction of large quantities of bone material for analysis (Collins et al., 2009; Gotherstrom et al., 2002; Haynes et al., 2002; Hedges, 1992; Kirchner et al., 1997; Schwarz et al., 2009; Weiner and Bar-Yosef, 1990).

Recently, Fourier-transform Infrared spectroscopy (FTIR spectroscopy) has proven to be an advantageous technique since it provides a fast, convenient, and accurate semi-quantitative approach to investigate mineralized tissues. Moreover, FTIR has been successfully applied to in-depth studied on the alteration of how teeth and bones that undergo diagenetic processes (Dal Sasso et al., 2016;

Hedges, 2002; Kendall et al., 2018; Kontopoulos et al., 2019; Lebon et al., 2014). Infrared absorption bands are characteristic of specific types of chemical bonds, enabling the descriptions of variations on the structural properties of organic and inorganic components (Carden and Morris, 2000) assessing bone quality and exploring sample variability. In this work, we intend to propose a pre-screening method investigating the relationship between the broadening of the FTIR phosphate peak (Dal Sasso et al., 2018) with the amount of collagen contained in the bone. As it is closely related to atomic order/disorder in the crystal lattice, widening of phosphate peak could be a useful new tool for predicting the absence/presence of DNA through the evaluation of minimal changes in the bioapatite chemical composition and structure. Since the survival of DNA is bounded by multiple intercorrelations, the study is integrated with the examination of microstructural modifications induced by diagenesis on the secondary structure of the preserved collagen (Chadefaux et al., 2009; Fredericks et al., 2012; Leskovar et al., 2020). The organic preservation was investigated by analyzing the infrared spectra of lyophilized collagen to determine its integrity, examining in detail its characteristic FTIR peaks (Doyle et al., 1975; Ishida and Griffiths, 1993; Vyskočilová et al., 2019). Results from FTIR analysis were compared with aDNA yields and stable isotopic data obtained by validating the collagen quality and the potential of genetic extraction. This research idea aims at improving the selection process of bone samples for different fields of interest, such as stable isotope and radiocarbon analyses by using predictive parameters based on micro-invasive FTIR spectroscopy.

2. MATERIALS AND METHODS

2.1. Bone specimens

In this study, samples were taken from human skeletal remains covering different chronologies and presenting different states of preservation from the Anthropology Museum of Padua University, Italy, in particular petrous bones (n=19) and tooth roots (n=22) and archaeological excavation of the early Middle Ages (n=3). The analyzed bones came from graves in Northern and Southern Italy (Covoloni del Broion, Salapia, Ordona, San Giovanni Rotondo and Desenzano del Garda). Covoloni del Broion cave stratigraphy, highlighted by systematic excavations, was made up of thin silty levels of terrain interspersed in the stalagmite concretions -incorporating human skeletal remains of Bronze Age (Saupe et al., 2021), followed by a series of silty levels with abundant crushed stone partially concreted (Ligabue, 1973). Dauna necropolis of Salapia (Corrain et al., 1972) and Ordona are located at 10 and 20 km from contemporary Cerignola (Foggia, Italy), respectively, while there are no archaeological record on nearby San Giovanni Rotondo. The necropolis of Ordona was the subject of two archaeological campaigns in 1978 and 1981 (Corrain, 1986) and is characterized by a chronological continuity in the burials that includes the Iron Age up to the Roman Empire. After a period of neglect, the use of the necropolis is resumed in the Middle Ages. Despite the lack of

archaeological information, recent studies have confirmed the attribution to the Iron Age of most of the graves (Aneli et al., 2022). Three bone elements of early medieval origin come from the site of the Church of San Lorenzo in Desenzano del Garda (Brescia, Italy) in agricultural territory (more information can be found in (Canci et al., 2012). Ancient samples are compared with animal fresh bone. The list of bone samples sequenced and analyzed by FTIR spectroscopy and Mass spectrometry is reported in **Table.1** (reported at the end of Chapter).

2.2. Micro-sampling for Infrared analysis

The external of all bone and root surface – encrusted by soil sediments – was mechanically removed by means of low-speed micro-drill, equipped with an abrasive round point of 2.4 mm using Dremel micro-tool. The powdered material from the tooth roots and petrous bones was extracted with a round engraver bit (2.4 mm diameter) after removal of the external deposits, while in other occasion for genetic analysis a complete root was chosen and cut with a cutting disc wheel (diameter 20 mm). Drill bits used for sampling operations were sterilized to avoid external contamination, operating the procedure for standard sampling of bone of the ICMP (International Commission on Missing Persons - ICMP 2015) was followed.

2.3. Sample preparation and IR spectra collected

Enough bone material was extracted from the petrous bones and tooth roots (20-30 mg) in order to perform the FTIR analysis. Pellets of bone powder and lyophilized collagen have been prepared maintaining a ratio 1:100 mg of sample/KBr with constant grinding for 2 minutes. Powder obtained was then pressed, using a hydraulic press, under 11 tons/cm^2 pressure, and a 12 mm in diameter and 1.5 mm of thickness transparent pellet was obtained. Spectrum was collected with a Nicolet 380 FTIR spectrometer equipped with a DTGS detector; 128 scans were acquired, in the range from 4000 to 400 cm^{-1}, with a spectral resolution of 4 cm^{-1}. Spectral analysis was performed using Omnic 9 software (Thermo Scientific).

2.3.1. Bone spectra analyses

In spectral analysis, the peak heights are inferred from baselines defined by two points calculated as local minimum ranging in selected regions of spectrum. In spectral analysis, the peak heights of Amide I and ν_3, ν_4 phosphate were selected and inferred from baselines defined by two points calculated as local minimum ranging in selected regions of spectrum 2000–1800/1400–1200 cm^{-1}, 1400–1200/900–750, respectively. The ν_4 (PO_4^{3-}) vibrational band shows two separate peaks at 604 and 565 cm^{-1} related to the atomic disorder and/or the crystalline size of the bioapatite (Dal Sasso et al., 2016). The height percentage used to calculate the width of peak at 604 cm^{-1} at the 85% (FW85%) was considered for each spectrum in order to monitor the physical-chemical variations of bioapatite

(Dal Sasso et al., 2018). To define **FW85%** parameter, the baseline has been defined by two points in the region 850-620 and 510-470 cm^{-1}. FW85% is inversely related to the Splitting Factor (IRSF) that quantifies the extent of splitting of the two peaks calculated with the sum of two peak intensities of ν_4 (PO_4^{3-}) divided by the intensity of the valley between them (Weiner and Bar-Yosef, 1990). IRSF is widely used in literature to monitor the degree of recrystallization of the bioapatite as higher values of IRSF show a more ordered crystal structure with an increase in crystal size and vice versa, determined by comparing IRSF values with other analytical techniques (Psycharis et al., 2001; Stathopoulou et al., 2008). Mineral-to-matrix ratios were calculated by Amide I (1660 cm^{-1}) divided by main phosphate peak intensity at 1035 cm^{-1} (ν_3 (PO_4^{3-})) (**AmI/PO_4^{3-}**). For a more detailed description, see **Figure.1**, **Table.2 and Table.3**.

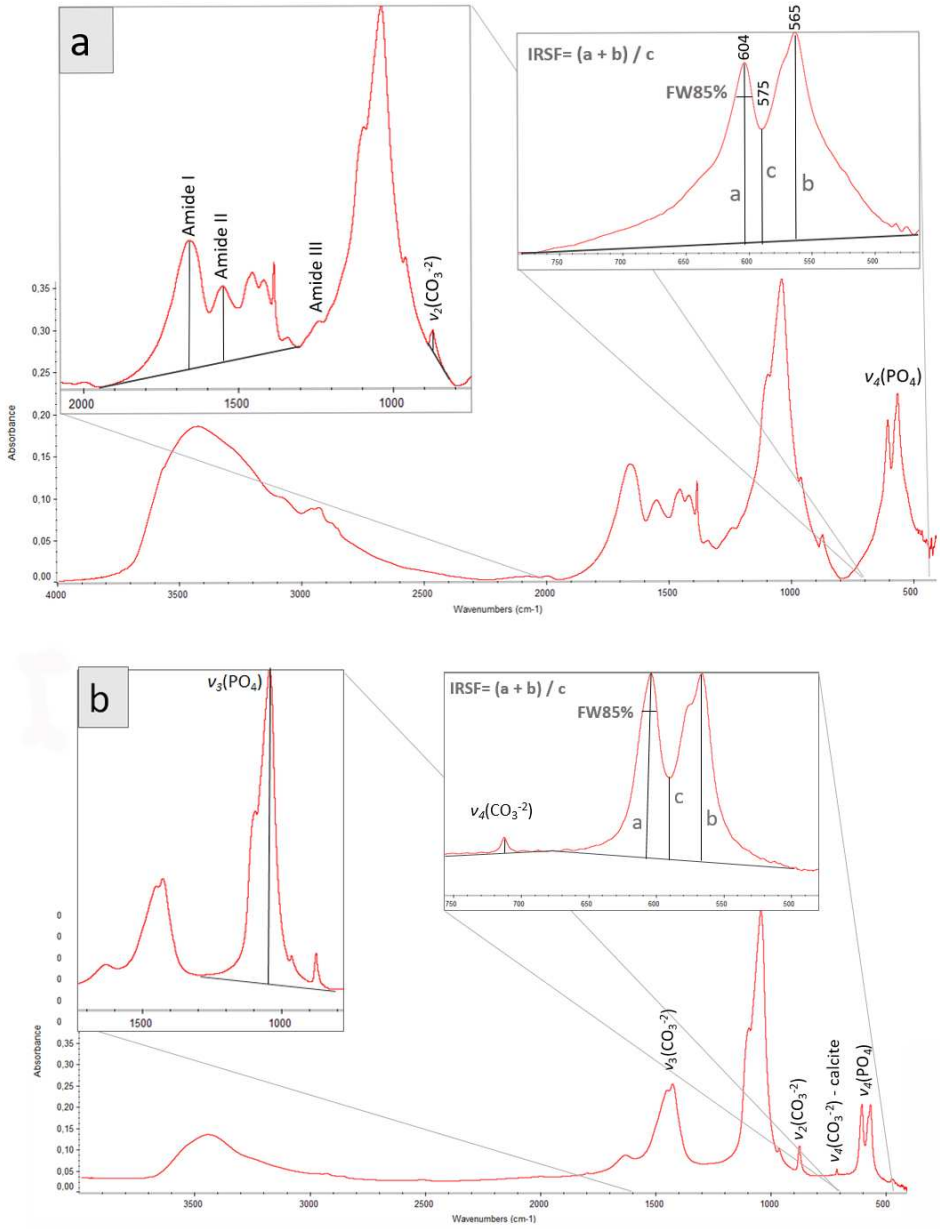

Figure 1. *Spectrum of the bone samples shows the characteristic band distribution that can be attributable to the amide, phosphates and carbonate groups.*

Table 2. *Wavenumbers and baselines adopted to calculate peaks intensity and width of tooth and bone sample properties by FTIR spectroscopy.*

Vibrational mode and functional group	Wavenumber (cm^{-1})	Baseline (cm^{-1})
AmideI ν (C=O) stretch	1660	2000–1800/1400–1200
$\nu_3(PO_4^{3-})$ antisymmetric stretching	1035	1400–1200/900–750
$\nu_4(PO_4^{3-})$ bend	604	850–620/510–470

Table 3. *List of parameters used.*

Parameters	Wavenumber (cm^{-1})	Characterization
FW85%	width at 85% of the height of the 604 cm^{-1} peak	atomic order/disorder
AmI/PO_4^{3-}	1660 cm^{-1}/1035 cm^{-1}	amount of amide on phosphate
1660:1690 cm^{-1} ratio	1660 cm^{-1}/1690 cm^{-1}	collagen quality

2.3.2. Lyophilized bone collagen spectra analyses

Collagen cross-links provide the fibrillary collagen matrices with properties such as tensile strength and viscoelasticity as described in (Paschalis et al., 2001); these properties can be modified by diseases but also by interaction with the burial environment (Chadefaux et al., 2009). The ratio of peak areas of the mature non-reducible (interfibrillar) cross-links and immature reducible (intrafibrillar) cross-links of Amide I sub-underlying of ~1660 and ~1690 cm^{-1} (**1660:1690**) provides a semi-quantitative measure of the enzymatic cross-linking in the organic matrix. According to the literature (Paschalis et al., 2001), the degree of cross-linking of collagen from immature to mature has been calculated from 1660:1690 cm^{-1} area ratio which informes on the collagen maturity. The majority of scientific community deems it is a valid parameter (Barth et al., 2011; Chadefaux et al., 2009; De Campos Vidal and Mello, 2011; Mieczkowska et al., 2015) that represents collagen maturity although its validation is still controversial, above all for what concerns the overlap structural water absorption interference (Farlay et al., 2011; Stani et al., 2020) at ~ 1645 cm^{-1} (H-O-H bending). For our study, this parameter, which can be considered an indicator of the collagen alteration degree (Barth et al., 2011), was estimated by employing a combination of second-derivative spectra and peak finding to determine underlying peaks of Amide I region at ~1660 and ~1690 cm^{-1} (Paschalis et al., 2011, 2004, 2003). The sub-underlying area ratio (1660:1690) was calculated for all collagen samples.

The presence of sub-underlying signals of Amide I to calculate the denaturation state of collagen have been estimated by employing of second-derivative analysis (calculated using Savitsky-Golay smoothing technique with 9 smoothing points) (Savitzky and Golay, 1964). OMNIC® software was used to convert the FTIR absorbance spectra into second-derivatives to determine the number and position of components corresponding to the Amide I region for the next curve-fitting process of

collagen spectra. For the average collagen spectrum, eight sub-components at 1737-1722-1703-1694-1660-1643-1632-1612 cm^{-1} were found. These sub-bands were chosen according to known peak-position from literature (Arrondo et al., 1993; Ganim et al., 2008; Schmidt et al., 2017; Titus et al., 2018) (**Figure.2; Table.3**).

In order to start the iteration of the curve fit, MagicPlot software was chosen for nonlinear fit, plotting and data analysis. For each collagen spectrum, a baseline, the region between 1800-1300 cm^{-1} of the selected spectrum and a Gaussian function were used. Sub-components position were selected setting the peak width at 16 cm^{-1} (4× the sensitivity of the tool) measuring principal areas of ~ 1660 and ~ 1690 cm^{-1}.

2.4. Bone collagen extraction and isotopic analysis

500 mg was extracted from 17 archaeological sub-samples and one fresh animal bone previously selected (**Table.1**). The samples preparation was carried out in the Mass Spectrometry lab of stable isotopes of the Department of Geoscience (University of Padova) using the standard laboratory protocol based upon Privat et al. (2002) and Richards and Hedges, (1999). Bone fragments were demineralized in c. 10 ml 0.5M aq. HCl and left in the fridge (4°C) for several days. The samples were then gelatinized in pH 3 water at 75°C in the oven for 48 hours. Finally, the collagen was lyophilised before being weighed for isotopic analysis. Isotope analyses were performed using a Thermo Scientific Delta V Advantage Isotope Ratio Mass Spectrometer in continuous flow mode coupled to a Flash 2000 Elemental Analyzer and a ConFlo IV interface. The isotopic composition of carbon ($\delta^{13}C$) and nitrogen ($\delta^{15}N$) was calibrated respectively on the VPDB and AIR scales, using international standards (CH-6 and CH-7 for carbon, N-1 and N-2 for nitrogen and both UREA).

2.5. Ancient DNA extraction

Minimally destructive sampling of ancient human remains was applied for 41 bone samples coming from Grottina Covoloni del Broion, Salapia, San Giovanni Rotondo and Ordona (**Table.1**) using protocols that will allow for different types of biomolecular or chemical analysis. The first layer of petrous bone was removed with a sterilized drill bite to avoid exogenous contamination and a core of 10 mm was extracted from the *cochlea*. The drill bits and core drill were sterilized in between samples with 6% (w/v) bleach followed by distilled water and then ethanol rinse, while tooth roots were removed with a sterile drill wheel. The root and the petrous portions were soaked in 6% (w/v) bleach for 5 min. Samples were rinsed three times with 18.2 MXcm H2O and soaked in 70% (v/v) Ethanol for 2 min. The tubes were shaken during the procedure to dislodge particles. One the sample has been weighed, 20× ml of ETDA for sample mass and 0.5× of Proteinase K for sample mass were added into PCR-clean 5 ml or 15 ml conical tubes (Eppendorf) along with the samples inside the IIB hood

and the tubes were incubated 72 hours on a slow shaker at room temperature. The DNA extracts from bone samples were concentrated to 250 µl using the Vivaspin Turbo 15 (Sartorius) and purified in large volume columns (High Pure Viral Nucleic Acid Large Volume Kit, Roche) using 2.5 ml of PB buffer, 1 ml of PE buffer, and 100 µl of EB buffer (MinElute PCR Purification Kit, QIAGEN).

For the elution of the endogenous DNA, the silica columns were transferred to a collection tube to dry and followed in 1.5 ml DNA lo-bind tubes (Eppendorf) to elute. The samples were incubated with 100 µl EB buffer at 37 °C for 10 min and centrifuged at 13,000 rpm for 2 minutes. After centrifugation, the silica columns were removed and the samples were stored at -20 °C. Only one extraction was performed per extraction for screening and 30 ml used for libraries. Laboratory extraction was performed in dedicated ancient DNA laboratories at the Estonian Biocentre, Institute of Genomics, University of Tartu, Tartu, Estonia. The library quantification and sequencing were performed at the Core Facility of the Institute of Genomics, Tartu, Estonia. Complete procedures and DNA analyses are detailed in Aneli et al., 2022 and Saupe et al., 2021.

2.6. Statistical analysis

Linear regression (function *lm*) in R free statistical software (4.1.2 version, available at www.r-project.org) allows to model a continuous variable Y as a mathematical function of one or more X variables, so that it can be used the regression model with X known to predict the Y variable (Lane, 2003).

$$Y = \beta_1 + \beta_2 X + \epsilon$$

(linear model formula)

β_1 is the intercept and β_2 is the slope and are called regression coefficients, ϵ is the error term, Y is the regression model not predictive.

The contribution of AmI/PO_4^{3-} and FW85% parameters measured for each bone IR spectra were simultaneously combined with endogenous aDNA yields to define a predictive model by establishing which parameter can be predictive. The data were integrated with other 96 infrared measures reported in **Chapter 3** of the thesis. An endogenous DNA yield of 50% and 100% was attributed to the samples for which no genetic extraction was carried out such as modern bones from the Tedeschi collection (Carrara et al., 2018) and fresh bones of animals of several taxa (Dal Sasso et al., 2018). A null value was attributed to highly diagenized osteological samples from the archaeological site of Al-Khiday (Central Sudan) (Usai et al., 2014, 2010) for which aDNA extraction was reported to have failed in previous attempts.

3. RESULTS

3.1. Measurements of the infrared bone spectrum

Bone is a composite material made up of a mineral phase (carbonate-hydroxylapatite), collagenous phase, non-collagenous proteins, lipids, and water (Boskey, 2001). IR absorption bands are characteristic of specific types of chemical bonds, enabling descriptions of variation of structural properties of organic and inorganic components (Carden and Morris, 2000; Goormaghtigh et al., 2006). The spectrum of the bone samples shows the characteristic band distribution that can be attributable to the amide, phosphates and carbonate groups (**Figure.1**)

The relative amounts of these constituents are variable due to conservation status. In Infrared spectra, water has strong absorbance and normally constitutes an interference factor for IR observations. Chemical bonds of molecular framework radiated by infrared radiation undergo specific vibrations with a variety of motions at specific wavenumber. All bone FTIR spectra display a broad band in the range 3500-2500 cm^{-1} generated by ν_1 and ν_3 stretching modes of hydrogen-bonded (H–O–H bonds) water molecules and broad resonance doublet of hydrogen bond at ~ 3310-3270 cm^{-1} and at ~ 3100-3030 cm^{-1} related to Amide A and Amide B (N–H stretching), proteins insensitive to the conformation of the polypeptide backbone (Barth, 2007).

Vibrational modes are evident in the region between 2961-2925 cm^{-1} and between 2875-2850 cm^{-1}, corresponding to the organic compounds of the bone system, attributed to asymmetric stretching C–H_3 and C–H_2 bonds in the aliphatic chains of collagen, respectively (Barth, 2007).

Moving the focus in the fingerprint region (1800-400 cm-^{1}), specific collagen absorption bands were identified and related to C=O stretching vibrations, N–H bonding and to the coupling of the C–N stretch to the N–H bend, corresponding to Amide I (1660 cm^{-1}), Amide II (1550 cm^{-1}) and Amide III (at 1340-1240 cm^{-1}), respectively. It is good to point out that in Amide I region the superimposed contribution of ν_2 (O–H) at 1645 cm^{-1} bending mode of structural water occurs (Socrates, 2004).

The osteological bone powders show the $\nu_3(CO_3^{2-})$ asymmetric stretching peak at 1452-1425 cm^{-1} and the $\nu_2(CO_3^{2-})$ out of plane bend vibration, at 875 cm^{-1} (Fleet, 2009). For all spectra, the contribution of the $\nu_3(PO_4^{3-})$ band at 1090 cm^{-1}, due to the asymmetric vibration of the phosphate ions, appears as a shoulder, more or less evident indicating a less ordered bioapatite in the crystal lattice (Dal Sasso et al., 2018; Pleshko et al., 1992). Absorption bands at 1035, 604 and 565 cm^{-1} correspond to the asymmetric stretching and bending (ν_3, ν_4) of phosphates, while a weak band at 960 cm^{-1} is attributed to the $\nu_1(PO_4^{3-})$ symmetric stretching mode. About half of the samples analysed show the $\nu_4(CO_3^{2-})$ peak of in-plane bending vibrational mode at 712 cm^{-1} which is characteristic of secondary calcite.

Petrous bones infrared spectra of Covoloni del Broion samples (BR-01; BR-03; BR-04; BR-05) show a particular band of rather intense absorption at 1600 cm^{-1}, followed by a decrease in intensity in the region of Amide I and Amide II. This strong absorption, not found in the other osteoarchaeological

spectra, could be linked to the double bonds of aromatic compounds (C=C bonding) (Nandiyanto et al., 2019) probably due to sample preparation.

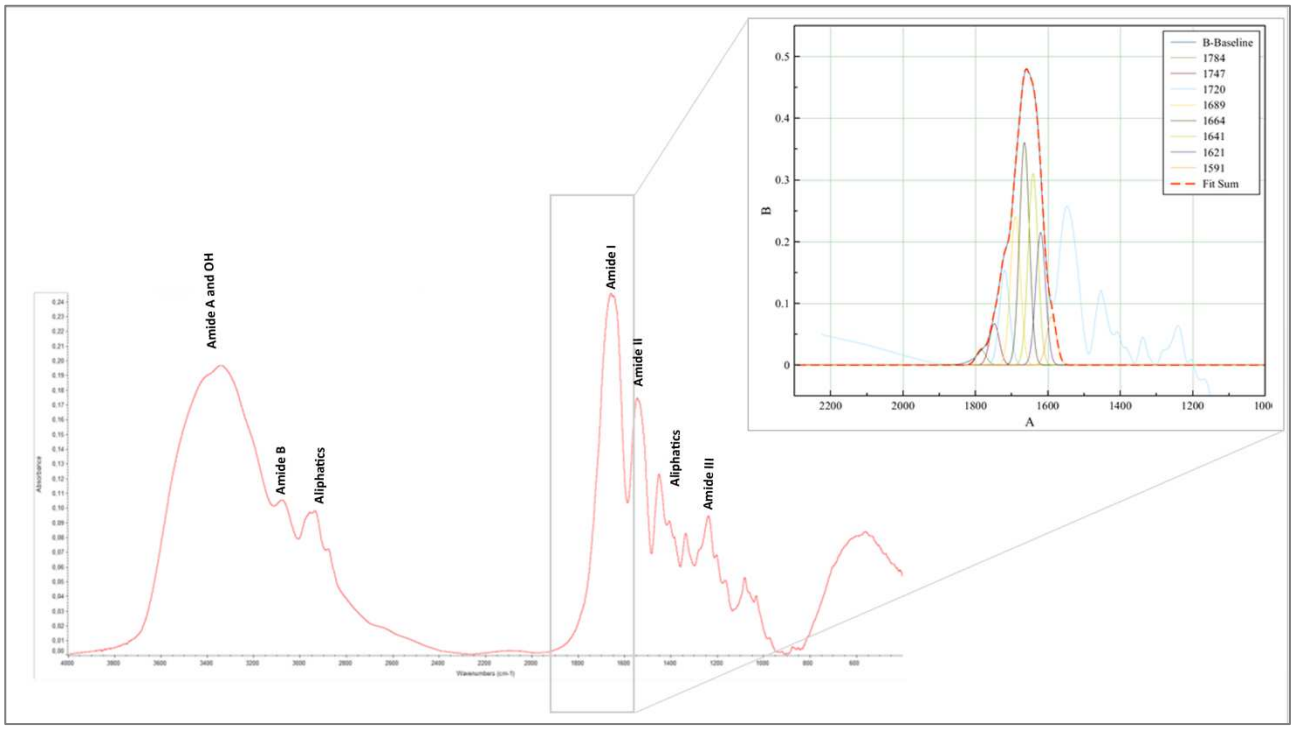

Figure 2. *Characteristics band distribution of collagen in IR spectra.*

3.2. FTIR measurement of bone collagen and quality control (QC)

Spectrum shows the characteristic band distribution of collagen in the regions 3500-2800 cm^{-1} and 1800-800 cm^{-1} (**Figure.2**).

All spectra presented absorptions at 3500-3300 cm^{-1} related to Amide A (N–H str.) and O–H vibrations, at 3079 cm^{-1} assigned to Amide B (N–H str.) while those at ~ 2970 and 2935 cm^{-1} are assigned to long-chain linear aliphatic compounds (CH str. and CH_3 str.) (Martínez Cortizas and López-Costas, 2020; Nandiyanto et al., 2019). Characteristic absorption bands of Amide I (1660 cm^{-1}), Amide II (1550 cm^{-1}), 1450-1407-1385-1338 cm^{-1} aliphatic compounds and aromatic at 1280 cm^{-1}, Amide III (1238 cm^{-1}) and carbohydrates from 1200 cm^{-1} to 874 cm^{-1} were found (Barth, 2007; Figueiredo et al., 2012; Nandiyanto et al., 2019; Stani et al., 2020). The sample DSL17-D_coll shows a strong vibrational mode at 3464 cm^{-1} (–OH groups str.) while other samples show a shift in the Amide I region from 1660 to 1645 cm^{-1}, which is evident in DLS25-dent_coll and DSL17-D_coll. All collagen spectra of archaeological bones show weakly defined shoulder at 1720 cm^{-1}, whereas this is not present in the collagen extracted from fresh animal bone (Oxbone_coll). Isotope data were collected for all collagen samples analysed by FTIR. Stable carbon ($\delta^{13}C$) and nitrogen ($\delta^{15}N$) isotopic analyses of collagen, preserved in subfossil and archaeological bone and tooth, is a powerful tool for environmental and palaeodiet reconstruction (Harrison and Katzenberg, 2003) and there is a growing

interest for protein available in hard tissue archives. The quality of measurements of carbon and nitrogen isotopes in ancient samples is confirmed by a robust protocol for the detection of contaminants and diagenesis (Talamo et al., 2021).

The most widely quality control (QC) criterion utilized is the atomic ratio (C:N), for which ancient samples have an acceptable quality when ranging between 2.9 and 3.6 (Ambrose, 1990). According to Guiry and Szpak (2020) this QC method, also widely used on modern samples, is not directly transferable to analyses of fresh collagen, thus establishing a new range for mammals modern collagen between 3.00 and 3.28. Covoloni del Broion, Ordona, Salapia and San Giovanni Rotondo samples provided values falling approximately within the range of well-preserved collagen, estimated by Ambrose (1990) while samples coming from Desenzano del Garda (signed DSL) shows an off-scale C:N ratio, resulting in bad preservation which invalidated the analyses.

The area ratio of Amide I underlying bands, 1660:1690 cm^{-1} (R^2= 0.99) ranges from 1.57 to 1.51, 1.58 to 1.36, 1.65 to 1.49 for Grottina Covolini del Broion, Ordona and Salapia, respectively while from 1.99 to 1.19 for fresh animal bone, San Giovanni Rotondo and Desenzano del Garda, respectively. All data are reported in **Table.1**

3.3 Infrared data samples and DNA yields

Mineral to matrix (AmI/PO_4^{3-}) indicator for determining the amount of preserved collagen (Naito et al., 2020) has been measured obtaining high variation among specimens with values ranging from 0.11 to 0.40 for Grottina Covoloni del Broion samples and 0.05 to 0.19, 0.07 to 0.40 and 0.10 to 0.27 for Ordona, Salapia and San Giovanni Rotondo, respectively. AmI/PO_4^{3-} values range from 0.08 to 0.33 and 0.39 for Desenzano samples and ox bone, respectively.

AmII/PO_4^{3-} parameter shows values ranging from 0.10 to 0.25, 0.06 to 0.17, 0.04 to 0.14 and from 0.07 to 0.16 for Covoloni del Broion, Ordona, Salapia and San Giovanni Rotondo, respectively, while values from 0.07 to 0.20 and 0.24 were obtained for DLS and ox bone (**Table.1**).

The results in **Table.1** show a few variations of FW85% values for the less crystalline bioapatite, ranging from 8.67 to 12.51 cm-1 and the corresponding minimal variation of IRSF values, ranging from 5.13 to 3.01. Dividing the samples by origin, variations that are more significant are evident in Desenzano samples with values from 9.40 to 11.21 of FW85% with IRSF from 4.25 to 3.14 and for S.Giovanni Rotondo with FW85% from 9.28 to 11.44 and IRSF of 3.81 to 3.27. The set of cave samples are divided into two sub-sets with variations in phosphate peak width ranging from 9.40 to 10.46 and IRSF between 3.76 and 3.41 while the other sub-set comprises values ranging from 11.20 to 12.25 and 3.53 and 2.99 for FW85% and IRSF, respectively. A high variability in the yield of ancient DNA (%) is instead shown in all sequenced samples. For Broion samples, endogenous DNA yields range from 0.07% to 17.32%, while more significant amounts were found in the Oderzo

samples between 0.69% and 43.75%. High yields of sequencable endogenous DNA were not obtained for Salapia and S. Giovanni Rotondo, providing aDNA content ranging from 0.15 to 4.71% and 3.72 and 6.39%, respectively. Of these samples, only SGR002 (SG.Rotondo) and SAL003 (Salapia) contain high yields of aDNA with 17.79% and 28.16% respectively. All the data obtained from the genetic analysis are shown in **Table.1.**

4. DISCUSSION

The state of conservation of osteo-archaeological human samples here analysed was investigated by considering a larger set of samples previously characterized (see **Chapter 3**), taking into account a large and diverse chronology and origin with high variability in terms of chemical composition due to different diagenetic processes occurring within the burial environments. The same bone elements, such as tooth roots and petrous bones, were examined for each individual selected necropolis, thus defining a range of homogeneous samples. The width at 85% of the height of the $\nu_4(PO_4^{3-})$ vibrational mode at 604 cm^{-1} (FW85%) was chosen as it produces a well-defined signal being less influenced by other overlapping vibrational modes, describing the variation of bioapatite structural properties through IR spectroscopy (Dal Sasso et al., 2018).

This parameter is useful to track subtle variations in terms of physico-chemical properties in well-preserved samples as well as in highly diagenized samples. AmI/PO_4^{3-} is widely utilized for quantifying the presence of collagen preserved in the bone system, nevertheless water signal overlaps to the amide band. The width of the phosphate peak indirectly depends on the collagen content, as also confirmed by the AmI/PO_4^{3-} - FW85% correlation since the presence of collagenous material in the hierarchical structure of the bone plays an important role in controlling the crystallite size of bioapatite after the death of the individual. The data provided by these two parameters gave way to effectively determine the degree of alteration of the examined samples (**Figure.3**).

Samples exhibiting high values of FW85% (higher than 8.96 to 12.25) and organic matter content (higher than 0.04 to 0.40) were considered promising for DNA extraction. All selected samples produced sequenceable aDNA but with high variability in terms of endogenous DNA yields.

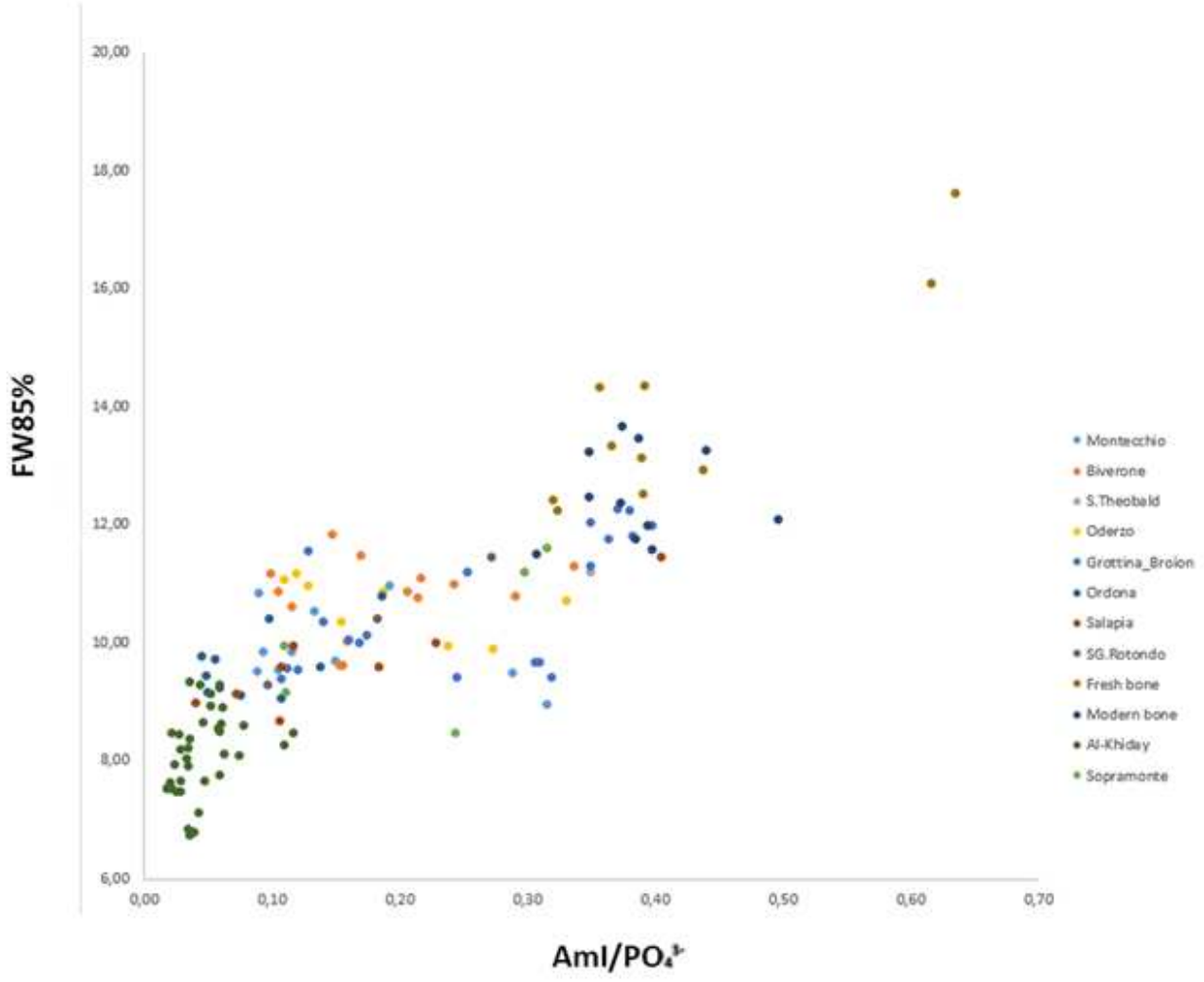

Figure 3. *IR data, reported in Chapter 3, show correlation between AmI/PO_4^{3-} and FW85%.*

Besides FTIR spectroscopy, collagen quality is commonly assessed by analysis of the atomic ratio C:N. Despite the same environmental conditions, genetic analysis revealed that similar samples for conservation status produced different yields of aDNA showing a good C:N range. Conversely, different aDNA yields and atomic ratio values were obtained from samples coming from different archaeological sites, characterized by a different state of conservation for burial conditions. To ascertain whether the quality of the collagen was indicative of the amount of extractable aDNA, a spectral analysis was performed of the collagen extracted from the selected samples that had a good atomic ratio range and amount of aDNA stored in the samples. Three medieval samples from the Desenzano del Garda site and characterized by a poor C:N ratio were selected for qualitative comparison of the secondary structure of collagen to determine the effectiveness of 1660:1690 cm^{-1} infrared parameter. Collagen secondary structure by means an Amide I band deconvolution (sensitive to the backbone of secondary structure (Ohi, 2009) through the two sub-bands ratio was investigated (1660:1690). The collagen Type I consists of a triple helix made of the repetitious amino acid

sequence glycine-X-Y, where X and Y are frequently proline or hydroxyproline and constitute the primary structure, called α-chains (Fietzek and Kuehn, 1976; Prockop and Williams, 1982). Polypeptide chains are linked to other chains via hydrogen bonds, which lead to the formation of the secondary structure which can be of two types: α-helix and β-sheets (Ohi, 2009). FTIR spectroscopy provides information on the secondary structure content of proteins, unlike NMR spectroscopy which provides information on the tertiary structure.

The characteristic bands found in the infrared spectra of proteins and polypeptide chains are Amide I and Amide II. These arise from the amide bonds that connect amino acids. The absorption of the Amide I band is associated with the stretching vibrations of the C=O bond of the amide, the absorption of the Amide II band is mainly associated with the bending vibrations of the N–H bond. Since both C=O and N–H bonds are involved in the hydrogen bonding that occurs between the different elements of the secondary structure, the positions of both the Amide I and Amide II bands are sensitive to the content of the secondary structure of a protein (DeFlores et al., 2009). Studies with proteins of known structure were used to systematically correlate the shape of the Amide I band to the proteins secondary structure (Byler and Susi, 1986; Surewicz and Mantsch, 1988). Molecular spectroscopy is useful to infer the quality of collagen (Chadefaux et al., 2009; France et al., 2014; Ishida and Griffiths, 1993; Kontopoulos et al., 2020; Riaz et al., 2018; Singh et al., 1993; Unal et al., 2016) and to that end, Amide I (1660 cm^{-1}) vibrational band is uniquely informative. The changes in Amide I profile with denaturation may reflect alterations in the collagen secondary structure, specifically a transition from ordered to less-ordered structure. The main degradation processes of collagen are oxidation, hydrolysis, and denaturation (Collins et al., 2002). Oxidation takes place on the basic amino acids, causing the weakening of physical stability of bone (Shoulders and Raines, 2009) while the collagen chain is split into shorter fragments, consisting of acid amides or keto acid derivatives during the denaturation.

Underground water, containing H_3O^+ and OH^- ions, may cause the breakdown of peptide bonds with N-terminal and C-terminal residues bringing to alkali or acidic hydrolysis of collagen (Rivas-Vela et al., 2021).

All FTIR spectra of lyophilized collagen from archaeological bones show a weak shoulder at ~ 1720 cm^{-1}, associated with acidic carboxyl groups, probably linked to a process of hydrolysis by humic soil (Derrick, 1991; Piccolo and Stevenson, 1982). The increase of protein carboxylation is associated with collagen protein fragmentation and aggregation. In this perspective, the FTIR band ratio 1660:1690 cm^{-1}, related to maturity of collagen cross-links (Paschalis et al., 2001), is negatively correlated to collagen quality: the lower is the ratio between these bands areas, the higher is the degenerative state of the collagen (Martínez Cortizas and López-Costas, 2020). This negative correlation, associated with the presence of the vibrational band of the carbonyl compounds, was

confirmed by comparison between 1660:1690 cm^{-1} ratio and the C:N data (**Figure.4**). Furthermore, from the results obtained from the set of available samples, it is possible to estimate that below a certain threshold of about 1.30 (1660:1690 area ratio) (**Figure.4**), the value of 1660:1690 cm^{-1} ratio indicates an instability in the collagen structure and consequently a poor conservation.

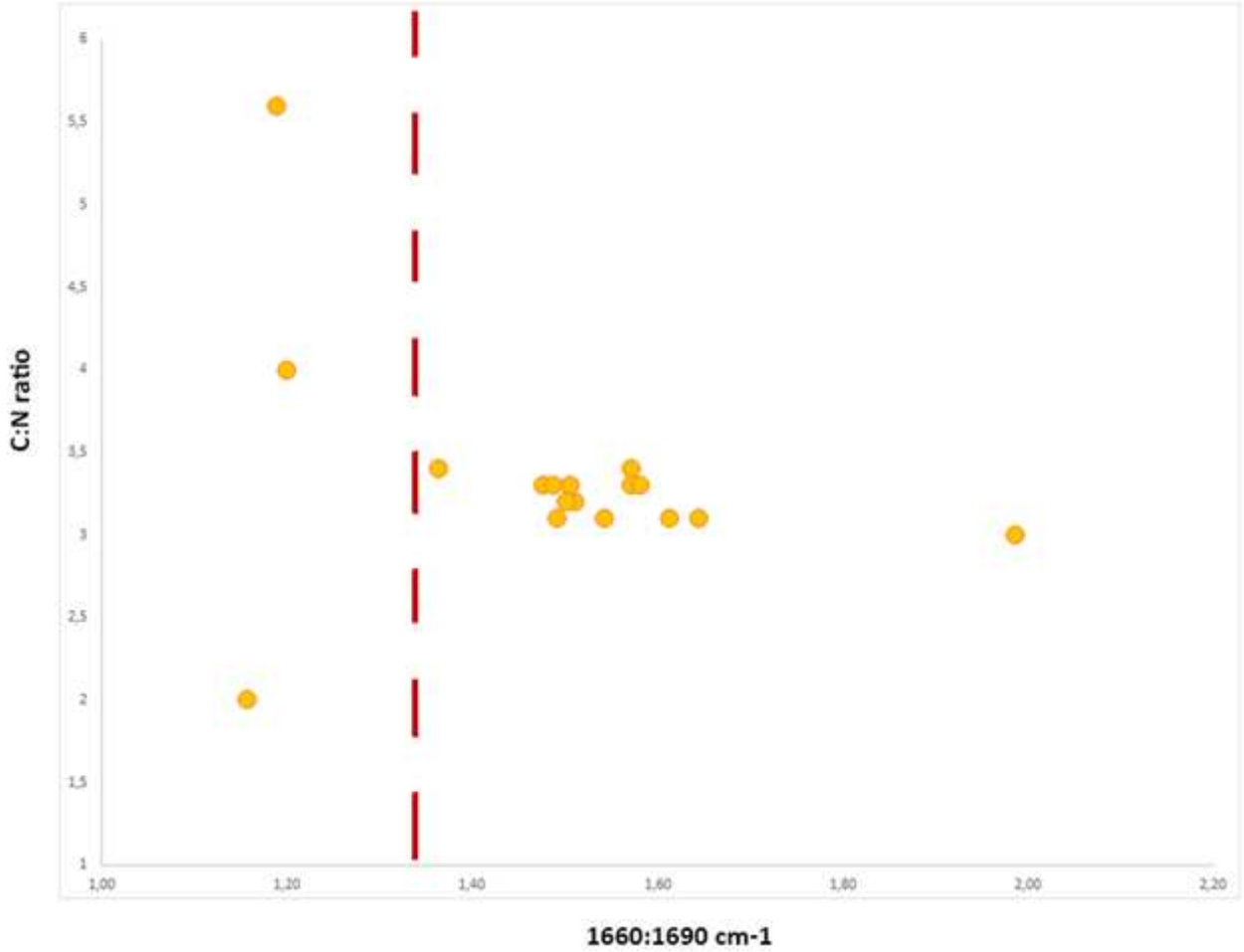

Figure 4. *Comparison between C:N ratio and 1660:1690 cm^{-1}. The determination of the conservative state of collagen through the spectral analysis of the sub-bands of Amide I (1660: 1690 cm^{-1}) well reflect the state of alteration of the secondary structure of collagen. Below the threshold (red line in the figure) of about 1.30 are the poorly-preserved collagen samples.*

As the deterioration proceeds, covalent bonds are cleaved and collagen molecules become unstable leading to the gelatinisation and the irreversible denaturation of polypeptide chains (Cappa et al., 2020). Secondary collagen structure could be used as a pre-screening for paleo-nutritional studies or radiocarbon dating while no correlation was found with aDNA yields.

Evaluating the spectra of archaeological bone, the shoulder linked to the denaturation of collagen (~ 1720 cm^{-1}) is not visible probably due to the overlaps with other vibrational bands of organic substances from the archaeological context (humic acids, fulvic acids) (Weiner, 2010). Furthermore,

the application of the ratio between the two amide sub-bands at 1660 cm^{-1} and 1690 cm^{-1} as pre-screening method directly on bone spectra is not straightforward due to the band overlaps.

For the careful selection of samples intended for analysis of stable isotopes, it is possible to evaluate the suitability of the finds by means of IR spectroscopy in two distinct phases. Initially, it will be possible to examine the state of conservation of the find by evaluating the state of alteration of the organic and inorganic component through infrared measurements of the bone sample and subsequently, if the quantity of collagen is considered sufficient, it will be possible to deepen the pre-screening by analyzing the structure polypeptide before laboratory analysis.

Linear regression analysis (lm), a predictive modelling technique that estimates the relationship between two or more variables (Gaurav and Patel, 2020; Kassambara, 2018), was performed. Regression analysis was chosen in that, focusing on the relationship between a dependent (aDNA) variable and one or more independent variables (parameters used) the dependent variable is assumed the effect by one or more independent variables.

The predictor value obtained was used to determine which parameters have powerful predictive bound to aDNA presence/absence. The model in R was created using aDNA values as target variable and AmI/PO_4^{3-} and FW85% as predictive variables. In **Figure.5** are reported the values of the intercept and the slope (coefficients) for endogenous aDNA and p-values for coefficients. The p-value for AmI/PO_4^{3-} is $9.09*10^{-16}$ indicating that this parameter is probably a good addition to the model with $R^2 = 0.4952$. The linear model was adopted with the peak width (FW85%) as it corresponded to a better dependent variable and overall result: the model has a p-value of $< 2*10^{-16}$ with $R^2 = 0.5754$. The contribution of multiple factors simultaneously was investigated by adopting a multiple linear model obtaining an $R^2 = 0.5763$ with a p-value of $< 2.2*10^{-16}$ (**Figure.5**). Among the variables considered, AmI/PO_4^{3-} is not representative of the model (p-value 0.651) unlike the more indicative variable FW85% (p-value $5.16*10^{-5}$).

The overall quality of the linear regression fit can be assessed using Residual Standard Error (RSE), displayed in the model summary (**Figure.5**). RSE provides an absolute measure of pattern in the data that cannot be explained by the model. The model with FW85% showed a RSE = 20.69 with respect to the AmI/PO_4^{3-} model with RSE = 22.55, indicating that the first model fits better the data.

This combination suggests that, if the parameters related to the organic and inorganic components are evaluated at the same time, the width of the phosphate peak has greater predictive probabilities. The parameter FW85% has a high potential probably determined by its property of being indirectly linked to the quantity of carbonate in the crystal lattice as well as sensitive to the presence of collagen with which bioapatite is intimately connected.

The multiple linear model here obtained is described by the formula % endogenous aDNA= a*Am/PO_4^{3-} + b*FW85% +β1, with β1= -90.708, a= 14.791, b=10.276. In conclusion, from multiple linear regression, the infrared parameter of phosphate peak width (FW85%) have greater predictive power for the presence/absence of DNA, thus being a promising tool to evaluate the potential of extraction of genetic molecules from archaeological bone samples.

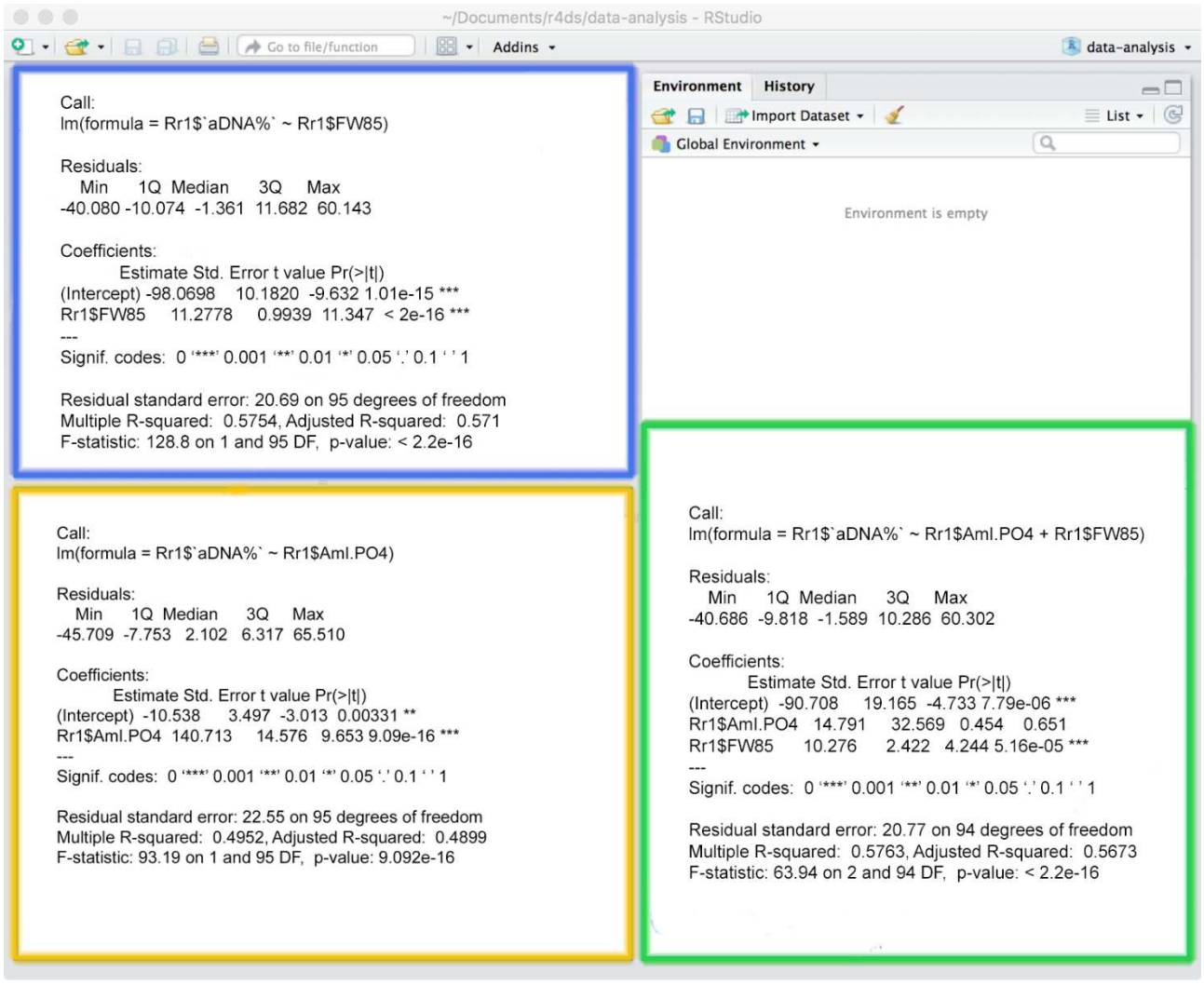

Figure 5. *Simple (yellow and blue) and multiple (green) regression model (function lm) in R utilizing AmI/PO_4^{3-} and FW85% as coefficients.*

5. CONCLUSION

The results here discussed show that the predictive model is functional with the indicated parameters but the quality/quantity of aDNA cannot be determined by means of the structural analysis of the organic component therein discussed. The quantity/quality of endogenous DNA may not be predictable because of the influence of environmental local factors. Measuring the FTIR signals of Amide I and phosphate peak width may predict the occurrence of DNA but it is still difficult to fully describe a multifactorial process.

At the end of the work, two questions arise that we hope can be resolved:

Is collagen degradation different from DNA degradation?

Is there a chemical control over DNA storage?

Results confirm a more intimate association with inorganic components and the effect of DNA on the crystal growth of bioapatite and its morphology was examined in previous studies of Okazaki et al. (2001) and Grunenwald et al. (2014).

Works suggest that there is an *in vivo* affinity binding between apatite crystal and DNA molecules and that its concentration has an impact on the growth of bio-crystal while its survival in calcified bone may is determined by phosphate-promoted molecule desorption, suggesting an ion exchange process between phosphate anionic groups of DNA backbone and labile non-apatitic hydrogen-phosphate ions released from the hydrated layer synthetic biomimetic apatite.

The work of De Campos Vidal and Mello (2011) claims four possible options for the *post-mortem* conservation of DNA: two hypotheses related to the maturation of collagen and inorganic compound *in vivo*, and two hypotheses related to diagenetic phenomena where the most accredited hypothesis is that aDNA can survive closely linked to the mineral phases. These statements are not opposed to the results we obtained as the monitoring of the inorganic component has a higher value than that of the organic component, although we believe that the latter should not be neglected for isotope and radiocarbon analyzes.

Future studies will be needed to investigate the relationship between DNA and hydroxylapatite and how this chemical interaction is functional to its preservation during diagenetic processes useful for both molecular-anthropological and forensic studies.

ACKNOWLEDGEMENT

We thank Dr. Isabella Tessari, laboratory technician to the Department of Biology of the Padua University and the laboratory technicians Drs. Lisa Santello and Carlotta Betto to the Department of Geosciences of Padua University for their valuable technical support. In addition, an acknowledgement goes to Dr. Christophe Snoeck Ph.D., researcher at Vrije Universiteit Brussel and the head of the Brussels Bioarchaeology Lab (BB-LAB) and to Dr. Ina Reiche Ph.D., research director of French National Centre for Scientific Research-CNRS for their valuable scientific advices.

FUNDING ACQUISITION

Starting Grant Stars (STARS-StG) “Paleopathology and molecular analysis in times of transition: understanding diet and health in Early Middle Ages (UDHEMA)”, Principal Investigator (PI): Maurizio Marinato.

Ph.D. scolarship, Fondazione Cassa di Risparmio di Padova e Rovigo (CARIPARO) and UniPD PRID 2019 to Luca Pagani.

AUTHORS CONTRIBUTION

C.S., G.A., L.P., conceptualization; C.S., M.M., G.D.S., T.S., S.A., C.L.S. data curation; formal analyses; L.N., G.A., L.P., validation; G.A., M.M., L.P. funding acquisition; N.C., M.R., resources; G.A., L.N., L.P., supervision; C.S., M.M., T.S., S.A., investigation; C.S., writing-original draft; all authors writing-review and editing.

ETHICAL STATEMENT

Human osteological remains, analyzed for this manuscript, are preserved in the collections of the Museum of Anthropology, Padua University. Bone remains come from archaeological campaigns of necropolis and caves carried out during the 19^{th}-20^{th} century and from exhumations, cemetery ossuaries or are acquired by Hospitals or other charitable organizations, the latter mostly collected by Prof. Enrico Tedeschi (1860-1931). The Museum of Anthropology also houses an important ancient osteological collection of about 200 human skeletons coming from the archaeological site of Al-Khiday (Central Sudan) covering a wide chronological span (7000 BCE to the IV-VI c.BCE). Other bone elements analyzed were provided by the Archaeological Superintendence of the Metropolitan Area of Venice and for the Provinces of Belluno, Padua and Treviso and by the Superintendence of Cultural Heritage of Trento.

DECLARATION OF COMPETING INTEREST

The authors do not have a conflict of interest to declare and the manuscript has not been published or presented to another journal.

Table 1. *Information on the included samples and spectroscopic, isotope and genetic results.*

SAMPLE	CHRONOLOGY	ARCHAEOLOGICAL SITE	SKELETAL ELEMENT	C:N ratio	aDNA (%) endogenous	1660 cm-1	16690 cm-1	1660:1690 cm^{-1}	Am/PO_4^{3-}	FW85%	IRSF
BR-03_ PbGCB O-Ce	Bronze Age	Grottina Covoloni Broion, Italy	petrous bone		5.44%				0.12	9.54	3.68
BR-04_PbGCB O-Ce	Bronze Age	Grottina Covoloni Broion, Italy	petrous bone		4.90%				0.14	10.36	3.41
BR-01_Pb GCB O:38	Bronze Age	Grottina Covoloni Broion, Italy	petrous bone		7.81%				0.11	9.57	3.66
BR-05_ PbGCB	Bronze Age	Grottina Covoloni Broion, Italy	petrous bone		7.14%				0.16	10.04	3.41
BR-25_GCB	Bronze Age	Grottina Covoloni Broion, Italy	tooth		0.09%				0.25	11.2	3.18
BR-06_GCB	Bronze Age	Grottina Covoloni Broion, Italy	tooth		0.88%				0.31	9.67	3.56
BR-07_GCB	Bronze Age	Grottina Covoloni Broion, Italy	tooth		0.13%				0.35	12.02	3.00
BR-09_GCB	Bronze Age	Grottina Covoloni Broion, Italy	tooth		0.13%				0.13	11.54	3.53
BR-16_ GCB 4B+4S 5S	Bronze Age	Grottina Covoloni Broion, Italy	tooth		0.42%				0.31	9.65	3.10
BR-22_ GCB 7-353	Bronze Age	Grottina Covoloni Broion, Italy	tooth	3.4	0.07%	15.8	10.0	1.57	0.11	9.39	3.64
BR-24_ GCB 53	Bronze Age	Grottina Covoloni Broion, Italy	tooth		3.86%				0.40	11.97	3.09
BR-10_GCB 223	Bronze Age	Grottina Covoloni Broion, Italy	tooth		0.30%				0.32	9.41	3.70
BR-27_GCB 275	Bronze Age	Grottina Covoloni Broion, Italy	tooth		0.22%				0.38	12.24	2.99
BR-13_ GCB274	Bronze Age	Grottina Covoloni Broion, Italy	tooth		0.25%				0.37	12.25	2.99
BR-17_GCB 629	Bronze Age	Grottina Covoloni Broion, Italy	tooth		0.35%				0.17	9.98	3.29
BR-21_GCB 807	Bronze Age	Grottina Covoloni Broion, Italy	tooth		1.52%				0.24	9.4	3.76
BR-19_ GCB 909	Bronze Age	Grottina Covoloni Broion, Italy	tooth	3.3	17.32%	15.8	10.0	1.57	0.17	10.12	3.37
BR-12_ GCB O:156	Bronze Age	Grottina Covoloni Broion, Italy	tooth		0.70%				0.38	11.8	3.11
BR-14_ GCB SeH6 Str248	Bronze Age	Grottina Covoloni Broion, Italy	tooth	3.2	9.40%	15.8	10.4	1.51	0.36	11.74	3.05
BR-18_GCB 221	Bronze Age	Grottina Covoloni Broion, Italy	tooth		0.85%				0.35	11.28	3.13
ORD001_PB	Iron Age	Ordona, Italy	petrous bone	3.3	2.14%	12.0	7.6	1.58	0.08	9.09	4.07
ORD004_PB	Iron Age	Ordona, Italy	petrous bone	3.1	5.50%	14.4	9.3	1.54	0.10	10.39	3.87
ORD006_PB	Iron Age	Ordona, Italy	petrous bone	3.3	31.42%	9.2	6.2	1.48	0.05	9.43	4.39
ORD009_PB	Iron Age	Ordona, Italy	petrous bone	3.1	42.27%	11.2	7.5	1.49	0.05	9.75	4.61
ORD010_PB	11thc.AD	Ordona, Italy	petrous bone	3.4	43.75%	7.1	5.2	1.36	0.14	9.59	4.16
ORD011_PB	Iron Age	Ordona, Italy	petrous bone	3.3	14.73%	8.2	5.5	1.51	0.19	10.78	3.51
ORD012_PB	Iron Age	Ordona, Italy	petrous bone		0.09%				0.05	9.16	4.36
ORD014_PB	Iron Age	Ordona, Italy	petrous bone	3.2	15.55%	11.7	7.8	1.50	0.11	9.04	4.37
ORD018_PB	Iron Age	Ordona, Italy	petrous bone		0.69%				0.06	9.23	4.13
ORD019_PB	Iron Age	Ordona, Italy	petrous bone		3.69%				0.06	9.7	4.29
SAL001_T	Iron Age	Salapia, Italy	tooth		4.71%				0.18	9.58	3.73
SAL003_PB	Iron Age	Salapia, Italy	petrous bone	3.3	28.16%	9.8	6.6	1.49	0.12	9.94	4.11
SAL005_PB	Iron Age	Salapia, Italy	petrous bone		0.99%				0.11	9.59	4.45
SAL007_PB	Iron Age	Salapia, Italy	petrous bone		3.39%				0.11	8.67	5.13
SAL008_T	Iron Age	Salapia, Italy	tooth		1.31%				0.23	9.99	3.67
SAL010_PB	Iron Age	Salapia, Italy	petrous bone		2.74%				0.04	8.96	5.12
SAL011_T	Iron Age	Salapia, Italy	tooth	3.1	3.85%	10.0	6.1	1.65	0.40	11.44	3.22
SAL012_T	Iron Age	Salapia, Italy	tooth		0.15%				0.07	9.12	4.58

REFERENCES

Allentoft, M.E., Collins, M., Harker, D., Haile, J., Oskam, C.L., Hale, M.L., Campos, P.F., Samaniego, J.A., Gilbert, M.T.P., Willerslev, E., Zhang, G., Scofield, R.P., Holdaway, R.N., Bunce, M., 2012. The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. Proc. R. Soc. B Biol. Sci. 279, 4724–4733. https://doi.org/10.1098/rspb.2012.1745

Ambrose, S.H., 1990. Preparation and characterization of bone and tooth collagen for isotopic analysis. J. Archaeol. Sci. 17, 431–451. https://doi.org/10.1016/0305-4403(90)90007-R

Aneli, S., Saupe, T., Montinaro, F., Solnik, A., Molinaro, L., Scaggion, C., Carrara, N., Raveane, A., Kivisild, T., Metspalu, M., Scheib, C.L., Pagani, L., 2022. The Genetic Origin of Daunians and the Pan-Mediterranean Southern Italian Iron Age Context. Mol. Biol. Evol. 39, 1–16. https://doi.org/10.1093/molbev/msac014

Arrondo, J.L.R., Muga, A., Castresana, J., Goñi, F.M., 1993. Quantitative studies of the structure of proteins in solution by fourier-transform infrared spectroscopy. Prog. Biophys. Mol. Biol. 59, 23–56. https://doi.org/10.1016/0079-6107(93)90006-6

Balzeau, A., 2015. Comparative aspects of temporal bone pneumatization in some African fossil hominins. BMSAP 27, 135–141. https://doi.org/10.1007/s13219-015-0126-5

Barth, A., 2007. Infrared spectroscopy of proteins. Biochim. Biophys. Acta - Bioenerg. 1767, 1073–1101. https://doi.org/10.1016/j.bbabio.2007.06.004

Barth, H.D., Zimmermann, E.A., Schaible, E., Tang, S.Y., Alliston, T., Ritchie, R.O., 2011. Characterization of the effects of x-ray irradiation on the hierarchical structure and mechanical properties of human cortical bone. Biomaterials 32, 8892–8904. https://doi.org/10.1016/j.biomaterials.2011.08.013

Bosshardt, D.D., Selvig, K.A., 1997. Dental cementum: The dynamic tissue covering of the root. Periodontol. 2000 13, 41–75. https://doi.org/10.1111/j.1600-0757.1997.tb00095.x

Brace, C.L., Smith, S.L., Hunt, K.D., 1991. What big teeth you had, Grandma! Human tooth size, past and present, in: M.K. Kelley and C.S. Larsen (Ed.), Advances in Dental Anthropology. Wiley-Liss, New York, pp. 33–57.

Brown, K.E., Winter, B.J., Shen, C., Yang, D.Y., 2016. Developing Minimally Destructive Protocols for DNA Analysis of Museum Collection Bone Artifacts. Saa 1.

Byler, D.M., Susi, H., 1986. Examination of the secondary structure of proteins by deconvolved FTIR spectra. Biopolymers 25, 469–487. https://doi.org/10.1002/bip.360250307

Campos, P.F., Craig, O.E., Turner-Walker, G., Peacock, E., Willerslev, E., Gilbert, M.T.P., 2012. DNA in ancient bone - Where is it located and how should we extract it? Ann. Anat. 194, 7–16. https://doi.org/10.1016/j.aanat.2011.07.003

Canci, A., Chavarria, A., Marinato, M., 2012. Il cimitero della chiesa altomedievale di San Lorenzo di Desenzano (Bs). VI Congr. degli Archeol. Mediev. Ital. Note di Bi, 452–455.

Cappa, F., Paganoni, I., Carsote, C., Badea, E., Schreiner, M., 2020. Studies on the effects of mixed light-thermal ageing on parchment by vibrational spectroscopy and micro hot table method. Herit. Sci. 8, 1–12. https://doi.org/10.1186/s40494-020-0353-z

Carden, A., Morris, M.D., 2000. Application of vibrational spectroscopy to the study of mineralized tissues (review). J. Biomed. Opt. 5, 259. https://doi.org/10.1117/1.429994

Carrara, N., Scaggion, C., Holland, E., 2018. The Tedeschi collection: A collection of documented and undocumented human skeletal remains at the Museum of Anthropology, Padua University (Italy). Am. J. Phys. Anthropol. 166. https://doi.org/10.1002/ajpa.23471

Chadefaux, C., Hô, A. Le, Bellot-gurlet, L., Ina, R., 2009. Curve-Fitting Micro-Atr-Ftir Studies of the Amide I and Ii Bands of Type I Collagen in. e-PRESERVATIONScience 129–137.

Collins, M.J., Nielsen-Marsh, C.M., Hiller, J., Smith, C.I., Roberts, J.P., Prigodich, R. V., Wess, T.J., Csapò, J., Millard, A.R., Turner-Walker, G., 2002. The survival of organic matter in bone: A review. Archaeometry 44, 383–394. https://doi.org/10.1111/1475-4754.t01-1-00071

Collins, M.J., Penkman, K.E.H., Rohland, N., Shapiro, B., Dobberstein, R.C., Ritz-Timme, S., Hofreiter, M., 2009. Is amino acid racemization a useful tool for screening for ancient DNA in bone? Proc. R. Soc. B Biol. Sci. 276, 2971–2977. https://doi.org/10.1098/rspb.2009.0563

Collins, M.J., Riley, M.S., Child, A.M., Turner-Walker, G., 1995. A Basic Mathematical Simulation of the Chemical Degradation of Ancient Collagen. J. Archaeol. Sci. 22, 175–183. https://doi.org/10.1006/jasc.1995.0019

Corrain, C., 1986. Resti scheletrici umani da Ordona (Foggia), Secc. VII–IV aC. Les Tombes Dauniennes.Etudes Philol. d'Arch_eologie d'Histoire Ancienne Publi_ees par l'Institut Hist. Belge Rome. Bruxel 24, 787–807.

Corrain, C., Capitanio, M., Erspamer, G., 1972. I resti scheletrici della ne-cropoli di Salapia (Cerignola), secoli IX-III aC. Padova Soc. Coop. Tipogr.

Dal Sasso, G., Asscher, Y., Angelini, I., Nodari, L., Artioli, G., 2018. A universal curve of apatite crystallinity for the assessment of bone integrity and preservation. Sci. Rep. 8, 1–13. https://doi.org/10.1038/s41598-018-30642-z

Dal Sasso, G., Lebon, M., Angelini, I., Maritan, L., Usai, D., Artioli, G., 2016. Bone diagenesis variability among multiple burial phases at Al Khiday (Sudan) investigated by ATR-FTIR spectroscopy. Palaeogeogr. Palaeoclimatol. Palaeoecol. 463, 168–179. https://doi.org/10.1016/j.palaeo.2016.10.005

De Campos Vidal, B., Mello, M.L.S., 2011. Collagen type I amide I band infrared spectroscopy.

Micron 42, 283–289. https://doi.org/10.1016/j.micron.2010.09.010

DeFlores, L.P., Ganim, Z., Nicodemus, R.A., Tokmakoff, A., 2009. Amide I′−II′ 2D IR Spectroscopy Provides Enhanced Protein Secondary Structural Sensitivity. J. Am. Chem. Soc. 131, 3385–3391. https://doi.org/10.1021/ja8094922

Demes, B., Creel, N., 1988. Bite force, diet, and cranial morphology of fossil hominids. J. Hum. Evol. 17, 657–670. https://doi.org/10.1016/0047-2484(88)90023-1

Der Sarkissian, C., Allentoft, M.E., Ávila-Arcos, M.C., Barnett, R., Campos, P.F., Cappellini, E., Ermini, L., Fernández, R., da Fonseca, R., Ginolhac, A., Hansen, A.J., Jónsson, H., Korneliussen, T., Margaryan, A., Martin, M.D., Víctor Moreno-Mayar, J., Raghavan, M., Rasmussen, M., Velasco, M.S., Schroeder, H., Schubert, M., Seguin-Orlando, A., Wales, N., Gilbert, M.T.P., Willerslev, E., Orlando, L., 2015. Ancient genomics. Philos. Trans. R. Soc. B Biol. Sci. 370. https://doi.org/10.1098/rstb.2013.0387

Derrick, M., 1991. Evaluation of the state of degradation of Dead Sea Scroll samples using FT-IR spectroscopy. AIC B. Pap. Gr. Annu. 10, 1–16.

Dey, P., 2020. Bone Mineralisation, in: Intech. pp. 137–144.

Doden, E., Halves, R., 1984. On the functional morphology of the human petrous bone. Am. J. Anat. 169, 451–462. https://doi.org/10.1002/aja.1001690407

Doyle, B.B., Bendit, E.G., Blout, E.R., 1975. Infrared spectroscopy of collagen and collagen-like polypeptides. Biopolymers 14, 937–957. https://doi.org/10.1002/bip.1975.360140505

Farlay, D., Duclos, M.E., Gineyts, E., Bertholon, C., Viguet-Carrin, S., Nallala, J., Sockalingum, G.D., Bertrand, D., Roger, T., Hartmann, D.J., Chapurlat, R., Boivin, G., 2011. The ratio 1660/1690 cm-1 measured by infrared microspectroscopy is not specific of enzymatic collagen cross-links in bone tissue. PLoS One 6. https://doi.org/10.1371/journal.pone.0028736

Fietzek, P.P., Kuehn, K., 1976. The primary structure of collagen. Int. Rev. Connect. Tissue Res. Vol. 7, 1–60. https://doi.org/10.1016/b978-0-12-363707-9.50007-1

Figueiredo, M.M., Gamelas, J.A.F., Martins, A.G., 2012. Characterization of Bone and Bone-Based Graft Materials Using FTIR Spectroscopy. Infrared Spectrosc. - Life Biomed. Sci. https://doi.org/10.5772/36379

Fleet, M.E., 2009. Infrared spectra of carbonate apatites: ν2-Region bands. Biomaterials 30, 1473–1481. https://doi.org/10.1016/j.biomaterials.2008.12.007

Florencio-Silva, R., Sasso, G.R. da S., Sasso-Cerri, E., Simões, M.J., Cerri, P.S., 2015. Biology of Bone Tissue: Structure, Function, and Factors That Influence Bone Cells. Biomed Res. Int. 2015, 1–17. https://doi.org/10.1155/2015/421746

France, C.A.M., Thomas, D.B., Doney, C.R., Madden, O., 2014. FT-Raman spectroscopy as a

method for screening collagen diagenesis in bone. J. Archaeol. Sci. 42, 346–355. https://doi.org/10.1016/j.jas.2013.11.020

Fratzl, P., Gupta, H.S., Paschalis, E.P., Roschger, P., 2004. Structure and mechanical quality of the collagen-mineral nano-composite in bone. J. Mater. Chem. 14, 2115–2123. https://doi.org/10.1039/b402005g

Fredericks, J.D., Bennett, P., Williams, A., Rogers, K.D., 2012. FTIR spectroscopy: A new diagnostic tool to aid DNA analysis from heated bone. Forensic Sci. Int. Genet. 6, 375–380. https://doi.org/10.1016/j.fsigen.2011.07.014

Gamba, C., Jones, E.R., Teasdale, M.D., McLaughlin, R.L., Gonzalez-Fortes, G., Mattiangeli, V., Domboróczki, L., Kővári, I., Pap, I., Anders, A., Whittle, A., Dani, J., Raczky, P., Higham, T.F.G., Hofreiter, M., Bradley, D.G., Pinhasi, R., 2014. Genome flux and stasis in a five millennium transect of European prehistory. Nat. Commun. 5, 5257. https://doi.org/10.1038/ncomms6257

Ganim, Z., Chung, H.S., Smith, A.W., Deflores, L.P., Jones, K.C., Tokmakoff, A., 2008. Amide I two-dimensional infrared spectroscopy of proteins. Acc. Chem. Res. 41, 432–41. https://doi.org/10.1021/ar700188n

Gaurav, K.A., Patel, L., 2020. Machine Learning With R. https://doi.org/10.4018/978-1-7998-2718-4.ch015

Goormaghtigh, E., Ruysschaert, J.-M., Raussens, V., 2006. Evaluation of the Information Content in Infrared Spectra for Protein Secondary Structure Determination. Biophys. J. 90, 2946–2957. https://doi.org/10.1529/biophysj.105.072017

Gotherstrom, A., Collins, M.J., Angerbjorn, A., Liden, K., 2002. Bone preservation and DNA amplification. Archaeometry 44, 395–404. https://doi.org/10.1111/1475-4754.00072

Grunenwald, A., Keyser, C., Sautereau, A.M., Crubézy, E., Ludes, B., Drouet, C., 2014. Adsorption of DNA on biomimetic apatites: Toward the understanding of the role of bone and tooth mineral on the preservation of ancient DNA. Appl. Surf. Sci. 292, 867–875. https://doi.org/10.1016/j.apsusc.2013.12.063

Guiry, E.J., Szpak, P., 2020. Quality control for modern bone collagen stable carbon and nitrogen isotope measurements. Methods Ecol. Evol. 11, 1049–1060. https://doi.org/10.1111/2041-210X.13433

Hansen, H.B., Damgaard, P.B., Margaryan, A., Stenderup, J., Lynnerup, N., Willerslev, E., Allentoft, M.E., 2017. Comparing ancient DNA preservation in petrous bone and tooth cementum. PLoS One 12, 1–18. https://doi.org/10.1371/journal.pone.0170940

Harney, É., Cheronet, O., Fernandes, D.M., Sirak, K., Mah, M., Bernardos, R., Adamski, N.,

Broomandkhoshbacht, N., Callan, K., Lawson, A.M., Oppenheimer, J., Stewardson, K., Zalzala, F., Anders, A., Candilio, F., Constantinescu, M., Coppa, A., Ciobanu, I., Dani, J., Gallina, Z., Genchi, F., Nagy, E.G., Hajdu, T., Hellebrandt, M., Horváth, A., Király, Á., Kiss, K., Kolozsi, B., Kovács, P., Köhler, K., Lucci, M., Pap, I., Popovici, S., Raczky, P., Simalcsik, A., Szeniczey, T., Vasilyev, S., Virag, C., Rohland, N., Reich, D., Pinhasi, R., 2021. A minimally destructive protocol for DNA extraction from ancient teeth. Genome Res. 31, 472–483. https://doi.org/10.1101/gr.267534.120

Harrison, R.G., Katzenberg, M.A., 2003. Paleodiet studies using stable carbon isotopes from bone apatite and collagen: Examples from Southern Ontario and San Nicolas Island, California. J. Anthropol. Archaeol. 22, 227–244. https://doi.org/10.1016/S0278-4165(03)00037-0

Haynes, S., Searle, J.B., Bretman, A., Dobney, K.M., 2002. Bone preservation and ancient DNA: The application of screening methods for predicting DNA survival. J. Archaeol. Sci. 29, 585–592. https://doi.org/10.1006/jasc.2001.0731

Hedges, R.E.M., 2002. Bone diagenesis: an overview of processes. Archaeometry 44, 319–328. https://doi.org/10.1111/1475-4754.00064

Hedges, R.E.M., 1992. A Review of Current Approaches in the Pretreatment of Bone. Radiocarbon 34, 279–291.

Ishida, K.P., Griffiths, P.R., 1993. Comparison of the amide I/II intensity ratio of solution and solid-state proteins sampled by transmission, attenuated total reflectance, and diffuse reflectance spectrometry. Appl. Spectrosc. 47, 584–589. https://doi.org/10.1366/0003702934067306

Kassambara, K., 2018. Machine Learning Essentials: Practical Guide in R, 1° edition. ed. CreateSpace Independent Publishing Platform.

Kendall, C., Eriksen, A.M.H., Kontopoulos, I., Collins, M.J., Turner-Walker, G., 2018. Diagenesis of archaeological bone and tooth. Palaeogeogr. Palaeoclimatol. Palaeoecol. 491, 21–37. https://doi.org/10.1016/j.palaeo.2017.11.041

Kirchner, M.T., Edwards, H.G.M., Lucy, D., Pollard, A.M., 1997. Ancient and modern specimens of human teeth: A fourier transform Raman spectroscopic study. J. Raman Spectrosc. 28, 171–178. https://doi.org/10.1002/(sici)1097-4555(199702)28:2/3<171::aid-jrs63>3.0.co;2-v

Kontopoulos, I., Penkman, K., McAllister, G.D., Lynnerup, N., Damgaard, P.B., Hansen, H.B., Allentoft, M.E., Collins, M.J., 2019. Petrous bone diagenesis: a multi-analytical approach. Palaeogeogr. Palaeoclimatol. Palaeoecol. 518, 143–154. https://doi.org/10.1016/j.palaeo.2019.01.005

Kontopoulos, I., Penkman, K., Mullin, V.E., Winkelbach, L., Unterländer, M., Scheu, A., Kreutzer, S., Hansen, H.B., Margaryan, A., Teasdale, M.D., Gehlen, B., Street, M., Lynnerup, N.,

Liritzis, I., Sampson, A., Papageorgopoulou, C., Allentoft, M.E., Burger, J., Bradley, D.G., Collins, M.J., 2020. Screening archaeological bone for palaeogenetic and palaeoproteomic studies. PLoS One 15, 1–17. https://doi.org/10.1371/journal.pone.0235146

Lebon, M., Zazzo, A., Reiche, I., 2014. Screening in situ bone and teeth preservation by ATR-FTIR mapping. Palaeogeogr. Palaeoclimatol. Palaeoecol. 416, 110–119. https://doi.org/10.1016/j.palaeo.2014.08.001

Leskovar, T., Zupanič Pajnič, I., Geršak, Ž.M., Jerman, I., Črešnar, M., 2020. ATR-FTIR spectroscopy combined with data manipulation as a pre-screening method to assess DNA preservation in skeletal remains. Forensic Sci. Int. Genet. 44, 102196. https://doi.org/10.1016/j.fsigen.2019.102196

Ligabue, G., 1973. GROTTINA DEI COVOLONI DEL BROION, in: Preistoria Alpina. pp. 263–264.

Martínez Cortizas, A., López-Costas, O., 2020. Linking structural and compositional changes in archaeological human bone collagen: an FTIR-ATR approach. Sci. Rep. 10, 1–14. https://doi.org/10.1038/s41598-020-74993-y

Mays, S., Edlers, J., Humphrey, L., White, W., Marshall, P., 2013. Science and the Dead: A Guideline for the Destructive Sampling of Archaeological Human Remains for Scientific Analysis.

Mieczkowska, A., Mansur, S.A., Irwin, N., Flatt, P.R., Chappard, D., Mabilleau, G., 2015. Alteration of the bone tissue material properties in type 1 diabetes mellitus: A Fourier transform infrared microspectroscopy study. Bone 76, 31–39. https://doi.org/10.1016/j.bone.2015.03.010

Naito, Y.I., Yamane, M., Kitagawa, H., 2020. A protocol for using attenuated total reflection Fourier-transform infrared spectroscopy for pre-screening ancient bone collagen prior to radiocarbon dating. Rapid Commun. Mass Spectrom. 34. https://doi.org/10.1002/rcm.8720

Nandiyanto, A.B.D., Oktiani, R., Ragadhita, R., 2019. How to read and interpret ftir spectroscope of organic material. Indones. J. Sci. Technol. 4, 97–118. https://doi.org/10.17509/ijost.v4i1.15806

Ohi, M.D., 2009. EM Analysis of Protein Structure. Encycl. Life Sci. 662–666. https://doi.org/10.1002/9780470015902.a0021885

Okazaki, M., Yoshida, Y., Yamaguchi, S., Kaneno, M., Elliott, J.C., 2001. Affinity binding phenomena of DNA onto apatite crystals. Biomaterials 22, 2459–2464. https://doi.org/10.1016/S0142-9612(00)00433-6

Orlando, L., Allaby, R., Skoglund, P., Der Sarkissian, C., Stockhammer, P.W., Ávila-Arcos, M.C., Fu, Q., Krause, J., Willerslev, E., Stone, A.C., Warinner, C., 2021. Ancient DNA analysis. Nat.

Rev. Methods Prim. 1. https://doi.org/10.1038/s43586-020-00011-0

Oxilia, G., Bortolini, E., Martini, S., Papini, A., Boggioni, M., Buti, L., Figus, C., Sorrentino, R., Townsend, G., Kaidonis, J., Fiorenza, L., Cristiani, E., Kullmer, O., Moggi-Cecchi, J., Benazzi, S., 2018. The physiological linkage between molar inclination and dental macrowear pattern. Am. J. Phys. Anthropol. 166, 941–951. https://doi.org/10.1002/ajpa.23476

Parker, C., Rohrlach, A.B., Friederich, S., Nagel, S., Meyer, M., Krause, J., Bos, K.I., Haak, W., 2020. A systematic investigation of human DNA preservation in medieval skeletons. Sci. Rep. 10, 18225. https://doi.org/10.1038/s41598-020-75163-w

Paschalis, E.P., Burr, D.B., Mendelsohn, R., Hock, J.M., Boskey, A.L., 2003. Bone Mineral and Collagen Quality in Humeri of Ovariectomized Cynomolgus Monkeys Given rhPTH(1-34) for 18 Months. J. Bone Miner. Res. 18, 769–775. https://doi.org/10.1359/jbmr.2003.18.4.769

Paschalis, E.P., Mendelsohn, R., Boskey, A.L., 2011. Infrared assessment of bone quality: A review. Clin. Orthop. Relat. Res. 469, 2170–2178. https://doi.org/10.1007/s11999-010-1751-4

Paschalis, E.P., Shane, E., Lyritis, G., Skarantavos, G., Mendelsohn, R., Boskey, A.L., 2004. Bone fragility and collagen cross-links. J. Bone Miner. Res. 19, 2000–2004. https://doi.org/10.1359/JBMR.040820

Paschalis, E.P., Verdelis, K., Doty, S.B., Boskey, A.L., Mendelsohn, R., Yamauchi, M., 2001. Spectroscopic characterization of collagen cross-links in bone. J. Bone Miner. Res. 16, 1821–1828. https://doi.org/10.1359/jbmr.2001.16.10.1821

Piccolo, A., Stevenson, F.J., 1982. Infrared spectra of Cu2+ Pb2+ and Ca2+ complexes of soil humic substances. Geoderma 27, 195–208. https://doi.org/10.1016/0016-7061(82)90030-1

Pleshko, N.L., Boskey, A.L., Mendelsohn, R., 1992. An FT-IR microscopic investigation of the effects of tissue preservation on bone. Calcif. Tissue Int. 51, 72–77. https://doi.org/10.1007/BF00296221

Privat, K.L., O'connell, T.C., Richards, M.P., 2002. Stable Isotope Analysis of Human and Faunal Remains from the Anglo-Saxon Cemetery at Berinsfield, Oxfordshire: Dietary and Social Implications. J. Archaeol. Sci. 29, 779–790. https://doi.org/10.1006/jasc.2001.0785

Prockop, D.J., Williams, C.J., 1982. Structure of the Organic Matrix: Collagen Structure (Chemical), in: Biological Mineralization and Demineralization. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 161–177. https://doi.org/10.1007/978-3-642-68574-3_8

Rathmann, H., Reyes-Centeno, H., 2020. Testing the utility of dental morphological trait combinations for inferring human neutral genetic variation. Proc. Natl. Acad. Sci. 117, 10769–10777. https://doi.org/10.1073/pnas.1914330117

Reznikov, N., Bilton, M., Lari, L., Stevens, M.M., Kröger, R., 2018. Fractal-like hierarchical

organization of bone begins at the nanoscale. Science (80-.). 360. https://doi.org/10.1126/science.aao2189

Riaz, T., Zeeshan, R., Zarif, F., Ilyas, K., Muhammad, N., Safi, S.Z., Rahim, A., Rizvi, S.A.A., Rehman, I.U., 2018. FTIR analysis of natural and synthetic collagen. Appl. Spectrosc. Rev. 53, 703–746. https://doi.org/10.1080/05704928.2018.1426595

Richards, M.P., Hedges, R.E.M., 1999. Stable Isotope Evidence for Similarities in the Types of Marine Foods Used by Late Mesolithic Humans at Sites Along the Atlantic Coast of Europe. J. Archaeol. Sci. 26, 717–722. https://doi.org/10.1006/jasc.1998.0387

Rivas-Vela, C.I., Amaya-Llano, S.L., Castaño-Tostado, E., Castillo-Herrera, G.A., 2021. Protein Hydrolysis by Subcritical Water: A New Perspective on Obtaining Bioactive Peptides. Molecules 26, 6655. https://doi.org/10.3390/molecules26216655

Rohland, N., Hofreiter, M., 2007. Ancient dna extraction from bones and teeth. Nat. Protoc. 2, 1756–1762. https://doi.org/10.1038/nprot.2007.247

Rollo, F., Ubaldi, M., Marota, I., Luciani, S., Ermini, L., 2002. DNA Diagenesis: Effect of Environment and Time on Human Bone. Anc. Biomol. 4, 1–7. https://doi.org/10.1080/13586120290018473

Saupe, T., Montinaro, F., Scaggion, C., Carrara, N., Kivisild, T., D'Atanasio, E., Hui, R., Solnik, A., Lebrasseur, O., Larson, G., Alessandri, L., Arienzo, I., De Angelis, F., Rolfo, M.F., Skeates, R., Silvestri, L., Beckett, J., Talamo, S., Dolfini, A., Miari, M., Metspalu, M., Benazzi, S., Capelli, C., Pagani, L., Scheib, C.L., 2021. Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula. Curr. Biol. 31, 2576-2591.e12. https://doi.org/10.1016/j.cub.2021.04.022

Savitzky, A., Golay, M.J.E., 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Anal. Chem. 36, 1627–1639. https://doi.org/10.1021/ac60214a047

Schmidt, F.N., Zimmermann, E.A., Campbell, G.M., Sroga, G.E., Püschel, K., Amling, M., Tang, S.Y., Vashishth, D., Busse, B., 2017. Assessment of collagen quality associated with non-enzymatic cross-links in human bone using Fourier-transform infrared imaging. Bone 97, 243–251. https://doi.org/10.1016/j.bone.2017.01.015

Schwarz, C., Debruyne, R., Kuch, M., McNally, E., Schwarcz, H., Aubrey, A.D., Bada, J., Poinar, H., 2009. New insights from old bones: DNA preservation and degradation in permafrost preserved mammoth remains. Nucleic Acids Res. 37, 3215–3229. https://doi.org/10.1093/nar/gkp159

Shoulders, M.D., Raines, R.T., 2009. Collagen structure and stability. Annu. Rev. Biochem. 78, 929–958. https://doi.org/10.1146/annurev.biochem.77.032207.120833

Singh, B.R., DeOliveira, D.B., Fu, F.-N., Fuller, M.P., 1993. Fourier transform infrared analysis of amide III bands of proteins for the secondary sctructure estimation. Biomol. Spectrosc. III 1890, 47–55. https://doi.org/10.1117/12.145242

Sirak, K., Fernandes, D., Cheronet, O., Harney, E., Mah, M., Mallick, S., Rohland, N., Adamski, N., Broomandkhoshbacht, N., Callan, K., Candilio, F., Lawson, A.M., Mandl, K., Oppenheimer, J., Stewardson, K., Zalzala, F., Anders, A., Bartík, J., Coppa, A., Dashtseveg, T., Évinger, S., Farkaš, Z., Hajdu, T., Bayarsaikhan, J., McIntyre, L., Moiseyev, V., Okumura, M., Pap, I., Pietrusewsky, M., Raczky, P., Šefčáková, A., Soficaru, A., Szeniczey, T., Szőke, B.M., Van Gerven, D., Vasilyev, S., Bell, L., Reich, D., Pinhasi, R., 2020. Human auditory ossicles as an alternative optimal source of ancient DNA. Genome Res. 30, 427–436. https://doi.org/10.1101/gr.260141.119

Sosa, C., Vispe, E., Núñez, C., Baeta, M., Casalod, Y., Bolea, M., Hedges, R.E.M., Martinez-Jarreta, B., 2013. Association between ancient bone preservation and DNA yield: A multidisciplinary approach. Am. J. Phys. Anthropol. 151, 102–109. https://doi.org/10.1002/ajpa.22262

Squires, K., Booth, T., Roberts, C.A., 2019. The Ethics of Sampling Human Skeletal Remains for Destructive Analyses, in: K., Squires, D., Errickson, Márquez-Grant, N. (Ed.), Ethical Approaches to Human Remains. Springer International Publishing, Cham, pp. 265–297. https://doi.org/10.1007/978-3-030-32926-6_12

Stani, C., Vaccari, L., Mitri, E., Birarda, G., 2020. FTIR investigation of the secondary structure of type I collagen: New insight into the amide III band. Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 229, 118006. https://doi.org/10.1016/j.saa.2019.118006

Surewicz, W.K., Mantsch, H.H., 1988. New insight into protein secondary structure from resolution-enhanced infrared spectra. Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol. 952, 115–130. https://doi.org/10.1016/0167-4838(88)90107-0

Talamo, S., Fewlass, H., Maria, R., Jaouen, K., 2021. “Here we go again”: the inspection of collagen extraction protocols for 14C dating and palaeodietary analysis. Sci. Technol. Archaeol. Res. 7, 62–77. https://doi.org/10.1080/20548923.2021.1944479

Titus, J., Ghimire, H., Viennois, E., Merlin, D., Unil Perera, A.G., 2018. Protein secondary structure analysis of dried blood serum using infrared spectroscopy to identify markers for colitis screening. J. Biophotonics 11. https://doi.org/10.1002/jbio.201700057

Turner-Walker, G., Syversen, U., 2002. Quantifying histological changes in archaeological bones using BSE-SEM image analysis. Archaeometry 44, 461–468. https://doi.org/10.1111/1475-4754.t01-1-00078

Unal, M., Jung, H., Akkus, O., 2016. Novel Raman Spectroscopic Biomarkers Indicate That Postyield Damage Denatures Bone's Collagen. J. Bone Miner. Res. 31, 1015–1025. https://doi.org/10.1002/jbmr.2768

Usai, D., Salvatori, S., Iacumin, P., Di Matteo, A., Jakob, T., Zerboni, A., 2010. Excavating a unique pre-Mesolithic cemetery in Central Sudan. Antiquity 84, 16–18.

Usai, D., Salvatori, S., Jakob, T., David, R., 2014. The Al Khiday Cemetery in Central Sudan and its "Classic/Late Meroitic" Period Graves. J. African Archaeol. 12, 183–204. https://doi.org/10.3213/2191-5784-10254

Vyskočilová, G., Ebersbach, M., Kopecká, R., Prokeš, L., Příhoda, J., 2019. Model study of the leather degradation by oxidation and hydrolysis. Herit. Sci. 7, 1–13. https://doi.org/10.1186/s40494-019-0269-7

Weiner, S., 2010. Microarchaeology. Beyond the visible archaeological record., first edit. ed. Cambridge University Press, UK.

Weiner, S., Bar-Yosef, O., 1990. States of preservation of bones from prehistoric sites in the Near East: A survey. J. Archaeol. Sci. 17, 187–196. https://doi.org/10.1016/0305-4403(90)90058-D

CHAPTER 5

Life and death of Saint Theobald (1033-1066 AD) from chemical and forensic-anthropological analysis

Cinzia Scaggion[1]*, Cicero Moraes[2], Luca Bezzi[2], Alberto Zanatta[3], EugeniaD'Atanasio[4], Luca Nodari[5], Luca Pagani[6,7], Gilberto Artioli[1], Nicola Carrara[3]

[1]Department of Geosciences, University of Padua, Italy
[2]Arc-Team srl, Trento, Italy
[3]University Museums Centre - CAM, University of Padua, Italy
[4]Sapienza, University of Rome, Italy
[5]Italian National Research Council-CNR, ICMATE, Italy
[6]Department of Biology, University of Padua, Italy
[7]Institute of Genomics, University of Tartu, Estonia

*corresponding author, e-mail address: cinzia.scaggion@phd.unipd.it (C. Scaggion)

Keywords: repiques; FTIR spectroscopy; bone preservation; leprosy; aDNA; facial reconstruction

ABSTRACT

Saint Theobald (Saint-Thibaut, 1033-1066 Provins, France) was an important medieval historical figure of Catholicism. The relevance of this Saint is attested by the statue dedicated to him (n.139) and located above Bernini's Colonnade in piazza San Pietro, Vatican City. Two hundred years before Saint Francis of Assisi (1181-1226), Theobald, coming from a noble family, was the first to embrace poverty and lived as a hermit and pilgrim. According to historical records, he died of leprosy (Mycobacterium leprae), a disease widespread in Europe at the time, at the age of 33. Some of his bones found their way from France as holy relics to the parish church of Saint Giovanni Battista in Badia Polesine, Italy. With the discovery of his remains, a multidisciplinary investigation was conducted to go beyond authentication and historical analysis to gain as much information as possible about his life, death and appearance. Results enabled us to confidently associate the human remains with Saint Theobald. Furthermore, forensic 3D cranio-facial reconstruction, in concordance with anthropological analysis, allowed us to reliably reconstruct the face, appearance and features of Theobald.

INTRODUCTION

The cult of relics in Europe, where parts of bodies and personal objects of Catholic saints (e.g. rags, bones, dried blood etc.) played an important role in the popular religion, has its first tangible manifestations from the middle of the 2nd century to the Middle Ages. Relics epitomized the intimate intersection between religion and magic that provoked dramatic oppositions throughout history. In the 15th century AD the protestants and their heretical precursors tried to abolish the veneration of relics (later legitimized with the Council of Trent 1545-1563) (Walsham, 2010). Later, during the French Revolution and Napoleonic era, religious goods were confiscated because of anti-religious ideas which led to the destruction of all things relating to the practice of Catholicism (leading to looting or destruction of shrines and their contents) (Hahn, 2017). Nowadays, although many relics are being

preserved in churches, abbeys, or museums, their authenticity is questioned as much as their religious meaning. The stealing of relics by bone traders was common practice in the Medieval Ages and after, which resulted in large number of counterfeit holy bones circulating as commercial goods, often leading to cases of bone falsification, thefts etc. (Geary, 1991; Peter Manseau, 2009).

Relics were sought after by tomb robbers thanks to the spiritual blessings which were attributed to them. This included the reduction of years in Purgatory for sinners and material advantages. Therefore, relics were a good reason to pillage Holy Land sepulchres and Paleochristian sepulchres in Rome. Shrines can often contain bones of similar or very different age than those of the Saint they are ascribed to, or they enclose a mixed assemblage of skeletal elements of diverse origin and age and sometimes even species (Nilsson et al., 2010). The study of the skeletal remains of Saints was typically restricted to the counting and identification of bones, sex-age determination and approximate palaeopathological analyses (Zampini and Corrain, 1975). In most cases, neither ancient DNA analysis nor absolute dating were attempted.

The aim of our project is not only to test authenticity and study the historical context in general, but also to extract as much information as possible on his life, death and appearance. The recognition of the skeletal remains attributed to Saint Theobald (1033-1066) located in the church of Saint Giovanni Battista in Badia Polesine, province of Rovigo, Italy, is the tenth in history and the second conducted by the University of Padua, after a project led by the anthropologist Cleto Corrain in 1972 (Zampini and Corrain, 1975). The remains of Saint Theobald are of particular interest from a scientific point of view as they are associated with a substantial series of historical and biographical events surrounding his remains. Saint Theobald (Saint-Thibaut) was born in 1033 in Provins, France from a noble family. Two hundred years before Saint Francis of Assisi (1181-1226), Theobald embraced poverty and lived as a hermit and pilgrim. On his way to the Holy Land he founded the Camaldolese community of St. Romuald in North-Western Italy. According to the historical record, he died of leprosy a disease also known as Hansen's disease and caused by the bacterium *Mycobaterium leprae*, which was widespread in Europe at the time, in 1066 at the age of 33. His canonization took place just seven years after his death by Pope Alexander II (1010/15-1073). In 1074, his body was moved from Vicenza cathedral to the Abbey of Vangadizza in Badia Polesine. On this occasion, some of the bones found their way to France as holy relics, probably via his brother Arnolf. After Napoleon's arrival in Italy in 1796, the relics were transferred to the parish church of Saint Giovanni Battista in Badia Polesine and have since been kept in a Carrara marble urn, which was placed in an altar dedicated to him. According to historical sources Saint Theobald's body never underwent *post-mortem* alterations related to the burial environment(Soffiantini, 2014). Our multidisciplinary investigation of his remains involved the following: (1) physical-anthropological analyses to define the biological profile of the Saint; (2)

paleopathological analyses to search for macroscopic traces of the reported disease; (3) micro-sampling of the petrous bones to investigate bioapatite recrystallization and collagen alteration through Fourier-transform infrared spectroscopy (FTIR), for the purposes of radiocarbon dating; (4) genomic analysis to define the Saint's genetic profile and detect *Mycobacterium leprae* or other pathogens afflicting the Saint; (5) the roots of the second lower molar were sampled for radiocarbon dating to confirm the age of the remains; (6) the cranium was imaged at high resolution using advanced treatment of photogrammetric techniques such as SFM (structure from motion) and MVS (multiple-view stereovision), to build a 3D virtual model of the skull, which will allow for the reconstruction of the Saint's face.

METHODS

Once the box containing the relics was reopened, in the presence of a canonical Commission, established by Bishop Pierantonio Pavanello, their inspection started and all the appropriate reliefs were carried out which will be used for the reconstruction of the face. First, this document examines osteological remains according to current anthropological methods.

Age determination

To classify the adult skeleton, the recommendations suggested by Ferembach et al. (1980) and Buckberry and Chamberlain, (2002) were followed, and the multifactorial approach was preferred (Lovejoy et al., 1985a). Then, the degree of cranial sutures synostosis (Meindl and Lovejoy, 1985), dental wear (Brothwell, 1981), morphological changes of the pubic symphysis (Meindl et al., 1985; Krogman and Iscan, 1986), and metamorphosis of auricular surface of the ilium (Lovejoy et al., 1985b) were employed.

Sex determination

With adult skeleton the classic criteria for sex and age diagnosis proposed by Ferembach et al. (1980), Sjøvold, (1990), Buikstra and Ubelaker, (1994), Bass, (1995) and Walker, (2005) were adopted.

Palaeopathological methodology

According to historical sources, the Saint died of leprosy. To palaeopathological signs identification ofleprosy-related lesions caused by *Mycobacterium leprae*, diagnosis of pathological conditions were determined following the recommendations of Ortner, (2003); Capasso et al. (1999); Baxarias J. and Herrerin, (2008); Robbins et al. (2009); Köhler et al. (2017); Costa et al., 2009; Møller-Christensen, (1961); Fornaciari and Giuffra, (2009); Armelagos et al. (2014).

Microsampling

To increase the potential for pathogenic aDNA retrieval among the available osteological samples, petrous bone and a tooth were collected. Despite being indicated in a recent article that petrous bones do not contain bloodborne pathogens (Margaryan et al., 2018), both were sampled in order to perform a comparative analysis between the two anatomical elements.

Teeth are more exposed than petrous bones to bacteria during a lifetime due to rich dental vascularization but also more directly to exogenous contamination (Hansen et al., 2017).

Inspired by previous work (Sirak et al., 2017), we have applied an alternative method for extracting a micro-core (20-30 mg) from the bone for Infrared spectroscopy (FTIR) and a micro-core of 100-150 mg for palaeopathological molecular analysis. The procedure of micro-cores extraction was carried out on both petrous bones of the Saint's cranium with Dremel Micro Rotary drilling using two different micro-tools. The external bone surfaces, more exposed to the external biodeteriogenous agents (diagenesis), were mechanically removed by means of low-speed micro-drill, equipped with an abrasive round point of 2.4 mm. In complete cranium (where the petrous bone is not easily accessible) a micro-core with an internal diameter of 1.8 mm was used.

For sampling procedure particular attention was paid in order to avoid overheating the bones during the mechanical extraction process. The sample fragment was thus extracted without being contaminated and stored in sterilized Eppendorf. The ICMP (International Commission on Missing Persons ICMP 2015) standard operating procedure for bone sampling was followed to disinfect the drill bit to avoid external contamination, thus invalidating the analyses. A control sample with no osteological lesions was taken from the lower second molar root on the right mandible. The roots of the left M_2 molar were cut with a cutting disc wheel (diameter 20 mm) obtaining a sample of 2.856 g of which half intended for radiocarbon analysis and the other one palaeopathological molecular analysis.

Infrared Spectroscopy: sample preparation and spectrum analysis

From the petrous bones was extracted enough micro-core (20-30 mg) to execute FTIR analysis. One pellet has been prepared maintaining a ratio 1:100 mg of sample/KBr with constant grinding for 2 minutes. Powder obtained was then pressed, using a hydraulic press, under 11 tons/cm^2 pressure, and a 12 mm in diameter and 1.5 mm of thickness transparent pellet was obtained. Spectrum was collected with a Nicolet 380 FTIR spectrometer equipped with a DTGS detector; 128 scans were acquired, in the range from 4000 to 400 cm^{-1}, with a spectral resolution of 4 cm^{-1}. Spectral analysis was performed using Omnic 9 software (Thermo Scientific). To semi-quantifying of carbonate and amide presents in the spectrum, the ratios have been calculated dividing the intensity of absorption by the intensity of υ_3 phosphate peak at 1305 cm^{-1}. In spectral analysis, the peak heights are calculated from baselines defined by two points calculated as local minimum ranging in selected regions of spectrum. Amide, υ_3 phosphate and carbonate groups absorptions are 2000–1800/1400–1200 cm^{-1}, 1400–1200/900–750 cm^{-1} and 891–816 cm^{-1}, respectively. To define IRSF and FW85% values, the baseline has been defined by two points in the region 850-620 and 510-470 cm^{-1} to calculate the phosphate absorption heights in the region of spectrum at 604-565 cm^{-1} (see **Figure.1** and **Table.1**).

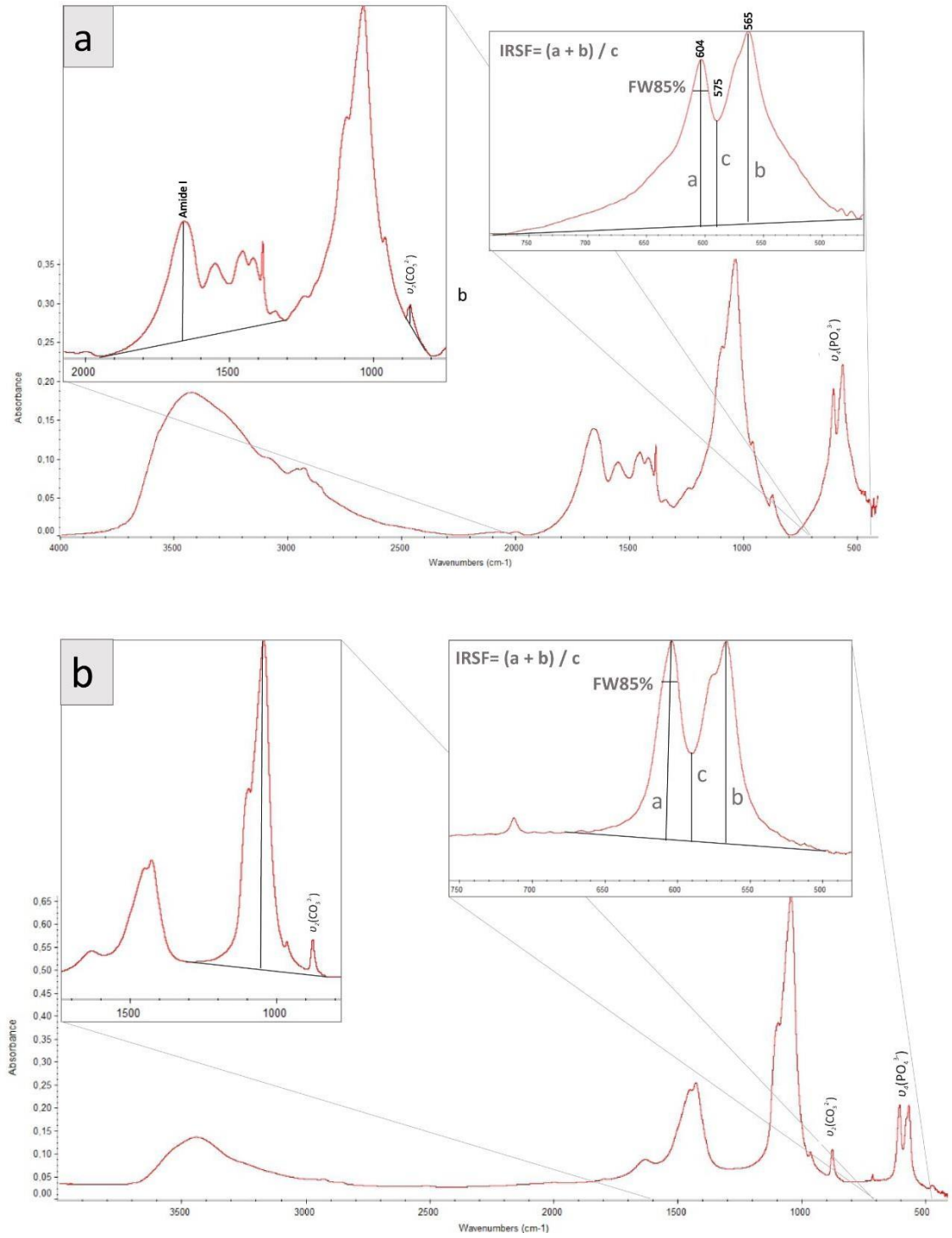

Figure 1. *Diagenetic parameters calculated from FTIR spectrum and referring to: bone crystallinity (IRSF); structural carbonate content in bioapatite (CO_3^{2-}/PO_4^{3-}); collagen preservation (AmI/ PO_4^{3-}). (adapted from references Dal Sasso et al., 2018)*

Table 1. *Wavenumbers and baselines adopted to calculate peaks intensity and width of petrous bone sample properties by FTIR spectroscopy.*

Vibrational mode and functional group	Wavenumber (cm^{-1})	Baseline (cm^{-1})
Amidel υ(C=O)stretch	1660	2000–1800/1400–1200
$\upsilon_3(PO_4^{3-})$ antisymmetric stretching	1035	1400–1200/900–750
$\upsilon_2(CO_3^{2-})$ out of plane bend vibration	875	891–816
$\upsilon_4(PO_4^{3-})$ bend	604	850–620/510–470
$\upsilon_4(PO_4^{3-})$ bend	565	850–620/510–470

The presents of sub-bands signals of Amide I to calculate the denaturation state of collagen, has been estimated by employing of second-derivative analysis (calculated using Savitsky-Golay smoothing technique (Savitzky and Golay, 1964) with 9 smoothing points) with baseline previously defined. OMNIC® software was used to convert the FTIR absorbance spectra into second-derivatives to determine the number and position of components corresponding to the Amide I region for the curve-fitting process. Sub-bands and shape were considered Gaussian and the baseline was defined by two points calculated as the local minimum in the region 1720-1840 and 1600-1590 cm^{-1}. The position of the sub-band components at 1699,1691,1685,1663,1652,1636 cm^{-1} and 1699,1693,1682,1662,1648 and 1636 cm^{-1} for fresh and ancient sample were fixed, whereas their band widths could be adjusted to perform the curve-fitting of the Amide I profile (**Figure.2**). The area of ~1660 and ~1690 cm^{-1} was calculated and normalized with phosphate peak at 1035 cm^{-1}.

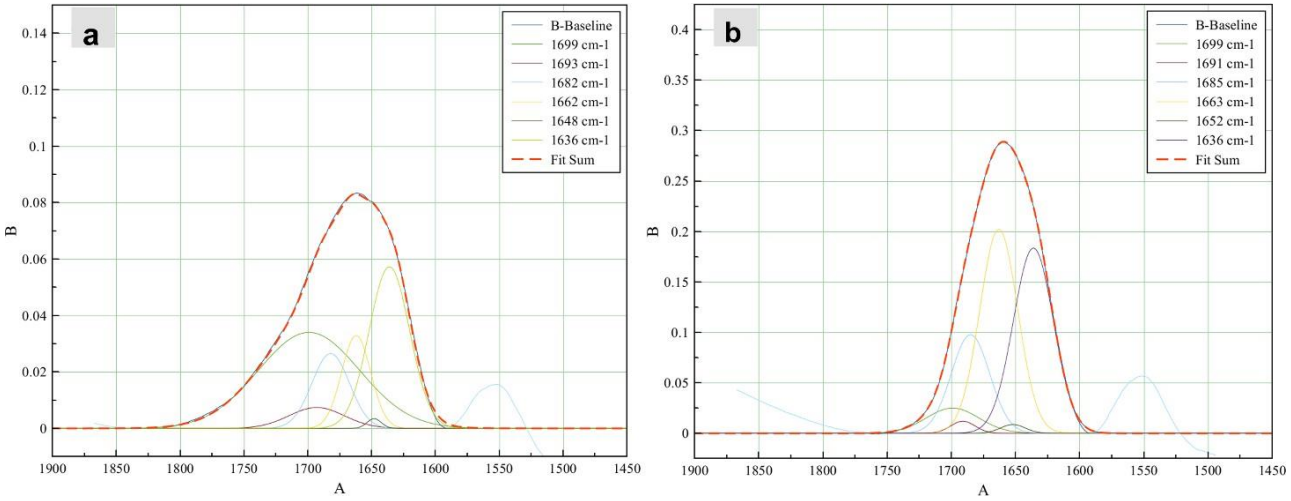

Figure 2. *Six sub-bands of Amide I and their positions were determined from the second derivatives of Saint Theobald (**a**) and from fresh bone (**b**). Sub-bands and shape were considered Gaussian. Two points calculated as the local minimum in the region 1720-1840 cm^{-1} and 1600-1590 cm^{-1}, defined the baseline. The position of the sub-band components of Amide I has been fixed, whereas their bandwidths were been adjusted to perform the curve fitting of the Amide I profile with a R^2=0.99.*

Photogrammetric acquisition and digital modelling

Photogrammetry techniques and, more precisely, Structure from Motion - Multiple-View Stereovision (SfM-MVS) (Bezzi et al., 2011), were used to create a high-resolution, three-dimensional virtual model of the skull, which became the basis for the 3D reconstruction of the face. Precise metrical models were generated using several photo sets, mainly taken with oblique photography, using a Nikon D800 Reflex (DSLR) camera, mounted on a tripod, with a CMOS sensor (36.3 Mp) and SIGMA 24 mm F 1.4 lens. Photo-acquisition was performed following the norms of convergent close-range Photogrammetry, adapting aperture value, shutter speed seconds and ISO to the low-light conditions of the church. Due to logistical problems, no graduated turntable, light box, or fixed light lamp were used. In general, when possible, photos were collected by maintaining the camera with a fixed angle both on the horizontal and the vertical axis of the object, while lighting remained constant during the image data capture. Metric and colorimetric values have been restored and calibrated thanks to fixed references used for the acquisition. Overall, during the data collection stage, 635 photos were taken. Once accomplished the work on the field, the acquired photographs were processed to generate a 3D model of the two bone elements (cranium and mandible) with the openMVG101 photogrammetric reconstruction open-source software, within ArcheOS100 Linux distribution. The polygonal lattice that defines the virtual object (3D mesh) was processed by MeshLab102, selecting the photographic sets from which the software extrapolated the dense point cloud (i.e. a set of points on the space defined by spatial coordinates and intensity values). Initially, the 3D scanned skull was imported, the mandible was segmented, adjusting its occlusion. The piece was aligned with the Frankfurt plane, a cephalometric plane that follows the virtual horizon, and touches the lowest part of orbitals (right and left) and highest part of the porios (right and left). The Frankfurt plane simulates the natural position of the head in a standing living people (Taylor, 2000). After this process, the facial reconstruction/approximation of Saint Theobald was carried out completely in a digital environment using the ForensicOnBlender95 add-on, a set of tools based on free and multiplatform software, available for use on Linux, Windows and Mac OS X operating systems. The machine used for this purpose was a desktop computer equipped with an Intel Core i9-9900K processor, with 64 GB of RAM, 8 GB nVidia GeForce 2070 GPU, 960 GB SSD storage, water cooling and Linux 3DCS operating system based on Linux Ubuntu© 20.04. The three-dimensional reconstruction started from the anthropological examination of the skull, from which the basic information on the age of death, gender and ancestry were obtained. Through the anthropological examination, the 3D model of the facial reference to guide the final steps of the reconstruction was chosen, following the protocol already used during previous forensic projects (Zanatta et al., 2018).

RESULTS AND DISCUSSION

Palaeopathological interpretation and differential diagnosis

It was possible to ascertain that the body preserved inside the urn was an adult male, who died around 30-35 years of age. Accordingly, the results of our anthropological analysis agree with what is known of the Saint, who died at 33 years old (for additional information see Supplementary Information). Molecular analysis to determine the presence of *M. leprae* yielded negative results. However, a false negative result cannot be excluded, as the DNA yield was low and the detection of *M. leprae* ancient DNA (aDNA) can be difficult even in individuals showing clear signs of the disease on their skeleton. For this reason, we relied exclusively on the observations of pathological lesions found on the skeleton to determine the presence of the disease. In literature, the osteoarchaeology characteristic lesions associated with leprosy can be found on nasal bones, nasal septum and *conchae bullosa*, hard palate, zygoma, maxilla, phalanges of hands and feet, metatarsals and metacarpals (Møller-Christensen, 1961; Ortner, 2003). Palaeopathological observation was focused on the cranium in absence of the distal portions of the skeleton, which are frequently involved in the disease. Some typical facial markers of leprosy can also occur in other diseases, such as syphilis, oral infections, maxillary sinusitis, leishmaniasis and tuberculosis(Köhler et al., 2017).

In syphilis cases, the most commonly affected areas are the cranial vault (in particular the external surface), tibiae, and nasal bones. Cranial vault lesions are produced by a combination of destruction and repairing of bone creating nodular cavity (*caries sicca*) (G. and Giuffra, 2009). Nasal bones, hard palate, turbinates and maxillary sinus are also involved with consequent remodelling with formation of sclerotic bone at the margins of the lytic lesions of the nasal septum and palate. One of the characteristic lesions of syphilis is the classical "*saber*" shape of the tibiae as well as osteomyelitis(Köhler et al., 2017; Møller-Christensen, 1961; Ortner, 2003). Based on our analyses of frontal bone and tibiae of the Saint showed no signs of nodular cavities, osteomyelitis, widespread periostitis or *saber* tibiae. On this basis, we could exclude syphilis as the cause of the lesions identified on the cranial vault of Saint Theobald.

Ante-mortem loss of upper incisors is frequent in leprosy (Møller-Christensen, 1961). In this case and those from the Copper Age Necropolis Abony-Turjányos dűlő site, located in Central Hungary (Köhler et al., 2017) or in some cases of Medieval graves in Denmark (Boldsen, 2009; Møller-Christensen, 1961), this phenomenon has not been observed. The severe atrophy of the alveoli, porosity of the hard palate and mandibular periostitis observed on the remains of the Saint may have been caused by an oral infection, while the circular depression in the infraorbital region of the maxilla could be the result of a rhinomaxillary sinusitis. Another disease that could involve the *splanchnocranium* (facial skeleton) is a parasitic infectious disease carried by sand flies of genus *Leishmania*. Sand flies

infections can lead to *Mucocutaneous leishmaniasis* which causes destruction of the nasal septum, palate, mouth or throat and particularly causes severe breathing problems and profound aesthetic damage (Köhler et al., 2017; Ortner, 2003; Robbins et al., 2009). The parasitic infection does not cause injury to long bones and leads to destruction of the facial bones without involving the piriform aperture which remains a peculiarity specific to leprosy (Costa et al., 2009).

Lastly in cases of tuberculosis, the locations in which lesions are known to arise mainly involve the vertebrae and ribs. In adult individuals, the long bones are most often affected at the metaphyseal and epiphyseal levels. The skull is rarely involved but can carry lesions on parietal and frontal bones in sub-adult cases (Ortner, 2003). Destruction caused by abscesses on the petrous bone and the mastoid process, as well as the inferior orbital margin and the zygomatic bone, can also occur in sub-adults with a periostitic reaction. In tuberculous meningitis, the cranial vault is the most frequent site of lytic lesions, but it can also cause apposition and remodelling of endocranial tissue and can affect the nasal cavity (Ortner, 2003; Spekker et al., 2020). The infection can also extend to the nasal bones and maxilla. The post-cranial lesions can affect the joints. The hip is the region most affected by tuberculosis. It too is mainly affected in childhood. Erosion, periostitis and osteophytic growth normally focus on the acetabulum, head and femoral neck, leading to the dislocation of the latter resulting in ankylosis. St. Theobald is an adult and his osteological remains do not demonstrate any lesions consistent with tuberculosis with the exception of lesions attributable to nutritional stress and occupational markers such as: porosities of the large wings of the sphenoid (Armelagos et al., 2014), bilateral *cribra femora* (Garcia et al., 2002), Schmörl's hernia, vertebral osteoarthritis of the upper articular facets, bilateral pectineus enthesopathy of coxals, osteochondritis dissecans on the head of femur (Baxarias J. and Herrerin, 2008; Capasso et al., 1999).

Infrared Spectroscopy analysis as bone quality indicator for radiocarbon dating

Pre-screening of bone quality and its organic component can contribute to reducing the costs of radiocarbon analyses. Radiocarbon dating (^{14}C dating) from Type I collagen (Gâza et al., 2019; Gelse et al., 2003) of bones, provides important information in archaeological contexts, especially when dealing with funerary deposits lacking archaeological artefacts that can be used for relative dating. Diagenetic changes in bone and the resulting loss of collagen (Collins et al., 2002) content can give inaccurate radiocarbon results (Calabrisotto et al., 2013; Hedges, 1992; Hedges and Law, 1989; Nielsen-Marsh and Hedges, 2000). In addition, ^{14}C dating with low amounts and poor quality of collagen can be time consuming and associated with the risk of getting no results.

For this purpose, the exploration of sample variability was based on Fourier-transform infrared spectroscopy (FTIR), a technique used for assessing the chemical composition and mineral alteration of ancient bone samples at molecular level (Paschalis et al., 2011). Infrared spectroscopy is an advantageous technique since it provides a fast, convenient, and accurate semi-quantitative approach to investigate mineralized tissues. IR absorption bands are characteristic of specific types of chemical bonds, enabling descriptions of variation of structural properties of organic and inorganic components (Carden and Morris, 2000; Goormaghtigh et al., 2006; Pelton and McLean, 2000).

Data collection

The spectrum analysed (see **Figure.3**) shows signals that can be attributable to the amide, phosphates and carbonate groups. According to the literature (Singh et al., 1993; Stani et al., 2020; Weiner, 2010), the absorption at ~3080 cm^{-1} corresponds to Amide B and the ones at 1660, 1550, 1340 and 1240 cm^{-1} correspond to the Amide I-II-III vibrational bands respectively. The absorption at 875 cm^{-1} corresponds to the bending vibrational modes of carbonate B-type and at 1035, 604 and 565 cm^{-1} the stretching and bending of phosphates. To monitor the state of sample conservation several ratios were considered to provide a semi-quantitative measure of bone integrity based upon the height of phosphate, B-type carbonate and amide absorption bands. The peak heights are calculated from the baselines which are defined by two points calculated as the local minimum ranging in selected regions of the spectrum, reported in **Table.1**, **Figure.1** and defined in *Methods*. The ratio between peak intensity of amide and phosphate was calculated on phosphate peak at 1035 cm^{-1} assigned to the υ_3 (PO_4^{3-}) vibrational intensity. The mineral to matrix (AmI/PO_4^{3-}) ratio has been calculated to provide a more pertinent description as to whether the bone tissue is diagenetically altered or not since it considers the main components of the bone (organic and inorganic phase) and it is well known that its preservation depends on the quantity and quality of both (Boskey et al., 2003; Donnelly, 2011; Figueiredo et al., 2012; Naito et al., 2020).

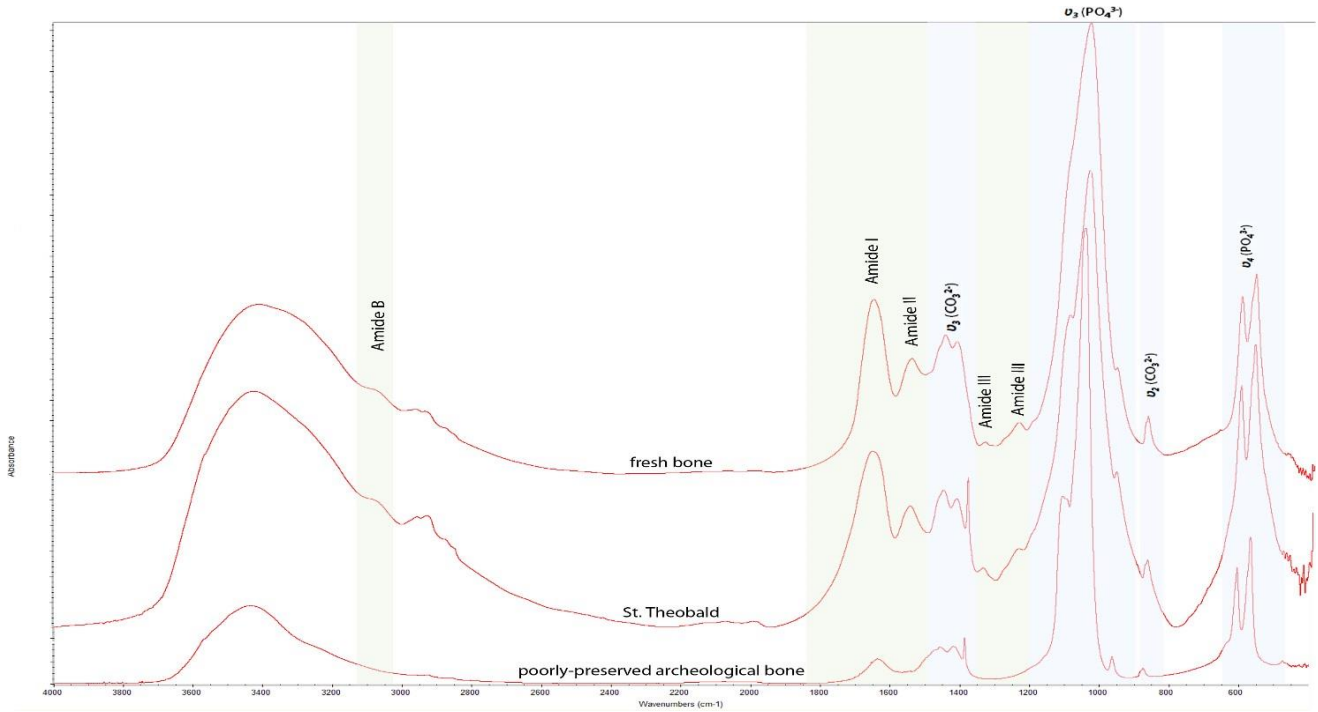

Figure 3. *Spectra measured in FTIR shows signals attributable to amide, carbonate and phosphate functional groups of bone. In this spectrum are evident organic (green) and inorganic (blue) components that make up the bone tissue. St. Theobald's spectrum is compared with fresh bone and poorly preserved archaeological bone samples. The figure clearly shows that St. Theobald's spectrum is more similar to a fresh bone than a strongly diagenized one.*

Bioapatite crystallinity indicators as splitting factor (IRSF), as defined by Weiner and Bar-Yosef (1990), and the width of the 604 cm^{-1} phosphate peak (defined as full width at 85% of maximum height - FW85%, defined by Dal Sasso et al.,(2018) was chosen and calculated. The IRSF is a widely used parameter to describe crystallinity variations together with of the phosphate (FW85%) peak at 604 cm^{-1} (Dal Sasso et al., 2018). They are commonly used to determine the general state of preservation of ancient bone.

The structural B-type carbonate content of bioapatite is monitored through the CO_3^{2-}/PO_4^{3-} parameter, calculated by dividing the intensity of the band at 875 cm^{-1} $\upsilon_2(CO_3^{2-})$ vibrational mode by the intensity of the $\upsilon_3(PO_4^{3-})$ vibrational mode (Grunenwald et al., 2014). This parameter provides the significant occurrence of carbonate ions in the hydroxylapatite structure and gives important information to reflect the level of carbonate substitution in the bioapatite crystal.

The carbonate ion is known to occupy two different positions in the hydroxylapatite bone structure in the hydroxide lattice position and that of phosphate (Sønju Clasen and Ruyter, 1997). The B-type carbonate (substitution of phosphate) is much more abundant than A-type (substitution of hydroxide) and therefore is assumed to produce reliable results in monitoring the carbonate content of bioapatite (Wopenka and Pasteris, 2005). The absorption υ_2 region band is attributed to carbonate B-type and is considered more useful for our study as it better reflects the real portions of carbonate principal

components without overlapping. In fact, υ_3 carbonate doublet at 1450-1423 cm^{-1} is more complex to analyze due to the broadening of both types and, furthermore, they tend to overlap others absorption bands of organic functional groups (Fleet, 2009; Rey et al., 1989). Molecular spectroscopy is useful to help infer the quality of collagen (Chadefaux et al., 2009; France et al., 2014; Ishida and Griffiths, 1993; Kontopoulos et al., 2020; Riaz et al., 2018; Singh et al., 1993). As described by Paschalis et al., (2001) (Paschalis et al., 2001), collagen cross-links provides the fibrillar collagen matrices with properties such as tensile strength and viscoelasticity; properties modified by diseases but also by interaction with the burial environment (Chadefaux et al., 2009). The ratio of peak areas of the mature non-reducible (interfibrillar) cross-links and immature reducible (intrafibrillar) cross-links of Amide I sub-bands of 1660 and 1690 cm^{-1} (1660:1690 cm^{-1}) provides a semi-quantitative measure of the enzymatic cross-linking in the organic matrix. According to the literature (Paschalis et al., 2001), the degree of cross-linking of collagen from immature to mature has been calculated from 1660: 1690 cm^{-1} area ratio which determines on collagen maturity. It provides a semi-quantitative measure of the cross-linking profile in the collagen matrix; this ratio has been correlated to the relative amounts of enzymatic cross-links, specifically mature non-reducible (interfibrillar) cross-links and immature reducible (intrafibrillar) cross-links. The majority of scientific community deems it is a valid parameter (Barth et al., 2011; Chadefaux et al., 2009; Mieczkowska et al., 2015; Paschalis et al., 2001; Sanden et al., 2021) that represents collagen maturity although its validation is still controversial above all the overlap structural water absorption interference (Farlay et al., 2011) at ~ 1645 cm^{-1} (H-O-H bending). For our study, this parameter can be considered an indicator of the degree of collagen alteration degree (Barth et al., 2011) and it was estimated by employing a combination of second-derivative spectra and peak finding to determine sub-peaks of Amide I region at ~1660 and ~1690 cm^{-1} (Paschalis et al., 2011, 2004, 2003). With band decomposition (see *Methods*), based on second derivative analysis, were found in the Amide I profile of collagen the six major known peak-positions components linked with aromatic ring vibrations of protein: 1699,1691,1685,1663,1652,1636 cm^{-1} to fit the Amide I band (Paschalis et al., 2011;Bryan et al., 2007; Chadefaux et al., 2009; Schmidt et al., 2017; Byler and Susi, 1986), while it was observed a shift of some sub-bands in ancient sample and there were determined the following sub-components at 1699,1693,1682,1662,1648,1636 cm^{-1}.

Infrared spectrum analysis results

St. Theobald's infrared spectrum was compared with that of a well-preserved fresh animal bone (**Figure.4**).

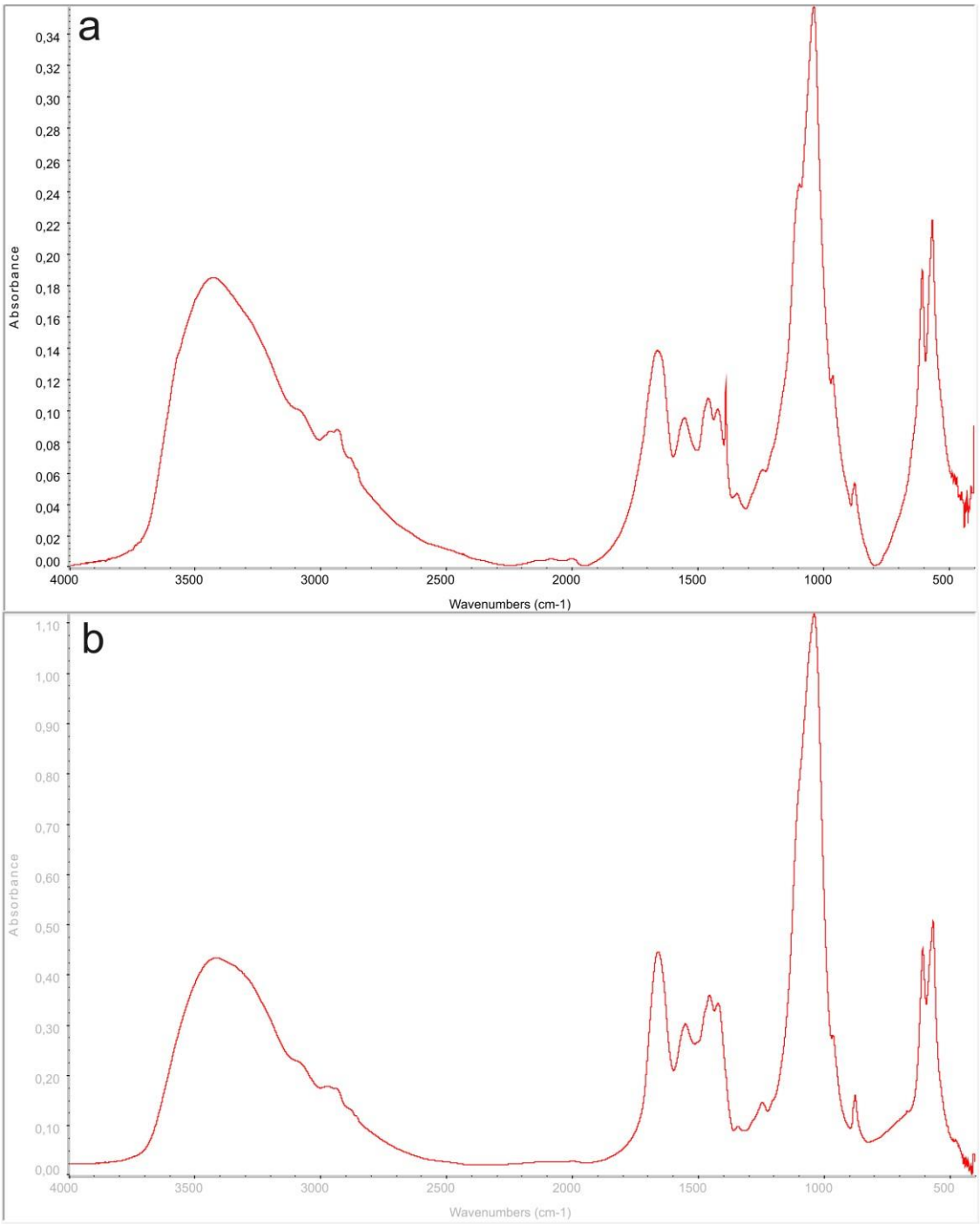

Figure 4. *IR spectra in comparison to observe diagenetic alternations (**a**) St. Theobald; (**b**) animal fresh bone.*

Based on FTIR spectra comparison we could observe that our samples showed noticeably different peaks which can be associated with altered bone.

We analysed the spectra and found peptide prominent bands of both Amide B (3080 cm^{-1}), Amide I (1660 cm^{-1}) and Amide II (1550 cm^{-1}) in addition to smaller peaks corresponding to Amide III (1240 cm^{-1}) and $\delta(CH_2)$ side chain vibration of collagen at 1341 cm^{-1} (Belbachir et al., 2009; Camacho et al., 2001; Mehra and Chadha, 2020; Singh et al., 1993) (see **Figure.3**).

The area ratio of Amide I underlying bands, 1660:1690 cm^{-1}, is 2.34 and 24.52 for St. Theobald and fresh bone with a model fit of R^2 =0.99, respectively. Values of mature non-reducible (interfibrillar) cross-links profile at ~1660 cm^{-1} are 2.77 and 6.96 while immature reducible (intrafibrillar) cross-links profile at ~1690 cm^{-1} are 1.18 and 0.28.

Mineral to matrix (AmI/PO_4^{3-}) indicator for determining the levels of preserved collagen(Naito et al., 2020) has been measured obtaining 0.35 and 0.39 values for St. Theobald and the ox bone (**Table.2**).

Table 2. *Results obtained for FTIR parameters calculated for all samples, i.e. the width at 85% of the height of the 604 cm^{-1} peak (FW85%), the carbonate to phosphate band intensity ratio (CO_3^{2-}/PO_4^{3-}) and amide to phosphate ratio (AmI/PO_4^{3-}), the infrared splitting factor (IRSF) and area ratio of Amide I underlying bands (1660:1690).*

BONE SAMPLE	AmI/PO_4^{3-}	FW85%(604 cm^{-1})	IRSF	CO_3^{2-}/PO_4^{3-}	1660 cm^{-1}	1690 cm^{-1}	1660:1690
St. Theobald	0.35	11.2	3.4	0.06	2.77	1.18	2.34
Fresh bone	0.39	12.5	3.0	0.07	6.96	0.28	24.52

Both spectra contain peaks corresponding to inorganic components of bones, namely carbonated – hydroxylapatite. We analysed the infrared spectrum from the St. Theobald sample and found that the increasing sharpening of the bands attributed to the signals related to bone bioapatite with a very pronounced sign at 1094 cm^{-1} (P=O symmetric bond stretching) and a less evident shoulder at ~638 cm^{-1} and at 3568 cm^{-1} assigned to O-H stretch of phosphates group(Pluduma et al., 2018; Stanislavov et al., 2018). Upon first observation of the spectral profiles, we found an increase in mineral maturity with a consequent increase in the size of the crystallites (defined as being inversely proportional to the FW85% of 604 cm^{-1} peak and proportional to the IRSF value), which is usually due to the loss of organic components(Farlay et al., 2010; Piga et al., 2016).

show a few variations of FW85% values, ranging from 12.5 to 11.2 and correspond to a minimal variation of IRSF values, ranging from 3.0 to 3.4 for fresh bone and St. Theobald, respectively. The CO_3^{2-}/PO_4^{3-} parameter, calculated for all samples, shows 0.07 value for fresh bone and 0.06 value for St. Theobald.

The bioapatite crystallinity helps to identify well-preserved bone with a similar IRSF value to fresh bone at around 2.7±0.2 (Asscher et al., 2012). The bioapatite samples analysed in this case of study do not show a high variability in terms of chemical and structural properties, and, therefore, do not seem to have suffered strong alteration due to recrystallization. Comparing the crystallinity index (IRSF) with the CO_3^{2-}/PO_4^{3-} ratio, we observed a negative correlation that can be associated with the loss of structural carbonate for archaeological bones occurring in the recrystallization processes (Hedges, 2002; Salesse et al., 2014). IRSF and CO_3^{2-}/PO_4^{3-} display a strong non-linear correlation with recrystallization indicating the loss of carbonate from the bioapatite structure resulting in a more stable

and less soluble mineral phase. In fact, based on our data, the decreased CO_3^{2-} content in the samples might have promoted the formation of a more stable mineral phase, i.e. crystal size and lattice perfection (Salesse et al., 2014).

Each of the amide bands mentioned above are found on the FTIR spectra of all the samples, and no obvious band shift was identified when compared to the spectrum of fresh collagen in the control animal bone sample. There is not an apparent difference between the fresh collagen and archaeological degradation products as directly observed from FTIR spectra.

Analyzing the ratio of peak areas of the sub-bands at ~1660 cm^{-1} and ~1690 cm^{-1} respectively, it has been shown that during collagen denaturation the relative intensity of the former sub-band decreases while the latter increases(Barth et al., 2011; Paschalis et al., 2004). Our analyses indicated that the denatured collagen has an inverse correlation with the strength of bone and a positive correlation with its elastic component.

In conclusion, our results indicate that the collagen contained in St. Theobald's remains are in a good state of preservation. Although the organic component has undergone denaturation (observable through cross-links ratio) of proteins resulting in the loss of water, which was possibly caused by microbial attack during the decomposition of the Saint's body (Xu, 2009), we could not observe a substantial reduction in the collagen content. The preservation of the organic structure is in strong relationship with the amount of collagen monitored through mineral to matrix (AmI/PO_4^{3-}) which ratio is indicative of a good amount of archaeological bone organic matter preservation (Naito et al., 2020). The amount of collagen preserved in the archaeological sample is comparable to that of the fresh bone, while the from analysis of cross-links parameters shows its quality of preservation where it is evident that a denaturation has occurred. The analysis at ^{14}C confirmed what had previously been estimated: the quantity of the collagen available was found to be equal to 9.7%, indicative of a good state of preservation. From a dental root sample at ^{14}C, it was possible to establish that the antiquity of the individual can be easily overlapped with historical data. In fact, radiocarbon records place the death of the individual between 988 AD-1163 AD, with a confidence level of 95%.

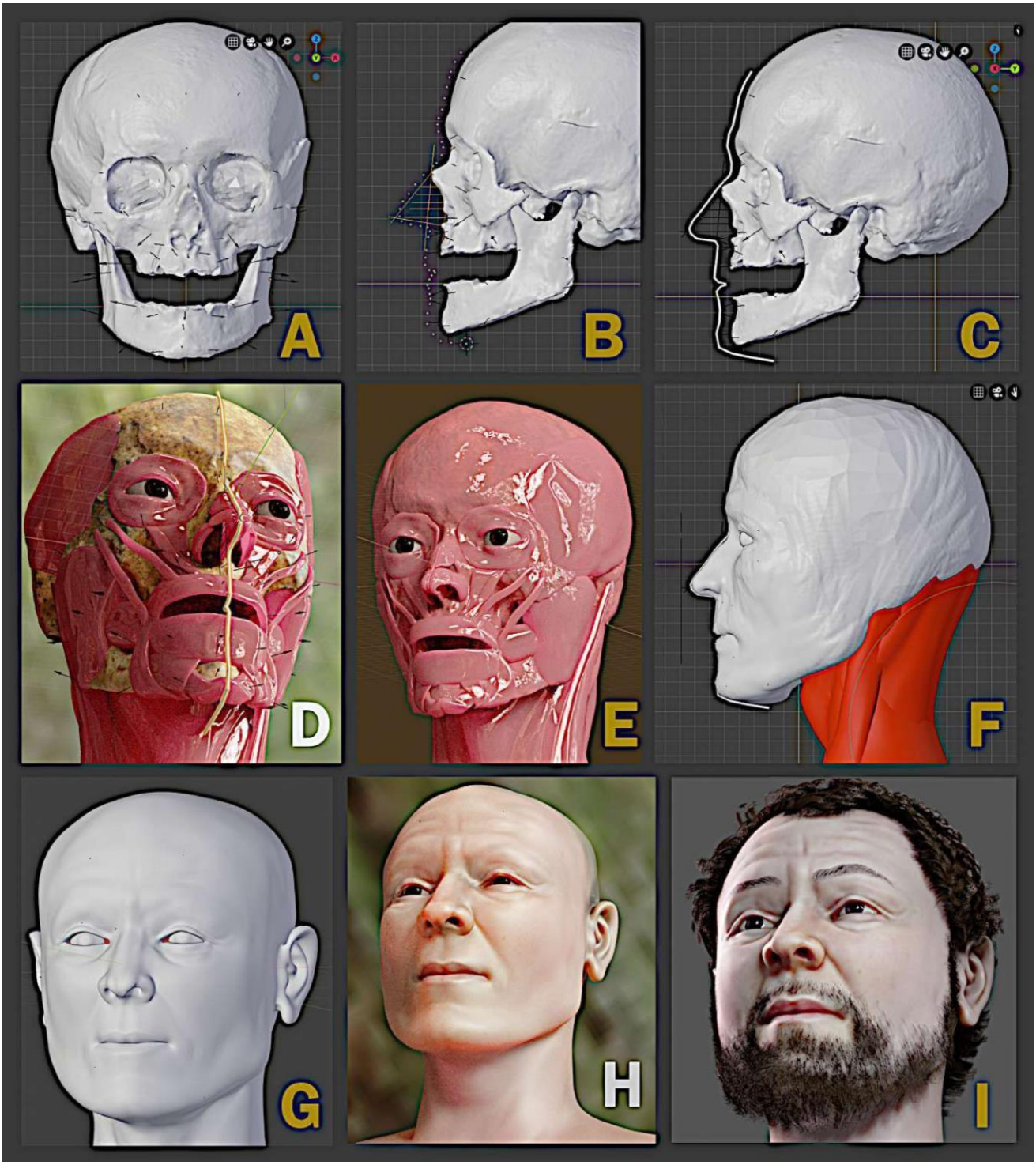

Figure 5. *Figure 5 from* ***a*** *to* ***i*** *shows the different procedures performed for facial reconstruction adopted.*

Forensic facial reconstruction is a technique for approximating a face from a skull with the aim of identifying or recognizing the reconstructed individual (Taylor, 2000; Wilkinson, 2004). Although the term reconstruction is widely used, the definition of approximation is more appropriately describes the process, since the final product results in a face that resembles the individual, thanks to its structural compatibility (Prag and Neave, 1997), providing enough material for a recognition that can be refined

after identification, by DNA, dental documents and other approaches (Baldasso et al., 2021). A series of soft tissue thickness markers have been defined (De Greef et al., 2006) according to the individual's age (**Figure.5a**). The soft tissue thickness markers function as small pins that indicate the limits of the face in the places where they are positioned. In this case, 31 markers were distributed, 10 central and 21 symmetrical, that is, 21 points to the left of the face and 21 points to the right, totaling 52 effective markers. These distances are averaged by ultrasound in 37 individuals of European ancestry and age in a range of 30-39 years old, with an average body mass index (20-25). Although these markers cover a considerable part of the face, some regions, such as the nose and lips, need projections in order to be approximated. The nose projection was done by crossing two techniques: Lebedinskaya and Manchester Method. The use of both techniques provides a clearer limit on the dimensions of the nose, avoiding, for example, possible distortions when projected on individuals with facial deformities. If there is a discrepancy between the two strokes, a drawing is chosen to be in the center of the two projections. It is also used to trace the lips using the sockets and teeth as a reference (**Figure.5b**) (Moraes et al., 2014). In order to facilitate the sculpting work, a profile line is drawn, which will serve as a limit for the posterior facial sculpture (**Figure.5c**). The eyeballs and the main muscles of the face, available in the software used, are distributed along the structure and suitable for it (**Figure.5d-5e**). Then the digital sculpture process begins, where an object called 3D mesh, a that resembles a mass is added to the scene and receives deformations in order to adapt to the face limits defined by the soft tissue markers, projections and profile line (**Figure.5f**) A 3D mesh is an entity formed by vertices, edges and faces that can be generated from a small cube to complex shapes such as a human face, a tree or an entire city in a 3D scene. The digital sculpture is covered by a simplified mesh and it is finished with the modeling of facial details (**Figure.5g**). The face mesh receives digital pigmentation and the eyes are adjusted according to the individual's characteristics as defined by the study of historical sources and the anthropological analysis, which estimated the ethnicity of the Saint's body (**Figure.5h**). The beard and hair are added using a generic setting (**Figure.5i**). It is noticeable that both the hair and the beard are short and uncombed as the arrangement of the strands contains a slight random alignment. The costumes are modeled after the historical survey, the scene received a series of lamps with values compatible with phosphorescent lights of 20 to 40 watts, distributed in order to favor the outline of the facial mass, then digital media such as images and videos were generated and the reconstruction process finished creating the final forensic facial reconstruction (**Figure.6**).

Figure 6. *Result of the forensic-facial 3D reconstruction. The individual's characteristics are defined by anthropological studies and historical sources.*

CONCLUSIONS

From the results obtained we can establish that the remains found in the parish church of Saint Giovanni Battista in Badia Polesine (Rovigo, Italy) is that of a man of 30-35 years, who probably died of leprosy around the 10^{th}-11^{th} century. The lesions 1 and 2 on the frontal bone (reported in *Supplementary Information,* ***Figure.S2***), the severe alveolar atrophy of the maxilla and mandible with consequent periostitic activity, the asymmetric atrophy of the pyriform aperture, the severe porosity of the nasal bones, the hard palate and the rhinomaxillary infection are all signs indicative of Hansen's disease. The porosities of the large wings of the sphenoid and the *cribra femora* suggest that the

individual was affected by nutritional deficiency. *Cribra femoris* is caused by systemic physiological stress since it is positively correlated with growth stunting and is more common in populations of low socioeconomic status with high rates of infectious disease (Garcia et al., 2002). It could be hypothesized that he had been suffering from a less aggressive form of leprosy such as the tuberculoid one. In conclusion, we can assume that the remains belong to the figure of Saint Theobald.

ACKNOWLEDGEMENTS

The study was strongly desired by the Ecclesiastical and local Community: the investigations are requested by Lions Club Badia Adige Po, CRAB-Circolo Ricreativo Anziani Badia, Workers' Society, Inter Club Association, Air Force Weapons Association, Association of Friends of Saint Theobald, Proloco, Center Youth Tourism and by the Diocese of Adria-Rovigo (Northeast Italy). We thank Dr. Christiana L. Scheib, Dr. Tina Saupe and Dr. Meriam Guellil of Institute of Genomics of Tartu University for valuable contributions for genetic analysis and interpretation of data. In addition, an acknowledgement goes to Dr. Christophe Snoeck Ph.D., researcher at Vrije Universiteit Brussel and the head of the Brussels Bioarchaeology Lab (BB-LAB) and to Dr. Ina Reiche Ph.D., research director of French National Centre for Scientific Research-CNRS for their valuable scientific advices.

ETHICS STATEMENT

No animal used was sacrificed for research purposes. Fresh bone was obtained from a slaughterhouse while archaeological specimen was provided through the approval by the ecclesiastical curia for the recognition of the Saint.

CONFLICT OF INTEREST

The authors do not have a conflict of interest to declare.

AUTHOR INFORMATIONS

Affiliations

Department of Geosciences, University of Padua, Italy
Cinzia Scaggion, Gilberto Artioli

Arc-Team srl, Trento, Italy
Cicero Moraes, Luca Bezzi

University Museums Centre - CAM, University of Padua, Italy
Alberto Zanatta, Nicola Carrara

Institute of Genomics, University of Tartu, Estonia
Meriam Guellil, Luca Pagani

Sapienza, University of Rome, Italy
Eugenia D'Atanasio

Italian National Research Council-CNR, ICMATE, Italy
Luca Nodari

Department of Biology, University of Padua, Italy
Luca Pagani

CONTRIBUTIONS

N.C. designed the project.; C.S., A.Z., N.C. carried out the anthropological studies; C.S., A.Z. carried out palaeopathological studies; C.S. analyzed palaeopathological data; C.S. carried out the FTIR analysis and C.S. and L.N. analysed the FTIR data; L.B. acquired photogrammetric data and C.M. carried out 3D facial reconstruction; M.G. carried out the aDNA investigation and E.D. genetic analysis; L.P., A.Z. funding acquisition for genetic analysis and radiocarbon dating; C.S. wrote original draft; M.G.,C.M.,G.A.,L.P. contributed to the manuscript; A.Z.,L.B.,C.M.,L.N.,G.A., supervision, all authors have written-reviewed & ended;

REFERENCES

Armelagos, G.J., Sirak, K., Werkema, T., Turner, B.L., 2014. Analysis of nutritional disease in prehistory: The search for scurvy in antiquity and today. Int. J. Paleopathol. 5, 9–17. https://doi.org/10.1016/j.ijpp.2013.09.007

Asscher, Y., Regev, L., Weiner, S., Boaretto, E., 2012. Atomic disorder in Fossil TooThand bone mineral: An FTIR study using the grinding curve method. ArcheoSciences 35, 135–141. https://doi.org/10.4000/archeosciences.3062

Baldasso, R.P., Moraes, C., Gallardo, E., Stumvoll, M.B., Crespo, K.C., Strapasson, R.A.P., Oliveira, R.N., 2021. 3D forensic facial approximation: Implementation protocol in a forensic activity. J. Forensic Sci. 66, 383–388. https://doi.org/10.1111/1556-4029.14587

Barth, H.D., Zimmermann, E.A., Schaible, E., Tang, S.Y., Alliston, T., Ritchie, R.O., 2011. Characterization of the effects of x-ray irradiation on the hierarchical structure and mechanical properties of human cortical bone. Biomaterials 32, 8892–8904. https://doi.org/10.1016/j.biomaterials.2011.08.013

Bass, W.M., 1995. Human Osteology: a Laboratory and Field Manual, 4th ed. ed. Columbia: Missouri Archaeological Society, University of Missouri.

Baxarias J., Herrerin, J., 2008. The handbook atlas of paleopathology. Zaragoza: Editorial Pórtico, Zaragoza.

Belbachir, K., Noreen, R., Gouspillou, G., Petibois, C., 2009. Collagen types analysis and differentiation by FTIR spectroscopy. Anal. Bioanal. Chem. 395, 829–837. https://doi.org/10.1007/s00216-009-3019-y

Bezzi, A., Bezzi, L., Ducke, B., 2011. Computer Vision e Structure from Motion, nuove metodologie per la documentazione archeologica tridimensionale: un approccio aperto. ARCHEOFOSS open source, Free Softw. e open format 103–111.

Boldsen, J.L., 2009. Leprosy in Medieval Denmark — Osteological and epidemiological analyses

Author (s): Jesper L . Boldsen Source : Anthropologischer Anzeiger , Vol . 67 , No . 4 , Highlights in Physical Anthropology : Eight Invited Reviews (December 2009), pp . 407-425. Anthropol. Anzeiger 67, 407–425.

Boskey, A.L., Moore, D.J., Amling, M., Canalis, E., Delany, A.M., 2003. Infrared Analysis of the Mineral and Matrix in Bones of Osteonectin-Null Mice and Their Wildtype Controls. J. Bone Miner. Res. 18, 1005–1011. https://doi.org/10.1359/jbmr.2003.18.6.1005

Brothwell, D., 1981. Digging up Bones. Oxford University Press, Ithaca, New York.

Bryan, M.A., Brauner, J.W., Anderle, G., Flach, C.R., Brodsky, B., Mendelsohn, R., 2007. FTIR studies of collagen model peptides: Complementary experimental and simulation approaches to conformation and unfolding. J. Am. Chem. Soc. 129, 7877–7884. https://doi.org/10.1021/ja071154i

Buckberry, J.L., Chamberlain, A.T., 2002. Age estimation from the auricular surface of the ilium: A revised method. Am. J. Phys. Anthropol. 119, 231–239. https://doi.org/10.1002/ajpa.10130

Buikstra, J.E., Ubelaker, D.H., 1994. Standards for Data Collection From Human Skeletal Remains. Fayetteville: Arkansas Archeological Survey Research Series No.44.

Byler, D.M., Susi, H., 1986. Examination of the secondary structure of proteins by deconvolved FTIR spectra. Biopolymers 25, 469–487. https://doi.org/10.1002/bip.360250307

Calabrisotto, C.S., Fedi, M.E., Caforio, L., Bombardieri, L., Mandò, P.A., 2013. Collagen Quality Indicators for Radiocarbon Dating of Bones: New Data on Bronze Age Cyprus. Radiocarbon 55, 472–480. https://doi.org/10.1017/S003382220005760X

Camacho, N.P., West, P., Torzilli, P.A., Mendelsohn, R., 2001. FTIR microscopic imaging of collagen and proteoglycan in bovine cartilage. Biopolym. - Biospectroscopy Sect. 62, 1–8. https://doi.org/10.1002/1097-0282(2001)62:1<1::AID-BIP10>3.0.CO;2-O

Capasso, L., Kennedy, K.A.R., Wilczak, C.A., 1999. Occupational Markers.

Carden, A., Morris, M.D., 2000. Application of vibrational spectroscopy to the study of mineralized tissues (review). J. Biomed. Opt. 5, 259. https://doi.org/10.1117/1.429994

Chadefaux, C., Le Hô, A., Bellot-gurlet, L., Reiche, I., 2009. Curve-Fitting Micro-Atr-Ftir Studies ofthe Amide I and Ii Bands of Type I Collagen in. e-PRESERVATIONScience 129–137.

Collins, M.J., Nielsen-Marsh, C.M., Hiller, J., Smith, C.I., Roberts, J.P., Prigodich, R. V., Wess, T.J., Csapò, J., Millard, A.R., Turner-Walker, G., 2002. The survival of organic matter in bone: A review. Archaeometry 44, 383–394. https://doi.org/10.1111/1475-4754.t01-1-00071

Costa, M.A., Matheson, C., Iachetta, L., Llagostera, A., Appenzeller, O., 2009. Ancient Leishmaniasis in a Highland Desert of Northern Chile. PLoS One 4, e6983. https://doi.org/10.1371/journal.pone.0006983

Dal Sasso, G., Asscher, Y., Angelini, I., Nodari, L., Artioli, G., 2018. A universal curve of apatite crystallinity for the assessment of bone integrity and preservation. Sci. Rep. 8, 1–13.

https://doi.org/10.1038/s41598-018-30642-z

De Greef, S., Claes, P., Vandermeulen, D., Mollemans, W., Suetens, P., Willems, G., 2006. Large-scale in-vivo Caucasian facial soft tissue thickness database for craniofacial reconstruction. Forensic Sci. Int. 159, S126–S146. https://doi.org/10.1016/j.forsciint.2006.02.034

Donnelly, E., 2011. Methods for assessing bone quality: A review. Clin. Orthop. Relat. Res. 469, 2128–2138. https://doi.org/10.1007/s11999-010-1702-0

Farlay, D., Duclos, M.E., Gineyts, E., Bertholon, C., Viguet-Carrin, S., Nallala, J., Sockalingum, G.D., Bertrand, D., Roger, T., Hartmann, D.J., Chapurlat, R., Boivin, G., 2011. The ratio 1660/1690 cm-1 measured by infrared microspectroscopy is not specific of enzymatic collagen cross-links in bone tissue. PLoS One 6. https://doi.org/10.1371/journal.pone.0028736

Farlay, D., Panczer, G., Rey, C., Delmas, P.D., Boivin, G., 2010. Mineral maturity and crystallinity index are distinct characteristics of bone mineral. J. Bone Miner. Metab. 28, 433–445. https://doi.org/10.1007/s00774-009-0146-7

Ferembach, D., Schwindezky, I., Stoukal, M., 1980. Recommendations for age and sex diagnosis of skeletons. J. Hum. Evol. 9, 517-549.

Figueiredo, M.M., Gamelas, J.A.F., Martins, A.G., 2012. Characterization of Bone and Bone-Based Graft Materials Using FTIR Spectroscopy. Infrared Spectrosc. - Life Biomed. Sci. https://doi.org/10.5772/36379

Fleet, M.E., 2009. Infrared spectra of carbonate apatites: ν2-Region bands. Biomaterials 30, 1473–1481. https://doi.org/10.1016/j.biomaterials.2008.12.007

France, C.A.M., Thomas, D.B., Doney, C.R., Madden, O., 2014. FT-Raman spectroscopy as a method for screening collagen diagenesis in bone. J. Archaeol. Sci. 42, 346–355. https://doi.org/10.1016/j.jas.2013.11.020

G., F., Giuffra, V., 2009. Lezioni di paleopatologia. ECIG- Edizioni Culturali Internazionali Genova, Genova.

Garcia, E., Berrocal, M.I., Baxarias, J., Campillo, D., Subirà, M.E., 2002. Cribra and trace elements in the Prat de la Riba Necropolis (Tarragona, Spain, 3rd-5th centuries AD). Antropol. Port. 19, 71–83. https://doi.org/10.14195/2182-7982_19_7

Gâza, O., Enachescu, M., Tută, C.S., Stavarache, C., Iovu, H., 2019. NMR characterization of bone collagen used for14c dating of osteological material. Rom. Reports Phys. 71, 1–7.

Geary, P.J., 1991. Furta Sacra: Thefts of Relics in the Central Middle Ages, Revisited. ed. Princeton University Press, New Jearsy.

Gelse, K., Pöschl, E., Aigner, T., 2003. Collagens - Structure, function, and biosynthesis. Adv. Drug Deliv. Rev. 55, 1531–1546. https://doi.org/10.1016/j.addr.2003.08.002

Goormaghtigh, E., Ruysschaert, J.-M., Raussens, V., 2006. Evaluation of the Information Content in

Infrared Spectra for Protein Secondary Structure Determination. Biophys. J. 90, 2946–2957. https://doi.org/10.1529/biophysj.105.072017

Grunenwald, A., Keyser, C., Sautereau, A.M., Crubézy, E., Ludes, B., Drouet, C., 2014. Revisiting carbonate quantification in apatite (bio)minerals: A validated FTIR methodology. J. Archaeol. Sci. 49, 134–141. https://doi.org/10.1016/j.jas.2014.05.004

Hahn, C., 2017. The Reliquary Effect: Enshrining the Sacred Object. Reaktion Books Ltd, London.

Hansen, H.B., Damgaard, P.B., Margaryan, A., Stenderup, J., Lynnerup, N., Willerslev, E., Allentoft, M.E., 2017. Comparing ancient DNA preservation in petrous bone and tooth cementum. PLoS One 12, 1–18. https://doi.org/10.1371/journal.pone.0170940

Hedges, R.E.M., 2002. Bone diagenesis: an overview of processes. Archaeometry 44, 319–328. https://doi.org/10.1111/1475-4754.00064

Hedges, R.E.M., 1992. A Review of Current Approaches in the Pretreatment of Bone. Radiocarbon 34, 279–291.

Hedges, R.E.M., Law, I.A., 1989. The radiocarbon dating of bone. Appl. Geochemistry 4, 249–253. https://doi.org/10.1016/0883-2927(89)90025-5

Ishida, K.P., Griffiths, P.R., 1993. Comparison of the amide I/II intensity ratio of solution and solid-state proteins sampled by transmission, attenuated total reflectance, and diffuse reflectance spectrometry. Appl. Spectrosc. 47, 584–589. https://doi.org/10.1366/0003702934067306

Köhler, K., Marcsik, A., Zádori, P., Biro, G., Szeniczey, T., Fábián, S., Serlegi, G., Marton, T., Donoghue, H.D., Hajdu, T., 2017. Possible cases of leprosy from the Late Copper Age (3780-3650 cal BC) in Hungary. PLoS One 12, e0185966. https://doi.org/10.1371/journal.pone.0185966

Kontopoulos, I., Penkman, K., Mullin, V.E., Winkelbach, L., Unterländer, M., Scheu, A., Kreutzer, S., Hansen, H.B., Margaryan, A., Teasdale, M.D., Gehlen, B., Street, M., Lynnerup, N., Liritzis, I., Sampson, A., Papageorgopoulou, C., Allentoft, M.E., Burger, J., Bradley, D.G., Collins, M.J., 2020. Screening archaeological bone for palaeogenetic and palaeoproteomic studies. PLoS One 15, 1–17. https://doi.org/10.1371/journal.pone.0235146

Krogman, W.M., Iscan, M.Y., 1986. The Human Skeleton in Forensic Medicine, second edi. ed. Charles C. Thomas, Springfield.

Lovejoy, C.O., Meindl, R.S., Mensforth, R.P., Barton, T.J., 1985a. Multifactorial determination of skeletal age at death: A method and blind tests of its accuracy. Am. J. Phys. Anthropol. 68, 1–14. https://doi.org/10.1002/ajpa.1330680102

Lovejoy, C.O., Meindl, R.S., Pryzbeck, T.R., Mensforth, R.P., 1985b. Chronological metamorphosis of the auricular surface of the ilium: A new method for the determination of adult skeletal age at

death. Am. J. Phys. Anthropol. 68, 15–28. https://doi.org/10.1002/ajpa.1330680103

Margaryan, A., Hansen, H.B., Rasmussen, S., Sikora, M., Moiseyev, V., Khoklov, A., Epimakhov, A., Yepiskoposyan, L., Kriiska, A., Varul, L., Saag, L., Lynnerup, N., Willerslev, E., Allentoft, M.E., 2018. Ancient pathogen DNA in human teeth and petrous bones. Ecol. Evol. 8, 3534–3542. https://doi.org/10.1002/ece3.3924

Mehra, S., Chadha, P., 2020. Alterations in structure of biomolecules using ATR-FTIR and histopathological variations in brain tissue of Channa punctatus exposed to 2Naphthalene sufonate. Toxicol. Res. (Camb). 9, 530–536. https://doi.org/10.1093/toxres/tfaa052

Meindl, R.S., Lovejoy, C.O., 1985. Ectocranial suture closure: A revised method for the determination of skeletal age at death based on the lateral-anterior sutures. Am. J. Phys. Anthropol. 68, 57–66. https://doi.org/10.1002/ajpa.1330680106

Meindl, R.S., Lovejoy, C.O., Mensforth, R.P., 1985. A revised method of age determination using the os pubis, with a review and tests of accuracy of other current methods of pubic symphyseal aging. Am. J. Phys. Anthropol. - Wiley Online Libr. 45, 29–45.

Mieczkowska, A., Mansur, S.A., Irwin, N., Flatt, P.R., Chappard, D., Mabilleau, G., 2015. Alteration of the bone tissue material properties in type 1 diabetes mellitus: A Fourier transform infrared microspectroscopy study. Bone 76, 31–39. https://doi.org/10.1016/j.bone.2015.03.010

Møller-Christensen, V., 1961. Bone Changes in Leprosy. Copenhagen: Munksgaard, Denmark.

Moraes, C.A. da C., Dias, P.E.M., Melani, R.F.H., 2014. Demonstration of protocol for computer-aided forensic facial reconstruction with free software and photogrammetry. J. Res. Dent. 2, 77. https://doi.org/10.19177/jrd.v2e1201477-90

Naito, Y.I., Yamane, M., Kitagawa, H., 2020. A protocol for using attenuated total reflection Fourier-transform infrared spectroscopy for pre-screening ancient bone collagen prior to radiocarbon dating. Rapid Commun. Mass Spectrom. 34. https://doi.org/10.1002/rcm.8720

Nielsen-Marsh, C.M., Hedges, R.E.M., 2000. Patterns of diagenesis in bone I: The effects of site environments. J. Archaeol. Sci. 27, 1139–1150. https://doi.org/10.1006/jasc.1999.0537

Nilsson, M., Possnert, G., Edlund, H., Budowle, B., Kjellström, A., Allen, M., 2010. Analysis of the Putative Remains of a European Patron Saint–St. Birgitta. PLoS One 5, e8986. https://doi.org/10.1371/journal.pone.0008986

Ortner, D.J., 2003. Identification of Pathological Conditions in Human Skeletal Remains, Identification of Pathological Conditions in Human Skeletal Remains. Elsevier. https://doi.org/10.1016/B978-0-12-528628-2.X5037-6

Paschalis, E.P., Burr, D.B., Mendelsohn, R., Hock, J.M., Boskey, A.L., 2003. Bone Mineral and Collagen Quality in Humeri of Ovariectomized Cynomolgus Monkeys Given rhPTH(1-34) for

18 Months. J. Bone Miner. Res. 18, 769–775. https://doi.org/10.1359/jbmr.2003.18.4.769

Paschalis, E.P., Mendelsohn, R., Boskey, A.L., 2011. Infrared assessment of bone quality: A review. Clin. Orthop. Relat. Res. 469, 2170–2178. https://doi.org/10.1007/s11999-010-1751-4

Paschalis, E.P., Shane, E., Lyritis, G., Skarantavos, G., Mendelsohn, R., Boskey, A.L., 2004. Bone fragility and collagen cross-links. J. Bone Miner. Res. 19, 2000–2004. https://doi.org/10.1359/JBMR.040820

Paschalis, E.P., Verdelis, K., Doty, S.B., Boskey, A.L., Mendelsohn, R., Yamauchi, M., 2001. Spectroscopic characterization of collagen cross-links in bone. J. Bone Miner. Res. 16, 1821–1828. https://doi.org/10.1359/jbmr.2001.16.10.1821

Pelton, J.T., McLean, L.R., 2000. Spectroscopic methods for analysis of protein secondary structure. Anal. Biochem. 277, 167–176. https://doi.org/10.1006/abio.1999.4320

Peter Manseau, 2009. Rag and Bone: a journey among the World's Holy Dead, fist editi. ed. Henry Holt&Company, New York.

Piga, G., Gonçalves, D., Thompson, T.J.U., Brunetti, A., Malgosa, A., Enzo, S., 2016. Understanding the Crystallinity Indices Behavior of Burned Bones and Teeth by ATR-IR and XRD in the Presence of Bioapatite Mixed with Other Phosphate and Carbonate Phases. Int. J. Spectrosc. 2016, 1–9. https://doi.org/10.1155/2016/4810149

Pluduma, L., Gross, K.A., Rey, C., Ubelis, A., Berzina, A., 2018. Production and Characterization of Oxyhydroxyapatites. Key Eng. Mater. 762, 48–53. https://doi.org/10.4028/www.scientific.net/KEM.762.48

Prag, J., Neave, R.A.H., 1997. Making Faces: Using Forensic and Archaeological Evidence, in: Wilkinson, C. (Ed.), The History of Facial Reconstruction. College Station: Texas A&M University Press, London.

Rey, C., Collins, B., Goehl, T., Dickson, I.R., Glimcher, M.J., 1989. The carbonate environment in bone mineral: A resolution-enhanced fourier transform infrared spectroscopy study. Calcif. Tissue Int. 45, 157–164. https://doi.org/10.1007/BF02556059

Riaz, T., Zeeshan, R., Zarif, F., Ilyas, K., Muhammad, N., Safi, S.Z., Rahim, A., Rizvi, S.A.A., Rehman, I.U., 2018. FTIR analysis of natural and synthetic collagen. Appl. Spectrosc. Rev. 53, 703–746. https://doi.org/10.1080/05704928.2018.1426595

Robbins, G., Mushrif Tripathy, V., Misra, V.N., Mohanty, R.K., Shinde, V.S., Gray, K.M., Schug, M.D., 2009. Ancient skeletal evidence for leprosy in India (2000 B.C.). PLoS One 4, 1–8. https://doi.org/10.1371/journal.pone.0005669

Salesse, K., Dufour, E., Lebon, M., Wurster, C., Castex, D., Bruzek, J., Zazzo, A., 2014. Variability of bone preservation in a confined environment: The case of the catacomb of Sts Peter and

Marcellinus (Rome, Italy). Palaeogeogr. Palaeoclimatol. Palaeoecol. 416, 43–54. https://doi.org/10.1016/j.palaeo.2014.07.021

Sanden, K.W., Böcker, U., Ofstad, R., Pedersen, M.E., Høst, V., Afseth, N.K., Rønning, S.B., Pleshko, N., 2021. Characterization of collagen structure in normal, wooden breast and spaghetti meat chicken fillets by ftir microspectroscopy and histology. Foods 10. https://doi.org/10.3390/foods10030548

Savitzky, A., Golay, M.J.E., 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Anal. Chem. 36, 1627–1639. https://doi.org/10.1021/ac60214a047

Schmidt, F.N., Zimmermann, E.A., Campbell, G.M., Sroga, G.E., Püschel, K., Amling, M., Tang, S.Y., Vashishth, D., Busse, B., 2017. Assessment of collagen quality associated with non-enzymatic cross-links in human bone using Fourier-transform infrared imaging. Bone 97, 243–251. https://doi.org/10.1016/j.bone.2017.01.015

Singh, B.R., DeOliveira, D.B., Fu, F.-N., Fuller, M.P., 1993. Fourier transform infrared analysis of amide III bands of proteins for the secondary structure estimation, in: Nafie, L.A., Mantsch, H.H. (Eds.), Biomolecular Spectroscopy III. pp. 47–55. https://doi.org/10.1117/12.145242

Sirak, K.A., Fernandes, D.M., Cheronet, O., Novak, M., Gamarra, B., Balassa, T., Bernert, Z., Cséki, A., Dani, J., Gallina, J.Z., Kocsis-Buruzs, G., Kővári, I., László, O., Pap, I., Patay, R., Petkes, Z., Szenthe, G., Szeniczey, T., Hajdu, T., Pinhasi, R., 2017. A minimally-invasive method for sampling human petrous bones from the cranial base for ancient DNA analysis. Biotechniques 62, 283–289. https://doi.org/10.2144/000114558

Sjøvold, T., 1990. Estimation of stature from long bones utilizing the line of organic correlation. Hum. Evol. 5, 431–447. https://doi.org/10.1007/BF02435593

Soffiantini, G., 2014. Le reliquie di San Teobaldo. Storia e “segni” dalla morte ai nostri giorni.

Sønju Clasen, A.B., Ruyter, I.E., 1997. Quantitative Determination of Type A and Type B Carbonate in Human Deciduous and Permanent Enamel by Means of Fourier Transform Infrared Spectrometry. Adv. Dent. Res. 11, 523–527. https://doi.org/10.1177/08959374970110042101

Spekker, O., Hunt, D.R., Paja, L., Molnár, E., Pálfi, G., Schultz, M., 2020. Tracking down the White Plague: The skeletal evidence of tuberculous meningitis in the Robert J. Terry Anatomical Skeletal Collection. PLoS One 15, e0230418. https://doi.org/10.1371/journal.pone.0230418

Stani, C., Vaccari, L., Mitri, E., Birarda, G., 2020. FTIR investigation of the secondary structure of type I collagen: New insight into the amide III band. Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 229, 118006. https://doi.org/10.1016/j.saa.2019.118006

Stanislavov, A.S., Sukhodub, L.F., Sukhodub, L.B., Kuznetsov, V.N., Bychkov, K.L., Kravchenko, M.I., 2018. Structural features of hydroxyapatite and carbonated apatite formed under the

influence of ultrasound and microwave radiation and their effect on the bioactivity of the nanomaterials. Ultrason. Sonochem. 42, 84–96. https://doi.org/10.1016/j.ultsonch.2017.11.011

Taylor, K.T., 2000. Forensic Art and Illustration. CRC Press. https://doi.org/10.1201/9781420036954

Walker, P.L., 2005. Greater sciatic notch morphology: Sex, age, and population differences. Am. J. Phys. Anthropol. 127, 385–391. https://doi.org/10.1002/ajpa.10422

Walsham, A., 2010. Introduction: Relics and Remains. Past Present 206, 9–36. https://doi.org/10.1093/pastj/gtq026

Weiner, S., 2010. Microarchaeology. Beyond the visible archaeological record, first edit. ed. Cambridge University Press, UK.

Weiner, S., Bar-Yosef, O., 1990. States of preservation of bones from prehistoric sites in the Near East: A survey. J. Archaeol. Sci. 17, 187–196. https://doi.org/10.1016/0305-4403(90)90058-D

Wilkinson, C., 2004. The history of facial reconstruction, in: Forensic Facial Reconstruction. Cambridge University Press, Cambridge, pp. 39–68. https://doi.org/10.1017/CBO9781107340961.004

Wopenka, B., Pasteris, J.D., 2005. A mineralogical perspective on the apatite in bone. Mater. Sci. Eng. C 25, 131–143. https://doi.org/10.1016/j.msec.2005.01.008

Xu, Y., 2009. Thermal stability of collagen triple helix., 1st ed, Methods in enzymology. Elsevier Inc. https://doi.org/10.1016/s0076-6879(09)66009-2

Zampini, P.L., Corrain, C., 1975. S. Teobaldo nella storia e nelle tradizioni popolari: tentativo di ricostruzione antropologica sulla base dei resti scheletrici conservati in Badia Polesine, nella chiesa arcipretale. Atti e Mem. del Sodal. Vangadiciense 1, 95/118.

Zanatta, A., Bezzi, L., Carrara, N., Moraes, C., Thiene, G., Zampieri, F., 2018. New technique in facial reconstruction: the case of Giovanni Battista Morgagni. Anthropol. Anzeiger 75, 89–100. https://doi.org/10.1127/anthranz/2018/0818

Supplementary Information

Life and death of Saint Theobald (1033-1066 AD) from the chemical and forensic-anthropological analysis

Cinzia Scaggion[1]*, Cicero Moraes[2], Luca Bezzi[2], Alberto Zanatta[3], Meriam Guellil[4], Eugenia D'Atanasio[5], Luca Nodari[6], Luca Pagani[7,4], Gilberto Artioli[1], Nicola Carrara[3]

[1]Department of Geosciences, University of Padua, Italy
[2]Arc-Team srl, Trento, Italy
[3]University Museums Centre - CAM, University of Padua, Italy
[4]Institute of Genomics, University of Tartu, Estonia
[5]Sapienza, University of Rome, Italy
[6]Italian National Research Council-CNR, ICMATE, Italy
[7]Department of Biology, University of Padua, Italy

*corresponding author, cinzia.scaggion@phd.unipd.it

Supplementary Results
Anthropological and Palaeopathological observations
Molecular investigations for ancient microbial infection and genetics

Supplementary Figures
Figures S1-S2-S3-S4

Supplementary Tables
Table S1

Anthropological and Palaeopathological observations

The osteological remains inside the urn were in good state of conservation and belong to a male individual: cranium, mandible, right scapula and ulna, n.3 right ribs, n.1 left rib, n.3 thoracic vertebrae (T_6-T_8-T_{11}), n.1 lumbar vertebra (L_2), n.2 coxal bones, n.2 femurs (right and left), n.2 tibiae (right and left) and n.1 left fibula (**Figure.S1**).

Figure S1. *Finding of the remains of the Saint after opening the urn. The skeletal appeared not complete and in good state of preservation.*

Sex determination was defined by morphological analysis all skeletal element found. Although the remains are incomplete, analyzing the anatomical districts indicators of the sex, such as coxal bones, femurs, cranium and mandible, the skeleton revealed typical male features. From the anthropometric analysis of the long bones, the diameter of the femoral heads, having a value of 48 mm (male range), was significant and diagnostic for the sex determination.

To determine the age of death, many skeletal age markers were analyzed. Analyzing the anatomical elements present of appendicular skeleton, starting from long bones, showed skeletal maturity with épiphyses and diaphyses fused, free from signs of degeneration, attributable to an adult individual between 25-40 years old. The diagnosis of sex and age of death was more accurately determined by visualizing the pelvis that showed male anthropometric characters: narrow ischial notch, intermediate grade pubic angle, narrow pubic bone body and absent ventral arch. From auricular surface (sacroiliac joint) the reputable age (derived from its degeneration) is around 30-34 years; from pubic symphysis around 29-35 years. Cranium and robust mandible confirm the sex diagnosis: pronounced glabella with robust superciliary arch, supraorbital margin rounded and rectangular orbits; robust mastoid process and nuchal crest middle pronounced, ascendant mandibular branches with vertical trend and eminence square chin rest. From dentition, was found *in situ* a second lower left molar tooth (LM_2), very worn. Even dental wear is indicative of age and its wear degree falls back into the range 45-55 years. Age evaluation can become very complex as the aging of the skeleton and worn teeth do not always correspond to the chronological age at the time of death, such as in this case of study. From several age ranges observed, it was evaluated that the age of individual was between 30-35 years at the time of death. By measuring the preserved long bones it was estimated that the individual's height was around 173.6 and 174.6 cm according to anthropometric indices (Sjøvold, 1990; Trotter and Gleser, 1952). Some historical sources report that Saint Theobald fell ill with Leprosy meanwhile dedicated himself to the poorest and most needy. In Europe with the advent of the Middle Ages there were numerous cases: in the 9th century there was a high peak in the frequency of the disease until 1200-1300 AD. and, given its high diffusion, 19.000 leprosaria were built (G. and Giuffra, 2009). Leprosy, also known as Hansen's disease, is a chronic and mutilating infectious disease that spreads through saliva drops or prolonged direct contact with the lesions of affected individuals. The disease affects both genders equally, it is not fatal and has a chronic and disfiguring course. The localization of the lesions characteristic of lepers are mainly located on the face, hands and feet. Lepers take on a distinctive appearance with nodules and skin ulcers on the face (*lepromas*), destruction of the nasal bones resulting in enlargement due to the discharge of purulent material, eye inflammation and chronic laryngitis. As the disease progresses, fingers and toes are amputated. Death can usually occur from complications not directly induced by the disease: the injuries expose the patient to infections,

compromising and further aggravating the state of health. In the absence of the bones that make up the hands and feet, traces of the pathognomonic signs (i.e. indicative of the disease) were looked for by first examining the cranium. Four erosion lesions were observed on frontal bone and one less obvious two involves zygomatic bones (**Figure.S2**).

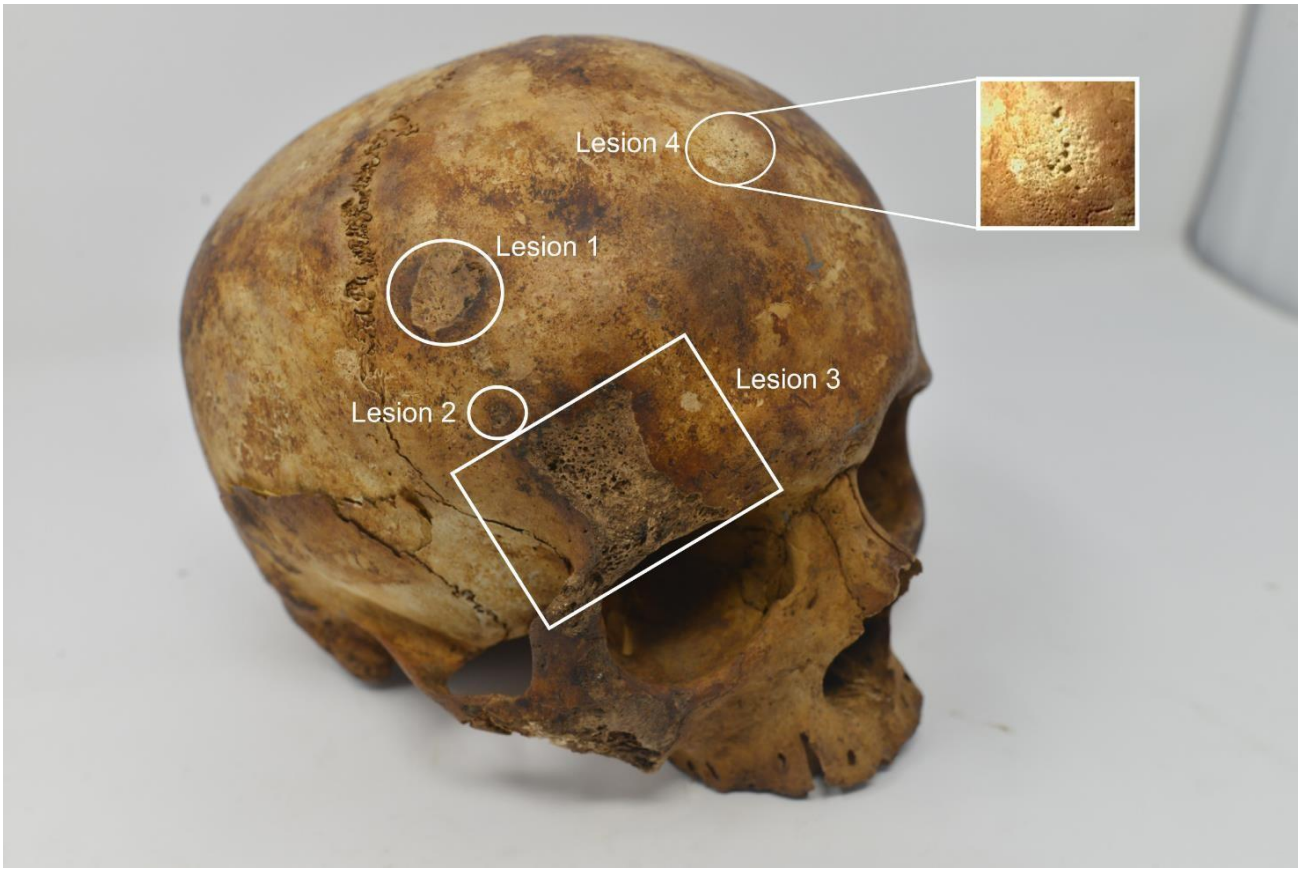

Figure S2. *Skull lesions observed on frontal bone and circular depression with an osteolytic in the infraorbital region of the maxilla (white arrow).*

Lesion 1 located on the high in the right portion of the frontal with circular shape (10x12 mm) showed rounded margins, slight porosity and followed by a second smaller circular lesion (Lesion 2) with dimension (4x4 mm), probably caused by ulcers. Lesion 3 was placed at the level of the right superciliary arch having a rectangular shape (30x18 mm) and extending along the right zygomatic bone. Analyzing the margins of the lesion 3, it seems to be caused by *post-mortem* alterations as the erosion of the cortical bone has well-defined margins without bone reaction. The mid-frontal portion of the skull was characterized by pseudo-circular macro-porosity (*porotic hyperostosis*) (Lesion 4) while a circular depression -characterized by small osteolytic lesions- was present in the infraorbital region of the maxilla. Porosity was also observed on the nasal bones and on the hard palate, slight atrophy of the pyriform opening at the level of the anterior nasal spine (right maxilla) while has been identified an osteophyte (3 mm) at the level of the anterior nasal spine of the left maxilla. Maxilla and mandible show significant alveolar resorption (**Figure S3**).

Figure S3. *Porosity is visible in nasal bone and hard palate with severe alveolar resorption of maxilla (**a**,**b**); slight atrophy of the pyriform opening (right maxilla) and an osteophyte (3 mm) on the left (**c**); on the mandible it was observed proliferative periostitis (woven) on the mental protuberance and in a portion of the mandibular body, inferior view (**d**).*

On the latter was observed proliferative periostitis (woven) on the mental protuberance and in a portion of the mandibular body (left side), see **Figure.S3**. A slight macro-porosity in the left greater wing of the sphenoid bone was observed, less evident in the right side. In the postcranial *cribra femora* were detected in both of anterior *collum femoris* (Göhring, 2021), while head of left femur showed a osteochondritis dissecans (a type of aseptic necrosis, benign and non-inflammatory) (Baxarias J. and Herrerin, 2008). Bilateral enthesopathies in the pectineal line on pubis areas (**Figure.S4**) and apposition of new bone tissue in the dorsal body of the pubis on the left coxal bone were observed. In

the spinal column, lumbar vertebra (L_2) is characterized by Schmörl's hernia, osteoarthritis on the upper articular facets while the thoracic vertebra (T_8) presents calcification of yellow ligament and slight rotation of the vertebral body to the right.

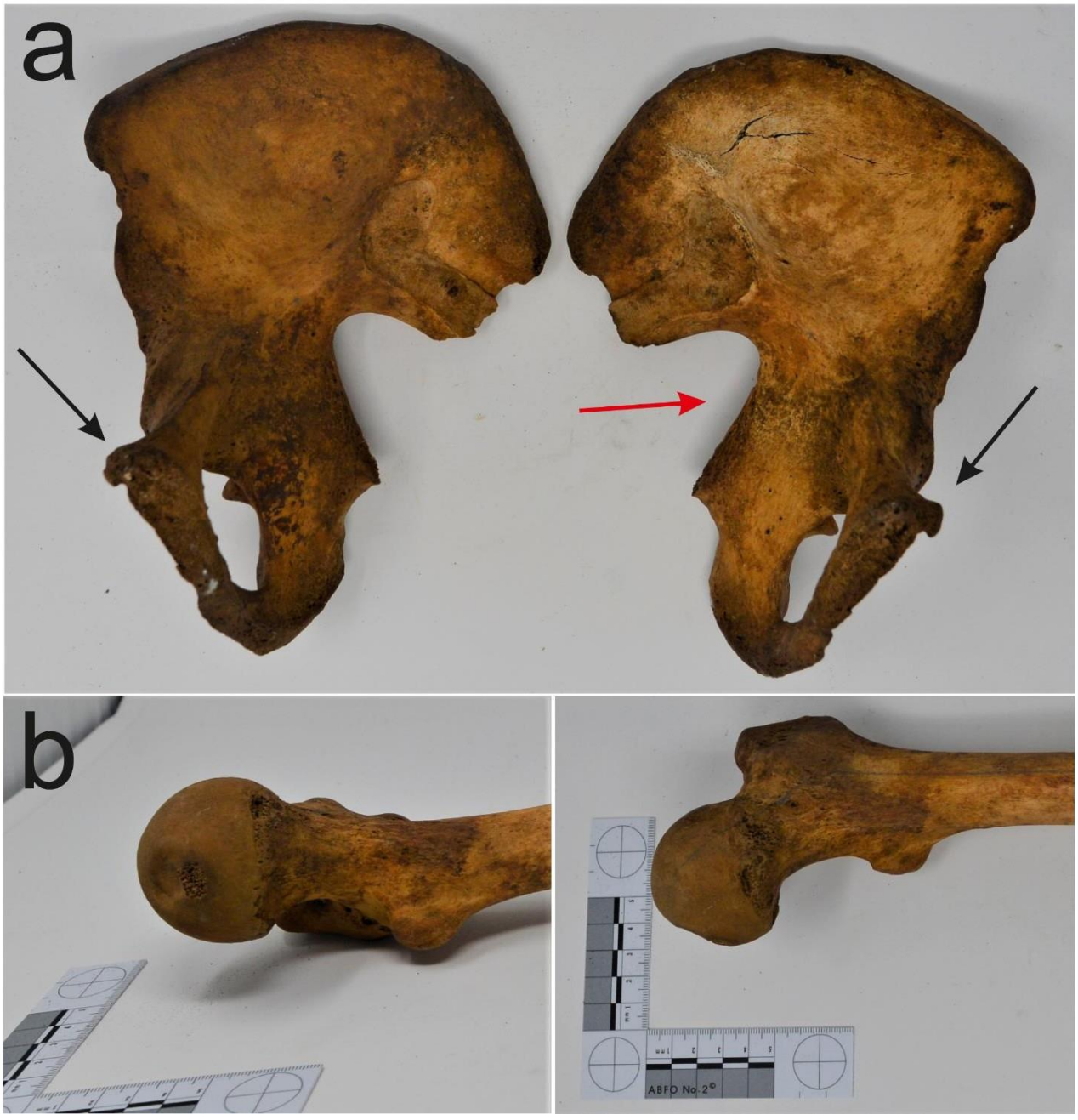

Figure S4. *Bilateral enthesopathies in the pectineal line on pubis areas (black arrows) of coxal bones and apposition of new bone tissue in the dorsal body of the pubis on the left (red arrow) (**a**); head of left femur showed an osteochondritis dissecans and cribra femora on collum femoris (**b**).*

Molecular investigation for ancient microbial infection and genetics

Quantification and statistical analysis

Mapping

Before mapping, the sequences of the adapters, indexes, and poly-G tails occurring due to the specifics of the NextSeq 500 technology were cut from the ends of DNA sequences using cutadapt-1.11 (Martin, 2011). Sequences shorter than 30 bp were also removed with the same program to avoid random mapping of sequences from other species. The sequences were aligned to the reference sequence

GRCh37 (hs37d5) using Burrows- Wheeler Aligner (BWA 0.7.12) (Li and Durbin, 2009) and the command mem with re-seeding disabled. After alignment, the sequences were converted to BAM format and only sequences that mapped to the human genome were kept with samtools 1.3 (Li et al., 2009). Afterwards, the data from different flow cell lanes were merged and duplicates were removed using picard 2.12 (http://broadinstitute.github.io/picard/index.html). Indels were realigned using GATK 3.5 (McKenna et al., 2010) and reads with a mapping quality less than 10 were filtered out using samtools 1.3 (Li et al., 2009).

aDNA authentication

As a result of degrading over time, aDNA can be distinguished from modern DNA by certain characteristics: short fragments and a high frequency of C=>T substitutions at the 5' ends of sequences due to cytosine deamination. The program mapDamage2.0 (Jónsson et al., 2013) was used to estimate the frequency of 5' C=>T transitions. Rates of contamination were estimated on mitochondrial DNA by calculating the percentage of non-consensus bases at haplogroup-defining positions as detailed in (Jones et al., 2017). Each sample was mapped against the RSRS downloaded from phylotree.org and checked against haplogroup-defining sites for the sample-specific haplogroup.

Samtools 1.9 (Li et al., 2009) option *stats* was used to determine the number of final reads, average read length, average coverage etc. The average endogenous DNA content (proportion of reads mapping to the human genome) was 0.08% and the average human coverage was 0.002x. Based on the average endogenous DNA content and the average human coverage, the individual was excluded from further genome-wide analysis.

Sample ID	Raw FASTQ Reads	Reads Mapping to hg19	Mapped Duplicates	% hg19 Dup	% Unique Human	Average length	Average human Covager [X]
TEO001	166,346,761	182,102	56,319	30.93%	0.08%	47	0.002

Table S1. *Table of results obtained from St. Theobald's remains.*

Calculating genetic sex estimation

Genetic sex was calculated using the methods described in (Skoglund et al., 2013), estimating the fraction of reads mapping to Y chromosome out of all reads mapping to either X or Y chromosome. However, the number of X and Y chromosome reads was limited to a total of 2,216 reads and the genetic sex could not be assigned.

Determining mtDNA haplogroups

For the determination of the mitochondrial DNA (mtDNA) haplogroup, the reads were re-aligned to the reference sequence RSRS. After the removal of duplicates, the number of mapped reads to the reference sequences was too low to estimate the mtDNA haplogroup for this individual.

Metagenomic analysis

Based on both runs we could not find evidence for the presence of *Mycobacterium leprae*, the causative pathogen for leprosy. However, as the bacterium is not always detectable in teeth upon infection it cannot be excluded that our lack of hits for *M. leprae* is a false negative which was either caused by preservation issues or insufficient perimortem bacterial colonization of the vascularized tooth. We also screened the datasets for a panel of human pathogens but could not identify any with certainty due to suboptimal DNA preservation.

Metagenomic Screening Methods

Sequencing data for both shotgun sequencing runs were quality checked using FastQC (Bioinformatics) and MultiQC (Ewels). Raw sequencing datasets were trimmed and filtered using cutadapt(Martin, 2011) (-m 30 –nextseq-trim=20 –times 3 -e 0.2 –trim-n) and deduplicated using ParDRe(González-Domínguez and Schmidt, 2016).

Deduplicated files were merged by sequencing run and analysed using KrakenUniq(Breitwieser et al., 2018). We used a custom database made up of complete genomes and chromosome level assemblies of bacteria, viruses, archaea and protozoa. The sequences were all dusted, i.e. we removed low complexity regions during database building. Additionally, the database contained the human genome, the NCBI Viral Neighbor database and the contaminant databases UniVec and EmVec. We computed a heatmap using plotly (Plotly Tecthologies Inc 2015), pandas (McKinney and Others, 2010), matplotlib (Hunter, 2007) and numpy (van der Walt et al., 2011) based on a custom E-value calculated as follows: $(\frac{K}{R}) * C$ (Guellil et al., 2021) (K is the kmer count, R is the read count and C is the coverage). The E-value cut-off for further inspection was 0.001.

DNA methods details

All of the laboratory work was performed in dedicated ancient DNA laboratories at the Estonian Biocentre, Institute of Genomics, University of Tartu, Tartu, Estonia. The library quantification and sequencing were performed at the Estonian Biocentre Core Laboratory. The main steps of the laboratory work are detailed below.

DNA extraction

For the study, DNA was extracted from a root of one tooth. The root was soaked in 6% (w/v) bleach for 5 minutes. Samples were rinsed three times with 18.2 MΩcm H_2O and soaked in 70% (v/v) Ethanol for 2 minutes. The tubes were shaken during the procedure to dislodge particles. The samples were transferred to a clean paper towel on a rack inside a class IIB hood with the UV light on and allowed to dry for two to three hours. Afterwards, the samples were weighed to calculate the accurate volume of EDTA (20x EDTA [µl] of sample mass [mg]) and Proteinase K (0.5x Proteinase K [µl] of sample mass [mg]). EDTA and Proteinase K were added into PCR-clean 5 ml or 15 ml conical tubes

(Eppendorf) along with the samples inside the IIB hood and the tubes were incubated 72 h on a slow shaker at room temperature.

The DNA extract was concentrated to 250 µl using the Vivaspin® Turbo 15 (Sartorius) and purified in a large volume column (High Pure Viral Nucleic Acid Large Volume Kit, Roche) using 2.5 ml of PB buffer, 1 ml of PE buffer and 100 µl of EB buffer (MinElute PCR Purification Kit, QIAGEN). For the elution of the endogenous DNA, the silica column was transferred to a collection tube to dry and followed in 1.5 ml DNA low-bind tubes (Eppendorf) to elute. The sample was incubated with 100 µl EB buffer at 37 C for 10 minutes and centrifuged at 13,000 rpm for two minutes. After centrifugation, the silica column was removed and the sample was stored at -20 C. Only one extraction was performed for screening and 30 µl was used for libraries.

Library preparation

Sequencing library was built using NEBNext® DNA Library Prep Master Mix Set for 454™ (E6070, New England Biolabs) and Illumina-specific adaptors (Meyer and Kircher, 2010) following established protocols (Malaspinas et al., 2014; Meyer and Kircher, 2010; Orlando et al., 2013). The end repair module was implemented using 18.75 µl of water, 7.5 µl of buffer and 3.75 µl of enzyme mix, incubating at 20 °C for 30 minutes. The sample was purified using 500 µl PB and 650 µl of PE buffer and eluted in 30 µl EB buffer (MinElute PCR Purification Kit, QIAGEN). The adaptor ligation module was implemented using 10 µl of buffer, 5 µl of T4 ligase and 5 µl of adaptor mix (Meyer and Kircher, 2010), incubating at 20 °C for 15 minutes. The sample was purified as in the previous step and eluted in 30 µl of EB buffer (MinElute PCR Purification Kit, QIAGEN). The adaptor fill-in module was implemented using 13 µl of water, 5 µl of buffer and 2 µl of Bst DNA polymerase, incubating at 37 °C for 30 and at 80 °C for 20 minutes. The library was amplified using the following PCR set up:

50µl DNA library, 1X PCR buffer, 2.5mM MgCl2, 1 mg/ml BSA, 0.2µM inPE1.0, 0.2mM dNTP each, 0.1U/µl HGS Taq Diamond and 0.2µM indexing primer. Cycling conditions were: 5' at 94C, followed by 18 cycles of 30 seconds each at 94C, 60C, and 68C, with a final extension of 7 minutes at 72C. The sample was purified and eluted in 35 µl of EB buffer (MinElute® PCR Purification Kit, QIAGEN). Three verification steps were implemented to make sure library preparation was successful and to measure the concentration of dsDNA/sequencing library -fluorometric quantitation (Qubit, Thermo Fisher Scientific), parallel capillary electrophoresis (Fragment Analyser, Agilent Technologies) and qPCR.

DNA sequencing

DNA was sequenced using the Illumina NextSeq500/550 High-Output single-end 75 cycle kit.

REFERENCES

Baxarias J., Herrerin, J., 2008. The handbook atlas of paleopathology. Zaragoza: Editorial Pórtico, Zaragoza.

Breitwieser, F.P., Baker, D.N., Salzberg, S.L., 2018. KrakenUniq: confident and fast metagenomics classification using unique k-mer counts. Genome Biol. 19, 198. https://doi.org/10.1186/s13059-018-1568-0

G., F., Giuffra, V., 2009. Lezioni di paleopatologia. ECIG- Edizioni Culturali Internazionali Genova, Genova.

Göhring, A., 2021. Allen's fossa—An attempt to dissolve the confusion of different nonmetric variants on the anterior femoral neck. Int. J. Osteoarchaeol. 1–10. https://doi.org/10.1002/oa.2968

González-Domínguez, J., Schmidt, B., 2016. ParDRe: faster parallel duplicated reads removal tool for sequencing studies: Table 1. Bioinformatics 32, 1562–1564. https://doi.org/10.1093/bioinformatics/btw038

Guellil, M., Keller, M., Dittmar, J., Inskip, S., Cessford, C., Solnik, A., Kivisild, T., Metspalu, M., Robb, J., Scheib, C., 2021. An Invasive Haemophilus Influenzae Serotype B Infection in an Anglo-Saxon Plague Victim. In Prep.

Hunter, J.D., 2007. Matplotlib: A 2D Graphics Environment. Comput. Sci. Eng. 9, 90–95. https://doi.org/10.1109/MCSE.2007.55

Jones, E.R., Zarina, G., Moiseyev, V., Lightfoot, E., Nigst, P.R., Manica, A., Pinhasi, R., Bradley, D.G., 2017. The Neolithic Transition in the Baltic Was Not Driven by Admixture with Early European Farmers. Curr. Biol. 27, 576–582. https://doi.org/10.1016/j.cub.2016.12.060

Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., Orlando, L., 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics 29, 1682–1684. https://doi.org/10.1093/bioinformatics/btt193

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Malaspinas, A.-S., Lao, O., Schroeder, H., Rasmussen, M., Raghavan, M., Moltke, I., Campos, P.F., Sagredo, F.S., Rasmussen, S., Gonçalves, V.F., Albrechtsen, A., Allentoft, M.E., Johnson, P.L.F., Li, M., Reis, S., Bernardo, D. V., DeGiorgio, M., Duggan, A.T., Bastos, M., Wang, Y., Stenderup, J., Moreno-Mayar, J.V., Brunak, S., Sicheritz-Ponten, T., Hodges, E., Hannon, G.J., Orlando, L., Price, T.D., Jensen, J.D., Nielsen, R., Heinemeier, J., Olsen, J., Rodrigues-Carvalho, C., Lahr, M.M., Neves, W.A., Kayser, M., Higham, T., Stoneking, M., Pena, S.D.J., Willerslev, E., 2014. Two ancient human genomes reveal Polynesian ancestry among the indigenous Botocudos of Brazil. Curr. Biol. 24, R1035–R1037. https://doi.org/10.1016/j.cub.2014.09.078

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 10. https://doi.org/10.14806/ej.17.1.200

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303. https://doi.org/10.1101/gr.107524.110

McKinney, W., Others, 2010. Data structures for statistical computing in python, in: van der Voort, S., Millman, J. (Eds.), Proceedings of the 9th Python in Science Conference. pp. 445:51–56.

Meyer, M., Kircher, M., 2010. Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing. Cold Spring Harb. Protoc. 2010, pdb.prot5448-pdb.prot5448. https://doi.org/10.1101/pdb.prot5448

Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., Stiller, M., Schubert, M., Cappellini, E., Petersen, B., Moltke, I., Johnson, P.L.F., Fumagalli, M., Vilstrup, J.T., Raghavan, M., Korneliussen, T., Malaspinas, A.-S., Vogt, J., Szklarczyk, D., Kelstrup, C.D., Vinther, J., Dolocan, A., Stenderup, J., Velazquez, A.M. V., Cahill, J., Rasmussen, M., Wang, X., Min, J., Zazula, G.D., Seguin-Orlando, A., Mortensen, C., Magnussen, K., Thompson, J.F., Weinstock, J., Gregersen, K., Røed, K.H., Eisenmann, V., Rubin, C.J., Miller, D.C., Antczak, D.F., Bertelsen, M.F., Brunak, S., Al-Rasheid, K.A.S., Ryder, O., Andersson, L., Mundy, J., Krogh, A., Gilbert, M.T.P., Kjær, K., Sicheritz-Ponten, T., Jensen, L.J., Olsen, J. V., Hofreiter, M., Nielsen, R., Shapiro, B., Wang, J., Willerslev, E., 2013. Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. Nature 499, 74–78. https://doi.org/10.1038/nature12323

Sjøvold, T., 1990. Estimation of stature from long bones utilizing the line of organic correlation. Hum. Evol. 5, 431–447. https://doi.org/10.1007/BF02435593

Skoglund, P., Storå, J., Götherström, A., Jakobsson, M., 2013. Accurate sex identification of ancient human remains using DNA shotgun sequencing. J. Archaeol. Sci. 40, 4477–4482. https://doi.org/10.1016/j.jas.2013.07.004

Trotter, M., Gleser, G.C., 1952. Estimation of stature from long bones of American Whites and Negroes. Am. J. Phys. Anthropol. 10, 463–514. https://doi.org/10.1002/ajpa.1330100407

van der Walt, S., Colbert, S.C., Varoquaux, G., 2011. The NumPy Array: A Structure for Efficient Numerical Computation. Comput. Sci. Eng. 13, 22–30. https://doi.org/10.1109/MCSE.2011.37

CONCLUSIONS AND FUTURE PERSPECTIVES

The main aim of this project was to survey skeletal tissues, in particular petrous bones and tooth roots, made up of mineralized collagen, as a pre-screening method using infrared spectroscopy (FTIR) to assess bone quality for molecular analyzes.

A large set of archaeological samples were selected characterized by different origins, chronology and different states of preservation because of different diagenetic pathways due to the different environmental conditions to which they were exposed. This approach aimed to overcome a major problem of identifying useful genetic material in archaeological bone collections without resorting to demanding and expensive laboratory studies.

To achieve this goal, the methods of archaeological documentation were first evaluated. The acquisition of archaeological data in bone finds can in some cases (e.g, X-ray tomography and synchrotron) represent a problem. X-rays can achieve information retrieval in particular techniques of highly degradable molecules at natural radiation doses. **Chapter 1** discusses the accuracy, *pros* and *cons* of generating 3D surfaces, generated by photogrammetry and micro-CT, on two samples of highly degraded small teeth. The small size, the very irregular areas and the surfaces characterized by micro-cracks in the enamel and in the cementum, made it possible to evaluate the potential as well as the two limits of the techniques.

Research has shown that the 3D photogrammetric models of the outer surface are accurate and differ by several hundred microns from the corresponding tomography even for small objects with very irregular surfaces.

Although photogrammetric models generate a significantly smaller point cloud than that originated by micro-CT, the model surfaces provide sufficient details, clear and well resolved, for morphological analysis of the findings.

Another element that can affect the quality of the model is the resizing of the object.

Incorrect scaling, especially in small objects, can lead to invalid morphometric measurements that distort the data. Micro-CT scanning provides automatic scaling as pixel size information is provided in the log file while automatic model scaling does not occur with manual photogrammetry. In the photogrammetric software used to process the images and create the digital model, the dimensions of the object are instead reported manually using a digital caliper. Considering the high spatial resolution of micro-CT (pixel size of c.a. 8 μm), the estimated error between the manually reported measurement and the micro-CT is 0.5 mm for both samples, which can be considered minimal.

It was found that the areas covered with dentin and enamel create reflective and translucent surfaces that can affect the results by causing a decrease in the accuracy of the model caused by the generation of noisy points.

Although the investigation was conducted with a limited number of samples, it can still be said that the photogrammetric technique, inexpensive, is sufficiently accurate and can be considered a valid alternative to X-ray imaging to provide 3D information with an additional advantage of providing texture rendering.

To develop a minimally invasive method for bone selection, a standard preparation method was sought to improve the ability to achieve data accuracy and data replicability. Through the preliminary micro-CT analysis, it was possible to obtain the standard morphometric indices of the petrous bone: bone surface, trabecular thickness and trabecular separation. As reported in Chapter 2, these parameters were used to verify the highest bone density area of the sample type, aimed at precise extraction of the powder for FTIR analysis, without damaging the very precious *otic capsule* for genetic analysis.

Bone powder extracted from different points of the sample at different depths was analyzed by IR spectroscopy to evaluate if there were detectable differences. The data collected and illustrated in **Chapter 2** show that there is no variation between sampling position and depth: the differences are due to statistical fluctuations caused by the limited noise sensitivity of the instrument used. This result defines that diagenesis acts uniformly on the entire sample while sample preparation has an evident impact. A uniform preparation method has been established for all samples from the results obtained by applying it to the entire set of samples obtaining accuracy, repeatability and reproducibility, as well as the reliability of the spectroscopic data. The standardization of the preparation method allows accurate analysis of the infrared spectra in order to evaluate the conservation status of the samples, aimed at the possibility of molecular extraction.

A precise evaluation of the diagenetic bone alteration was carried out through an analytical approach applied to the study of the finds and their burial environment, investigating type and extent, as well as their dependence on the climatic and local conditions of burial.

The development of a pre-screening method first involved the evaluation of a broad model of universal diagenetic alteration of osteological samples (discussed in **Chapter 3**), comparing the behavior of the organic and inorganic phase on a large and diverse set of finds. From extremely altered archaeological samples to fresh and modern bones, the sensitivity and efficacy of the infrared phosphate peak width (FW85%) was evaluated including it as the most predictive parameter on the degree of alteration of the bone system.

These characteristics determined the choice of using the peak width to monitor bioapatite crystallinity. Phosphate peak width was shown to effectively describe the chemical environment of

bioapatite crystallites by monitoring minimal changes in the chemical composition and structure of crystal lattice.

As evidenced by the data obtained in **Chapter 3**, the parameter FW85% is not only indirectly indicative of the quantity of carbonate in the crystal lattice but is sensitive to the presence of collagen with which the recrystallization of bioapatite is intimately connected.

Furthermore, FW85% provides a reliable evaluation of the osteological material properties even in subsamples coming from the same necropolis with minimal variations of FW85% assuming different biomechanical and chemical properties.

The survey carried out gave the opportunity to identify the potential of some IR parameters compared to others widely used in the literature. The sensitivity of the phosphate peak width, for the purposes of research, was evaluated using it for the selection of less diagenized and less altered samples useful for stable isotope analysis or the extraction of ancient DNA. Although the bone samples can undergo different levels of alteration (e.g. due to particular environmental and burial conditions), the analyzed data shows that diagenesis acts according to a generalized model, progressively losing collagen, a structural carbonate with consequent strong recrystallization of the apatite characterized by a higher structural order. From the research carried out, it was possible to confirm that diagenesis produces the same effects despite the fact that different locations, types of bone tissue, chronology and diagenetic pathways cause alterations of different entities and degree of intensity.

The comparison between the AmI/PO_4^{3-} and AmII/PO_4^{3-} absorption bands (illustrated in **Chapter 3**), characteristic of the peptide bond, was used to monitor the presence/absence of the organic matter distinguishing the altered samples that present only structural or weakly bound water. Samples containing conspicuous or small quantities of collagenous material, from results obtained, follow a single trend that can be reasonably approximated by a linear regression model.

Results show that the parameter used mainly in literature, AmI/PO_4^{3-}, is a very useful tool for quantifying the presence of collagen preserved in the bone system but is not very reliable for those highly degraded samples which, due to environmental and burial conditions, are free of organic matter and contain mainly structural water. Archaeological bones can incorporate water within the bone structure, the signal of which overlaps the amide band. In this regard, AmI/PO_4^{3-} - AmII/PO_4^{3-} could be a useful and more sensitive comparison between two parameters for discriminating well samples with collagen from those without. On these promising bases, a pre-screening method was developed with particular regard to the organic component as it is still controversial whether the organic

component protects the survival of ancient DNA rather than the inorganic component or vice versa, for samples destined for genetic, isotope and radiocarbon analyzes.

The data provided by parameters used gave way to effectively determine the degree of alteration of the examined samples. Samples which had high values of FW85% and of organic matter content were considered promising for DNA extraction. Their predictive values for determining the presence/absence of genetic molecules were evaluated using a multiple linear model (lm). In **Chapter 4,** the results showed that, among the variables considered, AmI/PO_4^{3-} is not representative of the model unlike the more indicative variable, FW85%.
The FW85% parameter has a high potential probably determined by its previously specified intrinsic properties.
In conclusion, from this model, the infrared parameter of the phosphate peak width (FW85%) could have a greater predictive power for the presence/absence of DNA and in the future it can be used to evaluate the extraction potential of sequenceable genetic molecules from archaeological bone samples.

All selected samples produced sequenceable aDNA but with high variability in terms of endogenous DNA yields. To ascertain whether the quality of the collagen was indicative of the amount of extractable aDNA, a multi-analytical analysis by Infrared spectroscopy and Mass Spectrometry of the lyophilized collagen extracted from bone samples previously sequenced was carried out.

From the spectral data it emerged that the altered amide I signal has a much more marked shoulder (~ 1720 cm-1) on highly diagenized samples and this alteration affects the good isotope yield (C: N atomic ratio) resulting in a worse quality of the collagen. In fact, the shoulder is a signal of increased denaturation of the carbonyl of proteins. This alteration was confirmed by analyzing from the IR spectra the amide I sub-bands at 1660 cm^{-1} and 1690 cm^{-1} related to the maturity of collagen cross-links (Paschalis et al., 2001).

The values 1660: 1690 cm-1 are negatively correlated to the quality of the collagen: the lower the ratio between these areas of the amide I bands, the greater degenerative state of the collagen confirmed by the presence of the vibrational band of the carbonyl compounds (~ 1720 cm -1) and by the atomic ratio C: N (Martínez Cortizas and López-Costas, 2020).

For the careful selection of samples intended for analysis of stable isotopes, it is possible to evaluate the suitability of the finds by means of IR spectroscopy in two distinct phases. Initially, the state of conservation of the find can be examined by evaluating the state of alteration of the organic and inorganic component by means of infrared measurements of the bone sample and subsequently, if the

quantity of collagen is considered sufficient, it is possible to deepen the pre-screening by analyzing the structure polypeptide before laboratory analysis.
From the research performed in **Chapter 4**, it can be concluded that the collagen structure could be used as a pre-screening for paleonutritional studies or radiocarbon dating while no correlation was found with aDNA yields.
The quantity/quality of endogenous DNA may not be predictable due to the influence of local environmental factors while the FTIR signal measurement of the phosphate peak width can predict the presence/absence of DNA, but it is still difficult to fully describe what it is a multifactorial process.

The case of study reported in **Chapter 5** multidisciplinary analyzes were carried out on the medieval sacred remains of the body of Saint Theobald, which, according to historical sources, has never undergone *post-mortem* alterations linked to the funerary environment (Soffiantini, 2014).
The investigations carried out on its remains were of a physical, anthropological and paleopathological type and infrared spectral analysis were carried out to determine its state of conservation and to define whether the sample was suitable for radiocarbon analysis. Finally, the features of the saint's face were reconstructed through the elaboration of a 3D virtual model of the forensic skull.
The investigation methodologies applied to the bone material in **Chapters 4** and **5**, gave way to evaluate the state of conservation of the skeletal remains, which were in an excellent state of conservation. In conclusion, the pre-screening analysis indicates that the collagen contained in the remains of the saint are in a good state of preservation although the organic component has undergone denaturation (observable through 1660 cm^{-1} and 1690 cm^{-1} sub-bands analysis) of proteins. The amount of collagen preserved in the archaeological sample is comparable to that of the fresh bone, while analysis of amide I sub-bands shows its quality of preservation where it is evident that a denaturation has occurred.
The analysis at ^{14}C confirmed what had previously been estimated: the quantity of the collagen available was found to be equal to 9.7%, indicative of a good state of preservation.

This research has highlighted the importance and potential of a multidisciplinary approach applied to samples of archaeo-anthropological and molecular interest.

Despite the failure to develop a pre-screening method on archaeological bones that could accurately determine the amount of preserved and sequenceable genetic material, a general diagenesis model that accurately describes the alteration of samples for their careful evaluation and selection was outlined.

More sensitive and precise parameters to describe the state of conservation of the osteological finds were identified and, through their application, the hypothesis that the ancient DNA, preserved in the archaeological bones, is preferentially linked to the inorganic component than to the organic one was confirmed and strengthened. In addition, a pre-screening method was devised to identify suitable samples for isotopic, radiocarbon and genetic analysis. In particular for latter case, the research carried out has identified a parameter that provides a predictive response in terms of the presence/absence of aDNA.

Taking into account the results here obtained, the research could be extended by investigating more deeply the *post-mortem* DNA-bioapatite relationship and if there is a chemical interaction functional to its conservation during diagenetic processes. This aspect could consequently lead to shed light on the mechanisms of collagen preservation and whether the diagenetic processes involved can similarly affect the preservation of DNA as well.

Future studies, useful both in molecular-anthropological and forensic fields, could clarify where and how ancient DNA is preserved in bone and how diagenetic alterations affect its preservation. The research could also be extended to paleoanthropological finds, such as Neanderthals and other archaic human relatives.

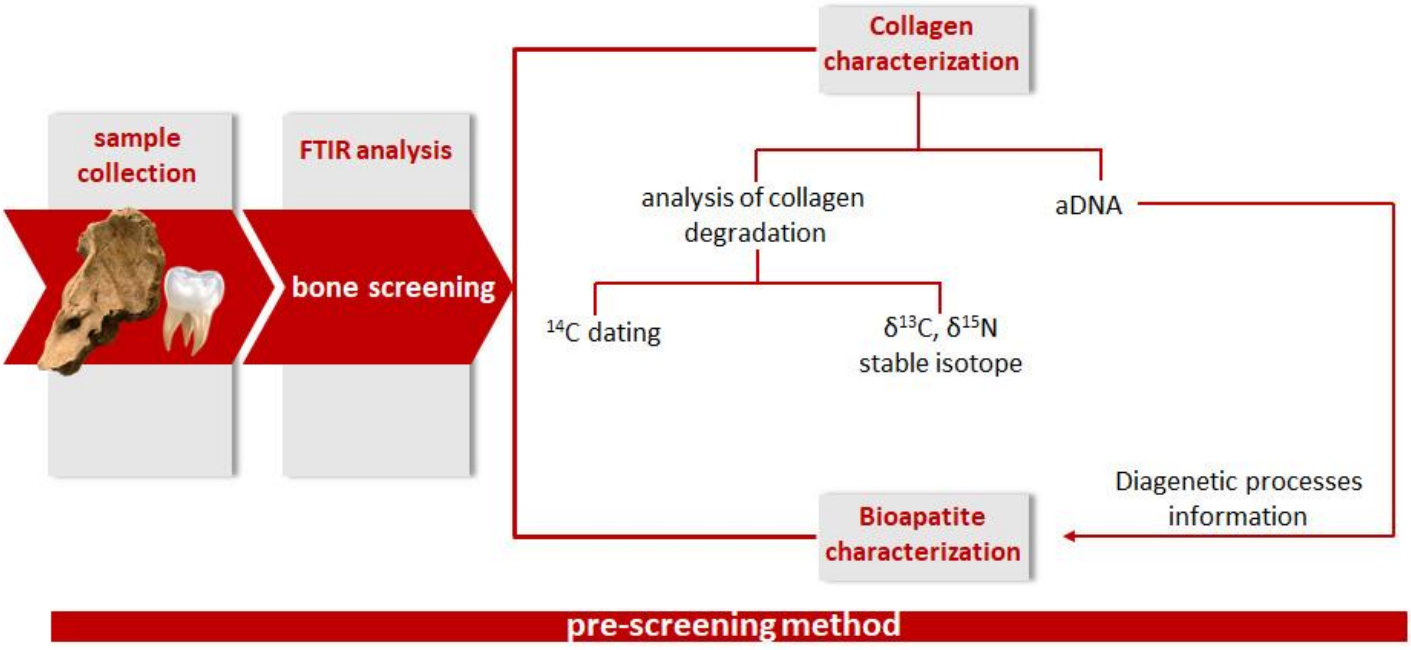

***Figure 1**. Flowchart showing bone pre-screening method*

REFERENCES

Alciati, G., Capitanio, M., Tommaseo Ponzetta, M., 1996. Il Museo di Antropologia ed Etnologia, in: Gregolin C. (a cura di) (Ed.), I Musei, Le Collezioni Scientifiche e Le Sezioni Antiche Delle Biblioteche. Univeristà di Padova, Centro Interdipartimentale di Servizi Musei Scientifici, pp. 89–100.

Allentoft, M.E., Collins, M., Harker, D., Haile, J., Oskam, C.L., Hale, M.L., Campos, P.F., Samaniego, J.A., Gilbert, M.T.P., Willerslev, E., Zhang, G., Scofield, R.P., Holdaway, R.N., Bunce, M., 2012. The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. Proc. R. Soc. B Biol. Sci. 279, 4724–4733. https://doi.org/10.1098/rspb.2012.1745

Analytical Methods Committee, 2006. Representing data distributions with kernel density estimates. AMC Tech. Br. 4, 2.

Berna, F., Matthews, A., Weiner, S., 2004. Solubilities of bone mineral from archaeological sites: The recrystallization window. J. Archaeol. Sci. 31, 867–882. https://doi.org/10.1016/j.jas.2003.12.003

Boskey, A.L., Moore, D.J., Amling, M., Canalis, E., Delany, A.M., 2003. Infrared Analysis of the Mineral and Matrix in Bones of Osteonectin-Null Mice and Their Wildtype Controls. J. Bone Miner. Res. 18, 1005–1011. https://doi.org/10.1359/jbmr.2003.18.6.1005

Bravenec, A.D., Ward, K.D., Ward, T.J., 2018. Amino acid racemization and its relation to geochronology and archaeometry. J. Sep. Sci. 41, 1489–1506. https://doi.org/10.1002/jssc.201701506

Bürgi, T., 2011. Attenuated total reflection infrared (ATR-IR) spectroscopy, modulation excitation spectroscopy (MES), and vibrational circular dichroism (VCD). Biointerface Charact. by Adv. IR Spectrosc. 115–144.https://doi.org/10.1016/B978-0-444-53558-0.00005-9

Canci, A., Chavarria, A., Marinato, M., 2012. Il cimitero della chiesa altomedievale di San Lorenzo di Desenzano (Bs). VI Congr. degli Archeol. Mediev. Ital. Note di Bi, 452–455.

Carrara, N., Scaggion, C., Holland, E., 2018. The Tedeschi collection: A collection of documented and undocumented human skeletal remains at the Museum of Anthropology, Padua University (Italy). Am. J. Phys. Anthropol. 166. https://doi.org/10.1002/ajpa.23471

Cipriano, S., Sandrini, G.., 2001. La banchina fluviale di Opitergium, in: EUT Edizioni Università di Trieste (Ed.), Antichità Altoadriatiche XLVI (2001), Strutture Portuali e Rotte Marittime Nell'Adriatico Di Età Romana. Trieste, pp. 454–455.

Cipriano, S., Sandrini, G.M., 2015. Redditur enim terrae corpus (Cic., De.Leg.,2,56) Necropoli di Oderzo tra media e tarda età Imperiale., in: Federica Rinaldi e Alberto Vigoni (Ed.), Le Necropoli Della Media e Tardaetà Imperiale (Iii-Iv Secolo d.c.) a Iulia Concordia e Nell'arco Altoadriatico.Organizzaizone Spaziale, Aspetti Monumentali e Sociali. Grafiche Turato Edizioni, Rubano (PD), pp. 225–244.

Collins, M.J., Nielsen-Marsh, C.M., Hiller, J., Smith, C.I., Roberts, J.P., Prigodich, R. V., Wess, T.J., Csapò, J., Millard, A.R., Turner-Walker, G., 2002. The survival of organic matter in bone: A review. Archaeometry 44, 383–394. https://doi.org/10.1111/1475-4754.t01-1-00071

Corrain, C., 1986. Resti scheletrici umani da Ordona (Foggia), Secc. VII–IV aC. Les Tombes Dauniennes.Etudes Philol. d'Arch_eologie d'Histoire Ancienne Publi_ees par l'Institut Hist. Belge Rome. Bruxel 24, 787–807.

Corrain, C., Capitanio, M., Erspamer, G., 1972. I resti scheletrici della ne-cropoli di Salapia (Cerignola), secoli IX-III aC. Padova Soc. Coop. Tipogr.

Croce da Villa, P., 1984. La necropoli tardo-antica di Biverone di S. Stino di Livenza, in: Archeologia Veneta. pp. 291–292.

Dal Sasso, G., Lebon, M., Angelini, I., Maritan, L., Usai, D., Artioli, G., 2016. Bone diagenesis variability among multiple burial phases at Al Khiday (Sudan) investigated by ATR-FTIR spectroscopy. Palaeogeogr. Palaeoclimatol. Palaeoecol. 463, 168–179. https://doi.org/10.1016/j.palaeo.2016.10.005

Dal Sasso, G., Maritan, L., Salvatori, S., Mazzoli, C., Artioli, G., 2014a. Discriminating pottery production by image analysis: a case study of Mesolithic and Neolithic pottery from Al Khiday (Khartoum, Sudan). J. Archaeol. Sci. 46, 125–143. https://doi.org/10.1016/j.jas.2014.03.004

Dal Sasso, G., Maritan, L., Usai, D., Angelini, I., Artioli, G., 2014b. Bone diagenesis at the micro-scale: Bone alteration patterns during multiple burial phases at Al Khiday (Khartoum, Sudan) between the Early Holocene and the II century AD. Palaeogeogr. Palaeoclimatol. Palaeoecol. 416, 30–42. https://doi.org/10.1016/j.palaeo.2014.06.034

Donnelly, E., 2011. Methods for assessing bone quality: A review. Clin. Orthop. Relat. Res. 469, 2128–2138. https://doi.org/10.1007/s11999-010-1702-0

Drusini, A., Rippa Bonati, M., Furlan, C., 1987. Oderzo: necropoli in via Spinè in: Riviste e Annali Archeologia. Quad. di Archeol. del Veneto 81.

Figueiredo, M.M., Gamelas, J.A.F., Martins, A.G., 2012. Characterization of Bone and Bone-Based Graft Materials Using FTIR Spectroscopy. Infrared Spectrosc. - Life Biomed. Sci. https://doi.org/10.5772/36379

Gadioli, G., Scaggion, C., Carrara, N., 2018. Anthropological analysis and paleo-demographic study of human skeletal remains from the late ancient necropolis of Biverone (4th-5th c.AD), San Stino di Livenza (Venice, Italy). Anthropol. Rev. 81. https://doi.org/10.2478/anre-2018-0006

Gamba, C., Jones, E.R., Teasdale, M.D., McLaughlin, R.L., Gonzalez-Fortes, G., Mattiangeli, V., Domboróczki, L., Kővári, I., Pap, I., Anders, A., Whittle, A., Dani, J., Raczky, P., Higham, T.F.G., Hofreiter, M., Bradley, D.G., Pinhasi, R., 2014. Genome flux and stasis in a five

millennium transect of European prehistory. Nat. Commun. 5, 5257. https://doi.org/10.1038/ncomms6257

Gilroy, A.M., MacPherson, B.R., Ross, R.M., Schuenke, M., Schulte, E., Schumacher, U., 2008. Temporal Bone & Ear, in: Thieme (Ed.), Atlas of Anatomy. Worcester, MA, Lexington, KY, and Houston, TX, p. 656.

Hansen, H.B., Damgaard, P.B., Margaryan, A., Stenderup, J., Lynnerup, N., Willerslev, E., Allentoft, M.E., 2017. Comparing ancient DNA preservation in petrous bone and tooth cementum. PLoS One 12, 1–18. https://doi.org/10.1371/journal.pone.0170940

Hedges, R.E.M., 2002. Bone diagenesis: an overview of processes. Archaeometry 44, 319–328. https://doi.org/10.1111/1475-4754.00064

Hedges, R.E.M., Millard, A.R., 1995. Bones and Groundwater: Towards the Modelling of Diagenetic Processes. J. Archaeol. Sci. 22, 155–164. https://doi.org/10.1006/jasc.1995.0017

Johnson, B.J., Miller, G.H., 1997. Archaeological applications of amino acid racemization. Archaeometry 39, 265–287.

Klont, B., Damen, J.J.M., ten Cate, J.M., 1991. Degradation of bovine incisor root collagen in an in vitro caries model. Arch. Oral Biol. 36, 299–304. https://doi.org/10.1016/0003-9969(91)90100-9

Kontopoulos, I., Penkman, K., Mullin, V.E., Winkelbach, L., Unterländer, M., Scheu, A., Kreutzer, S., Hansen, H.B., Margaryan, A., Teasdale, M.D., Gehlen, B., Street, M., Lynnerup, N., Liritzis, I., Sampson, A., Papageorgopoulou, C., Allentoft, M.E., Burger, J., Bradley, D.G., Collins, M.J., 2020. Screening archaeological bone for palaeogenetic and palaeoproteomic studies. PLoS One 15, 1–17. https://doi.org/10.1371/journal.pone.0235146

Lebon, M., Reiche, I., Gallet, X., Bellot-Gurlet, L., Zazzo, A., 2016. Rapid quantification of bone collagen content by ATR-FTIR spectroscopy. Radiocarbon 58, 131–145. https://doi.org/10.1017/RDC.2015.11

Ligabue, G., 1973. GROTTINA DEI COVOLONI DEL BROION, in: Preistoria Alpina. pp. 263–264.

Luhmann, T., Robson, S., Kyle, S., Harley, I., 2006. Introducción, Close Range Photogrammetry: Principles, techniques and applications.

Manning, D.A.C., 2000. Carbonates and oxalates in sediments and landfill: monitors of death and decay in natural and artificial systems. J. Geol. Soc. London. 157, 229–238. https://doi.org/10.1144/jgs.157.1.229

Margariti, C., 2019. The application of FTIR microspectroscopy in a non-invasive an non-destructive way to the study and conservation of mineralised excavated textiles. Herit. Sci. 7, 1–14. https://doi.org/10.1186/s40494-019-0304-8

Martínez Cortizas, A., López-Costas, O., 2020. Linking structural and compositional changes

in archaeological human bone collagen: an FTIR-ATR approach. Sci. Rep. 10, 1–14. https://doi.org/10.1038/s41598-020-74993-y

Müller, K., Reiche, I., 2011. Differentiation of archaeological ivory and bone materials by micro-PIXE/PIGE with emphasis on two Upper Palaeolithic key sites: Abri Pataud and Isturitz, France. J. Archaeol. Sci. 38, 3234–3243. https://doi.org/10.1016/j.jas.2011.06.029

Nielsen-Marsh, C.M., Hedges, R.E.M., 2000. Patterns of diagenesis in bone I: The effects of site environments. J. Archaeol. Sci. 27, 1139–1150. https://doi.org/10.1006/jasc.1999.0537

Nielsen-Marsh, C.M., Smith, C.I., Jans, M.M.E., Nord, A., Kars, H., Collins, M.J., 2007. Bone diagenesis in the European Holocene II: taphonomic and environmental considerations. J. Archaeol. Sci. 34, 1523–1531. https://doi.org/10.1016/j.jas.2006.11.012

Obermair, H., Brandstätter, Klaus. Curzel, E., 2006. Dom- und Kollegiatstifte in der Region Tirol – Südtirol – Trentino in Mittelalter und Neuzeit. Collegialità ecclesiastica nella regione trentino- tirolese dal medioevo all'età moderna. Innsbruck.

Ostwald, W., 1885. Lehrbuch der allgemeinen chemie, vol 2, pt. ed. Leipzig, W. Engelmann.

Parker, C., Rohrlach, A.B., Friederich, S., Nagel, S., Meyer, M., Krause, J., Bos, K.I., Haak, W., 2020. A systematic investigation of human DNA preservation in medieval skeletons. Sci. Rep. 10, 18225. https://doi.org/10.1038/s41598-020-75163-w

Paschalis, E.P., Mendelsohn, R., Boskey, A.L., 2011. Infrared assessment of bone quality: A review. Clin. Orthop. Relat. Res. 469, 2170–2178. https://doi.org/10.1007/s11999-010-1751-4

Piga, G., Gonçalves, D., Thompson, T.J.U., Brunetti, A., Malgosa, A., Enzo, S., 2016. Understanding the Crystallinity Indices Behavior of Burned Bones and Teeth by ATR-IR and XRD in the Presence of Bioapatite Mixed with Other Phosphate and Carbonate Phases. Int. J. Spectrosc. 2016, 1–9. https://doi.org/10.1155/2016/4810149

Pinhasi, R., Fernandes, D.M., Sirak, K., Cheronet, O., 2019. Isolating the human cochlea to generate bone powder for ancient DNA analysis. Nat. Protoc. 14, 1194–1205. https://doi.org/10.1038/s41596-019-0137-7

Privat, K.L., O'connell, T.C., Richards, M.P., 2002. Stable Isotope Analysis of Human and Faunal Remains from the Anglo-Saxon Cemetery at Berinsfield, Oxfordshire: Dietary and Social Implications. J. Archaeol. Sci. 29, 779–790. https://doi.org/10.1006/jasc.2001.0785

Reiche, I., Favre-Quattropani, L., Vignaud, C., Bocherens, H., Charlet, L., Menu, M., 2003. A multi-analytical study of bone diagenesis: the Neolithic site of Bercy (Paris, France). Meas. Sci. Technol. 14, 1608–1619. https://doi.org/10.1088/0957-0233/14/9/312

Reiche, I., Vignaud, C., Menu, M., 2002. The crystallinity of ancient bone and dentine: New insights by transmission electron microscopy. Archaeometry 44, 447–459. https://doi.org/10.1111/1475-4754.00077

Richards, M.P., Hedges, R.E.M., 1999. Stable Isotope Evidence for Similarities in the Types of

Marine Foods Used by Late Mesolithic Humans at Sites Along the Atlantic Coast of Europe. J. Archaeol. Sci. 26, 717–722. https://doi.org/10.1006/jasc.1998.0387

Rigoni, M., Bruttomesso, A., 2011. La necropoli dell'ospedale di Montecchio Maggiore, in: Rigoni, M., Bruttomesso, A. (Eds.), Materiali Di Età Longobarda Nel Museo «G. Zannato» Di Montecchio Maggiore. Vol.1. All'Insegna del Giglio, Firenze, pp. 7–107.

Rollo, F., Ubaldi, M., Marota, I., Luciani, S., Ermini, L., 2002. DNA Diagenesis: Effect of Environment and Time on Human Bone. Anc. Biomol. 4, 1–7. https://doi.org/10.1080/13586120290018473

Salvatori, S., Usai, D., 2019. The Mesolithic and Neolithic in Sudan, in: Handbook of Ancient Nubia. De Gruyter, pp. 171–194. https://doi.org/10.1515/9783110420388-009

Salvatori, S., Usai, D., Zerboni, A., 2011. Mesolithic Site Formation and Palaeoenvironment Along the White Nile (Central Sudan). African Archaeol. Rev. 28, 177–211. https://doi.org/10.1007/s10437-011-9095-3

Salvo, L., Cloetens, P., Maire, E., Zabler, S., Blandin, J.J., Buffière, J.Y., Ludwig, W., Boller, E., Bellet, D., Josserond, C., 2003. X-ray micro-tomography an attractive characterisation technique in materials science. Nucl. Instruments Methods Phys. Res. Sect. B Beam Interact. with Mater. Atoms 200, 273–286. https://doi.org/10.1016/S0168-583X(02)01689-0

Scaggion, C., Carrara, N., 2016. New studies on human skeletal remains from the ancient Her- donia (southeast Italy). Evidences of tuberculosis and brucello- sis : two diseases connected with farm animals 12, 85–96.

Shemesh, A., 1990. Crystallinity and diagenesis of sedimentary apatites. Geochim. Cosmochim. Acta 54, 2433–2438. https://doi.org/10.1016/0016-7037(90)90230-I

Smith, C.I., Chamberlain, A.T., Riley, M.S., Stringer, C., Collins, M.J., 2003. The thermal history of human fossils and the likelihood of successful DNA amplification. J. Hum. Evol. 45, 203–217. https://doi.org/10.1016/S0047-2484(03)00106-4

Smith, C.I., Nielsen-Marsh, C.M., Jans, M.M.E., Collins, M.J., 2007. Bone diagenesis in the European Holocene I: patterns and mechanisms. J. Archaeol. Sci. 34, 1485–1493. https://doi.org/10.1016/j.jas.2006.11.006

Soffiantini, G., 2014. Le reliquie di San Teobaldo. Storia e "segni" dalla morte ai nostri giorni.

Thompson, T.J.U., Islam, M., Piduru, K., Marcel, A., 2011. An investigation into the internal and external variables acting on crystallinity index using Fourier Transform Infrared Spectroscopy on unaltered and burned bone. Palaeogeogr. Palaeoclimatol. Palaeoecol. 299, 168–174. https://doi.org/10.1016/j.palaeo.2010.10.044

Trueman, C.N.G., Behrensmeyer, A.K., Tuross, N., Weiner, S., 2004. Mineralogical and compositional changes in bones exposed on soil surfaces in Amboseli National Park, Kenya: Diagenetic mechanisms and the role of sediment pore fluids. J. Archaeol. Sci. 31, 721–739. https://doi.org/10.1016/j.jas.2003.11.003

Turner-Walker, G., 2007. The Chemical and Microbial Degradation of Bones and Teeth. Adv. Hum. Palaeopathology 3–29. https://doi.org/10.1002/9780470724187.ch1

Tuross, N., 1994. The biochemistry of ancient DNA in bone. Experientia 50, 530–535. https://doi.org/10.1007/BF01921721

Usai, D., 2016. A Picture of Prehistoric Sudan. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199935413.013.56

Usai, D., Maritan, L., Dal Sasso, G., Artioli, G., Salvatori, S., Jakob, T., Salviato, T., 2017. Late pleistocene/early holocene evidence of prostatic stones at al khiday cemetery, Central Sudan. PLoS One 12, 1–12. https://doi.org/10.1371/journal.pone.0169524

Usai, D., Salvatori, S., 2019. The Mesolithic period on the White Nile region and the al-Khiday sites. Azania 54, 445–467. https://doi.org/10.1080/0067270X.2019.1691846

Usai, D., Salvatori, S., 2005. THE IsIAO ARCHAEOLOGICAL PROJECT IN THE EL SALHA AREA (OMDURMAN SOUTH, SUDAN): RESULTS AND PERSPECTIVES. Africa Riv. Trimest. Di Stud. e Doc. Dell'Istituto Ital. per l'Africa e l'Oriente 60(3/4), 474–493.

Usai, D., Salvatori, S., Iacumin, P., Di Matteo, A., Jakob, T., Zerboni, A., 2010. Excavating a unique pre-Mesolithic cemetery in Central Sudan. Antiquity 84, 16–18.

Usai, D., Salvatori, S., Jakob, T., David, R., 2014. The Al Khiday cemetery in central Sudan and its "classic/late Meroitic" period graves. J. African Archaeol. 12, 183–204. https://doi.org/10.3213/2191-5784-10254

van Klinken, G.J., Hedges, R.E.M., 1995. Experiments on Collagen-Humic Interactions: Speed of Humic Uptake, and Effects of Diverse Chemical Treatments. J. Archaeol. Sci. 22, 263–270. https://doi.org/10.1006/jasc.1995.0028

Wadsworth, C., Procopio, N., Anderung, C., Carretero, J.-M., Iriarte, E., Valdiosera, C., Elburg, R., Penkman, K., Buckley, M., 2017. Comparing ancient DNA survival and proteome content in 69 archaeological cattle tooth and bone samples from multiple European sites. J. Proteomics 158, 1–8. https://doi.org/10.1016/j.jprot.2017.01.004

Wagermaier, W., Fratzl, P., 2012. Collagen, in: Matyjaszewski, K., Möller, M. (Eds.), Polymer Science: A Comprehensive Reference. Elsevier Science, pp. 35–55.

Wegst, U.G.K., Bai, H., Saiz, E., Tomsia, A.P., Ritchie, R.O., 2015. Bioinspired structural materials. Nat. Mater. 14, 23–36. https://doi.org/10.1038/nmat4089

Weiner, S., Bar-Yosef, O., 1990. States of preservation of bones from prehistoric sites in the Near East: A survey. J. Archaeol. Sci. 17, 187–196. https://doi.org/10.1016/0305-4403(90)90058-D

Williams, M.A.J., Usai, D., Salvatori, S., Williams, F.M., Zerboni, A., Maritan, L., Linseele, V., 2015. Late Quaternary environments and prehistoric occupation in the lower White Nile valley, central Sudan. Quat. Sci. Rev. 130, 72–88. https://doi.org/10.1016/j.quascirev.2015.03.007

Zerboni, A., 2011. Micromorphology reveals in situ Mesolithic living floors and archaeological features in multiphase sites in central Sudan. Geoarchaeology 26, 365–391. https://doi.org/10.1002/gea.20355

APPENDIX

The Genetic Origin of Daunians and the Pan-Mediterranean Southern Italian Iron Age Context

Serena Aneli,[*,†,1] Tina Saupe,[†,2] Francesco Montinaro,[2,3] Anu Solnik,[4] Ludovica Molinaro,[2] Cinzia Scaggion,[5] Nicola Carrara,[6] Alessandro Raveane,[7] Toomas Kivisild,[2,8] Mait Metspalu,[2] Christiana L. Scheib,[‡,2,9] and Luca Pagani[‡,1,2]

[1]Department of Biology, University of Padua, Padova, Italy
[2]Estonian Biocentre, Institute of Genomics, University of Tartu, Tartu, Estonia
[3]Department of Biology-Genetics, University of Bari, Bari, Italy
[4]Core Facility, Institute of Genomics, University of Tartu, Tartu, Estonia
[5]Department of Geosciences, University of Padua, Padova, Italy
[6]Anthropology Museum, University of Padova, Padova, Italy
[7]Laboratory of Hematology-Oncology, European Institute of Oncology IRCCS, Milan, Italy
[8]Department of Human Genetics, KU Leuven, Leuven, Belgium
[9]St John's College, Cambridge, United Kingdom

*Corresponding author: E-mail: serena.aneli@gmail.com.
[†]These authors contributed equally to this work.
[‡]Senior authors.

Associate editor: Maria C. Ávila-Arcos

Abstract

The geographical location and shape of Apulia, a narrow land stretching out in the sea at the South of Italy, made this region a Mediterranean crossroads connecting Western Europe and the Balkans. Such movements culminated at the beginning of the Iron Age with the Iapygian civilization which consisted of three cultures: Peucetians, Messapians, and Daunians. Among them, the Daunians left a peculiar cultural heritage, with one-of-a-kind stelae and pottery, but, despite the extensive archaeological literature, their origin has been lost to time. In order to shed light on this and to provide a genetic picture of Iron Age Southern Italy, we collected and sequenced human remains from three archaeological sites geographically located in Northern Apulia (the area historically inhabited by Daunians) and radiocarbon dated between 1157 and 275 calBCE. We find that Iron Age Apulian samples are still distant from the genetic variability of modern-day Apulians, they show a degree of genetic heterogeneity comparable with the cosmopolitan Republican and Imperial Roman civilization, even though a few kilometers and centuries separate them, and they are well inserted into the Iron Age Pan-Mediterranean genetic landscape. Our study provides for the first time a window on the genetic make-up of pre-Roman Apulia, whose increasing connectivity within the Mediterranean landscape, would have contributed to laying the foundation for modern genetic variability. In this light, the genetic profile of Daunians may be compatible with an at least partial autochthonous origin, with plausible contributions from the Balkan peninsula.

Key words: **ancient DNA, Dauni. Iron Age, Southern Italy.**

Introduction

The Mediterranean's Iron Age populations, between 1100 and 600 BCE, lived in a time of previously unprecedented connectivity (Hodos 2020). Although the technological advances in seafaring had allowed great opportunities for long-distance mobility just a few millennia earlier, it was during the Iron Age that the cosmopolitan role of Mare Nostrum arose, by promoting the spreading of cultures, goods, languages, technical advances as well as heterogeneous ancestral genetic components coming from far and wide (Abulafia 2011; Hodos 2020). Notable examples of such distant connections are the Greek and Phoenician settlements across the Central and Western Mediterranean shores beginning from the 9th and 8th centuries (Fernandes et al. 2020; Marcus et al. 2020).

The Iron Age is conventionally marked as beginning 950 BCE in the Italian peninsula and its Tyrrenian islands (Nijboer 2006; Hodos 2020; Aneli et al. 2021). These areas too joined the cosmopolitan wave sweeping across Southern Europe, by hosting numerous trading posts along its shores. A patchwork of communities appeared in this period within the Italian

Article

borders, each characterized by unique and well-defined cultures and identities, which were later encapsulated and blurred by the Roman colonization. The shift from Republican to Imperial Rome, with the consequent inclusion of non-Roman civilizations from and beyond the Italian peninsula, has already been shown to be connected with a major eastward genetic shift in Central Italian samples (Antonio et al. 2019). Genetic material from Imperial Romans from Central Italy is indeed the first to colocalize with contemporary Italians in a Principal Component space, pointing to this time period (after 27 BCE) as a crucial one in shaping contemporary Italian genetic makeup. Whether such a shift can also be observed in the rest of the Italian peninsula and how the Republic to Imperial transition may have impacted the genetic landscape of local, pre-Roman populations remains an open question.

Despite the numerous written records and archaeological findings, questions about Iron Age populations, their origins, and mutual relationships remain unanswered. Among the many groups occupying Italy in the Iron Age, the Daunians, a Iapygian population from northern Apulia, were first mentioned in the 7–6th century BCE (Lombardo 2014; Norman 2016). Similarly to their neighboring populations, Peucetians and Messapians (living in central and southern Apulia, respectively), the name of the Daunians comes from ancient Greek documents and, given the absence of written Daunian records, the scant information we have on their social, political, and religious life are wholly reliant upon the material record, such as their one-of-a-kind stelae (Norman 2016). For instance, we know that they were mainly farmers, animal breeders, horsemen, and maritime traders with an established trade network extending across the sea with Illyrian tribes (Tagliente 1986; Kirigin et al. 2010; Norman 2016). A fascinating aspect of this population, as opposed to their neighbors in Apulia, was their tenacious resistance to external influences. For instance, they did not acquire either social or cultural Hellenic elements and no Greek alphabet inscriptions have been found in their settlements. Indeed, they retained a strong cultural identity and political autonomy until the Roman arrival in the late 4th–early 3rd century BCE (Norman 2016).

Despite the extensive archaeological literature, their origin has been lost to time and, as early as in the Hellenistic period, various legends already existed connecting them to either Illyria (an ancient region broadly identifiable with the Balkan peninsula), Arkadia (present-day Peloponnesus), or Crete (Lombardo 2014; Norman 2016).

To shed light on this population and provide a glance of the Iron Age genetic landscape of the Southern Italian peninsula, we collected human remains from three Iron Age necropoleis geographically located within the area historically inhabited by Daunians: *Ordona* (ancient Herdonia), one of the largest settlements estimated at 600 hectares (Norman 2016), *Salapia*, and *San Giovanni Rotondo* (fig. 1A). Here, we provide for the first time a genetic investigation of Iron Age Apulian individuals associated with the Daunian culture, offering insight into the genetic landscape of pre-Roman Southern Italy.

Results

We extracted DNA from 34 human skeletal remains (petrous bone = 23 and teeth = 11) from three necropoleis (Ordona = 19; Salapia = 12; San Giovanni Rotondo = 3; hereafter ORD, SAL, and SGR, respectively) at the Ancient DNA Laboratory of the Institute of Genomics, University of Tartu in Estonia (supplementary data 1A, Supplementary Material online), processing between 250 and 1,200 mg of bone material for each sample. All three necropoleis are geographically located less than 50 km from each other in modern Apulia, Southeastern Italy. Ordona and Salapia are archaeologically dated to the Daunian period (6th–3rd century BCE) except one individual (ORD010) archaeologically dated to the Mediaeval time period (which in Europe dates approximately from the 5th to the late 15th century CE) (fig. 1B). Based on the museum record, the samples from the San Giovanni Rotondo necropole have been archaeologically inferred to be from the Iron Age period. After excluding 13 samples due to low DNA concentration (<0.5 ng/μl) and screening 21 libraries at a low depth (±20M reads per library), we additionally sequenced to higher depth 16 libraries with endogenous DNA between 1.81% and 38.82% and mitochondrial DNA-based (mtDNA) contamination estimated at less than 2.89% (supplementary data 1A, Supplementary Material online).

The sequencing runs were merged resulting in 16 individuals for genome-wide analysis: eight from Ordona, five from Salapia, and three from San Giovanni Rotondo. The final data set includes individuals with an average genomic coverage between 0.031 and 0.995× and a number of single-nucleotide polymorphisms (SNPs) overlapping with Human Origins 1240k between ∼40,000 and ∼810,000 (supplementary data 1A, Supplementary Material online).

Out of those 16 individuals, we selected ten individuals based on their proximity within the principal component analysis (PCA) space (fig. 1C) for radiocarbon dating and estimated their age between 1157 and 275 calBCE with a median date of 521 calBCE (supplementary data 1B, Supplementary Material online). The radiocarbon dates confirm the archaeological dates of individuals from Salapia and Ordona as well as the Iron Age affiliation of the two San Giovanni Rotondo samples (SGR002 and SGR003). Two additional samples (ORD010 and SGR001) with a shift toward the Near East in the PCA (fig. 1C) were radiocarbon dated to 995–1156 calCE (95.4%) and 670–774 calCE (95.4%), respectively (supplementary data 1B, Supplementary Material online). Based on their dates, the samples were used as an external control for further analyses focusing on Iron Age Apulia (IAA).

We determined the mitochondrial DNA (mtDNA) haplotype of each individual and the Y chromosome (Ychr) haplogroup for nine male individuals (see Materials and Methods and supplementary data 1A, C, and D, Supplementary Material online). We found that the mtDNA haplotypes mostly belong to the mtDNA lineages H1, H5, K1, and U5, haplotypes found in previous studies of individuals from this time period in the Italian Peninsula (Antonio et al. 2019). Besides the Ychr haplogroup R1b, which is the most frequent

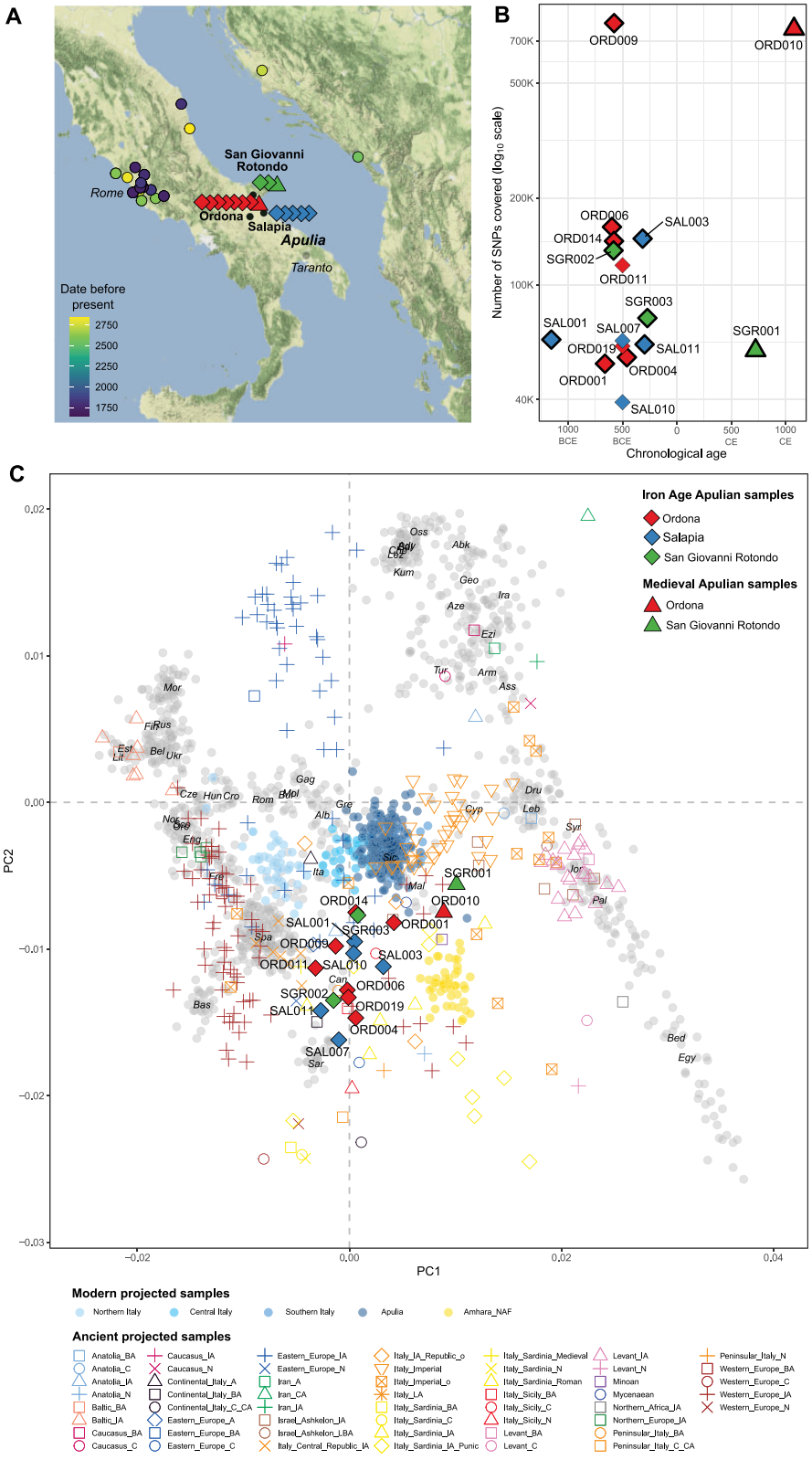

Fig. 1. Geographical location, dating, and genetic structure. (*A*) Geographical location of the newly generated Apulian ancient individuals (green, red, and blue diamonds and triangles) and other published Iron Age samples from Italy and surrounding areas (color scale mirroring their "years before present" dates). (See also table 1 and supplementary data 1 and 2, Supplementary Material online.) (*B*) Dating and number of SNPs from 1240K set covered. Samples with thick edges have been radiocarbon dated (mean presented), whereas for the others, we represented their archaeological dating (table 1 and supplementary data 1*B*, Supplementary Material online). (*C*) PCA of newly generated individuals with reference ancient individuals projected onto the genetic variation inferred from modern-day Western Eurasian populations (supplementary data 2 and 3, Supplementary Material online). For the ancient reference populations which have not been attributed to the Iron Age, we plotted their centroids.

haplogroup during the Bronze Age in the Italian Peninsula and on the islands Sardinia and Sicily (Allentoft et al. 2015; Haak et al. 2015; Antonio et al. 2019; Fernandes et al. 2020; Saupe et al. 2021), we found the Ychr lineages I1-M253, I2d-M223, and J2b-M241. The haplogroup I2d-M223 was one of the main Y chromosome lineages in Western Europe until the Late Neolithic whereas J2b-M241 first appears in the Bronze Age (Allentoft et al. 2015; Mathieson et al. 2015; Schuenemann et al. 2017; Fernandes et al. 2020; Marcus et al. 2020). We found one Early Mediaeval individual (SGR001) belonging to haplogroup I1-M253, which is common in Northern Europe and previously also detected in a 6th Century Langobard burial from North Italy (Amorim et al. 2018).

We used READ (Monroy Kuhn et al. 2018) to identify pairs with first or second degree of genetic relatedness from autosomal data (see Materials and Methods and supplementary data 1*E–H*, Supplementary Material online, for details, supplementary fig. 11, Supplementary Material online). We found one first-degree relationship between the two female individuals ORD001 (age: 9–11 years, 779–544 calBCE; 95.4%) and ORD009 (age: 40–45 years, 749–406 calBCE; 95.4%) from Ordona sharing the identical mtDNA haplogroup H5c and indicating a mother–daughter or full sibling relationship (supplementary data 1 and fig. 1*A* and *B*, Supplementary Material online).

The Making of Modern Italians

To explore the genetic make-up of the IAA population, we performed a PCA projecting the ancient individuals onto the genetic variation of modern Eurasian samples (fig. 1*C* and supplementary data 2 and 3, Supplementary Material online). Our samples are largely scattered between modern peninsular Italians and Sardinians, and, in contrast to what was generally described (Schiffels et al. 2016; Saag et al. 2019) for other European Iron Age populations (e.g., Northern_Europe_IA, Western_Europe_IA, and Levant_IA in fig. 1*C*), they are still clearly distant from the genetic variability of modern-day inhabitants of Apulia. The downward shift of Iron Age Apulians from the present-day ones is further confirmed by the significantly negative *f4(Modern Apulians, IAA; X, Mbuti)*, where *X* is a Neolithic/Chalcolithic/Copper Age population (fig. 2*A* and supplementary data 4, Supplementary Material online). Within the heterogeneity reported by the PCA, which does not mirror the archaeological sites, the two Mediaeval individuals are shifted toward modern Middle Eastern and Caucasus populations (ORD010 and SGR001), whereas the others are stretched along the PC2. This pattern partially mirrors the chronological date with the most recent being more similar to present-day Southern Europeans and is further strengthened when considering the PC3 distribution (supplementary fig. 2, Supplementary Material online). Three samples located at the bottom of the PCA (ORD004, ORD019, SAL007) and one (SAL010) falling in the middle did not include modern Apulians among the top 25 results of an *f3* outgroup analysis (supplementary fig. 3, Supplementary Material online). All of them showed an affinity to Copper and Bronze Age Italians (Saupe et al. 2021) as well as the Aegean and the Mediterranean worlds (including Minoans, Greece, Croatians, and Gibraltar). A similar distribution is mirrored in the multidimensional scaling (MDS) built from the *f3* outgroup measures, where the oldest IAA individual (SAL001: 1260–1048 calBCE; 95.4%) lies farthest from the modern samples, whereas the Mediaeval ones (ORD010 and SGR001) are the closest (supplementary fig. 4, Supplementary Material online). An additional MDS plot is reported in supplementary figure 5, Supplementary Material online: There, we computed the median distances from the centroid of the IAA samples (excluding the two Mediaeval samples), the Roman Imperials and the Republicans, obtaining 0.0093, 0.0140, and 0.0101, respectively. Notably, the distance distributions were not statistically different, thus suggesting that the extent of genetic heterogeneity of IAA individuals is comparable to that of the cosmopolitan Republican and Imperial Roman samples, the latter having experienced a remarkable degree of admixture with foreign Mediterranean groups, pushed by strong geopolitical contingencies (Antonio et al. 2019). We also investigated whether the coverage was somehow responsible for the sample scattering by computing the Pearson correlation between the coverage itself and the distance from the IAA centroid. We found out that the two measures are not correlated ($R = -0.17$, $P = 0.57$), thus showing that most of the sample variation was not due to technical reasons (supplementary fig. 6, Supplementary Material online).

The peculiar positioning of the IAA individuals casts doubt on when the major population shift resulting in modern Italian genetic composition took place. The shift toward the modern Italian genetic variability can be seen with the Republican-Imperial Roman samples (Antonio et al. 2019), the latter being more "similar" to modern Italians (fig. 1*C*). Whether Apulian individuals dating back to the Republican or Imperial phase would also show a repositioning toward modern genetic variability remains an open question, although the later Mediaeval samples of this study point to population changes having taken place after the Iron Age in the area.

The Pan-Mediterranean Genetic Landscape of IAA

The geographic location of Apulia, a narrow peninsula stretching out in the sea at the South of Italy, has made this region an important Mediterranean crossroads connecting Western Europe, the Balkans, the Aegean, and Levant worlds. This is reflected in the PCA where IAA individuals are closely related to other Iron Age populations from the Mediterranean and surrounding areas (e.g., Montenegro, Bulgaria, and Sardinia) (fig. 1*C* and supplementary fig. 4, Supplementary Material online). Nomadic or cosmopolitan groups scatter like IAA: three Punic individuals from Sardinia (Italy_Sardinia_IA_Punic; Marcus et al. 2020), three Moldova Scythians already reported to be genetically similar to Southern Europeans (Krzewińska et al. 2018), Spanish individuals from the Hellenistic and the Romans periods (Olalde et al. 2019), and an individual from the 12th century Iron Age Ashkelon (Feldman, Master, et al. 2019) which clusters with ORD001.

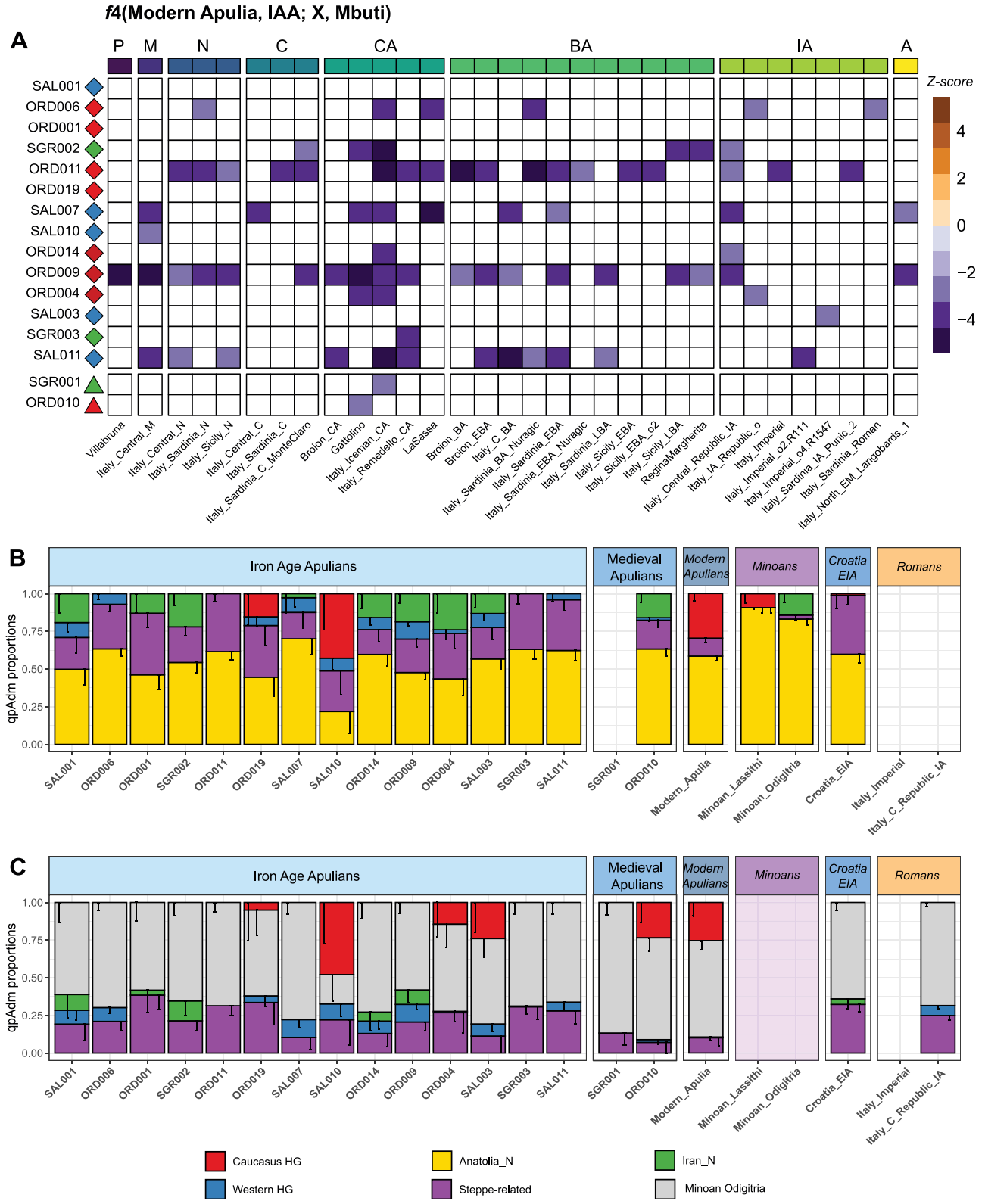

Fig. 2. Genetic relationship of Iron Age and Middle Age Apulians and their ancestral composition. (*A*) Heatmap representing the *Z* score values of *f*4(Modern Apulia, IAA; X, Mbuti), where X is an ancient Italian population. We also added the two Middle Age samples for comparison. Tests with *Z* scores between −3 and 3 or with less than 5,000 SNPs were not included (P, Palaeolithic; M, Mesolithic; N, Neolithic; C, Chalcolithic; CA, Copper Age; BA, Bronze Age; IA, Iron Age; A, Antiquity). (*B*) qpAdm proportions of ancient Apulian samples and other reference populations (Modern Apulians, Minoans, Croatia_EIA, and the Roman individuals from the Republican and the Imperial period) using "base" sources: Caucasus and Western hunter-gatherer component (HG), Anatolia Neolithic (Anatolia_N), Iranian Neolithic (Iran_N), Steppe-related ancestry. (*C*) qpAdm proportions using also the Minoans as source (Materials and Methods). In (*B*) and (*C*), we plotted the model with the highest number of sources and *P* value (complete results may be seen in supplementary fig. 7, Supplementary Material online).

In order to shed light onto the genetic composition of the IAA individuals, we modeled them as a combination of the main ancestries documented across Western Europe at that time: Western Hunter-Gatherers (WHG), Anatolian Neolithic (AN), Steppe-related and, interchangeably, Caucasus Hunter-Gatherers (CHG), or Iranian Neolithic (IN) using the qpWave/qpAdm framework (fig. 2B and supplementary fig. 7A, Supplementary Material online, and Materials and Methods). Broadly, the contributions of such ancestries to the genetic variability of ancient European populations vary according to their geographical positions: in particular, northernmost locations received higher proportions of WHG, Steppe-related ancestry and, consequently, CHG ancestries, whereas Southern European groups carried variable Iranian Neolithic or CHG traces (Lazaridis et al. 2014). In view of this, we observed that although the IAA individuals could generally be modeled as a two-way admixture between AN and Steppe (0.63 ± 0.08 and 0.37 ± 0.08, respectively), the alternative model AN + CHG/IN could also fit for a subset of them, particularly in case of the samples ORD004, ORD010, and SAL010 with higher or comparable *P* values (supplementary fig. 7A, Supplementary Material online, first row with two sources). When three or four sources were tested, the presence of WHG ancestry in the majority of our individuals emerges and together with AN, Steppe, and CHG/IN, forms a supported model for IAA samples (fig. 2B and supplementary fig. 7A and data 5, Supplementary Material online). Notably, for the individuals stretching downward in the PCA (ORD004, ORD019, SGR002, and the Mediaeval ORD010) a three-way admixture involving AN, Steppe, and CHG/IN is generally preferable. To better understand the putative contribution of more recent populations, we modeled our samples with base sources (WHG, AN, Steppe-related, and CHG/IN) and, alternatively, Minoans (Odigitria and Lasithi), Amhara_NAF, and Roman Republicans (fig. 2C and supplementary fig. 7B–D, Supplementary Material online, and Materials and Methods). Amhara_NAF can be used as a proxy for the non-African component in modern Ethiopian individuals that was tentatively linked to the Sea People, a Bronze Age nomadic seafaring population (Feldman, Master, et al. 2019; Molinaro et al. 2019). Together with Minoans and Roman Republicans, this component can be broadly modeled as a Pan-Mediterranean population (constituted by AN and IN/CHG components) with the addition of WHG and Steppe-related ancestry in Roman Republicans. When modeled also with Minoans and Amhara_NAF, which roughly proxies the same ancestral signature, the majority of the samples required an additional CHG/IN contribution (two-way admixtures in supplementary fig. 7B and C, Supplementary Material online) as well as Steppe-related and WHG. We further observed that, as previously seen, the WHG contribution is less clear in those samples stretching downward in the PCA. Although the CHG/IN additional contribution may simply proxy the presence of Steppe-related ancestry in IAA, the absence of which in Minoans has already been reported (Lazaridis et al. 2017), the same cannot be said about Roman Republicans (two-way admixtures in supplementary fig. 7D, Supplementary Material online), which harbored a considerable amount of Steppe component (Antonio et al. 2019). However, this signature is not confirmed with *f*4 analyses (supplementary fig. 8B and C, Supplementary Material online), where just Mycenaean groups report less CHG ancestry than our samples.

Even taking into consideration some uncertainty, especially for the samples with lower coverage, the broader picture emerging from qpAdm analyses is the pervasive presence of coeval Italian ancestries (Rome Republicans), as well as previous Bronze Age sources (Minoans and Sea People) spread far and wide across the Mediterranean Sea. In such a melting pot scenario, the genetic heterogeneity of IAA individuals, who lived in close temporal and geographical proximity, stand out (fig. 1C and supplementary figs. 4 and 9, Supplementary Material online). An example of the cosmopolitan nature of Iron Age Mediterranean is the first-degree relationship between ORD009 and ORD001, whose positions in the PCA strikingly differ, with the individual ORD001 being stretched toward Middle Eastern and Caucasus modern populations, potentially as a consequence, in case of a mother–daughter relationship, of the foreign origin of the father (fig. 1C and supplementary fig. 4, Supplementary Material online). *F*4 analyses in the form *f4(ORD009, ORD001; X, Mbuti)* report a slightly significant (*Z* score between −2 and −3) excess of Greece_N, Portugal_LN_C, Lebanon_Roman, and Italy_Sicily_EBA in ORD001, which may explain its eastward shift (supplementary fig. 10A, Supplementary Material online). Moreover, ORD001, together with SGR002, but not ORD009 harbored more CHG when compared with Lebanon_Hellenistic and Lebanon_IA3 samples, respectively (supplementary fig. 10C, Supplementary Material online).

We also investigated whether the PCA scattering was due to varying African or Levantine contributions with *f*4(*Rome Republican, IAA, Levant_N/YRI, Mbuti)* and tried the same on Mediaeval ancient Apulians (ORD010 and SGR001). However, none of the tested ancient Apulians shows a significant excess of YRI ancestry when compared with the contemporary Roman Republicans, even though ORD014, SAL007, and SAL011 show negative *f*4 values with a *Z* score between −2 and −3 (supplementary fig. 10B, Supplementary Material online).

The Origin of Daunians

The apparent genetic heterogeneity of IAA individuals connected to the Daunian population and the cosmopolitan landscape of Iron Age Mediterranean populations hinder a full reconstruction of the demographic processes leading to the Daunians. Nevertheless, a few milestones can be spotted.

When we performed *f*3 analyses to investigate the nearest possible source for each IAA individual using Minoans, Iron Age Croatians, and the local Roman Republicans (fig. 3A), we found that none of the IAA individuals shows higher affinities with Minoans. Three of them, which clustered close to modern Italians in the PCA (ORD001, ORD014, and SGR003, fig. 1C), show higher affinity with the Iron Age Croatian sample (ORD004 followed this pattern too, but with lower *f*3 values). However, the remaining majority are closest to the Roman Republicans, which can be interpreted as

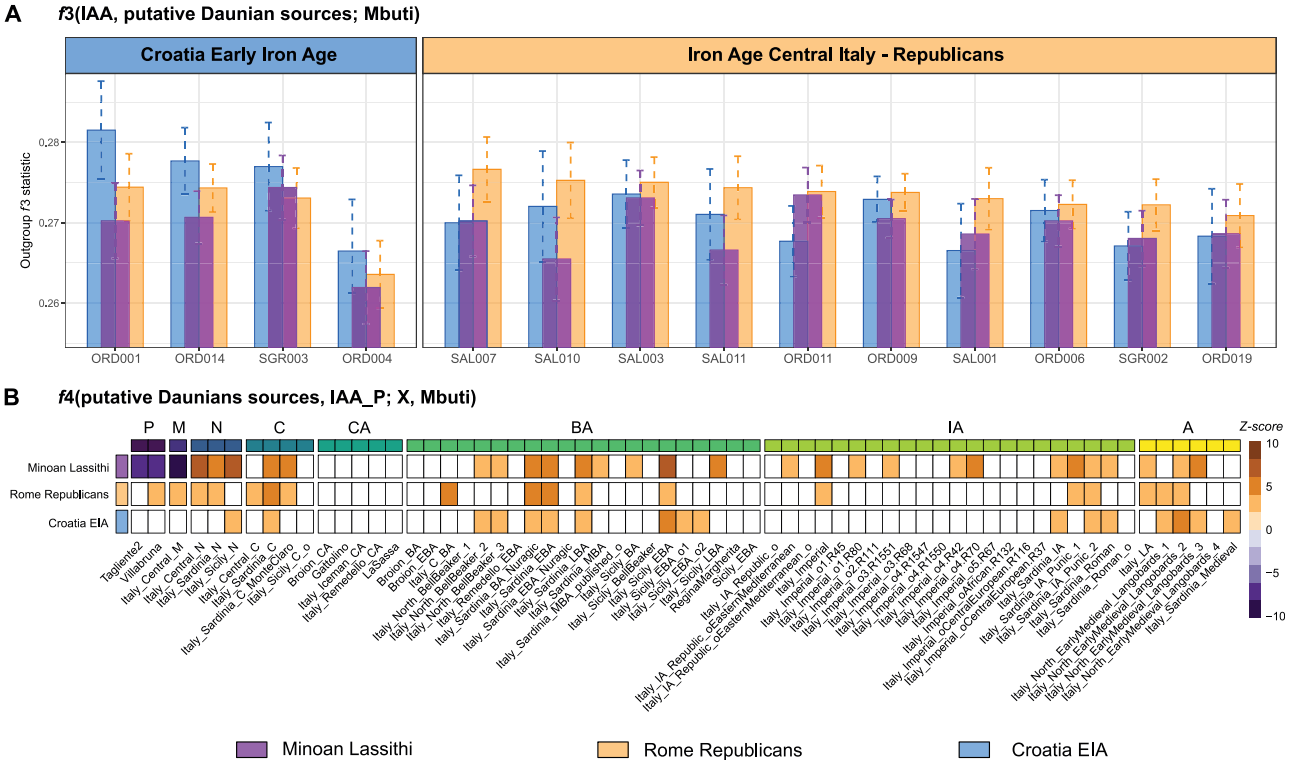

FIG. 3. Genetic affinities of Iron Age Apulian samples with the putative populations of origin: Minoans (Minoan_Lassithi), Illyrians (here proxied by the Croatia_EIA individual), and the Roman Republicans (here proxying the autochthonous Iron Age Italian ancestry). (*A*) Outgroup *f*3 statistics of IAA samples compared with the putative Daunian sources. Samples have been sorted and grouped according to the source whose *f*3 values were higher. (*B*) Heatmap showing the *Z* score values of *f*4(putative sources, IAA_P; X, Mbuti), where X is an ancient Italian population and IAA_P is the entire set of Iron Age Apulian samples taken together. Tests with *Z* scores between −3 and 3 or with less than 5,000 SNPs were not included (P, Palaeolithic; M, Mesolithic; N, Neolithic; C, Chalcolithic; CA, Copper Age; BA, Bronze Age; IA, Iron Age; A, Antiquity).

representative of local Iron Age peninsular Italy ancestry, as also indicated by our MDS results.

Moreover, the WHG contribution, which was a necessary component to explain IAA and Roman Republicans according to qpAdm output, is absent from Minoans and Iron Age Croatians, thus making it a putative signature of an at least partial autochthonous origin (fig. 2*B* and *C* and supplementary fig. 7, Supplementary Material online). These results are confirmed for Minoans, but not for Croatia_EIA, by the *f*4 analyses (fig. 3*B* and supplementary fig. 8*A* and data 6, Supplementary Material online). Indeed, the significantly negative *f*4(*Minoans, IAA; WHG, Mbuti*) (fig. 3*B* and supplementary fig. 8*A*, Supplementary Material online), as well as the *f*4(*Greece_Minoan_Lassithi/Greece_BA_Mycenaean/Greece_BA_Mycenaean_Pylos, IAA; X, Mbuti*), reported a significant excess of WHG ancestry (*X* being Tagliente2, Loschbour, Italy_Central_M, Switzerland_Bichon, and so on) in all IAA individuals, with the exception of ORD004, SAL001, and SGR003 (also the Mediaeval samples ORD010 and SGR001 deviate from this pattern, supplementary data 6, Supplementary Material online). An excess of Bronze Age Steppe-related component, which at that time was already present along the Italian peninsula (Saupe et al. 2021), was clearly present in SGR002, ORD006, and ORD009 by the *f*4(*IAA, Greece_Minoan_Lassithi; X, Mbuti*) (supplementary data 6, Supplementary Material online).

The excess of WHG ancestry, tentatively suggesting a local origin, is somewhat blurred by the genetic similarity of the two most probable sources—Illyrians (Croatia EIA) and an autochthonous one (Roman Republicans), which together make part of the same Mediterranean continuum. Indeed, although no significantly negative *f*4(*Croatia_EIA, IAA; X, Mbuti*) values have been found (with the exception of *X* being Italy_C_BA, Italy_Iceman_CA, Portugal_MBA, and Anatolia_MLBA in SAL007, which retains the highest proportion of AN in qpAdm in fig. 2*B* and supplementary fig. 7, Supplementary Material online), the same pattern is obtained when IAA individuals are compared with Roman Republicans (*f*4(*Roman Republicans, IAA; X, Mbuti*), supplementary data 6, Supplementary Material online). Imperial individuals from a few centuries later (*f*4(*Rome_Imperial, IAA; X, Mbuti*)) show quite the opposite pattern: no positive results have been produced with the exception of ORD004 (X = Lebanon_Roman), SAL001 (X = Italy_Sardinia_Roman_o), SAL010 (X = Russia_LateMaikop), and ORD010. Conversely, the negative *f*4 values point toward WHG, Neolithic, and Bronze Age Steppe-related ancestries (supplementary data 6, Supplementary Material online).

Another signal coming from qpAdm analyses is the apparent excess of CHG ancestry in IAA; however, the predominant contribution of CHG to the Steppe-related ancestry that, by the Iron Age, had already spread to the Mediterranean area

makes it hard to properly detect a CHG signature independent from the Steppe wave, possibly brought by pan-Mediterranean influxes. When directly investigated with an *f4* framework, IAA shows generally more CHG than Mycenaean, less CHG than contemporary Croatian_EIA and, in some cases (ORD019, SGR002, and the Mediaeval SGR001 with *Z* scores higher than 2) more CHG than older Croatian samples (_N = Neolithic and _MN = Middle Neolithic) (supplementary fig. 8, Supplementary Material online).

Discussion

The new genomic sequences of Daunian samples reveal that Iron Age (pre-Roman) Southern Italy (Apulia) can be placed within a Pan-Mediterranean genetic continuum that stretches from Crete (Minoans; Lazaridis et al. 2017) and the Levant (Sea People; Feldman, Master, et al. 2019; Molinaro et al. 2019) to the Republican Rome and the Iberian Peninsula (Antonio et al. 2019), mainly composed by AN and IN/CHG genetic features with the addition of WHG and Steppe-related influences in Continental Italy. Pre-Roman Italian populations, being part of this broader landscape, are not directly superimposable with contemporary Italians which instead seem to be influenced by the homogenizing effect of Imperial Roman and late Antiquity events.

Within the described Pan-Mediterranean landscape, the IAA/Daunians show a sizeable heterogeneity, comparable to the one of the broader and cosmopolitan Republican and even more so Imperial Roman civilization, and the highest genetic affinity to Republican Romans and Iron Age Croatians, whereas Minoans and other Iron Age Greek samples show absent or reduced WHG contribution when compared with IAA. This makes a Cretan or Arkadian origin less likely, even though some tales have connected them with the Greek hero Diomedes and many ancient historians have claimed such origins for their neighbor Messapians and Peucetians.

The Daunians maintained strong commercial and political relations with the Illyrian people, controlling together the area spanning from the Dalmatia to the Gargano peninsula (Kirigin et al. 2010) and had many cultural affinities with them (Nava and Descoeudres 1990). The material culture, involving peculiar anthropomorphic statue stelae, has provided some information on Daunian culture and may also help in unraveling their mysterious origin. In particular, the forearm decorations on a female stela have been interpreted as tattoos and, whereas tattooing practices were considered barbarian among the Greeks (Jones 1987), they were customary in populations from Tracia and Illyria and, more generally, among the women of status from the Balkans (Norman 2011, 2016).

It is not clear whether these connections indicate a movement of people or a sharing of cultural ideas and a conclusive answer to the origin of the Daunians remains elusive. From a parsimony perspective, the genetic results point to an autochthonous origin (e.g., a genetic continuity of Daunians with the population that inhabited the area prior to the examined historical period), here mainly marked by the presence of WHG signature, although we cannot exclude additional influences from Croatia (ancient Illyria), as described by available historical sources and by the material remains (De Juliis 1988; Norman 2016).

Materials and Methods

Samples Information and Ethical Statement

The ancient human remains analyzed in this work belong to the collections of the Museum of Anthropology, University of Padua and have been found and cataloged by different archaeological campaigns carried out during the 20th century. Given the samples were kept at the Museum of Anthropology of the University of Padova, we informed the territorially competent Soprintendenza (Soprintendenza Archeologia, Belle Arti e Paesaggio per l'Area metropolitana di Venezia e per le Province di Belluno, Padova e Treviso) about our sampling strategy and scientific research project, which was approved and our communication assigned the following protocol numbers: prot. 27903 (September 10, 2020) and prot. 29131 (November 23, 2020).

Ordona

The necropolis of Herdonia (today's Ordona, within the Apulian Foggia province in Italy) was studied by different archaeological campaigns interested in the Daunians, Roman as well as the Mediaeval settlements in 1978 and 1981 (Corrain 1986). The human remains collected on such occasions were later extensively cataloged and studied and new paleopathological evidence was also brought back to light (Scaggion and Carrara 2016). Inhabited starting from the Neolithic period, Herdonia became an important Daunian center from the 6th century BCE.

For this study, in collaboration with the University of Padova and Museum of Anthropology of Padova, samples of human remains (petrous bone = 19) were taken. The samples ORD001, ORD004, ORD006, ORD009, ORD010, and ORD014 were dated at ^{14}CHRONO Centre for Climate, the Environment, and Chronology in Belfast, United Kingdom (supplementary data 1*B*, Supplementary Material online).

Salapia

The necropolis of Salapia, an ancient town located 10 km from contemporary Cerignola, within the Apulian Foggia province in Italy. The osteological samples were brought by Prof. Santo Tiné to the Museum of Anthropology of the University of Padova with no further information on the archaeological context, and osteological studies were carried out by Cleto Corrain and colleagues in 1971 (Corrain et al. 1972). From this site, samples of 12 human remains (petrous bone = 5, teeth = 7) were taken. The samples SAL001, SAL003, and SAL011 were dated at ^{14}CHRONO Centre for Climate, the Environment, and Chronology in Belfast, United Kingdom (supplementary data 1*B*, Supplementary Material online).

San Giovanni Rotondo

The San Giovanni Rotondo samples come from the osteo-archaeological collection of the Museum of Anthropology of

the University of Padova and are not associated to any further record with the exception of a broad "Iron Age" archaeological label and may be part of the samples brought to the Museum by Prof. Santo Tiné in the 1960s. From this site, samples of human remains (petrous bone = 3) were taken. The samples SGR001, SGR002, and SGR003 were dated at ^{14}CHRONO Centre for Climate, the Environment, and Chronology in Belfast, United Kingdom (supplementary data 1*B*, Supplementary Material online).

DNA Extraction

In all ancient DNA work, we strive to reduce the impact of destructive sampling and retain as much information for future researchers as possible. We sample individuals with multiple elements present, ideally an antimere to the sample taken (when possible); we take detailed photographs of material prior to sampling; we sample the minimum amount that is necessary and we use protocols that will allow for different types of biomolecular or chemical analysis (either simultaneously or in the future).

All of the laboratory work was performed in dedicated ancient DNA laboratories at the Estonian Biocentre, Institute of Genomics, University of Tartu, Tartu, Estonia. The library quantification and sequencing were performed at the Core Facility of the Institute of Genomics, Tartu, Estonia. The main steps of the laboratory work are detailed below.

In total 34 samples from human remains were extracted for DNA analysis (supplementary data 1, Supplementary Material online). The first layer of pars petrous was removed with a sterilized drill bite to avoid exogenous contamination. A 10-mm core of the inner ear was sampled from the pars petrous. The drill bits and core drill were sterilized in between samples with 6% (w/v) bleach followed by distilled water and then ethanol rinse. Root portions of teeth were removed with a sterile drill wheel.

The root and the petrous portions were soaked in 6% (w/v) bleach for 5 min. Samples were rinsed three times with 18.2 MΩcm H_2O and soaked in 70% (v/v) Ethanol for 2 min. The tubes were shaken during the procedure to dislodge particles. The samples were transferred to a clean paper towel on a rack inside a class IIB hood with the UV light on and allowed to dry for 2–3 h.

Afterward, the samples were weighed to calculate the accurate volume of EDTA (20× EDTA [µl] of sample mass [mg]) and Proteinase K (0.5× Proteinase K [µl] of sample mass [mg]). EDTA and Proteinase K were added into PCR-clean 5 ml or 15 ml conical tubes (Eppendorf) along with the samples inside the IIB hood and the tubes were incubated 72 h on a slow shaker at room temperature.

The DNA extracts (of root portions and pars petrous portions) were concentrated to 250 µl using the Vivaspin Turbo 15 (Sartorius) and purified in large volume columns (High Pure Viral Nucleic Acid Large Volume Kit, Roche) using 2.5 ml of PB buffer, 1 ml of PE buffer, and 100 µl of EB buffer (MinElute PCR Purification Kit, QIAGEN). For the elution of the endogenous DNA, the silica columns were transferred to a collection tube to dry and followed in 1.5 ml DNA lo-bind tubes (Eppendorf) to elute. The samples were incubated with 100 µl EB buffer at 37 °C for 10 min and centrifuged at 13,000 rpm for 2 min. After centrifugation, the silica columns were removed and the samples were stored at −20 °C. Only one extraction was performed per extraction for screening and 30 µl used for libraries.

Library Preparation

Sequencing libraries were built using NEBNext DNA Library Prep Master Mix Set for 454 (E6070, New England Biolabs) and Illumina-specific adaptors (Meyer and Kircher 2010) following established protocols(Meyer and Kircher 2010; Orlando et al. 2013; Malaspinas et al. 2014). The end repair module was implemented using 18.75 µl of water, 7.5 µl of buffer, and 3.75 µl of enzyme mix, incubating at 20°C for 30 min. The samples were purified using 500 µl PB and 650 µl of PE buffer and eluted in 30 µl EB buffer (MinElute PCR Purification Kit, QIAGEN). The adaptor ligation module was implemented using 10 µl of buffer, 5 µl of T4 ligase, and 5 µl of adaptor mix (Meyer and Kircher 2010), incubating at 20°C for 15 min. The samples were purified as in the previous step and eluted in 30 µl of EB buffer (MinElute PCR Purification Kit, QIAGEN). The adaptor fill-in module was implemented using 13 µl of water, 5 µl of buffer, and 2 µl of Bst DNA polymerase, incubating at 37°C for 30 min and at 80°C for 20 min. Libraries were amplified using the following PCR set up: 50 µl DNA library, 1× PCR buffer, 2.5 mM MgCl2, 1 mg/ml BSA, 0.2 µM inPE1.0, 0.2 mM dNTP each, 0.1U/µl HGS Taq Diamond, and 0.2 µM indexing primer. Cycling conditions were: 5′ at 94 °C, followed by 18 cycles of 30 s each at 94, 60, and 68 °C, with a final extension of 7 min at 72 °C. The samples were purified and eluted in 35 µl of EB buffer (MinElute PCR Purification Kit, QIAGEN). Three verification steps were implemented to make sure library preparation was successful and to measure the concentration of dsDNA/sequencing libraries—fluorometric quantitation (Qubit, Thermo Fisher Scientific), parallel capillary electrophoresis (Fragment Analyser, Agilent Technologies), and qPCR.

DNA Sequencing

DNA was sequenced using the Illumina NextSeq500/550 High-Output single-end 75 cycle kit. As a norm, 15 samples were sequenced together on one flow cell; additional data was generated for 16 samples to increase coverage (supplementary data 1, Supplementary Material online).

Mapping

Before mapping, the sequences of the adapters, indexes, and poly-G tails occurring due to the specifics of the NextSeq 500 technology were cut from the ends of DNA sequences using cutadapt-1.11 (Martin 2011). Sequences shorter than 30 bp were also removed with the same program to avoid random mapping of sequences from other species. The sequences were aligned to the reference sequence *GRCh37 (hs37d5)* using Burrows-Wheeler Aligner (BWA 0.7.12) (Li and Durbin 2009) and the command mem with reseeding disabled.

After alignment, the sequences were converted to BAM format and only sequences that mapped to the human genome were kept with samtools 1.3 (Li et al. 2009). Afterward,

the data from different flow cell lanes were merged and duplicates were removed using picard 2.12 (https://github.com/broadinstitute/picard/tree/2.12.1, last accessed January 24, 2022). Indels were realigned using GATK 3.5 (McKenna et al. 2010) and reads with a mapping quality less than 10 were filtered out using samtools 1.3 (Li et al. 2009).

aDNA Authentication

As a result of degradation over time, aDNA can be distinguished from modern DNA by certain characteristics: short fragments and a high frequency of C=>T substitutions at the 5′ ends of sequences due to cytosine deamination. The program mapDamage2.0 (Jónsson et al. 2013) was used to estimate the frequency of 5′ C=>T transitions. Rates of contamination were estimated on mitochondrial DNA by calculating the percentage of nonconsensus bases at haplogroup-defining positions as detailed in (Jones et al. 2017). Each sample was mapped against the *RSRS* downloaded from phylotree.org and checked against haplogroup-defining sites for the sample-specific haplogroup (supplementary data 1A, Supplementary Material online).

Samtools 1.9 (Li et al. 2009) option *stats* was used to determine the number of final reads, average read length, and average coverage. The final average endogenous DNA content of all individuals (proportion of reads mapping to the human genome) was 10.39% (0.09–38.82%) (supplementary data 1A, Supplementary Material online).

Calculating Genetic Sex Estimation

Genetic sex was calculated using the methods described in (Skoglund et al. 2013), estimating the fraction of reads mapping to Y chromosome out of all reads mapping to either X or Y chromosome. Additionally, sex was determined using a method described in (Fu et al. 2016), calculating the X and Y ratio by the division of the coverage by the autosomal coverage. Here, the sex was calculated for samples with a coverage $>0.01\times$ and only reads with a mapping quality >10 were counted for the autosomal, X, and Y chromosomes (supplementary data 1A, Supplementary Material online).

Determining mtDNA Haplogroups

Mitochondrial DNA haplogroups were determined using HaploGrep2 on the command line. For the determination, the reads were realigned to the reference sequence RSRS and the parameter -rsrs were given to estimate the haplogroups using HaploGrep2 (van Oven and Kayser 2009; Weissensteiner et al. 2016) (supplementary data 1A, Supplementary Material online). Subsequently, the identical results between the individuals were checked visually by aligning mapped reads to the reference sequence using samtools 0.1.19 (Li et al. 2009) command *tview* and confirming the haplogroup assignment in PhyloTree. Additionally, private mutations were noted for further kinship analysis. The polymorphisms were estimated using the online platform of HaploGrep2. Here, the variant calling files (vcf) were uploaded to the online platform and the known polymorphisms in the RSRS were converted to rCRS (supplementary data 1D, Supplementary Material online).

Y Chromosome Variant Calling and Haplotyping

A total of 161,140 binary Y chromosome SNPs that have been detected as polymorphic in previous high coverage whole Y chromosome sequencing studies (Hallast et al. 2015; Karmin et al. 2015; Poznik et al. 2016) were called in nine male individuals with more than $0.01\times$ autosomal coverage using ANGSD-0.916 (Hallast et al. 2015; Karmin et al. 2015; Poznik et al. 2016) –doHaploCall option. A subset of 826 sites had at least one of the nine individuals with a derived allele (supplementary data 1C, Supplementary Material online). Basal haplogroup affiliations of each sample were determined by assessing the proportion of derived allele calls in a set of primary haplogroup defining internal branches, as defined in (Karmin et al. 2015), using 1,677 informative sites.

Kinship Analysis

All newly generated individuals with an average human coverage more than $0.03\times$ were selected to assess kinship relationships up to the third degree (Saupe et al. 2021; Wang et al. 2021; Žegarac et al. 2021). We divided the individuals in different groups regarding their geographical area and resulting relationships. First, we were called all individuals using ANGSD-0.916 (Korneliussen et al. 2014) command –doHaploCall to sample a random base for the positions that are present at MAF >0.1 in the 1000 Genomes GBR population (1000 Genomes Project Consortium et al. 2015) giving a total of 4,045,514 SNPs for autosomal kinship analysis. The ANGSD output files were converted to .tped format and used as an input for kinship analyses with READ (Monroy Kuhn et al. 2018), known to be reliable when working with sequencing depth in the order of $0.1\times$ and routinely used in aDNA studies. First, we tested all individuals together and found one first degree relationship between ORD001 and ORD009 and two second degree relationships between ORD004, SAL007, and SAL011. We tested those relationships using all individuals from Ordona including SAL007 and SAL011 and all individuals from Salapia including ORD004. We compared the estimated mtDNA haplogroups, radiocarbon dates, genetic sexes, average human coverages, and the age of death from the individuals to confirm or dismiss relationships (supplementary data 1 and fig. 1, Supplementary Material online). Based on the outputs, we dismissed the possible second degree relationships.

Data Set Preparation and Preprocessing for Autosomal Analysis

We assembled a genome-wide data set of ancient and modern samples by converting the following data sets, where needed, in PLINK format using *convertf* from the EIGENSOFT 8.0.0 (Patterson et al. 2006) and merging them together with PLINK 1.9 (Chang et al. 2015): 1) the ancient individuals and the Mbuti Congolese individuals (HGDP; Patterson et al. 2012) from the "1240K" data set from Dr David Reich laboratory (https://reich.hms.harvard.edu/allen-ancient-dna-resource-aadr-downloadable-genotypes-present-day-and-ancient-dna-data, version 44.3 release, last accessed January 21, 2021), 2) the modern samples from "1240K+HO" of the same release, 3) the Italian Chalcolithic and Bronze Age

samples produced by a recent study (Saupe et al. 2021), 4) the genotypes of an Italian hunter-gatherer from the Late Epigravettian site of Riparo Tagliente dated around 17kya (Bortolini et al. 2021), 5) additional genome-wide data of 129 modern-day Italian samples covering the entire country (Raveane et al. 2019), 6) genome-wide data of 217 present-day Apulian individuals (Sallustio et al. 2015), of which 41 individuals (unrelated and filtered for PCA outliers) were used for qpadm and *f*3 analyses, and 7) 48 haploid genomes representing the Eurasian component of modern Ethiopians (here called Amhara_NAF for "Non African Fraction"; Molinaro et al. 2019). We excluded the ancient samples showing issues in their group assignment (i.e., "*Ignore*" in the "1240K" data set) or in their DNA quality (e.g., those not containing the string "PASS," with the exception of the *Iceman* individual, or those with less than 5,000 SNPs) and one of a pair of related samples. We further refined the list of ancient samples by retaining only those coming from geographical regions relevant for this study (supplementary data 2, Supplementary Material online) (Keller et al. 2012; Gamba et al. 2014; Lazaridis et al. 2014, 2016, 2017; Allentoft et al. 2015; Günther et al. 2015; Jones et al. 2015, 2017; Mathieson et al. 2015, 2018; Olalde et al. 2015, 2018, 2019; Broushaki et al. 2016; Cassidy et al. 2016; Fu et al. 2016; Hofmanová et al. 2016; Kılınç et al. 2016; Martiniano et al. 2016, 2017; Omrak et al. 2016; Schiffels et al. 2016; González-Fortes et al. 2017, 2019; Haber et al. 2017, 2019, 2020; Lipson et al. 2017; Saag et al. 2017, 2019; Schuenemann et al. 2017; Unterländer et al. 2017; van den Brink et al. 2017; Amorim et al. 2018; de Barros Damgaard, Marchi, et al. 2018; de Barros Damgaard, Martiniano, et al. 2018; Fregel et al. 2018; Harney et al. 2018; Krzewińska et al. 2018; Mittnik et al. 2018; Valdiosera et al. 2018; Veeramah et al. 2018; Zalloua et al. 2018; Antonio et al. 2019; Brace et al. 2019; Feldman, Fernández-Domínguez, et al. 2019; Feldman, Master, et al. 2019; Järve et al. 2019; Narasimhan et al. 2019; Sikora et al. 2019; Villalba-Mouco et al. 2019; Wang et al. 2019; Agranat-Tamir et al. 2020; Brunel et al. 2020; Fernandes et al. 2020; Furtwängler et al. 2020; Gokhman et al. 2020; Marcus et al. 2020; Margaryan et al. 2020; Rivollat et al. 2020; Skourtanioti et al. 2020). The same geographical filter was applied on the modern samples from "1240K+HO" and only those flagged as "*PASS (genotyping)*" have been selected in order to minimize variation due to different genotyping techniques (supplementary data 3, Supplementary Material online) (Patterson et al. 2012; Lazaridis et al. 2014, 2016; Biagini et al. 2019; Jeong et al. 2019). In case of duplicated samples, we selected the item showing the highest number of SNPs among the 1240K SNP set. Finally, we added the newly generated ancient Apulian samples to the merge.

We used two different SNPs set for the following analyses. The autosomal "1240K" SNP set (1,150,639 SNPs) was used for the analyses involving just ancient samples, such as the *f*-statistics and qpAdm analyses (see below). Conversely, the autosomal "HO" SNP set was used for those analyses relying also on modern populations (e.g., PCA and ADMIXTURE). In the latter case, we further refine the "HO" SNP set by removing monomorphic SNPs and those with more than 5% missing count in modern populations with PLINK, thus obtaining 503,062 SNPs.

Principal Component Analysis

We performed PCA using the program smartpca implemented in EIGENSOFT software 8.0.0 (Patterson et al. 2006) with the parameters "*lsqproject*" and "*shrinkmode*." In particular, we projected the ancient samples' genetic variation onto the principal components inferred from the modern samples of the "1240K+HO" data set. Modern Italian samples coming from Raveane et al. (2019), the modern Apulians from Sallustio et al. (2015), and the Amhara_NAF were also projected.

Admixture

We used the same data set of modern and ancient individuals presented in the PCA to explore the genetic components of each individual. For the analysis, we exploited the model-based algorithm implemented in ADMIXTURE (Alexander et al. 2009) projecting ancient individuals (-P flag) into the genetic structure calculated on the modern data set, due to missing data in the ancient individuals. We performed unsupervised Admixture for $K \in \{2.10\}$ for modern individuals and used the "per-cluster" inferred allele frequencies to project the ancient individuals. We visualized the Q output using R 3.6 (R Core Team 2021).

*f*4 Statistics

We performed a series of *f*4 statistics using the program *qpDstat* (option f4mode) implemented in the software ADMIXTOOLS 7.0.2 (Patterson et al. 2012) in different forms. *f*4 statistics are computed for a set of four different populations (W, X, Y, Z) and the sign of the resultant Z score indicates the direction of the gene flow: if it is positive, the gene flow occurred either between W and Y or X and Z, whereas if it is negative the gene flow occurred either between W and Z or X and Y). We grouped together ancient published samples from the same cultural/archaeological assemblages (labels used for *f*3 and *f*4 analyses can be found in the column "Analysis.Label" of supplementary data 2, Supplementary Material online). Conversely, ancient Apulian samples were both analyzed as individual samples and grouped together according to their archaeological age (Iron Age samples were grouped together under the acronym "IAA," Iron Age Apulians). About 10 individuals from the Mbuti population from Congo coming from the "1240K" data set were used as an outgroup (population Z).

Outgroup *f*3 Statistics

To investigate genetic affinities between ancient Apulian samples with other ancient and modern human populations, we performed a series of Outgroup *f*3 analyses in the form *f*3(X, Y; outgroup = Mbuti) using the qp3Pop function implemented in ADMIXTOOLS 7.0.2 (Patterson et al. 2012). As above-described, we used the "Analysis.Label" groups for the ancient samples, whereas we used the "Group.Label" already present in the "1240K+HO" annotation file for modern samples.

Table 1. Archaeological Information, Dating, Genome Coverage, Genetic Sex, mtDNA, and Y Chromosome Haplogroups of the Individuals of This Study Selected for Genome-Wide Analysis.

Individual	Site	Date	Genome Coverage	Gen. Sex	mtDNA HG	Y Chromosome HG	No. SNPs
ORD001	Ordona	2516 ± 25 BP; 779–544 calBCE	0.0396	XX	H5c	—	53,277
ORD004	Ordona	2374 ± 21 BP; 536–384 calBCE	0.0429	XX	U8b1b1	R1b-M269	56,070
ORD006	Ordona	2476 ± 29 BP; 770–425 calBCE	0.12	XY	H + 16291T	—	158,562
ORD009	Ordona	2433 ± 27 BP; 749–406 calBCE	0.995	XX	H5c	—	805,966
ORD010	Ordona	984 ± 25 BP; 995–1156 calCE	0.91	XX	X2i	—	769,336
ORD011	Ordona	Not dated	0.0889	XY	H1e	R1b-P312	116,800
ORD014	Ordona	2438 ± 25 BP; 749–408 calBCE	0.108	XY	I5a2 + 16086C	J2b2-L283	141,736
ORD019	Ordona	Not dated	0.0484	XY	T2e	I2d-Z2093/Y3670	61,292
SAL001	Salapia	2946 ± 30 BP; 1260–1048 calBCE	0.049	XY	H1 + 16189!	J2b-M241	64,574
SAL003	Salapia	2287 ± 24 BP; 401–354 calBCE	0.113	XX	K1a2a	—	144,490
SAL007	Salapia	Not dated	0.0529	XX	K1a + 195T!	—	64,034
SAL010	Salapia	Not dated	0.031	XY	U5a1	J2b-M241	39,174
SAL011	Salapia	2241 ± 23 BP; 389–206 calBCE	0.0496	XY	U5b1	I2d-M223	62,155
SGR001	San Giovanni Rotondo	1285 ± 23 BP; 670–774 calCE	0.044	XY	U3a	I1-M253	59,213
SGR002	San Giovanni Rotondo	2451 ± 22 BP; 750–415 calBCE	0.102	XY	U5b1d1	R1b-M269	131,437
SGR003	San Giovanni Rotondo	2205 ± 26 BP; 368–176 calBCE	0.0572	XX	H1 + 16311T!	—	76,687

NOTE.—Dates in BP are raw radiocarbon dates, calibrated dates are the 95.4% probability and were calibrated using IntCal20 (Reimer et al. 2020). See also supplementary data 1, Supplementary Material online. Gen., genetic; HG, haplogroup.

We also computed the distance matrices from the Outgroup *f*3 results comparing ancient Apulian samples (both Iron and Middle Age samples) with other published ancient samples by subtracting the *f*3 values from 1 and we performed a classical (metric) MDS with default options using the function *cmdscale* in the stats package of R (version 4.0.4; R Core Team 2021).

qpAdm

We exploited the qpWave/qpAdm framework within the software ADMIXTOOLS 7.0.2 to investigate the genetic composition of ancient Apulian samples and other ancient samples of interest with respect to human ancestries that contributed to coeval European population in the Mediterranean area. In particular, we tested a number of left sources from 2 to 5 in combination with a list of reference sources (right). We used the options "allsnps = YES" and "inbreed = NO" (the last option was used because some populations were composed of just one haploid individual). We tested four different sets of left sources. The "base" set included Western hunter-gatherers (*Loschbour_published.DG*), Anatolian Neolithic (26 samples flagged as "Anatolia_N" in the Analysis.Label column of supplementary data 2, Supplementary Material online), Iranian Neolithic (samples Iran_GanjDareh_N), Caucasus hunter-gatherers (KK1.SG and SATP.SG), and Yamnaya (10 Russia_Samara_EBA_Yamnaya samples). To this set of base sources, we separately added as left sources: Minoans (both Greece_Minoan_Lassithi and Greece_Minoan_Odigitria, but taken individually), Amhara_NAF, and Roman Republican samples (Italy_Central_Republic_IA). In each run, we consistently used the same set of right sources (if present in the left set, they were removed from the right): Russia_Ust_Ishim_HG_published.DG (which was always kept as the first of the list), Russia_Kostenki14, Russia_AfontovaGora3, Eastern hunter-gatherers (I0061 = Russia_HG_Karelia, I0124 = Russia_HG_Samara, I0211 = Russia_HG_Karelia, and UzOO77 = Russia_EHG), Spain_ElMiron, Belgium_UP_GoyetQ116_1_published_all, Levant_N (Israel_PPNB), Russia_MA1_HG.SG, Israel_Natufian_published, Czech_Vestonice16, and Caucasus hunter-gatherers.

Ethical Statement

All ancient samples newly sequenced for this study are kept at the Museum of Anthropology of the University of Padova, hence we informed the territorially competent Soprintendenza (Soprintendenza Archeologia, Belle Arti e Paesaggio per l'Area metropolitana di Venezia e per le Province di Belluno, Padova e Treviso) about our sampling strategy and scientific research project, which was approved and our communication assigned the following protocol numbers: prot. 27903 (September 10, 2020) and prot. 29131 (November 23, 2020). We sampled individuals with multiple elements present, ideally an antimere to the sample taken (when possible) and took detailed photographs of material prior to performing the micro destructive sampling. Furthermore, we sampled the minimum amount that is necessary and we used protocols that allow for different types of biomolecular or chemical analysis (either simultaneously or in the future) to be carried out on the same bone material.

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

The study was supported by the Estonian Research Council grant PUT (PRG243) (A.S., T.K., M.M., C.L.S); the European Union through the European Regional Development Fund (Project No. 2014-2020.4.01.16-0030) (C.L.S., M.M., T.S., F.M., L.M.); the European Regional Development Fund (Project No. 2014-2020.4.01.15-0012) (M.M.); and UniPd PRID 2019 (S.A., L.P.). The authors would like to thank Dr Francesco Bertolini

for facilitating the research of A.R. in the last stage of the manuscript preparation.

Author Contributions

Conceptualization: L.P., S.A., and C.L.S. Formal analysis: S.A. and T.S. Funding acquisition: M.M. and L.P. Investigation: T.S. (aDNA data generation) and A.S. (quality control). Project administration: L.P. and C.L.S. Resources: C.L.S., N.C., and L.M. Supervision: L.P. and C.L.S. Visualization: S.A. and T.S. Writing—original draft: S.A., T.S., and L.P. Writing—review and editing: all authors.

Data Availability

The DNA sequences generated during this study are available at the European Nucleotide Archive (ENA) at the accession number PRJEB47831. The data are also available in PLINK format through the data depository of the EBC-IG (http://www.ebc.ee/free_data/AneliSaupe_2021/).

References

1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, et al. 2015. A global reference for human genetic variation. *Nature* 526(7571):68–74.

Abulafia D. 2011. The great sea: a human history of the Mediterranean. London: Allen Lane.

Agranat-Tamir L, Waldman S, Martin MAS, Gokhman D, Mishol N, Eshel T, Cheronet O, Rohland N, Mallick S, Adamski N, et al. 2020. The genomic history of the Bronze Age Southern Levant. *Cell* 181(5):1146–1157.e11.

Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 19(9):1655–1664.

Allentoft ME, Sikora M, Sjögren K-G, Rasmussen S, Rasmussen M, Stenderup J, Damgaard PB, Schroeder H, Ahlström T, Vinner L, et al. 2015. Population genomics of Bronze Age Eurasia. *Nature* 522(7555):167–172.

Amorim CEG, Vai S, Posth C, Modi A, Koncz I, Hakenbeck S, La Rocca MC, Mende B, Bobo D, Pohl W, et al. 2018. Understanding 6th-century barbarian social organization and migration through paleogenomics. *Nat Commun*. 9(1):3547.

Aneli S, Caldon M, Saupe T, Montinaro F, Pagani L. 2021. Through 40,000 years of human presence in Southern Europe: the Italian case study. *Hum Genet*. 140(10):1417–1431.

Antonio ML, Gao Z, Moots HM, Lucci M, Candilio F, Sawyer S, Oberreiter V, Calderon D, Devitofranceschi K, Aikens RC, et al. 2019. Ancient Rome: a genetic crossroads of Europe and the Mediterranean. *Science* 366(6466):708–714.

Biagini SA, Solé-Morata N, Matisoo-Smith E, Zalloua P, Comas D, Calafell F. 2019. People from Ibiza: an unexpected isolate in the Western Mediterranean. *Eur J Hum Genet*. 27(6):941–951.

Bortolini E, Pagani L, Oxilia G, Posth C, Fontana F, Badino F, Saupe T, Montinaro F, Margaritora D, Romandini M, et al. 2021. Early Alpine occupation backdates westward human migration in Late Glacial Europe. *Curr Biol*. 31(11):2484–2493.e7.

Brace S, Diekmann Y, Booth TJ, van Dorp L, Faltyskova Z, Rohland N, Mallick S, Olalde I, Ferry M, Michel M, et al. 2019. Author correction: ancient genomes indicate population replacement in Early Neolithic Britain. *Nat Ecol Evol*. 3(6):986–987.

Broushaki F, Thomas MG, Link V, López S, van Dorp L, Kirsanow K, Hofmanová Z, Diekmann Y, Cassidy LM, Díez-Del-Molino D, et al. 2016. Early Neolithic genomes from the eastern Fertile Crescent. *Science* 353(6298):499–503.

Brunel S, Bennett EA, Cardin L, Garraud D, Barrand Emam H, Beylier A, Boulestin B, Chenal F, Ciesielski E, Convertini F, et al. 2020. Ancient genomes from present-day France unveil 7,000 years of its demographic history. *Proc Natl Acad Sci U S A*. 117(23):12791–12798.

Cassidy LM, Martiniano R, Murphy EM, Teasdale MD, Mallory J, Hartwell B, Bradley DG. 2016. Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. *Proc Natl Acad Sci U S A*. 113(2):368–373.

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 4:7.

Corrain C. 1986. Resti scheletrici umani da Ordona (Foggia), Secc. VII–IV aC. In: Iker R, editor. Ordona VII/2. Les Tombes Dauniennes. Vol. 24. Etudes de Philologie, d'Archéologie et d'Histoire Ancienne Publiées par l'Institut Historique Belge de Rome. Bruxelles, Rome: Academia Belgica. p. 787–807.

Corrain C, Capitanio M, Erspamer G. 1972. I resti scheletrici della necropoli di Salapia (Cerignola), secoli IX-III aC. Padova: Società Cooperativa Tipografica.

de Barros Damgaard B, Marchi N, Rasmussen S, Peyrot M, Renaud G, Korneliussen T, Moreno-Mayar JV, Pedersen MW, Goldberg A, Usmanova E, et al. 2018. 137 ancient human genomes from across the Eurasian steppes. *Nature* 557(7705):369–374.

de Barros Damgaard P, Martiniano R, Kamm J, Víctor Moreno-Mayar J, Kroonen G, Peyrot M, Barjamovic G, Rasmussen S, Zacho C, et al. 2018. The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science* 360(6396):eaar7711.

De Juliis EM. 1988. Gli Iapigi: storia e civiltà della Puglia preromana. Milano: Longanesi & C.

Feldman M, Fernández-Domínguez E, Reynolds L, Baird D, Pearson J, Hershkovitz I, May H, Goring-Morris N, Benz M, Gresky J, et al. 2019. Late Pleistocene human genome suggests a local origin for the first farmers of central Anatolia. *Nat Commun*. 10:1218.

Feldman M, Master DM, Bianco RA, Burri M, Stockhammer PW, Mittnik A, Aja AJ, Jeong C, Krause J. 2019. Ancient DNA sheds light on the genetic origins of early Iron Age Philistines. *Sci Adv*. 5(7):eaax0061.

Fernandes DM, Mittnik A, Olalde I, Lazaridis I, Cheronet O, Rohland N, Mallick S, Bernardos R, Broomandkhoshbacht N, Carlsson J, et al. 2020. The spread of steppe and Iranian-related ancestry in the islands of the western Mediterranean. *Nat Ecol Evol*. 4(3):334–345.

Fregel R, Méndez FL, Bokbot Y, Martín-Socas D, Camalich-Massieu MD, Santana J, Morales J, Ávila-Arcos MC, Underhill PA, Shapiro B, et al. 2018. Ancient genomes from North Africa evidence prehistoric migrations to the Maghreb from both the Levant and Europe. *Proc Natl Acad Sci U S A*. 115(26):6774–6779.

Fu Q, Posth C, Hajdinjak M, Petr M, Mallick S, Fernandes D, Furtwängler A, Haak W, Meyer M, Mittnik A, et al. 2016. The genetic history of Ice Age Europe. *Nature* 534(7606):200–205.

Furtwängler A, Rohrlach AB, Lamnidis TC, Papac L, Neumann GU, Siebke I, Reiter E, Steuri N, Hald J, Denaire A, et al. 2020. Ancient genomes reveal social and genetic structure of Late Neolithic Switzerland. *Nat Commun*. 11(1):1915.

Gamba C, Jones ER, Teasdale MD, McLaughlin RL, Gonzalez-Fortes G, Mattiangeli V, Domboróczki L, Kővári I, Pap I, Anders A, et al. 2014. Genome flux and stasis in a five millennium transect of European prehistory. *Nat Commun*. 5(1):5257.

Gokhman D, Nissim-Rafinia M, Agranat-Tamir L, Housman G, García-Pérez R, Lizano E, Cheronet O, Mallick S, Nieves-Colón MA, Li H, et al. 2020. Differential DNA methylation of vocal and facial anatomy genes in modern humans. *Nat Commun*. 11(1):1189.

González-Fortes G, Jones ER, Lightfoot E, Bonsall C, Lazar C, Grandal-d'Anglade A, Garralda MD, Drak L, Siska V, Simalcsik A, et al. 2017. Paleogenomic evidence for multi-generational mixing between Neolithic farmers and mesolithic hunter-gatherers in the lower Danube Basin. *Curr Biol*. 27(12):1801–1810.e10.

González-Fortes G, Tassi F, Trucchi E, Henneberger K, Paijmans JLA, Díez-del-Molino D, Schroeder H, Susca RR, Barroso-Ruíz C, Bermudez FJ, et al. 2019. A western route of prehistoric human migration from Africa into the Iberian Peninsula. *Proc Biol Sci*. 286(1895):20182288.

Günther T, Valdiosera C, Malmström H, Ureña I, Rodriguez-Varela R, Sverrisdóttir ÓO, Daskalaki EA, Skoglund P, Naidoo T, Svensson EM,

et al. 2015. Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. *Proc Natl Acad Sci U S A*. 112(38):11917–11922.
Haak W, Lazaridis I, Patterson N, Rohland N, Mallick S, Llamas B, Brandt G, Nordenfelt S, Harney E, Stewardson K, et al. 2015. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* 522(7555):207–211.
Haber M, Doumet-Serhal C, Scheib C, Xue Y, Danecek P, Mezzavilla M, Youhanna S, Martiniano R, Prado-Martinez J, Szpak M, et al. 2017. Continuity and admixture in the last five millennia of Levantine history from ancient Canaanite and present-day Lebanese genome sequences. *Am J Hum Genet*. 101(2):274–282.
Haber M, Doumet-Serhal C, Scheib CL, Xue Y, Mikulski R, Martiniano R, Fischer-Genz B, Schutkowski H, Kivisild T, Tyler-Smith C. 2019. A transient pulse of genetic admixture from the crusaders in the near east identified from ancient genome sequences. *Am J Hum Genet*. 104(5):977–984.
Haber M, Nassar J, Almarri MA, Saupe T, Saag L, Griffith SJ, Doumet-Serhal C, Chanteau J, Saghieh-Beydoun M, Xue Y, et al. 2020. A genetic history of the Near East from an aDNA time course sampling eight points in the past 4,000 years. *Am J Hum Genet*. 107(1):149–157.
Hallast P, Batini C, Zadik D, Maisano Delser P, Wetton JH, Arroyo-Pardo E, Cavalleri GL, de Knijff P, Destro Bisol G, Dupuy BM, et al. 2015. The Y-chromosome tree bursts into leaf: 13,000 high-confidence SNPs covering the majority of known clades. *Mol Biol Evol*. 32(3):661–673.
Harney É, May H, Shalem D, Rohland N, Mallick S, Lazaridis I, Sarig R, Stewardson K, Nordenfelt S, Patterson N, et al. 2018. Publisher correction: ancient DNA from Chalcolithic Israel reveals the role of population mixture in cultural transformation. *Nat Commun*. 9(1):3913.
Hodos T. 2020. The archaeology of the Mediterranean Iron Age: a globalising world c.1100–600 BCE. Cambridge: Cambridge University Press.
Hofmanová Z, Kreutzer S, Hellenthal G, Sell C, Diekmann Y, Díez-Del-Molino D, van Dorp L, López S, Kousathanas A, Link V, et al. 2016. Early farmers from across Europe directly descended from Neolithic Aegeans. *Proc Natl Acad Sci U S A*. 113(25):6886–6891.
Järve M, Saag L, Scheib CL, Pathak AK, Montinaro F, Pagani L, Flores R, Guellil M, Saag L, Tambets K, et al. 2019. Shifts in the genetic landscape of the Western Eurasian Steppe associated with the beginning and end of the Scythian dominance. *Curr Biol*. 29(14):2430–2441.e10.
Jeong C, Balanovsky O, Lukianova E, Kahbatkyzy N, Flegontov P, Zaporozhchenko V, Immel A, Wang C-C, Ixan O, Khussainova E, et al. 2019. The genetic history of admixture across inner Eurasia. *Nat Ecol Evol*. 3(6):966–976.
Jones CP. 1987. Stigma: tattooing and branding in Graeco-Roman antiquity. *J Rom Stud*. 77:139–155.
Jones ER, Gonzalez-Fortes G, Connell S, Siska V, Eriksson A, Martiniano R, McLaughlin RL, Gallego Llorente M, Cassidy LM, Gamba C, et al. 2015. Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nat Commun*. 6:8912.
Jones ER, Zarina G, Moiseyev V, Lightfoot E, Nigst PR, Manica A, Pinhasi R, Bradley DG. 2017. The Neolithic transition in the Baltic was not driven by admixture with early European farmers. *Curr Biol*. 27(4):576–582.
Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L. 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* 29(13):1682–1684.
Karmin M, Saag L, Vicente M, Wilson Sayres MA, Järve M, Talas UG, Rootsi S, Ilumäe A-M, Mägi R, Mitt M, et al. 2015. A recent bottleneck of Y chromosome diversity coincides with a global change in culture. *Genome Res*. 25(4):459–466.
Keller A, Graefen A, Ball M, Matzas M, Boisguerin V, Maixner F, Leidinger P, Backes C, Khairat R, Forster M, et al. 2012. New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. *Nat Commun*. 3:698.
Kılınç GM, Omrak A, Özer F, Günther T, Büyükkarakaya AM, Bıçakçı E, Baird D, Dönertaş HM, Ghalichi A, Yaka R, et al. 2016. The demographic development of the first farmers in Anatolia. *Curr Biol*. 26(19):2659–2666.
Kirigin B, Miše M, Barbarić V. 2010. Palagruža – the Island of Diomedes: summary excavation report 2002–2008. *Hesperia*. 25:65–91.
Korneliussen TS, Albrechtsen A, Nielsen R. 2014. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* 15:356.
Krzewińska M, Kılınç GM, Juras A, Koptekin D, Chyleński M, Nikitin AG, Shcherbakov N, Shuteleva I, Leonova T, Kraeva L, et al. 2018. Ancient genomes suggest the eastern Pontic-Caspian steppe as the source of western Iron Age nomads. *Sci Adv*. 4(10):eaat4457.
Lazaridis I, Mittnik A, Patterson N, Mallick S, Rohland N, Pfrengle S, Furtwängler A, Peltzer A, Posth C, Vasilakis A, et al. 2017. Genetic origins of the Minoans and Mycenaeans. *Nature* 548(7666):214–218.
Lazaridis I, Nadel D, Rollefson G, Merrett DC, Rohland N, Mallick S, Fernandes D, Novak M, Gamarra B, Sirak K, et al. 2016. Genomic insights into the origin of farming in the ancient Near East. *Nature* 536(7617):419–424.
Lazaridis I, Patterson N, Mittnik A, Renaud G, Mallick S, Kirsanow K, Sudmant PH, Schraiber JG, Castellano S, Lipson M, et al. 2014. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* 513(7518):409–413.
Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14):1754–1760.
Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and samtools. *Bioinformatics* 25(16):2078–2079.
Lipson M, Szécsényi-Nagy A, Mallick S, Pósa A, Stégmár B, Keerl V, Rohland N, Stewardson K, Ferry M, Michel M, et al. 2017. Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature* 551(7680):368–372.
Lombardo M. 2014. Iapygians: the indigenous populations of Ancient Apulia in the fifth and fourth centuries B.C.E. In: Carpenter T, Lynch K, Robinson E, editors. The Italic people of ancient Apulia: new evidence from pottery for workshops, markets, and customs. Cambridge: Cambridge University Press. p. 36–68.
Malaspinas A-S, Lao O, Schroeder H, Rasmussen M, Raghavan M, Moltke I, Campos PF, Sagredo FS, Rasmussen S, Gonçalves VF, et al. 2014. Two ancient human genomes reveal Polynesian ancestry among the indigenous Botocudos of Brazil. *Curr Biol*. 24(21):R1035–R1037.
Marcus JH, Posth C, Ringbauer H, Lai L, Skeates R, Sidore C, Beckett J, Furtwängler A, Olivieri A, Chiang CWK, et al. 2020. Genetic history from the Middle Neolithic to present on the Mediterranean island of Sardinia. *Nat Commun*. 11(1):939.
Margaryan A, Lawson DJ, Sikora M, Racimo F, Rasmussen S, Moltke I, Cassidy LM, Jørsboe E, Ingason A, Pedersen MW, et al. 2020. Population genomics of the Viking world. *Nature* 585(7825):390–396.
Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *Embnet J*. 17(1):10–12.
Martiniano R, Caffell A, Holst M, Hunter-Mann K, Montgomery J, Müldner G, McLaughlin RL, Teasdale MD, van Rheenen W, Veldink JH, et al. 2016. Genomic signals of migration and continuity in Britain before the Anglo-Saxons. *Nat Commun*. 7:10326.
Martiniano R, Cassidy LM, Ó'Maoldúin R, McLaughlin R, Silva NM, Manco L, Fidalgo D, Pereira T, Coelho MJ, Serra M, et al. 2017. The population genomics of archaeological transition in west Iberia: investigation of ancient substructure using imputation and haplotype-based methods. *PLoS Genet*. 13(7):e1006852.
Mathieson I, Alpaslan-Roodenberg S, Posth C, Szécsényi-Nagy A, Rohland N, Mallick S, Olalde I, Broomandkhoshbacht N, Candilio F, Cheronet O, et al. 2018. The genomic history of southeastern Europe. *Nature* 555(7695):197–203.
Mathieson I, Lazaridis I, Rohland N, Mallick S, Patterson N, Roodenberg SA, Harney E, Stewardson K, Fernandes D, Novak M, et al. 2015. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* 528(7583):499–503.
McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 20(9):1297–1303.

Meyer M, Kircher M. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc*. 2010(6):db.prot5448.
Mittnik A, Wang C-C, Pfrengle S, Daubaras M, Zariņa G, Hallgren F, Allmäe R, Khartanovich V, Moiseyev V, Tõrv M, et al. 2018. The genetic prehistory of the Baltic Sea region. *Nat Commun*. 9(1):442.
Molinaro L, Montinaro F, Yelmen B, Marnetto D, Behar DM, Kivisild T, Pagani L. 2019. West Asian sources of the Eurasian component in Ethiopians: a reassessment. *Sci Rep*. 9(1):18811.
Monroy Kuhn JM, Jakobsson M, Günther T. 2018. Estimating genetic kin relationships in prehistoric populations. *PLoS One* 13(4):e0195491.
Narasimhan VM, Patterson N, Moorjani P, Rohland N, Bernardos R, Mallick S, Lazaridis I, Nakatsuka N, Olalde I, Lipson M, et al. 2019. The formation of human populations in South and Central Asia. *Science* 365(6457):eaat7487.
Nava ML, Descoeudres JP. 1990. Greek and Adriatic influences in Daunia in the Early Iron Age. In: Descoeudres JP, editor. Greek colonists and native populations. Canberra: Humanities Research Centre; Oxford: Clarendon Press. p. 560–578.
Nijboer AJ. 2006. La cronologia assoluta dell'età del Ferro nel Mediterraneo, dibattito sui metodi e sui risultati. In: Bartoloni G, Delpino F, editors. Oriente e Occidente: Metodi e Discipline a Confronto. Pisa: Istituti Editoriali e Poligrafici Internazionali. p. 527–556.
Norman C. 2011. The tribal tattooing of Daunian women. *Eur J Archaeol*. 14(1–2):133–157.
Norman C. 2016. Daunian women: costume and actions commemorated in stone. In: Budin SL, Turfa JM, editors. Women in antiquity. Real women across the Ancient World. London; New York: Routledge, Taylor & Francis Group. p. 901–912.
Olalde I, Brace S, Allentoft ME, Armit I, Kristiansen K, Booth T, Rohland N, Mallick S, Szécsényi-Nagy A, Mittnik A, et al. 2018. The Beaker phenomenon and the genomic transformation of northwest Europe. *Nature* 555(7695):190–196.
Olalde I, Mallick S, Patterson N, Rohland N, Villalba-Mouco V, Silva M, Dulias K, Edwards CJ, Gandini F, Pala M, et al. 2019. The genomic history of the Iberian Peninsula over the past 8000 years. *Science* 363(6432):1230–1234.
Olalde I, Schroeder H, Sandoval-Velasco M, Vinner L, Lobón I, Ramirez O, Civit S, García Borja P, Salazar-García DC, Talamo S, et al. 2015. A common genetic origin for early farmers from Mediterranean Cardial and Central European LBK cultures. *Mol Biol Evol*. 32(12):3132–3142.
Omrak A, Günther T, Valdiosera C, Svensson EM, Malmström H, Kiesewetter H, Aylward W, Storå J, Jakobsson M, Götherström A. 2016. Genomic evidence establishes Anatolia as the source of the European Neolithic gene pool. *Curr Biol*. 26(2):270–275.
Orlando L, Ginolhac A, Zhang G, Froese D, Albrechtsen A, Stiller M, Schubert M, Cappellini E, Petersen B, Moltke I, et al. 2013. Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* 499(7456):74–78.
Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. 2012. Ancient admixture in human history. *Genetics* 192(3):1065–1093.
Patterson N, Price AL, Reich D. 2006. Population structure and eigenanalysis. *PLoS Genet*. 2(12):e190.
Poznik GD, Xue Y, Mendez FL, Willems TF, Massaia A, Wilson Sayres MA, Ayub Q, McCarthy SA, Narechania A, Kashin S, et al. 2016. Punctuated bursts in human male demography inferred from 1,244 worldwide Y-chromosome sequences. *Nat Genet*. 48(6):593–599.
R Core Team. 2021. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing. Available from: https://www.R-project.org/.
Raveane A, Aneli S, Montinaro F, Athanasiadis G, Barlera S, Birolo G, Boncoraglio G, Di Blasio AM, Di Gaetano C, Pagani L, et al. 2019. Population structure of modern-day Italians reveals patterns of ancient and archaic ancestries in Southern Europe. *Sci Adv*. 5(9):eaaw3492.
Reimer PJ, Austin WEN, Bard E, Bayliss A, Blackwell PG, Ramsey CB, Butzin M, Cheng H, Lawrence Edwards R, Friedrich M, et al. 2020. The IntCal20 Northern Hemisphere radiocarbon age calibration curve (0–55 cal kBP). *Radiocarbon* 62(4):725–757.
Rivollat M, Jeong C, Schiffels S, Küçükkalıpçı İ, Pemonge M-H, Rohrlach AB, Alt KW, Binder D, Friederich S, Ghesquière E, et al. 2020. Ancient genome-wide DNA from France highlights the complexity of interactions between Mesolithic hunter-gatherers and Neolithic farmers. *Sci Adv*. 6(22):eaaz5344.
Saag L, Laneman M, Varul L, Malve M, Valk H, Razzak MA, Shirobokov IG, Khartanovich VI, Mikhaylova ER, Kushniarevich A, et al. 2019. The arrival of Siberian ancestry connecting the Eastern Baltic to Uralic speakers further East. *Curr Biol*. 29(10):1701–1711.e16.
Saag L, Varul L, Scheib CL, Stenderup J, Allentoft ME, Saag L, Pagani L, Reidla M, Tambets K, Metspalu E, et al. 2017. Extensive farming in Estonia started through a sex-biased migration from the Steppe. *Curr Biol*. 27(14):2185–2193.e6.
Sallustio F, Cox SN, Serino G, Curci C, Pesce F, De Palma G, Papagianni A, Kirmizis D, Falchi M, Schena FP, et al. 2015. Genome-wide scan identifies a copy number variable region at 3p21.1 that influences the TLR9 expression levels in IgA nephropathy patients. *Eur J Hum Genet*. 23(7):940–948.
Saupe T, Montinaro F, Scaggion C, Carrara N, Kivisild T, D'Atanasio E, Hui R, Solnik A, Lebrasseur O, Larson G, et al. 2021. Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula. *Curr Biol*. 31(12):2576–2591.e12.
Scaggion C, Carrara N. 2016. New studies on human skeletal remains from the ancient Herdonia (southeast Italy). Evidences of tuberculosis and brucellosis: two diseases connected with farm animals. *Antrocom* 12:85–96.
Schiffels S, Haak W, Paajanen P, Llamas B, Popescu E, Loe L, Clarke R, Lyons A, Mortimer R, Sayer D, et al. 2016. Iron Age and Anglo-Saxon genomes from East England reveal British migration history. *Nat Commun*. 7:10408.
Schuenemann VJ, Peltzer A, Welte B, van Pelt WP, Molak M, Wang C-C, Furtwängler A, Urban C, Reiter E, Nieselt K, et al. 2017. Ancient Egyptian mummy genomes suggest an increase of Sub-Saharan African ancestry in post-Roman periods. *Nat Commun*. 8:15694.
Sikora M, Pitulko VV, Sousa VC, Allentoft ME, Vinner L, Rasmussen S, Margaryan A, de Barros Damgaard P, de la Fuente C, Renaud G, et al. 2019. The population history of northeastern Siberia since the Pleistocene. *Nature* 570(7760):182–188.
Skoglund P, Storå J, Götherström A, Jakobsson M. 2013. Accurate sex identification of ancient human remains using DNA shotgun sequencing. *J Archaeol Sci*. 40(12):4477–4482.
Skourtanioti E, Erdal YS, Frangipane M, Balossi Restelli F, Yener KA, Pinnock F, Matthiae P, Özbal R, Schoop U-D, Guliyev F, et al. 2020. Genomic history of Neolithic to Bronze Age Anatolia, Northern Levant, and Southern Caucasus. *Cell* 181(5):1158–1175.e28.
Tagliente M. 1986. I signori dei cavalli nella Daunia di età arcaica. *Annali Della Facoltà di Lettere e Filosofia Dell'Università Degli Studi di Perugia*. 23:1985–1986.
Unterländer M, Palstra F, Lazaridis I, Pilipenko A, Hofmanová Z, Groß M, Sell C, Blöcher J, Kirsanow K, Rohland N, et al. 2017. Ancestry and demography and descendants of Iron Age nomads of the Eurasian Steppe. *Nat Commun*. 8:14615.
Valdiosera C, Günther T, Vera-Rodríguez JC, Ureña I, Iriarte E, Rodríguez-Varela R, Simões LG, Martínez-Sánchez RM, Svensson EM, Malmström H, et al. 2018. Four millennia of Iberian biomolecular prehistory illustrate the impact of prehistoric migrations at the far end of Eurasia. *Proc Natl Acad Sci U S A*. 115(13):3428–3433.
van den Brink ECM, Beeri R, Kirzner D, Bron E, Cohen-Weinberger A, Kamaisky E, Gonen T, Gershuny L, Nagar Y, Ben-Tor D, et al. 2017. A Late Bronze Age II clay coffin from Tel Shaddud in the Central Jezreel Valley, Israel: context and historical implications. *Levant* 49(2):105–135.
van Oven M, Kayser M. 2009. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat*. 30(2):E386–E394.
Veeramah KR, Rott A, Groß M, van Dorp L, López S, Kirsanow K, Sell C, Blöcher J, Wegmann D, Link V, et al. 2018. Population genomic analysis of elongated skulls reveals extensive female-biased

immigration in Early Medieval Bavaria. *Proc Natl Acad Sci U S A*. 115(13):3494–3499.
Villalba-Mouco V, van de Loosdrecht MS, Posth C, Mora R, Martínez-Moreno J, Rojo-Guerra M, Salazar-García DC, Royo-Guillén JI, Kunst M, Rougier H, et al. 2019. Survival of Late Pleistocene hunter-gatherer ancestry in the Iberian Peninsula. *Curr Biol*. 29(7):1169–1177.e7.
Wang C-C, Reinhold S, Kalmykov A, Wissgott A, Brandt G, Jeong C, Cheronet O, Ferry M, Harney E, Keating D, et al. 2019. Ancient human genome-wide data from a 3000-year interval in the Caucasus corresponds with eco-geographic regions. *Nat Commun*. 10(1):590.
Wang RJ, Al-Saffar SI, Rogers J, Hahn MW. 2021. Human generation times across the past 250,000 years. bioRxiv 2021.09.07.459333. doi: https://doi.org/10.1101/2021.09.07.459333.
Weissensteiner H, Pacher D, Kloss-Brandstätter A, Forer L, Specht G, Bandelt H-J, Kronenberg F, Salas A, Schönherr S. 2016. HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res*. 44(W1):W58–W63.
Zalloua P, Collins CJ, Gosling A, Biagini SA, Costa B, Kardailsky O, Nigro L, Khalil W, Calafell F, Matisoo-Smith E. 2018. Ancient DNA of Phoenician remains indicates discontinuity in the settlement history of Ibiza. *Sci Rep*. 8(1):17567.
Žegarac A, Winkelbach L, Blöcher J, Diekmann Y, Krečković Gavrilović M, Porčić M, Stojković B, Milašinović L, Schreiber M, Wegmann D, et al. 2021. Ancient genomes provide insights into family structure and the heredity of social status in the early Bronze Age of southeastern Europe. *Sci Rep*. 11(1):10072.

Article

Current Biology

Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula

Highlights

- 22 genomes from Northeastern and Central Italy dated between 3200 and 1500 BCE
- Arrival of Steppe-related ancestry in the central Italian Peninsula by 1600 BCE
- Close patrilineal kinship patterns within commingled Chalcolithic cave burials
- Roman Imperial period had a stronger effect on phenotype shifts than the Bronze Age

Authors

Tina Saupe, Francesco Montinaro, Cinzia Scaggion, ..., Cristian Capelli, Luca Pagani, Christiana L. Scheib

Correspondence

tina.saupe@ut.ee (T.S.), cls83@ut.ee (C.L.S.)

In brief

Using ancient DNA sequences from the Italian Peninsula, Saupe et al. explore the genetic history of the Italian Peninsula through the Chalcolithic/Bronze Age transition. The time transect shows Steppe-related ancestry arrived in Central Italy by 1600 BCE and possibly precipitated shifts in kinship patterns, but not phenotypes.

Saupe et al., 2021, Current Biology *31*, 1–16
June 21, 2021
https://doi.org/10.1016/j.cub.2021.04.022

CellPress

Current Biology

CellPress
OPEN ACCESS

Article

Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula

Tina Saupe,[1,26,27,*] Francesco Montinaro,[1,2,26] Cinzia Scaggion,[3] Nicola Carrara,[4] Toomas Kivisild,[1,5] Eugenia D'Atanasio,[6] Ruoyun Hui,[7] Anu Solnik,[8] Ophélie Lebrasseur,[9,10] Greger Larson,[12] Luca Alessandri,[11] Ilenia Arienzo,[12] Flavio De Angelis,[13] Mario Federico Rolfo,[14] Robin Skeates,[15] Letizia Silvestri,[14] Jessica Beckett,[16] Sahra Talamo,[17,18] Andrea Dolfini,[19] Monica Miari,[20] Mait Metspalu,[1] Stefano Benazzi,[21] Cristian Capelli,[22,23,26] Luca Pagani,[1,24,26] and Christiana L. Scheib[1,25,26,*]

[1]Estonian Biocentre, Institute of Genomics, University of Tartu, Riia 23B, Tartu 51010, Estonia
[2]Department of Biology-Genetics, University of Bari, Via E. Orabona, 4, Bari 70124, Italy
[3]Department of Geosciences, University of Padova, Via Gradenigo 6, Padova 35131, Italy
[4]Museum of Anthropology, University of Padova, Palazzo Cavalli, via Giotto 1, Padova 35121, Italy
[5]Department of Human Genetics, KU Leuven, Leuven, Herestraat 49 3000, Belgium
[6]Institute of Molecular Biology and Pathology, CNR, Piazzale Aldo Moro 5, Rome 00185, Italy
[7]McDonald Institute for Archaeological Research, University of Cambridge, Downing Street, Cambridge CB2 3ER, UK
[8]Core Facility, Institute of Genomics, University of Tartu, Riia 23B, Tartu 51010, Estonia
[9]Department of Archaeology, Classics and Egyptology, University of Liverpool, 12-14 Abercromby Square, Liverpool L69 7WZ, UK
[10]Palaeogenomics & Bio-Archaeology Research Network, School of Archaeology, University of Oxford, 1 South Parks Road, Oxford OX1 3TG, UK
[11]Groningen Institute of Archaeology, University of Groningen, Poststraat 6, Groningen 9712, the Netherlands
[12]Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Vesuviano, Via Diocleziano 328, Naples 80125, Italy
[13]Centre of Molecular Anthropology for Ancient DNA Studies, Department of Biology, University of Rome "Tor Vergata," Via della Ricerca Scientifica 1, Rome 00133, Italy
[14]Department of History, Culture and Society, University of Rome "Tor Vergata," Via Columbia 1, Rome 00133, Italy
[15]Department of Archaeology, Durham University, Lower Mountjoy, South Road, Durham DH1 3LE, UK
[16]Independent scholar, Cagliari, Italy
[17]Department of Chemistry "Giacomo Ciamician," University of Bologna, Via Selmi 2, Bologna 40126, Italy
[18]Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, Leipzig 04103, Germany
[19]School of History, Classics and Archaeology, Newcastle University, Newcastle upon Tyne NE1 7RU, UK
[20]Superintendency of Archeology, Fine Arts and Landscape for the metropolitan city of Bologna and the provinces of Modena, Reggio Emilia and Ferrara, Comune di Bologna, Sede Via Belle Arti n. 52, Bologna 40126, Italy
[21]Department of Cultural Heritage, University of Bologna, Via degli Ariani, 1, Ravenna 40126, Italy
[22]Department of Zoology, University of Oxford, 11a Mansfield Road, Oxford OX1 3SZ, UK
[23]Dipartimento di Scienze Chimiche, della Vita e della Sostenibilità Ambientale, University of Parma, Parco Area delle Scienze 17/A, Parma 43124, Italy
[24]Department of Biology, University of Padova, Via U. Bassi, 58/B, Padova 35122, Italy
[25]St. John's College, University of Cambridge, St. John's Street, Cambridge CB2 1TP, UK
[26]These authors contributed equally
[27]Lead contact
*Correspondence: tina.saupe@ut.ee (T.S.), cls83@ut.ee (C.L.S.)
https://doi.org/10.1016/j.cub.2021.04.022

SUMMARY

Across Europe, the genetics of the Chalcolithic/Bronze Age transition is increasingly characterized in terms of an influx of Steppe-related ancestry. The effect of this major shift on the genetic structure of populations in the Italian Peninsula remains underexplored. Here, genome-wide shotgun data for 22 individuals from commingled cave and single burials in Northeastern and Central Italy dated between 3200 and 1500 BCE provide the first genomic characterization of Bronze Age individuals (n = 8; 0.001–1.2× coverage) from the central Italian Peninsula, filling a gap in the literature between 1950 and 1500 BCE. Our study confirms a diversity of ancestry components during the Chalcolithic and the arrival of Steppe-related ancestry in the central Italian Peninsula as early as 1600 BCE, with this ancestry component increasing through time. We detect close patrilineal kinship in the burial patterns of Chalcolithic commingled cave burials and a shift away from this in the Bronze Age (2200–900 BCE) along with lowered runs of homozygosity, which may reflect larger changes in population structure. Finally, we find no evidence that the arrival of Steppe-related ancestry in Central Italy directly led to changes in frequency of 115 phenotypes present in the dataset, rather that the post-Roman

Please cite this article in press as: Saupe et al., Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.04.022

Imperial period had a stronger influence, particularly on the frequency of variants associated with protection against Hansen's disease (leprosy). Our study provides a closer look at local dynamics of demography and phenotypic shifts as they occurred as part of a broader phenomenon of widespread admixture during the Chalcolithic/Bronze Age transition.

INTRODUCTION

The Italian Chalcolithic (or Copper Age [CA]; 3600–2200 BCE), the period between the Late Neolithic (N) (7000–3600 BCE) and the Bronze Age (BA) (2200–900 BCE), is characterized by the development of new tools from different metallic sources and was followed by major cultural transformations, including that of burial practice—from an emphasis on the collective to the individual and of personal, prestige grave goods.[1,2] Ancient DNA (aDNA) studies have highlighted the occurrence of major shifts in the genetic profiles of populations coinciding with material culture changes, such as from hunting-gathering to farming.[3–6] At the beginning of the transition from the Chalcolithic to the BA ~5,000 years ago, people from the Eurasian Steppe arrived in Europe, resulting in further admixing with local populations.[7–10] Although these events have been extensively studied in most of Europe[4,11,12] and a number of studies on ancient genomes from the Italian Peninsula, Sardinia, and Sicily have been recently published,[4,7,9,13–17] the demographic dynamics of the Chalcolithic/Early BA in the Italian Peninsula are still not well characterized. Though previous studies place the arrival of a Steppe-related ancestry component in Northern Italy[9] and in Sicily[16] after 2300 BCE, a chronological gap from 1900 to 900 BCE is present, and little is known about the spread of Steppe-related ancestry in Central Italy. In addition, the available data show an Iranian N-related component detected in Sardinia after 900 BCE,[16,17] although affinities to Caucasus hunter-gatherers (CHG) and Iran N farmers are present in Central Italian N individuals[13] and in Middle BA Sicily,[16] at a lower proportion than modern Italians.[11] However, although the BA CHG affinity in Sicily is supported by $f4$ statistics, the evidence for the N CHG influx is less robust. Furthermore, with a few exceptions,[18,19] previous surveys have focused primarily on describing ancestral relationships or inferring movement and mixtures of populations at the expense of questions focusing on the social dynamics associated with these events, e.g., evaluating the kinship structure in prehistoric society.

aDNA is proving a useful tool for helping to infer past social structures and reproductive behavior (reviewed in Racimo et al.[20]). In N Europe, several studies have detected a widespread cultural connection of patrilineal social organization[21,22] as well as for the BA transition with large-scale, sex-biased migrations,[8,23,24] local patterns of patrilocality and female exogamy,[18,25] and the influence of cultural diffusion versus migration.[9] Although the social implications of these changes are still debated,[26] cultural shifts can have an effect on adaptation (e.g., a change in technology leads to a change in diet, leading to selective pressure on metabolism genes). So far, the social structure(s) in Central Italy during the Chalcolithic/BA transition and whether shifts in cultural practices (kinship, patrilocality, and exogamy) correlate with the introduction of Steppe-related ancestry remains unexplored. This may be partially due to the fact that, although there exists a wide variety of burial practices in the Chalcolithic period in Italy, they are often characterized by collective depositions of commingled remains.[2] This has made the anthropological analysis of the burial populations and interpretations regarding kinship and social structure difficult; however, high-throughput aDNA sequencing allows for the genetic screening of large numbers of skeletal samples and reconstruction of individuals from disarticulated remains.

In addition to reshaping our understanding of the demographic history of the European continent, analyses of ancient genomes from Europe have recently called into question hypotheses regarding the time depth of selection on phenotypic traits in Europe. For example, aDNA has revealed that Mesolithic hunter-gatherers in Europe could have dark skin and blue eyes (a combination rarely seen today)[27–29] and that selection on skin pigmentation occurred in the last 5,000 years.[30,31] Other recent work has suggested that selection within genes related to fatty acid metabolism and starch digestion did not take place during the transition to agriculture but instead initiated closer to 2000 BCE following the introduction of Steppe-related ancestry into Europe,[31] with the ancestry component itself a possible driver. Another open question, that of the role of pathogens in shaping human genomes, is now starting to be explored using aDNA. One particular pathogen, Hansen's disease (leprosy), is first seen in paleopathological evidence in the Mediterranean dating to the BA[32] and is noted in Central and Northeast Italy by 300 BCE.[33] The disease may have been later spread by Roman military movements[34] and increased to high numbers in Europe in the Early Medieval Period, but declined by the 15th century CE, and the role of human genetic adaptation in this decline is unknown. There are a number of genetic loci that have been implicated in the manifestation and progression of the infection,[35–38] including one recently discovered using aDNA.[39]

Here, using an aDNA approach, we investigated the diversity of ancestry components prior to and through the Chalcolithic/BA transition in Northeastern and Central Italy and whether shifts in Steppe-related ancestry correlate with changes in inferred social structure and/or phenotypic traits.

RESULTS

We extracted DNA from 51 skeletal elements (teeth = 37, petrous bones = 10, and additional powder from petrous bones = 4) at the Ancient DNA Laboratory of the Institute of Genomics, University of Tartu in Estonia. The human remains are from one necropolis (Necropoli di Gattolino; hereafter, "Gattolino") and three cave sites located in Northeastern (Grottina dei Covoloni del Broion: "Broion") and Central Italy (Grotta La Sassa: "La Sassa" and Grotta Regina Margherita: "Regina Margherita"; Figures 1A and S1). After screening 47 libraries at a low depth (±20M reads per library), we identified and sequenced 20 libraries with more than 4% endogenous human DNA and mtDNA-based contamination estimates less than 1.44% (Data S1A and S1B). For the disarticulated

Please cite this article in press as: Saupe et al., Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.04.022

CellPress
OPEN ACCESS

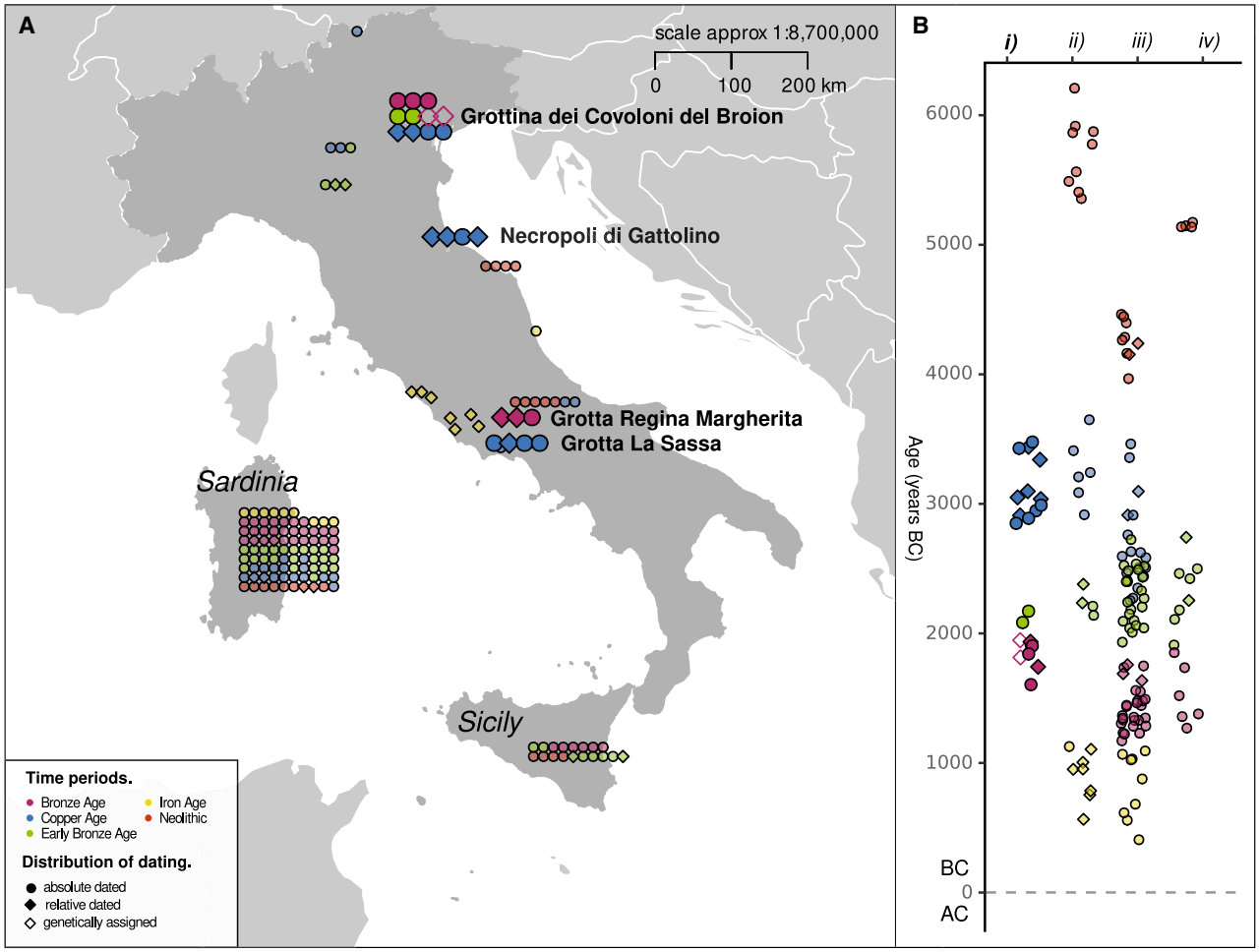

Figure 1. Geographical location of samples and relative or absolute dating

(A) Map of the geographical location of selected published (smaller, transparent) and newly generated samples from the Italian Peninsula, Sardinia, and Sicily included in this study. The titled locations are here newly reported. See also Table 1 and Data S1.

(B) Distribution of relative and absolute dating and genetic assignment from newly generated samples (i) and published (transparent; order: [ii] Italian Peninsula; [iii] Sardinia; and [iv] Sicily).

See also Data S1A–S1D.

remains in the cave sites, we calculated the pairwise mismatch rate of SNPs (P0) on pseudo-haploid data as implemented in READ[40] to identify genetically identical samples and attempt to calculate a minimum number of individuals (see STAR Methods and Data S1B for details).

The sequences of identical samples were merged together, leaving 22 unique individuals: eleven from Broion (Italy_Broion_CA = 4, Italy_Broion_EBA = 2, and Italy_Broion_BA = 5), four from Gattolino (Italy_Gattolino_CA), three from Regina Margherita (Italy_ReginaMargherita_BA), and four from La Sassa (Italy_LaSassa_CA). The final data are composed of individuals with endogenous DNA between 0.48% and 48.87%, average genomic coverage between 0.0016× and 1.24×, and estimated contamination rates of 0.00%–1.44% (mtDNA-based) and 0.45%–1.98% (X-chromosome-based in males only; Table 1; Data S1A and S1B).

In order to characterize the timing of genetic shifts related to the presence of ancestry from the Steppe and to assign individuals in commingled contexts within chronological space, we relied on two forms of temporal assignment: archaeological evidence (e.g., ceramic fragments, a significant amount of metallic tools found with the human remains, and geological layers; STAR Methods) and direct radiocarbon dating of 12 of the 22 individuals here sequenced (Figure 1B; Data S1C; STAR Methods). Of the ten undated samples, the age of BRC013 can be inferred as ± the average reproductive span of a woman (30 years) to a directly dated sample (BRC022) from the first-degree relatedness (likely brothers), although for the other nine (BRC007, BRC011, BRC024, LSC011A1, GCP002A1, GCP004A1, GLR001A1, GLR002A1, and GLR004), we assigned them to the most parsimonious group, considering both archaeological and genetic information.

Genetic structure of Italy from the N to the Iron Age

To compare Chalcolithic and BA individuals from Italy to other ancient and contemporary European populations, we performed

Please cite this article in press as: Saupe et al., Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.04.022

Table 1. Archaeological information, genome coverage, genetic sex, mtDNA, and Y chromosome haplogroups of the individuals of this study

Individual	Site	Date	Genome coverage	Gen. sex	mtDNA HG	Y chromosome HG	No. of SNPs
BRC001/023	Broion	4430 ± 40 BP; 3313–2934 cal BCE	0.172	XY	J2a1a1	G2a3-F1193-F2291	231,702
BRC002	Broion	failed C14 dating	0.143	XY	N1a1a1a1	R1b1-DF90	200,133
BRC003	Broion	3239 ± 31 BP[a]; 1532–1452 cal BCE	1.235	XY	U4a2f	R1b1'5-P312	920,226
BRC007/019	Broion	3272 ± 29 BP; 1608–1502 cal BCE	0.189	XX	K1a1b1	–	259,050
BRC010/018	Broion	3532 ± 35 BP; 1926–1775 cal BCE	0.360	XX	H+16291T	–	437,478
BRC011	Broion	not dated	0.0025	XY	T2c1+146T!	–	3,606
BRC013	Broion	failed C14 dating	0.052	XY	H5a1[b]	G2a3-F705	70,071
BRC015	Broion	not dated	0.0016	XY	T2c1e	–	2,145
BRC022	Broion	4489 ± 41 BP; 3334–3100 cal BCE	0.145	XY	H5a1[b]	G2a3-F1193	188,166
BRC024	Broion	not dated	0.0144	XY	HV0a	R1	19,590
BRC030	Broion	3502 ± 41 BP; 1886–1751 cal BCE	0.110	XX	K1a4a1e	–	145,106
GCP002	Regina Margherita	failed C14 dating	0.138	XY	U5b2b5	G2	181,152
GCP003	Regina Margherita	3277 ± 29 BP; 1608–1504 cal BCE	0.138	XX	H3am	–	180,991
GCP004	Regina Margherita	not dated	0.0052	XX	U5b2b3	–	7,278
GLR001	Gattolino	not dated	0.214	XX	K2b1b	–	262,814
GLR002	Gattolino	not dated	0.129	XY	J1c3e1	I2d-L623/M223	163,883
GLR003	Gattolino	4829–4627 BP; 2874–2704 cal BCE	0.008	XX	H10d	–	10,863
GLR004	Gattolino	not dated	0.122	XY	H3am	I2b-M26	157,772
LSC002/004	La Sassa	4091 ± 29 BP: 2840–2575 cal BCE	1.014	XY	H1bv1[b]	J2a7-Z2397	903,958
LSC005/013	La Sassa	4097 ± 39 BP; 2847–2575 cal BCE	0.1323	XX	H1e5a	–	176,120
LSC007	La Sassa	not dated	0.0029	XY	H1bv1[b]	–	4,033
LSC011	La Sassa	4073 ± 37 BP; 2837–2498 cal BCE	0.063	XY	J1c1	J2a-M410 (J2a7-Z2397)	83,596

Dates in BP are raw radiocarbon dates; calibrated dates are the 68.3% probability and were calibrated using IntCal20[41] and either CALIBREV8.2[42] or OxCal 4.4[43] (GLR003). See also Data S1. Gen., genetic; HG, haplogroup.

[a]Sample failed C14 dating, but a genetically identical sample was able to be dated

[b]Indicates identical haplotype

Please cite this article in press as: Saupe et al., Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.04.022

CellPress
OPEN ACCESS

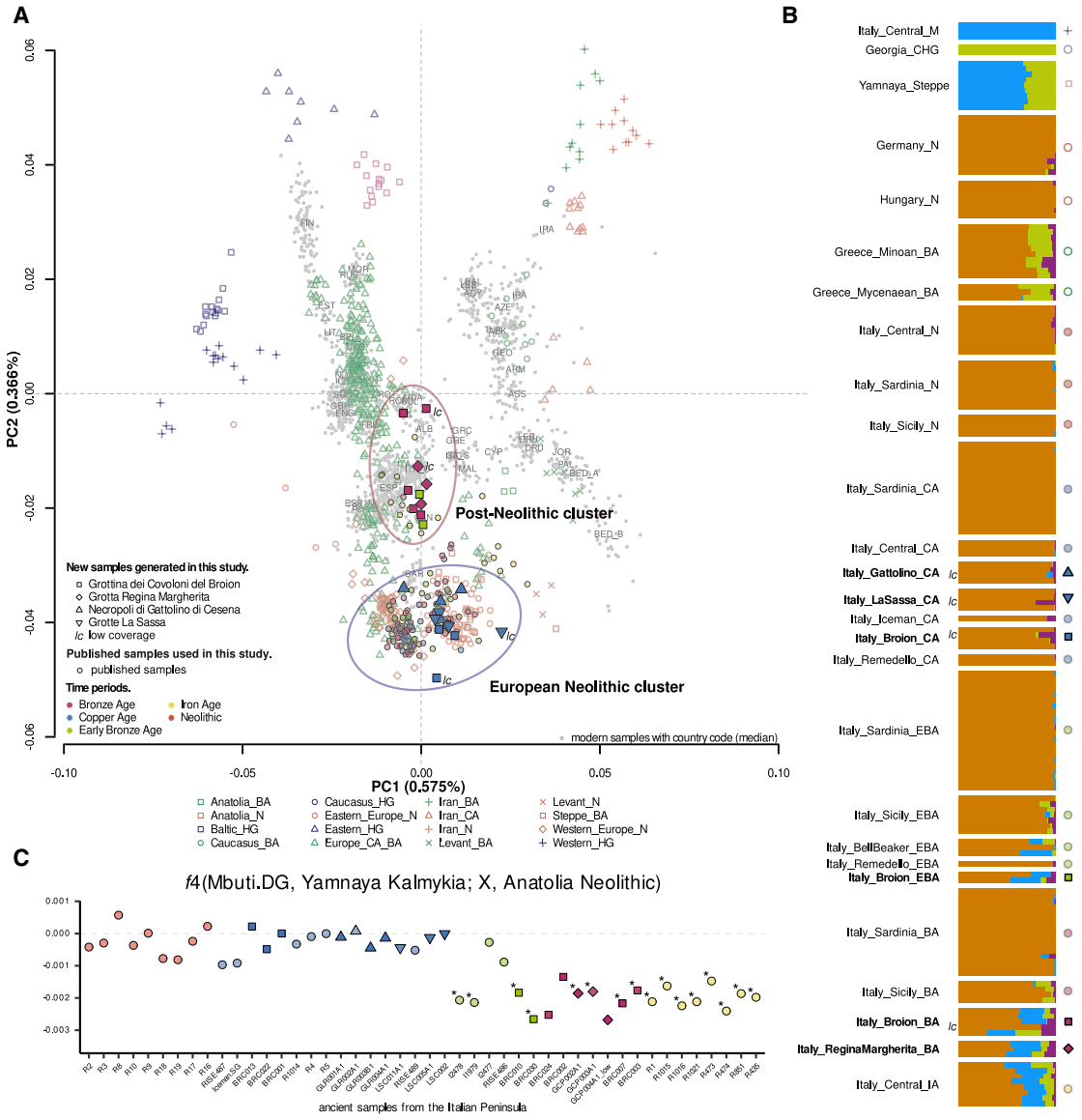

Figure 2. Overview of genetic structure

(A) Principal component analysis (PCA) of newly generated individuals with previously published data projected onto the variation from present-day populations (Data S1D).

(B) DyStruct analysis of newly reported ancient individuals (bold) numerically ordered at K = 4 together with a subset of key ancient Eurasian populations spanning from Neolithic to Iron Age. The icons on the right of the graph indicate the symbols of the population as shown on the PCA in (A). Samples with *lc* indicate low coverage samples generated in this study. See also Figures S2–S4A.

(C) Analysis of the Steppe-related ancestry components in selected published (transparent) and newly generated ancient samples (X) from the Italian Peninsula using *f*4 statistics in form *f*4(Mbuti.DG, Yamnaya Kalmykia; X, Anatolia Neolithic) (samples with * have a *Z* score less than or equal to −3). Tests with less than 5,000 SNPs were not included.

See also Data S2B and Figure S4B.

principal-component analysis (PCA), projecting ancient samples onto the genetic variation of 1,471 present-day individuals from Eurasia (Figures 2A and S4A).[44,45] Our newly generated samples scatter into two main clusters: European N (EN) (blue circle) encompassing all the Chalcolithic samples (Italy_LaSassa_CA, Italy_Gattolino_CA, and Italy_Broion_CA) and post-Neolithic (PN) (red circle), including Early Bronze and BA samples from Grotta Regina Margherita and Broion. Interestingly, most of the N, Chalcolithic, and BA Italian samples available from the literature fall within the EN cluster, although PN is mostly featuring Iron Age (IA) samples together with a few published BA individuals. This can be reconciled by the fact that most of the Italian BA samples available to date come from Sardinia and Sicily, two Mediterranean islands for which a reduced Steppe-related ancestry component has already been reported.[6,11,16,17,46] This is confirmed by DyStruct analysis (Figures 2B and S2), which

Please cite this article in press as: Saupe et al., Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.04.022

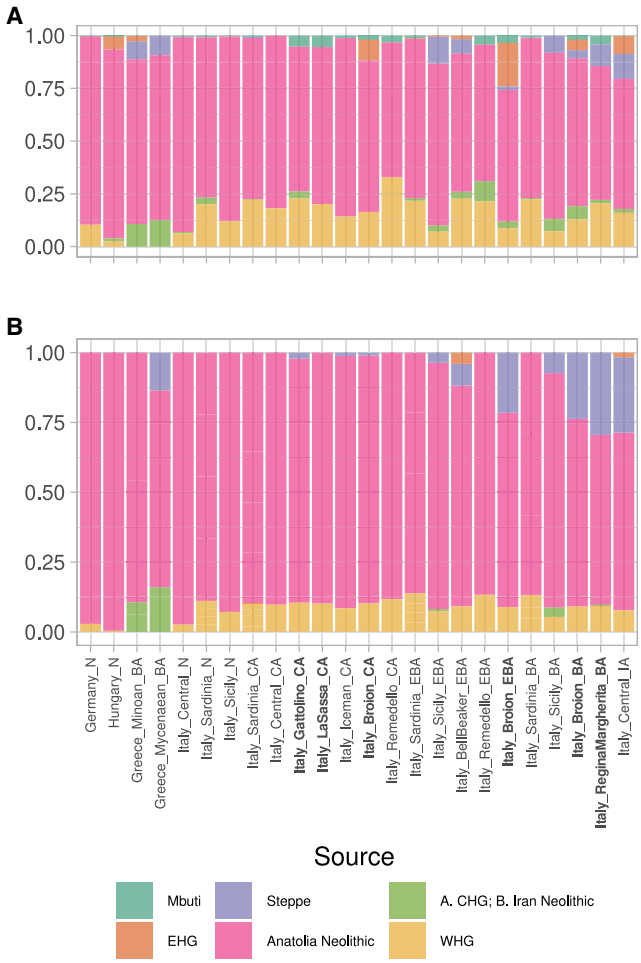

Figure 3. Ancestry composition of European and Italian ancient samples

(A) We used Chromopainter in the unlinked mode to reconstruct the genome of European (for which only a subset of relevant samples was analyzed) and Italian ancient individuals as a combination of six main putative sources. See also Figure S6.

(B) The same framework has been used on a vector composed of 4,950 different $f4$ in the form $f4$(X, Y; Target, Mbuti.DG). Mbuti.DG has not been used as a source.

See also Data S2 and S3.

also shows a high Anatolian N-like component in Sardinian and Sicilian individuals from the Chalcolithic and BA.

A separation within the EN cluster (Figure 2A) clearly differentiates Anatolian and Eastern Europe N (right) from Western Europe N, defined as samples west of Germany (left, toward western hunter-gatherer [WHG]; Data S1D). A similar separation has already been reported and interpreted as a difference in WHG proportion of these samples.[17,47–49] We note that most of our Chalcolithic individuals fall on the right side of the cluster (Anatolia and Eastern Europe).

We further explored the affinity between Italian Chalcolithic samples and WHG by computing $f4$(Mbuti.DG, Italy_Mesolithic.SG; Italy_Sardinia_N, X) (Data S2H) where X is either Italy_N, Italy_LaSassa_CA, Italy_Gattolino_CA, or Italy_Broion_CA and negative Z scores (min = −12.942; max = −3.026) indicate that Italy_Central_Mesolithic shares more with Italy_Sardinia_N than with peninsular Italian Chalcolithic. At the same time, we also tested $f4$(Mbuti.DG, Anatolia_N; Italy_Sardinia_N, X), for which only the comparison with Italy_Gattolino_CA yields significantly negative values (Z score = −3.753), suggesting that either an imbalance in WHG and Anatolia_N components between the two sides of the EN cluster or structure within the EN component[9,50] may explain our results. We note however that outgroup $f3$ statistics in the form $f3$(Mbuti.DG; Italy_CA, Y), where Italy_CA are all Chalcolithic samples from the Italian Peninsula and Y are members of the EN clusters may have no power to discriminate the observations from the PCA analysis (Data S1K). Some studies have detected an influence from groups related to hunter-gatherers from Georgia and Iran starting from the N;[13,51] however, our $f4$ of the form $f4$(Mbuti.DG, Georgia Kotias; Italy_Sardinia_N, X) is not significant for any of the pre-BA Italian groups (Data S2H).

We have evaluated the possibility of continuity at the interface between N/Chalcolithic and BA, using the qpWave/qpAdm framework (using the option "allsnps=YES"; Data S3G and S3H). When using two putative sources, all the target BA groups from North and Central Italy presented here support a scenario in which Chalcolithic-like individuals received a contribution of Steppe-related ancestry, possibly through Late N/Chalcolithic groups from the north, such as Germany Bell Beaker, France Middle N, and Italian Chalcolithic sources (Data S3G and S3H). Model-based clustering analysis (DyStruct [Figures 2B and S2] and ADMIXTURE [Figure S3])[52,53] of selected ancient and present-day individuals from Eurasia spanning from the Mesolithic to the IA points to the presence of a Steppe-related ancestry component in BA individuals from Italian Peninsula as the main difference from the Sardinian and Sicilian individuals, explaining the distribution of individuals between PN and EN in the PCA (Figure 2A).

We tested this through $f4$(Mbuti.DG, Yamnaya Kalmykia; X, Anatolia_N) (Figures 2C and S4B; Data S2B)[7,9,13,15–17,54] with X being Italian post-Mesolithic (only considering tests when at least 5,000 SNPs were available) and show that the only individuals with a significant enrichment for Steppe-related ancestry components are included within the newly generated Early BA and BA, and in published Bell Beaker (I2478 and I1979) or Italian IA individuals. Contrary to what has been previously reported for other Chalcolithic to BA transitions in Europe,[23] we also noted through outgroup $f3$ tests in form $f3$(Italian_CA/Italian_EBA_BA, ancient; Mbuti.DG) (Data S2K) that populations associated with Steppe-related ancestry did not leave a male-biased signature in Italy, which, if at all, can instead be seen through the contribution of pre-existing N groups (Figure S5).

We recapitulated the emerging picture using two orthogonal methods based on copying affinity, Chromopainter/NNLS[55] and SOURCEFIND,[56,57] and on the comparison of multiple $f4$ pairs (Figure 3). Concerning the novel samples, both approaches show an overall consistency with DyStruct (Figures 2B and S2), highlighting the post-N increase of European hunter-gatherer (HG)-related components[10,15] and the arrival of Steppe-related ancestry, with the only exception being that of Sardinia in the Early BA. Additionally, the Anatolian N and WHG proportions reported for the Italian N and Chalcolithic samples are similar

Table 2. Comparison of pairwise estimates of Chalcolithic genomes

ID1	ID2	Non-normalized P0	SE	READ relationship	*Z* upper	*Z* lower	IBD1	IBD2	Plink Fxy	Degree	Relationship
BRC013	BRC022	0.255	0.003	first degree	6.23	−13.73	0.44	0.22	0.22	first	sibling
LSC002/4	LSC011	0.266	0.002	first degree	3.37	−33.09	0.71	0.06	0.21	first	parent-offspring
LSC007	LSC002/4	0.244	0.005	first degree	4.46	−6.86					
BRC001	BRC013	0.298	0.003	second degree	2.94	−7.74	0.37	0.03	0.11	second	avuncular or grandparent-grandchild
BRC001	BRC022	0.323	0.002	unrelated	NA	−6.87	0.21	0.04	0.07	third	1st cousin

Kinship coefficient, Fxy, was estimated in PLINK 1.9.0[59] as 1/2 of the reported PIHAT value. PLINK-genome analyses used an input file of 16 imputed genomes and a total of 5,526,356 variants with MAF > 0.05 and GP > 0.99. Reported are pairs of individuals in which at least one was >0.1× coverage and who shared more than 12% of IBD. READ does not calculate third-degree relationships, rather categorizes them as "unrelated." LSC007 was too low coverage for imputation; thus, the pair with LSC002/4 was not estimated using PLINK. See also Data S4.

regardless of their position within the EN cluster identified in the PCA (Figure 2A).

We determined mitochondrial DNA (mtDNA) haplotypes for all newly generated samples (Table 1; Data S1B and S1E) and Y chromosome (Ychr) haplogroups for ten males with Ychr coverage >0.01× (Table 1; Data S1B and S1F). Consistent with the previously reported co-spread of Steppe-related ancestry and Ychr haplogroup R1,[7,10] we observed that three out of the four Italian BA males for which a Ychr haplogroup could be determined belong to haplogroup R1 and two of those were of the R1b lineage (Table 1; Data S1B and S1F). This haplogroup does not appear in the Chalcolithic samples. The two Italian R1b lineages belong to the L11 subset of R1b, which is common in modern Western Europe[58] and in ancient male Bell-Beaker burials,[9] rather than to the R1b-Z2105 varieties found in the ancient genomes from the Steppe-Belt of Russia.

Structure and mobility in the Chalcolithic and BA

The sites examined in this study include one necropolis with single burials (Gattolino, 2874–2704 BCE), and the rest are commingled cave burials (Broion, La Sassa, and Regina Margherita). Samples in this study from La Sassa are restricted to the Chalcolithic period (2850–2499 BCE) and from Regina Margherita to the BA (1609–1515 BCE). The Broion site is the only one that spans both phases: the Chalcolithic (3335–2936 BCE) and the BA (1923–1451 BCE). To better understand the site usage, whether differences are present between cave and cemetery burials and whether shifts in burial behavior occurred concurrently with shifts in genetic ancestry, we analyzed uniparental markers (informative of maternal versus paternal lineage diversity and mobility), genetic kinship (informative of mobility of family structure), and runs of homozygosity (ROHs) (which can indicate the size and homogeneity of a population).

In total, the Chalcolithic and BA sample sizes are approximately equal (n = 12 and 10, respectively), and the total number of males and females does not differ from expectations (binomial; p = 0.143); however, the ratio of males to females (7:1) within the Chalcolithic cave burials does differ from the expected value (binomial; p = 0.038), although it does not within the BA layer of Broion and Regina Margherita (5:5; binomial; p = 0.50) or in the Chalcolithic necropolis of Gattolino (2:2; binomial; p = 0.5). These results indicate a slight bias in the Chalcolithic cave burials toward males.

To identify whether first- and/or second-degree genetic kinship relationships existed among individuals within or between the four sites or between the new genomes and published ancient datasets (Data S4), we utilized two methods. First, for initial determination of kinship degrees, we utilized a pairwise mismatch estimation on the pseudo-haploid data (READ).[40] We ran the analysis on all of the newly reported ancient genomes together as one group as well as in groups by site and/or chronology (separated into N/Chalcolithic and BA) and found consistent results, regardless of grouping (Data S4A). To distinguish relationship type within degree (e.g., first degree, full siblings versus parent-offspring), we used IBD analysis as implemented in Plink-1.9[59] on imputed genotypes (STAR Methods). Genotypes were imputed using a pipeline detailed in Hui et al.[60] and STAR Methods, and in the case of close (1 to 2) degrees of relationship, both methods provided consistent results (Table 2).

We found no relationships among the seven tested BA individuals, between any of the newly presented sites, or with published data (Data S4); however, the small sample size suggests caution in inferring patterns of general validity from these individuals. For the twelve Chalcolithic individuals, close kinship relationships were detected at both cave burial sites: La Sassa and Broion and all relationships were detected between males (Tables 1 and 2). At La Sassa, the two males LSC002/004 and LSC011 have an identical Ychr haplotype (J2a-M410/J2a7-Z2397; Table 1; Data S1B and S1F), different mtDNA haplotypes (H1bv1 and J1c1; Table 1; Data S1B and S1E), a first-degree relationship, and a proportion of IBD (PI_HAT) values consistent with parent-offspring (Table 2), the summary of which indicates a father-son relationship. The nature of calibration curves for radiocarbon dating prevents the exact estimation of who is the father; however, the radiocarbon dates do not reject the relatedness inferred from aDNA (Data S1C; STAR Methods). Very low coverage sample LSC007 appears to have a first-degree relationship with LSC002/004 (Data S1B; Figure S4), and they share an identical mtDNA haplotype (H1bv1). LSC007 was too low coverage to assess the Ychr.

Please cite this article in press as: Saupe et al., Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.04.022

The chronologically contemporaneous female LSC005/013 (H1e5a) does not have any detectable close genetic kinship relationships, clusters separately from the males on the PCA, and has a strontium isotope signature that falls outside the range determined for the rest of the sampled teeth in La Sassa (n = 27; STAR Methods). To test whether LSC005/013 was genetically more similar to another population than the other La Sassa individuals, we tested *f*4(Mbuti.DG, LSC005A1_LSC013; Italy_LaSassa_CA, Other Sample/Population). There are positive non-zero *Z* scores for a few contemporaneous populations, however, nothing above the significance threshold (Data S2G). The summary of evidence indicates that she may not have grown up in the same local area as the other La Sassa individuals, but her genetic affinities require higher coverage and more comparative samples to be certain.

At Broion, all the individuals directly dated to the Chalcolithic (BRC001, BRC013, and BRC022) and/or that fall in the EN cluster (BRC011; Figure 2A) are males, and all of the directly dated individuals show first- and second-degree relationships (Tables 1 and 2; Data S1). These three share the Ychr G2a-P15 marker and could have the exact same haplotype (G2a3-F1193-F2291; Table 1; Data S1B and S1F); however, due to differences in coverage, the terminal branch markers are not covered in all three individuals (Data S1F). BRC013 and BRC022 share an mtDNA haplotype (H5a1; Data S1B) and have a first-degree relationship although BRC001/023 (J2a1a1) has a second-degree relationship with BRC013, but not BRC022 (Table 2). The PI_HAT values support the first-degree relationship between BRC013 and BRC022 as full siblings as well as the differing second-degree relationship between BRC013, BRC022, and BRC001/023 (Table 2). Given these values, the most parsimonious scenario is that BRC013 and BRC022 are brothers and BRC013 is the grandfather of BRC001/023; the radiocarbon dates do not reject this scenario (STAR Methods; Data S1C).

Among the 22 novel individuals, we only found two cases of matching mtDNA (Data S1B and S1E) among a variety of haplogroups associated with the N transition N1, H, J, and K,[6,10] indicating a lack of detectable structure at the mitochondrial level, which could be consistent with a larger maternal population size, exogamy, and/or with a patrilocal kinship structure across both time periods. To further explore these scenarios, we analyzed ROHs (STAR Methods) in the ancient and selected modern populations (TSI; Figures 4A–4C; Data S5C) with hapROH, which is a method to detect ROH segments in low-coverage genomes using a haplotype reference panel (STAR Methods).[61] We checked that differences in coverage did not systematically bias estimates (Figure 4D) and checked against imputation and sequencing errors (STAR Methods; Figure S7; Data S5B). We calculated segment number and length in four length categories: <1.6 centimorgan (cM); >1.6 cM; >4 cM; and >8 cM (Data S5C) and focus on the greater than 1.6 cM length category (which includes 4 and 8 cM segments), which is informative regarding recent consanguinity and is the most reliably inferred (STAR Methods; Figure S5; Data S5).

Estimates of ROH > 1.6 cM for the ancient samples after the Mesolithic[13] fall within the range of values obtained from modern Italian (TSI) individuals, suggesting similar levels of endogamy (Figures 4A and 4B; Data S5C); however, there is a significant (two-tailed t test; p = 0.0003) difference between the lengths of >1.6-cM segments in Italian N (this study and Antonio et al.[13]) and Italian BA (newly reported) and the number of >1.6-cM segments (two-tailed t test; p = 0.0001; Figure 4C; Data S5D), which is consistent with either larger effective population sizes of the BA or as the result of added diversity following an admixture event with the local Chalcolithic populations. Within the La Sassa site, individual LSC002/4 has the highest total length of >1.6-cM segments and is the only one with detected segments >8 cM (Data S5C).

Shifts in phenotypic features of ancient Italians

To determine whether the shifts in ancestry components through time corresponded to any shifts in phenotypes, we imputed genetic markers related to 115 phenotypes associated with metabolism, immunity, and pigmentation in the ancient samples presented here and in previously published studies (Data S6A–S6D; STAR Methods).[7,13,46,64,65] We analyzed a total of 332 ancient individuals (16 presented here for the first time and 316 from the literature) grouped by population: Italian Mesolithic (n = 3),[13] Italian N/Chalcolithic (n = 52; this study and Antonio et al.[13]), Italian BA (n = 60; this study, Fernandes et al.,[16] and Marcus et al.[17]), Italian IA/Modern (n = 133),[13] Near East N/Chalcolithic (n = 41),[46,64] Near East BA (n = 18),[7,46,65] and Yamnaya (n = 18; Data S6A).[4,7,64] For groups with a sample size larger than three, we calculated the frequency of the effective allele for each phenotypic variant in each population and then performed an ANOVA test to analyze shifts in the allele frequency. We compared both between Italian and non-Italian groups from different periods (Data S6B) as well as between groups within Italy, grouping the Italian individuals into 12 cohorts based both on period and geographic location (Data S6C). For both tests, we applied a Bonferroni's correction to set the significance threshold and used a Tukey test to determine the significant pairs (STAR Methods).

Eleven variants were significant when comparing the Italian groups with Near Eastern and Yamnaya populations and eight in the intra-Italy test, with four significant in both tests (Table 3; Data S6B and S6C). Although these results should be interpreted with caution due to the small sample size, some potentially interesting results emerge. For the variants that are significant in both tests (*TLR1* [rs5743618], *TNF* [rs1800629], *HLA* [rs3135388], and *SLC45A2* [rs16891982]), the signal is driven almost entirely by the post-Roman Republic Central Italy sample group (Cen_postRep), which includes Roman, Late Antiquity, and Medieval individuals. There is no detectable difference between the Italian BA samples presented here and the Italian N/Chalcolithic groups, despite the additional Steppe-related ancestry. Three out of four of the variants highlighted here are linked to protection and susceptibility to Hansen's disease (leprosy). The *HLA*-related variant (rs3135388), indicated in the susceptibility to physical manifestations of the disease in a Danish medieval population,[39] was significantly different between post-BA Italians and N/Chalcolithic Near East, BA Italians, and Yamnaya. The statistical significance of this variant in this test is probably due to the low frequency of the protective allele in central Italians from the IA onward (Data S6C). Another variant (*TNF*—rs1800629)[66] also seems to decrease in frequency of the protective allele (Table 3; Data S6C). Both results are consistent with the rise in frequency of this disease in the historical and archaeological record in Europe. The other variant, *TLR1* (rs5743618), which has been linked to both protection against leprosy[38,67] and an increased susceptibility to

Please cite this article in press as: Saupe et al., Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.04.022

CellPress
OPEN ACCESS

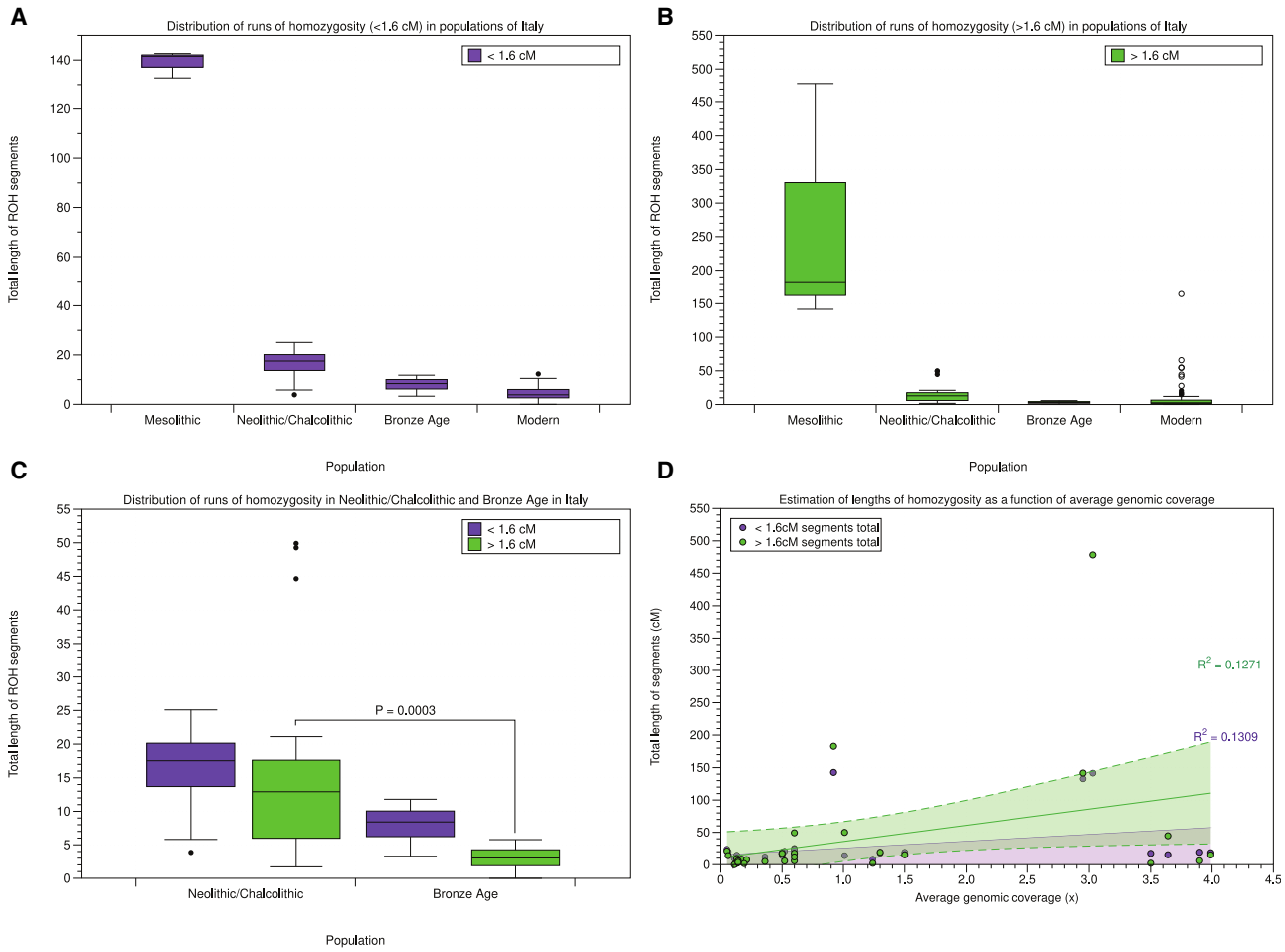

Figure 4. Runs of homozygosity in ancient and modern populations of the Italian Peninsula

(A) Distribution of total length of segments <1.6 centimorgan (cM) in Mesolithic,[13] Neolithic/Chalcolithic (Antonio et al.[13] and this study), Bronze Age (this study), and Modern Italians (TSI).[62]

(B) Distribution of total length of segments >1.6 cM in Mesolithic,[13] Neolithic/Chalcolithic (Antonio et al.[13] and this study), Bronze Age (this study), and Modern Italians (TSI).[62]

(C) Distribution of the total length of segments <1.6 cM (purple) and >1.6 cM (green) in Neolithic/Chalcolithic (Antonio et al.[13] and this study) and Bronze Age (this study).

(D) Total lengths of segments <1.6 cM (purple) and >1.6 cM (green) in all ancient samples (Antonio et al.[13] and this study) plotted against the average genomic coverage. Linear regression fitted by Datagraph (http://visualdatatools.com),[63] solid line is the fitted line, and shaded area represents 95% confidence interval. R^2 values are color coded to match the legend and reported in the lower right-hand corner.

See also Data S5.

tuberculosis in Asian populations,[68] shows a significant result in the other direction, driven by higher frequency only in the post-Roman Republic Central Italy group (Table 3; Data S6C).

The fourth variant significant in both tests (rs16891982 in *SLC45A2* gene) is implicated in hair and eye pigmentation. In terms of physical appearance, both the Chalcolithic and BA Italy groups have imputed phenotypes more similar to IA and Later Romans than to earlier populations in Italy and the Near East. The three previously published Mesolithic individuals from Italy[13,27,29] are predicted to have dark skin, dark hair, and blue eyes (Data S6D), although most of the other samples have predicted intermediate skin pigmentation, brown hair, and brown eyes; however, individuals with blue eyes paired with either dark or blond hair are also predicted in all time periods except in the N individuals from Central Italy (Data S6D).[13] The variant rs16891982, linked to darker eyes and hair, shows a significant difference between post-BA Italy group and previous groups from Italy, with the frequency decreasing in Central Italy starting in the newly reported Chalcolithic individuals and with the lowest values observed for the newly reported BA individuals and the post-Roman Republic Central Italy group. This difference is particularly notable compared to N Central Italy and Sardinian groups prior to the BA (Data S6C).

DISCUSSION

The newly generated genomes provide a more detailed description of demographic dynamics of later Italian prehistory in a

Please cite this article in press as: Saupe et al., Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.04.022

Table 3. List of significant phenotypic variants

SNP	Gene	Phenotype	Comparison between Italy, Near East, and Yamnaya significance	Significant pairs of groups	Intra-Italy comparison significance	Significant pairs of groups
rs2167079	*NR1H3, ACP2*	HDL +	y	Yam[a] versus all except NE_BA		
rs174546	*FADS1/ FADS2*	LDL +	y	Ita_postBA versus NE_NeoCA and Yam[a]		
rs174570	*FADS1*vs.3	LDL +;HDL +;TG −	y	Yam versus NE_BA and Ita_postBA[a]; Ita_NeoCA versus Ita_postBA		
rs5743618	*TLR1*	protective factor for leprosy	y	Ita_postBA[a] versus Ita_NeoCA, NE_NeoCA, Ita_BA	y	Cen_postRep[a] versus Cen_NeoCA, Sic_Neo, Sar_CA, Sic_BA, Sar_BANur, Sar_IAPun, Cen_IARep
rs1800629	*TNF*	protective factor for leprosy	y	Ita_postBA versus Ita_NeoCA and NE_NeoCA[a]	y	Cen_postRep versus CA[a], Sar_BANur
rs3135388	*HLA*	protective factor for leprosy	y	Ita_postBA versus NE_NeoCA[a], Ita_BA, Yam	y	Sar_BANur[a] versus Cen_IARep, Cen_postRep; Sar_CA versus Cen_IARep
rs4251545	*IRAK4*	risk factor for Gram-positive infection			y	not significant pairs
rs13119723	intronic	protective factor for celiac disease	y	Yam[a] versus all Ita		
rs2187668	*HLA-DQA1*	risk factor for celiac disease and gluten intolerance	y	NE_NeoCA and Yam[a] versus all Ita		
rs7775228	*HLA-DRA*	risk factor for celiac disease and gluten intolerance			y	BA[a] versus Sar_Neo and all Cen
rs4713586	*HLA-DRA*	risk factor for celiac disease and gluten intolerance			y	BA versus all except CA[b]
rs1050152	*SLC22A4*	risk factor for Crohn's disease	y	NE_NeoCA versus Ita_BA[a] and Ita_postBA		
rs16891982	*SLC45A2*	eye color: brown; hair color: black versus nonvs.black	y	Ita_postBA versus NE_NeoCA, Ita_NeoCA[a], Ita_BA	y	Cen_postRep versus Cen_NeoCA[a], Sar_CA, Sar_BANur
rs28777	*SLC45A2*	hair color: AAvs.black/CCvs.red; skin color: dark	y	Ita_postBA versus Ita_NeoCA[a], Ita_BA		
rs1426654	*SLC24A5*	skin color: intermediate			y	Cen_NeoCA[a] versus all

See also Data S6. BA, Bronze Age; CA, Chalcolithic; Cen, Central Italians; IA, Iron Age; Ita, Italy; NE, Near East; Neo, Neolithic; Nur, Nuragic; Pun, Punic; Rep, Republic; Sar, Sardinia; Sic, Sicily; SNP, variantID; Yam, Yamnaya.

[a]The groups with the highest frequency of the effective allele in the significant pairs

[b]Frequency approaching 1 in most groups, except BA and CA

European context.[7,9,11,13] The split among European Early N individuals into two groups observed in the PCA and already reported in Marcus et al.[17] separates Sardinia N from mainland Italy, coupled with the higher affinity of the Sardinia N to both Anatolia_N and WHG, raise the possibility of population structure within the EN component, although deeper analysis, including high-coverage ancient genomes, is needed to dissect subtle differences in ancestry.

Our analyses show the expected signature of peri- and post-BA movements from Steppe-related populations across Italy: absent in Italian individuals from the N and Chalcolithic, emerging in the Early BA (Italian Bell Beaker [I2478: 2195–1940 calBCE],[9] Italian Remedello [RISE486: 2134–1773 calBCE],[7,13] and Broion [BRC010: 1952–1752 calBCE (95.4%)]) and increasing through time in the individuals from Broion and Regina Margherita (GCP003: 1626–1497 calBCE [95.4%]). These samples confirm the date of arrival in Northern Italy to at least ~2000 BCE and its presence in Central Italy by 4 centuries later, although denser sampling strategies are needed to assess the dynamics of this spread. Our qpAdm results suggest that the Steppe-related ancestry component could have arrived through Late N/Bell Beaker groups from Central Europe, though what remains unknown due to small sample size and limited geographical and chronological distribution is whether there were multiple Steppe population sources and the exact timing and diffusion of this ancestry component through the Italian Peninsula. The R1b subtype found in BA Broion is a lineage found in both ancient Sicilian samples and Italian Bell Beakers. Together with the autosomal affinity of North and Central Italian BA groups with Late N Germany, the Ychr data point to a possibly Northern-, transalpine-, and potentially Bell-Beaker-associated source of the Italian Steppe-related ancestry.

The importance of male kinship structures in the interface between the Chalcolithic and BA has also been explored using our autosomal data. It has long been assumed that the commingled cave burials of the Chalcolithic included some form of kinship structure; however, it was not possible to directly reveal it before the advent of aDNA. Here, we see the pattern that, in the Chalcolithic period, these locations were preferentially used to bury closely related male individuals, though the social significance of this fact is not clear. Although the Chalcolithic populations of Italy utilized natural burial chamber spaces, rock-cut tombs, and trench graves more than building megalithic monuments of the kind seen along the Atlantic Façade in an earlier time period, it appears that the importance of burying related males together is a shared feature. The genetic evidence shown here is consistent with an emphasis on patrilineal descendancy and patrilocality for these burial rites to the populations at both La Sassa and Broion, an emphasis that disappears in the BA but is also not present at the single-burial style Chalcolithic cemetery of Gattolino (possibly due to small sample size). It is important to note that these sites do not represent a random and unbiased sampling of the local populations, rather a snapshot of one particular ritual aspect of these societies; thus, it cannot be inferred whether patrilocality and patrilineality were generally practiced or whether these patterns changed over time. More sampling for genetics and isotopes is needed to reconstruct the general population structure and inter-community relationships.

The arrival of the Steppe-related ancestry does not seem to have affected the frequency patterns of any of the phenotypes assessed in this work. Rather, the biggest changes seem to have occurred with or after the Roman Imperial period. The decrease in alleles associated with a protection against leprosy after the IA is potentially interesting, given the increase in manifestations of the disease in the European bioarchaeological and historical record from the 3rd to 4th millennium BCE[69] until its decline around the 1st millennium CE.[70] It is still not clear exactly how these variants interact with the disease and other pathogenic mycobacterial infections; thus, more work is needed on the clinical side before the full evolutionary history can be determined. It is also important to note that, as we did not test all possible phenotypes but only a small subset, our results are likely to be not the only phenotypic differences and more work must be done to fully understand the complex relationships between evolutionary mechanisms and human genes. Fortunately, whole ancient genomes like those generated in this study provide an invaluable resource that can be revisited in light of advancements in all areas of biology and genetics.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Grottina dei Covoloni del Broion
  - Grotta La Sassa
  - Necropoli di Gattolino di Cesena
  - Radiocarbon method and date of Gattolino
  - Grotta Regina Margherita
  - Sample B297 comes from Sounding D, Context 32
  - Samples B152 and B154 come from Sounding E, Context 60
- METHOD DETAILS
  - DNA extraction
  - Library preparation
  - DNA sequencing
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Mapping
  - aDNA authentication
  - Calculating genetic sex estimation
  - Determining mtDNA haplogroups
  - Y chromosome variant calling and haplotyping
  - Kinship analysis and identical samples
  - Identifying kinship in new samples
  - Preparing the datasets for autosomal analysis
  - Principal component analysis
  - DyStruct
  - $f4$ statistics
  - Outgroup f3 statistics
  - Admixture analysis
  - Chromopainter/NNLS and SourceFind

- f4 NNLS analysis
- qpAdm
- Genome imputation
- Runs of Homozygosity
- Predicting imputation errors
- Introducing imputation errors into modern genomes
- Detecting ROH segments in ancient genomes
- Phenotype prediction

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2021.04.022.

ACKNOWLEDGMENTS

This work is supported by European Union through the European Regional Development Fund (project no. 2014–2020.4.01.16–0030; C.L.S., T.S., F.M., and M.M.) and the Estonian Research Council personal research grant (PRG243; C.L.S. and M.M.). This research was additionally supported by the European Union through the European Regional Development Fund (Project No. 2014-2020.4.01.15-0012; M.M.) and European Union through Horizon 2020 grant no. 810645 (M.M.). C.C. was supported for this project by St. Hugh's College via a Jackie Lambert Research Grant. E.D. was supported by Sapienza University of Rome fellowship "borsa di studio per attività di perfezionamento all'estero 2017." L.P. was supported by UniPd PRID 2019 and by European Regional Development Fund project no. 2014-2020.4.01.16-0024, MOBTT53. R.H. was supported by the Wellcome Trust (award no. 2000368/Z/15/Z). T.K. is supported by KU Leuven startup grant STG/18/021. S.T. is funded by the European Research Council under the European Union's Horizon 2020 Research and Innovation Programme (grant agreement no. 803147-951 RESOLUTION; https://site.unibo.it/resolution-erc/en). S.B. is supported by ERC no. 724046 – SUCCESS; http://www.erc-success.eu/. R.S. is funded by the British Academy (SG140575). L.A. was supported by the Netherlands Organisation for Scientific Research (NWO) Free Competition grant 360-61-060. A.D. contributed sequencing funds via the Faculty Research Fund of the Faculty of Humanities and Social Sciences, Newcastle University. We thank the other members of the Estonian Biocentre and aDNA group, including Lehti Saag, for their expertise and Paula Reimer and the 14Chrono Centre at Queen's University Belfast for assistance with the radiocarbon dates.

AUTHOR CONTRIBUTIONS

Conceptualization, C.L.S., L.P., C.C., T.S., and F.M.; investigation, T.S., G.L., and O.L. (aDNA); I.A. (isotopes); and S.T. (radiocarbon); resources, N.C., C.S., L.P., C.C., S.B., M.F.R., F.D.A., L.S., L.A., and R.S.; formal analysis, T.S., F.M., C.L.S., T.K., E.D., R.H., and L.P.; writing – original draft, T.S., F.M., L.P., and C.L.S.; writing – review & editing, all authors; funding acquisition, A.D., L.P., and M.M.; supervision, C.L.S., L.P., and F.M.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 10, 2020
Revised: November 28, 2020
Accepted: April 9, 2021
Published: May 10, 2021

REFERENCES

1. Pearce, M. (2019). The "copper age"—a history of the concept. J. World Prehist. *32*, 229–250.
2. Dolfini, A. (2020). From the Neolithic to the Bronze Age in Central Italy: settlement, burial, and social change at the dawn of metal production. J. Archaeol. Res. *28*, 503–556.
3. Skoglund, P., Malmström, H., Raghavan, M., Storå, J., Hall, P., Willerslev, E., Gilbert, M.T.P., Götherström, A., and Jakobsson, M. (2012). Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. Science *336*, 466–469.
4. Mathieson, I., Alpaslan-Roodenberg, S., Posth, C., Szécsényi-Nagy, A., Rohland, N., Mallick, S., Olalde, I., Broomandkhoshbacht, N., Candilio, F., Cheronet, O., et al. (2018). The genomic history of southeastern Europe. Nature *555*, 197–203.
5. Hofmanová, Z., Kreutzer, S., Hellenthal, G., Sell, C., Diekmann, Y., Díez-Del-Molino, D., van Dorp, L., López, S., Kousathanas, A., Link, V., et al. (2016). Early farmers from across Europe directly descended from Neolithic Aegeans. Proc. Natl. Acad. Sci. USA *113*, 6886–6891.
6. Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S., Kirsanow, K., Sudmant, P.H., Schraiber, J.G., Castellano, S., Lipson, M., et al. (2014). Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature *513*, 409–413.
7. Allentoft, M.E., Sikora, M., Sjögren, K.-G., Rasmussen, S., Rasmussen, M., Stenderup, J., Damgaard, P.B., Schroeder, H., Ahlström, T., Vinner, L., et al. (2015). Population genomics of Bronze Age Eurasia. Nature *522*, 167–172.
8. Olalde, I., Mallick, S., Patterson, N., Rohland, N., Villalba-Mouco, V., Silva, M., Dulias, K., Edwards, C.J., Gandini, F., Pala, M., et al. (2019). The genomic history of the Iberian Peninsula over the past 8000 years. Science *363*, 1230–1234.
9. Olalde, I., Brace, S., Allentoft, M.E., Armit, I., Kristiansen, K., Booth, T., Rohland, N., Mallick, S., Szécsényi-Nagy, A., Mittnik, A., et al. (2018). The Beaker phenomenon and the genomic transformation of northwest Europe. Nature *555*, 190–196.
10. Haak, W., Lazaridis, I., Patterson, N., Rohland, N., Mallick, S., Llamas, B., Brandt, G., Nordenfelt, S., Harney, E., Stewardson, K., et al. (2015). Massive migration from the steppe was a source for Indo-European languages in Europe. Nature *522*, 207–211.
11. Raveane, A., Aneli, S., Montinaro, F., Athanasiadis, G., Barlera, S., Birolo, G., Boncoraglio, G., Di Blasio, A.M., Di Gaetano, C., Pagani, L., et al. (2019). Population structure of modern-day Italians reveals patterns of ancient and archaic ancestries in Southern Europe. Sci. Adv. *5*, eaaw3492.
12. Lazaridis, I., Mittnik, A., Patterson, N., Mallick, S., Rohland, N., Pfrengle, S., Furtwängler, A., Peltzer, A., Posth, C., Vasilakis, A., et al. (2017). Genetic origins of the Minoans and Mycenaeans. Nature *548*, 214–218.
13. Antonio, M.L., Gao, Z., Moots, H.M., Lucci, M., Candilio, F., Sawyer, S., Oberreiter, V., Calderon, D., Devitofranceschi, K., Aikens, R.C., et al. (2019). Ancient Rome: a genetic crossroads of Europe and the Mediterranean. Science *366*, 708–714.
14. Keller, A., Graefen, A., Ball, M., Matzas, M., Boisguerin, V., Maixner, F., Leidinger, P., Backes, C., Khairat, R., Forster, M., et al. (2012). New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. Nat. Commun. *3*, 698.
15. Fu, Q., Posth, C., Hajdinjak, M., Petr, M., Mallick, S., Fernandes, D., Furtwängler, A., Haak, W., Meyer, M., Mittnik, A., et al. (2016). The genetic history of Ice Age Europe. Nature *534*, 200–205.
16. Fernandes, D.M., Mittnik, A., Olalde, I., Lazaridis, I., Cheronet, O., Rohland, N., Mallick, S., Bernardos, R., Broomandkhoshbacht, N., Carlsson, J., et al. (2020). The spread of steppe and Iranian-related ancestry in the islands of the western Mediterranean. Nat. Ecol. Evol. *4*, 334–345.
17. Marcus, J.H., Posth, C., Ringbauer, H., Lai, L., Skeates, R., Sidore, C., Beckett, J., Furtwängler, A., Olivieri, A., Chiang, C.W.K., et al. (2020). Genetic history from the Middle Neolithic to present on the Mediterranean island of Sardinia. Nat. Commun. *11*, 939.
18. Mittnik, A., Massy, K., Knipper, C., Wittenborn, F., Friedrich, R., Pfrengle, S., Burri, M., Carlichi-Witjes, N., Deeg, H., Furtwängler, A., et al. (2019). Kinship-based social inequality in Bronze Age Europe. Science *366*, 731–734.

19. Schroeder, H., Margaryan, A., Szmyt, M., Theulot, B., Włodarczak, P., Rasmussen, S., Gopalakrishnan, S., Szczepanek, A., Konopka, T., Jensen, T.Z.T., et al. (2019). Unraveling ancestry, kinship, and violence in a Late Neolithic mass grave. Proc. Natl. Acad. Sci. USA *116*, 10705–10710.

20. Racimo, F., Sikora, M., Vander Linden, M., Schroeder, H., and Lalueza-Fox, C. (2020). Beyond broad strokes: sociocultural insights from the study of ancient genomes. Nat. Rev. Genet. *21*, 355–366.

21. Sánchez-Quinto, F., Malmström, H., Fraser, M., Girdland-Flink, L., Svensson, E.M., Simões, L.G., George, R., Hollfelder, N., Burenhult, G., Noble, G., et al. (2019). Megalithic tombs in western and northern Neolithic Europe were linked to a kindred society. Proc. Natl. Acad. Sci. USA *116*, 9469–9474.

22. Scheib, C.L., Hui, R., D'Atanasio, E., Wohns, A.W., Inskip, S.A., Rose, A., Cessford, C., O'Connell, T.C., Robb, J.E., Evans, C., et al. (2019). East Anglian early Neolithic monument burial linked to contemporary Megaliths. Ann. Hum. Biol. *46*, 145–149.

23. Saag, L., Varul, L., Scheib, C.L., Stenderup, J., Allentoft, M.E., Saag, L., Pagani, L., Reidla, M., Tambets, K., Metspalu, E., et al. (2017). Extensive farming in Estonia started through a sex-biased migration from the Steppe. Curr. Biol. *27*, 2185–2193.e6.

24. Martiniano, R., Cassidy, L.M., Ó'Maoldúin, R., McLaughlin, R., Silva, N.M., Manco, L., Fidalgo, D., Pereira, T., Coelho, M.J., Serra, M., et al. (2017). The population genomics of archaeological transition in west Iberia: Investigation of ancient substructure using imputation and haplotype-based methods. PLoS Genet. *13*, e1006852.

25. Knipper, C., Mittnik, A., Massy, K., Kociumaka, C., Kucukkalipci, I., Maus, M., Wittenborn, F., Metz, S.E., Staskiewicz, A., Krause, J., and Stockhammer, P.W. (2017). Female exogamy and gene pool diversification at the transition from the Final Neolithic to the Early Bronze Age in central Europe. Proc. Natl. Acad. Sci. USA *114*, 10083–10088.

26. Furholt, M. (2019). Re-integrating archaeology: a contribution to aDNA studies and the migration discourse on the 3rd millennium BC in Europe. Proc. Prehist. Soc. *85*, 115–129.

27. Olalde, I., Allentoft, M.E., Sánchez-Quinto, F., Santpere, G., Chiang, C.W.K., DeGiorgio, M., Prado-Martinez, J., Rodríguez, J.A., Rasmussen, S., Quilez, J., et al. (2014). Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. Nature *507*, 225–228.

28. Brace, S., Diekmann, Y., Booth, T.J., van Dorp, L., Faltyskova, Z., Rohland, N., Mallick, S., Olalde, I., Ferry, M., Michel, M., et al. (2019). Ancient genomes indicate population replacement in Early Neolithic Britain. Nat. Ecol. Evol. *3*, 765–771.

29. Saag, L., Laneman, M., Varul, L., Malve, M., Valk, H., Razzak, M.A., Shirobokov, I.G., Khartanovich, V.I., Mikhaylova, E.R., Kushniarevich, A., et al. (2019). The arrival of Siberian ancestry connecting the Eastern Baltic to Uralic speakers further East. Curr. Biol. *29*, 1701–1711.e16.

30. Wilde, S., Timpson, A., Kirsanow, K., Kaiser, E., Kayser, M., Unterländer, M., Hollfelder, N., Potekhina, I.D., Schier, W., Thomas, M.G., and Burger, J. (2014). Direct evidence for positive selection of skin, hair, and eye pigmentation in Europeans during the last 5,000 y. Proc. Natl. Acad. Sci. USA *111*, 4832–4837.

31. Mathieson, S., and Mathieson, I. (2018). FADS1 and the timing of human adaptation to agriculture. Mol. Biol. Evol. *35*, 2957–2970.

32. Roberts, C.A., Lewis, M.E., and Manchester, K. (2002). The past and present of leprosy: archaeological, historical, palaeopathological and clinical approaches. In Proceedings of the 3rd International Congress on the Evolution and Palaeoepidemiology of the Infectious Diseases, ICEPID, 26-31 July 1999, University of Bradford, C.A. Roberts, et al., eds. (Archaeopress).

33. Mariotti, V., Dutour, O., Belcastro, M.G., Facchini, F., and Brasili, P. (2005). Probable early presence of leprosy in Europe in a Celtic skeleton of the 4th-3rd century BC (Casalecchio di Reno, Bologna, Italy). Int. J. Osteoarchaeol. *15*, 311–325.

34. Kalisch, P.A. (1975). An overview of research on the history of leprosy. Part 1. From Celsus to Simpson, Circa. 1 AD Part 2. From Virchow to Møller-Christense, 1845-1973. Int. J. Lepr. Other Mycobact. Dis. *43*, 129–144.

35. Schuring, R.P., Hamann, L., Faber, W.R., Pahan, D., Richardus, J.H., Schumann, R.R., and Oskam, L. (2009). Polymorphism N248S in the human Toll-like receptor 1 gene is related to leprosy and leprosy reactions. J. Infect. Dis. *199*, 1816–1819.

36. Wong, S.H., Gochhait, S., Malhotra, D., Pettersson, F.H., Teo, Y.Y., Khor, C.C., Rautanen, A., Chapman, S.J., Mills, T.C., Srivastava, A., et al. (2010). Leprosy and the adaptation of human toll-like receptor 1. PLoS Pathog. *6*, e1000979.

37. Sapkota, B.R., Macdonald, M., Berrington, W.R., Misch, E.A., Ranjit, C., Siddiqui, M.R., Kaplan, G., and Hawn, T.R. (2010). Association of TNF, MBL, and VDR polymorphisms with leprosy phenotypes. Hum. Immunol. *71*, 992–998.

38. Misch, E.A., Macdonald, M., Ranjit, C., Sapkota, B.R., Wells, R.D., Siddiqui, M.R., Kaplan, G., and Hawn, T.R. (2008). Human TLR1 deficiency is associated with impaired mycobacterial signaling and protection from leprosy reversal reaction. PLoS Negl. Trop. Dis. *2*, e231.

39. Krause-Kyora, B., Nutsua, M., Boehme, L., Pierini, F., Pedersen, D.D., Kornell, S.-C., Drichel, D., Bonazzi, M., Möbus, L., Tarp, P., et al. (2018). Ancient DNA study reveals HLA susceptibility locus for leprosy in medieval Europeans. Nat. Commun. *9*, 1569.

40. Monroy Kuhn, J.M., Jakobsson, M., and Günther, T. (2018). Estimating genetic kin relationships in prehistoric populations. PLoS ONE *13*, e0195491.

41. Reimer, P.J., Austin, W.E.N., Bard, E., Bayliss, A., Blackwell, P.G., Ramsey, C.B., Butzin, M., Cheng, H., Lawrence Edwards, R., Friedrich, M., et al. (2020). The IntCal20 Northern Hemisphere Radiocarbon Age Calibration Curve (0–55 cal kBP). Radiocarbon *62*, 725–757.

42. Stuiver, M., Reimer, P.J., and Reimer, R.W. (2021). CABLIB 8.2. http://calib.org/calib/.

43. Ramsey, C.B. (2009). Bayesian analysis of radiocarbon dates. Radiocarbon *51*, 337–360.

44. Patterson, N., Price, A.L., and Reich, D. (2006). Population structure and eigenanalysis. PLoS Genet. *2*, e190.

45. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. *38*, 904–909.

46. Lazaridis, I., Nadel, D., Rollefson, G., Merrett, D.C., Rohland, N., Mallick, S., Fernandes, D., Novak, M., Gamarra, B., Sirak, K., et al. (2016). Genomic insights into the origin of farming in the ancient Near East. Nature *536*, 419–424.

47. Lipson, M., Szécsényi-Nagy, A., Mallick, S., Pósa, A., Stégmár, B., Keerl, V., Rohland, N., Stewardson, K., Ferry, M., Michel, M., et al. (2017). Parallel palaeogenomic transects reveal complex genetic history of early European farmers. Nature *551*, 368–372.

48. Brunel, S., Bennett, E.A., Cardin, L., Garraud, D., Barrand Emam, H., Beylier, A., Boulestin, B., Chenal, F., Ciesielski, E., Convertini, F., et al. (2020). Ancient genomes from present-day France unveil 7,000 years of its demographic history. Proc. Natl. Acad. Sci. USA *117*, 12791–12798.

49. Rivollat, M., Jeong, C., Schiffels, S., Küçükkalıpçı, İ., Pemonge, M.-H., Rohrlach, A.B., Alt, K.W., Binder, D., Friederich, S., Ghesquière, E., et al. (2020). Ancient genome-wide DNA from France highlights the complexity of interactions between Mesolithic hunter-gatherers and Neolithic farmers. Sci. Adv. *6*, eaaz5344.

50. Shennan, S. (2018). The First Farmers of Europe: An Evolutionary Perspective (Cambridge University Press).

51. Sarno, S., Boattini, A., Pagani, L., Sazzini, M., De Fanti, S., Quagliariello, A., Gnecchi Ruscone, G.A., Guichard, E., Ciani, G., Bortolini, E., et al. (2017). Ancient and recent admixture layers in Sicily and Southern Italy

trace multiple migration routes along the Mediterranean. Sci. Rep. *7*, 1984.

52. Joseph, T.A., and Pe'er, I. (2019). Inference of population structure from time-series genotype data. Am. J. Hum. Genet. *105*, 317–333.
53. Alexander, D.H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Res. *19*, 1655–1664.
54. Sikora, M., Carpenter, M.L., Moreno-Estrada, A., Henn, B.M., Underhill, P.A., Sánchez-Quinto, F., Zara, I., Pitzalis, M., Sidore, C., Busonero, F., et al. (2014). Population genomic analysis of ancient and modern genomes yields new insights into the genetic ancestry of the Tyrolean Iceman and the genetic structure of Europe. PLoS Genet. *10*, e1004353.
55. Lawson, D.J., Hellenthal, G., Myers, S., and Falush, D. (2012). Inference of population structure using dense haplotype data. PLoS Genet. *8*, e1002453.
56. Chacón-Duque, J.-C., Adhikari, K., Fuentes-Guajardo, M., Mendoza-Revilla, J., Acuña-Alonzo, V., Barquera, R., Quinto-Sánchez, M., Gómez-Valdés, J., Everardo Martínez, P., Villamil-Ramírez, H., et al. (2018). Latin Americans show wide-spread Converso ancestry and imprint of local Native ancestry on physical appearance. Nat. Commun. *9*, 5388.
57. Ongaro, L., Scliar, M.O., Flores, R., Raveane, A., Marnetto, D., Sarno, S., Gnecchi-Ruscone, G.A., Alarcón-Riquelme, M.E., Patin, E., Wangkumhang, P., et al. (2019). The genomic impact of European colonization of the Americas. Curr. Biol. *29*, 3974–3986.e4.
58. Kivisild, T. (2017). The study of human Y chromosome variation through ancient DNA. Hum. Genet. *136*, 529–546.
59. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience *4*, 7.
60. Hui, R., D'Atanasio, E., Cassidy, L.M., Scheib, C.L., and Kivisild, T. (2020). Evaluating genotype imputation pipeline for ultra-low coverage ancient genomes. Sci. Rep. *10*, 18542.
61. Ringbauer, H., Novembre, J., and Steinrücken, M. (2020). Human parental relatedness through time - detecting runs of homozygosity in ancient DNA. bioRxiv. https://doi.org/10.1101/2020.05.31.126912.
62. The 1000 Genomes Project Consortium (2015). A global reference for human genetic variation. Nature *526*, 68–74.
63. MacAskill, M.R. (2012). DataGraph 3.0. J. Statist. Softw. Softw. Rev. *47*, 1–9.
64. Mathieson, I., Lazaridis, I., Rohland, N., Mallick, S., Patterson, N., Roodenberg, S.A., Harney, E., Stewardson, K., Fernandes, D., Novak, M., et al. (2015). Genome-wide patterns of selection in 230 ancient Eurasians. Nature *528*, 499–503.
65. Damgaard, P.B., Marchi, N., Rasmussen, S., Peyrot, M., Renaud, G., Korneliussen, T., Moreno-Mayar, J.V., Pedersen, M.W., Goldberg, A., Usmanova, E., et al. (2018). 137 ancient human genomes from across the Eurasian steppes. Nature *557*, 369–374.
66. Zhang, F.-R., Huang, W., Chen, S.-M., Sun, L.-D., Liu, H., Li, Y., Cui, Y., Yan, X.-X., Yang, H.-T., Yang, R.-D., et al. (2009). Genomewide association study of leprosy. N. Engl. J. Med. *361*, 2609–2618.
67. Johnson, C.M., Lyle, E.A., Omueti, K.O., Stepensky, V.A., Yegin, O., Alpsoy, E., Hamann, L., Schumann, R.R., and Tapping, R.I. (2007). Cutting edge: a common polymorphism impairs cell surface trafficking and functional responses of TLR1 but protects against leprosy. J. Immunol. *178*, 7520–7524.
68. Qi, H., Sun, L., Wu, X., Jin, Y., Xiao, J., Wang, S., Shen, C., Chu, P., Qi, Z., Xu, F., et al. (2015). Toll-like receptor 1(TLR1) gene SNP rs5743618 is associated with increased risk for tuberculosis in Han Chinese children. Tuberculosis (Edinb.) *95*, 197–203.
69. Köhler, K., Marcsik, A., Zádori, P., Biro, G., Szeniczey, T., Fábián, S., Serlegi, G., Marton, T., Donoghue, H.D., and Hajdu, T. (2017). Possible cases of leprosy from the Late Copper Age (3780-3650 cal BC) in Hungary. PLoS ONE *12*, e0185966.
70. Donoghue, H.D., Marcsik, A., Matheson, C., Vernon, K., Nuorala, E., Molto, J.E., Greenblatt, C.L., and Spigelman, M. (2005). Co-infection of Mycobacterium tuberculosis and Mycobacterium leprae in human archaeological samples: a possible explanation for the historical decline of leprosy. Proc. Biol. Sci. *272*, 389–394.
71. Behar, D.M., van Oven, M., Rosset, S., Metspalu, M., Loogväli, E.-L., Silva, N.M., Kivisild, T., Torroni, A., and Villems, R. (2012). A "Copernican" reassessment of the human mitochondrial DNA tree from its root. Am. J. Hum. Genet. *90*, 675–684.
72. McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A.R., Teumer, A., Kang, H.M., Fuchsberger, C., Danecek, P., Sharp, K., et al.; Haplotype Reference Consortium (2016). A reference panel of 64,976 haplotypes for genotype imputation. Nat. Genet. *48*, 1279–1283.
73. Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb. Protoc. *2010*, pdb.prot5448.
74. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J. *17*, 10–12.
75. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.
76. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.
77. DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. *43*, 491–498.
78. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., and Orlando, L. (2013). mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics *29*, 1682–1684.
79. Jones, E.R., Zarina, G., Moiseyev, V., Lightfoot, E., Nigst, P.R., Manica, A., Pinhasi, R., and Bradley, D.G. (2017). The Neolithic transition in the Baltic was not driven by admixture with Early European farmers. Curr. Biol. *27*, 576–582.
80. Korneliussen, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics *15*, 356.
81. Skoglund, P., Storå, J., Götherström, A., and Jakobsson, M. (2013). Accurate sex identification of ancient human remains using DNA shotgun sequencing. J. Archaeol. Sci. *40*, 4477–4482.
82. Weissensteiner, H., Pacher, D., Kloss-Brandstätter, A., Forer, L., Specht, G., Bandelt, H.-J., Kronenberg, F., Salas, A., and Schönherr, S. (2016). HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. Nucleic Acids Res. *44*, W58–W63.
83. Quinlan, A.R. (2014). BEDTools: the Swiss-Army tool for genome feature analysis. Curr. Protoc. Bioinformatics *47*, 11.12.1–11.12.34.
84. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575.
85. Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al.; 1000 Genomes Project Analysis Group (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156–2158.
86. Browning, B.L., and Browning, S.R. (2016). Genotype imputation with millions of reference samples. Am. J. Hum. Genet. *98*, 116–126.
87. R Development Core Team (2013). R: A language and environment for statistical computing (R Foundation for Statistical Computing).
88. Link, V., Kousathanas, A., Veeramah, K., Sell, C., Scheu, A., and Wegmann, D. (2017). ATLAS: analysis tools for low-depth and ancient samples. bioRxiv. https://doi.org/10.1101/105346.

89. Chaitanya, L., Breslin, K., Zuñiga, S., Wirken, L., Pośpiech, E., Kukla-Bartoszek, M., Sijen, T., Knijff, P., Liu, F., Branicki, W., et al. (2018). The HIrisPlex-S system for eye, hair and skin colour prediction from DNA: introduction and forensic developmental validation. Forensic Sci. Int. Genet. *35*, 123–135.

90. Walsh, S., Liu, F., Wollstein, A., Kovatsi, L., Ralf, A., Kosiniak-Kamysz, A., Branicki, W., and Kayser, M. (2013). The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA. Forensic Sci. Int. Genet. *7*, 98–115.

91. Fedele, F. (2013). I Covoloni del Broion (colli Berici, VI). In L'età del Rame. La pianura Padana e le Alpi al tempo di Ötzi., R. De Marinis, ed. (Compagnia della Stampa), pp. 450–456.

92. Alessandri, L., and Rolfo, M.F. (2016). L'utilizzo delle cavità naturali nella media età del Bronzo: nuovi dati dal Lazio meridionale.*CVII* (Bollettino della Unione Storia ed Arte), pp. 109–126.

93. Anzidei, A.P., and Carboni, G. (2013). L'eneolitico recente e finale nel Lazio centro-meridionale: una puntualizzazione sullo sviluppo e la durata di alcuni aspetti culturali sulla base delle più recenti datazioni radiometriche. In Cronologia Assoluta e Relativa Dell'età Del Rame in Italia, C. Genick, ed. (QuiEdit), pp. 98–118.

94. Alessandri, L. (2019). The early and Middle Bronze Age (1/2) in South and central Tyrrhenian Italy and their connections with the Avellino eruption: An overview. Quat. Int. *499*, 161–185.

95. Alessandri, L., Baiocchi, V., Del Pizzo, S., Rolfo, M.F., and Troisi, S. (2019). Photogrammetric survey with fisheye lens for the characterization of the La Sassa cave. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. *42-2/W9*, 25–32.

96. Anzidei, A.P., and Carboni, G. (2006). Rinaldone e Gaudo in un territorio di confine: il Lazio centro-meridionale. In Atti Del VII Incontro Di Studi Preistoria e Protostoria in Etruria. Pastori e Guerrieri Nell'Etruria Del IV e III Millennio a.C. La Civiltá Di Rinaldone a 100 Anni Dalle Prime Scoperte, N. Catacchio, ed. (Centro Studi di Preistoria e Archeologia), pp. 174–192.

97. Lovejoy, C.O. (1985). Dental wear in the Libben population: its functional pattern and role in the determination of adult skeletal age at death. Am. J. Phys. Anthropol. *68*, 47–56.

98. Rogers, T.L. (2005). Determining the sex of human remains through cranial morphology. J. Forensic Sci. *50*, 493–500.

99. Talamo, S., and Richards, M. (2011). A comparison of bone pretreatment methods for AMS dating of samples > 30,000 BP. Radiocarbon *53*, 443–449.

100. Longin, R. (1971). New method of collagen extraction for radiocarbon dating. Nature *230*, 241–242.

101. Brown, T., Nelson, D., Vogel, J., and Southon, J. (1988). Improved collagen extraction by modified Longin method. Radiocarbon *30*, 171–177.

102. Brock, F., Ramsey, C., and Higham, T. (2007). Quality assurance of ultrafiltered bone dating. Radiocarbon *49*, 187–192.

103. van Klinken, G.J. (1999). Bone collagen quality indicators for palaeodietary and radiocarbon measurements. J. Archaeol. Sci. *26*, 687–695.

104. Kromer, B., Lindauer, S., Synal, H.-A., and Wacker, L. (2013). MAMS – a new AMS facility at the Curt-Engelhorn-Centre for Achaeometry, Mannheim, Germany. Nucl. Instrum. Methods Phys. Res. B *294*, 11–13.

105. Korlević, P., Talamo, S., and Meyer, M. (2018). A combined method for DNA analysis and radiocarbon dating from a single sample. Sci. Rep. *8*, 4127.

106. Angle, M., Catracchia, F., Cavazzuti, C., Celletti, P., Malorgio, M., and Mancini, D. (2010). La grotta Regina Margherita a Collepardo (Frosinone). In Lazio e Sabina. Atti del Convegno: Sesto incontro di studi sul Lazio e la Sabina, G. Ghini, ed. (Atti del Convegno. Sesto Incontro di Studi sul La-zio e la Sabina), pp. 381–396.

107. Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., Stiller, M., Schubert, M., Cappellini, E., Petersen, B., Moltke, I., et al. (2013). Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. Nature *499*, 74–78.

108. Malaspinas, A.-S., Lao, O., Schroeder, H., Rasmussen, M., Raghavan, M., Moltke, I., Campos, P.F., Sagredo, F.S., Rasmussen, S., Gonçalves, V.F., et al. (2014). Two ancient human genomes reveal Polynesian ancestry among the indigenous Botocudos of Brazil. Curr. Biol. *24*, R1035–R1037.

109. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.

110. van Oven, M., and Kayser, M. (2009). Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum. Mutat. *30*, E386–E394.

111. Hallast, P., Batini, C., Zadik, D., Maisano Delser, P., Wetton, J.H., Arroyo-Pardo, E., Cavalleri, G.L., de Knijff, P., Destro Bisol, G., Dupuy, B.M., et al. (2015). The Y-chromosome tree bursts into leaf: 13,000 high-confidence SNPs covering the majority of known clades. Mol. Biol. Evol. *32*, 661–673.

112. Karmin, M., Saag, L., Vicente, M., Wilson Sayres, M.A., Järve, M., Talas, U.G., Rootsi, S., Ilumäe, A.-M., Mägi, R., Mitt, M., et al. (2015). A recent bottleneck of Y chromosome diversity coincides with a global change in culture. Genome Res. *25*, 459–466.

113. Poznik, G.D., Xue, Y., Mendez, F.L., Willems, T.F., Massaia, A., Wilson Sayres, M.A., Ayub, Q., McCarthy, S.A., Narechania, A., Kashin, S., et al.; 1000 Genomes Project Consortium (2016). Punctuated bursts in human male demography inferred from 1,244 worldwide Y-chromosome sequences. Nat. Genet. *48*, 593–599.

114. Broushaki, F., Thomas, M.G., Link, V., López, S., van Dorp, L., Kirsanow, K., Hofmanová, Z., Diekmann, Y., Cassidy, L.M., Díez-Del-Molino, D., et al. (2016). Early Neolithic genomes from the eastern Fertile Crescent. Science *353*, 499–503.

115. Harney, É., May, H., Shalem, D., Rohland, N., Mallick, S., Lazaridis, I., Sarig, R., Stewardson, K., Nordenfelt, S., Patterson, N., et al. (2018). Ancient DNA from Chalcolithic Israel reveals the role of population mixture in cultural transformation. Nat. Commun. *9*, 3336.

116. Günther, T., Malmström, H., Svensson, E.M., Omrak, A., Sánchez-Quinto, F., Kılınç, G.M., Krzewińska, M., Eriksson, G., Fraser, M., Edlund, H., et al. (2018). Population genomics of Mesolithic Scandinavia: Investigating early postglacial migration routes and high-latitude adaptation. PLoS Biol. *16*, e2003703.

117. Mittnik, A., Wang, C.-C., Pfrengle, S., Daubaras, M., Zariņa, G., Hallgren, F., Allmäe, R., Khartanovich, V., Moiseyev, V., Tõrv, M., et al. (2018). The genetic prehistory of the Baltic Sea region. Nat. Commun. *9*, 442.

118. van den Brink, E.C.M., Beeri, R., Kirzner, D., Bron, E., Cohen-Weinberger, A., Kamaisky, E., Gonen, T., Gershuny, L., Nagar, Y., Ben-Tor, D., et al. (2017). A Late Bronze Age II clay coffin from Tel Shaddud in the Central Jezreel Valley, Israel: context and historical implications. Levant *49*, 105–135.

119. Valdiosera, C., Günther, T., Vera-Rodríguez, J.C., Ureña, I., Iriarte, E., Rodríguez-Varela, R., Simões, L.G., Martínez-Sánchez, R.M., Svensson, E.M., Malmström, H., et al. (2018). Four millennia of Iberian biomolecular prehistory illustrate the impact of prehistoric migrations at the far end of Eurasia. Proc. Natl. Acad. Sci. USA *115*, 3428–3433.

120. Fregel, R., Méndez, F.L., Bokbot, Y., Martín-Socas, D., Camalich-Massieu, M.D., Santana, J., Morales, J., Ávila-Arcos, M.C., Underhill, P.A., Shapiro, B., et al. (2018). Ancient genomes from North Africa evidence prehistoric migrations to the Maghreb from both the Levant and Europe. Proc. Natl. Acad. Sci. USA *115*, 6774–6779.

121. Jones, E.R., Gonzalez-Fortes, G., Connell, S., Siska, V., Eriksson, A., Martiniano, R., et al. (2015). Upper Palaeolithic genomes reveal deep roots of modern Eurasians. Nat. Commun. *6*.

122. Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient admixture in human history. Genetics *192*, 1065–1093.

123. Busby, G.B.J., Hellenthal, G., Montinaro, F., Tofanelli, S., Bulayeva, K., Rudan, I., et al. (2015). The role of recent admixture in forming the contemporary West Eurasian genomic landscape. Curr. Biol. *25*, 2878–2526.

124. Browning, S.R., Browning, B.L., Zhou, Y., Tucci, S., and Akey, J.M. (2018). Analysis of human sequence data reveals two pulses of archaic Denisovan admixture. Cell *173*, 53–61.e9.

125. Gamba, C., Jones, E.R., Teasdale, M.D., McLaughlin, R.L., Gonzalez-Fortes, G., Mattiangeli, V., Domboróczki, L., Kővári, I., Pap, I., Anders, A., et al. (2014). Genome flux and stasis in a five millennium transect of European prehistory. Nat. Commun. *5*, 5257.

126. Narasimhan, V., Danecek, P., Scally, A., Xue, Y., Tyler-Smith, C., and Durbin, R. (2016). BCFtools/RoH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data. Bioinformatics *32*, 1749–1751.

Please cite this article in press as: Saupe et al., Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.04.022

CellPress
OPEN ACCESS

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Ancient skeletal element	This paper	BRC001
Ancient skeletal element	This paper	BRC002
Ancient skeletal element	This paper	BRC003
Ancient skeletal element	This paper	BRC007
Ancient skeletal element	This paper	BRC010
Ancient skeletal element	This paper	BRC011
Ancient skeletal element	This paper	BRC013
Ancient skeletal element	This paper	BRC015
Ancient skeletal element	This paper	BRC022
Ancient skeletal element	This paper	BRC024
Ancient skeletal element	This paper	BRC030
Ancient skeletal element	This paper	GCP002
Ancient skeletal element	This paper	GCP003
Ancient skeletal element	This paper	GCP004
Ancient skeletal element	This paper	GLR001
Ancient skeletal element	This paper	GLR002
Ancient skeletal element	This paper	GLR003
Ancient skeletal element	This paper	GLR004
Ancient skeletal element	This paper	LSC002
Ancient skeletal element	This paper	LSC005
Ancient skeletal element	This paper	LSC007
Ancient skeletal element	This paper	LSC011
Chemicals, peptides, and recombinant proteins		
Sodium Hypochlorite solution (15%)	N/A	CAS:7681-52-9
0.5 M EDTA pH 8.0	Fisher Scientific	Cat# BP24821
Ethanol 96%	Chemlab	Cat# CL00.0507.1000
dNTP Mix (25mM each)	Thermo Fisher Scientific	Cat# R1122
dNTP Mix (10mM each)	Thermo Fisher Scientific	Cat# R0192
BSA	Thermo Fisher Scientific	Cat# B14
HGS Diamond Taq	Eurogentec	Cat# TAQ-I011-5000+
Critical commercial assays		
MinElute PCR Purification Kit	QIAGEN	Cat# 28006
High Pure Viral Nucleic Acid LargeVolume Kit	Roche	Cat# 5114403001
NEBNext DNA Library Prep Master Mix Set 454	New England Biolabs	Cat# E6070L
Qubit dsDNA BR Assay Kit	Thermo Fisher Scientific	Cat# Q32853
Qubit dsDNA HS Assay Kit	Thermo Fisher Scientific	Cat# Q32854
Fragment Analyzer High Sensitivity NGS Fragment Analysis Kit	Agilent	Cat# DNF-474-0500
D1000 ScreenTape	Agilent	Cat# 5067-5582
D1000 Reagents	Agilent	Cat# 5067-5583
High Sensitivity D1000 ScreenTape	Agilent	Cat# 5067-5584

(*Continued on next page*)

Please cite this article in press as: Saupe et al., Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.04.022

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
High Sensitivity D1000 Reagents	Agilent	Cat# 5067-5585
KAPA Library Quantification Kit	Roche	Cat# KK4835
Deposited data		
Human reference genome NCBI build 37, GRCh37	Genome Reference Consortium	https://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/human/
Mitochondrial DNA reference genome, cRSRS	Behar et al.[71]	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3322232/#!po=70.8333
Compiled modern and ancient comparison dataset (including restricted access samples) 1240K and Human Origins	N/A	https://reichdata.hms.harvard.edu/pub/datasets/amh_repo/curated_releases/index_v42.4.html
1000 Genomes Project Phase 3	The 1000 Genomes Project Consortium[62]	https://www.internationalgenome.org/category/phase-3/
Haplotype Reference Consortium	McCarthy et al.[72]	http://www.haplotype-reference-consortium.org/
Italian aDNA data	This paper	http://www.ebc.ee/free_data; http://www.ebi.ac.uk/ena/data/view/PRJEB37660; ENA: PRJEB37660
Oligonucleotides		
NEBNext Multiplex Oligos for Illumina	New England Biolabs	Cat# E7335
IS1_adapter.P5 A*C*A*C*TCTTTCCCTACACG ACGCTCTTCCG*A*T*C*T	Meyer and Kircher[73]	Eurofins
IS2_adapter.P7 G*T*G*A*CTGGAGTTCAGACGTG TGCTCTTCCG*A*T*C*T	Meyer and Kircher[73]	Eurofins
IS3_adapter.P5+P7 A*G*A*T*CG GAA*G*A*G*C	Meyer and Kircher[73]	Eurofins
Software and algorithms		
cutadapt	Martin[74]	https://cutadapt.readthedocs.io/en/stable/#
Burrow-Wheeler Aligner (BWA 0.7.12)	Li and Durbin[75]	http://bio-bwa.sourceforge.net/
Samtools 1.3, 1.6, 1.9	Li et al.[76]	http://samtools.sourceforge.net/
Picard 2.12, 1.54	N/A	http://broadinstitute.github.io/picard/
GATK 3.5	DePristo et al.[77]	https://gatk.broadinstitute.org
mapDamage 2.0	Jónsson et al.[78]	https://ginolhac.github.io/mapDamage/
mtDNA contamination algorithm	Jones et al.[79]	N/A
ANGSD-0.916	Korneliussen et al.[80]	http://www.popgen.dk/angsd/index.php/ANGSD
sex determination algorithm #1	Skoglund et al.[81]	https://github.com/pontussk/ry_compute
sex determination algorithm #2	Fu et al.[15]	N/A
HaploGrep2	Weissensteiner et al.[82]	https://haplogrep.i-med.ac.at/category/haplogrep2/
BEDTools	Quinlan[83]	https://bedtools.readthedocs.io/en/latest/
PLINK v1.90b3.27	Purcell et al.[84]	https://www.partners.org/purcell/plink/
EIGENSOFT 7.2.0	Patterson et al.[44]	https://github.com/DReichLab/EIG
ADMIXTOOLS	Price et al.[45]	https://github.com/DReichLab/AdmixTools
ADMIXTURE	Alexander et al.[53]	https://bio.tools/admixture
ChromoPainter/NNLS pipeline	Lawson et al.[55]	https://people.maths.bris.ac.uk/~madjl/finestructure-old/chromopainter_info.html
SourceFind	Chacón-Duque et al.[56]	https://people.maths.bris.ac.uk/~madjl/finestructure/sourcefind.html

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
READ	Monroy Kuhn et al.[40]	https://bitbucket.org/tguenther/read
VCFtools	Danecek et al.[85]	http://vcftools.sourceforge.net/
Datagraph	MacAskill[63]	https://visualdatatools.com/
Beagle 4.1, 5.0	Browning and Browning[86]	https://faculty.washington.edu/browning/beagle/b4_1.html
R 3.6	R Development Core Team[87]	http://www.R-project.org/
ATLAS v0.9.0	Link et al.[88]	https://bitbucket.org/wegmannlab/atlas/wiki/Home
hapROH	Ringbauer et al.[61]	https://pypi.org/project/hapROH/
HIrisPlex-S	Chaitanya et al.[89] and Walsh et al.[90]	https://hirisplex.erasmusmc.nl/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact: Tina Saupe (tina.saupe@ut.ee).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The accession number for the DNA sequences reported in this paper is ENA:PRJEB37660 (https://www.ebi.ac.uk/ena/data/view/PRJEB37660). The data are also available through the data depository of the EBC (http://evolbio.ut.ee).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The DNA was extracted from 47 bone and tooth samples in 51 extracts, 51 double-stranded, single-indexed libraries generated: 30 from Grottina dei Covoloni del Broion, 13 from Grotta La Sassa, four from the Necropoli di Gattolino di Cesena and four from Grotta Regina Margherita. Forty-nine libraries were NGS screened at low depth and 20 sequenced to higher coverage and analyzed. More detailed information about the archaeological sites of this study are given below.

Grottina dei Covoloni del Broion

Tina Saupe & Cinzia Scaggion

The Grottina dei Covoloni del Broion is part of the oriental rocky wall complex of Colli Berici. It is close to the famous cave Grotta Broion in the province of Vicenza in Northern Italy where among Paleolithic sediments, the Italian geologist Piero Leonardi found significant traces of Neanderthal. This prehistoric shelter is located in Vallà di Lumignano in the municipality of Longare and was used as a funerary site. Through the analysis of the palaeosurface, the site seems to have been used for a long time covering many generations belonging to the Chalcolithic confirmed by the discovery of Eneolithic artifacts.

Grottina dei Covoloni del Broion was discovered in 1973 and systematic excavations were conducted of the different geological layers and of parts of the Grottina during four archeological campaigns between 1973 and 1977. The archaeological excavation of 1977 was entirely reserved for the study of funerary depositions. The inner part of the cave was intended for sepulchral use such as was evidenced by the discovery of many human osteological findings incorporated in the calcareous sediment. The archaeologists had also found ceramic fragments and many flint artifacts: a flat blade of dagger, a significant amount of arrow cusps, blades, two flint cores, small discoid beads and a small calcite tube from a bracelet or necklace, and other personal ornaments found in the same layer. The human remains include bones and teeth under different states of conservation and distribution.[91] For this study, in collaboration with the University of Padova and Museum of Anthropology of Padova, samples of 26 human remains (pars petrous bone = 4, teeth = 22) were taken with unknown morphological background of age and sex. Samples BRC001, BRC002, BRC003, BRC007, BRC010, BRC013, BRC022, and BRC030 were dated at ^{14}CHRONO Centre for Climate, the Environment, and Chronology in Belfast, UK (Data S1C).

Grotta La Sassa

Luca Alessandri, Ilenia Arienzo, Flavio De Angelis and Mario Federico Rolfo

The Grotta La Sassa (National Cave Cadastre id: LA 2001) was discovered in 2015 during a survey of the Ausoni Mountains natural caves carried out by two speleological groups: Gruppo Grotte Castelli Romani and Speleo Club Roma.[92] Several archaeological surveys have been carried out in the cave, from 2016 to the present day by the Groningen Institute of Archaeology and the University of

Please cite this article in press as: Saupe et al., Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.04.022

Rome Tor Vergata. The cave is in the valley close to the modern town of Sonnino (Latina, Italy). The coordinates of the entrance are: WGS84, UTM 33T, 352627E, 4587452N. The cave yields an impressive stratigraphy ranging from the late Pleistocene to the Second World War, when the cave was used as a hiding shelter by the locals. Hundreds of disarticulated human bones and teeth scattered in Room 1 (soundings N, WD and L) and Room 2 were collected (Figure S1A). They have been radiocarbon dated to the Copper Age and the Early Bronze Age (for the Italian chronology, see Anzidei and Carboni[93] and Alessandri[94]). The sampled bones pertaining to this work come from the stratigraphic units (SUs) 19, 31 and 36, in Room 2 (Figure S1B).

The three radiocarbon dates that come from human bones in SU 19 set the layer in the Copper Age/Chalcolithic. SUs 31 and 36 belong to a group of layers which contain Middle Bronze Age potsherds, possibly reworked in Medieval times (Figure S1). However, the four radiocarbon dates which have been obtained from these samples are all tightly clustered around 4,050 BP (Figure S1D) which, in Italian chronology, falls into the Late Copper Age. No pottery items were found in association with these bones, despite the only Copper Age potsherds collected in the cave (Figure 13 in Alessandri et al.[95]) belonging to the Gaudo funeral facies,[96] which is characteristic of Southern Italy.

The anthropological samples were recruited for micro-sampling at the University of Rome "Tor Vergata" and the sampling process was performed at Oxford University afterward. The visual preservation status was the driving inclusion criterion for the recruitment.

The osteological characteristics of the selected bones suggested that the fragments could belong to different individuals, even though the molecular analyses later showed that the bone fragments pertaining to LSC002 and LSC004 come from the same individual. These bone fragments consist of diagenetically damaged petrous bones and their siding are not reliably assessed by morphological evaluation.

LSC005 is a lower second molar tooth pertaining to an adult individual, whose age at death could be set as 35-40 years old according to classical osteological methods.[97]

LSC007 is an upper first molar related to a young adult whose age at death should be between 15 and 20 years old according to the lack of enamel wear.[97]

LSC011 refers to a lower first molar dental element carried by a mandible of an adult male individual whose age at death should be set up to 40 years old.[97,98]

The samples are already submitted to multi-isotope analysis for unravelling diet and mobility of people buried in Grotta La Sassa. The mobility of the ancient people buried in La Sassa is currently under evaluation and the data will be conclusive shortly with the publication of a dedicated paper in preparation. For the purpose of the present paper, the punctual data pertaining to the sampled teeth are reported. Remarkably, LSC005 returns an enamel 87Sr/86Sr ratio of 0.709549, which is out of the range determined for the rest of 27 sampled teeth from the cave. The LSC011 and LSC007 samples return an enamel 87Sr/86Sr ratio of 0.709308 and 0.708800, respectively, that are tuned with the data recovered from the rest of the teeth in La Sassa.

Samples LSC002/004, LSS005/013 and LSC011 were dated at ^{14}CHRONO Centre for Climate, the Environment, and Chronology in Belfast, UK (Data S1C).

Necropoli di Gattolino di Cesena

Monica Miari

The Necropoli of Gattolino is located in the municipality of Cesena (Emilia-Romagna region, central Italy). The four graves were probably part of a larger Copper Age necropolis. The graves were all single, NW-SE oriented and the inhumations were in a lying supine position. The burials included a vessel placed at the foot of the body in addition to flint artifacts, in particular arrow-heads. Based on the materials, the necropolis highlights important contacts with central and southern Italy - with particular regard to the Laterza *facies* - and seems to be framed from a recent phase of the Copper Age to the beginning of the Bronze Age, due to the presence in the Grave 1 of beads in Sicilian amber and silver.

Radiocarbon method and date of Gattolino

Sahra Talamo

The Gattolino bone collagen was extracted at the Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology (MPI-EVA) in Leipzig, Germany, following the pretreatment procedures in Talamo and Richards[99] (MPI-Code: S-EVA). The outer surface of the bone sample is first cleaned by a shot blaster and then 500 mg of the whole bone is taken. The samples are then decalcified in 0.5M HCl at room temperature until no CO_2 effervescence is observed. 0.1M NaOH is added for 30 minutes to remove humics. The NaOH step is followed by a final 0.5M HCl step for 15 minutes. The resulting solid is gelatinized following Longin[100] at pH 3 in a heater block at 75°C for 20h. The gelatine is then filtered in an Eeze-Filter (Elkay Laboratory Products (UK)) to remove small (> 80 mm) particles. The gelatine is then ultrafiltered[101] with Sartorius "VivaspinTurbo" 30 KDa ultrafilters. Prior to use, the filter is cleaned to remove carbon containing humectants.[102] The samples are lyophilized for 48 hours. C:N atomic ratios, and collagen yields were measured to determine the extent of collagen preservation. Bones with > 1% weight collagen and C:N ratios in the range 2.9–3.6 are passing the evaluation criteria for collagen to proceed with the AMS analysis.[103] Samples were graphitised and dated by AMS at the Mannheim facility (laboratory code MAMS[104]).

In order to monitor contamination introduced during the pretreatment stage, a sample from a cave bear bone, kindly provided by D. Döppes (MAMS, Germany), was extracted along with the batch of La Ferrassie samples.[105] The Gattolino collagen sample passed the evaluation criteria for good quality collagen and is reported in Data S1C. The Radiocarbon date of Gattolino 3 was calibrated using IntCal20;[41] with the OxCal 4.4 program.[43] The C:N ratio and the amount of collagen extracted (%Coll, > 30 kDa fraction) are reported.

Please cite this article in press as: Saupe et al., Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.04.022

Grotta Regina Margherita

Robin Skeates

Grotta Regina Margherita is situated in the modern municipality of Collepardo (Frosinone province, Westcentral Italy). It is a large limestone karst cave. It is located at an altitude of 480 m in the Fiume valley, and opens about 30 m above the bottom of a gorge and below the hill on which the Medieval village of Collepardo now sits. It is the largest known mortuary cave in Middle Bronze Age Central Italy (c. 1650-1400 BCE). Essentially, the cave can be divided into two contrasting zones: the twilight Entrance Hall and the dark Interior Hall. Since 2008,[106] our team has excavated and compared soundings in the archaeological deposits found in diverse areas of the cave. The three aDNA samples of human petrous bone come from two different soundings in the Interior Hall. In this part of the cave, we found clear evidence for primary and secondary mortuary rites having been performed in the Middle Bronze Age in well-defined spaces enclosed by speleothems- distinguished archaeologically by concentrations of human remains and body ornaments. This mortuary activity in the cave has been radiocarbon dated to between c. 1650 and 1450 cal BC.

Sample B297 comes from Sounding D, Context 32

Sounding D lies in a relatively well-defined sunken space delimited by large speleothems. Here, a rich and relatively extensive mortuary deposit has been identified, comprising a dense and compact 'carpet' of human bones and associated artifacts (Context 32), embedded in a fine cave loam and patches of calcite crust. Finds include: human and animal bones, pottery fragments, a clay spindle whorl, two fragments of a small bronze ornament made of a cylindrical spiral of wire to which is attached a fragment of a faïence bead, a disc-shaped bead or button of mother-of-pearl, a quadrangular piece of sandstone, with wear traces from use in sharpening or smoothing, and some pieces of charcoal. The human remains recovered so far (from the upper two thirds of the deposit) belong to at least 19 individuals, with ten adults (including at least three young adults and one older adult of 40-60 years) and 9 sub-adults (one fetus/perinatal individual, one infant of 18 months, one infant of approximately 2 years, one child of approximately 3-6 years, one child of approximately 8 years, one older child of 10 years, and three adolescents). There are at least two females and one male adult. Despite the generally disarticulated (and also somewhat fragmented) state of the human remains, a few of the bones (including some phalanges) were found in articulation, and all bones from the skeleton (including many small and fragile bones) were represented, suggesting the original deposition of whole bodies in this area. However, the significant under-representation of long bones (particularly the lower limbs-tibiae and femurs) hints at the successive removal of large bones from this area, which might then have been redeposited elsewhere in or beyond the cave. The presence of mineralised breaks and calcite accretions covering breakage points also indicates that most of the bones were fragmented prior to the formation of the calcite, quite possibly unintentionally in the Bronze Age during the course of primary and secondary mortuary rites, which added to and disturbed earlier mortuary deposits in this area.

We obtained 10 radiocarbon determinations on human bones from Sounding D, to gain some idea of the period of time over which these mortuary deposits were formed, selecting only left heels (astragali) to ensure that we were dating different individuals. The basic time span provided by these determinations, at the 68 per cent probability level, is c. 1600-1450 cal BCE. Using Bayesian chronological modeling, at the same probability level, the period of activity can be narrowed down to 1-60 years, falling within 1545-1480 cal BCE.

Samples B152 and B154 come from Sounding E, Context 60

Sounding E comprises a small sunken space situated among a group of speleothems. It is the innermost area excavated in the cave. The deposits in this area are loose, having been heavily disturbed by the installation of the tourist walkway. Context 60 is the uppermost layer of these deposits. Numerous concrete human bones, a few pottery fragments, and some charcoal and ashes were found here. A minimum number of 7 individuals is represented by the human bones, with three adults and four sub-adults (one infant of less than one year, one child of around 6 years, one child of around 9 years, and one adolescent). There is at least one female adult and one male adult. Exceptionally, a humerus and an ulna were found in anatomical connection here. This, together with the relatively high average frequency of representation of skeletal elements in this area (50 per cent), indicates the successive primary deposition of whole bodies in this area.

A human left astragalus (SUERC-78150) from Sounding E has a radiocarbon date of c. 1500–1450 cal BC, which is contemporary with some of the dated left astragali from Sounding D.

Samples GCP002 and GCP003 were dated at ^{14}CHRONO Centre for Climate, the Environment, and Chronology in Belfast, UK (Data S1C).

METHOD DETAILS

All of the laboratory work was performed in dedicated ancient DNA laboratories at the Estonian Biocenter, Institute of Genomics, University of Tartu, Tartu, Estonia. The library quantification and sequencing were performed at the Estonian Biocenter Core Laboratory. The main steps of the laboratory work are detailed below.

DNA extraction

In total 47 samples from human remains were extracted for DNA analysis. The four petrous bones from Broion cave were sampled twice, giving a total of 51 extracts.

Please cite this article in press as: Saupe et al., Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.04.022

The first layer of pars petrous was removed with a sterilized drill bite to avoid exogenous contamination. Bone powder and a 10 mm core of the inner ear were sampled from the pars petrous. The drill bits and core drill were sterilized in between samples with 6% (w/v) bleach followed by distilled water and then ethanol rinse. Root portions of teeth were removed with a sterile drill wheel.

The root and the petrous portions were soaked in 6% (w/v) bleach for 5 minutes. Samples were rinsed three times with 18.2 MΩcm H_2O and soaked in 70% (v/v) Ethanol for 2 minutes. The tubes were shaken during the procedure to dislodge particles. The samples were transferred to a clean paper towel on a rack inside a class IIB hood with the UV light on and allowed to dry for two to three hours.

Afterward, the samples were weighed to calculate the accurate volume of EDTA (20x EDTA [μl] of sample mass [mg]) and Proteinase K (0.5x Proteinase K [μl] of sample mass [mg]). EDTA and Proteinase K were added into PCR-clean 5 mL or 15 mL conical tubes (Eppendorf) along with the samples inside the IIB hood and the tubes were incubated 72 h on a slow shaker at room temperature. 1000 μl of EDTA and 25 μl of Proteinase K were added to the bone powder of the pars petrous and were incubated 24 h on a slow shaker at room temperature.

The DNA extracts (of root portions, pars petrous portions and bone powder) were concentrated to 250 μl using the Vivaspin Turbo 15 (Sartorius) and purified in large volume columns (High Pure Viral Nucleic Acid Large Volume Kit, Roche) using 2.5 mL of PB buffer, 1 mL of PE buffer and 100 μL of EB buffer (MinElute PCR Purification Kit, QIAGEN). For the elution of the endogenous DNA, the silica columns were transferred to a collection tube to dry and followed in 1.5 mL DNA lo-bind tubes (Eppendorf) to elute. The samples were incubated with 100 μl EB buffer at 37 C for 10 minutes and centrifuged at 13,000 rpm for two minutes. After centrifugation, the silica columns were removed and the samples were stored at −20 C. Only one extraction was performed per extraction for screening and 30 μl used for libraries.

Library preparation

Sequencing libraries were built using NEBNext DNA Library Prep Master Mix Set for 454 (E6070, New England Biolabs) and Illumina-specific adaptors[73] following established protocols.[73,107,108] The end repair module was implemented using 18.75 μL of water, 7.5 μL of buffer and 3.75 μL of enzyme mix, incubating at 20°C for 30 minutes. The samples were purified using 500 μL PB and 650 μL of PE buffer and eluted in 30 μL EB buffer (MinElute PCR Purification Kit, QIAGEN). The adaptor ligation module was implemented using 10 μL of buffer, 5 μL of T4 ligase and 5 μL of adaptor mix,[73] incubating at 20°C for 15 minutes. The samples were purified as in the previous step and eluted in 30 μL of EB buffer (MinElute PCR Purification Kit, QIAGEN). The adaptor fill-in module was implemented using 13 μL of water, 5 μL of buffer and 2 μL of Bst DNA polymerase, incubating at 37°C for 30 and at 80°C for 20 minutes. Libraries were amplified using the following PCR set up: 50 μL DNA library, 1X PCR buffer, 2.5mM MgCl2, 1 mg/ml BSA, 0.2 μM inPE1.0, 0.2mM dNTP each, 0.1U/μl HGS Taq Diamond and 0.2 μM indexing primer. Cycling conditions were: 5′ at 94C, followed by 18 cycles of 30 s each at 94C, 60C, and 68C, with a final extension of 7 minutes at 72C. The samples were purified and eluted in 35 μL of EB buffer (MinElute PCR Purification Kit, QIAGEN). Three verification steps were implemented to make sure library preparation was successful and to measure the concentration of dsDNA/sequencing libraries - fluorometric quantitation (Qubit, Thermo Fisher Scientific), parallel capillary electrophoresis (Fragment Analyzer, Agilent Technologies) and qPCR.

DNA sequencing

DNA was sequenced using the Illumina NextSeq500/550 High-Output single-end 75 cycle kit. As a norm, 15 samples were sequenced together on one flow cell; additional data was generated for 20 samples to increase coverage (Data S1).

QUANTIFICATION AND STATISTICAL ANALYSIS

Mapping

Before mapping, the sequences of the adapters, indexes, and poly-G tails occurring due to the specifics of the NextSeq 500 technology were cut from the ends of DNA sequences using cutadapt-1.11.[74] Sequences shorter than 30 bp were also removed with the same program to avoid random mapping of sequences from other species. The sequences were aligned to the reference sequence GRCh37 (hs37d5) using Burrows- Wheeler Aligner (BWA 0.7.12)[75] and the command mem with re-seeding disabled.

After alignment, the sequences were converted to BAM format and only sequences that mapped to the human genome were kept with samtools 1.3.[76] Afterward, the data from different flow cell lanes were merged and duplicates were removed using picard 2.12 (http://broadinstitute.github.io/picard/index.html). Indels were realigned using GATK 3.5[109] and reads with a mapping quality less than 10 were filtered out using samtools 1.3.[76]

aDNA authentication

As a result of degrading over time, aDNA can be distinguished from modern DNA by certain characteristics: short fragments and a high frequency of C = > T substitutions at the 5′ ends of sequences due to cytosine deamination. The program mapDamage2.0[78] was used to estimate the frequency of 5′ C = > T transitions. Rates of contamination were estimated on mitochondrial DNA by calculating the percentage of non-consensus bases at haplogroup-defining positions as detailed in Jones et al.[79] Each sample was mapped against the RSRS downloaded from phylotree.org and checked against haplogroup-defining sites for the sample-specific haplogroup (Data S1).

Samtools 1.9[76] option *stats* was used to determine the number of final reads, average read length, average coverage etc. The average endogenous DNA content (proportion of reads mapping to the human genome) was 10.77% (0.48 - 48.87%) (Data S1A and S1B).

Calculating genetic sex estimation

Genetic sex was calculated using the methods described in Skoglund et al.,[81] estimating the fraction of reads mapping to Y chromosome out of all reads mapping to either X or Y chromosome. Additionally, sex was determined using a method described in Fu et al.,[15] calculating the X and Y ratio by the division of the coverage by the autosomal coverage. Here, the sex was calculated for samples with a coverage > 0.01 × and only reads with a mapping quality > 10 were counted for the autosomal, X, and Y chromosome (Data S1A and S1B).

Determining mtDNA haplogroups

Mitochondrial DNA haplogroups were determined using Haplogrep2 on the command line. For the determination, the reads were realigned to the reference sequence RSRS and the parameter -rsrs were given to estimate the haplogroups using Haplogrep2[82,110] (Data S1B). Subsequently, the identical results between the individuals were checked visually by aligning mapped reads to the reference sequence using samtools 0.1.19[76] command *tview* and confirming the haplogroup assignment in PhyloTree. Additionally, private mutations were noted for further kinship analysis (Data S1E). The polymorphisms were estimated using the online platform of haplogrep2. Here, the variant calling files (*vcf.*) were uploaded to the online platform and the known polymorphism in the RSRS were converted to rCRS (Data S1E).

Y chromosome variant calling and haplotyping

A total of 138,425 binary Y chromosome SNPs that have been detected as polymorphic in previous high coverage whole Y chromosome sequencing studies[111–113] were called in Chalcolithic/Bronze Age samples using ANGSD-0.916[80] command–doHaploCall. Only ten individual samples that had more than 0.01 × Y chromosome variant coverage were kept for further analyses. A subset of 105,691 sites yielded a call in at least one of the samples and in case of 2,480 sites at least one of the ten samples carried a derived allele. Basal haplogroup affiliations of each sample were determined by assessing the proportion of derived allele calls (pD) in a set of primary (A, B, C...T) haplogroup defining internal branches, as defined in Karmin et al.,[112] using 1,677 informative sites. In case of all ten samples the primary haplogroup could be determined unambiguously (pD > 0.85) with the support of at least 4 variants in derived state considering the pD values outside the path connecting the root of the Y chromosome tree and the respective haplogroup were observed in the range of 0-0.037. Further detailed sub-haplogroup assignments within the phylogeny of the primary haplogroup were determined on the basis of mapping the derived allele calls to the internal branches of a tree based on modern high coverage genomes and highlighting the marker tagging the branch with the lowest derived allele frequency (Data S1B and S1F).

Kinship analysis and identical samples

Preparing data for analysis

First, all newly generated samples were called with ANGSD-0.916[80] command–doHaploCall, sampling a random base for the positions that are present at MAF > 0.1 in the 1000 Genomes GBR population[62] giving a total of 4,446,224 SNPs for autosomal kinship analysis.

For the comparison with published studies, we used the 1240k panel SNPs only and select populations (Data S4) were retained from the combined dataset using plink–keep and converted to .tped. We chose two sets of contemporaneous samples (CA and BA) and only samples with > 0.1x coverage were included. The ANGSD output files were converted to .tped format, which was used as an input for kinship analyses with READ.[40]

Identifying genetically identical individuals

When using this approach the coverage, sample size and population diversity is important. Recommended minimum coverage combined between compared individuals is 0.1x[40] for accurate relationship estimation and too small a sample size will shift the estimated P0 upward leading to false negatives (not detecting 1st or 2nd degree relatives) and too much population diversity, i.e., analyzing completely different populations together (separated by too much time, distance etc.) will shift the P0 values lower, leading to false positives. Keeping this in mind, we first used the tool to identify the redundant sampling of identical individuals within sites. All samples over 0.1x were first run and any identical samples merged. Then one sample of lower than 0.1x was tested against the newly merged set starting with the highest coverage (e.g., 0.08x then 0.06x) and each time if the low coverage sample was estimated to be identical to any other sample, it was merged. If it was not, it was removed from the pool. Only if the estimated mtDNA haplotype and the READ analysis matched (showed identical) were they merged. This process was repeated until all samples with over 100,000 reads mapping to hg19 had been tested. We also checked the method using samples BRC027-29, which were powders taken from the same petrous bones as BRC001-3. The approach correctly identified the matches in these three cases. Using this approach we estimate the minimum/maximum number of individuals per site to be: Broion (12-15) and La Sassa (4-9).

Identifying kinship in new samples

Once identical individuals were identified and merged, to assess kinship relationships up to the 3rd degree, we divided the samples several ways according to geography and time. First all sites were analyzed together regardless of time period, then all sites

separated into Chalcolithic and Bronze Age based on C14 dates and genetic clustering on the PCA (EN versus PN clusters), then each site separately with Broion also run in two ways: 1) all samples together and 2) separated into Chalcolithic and Bronze Age sets. The mean P0 values are consistent across groupings (Data S4A) when assessing the results output, only pairs in which both the lower and upper Z score are greater than 4 are reported in Table 2. The full list of comparative genomes is listed in Data S4B, newly reported CA and BA results in Data S4C and S4D. Our main analysis is restricted to samples over 0.05x; however, we did run some tests on individuals over 0.0005x as part of the exploration of identical samples, these results are in Data S4E and S4F. We also compared our newly generated genomes (> 0.05x) to published CA and BA samples as listed in Data S4B, restricting analysis to published genomes over 0.1x. Results are listed in Data S4G–S4J, with no new confirmed close kinship relationships to report.

Imputed genomes (see section Genome imputation) were used to study in further detail cases of close degree of genetic relatedness detected with READ. We used the–genome function of PLINK 1.9.0[59] to estimate pairwise proportions of IBD1 and IBD2 that are informative, for example, for distinguishing parent-offspring from sibling relationships (Table 2).

Preparing the datasets for autosomal analysis

For the chronological sample assignment for the study of the arrival of Steppe-related ancestry component in Italy, comparative ancient samples from the Italian Peninsula, Sardinia, and Sicily broadly dated from the Neolithic to the Iron Age were added to the dataset along with the newly generated ancient samples from this study.[7,9,13,14,16,17] We used the following ranges for the time periods: Neolithic ((N), 7000 - 3500 BCE), Copper Age/Chalcolithic ((CA), 3500 - 2200 BCE), Early Bronze Age ((EBA), 2200 - 1700 BCE), Bronze Age ((BA), 1700 - 900 BCE), and Iron Age ((IA), 900 - 200 BCE) and grouped samples using the published relative and absolute dates. Because of the different chronological association of the transition from the Copper Age to the Bronze Age in Sardinia, ancient samples from the Nuragic culture were grouped with other ancient samples from Sardinia dated to the Bronze Age and ancient samples from the Punic culture were grouped with samples dated to the Iron Age (see also Data S1D).

20 individuals were sequenced in additional runs to an average genomic coverage between 0.13 and 1.24x. Sequences were realigned using the same process as previously described and identical individuals were merged using samtools 1.6[76] command *merge* and duplicates were removed using picard tools 1.54 (http://broadinstitute.github.io/picard/index.html).

In total, 22 individuals were selected for genomic analysis (Table 1; Data S1). Autosomal variants were called with ANGSD-0.916[80] command–doHaploCall calling all the positions in the Lazaridis et al., 2016 aDNA dataset.[46] For the analysis, a dataset of ancient and present-day individuals from David Reich Lab (https://reich.hms.harvard.edu/allen-ancient-dna-resource-aadr-downloadable-genotypes-present-day-and-ancient-dna-data, release: 1st of March 2020) and the dataset from Fernandes et al.[16] and Marcus et al.[17] were merged using PLINK v1.90[59] (Table S4). We performed all subsequent analysis on autosomal data. However, we excluded published ancient samples from Sardinia and Sicily which had been annotated as low coverage or contaminated samples and *group_label_sampleID*.

We created two different datasets maximizing SNP coverage in 1) *modern* (1240K +HO) and 2) *ancient* samples (1240K). The *modern* dataset was used to perform principal component analysis, DyStruct, and ADMIXTURE. The *ancient* dataset was used for outgroup *f*3 statistics, *f*4 statistics, Chromopainter/NNLS, *f*4 NNLS analysis and qpAdm.

Principal component analysis

Plink files were converted to EIGENSTRAT format using the program convertf from the EIGENSOFT 7.2.0 package with the parameter *familynames:NO*.[44,45] PCA was performed using the program smartpca with the parameter *autoshrink:YES*, projecting ancient individuals onto the components constructed based on the modern genotypes. The PCA was visualized using R-3.6.[87] The sampled individuals were projected on top of the present-day individuals in the David Reich dataset. A subset of ancient individuals from Eurasia were projected in groups (Figures 2A and S4; Data S1D).[4–7,9,10,12–15,23,24,46,47,64,65,114–121]

DyStruct

We modeled individuals as a mixture of different ancestral components by means of DyStruct, which takes into consideration temporal dynamics in the model based clustering method similar to admixture.[52] In detail, we performed five independent runs for K ancestral component K Î {2..10}, of 50 epochs each, using the same set of modern and ancient individuals used in PCA and ADMIXTURE analysis, and assigning them to time-interval of 1000 years, assuming generations of 25 years (Figures 2B and S2). To ease the visualization in Figure 2B, we focused only on Italian individuals together with a small subset of key Eurasian populations.

*f*4 statistics

For the *f*4 statistics, the same dataset as for the PCA was used and the target groups with 798 individuals from 191 ancient groups from Europe, Caucasus, Near East (Data S2). We performed *f*4 statistics in the form *f*4(Mbuti.DG, site1; site2, Y) for each studied site, whereby site 1 + 2 are the newly studied sites and Y is one of the comparative ancient groups/population (Data S2A). Additionally, we performed *f*4 statistics in form *f*4(Mbuti.DG, sample1; sample2, Y) as well as *f*4(Mbuti.DG, sample1; Y1, Y2) for each studied sample (Data S2C–S2F). We used ancient individuals from different context: Western Hunter-Gatherers, Eastern Hunter-Gatherers, Yamnaya Steppe/ Pontic Steppe, European Neolithic farmers, Anatolia Neolithic, and ancient Italian individuals from previously studies.[7,9,13,14,16,17] We used the program qpDstat with the option *f4mode:YES* from the software ADMIXTOOLS 4.1.[122]

Additionally, we performed *f*4 statistics in form *f*4(Mbuti.DG, Russia_EBA_Yamnaya_Kalmykia.SG/Samara, X, Anatolia_N) to test the affinity of all ancient Italian individuals spanning from the Neolithic to the Iron Age using qpDstats (Data S2B). The results of the

Please cite this article in press as: Saupe et al., Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.04.022

*f*4 statistics in form *f*4(Mbuti.DG, Russia_EBA_Yamnaya_Kalmykia.SG, X, Anatolia_N) were visualized using R 3.6[87] (See also Figures 2C and S4B).

Outgroup f3 statistics

For the study of X chromosome versus autosomes, outgroup *f*3 statistics in the form *f*3(Italian_CA/Italian_EBA_BA,X; Mbuti.DG) were performed with Mbuti.DG as the outgroup and the same subset of the ancient population previously described was chosen. For the analysis, the autosome and X chromosome positions available in Lazaridis et al.[46] ancient dataset were selected and the data was converted to EIGENSTRAT format using the program convertf from EIGENSOFT 7.2.0[44] with the parameter *familynames:NO*. To understand the sex-specific patterns detected in European Bronze Age populations, the newly generated samples were grouped according to the clusters seen in the PCA (Figure 2A) - Italy_CA (n = 12) and Italy_EBA_BA (n = 8). Published ancient individuals from Italian Peninsula were added to the groups following Italy_CA: Italy_C.SG (R1014.SG, R4.SG, R5.SG = 3),[13] Italy_North_Remedello_C.SG (RISE487.SG, RISE4879.SG = 2)[7] and Italy_EBA_BA: Italy_North_Remedello_EBA.SG (RISE486.SG = 1), Italy_North_BellBeaker (I1979, I2477, I2478 = 3).[9] Outgroup *f*3 statistics were computed using ADMIXTOOLS 1.1[122] option qp3Pop (See also Data S1K, Figure S5).

Additionally, we performed outgroup *f*3 statistics in form *f*3(Italy_CA, X; Mbuti.DG) to explore the genetic relationships of Italy_CA and peri-Neolithic groups (X) (Data S1K). The group Italy_CA contains following samples from the Italian Peninsula dated to the Chalcolithic: Italy_Broion_CA, Italy_Gattolino_CA, Italy_LaSassa_CA, Italy_C.SG, Italy_North_MN_Iceman_contam.S, and Italy_North_Remedello_C.SG.

Admixture analysis

We exploited the model-based algorithm implemented in ADMIXTURE[53] projecting ancient individuals (-P flag) into the genetic structure calculated on the modern dataset, due to missing data in the ancient samples. In detail, we performed unsupervised Admixture for K Î {2..10} for modern samples, and used the "per-cluster" inferred allele frequencies to project the ancient samples. We visualized the Q output using R 3.6[87] (Figure S3).

Chromopainter/NNLS and SourceFind

We reconstructed the ancestry of each individual using the Chromopainter (CP)/NNLS framework[55] and SourceFind (Figures 3 and S6). First, in order to obtain information from the highest number of markers, we painted all the Italian and a selection of European ancient individuals, using the unlinked mode, together with Loschbour_published.DG (Western European Hunter-Gatherer), I0061 (Eastern European Hunter-Gatherer), I0707 (Anatolia Neolithic), KK1.SG (Caucasus Hunter-Gatherer), HGDP00982 (Mbuti), I0443 (Yamnaya herders), against a set of 1,260 modern individuals (donors). We used 0.0002 and 318 as M and n parameters.[123] The resulting copying vectors, summarizing the number of markers inherited from each modern individual, were then pooled according to donors' affiliation and normalized to sum 1. Finally, we reconstructed each target individual copying vector as a mixture of different proportions of the putative surrogate sources, taking advantage of a slightly modified version of the nnls function in the "nnls" package in R software, and implemented in GlobeTrotter. In addition, in order to reduce the noise generated by the relatively low number of markers, the same analysis has been repeated using the putative source average copying vectors through all the ChromoPainter runs (Figure S6).

Additionally, the same analysis has been performed using SourceFind rather than the NNLS.[56,57] In detail, for each sample, we performed 10 runs of 5 million iterations thinned by 10,000 and discarding the first 50,000 iterations. The expected number of sources was set to 3 with a maximum of 4 possible sources. Only the iteration characterized by the highest Likelihood was shown in a barplot (Figure S6).

f4 NNLS analysis

In order to provide an additional description of ancestral composition of the analyzed samples we carried out a NNLS analysis using different *f*4 statistics vectors as a proxy of relations among different ancient groups. In detail, for any given target population, an *f*4 analysis in the form *f*4(X,Y, Target, Mbuti.DG) has been performed, where X, Y and target belong to one of 100 ancient groups from previously and newly genotyped data. In doing so, we obtain for each target a vector of 4,950 D Statistics; and we used the vectors for two different sets of putative sources to infer the ancestral proportions of a selection of European and Italian ancient samples, using the same nnls function of the Chromopainter/NNLS analysis (See also Figure S6). We used the following sets of sources:

Set 1: Luxembourg_Loschbour_published.DG, Anatolia_N, Iran_Tepe_Abdul_Hosein_N.SG, Russia_EHG
Set 2: Luxembourg_Loschbour_published.DG, Anatolia_N, Iran_Tepe_Abdul_Hosein_N.SG, Russia_Yamnaya_Kalmykia.SG

qpAdm

In order to describe the ancestry of Italian target individuals as a combination of prehistoric groups known to have played a major role in European demography, we harnessed the qpWave/qpAdm framework (Data S3). In details, for each sample, given a set of left and right populations, and for the number of left (sources) K Î {2..4} we:

a) Evaluated if the right samples can be used to significantly discriminate the provided sources, harnessing qpWave. Right populations present in sources were excluded for that specific test.
b) If step a) provided a significant p value < 0.01, we used qpWave to evaluate if the target sources could be described as a combination of K sources.
c) If b) provided a p value > 0.01, we used qpAdm to describe the target as a mixture of the K sources employed for that specific experiment. Given the high rate of missingness in some of our tested samples we used the option "allsnps=YES"

We used three different sets of left groups which we called "pre-Bronze Age," "post-Bronze Age," and "Continuity Neolithic"
pre-Bronze Age left panel:
Anatolia_N, Luxembourg_Loschbour_published.DG,Russia_EHG, Georgia_Kotias.SG
post-Bronze Age left panel:
Anatolia_N, Russia_Yamnaya_Samara, Luxembourg_Loschbour_published.DG
Continuity Neolithic left panel:
Russia_EBA_Yamnaya_Samara, Italy_Mesolithic.SG, Italy_C.SG, LaSassa, Broion_CA, Iberia_C, Germany_BellBeaker, France MN
In all cases we used the following right populations:
Russia_AfontovaGora3, Russia_EHG, Iberia_ElMiron, Belgium_GoyetQ116_1_published, Russia_Kostenki14, Jordan_PPNB, Russia_MA1_HG.SG, Israel_Raqefet_M_Natufian, Ust_Ishim.DG, Czech_Vestonice16, Georgia_Kotias.SG

Genome imputation

Following Hui et al.,[60] genotype likelihoods were first updated with BEAGLE 4.1[86] from genotype likelihoods produced by ATLAS[88] in Beagle -gl mode, followed by imputation in Beagle -gt mode with BEAGLE 5[124] from sites where the genotype probability (GP) of the most likely genotype reaches 0.99. To balance between imputation time and accuracy, we used 503 Europeans genomes in 1000 Genomes Project Phase 3[62] as the reference panel in Beagle -gl step, and 27,165 genomes (except for chromosome 1, where the sample size is reduced to 22,691 due to a processing issue in the release) from the Haplotype Reference Consortium (HRC)[72] in the Beagle -gt step. A second GP filter (MAX(GP) > = 0.99) was applied after imputation. Both new and published genomes were processed individually in the -gl step; in the -gt step, they were jointly imputed in the following groups:

All new genomes sequenced in this study; Mesolithic genomes (n = 3), Neolithic and Chalcolithic genomes (n = 13), Iron Age and later genomes (n = 11) from Antonio et al.;[13] Bronze Age Steppe: RISE509, RISE511, RISE547, RISE548, RISE550, RISE552 from Allentoft et al.;[7] EBA1 and EBA2 from Damgaard et al.;[65] Anatolia Neolithic: Bar31, Bar8, Klei10, Pal7, Rev5 from Hofmanová et al.;[5] WHG: KO1 from Gamba et al.;[125] Bichon from Jones et al.;[121] La Braña from Olalde et al.;[27] EHG: Sidelkino from Damgaard et al.;[65] CHG: KK1 and SATP from Jones et al.[121]

Runs of Homozygosity

We used hapROH[61] to detect runs of homozygosity (ROH) in ancient genomes. Using information from a reference panel, hapROH has been shown to work for genomes with more than 400K of the 1240K SNPs panel covered at an error rate lower than 3% in pseudo-haploid genotypes.[61] We note that the requirement is broadly in line with the imputation accuracy we get from coverages as low as 0.05x, where ~60% of common variants (MAF greater than or equal to 0.05) in the HRC panel are recovered with an accuracy greater than 0.95 in diploid genotypes.[60] Among common variants in the HRC panel, 853,159 overlap with the 1240K SNPs panel.

Nevertheless, imputation errors are not evenly distributed across the genome; the accuracy varies with the genotypes, minor allele frequency and local recombination rate, which may bias results in downstream analysis. We introduced random genotype errors into modern high-coverage genomes following the pattern observed in imputed ancient genomes. We then compare the total amount of ROH detected in the original genomes and the genomes with simulated imputation errors to explore whether hapROH can be used on imputed genomes.

Predicting imputation errors

A logistic regression model was developed to predict imputation errors from the features of the variant. Imputation accuracy has been shown to vary according to the true underlying genotypes (heterozygous sites have lower accuracy than homozygous sites) and minor allele frequencies (rarer variants have lower accuracy).[24,60,125] Since recombination breaks down linkage disequilibrium, the local recombination rate will also affect imputation accuracy via the number of linked sites. Finally we also included the substitution type (transition versus transversion) among the explanatory variables, because *post-mortem* damage in aDNA might be erroneously interpreted as C to T (or G to A) mutations. Considering that allele frequencies influence both imputation accuracy and ROH detection, a log-transformed MAF term was added in addition to the untransformed MAF to capture its effect more accurately. The recombination rate was also log-transformed. The header of Data S5 lists the explanatory variables and the transformation applied to each of them.

We down-sampled the 20x Neolithic Hungarian genome NE1 to 0.05x and ran it through the same imputation pipeline. The imputed genotypes were then compared to the confidently called genotypes in the original 20x genome to calculate the error rate. Because NE1 is older than all the genomes sequenced in this study, we expect to establish a lower bound of the performance of imputation from its result.

Please cite this article in press as: Saupe et al., Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.04.022

We fitted three separate multinomial logistic regression models to predict the imputed genotypes when the true genotype is homozygous for the reference allele (0/0), homozygous for the alt allele (1/1), or heterozygous (0/1). The outcome is a categorical variable with four values: missing genotypes (not passing the max(GP) greater than or equal to 0.99 post-imputation filter), 0/0, 0/1, and 1/1. Data S5A shows the coefficients and intercept for each model.

We then used the model to simulate the effect of imputation by introducing random genotype errors into a VCF file with high-quality genotypes. For each variant, we applied the softmax function on the scores for the four outcome categories (./., 0/0, 0/1, 1/1) to obtain the probability distribution before randomly drawing an ''imputed'' genotype. This allows us to quickly generate large numbers of ''pseudo-imputed'' genomes where the true genotypes are known for assessing the effect on downstream analysis without the lengthy process of down-sampling and imputation. Data S5B shows that the error pattern after imputation is well approximated by this approach when applied onto the original 20x NE1 genome.

Introducing imputation errors into modern genomes

We used the same model to introduce random error into 107 modern Italian (TSI) genomes from the 1000 Genomes Project Phase 3.[62] We extracted them from the HRC panel so that the list of variants is identical to what we get after imputation with HRC as the reference panel. We then used hapROH to detect ROH segments in both the original genomes and the genomes after introducing imputation errors. Except for switching from haploid to diploid mode, we kept the default parameters unchanged. We also used the default reference panel, 1000 Genomes Project data at 1240K sites, after masking out the TSI population. The relevant parameters to the hapsb_ind function are:

h5_path1000 g = 'all1240/chr', e_model = 'diploid_gt', p_model = 'EigenstratUnpacked', post_model = 'Standard', delete = False, n_ref = 2504, exclude_pops = ['TSI'], readcounts = True, random_allele = False, roh_in = 1, roh_out = 20, roh_jump = 300, e_rate = 0.01, e_rate_ref = 0.0, cutoff_post = 0.999, max_gap = 0, roh_min_l = 0.01

Figure S7A compares the total length (top) and number (bottom) of ROH tracks detected in the original genomes and the genomes with simulated imputation errors. In general the correlation is high for ROH segments longer than 1.6cM, although the total length is under-estimated in the presence of simulation errors when we filter for ROH segments longer than 4cM. This is most likely caused by longer ROH segments being broken down into shorter ones due to erroneous genotypes. Nevertheless outlier individuals with long ROH segments still stand out. The total number of ROH segments is reproduced less accurately in the simulated genomes than the total length of ROH segments.

In addition we tested a hidden Markov model-based ROH detection algorithm implemented in BCFtools[126] on the original and simulated genomes. Because this hidden Markov model-based algorithm takes allele frequencies in the population into account, we randomly divided the 107 individuals into 9 groups in order to 1) better match the sample size in this study and 2) reduce the running time of the analysis. Example command:

bcftools roh -G 30 -I -m genetic_map_chr{CHROM}_combined_b37.txt < input.vcf.gz > -o < output > -O r

Only ROH segments with a quality above 20 on the phred scale were retained. The result is similar to that from hapROH although the correlation is slightly lower (Figure S7B), supporting that the total length of ROH segments longer than 1.6cM can be reliably recovered in imputed genomes.

We also observed a strong correlation between the total ROH lengths detected by hapROH and BCFtools in the original genomes, which drops from ~0.97 to ~0.8 after imputation errors are introduced (Figure S7C). We chose to use hapROH on the ancient genomes out of two considerations. First, hapROH appears more robust in the presence of imputation errors (Figures S7A and S7B). Moreover, unlike BCFtools which processes a group of individuals together and utilizes allele frequencies in the population, hapROH examines one individual at a time. In this way it is less likely for our ROH result to be biased by the differences in sequencing coverage.

Detecting ROH segments in ancient genomes

We ran hapROH with the same setting as above on the imputed Chalcolithic and Bronze Age genomes generated in this study and the published Italian Mesolithic, Neolithic and Chalcolithic genomes (Data S1; Figure 4). For the 107 TSI genomes, we used results from the original genomes without the simulated errors.

Phenotype prediction

For the 41 HIrisPlex-S set of SNPs we selected 2 Mb around the informative variants, merging the regions on the same chromosome, with the exception of the variants on chromosome 15, which have been analyzed in two different regions since the distance between the two nearest SNPs was about 20 Mb. We finally selected 10 regions from 9 autosomes, spanning from about 1.5 Mb to 6 Mb. For the other phenotypic informative markers (diet, immunity and diseases), we selected 2 Mb around each variant and merged the overlapping region, for a total of 48 regions from 17 autosomes and X chromosome.

We called the variants using ATLAS v0.9.0[88] task = call and method = MLE commands at positions with a minimum allele frequency (MAF) greater than or equal to 0.1% in the reference panel, that has been selected according to the different components of the samples: 1) Europeans from 1000 Genomes (EUR),[62] for our Chalcolithic and Bronze Age Italians, for the pre-Imperial Romans from Antonio et al.[13] for pre-Nuragic and Nuragic Sardinians from Fernandes et al.[16] and Marcus et al.[17] and for Yamnaya;[4,7,64] 2) EUR plus the MANOLIS (EUR-MNL) set from Greece and Crete extracted from the Haplotype Reference Consortium (HRC)[72] (accessed at: http://www.haplotype-reference-consortium.org/) for the ancient Near Easterns, for non-Sardinian western Mediterraneans from

Please cite this article in press as: Saupe et al., Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.04.022

Fernandes et al.,[16] for post-Nuragic Sardinians from Marcus et al.[17] and for Imperial and Later Romans from Antonio et al.[13] (Table S4). After calling the variants separately for each sample, we merged them in one VCF file per region. We used the merged VCFs as input for the first step of our imputation pipeline[60] (genotype likelihood update), performed with Beagle 4.1 -gl command[86] using the same panels as before as reference. We then discarded the variants with a genotype probability (GP) less than 0.99 and imputed the missing genotype with the -gt command of Beagle 5.0[124] using the HRC as a reference panel for all groups of samples. We then discarded the variants with a GP < 0.99 and used the remaining SNPs to perform the phenotype prediction. Two markers of the HIrisPlex-S set, namely the rs312262906 indel and the rare (MAF = 0 in the HRC) rs201326893 SNP, were not analyzed because of the difficulties in the imputation of such variants. Results are reported in Data S6A–S6D.

We performed this analysis on our Chalcolithic and Bronze Age individuals and on published ancient samples from the Near East,[7,46,64,65] Italy[13,16,17] and Yamnaya population,[4,7,64] with a coverage greater than or equal to 0.05x. First, we grouped the individuals in six groups based on their location and time, namely two groups from Near East (Neolithic/Chalcolithic and Bronze Age), three from Italy (Neolithic/Chalcolithic, Bronze Age, post-Bronze Age) and one Yamnaya (Data S6A). We compared the groups performing an ANOVA test and, for the significant variants, we performed a Tukey test to identify the significantly different pairs of groups (Data S6B). We then analyzed the difference within Italy by creating 13 local groups (excluding 7 Sardinian individuals with an uncertain dating) (Data S6A) and comparing the 12 groups larger than 3 with an ANOVA test (Data S6C). For both comparisons, we used a Bonferroni's correction on an alpha value of 0.01 for the number of tested SNPs to set the significance threshold. We performed our phenotype analysis in a set of 332 ancient individuals, composed of the 16 Italian Bronze and Chalcolithic individuals reported here for the first time and 316 ancient Italians, Near Easterns and Yamnaya from the literature (Data S6A). Sample-by-sample phenotype prediction and genotype at the selected phenotype informative SNPs, reported as number of effective alleles (0, 1 or 2) are shown in Data S6D.