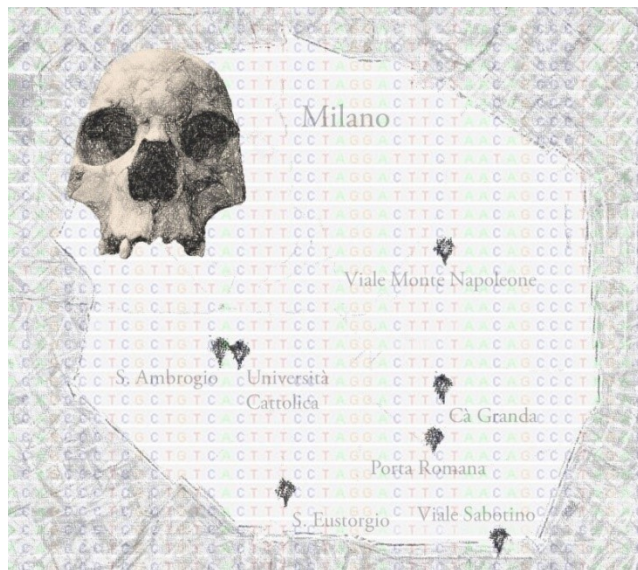*UNIVERSITÀ DEGLI STUDI DI PAVIA*
Dipartimento di Biologia e Biotecnologie
"Lazzaro Spallanzani"

Genetic and anthropological investigations on ancient skeletal remains recovered in Milan, from the Roman period to the Contemporary Age

Barbara Bertoglio

Dottorato di Ricerca in
Genetica, Biologia Molecolare e Cellulare
XXX Ciclo – A.A. 2014-2017

UNIVERSITÀ DEGLI STUDI DI PAVIA

Dipartimento di Biologia e Biotecnologie
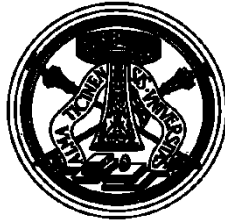
"Lazzaro Spallanzani"

Genetic and anthropological investigations on ancient skeletal remains recovered in Milan, from the Roman period to the Contemporary Age

Barbara Bertoglio

Supervised by Prof. Antonio Torroni
Prof. Carlo Previderè
Prof. Cristina Cattaneo

Dottorato di Ricerca in
Genetica, Biologia Molecolare e Cellulare
XXX Ciclo – A.A. 2014-2017

ABSTRACT

The Milan project arises to describe the life and evolution of the populations living in the urban area of Milan across two millennia (from the Roman Imperial Age to Contemporary times) through a multidisciplinary approach involving archaeology, anthropology and genetics.

Archaeological and anthropological investigations revealed over time the great potential of this skeletal collection, supplying an interesting portrait of the demographic conditions and the biological evolution of people living in different historical periods. Considering the intriguing results and the improvement of technologies to study ancient DNA, the Milan project would enrichen the profile of these populations in terms of genetic evolution, origin and variations, identifying possible parental relationships between individuals and genetic relations with people living in border areas.

This research project represented an articulated pilot study carried out to investigate this ancient material combining anthropology and genetics. In particular, this was the first attempt performed to assess the feasibility of reconstructing the history of Milan through DNA analysis.

Genetic analyses were carried out to enrichen anthropological and archaeological data, giving information about the genetic sex and the geographic origin. Because no genetic studies were performed previously on the skeletal collection of the Milan population, these analyses were also performed to evaluate the possibility to recover DNA from samples coming from different archaeological contexts. To this purpose, ten intriguing skeletal remains excavated from seven different sites and belonging to different historical periods (from the Roman Imperial Age to the Contemporary Age) were selected and submitted to osteological analyses. From each individual a bone sample was collected for the following genetic studies. In particular, DNA extraction, genetic sex typing and mitochondrial DNA analysis were performed applying protocols and procedures commonly used in cases of highly degraded samples, such as ancient remains. For this reason, mitochondrial DNA studies were carried out applying two different protocols in order to recover the complete mitogenome (mitoTiling protocol) and enrichen samples in sequences of interest through mtDNA capture (PEC protocol). Finally, to evaluate how the different archaeological contexts could influence bone tissue appearance and to verify the presence of a correlation with DNA content, macroscopic and microscopic analyses were performed from the same bone specimens submitted to genetic analyses.

The results obtained suggested the great potential of the approach applied, providing intriguing results concerning the biological profiles of each individual. In particular, anthropology provided interesting data, specifically about stress markers and pathology (infectious, degenerative, metabolic, congenital diseases and tumour), which highlighted the difficult living conditions and the situation of

extreme poverty of people during these historical periods. On the other hand, genetics was able to enrichen biological profiles supplying interesting information about the geographic origin, also for some of the cases for which the state of preservation of the bone samples did not allow the anthropological estimation. For this reason, the methods used to evaluate genetic sex and mitochondrial DNA were proved to be efficient in the recovery of these information from such ancient skeletal remains. In particular, mito Tiling approach was shown to be a suitable protocol for the recovery of consistent genetic information from well-preserved samples; on the other hand, PEC protocol was proven to be very efficient in the enrichment of those samples which gave no results with the first approach or in the increase of the coverage of the other samples. To this purpose, the PEC approach is suggested to be a useful method to recover successfully genetic information from samples characterised by different taphonomical conditions.

Moreover, genetic analyses pointed out degraded DNA in the bone samples characterised by very short fragments length, low amounts of DNA and peculiar damage patterns. Both time since deposition and environmental conditions influenced DNA preservation in the bone samples. Similar results were obtained only from the microscopic analysis of calcified and decalcified bone thin sections, with an increase of bone degradation over time and differences ascribable to environmental factors. In spite of this similarity, no relation between bone structure appearance and genetic data was detected, even if the small sample size limits the final conclusion about a possible correlation. Moreover, differences identified between macroscopic and microscopic preservation highlighted the importance to include the histological analysis in the evaluation of bone tissue appearance. In particular, decalcified thin sections are suggested to be a valid tool to examine the diagenetic changes at the microscopic level, also in archaeological samples, where microbial activity can hide bone structures in calcified sections, thus preventing a correct analysis.

Finally, these satisfactory preliminary results supported the possibility to carry on further analyses on this skeletal population. For this reason, the most intriguing samples will be further analysed in the Max Planck Institute for the Science of Human History in Jena (Germany), applying specific protocols and techniques used in the ancient DNA field, in order to investigate the DNA content, both human and microbial, and to assess the feasibility to analyse other interesting genetic markers, such as nuclear DNA. In addition, the new data could confirm the paleopathological diagnosis and can get interesting information about ancient microorganisms genomes. As a future perspective, an increase in the number of individuals collected from each site could be planned to be submitted to the set of specific genetic procedures thus providing a portrait of the evolution of the Milan population over time. Moreover, the analysis of further samples coming from the different archaeological contexts could supply interesting information concerning the preservation of bone tissue in the different conditions and verify its relation with the DNA content.

CONTENTS

1. INTRODUCTION

1.1 THE MILAN PROJECT

The Milan project (coordinators: Dott.ssa AnnaMaria Fedeli, Soprintendenza Archeologica della Lombardia; Prof. Cristina Cattaneo, LABANOF, Laboratorio di Antropologia ed Odontologia Forense, Dipartimento di Scienze Biomediche per la Salute, Università degli Studi di Milano) arises to describe the life and evolution of the populations living in the urban area of Milan across two millennia (from the Roman Imperial Age to Contemporary times) through a multidisciplinary approach involving archaeology, anthropology and genetics.

1.1.1 The scenario

During the last 30 years, several archaeological excavations were performed in the Milan area.

Among these, great attention was given to nine sites located in the city center and in its outskirts, especially for the archaeological context and the peculiarity of the ancient remains discovered.

Four of them are located in proximity or within important academic or religious landmarks, such as Università Cattolica del Sacro Cuore and Università degli Studi di Milano and two cathedrals, S. Eustorgio and S. Ambrogio. Three are placed in Via Monte Napoleone, Corso Porta Romana and Viale Sabotino, which represent interesting area, especially when contextualized into the past (Figure 1); the last two are located in Piazza del Kuerc, Bormio, and in Bolgare – St. Chierico, two regions of historical interest placed outside Milan.

During the excavation procedures, the detection of several burials suggested the cemetery use of each area in different stages of the settlement.

The analysis of the ground layers and radiocarbon dating of bone samples showed a use of each site in different historical periods, distributed across two millennia, from the I century AD to XVIII century AD (Table 1 and Figure 2).

From these sites over 1,000 skeletons covering the time span were recovered, representing a cultural asset, interesting both archaeology and anthropology to study the life and evolution of ancient populations in Milan.

A more ancient skeleton, which was included in the project, is represented by the skeletal remains found in the outskirts of Milan, in Parabiago, in a tomb which has been dated to the Roman Imperial Age (specifically from the second half of the I century AD to the II century AD). The ceramic plate buried with the corpse bore the name ATILIA.

Finally, to the archaeological remains, the modern skeletal collection coming from Cimitero Maggiore was added. It involves remains belonging to people who died in the XX century, representing an important anthropological source useful to trace and record the biological changes of our population during the last years.

Figure 1. The archaeological sites excavated in Milan in the last years (modified from www.google.it/maps).

Historical period	Age	Site	Burial practice
Roman Imperial Age	I-II AD	Parabiago	Inhumations, cremations
Roman Imperial Age	I–II AD	Porta Romana	Inhumations, cremations
Roman Imperial Age – Late Antiquity	I–V AD	Catholic University	Inhumations, cremations
Late Antiquity	III – VI AD	S. Eustorgio	Inhumations
Late Antiquity-Medieval	V – XV AD	S. Ambrogio	Inhumations
Early Medieval	VII–IX AD	Bolgare	Inhumations
Late Medieval	XIV AD	Bormio, Piazza del Kuerc	Inhumations
Late Medieval	XV - XVI AD	Via Monte Napoleone	Mass grave
Modern Age	XVII AD	Viale Sabotino	Mass grave
Modern Age	XVII-XVIII AD	Cà Granda	Ossuary
Contemporary Age	XX AD	Cimitero Maggiore	Inhumations

Table 1. Historical periods and information about the sites investigated in the Milan area.
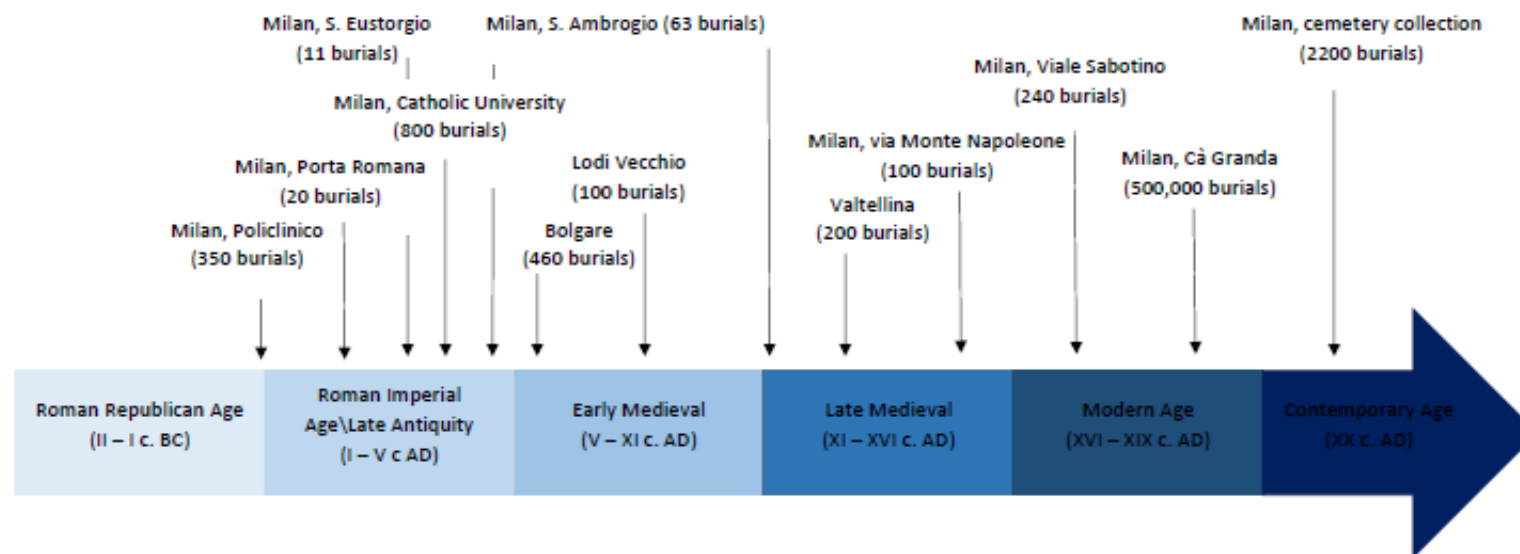
Figure 2. The Milan skeletal collection: the arrow shows the different historical periods from the Roman Republican Age to the Contemporary Age; for each period the skeletal material available is reported.

1.1.2 The archaeological and anthropological studies

The archaeological analyses which gave birth to the myriad of skeletons of each different period documented, layer after layer, the origin and development of these sites and the human activities that characterized each period. Particular attention was given to the discovered *necropolis*.

The analysis of the ground layers, together with radiocarbon dating, could localize in terms of time and space each area. In particular, depending on the historical period, the position inside or outside the city walls could be determined, supplying important data to evaluate the kind of cemetery destination (Figure 3). In addition, archaeological examinations of the orientation and position of the graves, burial practises and position of the remains, as well as possible grave goods, gave further information about the ritual, the relevance and the status of the dead.

The most representative type of burial is inhumation, generally of a single person. Only in some cases, cremations and multiple burials were detected, the last probably caused by the reuse of the graves in following periods. The only exception was represented by the mass graves discovered in Viale Sabotino and in Via Monte Napoleone and the ossuary of Cà Granda.

At the same time, the skeletal populations recovered during the archaeological excavations were studied by anthropologists.

Anthropology applied the knowledge of human osteology to interpret human bones, providing information about "the life, the death and the postlife history of a specific individual" (Dirkmaat *et al*, 2008). The analysis of bones and teeth could supply data to build the biological profile of each subject; this included ancestry, sex, age at death and stature, but also all those signs due to occupational stress, pathology and trauma. In fact, many situations of physical stress leave a mark on bone tissue during life, causing a variation of the normal appearance of the skeleton. These variations are represented by abnormal bone formation or destruction, abnormal bone density, abnormal bone size and shape, in cases of pathology or *antemortem* trauma (Ortner, 2003), and by morphological variations or pathology related to work activities in specific anatomical regions, in cases of occupational stress (Capasso *et al*, 1999).

The aim of the anthropological analyses was to "bring back to life" each individual not only in their physical appearance, but also in their activities, private and social daily life, diet habits, pathology and the *perimortem* period (Cattaneo *et al*, 2002).

All these data revealed the profile and evolution of the population in time according to sex, age, mortality, life expectancy, ancestry, occupation and health, supplying an interesting portrait of the demographic condition of people living in different historical periods.

In particular, the first anthropological studies carried out on the skeletal remains showed the great potential of this skeletal collection, especially from the pathological point of view.

A survey through time gave the evidence of the first cases of syphilis in the III-V centuries AD (Cattaneo *et al*, 1996), the appearance of tuberculosis only in the dark ages, and a distinct worsening of the nutritional conditions in early medieval times (Cattaneo and Gibelli, 2015; Cattaneo *et al*, 2015b; Sguazza *et al*, 2015; Caruso *et al*, 2017b). The interest increased with the discovery of a mass grave revealing the burial of individuals who died during the Manzoni's plague (1600) and the Ossuary of the Cà Granda (Ospedale Maggiore, the largest medieval hospital in Europe), where 500,000 bodies are estimated to have been buried (Caruso *et al*, 2013a; Caruso *et al*, 2013b; Cattaneo and Gibelli, 2015; Sguazza *et al*, 2016a; Sguazza *et al*, 2016b).

To date a considerable number of skeletal remains have been analysed by anthropologists with a clear definition of biological profiles and anthropological characteristics of the population.

1.1.3 The genetic approach

The Milan project would enrichen the history of the skeletal populations through DNA analysis, adding genetic data to those acquired by archaeologists and anthropologists. The aim was to assess the feasibility of reconstructing through DNA the history of the city in order to retrace the profile of populations coming from different historical periods also in terms of genetic evolution, origin and variations, identifying possible parental relationships between individuals and genetic relations with people living in the border areas.

To this purpose the study of sequences and regions of interest, such as mitochondrial DNA, Y-chromosome and autosomal SNPs, are considered to obtain information about evolution, migration, diet and pathology. In addition, the analysis of DNA phenotyping markers could be added, supplying interesting data about the appearance of individuals, such as hair and eye colour and some traits of the face, to extend the individual genetic profile.

In the following paragraphs the archaeological sites were described, highlighting the peculiarities and evolution in time of each skeletal population.
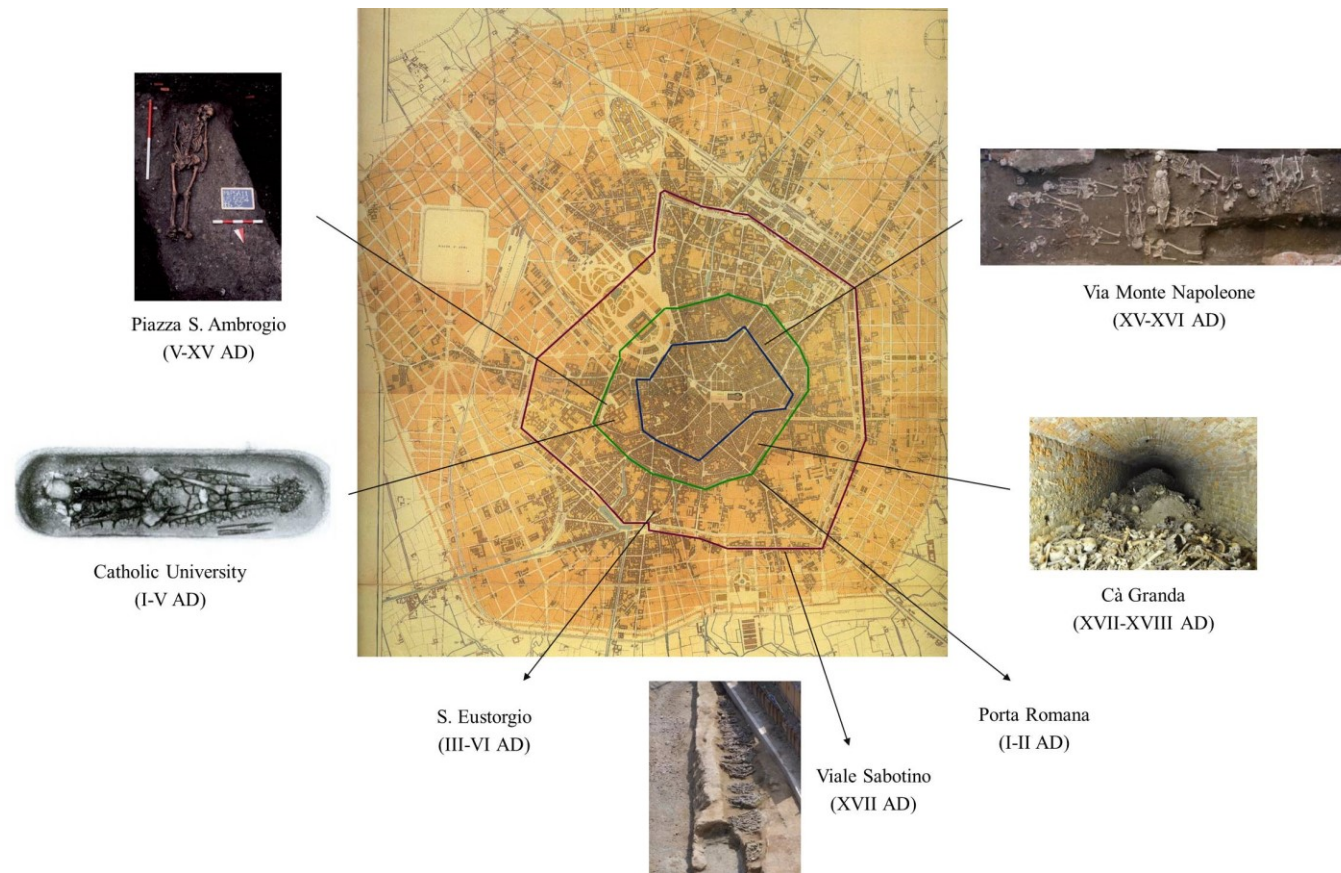
Figure 3. Localization of the archaeological sites in the ancient Milan (blue: the Roman walls; green: the Medieval walls; purple: the Spanish walls (modified from Piano Beruto, 1889; www.wikipedia.it).

1.1.4 The archaeological sites

Parabiago

In the 1999 during building construction carried out in the proximity of Piazza Marconi, twenty burials dated to the second half of the I century AD and to the II century AD were discovered in an area closed to a region where further nine graves, dated to a later period (III-IV centuries AD), were found in the 1933. Both incinerations and cremations, as well as inhumations, were detected. Among these a peculiar inhumation was excavated in the grave 6 (Figure 4): a skeleton with grave goods located near the skull and the feet and a ceramic plate with the name of the deceased, *Atilia*. Moreover, close to the body nails were found suggesting the burial inside a wood coffin (Excavation Report, Soprintendenza Archeologica della Lombardia, 1999). The skeletal remains recovered belonged to a young woman (17 years old), with typical mongolid features. It has been suggested that she represented perhaps a female servant coming from the Orient.

Figure 4. Grave 6: the appearance of the grave before the excavation procedures (on the left) and the skeletal remains discovered (on the right) (Excavation Report, Soprintendenza Archeologica della Lombardia, 1999).

Corso Porta Romana (Milan)

Corso Porta Romana corresponds to an ancient road path connecting the city with the capital. During building construction in the 1999, the evidence of a stratigraphic sequence from the XVIII century AD to the Golasecca period was detected.

Several stages of usage were identified. In the first period, the archaeological evidences suggested an agricultural use of the area, as well as a structural function, followed by a cemetery destination of the site, during the first Imperial Age. Most of the burials were interested by removal action during the past and recent ages, making difficult their recovery and identification (Excavation Report, Soprintendenza Archeologica della Lombardia, 1999).

Both inhumations and incinerations were identified. In particular, twenty individuals were recovered, eleven were buried and nine cremated (Chiti, 2002).

The anthropological analyses revealed a uniform population, characterized by males and females, children and adults, without ethnic diversity (all the individuals showed European traits). Interesting differences were observed between buried and cremated subjects.

In fact, occupational stress markers, especially in the spinal column, were detected in the first group, affecting both females and males, young and adult people, suggesting an involvement in agricultural work. In addition, the evidence of dental wear and caries, indicated a poor and incomplete diet, as revealed also by the detection in some individuals of anaemia, due to iron deficiency, and hypoplasia, caused by nutritional or pathological stress which occurred during the growth.

On the other hand, no pathology was identified in the cremated population, suggesting a different lifestyle, most likely associated to the belonging to another social status, and the possible use of distinct burial practises for persons coming from different classes (Chiti, 2002).

Catholic University (Milan)

Belonging to a longer period, from the Roman Imperial Age to Late Antiquity, the Catholic University represents one of the most precious sites in Milan, providing a sight of the history in the South-Western suburbs of the city. Located close to the city walls, it was characterized by farming activities and buildings in the first two centuries. During this period, some inhumations and cremations involved the area, but only from the III century AD it acquired only the burial destination. In 3,500 mq, about 800 burials, belonging to the time lapse between the I and V centuries, were brought to light, making them one of the most ancient and numerous populations of the North Italy. The long period can be divided in three attending stages: a first stage (Period A), a restoration stage (Period B) and a state of neglect (period C) (Cortese, 2011; Sannazaro, 2011).

The disposition of the graves was messy and confused; different orientations and types of burial were detected, with intersections and overlapping of the graves, especially in areas characterized by a greater density of use.

Both cremations and inhumations were identified (the inhumations were prevalent), but no differences were observed between buried and cremated individuals, suggesting no factors correlated with the type of burial used (Barlassina, 2011; Sannazaro, 2011).

The bone remains revealed a social change in the population in the V century AD; in fact, ethnicity diagnosis identified the presence of individuals coming from the sub-Saharan Africa, as well as subjects of European origin. In addition, a variety in the diet was supposed, as highlighted by a concentration of zinc and magnesium in bones, indicative of a diet rich of zinc-containing foods such as cheese, dried fruit and legumes (Cattaneo *et* al, 2011; Cattaneo and Gibelli, 2015).

Anthropological results defined an high childhood mortality, especially caused by infectious diseases, and a life expectancy of 28-30 years.

Several pathology were detected, among which infectious disease are poorly represented, with the exception of syphilis, detected in both adults and children. A particular characteristic was a dental "V" shape wear observed in some females, probably associated to spinning practices (Cattaneo *et* al, 2011; Cattaneo and Gibelli, 2015).

A peculiarity of the site was the identification of a burial of a young woman with an age range between 24-31 years old, who died probably from multiple myeloma. Her peculiarities were the personal decorations and the abundance of grave goods. She was buried with her clothing in a sarcophagus in *serizzo* (stone); in the forehead she dressed a crown with ivy leaves obtained from amber and hair were gathered in a hairstyle hold by a gold reticulum. In addition, an ivory fan and a distaff were arranged along the sides and grapes, perfumed resin and flowers decorated the grave. The accuracy and wealth of the burial suggested a high ranking social status (Cattaneo *et al*, 1994; Rossignani *et al*, 2005), (Figure 5).
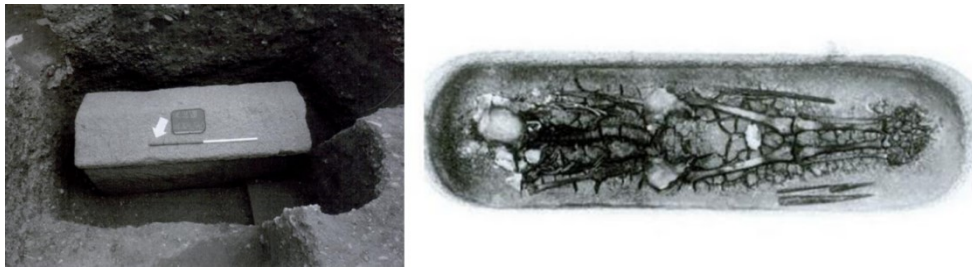
Figure 5. The sarcophagus inside the grave (left) and after its opening (right) (Rossignani *et al*, 2005).

Another particular grave belonged to a young individual (20-25 years old), affected by cranial deformities and diseases (rickets). The anthropological analyses highlighted the African origin of the subject, the only one in the *necropolis* (Cattaneo *et* al, 2011). This case represents a singular burial, especially for the location of the skull between the legs; this ritual probably occurred after the decomposition of soft tissue to prevent the return of that individual as a ghost (Sannazaro, 2011).

S. Eustorgio (Milan)

S. Eustorgio was a small archaeological site, in which eleven inhumations, dated back to the III-VI century AD, were brought to light. Unfortunately, the damaged state of the bone samples prevented to perform a complete building of the biological profile of each individual.

The analyses carried out on the well preserved bone districts allowed to identify a population composed especially by old males (age range between 40 and 60 years old), with European features.

The peculiarity was the absence of diseases, especially osteoarthritis, curious considering the age of the subjects and indicative of not hard occupational activities. Moreover, the lack of particular pathology on teeth, as caries and hypoplasia, pointed at a complete diet, low in sugars.

The only two particular pathology identified in two different individuals were a congenital disease characterized by the lack of fusion of the neural arch of the second cervical vertebra, with possible implications in the mobility of the upper and lower limbs, and an osteoma on a tibia (Figure 6), (Anthropological Report, Soprintendenza Archeologica della Lombardia, 1999).
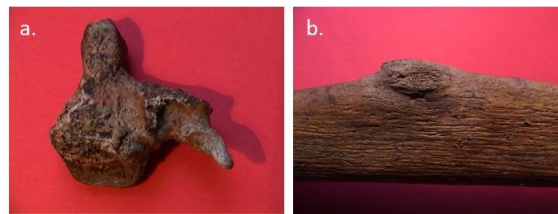
Figure 6. The two pathological conditions identified in the population recovered in S. Eustorgio: a. congenital deformity of the second cervical vertebra, characterized by the lack of fusion of the neural arch; b. osteoma on a tibia.

Piazza S. Ambrogio (Milan)

Located near the homonym cathedral, the site showed a context similar to that observed in the Catholic University.

Archaeological analyses of the ground layers highlighted different usages of the area. Starting from quarrying activities in a time lapse between the I and III centuries AD, it became an area with burial purposes in the Late Antiquity (IV-V centuries AD). The cemetery expansion was carried on during the V-VI centuries AD and in the High Middle Ages, until the Late Medieval period.

Inhumation was the type of burial most represented; generally, bodies were buried in soil, with a supine position of the corpse, but also burials in amphora, wooden coffin, brickwork and "alla cappuccina" were detected (Figure 7), (De Vanna, 2015a).

Figure 7. Piazza S. Ambrogio, some of the different burial practises observed in the *necropolis*: a. burial in earth graves; b. burial in wooden coffin; c. burial in coffin; d. burial "alla cappuccina" (De Vanna, 2015a).

Sixty-three individuals were identified during the archaeological excavation. The skeletal remains were recovered in two different area, belonging to different historical periods: the first one, the most extended with 56 individuals, was dated back to the V-VI centuries AD, the second one included an isolated group of 7 subjects belonging to the Medieval period, in a time lapse between the 1290 and 1430.

The fragmented state of the recovered skeletal remains made difficult the anthropological analyses, sometimes preventing the building of a complete biological profile of the subjects. The most ancient and numerous population of the site was composed by males and females, both adults and children. Paleopathological analyses revealed the presence of several diseases, among which deficiency diseases were the most represented, especially among the children. In addition, *antemortem* trauma and osteoarthritis were detected, the last observed also in young males, suggesting the involvement of young people in working activities (Cattaneo *et al*, 2015a).

The analyses were more difficult for the second small population of individuals, in which only in three cases the good state of preservation allowed the building of the biological profile. This group, which was constituted by males with different ages at death (18-20, 26-45 and 17-22 years old respectively), included two particular subjects.

The first one, buried in grave 23, was a young man characterized by several *perimortem* trauma on both forearms and legs, on the forth lumbar vertebra and the occipital region of the skull, probably associated with his death. The scenario is compatible with the torture of the wheel. Moreover the pathological analysis revealed the possible presence of congenital diseases that probably altered his features, suggesting the possibility that he was considered "different" by the community, maybe marginalized and subjected to discrimination (Cattaneo *et al*, 2015a).

The second interesting case was a young man, buried in grave 49, with morphological features on the skull characteristics of people of African origin. The anthropological analyses revealed, in addition, typical occupational stress markers

on the clavicles, the right shoulder girdle and on the metatarsal bones, indicating the possible dedication of this man to hard activities (dragging and transferring of heavy weights), (Cattaneo *et al*, 2015a).

Bolgare

The archaeological excavations carried out in Bolgare – St. Chierico revealed the presence of an important medieval necropolis with a great number of burials: 284 graves with approximately 460 individuals were detected in an area of about 1,500 square metres (Figure 8), (Fortunati and Ghiroldi, 2006).

Figure 8. The *necropolis* of Bolgare – St. Chierico (Sguazza *et al*, 2015).

The graves, almost all of them with the original covering, contained a single individual in 116 cases, while in the remaining from 2 to 8 subjects, which were buried at the same or in different times, were detected.

The archaeological analyses identified an uninterrupted use of the necropolis for two centuries, between the pre-Lombard and the Medieval periods. In particular, a mixture between the local and the Lombard populations were observed, confirmed also by the anthropological studies (Fortunati and Ghiroldi, 2006).

The population was heterogeneous both for sex and age, with a prevalence of adult individuals. The anthropological investigations suggested an agricultural economy and an uniformity in the distribution of pathology, among which degenerative disorders were the most common, probably associated to heavy activities due to agricultural works.

Peculiar results were obtained from the ethnical diagnosis, which revealed two main groups. Specifically, individuals with northern or eastern European features, associated to the Lombard population, and individuals with central European or

Mediterranean characteristics, typical of local people, were identified. In addition, differences between the groups were observed in stature and, in particular, males and females belonging to the first group had an average stature greater than that of the entire sample. These results confirmed a Lombard migration and especially the burial of Lombard people in Bolgare. Moreover, the observation of skulls with local and northern or eastern features highlighted the integration between the two populations (Sguazza *et al*, 2015; Mazzucchi *et al*, 2006).

Bormio, Piazza del Kuerc

The archaeological site located in Piazza Cavour, Bormio, was characterized by several phases of settlement, from Prehistory to the Modern Age, among which the most preserved belonged to the High Middle Ages and to the second Iron Age. The most ancient human presence in this area was identified by building traces which were destroyed by fire and dated back to the protohistoric period. To these, the archaeological excavations revealed the cemetery area of the church dedicated to SS. Gervasio and Protasio, used probably until1376 (De Vanna, 2015b).

The recovered skeletal remains belonged to 24 adults subjects, with an age range between 20 and more than 80 years at death. Of particular interest were the presence of old individuals also in the past and the difference observed in stature between females and males, about 158 cm and 170 cm respectively (Cattaneo *et al*, 2015b).

The paleopathological analyses identified in the population individuals with lesions caused by repetitive trauma. Two particular cases were represented by a woman with several compound and remodelled fractures, probably related to a pathological condition, and a man with fractures occurred in different times, associated maybe to military events (Figure 9a). Moreover, the most interesting finding was the detection of metastatic bone lesions in two young individuals (Figure 9b). Considering the several genetic and environmental factors that are involved in the onset of carcinoma, the determination of the aetiology of these lesions was difficult. However, among the types of cancer which metastasize to bones, breast and lung tumours were suggested. In particular, the last could be related to cancerogenic environmental factors, such the natural emission of radon from soil or other substances during mining activities (Cattaneo *et al*, 2015b; Caruso *et al*, 2017b).

Finally, the anthropological results revealed a population involved in hard activities, characterized by a poor diet and scarce health conditions (Cattaneo *et al*, 2015b).
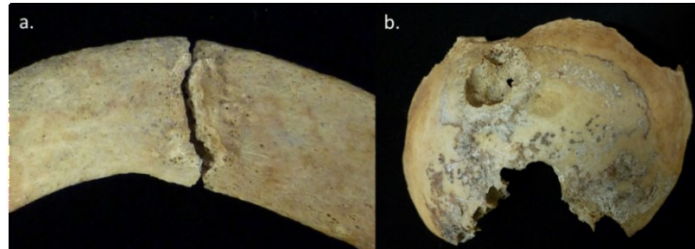
Figure 9. Bone fracture with periostitis on the first left rib (a.) and metastatic lesions on the cranial vault (b.) observed in two individuals of the Bormio archaeological site (Cattaneo *et al*, 2015b).

Via Monte Napoleone (Milan)

Via Monte Napoleone traces the ancient route of the city walls built during the Emperor Massimiliano.

The archaeological excavations, occurred in a building located in via Monte Napoleone 12 during reconditioning works of the basement in the 2011, interested an area of 1,500 m^2.

In this section six stages of usage, from the Roman Age (III century AD) until the '70s of the 1900, were recorded. The archaeological evidences pointed three different areas and usages of the site: the North area, designated to a residential use; the area located in the centre and south of the site, assigned to working activities; the East courtyard, used as cemetery, probably related to the church S. Andrea alla Pusterla Nuova, located between via Monte Napoleone and S. Andrea. The analysis of archive documents suggested the cemetery destination of the area in a period between the 1478 and 1547.

The examination of the *necropolis* showed two different time of burials.

The first one, which occurred during the third stage of usage of the site (1400-1500), is associated to a small group of inhumations, some of which showing bone elements with a characteristic green colour, most likely related to production scraps containing copper (Figure 10).

Figure 10. One of the skeletal remains recovered showing bone elements with a characteristic green colour (Pisacreta, 2014).

17

The most recent instead was a mass grave belonging to the early decades of the 1500s, that included inhumations of individuals buried in a messy position, one over the other, that upset the first burials, suggesting a death related to a possible epidemic event (Figure 11) (Excavation Report, Soprintendenza Archeologica della Lombardia, 2011).

Overall 105 individuals were recovered, including adults and children, females and males. The anthropological analyses pointed out the presence of nutritional deficiencies and especially degenerative diseases, as indicated by osteoarthritis signs on the vertebral column, observed also in young people, evidence of the raising and transferring of heavy weights. Moreover, *antemortem* fractures were recorded in different skeletal districts in men, as well as women, suggesting their involvement in hard activities (Pisacreta, 2014).

Figure 11. Photomap of the burial site in Via Monte Napoleone (Pisacreta, 2014).

A peculiar characteristic of this population was the presence of a particular dental wear observed in the superior central incisors in females, indicating possibly sartorial activity and/or leather and textile manufacturing, the first evidence of the activity in the ancient Milan (Pisacreta, 2014; Cattaneo and Gibelli, 2015).

Viale Sabotino (Milan)

Dated back to the XVII century AD, Viale Sabotino is the only site set outside the city walls. The archaeological evidences revealed a first agricultural use, followed by the excavation of channels during the Spanish dominion and later stages.

Close to the city walls, in the XVII century AD, a grave was dug and filled up in a short period. The grave (length 46 m, width 2.5 m and depth 0.70-1 m) named grave 2 showed 240 individuals divided in 16 groups spaced out by ruins (Figure 12).

The corpses, organized in anatomical connection, were placed in mounds on top of each other in a disorganized fashion, nearly unnatural, suggesting that they were thrown in the grave.

Figure 12. Viale Sabotino: photomap of the grave 2 (Caruso *et al*, 2013b).

The burial far away the city, the position of the remains and the urgency to release the bodies proposed a death caused by an epidemic event, probably the epidemic plague occurred in this period, between the 1629 and 1632, and described by Manzoni in his famous book "*I Promessi Sposi*" (Marsden and Pagani, 2008; Caruso *et al*, 2013a; Caruso *et al*, 2013b).

The population, which was composed fairly by sex and age, showed an high mortality during the first childhood and adult age, in agreement with other contemporary populations, and during adolescence, probably due to nutritional deficiencies and pregnancy in females.

The paleopathological analyses revealed the poor health conditions of the population. In fact, congenital, nutritional, infectious diseases and metabolic disorders were observed on the skeletal remains, as well as degenerative pathology and trauma, suggesting the involvement in hard activities (Caruso *et al*, 2013b).

This confirmed the difficult life conditions during the end of 1500 and the beginning of 1600 in Milan, marked out by epidemic events, dearth and conflicts. The malnutrition observed on the skeletal remains and the bad hygienic conditions associated to the poverty and high urban concentration of this period exposed the population to infectious diseases. In fact, tuberculosis and syphilis were detected on the skeletal remains of some individuals. Unfortunately, no evidence of the plague was identified on bone samples due to the severe progress of the disease.

Only biomolecular analyses allowed the identification of the pathogen *Yersinia pestis* in the population recovered in the mass grave, confirming the death caused by this epidemic event (Caruso *et al*, 2013a; Caruso *et al*, 2013b).

Cà granda (Milan)

Cà Granda was the ancient greater hospital in Milan, built for the will of Francesco Sforza in the second half of the XV century AD. The hosted patients were usually people coming from the needy and poor classes, suffering of severe diseases, generally syphilis, plague and all types of fever. The hospital, characterized by two symmetrical parts divided by a courtyard, included also a church and a cemetery, which was built only in the 1634 (Cosmacini, 1999; Riva and Mazzoleni, 2012).

The ossuary, located under the chapel, consisted of fifteen cells, only nine with skeletal remains. Each cell had a manhole for the access, an opening of square

shape (50 cm per side) with a manhole cover, through which the corpses were gone down (Carlessi and Kluzer, 2011; Carlessi and Kluzer, 2013).

Nowadays, the grave shows up as a mass of commingled bone elements, not jointed anatomically due to the burial occurred in a no neatly manner.

The archaeological and anthropological inspections started in 2010-2011 and it is currently being carried on. The first examinations suggested a use of the grave as place of primary deposition for a long period of time, in which other burial occurred disturbing the previous inhumations, confirming the historical sources. Each cell contains bone elements in different state of preservation and amounts, mixed with ruins and other materials (Figure 13), (Sguazza, 2015).

Figure 13. Cà Granda: photomap of two tunnels of the ossuary (Cattaneo *et al*, 2013).

Approximately, two millions of skeletal elements belonging to more than 500,000 hospitalized individuals have been estimated, making the Ospedale Maggiore (Cà Granda) a cultural resource of inestimable value due to its antiquity, numbers and peculiarity of the context (Sguazza, 2015).

The particular high number of commingled bone samples could be attributed to the presence of water within the tunnels in that period, due to the proximity of the canal as reported in a relation by Cristoforo Monforti (1694), which probably caused the dispersion of the skeletal remains, and especially to the human activity to empty the grave and create more space for new corpses.

The anthropological analyses revealed an heterogeneous population by sex and age, characterized by an high mortality in the period between birth and the first month of life (Sguazza, 2015). This was in agreement with the demographic condition of that period and, in particular, the recovery of remains belonging to foeti and newborn babies proved the hospitalization of pregnant women and unacknowledged children started from the 1671, as indicated by Cosmacini and Reggiani (Cosmacini, 1999; Reggiani, 2014).

The skeletal remains portrayed the pain of these people, revealing skeletal deformities caused by trauma or other pathology, such as degenerative, infectious and autoimmune diseases and metabolic disorders. To these, congenital diseases, as

the harmonic dwarfism and genetic limbs deformation, and lead poisoning were added (Sguazza, 2015; Sguazza *et al*, 2016a; Sguazza *et al*, 2016b).

In particular, the detection of infectious diseases (osteomyelitis and syphilis) suggested the inadequate hygienic conditions and especially the poverty characterizing this period.

Generally, the paleopathological analyses highlighted a population suffering for several pathology, associated in some cases to disability, as well as *antemortem* trauma, that could limited the motion of the involved articulations, making difficult sometimes daily activities (Sguazza, 2015).

Cimitero Maggiore (Milan)

The skeletal collection from Cimitero Maggiore, Milan, includes more than 2,000 individuals, buried between 1990 and 1998 in wooden coffins in the cemetery, exhumed and re-buried twice in 15 years and finally moved to ossuary for further 5 years. Owing to the missed claim from relatives, the remains were acquired by the Laboratory of Forensic Anthropology and Odontology (LABANOF), University of Milan, for scientific purposes, according to the Article 43 of the DPR n. 285 of September 10th, 1990 of the National Police Mortuary Regulation.

The skeletal population is constituted by approximately 150 infants, with an age range between 0 and 3 years, and 2,200 adults, both females and males. The availability of individual information such as gender, age, ethnicity, date of death and the modernity of the samples make this population an useful resource to validate the anthropological methods used in the building of the biological profile from skeletal remains.

The first analyses carried out on a sample of the collection revealed a population composed especially by old subjects, the 80%, in fact, was older than 60 years. Consequently, the pathological diagnosis of the skeletal remains identified diseases typical of the old contemporary population, in particular degenerative joint disease. Furthermore, some pathology common and increasing in these last years, such as metastatic tumours, were recognized.

In addition, trauma analysis revealed both *antemortem* and *perimortem* lesions, the last one observed only in a small group of skeletons (Cappella, 2014).

1.2 ANTHROPOLOGY

1.2.1 An overview

Anthropology, from the Greek *anthropos* (human) and *logos* (science), is the science that studies the human being from a biological, cultural and physical point of view. It includes physical anthropology which deals with the biological characteristics of human, with particular interest for human evolution and variability (Kranioti and Paine, 2011).

Physical anthropology applies the knowledge of human osteology to interpret human bones.

The aim is to describe each individual in terms of physical appearance, providing data also about habits, diet, health and last moments before death.

These information can be acquired by studying the morphology and the shape of bones and teeth. In fact, the skeleton preserves in its structure the traits and the signs related to ancestry, sex, age at death and the life conditions which modified its appearance. These markers and features can be generated during skeletal formation and growth under the influence of several factors, as the action of genes and environment, as well as after particular physical stress and pathological or traumatic events. The dynamism which characterises the bone tissue is due to the modelling and re-modelling operated by bone cells, which modify its shape during the lifetime (White and Folkens, 2005).

This variability is on the basis of the anthropological analyses of skeletal remains.

After having verified the human origin of the remains in study, the anthropologist searches for those specific signs which allow to trace a biological profile of the individual, starting from ancestry estimation, sex diagnosis and the estimation of age at death.

For this purpose, particular attention is given to those features on the skeleton which vary among people coming from different geographic regions and, specifically, both morphology and metric features of the skull in its entirety and of specific parts, including teeth, nasal aperture, orbits, palate, mandible and maxillary bones (Figure 14) are evaluated (Bass, 1995; Cattaneo and Grandi, 2004; Hefner, 2009; Slice and Ross, 2009; Stull *et al*, 2014).

Similarly, in sex diagnosis precise traits visible in the skull and pelvis, the most significant bone districts, are considered to identify the sex of the person (Figures 15-16) (Cattaneo and Grandi, 2004; White and Folkens, 2005).

Finally, to estimate age the variability existing in terms of size and shape between skeletal remains belonging to children and adults is investigated. Keeping in mind the osteogenesis processes for bone formation, age estimation in subadults could be performed considering teeth eruption and formation (Ubelaker 1999; AlQahtani *et al*, 2010), length of diaphyses not fused (Figure 17) (Fazekas and

Kosà, 1978; Scheuer & Black, 2000) and evaluation of the fusion state of the epiphyses (Scheuer & Black, 2000; Schafer *et al*, 2008). Differently, in adults, age is easily and mainly estimated considering the state of those articulation, whose wear rate advanced in an homologous way in the general population (Cunha *et al*, 2009). Among these, the sternal end of the IV rib (Işcan's method: Işcan *et al*, 1984; Işcan *et al*, 1985), the pubic symphysis (Suchey-Brooks' method: Brooks and Suckey, 1990, Figure 18) and the auricular surface of the ilium (Lovejoy's method: Lovejoy *et al*, 1985) are the most considered joints.

Figure 14. Skulls belonging to individuals from two different geographic areas: a. Europe; b. Africa. European people are characterized by a long and narrow nasal aperture, with a sharp nasal sill (the inferior edge of the nasal aperture) and prominent nasal spine. African individuals, differently, have a greater nasal aperture and a poor nasal sill (modified from Cattaneo and Grandi, 2004).

Figure 15. Comparison between the skull of a female (on the left) and a male (on the right). Females have a minimal prominent glabella (1), thin and sharp supraorbital margin (2), a zygomatic arch which does not extend posterior to the external auditory meatus (3), small and oblique mastoid processes (4) and no gonial eversion (6); on the other hand, males show a protruding glabella (1), round and thick supraorbital margin (2), a zygomatic arch which extends posterior to the external auditory meatus (3), massive, large and verticalized mastoid processes (4) and gonial eversion (6) (Cattaneo and Grandi, 2004).

Figure 16. Comparison between pelvis of a female (on the left) and a male (on the right). Females have a wider pelvic inlet with lower iliac bones (A); males show a tight pelvic inlet with high iliac bones (B) (Cattaneo and Grandi, 2004).

Figure 17. Comparison of the length from femurs belonging to children with different ages (modified from Cattaneo and Grandi, 2004).

Figure 18. Suchey-Brooks' method for age estimation in adults: the method considers the appearance of the surface of the pubic symphysis. Six stages are identified, each one associated to specific features observed on the surface. In particular, the figure shows the gradual transition from surfaces showing the characteristics typical of young individuals (in the first stages), such as dense bone with billowing (small crests separated by small cavities), to flat surfaces typical of older subjects (in the last stages), in which billowing disappeared and deteriorations signs, as porosity, appear (modified from Cattaneo and Grandi, 2004).

The study of skeletal remains proceeds subsequently with pathological and trauma analyses.

Each bone element during life is in contact with soft tissue, including tendons, ligaments, muscles, blood vessels, and carries in its morphology their signature. After death and decomposition processes when soft tissues disappeared, bones preserve those information suitable to describe the individual's life.

Some diseases and stress suffered during life could leave marks on bone tissue. In this case, bone tissue reacts in two different manners, destroying or producing new bone tissue (Figure 19).

Similarly, traumatic events show peculiar features depending on the time of occurrence: during life (*antemortem*), closeness to death (*perimortem*) or after death (*post mortem*). Fractures which occur during life activate the healing process and a bone reaction could be visible on the involved bone (Ortner, 2003); differently, trauma which occur closeness to or after death could be identified on the basis of the shape of the lesions and of specific characteristics of the fractured edges (Figure 20).
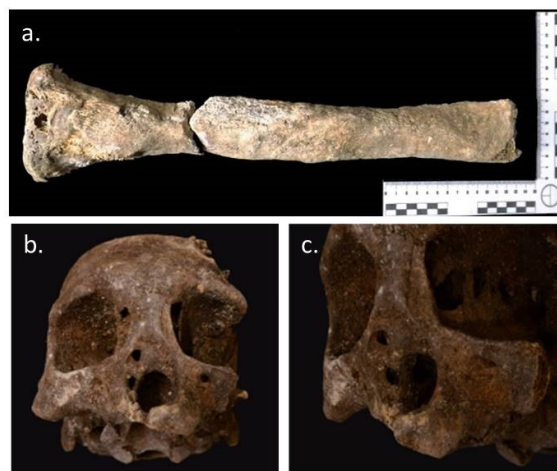
Figure 19. Examples of pathological conditions: a. osteomyelitis in a tibia (example of proliferative disease); b. potential syphilis in a skull (example of lytic disease). In c. a detail of the lesions in proximity of the nasal aperture (Sguazza, 2015).

Figure 20. Examples of trauma: a. antemortem trauma in a rib, in which bone callus is a marker of the ongoing healing process; b. perimortem fracture in a tibia, characterized by smooth and sharp edges with a obtuse or acute angles and a colour similar to the rest of the bone, indicative of a trauma occurred closeness to death; c. post mortem fractures in a humerus, characterized by irregular and jagged edges with a right angle and a colour of the margins lighter than the external cortical surface (Moraitis and Spiliopoulou, 2006).

Macroscopic analysis of the skeletal remains is supported also by microscopic investigations. In fact, the study of bone structure at microscopic level is a valid

help for different purposes, such as species diagnosis, to discriminate human and animal bones, age estimation at death, from bones and teeth sections (Kerley and Ubelaker, 1978; Gustafson, 1950), and in trauma analysis.

In particular, concerning trauma, the possibility to investigate bone callus and fractures microscopically could supply crucial information about the time in which the lesion occurred and the definition of its vitality, if it was caused at the time of death or in a following period (Cattaneo *et al*, 2010). This is very important in the determination of the cause of death, especially in those cases in which the macroscopic analysis was not sufficient for the diagnosis.

Similarly, in species discrimination the study of the microscopic structure of the bone components could help in the identification of human remains, especially when these are highly fragmented.

Unfortunately, the possibility to get answer and to obtain a complete biological profile depends on the state of preservation of the remains which are studying. Time since deposition, as well as environmental factors, including scavenger activities, can influence the preservation of those bone districts of interest for the anthropological diagnoses, such as those for age estimation as reported by Cappella and colleagues (2017). Moreover, they can alter bone morphology preventing the identification of important signs necessary to identify, for example, *perimortem* fractures (Cappella *et al*, 2014) or pathological markers.

These changes affect not only the macroscopic appearance of bones, but also the microscopic structure of bone tissue. These alterations prevent the possibility to perform analyses and get successful data from bone sections.

The study, both at the macroscopic and microscopic level, of skeletal remains recovered from different burials and with different times since deposition could provide important data concerning the involvement of these several factors on the preservation of the structure and the appearance of bone tissue.

Gross bone and histological appearance were already investigated in genetic studies of archaeological remains to find a possible correlation between bone tissue features and DNA preservation. This is based on the belief that as bones weather, similarly the organic molecules, among which DNA, degrade. In fact, the same environmental and timing factors which influence bone degradation are involved in DNA preservation, but few studies described this overlapping (Misner *et al*, 2009).

In some studies, authors found a correlation between microscopic tissue preservation and DNA recovery: specimens characterised by well-preserved structure provided also good genetic data (Haynes *et al*, 2002; Misner *et al*, 2009). However, in other cases no relation was detected (Misner *et al*, 2009), highlighting contradictory conclusions. For this reason, histological and macroscopic bone appearance are not adopted so far as screening methods for the choice of bone samples to submit to DNA analyses.

In the following paragraphs macroscopic and microscopic structure of bones are described followed by a description of the changes which could influence their appearance.

1.2.2 Bone histology

Bone is a specialized connective tissue, whose cells are specialized in the synthesis and secretion of bone matrix within which they are subsequently enclosed.

Bone matrix is formed by two components: inorganic (65% of its dry weight), composed especially by calcium and phosphorus, in the form of hydroxyapatite crystals, and organic (35% of its dry weight), constituted only by collagen fibers, especially collagen type I. In the bone matrix an amorphous substance is also present. It is composed by proteoglycans, characterized by chondroitin-sulfate and keratan-sulfate chains, and glycoproteins typical of bones such as osteocalcin and osteopontin (Gartner and Hiatt, 2002).

1.2.2.1 Bone macroscopic structure

A simple model useful to describe the bone macroscopic structure comes from long bones, such as humerus, femur and tibia. Each long bone is composed by a central part, diaphysis, and two extremities, epiphyses, connected together by a transition zone, called metaphysis.

Cutting longitudinally a diaphysis of a femur, two components can be identified: a thicken structure that characterized the external part, called compact or cortical bone, and a porous structure which delimit the medullary cavity, organized in trabeculae and bone laminae, called spongy or trabecular bone. In particular, trabeculae and spicules originate from the inner surface of the compact bone to the medullary cavity.

While the diaphysis is composed especially by compact bone, the epiphyses and metaphysis are composed mostly by trabecular bone with a thin layer of the cortical one (Figure 21).

The external surface of bones is surrounded by periosteum, except in the regions where muscles and tendons insert in, in the joint surface and sesamoid bones. The periosteum is a dense connective tissue, not calcified, irregular, characterized by a great amount of collagen, anchored to the bone tissue by the Sharpey's fibers. Two parts are recognized: an external fibrous portion, which conveys blood vessels and nerves, and an inner cellular part, rich in osteoprogenitor cells.

The medullary cavity of the diaphysis and cavities of the trabecular bone are surrounded by a thin cellular layer, composed by osteoprogenitor cells and osteoblasts, called endosteum (Weiss, 1983; Gartner and Hiatt, 2002; Martini *et al*, 2010).
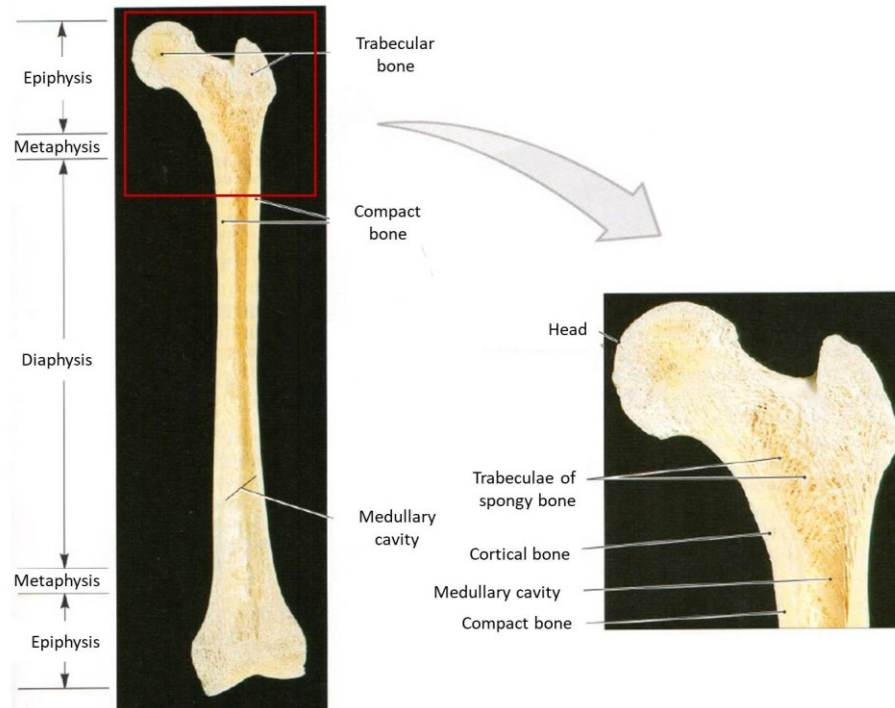
Figure 21. Macroscopic structure of a long bone (modified from Martini *et al*, 2010).

1.2.2.2 Bone microscopic structure

Adult compact and trabecular bone tissues are lamellar. Lamellae look as a series of parallel or concentric layers, similar to the layers of plywood.

Each lamella is 3-7 µm of thickness and the collagen fibers are in parallel oriented. Microscopic view of histological bone sections using polarized light allows to discriminate the lamellar structure of bone as a succession of dark and light layers, due to the different orientation of the collagen fibers between consecutive lamellae.

Between compact and spongy bone lamellae, small cavities, called lacunae, are visible and connected together by canaliculi, thin tubular channels. Osteocytes, the mature bone cells, and the corresponding cytoplasmic processes are located in lacunae and canaliculi respectively.

Cellular processes are connected by gap junctions with those of osteocytes localized in adjacent lacunae, supplying a way to exchange nutrients and waste products.

In compact bone, lamellae could constitute four different structures (Figure 22):

- external circumferential lamellae, below the periosteum, which characterize the outermost part of the diaphysis and, together with the Sharpey's fibers, anchor the periosteum to the bone tissue;
- internal circumferential lamellae, that delimit the medullary cavity; bone trabeculae originate from them to the medullary cavity, interrupting the endosteum surface of the inner lamellar layer;
- osteon or Haversian system, characterized by concentric lamellae surrounding a vascular channel; osteon size is different as different is the number of lamellae of each one;
- interstitial lamellae, osteon residues surrounding by cement lines, which "fill" the space among osteons

(Weiss, 1983; Gartner and Hiatt, 2002; Martini *et al*, 2010).
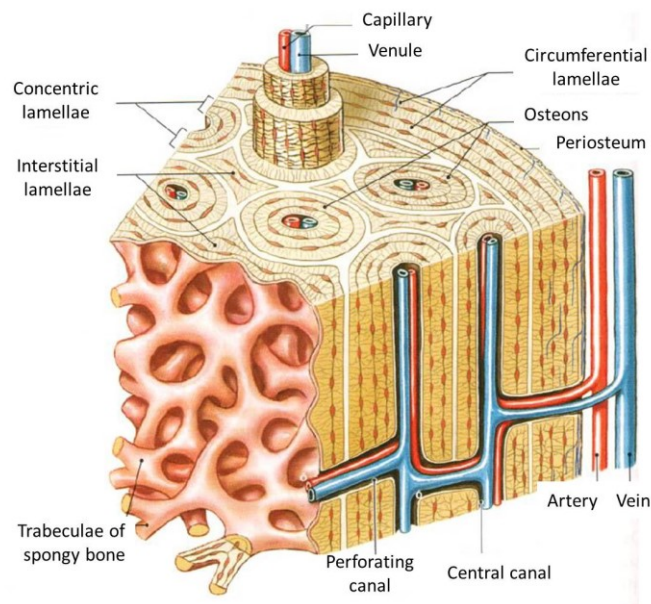
Figure 22. Bone microscopic structure (modified from Martini *et al*, 2010).

Osteon or Haversian system

In the diaphysis, osteons are parallel to the long axis of the bone, thereby in a transversal section the osteons are cut transversally.

Osteons are crossed along the entire axis by central or Haversian canals, surrounding by concentric lamellae, among which lacunae with osteocytes are visible (Figure 23).

Each canal, characterized by a diameter of 30-70 μm, is delimited by a layer of osteoblasts and osteoprogenitor cells and included blood vessels, nerves and connective tissue. The blood vessels come from the periosteum or bone marrow, reaching in the last case the Haversian canal through the Volkmann's canals,

Figure 23. Osteon with erythrocytes identifiable inside the canal and osteocytes visible in the lacunae surrounding the Haversian canal (Hematoxylin-Eosin staining, 630x).

transverse or oblique canals, which differs from the first one for the lack of concentric lamellae.

Externally, a cement line, 1-2 µm of thickness, composed by a calcified amorphous substance with few collagen fibrils, surrounds each osteon. In particular, two types are identified: the arrest lines, smooth and uniform, and the reversal lines, irregular or dentate. The arrest lines occur when the bone formation resumes after a block; the reversal lines, on the contrary, form when the bone formation follows the bone removal or reabsorption. In summary, the cement line typology gives evidence of the history of a specific bone region.

During osteon formation, the first lamella which is synthetized is that close to the cement line. After the deposition of further lamellae, the osteon thickness increases, while the canal lumen decreases. Each osteon could have a total of 4-20 lamellae (Weiss, 1983; Gartner and Hiatt, 2002; Martini *et al*, 2010).
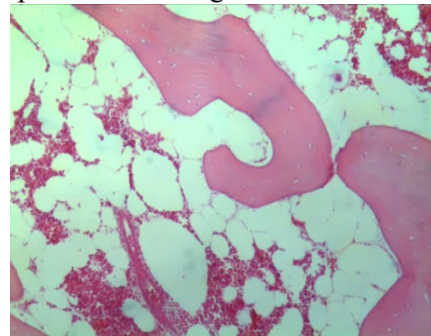
The trabecular bone

Mature trabecular bone is composed by bone trabeculae and intertrabecular spaces containing bone marrow (Figure 24). The majority of trabeculae are characterized by a thickness lower than 0.2 mm and lacks blood vessels. Differently, trabeculae with a thickness greater than 0.2 mm generally contain a structure similar to osteons in the central region.

Figure 24. Bone trabeculae with bone marrow spaces containing blood components (Hematoxylin-Eosin staining, 100x).

In fact, each bone trabecula is composed by trabecular packet, a mosaic of angular segments, each one characterized by lamellae comparable from a functional point of view to osteons. As osteon is the structural unit of the compact bone so the trabecular packets represent the structural unit of the trabecular bone.

The trabecular packet has a shape similar to an half-moon with a radius equal to 600 µm, a thickness of 50 µm and a length of 1 mm. As in the compact bone, a cement line keep the trabecular packets together (Weiss, 1983; Gartner and Hiatt, 2002; Martini *et al*, 2010).

The bone marrow

Bone marrow is localized in the medullary cavity of long bones and in the spaces between trabeculae of the trabecular bone, separated from bone tissue by endosteum. It is a greatly vascularized tissue.

Since the fifth month of life, it is responsible of the production of the blood cellular components (hematopoiesis) and their passage to the circulatory system.

Two different types could be identified: red bone marrow, for the great amount of erythrocytes, and yellow bone marrow, characterized by an accumulation of adipose tissue instead of the hematopoietic one which regresses. Red and yellow bone marrow could be interconverted, depending on the request for changes in the hematopoiesis. Since twenty years from birth, the diaphysis of long bones have only yellow bone marrow and in adults the hematopoietic tissue is limited to the skull, clavicles, vertebrae, ribs, sternum and pelvis.

Blood vessels localized in the bone marrow come from arteries which perforate the diaphysis of long bones through nutrient foramen and give rise to small vessels, which spread peripherally and generate branches both to the central part of the cavity and to the cortical bone. The last ones branch out and go in Haversian and Volkmann's canals of the compact bone. The branches to the medullary cavity generate sinusoids, with diameter equal to 45-80 μm, covered by endothelial cells and surrounded by thin reticular fibers and a great number of adventitial reticular cells. The sinusoids flow into a central longitudinal vein, which end into veins coming out the bone by nutrient canals (Weiss, 1983; Gartner and Hiatt, 2002; Martini *et al*, 2010).

The bone cells

Within bone tissue we can identify four cells, following described:

OSTEOPROGENITOR CELLS: localized in the endosteum, in the inner layer of periosteum and Haversian canals, osteoprogenitor cells are able to divide by mitosis and differentiate in mature bone cells. Two different types could be distinguished: pre-osteoblasts, which give rise to osteoblasts, and pre-osteoclasts, which differentiate in osteoclasts.

OSTEOBLASTS: localized in the outer and inner bone surface, where new bone is synthetized, they are involved in the formation and calcification of bone tissue and synthesis and secretion of the non-mineralized bone matrix, called osteoid. After the release of the secretion products, cells surrounded by the synthetized bone matrix are called osteocytes.

OSTEOCYTES: derived from osteoblasts which secreted bone matrix around them, osteocytes adapt themselves to the lacuna. Young osteocytes, visible near the bone surface, are localized in rounded lacunae, while the oldest ones stay in oval lacunae.

OSTEOCLASTS: multinucleated cells (from 2 to 50 nuclei), with a diameter equal to 20-100 μm and acidophilic cytoplasm, involved in bone reabsorption. Osteoclasts are localized in the so-called Howship's lacuna, a small depression, demonstration of an ongoing reabsorption process.

Osteoblasts and osteoclasts activities have to be equilibrated: if osteoclasts remove calcium salts faster than osteoblasts deposit them, bone becomes more fragile; differently, if osteoblasts activity is predominant, bone becomes more robust and massive (Weiss, 1983; Gartner and Hiatt, 2002; Martini *et al*, 2010).

1.2.3 Bone taphonomy and diagenesis

As during life, modifications of bone tissue can occur also after death.

In fact, during the decomposition process, even after the complete skeletonization, a body is continuously subjected to transformations and alterations caused by the context where it is. We name "taphonomy", all the transformation that occur after death induced by scavengers, human and environmental factors (water, temperature, soil characteristics, sunlight exposure…) (Dirkmaat *et al*, 2008).

These changes can modify the macroscopic appearance of the skeletonised remains. In fact, soil can cause colour alterations (Figure 25), while the exposure to the sun for long periods can determine bleaching and cracking in the bone surface. To these, other factors, as water and fire, affect bone structure, causing scratching in the first case and deformations and alterations in the second one.

Figure 25. Taphonomic changes in skeletal remains at the macroscopic level: a. skull buried in soil (5 years); b. skull decomposed on the surface (8 months) (Cattaneo and Grandi, 2004).

Taphonomic alterations occur not only at the macroscopic level, but they involve also the microscopic structure of bone tissue, as reported in a recently published paper (Cappella *et al*, 2015). In particular, the time after death and the different taphonomic conditions influenced the presence and the appearance of bone blood components. In fact, erythrocytes could not show their typical shape

just after a week of decomposition in air; from this time only nonspecific accumulations were detectable by Hematoxylin-Eosin staining (Figure 26a). This was confirmed through immunohistochemical analysis searching the antigen specific for red blood cells, Glycophorin A. In addition, generic debris derived from botanical or microorganisms contaminations were observed inside bone canals and spaces (Figure 26b), indicating further alterations caused by environment and time since death.
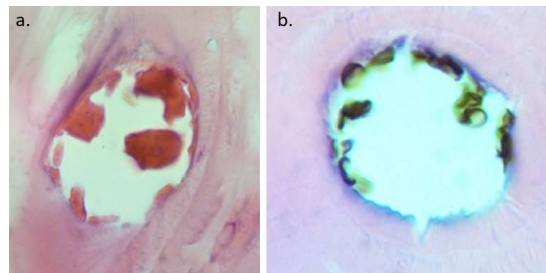
Figure 26. Taphonomic changes at the microscopic level: a. blood components in an Haversian canal as nonspecific accumulations after six weeks of decomposition in air (Hematoxilin-Eosin staining, 630x); b. presence of generic debris (possible microorganisms) in an Haversian canal after five weeks of decomposition in air (Hematoxilin-Eosin staining, 630x) (Cappella *et al*, 2015).

In archaeological context bones might be subjected to marked and complex modifications which change the structure and chemical composition due to the environment where skeletal remains were buried. This alteration is named diagenesis (Guarino *et al*, 2006). Diagenesis included all the transformations that occur during the burial. Specifically, it is the accumulation of chemical, physical and biological events which modify the bone tissue and alter the original chemical composition and structural appearance, influencing its fate in terms of preservation or destruction (Wilson and Pollard, 2002).

During burial, bones are in contact with the surrounding environment, including sediments, soils, pore water. Taking into account the bone tissue composition, characterized by the organic component, constituted principally by collagen, and the inorganic fraction, composed by hydroxyapatite crystals, several alterations might be occurred. Among these, erosion, precipitation, partial or complete dissolution, recrystallization, ion uptake by resorption and diffusion, hydrolysis and repolymerization.

Bone preservation varies and is influenced by environmental conditions, as temperature, groundwater and sediment composition, mechanical pressure, soil hydrology and pH, biological factors and particle transport (Reiche *et al*, 2003). At the same time, it is shown that different elements of the skeleton could be characterized by different level of resistance to destruction in soil (Lambert *et al*, 1982).

Several studies were performed to understand the diagenesis process, the influence of environmental conditions on bone tissue and the effect of diagenesis on bones (Reiche *et al*, 2003). In particular, Collins *et al* (2002) proposed three mechanisms for bone deterioration: chemical deterioration of the organic fraction, chemical deterioration of the mineral phase and microorganisms attack.

Specifically, collagen is lost by chemical hydrolysis and consequent porosity. This is an uncommon and slow mechanism, which occurs in environments that are stable for the mineral component from a geochemical point of view. Generally, most environment are not in thermodynamic equilibrium and cause chemical deterioration of the mineral phase, governed by dissolution/re-precipitation (Collins *et al*, 2002). The rate of mineral loss is conditioned by the concentration gradient of the surrounding environment, which in turn depends on the soil water composition. In fact, in soil characterized by neutral pH the concentrations of calcium and phosphate are close to saturation with the bone mineral content and thereby the dissolution rate is slow. Differently, dissolution rate increases due to change of environmental conditions owing to a decrease in pH or recharge with new water (Hedges, 2002).

Because a mutual protection exists between the organic and mineral components, which confers stability to the two phases during burial, when it is broken, it results in a rapid degradation of bone tissue (Trueman and Martin, 2002). In particular, mineral dissolution causes a major exposure of the collagen to chemical and microbial deterioration.

Microorganisms destruction is the most common cause of bone deterioration. The main agents responsible of bone alteration are fungi and bacteria, as demonstrated by Marchiafava *et al* (1974), which demineralise bone producing the common alterations known as tunnels or borings (Dixon *et al*, 2008). In particular, bacteria produce collagenases, which destroy bone collagen, generating focal destructions within bone tissue. Microbial activity is influenced by several factors, such as temperature, oxygen and water availability and the presence of inhibitors for the growth of microorganisms (Collins *et al*, 2002).

To evaluate bone preservation, macroscopic and microscopic investigations could be performed. At the macroscopic point of view, one of the first and important studies in which bone appearance was classified on the basis of bone surface preservation was that performed by Beherensmayer (1978). In particular, Beherensmayer identified six stages, from 0 to 5, evaluating the presence of soft tissue, cracking (cracks on the bone surface), flaking (cortical exfoliation), the appearance of the cortical surface and the exposure of spongy bone.

Concerning microscopic analysis, Hedges and colleagues (1995) proposed the Histological Index, still used to describe the preservation of bone tissue sections. Specifically, it is a value, ranging from 0 to 5, according from bad preserved tissue, where the internal structure is lost and several holes (tunnels) are observed, to well preserved sections, in which all the bone features are visible (Figure 27).
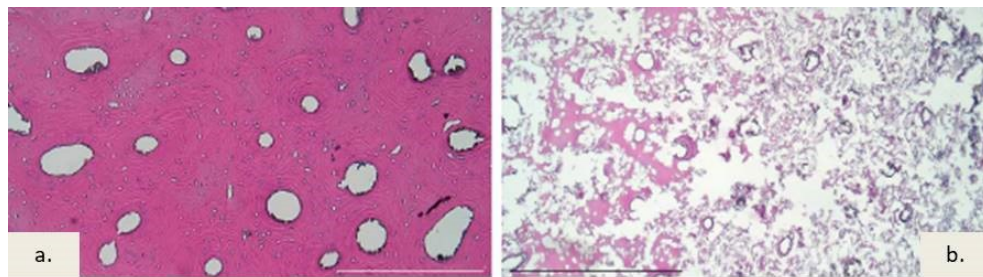
Figure 27. Microscopic view of decalcified thin bone sections: a. well-preserved bone tissue; b. severely degraded tissue (Caruso *et al*, 2017a).

1.3 ANCIENT DNA ANALYSIS

The field of ancient DNA (aDNA) research is a young area of interest that developed quickly in the last decades. It involves all genetic studies carried out on archaeological and historical specimens.

These include any type of preserved tissue, among which bones and teeth are the most common. In addition, mummified or permafrost tissues, extinct animals, plant and insect remains, coprolites (vertebrates fossil excrements) are other examples of ancient DNA sources.

The first successful ancient DNA study was carried out using a dried muscle collected from a museum specimen of a quagga, an extinct zebra sub-species, by Higuchi and colleagues in 1984 (Higuchi *et al*, 1984), amplifying the extracted molecules through molecular cloning. In 1985, Pääbo confirmed the possibility to recover DNA sequence information from ancient samples analysing dried skin tissues from an ancient Egyptian mummy (Pääbo, 1985). In both cases, the genetic analyses revealed the very limited amount and the low molecular weight (< 500 bp) of DNA molecules recovered from the specimens, pointing out the degraded state of preservation of ancient DNA and the low number of endogenous DNA fragments contained in the sample.

The development of the Polymerase Chain Reaction (PCR) allowed to overcome these problems, amplifying exponentially a specific fragment of interest starting from a very limited number of DNA molecules (Pääbo, 1989; Pääbo *et al*, 2004). Therefore, the amplification reaction could be performed multiple times from the same sample in order to verify the repeatability and reproducibility of the results (Pääbo *et al*, 2004). Moreover, DNA characterization carried out on different DNA extracts from the same and/or different bone elements, as well as the analysis of the same extract in a different laboratory, could assess that the obtained sequences did not contain errors and that they came really from the original biological sample and not from a contamination.

In the years following the introduction of the PCR technique, the number of studies carried out on ancient specimens increased considerably. Following the great advantage related to the possibility to recover ancient DNA sequences in multiple copies, two major problems of working with ancient DNA were highlighted, the molecular damage of ancient DNA fragments and the exogenous contamination.

In fact, the application of new molecular approaches to the first aDNA studies evidenced two incorrect positions in the original quagga sequence (Pääbo and Wilson, 1988) and the issue of modern DNA contamination (Pääbo, 1989) in ancient samples (Pääbo *et al*, 2004). Furthermore, DNA sequences were obtained from million years old samples of plants (Golenberg *et al*, 1990), dinosaurs (Woodward *et al*, 1994) and amber inclusions (DeSalle *et al*, 1992; Cano *et al*, 1993; Poinar *et al*, 1993) but the results were then proved to be non-authentic (Willerslev and Copper, 2005).

This led to the improvement of methods to recover reliable information from ancient specimens. In particular, criteria of authenticity and guidelines were recommended, extraction methods were improved and new techniques, such as next generation sequencing (NGS), were introduced.

Moreover, the advances in high-throughput sequencing technologies shifted attention of researchers from single genetic loci to genome wide analyses, generating complete genome sequences of the Neanderthal man and other hominins genomes (Stoneking and Krause, 2011), thus supplying relevant genetic data useful to understand the evolutionary and migration events occurred in the past.

In addition to the studies performed on ancient human remains, other molecular investigations have been carried out on ancient human pathogen genomes and on host-pathogen interaction changes to study the diversity and evolution of human-associated microorganisms (Hagelberg *et al*, 2015; Krause and Pääbo, 2016). To these, further analyses have been performed to understand the evolutionary process and phylogenetic relationships between extinct organisms and present species, as well as outline the genetic history of domestication of animals, such as pigs and cattle (Rizzi *et al*, 2012; Krause and Pääbo, 2016).

1.3.1 Molecular damage

In living organisms, the DNA molecules continuously are undergoing to chemical modifications, which are repaired by enzymatic mechanisms able to maintain the integrity of the genome (Lindahl, 1993).

After death, these repair processes stop to function, causing an accumulation of damage and instability in the genome. In addition, other factors contribute to DNA degradation, such as catabolic enzymes like lysosomal nucleases, which are released after the breakage of the cellular compartments, and microorganisms, which colonized the decomposed tissue. All these events could lead to a reduced chance to recover DNA for the following genetic typing.

Only under particular conditions, as frozen tissue or very quickly dehydrated tissue, the enzymatic and microbial degradation processes could be limited, although other additional chemical modification could occur, such as oxidation and hydrolysis (Pääbo *et al*, 2004; Dabney *et al*, 2013b).

The degradation process of the genetic substrate, as result of chemical modifications and enzymatic activities, is more influenced by environmental conditions than the time elapsed since deposition (Alaeddini *et al*, 2010). In fact, as reported by the authors, the presence or absence of water as moisture, dry conditions, variations in temperature and pH are some of the factors that are involved in DNA preservation.

Three different lesions generally are identified as characterizing ancient DNA: fragmentation, blocking lesions and miscoding lesions.

Fragmentation

DNA extracted from ancient specimens is characterized by a short fragment length, generally lower than 500 bp (Pääbo, 1989; Hofreiter *et al*, 2001b; Green *et al*, 2009). Both enzymatic and nonenzymatic processes are suggested to cause a reduction in size (Pääbo *et al*, 2004). Enzymatic processes are originated by nucleases, released from the organism or by microorganisms which colonised the body. These degradative enzymes start degrading DNA after death and their activity depends on several factors, such as the presence of specific cations (Mg^{2+} and Ca^{2+}), temperature and pH (Alaeddini *et al*, 2010). On the other hand, nonenzymatic process creates single strand breaks by an hydrolytic depurination, followed by β elimination. In detail, the cleavage of the N-glycosil bond between the sugar and guanine or adenine residues produces an abasic site, and the subsequent β elimination the fragmentation of the strand (Figure 28) (Lindhal and Nyberg, 1972; Pääbo and Wilson, 1991; Lindahl, 1993).
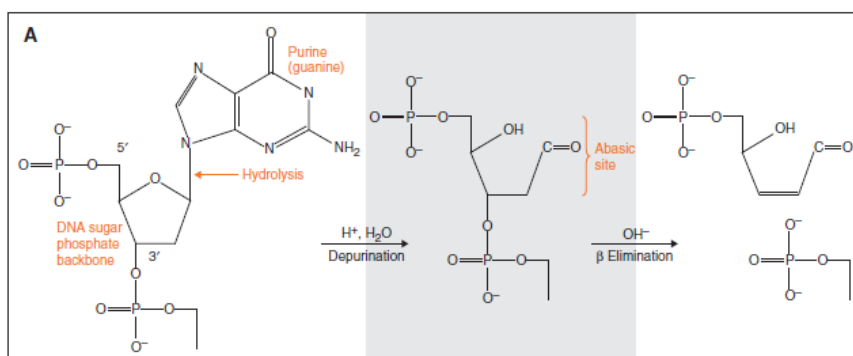
Figure 28. Fragmentation of a DNA strand through hydrolytic depurination, followed by β elimination (Dabney *et al*, 2013b).

The development of high-throughput sequencing techniques allowed to better investigate the fragmentation pattern and specifically the strand breaks (Dabney *et al*, 2013b). Indeed, studies carried out on ancient animal species and human specimens suggested an overrepresentation of purines (adenine and guanine) next to 5' ends (Briggs *et al*, 2007), and in particular guanine was revealed to be more represented than adenine (Orlando *et al*, 2011). Probably, this finding is due to a resonance structure in guanine that could reduce the activation energy required to break the N-glycosil bond (Overballe-Petersen *et al*, 2012).

Recent methods used in library preparation permitted to explore both 5' and 3' ends, confirming that purines, especially guanine, are overrepresented in the proximity of both ends in DNA strands obtained from ancient samples (Meyer *et al*, 2012).

Blocking lesions

As fragment breaks, other DNA modifications could block the elongation of DNA strand operated by DNA polymerases. Among these, nucleotide modifications and cross-links between DNA strands, between different DNA fragments or between DNA and other molecules.

Nucleotide modifications are originated by oxidative damages, caused by free radicals, such as peroxide radicals, hydrogen peroxide and hydroxy radicals, which create lesions on both bases and deoxyribose residues. In particular, the double bonds of pyrimidines and purines and the chemicals bonds of the deoxyribose residues are the target of oxidative attack, leading to base and sugar ring fragmentation (Pääbo *et al*, 2004; Dabney *et al*, 2013b).

Analyses performed on 11 samples of bones and soft tissues coming from different preservation conditions (permafrost and nonpermafrost) showed the presence of oxidative products of pyrimidines in all samples. In particular, samples with an high amount of oxidative damage, that blocks polymerase, gave no amplification results (Hoss *et al*, 1996).

The second type of blocking lesions are the cross-link between nucleobases. In a previous study (Pääbo, 1989), Pääbo detected intermolecular cross-links in ancient samples using electron microscopy. In addition, a following research (Poinar *et al*, 1998) identified products of the Maillard reaction, which caused cross-links between macromolecules, in coprolites (ancient fecal specimens) by gas chromatography-mass spectroscopy.

Miscoding lesions

While fragmentation and blocking lesions avoid the elongation of the DNA strand, on the other hand miscoding lesions allow DNA amplification which can result in the insertion of an incorrect base. This is due to the modification of the primary structure of a nucleobase, caused by hydrolytic deamination, originated by a DNA polymerase misreading.

The main target of this error is cytosine, which after the deamination event becomes uracil (Figure 29). In this case, during DNA replication, the DNA polymerase incorporates adenine instead of guanine and subsequently thymine instead of cytosine, leading to a G→A or C→T substitutions, depending on the strand considered (Pääbo *et al*, 2004; Dabney *et al*, 2013).
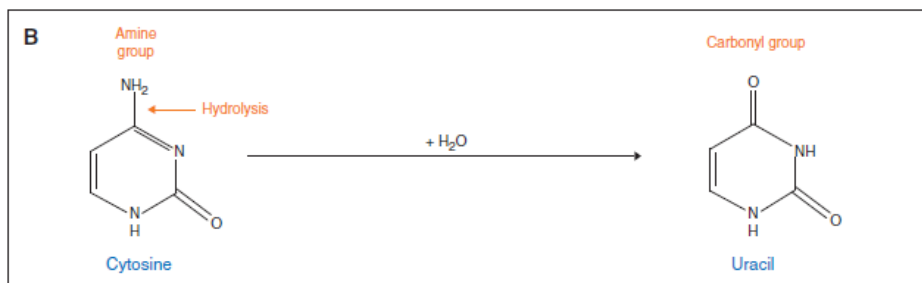
Figure 29. Deamination of cytosine to uracil (Dabney *et al*, 2013b).

Experimental data suggested that deamination is the main cause of miscoding lesions (Pääbo, 1989; Hansen et al, 2001) and, in particular, Hofreiter *et al* (2001a) indicated the C→T as the major source of substitutions occurring in ancient DNA molecules.

New generation sequencing analyses confirmed these results (Stiller *et al*, 2006; Gilbert *et al*, 2007) and, in particular, highlighted a different distribution of C→T and G→A substitutions in the DNA strands. In fact, C→T were elevated at the 5' end and decreased toward the 3' end, while G→A were limited at the 5' end and increased at 3' one (Figure 30), (Briggs *et al*, 2007).
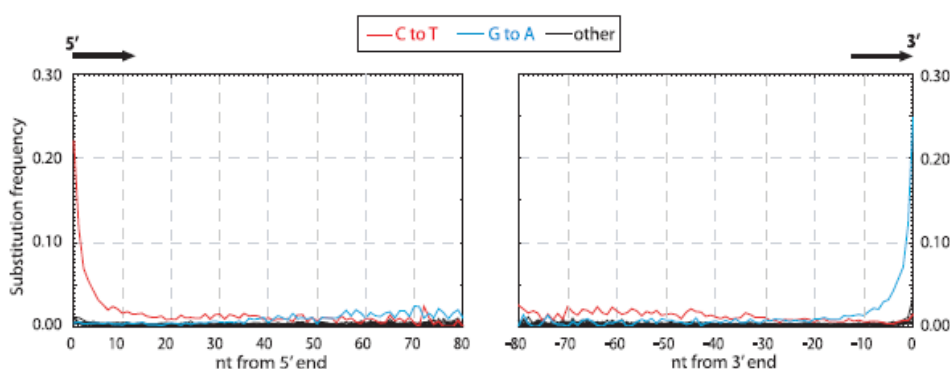
Figure 30. Distribution pattern of C→T and G→A substitutions in the DNA strand. C→T substitutions are very represented at the 5' end and decreased after the first ten nucleotides toward the 3' end, while the number of G→A substitutions are limited at the 5' end and raised rapidly ten nucleotides the 3'end (Briggs *et al*, 2007).

Moreover, considering sample preparation according to the new protocols used for sequencing, G→A substitutions were suggested to originate from C→T substitutions located on the 5' end during template replication. This implied that the major chemical modification in ancient DNA was represented by cytosine deamination (Briggs *et al*, 2007; Brotherton *et al*, 2007).

Further studies carried out using new technologies identified that C→T modifications took place in both ends with a similar frequency (Orlando *et al*,

2011; Meyer *et al*, 2012). Furthermore, taking into account the main localization of deamination damages at the strand ends and the deamination rate faster in single-strand than in double-strand molecules (Lindahl, 1993), it was suggested that cytosine deamination was related to the occurrence of single-strand breaks in ancient DNA (Briggs *et al*, 2007; Orlando *et al*, 2011).

1.3.2 The problem of exogenous DNA contamination

The development of PCR simplified the recovery of ancient DNA sequences, but at the same time made difficult the analyses amplifying also exogenous DNA coming from contamination during sampling or laboratory procedures (Willerslev and Cooper, 2005).

Contamination of ancient DNA by modern DNA is quite common (Kolman and Tuross, 2000; Cooper *et al*, 2001). In particular, in biological specimens from which a very limited amount of endogenous DNA is recovered, the amplified DNA could be originated by an exogenous contamination, erroneously confused as authentic (Pääbo, 1989).

The first results about ancient remains (Pääbo, 1985; Hauswirth *et al*, 1994) were subsequently considered as possible contaminants (Willerslev and Cooper, 2005), as well as other studies performed on old samples of plants (Golenberg *et al*,1990), dinosaurs (Woodward *et al*, 1994) and amber inclusions (DeSalle *et al*, 1992; Cano *et al*, 1993; Poinar *et al*, 1993) are now supposed to originate from contamination or artifacts (Willerslev and Copper, 2005).

Exogenous contamination could be introduced before working in aDNA laboratory or during aDNA analyses.

The first cause of contamination is represented by microorganisms which invade and degrade the cellular components during the *post mortem* interval.

During the excavation procedures, samples could be contaminated by operators, as well as in the course of cleaning/washing steps and morphological analyses of the skeletal remains. To this regards, Gilbert and co-workers (Gilbert *et al*, 2005; Gilbert *et al*, 2006) suggested that contamination could not be limited to the bone surface, but involved also the inner part, depending on the bone preservation and porosity, especially during handling and washing phases. Moreover, when stored in museums, samples could be handled by the personnel, treated or stored in unsuitable conditions promoting DNA contamination.

Finally, in an aDNA laboratory, exogenous contaminants could be introduced in every laboratory step, in particular during DNA extraction and amplification. In this case, not only the operator could be the carrier of the exogenous DNA in the sample, but also amplified products disseminated in reagents and on the surface of laboratory equipment may represent another important source of contamination (Caramelli, 2009).

In addition, Willerslev and Copper (2005) suggested to consider reagents and tools coming from manufacturers as possibly contaminated by human and

41

microbial DNA; thereby, specific decontamination procedures of such material are strongly suggested before use.

Contamination is a relatively minor problem in modern DNA laboratories, where the consistent amount of DNA which can be extracted from the samples (for example cultured cells) is able to overlay the low amount of contaminants. The situation is different in the ancient DNA field where, on the opposite, the very limited amount of DNA recovered from ancient specimens could be masked by the modern contamination.

1.3.3 How to study ancient DNA

Considering the low amount of endogenous DNA, the state of preservation, the molecular damage and modern contamination, the study of DNA extracted from ancient samples is a difficult task and requires specific protocols and guidelines in order to achieve reliable results.

In the last years, ancient DNA researchers improved their skills in order to find out new methods and techniques able to overcome the challenges of this analysis.

In the following paragraphs are summarised the solutions achieved which consisted in:
1.3.3.1 the selection of the sample;
1.3.3.2 going beyond the problem of contamination;
1.3.3.3 methodological approaches in aDNA studies;
1.3.3.4 authentication of ancient DNA.

1.3.3.1 The selection of the sample

The selection of the sample is a relevant step of the analytical strategy in aDNA studies. Many different studies have proved that bones and teeth are good sources of ancient DNA molecules (Key *et al*, 2017). However, taking into account the *post mortem* degradation of DNA and the influence of environmental conditions on the DNA survival, such as temperature, pH, oxygen availability, exposure to water and thermal history of a sample (Pinhasi *et al*, 2015), it is important to determine the most suitable bone/tooth district and portion for ancient DNA analyses.

In literature several studies relate different methods (amino acid racemisation, crystallinity index analysis, collagen content, bone histology, thermal age calculation and cytosine to uracil deamination patterns) to the survival of DNA in bone (Poinar *et al*, 1996; Götherström *et al*, 2002; Haynes *et al*, 2002; Smith *et al*, 2003; Collins *et al*, 2009; Schwarz *et al*, 2009).

Unfortunately the resulting data are ambiguous and no reliable methods are nowadays suggested for a relevant sample screening (Gilbert *et al*, 2005).

Bone samples

In cases in which there is the possibility of choosing which bone should be submitted to the genetic analysis, it is important to know which one could give the best result.

As already described, in all bones two components are identified: the thicken structure, that characterizes the outer part of bone shafts and the external surface of bones, called compact or cortical bone, and the porous, lightweight structure, arranged in thin spicules, called spongy or trabecular bone. Both the components share the same molecular and cellular structures (White and Folkens, 2005).

It is clear that long, flat and short bones differ from each other for the relationship between these two structures, but also in the same bone element we can find differences among the portions.

Most of the information about which skeletal element is the most suitable for genetic analyses come from the forensic field. Prinz and co-workers in the "Recommendation regarding the role of forensic genetics for disaster victim identification (DVI)" (2007) suggested the sampling of long bones of the legs for identification purposes in case of mass disasters.

According to the data gathered from studies on the genetic identification of war victims, dense, weight bearing bones, especially the long bones of the lower limbs, such as femur and tibia are considered as the districts showing the greatest probability of success for DNA analysis (Edson et al, 2004; Leney, 2006; Miloš et al, 2007; Mundorff et al, 2009). In particular, Kaiser and colleagues (2008) pinpointed the compact bone, especially the middle third of the cross section, as the best region of the diaphysis of long bones with the high-quality DNA.

Later, Edson and colleagues (2009) identified the temporal bone of the skull as the cranial bone associated with the greatest success rate, similar to that of long bones, and suggested to sample petrous bone, the inner part of the temporal region. This particular bone sample seems to be more resistant to damage and more likely to remain intact also in not well-preserved skeletal remains.

Recently, Gamba *et al* (2014) and Rasmussen *et al* (2014) confirmed these results, proving the possibility to get high amounts of endogenous DNA from petrous bone in archaeological specimens. More specifically, Pinhasi *et al* (2015) indicated the dense part of the otic capsule as the best region for sampling and Sirak and colleagues (2017) developed a new method for reaching this part preventing the damage of the surrounding bone when the complete skull was retrieved.

Tooth sample

Teeth are fastened in the jawbones and this protect them from environmental conditions. In addition, their chemical and biological compositions make them an important source of DNA for ancient DNA studies (Higging and Austin, 2013).

Each tooth is characterized by two parts: the crown, exposed inside the mouth, and the root, located inside the jawbone. The pulp is contained in the inner portion

of the tooth, surrounded by dentine and enamel, in the crown, and cementum, in the root.

The pulp is highly vascularised and innervated and includes several cells, among which odontoblasts and fibroblasts are the most numerous. Differently, the dentine, composed by minerals, in the form of hydroxyapatite crystals, water and collagen type I, does not contain nucleated cell bodies. However, the odontoblasts localised in the pulp propagate their processes in the dentine tissue passing through it, inside dentinal tubules, from the inner to the outer surface.

The enamel, which surrounds the crown, is the hardest tissue of the human body. It is acellular and composed of 96% of inorganic material. On the other hand, cementum, characterized by organic and inorganic components, is characterised by cellular or acellular regions depending on the presence or absence of biological components in the root portion; specifically, the cellular cementum is located on the tip of the root (Gartner and Hiatt, 2002).

Taking into account tooth composition, we can argue that pulp and cementum are sources of nuclear DNA, while cementum and pulp, together with dentine are sources of mitochondrial DNA (Higging and Austin, 2013).

Even if tooth composition is the same for each dental type, molar teeth are the preferred samples for genetic analyses, due to the great pulp volume and root area (Higging and Austin, 2013). In particular, literature studies showed that between the two parts forming a tooth, the root supplied more DNA than crown, which is composed mainly by enamel (Gaytmenn and Sweet, 2003; Dobberstein *et al*, 2008; Higgins *et al*, 2011).

Moreover, during sampling, it is important to select intact samples, not affected by dental pathology, such as caries and periodontitis, or other taphonomical signs, which could decrease the amount of endogenous DNA and increase the fraction originated by contaminants (Higging and Austin, 2013).

DNA preservation in bone/tooth samples

To date, the biochemical processes that lie on DNA preservation on bone/tooth tissues are not completely understood (Damgaard *et al*, 2015).

Some studies proved the preservation of ancient DNA both associated with the mineral component, the hydroxyapatite crystals, and the collagen content of bone and tooth samples (Schwarz *et al*, 2009; Campos *et al*, 2012).

Korlević and colleagues (2015) reported in another research that DNA association with the mineral component results from the interaction of positive charge of the calcium ions and the negative charge of the phosphate groups of the DNA molecules (Lindahl, 1993). Moreover, hydroxyapatite is shown to inactivate nucleases (Brundin *et al*, 2013) and decrease the depurination events of the bound DNA compared to free DNA (Lindahl, 1993).

Less understood is the association DNA-collagen and how this interaction plays a role in DNA preservation in ancient specimens. Campos *et al* (2012) and Svintradze *et al* (2008) suggested a possible protective role of the DNA-collagen

complexes, which are generated in aqueous solution spontaneously, in *in vitro* experiments.

These findings supported the choice of an extraction method that involved the decalcification and lysis steps performed together, such as that published by Dabney *et al* (2013a), rather than the two steps carried out separately. In fact, in this last case usually the supernatant obtained after the decalcification of the bone sample is discarded and only the pellet (the collagen fraction) is carried on to the lysis step, losing obviously an unknown amount of genetic substrate.

In addition, considering the short length of aDNA fragments, methods characterized by few washing steps and core molecular chemistry for decalcification, lysis and purification phases are preferred (Key *et al*, 2017).

1.3.3.2 Going beyond the problem of contamination

As discussed previously, exogenous contamination could occur in at least three stages, namely during the *post mortem* interval, prior the handling in the aDNA facilities and inside aDNA laboratories.

Two main approaches could be carried out to overcome the problem of contamination, in particular the prevention of exogenous contamination before sampling and inside the aDNA facilities and the decontamination of ancient specimens before laboratory analyses.

To prevent modern contamination before sampling and handling in ancient DNA structures, it is important to use specific precautions during excavation procedures, sample cleaning and morphological analyses. Specifically, it is strongly suggested to wear protective clothing during sample recovery in order to reduce the possibility to supply modern DNA from the operators to the specimen. To this purpose, cooperation among geneticists, archaeologists and anthropologists is required to highlight the importance to apply specific operating procedures and to get as much information as possible about the burial environment and sample storage in order to identify possible contaminants during the following analyses.

In addition, specific precautions have to be taken when the samples are analysed in aDNA laboratories. To this purpose, guidelines and criteria of authenticity were developed to increase the ability to obtain reliable sequence data from ancient specimens.

To this aim, in 1989, Pääbo firstly defined three major points: a. the analysis of a control extract carried out in parallel with the ancient samples to control potential contamination in reagents and solutions; b. the analysis of more than one extract per sample, which have to provide the same sequence; c. the presence of an inverse correlation between amplification efficiency and size of the DNA fragment (Pääbo 1989; Pääbo *et al*, 2004).

Later, these guidelines were developed and extended (e.g. Lindahl, 1993; Handt *et al*, 1994; Cooper and Poinar, 2000; Hofreiter *et al*, 2000; Pääbo *et al*, 2004), as some aspects of contamination became obvious and technologies

improved. Among these, the spatial separation of ancient DNA facilities from modern laboratories, bleach and UV irradiation as decontamination procedures and replication of the experiments within a laboratory and, in particular cases, in a second laboratory were included (Green *et al*, 2009).

After the introduction of high-throughput sequencing, some criteria were revised, such as the independent replication of results within the same or a different aDNA facility, which became not practical due to time, costs and sample material limitations (Green *et al*, 2009).

Currently, ancient DNA laboratories must be isolated from other molecular biology labs and, in particular, sample handling, DNA extraction and PCR reaction set up procedures must be performed in facilities separated from post-PCR area. In addition, all the personnel working on aDNA has to operate from ancient to modern lab and not viceversa.

Moreover, particular care should be taken inside an ancient DNA facility, named "clean room". In fact, the access to this area should be allowed to personnel wearing specific protective clothing and each work space, including surface and devices, chemicals and consumables, should be decontaminated using sodium hypochlorite and UV irradiation (Pääbo *et* al, 2004; Willerslev and Cooper, 2005).

Furthermore, negative controls have to be set up to monitor potential contaminations within reagents and extraction and amplification reactions have to be prepared under laminar flow hoods provided with UV lamps for the decontamination process (Caramelli, 2009).

In addition, attention has to be given to the introduction of reagents and consumables in the clean room. In these cases, specific decontamination involving bleach treatment and UV irradiation has to be performed to prevent the carry-over of contamination from the outside to the inside.

More difficult is the decontamination procedure of bone/tooth samples.

Nowadays, different methods are suggested to remove exogenous contamination from skeletal specimens before pulverization and DNA extraction: 1) washing the surface of bone/tooth; 2) mechanical removal of few millimetres of the sample surface; 3) washing the surface with acid; 4) collecting bone/tooth powder from the interior of the specimen; 5) UV irradiation of the sample; 6) exposing bone/tooth to highly concentrated ethanol or different concentrations of bleach; 7) treatment of the surface with a phosphate buffer; 8) treatment of bone with Proteinases or 9) a combination of these (Kemp and Smith, 2005; Korlević *et al*, 2015; Boessenkool *et al*, 2017).

All these methods assumed that contamination was located on the sample surface, but it was shown that after bleach and UV treatments modern contaminants were not removed completely. This is probably caused by bone and tooth porosity, which represents the access point for contamination (Gilbert *et al*, 2005, Rizzi *et al*, 2012).

Furthermore, in the last years, a pre-digestion method was proposed (Orlando *et al*, 2011; Damgaard *et al*, 2015; Gamba *et al*, 2016). Based on the hypothesis that contaminant DNA is located on the surface and endogenous DNA is more

protected in bone's microniches, treating for a short time bone powder with a digestion buffer removes exogenous DNA and enriches the following DNA extract for the endogenous content (Damgaard *et al*, 2015). Later, Boessenkool and co-workers (2017) indicated that a combination of bleach treatment and predigestion of bone powder improved the recovery of endogenous DNA.

Among all the methods, generally surface removal, bleach treatment and/or UV exposition are the most used, as reported by Kemp and Smith (2005) and in many other papers (Lazaridis *et al*, 2014; Fehren-Schmitz *et al*, 2015; Heupink *et al*, 2016; Llamas *et al*, 2016; Valverde *et al*, 2016; Posth *et al*, 2017; Schuenemann *et al*, 2017; Suppersberger Hamre *et al*, 2017).

1.3.3.3 Methodological approaches in aDNA studies

In the pre-NGS era, the conventional methodology applied to ancient DNA studies was PCR and Sanger sequencing, as main tools. Even though Sanger sequencing was considered the gold standard, it was characterized by a low-throughput which resulted expensive in large-scale sequencing. In addition, the pre-sequencing procedures of amplification and bacterial cloning were time consuming and showed a low efficiency (Rizzi *et al*, 2012).

Genetic analysis of ancient specimens is, in general, a difficult task mainly for the features of the DNA that can be recovered. In fact, ancient DNA is a mixture of endogenous DNA from the sample (usually human), bacterial, fungal and environmental DNA introduced during the deposition, and eventually DNA contaminants, coming from sample collection and manipulation. In addition, it is characterized by a reduced amount of genetic substrate showing a low molecular weight and chemical modifications in the primary structure which could lead to blocking and miscoding lesions (Stoneking and Krause, 2011).

To this purpose, in the following years protocols to improve sequencing methods and quality data were developed.

In particular, treatment of ancient samples with N-phenacylthiazolium bromide (N-PTB) has been proved to break the products of the Maillard reaction, which causes cross-links between macromolecules as DNA and proteins (Dabney *et al*, 2013b). Through this reaction it was possible to get genetic information from coprolites (Dabney *et al*, 2013b), even if some following studies highlighted a reduced success after the treatment to improve the amplification yield (Rohland and Hofreiter, 2007; Binlanden and Willerslev, 2010).

Furthermore, to reduce the number of misleading results caused by cytosine deamination, a pre-treatment with uracil-N-glycosylase (UDG) and *Escherichia coli* endonuclease VIII was suggested (Briggs *et al*, 2010; Rohland *et al*, 2015). The reaction catalyses the removal of uracil bases and the creation of an abasic site, which is cleaved on both sides by Endonuclease VIII, improving sequencing quality without causing a reduction in DNA recovery (Dabney *et al*, 2013b).

However, the next step in ancient DNA analysis was reached when next generation sequencing (NGS) technologies were developed. This new approach to DNA sequencing allowed to increase the number of bases that could be sequenced per run, thus cutting time and costs (Rizzi *et al*, 2012).

Next Generation Sequencing technology

Several NGS platforms are now available from different companies, among which NextSeq and HiSeq Series by Illumina and Ion PGM and Ion S5 systems by ThermoFisher Scientific are the most popular.

The NGS approach requires that the samples have to be prepared for sequencing through three main steps: library preparation, library amplification and sequencing.

Library preparation consists in ligation at both ends of DNA adapters, short oligonucleotides of known sequence, necessary for the subsequent amplification. To this purpose, 5' and 3' ends are converted in blunt ends using DNA polymerase and polynucleotide kinase, creating polished and phosphorylated extremities amenable for adapter ligation (Rizzi *et al*, 2012), (Figure 31).

Furthermore, during this step index sequences are added. The different index combination for each sample is able to identify the sample source of each sequence after sequencing and during sample analysis.
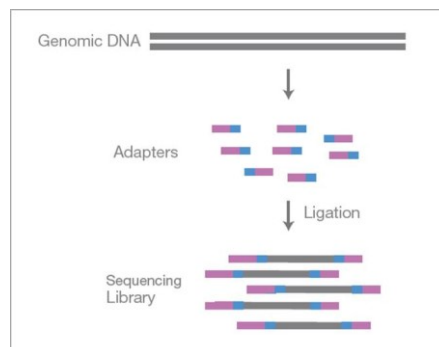
Figure 31. Library preparation of DNA fragments: amenable extremities are created through blunt end repair, generating polished and phosphorylated ends. Adapters sequences, composed by universal adapters (in violet) and indexes (in blue), are subsequently ligated to each fragment end to obtain double-stranded libraries for sequencing (modified from www.illumina.com).

Library amplification is carried out using primers complementary to a portion of the adapter sequence. It is performed by two different procedures depending on the platform used: emulsion PCR for Ion Torrent technologies and bridge amplification for Illumina systems.

In emulsion PCR single-stranded DNA fragments are bounded to a bead and amplified within a water-in-oil droplet with PCR reagents (Figure 32_1). In the

Illumina systems, amplification is performed through bridge amplification in a glass slide, named flow cell, containing probes complementary to adapters (Figure 32_2).
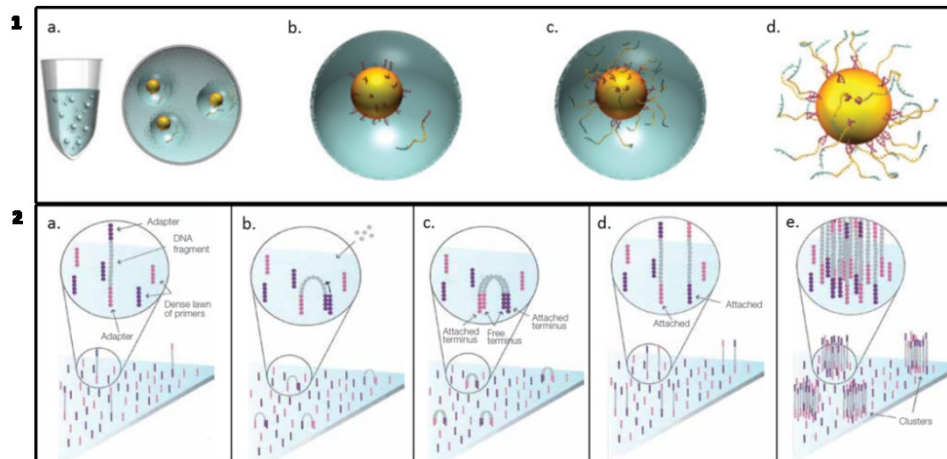
Figure 32. Library amplification. 1. Emulsion PCR adopted in Ion Torrent technologies: a. annealing of single-stranded DNA libraries to DNA capture beads; b. beads and PCR reagents in a water-in-oil droplet; c. clonal amplification within the water-in-oil droplet; d. breakage of the water-in-oil droplet and enrichment for DNA positive beads (modified from Mardis, 2008). 2. Bridge amplification used by Illumina systems: a. ligation of single-stranded DNA libraries to probes complementary to the adapter sequence on the flow cell; b. bridge amplification: bending of DNA fragment and ligation to the complementary probe on the flow cell and start of the amplification; c. enzymes incorporate nucleotides to create double-stranded fragments; d. denaturation of the double-stranded molecules, generating single-stranded fragments and restarting of the bridge amplification; e. generation of dense clusters on the flow cell, each one derived by a single DNA fragment (modified from www.illumina.com).

The last step involves sequencing. Ion Torrent sequencing technology is based on the change of pH of the solution during nucleotide incorporation. In particular, when a base is added and incorporated, an hydrogen ion is released. The charge of this ion modifies the pH of the solution, which will be detected by the ion sensor of the system (Figure 33) (www.thermofisher.com).
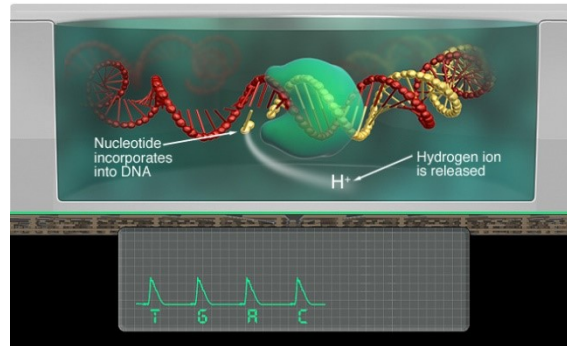
Figure 33. Ion Torrent sequencing technology: when a nucleotide is incorporated in the strand, an hydrogen ion is releasd and the change of pH of the solution records by the ion sensor of the system (www.thermofisher.com).

Differently, Illumina systems use a sequencing-by-synthesis (SBS) technology, using reversible terminator-bound dNTPs, that are modified nucleotides. After the incorporation on the DNA strand, the single base is excited by a laser and the generated signal detected by a CCD camera moving along the flow cell (Figure 34) (www.illumina.it).
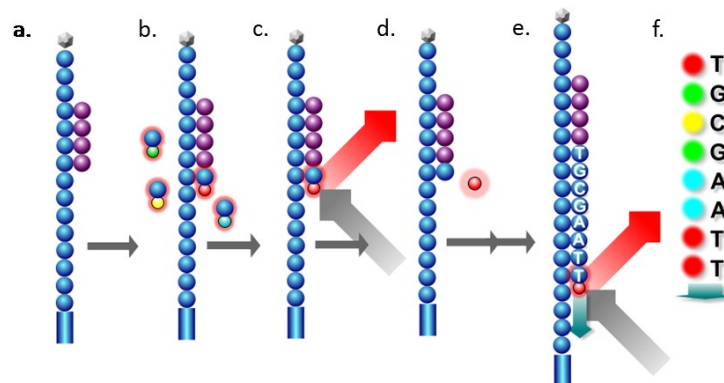
Figure 34. Sequencing-by-synthesis (SBS) system adopted by Illumina: a. ligation of the sequencing primer to the DNA fragment; b. addition of fluorescent labelled ddNTPs and base incorporation; c. excitation operated by the laser and light emission captured by the CCD camera; d. cleavage of the fluorophore; e. following sequencing cycles; f. record of the light emission by the software.

Sequencing approaches in ancient DNA fields

After library preparation, samples can be studied using two main different strategies, depending on the purpose of the research and the availability of the material: shotgun sequencing and DNA capture.

Shotgun sequencing

Shotgun sequencing is performed when the aim of the research is to identify all the species present in the DNA extract obtained from an ancient specimen. In this case no *a priori* selection has to be carried out.

Considered the type of sample analysed, the quantity of endogenous DNA could be low, also due to microorganisms contamination.

To determine the species to which a sequence belongs, the data are compared to sequence databases. In addition, to recognize specifically the recovered fragments, samples have to be sequenced deeply in order to obtain many copies from each DNA region (Rizzi *et al*, 2012).

DNA capture

The DNA capture approach allows to isolate the sequences of interest using specific probes, thus enriching the DNA extract only in those fragments which have been selected for the study.

To this aim, the primer extension capture (PEC) method uses biotinylated primers to capture the regions of interest. After the binding with the target DNA, the complementary strand is synthetized by a DNA polymerase. Subsequently, using streptavidin-coated beads the target molecules are isolated and eluted during a melting step. The recovered fragments are finally amplified and sequenced (Briggs *et al*, 2009).

To this purpose, Maricic and co-workers (2010) suggested a protocol to create PEC probes by fragmenting and biotin labelling modern mitochondrial DNA obtained through long PCR amplifications.

Other available methods employ biotinylated probes longer than a primer to capture the sequences of interest. Differently to PEC, in this case no extension of the probes operated by DNA Polymerases is required, thus allowing the recovery of the real information about miscoding lesions at the fragment ends (Rizzi *et al*, 2012).

To this aim, two different approaches are suggested: the first one is called "in-solid capture" and the second one "in-liquid capture". According to these methods, synthetized or amplified biotinylated probes are bound to the surface of an array (in-solid capture) or are dispersed in solution (in-liquid capture). The target regions bind the bait and the non-hybridized material is washed away. Subsequently, captured DNA is eluted and sequenced (Stoneking and Krause, 2011).

Ancient DNA markers

At the beginning of aDNA analyses, most of the studies were focused on mitochondrial DNA. Then, following the progress in high-throughput sequencing additional DNA sequences were identified and studies involving nuclear DNA, as well as pathogen and microbial genomes, were targeted.

Mitochondrial DNA

Mitochondrial DNA is a circular, double-stranded molecule, composed by 16,569 base pairs (bp). It is localised inside mitochondria, cellular organelles surrounded by two membranes and characterized by an independent genome with own replication, transcription and transduction processes. Because every cell may contain several mitochondria, each one with a variable number of mitochondrial genomes, a great number of mtDNA copies can be obtained from the analysis of a biological sample.

The first complete sequence, named *Cambridge Reference Sequence* (CRS), was published in 1981 by Anderson and colleagues (Anderson *et al*, 1981). Few errors were identified in the following resequencing of this genome by Andrews and colleagues (1999), which defined the current reference sequenced named *revised Cambridge Reference Sequence* (rCRS).

The two strands, named H (heavy) and L (light), have different base composition, guanine-rich, the H-strand, and cytosine-rich, the L-one.

It contains 37 genes, all of them important for mitochondria functions. In particular, thirteen codify enzymes involved in the oxidative-phosphorilation system, twenty-two encode for transfer RNAs (tRNA) and the remaining two for ribosomal RNAs, involved in the production of mitochondrial proteins (Anderson *et al*, 1981; Wallace, 1994; DiMauro and Schon, 2003).

In addition, non-coding regions are identified distributed within the genome. Among these, the most important is the control region or mtDNA displacement loop (D-loop), localized between the nucleotide positions 16,024 to 576, implicated in the regulation of mtDNA replication and transcription. It is constituted by three regions, called hypervariable sequence I, hypervariable sequence II and hypervariable sequence III (HVI, HVII and HVIII) (Anderson *et al*, 1981; Lutz *et al*, 1997; Lutz *et al*, 1998).

One of the peculiarity of mitochondrial genome is the lack of recombination and the transmission in a non-mendelian fashion, along the maternal lineage, without paternal contribution. Moreover, it is characterized, in general, by a mutation rate higher than nuclear DNA, which reaches ten-fold greater in the control region (Brown *et al*, 1979; Pakendorf and Stoneking, 2005; Howell *et al*, 2007; van Oven and Kayser, 2009).

These characteristics make mitochondrial DNA very informative for evolutionary studies.

In fact, the variation of the sequence is only due to the accumulation of mutations inherited by maternal lineage, generating over the time monophyletic units named haplogroups. Haplogroup is a group of mtDNA mutations derived from a common female ancestor (Torroni *et al*, 1993).

The mutation rate is such that mtDNA genome records the genealogical history of an individual and also the migration history of the women which transmitted it through generations. Taking into account that the molecular differentiation process is quite fast and took place primarily during and after human

movements into different region of the world, subsets of mtDNA variation could be limited to specific geographic areas and populations (Torroni *et al*, 2006).

These features make mtDNA a useful and very informative tool to trace the history and migration of female ancestor and to identify where and when such variation occurred thus trying to build the dispersal of Homo sapiens.

Nuclear DNA

Human nuclear genome contains 3.2 billion base pairs, assembled in 46 chromosomes, among which 44 autosomes and two sex chromosomes, X and Y.

The majority of adult human cells, named somatic cells, are diploid and have two copies of each autosome and two sex chromosomes (XX or XY). Gametes, differently, are haploid and are characterized by 23 chromosomes, one for each autosome and one sex chromosome, X or Y.

It is composed by genes and gene related sequences, such as introns, gene fragments and pseudogenes, and intergenic DNA, localized among genes and with no known function. Among these there are genome-wide or interspersed repeats (SINEs, LINEs, LTRs and transposons) and tandemly repeated DNA.

Nuclear DNA is inherited by both parents and is submitted to recombination during meiosis in the process of gamete formation.

One of the targets studied in the ancient DNA field are Single Nucleotide Polymorphisms (SNPs), which represent the most widespread genetic variations observed between individuals. A SNP is a difference in a nucleotide position in a specific locus of the genome, caused by mutational events and transmitted to the descendants. The frequency of each allele varies in different populations depending on selection, genetic drift and migration. Generally, most of them are biallelic, that is two variants exist at that locus in a particular population. In some cases, specifically when new mutations are introduced in a specific site, multiple alleles can be present (Brown, 2008).

Y chromosome

Y-chromosome is one of the smallest chromosomes of the human genome, characterised by 60 Mb of length and with an important role in sex determination. It is composed by two parts: the pseudo-autosomal (PAR) and the Y-specific regions. The pseudo-autosomal region, which amounts to nearly 3 Mb of the total length, is the only region that undergo to recombination with the X chromosome during meiosis; the Y-specific region, known as Male Specific region of Y chromosome (MSY) or Non-Recombining region of Y chromosome (NRY), instead, is transmitted without recombination along the paternal lineage. Thereby, the escaping from recombinational events and its paternal inheritance make Y chromosome similar to mitochondrial DNA and thus very informative for phylogenetic studies and evolutionary process along the male lineage (Jobling and Tyler-Smith, 2003). Two types of polymorphisms are used in genetic studies: Single Nucleotide Polymorphisms (SNPs) and Short Tandem Repeats (STRs),

where the first are more numerous than the second ones (Calafell and Larmuseau, 2017).

The combination of allelic states of markers along a chromosome is known as haplotype. Because no recombination occurs, haplotypes are transmitted intact from a generation to another; the only change is characterised by mutational events which take place through generations. Mutations accumulated over time generated, as for mitochondrial genome, monophyletic units called haplogroups, characterised by the same mutations, in the same positions and order (Jobling and Tyler-Smith, 2003). Haplogroups that share some alleles prove their origin from a common ancestor, while different alleles observed in two haplogroups demonstrate that they diverged over time by the accumulation of different mutations.

These features make Y chromosome, as mitochondrial DNA, an interesting tool to trace the history and the migration events of populations during the past.

1.3.3.4 Authentication of ancient DNA

The possibility to discriminate endogenous DNA from present-day human contamination is more difficult when ancient anatomically modern human samples (that are humans with skeletons similar to the present-day humans, Reich *et al*, 2010) are analysed compared to archaic hominins specimens, such as Neanderthals. In fact, present-DNA human contamination is much closer from a genetic point of view, thereby more similar, to ancient anatomically modern humans than archaic individuals. Although the compliance with specific recommendations can prevent or, at least, reduce contamination, ancient DNA sequences have to be analysed in order to confirm their authenticity (Key *et al*, 2017).

To this regards, the introduction of NGS technology in aDNA studies had two main advantages. First, the possibility to recover very short molecules. This was unfeasible by the conventional direct PCR approach, which required a portion of the preserved fragment as priming site. Second, the opportunity to sequence the entire molecule, thus identifying the C$\rightarrow$T and G$\rightarrow$A changes at the fragments ends already described before (Stoneking and Krause, 2011; Krause and Pääbo, 2016). This generates the peculiar damage pattern typical of ancient DNA, not present in modern contamination (Krause *et al*, 2010).

Therefore, fragmentation and damage patterns could be used as quantitative measures to identify endogenous DNA. In particular, the damage pattern could be adopted to take off undamaged molecules and enrichen the data set of endogenous DNA (Key *et al*, 2017).

However, damage pattern is not sufficient for aDNA authentication, because samples could be contaminated with ancient DNA in the burial environment, by past manipulation or cross-contamination in aDNA facilities. To this purpose, two additional methods are suggested.

The first one evaluates the percentage of mitochondrial DNA (mtDNA) sequences coming from a different source. In this case, following the mitochondrial

genome sequence has been obtained, mtDNA contamination is identified by comparing each position of each fragment to the obtained sequence and to a dataset of mitochondrial genomes, characterizing potential modern contaminants (Fu *et al*, 2013b; Renaud *et al*, 2015).

The second method uses the genetic sex determination. Two approaches are suggested: **a.** the calculation of the ratio between the sequences mapped on the Y chromosome and those mapped on both sex chromosomes in order to compute a confidential interval used to distinguish male and female individuals; in this case, to get solid results a number of human endogenous sequences greater than 10,000 is needed (Skoglund *et al*, 2013). **b.** the evaluation of the X ratio, computing the mean coverage on the X chromosome, normalized by the mean coverage on the autosomes. In such case, an X ratio about equal to one is indicative for female individuals, having two X chromosomes, while a coverage equivalent to half of the autosomes coverage is predictive for males, which carry only one X chromosome (Mittnik *et al*, 2016). In both cases, results different from the expected suggest possible female contamination in males or differently male contamination in female (Key *et al*, 2017).

Additional methods take into account the rate of heterozygous loci on the haploid X chromosome in males, the autosomal contamination estimate or plot genomic data to identify possible changes suggestive of modern contamination (Key *et al*, 2017).

2. AIM OF THE RESEARCH

In the last thirty years, the archaeological and anthropological analyses carried out on the Milan skeletal population revealed over time the great potential of this archaeological collection, supplying an interesting portrait of the demographic conditions and biological evolution of people living in different historical periods. Considering the intriguing data and the improvement of technologies for ancient DNA studies, the Milan project would enrichen the profile of these populations also in terms of genetic evolution, origin and variations, identifying possible parental relationships between individuals and genetic relations with people living in border areas.

The present research project arises as a collaboration between three laboratories – Laboratorio di Antropologia ed Odontologia Forense (LABANOF), Dipartimento di Scienze Biomediche per la Salute (Università di Milano); Laboratorio di Genetica Forense, Dipartimento di Sanità Pubblica, Medicina Sperimentale e Forense, Unità di Medicina Legale (Università di Pavia); Laboratorio di Genetica Umana e Animale, Dipartimento di Biologia e Biotecnologie (Università di Pavia) – and represents an articulated pilot study carried out to investigate the Milan skeletal population combining anthropology and genetics. In particular, this is the first attempt performed to assess the feasibility of reconstructing the history of Milan through DNA, becoming part of the small group of studies which would trace the evolution of a city over time through a multidisciplinary approach.

Two main topics are identified:

I.   the first analyses on the skeletal collection of Milan
II.   a further experience in the aDNA field

corresponding to two international experiences carried out at the Institute of Legal Medicine of Innsbruck (GMI), Medical University of Innsbruck (Austria), under the supervision of Prof. Walther Parson, and at the Max Planck Institute for the Science of Human History in Jena (Germany), under the supervision of Prof. Wolfgang Haak.

In the first part of the project, genetic investigations will be carried out to enrichen the anthropological data and archaeological information already achieved, focusing on genetic sex and geographic origin as a first approach. As no genetic studies were performed previously on the skeletal collection of Milan, these analyses are also performed to evaluate the possibility to recover DNA, and especially this information, from samples coming from different archaeological sites and historical periods.

To this purpose a subset of ten individuals from seven of the wider and most interesting sites, which were dated back from the Roman Imperial Age to the Contemporary Age, will be selected choosing among the most intriguing skeletal remains on the basis of anthropological and/or archaeological information. Among these, two individuals coming from Cimitero Maggiore are included as representative people of the contemporary local population. Information on the

biological profile was obtained from previous studies performed and published on the single populations; only pathology will be revisited.

Bone samples will be collected from each individual and studied at the Institute of Legal Medicine of Innsbruck (GMI), Medical University of Innsbruck (Austria), where specific methods and techniques dedicated to cases of ancient and/or severely degraded DNA will be applied. In particular, three main steps will characterise the study: DNA extraction, genetic sex typing and mitochondrial DNA analysis using the next generation sequencing technology.

In addition, to evaluate how the different temporal and environmental conditions of each archaeological context influenced bone tissue preservation, macroscopic and microscopic analysis will be carried out from the same bone elements submitted to genetic investigations. In particular, calcified and decalcified thin bone sections will be performed to investigate the microscopic appearance of the bone tissue. The information obtained are also compared to genetic data, and specifically DNA recovery, to evaluate a possible relation between macroscopic and microscopic bone appearance and DNA content, as well as to verify if these approaches could be considered reliable methods to predict DNA preservation in bones.

In the second part of the project, another international experience will be carried out at the Max Plank Institute for the Science of Human History in Jena (Germany) in order to get additional skills in ancient DNA analyses. The aim is to learn the ancient DNA techniques able to recover mitochondrial and nuclear DNA information from highly degraded samples, as well as bioinformatics tools essential in this field to recognize and analyse aDNA sequences. Great attention will also be given to laboratory setup and organization, as well as to sample processing.

3. PART 1:

The first analyses on the skeletal collection from Milan

3.1 MATERIALS AND METHODS

In the first part of the project, the research activity was focused on the study of skeletal remains recovered during the archaeological excavations in the Milan area using a multidisciplinary approach involving both anthropology and genetics. In particular, analyses were carried out to enrichen the anthropological and/or archaeological information, as well as to investigate bone tissue and DNA preservation in samples coming from different archaeological sites and belonging to different historical periods.

Anthropological data on the skeletons of all burial sites was recovered and the ten individuals on which to perform the pilot study were selected according to availability. Information on the biological profile was obtained from previous studies performed and published on the single populations. Then, samples were selected for anthropological macroscopic and microscopic bone preservation analyses. Finally, genetic studies were carried out.

Anthropological and histological analyses were performed at the Laboratorio di Antropologia e Odontologia Forense (LABANOF), Università di Milano; genetic studies were carried out at the Institute of Legal Medicine, Medical University of Innsbruck (Austria), under the supervision of Prof. Walther Parson.

3.1.1 The skeletal samples

Eight "intriguing" skeletal remains coming from six archaeological sites were selected in order to analyse samples from different contexts and historical periods. In addition two males from Cimitero Maggiore of Milan were added as representative people of the contemporary local population.

The archaeological information of the samples is summarised in Table 2.

From each individual a bone sample was collected for the histological and genetic analyses. In particular, a small fragment from the diaphysis of long bones (femur, tibia or humerus) was preferred when available.

ID	Individual	Archaeological site	Historical period	Peculiarity	Skeletal element
MI UC	MI UC VII 1992 US 5890	Milan, Catholic University	I-V AD	Occupational disease marks	Left femur
BG-BO-SC	BG-BO-SC T156 A	Bolgare	VII-IX AD	Member of local people	Right femur
MI PSA T49	MI PSA08 T49	Milan, Piazza S. Ambrogio	XIII-XV AD	African traits	Left femur
MI PSA T23	MI PSA07 T23	Milan, Piazza S. Ambrogio	XIII-XV AD	*Perimortem* traumas	Left femur
BO.K	BO.K.05 T8	Bormio, Piazza del Kuerc	XIV AD	Neoplastic disease	Left humerus
MI SAB	MI SAB 05 G4 IND S	Milan, Viale Sabotino	XVII AD	Pestis	Left tibia
MI CG Q/3	MI CG Q/3	Milan, Cà Granda	XVII-XVIII AD	Ankyloses	Left femur
MI CG O 2846	MI CG O2846	Milan, Cà Granda	XVII-XVIII AD	Syphilis	Occipital bone
MI CM 39	MI CM CASSA 39	Milan, Cimitero Maggiore	XX AD	Contemporary sample	Right femur
MI CM 33	MI CM CASSA 33	Milan, Cimitero Maggiore	XX AD	Contemporary sample	Right femur

Table 2. Archaeological information of the ten individuals analysed in the first part of this project.

3.1.2 Anthropological analyses

As mentioned before, information on the biological profile was acquired from previous published studies according to the methods briefly stated below, whereas the anthropological macroscopic and microscopic testing for preservation was performed ex novo from samples taken from the selected skeletons.

Biological profile

1. Sex diagnosis

The sex diagnosis was performed through the analysis of morphological parameters of specific districts, among which the most significant are the skull and pelvis, the last with a level of reliability better (98%) than the skull (80%).

In Table 3 the criteria considered are reported. When possible, all of which were evaluated, depending on the state of preservation of the bone elements, taking into account their reliability. To this regards, when the Phenice's triade could be examined, it was preferred to the other features visible on the pelvis (White and Folkens, 2005).

Skull	*Pelvis*
Glabella	Pelvic inlet
Eyebrow arches	Subpubic angle
Supraorbital ridge	Iliac crest
Zygomatic arch	Preauricular sulcus
Mastoid process	Phenice's triade
Nuchal crest	Sacrum: general appearance
Mental eminence	Sacrum: auricular surface
Gonial eversion	

Table 3. Morphological parameters evaluated on skull and pelvis for sex diagnosis.

In addition, the metric analysis of the diameter of radius, humerus and femur were calculated; in the females the measurements are lower (≤ 21 mm, ≤ 43 mm e ≤ 43 mm respectively) than males (≥ 23 mm, ≥ 47 mm e ≥ 48 mm respectively).

2. Age estimation

Age estimation is performed by different methods in adults and subadults.

In young subjects, age estimation was carried out evaluating teeth eruption and formation (Ubelaker 1999; AlQahtani *et al*, 2010), the length of diaphyses not

fused (Fazekas and Kosà, 1978; Scheuer & Black, 2000) and the fusion state of the epiphyses (Scheuer & Black, 2000; Schafer *et al*, 2008).

Differently, in adults, age at death was estimated evaluating those articulations, whose wear rate advances in a similar way in all adult individuals. This rate is related to the age at death of the subject (Cattaneo and Grandi, 2004). In particular, among the published methods, three were applied: the Suchey-Brooks' method, which considers the pubic symphysis (Brooks and Suchey, 1990), the Lovejoy's method, which analyses the auricular surface of the ilium (Lovejoy *et al*, 1985) and the Işcan's estimation, which takes into account the appearance of the sternal end of the fourth rib (Işcan *et al*, 1984; Işcan *et al*, 1985).

Unfortunately, not in all cases each method was used, due to the bad preservation of the bone district required.

3. *Ancestry*

Anthroposcopic and anthropometric methods were applied.

Anthroposcopic method describes non-metric features visible in the different skeletal districts; these characteristics are more detectable and frequent in a specific population rather than another. Especially, dental and cranial shape and morphology, as well as presence of particular non-metric characteristics, were evaluated. Differently, anthropometric method examines the metric features, such as bone lengths and diameters, and calculates specific indices in order to describe more accurately the bone shapes (Cattaneo and Grandi, 2004).

In addition, the software Fordisc 3.1 was applied. The software used cranial measurements and morphology to estimate the ethnic group of belonging, on the basis of a database of known population samples.

4. *Stature*

The height of the individuals was estimated through the application of formulae that take into account the maximum length of the long bones of the upper and lower limbs. The most used are those of Trotter and Gleser (1952, 1977).

5. *Pathology*

The skeletal remains in studying were subsequently analysed in order to identify possible marks useful in the building of the biological profile, such as pathological, stress markers or bone trauma that could give information about the life condition of a person. In fact, physical and pathological stress experienced in life can leave a mark on bones and be observed *post mortem* (Ortner, 2003). In the course of this research project pathology of the 10 skeletons was re-examined.

3.1.3 Macroscopic taphonomical analysis

The ten bone elements selected for the following histological and genetic investigations were first analysed to evaluate the state of preservation at a macroscopic point of view, applying the bone weathering staging proposed by Behrensmeyer (1978).

Behrensmeyer identified six stages, from 0 to 5, taking into account some parameters and characteristics, such as cracking, flaking, persistence of skin, ligaments, muscles, to describe the general appearance and preservation of the skeletal remains. Stage 0 was associated to bone samples very recent, showing greasiness, no cracks or flakes and soft tissues still present. The following phases involved the absence of soft tissue residuals and generation of cracking and flaking which become more consistent over time, associated to rough area and weathered compact bone. The last stage is characterised by bones very fragile and easily prone to break.

In the table below the five stages are described in-depth.

Stage	*Description*
Stage 0	The surface shows no signs of cracking or flaking. Usually, bone is still greasy, the medullary cavity contains tissues and skin, muscles and ligaments may cover part or all of the bone surface.
Stage 1	The bone shows cracking parallel to the fiber structure (longitudinal in long bones); the joint surfaces may show mosaic cracking of covering tissue and in the bone itself. Skin and other tissues may still be present.
Stage 2	Concentric thin layers of bone show flaking, usually associated with cracks. Flakes are long and thin, with one edge attached to the bone, at the beginning; deeper and more extensive flaking follows, until the majority of the external bone is gone. Crack edges are usually angular. Traces of ligaments, cartilage and skin may be still present.
Stage 3	The surface of compact bone is characterized by rough areas, resulting in a fibrous texture. In these areas, the external bone surface has been removed and gradually these zones extend to cover the entire bone surface. Weathering is not deeper than 1.0-1.5 mm and the fibers are still firmly attached to each other. In cross sections, crack edges appear rounded. Soft tissues rarely present.
Stage 4	The bone surface is fibrous and rough. Small and large splinters occur and may be loose enough to fall away from the bone. Weathering penetrates into inner cavities. Cracks are open and have splintered or rounded edges.
Stage 5	The bone is falling apart in situ, with large splinters lying around what remains of the whole, which is fragile and easily broken by moving. The spongy tissue is usually exposed, when present, and may outlast all traces of the former more compact, outer parts of the bones.

Table 4. Bone weathering stages as described by Behrensmeyer (1978).

3.1.4 Microscopic analysis

To evaluate bone preservation at the microscopic point of view, histological bone sections were prepared using two different approaches. Specifically, calcified and decalcified thin sections were prepared from each long bones underwent to the specified genetic analyses.

To this purpose, a 5 mm thick semi-circular section was cut transversely with an handsaw at the diaphysis of long bones and the same section divided in two pieces to get a quarter of a circular diaphysis section. The two fragments were submitted to the calcified and decalcified approaches respectively.

3.1.4.1 Calcified thin sections

Thick sections were levelled on a side using a Struers Dap-7 grinding wheel, with abrasive discs of grainsize between 180 and 4000, to obtain flat fragments. After the mounting of the smooth side on a glass slide with Pertex resin (Histolab Products AB), samples were polished again on the other side to get thin sections suitable for microscopic analysis.

3.1.4.2 Decalcified thin sections

To prepare decalcified thin sections, samples were fixed in formalin (v/v, pH 7-7.6), ratio formalin/sample 20:1, for 24 hours and, after rinsing for an hour using tap water, decalcified at room temperature with Histo-Decal, containing 14% hydrochloric acid (Histo-Line Laboratories, Milan).

The complete decalcification of the bone sample was checked manually, evaluating the flexibility of the specimen. In fact, a well decalcified bone tissue bends and compresses easily.

At the end of the decalcification process, a rinsing for 1-3 hours was performed to remove every residual of the decalcifying solution and samples were handled to prepare histological sections.

At first, decalcified bone fragments were dehydrated using an increasing alcohol series to remove water, submerging samples in the following alcohol concentrations: 70% ethanol for 24 hours, 80% ethanol for 24 hours, 90% ethanol for 24 hours, 95% ethanol for 24 hours, 95% ethanol for 12 hours, 100% ethanol for 24 hours, 100% ethanol for 12 hours.

Subsequently, the last absolute ethanol was substituted by a solvent of paraffin, xylene, in the clarification phase, through two passages of 24 hours. Subsequently, samples were embedded in fused Paraplast twice, for 24 and 12 hours respectively.

After the inclusion in paraffin blocks, each sample was cut in 5 μm sections and stained with Hematoxylin-Eosin.

Before staining, the inclusion medium (paraffin) was removed and slides re-hydrated using the following series: xylene for 10 minutes (x2), 50% xylol-50% absolute ethanol for 10 minutes, 100% ethanol for 5 minutes (x2), 95% ethanol for 5 minutes (x2), 90% ethanol for 5 minutes, 80% ethanol for 5 minutes, 70% ethanol for 5 minutes, distilled water for 10 minutes.

After hydrating, bone sections were stained firstly in Hematoxylin for 12 minutes and, after a rinsing using tap water for 12 minutes, in Eosin (1% in distilled water) for 6 minutes.

The obtained stained slides were subsequently dehydrated in 70% ethanol for 30 seconds, 80% ethanol for 30 seconds, 90% ethanol for 60 seconds, 95% ethanol for 60 seconds, 100% ethanol for 2-3 minutes and clarified through two passages in xylol for 10 and 5 minutes.

Finally, slides were mounted and visualised under light microscope.

Sample slides were observed and analysed using the polarizing optical microscope Zeiss Axio Scope.A1. Image acquisition was carried out at different magnifications with the camera system True Chrome Hd II and the software ISCapture version 3.6.7.

The appearance of the bone tissue was examined and classified according to the Oxford Histological Index described by Hedges and colleagues (1995), as already applied by Caruso *et al* (2017a). In particular, the state of preservation of each section was evaluated and a score was assigned considering the amount of destructive foci caused by microbial activity and the possibility to recognize bone structures such as osteons and bone lamellae. Six scores were identified by the authors, as reported in the table below.

Score	% of unaltered bone	Description
0	< 5	No original feature identifiable, other than Haversian canals
1	< 15	Small areas of well-preserved bone present
2	< 33	Clear lamellate structure preserved between destructive foci
3	> 67	Clear preservation of some osteocyte lacunae
4	> 85	Minor amount of destructive foci
5	> 95	Very well preserved, indistinguishable from fresh bone

Table 5. Summary of the scores used to evaluate bone tissue preservation as reported by Hedges *et al* (1995).

3.1.5 Genetic analyses

After the anthropological and macro/microscopic assays have been performed, the bone samples were submitted to DNA analysis. The task was to identify and evaluate the best genetic approach for recovering as much information possible from such ancient samples. Three main analytical steps were characterized in this study: DNA extraction, genetic sex typing through a new method able to get results also from degraded DNA and mitochondrial DNA analysis using the NGS technology.

3.1.5.1 Decontamination and pulverization

Before carrying out the extraction, each sample was submitted to a decontamination process to remove any external contamination potentially located on the surface.

Specifically, each bone fragment was submitted to a physical cleaning removing 1-2 mm of the external bone surface, followed by a chemical cleaning (three consecutive washing with 5% bleach, deionized water and absolute ethanol respectively) and UV irradiation for 30 minutes per side.

After the decontamination process, the bone samples were reduced to fine powder using the Mixer Mill MM 400 (Retsch), setting a vibrational frequency at 25 Hz.

For each one, 50 mg of bone powder were prepared for the following DNA extraction.

3.1.5.2 DNA extraction

DNA extraction was carried out as reported in Dabney *et al* (2013a), with some modifications.

The protocol provided an overnight extraction at 37°C adding to each sample 1 ml of extraction buffer composed by 0.45 M EDTA and 0.25 mg/ml Proteinase K.

The remaining bone powder was pelleted, the supernatant mixed to a binding buffer contained 5 M guanidine hydrochloride, 40% isopropanol, 0.05% Tween 20 and 90 mM sodium acetate pH 5.2 and the overall solution transferred to a specific binding apparatus in 50 ml Falcon tubes. Each tube included an extension Reservoir removed from the Zymo-Spin™ V column (Zymo Research), necessary to collect the huge amount of the binding solution, joined with a MinElute Column (MinElute PCR Purification kit, Qiagen).

After centrifugation, the column was removed from the reservoir and placed in a new collection tube. DNA bound to the silica membrane was purified through two consecutive washing steps with 750 µl of PE buffer from MinElute PCR Purification kit (Qiagen) and finally eluted using 25 µl EB buffer (MinElute PCR

Purification kit, Qiagen). The elution step was repeated to obtain a final volume of 50 μl.

The same bone samples were then submitted to a second extraction procedure following the same experimental conditions described above, except for the incubation time which was extended to two days at 37°C.

Extraction negative controls were performed for each batch of samples.

A 2μl aliquot of the extracted DNA were used to quantify both nuclear (60 bp autosomal probe) and mitochondrial DNA (143 bp and 60 bp probes) in the Applied Biosystems 7500 Fast Real-Time PCR System using an internal laboratory protocol which has not been yet published. In particular, for each sample duplicate tests were performed.

3.1.5.3 Genetic sex typing

Genetic sex typing was performed following the TriXY protocol as reported in Madel *et al* (2016).

In detail, three short sequences (< 50bp), localised one on the X chromosome (marker X44, chromosome position: 11,864,333-11,864,376) and two on the Y chromosome (markers Y44 and Y47, chromosome positions: 19,755,485-19,755,528 and 19,568,322-19,568,368 respectively), were amplified through real-time PCR in the Rotor Gene Q 5 plex HRM instrument (Qiagen) using the Type-it HRM PCR kit (Qiagen).

The amplification reaction included 1x HRM PCR Master Mix (Qiagen), 16 mM tetrapropylammonium chloride (TPrACl), 12 mM ammonium sulfate, 5 μg non-acetylated bovine serum albumin, 200 nM of the forward and reverse primers targeting the X44 and Y44 markers, 300 nM of the forward and reverse primers targeting the Y47 marker and up to 4 μl genomic DNA depending on sample concentration, in a total volume of 20 μl.

The PCR conditions involved a denaturation at 95°C for 5 minutes followed by 47 cycles at 95°C for 15 seconds, 60°C for 30 seconds, 72°C for 20 seconds.

PCR amplification was followed by High Resolution Melting (HRM) analysis evaluating the change of the fluorescence signal during increasing temperature steps (temperature range from 65°C to 90°C, increasing of 0.1°C for each step).

Data were analysed using the Rotor-Gene Q Series Software version 2.3.1 (Qiagen).

Following the recommendations by Madel and colleagues (2016) who described that TriXY could provide reliable sexing information for DNA input range between 100 ng and 12 pg, samples with a low amount of DNA (less than 3 pg/μl) were not submitted to the analysis.

For each sample, triplicate tests were performed using at least 12 pg of nuclear DNA (2-4 μl) from the overnight extracts.

Extraction negative controls were included in the analysis in order to verify the absence of exogenous contamination. Moreover, negative controls of the reaction were performed.

3.1.5.4 Mitochondrial DNA analysis

Mitochondrial genome studies were carried out applying two different approaches:
- Primer Extension Capture (PEC)
- Mito Tiling protocol.

In both cases the Ion Torrent technology was used and the workflow was based on four main steps: library and template preparation, sequencing and data analysis.

Moreover, to check whether exogenous contamination was present, negative extraction controls were submitted to the same steps of the samples.

Library preparation

Before sequencing, each sample was converted into libraries to allow the correct identification of the sequenced fragments in the downstream analyses. During library preparation, adapters and unique barcodes are added to each fragment end, as shown in Figure 35.
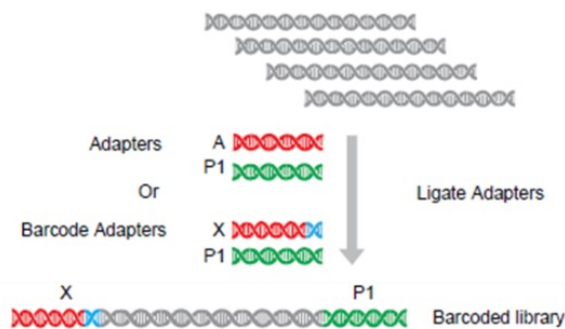
Figure 35. Library preparation: adapters and barcode adapters are added at each fragment end. Barcode adapters are specific and unique for each sample, a sort of tag that allows to define during the downstream analyses the origin of each sequenced fragment (modified from www.termofisher.com).

Two different protocols were adopted for the two approaches.
Below, the protocols and the specific procedures used for mtDNA analysis were described.

PEC protocol

The PEC protocol allows to enrichen samples in sequences of interest, by capturing and amplifying them. It is generally applied to highly degraded samples with a low amount of endogenous DNA, from which no satisfactory results have been obtained through the standard procedures.

For this reason, based on quantitation results, it was applied to the DNA extracts with the minor amount of DNA.

The core of the procedure lays in the isolation of the DNA fragments, previously converted into libraries, using human specific biotin-labelled primers, which are subsequently extended by DNA polymerase and captured by strepdavidin-coated magnetic beads (Figure 36).
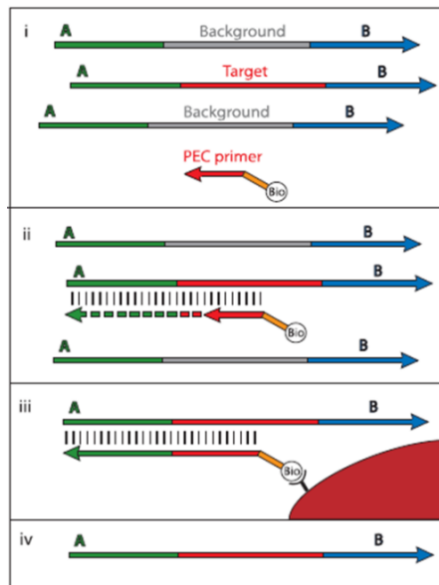
Figure 36. PEC protocol: i. biotin-labelled primers are added to DNA libraries; ii. DNA Polymerase performs primer elongation, generating double-stranded molecules; iii. not bound fragments are washed away through consecutive washing steps and the primer-target complexes captured by streptavidin-coated beads; iv. Finally, the target of interest is eluted (Briggs *et al*, 2009).

In this study library preparation and capture were carried out following the protocol published by Eduardoff and colleagues (2017).

Specifically, libraries were prepared using the IonXpress Fragment Library kit (ThermoFisher Scientific), as described by the manufacturer. Considering that DNA from ancient samples could be highly degraded and modified, no further fragmentation was performed and DNA extracts were immediately submitted to the end repair process in order to get blunt ends amenable for the following adapter and barcode ligation. DNA libraries were subsequently amplified for 10 cycles in a

reaction mix composed by DNA Polymerase and primer mixes from the same library kit.

Primer Extension Capture was performed using human specific mitochondrial DNA sequences, which covered the mtDNA control region and a region localised around position 8,342 of the same genome.

In the first step, the annealing of biotinylated primers with the corresponding complementary sequences and primer elongation were carried out. To this purpose, 25 µl of DNA library were added to 5 µl of 10x AmpliTaq Gold Buffer, 5 µl of $MgCl_2$ (25 mM), 2 µl BSA (2.5 mg/ml), 1 µl dNTPs (10mM), 5 µl of the PEC Primer mix (1 µM) and 1.2 µl AmpliTaq Gold DNA Polymerase (5U/ µl ThermoFisher Scientific), to a final reaction volume of 50 µl. After an incubation at 95°C for 12 minutes, 57°C for 1 minute, 72°C for 3 minutes, the reaction was blocked adding 10 µl 0.5 M EDTA to each sample and putting tubes on ice.

Before carrying on with primer-target complexes capture, 30 µl per sample of Dynabeads MyOne Streptavidin C1 (ThermoFisher Scientific) were washed three times with 2xBW Buffer and resuspended in 60 µl 2xBW Buffer. 60 µl of the washed Dynabeads were added to each sample and incubated for 30 minutes at room temperature to promote the interaction between biotin-labelled primers and streaptavidin-coated magnetic beads.

To purify the captured fragments, six washing steps were carried out using a magnet to pellet beads as follows: three washes with Wash1 buffer (1x BW Buffer, 0.1% SDS, 0.01% Tween 20), two washes with Wash2 buffer (2x SSC, 0.1% SDS, 0.01% Tween 20), one wash with 65°C heated Wash2 buffer, shaking samples mildly at 65°C for 2 minutes.

Finally, DNA was eluted adding 25 µl of low TE to the beads and keeping tubes 3 minutes at 95°C before transferring the supernatant to a new tube.

The last step of the PEC protocol involved the amplification of the captured fragments for 14 cycles in a reaction mix composed by DNA Polymerase and primer solutions from the library preparation kit. Amplified DNA was subsequently purified using Agencourt AMPure XP (Beckman Coulter) following manufacturers' instructions.

To check the success of the PEC reaction, amplified DNA was analysed in the 2100 Bioanalyzer (Agilent Technologies) using the Agilent High Sensitivity DNA kit (Agilent Technologies), according to manufacturers' protocol.

In agreement with the protocol, a second round of capture, capture purification and amplification was performed on the DNA samples as described before. The only change was in the annealing temperature of the biotinylated primers, which was set up to 59°C.

Before setting up the sequencing reaction, DNA was quantified through qPCR.

Mito Tiling protocol

Mito Tiling protocol allows to get information of the complete mitochondrial sequence from extremely degraded samples. To this purpose, two pools of primers are used to generate overlapping fragments, covering all the positions of the entire mitogenome.

Differently to PEC protocol, an initial amplification of each fragment was performed, followed by barcoded adapter ligation and purification of the obtained libraries.

The protocol has been developed in the Forensic Genetics lab of the Innsbruck University but not yet published. For this reason, it cannot be described in detail in this section as it has to be considered confidential.

Before proceeding with sample preparation for sequencing, DNA was quantified through qPCR.

Template preparation

Based on the platform used for sequencing, different methods are available to amplify libraries. To analyse mitochondrial DNA from PEC and Mito Tiling protocols, Ion Torrent technology was adopted.

Ion Torrent developed the Ion OneTouch 2 System (ThermoFisher Scientific) to carry out library amplification and enrichment for different Ion Torrent sequencing systems. It includes two modules, the Ion OneTouch 2 Instrument, which performs emulsion PCR to amplify samples using biotinylated primers, and the Ion OneTouch ES (Enrichment System), to capture amplified fragments by streptavidin-coated beads.

In this study the Ion PGM Hi-Q OT2 kit (ThermoFisher Scientific) was used for sample preparation, following manufacturers' instructions.

Ion OneTouch 2 Instrument

Ion OneTouch 2 Instrument carried out library amplification by emulsion PCR.

In emulsion PCR, amplification occurs within a micro-reactor, containing all the PCR reagents, including a bead to capture library DNA fragments and biotinylated primers. Ideally, each droplet encloses a bead with a single DNA fragment.

Amplification reaction involved the common denaturation, annealing and elongation steps to create clonal beads.

When PCR is completed, the emulsion is broken using a breaking solution and the solution centrifuged to recover empty, clonal or non-clonal beads, which are separated in the following enrichment phase.

To perform all these steps, the Ion OneTouch 2 includes three technologies:

- a reaction filter, which generates millions of micro-reactors, where amplification occurs, containing beads, library fragments and PCR reagents (Figure 37_1);
- an integrated thermal cycler and amplification plate able to allow thermal cycling of the micro-reactors (Figure 37_2);
- an integrated centrifuge to recover the amplified solution (Figure 37_3).

Instrument set up and sample preparation were performed as reported in the manufacturers' instructions.
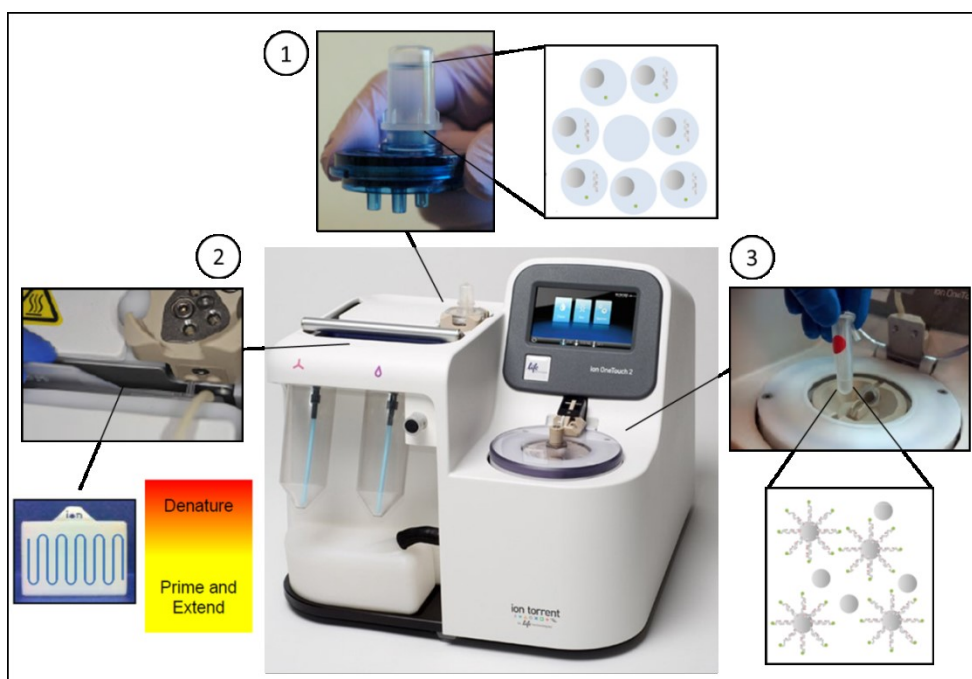
Figure 37. Ion OneTouch 2 Instrument: 1. the reaction filter, where beads are emulsified in micro-reactors, containing DNA fragment and PCR reagents; 2. the thermal cycler and amplification plate. Plate consists in wavy canals in which micro-reactors pass through. The route is such that on the top the high temperature denatures the fragments, while on the bottom a decrease of temperature allows annealing and elongation phases. 3. centrifuge with recovery tubes to recover, after the amplification process, empty, clonal or non-clonal beads (modified from www.thermofisher.com).
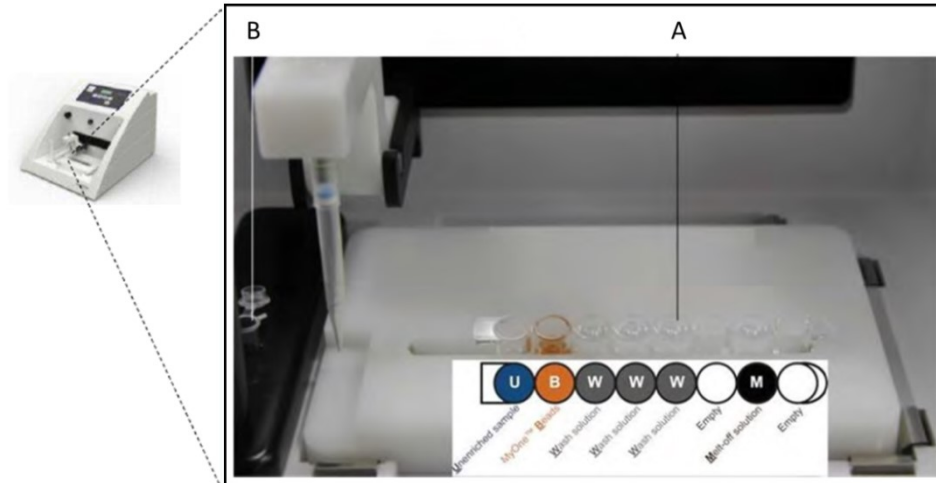
Ion OneTouch ES

After PCR, the recovered solution contained beads with amplified DNA on the surface or no DNA. To separate and recover beads carrying amplified DNA fragments, the Ion OneTouch Enrichment System was used. The instrument employed magnetic streptavidin-coated beads to capture clonal Ion Sphere Particles.

During the run, beads were added to the sample and, using a magnet, the supernatant was discarded. Following the same approach, three washing steps were carried out and, finally, clonal beads were recovered in a Melt-Off Solution to denature samples.

To neutralize the solution, the sample was transferred in a Neutralization Solution (Figure 38).


Figure 38. Ion OneTouch ES, particular of the instrument: A. 8-well strip used for sample enrichment (U: Unenriched sample; B: MyOne Beads; W: Wash Solution; M: Melt-Off Solution); B. Tip installed in the Tip Arm and 0.2 ml PCR tube containing the Neutralization Solution (modified from ThermoFisher Scientific).

Sequencing

The clonal beads recovered from enrichment were subsequently prepared using the Ion PGM Hi-Q Sequencing kit and sequenced in the Ion PGM Sequencer (ThermoFisher Scientific).

Ion Torrent technology sequenced DNA using a chip, containing millions of wells (Figure 39).
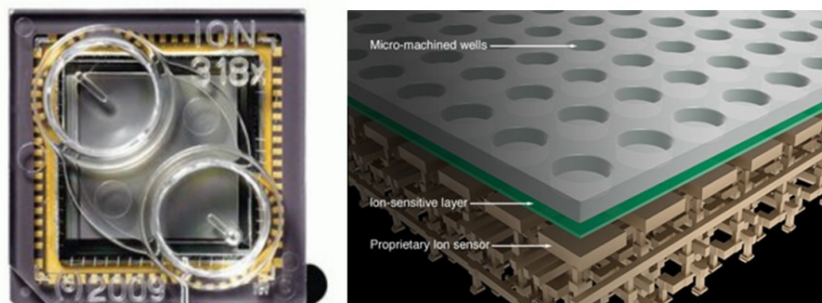
Figure 39. Ion Chip: on the left, Ion 318 Chip v2 used in this study; on the right, representation of the wells within the chip (modified from www.thermofisher.com).

When sample is loaded into the chip, clonal beads covered by millions of different amplified fragments flow across the chip, depositing into a well, which capture chemical information from sequencing and translate it into digital information or base calls. Specifically, the chip is floated with the four nucleotides and, if a nucleotide is incorporated into a fragment, an hydrogen ion (H^+) is released. This creates a change in the pH of the solution, which is measured by an ion sensor that converts it to a voltage. If two bases are incorporated the voltage is double. The voltage change is recorded, indicating that the nucleotide is incorporated and the base is called. The process is repeated for every incorporated nucleotide and verifies simultaneously in each well (www.thermofisher.com).

Sample preparation was performed as described in the Ion Torrent guide. Specifically, enriched clonal beads were mixed with the Control Ion Sphere Particles (ISPs) (a positive control to check the system), Sequencing Primers and polymerase and finally loaded in the Ion 318 Chip v2, following manufacturers' recommendations.

Run analyses were performed using the Torrent Suite Software version 5.0.5.

The created BAM files were subsequently loaded and aligned to the reference rCRS in the IGV software (version 2.3). Specific filters were applied for both library preparation protocols. In detail, the mapping quality threshold was set up to a predefined value and "duplicate and vendor failed reads" (that is alignments marked as duplicate reads and reads that failed the vendor quality check, respectively) were filtered in samples analysed with the mito Tiling approach; for the PEC protocol, the same settings were used, except the duplicate read filter, that was substituted by another filter, the secondary alignments filter, to show only the primary alignment. In fact, each reads could map to multiple positions; only one of the alignments is considered a primary alignment, while the others are the secondary ones.

Nevertheless mutations were manually checked taking into account some parameters (the coverage, presence of artifacts/deamination events and length).

Haplotypes were loaded on the EMPOP database (empop.online) for haplogroup attribution and sequences re-analysed looking for missing mutations. To confirm the results Phylotree database (www.phylotree.org) was checked.

3.2 RESULTS

3.2.1 Anthropological analyses

Information on the biological profile of each selected individual was obtained from previous studies performed and published on the single populations; only pathology will be revisited. In some cases some information could not be achieved due to the absence or the bad state of preservation of the bone district of interest.

Below, the anthropological data were reported describing each subject from the most ancient archaeological sites to the contemporary skeletal cemetery collection.

Milan, Catholic University

<u>MI UC</u>

A male individual (Figure 40a), 43-55 years old, 165-170 cm of height, of caucasoid ancestry (Sguazza, 2010).

Both humeri showed a maximum length lower than the expected for an adult male with these features (Figure 40b). In particular, the diaphysis were bowed and both heads displayed relevant bone proliferation and eburnation. In addition, osteoproliferation was observed in the glenoid fossa of the scapulae (Figure 40c). Different causes could be suggested in differential diagnosis: a. a serious trauma to the shoulders/arms; b. mucopolysaccharidosis, a metabolic disease (Ortner, 2003).

Bone remodelling owing to bone fractures was also observed in the diaphysis of the right ulna and in proximity of the sternal end of a left rib.

Osteoarthritis was observed in the vertebral column, from the VII to the XI dorsal vertebrae and from the I to the IV lumbar vertebrae, probably associated both to the age of the individual and repetitive stress of the distal trait of the column. Minor signs were recognized in ribs, pelvis and patellae.

In both calcanei the peroneal sulcus was identified, generally related to subjects who walk barefoot or in hard ground.

Finally, dental wear was present (Sguazza, 2010).

Figure 40. MI UC: a. the complete skeleton; b. comparison between the humerus of the individual, on the left, and an humerus of another subject: it is evident the small size of the first; c. humeri and scapulae: osteoproliferation is identifiable in the head of both humeri and in the glenoid fossa of the scapulae (Sguazza, 2010).

Bolgare

BG-BO-SC

A female individual, 42-87 years old and 158 cm of height.

Unfortunately, not the complete skeleton was retrieved, some bone elements were not present in the grave. The good state of preservation of the few recovered bones allowed to determine with certainty the sex from the pelvis, while age at death was estimated evaluating only the pubic symphysis. No information could be obtained from the skull due to the lack of the district of interest for the analyses (Figure 41) (La Ferla, 2005).

In both femurs exostosis was identified, suggesting a sitting position with extended legs, and inflammation at the flexor of the phalanges was observed as result of using thin objects tightened in the hand (Capasso *et al*, 1999; La Ferla, 2005).
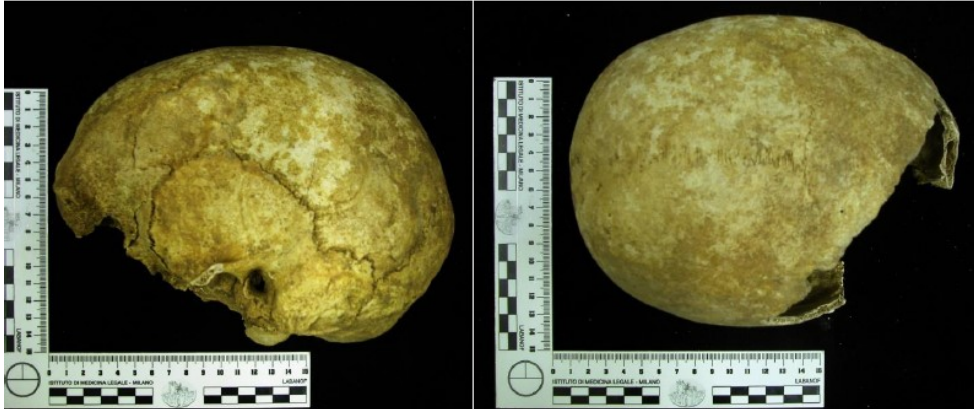
Figure 41. BG-BO-SC: the skull. The lack of specific districts of interest prevented ancestry estimation for this individual.

Milan, Piazza S. Ambrogio

<u>MI PSA T49</u>

A young man (17-22 years old), height 175 cm.

The morphological features observed on the skull and the metric analyses suggested an individual of north African origin (the only one in the entire *necropolis*) (Figure 42a).

The anthropological analyses revealed a blunt trauma, occurred probably 1-2 years before death, and hypoplasia, in both the superior canines, caused by a biological stress occurred during the childhood, hypothetically associated with a malnutrition status.

Moreover, typical occupational stress markers were observed on the clavicles, in the inferior medial surface, ascribable to an overloaded of the upper limbs, and on the right shoulder girdle; in addition, on the head of metatarsal bones an accessory articular facet was detected, probably associated to an hyperextension of the toes for consecutive periods (Figure 42b). All these markers pointed out a specific posture maybe to fix the feet to transfer heavy weights.

These features indicated the possible dedication of this man to hard activities, most likely correlated to the dragging and transferring of heavy weights, suggesting a working activity (Cattaneo *et al*, 2015a).
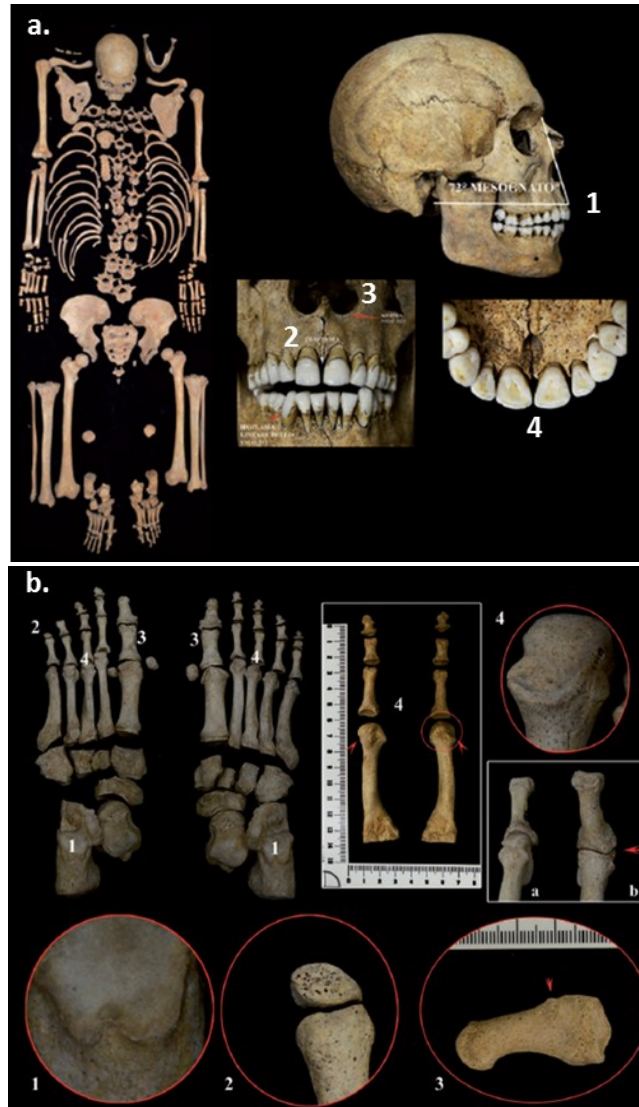
Figure 42. MI PSA T49: **a.** on the left, the skeletal remains of the individual; on the right, the skull, in lateral view, and two details of the teeth: 1. some prognathism; 2. presence of diastema, the space between the two superior central incisors; 3. absence of the nasal sill, the inferior edge of the nasal aperture; 4. detail of the palate. **b.** Occupational stress markers on the feet: 1. anomaly in the articular facet of the calcaneus; 2. a detail of the phalanxes of the V left toe with signs of bone reabsorption; 3. a detail of the normal view of the proximal phalanx of the I left toe with alteration caused by a prolonged and repeated hyperflexion of the feet toes; 4. III metatarsal bone with a detail of the accessory articular facet and particular of the joint with the proximal phalanx (4a. the normal joint; 4b. the joint observed in the individual). All these markers suggested a posture kept maybe to fix the feet to transfer heavy weights (Cattaneo *et al*, 2015a).

<u>MI PSA T23: a victim of torture</u>

A Caucasoid young man (18-20 years old), height 170 cm, with syndromic characteristics.

The position of the skeleton suggested that he was buried wrapped in a shroud. Moreover, the recovery of two iron buckles on the side of each femur indicated the possible presence of a belt used to clamp the tissue around the corpse (Figure 43). Radiocarbon analysis was performed, stating a dating between the 1290 and 1430.

**Figure 43.** MI PSA T23: the burial (De Vanna, 2015a).

The anthropological studies revealed several fractures characterized by *perimortem* features, probably associated with the cause of his death. In particular, the analyses identified blunt force trauma on both the forearms and legs, as well as sharp force trauma on the fourth lumbar vertebra and the occipital region of the skull (Figure 44b). The observations indicated a repeated traumatism, caused by different tools, occurred close to death and in a limited time lapse.

This scenario is compatible with the hypothesis of a torture, such as that of the wheel, in which the subject, tied to a wood structure (a wheel for example), was submitted to the fracture of the upper and lower limbs. Since the limbs fractures didn't caused death, the individual was stabbed to the thorax or the abdomen. In this case, the trauma observed on the skull could be occurred *post-mortem*, in a period close to death, maybe to disfigure the corpse (Cattaneo *et al*, 2015a).

The torture of the wheel was reserved to those individuals guilty of crimes against the community or for having killed the landlord's wife or son or, again, to criminals guilty of murders and finally to poisoners.

Moreover, the pathological analysis revealed anomalies on vertebrae and ribs, as well as wormian bones and bone thickening on the skull, supporting the possible presence of congenital diseases that probably altered his features. This was corroborated by the observation of a torsion of the upper lateral incisors (Figure 44a), suggesting the possibility that he was considered "different" by the community, maybe marginalized and subjected to discrimination (Cattaneo *et al*, 2015a).
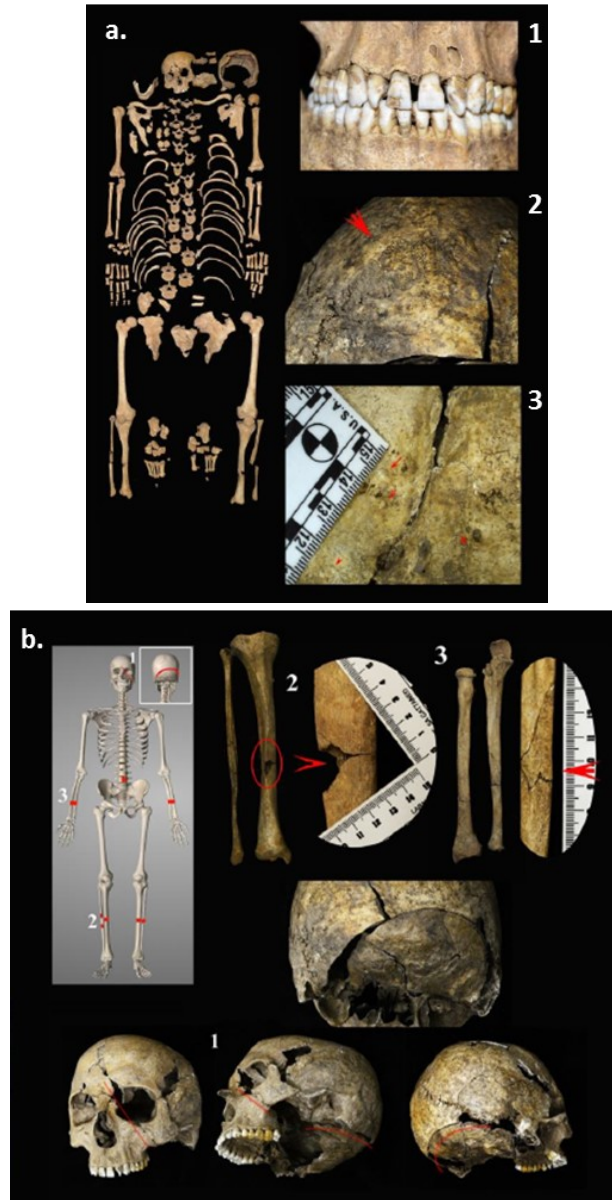
Figure 44. MI PSA T23: **a.** on the left, the skeletal remains of the individual; on the right, the skeletal anomalies observed on the skeletal remains: 1. a detail of the teeth showing the torsion of both upper lateral incisors; 2. a detail of the wormian bones observed on the lambdoid suture; 3. a detail of the frontal bone showing hyperostosis in the inner surface, usually associated to metabolic disorders. **b.** Localization of the *perimortem* trauma: 1. a detail of the sharp force trauma in the occipital region of the skull; 2. a detail of the blunt force trauma in the right tibia; 3. a detail of the blunt force trauma in the right forearm. (Cattaneo *et al*, 2015a)

Bormio, Piazza del Kuerc

<u>BO.K</u>

An adult female individual, 45-55 years old and 160-168 cm of height.

The entire skeleton was not recovered. Specifically, the lower limbs were absent, except few elements of the feet and the right fibula. Moreover, the skeletal remains were in some districts highly fragmented and fragile (Figure 45), but the state of preservation was good enough to allow however the identification of peculiar characteristics of this individual (Cattaneo *et al*, 2015b).

In particular, macroporosity, lipping and osteophytes observed in some bone elements, such as vertebrae, feet phalanges and the bones of the upper limbs, pinpointed the development of osteoarthritis. In addition, enthesopathies, alterations of the enthesis (the insertion of a tendon or ligament to a bone) due to mechanic stress, metabolic or inflammatory disorders, were identified in both humeri, radii and calcanei (Cattaneo *et al*, 2015b; Caruso *et al*, 2017b).

Lytic lesions, specifically cavitations, were identified in several bone districts (skull, right humerus, scapulae, ribs and pelvis). These lesions were highly destructive, with different sizes and shapes, affecting all bone layers (Figure 46a-b). Considering the pattern and the distribution a metastatic cancer was suggested. In particular, given the characteristics of the lesions and the sex of the individual, the breast cancer was hypothesized. However, other carcinoma could not be excluded (Caruso *et al*, 2017b). To support the diagnosis, the metastatic lesions were compared to those identified in two skeletal remains from Cimitero Maggiore. In particular, a female individual with a breast cancer and a male individual with a lung cancer were selected for the comparison. The woman showed numerous lesions on different bone regions (skull, scapulae, pelvis, ribs and vertebrae), characterised by different sizes and shapes, a remodelled border and denticulated morphology, in some cases surrounded by porosity, involving both the cortical and the cancellous bone. Differently, the man had a single lytic lesion in the iliac fossa, with an irregular shape, surrounded by a depression area, with sharply and denticulated edges. The similarity of the lesions between the two women in terms of size, shape and morphology of the lesions supported the first hypothesis of breast cancer (Caruso *et al*, 2017b).

In addition, bone remodelling associated to *antemortem* compound fractures were identified in six ribs and the right clavicle (Figure 46c). In this case, the observation of traumas in different bone districts was probably associable to the weakness related to the pathological condition of the subject (Caruso *et al*, 2017b).

85

**Figure 45.** BO.K: the skeletal remains of the individual.
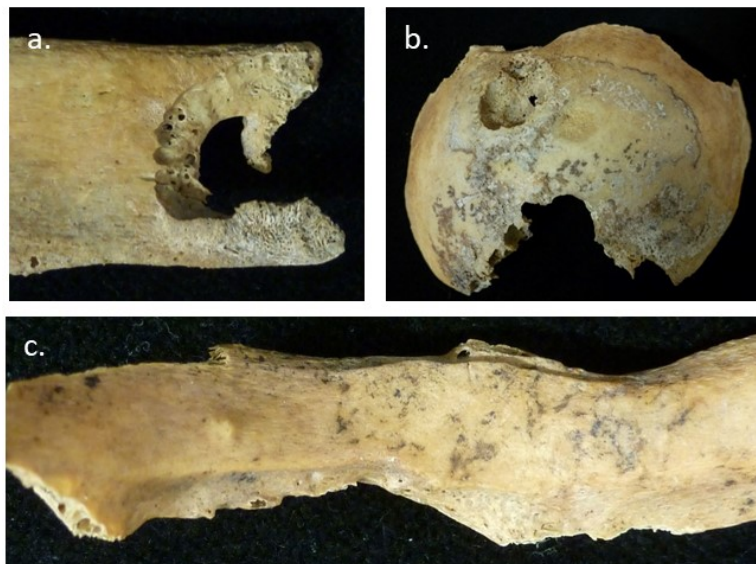
Figure 46. BO.K: a. lytic lesions on the cranial vault; b. lytic lesion on the sternal end of a rib; c. remodelled compound fracture on a left rib (Cattaneo *et al*, 2015b; Caruso *et al*, 2017b).

Milan, Viale Sabotino

MI SAB

A female individual, 45-55 years old, 172 cm of height (Figure 47a).

The paleopathological analyses revealed osteoarthritis, especially in both clavicles, and *cribra orbitalia*, a porotic hyperostosis in the orbital roof which was associated to chronic deficiency anemia, generally related to iron deficiency (Ortner, 2003; Saccomani, 2010). In addition, lytic-proliferative lesions on the cranial vault were indicative of a syphilis at the advanced stage (Figure 47b). Consequently, considered the status of the disease, this woman could be paralysed and/or suffering of mental disorders. In fact, one of the symptoms of advanced syphilis is mental illness (Fornaciari and Giuffra, 2009; Saccomani, 2010).

Teeth were worn and characterized by enamel hypoplasia, which could be caused by nutritional or pathological stress occurred during the childhood, that stopped dental growth (Saccomani, 2010).
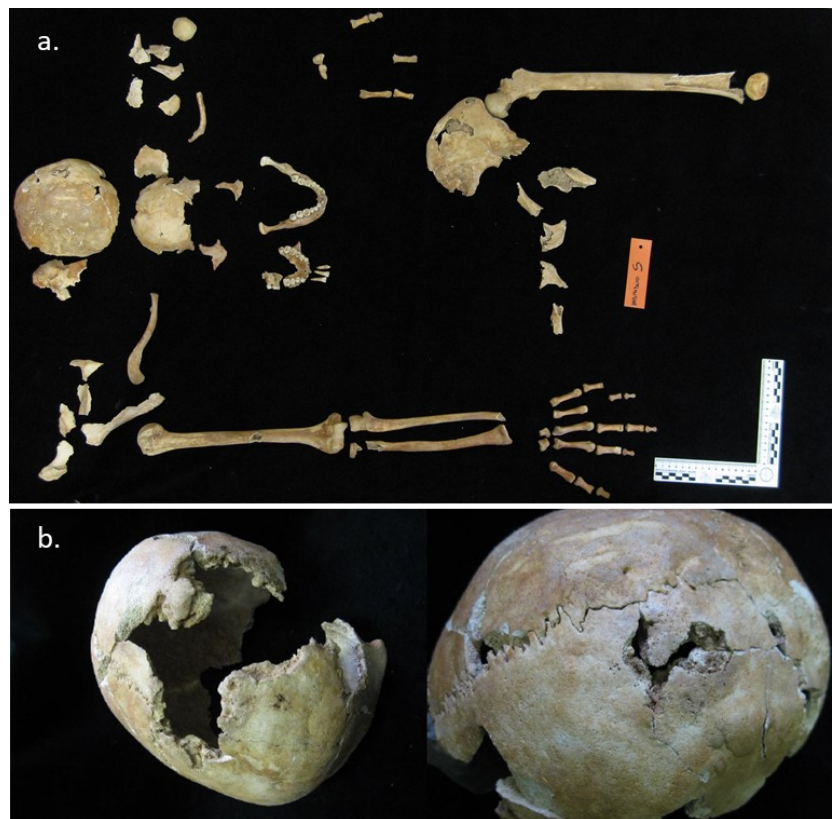
Figure 47. MI SAB: a. the skeletal remains; b. lytic and proliferative lesions associated to syphilis.

Milan, Cà Granda

Considering the deposition of the corpse, as well as environmental factors and human activities occurred in the past, the skeletal remains recovered during the excavations could not be ascribed to specific individuals due to the commingling of the several bone elements. Therefore, every single bone was recovered and named with a different code, except some rare cases in which the articulation of bones allowed the assignment to the same individual.

MI CG Q/3

A left femur anchylosed to the coxal bone belonged to an adult male individual, 30-45 years old (Figure 48). The coxo-femoral anchyloses could be associated to ankylosing spondylitis or post traumatic anchyloses. The radiological examination of the remains suggested a bone anchyloses after an *antemortem* trauma of this portion, confirming the second hypothesis (Sguazza, 2015).

Figure 48. MI CG Q/3: coxo-femoral anchyloses of a male individual of 30-45 years old (Sguazza, 2015).

MI CG O2846

A skull belonged to an adult female individual.

The paleopathological analyses identified lytic lesions in the cranial vault associated to syphilis at an advanced stage (Figure 49). Considered this status, the disease could have been involved also the nervous system, causing paralysis or mental disorders, which could bring to a deranged psychological asset of the person (Friedrich *et al*, 2014; Crozatti *et al*, 2015; Sguazza, 2015).

Figure 49. MI CG O2846: a skull belonged to a female individual possibly affected by syphilis. In the picture, from the left to the right, the skull in upper, lateral and frontal view (Sguazza, 2015).

Milan, Cimitero Maggiore

<u>MI CM 39</u>

A Caucasoid male individual, older than 65 years old (Figure 50).

The paleopathological analyses identified osteoarthritis spread in all the vertebral column. In particular, cervical vertebrae showed an initial compression of the vertebral body and deformation of the articular facets. In addition, evident osteophytes were observed in the lumbar vertebrae.

Additional signs related to osteoarthritis were recognized in the head of both humeri, femurs and in the glenoid fossa of the right scapula.

Figure 50. MI CM 39: the skeletal remains.

MI CM 33

A Caucasoid male individual, 49-73 years old (Figure 51).

No particular pathology was identified. The skeletal remains were characterized especially by osteoarthritis involving the entire vertebral column. Moreover, the ossification of the ileum-sacrum joint was detected.

Figure 51. MI CM 33: the skeletal remains.

3.2.2 Macroscopic taphonomical analysis

Macroscopic analysis was performed evaluating the general appearance and preservation of each bone element, especially considering bone diaphysis when possible, following the criteria defined by Behrensmeyer (1978).

As expected, skin, muscles, ligaments and greasiness were absent in all specimens.

In approximately half cases bone samples were complete and characterized by taphonomical lesions usually in proximity of the epiphyses, which showed an overall loss of the cortical bone leaving exposed the underlying spongy tissue. Differently, the remaining cases lacked some portions, generally one or both epiphyses.

Cracking and flaking were observed in the majority of the samples, as well as weathering on the cortical surface of the more damaged samples.

In the table below a short description and pictures of each sample are reported.

Sample	Historical period	Stage	Description	Picture
MI UC	I-V AD	1	Complete left femur, characterized by small erosion of the cortical surface on both extremities. The bone was well preserved, with no cracking and flaking macroscopically visible.	
BG-BO-SC	VII-IX AD	4	Right incomplete femur, the third distal end was lacking. Erosion of the cortical surface was identified at the extremities, where spongy tissue was also observed, especially on the head. Bone surface appeared widely rough and fibrous in texture. Large areas were characterised by a deeper weathering of the cortical surface. Cracking was visible. Bone appeared very fragile especially at the broken end.	

MI PSA T49	XIII-XV AD	1	Complete left femur, characterised by small erosions of the cortical surface and few spongy tissue exposed at the epiphyses. The bone was well preserved, with no cracking and flaking macroscopically visible.	
MI PSA T23	XIII-XV AD	3	Left femur, complete. Erosion of the cortical surface was visible at the extremities, where spongy tissue was also exposed. Bone surface was characterised by areas of rough and homogenously weathered compact bone. In these areas the external layered bone has been removed. Cracking was also observed.	
BO.K	XIV AD	2	Complete left humerus, showing erosion of the cortical surface in both epiphyses. Specifically, proximal end appeared more affected and spongy tissue was observed on both extremities. Cracking was visible on the head and small islets of flaking at the distal epiphysis.	
MI SAB	XVII AD	4	Left incomplete tibia, lacking the anterior portion of the proximal epiphysis and the third distal end. Cortical surface was eroded and spongy tissue exposed, especially at the extremities. Bone surface appeared widely rough and fibrous in texture, with large areas characterised by a deeper weathering of the cortical surface. Cracking was visible. Bone appeared very fragile especially at the proximal epiphysis.	

MI CG Q/3	XVII-XVIII AD	3	Left femur ankylosed to the coxal bone. Pubis and part of the ilium were absent. Erosion of the cortical surface was visible at the distal end and in the coxal bone, where spongy bone was exposed. Bone diaphysis had several areas of rough and weathered compact bone, with a fibrous texture. In these areas the external layered bone has been removed. Cracking was also visible.	
MI CG O2846	XVII-XVIII AD	4	Skull lacking of some portions in the occipital and left parietal bones. The cortical surface was eroded and spongy tissue exposed in many regions of the cranial bones. Bone surface was widely rough and fibrous in texture. Large areas were characterised by a deeper weathering of the cortical surface. Several cracks were observed and splinters were identified. Generally, the bone sample appeared very fragile.	
MI CM 39	XX AD	2	Complete right femur, with erosion of the compact tissue at the extremities, where spongy tissue was also exposed. Few cracking and flaking were visible, the last as islets in the posterior and lateral sides of the bone.	
MI CM 33	XX AD	1	Right femur, incomplete due to the lack of the proximal epiphysis. Cortical surface was eroded at both extremities, where spongy tissue was exposed. Cracking and flaking were not visible macroscopically.	

Table 6. Macroscopic bone preservation. In the table, stage, a short description and picture, included macro, are reported for each sample.

93

3.2.3 Microscopic analysis

The histological sections were evaluated and analysed at the optical microscope and the results summarised in two paragraphs:
- calcified thin sections (non-decalcified);
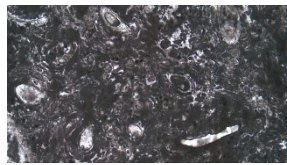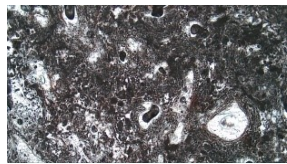- decalcified thin sections.
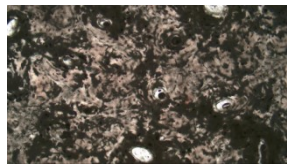
The state of preservation of each sample was correlated, analysing the microscopic appearance of each section. In particular, the possibility to recognize bone structures, such as osteons and lamellae, the amount of intact bone and the presence of destructive foci caused by microbial activity were considered in order to classify the bone tissue and assign the Histological Index described by Hedges and colleagues (1995) as reported by Caruso and co-workers (2017a).

3.2.3.1 Calcified thin sections
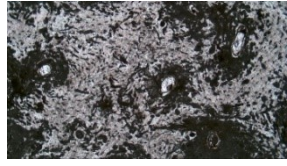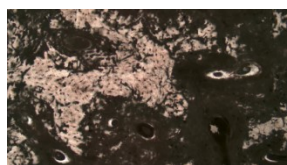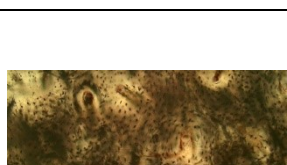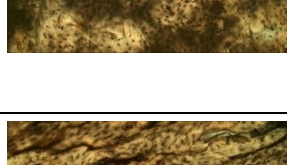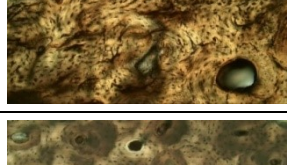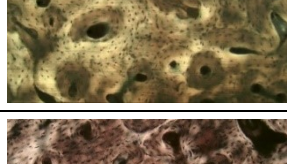
The majority of calcified thin sections were barely legible in the outer and inner portions, corresponding to the regions in contact with the outside and the medullary cavity respectively, as no bone structures were visible and all the tissue was destroyed or covered by tunnelling. The central region was generally less affected by microbial activity and osteons and bone lamellae usually were detectable, except for highly degraded samples.

On the contrary, well-preserved bone samples were characterised by intact bone tissue in the entire section, with bone structures similar to fresh thin sections.

In the table below, the bone tissue preservation is reported for each sample.

Sample	Historical period	OHI	Description	Picture
MI UC	I-V AD	0	The entire section was covered by tunnelling; no original structures were identifiable, only in some areas, Haversian canals were recognizable.	
BG-BO-SC	VII-IX AD	0	All the section was covered by tunnelling; only in some portions, some Haversian canals could be seen.	
MI PSA T49	XIII-XV AD	2	Clear bone lamellae were visible and well preserved between destructive foci. Less than 30% of the section had intact bone.	

MI PSA T23	XIII-XV AD	2	Clear lamellar structures were visible and preserved among pattern of destructive foci. Intact bone was evident in less than 25% of the section.	
BO.K	XIV AD	5	Well-preserved bone section, where osteons and bone lamellae were intact and not affected by tunnelling.	
MI SAB	XVII AD	1	Small areas of well-preserved bone structures present, but less than 20% of the section was characterised by intact bone. Many destructive foci were visible.	
MI CG Q/3	XVII-XVIII AD	3	The section was covered by tunnelling in both the periosteal and endosteal surfaces. The middle zone was more preserved with lamellar structures visible, even if contamination was present. Bone tissue well preserved in more than 50% of the section.	
MI CG O2846	XVII-XVIII AD	4	A small amount of destructive foci was present. More than 85% of the section was composed by intact bone, with well-preserved bone structures.	
MI CM 39	XX AD	5	Well preserved bone section, with no destructive foci identifiable.	
MI CM 33	XX AD	5	No destructive foci were present; all the section was characterised by well-preserved bone structures.	
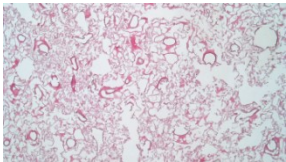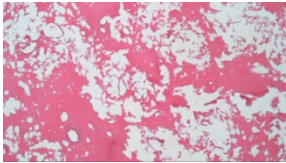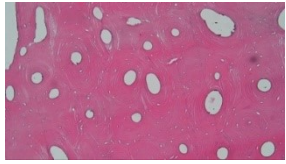
Table 7. Microscopic bone preservation. In the table the Oxford Histological Index (OHI) and a short description of the section are reported for each sample. Moreover, a representative picture for each section is included (magnification 100X).

95
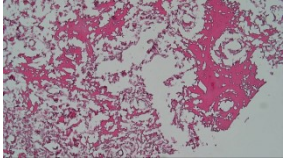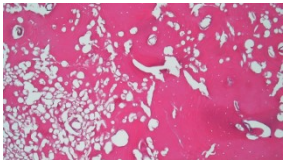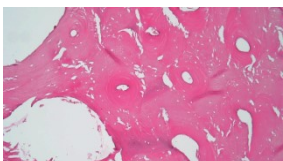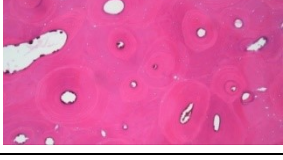
3.2.3.2 Decalcified thin sections

After removal of the mineral component, only the organic fraction was visible. Sections were characterised by different amounts of preserved collagen and destructive foci caused by microbial activity, marked as holes and tissue degradation. The outer and inner regions of the slides were in some samples preserved, contrary to what was observed in the calcified sections.

In the table below, the bone tissue preservation is reported for each sample.

Sample	Historical period	OHI	Description	Picture
MI UC	I-V AD	0	The majority of the section (>95%) was completely destroyed. No original bone structures were identifiable, only in few cases some canals were recognizable.	
BG-BO-SC	VII-IX AD	1	Small portions were preserved and few osteons were identifiable, but more than 80% of the section was destroyed. The outer and inner regions of the tissue were preserved.	
MI PSA T49	XIII-XV AD	2	The section was more preserved in the outer region. Clear lamellar structures were preserved between areas affected by taphonomic activity. Intact bone was visible in about 30% of the section.	
MI PSA T23	XIII-XV AD	2	Bone tissue was well-preserved in about 40% of the section. Lamellar structures were identifiable and preserved between area affected by taphonomic activity.	
BO.K	XIV AD	5	No destructive foci were visible. Well-preserved bone structures were observed in the entire section.	

MI SAB	XVII AD	1	The majority of the section was completely destroyed. Only small areas were well-preserved and some bone structures were identifiable even if affected by taphonomic activity.	
MI CG Q/3	XVII-XVIII AD	3	The section was well preserved (more than 70% of the tissue was characterised by intact bone). Bone structures appeared intact, even if several round destructive foci were visible .	
MI CG O2846	XVII-XVIII AD	3	More than 70% of the section was composed by intact bone tissue. Bone structures were generally completely intact but affected by several destructive foci.	
MI CM 39	XX AD	5	Well-preserved bone tissue in the entire section.	
MI CM 33	XX AD	5	Well-preserved bone tissue in the entire section.	

Table 8. Microscopic bone preservation. In the table the Oxford Histological Index (OHI) and a short description of the section are reported for each sample. Moreover, a representative picture for each section is included (magnification 100X).
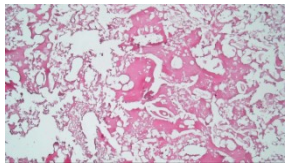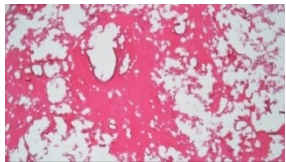
In addition, in some sections generic debris, possibly botanical and/or microbial contamination, were visible inside the bone canals. They appeared as brownish conglomerated (Figure 52a) or single cells with a biconcave disc shape similar to that of erythrocytes (Figure 52b-d), as previously noted by Cappella *et al* (2015).

97
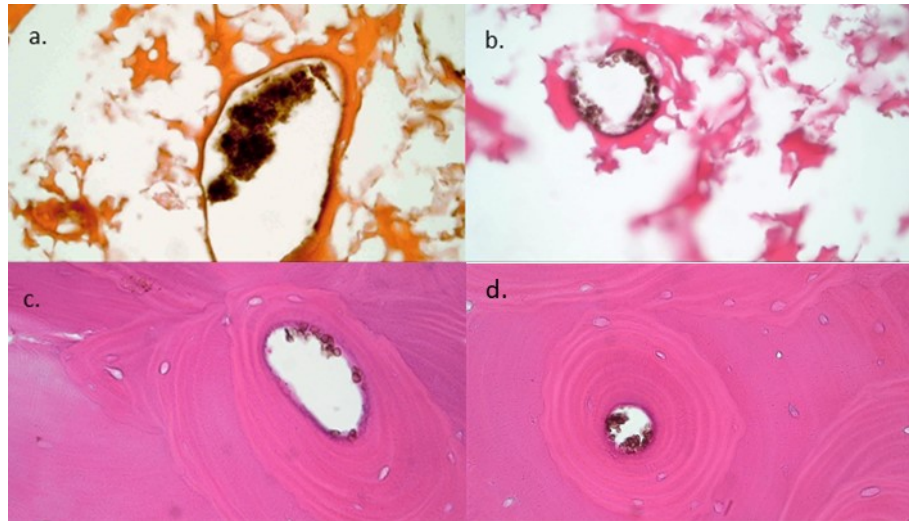
Figure 52. Examples of generic debris, possibly botanical and/or microbial contamination inside bone canals: a. MI PSA T49 (400X); b. MI PSA T49 (400X); c. MI CM 39 (400X).

3.2.4 Genetic analyses

3.2.4.1 DNA quantitation

DNA quantitation showed the amount of DNA recovered from each bone sample and evidenced at the same time the state of preservation of the genetic substrate. Specifically, good results were obtained for the short mitochondrial fragment (60 bp), while the longest one (143 bp) and the nuclear DNA probe were amplified only in recent samples (Tables 9-10).

No DNA was obtained from samples MI CG Q/3 and MI CG O2846, whose quantification amounts were comparable to the negative control (Ex0_2).

				Overnight incubation			
	Period	*Nuclear DNA pg/µl*		*mtGE/µl 143bp*		*mtGE/µl 60bp*	
		Rep 1	Rep 2	Rep 1	Rep 2	Rep 1	Rep 2
MI UC	I-V AD	1.016	0.829	0.258	0.772	4.463	4.026
BG-BO-SC	VII-IX AD	0.418	0.532	0.353	Undet	1.813	1.541
MI PSA T49	XIII-XV AD	2.713	3.120	5.112	4.240	70.952	82.628
MI PSA T23	XIII-XV AD	7.348	7.489	6.630	9.962	104.941	118.169
BO.K	XIV AD	24.181	30.492	46.779	56.610	145.701	175.457
Ex0_1		0.208	0.096	0.830	0.826	0.517	0.600
MI SAB	XVII AD	0.742	0.457	11.669	12.970	107.462	125.950
MI CG Q/3	XVII-XVIII AD	0.214	0.031	0.207	Undet	0.430	0.283
MI CG O2846	XVII-XVIII AD	0.272	0.110	Undet	Undet	0.676	0.208
MI CM 39	XX AD	22.658	24.456	70.279	78.175	303.000	252.468
MI CM 33	XX AD	19.092	27.899	167.619	186.566	457.254	498.465
Ex0_2		0.043	0.058	Undet	0.142	0.174	0.158

Table 9. DNA quantitation results for each replicate (Rep1 and Rep 2) of the bone samples extracted with the overnight incubation protocol. In detail, the amounts of nuclear DNA (pg/µl), 143 bp and 60 bp mitochondrial fragments (mitoGenomes/µl) are reported. (Ex0_1 and Ex0_2: extraction negative controls).

	Period	Two-days incubation					
		Nuclear DNA pg/µl		**mtGE/µl 143bp**		**mtGE/µl 60bp**	
		Rep 1	Rep 2	Rep 1	Rep 2	Rep 1	Rep 2
MI UC	I-V AD	0.853	2.079	0.478	0.368	3.521	2.723
BG-BO-SC	VII-IX AD	0.364	0.252	0.157	0.272	1.567	1.720
MI PSA T49	XIII-XV AD	3.554	3.089	5.006	5.224	83.574	86.165
MI PSA T23	XIII-XV AD	6.063	9.159	11.683	10.977	123.965	132.843
BO.K	XIV AD	27.084	32.461	62.971	74.845	196.482	236.401
Ex0_1		0.107	0.156	0.639	0.213	1.004	0.309
MI SAB	XVII AD	1.469	1.268	9.384	12.837	118.609	119.219
MI CG Q/3	XVII-XVIII AD	0.271	0.079	Undet	Undet	0.362	0.735
MI CG O2846	XVII-XVIII AD	0.018	0.035	Undet	Undet	0.797	0.770
MI CM 39	XX AD	57.777	26.234	65.583	73.503	249.193	251.699
MI CM 33	XX AD	31.871	29.099	190.643	211.727	508.645	578.763
Ex0_2		0.065	0.108	0.221	Undet	0.367	0.351

Table 10. DNA quantitation results for each replicate (Rep1 and Rep 2) of the bone samples extracted with the two-days incubation protocol. In detail, the amounts of nuclear DNA (pg/µl), 143 bp and 60 bp mitochondrial fragments (mitoGenomes/µl) are reported. (Ex0_1 and Ex0_2: extraction negative controls).

Even if several factors, especially environmental conditions of the burial, are involved in DNA preservation, the quantitation data obtained from the selected bones showed results according to the age of the samples: low DNA amounts were observed in the most ancient specimens and a greater DNA preservation was identified progressively in the most recent ones. Three samples diverged from this trend: two samples coming from the Cà Granda site (MI CG Q/3 and MI CG O2846), which gave no results, and sample BO.K, whose DNA quantities were similar to those observed in the two most recent samples (MI CM 39 and MI CM 33).

Moreover, the two extraction methods (overnight and two-days incubation protocols) were proved to have the same efficiency in DNA yield, by comparing the mean of the DNA amounts obtained for each probe using a paired two-sample Student's t-test (the critical t-values were -1.57, -1.30 and -1.36 corresponding to p-values of 0.14, 0.22 and 0.20 for the nuclear DNA, mtDNA 143 bp and mtDNA

60bp probes, respectively). This follows the demonstration that the differences in the values obtained by the replicates of each sample were statistically not significant as proven by comparing the DNA amounts obtained from each test for each probe using a paired two-sample Student's t-test (the critical t-values were: -1.65, -2.00 and -0.81 corresponding to p-values of 0.13, 0.08 and 0.44 for the nuclear DNA, mtDNA 143 bp and mtDNA 60bp probes, respectively, using the overnight protocol; 0.78, -1.92 and -1.66 corresponding to p-values of 0.46, 0.09 and 0.13 for the nuclear DNA, mtDNA 143 bp and mtDNA 60bp probes, respectively, using the two-days incubation protocol).

Figure 53 shows the graphical comparison between the replicates of each probe from the DNA extracts obtained using the overnight and the two-days incubation protocols. Specifically, the comparisons for the nuclear DNA, mitochondrial DNA 143 bp and 60 bp probes are reported in Figure A, B and C, respectively. The DNA amount for each probe for the overnight and two-days protocols are plotted in the x- and y-axes.

In particular, each value obtained from the overnight extraction protocol was similar to that obtained from the two-days incubation protocol, as they lie close to the line corresponding to the function x=y (that is: values from the first protocol are equal to the second one).

In the nuclear DNA figure, two outliers were identified. These corresponded to the comparison of the first replicate of samples MI CM 39 and MI CM 33. The analysis of the data evidenced that this difference between the values is statistically not significant, and as such was due only to the random error affecting the sampling process (i.e. the entire process which started with the specimen collection and ended with the DNA value output).
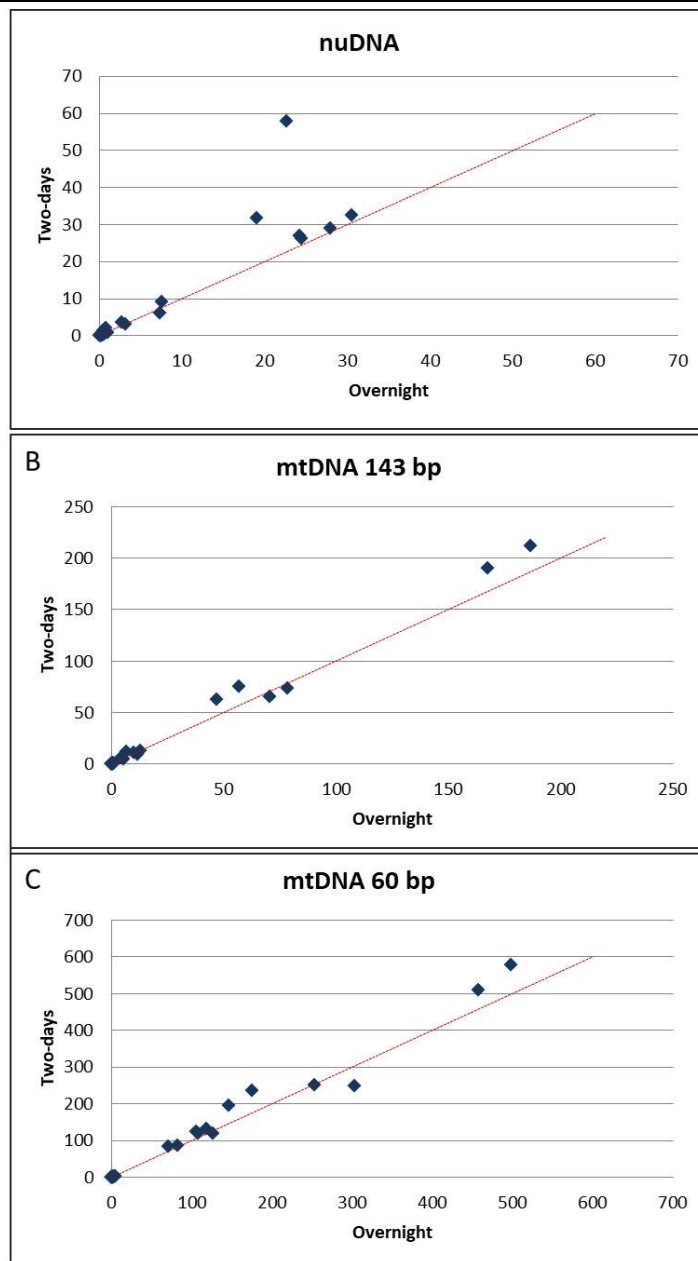
Figure 53. Graphic comparison between the two extraction methods for each probe (A: nuclear DNA probe; B: mitochondrial DNA 143 bp probe; C: mitochondrial DNA 60 bp probe). The DNA amount of each replicate test from the overnight and two-days incubation protocols are plotted on the x- and y-axes, respectively. In red the line corresponding to the function x=y; values that lie in proximity of the line pinpoint a similarity of the two methods.

3.2.4.2 Genetic sex typing

Considering the quantitation data, the genetic sex typing was carried out using the overnight extracts for the samples with nearly 3 pg/µl of DNA amount, in order to reach at least a quantity of 12 pg using the maximum volume of DNA extract required in the reaction (4 µl). For this reason, five samples (MI PSA T49, MI PSA T23, BO.K, MI CM 39 and MI CM 33) were subjected to the genetic sex typing using between 11.7 pg and 54.7 pg of input DNA.

Positive amplification results were achieved from all samples.

Specifically, HRM profiles revealed the expected three peaks for males (two for the Y and one for the X markers, respectively), with a 4°C spacing among them, and the X-female specific peak for females, as shown in Figure 54.
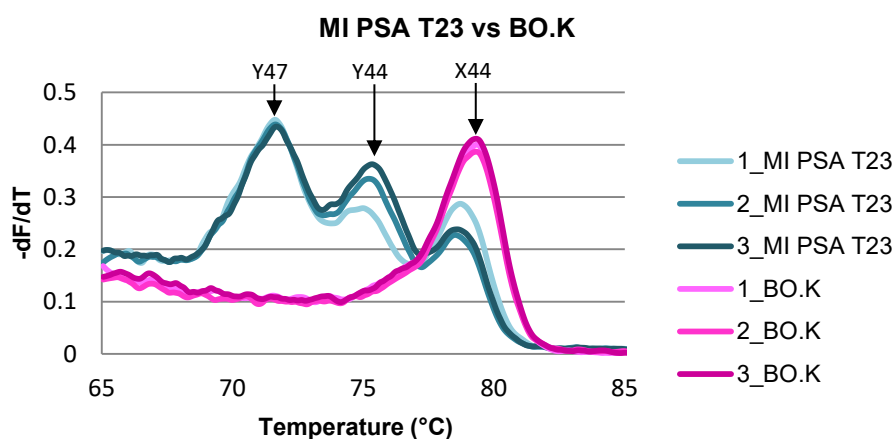
Figure 54. HRM profiles of the three replicates performed from samples MI PSA T23 (male, in blue) and BO.K (female, in pink). The negative first derivate of the T_m point and the temperature are plotted in the y- and x-axes, respectively. The male sample shows three peaks: the first two represent the Y-specific markers (Y47 and Y44) and the third one the X-specific sequence (X44). The female sample has only the X-specific characteristic as expected.

On the contrary sample MI PSA T49 (Figure 55) showed an HRM profile with only the two Y specific peaks. Even if the X-fragment was missing, the presence of the two peculiar Y-chromosome peaks was sufficient to attribute the sample to a male individual.
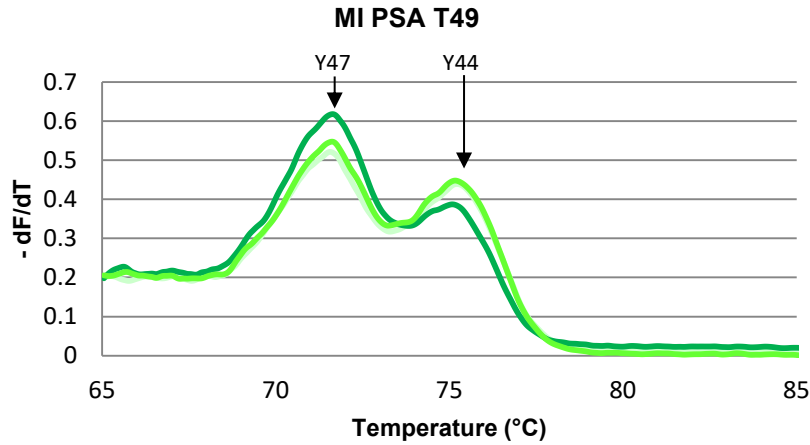
MI PSA T49

Figure 55. HRM profile of the three replicates from sample MI PSA T49. Only the two peaks of the Y specific markers can be seen.

In all the cases the genetic sex was in agreement with the anthropological diagnosis (Table 11).

Sample	Nuclear DNA (pg/µl)	Nuclear DNA used (pg/µl)	Anthropological diagnosis	Genetic diagnosis
MI PSA T49	2.916609	11.666	Male	XY
MI PSA T23	7.418931	22.257	Male	XY
BO.K	27.336642	54.673	Female	XX
MI CM 39	23.557059	47.114	Male	XY
MI CM 33	23.495678	46.991	Male	XY

Table 11. Genetic sexing results compared to the anthropological diagnosis. The average amount of nuclear DNA (overnight incubation protocol) recovered from the samples and that used in the reaction are reported.

The extraction negative controls and the negative controls of the reaction gave no results, confirming the absence of any possible external contamination.

3.2.4.3 Mitochondrial analysis

Mitochondrial DNA analysis was performed for each bone sample selected for this study using both methods. As no difference in DNA yield was statistically observed between the two extraction methods, the extracts were considered equivalent and submitted to genetic typing as two different aliquots of the same sample. In particular, samples incubated overnight were submitted to PEC protocol, while samples incubated two days were submitted to mito Tiling protocol. Interesting results were obtained in both cases.

Differences in the coverage were observed between the two approaches. In fact, mito Tiling protocol supplied information almost for the entire genome, even if few reads were recovered in some positions. On the other hand, the PEC protocol generated a good coverage especially in the control region and in a region localised around position 8,342, while few data on other positions were obtained (Figure 56).

MI CM 33_mT

MI CM 33_PEC
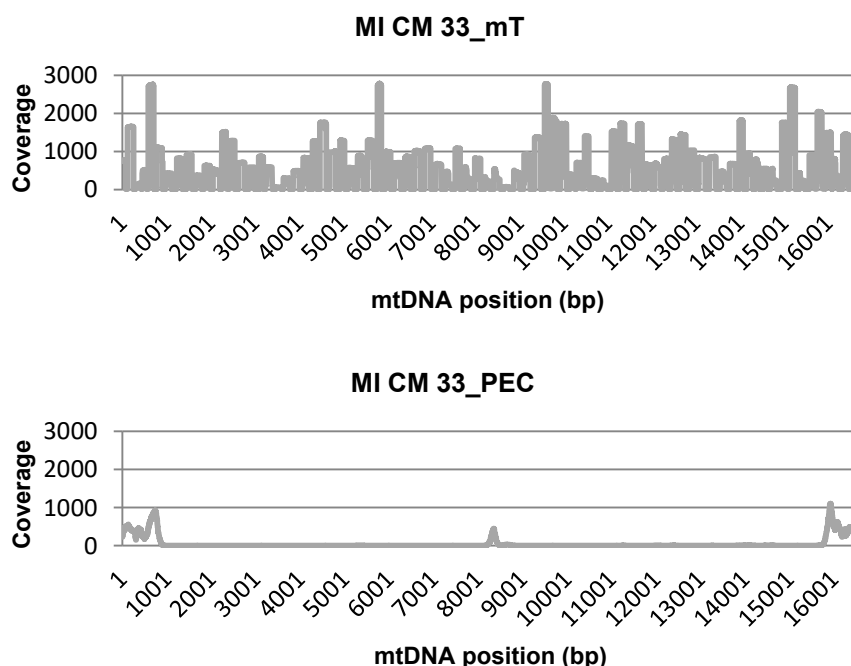
Figure 56. Mitochondrial DNA coverage of sample MI CM 33 obtained with the two methods (mtDNA: mito Tiling protocol; PEC: PEC protocol).

For most of the bone samples analysed in this study, it was not possible to recover a complete mtDNA sequence, except for sample MI PSA T49 using the mito Tiling protocol. Generally, some gaps of different sizes were observed as reported in Tables 12-13.

Extraction negative controls were submitted to both analyses and no results were obtained in each case.

Sample	Covered positions
MI UC_mT	11669-11704 15823-15843 16073-16107
BG-BO-SC_mT	1-80 118-256 456-545 903-935 991-1120 1236-1352 1444-1577 2231-2353 3446-3462 3471-3576 3597-3614 3657-3786 3876-4017 4315-4423 4617-4637 4929-5050 5103-5127 5777-5875 5969-6084 6294-6524 6599-6626 7232-7259 7537-7779 8619-8788 8878-8898 8901-9012 9569-9816 10288-10408 10465-10592 11077-11188 11290-11387 11453-11589 11686-11784 11849-11865 11886-11904 12051-12067 12073-12216 12439-12569 12867-12989 13074-13215 13317-13464 13544-13698 13770-13905 13981-14090 14147-14166 14288-14319 14346-14448 14531-14590 14776-14840 14953-15071 15118-15137 15346-15468 15579-15702 15823-15843 15860-15886 15945-16116 16227-16250 16335-16475 16520-16569
MI PSA T49_mT	1-16569
MI PSA T23_mT	4-1620 1669-9669 9742-9864 9892-10072 10139-11000 11089-11396 11453-13026 13064-15877 15945-16569
BO.K_mT	1-1364 1444-1998 2050-2557 2639-3004 3093-3105 3108-4625 4694-5055 5126-5512 5554-6098 6167-6329 6410-6553 6621-7503 7537-8343 8428-8798 8878-9250 9271-10617 10641-11000 11077-13037 13065-13916 13980-14448 14531-14644 14690-16569
MI SAB_mT	1-81 118-425 443-917 930-943 975-1156 1220-1352 1444-1803 1844-1990 2056-2206 2257-2366 2412-2786 2872-2996 3093-3105 3107-3232 3251-3408 3457-3617 3634-4017 4086-4224 4314-4423 4436-4625 4694-4841 4929-5055 5126-5883 5957-6098 6119-6315 6395-6587 6621-6766 6844-7332 7336-7518 7549-7669 7699-8131 8206-8562 8640-8798 8868-9023 9096-9250 9341-9490 9569-9863 9931-10072 10139-10595 10653-10780 10855-11000 11089-11201 11277-11396 11453-11600 11679-11796 11855-12003 12067-12141 12262-12391 12454-12776 12855-13021 13065-13217 13316-13737 13770-13944 13976-14132 14154-14303 14346-14448 14531-14644 14835-15469 15567-15703 15727-15886 15945-16288 16335-16569
MI CG Q/3_mT	1444-1462 1484-1497 2052-2068 2627-2643 4009-4029 4845-4864 5983-5998 6845-6859 9931-9948 10365-10380 10663-10677 11077-11097 11158-11175 11886-11904 12330-12350 12974-12989 13074-13127 13129-13215 13328-13341 13571-13590 13901-13916 14545-14565 15118-15134 15694-15714 15823-15843 16069-16116 16200-16222 16459-16475
MI CG O2846_mT	1220-1239 1444-1465 1686-1706 4810-4824 10063-10085 10192-10208 10662-10683 11077-11096 11158-11174 11850-11865 11889-11904 12883-12899 12975-12989 13327-13341 13666-13683 14147-14167 14367-14391 14545-14565 15823-15843 16076-16107
MI CM 39_mT	1-917 975-1158 1220-16569
MI CM 33_mT	1-1364 1414-3823 3867-4046 4086-4224 4300-6329 6395-7485 7537-7725 7745-8798 8878-10132 10139-13224 13287-13709 13763-14913 14944-15265 15331-16280 16323-16569

Table 12. mtDNA positions covered using the mito Tiling protocol.

Sample	Covered positions
MI UC_PEC	1-950 1098-1220 1604-1811 1818-1946 2091-2522 2893-2929 3425-3490 3584-3788 3910-4175 4177-4237 4301-4333 4597-5140 5151-5516 5567-5766 5866-6004 6368-6482 6700-6817 7076-7223 7371-7398 7414-7497 7550-7658 7676-8007 8059-8460 8499-8702 8716-8957 9407-9477 9480-9536 9618-9809 10768-11241 11369-11472 11582-11814 11915-12479 12512-12637 12854-13025 13027-13046 13265-13539 13580-13850 13878-14434 14456-14490 14509-14530 14676-15007 15133-15259 15579-15787 15803-16569
BG-BO-SC_PEC	1-306 320-492 511-524 654-814 3840-3894 4932-4947 5095-5236 5290-5431 6308-6445 7105-7236 7659-7814 8279-8407 8651-8730 8958-9143 9495-9667 10851-10931 10965-11205 11292-11394 12335-12374 12407-12417 12420-12499 12901-12946 13195-13667 13688-13878 14007-14154 14164-14350 14426-14533 14765-14769 14771-14810 15832-15989 16007-16272 16330-16387
MI PSA T49_PEC	1-964 1008-1339 1361-1587 1601-1829 2019-2591 2593-2605 2753-2909 3041-3105 3107-3167 3270-3916 3928-4439 4567-4679 5032-5498 5542-5823 5840-6052 6312-6704 7173-7290 7312-7342 7344-7514 7516-7545 7919-8002 8029-9026 9051-9212 9344-9859 10230-10421 10745-11392 11429-11537 11915-12194 12220-12696 12879-14872 15061-15115 15349-15577 15618-16569
MI PSA T23_PEC	1-1037 1507-1520 1593-2515 2769-2916 3083-3105 3107-3228 3242-3910 3914-5987 6048-6285 6306-6903 7017-7331 7342-9968 10015-10435 10461-10669 10703-15184 15244-16569
BO.K _PEC	1-998 1236-1364 1653-1683 1814-1948 2030-2592 3037-3105 3107-3166 3381-3655 4814-5700 6141-6267 6286-6754 6796-6912 7084-7214 7447-7557 7773-8830 9029-9244 9373-9471 9482-9794 9796-9905 9907-9922 10232-10312 10489-10612 10625-11397 11442-11573 11867-12590 12779-13514 13520-13991 14236-14797 14840-15043 15340-16569
MI SAB_PEC	1-1027 1059-1282 1560-2510 3360-3726 3757-4350 4387-4493 4826-5752 5865-5949 6126-6152 6154-6252 6277-6664 7086-7239 7267-7331 7339-7486 7567-9708 10083-10163 10178-10425 10530-10665 10709-11402 11464-11577 11595-11791 11843-12637 12651-15003 15250-16569
MI CG Q/3_PEC	40-103 961-978 6491-6506 7144-7160 7870-7893 9569-9653 10202-10215 14095-14110 15440-15459
MI CG O2846_PEC	512-525 831-844 4494-4509 4932-4947 6215-6236 6429-6524 12958-13025 13027-13034 13036-13086 14094-14110 15443-15461 16106-16226 16439-16453
MI CM 39_PEC	1-1020 1076-1186 1201-1234 1623-1684 1788-1945 2047-2056 2058-2457 2710-2868 3331-3613 3648-3775 4042-4289 4319-4481 4585-4614 5550-5645 5980-6481 6523-6702 7323-7353 7641-7965 8049-8858 8864-9218 9282-9454 9479-9494 9496-9676 9691-9789 9839-10056 10242-10268 10296-10448 10691-11389 11653-12200 12248-12605 12611-12744 12850-13388 13442-13692 13773-14875 14928-15050 15253-16569
MI CM 33_PEC	1-1088 1098-1428 1626-1763 1873-2492 2508-2611 3404-3568 3703-4443 4451-4576 4586-4797 4889-5498 5548-5877 6096-6220 6420-6479 6481-6645 6658-6868 7101-7234 7613-7816 7821-8015 8042-9353 9452-9525 9527-9531 9533-9545 9547-9826 10245-10420 10538-11378 11555-11861 11888-12651 12756-13637 13682-14802 14871-15021 15268-16569

Table 13. mtDNA positions covered using the PEC protocol.

In the following figures, the coverage obtained for each mtDNA nucleotide position has been reported.

For each sample, the plots describing the coverages for both mito Tiling (mT) (A1) and PEC protocols (B1), normalized to the highest coverage recorded among the samples (12,645x), are shown. In order to highlight the coverages using a more suitable scale, an enlarged detail of the two plots (A2 and B2) with different scales is also reported, on the bottom of each figure.

The mtDNA position in bp and the coverage are plotted in the x- and and y-axes, respectively.

The best results with mito Tiling protocol were obtained from sample MI PSA T49, where the complete sequence was achieved with a good coverage (mean coverage 959x), followed by MI CM 39, MI CM 33, MI PSA T23, BO.K and MI SAB.

Samples MI UC and BG-BO-SC, for which no successful coverage was achieved using the mito Tiling approach, gave better results using the PEC protocol.

Differently, no data were recovered from the two bone samples coming from the Cà Granda site (MI CG Q/3 and MI CG O2846) with both methods.
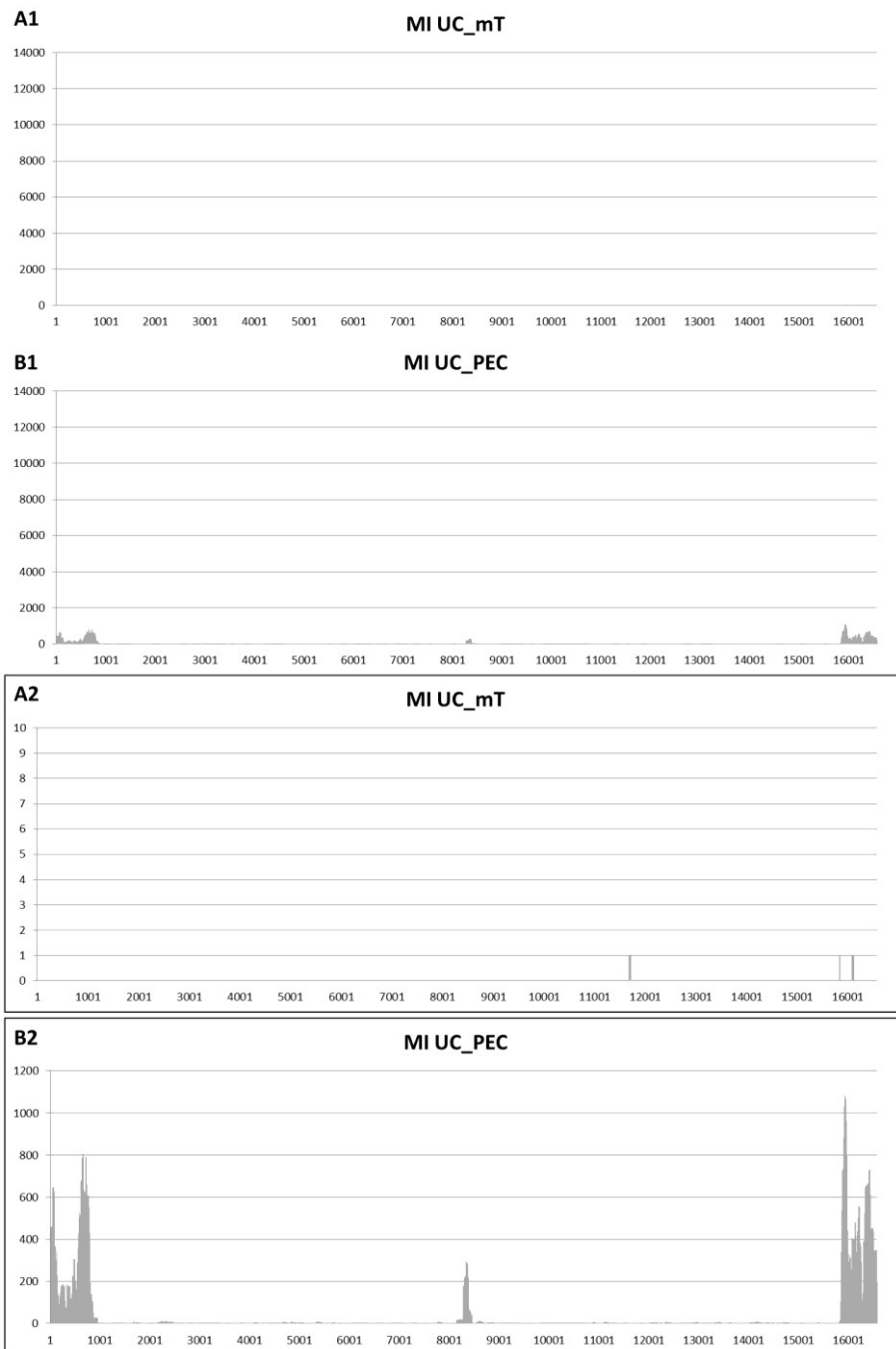
Figure 57. Mito Tiling (A) and PEC (B) coverages of the sample MI UC: A1 and B1. the coverages obtained normalised to the highest coverage recorded among the samples; A2 and B2. enlarged details of the coverages above.

109
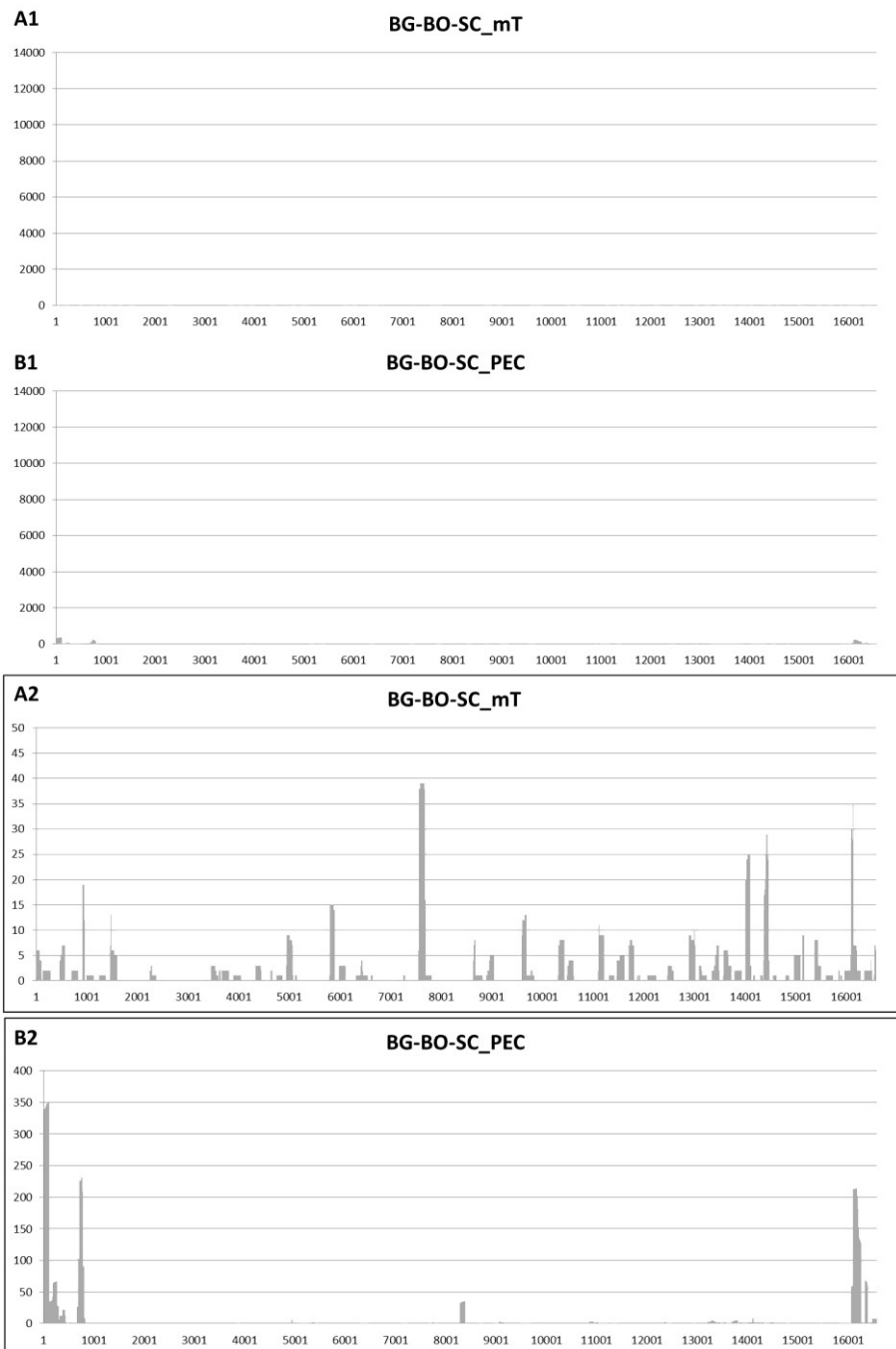
Figure 58. Mito Tiling (A) and PEC (B) coverages of the sample BG-BO-SC: A1 and B1. the coverages obtained normalised to the highest coverage recorded among the samples; A2 and B2. enlarged details of the coverages above.
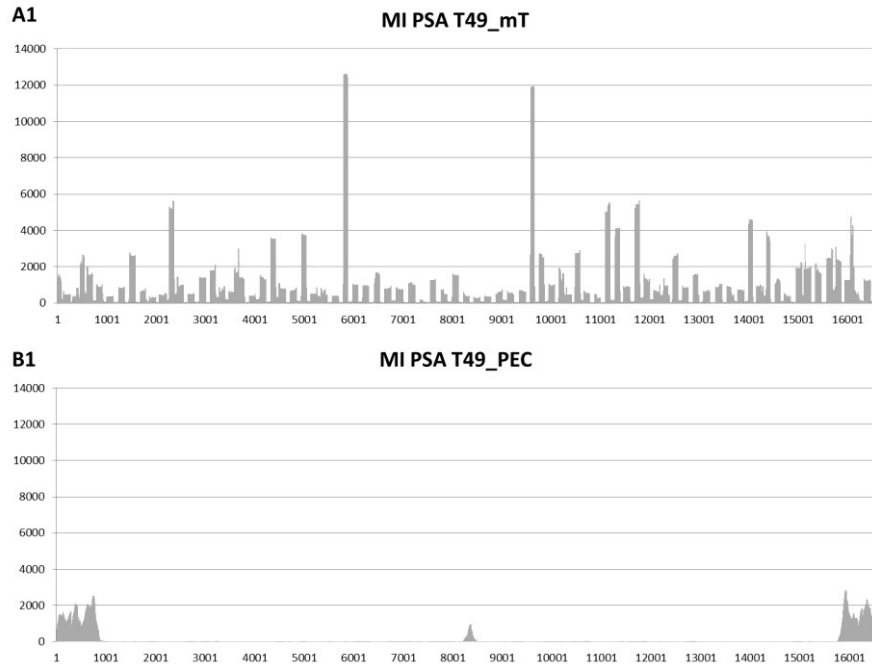
Figure 59. Mito Tiling (A) and PEC (B) coverages of the sample MI PSA T49: A1 and B1. the coverages obtained normalised to the highest coverage recorded among the samples; A2 and B2. enlarged details of the coverages above.
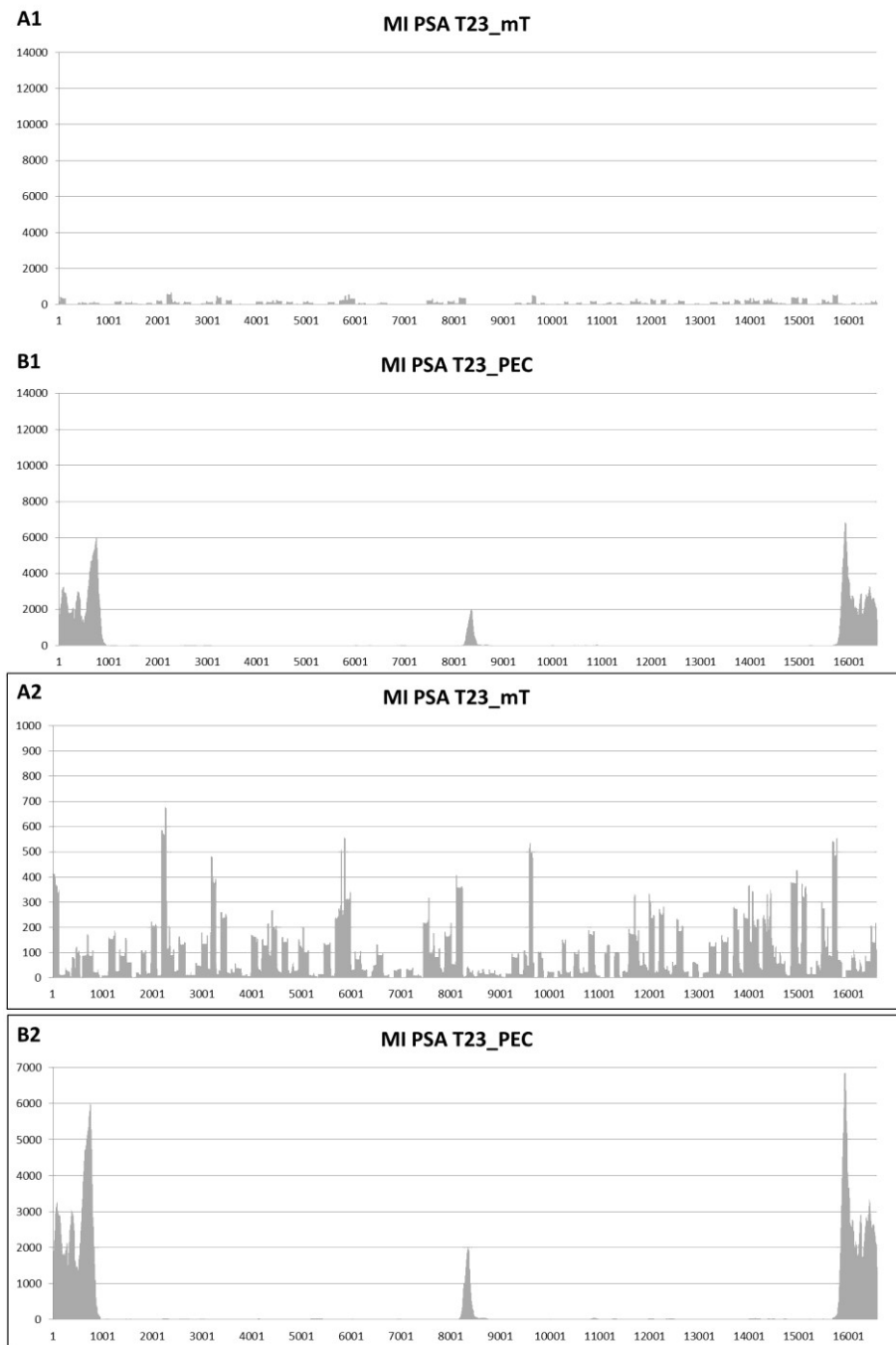
Figure 60. Mito Tiling (A) and PEC (B) coverages of the sample MI PSA T23: A1 and B1. the coverages obtained normalised to the highest coverage recorded among the samples; A2 and B2. enlarged details of the coverages above.
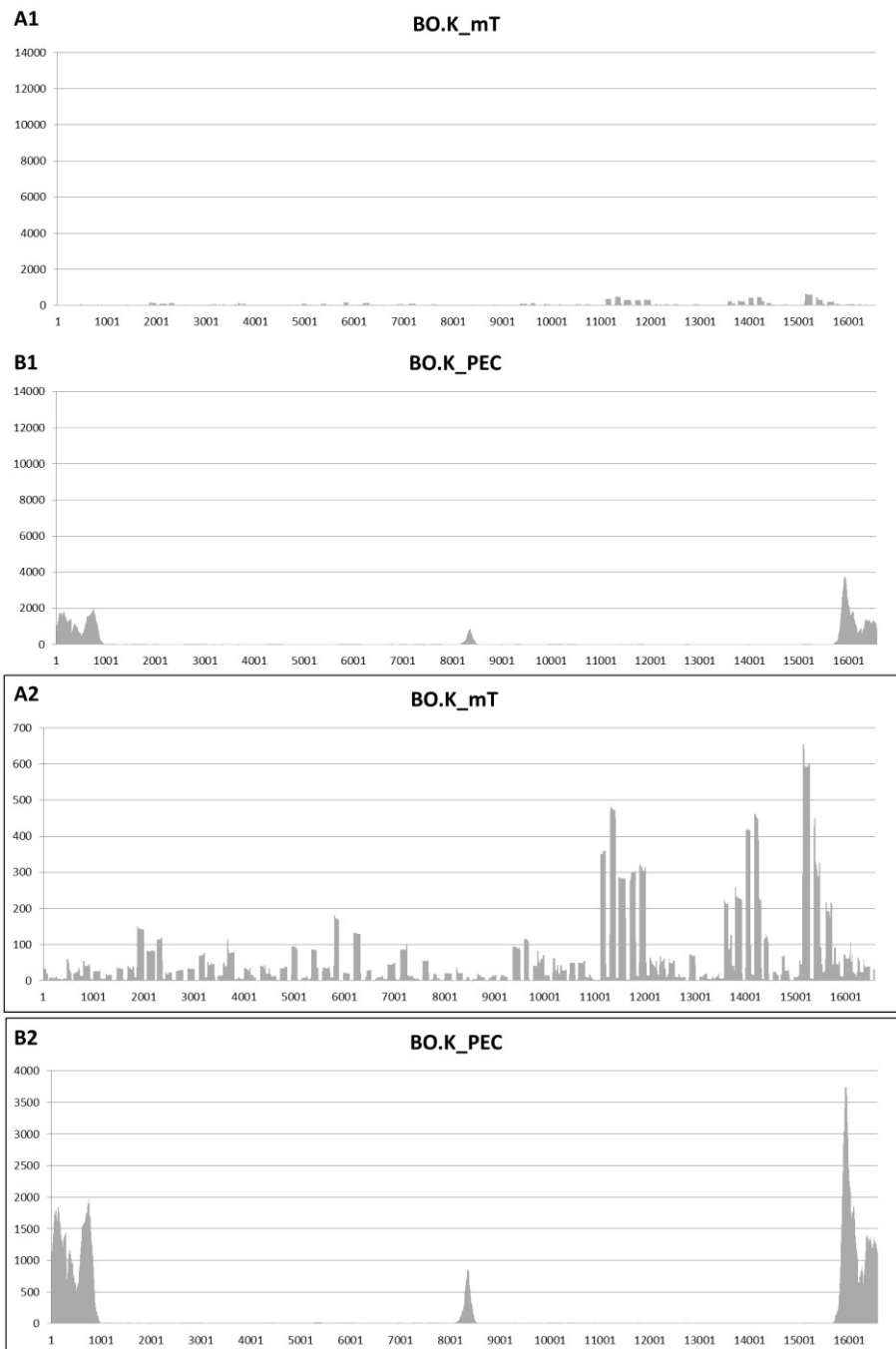
Figure 61. Mito Tiling (A) and PEC (B) coverages of the sample BO.K: A1 and B1. the coverages obtained normalised to the highest coverage recorded among the samples; A2 and B2. enlarged details of the coverages above.
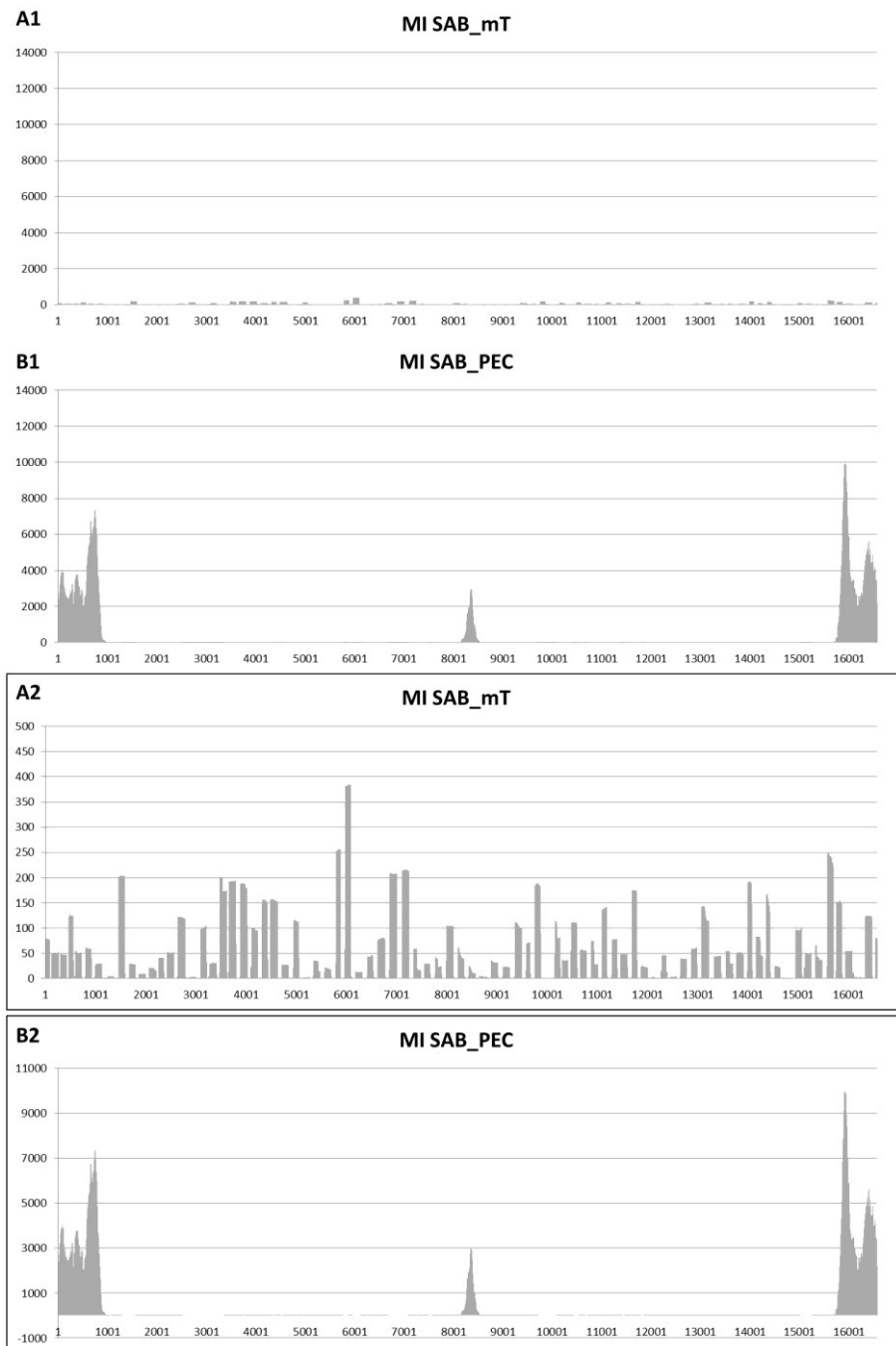
Figure 62. Mito Tiling (A) and PEC (B) coverages of the sample MI SAB: A1 and B1. the coverages obtained normalised to the highest coverage recorded among the samples; A2 and B2. enlarged details of the coverages above.

114
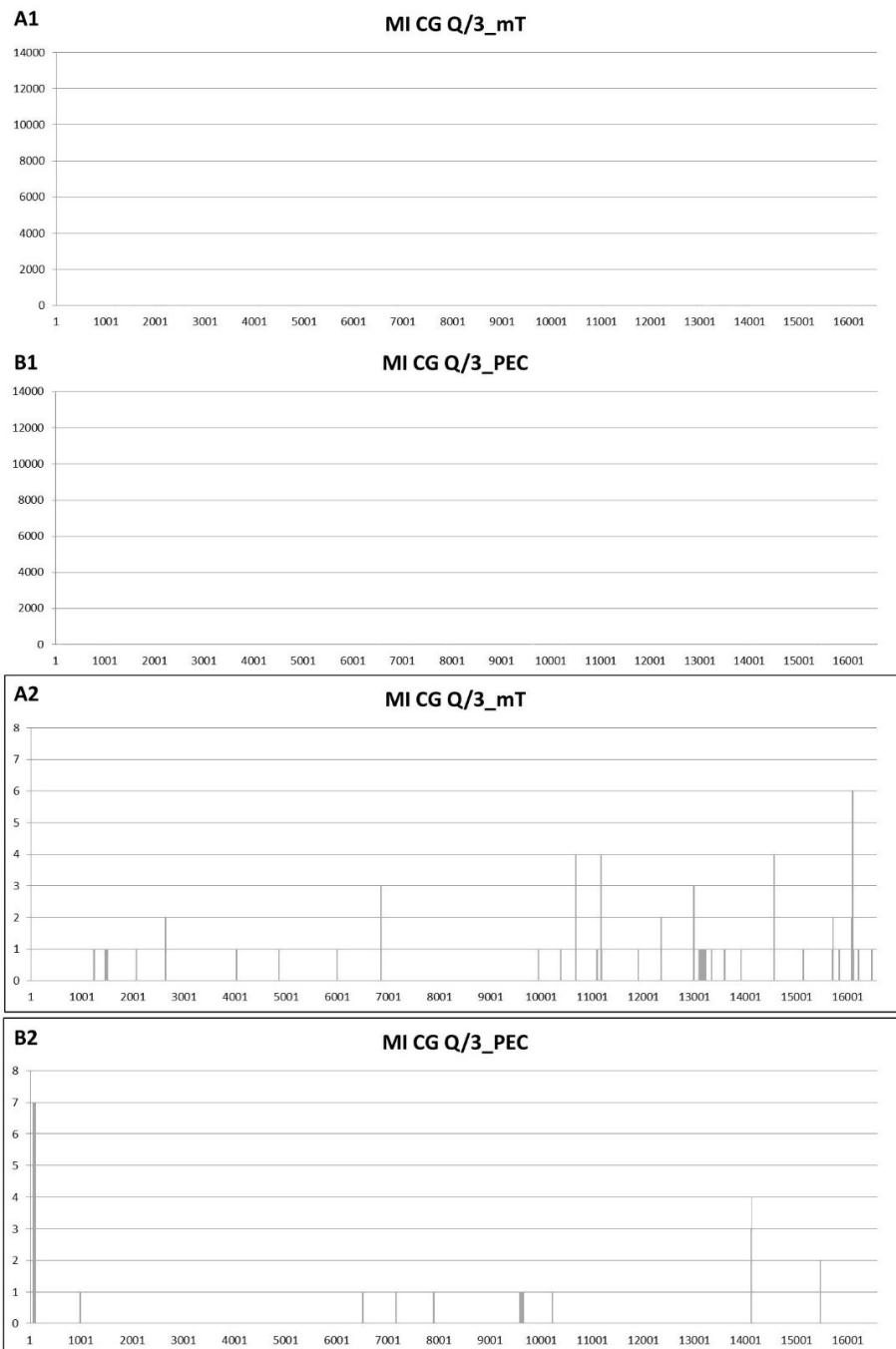
Figure 63. Mito Tiling (A) and PEC (B) coverages of the sample MI CG Q/3: A1 and B1. the coverages obtained normalised to the highest coverage recorded among the samples; A2 and B2. enlarged details of the coverages above.

115
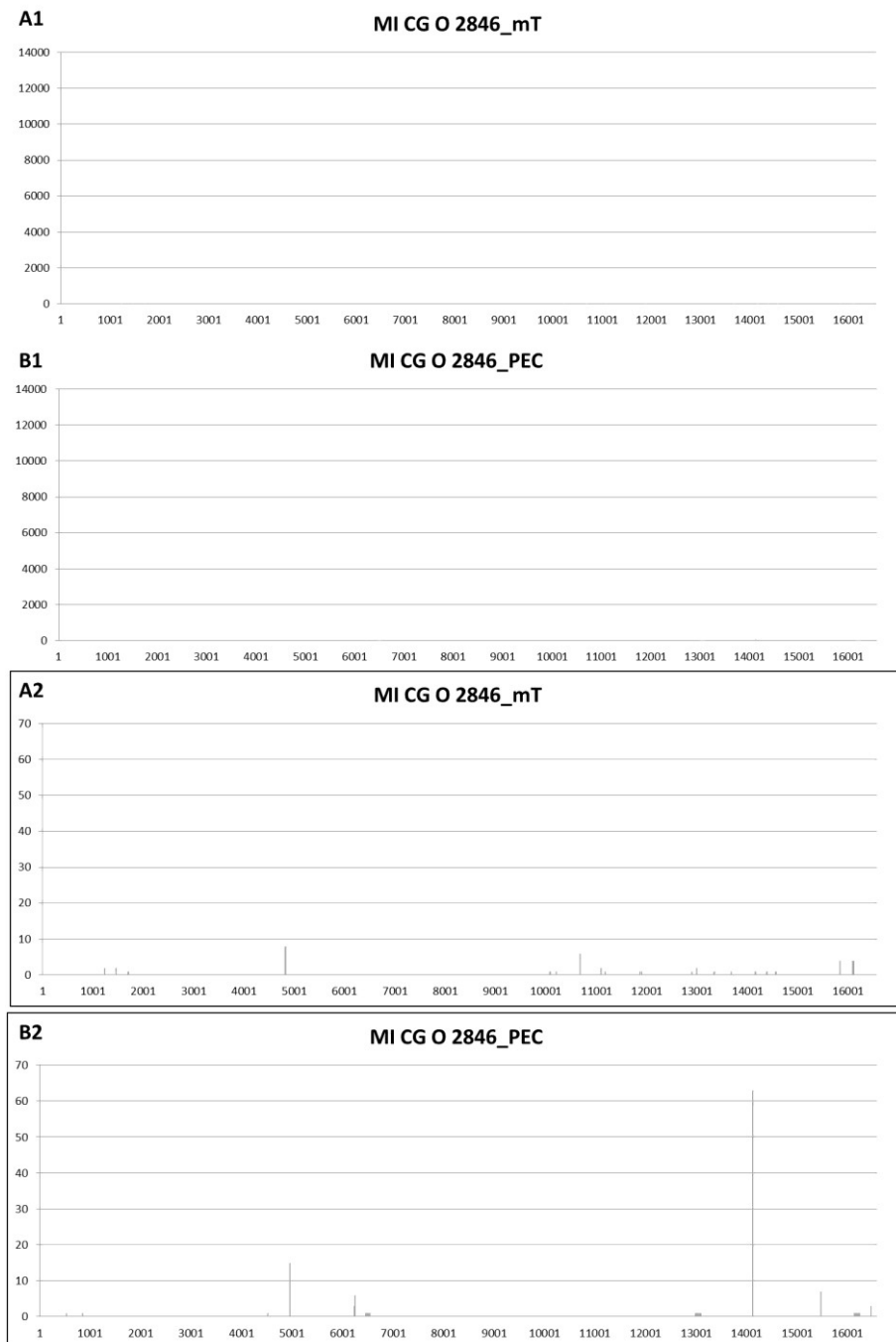
Figure 64. Mito Tiling (A) and PEC (B) coverages of the sample MI CG O2846: A1 and B1. the coverages obtained normalised to the highest coverage recorded among the samples; A2 and B2. enlarged details of the coverages above.

116
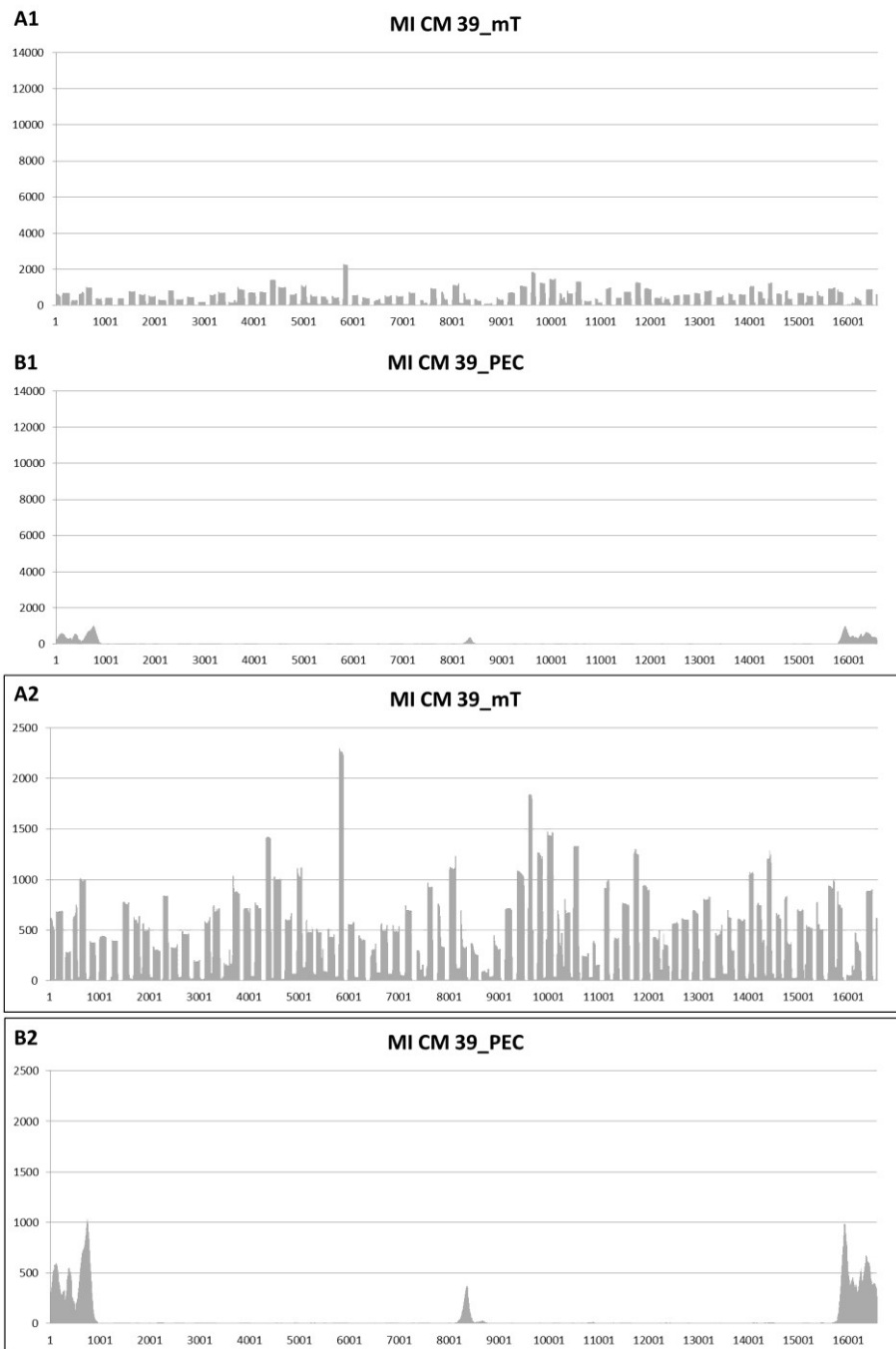
Figure 65. Mito Tiling (A) and PEC (B) coverages of the sample MI CM 39: A1 and B1. the coverages obtained normalised to the highest coverage recorded among the samples; A2 and B2. enlarged details of the coverages above.

117
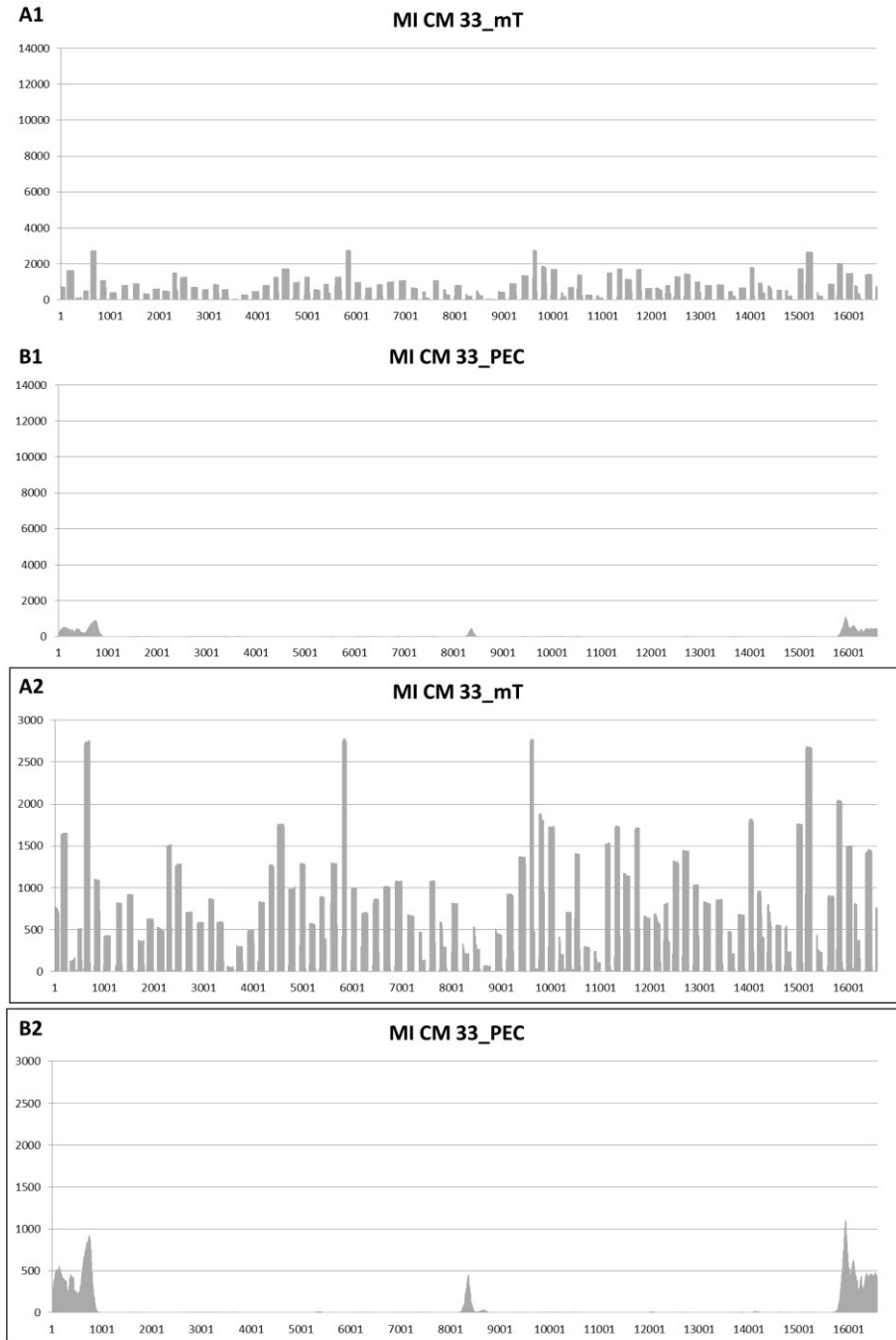
Figure 66. Mito Tiling (A) and PEC (B) coverages of the sample MI CM 33: A1 and B1. the coverages obtained normalised to the highest coverage recorded among the samples; A2 and B2. enlarged details of the coverages above.

118

Deamination events were detected in all samples. This condition complicated difficult data interpretation, especially for sample MI SAB (mito Tiling protocol), for which no haplotype could be extrapolated. Moreover, for some of the samples the low coverage did not allow the identification of mutations useful to attribute the corresponding haplogroup, as for samples MI UC (mito Tiling protocol), BG-BO-SC, MI CG Q/3 and MI CG O2846.

The haplotypes established for the same samples using the two protocols were compared and loaded on the EMPOP database for haplogroup attribution. Sequences were reanalysed looking for missing mutations, being aware that in many cases the coverage or the quality of the reads were too low to assign the corresponding mutation.

In Table 14, for each sample, the corresponding haplotypes and haplogroups are reported.

Phylogenetic analysis suggested the belonging of all individuals to the European population, and in particular Western, central and Eastern European, confirming the anthropological diagnosis. In fact, the analysis of morphological traits observed on the skull of those well-preserved remains suggested a Caucasoid origin. The only exception was sample MI PSA T49 whose skeletal remains were characterized by African traits, in contrast with the genetic result, which assigned this sample to haplogroup T2b. This haplogroup is indeed quite frequent in Europe (empop.online).

MI UC		BG-BO-SC		MI PSA T49		MI PSA T23		BO.K	
mT	*PEC*	*mT*	*PEC*	*mT*	*PEC*	*mT*	*PEC*	*mT*	*PEC*
-	N1b	-	-	T2b	T2b	H8	H8	J1c	J1c3
fail	73G	73G	73G	41T	41T	146C	146C	73G	73G
	152C	cv=0	263G	73G	73G	195C	195C	185A	185A
	263G	cv=0	279C	199C	199C	263G	263G	228A	228A
	315.1C	cv=0*	709A	263G	263G	309.1C	309.1C	263G	263G
	523del	cv=0*	750G	309.1C	309.1C	315.1C	315.1C	295T	295T
	524del	11719A	cv=0	315.1C	315.1C	524.1A	524.1A	315.1C	315.1C
	750G	15452A	cv=0*	319C	319C	524.2C	524.2C	462T	462T
	1719A	cv=0*	16126C	709A	709A	709A	709A	489C	489C
	4769G	cv=0*	16186T	750G	750G	750G	750G	750G	750G
	5471A			930A	930A	4769G	4769G	2706G	cv=0
	8251A			1438G	1438G	8860G	8860G	3202C	cv=0*
	8836G			1888A	cv=0	13101C	13101C	4048A	cv=0*
	8860G			2706G	cv=0	15326G	15326G	4063A	cv=0*
	11719A			4216C	4216C	16362C	16362C	4216C	cv=0*
	14766T			4769G	cv=0*	16519C	16519C	4769G	cv=0*
	16145A			4917G	cv=0*			7028T	cv=0
	16176G			5147A	5147A			10398G	cv=0
	16223T			7028T	cv=0			11251G	11251G
	16390A			8697A	8697A			11719A	cv=0*
	16519C			8860G	8860G			12612G	cv=0*
				10463C	cv=0			13708A	13708A
				11251G	11251G			cv=0	13934T
				11719A	cv=0			14173C	cv=0*
				11812G	cv=0			14766T	14766T
				13368A	13368A			14798C	cv=0*
				14233G	14233G			15326G	cv=0*
				14766T	14766T			15452A	15452A
				14905A	cv=0			16069T	16069T
				15326G	cv=0*			16126C	16126C
				15452A	15452A			16362C	16362C
				15607G	cv=0*				
				15928A	15928A				
				16126C	16126C				
				16294T	16294T				
				16296T	16296T				
				16304C	16304C				
				16519C	16519C				

MI SAB		MI CG Q/3		MI CG O2846		MI CM 39		MI CM 33	
mT	*PEC*	*mT*	*PEC*	*mT*	*PEC*	*mT*	*PEC*	*mT*	*PEC*
-	H1j3	-	-	-	-	U5a1c2a	U5a1c2a	U4c	U4
fail	263G	fail	73G	fail	fail	73G	73G	73G	73G
	309.1C		75A			183G	183G	195C	195C
	315.1C					263G	263G	263G	263G
	750G					-	309.1C	309.1C	309.1C
	5249C					315.1C	315.1C	315.1C	315.1C
	8161T					524.1A	524.1A	499A	499A
	8860G					524.2C	524.2C	750G	750G
	15326G					750G	750G	1438G	cv=0*
	16270T					1438G	cv=0	1811G	cv=0
	16519C					2706G	cv=0	2706G	cv=0*
						3197C	cv=0	4646C	4646C
						4769G	cv=0*	4769G	4769G
						7028T	cv=0*	4811G	cv=0*
						8860G	cv=0	5342T	5342T
						9477A	cv=0	5999C	cv=0*
						10544T	cv=0*	6047G	cv=0
						10792G	10792G	7028T	cv=0*
						11467G	cv=0	cv=0*	8860G
						11719A	11719A	10907C	cv=0*
						12308G	12308G	11332T	11332T
						12372A	12372A	11467G	cv=0*
						13617C	13617C	11719A	11719A
						14766T	14766T	12308G	12308G
						14793G	14793G	12372A	12372A
						15218G	cv=0	14620T	14620T
						15326G	15326G	14766T	14766T
						16192T	16192T	cv=0*	15326G
						16256T	16256T	15693C	15693C
						16270T	16270T	16179T	16179T
						16286T	16286T	16356C	16356C
						16320T	16320T	16512C	16512C
						16399G	16399G	16519C	16519C

Table 14. Haplotypes and haplogroups (in blue) for each sample. (mT: mito Tiling protocol; PEC: PEC protocol; cv=0: coverage 0; cv=0*: coverage eliminated from the sequence range because of the low quality of the reads).

121

3.2.5 Comparison between anthropological and genetic results

The results obtained from macroscopic and microscopic evaluations of the bone samples and DNA analyses are summarised in Table 15.

Since no statistical differences were observed in DNA recovery between the two replicates and between the two incubation protocols adopted for DNA extraction, only the results which were obtained from the overnight incubation are reported as mean of the two replicated tests. For mitochondrial DNA analysis, data from both approaches (PEC and mito Tiling) were considered to evaluate the degree of DNA preservation in each sample. In detail, samples were classified as follows:

- samples which gave a coverage greater than 1,000 with mito Tiling approach (labelled as "+++");

- samples which showed a coverage lower than 1,000 with mito Tiling approach (labelled as "++");

- samples which gave no successful results using mito Tiling protocol, but some data were obtained after capture enrichment (labelled as "+");

- samples which provided no information with both methods (labelled as "No data").

In order to compare the different analyses, the results were reported according to four categories:

- category 1 (green), well-preserved bone samples associated to stages 0-1 according to Behrensmeyer, stage 5 based on the Hedge's classification and "+++" for DNA analysis;

- category 2 (yellow), intermediate bone preservation associated to stage 2 according to Behrensmeyer, stage 4 based on the Hedge's classification and "++" for DNA analysis;

- category 3 (orange), low bone preservation associated to stages 3-4 according to Behrensmeyer, stages 3-2 based on the Hedge's classification and "+" for DNA analysis;

- category 4 (red), poorly-preserved bones associated to stages 5 according to Behrensmeyer, stages 1-0 based on the Hedge's classification and "No data" for DNA analysis.

Sample	Period	Macroscopic analysis	Microscopic analysis		DNA quantitation			Sex	mtDNA
			Calcified	Decalcified	Nuclear DNA pg/µl	mtGE/µl 143bp	mtGE/µl 60bp		
MI UC	I-V AD	1	0	0	0.92247	0.51514	4.24445	---	+
BG-BO-SC	VII-IX AD	4	0	1	0.47473	0.35307	1.67697	---	+
MI PSA T49	XIII-XV AD	1	2	2	2.91661	4.675924	76.79021	XY	+++
MI PSA T23	XIII-XV AD	3	2	2	7.41893	8.29649	111.55458	XY	++
BO.K	XIV AD	2	5	5	27.33664	51.69446	160.57898	XX	++
MI SAB	XVII AD	4	1	1	0.59942	12.31960	116.70593	---	++
MI CG Q/3	XVII-XVIII AD	3	3	3	0.12256	0.20681	0.35638	---	No data
MI CG O2846	XVII-XVIII AD	4	4	3	0.19095	Undet	0.44166	---	No data
MI CM 39	XX AD	2	5	5	23.55706	74.22729	277.73401	XY	+++
MI CM 33	XX AD	1	5	5	23.49568	177.09259	477.85938	XY	+++

Table 15. Comparison between osteological and genetic data (green: well-preserved bone samples; yellow: intermediate bone preservation; orange: low bone preservation; red: poorly-preserved bones). For DNA quantitation only the average amount of DNA extracted followed the overnight incubation protocol are shown because equal (i.e., statistically not different) to the two-days incubation protocol. To provide information about mtDNA, results obtained from both approaches are considered. ("No data": no information recovered with both approaches; +: low coverage in mito Tiling, but information obtained after enrichment; ++: coverage lower than 1000 in mito Tiling approach; +++: coverage greater than 1000 in mito Tiling approach).

Macroscopic analysis identified well-preserved bone structures in both the two most recent cemetery skeletal samples and in some ancient remains, such as those coming from Catholic University and S. Ambrogio. Similarly, a degraded macroscopic appearance was observed not only in ancient samples, but also in some recent bone elements belonging to the XVII-XVIII centuries AD.

On the opposite, microscopic evaluations of calcified and decalcified thin sections showed a degraded bone tissue in the most ancient samples and an increase of the preservation degree toward the most recent bones. Two cases diverged from this pattern: sample BO.K, where no destructive foci were visible and for which the sections were identical to the ones obtained for the contemporary bones, and sample MI SAB, which was severely affected by taphonomical activity that destroyed the majority of the bone tissue, leaving only few bone structures identifiable. Generally, calcified and decalcified thin sections showed similar scoring, except for few cases (BG-BO-SC and MI CG O2846) where however no great differences in tissue preservation were observed.

Comparing macroscopic and microscopic results no relation between the stages was observed, except for three cases as MI PSA T23, MI CG Q/3 and MI CM 33.

Concerning genetic data, quantitation results of the nuclear fragments showed very low values, especially for the most ancient samples and those coming from Cà Granda. The five bones which gave amounts greater than 3 pg were submitted to genetic sex diagnosis, giving successful results. Mitochondrial DNA analysis provided results in accordance with the age of the samples from very partial sequences for samples (MI UC and BG-BO-SC) to almost the complete genome (MI CM 33 and MI CM 39). The only exceptions were represented by samples MI PSA T49 and the two samples coming from the Cà Granda site (MI CG Q/3 and MI CG O2846) whose results diverged from this trend. Moreover, differences were observed between the quantitation data and the mitochondrial DNA results for samples MI PSA T49 and BO.K, for which indeed the mtDNA amounts produced sequences of different qualities.

Comparing genetic to macroscopic and microscopic results, no relation was observed except for the most recent cemetery samples.

4. PART 2:

A further experience in aDNA field

4.1 The Jena experience

The second part of this project was focused on the acquisition of training and specific skills on ancient DNA analysis. To this purpose, three months were spent at the Max Planck Institute for the Science of Human History in Jena (Germany), under the supervision of Prof. Wolfgang Haak, following trained personnel during the lab-work carried out on archaeological skeletal remains. Specifically, special care was given to how a laboratory dedicated to aDNA analysis has been set up and the protocols and methods established to recover information from highly degraded skeletal remains, such as those coming from the prehistoric period. A crucial part of the work was relative to the training on bioinformatic tools, essential for the authentication of the DNA results.

In an ancient DNA laboratory, particular attention was given to the problem of the contamination. To this purpose, specific operating procedures had been defined concerning the organization of the laboratory areas, the introduction of chemicals, consumables and samples in the lab, the training of the lab people and the decontamination procedures of the working areas and samples before pulverization (see sections *Laboratory Setup* and *Bone decontamination and pulverization*).

Specific protocols set up to recover also the small fragments contained in the aDNA samples (see section *DNA extraction*) were applied, as well as protocols to reduce the chemical modifications due to deamination events (see section *Library preparation*) and to enrichen samples in sequences of interest by DNA capture (see section *DNA capture*).

A further advantage comes from the introduction of the next generation sequencing technologies (see section *Shotgun sequencing*) which allowed the recovery even of short molecules whose sequence can be determined thus allowing the detection of the typical damage pattern (Stoneking and Krause, 2011; Krause and Pääbo, 2016).

Considering both the amount of data generated from these instruments and the difficulty to study ancient DNA, specific tools and software were developed in order to give a support in the analysis of the reads obtained and in the discrimination of authentic ancient DNA from modern and ancient contaminants (see sections *Analysis of the first data* and *DNA analysis*).

In the following paragraphs the procedures and protocols applied in the Jena laboratories are described.

4.1.1 Laboratory setup

The ancient DNA laboratories were located far away from other molecular biology laboratories to avoid contamination caused by modern samples. The pre-PCR ("clean room") and post PCR areas were physically separated and in the "clean room" specific measures were taken to avoid contamination of the samples. These included individual protection devices (IPD) such as head-to-toe suit, cuff, mask, glasses and multiple pairs of gloves. Different rooms were equipped, each one dedicated to a specific processing steps: sample collection, sample preparation, extraction and library preparation. The decontamination process involved bleaching, followed by UV exposure, of the surfaces, hoods and devices, chemicals and consumables which were introduced in the clean area, as well as UV decontamination of the room at the end of every batch of samples.

Before introducing samples inside the clean room, each falcon tube or plastic bag containing the sample was decontaminated using diluted sodium hypochlorite, followed by water and UV exposure for 15 minutes. Samples were subsequently transferred into a clean plastic bag and associated to an internal ID.

4.1.2 Decontamination and pulverization

Before pulverization, each samples was submitted to a decontamination process to remove any external contamination potentially located on the surface. Decontamination and sampling were performed in different ways for teeth and bones (femur or petrous bone), after a first UV radiation for 10 minutes per side.

Teeth were further cleaned using diluted sodium hypochlorite and cut in the neck, under the crown, using a dentist drill. Pulp and dentine were pulverized from both root and crown employing a dentist drill with a bits rotated at low speed. Root and crown powder were stored in different tubes.

The same approach was used to remove a thin layer of the external surface of the femur sample. A drill was used to collect the bone powder from the inner part of the samples.

Similarly, petrous bone was cleaned using diluted sodium hypochlorite and few millimetres of the external surface of the area of interest, containing the dense part of the otic capsule, were removed using a dentist drill. A small piece from that region was subsequently cut and pulverised employing a mill setting a vibrational frequency at 25 Hz.

About 50 mg of bone/tooth powder were collected from each sample and submitted to the following step.

4.1.3 DNA extraction

Considering the inverse correlation between the number of DNA fragments and their length, it is supposed that in highly degraded samples the majority of the information are located in short molecules. To this purpose, it is important to extract DNA using protocols which allow the recovery also of these fragments, especially when samples come from very ancient ages (Dabney *et al*, 2013a).

For this reason, Dabney and colleagues (2013a) improved a silica-based method, able to recover DNA fragments with molecular weight ≥ 30 bp.

The protocol involved an overnight extraction at 37°C adding to each sample 1 ml of the extraction buffer composed by 0.45 M EDTA and 0.25 mg/ml Proteinase K. The decalcification and lysis steps performed together allowed the recovery of the DNA from both the organic and the inorganic fractions, avoiding a loss of the genetic material eventually dissolved in the decalcification solution which is usually discarded after the first step.

The obtained supernatant, containing the genetic material, was mixed to a binding buffer containing 5 M guanidine hydrochloride, 40% isopropanol, 0.05% Tween 20 and 90 mM sodium acetate pH 5.2 and the overall solution transferred to High Pure Extender Assembly falcons from the High Pure Viral Nucleic Acid Large Volume kit (Roche) for the following purification (as reported by Korlević *et al*, 2015). The falcons tubes included a funnel, necessary to collect the huge amount of the binding solution, jointed to a silica spin column.

After centrifugation, the column was removed from the reservoir and place in a new collection tube. DNA bound to the silica membrane was purified through two consecutive washing steps and finally eluted using TET buffer (Tris-HCl 1mM- EDTA 5mM with 0.05% Tween 20).

4.1.4 Library preparation

Before sequencing, samples were converted into libraries. The aim was to immortalize DNA extracts, maintaining all the diversity, including bacterial DNA and contaminants, and allowed sample tracking and identification.

Samples were at first treated to avoid artificial base changes caused by miscoding lesions during DNA amplification, removing deaminated cytosines (uracils), following the partial UDG treatment protocol performed by Rohland and colleagues (2015), with some modifications (Figure 67).

The damage repair was carried out incubating a fraction of the DNA extract for 30 minutes at 37°C in a solution containing two main enzymes, uracil-DNA glycosylase (UDG) and endonuclease VIII, which catalysed respectively the removal of uracils, creating an abasic site, and the cleavage on both sides of the abasic site, leaving the 5'- and 3'-phosphate groups.

Because the damage pattern is useful to differentiate ancient DNA from modern contamination, the reaction was subsequently inhibited adding 3.6 µl

Uracil Glycosilase Inhibitor (UGI) and incubating samples for further 30 minutes at 37°C, to prevent a complete removal of the bases and preserve a small damage signal at the ends of fragments, necessary for the authentication of ancient molecules during sample analysis.

The next step of the protocol consisted in the blunt end repair with two enzymes, T4 polynucleotide kinase (PNK) and T4 DNA polymerase, necessary to create amenable extremities for adapters ligation.
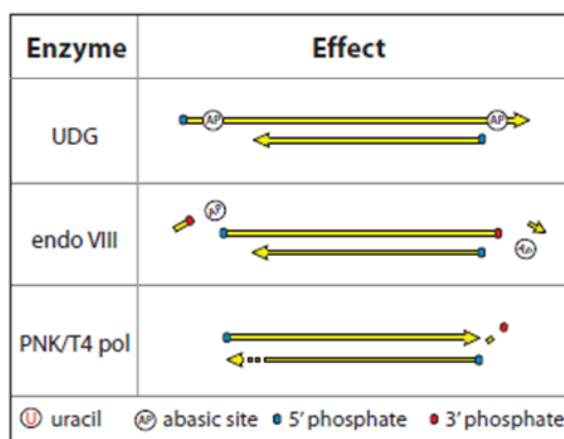
Figure 67. Representation of the UDG treatment with a specific focus on the enzymes involved: UDG (Uracil-DNA Glycosilase) removes uracils in the 5' and 3' protruding ends, creating an abasic site, which is cleaved on both sides by Endonuclease VIII, leaving phosphate groups on both ends. After the inhibition of the reaction, blunt ends are created using PNK (Polynucleotyde Kinase) to remove the 3'-phosphate group and phosphorylate the 5' end and T4 DNA Polymerase to fill in the 5' protruding ends and cleave the 3' ones (modified from Briggs *et al*, 2009).

To obtain double stranded libraries for Illumina sequencing, library preparation was performed as reported in Meyer and Kircher (2010) with some modifications (Figure 68).

To this purpose, two universal adapter sequences, named P5 and P7, were ligated to each fragment end. Because the adapters did not have the 5' phosphate, ligation occurred only in a single strand. Thereby, to remove nicks, a fill-in reaction was carried out after a purification step using a Bst polymerase, characterised by a strand-displacement activity. The enzyme was subsequently inactivated at 80°C.

The copies of the obtained libraries were measured through qPCR and amplified in a 100 µl PCR reaction with 10 cycles where indexes were added to create double-indexed library. Different barcodes combinations were used for each sample in order to identify the sample source of each sequence in the downstream analyses.

PCR products were subsequently purified and quantified again through qPCR and a second amplification was performed. After a further purification, sample

concentration was determined using the Agilent D1000 ScreenTapes assay (Agilent Technologies).
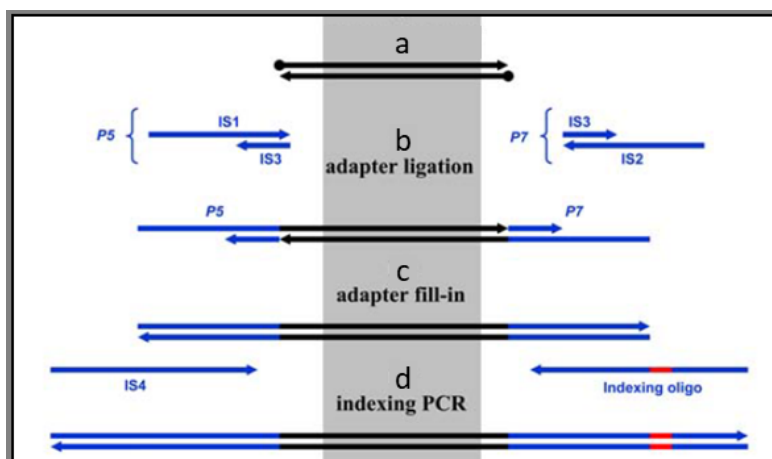
Figure 68. Representation of library preparation: a. fragments obtained after blunt end repair reaction; b. ligation of the two adapter sequences, P5 and P7; c. adapter fill-in to remove nicks; d. addition of indexes (in red) and full length adapters (in blue) (modified from Meyer and Kircher, 2010).

4.1.5 Shotgun sequencing

Libraries pool was subsequently submitted to shotgun sequencing using the HiSeq 4000 Sequencing System (Illumina). The aim was to get a picture of the molecular content of each skeletal element, including pathogens and contaminants.

Each library pool was diluted to 10 nM, pooled together in a single tube and prepared for sequencing as reported in the manufacturers' recommendations. Samples were identified thanks to the specific index combinations added during the previous step.

As all the Illumina's systems, HiSeq 4000 Sequencing System is characterized by a flow cell containing DNA probes complementary to libraries adapters. DNA libraries are bound to the probes and amplified, generating clonal clusters during a process called bridge amplification. For the HiSeq System, cluster amplification occurs in the c-Bot System, a separate instrument. After clustering, the patterned flow cell was loaded on the HiSeq 4000 Sequencing System to sequence samples.

Sequencing occurs through a sequencing by synthesis (SBS) technology, using reversible terminator-bound dNTPs. After the incorporation on the DNA strand, the single base is excited by a laser and the generated signal detected by a camera moving along the flow cell (www.illumina.com).

The peculiarity of HiSeq system lies on a new patterned flow cell technology. The flow has eight separated lanes, each one characterized by nanowells containing

DNA probes to capture the DNA libraries, to allow cluster generation inside each well and better resolution of the signal during imaging (Figure 69). In addition, a new chemistry promotes the binding and amplification of a single DNA strand within a well to ensure monoclonal cluster formation (www.illumina.com).
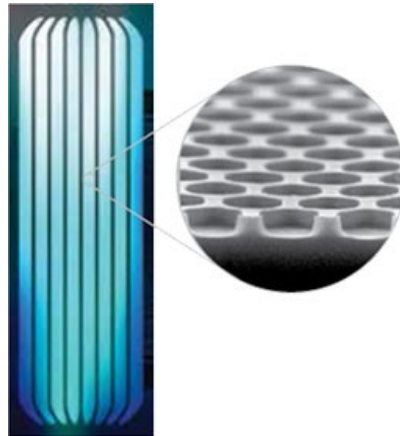
Figure 69. Patterned flow cell with billions of nanowells located on the surface (www.illumina.com).

Quality metrics data were generated by the Real-Time Analysis software of the sequencing system and examined using the Sequencing Analysis Viewer software (Illumina).

Moreover, the sequencing system created BCL files that were subsequently converted into FastQ files for the following analyses. During conversion, samples were "demultiplexed", that is reads with the same index combination were assigned to the corresponding sample.

4.1.6 Analysis of the first data

Shotgun data were analysed using Eager (Peltzer *et al*, 2016), a pipeline developed to process sequencing data from adapter clipping to the creation of fasta files. It is used for both modern and ancient DNA, even if in the analysis of this last DNA samples specific additional tools are selected.

It includes several tools to preprocess and map the reads, remove duplicates, evaluate contamination and damage and genotype samples. In addition, tools to authenticate and assess the quality of ancient DNA are included.

Eager needs fastQ or bam files and some information about samples, such as the organism type, the age of the dataset (modern or ancient), any DNA treatments (UDG treatment or capture data) and sequencing method (pair or single end data), as well as the reference genome.

The samples were analysed selecting the following tools for the different steps, applying the default settings used in ancient DNA analysis:

- FastQC (Andrew S, "FastQC: a quality control tool for high throughput sequence data", unpublished, 2010), to get information about the quality of raw data.
- AdapterRemoval, to clip adapter sequences and low quality bases from reads.
- BWA (Li and Durbin, 2009), to map reads to the reference.
- SAMtools (Li *et al*, 2009), to remove PCR duplicates.
- QualiMap, to obtain information about the average coverage and other statistics calculations.
- MapDamage, to calculate DNA damage, important for samples authentication.

Eager created a report in which several information for each sample were summarized. Among these, cluster factor, percentage of endogenous DNA and damage pattern are the most important.

Cluster factor gives a measure about the complexity of the sample. In particular, values approximately equal to 1.0 are associated to a good cluster factor and, especially, an high number of unique reads. It indicates how deep the samples have been sequenced.

The percentage of endogenous DNA specifies the content of endogenous DNA in a sample and the damage pattern gives the frequency of bases with damage in the first and second positions on both 5' and 3' ends.

4.1.7 DNA capture

Samples which gave good results after shotgun sequencing, especially percentage of endogenous DNA equal or greater than 0.1, cluster factor close to 1 and ancient damage pattern or an high mitochondrial/nuclear DNA ratio (for mtDNA capture), were submitted to DNA capture.

Human and mitochondrial DNA were captured using biotin labelled probes in two different experiments, named 1240k and mtDNA captures respectively, following the in-solution capture protocol described by Fu and colleagues (2013a), with some modifications.

To this purpose, the two different sets of probes which were characterized by human (1,240,000 SNPs) and mitochondrial specific sequences were selected and printed on distinct customized arrays.

The protocol consisted in an initial probe preparation to generate biotinylated libraries which were used for the following capture to isolate the DNA of interest using streptavidin beads.

According to the protocol by Fu and co-workers (Fu *et al*, 2013a), probes were cleaved from the array and submitted to the probe preparation procedure. The recovered fragments were composed by the oligonucleotide of interest (the probe) and the linker used to attach the sequence to the slide (Figure 70).

The first part involved the probe library preparation. To this purpose, the first library adapter, which was biotinylated and complementary to the linker, was ligated to the purified probes in a primer extension reaction carried out by Bst polymerase, generating a double-stranded fragment. A following blunt end repair was performed to remove the 3' protruding ends created by the polymerase and subsequently the second adapter was added. This was a double-stranded adapter, previously prepared, which bound the last two extremities of the fragment.

In the second part, several amplifications were carried out, among which a 3% low melting/1% high melting agarose gel with SYBR Safe DNA Gel Stain was performed in order to remove shorter fragments below the full-length probes.

To complete probe preparation, single-stranded DNA probes were generated.

A graphic representation of the protocol is reported in Figure 70.
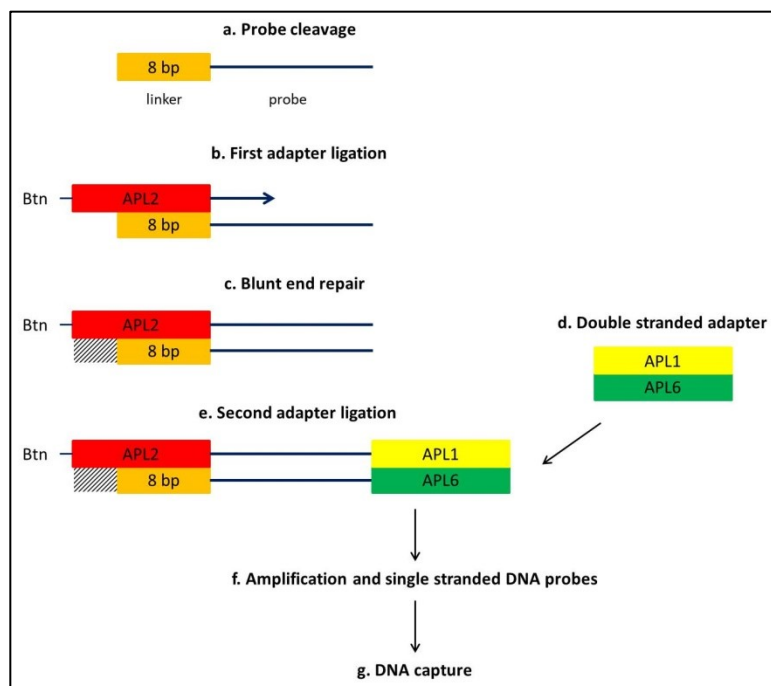
Figure 70. Probe library preparation protocol: **a.** Probes cleaved from the array, composed by the single strand oligonucleotide of interest (probe) and the linker used to attach the sequence to the slide; **b.** Ligation of the first adapter (Biotinylated, Btn) and synthesis of the second strand by Bst Polymerase; **c.** Blunt end repair carried out using T4 DNA Polymerase; **d.** Double stranded adapter generation **e.** Ligation of the second adapter to the probe; **f.** Amplification assays and single-stranded biotinylated DNA probes generation for the following DNA capture (**g.**).

Before capturing the DNA of interest, indexes used to mark each sample were blocked to prevent the binding to probes. After a denaturation step, the denatured libraries were added to the probe pool and to a pre-warmed (60°C) hybridization buffer and incubated at 65°C to promote DNA capture.

Captured DNA was subsequently purified twice, quantified to check the concentration and amplified. Samples were captured, purified and amplified again as already described and submitted to a reconditioning PCR to reduce heteroduplex.

The amplified product was subsequently pooled and prepared for the sequencing run performing a dilution of the pool to 10 nM.

Samples were sequenced in the HiSeq 4000 Sequencing System and the generated FastQ files analysed in Eager to look for some relevant statistics.

4.1.8 Analysis

Specific analyses were carried out after sample sequencing, which were different for nuclear and mitochondrial DNA. Here few examples of the analyses performed in the aDNA laboratory in Jena are described.

4.1.8.1 Nuclear DNA analyses

After 1240k capture, samples were genotyped and submitted to specific analyses to determine the genetic sex and to estimate X contamination. Moreover, PCA and admixture analyses could be performed to look at the phylogeny and investigate the genetic composition of each individual.

In the following paragraphs, genotyping, genetic sex determination and X contamination are described. Before starting with the analysis of the samples, at least two bases from each end were clipped in order to remove the observed damage pattern (C-T substitutions).

Genotyping

BAM files were processed to determine genotypes of the selected set of SNPs.

To genotype, samples were compared to a large data set of modern reference populations using a specific tools called SAMtools (Li *et al*, 2009).

Three output files were generated, reporting information about the genotype ("genotype file"), which gave a matrix of genotypes for each SNP per individual, SNP ("SNP file"), that included a list of the SNPs with the chromosome, position, reference and alternative alleles, and sample ("individual file"), that contained a list of individuals with sex and populations to which they belong.

Sex determination

Molecular sex was estimating using the approach described in Mittnik *et al* (2016).

The developed method compared the coverage on the X and Y chromosomes with the coverage on autosomes. Males, which have only one X chromosome, will give values equal to half of the total coverage on autosomes, while females should show the same. In the same way, the coverage on the Y chromosome should be half of the autosomal coverage in males, and equal to zero in females.

The calculations were performed using specific *scripts* to count the number of reads mapped on autosomes, X and Y chromosomes.

Subsequently, the ratios X /autosomal coverage and Y /autosomal coverage were determined and the results plotted in a 2D-scatter plot.

X contamination estimation

X contamination was estimated following the method described by Rasmussen and colleagues (2011).

The approach looked at reads mapping on the X chromosome. Reads with different bases at the same position could be associated to sequencing or mapping errors or to different individuals. In the last case sample is contaminated. In particular, contaminated samples should show a greater mismatch rate in more polymorphic loci.

To estimate contamination, polymorphic sites were used and their mismatch rate compared to the adjacent site, which were probably not polymorphic.

The analysis was performed using the ANGSD software (Korneliussen *et al*, 2014) to count how often each allele was seen in variable sites on the X chromosome and subsequently estimate the contamination.

The method gave a percentage of contamination for male individuals, which carried only one X chromosome and therefore one allele for each site.

4.1.8.2 Mitochondrial DNA analyses

Schmutzi

In the analysis of ancient DNA sequences, cytosine deamination and exogenous human contamination are the two main problems which make difficult the accurate assembly of the endogenous DNA to the reference mitogenome.

Schmutzi (Renaud *et al*, 2015) is an iterative approach developed to reconstruct the endogenous mtDNA and at the same time estimate the human contaminants. The endogenous sequence is inferred recognizing the endogenous nucleotides from the contaminants, providing a prior on contamination, deamination rate and length distribution of the fragments. On the other hand, contamination is predicted evaluating nucleotide differences between the endogenous DNA and a dataset of possible mitogenome contaminants. The base

calling and the identification of contaminants are repeated "until a stable contamination rate estimate is reached" (Renaud *et al*, 2015).

To this purpose, Schmutzi uses some internal programs: contDeam, endoCaller, mtCont and splitEndo. In the first iteration, contDeam estimates the deamination rate and the contamination prior, which are used by endoCaller to call the endogenous mitochondrial genome. Subsequently, contamination is re-estimated by mtCont, providing as input data the endogenous consensus sequence and the database of potential contaminants, and deamination rate and distribution of the fragment length are determined for both endogenous and contamination sequences by splitEndo. The outputs obtained from mtCont and splitEndo becomes the input of endoCaller to re-call the endogenous genome. This iterative steps are repated until a stable contamination rate is reached (Renaud *et al*, 2015). In Figure 71 a representation of this approach is described.
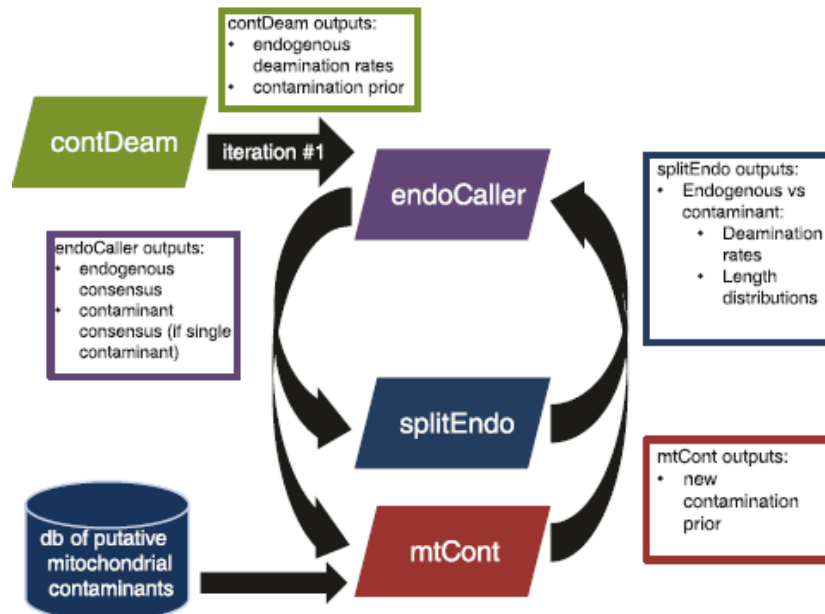
Figure 71. Schematic representation of the iterative approach performed by Schmutzi (Renaud *et al*, 2015).

Schmutzi creates for each iterative step files containing the contamination estimate and the endogenous mitochondrial consensus sequence. In particular, the endogenous file reports for each position of the mitogenome the base of the reference and that identified in the sample analysed, the quality and the average map quality, the coverage and the number of reads which support the base calling and finally the probability for each base.

The final files are usually considered for further genetic analyses to compare genomes and in phylogenetic studies.

5. DISCUSSION

The Milan project arises to describe the life and the evolution of the populations living in the urban area of Milan across two millennia (from the Roman Imperial Age to contemporary times) through a multidisciplinary approach involving archaeology, anthropology and genetics.

Archaeological and anthropological investigations revealed over time the great potential of this skeletal collection, supplying an interesting portrait of the demographic conditions and biological evolution of people living in different historical periods. A particular interest arose from paleopathological investigations which revealed the changes of the health status and the living conditions over time (Cattaneo and Gibelli, 2015), identifying also the appearance of pathology never detected in this area. Considering the intriguing results and the improvement of technologies to study ancient DNA, the Milan project would enrichen the profile of these populations also in terms of genetic evolution, origin and variations, identifying possible parental relationships between individuals and genetic relations with surrounding people.

This research activity was a pilot study on the Milan population and represented the first attempt to investigate this ancient material combining anthropology and genetics in order to assess the feasibility of reconstructing through DNA the history of the city through DNA analysis.

In the first part of the project, genetic investigations were carried out to enrichen anthropological data and archaeological information, focusing on genetic sex and on DNA markers useful to infer the most likely geographic origin of the samples. As no genetic studies had been previously performed on the skeletal collection of Milan, an additional purpose of these analyses was the evaluation of the DNA recovery rate from samples coming from different archaeological sites within the city and belonging to different historical periods.

In addition, to investigate how the different archaeological contexts could influence bone tissue appearance, macroscopic and microscopic analyses were performed, describing the state of preservation of each sample, as well as verifying a correlation with the DNA content.

To this purpose, ten intriguing skeletal remains excavated in seven different sites and belonging to different historical periods (from the Roman Imperial Age to the Contemporary Age) were selected and submitted to the osteological analyses. Among these, two individuals coming from the Cimitero Maggiore of Milan (Contemporary Age) were included as representative people of the contemporary local population. From each individual a bone sample was collected for the following analyses, selecting the diaphysis of the long bones of the lower limbs, when available, or of another long bone as suggested by other authors (Edson *et al*, 2004; Leney, 2006; Miloš *et al*, 2007; Prinz *et al*, 2007; Mundorff *et al*, 2009). The petrous bones (Gamba *et al*, 2014; Pinhasi *et al*, 2015; Rasmussen *et al*, 2014) could not be selected because it was no available or because it was difficult to sample for practical and anthropological reasons.

Concerning genetic analyses, taking all safety measures to avoid exogenous contamination, samples were processed using specific procedures to remove

contaminants on the surface of the bone samples, such as physical and chemical cleaning, and applying specific protocols generally used in the ancient DNA field and/or able to recover genetic information from degraded samples. These included an improved silica-based method, which could recover very fragmented and degraded DNA (molecular weight $\geq$ 30 bp), reducing the loss of the genetic material during DNA extraction and purification (Dabney *et al*, 2013a). In fact, performing a decalcification and a lysis steps together, the recovery of the DNA from both the organic and the inorganic fractions was carried out, without wasting molecules eventually dissolved in the decalcification solution which is usually discarded after the first step. In particular, each bone sample was extracted using basically the same extraction protocol, but incubating samples overnight and two-days respectively in order to verify if a greater DNA amount could be recovered with the second approach. Because statistically the two methods were demonstrated to have a similar DNA recovery rate, the extracts were considered equivalent and then submitted to genetic typing as two different aliquots of the same sample.

Taking into account the small fragment length and the low amount of DNA in ancient skeletal remains, the genetic analyses involved the study of short sequences to get information about genetic sex and mitochondrial DNA. In particular, short fragments localised on the X and Y chromosomes were analysed to determine sex (Madel *et al*, 2016) and small overlapping regions covering the entire mitogenome were amplified (mito Tiling protocol) to get mtDNA sequences. Moreover, to verify the possibility to enrichen samples in sequences of interest, especially for cases characterized by highly degraded DNA, mtDNA capture (PEC protocol) was carried out to recover the information useful to estimate the ethnic origin of the individuals. Even if these methods were designed to analyse very compromised samples in the forensic field such as mock crime scene and hair samples (Madel *et al*, 2016; Eduardoff *et al*, 2017), the ability of the protocols to provide successful results also from archaeological samples, analysing ancient bone specimens dated to the IV-V centuries AD and XII-XIII centuries AD, was proven by different authors (Madel *et al*, 2016; Eduardoff *et al*, 2017). In particular, in contrast to mito Tiling approach which provided information for almost the entire genome, PEC protocol was able to enrichen samples supplying data especially from the control region (positions 16,024-576). This feature is peculiar of the field for which this protocol was designed and applied. In fact, the hypervariable region is the target in forensic genetics due to its high variability and to the possibility to discriminate groups of individual belonging to different maternal lineages.

In order to get the genetic information, a set of primers specifically binding to the target regions was used. However, few additional data could be obtained outside the control region. In fact, information about some positions of the coding region could be obtained through: a. the unspecific binding of primers outside the CR; b. the capture of fragments belonging to the control region which even included few positions of the region adjacent the CR. In this last case, during

sequencing, data concerning the region close to the hypervariable one could be recovered (Eduardoff *et al*, 2017).

The haplotypes established from the same samples using the two protocols were compared and loaded on the EMPOP database for haplogroup attribution. Sequences were reanalysed looking for missing mutations, being aware that in many cases the coverage or the quality of the reads were too low to assign the corresponding mutation.

To verify the presence of a relation between DNA content and bone preservation, as well as to examine bone tissue appearance, the same bone elements submitted to genetic investigations were analysed through macroscopic and microscopic approaches. In particular, the evaluation of the appearance of the bone surface was performed following the criteria defined by Behrensmeyer (1978), focusing on the general state of bones and the presence of specific weathering markers, such as cracking and flaking. Moreover, calcified and decalcified bone thin sections were prepared and analysed in transmission light microscopy following the classification designed by Hedges and colleagues (1995), looking at the microscopic features of each section and specifically at the appearance of bone structures, the amount of intact bone and the presence of destructive foci caused by taphonomical activity.

The results can be summarised as follows, according to the biological and genetic profiles of the individuals selected for this study and the preservation of DNA and bone tissue of each collected sample. These are described and discussed in the following two paragraphs:

- Biological and genetic peculiarities of the skeletal samples

- DNA and bone tissue preservation in bone samples

Biological and genetic peculiarities of the skeletal samples

Anthropological and genetic data were compared in order to confirm the analyses performed on the skeletons, as well as to enrich the biological profile of each subject. In particular, each individual was described in terms of biological profile, including ancestry, sex, age at death, and providing further data about occupations, health status and when possible the circumstance and causes of death. Unfortunately, in some cases the state of preservation of the skeletal remains and/or DNA precluded the analysis and the evaluation of bone features and genetic sequences necessary for the diagnoses.

The individuals were all adults, six were males and four females as pinpointed by the morphological features observed especially on the skull and pelvis. Genetic sex diagnosis confirmed anthropological results for the five samples that were submitted to the analysis and that had the appropriate amount of nuclear DNA necessary to obtain reliable results (Madel *et al*, 2016). For the other samples the genetic sex could not be estimated due to the very low quantity of DNA recovered from the specimens.

Ancestry estimation could be performed for seven individuals: MI UC (Catholic University, I-V AD), MI PSA T23 and MI PSA T49 (Piazza S. Ambrogio, XIII-XV AD), BO.K (Bormio, Piazza del Kuerc, XIV AD), MI SAB (Viale Sabotino, XVII AD), CM 33 and CM 39 (Cimitero Maggiore, XX AD). In particular, genetic analyses were able to provide genetic information also from two skeletal remains whose geographic origin was uncertain by anthropological methods due to the bad state of preservation of the bone elements of interest. Only in three cases, sample BG-BO-SC (Bolgare, VII-IX AD) and the two samples coming from Cà Granda (XVII-XVIII AD), no results were obtained due to the low coverage that made the interpretation very complicated and prevented the identification of mutations useful to attribute the corresponding haplogroup.

Genetic data were acquired with both methods, mito Tiling and PEC protocols, for the majority of the samples, obtaining overlapping haplotypes, generally with a greater coverage using the second approach. This allowed to confirm mutations detected especially in the control region or to identify new mutations in those positions not recovered with the first approach. In two cases, samples MI SAB (Viale Sabotino, XVII AD) and MI UC (Catholic University, I-V AD), data were recovered only with the PEC protocol, suggesting that this approach was the best performing in the enrichment of highly degraded samples.

Many gaps were recorded in the mitochondrial sequences of most of the samples, analysed with both methods (mito Tiling and PEC). The only sample for which a complete mitogenome was obtained is sample MI PSA T49 (Piazza S. Ambrogio, XIII-XV AD) using the mito Tiling protocol.

Phylogenetic analysis suggested the belonging of all individuals to the European population, best fitting in haplogroups predominant in Western, central and Eastern Europe (empop.online), confirming anthropological data, except one case, MI PSA T49, which will be discussed later.

Moreover, anthropological analyses provided further interesting results, specifically concerning stress markers and pathology. In fact, in the majority of the archaeological skeletal remains (from the Roman Imperial Age to the Modern Age) signs associated to stress suffered during life and diseases (infectious, degenerative, metabolic, congenital diseases and tumour) were detected, highlighting the difficult living conditions of each period and the poverty that characterised these ages.

Below the biological profiles obtained for each individual combining anthropological and genetic results were reported and discussed.

MI UC

A male individual, 43-55 years old and 165-170 cm of height (Sguazza, 2010), whose skeletal remains were recovered in the Catholic University site, dated back to a long period, from the Roman Imperial Age to the Late Antiquity. The low amount of nuclear DNA recovered from the bone sample selected for the genetic analyses (average amount: 0.92-1.47 pg/µl) was not sufficient to evaluate the genetic sex, for which at least 3 pg/µl of nuclear DNA are needed (Madel *et al*, 2016).

The skull was classified as Caucasoid (Sguazza, 2010). The degraded state of the genetic material recovered from the sample selected did not allow to get mitochondrial DNA information with the mito Tiling protocol (very few sequences were recovered). Nevertheless, genetic data were obtained after DNA enrichment using the PEC protocol, allowing to identify mtDNA mutations useful to infer the peculiar haplogroup of this subject. This is one of the two samples where the PEC approach was proven to be efficient in those cases characterised by highly degraded DNA. Specifically, genetic analysis identified this samples as belonging to haplogroup N1b, very common in central and Eastern Europe (especially in Croatia, Hungary and Slovakia; empop.online) and in Middle East, enrichening the anthropological estimation.

The detection of osteoarthritis especially on the vertebral column and the peroneal sulcus in both calcanei pinpointed the practise of this individual to repetitive stress. Moreover, the detection of deformed humeri with a bowed diaphysis and relevant bone proliferation and eburnation in the head suggested possibly metabolic disorders and/or trauma (Ortner, 2003). All these data, together with the identification of remodelling fractures resulting from trauma occurred when he was alive, suggested the difficult living conditions of this subject.

BG

The skeletal remains recovered in the grave T156A in Bolgare-St. Chierico (VII-IX AD) belonged to a female individual, 42-87 years old and 158 cm of height. The complete skeleton was not recovered preventing the possibility to trace a complete biological profile. The exostosis detected in both femurs was ascribed to a sitting position and the inflammation at the flexor of the phalanxes suggested the use of thin objects tightened in the hand (Capasso *et al*, 1999; La Ferla, 2005).

The lack of the district of interest on the skull prevented ancestry estimation anthropologically. Similarly, the state of preservation of the genetic material did not allow the genetic sex diagnosis and the analysis of mitochondrial DNA.

In this case genetic data could be very interesting due to the peculiar information retrieved by the study of the population of Bolgare. In fact, anthropological analyses carried out on 460 individuals identified two main groups: individuals with northern or eastern European features, associated to the Lombard population, and individuals with central European or Mediterranean characteristics, typical of local people. Differences between the groups were observed also in stature: males and females belonging to the first group had an average stature greater than that of the entire sample. These results suggested a Lombard migration and especially the burial of Lombard people in Bolgare. Moreover, the observation of skulls with local and northern or eastern features highlighted the integration between the two populations (Sguazza *et al*, 2015). The possibility to provide genetic information about the origin of these individuals could confirm the anthropological data and the presence of Lombard people in Bolgare.

The preliminary genetic results obtained from skeletal remains coming from this archaeological site suggested a degraded state of preservation of the genetic substrate. This data was in agreement with the anthropological investigation of the bone tissue appearance, which highlighted a bad state of preservation of the skeletal remains. For this reason, further analyses are needed to evaluate whether the genetic information could be recovered from this population in order to verify the genetic origin of the two groups identified.

MI PSA T49

A young man (17-22 years old), height 175 cm, recovered in the grave 49 in Sant'Ambrogio and belonging to the Late Medieval period. Genetic sex results confirmed the anthropological diagnosis, even if the HRM profile obtained was composed only by the two Y specific peaks. The lack of the X chromosome peak was ascribed to dropout phenomena which can occur as described by Madel *et al* (2016) in cases of highly compromised materials, such as bones. Thereby, even if the X characteristic was missing, the presence of both Y specific peaks could be considered enough to assign the male genetic feature to this sample.

Anthropological analyses identified morphological and metric features typical of individuals of north African origin; diastema (the space between the two superior central incisors), prognathism and the absence of the nasal sill (the inferior edge of the nasal aperture) were some of the traits observed on the skull. The diagnosis was subsequently confirmed by the metric analysis carried out using the software ForDisc (Cattaneo *et al*, 2015a), which provided "Black" as result. By contrast, genetic analyses assigned this sample to haplogroup T2b, which is quite frequent in the European populations (empop.online). The same result was confirmed with both mito Tiling and PEC analyses, which were carried out using the two different DNA extracts (overnight and two-days incubation protocols). This unexpected but clear discrepancy between anthropological and genetic results

146

could be ascribed to a different ethnic origin of the two parents of this individual. In particular, an European origin of the mother is suggested, who transmitted this specific mitochondrial DNA to her son, and an African origin of the father, whose morphological features were observed on the skeletal remains of this subject. In this case, further analyses are needed to confirm the genetic results obtained, as well as to better understand the transmission of morphological features from the parents to their children.

This was not the only case of individuals of African origin in Milan; in fact, even during the excavation procedures occurred in the Catholic University (I-V AD), another African individual was identified, highlighting the multi-ethnic context of the city, which especially during the period as imperial capital city was a melting pot of people coming from different geographic areas (Cattaneo and Ravedoni, 1997; Cattaneo *et al*, 2001; Caporusso *et al*, 2007). Unfortunately, in this case no genetic analyses were performed and no genetic information are available about the haplotype and the haplogroup of this subject, which are very interesting for a comparison with the skeleton recovered in Piazza S. Ambrogio.

Another peculiarity of this subject is the detection of particular stress markers on the upper and lower limbs (feet), caused by repetitive stress suffered during life. In particular, the evidence of these signs on the inferior medial surface of both clavicles and on the right shoulder girdle suggested an overloaded of the upper limbs. On the other hand, the accessory articular facet observed on the head of the metatarsal bones was possibly associated to a hyperextension of the toes for consecutive periods. These features pinpointed the possible dedication of this person to hard activities, most likely correlated to the dragging and transferring of heavy weights, suggesting a possible involvement in working activities in a city construction site (Cattaneo *et al*, 2015a).

In addition, blunt trauma occurred probably 1-2 years before death were detected and hypoplasia was observed in both superior canine, suggesting a biological stress occurred during the childhood, hypothetically associated with a malnutrition status (Ortner, 2003).

MI PSA T23

These skeletal remains came from the same archaeological site (Piazza S. Ambrogio) and belonged to the same historical period (XIII-XV AD) of the "African" man, described before. The anthropological investigations revealed a Caucasoid young man (18-20 years old), with a height of 170 cm. These results were confirmed by the genetic analyses carried out on the left femur, which, in particular, ascribed this subject to haplogroup H8. Haplogroup H is one of the most frequent and most diverse in Europe (www.eupedia.com/europe/Haplogroup_H_mtDNA.shtml); specifically, H8 was found to be frequent in the South-Eastern Europe (empop.online).

Taphonomy seems to suggest that this individual was buried wrapped in a shroud, with two iron buckles on the side of each femur as indicating the possible presence of a belt used to clamp the tissue around the corpse. Anthropological

147

analyses revealed the presence of blunt force trauma on both forearms and legs, as well as sharp force trauma on the fourth lumbar vertebra and on the occipital region of the skull. The observed repetitive traumatism, at different body regions, provoked using different instruments, in a period close to death, was compatible with the hypothesis of a torture, and specifically the torture of the wheel, which was reserved to those individuals guilty of crimes against the community or for having killed the landlord's wife or son, or to criminals guilty of murders and finally to poisoners. In addition, the identification of anomalies on vertebrae and ribs supported the possible presence of congenital diseases, as well as the possibility that he was considered "different" by the community and therefore marginalized and subjected to discrimination (Cattaneo *et al*, 2015a).

BO.K

This is the second skeleton included in the study that comes from an archaeological site located outside Milan, in Bormio, dated back to the XIV century AD. The remains belonged to an adult female, 45-55 years old, 160-168 cm of height (Caruso *et al*, 2017b). The genetic analysis confirmed the morphological sex diagnosis, identifying only the X specific peak.

The entire skeleton was not recovered and the state of preservation of the bones did not allow the ancestry estimation. On the other hand, DNA extracted from the humerus submitted to the genetic analyses was well-preserved and allowed to provide further information about the geographical origin. The haplotypes obtained from the mito Tiling and the PEC protocols were basically equivalent assigning this sample to haplogroup j1c and j1c3, respectively. In particular, in this case, PEC protocol allowed to get a mutation not detected in the sequence recovered using the mito Tiling approach, because the low coverage of that specific position prevented its identification. In both cases, the haplogroups are predominant in the central and Eastern Europe and specifically j1c3 is more frequent in Romania and Bulgary (empop.online).

Paleopathological investigations identified osteoarthritis in some regions, especially vertebrae, feet phalanges and upper limbs, as well as enthesopathies (alteration of the enthesis, the insertion of tendons or ligaments to a bone), suggesting mechanical stress and in the second case, also, metabolic or inflammatory disorders (Cattaneo *et al*, 2015b; Caruso *et al*, 2017).

One of the most interesting piece of evidence was the identification of metastatic lesions, generally rare in archaeological populations where the low life expectancies reduced the risk of the onset. The particular pattern and distribution of the lesions suggested a metastatic tumour and, in particular, considering also the sex of the individual, breast cancer. However, other cancers could not be excluded as primary sites, even if further evidence supported the hypothesis. In fact, bone fractures were observed on six ribs and the right clavicle; the bone remodelling and the bone callus identified suggested *antemortem* trauma. In this case, a relation with a direct injury was not supposed and the trauma was probably related to the weakness of bones caused by the pathological condition. In fact, as reported in

Coleman (1997), bone fractures could occur, especially in ribs, in the 29% of individuals with breast cancer. To perform a more certain diagnosis, the pattern of the lesions observed on the skeletal remains were compared to two known tumour cases coming from the Contemporary Milan Skeletal Collection. In particular, a female individual with breast cancer and a male subject with lung cancer were selected for the comparison. The analyses identified a greater similarity with the woman than the man, in terms of size and shape, morphology and amount of metastatic lesions. In fact, the female skeleton showed multiple lesions characterised by different sizes and shapes, involving both the cortical and the cancellous bone. Differently, the male individual had a single lytic lesion surrounded by a depression area (Caruso *et al*, 2017b). In spite of the possible differences existing between the individuals from the Cimitero Maggiore of Milan and Bormio, metastatic cancer as possible diagnosis of the woman coming from the Valtellina site was supported by the contemporary population. This result confirmed how the study of modern skeletal individuals, such as those belonging to the collection of Milan, Cimitero Maggiore, whose information concerning the health status and the cause of death are available, represents an important source to investigate the appearance and the lesions that could be observed on bones, also at the microscopic level, in particular pathological status. This could help in the differential diagnosis of diseases not only in forensic cases, but also in ancient population studies.

MI SAB

A female individual, 45-55 years old, 172 cm of height (Saccomani, 2010). The skeleton was not complete and the fragmentation state of the skull did not allow ancestry estimation. Heavy DNA degradation resulted from the finding of frequent deamination events which made difficult data interpretation using the mito Tiling approach prevented the extrapolation of the corresponding haplogroup. On the other hand, as for sample MI UC, the PEC protocol was able to enrichen the sample, providing genetic data which allowed to assign this sample to haplogroup H1j3, predominant in North-Eastern Europe (empop.online).

The peculiarity of this individual lies on the site where she was found: Viale Sabotino (XVII century AD), a mass grave where 240 individuals were buried. The burial far away from the city and the position of the remains that suggested the urgency to release the bodies proposed a death caused by an epidemic event, probably the epidemic plague occurred between the 1629 and 1632 and described by Manzoni in his famous book "*I Promessi Sposi*" (Marsden and Pagani, 2008; Caruso *et al*, 2013b). The paleopathological analyses carried out on this population did not highlight any evidence of the plague in bones due to the severe progress of the disease. Only biomolecular analyses were performed so far, identifying the pathogen *Yersinia pestis* in the skeletal remains (Caruso *et al*, 2013b). To this purpose, further genetic analyses could be carried out to confirm the pathology, as well as to describe the genetic characteristics of ancient pathogens which marked our history in the past, as the plague during the XVII century.

The identification of osteoarthritis, as well as stress disorders as *cribra orbitalia* and hypoplasia, were suggestive of the indigence status and the difficult living conditions of this person. This was corroborated by the observation of lytic-proliferative lesions on the cranial vault, which suggested possibly a syphilis case at the advance stage: this woman could be paralysed and/or suffering of mental illness (Fornaciari and Giuffra, 2009; Saccomani, 2010). The paleopathological results complied with the health status recorded in the population recovered from the same archaeological site. In fact, congenital, nutritional, infectious diseases and metabolic disorders were detected on the skeletal remains, as well as degenerative pathology and trauma. All these observations confirmed the difficult life conditions that characterised this historical period included between the end of 1500 and the beginning of 1600, marked out by epidemic events, dearth and conflicts. In particular, the malnutrition and the bad hygienic conditions associated to the poverty and high urban concentration of this period increased the risk to the exposition of the population to infectious diseases (Caruso *et al*, 2013b).

MI CG Q/3 and O2846

The Cà Granda site was the most important hospital in Milan built during the XV century AD. The ossuary from which the selected samples came from was used for the burial of patients, during a period between the XVII and the XVIII centuries AD. Due to several factors, such as environmental factors and human activities, the site is characterised by a mass of commingled bone elements, estimated to belong nearly to 500,000 individuals (Sguazza, 2015). This explained why only two bones were available for the subjects selected from this site and included in this study.

MI CG Q/3 was a left femur anchylosed to the coxal bone, belonging to an adult male (30-45 years old). The anchyloses detected was ascribed to an *antemortem* trauma as evidenced by the radiological examination (Sguazza, 2015).

MI CG O2846 was a skull of an adult female individual, showing several lytic lesions in the cranial vault associated to syphilis at an advanced stage. As for the woman from Viale Sabotino, the pathology could have involved the nervous system and caused a paralysis or mental disorders, which could bring to a mental deterioration of the person (Friedrich *et al*, 2014; Crozatti *et al*, 2015; Sguazza, 2015).

The pathological state observed in these two bone elements is in agreement with the results obtained from the first anthropological analyses carried out on this skeletal material. In fact, degenerative, infectious and autoimmune diseases and metabolic disorders were detected. In particular, the identification of infectious diseases as osteomyelitis and syphilis suggested the bad hygienic conditions and the indigence of this historical period (Sguazza, 2015). Moreover, the results corroborated the burial of hospitalised persons, which needed care and medical treatments.

In both cases, the DNA recovered from the bone samples was severely degraded thus not allowing the possibility to get data about sex determination and genetic origin. The PEC protocol was unsuccessful as well as no data could be

recovered. In this case, further bone elements could be collected from the same site to verify whether the environmental conditions of the site were unsuitable for DNA preservation. However, the possibility to recover genetic information from such skeletal remains will allow to perform genetic analyses also concerning pathology, especially infectious diseases, in order to confirm the anthropological diagnosis and have data about ancient pathogen which could be compared with the previous and contemporary microorganisms for evolutionary studies.

MI CM 33 and 39

These skeletal remains recovered at the Cimitero Maggiore of Milan were representative of the contemporary local population (XX AD). They belonged to two adult males, both Caucasians, 49-73 years old MI CM 33 and over 65 years old MI CM 39, respectively. No particular diseases were detected, except osteoarthritis, especially on the vertebral column, which is a common evidence in old contemporary individuals.

This result was in agreement with the first analyses performed on a subset of this contemporary skeletal collection, which pinpointed degenerative diseases as predominant pathology detected in the individuals. However, the paleopathological analyses performed on this population identified other common pathology as tumour and metastasis, but also diseases that were typical of the past, such as tuberculosis and rickets, even if in lower percentages (Cappella, 2014).

Genetic analyses revealed the good state of preservation of the DNA, as expected from recent human remains and confirmed the anthropological diagnosis both for sex and ancestry. In particular, mitochondrial DNA analysis assigned sample MI CM 33 to haplogroup U4 and subhaplogroup U4c according to PEC and mito Tiling protocols, respectively. Mito Tiling approach detected the genetic information of the entire mitogenome thus identifying specific mutations in the coding region diagnostic of mtDNA haplogroups and subhaplogroups. Haplogroup U4c is predominant in Central Europe, and especially in Germany (empop.online). Similarly, mitochondrial DNA analyses carried out on sample MI CM 39 assigned it to haplogroups U5a1c2a with both methods. This haplogroup is very frequent in Eastern Europe, and particularly in Poland.

The preliminary results obtained highlighted the great potential of combining anthropological and genetics analyses on the Milan skeletal collection to trace the biological profiles of this population, providing intriguing results concerning the physical appearance of the individuals, occupations, pathology and the cause of death. Moreover, genetics is able to enrichen these profiles supplying interesting information about the geographic origin, also for those cases where the state of preservation of the bone samples of interest prevented the anthropological estimation. For this reason, the approaches used to investigate genetic sex and mitochondrial DNA were proven to be efficient in the recovery of genetic information from ancient skeletal remains excavated in different sites and belonging to different historical periods, confirming the results obtained by Madel

and colleagues (2016) and Eduardoff and co-workers (2017). In particular, comparing the two protocols used for mitochondrial DNA analysis, results suggested that the mito Tiling approach is a suitable protocol to recover a full genetic information from well-preserved samples; on the other hand, PEC protocol was proven to be very efficient in the enrichment of those samples which gave no results with the first approach or in the increase of the coverage of the others. To this purpose, the PEC approach is suggested to be a useful method to recover successfully genetic information from samples characterised by different taphonomical conditions.

Finally, these satisfactory preliminary results supported the possibility to carry on further analyses on this skeletal population. In particular, the analysis of a greater number of individuals coming from each site will allow to acquire interesting data concerning the genetic characterization, as well as evolution, of each population. In addition, considering the very interesting pathological portrait, the investigation of pathogens through genetics could confirm the pathological diagnosis and allow to investigate the genetic pattern of ancient microorganisms.

DNA and bone tissue preservation in bone samples

DNA extracted from all bones selected for this study showed the typical features of compromised and ancient samples (Pääbo *et* al, 2004; Briggs *et al*, 2007; Brotherton *et al*, 2007; Green *et al*, 2009; Sawyer *et al*, 2012; Dabney *et al*, 2013b). In fact, DNA characterised by short fragment length and peculiar damage pattern was identified at different levels depending on the sample, as expected for ancient specimens coming from different historical periods.

In particular, a highly fragmented DNA was detected from the quantitation results, which reported a good amount of the short 60 bp mitochondrial probe and a very low content of the long mtDNA and the nuclear DNA ones in most of the samples. Quantification results for the long mtDNA fragment (143 bp) and a very limited amount of nuclear DNA (average amount between 19 and 57 pg/µl) were scored only for the most recent bone samples (Late Medieval Period and Contemporary Age). The degradation state of the genetic material was confirmed by the following mitochondrial DNA analyses, whose results highlighted the presence of several gaps and low coverages in different nucleotide positions. Moreover, deamination events were recorded in all samples, complicating sequence interpretation, in some samples such as, for example, MI SAB (Viale Sabotino, XVII century AD).

Comparing quantitation and mitochondrial DNA analyses, comparable results were obtained: low amounts of DNA and poor mtDNA data, characterised by low coverage and several gaps, were recovered from the most ancient samples; while a greater DNA content and a better recovery of mitochondrial DNA information, were observed in the modern samples, highlighting the influence of time in the preservation of the genetic material.

To this pattern, some exceptions were observed. In fact, the two most recent samples selected from the Cà Granda site gave no results, after PEC enrichment and even if the bone sample selected (femur and occipital bone) are considered suitable elements for the genetic analysis (Edson *et al*, 2004; Leney, 2006; Miloš *et al*, 2007; Prinz *et al*, 2007; Edson *et al*, 2009; Mundorff *et al*, 2009). The severe degradation pattern of the genetic material did not allow any genetic analysis. In this case, environmental conditions could have influenced DNA preservation inside the osteological material. Cà Granda, indeed, is a mass grave, composed by tunnels located underground, where a "mass" of commingled bones was visible. Historical reports described the presence of water due to the proximity of a canal (Monforti,1694); thereby, samples could be submitted to a fluctuation of wet and dry conditions which may have altered the DNA preservation, increasing the decay of the genetic material as reported by Gilbert and colleagues (2005).

Another sample which gave surprising results after mito Tiling approach was the "African" man from S. Ambrogio (MI PSA T49, XIII-XV AD). From this sample in fact the complete mitogenome was recovered with a coverage (minimum 9x and maximum 12,645x) much higher than those of the two contemporary samples, which showed values included between 1x and about 2700x, as well as some gaps in the resulting sequence. However, it was not possible to obtain similar results from the second sample selected from the same site (MI PSA T23), confirming that differences in DNA preservation could be present in samples coming from different skeletal remains recovered in the same archaeological excavation. Moreover, the differences observed between the mitochondrial results and the quantitation data, as evidenced also in sample BO.K (Bormio Piazza del Kuerc, XIVAD), pinpointed the possible different preservation of the fragments selected to estimate the DNA amount compared to the complete sequence.

According to the results presented here, the selection of the relevant bone elements which may contain the best preserved DNA is a very important step in setting up the analytical strategy. In literature several studies tried to identify a method to predict DNA survival in bone using several approaches, such as amino acid racemisation, crystallinity index analysis, collagen content, bone histology, thermal age calculation and cytosine to uracil deamination patterns, (Poinar *et al*, 1996; Götherström *et al*, 2002; Haynes *et al*, 2002; Smith *et al*, 2003; Collins *et al*, 2009; Misner *et al*, 2009; Schwarz *et al*, 2009). Unfortunately nowadays the results produced from these different methods are contradictory and no reliable methods have been identified for sample screening (Gilbert *et al*, 2005).

In this study, the correlation between DNA content and bone preservation was tested evaluating both macroscopic and microscopic bone tissue appearance.

At the macroscopic level all the specimens were completely skeletonised and without remaining soft tissue. Bone samples were well preserved in the 50% of the cases (stages 1-2 following the criteria designed by Behrensmeyer, 1978), showing cracking and flaking limited to small portions of the bone diaphysis. In particular, the low weathering stages were associated both to ancient (Catholic University, I-V

5. DISCUSSION

AD; Piazza S. Ambrogio, XIII-XV AD; Bormio, Piazza del Kuerc, XIV AD) and more recent (Cimitero Maggiore, XX AD) samples. Similarly, taphonomic alterations were observed also in some modern cases (Viale Sabotino, XVII AD; Cà Granda, XVII AD), highlighting no relation with the historical periods. Moreover, only in half of the cases a relation with the DNA content was identified, suggesting that the method was not useful to predict DNA survival in bone samples, according to Misner and co-authors (2009). In particular, in this study the authors compared mitochondrial DNA quality and quantity to the skeletal and bone weathering, evidencing no relations between bone appearance and the DNA preservation. This is not an unexpected result; in fact, the authors argued that bone tissue and DNA preservation are influenced by different conditions (such as mechanical stress, that affect only bones and not DNA, or nuclease activities, which affect the genetic material and not bone tissue) or by the same factors which have a different effect on bone and DNA preservation (Misner *et al*, 2009).

Because macroscopic appearance was not correlated to DNA content, microscopic analyses were performed, investigating bone tissue preservation both on calcified and decalcified thin sections.

Calcified and decalcified thin sections analysis gave comparable results in the majority of the samples; only in few cases different scores were assigned. The few differences observed in the scoring were however probably ascribed to the different thickness of the slides. The greater thickness of calcified sections (100 µm), indeed, causes an accumulation of taphonomical alterations that limits the translucency and consequently does not allow a correct histological evaluation (Caruso *et al*, 2017a). On the other hand, decalcified sections are thinner (5 µm) and the removal of the mineral component allows to better investigate and analyse the remaining collagen fraction. This results in a well-preserved bone structure, where osteons and bone lamellae are intact, or in different levels of tissue degradation characterized by lytic lesions of several sizes, caused by microbial activity, dispersed among preserved lamellar structures. Moreover, the decalcification process of bone samples allowed to better detect the generalized destruction observed at the outer and inner layers of the sections, corresponding to the periosteal and endosteal surfaces, which appeared covered entirely by tunnelling in the most degraded calcified thin sections (scores 0-3). The removal of the mineral component confirmed in the majority of the cases the degradation of bone tissue in these portions, highlighting the destroying activity of microorganisms, which attack the body from the inside (the anaerobic gut microbiota) and the outside (aerobic bacteria and fungi) (Kontopoulos *et al*, 2016). On the other hand, the differences observed between the two sections confirmed once more the great potential of working on decalcified thin sections to examine bone tissue preservation, also in ancient specimens, allowing a better analysis of bone tissue, whose appearance could be hidden by microbial activity if the calcified approach was used (Caruso *et al*, 2017a).

The state of preservation of bone sections seems to be related to the historical periods, confirming the results already observed by Caruso and colleagues (2017a).

Generally, low OHI indexes (scores 0-1, following the Hedges' stadiation, 1995) were associated to ancient samples (Catholic University, I-V AD; Bolgare, VII-IX AD) and higher OHI scores (scores 4-5) to modern and contemporary specimens (Cà Granda, XVII AD; Cimitero Maggiore, XX AD). In the first case, the bone tissue was highly degraded and no structures could be visible; in the second case, osteons and lamellae are well-preserved and no or very few traces of microbial activities were detected.

Two exceptions to this trend were observed; in fact, samples BO.K (Bormio Piazza del Kuerc, XIV AD) and MI SAB (Viale Sabotino, XVII century AD) showed a well-preserved, similar to cemetery samples, and a highly degraded bone tissue respectively in both calcified and decalcified thin sections, confirming the influence of environmental factors on bone tissue preservation (Jans *et al*, 2002; Mello *et al*, 2017). Despite of the different location of these two archaeological sites, no influence of the soil composition could be hypothesized. In fact, sediment composition is nearly homogenous in the city and no particular differences in the geological structure, acidity, soil hydrology, biological factors and temperatures have been detected yet.

In spite of the influence of time on both microscopic appearance and DNA degradation, no relation was detected; thereby, as for macroscopic analysis, these results pinpointed the histological investigation of bone tissue appearance as a non-useful method to predict DNA survival.

Comparing macroscopic and microscopic data, several differences in bone preservation were detected; in particular, in some cases a well-preserved bone surface hid a severely degraded microscopic structure, highlighting a marked difference between bone surface appearance and collagen loss (Kontopoulos *et al*, 2016). These results supported the importance to not analyse bone preservation only at macroscopic level, but suggested histology as an important tool to examine bone tissue appearance and investigate taphonomical alterations in bone samples as already reported in several studies in literature (Jans *et al*, 2002; Kontopoulos *et al*, 2016).

No huge differences were observed between samples coming from the same archaeological sites, as well as no relation between the bone sample selected for the analyses and bone tissue preservation was detected.

The results obtained highlighted the influence of time and environment on both DNA and bone tissue degradation. Because no relation between DNA and bone preservation was detected, further researches are needed to confirm this data and to identify other useful markers to predict the survival of the genetic substrate in ancient bone samples. Moreover, the differences observed in macroscopic and microscopic analyses of bone appearance suggested the importance to introduce histology, especially of decalcified thin sections, to examine the preservation of the bone tissue, as already reported. In fact, decalcification allows to better analyse bone appearance removing possible taphonomical alterations which could alter the examination in calcified thin sections. However, the small sample size limits the

possibility to draw final conclusions concerning both bone weathering and its relation with the DNA content. In particular, considering the great potential of the skeletal material available, excavated from different archaeological sites and belonging to different historical periods, further analyses from more samples are needed to better investigate these issues.

Further prospectives

After the international experience carried out at the Max Planck Institute for the Science of Human History in Jena, more skills in the ancient DNA field were acquired. In particular, it was evident the difficult task to obtain reliable data from archaeological remains and the importance to set up both laboratory facilities and protocols in order to limit exogenous contamination and to recover the authentic genetic information.

To this purpose, the aDNA laboratory in Jena is set up positioning the ancient DNA facilities far away other molecular laboratories to avoid contamination with modern DNA, separating pre- and post-PCR areas and supplying the work space with high-tech positive air-pressure and UV systems for decontamination processes, following the guidelines reported in literature (Hofreiter *et al*, 2001b; Pääbo *et al*, 2004; and Willerslev and Copper, 2004). Similarly, training were carried out to provide specific information concerning the operating procedures for workin in the "clean room", which is the area dedicated to the study of aDNA from sampling to library preparation, including the wearing of protective clothing, decontamination procedures of consumables and chemicals, as well as of the room itself.

Sample processing is performed as reported in the most important and successful studies published in literature, starting from sample selection (Prinz *et al*, 2007; Higgins and Austin, 2013; Gamba *et al*, 2014; Pinhasi *et al*, 2014; Rasmussen *et al*, 2014) and decontamination (Lazaridis *et al*, 2014; Fehren-Schmitz *et al*, 2015; Heupink *et al*, 2016; Llamas *et al*, 2016; Valverde *et al*, 2016; Posth *et al*, 2017; Schuenemann *et al*, 2017; Suppersberger Hamre *et al*, 2017), DNA extraction (Dabney *et al*, 2013a), library preparation (Meyer *et al*, 2010; Rohland *et al*, 2015) and DNA capture (Fu *et al*, 2013a). These methods supported by next generation sequencing allow to recover information from short molecules, sequencing the entire fragment and therefore identifying the damage pattern at the extremities useful to distinguish ancient DNA from contaminants (Key *et al*, 2017).

The workflow involved a first analysis of the extracted DNA by shotgun sequencing to look into the DNA content of each specimens, in terms of endogenous DNA, microbial DNA and contaminants. Then, on the basis of the results obtained DNA capture is performed to enrichen samples in mitochondrial DNA, nuclear DNA or pathogen DNA. Moreover, considered the complexity of the results obtained after sequencing and the difficulties to study highly degraded material, another important step involves bioinformatics. To this purpose, specific tools and programs were designed and continuously revised and updated to elaborate results after next generation sequencing and DNA capture, able to discriminate ancient endogenous DNA from modern contaminants (Key *et al*, 2017).

For this reason, considered the degraded pattern observed in the samples characterised in this study, further analyses will be carried out on these skeletal remains at the Max Planck Institute for the Science of Human History in Jena to

further investigate the content of each selected samples, in terms of both human and microbial DNA, as well as contaminants, using a shotgun sequencing approach. Moreover, other markers different from mtDNA could be studied, such as human nuclear DNA and pathogen DNA, by in-solution capture approaches. This could allowed a better understanding of the genetic origin of these individuals, confirming, at the same time, the paleopathological diagnoses performed on the skeletal remains. To this aim, bioinformatic analyses will be carried out to confirm the ancient origin of the extracted DNA using the specific tools able to evaluate the damage pattern at the fragment ends and to discriminate the endogenous DNA from exogenous contaminants.

6. CONCLUSIONS

In conclusion, this research project represented an articulated pilot study carried out to investigate the Milan skeletal collection through a multidisciplinary approach, involving anthropology and genetics. The analyses performed on ten individuals recovered during excavation procedures from different archaeological sites and belonging to different historical periods highlighted the great potential of the proposed approach. The combination of anthropological and genetic analyses produced, in fact, intriguing results regarding the biological profile and pathology. In particular, anthropology provided interesting data about stress markers and pathology (infectious, degenerative, metabolic, congenital diseases and tumour) which suggested the difficult living conditions and the situation of extreme poverty of people during these historical periods. On the other hand, genetics was able to enrichen biological profiles supplying interesting information about the geographic origin, also for some of the cases for which the state of preservation of the bone samples did not allow the anthropological estimation. For this reason, the molecular approaches used to investigate genetic sex and mitochondrial DNA were proven to be efficient in the recovery of genetic information from ancient skeletal remains coming from different archaeological contexts. In particular, mito Tiling approach was shown to be a suitable protocol for the recovery of consistent genetic information from well-preserved samples; on the other hand, PEC protocol was proven to be very efficient in the enrichment of those samples which gave no results with the first approach or in the increase of the coverage of the others. To this purpose, the PEC approach is suggested to be a useful method to recover successfully genetic information from samples characterised by different taphonomical conditions. Considering the successful results obtained, the genetic approach used in this study may be applied to forensic casework for which the conventional approach gave no results due to DNA severe degradation.

Moreover, genetic analyses pointed out degraded DNA in the bone samples characterised by very short fragments length, low amounts of DNA and peculiar damage patterns. Time since deposition and environmental conditions influenced both DNA and bone tissue preservation, even if macroscopic bone appearance was not correlated to the historical period. In spite of this similarity, no relation between bone weathering and genetic data was detected, even if the small sample size of the selected sample limits the final conclusion about a possible relation. Moreover, differences detected between macroscopic and microscopic analyses highlighted the importance of including histological analysis in the evaluation of bone tissue preservation and taphonomical alterations. In particular, decalcified thin sections were considered a valid approach to examine microscopic bone structures, especially in archaeological bone samples, where microbial activity can hide the bone appearance in calcified sections preventing a correct analysis.

Finally, these satisfactory preliminary results supported the possibility to carry on further analyses on this skeletal population. For this reason, some of the most intriguing samples will be submitted, in a future perspective, to further genetic analyses in ancient DNA dedicated laboratory in Jena, applying specific procedures and protocols used in the ancient DNA field, in order to deeply investigate the

6. CONCLUSIONS

DNA content (both human and microbial) and assess the feasibility to analyse other interesting genetic markers, both human and microbial. The possibility to get additional information able to enrichen the biological profile of the individuals can be useful in providing also data about their pathological conditions. In addition, the new data could confirm the paleopathological diagnosis and get interesting information about ancient microorganisms. As a future perspective, an increase in the number of individuals collected from each site could be planned, to be submitted to the set of specific genetic procedures thus providing a portrait of the evolution of the Milan population over time. Moreover, the analysis of further samples coming from the different archaeological contexts could supply interesting information concerning the preservation of bone tissue in the different conditions and verify its relation with the DNA content.

REFERENCES

Alaeddini R, Walsh SJ, Abbas A. Forensic implications of genetic analyses from degraded DNA--a review. Forensic Sci Int Genet. 2010; 4: 148-157.

AlQahtani SJ, Hector MP, Liversidge HM. Brief communication: The London atlas of human tooth development and eruption. Am J Phys Anthropol. 2010; 142: 481-490.

Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG. Sequence and organization of the human mitochondrial genome. Nature. 1981; 290: 457-465.

Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Genet. 1999; 23: 147.

Barlassina S. Le tipologie funerarie. In Lusuardi Siena S, Rossignani MP, Sannazaro M (a cura di), L'abitato, la necropoli, il monastero. Evoluzione di un comparto del suburbio Milanese alla luce degli scavi nei cortili dell'Università Cattolica. Milano. 2011; pp. 86-93.

Bass WM. Human osteology: a laboratory and field manual. USA: Missouri Archaeological Society. 1995.

Behrensmeyer AK. Taphonomic and ecologic information from bone weathering. Paleobiology. 1978; 4.2: 150-162.

Binladen J, Willerslev E. Why study ancient DNA damage. J Nordic Archaeol Sci. 2010; 14: 11-14.

Boessenkool S, Hanghøj K, Nistelberger HM, Der Sarkissian C, Gondek AT, Orlando L, Barrett JH, Star B. Combining bleach and mild predigestion improves ancient DNA recovery from bones. Mol Ecol Resour. 2017; 17: 742-751.

Briggs AW, Stenzel U, Johnson PL, Green RE, Kelso J, Prüfer K, Meyer M, Krause J, Ronan MT, Lachmann M, Pääbo S. Patterns of damage in genomic DNA sequences from a Neandertal. Proc Natl Acad Sci U S A. 2007; 104: 14616-14621.

Briggs AW, Good JM, Green RE, Krause J, Maricic T, Stenzel U, Lalueza-Fox C, Rudan P, Brajkovic D, Kucan Z, Gusic I, Schmitz R, Doronichev VB, Golovanova LV, de la Rasilla M *et al*. Targeted retrieval and analysis of five Neandertal mtDNA genomes. Science. 2009; 325: 318-321.

Briggs AW, Stenzel U, Meyer M, Krause J, Kircher M, Pääbo S. Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. Nucleic Acids Res. 2010; 38: e87.

Brooks S, Suchey JM. Skeletal age determination based on the os pubis: A comparison of the Acsádi-Nemeskéri and Suchey-Brooks methods. Human Evolution. 1990; 5: 227–238.

Brotherton P, Endicott P, Sanchez JJ, Beaumont M, Barnett R, Austin J, Cooper A. Novel high-resolution characterization of ancient DNA reveals C > U-type base modification events as the sole cause of post mortem miscoding lesions. Nucleic Acids Res. 2007; 35: 5717-5728.

Brown TA. Genomi 3. EdiSES, III edizione. 2008.

Brown WM, George M Jr, Wilson AC. Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci U S A. 1979; 76: 1967-1971.

Brundin M, Figdor D, Sundqvist G, Sjögren U. DNA binding to hydroxyapatite: a potential mechanism for preservation of microbial DNA. J Endod. 2013; 39: 211-216.

Calafell F, Larmuseau MHD. The Y chromosome as the most popular marker in genetic genealogy benefits interdisciplinary research. Hum Genet. 2017;136: 559-573.

Campos PF, Craig OE, Turner-Walker G, Peacock E, Willerslev E, Gilbert MT. DNA in ancient bone - where is it located and how should we extract it? Ann Anat. 2012; 194: 7-16.

Cano RJ, Poinar HN, Pieniazek NJ, Acra A, Poinar GO Jr. Amplification and sequencing of DNA from a 120-135-million-year-old weevil. Nature. 1993; 363: 536-538.

Capasso L, Kennedy KAR, Wilczak CA. Atlas of occupational markers on human remains. Edigrafital S.P.A., Teramo. 1999.

Caporusso D, Donati MT, Masseroli S, Tibiletti T. Immagini di Mediolanum. Archeologia e storia di Milano dal V secolo a.C. al V secolo d.C. Milano. 2007.

Cappella A. Interpretation of trauma and taphonomy in a modern known skeletal population: implications for forensic anthropology. Tesi di Dottorato in Scienze Morfologiche. 2014.

Cappella A, Amadasi A, Castoldi E, Mazzarelli D, Gaudio D, Cattaneo C. The difficult task of assessing perimortem and postmortem fractures on the skeleton: a blind text on 210 fractures of known origin. J Forensic Sci. 2014; 59: 1598-1601.

Cappella A, Bertoglio B, Castoldi E, Maderna E, Di Giancamillo A, Domeneghini C, Andreola S, Cattaneo C. The taphonomy of blood components in decomposing bone and its relevance to physical anthropology. Am J Phys Anthropol. 2015; 158: 636-645.

Cappella A, Cummaudo M, Arrigoni E, Collini F, Cattaneo C. The Issue of Age Estimation in a Modern Skeletal Population: Are Even the More Modern Current Aging Methods Satisfactory for the Elderly? J Forensic Sci. 2017; 62: 12-17.

Caramelli D. Antropologia Molecolare. Manuale di base. Firenze University Press. 2009.

Carlessi M, Kluzer A. Il cuore dell'antico Ospedale Maggiore di Milano. I luoghi dell'Archivio e la chiesa della Beata Vergine Annunciata, Cinisello Balsamo, Milano: Silvana Editoriale. 2011.

Carlessi M, Kluzer A. Storia e identità del «sepulcrum magnum sub ecclesia annuntiationis» (1636-1696). I documenti e la lettura della fabbrica. In F. Vaglienti e C. Cattaneo, a cura di La popolazione di Milano dal Rinascimento. Milano: Biblioteca Francescana. 2013; pp. 83–136.

Caruso V, Gibelli D, Sassi F, Sguazza E, Ceresa Mori A, Bianucci R, Bramanti B, Haensch S, Cattaneo C. Le vittime della peste Manzoniana? XX Congresso degli Antropologi Italiani. Ferrara, 11-13 September 2013a. (Poster presentation).

Caruso V, Sguazza E, Sassi F, Gibelli D, Mori AC, Cattaneo C. Gli scheletri della fossa comune di viale Sabotino a Milano: le vittime della peste manzoniana?. *FOLD&R* FastiOnLine documents & research. 2013b; 285: 1-11.

Caruso V, Cummaudo M, Maderna E, Cappella A, Caudullo G, Scarpulla V, Cattaneo C. A comparative analysis of microscopic alterations in modern and ancient undecalcified and decalcified dry bones. Am J Phys Anthropol. 2017a. [Epub ahead of print].

Caruso V, Gibelli D, Castoldi E, Sconfienza LM, Sardanelli F, Cattaneo C. Metastatic Cancer in the Middle Age: The Possible Case of a Female Skeleton from Bormio (Italy). International Journal of Osteoarchaeology. 2017b. [Epub ahead of print].

Cattaneo C, Gelsthorpe K, Phillips P, Waldron T, Booth JR, Sokol RJ. Immunological diagnosis of multiple myeloma in a medieval bone. International Journal of Osteoarchaeology. 1994; 4: 1-2.

Cattaneo C, Binda M, Ravedoni C. Two cases of pre-Colombian syphilis in Roman Milan and Mediaeval Luni. Rivista CNR: Science and Technology for Cultural Heritage. 1996; 5: 1-5.

REFERENCES

Cattaneo C, Ravedoni C. Considerazioni preliminari sugli abitanti della Milano tardoantica. Studi di antropologia e paleopatologia, in La città e la sua memoria: Milano e la tradizione di Sant'Ambrogio, catalogo della mostra (Milano 1997). Milano. 1997; pp. 127-129.

Cattaneo C, Ravedoni C, Martino B, Mazzucchi A, Porta D, Binda M. Vita nella Milano romana: evidenze antropologiche e paleopatologiche provenienti dalla necropolis. In Sannazaro M (a cura di), La necropoli tardoantica, atti delle giornate di studio (Milano 25-26 gennaio 1999), Milano. 2001; pp. 59-66.

Cattaneo C, Ravedoni C, Martino B, Mazzucchi A, Porta A, Binda M. Vita nella Milano romana: evidenze antropologiche e paleopatologiche provenienti dalla necropoli. In Sannazaro M (a cura di), La Necropoli Tardoantica. Ricerche archeologiche nei cortili dell'Università cattolica. Milano. 2002; pp. 59-66.

Cattaneo C, Grandi M. Antropologia e odontologia forense. Guida allo studio dei resti umani. Monduzzi Editore. 2004.

Cattaneo C, Andreola S, Marinelli E, Poppa P, Porta D, Grandi M. The detection of microscopic markers of hemorrhaging and wound age on dry bone: a pilot study. Am J Forensic Med Pathol. 2010; 31: 22-26.

Cattaneo C, Mazzucchi A, Gibelli D. Le analisi antropologiche: stato dell'arte. In Lusuardi Siena S, Rossignani MP, Sannazaro M (a cura di), L'abitato, la necropoli, il monastero. Evoluzione di un comparto del suburbio Milanese alla luce degli scavi nei cortili dell'Università Cattolica. Milano. 2011; pp. 134-139.

Cattaneo C, Poppa P, Gibelli D, Sassi F, Porta D. "Sit tibi terra levis" Prime risultanze dall'analisi antropologica e paleopatologica del Sepolcreto dell'Ospedale Maggiore di Milano. In Vaglienti F, Cattaneo C. a cura di La popolazione di Milano dal Rinascimento. Milano: Biblioteca Francescana 2013.

Cattaneo C, Cotti C, Gibelli D, Mazzarelli D, Sguazza E. La necropoli di Piazza Sant'Ambrogio. Testimonianze della storia di Milano, fra carenze nutrizionali e violenza. In Fedeli AM, Pagani C (a cura di), Il volto di una piazza. Indagini archeologiche per la realizzazione del parcheggio in Piazza Sant'Ambrogio a Milano. Edizioni Et, Milano. 2015a; pp. 51-58.

Cattaneo C, Gibelli D. Valorizzazione del patrimonio scheletrico umano: una prospettiva su Milano. LANX. Rivista della Scuola di Specializzazione in Archeologia-Università degli Studi di Milano. 2015; 19: 129-136.

Cattaneo C, Gibelli D, Caruso V. Antichi popoli di Valtellina. Analisi paleobiologiche su resti scheletrici tra tardoantico, Medioevo e Rinascimento. In Mariotti V (a cura di), La Valtellina nei secoli. Studi e ricerche

archeologiche. Vol II. Ricerche e materiali archeologici. SAP Società Archeologica. 2015b.

Chiti M. Ricostruzione antropologica di Milano antica: la necropoli a inumazione e cremazione di Porta Romana (I secolo a.C. – III secolo d.C.). Tesi di Laurea in Scienze Naturali Anno. 2002.

Coleman RE. Skeletal complication of malignancy. Cancer. 1997; 80: 1588-1594.

Collins MJ, Nielsen–Marsh CM, Hiller J, Smith CI, Roberts JP, Prigodich RV, Weiss TJ, Csapò J, Millard AR, Turner–Walker G. The survival of organic matter in bone: a review. Archaeometry. 2002; 44: 383–394.

Collins MJ, Penkman KE, Rohland N, Shapiro B, Dobberstein RC, Ritz-Timme S, Hofreiter M. Is amino acid racemization a useful tool for screening for ancient DNA in bone? Proc Biol Sci. 2009; 276: 2971-2977.

Cooper A, Poinar HN. Ancient DNA: do it right or not at all. Science. 2000; 289: 1139.

Cooper A, Rambaut A, Macaulay V, Willerslev E, Hansen AJ, Stringer C. Human origins and ancient human DNA. Science. 2001; 292: 1655-1656.

Cortese C. Genesi e trasformazioni di un quartiere suburbano della Milano romana. In Lusuardi Siena S, Rossignani MP, Sannazaro M (a cura di), L'abitato, la necropoli, il monastero. Evoluzione di un comparto del suburbio Milanese alla luce degli scavi nei cortili dell'Università Cattolica. Milano. 2011; pp. 5-17.

Cosmacini G. La Ca' Granda dei Milanesi. Storia dell'Ospedale Maggiore. Roma - Bari: Laterza. 1999.

Crozatti LL, de Brito MH, Lopes BN, de Campos FP. Atypical behavioral and psychiatric symptoms: Neurosyphilis should always be considered. Autops Case Rep. 2015; 5: 43-47.

Cunha E, Baccino E, Martrille L, Ramsthaler F, Prieto J, Schuliar Y, Lynnerup N, Cattaneo C. The problem of aging human remains and living individuals: a review. Forensic science international. 2009; 193: 1-13.

Dabney J, Knapp M, Glocke I, Gansauge MT, Weihmann A, Nickel B, Valdiosera C, García N, Pääbo S, Arsuaga JL, Meyer M. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proc Natl Acad Sci U S A. 2013a; 110: 15758-15763.

Dabney J, Meyer M, Pääbo S. Ancient DNA damage. Cold Spring Harb Perspect Biol. 2013b; 5: a012567.

Damgaard PB, Margaryan A, Schroeder H, Orlando L, Willerslev E, Allentoft ME. Improving access to endogenous DNA in ancient bones and teeth. Sci Rep. 2015; 5: 11184.

DeSalle R, Gatesy J, Wheeler W, Grimaldi D. DNA sequences from a fossil termite in Oligo-Miocene amber and their phylogenetic implications. Science. 1992; 257: 1933-1936.

De Vanna L. La sequenza stratigrafica degli scavi archeologici di Piazza Sant'Ambrogio a Milano. In Fedeli AM, Pagani C (a cura di), Il volto di una piazza. Indagini archeologiche per la realizzazione del parcheggio in Piazza Sant'Ambrogio a Milano. Edizioni Et, Milano. 2015a; pp. 36-45.

De Vanna L. Scavi nei centri urbani della Valtellina. In Mariotti V (a cura di), La Valtellina nei secoli. Studi e ricerche archeologiche. Vol II. Ricerche e materiali archeologici. SAP Società Archeologica. 2015b.

DiMauro S, Schon EA. Mitochondrial respiratory-chain diseases. N Engl J Med. 2003; 348: 2656-2668.

Dirkmaat DC, Cabo LL, Ousley SD, Symes SA. New perspectives in forensic anthropology. Am J Phys Anthropol. 2008; Suppl 47: 33-52.

Dixon R, Dawson L, Taylor D. The experimental degradation of archaeological human bone by anaerobic bacteria and the implications for recovery of ancient DNA. Paper presented at the Ninth International Conference on Ancient DNA and Associated Biomolecules, Pompeii, Italy. 2008.

Dobberstein RC, Huppertz J, von Wurmb-Schwark N, Ritz-Timme S. Degradation of biomolecules in artificially and naturally aged teeth: implications for age estimation based on aspartic acid racemization and DNA analysis. Forensic Sci Int. 2008; 179: 181-191.

Edson SM, Ross JP, Coble MD, Parsons TJ, Barritt SM. Naming the Dead - Confronting the Realities of Rapid Identification of Degraded Skeletal Remains. Forensic Sci Rev. 2004; 16: 63-90.

Edson SM, Christensen AF, Barritt SM, Meehan A, Leney MD, Finelli LN. Sampling of the cranium for mitochondrial DNA analysis of human skeletal remains. Forensic Science International: Genetics Supplement Series. 2009; 2: 269-270.

Eduardoff M, Xavier C, Strobl C, Casas-Vargas A, Parson W. Optimized mtDNA Control Region Primer Extension Capture Analysis for Forensically Relevant Samples and Highly Compromised mtDNA of Different Age and Origin. Genes (Basel). 2017; 8.

Fazekas IG, Kosà F. Forensic fetal osteology. Budapest: Akadémiai Kiadó. 1978.

Fehren-Schmitz L, Llamas B, Lindauer S, Tomasto-Cagigao E, Kuzminsky S, Rohland N, Santos FR, Kaulicke P, Valverde G, Richards SM, Nordenfelt S, Seidenberg V, Mallick S, Cooper A, Reich D *et al.* A Re-Appraisal of the Early Andean Human Remains from Lauricocha in Peru. PLoS One. 2015; 10: e0127141.

Fornaciari G, Giuffra V. Lezioni di paleopatologia. Genova. 2009.

Fortunati M, Ghiroldi A. La necropoli di San Chierico di Bolgare. Notizie Archeologiche Bergomensi. 2006; 14: 87-135.

Friedrich F, Aigner M, Fearns N, Friedrich ME, Frey R, Geusau A. Psychosis in neurosyphilis -- clinical aspects and implications. Psychopathology. 2014; 47: 3-9.

Fu Q, Meyer M, Gao X, Stenzel U, Burbano HA, Kelso J, Pääbo S. DNA analysis of an early modern human from Tianyuan Cave, China. Proc Natl Acad Sci USA. 2013a; 110: 2223-2227.

Fu Q, Mittnik A, Johnson PLF, Bos K, Lari M, Bollongino R, Sun C, Giemsch L, Schmitz R, Burger J, Ronchitelli AM, Martini F, Cremonesi RG, Svoboda J, Bauer P *et al.* A revised timescale for human evolution based on ancient mitochondrial genomes. Curr Biol. 2013b; 23: 553-559.

Gamba C, Jones ER, Teasdale MD, McLaughlin RL, Gonzalez-Fortes G, Mattiangeli V, Domboróczki L, Kővári I, Pap I, Anders A, Whittle A, Dani J, Raczky P, Higham TF, Hofreiter M *et al.* Genome flux and stasis in a five millennium transect of European prehistory. Nat Commun. 2014; 5: 5257.

Gamba C, Hanghøj K, Gaunitz C, Alfarhan AH, Alquraishi SA, Al-Rasheid KA, Bradley DG, Orlando L. Comparing the performance of three ancient DNA extraction methods for high-throughput sequencing. Mol Ecol Resour. 2016; 16: 459-469.

Gartner LP, Hiatt JL. Istologia. Seconda edizione, EdISES. 2002.

Gaytmenn R, Sweet D. Quantification of forensic DNA from various regions of human teeth. J Forensic Sci. 2003; 48: 622-625.

Gilbert MTP, Rudbeck L, Willerslev E, Hansen AJ, Smith C, Penkman KEH, Prangenberg K, Nielsen-Marsh CM, Jans ME, Arthur P, Lynnerup N, Turner-Walker G, Biddle M, Kjølbye-Biddle B, Collins M. Biochemical and physical correlates of DNA contamination in archaeological human bones and teeth excavated at Matera, Italy. Journal of Archaeological Science. 2005; 32: 1053–1060.

Gilbert MTP, Hansen AJ, Willerslev E, Turner-Walker G, Collins M. Insights into the processes behind the contamination of degraded human teeth and

bone samples with exogenous sources of DNA. International Journal of Osteoarchaeology. 2006; 16: 156-164.

Gilbert MT, Binladen J, Miller W, Wiuf C, Willerslev E, Poinar H, Carlson JE, Leebens-Mack JH, Schuster SC. Recharacterization of ancient DNA miscoding lesions: insights in the era of sequencing-by-synthesis. Nucleic Acids Res. 2007; 35: 1-10.

Golenberg EM, Giannasi DE, Clegg MT, Smiley CJ, Durbin M, Henderson D, Zurawski G. Chloroplast DNA sequence from a miocene Magnolia species. Nature. 1990; 344: 656-658.

Götherström, A, Collins MJ, Angerbjörn A, Lidén K. Bone preservation and DNA amplification. Archaeometry. 2002; 44: 395-404.

Green RE, Briggs AW, Krause J, Prüfer K, Burbano HA, Siebauer M, Lachmann M, Pääbo S. The Neandertal genome and ancient DNA authenticity. EMBO J. 2009; 28: 2494-2502.

Guarino FM, Angelini F, Vollono C, Orefice C. Bone preservation in human remains from the Terme del Sarno at Pompeii using light microscopy and scanning electron microscopy. Journal of Archaeological Science. 2006; 33: 513–520.

Gustafson G. Age determination on teeth. J Am Dent Assoc. 1950; 41: 45-54.

Hagelberg E, Hofreiter M, Keyser C. Introduction. Ancient DNA: the first three decades. Philos Trans R Soc Lond B Biol Sci. 2015; 370: 20130371.

Handt O, Richards M, Trommsdorff M, Kilger C, Simanainen J, Georgiev O, Bauer K, Stone A, Hedges R, Schaffner W *et al*. Molecular genetic analyses of the Tyrolean Ice Man. Science. 1994; 264: 1775-1778.

Hansen A, Willerslev E, Wiuf C, Mourier T, Arctander P. Statistical evidence for miscoding lesions in ancient DNA templates. Mol Biol Evol. 2001; 18: 262-265.

Hauswirth WW, Dickel CD, Rowold DJ, Hauswirth MA. Inter- and intra-population studies of ancient humans. Experientia. 1994; 50: 585–591.

Haynes S, Searle JB, Bretman A, Dobney KM. Bone preservation and ancient DNA: the application of screening methods for predicting DNA survival. Journal of Archaeological Science. 2002; 29: 585-592.

Hedges REM, Millard AR, Pike AWG. Measurements and relationships of diagenetic alteration of bone from three archaeological sites. Journal of Archaeological Science. 1995; 22: 201-209.

Hedges REM. Bone diagenesis: an overview of processes. Archaeometry. 2002; 44: 319–328.

Hefner JT. Cranial nonmetric variation and estimating ancestry. J Forensic Sci. 2009; 54: 985-995.

Heupink TH, Subramanian S, Wright JL, Endicott P, Westaway MC, Huynen L, Parson W, Millar CD, Willerslev E, Lambert DM. Ancient mtDNA sequences from the First Australians revisited. Proc Natl Acad Sci U S A. 2016; 113: 6892-6897.

Higgins D, Kaidonis J, Austin J, Townsend G, James H, Hughes T. Dentine and cementum as sources of nuclear DNA for use in human identification. Australian Journal of Forensic Sciences. 2011; 43: 287-295.

Higgins D, Austin JJ. Teeth as a source of DNA for forensic identification of human remains: a review. Sci Justice. 2013; 53: 433-441.

Higuchi R, Bowman B, Freiberger M, Ryder OA, Wilson AC. DNA sequences from the quagga, an extinct member of the horse family. Nature. 1984; 312: 282-284.

Hofreiter M, Poinar HN, Spaulding WG, Bauer K, Martin PS, Possnert G, Pääbo S. A molecular analysis of ground sloth diet through the last glaciation. Mol Ecol. 2000; 9: 1975-1984.

Hofreiter M, Jaenicke V, Serre D, von Haeseler A, Pääbo S. DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. Nucleic Acids Res. 2001a; 29: 4793-4799.

Hofreiter M, Serre D, Poinar HN, Kuch M, Pääbo S. Ancient DNA. Nat Rev Genet. 2001b; 2: 353-359.

Höss M, Jaruga P, Zastawny TH, Dizdaroglu M, Pääbo S. DNA damage and DNA sequence retrieval from ancient tissues. Nucleic Acids Res. 1996; 24: 1304-1307.

Howell N, Elson JL, Howell C, Turnbull DM. Relative rates of evolution in the coding and control regions of African mtDNAs. Mol Biol Evol. 2007; 24: 2213-2221.

Işcan MY, Loth SR, Wright RK. Age estimation from the rib by phase analysis: white males. J Forensic Sci. 1984; 29: 1094-1104.

Işcan MY, Loth SR, Wright RK. Age estimation from the rib by phase analysis: white females. J Forensic Sci. 1985; 30: 853-863.

Jans MME, Kars H, Nielsen–Marsh CM, Smith CI, Nord AG, Arthur P, Earl N. In situ preservation of archaeological bone: a histological study within a multidisciplinary approach. Archaeometry. 2002; 44: 343-352.

Jobling MA, Tyler-Smith C. The human Y chromosome: an evolutionary marker comes of age. Nat Rev Genet. 2003; 4: 598-612.

Kaiser C, Bachmeier B, Conrad C, Nerlich A, Bratzke H, Eisenmenger W, Peschel O. Molecular study of time dependent changes in DNA stability in soil buried skeletal residues. Forensic Sci Int. 2008; 177: 32-36.

Kemp BM, Smith DG. Use of bleach to eliminate contaminating DNA from the surface of bones and teeth. Forensic Sci Int. 2005; 154: 53-61.

Kerley ER, Ubelaker DH. Revisions in the microscopic method of estimating age at death in human cortical bone. Am J Phys Anthropol. 1978; 49: 545-546.

Key FM, Posth C, Krause J, Herbig A, Bos KI. Mining Metagenomic Data Sets for Ancient DNA: Recommended Protocols for Authentication. Trends Genet. 2017; 33: 508-520.

Kolman CJ, Tuross N. Ancient DNA analysis of human populations. Am J Phys Anthropol. 2000; 111: 5-23.

Kontopoulos I, Nystrom P, White L. Experimental taphonomy: post-mortem microstructural modifications in Sus scrofa domesticus bone. Forensic Sci Int. 2016; 266: 320-328.

Korlević P, Gerber T, Gansauge MT, Hajdinjak M, Nagel S, Aximu-Petri A, Meyer M. Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. Biotechniques. 2015; 59: 87-93.

Korneliussen TS, Albrechtsen A, Nielsen R. ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics. 2014; 15: 356.

Kranioti E, Paine R. Forensic anthropology in Europe: an assessment of current status and application. J Anthropol Sci. 2011; 89: 71-92.

Krause J, Briggs AW, Kircher M, Maricic T, Zwyns N, Derevianko A, Pääbo S. A complete mtDNA genome of an early modern human from Kostenki, Russia. Curr Biol. 2010; 20: 231-236.

Krause J, Pääbo S. Genetic Time Travel. Genetics. 2016; 203: 9-12.

La Ferla M. Studio antropologico e paleopatologico di un campione della necropoli longobarda di Bolgare (BG). Tesi di Laurea in Scienze Naturali. 2005.

Lambert JB, Vlasak SM, Thometz AC, Buikstra JE. A comparative study of the chemical analysis of ribs and femurs in Woodland populations. Am J Phys Anthropol. 1982; 59: 289-294.

Lazaridis I, Patterson N, Mittnik A, Renaud G, Mallick S, Kirsanow K, Sudmant PH, Schraiber JG, Castellano S, Lipson M, Berger B, Economou C, Bollongino R, Fu Q, Bos KI *et al.* Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature. 2014; 513: 409-413.

Leney MD. Sampling skeletal remains for ancient DNA (aDNA): a measure of success. Historical Archaeology. 2006; 40: 31-49.

Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25: 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. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25: 2078-2079.

Lindahl T, Nyberg B. Rate of depurination of native deoxyribonucleic acid. Biochemistry. 1972; 11: 3610-3618.

Lindahl T. Instability and decay of the primary structure of DNA. Nature. 1993; 362: 709-715.

Llamas B, Fehren-Schmitz L, Valverde G, Soubrier J, Mallick S, Rohland N, Nordenfelt S, Valdiosera C, Richards SM, Rohrlach A, Romero MI, Espinoza IF, Cagigao ET, Jiménez LW, Makowski K *et al.* Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. Sci Adv. 2016; 2: e1501385.

Lovejoy CO, Meindl RS, Pryzbeck TR, Mensforth RP. 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. 1985; 68:15-28.

Lutz S, Weisser HJ, Heizmann J, Pollak S. A third hypervariable region in the human mitochondrial D-loop. Hum Genet. 1997; 101: 384.

Lutz S, Weisser HJ, Heizmann J, Pollak S. Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany. Int J Legal Med. 1998; 111: 67-77.

Madel MB, Niederstätter H, Parson W. TriXY-Homogeneous genetic sexing of highly degraded forensic samples including hair shafts. Forensic Sci Int Genet. 2016; 25: 166-174.

Marchiafava V, Bonucci E, Ascenzi A. Fungal osteoclasia: a model of dead bone resorption. Calcif Tissue Res. 1974; 14: 195-210.

Mardis ER. Next-generation DNA sequencing methods. Annu Rev Genomics Hum Genet. 2008; 9: 387-402.

Maricic T, Whitten M, Pääbo S. Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. PLoS One. 2010; 5: e14004.

Marsden I, Pagani C. Milano, viale Sabotino. Indagini archeologiche, in Notiziario della Soprintendenza per i beni archeologici della Lombardia, 2006. Edizioni Et, Milano. 2008; pp. 119-121.

Martini FH, Timmons MJ, Tallitsch RB. Anatomia umana. Quarta edizione, EdISES. 2010.

Mazzucchi A, Betti L, Steffenini D, Cattaneo C. La necropoli di San Chierico di Bolgare. Studi antropologici e paleopatologici. Notizie Archeologiche Bergomensi. 2006; 14: 203-272.

Mello RB, Silva MR, Alves MT, Evison MP, Guimarães MA, Francisco RA, Astolphi RD, Iwamura ES. Tissue Microarray Analysis Applied to Bone Diagenesis. Sci Rep. 2017; 7: 39987.

Meyer M, Kircher M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb Protoc. 2010; 2010: pdb.prot5448.

Meyer M, Kircher M, Gansauge MT, Li H, Racimo F, Mallick S, Schraiber JG, Jay F, Prüfer K, de Filippo C, Sudmant PH, Alkan C, Fu Q, Do R, Rohland N *et al.* A high-coverage genome sequence from an archaic Denisovan individual. Science. 2012; 338: 222-226.

Milos A, Selmanović A, Smajlović L, Huel RL, Katzmarzyk C, Rizvić A, Parsons TJ. Success rates of nuclear short tandem repeat typing from different skeletal elements. Croat Med J. 2007; 48: 486-493.

Misner LM, Halvorson AC, Dreier JL, Ubelaker DH, Foran DR. The correlation between skeletal weathering and DNA quality and quantity. J Forensic Sci. 2009; 54: 822-828.

Mittnik A, Wang CC, Svoboda J, Krause J. A Molecular Approach to the Sexing of the Triple Burial at the Upper Paleolithic Site of Dolní Věstonice. PLoS One. 2016; 11: e0163019.

Monforti C. Doc. del 15 maggio 1694. ASMi, Sanità, 127.

Moraitis K., Spiliopoulou C. Identification and differential diagnosis of perimortem blunt force trauma in tubular long bones. Forensic Science, Medicine, and Pathology. 2006; 2: 221-229.

Mundorff AZ, Bartelink EJ, Mar-Cash E. DNA preservation in skeletal elements from the World Trade Center disaster: recommendations for mass fatality management. J Forensic Sci. 2009; 54: 739-745.

Orlando L, Ginolhac A, Raghavan M, Vilstrup J, Rasmussen M, Magnussen K, Steinmann KE, Kapranov P, Thompson JF, Zazula G, Froese D, Moltke I, Shapiro B, Hofreiter M, Al-Rasheid KA *et al.* True single-molecule DNA sequencing of a pleistocene horse bone. Genome Res. 2011; 21: 1705-1719.

Ortner DJ. Identification of pathological conditions in human skeletal remains. 2nd edition. San Diego, CA: Academic Press. 2003.

Overballe-Petersen S, Orlando L, Willerslev E. Next-generation sequencing offers new insights into DNA degradation. Trends Biotechnol. 2012; 30: 364-368.

Pääbo S. Molecular cloning of Ancient Egyptian mummy DNA. Nature. 1985; 314: 644-645.

Pääbo S, Wilson AC. Polymerase chain reaction reveals cloning artefacts. Nature. 1988; 334: 387-388.

Pääbo S. Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. Proc Natl Acad Sci U S A. 1989; 86: 1939-1943.

Pääbo S, Wilson AC. Miocene DNA sequences - a dream come true? Curr Biol. 1991; 1: 45-46.

Pääbo S, Poinar H, Serre D, Jaenicke-Despres V, Hebler J, Rohland N, Kuch M, Krause J, Vigilant L, Hofreiter M. Genetic analyses from ancient DNA. Annu Rev Genet. 2004; 38: 645-679.

Pakendorf B, Stoneking M. Mitochondrial DNA and human evolution. Annu Rev Genomics Hum Genet. 2005; 6:165-183.

Peltzer A, Jäger G, Herbig A, Seitz A, Kniep C, Krause J, Nieselt K. EAGER: efficient ancient genome reconstruction. Genome Biol. 2016; 17: 60.

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 *et al.* Optimal Ancient DNA Yields from the Inner Ear Part of the Human Petrous Bone. PLoS One. 2015; 10: e0129102.

Pisacreta S. Montenapoleone 2011: studio antropologico di una popolazione della Milano rinascimentale. Tesi di Master di I livello in Bioarcheologia, Paleopatologia e Antropologia Forense. 2014.

Poinar HN, Cano RJ, Poinar GO. DNA from an extinct plant. Nature. 1993; 363: 677.

Poinar HN, Höss M, Bada JL, Pääbo S. Amino acid racemization and the preservation of ancient DNA. Science. 1996; 272: 864-866.

Poinar HN, Hofreiter M, Spaulding WG, Martin PS, Stankiewicz BA, Bland H, Evershed RP, Possnert G, Pääbo S. Molecular coproscopy: dung and diet of the extinct ground sloth Nothrotheriops shastensis. Science. 1998; 281: 402-406.

Posth C, Wißing C, Kitagawa K, Pagani L, van Holstein L, Racimo F, Wehrberger K, Conard NJ, Kind CJ, Bocherens H, Krause J. Deeply divergent archaic mitochondrial genome provides lower time boundary for African gene flow into Neanderthals. Nat Commun. 2017; 8: 16046.

Prinz M, Carracedo A, Mayr WR, Morling N, Parsons TJ, Sajantila A, Scheithauer R, Schmitter H, Schneider PM; International Society for Forensic Genetics. DNA Commission of the International Society for Forensic Genetics (ISFG): recommendations regarding the role of forensic genetics for disaster victim identification (DVI). Forensic Sci Int Genet. 2007; 1: 3-12.

Rasmussen M, Guo X, Wang Y, Lohmueller KE, Rasmussen S, Albrechtsen A, Skotte L, Lindgreen S, Metspalu M, Jombart T, Kivisild T, Zhai W, Eriksson A, Manica A, Orlando L *et al*. An Aboriginal Australian genome reveals separate human dispersals into Asia. Science. 2011; 334: 94-98.

Rasmussen M, Anzick SL, Waters MR, Skoglund P, DeGiorgio M, Stafford TW Jr, Rasmussen S, Moltke I, Albrechtsen A, Doyle SM, Poznik GD, Gudmundsdottir V, Yadav R, Malaspinas AS, White SSV *et al*. The genome of a Late Pleistocene human from a Clovis burial site in western Montana. Nature. 2014; 506: 225-229.

Reggiani F. Sotto le ali della colomba: famiglie assistenziali e relazioni di genere a Milano dall'età moderna alla Restaurazione. Roma: Viella. 2014.

Reich D, Green RE, Kircher M, Krause J, Patterson N, Durand EY, Viola B, Briggs AW, Stenzel U, Johnson PL, Maricic T, Good JM, Marques-Bonet T, Alkan C, Fu Q *et al*. Genetic history of an archaic hominin group from Denisova Cave in Siberia. Nature. 2010; 468: 1053-1060.

Reiche I, Favre-Quattropani L, Vignaud C, Bocherens H, Charlet L, Menu M. A multi-analytical study of bone diagenesis: the Neolithic site of Bercy (Paris, France). Measurement Science and Technology. 2003; 14: 1608–1619.

Renaud G, Slon V, Duggan AT, Kelso J. Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. Genome Biol. 2015; 16: 224.

Riva MA, Mazzoleni D. The Ospedale Maggiore Policlinico of Milan. Journal of Medicine and the Person. 2012; 10: 136–138.

Rizzi E, Lari M, Gigli E, De Bellis G, Caramelli D. Ancient DNA studies: new perspectives on old samples. Genet Sel Evol. 2012; 44: 21.

Rohland N, Hofreiter M. Comparison and optimization of ancient DNA extraction. Biotechniques. 2007; 42: 343-352.

Rohland N, Harney E, Mallick S, Nordenfelt S, Reich D. Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. Philos Trans R Soc Lond B Biol Sci. 2015; 370: 20130624.

Rossignani MP, Sannazaro M, Legrottaglie G. La Signora del sarcofago: una sepoltura di rango nella necropoli dell'Università cattolica: ricerche archeologiche nei cortili dell'Università cattolica. Vita e Pensiero. 2005.

Saccomani M. Indagine paleopatologica degli scheletri delle vittime della "peste del manzoni". Tesi di Laurea in Scienze Naturali. 2010.

Sannazaro M. La necropoli tardoantica. In Lusuardi Siena S, Rossignani MP, Sannazaro M (a cura di), L'abitato, la necropoli, il monastero. Evoluzione di un comparto del suburbio Milanese alla luce degli scavi nei cortili dell'Università Cattolica. Milano. 2011; pp. 73-85.

Schafer M, Black S, Scheuer L. Juvenile osteology: a laboratory and field manual. San Diego, CA: Academic Press. 2008.

Scheuer L, Black S. Developmental Juvenile Osteology. San Diego, CA: Academic Press. 2000.

Schuenemann VJ, Peltzer A, Welte B, van Pelt WP, Molak M, Wang CC, Furtwängler A, Urban C, Reiter E, Nieselt K, Teßmann B, Francken M, Harvati K, Haak W, Schiffels S *et al.* Ancient Egyptian mummy genomes suggest an increase of Sub-Saharan African ancestry in post-Roman periods. Nat Commun. 2017; 8: 15694.

Schwarz C, Debruyne R, Kuch M, McNally E, Schwarcz H, Aubrey AD, Bada J, Poinar H. New insights from old bones: DNA preservation and degradation in permafrost preserved mammoth remains. Nucleic Acids Res. 2009; 37: 3215-3229.

Sguazza E. L'età romana in Lombardia: quadro demografico e patologico. Tesi di Laurea Magistrale in Archeologia. 2010.

Sguazza E. Il sepolcreto dell'Ospedale Maggiore (Ca Granda) di Milano: indagini antropologiche di un singolare contesto di resti commisti. Tesi di Dottorato in Medicina e Scienze Umane. 2015.

Sguazza E, Mazzucchi A, Fortunati M, Cattaneo C. The necropolis of Bolgare (Lombardy, Italy): Anthropological and paleopathological features of a Lombard population. Homo. 2015; 66: 139-148.

Sguazza E, Gibelli D, Caligara M, Di Candia D, Paolo M, Galimberti PM, Cattaneo C. The role of toxicological analyses in anthropology: a case report on lead intoxication. Archaeometry. 2016a; 58: 152-158.

Sguazza E, Gibelli D, Galimberti PM, Cattaneo C. A Case of Elbow Synostosis in a Child of the Ancient Hospital of Milan. International Journal of Paleopathology. 2016b; 26: 69-72.

Sirak KA, Fernandes DM, Cheronet O, Novak M, Gamarra B, Balassa T, Bernert Z, Cséki A, Dani J, Gallina JZ, Kocsis-Buruzs G, Kővári I, László O, Pap I, Patay R *et al*. A minimally-invasive method for sampling human petrous bones from the cranial base for ancient DNA analysis. Biotechniques. 2017; 62: 283-289.

Skoglund P, Storå J, Götherström A, Jakobsson M. Accurate sex identification of ancient human remains using DNA shotgun sequencing. Journal of Archaeological Science. 2013; 40: 4477-4482.

Slice DE, Ross A. 3D-ID: Geometric Morphometric Classification of Crania for Forensic Scientists. 2009.

Smith CI, Chamberlain AT, Riley MS, Stringer C, Collins MJ. The thermal history of human fossils and the likelihood of successful DNA amplification. J Hum Evol. 2003; 45: 203-217.

Stiller M, Green RE, Ronan M, Simons JF, Du L, He W, Egholm M, Rothberg JM, Keates SG, Ovodov ND, Antipina EE, Baryshnikov GF, Kuzmin YV, Vasilevski AA, Wuenschell GE *et al*. Patterns of nucleotide misincorporations during enzymatic amplification and direct large-scale sequencing of ancient DNA. Proc Natl Acad Sci U S A. 2006; 103: 13578-13584.

Stoneking M, Krause J. Learning about human population history from ancient and modern genomes. Nat Rev Genet. 2011; 12: 603-614.

Stull KE, Kenyhercz MW, L'Abbé EN. Ancestry estimation in South Africa using craniometrics and geometric morphometrics. Forensic Sci Int. 2014; 245: 206.e1-7.

Suppersberger Hamre S, Ersland GA, Daux V, Parson W, Wilkinson C. Three individuals, three stories, three burials from medieval Trondheim, Norway. PLoS One. 2017; 12: e0180277.

Svintradze DV, Mrevlishvili GM, Metreveli N, Jariashvili K, Namicheishvili L, Skopinska J, Sionkowska A. Collagen-DNA complex. Biomacromolecules. 2008; 9: 21-28.

Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, Vullo CM, Wallace DC. Asian affinities and continental radiation of the four founding Native American mtDNAs. Am J Hum Genet. 1993; 53: 563-590.

Torroni A, Achilli A, Macaulay V, Richards M, Bandelt HJ. Harvesting the fruit of the human mtDNA tree. Trends Genet. 2006; 22: 339-345.

Trotter M, Gleser GC. Estimation of stature from long bones of American Whites and Negroes. American Journal of Physical Anthropology. 1952; 10: 463-514.

Trotter M, Gleser GC. Corrigenda to "estimation of stature from long limb bones of American Whites and Negroes. American Journal Physical Anthropology (1952). Am J Phys Anthropol. 1977; 47: 355-356.

Trueman CN, Martin DM. The long-term survival of bone: the role of bioerosion. Archaeometry. 2002; 44: 371–382.

Ubelaker DH. Human Skeletal Remains: Excavation, Analysis, Interpretation. 3rd ed. Washington, D.C.: Taraxacum. 1999.

Valverde G, Barreto Romero MI, Flores Espinoza I, Cooper A, Fehren-Schmitz L, Llamas B, Haak W. Ancient DNA Analysis Suggests Negligible Impact of the Wari Empire Expansion in Peru's Central Coast during the Middle Horizon. PLoS One. 2016; 11: e0155508.

van Oven M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum Mutat. 2009; 3: E386-394.

Wallace DC. Mitochondrial DNA sequence variation in human evolution and disease. Proc Natl Acad Sci U S A. 1994; 91: 8739-8746.

Weiss L. Histology: cell and tissue biology. Fifth edition, Elsevier Biomedical. 1983.

White T, Folkens P. The human bone manual. Academic Press. 2005.

Willerslev E, Cooper A. Ancient DNA. Proc Biol Sci. 2005; 272: 3-16.

Wilson LYN, Pollard AM. Here Today, Gone Tomorrow ? Integrated Experimentation and Geochemical Modeling in Studies of Archaeological Diagenetic Change Understanding Diagenetic Processes. Recent. 2002; 35: 644–651.

Woodward SR, Weyand NJ, Bunnell M. DNA sequence from Cretaceous period bone fragments. Science. 1994; 266: 1229-1232.

www.illumina.com

www.thermofisher.com

LIST OF ORIGINAL PUBLICATIONS

Cappella A, **Bertoglio B**, Castoldi E, Maderna E, Di Giancamillo A, Domeneghini C, Andreola S, Cattaneo C. The taphonomy of blood components in decomposing bone and its relevance to physical anthropology. Am J Phys Anthropol. 2015; 158(4): 636-645. (IF 2.552)

Fattorini P, Previderé C, Carboni I, Marrubini G, Sorçaburu-Cigliero S, Grignani P, **Bertoglio B**, Vatta P, Ricci U. Performance of the ForenSeq(TM) DNA Signature Prep kit on highly degraded samples. Electrophoresis. 2017; 38(8): 1163-1174. (IF 2.774)

Bertoglio B, Grignani P, Di Simone P, Polizzi N, Bertuglia C, Cattaneo C, Presciuttini S, Previderè C. Preliminary results of the genetic identifications of the Lampedusa 2013 shipwreck victims. Oral presentation. La Revue de Médicine Légale. 2017; 8(4): 184.

The Taphonomy of Blood Components in Decomposing Bone and Its Relevance to Physical Anthropology

Annalisa Cappella,[1]* Barbara Bertoglio,[1,3] Elisa Castoldi,[1] Emanuela Maderna,[1] Alessia Di Giancamillo,[2] Cinzia Domeneghini,[2] Salvatore Andreola,[1] and Cristina Cattaneo[1]

[1]*Dipartimento Di Scienze Biomediche per La Salute, LABANOF, Laboratorio Di Antropologia E Odontologia Forense, Sezione Di Medicina Legale, Università Degli Studi Di Milano, Milan 20133, Italy*
[2]*Department of Health, Animal Science and Food Safety, Università Degli Studi Di Milano, Milan 20133, Italy*
[3]*Fellow of the Genetics, Molecular and Cellular Biology Ph.D. program of the University of Pavia*

ABSTRACT

Objectives: The variation and persistence of blood components, in particular red blood cells (RBCs), within bone tissue during the decomposition process, especially at the early stages and in different taphonomic conditions, has never been thoroughly investigated, regardless of the fact that knowing how blood survives or degrades within bone could be of help in solving many anthropological issues, such as trauma analysis and interpretation.

Materials and Methods: This research investigated the influence of time and taphonomy on the persistence and detectability of blood components in parietal bone fragments (of different post mortem periods and taphonomic conditions) through histological (Hematoxilin and Eosin, HE) and immunohistochemical (Glycophorin A, GYPA) analyses.

Results: The immunohistochemical investigation for GYPA showed the presence of RBCs under the form of erythrocyte debris or residues otherwise morphologically unidentifiable using only HE staining. Hence, while well-defined RBCs can be detectable only in the first week of decomposition, afterward these structures can be detectable with certainty only by immunohistochemical analysis, which reveals discrete quantities of RBC residues also in dry bone (post mortem interval, or PMI, of 15 years), but not in archaeological samples, in which the greater PMI and the different taphonomic conditions together could be the answer behind such difference.

Discussion: This study highlights the usefulness and potential of immunohistochemical detection of GYPA in RBC investigation and gives a realistic idea of the persistence and detectability of erythrocytes in different osteological taphonomic conditions, in contrast to results reported by some authors in literature. Another important result concerns the detection of RBC residues in dry bone, which opens the way to the possible use of RBCs in trauma interpretation. Am J Phys Anthropol 000:000–000, 2015.   © 2015 Wiley Periodicals, Inc.

After death, the body is exposed to different endogenous and exogenous variables that affect the progression of postmortem modifications, but also the appearance and structure of body, tissues, cells, and biomolecules. In literature many articles show how the body changes and appears at different decomposition stages (fresh, bloated, decay, post-decay, and skeletal) and in different environments (Kovarik et al., 2005; O'Brien and Kuehner, 2007; Voss et al., 2008; Lee Goff, 2009), whereas a study already dealt with how bone varies in various taphonomic conditions: e.g. bleaching is associated with sun exposure, whereas rarefaction with mechanical erosion (Quatrehomme and Işcan, 1997).

Always referring to the microscopic and biomolecular level, another important factor is how the organic and inorganic components of bone tissue react and change over time after death; Dent et al. (2004) showed how protein degradation and loss of mineral hydroxyapatite take place and emphasized the link between bone preservation and burial environment.

Unlike the above mentioned macroscopic and microscopic changes on bone tissue caused by environmental factors, scavengers and microorganisms, which have already been explored by the anthropological literature (Marchiafava et al., 1974; Hackett, 1981; Jans et al., 2004), little is known about bone soft tissue degradation and its components, particularly the cellular ones. This is of crucial importance since specific cellular populations surviving in bone, on fractures, or within a particular lesion, may be diagnostic of the timing of a fracture, or may be markers of a specific disease. Archaeological and anthropological literature have not yet approached this issue in a critical fashion, meaning that sometimes gross misinterpretation of the cellular structure on archaeological or forensic bone can be seen or read.

Bone tissue is formed by a protein structure (mainly collagen), strengthened by a mineral constituent (hydroxyapatite), and other organic components (e.g. mucopolysaccharides and glycoproteins) (Goffer, 1980; Janaway, 1997; Hare, 1988). Like every other tissue in the body, the bone tissue is characterized by

vascularization and therefore it houses in its structures all the cells and typical components of the blood: erythrocytes, leukocytes, platelets, albumin, fibrinogen and other plasma proteins. These components are ubiquitous in the bone tissue: specifically, they can be observed at the periosteum (within the vessels located here), inside Haversian and Volkmann canals (transverse or oblique canals that connect the meaning functional bone units, the so-called osteons), located in both lamellar and trabecular bone, in vessels placed into bone marrow spaces where hematopoiesis takes place, and therefore where blood cells can be found at various stages of maturation.

All of these components may be useful for understanding microscopic changes occurring within bone tissue during the decomposition process and in relation to different environmental conditions in which a body decomposes (water, high or low temperature, humidity, air). In this regard it would be interesting to know which variations take place (e.g. in the persistence and morphological changes of blood components) and in which specific structures.

In literature histological knowledge on bone tissue blood and marrow components is limited to diagnostic histopathology or paleopathology (van den Berg et al., 1990; Thiele et al., 2000; Schultz, 2001; Thiele et al., 2001; Wain et al., 2001; von Hunnius et al., 2006; Wilkins et al., 2008; van der Merwe et al., 2010; Schultz, 2012). In the forensic field, Penttilä and Laiho (1981) investigated the morphological changes of blood components (such as erythrocytes, white blood cells, and platelets) in cadaver blood at different post-mortal intervals; their results highlighted modifications in shape and size over time. Differently, other studies focused on morphological changes of blood cells over time with the intent of aging them and supplying possible information useful in the estimation of time since death (Dokgöz et al., 2001; Chen and Cai, 2006; Strasser et al., 2007; Wu et al., 2009). Finally, attention was also given to possible mistakes in differential diagnosis between erythrocytes and spores

and pollens (Cappella et al., 2015). Nevertheless, the observation of changes of blood components within bone tissue with regard to decomposition processes has not yet been fully explored: for how long are well-preserved blood components observable in bone tissue of cadavers and still detectable over time, in which area of bone tissue are they mostly protected or unaltered, and again what information can be gained from their persistence in relation to different environmental conditions, all questions that still remain unanswered. In other words, when we look at something that looks like a cell in old or ancient bones, is it really what we think we are looking at? Regardless of the caution which should be adopted in answering this question, in literature authors have report the presence of erythrocytes or structures like-erythrocytes in ancient skeletal remains (Stout and Teitelbaum, 1976; Penttilä and Laiho, 1981; Maat and Baig, 1990; Maat, 1991; Schweitzer et al., 2007; Cappella et al., 2015), but in most of these cases such findings have not yet been demonstrated by the specific techniques that science owns.

In this perspective, the current study aims to examine the changes and persistence of blood components (with special attention to erythrocytes) in bone tissue in relation to different decomposition and taphonomic conditions over time, by using histology (Hematoxylin-Eosin staining, HE) and immunohistochemistry (Glycophorin A, GYPA). The observation concerns histological assessments conducted on bone samples from both well and badly preserved cadavers (early decomposition), as well as from modern and ancient skeletonized remains, in order to observe how blood components appear, if they are noticeable, and for how long their persistence can be detected in such conditions. The present study may contribute to clarify the use of some valid techniques as an aid to the understanding of taphonomy of microscopic bone structures in skeletal remains and acquiring new important information useful in the field of paleopathology and trauma analysis in human remains, possibly demonstrating also the presence of bleeding in trauma, an aspect that has never been investigated in this perspective.

MATERIALS AND METHODS

Bone samples

The study was conducted on a total of 32 samples of parietal bone fragments, which had been previously sampled from forensic cases from eight different cadavers (with no blood diseases or bone pathological conditions reported) at the Institute of Legal Medicine of Milan, according to Police Mortuary Regulations. The

TABLE 1. Samples and related state of decomposition and PMI

Sample n°	Decomposition degree	PMI
1	Fresh cadaver	<24 hrs
2	Fresh cadaver	<24 hrs
3	Fresh cadaver	<24 hrs
4	Fresh cadaver	<24 hrs
5	Putrefied cadaver	48< hrs < 72
6	Skeleton	20 yrs
7	Skeleton	20 yrs
8	Skeleton	400 yrs

TABLE 2. Experimental conditions for sample 1, 2, 3 and 4

Sample n°	Time of air decomposition (wks)							
	T0 (ctrl)	T1	T2	T3	T4	T5	T6	T7
1	<24 hrs	1 wk	2 wks	3 wks	4 wks	5 wks	6 wks	7 wks
2	<24 hrs	1 wk	2 wks	3 wks	4 wks	5 wks	6 wks	7 wks
3	<24 hrs	1 wk	2 wks	3 wks	4 wks	5 wks	6 wks	7 wks

Sample n°	Laboratory treatment			
	1	2	3	4
4	Ctrl	Freezing	Boiling	Maceration

While sample 1, 2, and 3 were subjected to decomposition in air, sample 4 was subjected to laboratory treatment for simulating different taphonomic conditions.
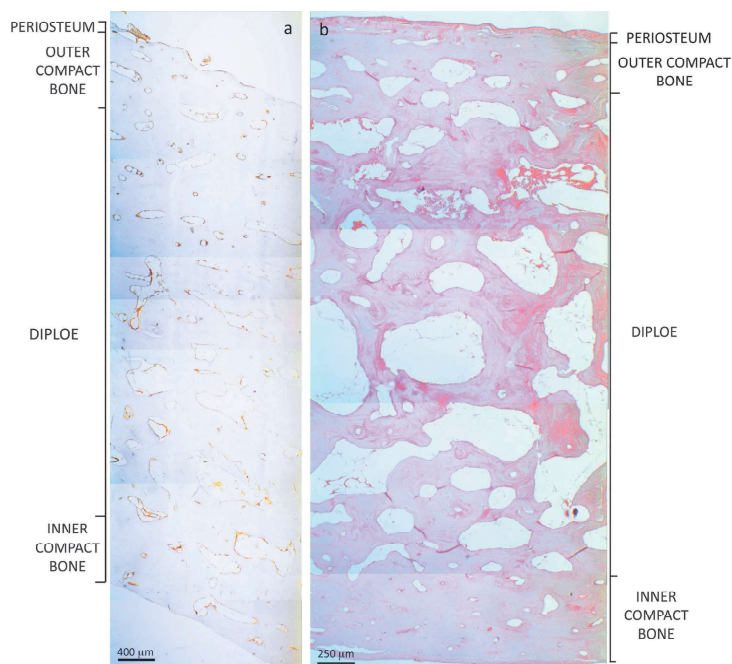
Fig. 1. Examples of thin parietal sections stained by anti-Glycophorin A (**a**) and Hematoxilin-Eosin (**b**) in which it is possible to observe the division of the different layers (several merged photomicrographs at 40× magnification). The score of RBC-related structures (such as isolated, agglomerated, organized in rouleaux, and/or eosinophilic accumulation) was performed in each vessel, canal, and marrow space of each layer within the entire section and then calculated as the amount of structures in which RBCs are localized. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

cadavers were characterized by different states of decomposition (early and advanced): the diverse post mortem intervals (PMI) as summarized in Table 1.

The parietal bone fragments from the three well-preserved cadavers (samples 1, 2 and 3, all with a PMI <24 hrs) were exposed to decomposition in air at standard environmental conditions (constant humidity and temperature), and then monitored weekly for the two following months (for a total of eight fragments per sample). Unlike the previous, sample 4 (derived from a well preserved cadaver with a PMI <24 hrs) was divided in fragments subjected to diverse laboratory procedures in order to simulate various taphonomic contexts, consisting in boiling (for 2 hrs), maceration (immersion in cold water for 1 month) and freezing (−4°C for a month). From every sample mentioned above, a small portion was used as a control for the evaluation of all the experiments, and therefore was not subjected to any experimental conditions. All these information are summarized in Table 2.

In addition, parietal fragments were taken also from human remains in which the decomposition process of skeletonization occurred in real conditions, in order to compare results obtained from actual circumstances and the experimental setting. The real conditions are represented by a cadaver at the early stages of putrefaction (sample 5, with a PMI ranging between 2 and 4 days) and in skeletonized remains of different PMI (sample 6,

7, and 8). In particular, samples 6 and 7 were obtained from two well-preserved modern skeletons (selected from the "Milano Skeletal Collection," a modern referenced skeletal population) (Cappella et al., 2014), while sample 8 was taken from a badly preserved ancient skeleton (1600–1630 AD).

Histological and immunohistochemical analysis

The same protocol was used for preparing all bone fragments for both histological and immunohistochemical analysis, consisting in fixation (in 10% Formalin for 24 hrs) and decalcification (in a decalcifying solution containing 14% hydrochloric acid, DECALC, Histo-Line Laboratories, Milan). The decalcification was performed at room temperature for a time dependent on the size and thickness of the fragment (ranging from 4 hrs for ancient bone fragments to 2 days for fresh bone samples). Once the optimal flexibility was achieved, the samples were washed up under running tap water for 24 hrs before proceeding with dehydration and embedding into paraffin. From each sample treated as described, 5 micron sections were obtained and subsequently stained with Hematoxylin and Eosin (HE) for standard histological analysis, as well as immunohistochemically stained for the identification of Glycophorin A (GYPA), a specific membrane antigen for erythrocytes. The latter was
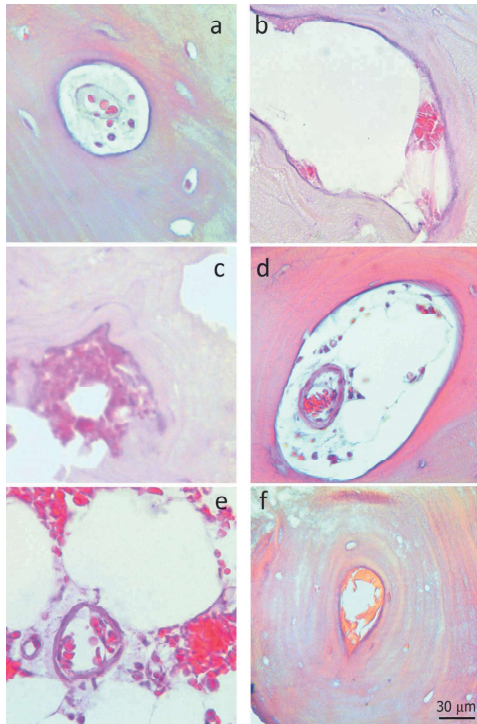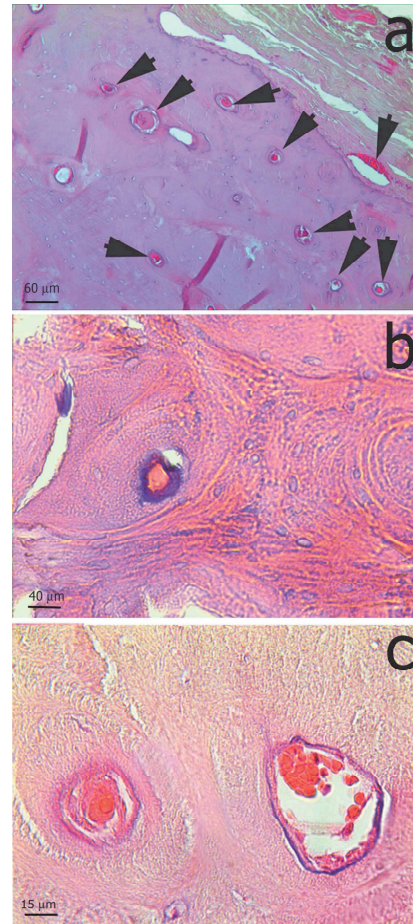
Fig. 2. Example of some parameters used as a score in the microscopic evaluation: **a**) Intracanal dispersed erythrocytes in a Haversian canal; **b**) Intracanal accumulated erythrocytes; **c**) Extracanal accumulated erythrocytes; **d**) Intravasal agglomerated erythrocytes into a Haversian canal; **e**) Intravasal erythrocytes in rouleaux into a marrow space; **f**) Intracanal aspecific accumulations. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Fig. 3. a) Sample 2 at T2, showing numerous Haversian canals full of blood components (100× magnification); in this case the histological analysis permits to highlight only a non-specific accumulation, as that shown in **b**) at a greater magnification (200×), in contrast with the well-defined erythrocytes visible at T0, as shown in c) (at 400× magnification). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

carried out using a monoclonal antibody (antihuman Glycophorin A, clone JC159, Dako-Dakopatts, Denmark) diluted 1:400 by a Dako Autostainers, as according to the literature (Luna, 1968; Cattaneo et al., 2010).

Microscopic evaluation

Histological analysis and immunohistochemistry allowed us to observe in detail the blood components of every single bone tissue structure localized in the different layers: periosteum, outer compact bone, diploe (spongy bone), and inner compact bone. Specifically, the analysis focused on the research of blood components within all structures such as Haversian and Volkmann canals, canals and vessels localized in bone trabeculae, and in bone marrow spaces (Fig. 2).

The score used for the evaluation (in structures of every layer) included different parameters concerning red blood cells and not merely their presence/absence: the erythrocytes, if present, were considered as dispersed/isolated, organized in rouleaux, agglomerated,
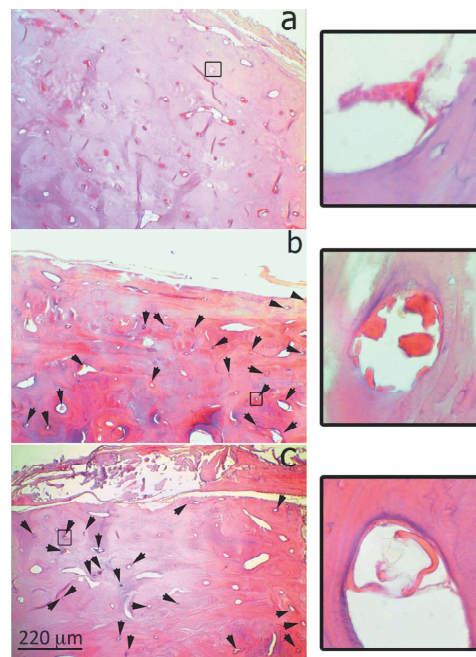
and as generic eosinophilic accumulation which could be (but not necessarily) related to decomposing RBCs; the presence of nucleated cells was also indicated, when recognizable. Moreover, as additional information, the position of cells with respect to the vessels and canals were inspected: RBCs were described as intra- or extra-vasal when concerning vessels, intra- or extra-canal if regarding Volkmann and Haversian canals, or indeterminate in case of impossibility in identifying their origin (Fig. 3).
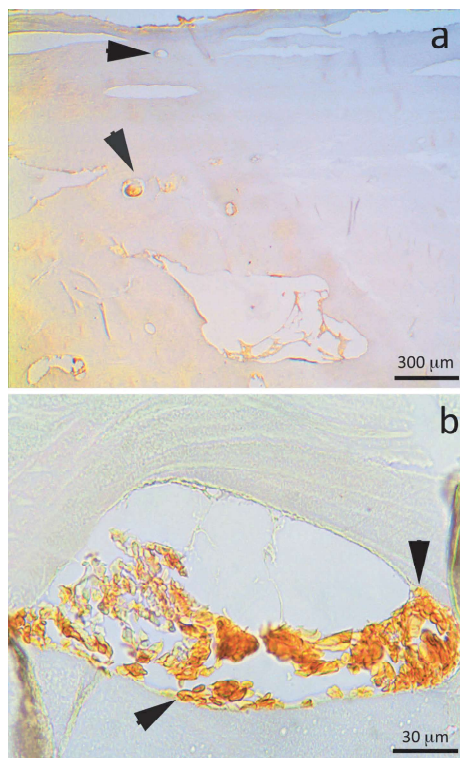
Fig. 4. a) Sample 2 at T3, displaying numerous Haversian canals still full of hematic components, indicated by black arrows (40× magnification). In this case the immunohistochemical analysis allows us to prove the presence of the erythrocytes in the nonspecific accumulation or in those cases where the RBC morphology is not clearly identifiable as such: in fact, the dark brown coloration means that the antigen GYPA is located into the canals and vessels, therefore proving the hematic origin also for those nonspecific accumulations highlighted by the histological analysis. In **b**), a Haversian canal of sample 3 at T4 at greater magnification, in which some preserved erythrocytes are still recognizable, as those indicated by black arrows (400×). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Fig. 5. a) Sample 1 at T1 shows numerous Haversian canals completely full of blood components, stained by HE (general view at 40× magnification). In this case, the canals do not need to be indicated by arrows as they are very well observable even without any help (an enlargement is visible on the right, at 630× of magnification). In **b**) a total view (40× magnification) of sample 1 at T6: despite the bone fragment has decomposed for about 1 month and a half, Haversian canals still have blood components inside in a high percentage out of the total of these structures, as for sample 1 at T7 (1 week later), shown at 40× of magnification in **c**). In both, b) and c) the nonspecific accumulations located in the canals are indicated by black arrows in order to help the observer in their identification, and an enlargement (630×) is proposed on the right, for a better visualization of the characteristics described. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Afterwards, a semi-quantitative assessment was possible by calculating the percentage of canals and bone marrow spaces containing red blood cells. In particular, a ratio between the total amount of structures in which blood elements in theory should have been present (e.g. vessels, marrow spaces, etc.) and the total number of structures where blood elements actually were present was calculated, in order to obtain percentages representing those structures in which RBCs or cellular hematic debris could still be detected. For all samples both histological and immunohistochemical evaluations were performed in order to confirm the results. When the histological observation did not allow for the certain identification of well-defined cells (scored as described

below), but only a strong eosinophil accumulation was detectable (referred to as hematic components by the authors, as confirmed by immunohistochemical analysis), the term "nonspecific accumulation" was used.

Finally, our first objective was to analyze how the features of red blood cells in bone tissue change in the monitored samples at the early stages of decomposition in air (for a maximum of 2 months). In this case the evaluation concerned the quantity of structures containing blood with respect to the total canals, vessels, and marrow spaces, but also the recognition of single/groups of defined red blood cells before their decomposition and disappearance. The same assessment was conducted also on all those samples in which the real decomposition process had already started a few days before (sample 5) and on those for which the natural skeletonization

process had occurred over a period of years or centuries (sample 6, 7, 8).

The subsequent examination of all samples by immunohistochemistry had the specific purposes of identifying red blood cell residues as fragments of RBCs when they are no longer morphologically recognizable, or, on the other hand, of dismissing other debris as non-RBC.

Microscopic observation was performed by an optical microscope (LEICA DMLB) at several magnifications (4×, 10×, 20×, 40×, 63×) with 10× eyepieces. By means of a photographic system (Eurekam 3.0, DV-3000), photographic surveys related to the most significant areas were captured.

RESULTS

Results from HE staining revealed well defined red blood cells only in control samples (1, 2, 3 at T0 and 4, not treated) as well as in sample 5 (fragment taken from putrefied cadaver with PMI between 48 and 72 hrs). In other words, red blood cells were no longer observable as cells after a week of decomposition in air: from this time onward solely nonspecific accumulations were noticeable by using HE staining (Fig. 4).

On the contrary, the immunohistochemical analysis adopted for the search of GYPA allowed for the detection of the presence of red blood cells even when degraded (no defined intact cells but only hematic debris), otherwise not detectable by the histological examination; thus, this analysis allowed us to confirm the hematic nature of most nonspecific accumulations located into canals and vessels stained by HE (Fig. 5).

Regarding the persistence of blood components (assessed as the percentage of bone tissue structures containing RBC), no great changes in terms of percentage were observed concerning the passing of time in the first 2 months of decomposition in air; the percentages obtained by the evaluation in control samples are similar to those found in samples at all experimental times and also after 2 months circa of decomposition in air (Fig. 8). The differences found concern mostly the abundance of coloration: if at very early times, like T0, T1, and T2, the canals are completely full of components, at a more advanced time the same structures show a lighter staining and little space is occupied by the components (Fig. 8).

Hematoxilin and Eosin revealed percentages between 60% and 80%, also confirmed by GYPA, which demonstrated percentages slightly higher, but very similar. As a matter of fact, the immunohistochemistry in general showed higher values compared to the common histological analysis, revealing in this sense not only the presence of entire cells, but also RBC fragments in samples at more advanced stages, not detectable by HE (percentages around 80% in cases where HE revealed just 60%) (Graph 1).

Similar results were found also in real putrefaction case (sample 5, PMI 48–72 hrs), where well-defined cellular elements as well as abundant nonspecific accumulations are represented by percentages greater than 60% in both histological and immunohistochemical analysis. Unlike the previous case, the detection of RBC in skeletonized human remains samples characterized by a higher postmortem interval (20 years, 6 and 7) was possible only by using the monoclonal antibody GYPA. Hematoxilin and Eosin did not allow to observe either well defined blood components nor nonspecific
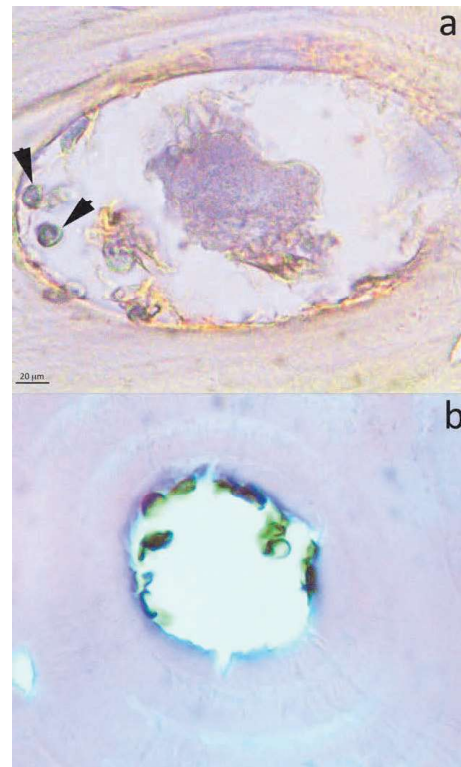
Fig. 6. a) Sample 2 (T5) stained by GYPA (630× original magnification). The black arrows in a Haversian canal indicate some erythrocyte like-structures; the dimension together with the absence of any immunostaining reaction indicate that a hematic origin is very unlikely and rather suggest microorganism contamination. **b**) Similar structures stained by HE (630× magnification) resembling erythrocytes whose antigenic nature was then not confirmed by GYPA staining (a). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

accumulation, but only generic debris (which could be hypothetically the result of putrefied products as well as both botanical and microorganism contamination, as displayed in Fig. 6 and 7), while the modern skeletal remains showed RBC residues in more than 10% of canals and marrow spaces (Fig. 8), the ancient skeleton (sample 8) shows a complete lack of any cellular debris, as confirmed by GYPA stained sections (Graph 1).

Concerning the results from the evaluation of samples subjected to lab treatments, both the analyses revealed a great similarity between the control and the frozen sample, which displayed both erythrocytes and nonspecific accumulations in consistent percentages (>50%), similar to what is seen in the control (60%). If on one hand the presence of blood components is maintained quite constant in the frozen sample (Fig. 9), on the other hand
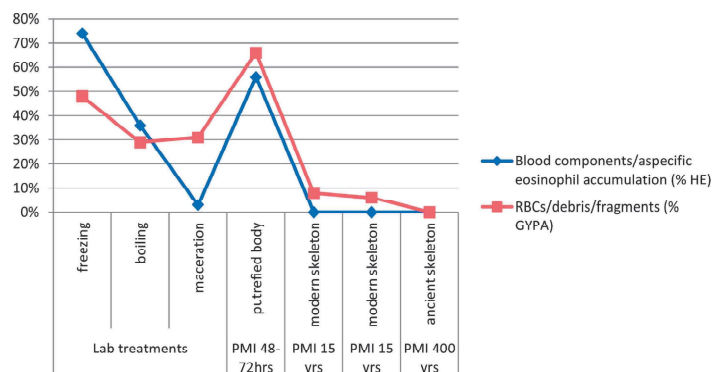
Graph 1. Percentages relative to the number of tissue structures (e.g. vessels, marrow spaces, etc.), among the total amount present, in which blood components and their debris could actually be detected. The slides of the cases subjected to laboratory treatment (Sample 4), the real putrefaction case (Sample 5), and the skeletonized ones (Samples 6–8) are considered: each sample has been analyzed in toto, relating the number of those structures in which blood components or RBC debris were present to the total number of bone and vascular structures observable (e.g. Haversian and Volkmann canals, vessels, and marrow spaces). The blue line refers to the histological analysis (HE), whereas the red one refers to the immunohistochemical staining (GYPA). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

the reduction of blood components (values close and under 30%) has been emphasized in the boiled and macerated sample. In this last case solely the immunohistochemical investigation stressed the RBC presence in canals and marrow spaces (Graph 1).

DISCUSSION

Given the poor knowledge concerning decomposition and taphonomic effects on blood components in the bone tissue, the present study focused on the acquisition of such information through histology and immunohistochemistry. Decalcified sections of parietal bone fragments, characterized by different states of decomposition and various postmortem intervals, were analyzed with the intent to observe differences due to various PMI and taphonomic variables. An evaluation of the presence/absence of erythrocytes or of their residues was considered in all canals and marrow spaces of the sections, and so their respective percentages. The immunohistochemical technique, by using antibodies for GYPA, made it possible to highlight data otherwise not noticeable in the sections stained by HE: in fact, it allowed us to observe whether the so-called nonspecific accumulations were of RBC origin or not. Obviously, they may contain also other cellular and hematic debris, which still have to be ascertained with further specific immunohistochemical investigations.

Results from both analyses demonstrated that cells appear well defined only if the period of time is shorter than 1 week (in control sample and in sample 5), a fact that suggests a certain influence of time on the preservation of morphology, as reported in a recent publication (Cappella et al., 2015). If the morphological preservation of RBCs is quite dependent on time (since the very early stages of decomposition), on the contrary the debris were seen to survive longer than expected: the percentage of bone tissue structures in which such cell remains are still detectable, especially by immunohistochemistry, is constant in the first 2 months of decomposition in air. The confirmation of these results comes also from the

assessments conducted on samples taken from a cadaver in putrefaction (PMI less than a week), representative of a real decomposition circumstance, in which biconcave red blood cells are still well identifiable and observed in considerable percentages of vessels and canals.
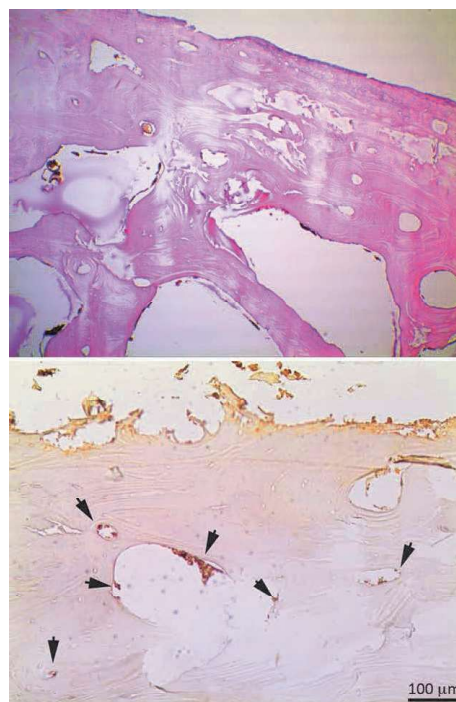
Also in modern skeletal remains with a PMI of 20 years RBCs are traceable (even if in very low quantities). In these last samples, in fact, despite the long time elapsed since death, a small percentage of canals and marrow spaces (10%) still contains RBCs, which were only discernable by immunohistochemistry; on the contrary, a total loss of these cells was observed in samples with PMI in the order of centuries (sample 8). This last account revealed divergent results from findings of some recent publications (Maat and Baig, 1990; Maat, 1991; Schultz, 2006; Schweitzer et al., 2007; Charlier et al., 2008; Setzer et al., 2013), where the identification of red blood cells even in ancient remains and through different techniques is frequently reported. Despite the present study did not investigate conditions equal to those reported in some publications, such as fossilization or mummification (conditions in which the bodies may be exceptionally well preserved), the authors observed great difficulties in recognizing well defined cells after a week of decomposition, and in most samples this was provided just because of the implementation of immunohistochemistry. The crucial point in this sense is obviously the technique used to demonstrate the presence of red blood cells, which has to be based not only on morphology but also on more satisfactory requirements at the biomolecular level. In this sense the suggestion is towards caution. In fact, the structures of botanical nature sometimes can be very tricky and may mimic the presence of erythrocytes, as reported by a recent work (Cappella et al., 2015), which shows the similarity between red blood cells and spores/pollen as well as providing important information about the survival and morphological modification of erythrocytes in vitro.

The results obtained from samples subjected to maceration, boiling and freezing showed that the normal
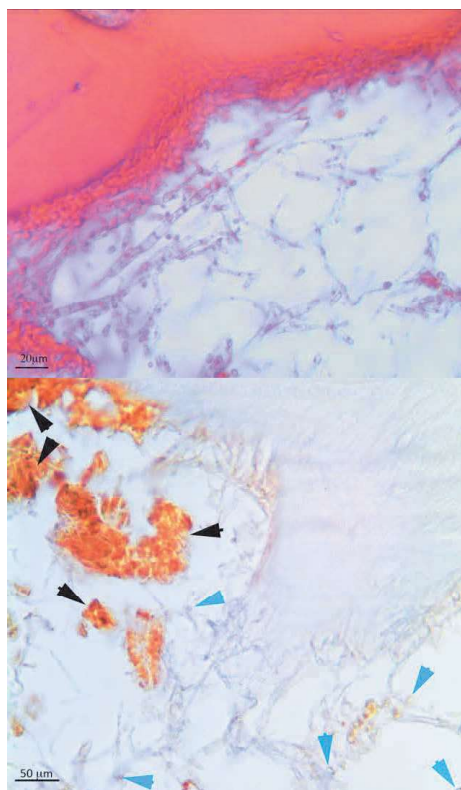
Fig. 7. The images show an example of microorganism contamination in a marrow space: the HE staining makes it difficult to interpret the section, whereas with the immunohistochemical technique it is possible to differentiate the hematic debris (stained in marrow and indicated by black arrows) from other non hematic origin materials/structures (light blue arrows). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Fig. 8. Sample 6, derived from cemetery skeletal remains (PMI of 15 years) stained by HE (left image at 40× magnification) and by GYPA staining (right image at 100× magnification). Only the immunohistochemical analysis permits to observe the presence of blood component residues (indicated by black arrows in the right image). The histological analysis, in fact, shows only the presence of some brown structures, which can resemble erythrocytes in terms of morphology and size, but, since they are not stained by Eosin as the components in all the other samples (staining includes nonspecific accumulation and defined blood components), their origin is supposed to be botanical or from some other contamination. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

laboratory practices, commonly required for cleaning up bones (maceration and boiling), remove most of the cell components of the tissue, even though some minimal quantity of debris can survive in it, therefore being detectable only by immunohistochemistry. On the contrary, the conservative procedures, such as freezing, revealed good conservation of cellular components, not only in terms of presence and quantity, but even in terms of quality (by preserving the morphology). Moreover, in the frozen sample the presence of some cracks caused by the thermal shock, which can cause migration of nonspecific accumulations or debris, was also detectable. The evaluation of the percentages of bone structures containing blood elements allowed, not only to verify the persistence of blood components, but also to highlight the sensitivity of the techniques exploited. In fact, in those cases in which the HE staining did not provide clear results, the immunohistochemical analysis showed

the presence even of the smallest blood group elements and hematic debris thus proving to be a reliable method for the study of hematic components also in decomposed material.

In general, the study was based on experimental conditions (control samples and samples from real conditions—putrefied cadaver and skeletal remains—apart) which simulate a process of decomposition, otherwise not easily comparable to what happens in the real decomposition of a body (as in a closed system like a cadaver): in the first case, it is the external lamina to be directly in contact with the surrounding environment; in the second, however, events concern the most superficial tissues that cover and protect the underlying bone tissue at the early stages of decomposition. This fact could give rise to some variations in the detection of blood components and in their survival in decomposed or decomposing bone
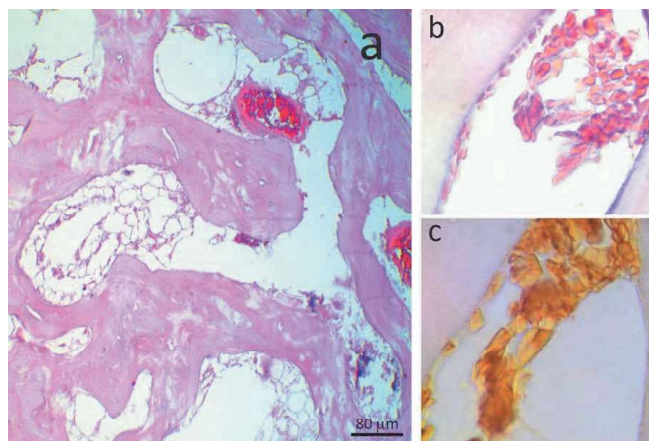
Fig. 9. a) Sample 4–2 (freezing treatment) stained by HE (400× original magnification). The erythrocytes are well localized everywhere in cortical bone in the Haversian Canals and into marrow spaces. The treatment seems to maintain constant the level of the components. **b**) and **c**): Images depicting the same structure (630× original magnification) stained by HE (b) and GYPA (c): despite the different staining technique, the same hematic elements are well-identifiable, therefore proving, thanks to GYPA, that the elements highlighted by HE are actually RBCs. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

tissue, although the potential of a technique such as immunohistochemistry in proving their presence and in being an optimal technique for the detection of blood components residues, even when they are just in traces, does not change. This study highlights the usefulness of immunohistochemical detection of GYPA in RBC investigation and gives an idea of the persistence of erythrocytes in different taphonomic and PMI conditions, often in contrast with results reported by some authors in literature, as mentioned above. The presence of RBC, as defined cells or as cellular residues, can be a possible marker not only in relation to PMI, taphonomic conditions, and decomposition, but also of bleeding or trauma in bone. This study is of course only a first approach to blood taphonomy in bone and it cannot be considered as a representative study for all the different conditions. Nevertheless it is a good starting point for future research, which needs to be conducted in order to acquire knowledge in this field. Nonetheless, the great capacity of the immunohistochemistry in searching for specific biomarkers in both fresh and dry bone tissue has emerged; the use of biomarkers such as RBC through immunohistochemistry could be crucial for forensic purposes such as the interpretation of PMI (analogous use of Luminol) or of trauma (for the first stages of healing or for the interpretation of vital lesions), all topics that still need to be deeply investigated and clarified in the anthropological field.

LITERATURE CITED

Cappella A, Castoldi E, Sforza C, Cattaneo C. 2014. An osteological revisitation of autopsies: Comparing anthropological findings on exhumed skeletons to their respective autopsy reports in seven cases. Forensic Sci Int 244:315.e1–315.e10.

Cappella A, Stefanelli S, Caccianiga M, Rizzi A, Bertoglio B, Sforza C, Cattaneo C. 2015. Blood or spores? A cautionary note on interpreting cellular debris on human skeletal remains. Int J Legal Med 129:919–926.

Cattaneo C, Andreola S, Marinelli E, Poppa P, Porta D, Grandi M. 2010. The detection of microscopic markers of hemorrhaging and wound age on dry bone. Am J Forensic Med Pathol 31:22–26.

Charlier P, Georges P, Bouchet F, Huynh-Charlier I, Carlier R, Mazel V, Richardin P, Brun L, Blondiaux J, Lorin de la Grandmaison G. 2008. The microscopic (optical and SEM) examination of putrefaction fluid deposits (PFD). Potential interest in forensic anthropology. Virchows Archiv 453:377–386.

Chen Y, Cai J. 2006. Membrane deformation of unfixed erythrocytes in air with time lapse investigated by tapping mode atomic force microscopy. Micron 37:339–346.

Dent BB, Forbes SL, Stuart BH. 2004. Review of human decomposition processes in soil. Environ Geol 45:576–585.

Dokgöz H, Arican N, Elmas I, Fincanci SK. 2001. Comparison of morphological changes in white blood cells after death and in vitro storage of blood for the estimation of postmortem interval. Forensic Sci Int 124:25–31.

Goffer Z. 1980. Archaeological chemistry: A sourcebook on the applications of chemistry to archaeology. New York: Wiley.

Hackett CJ. 1981. Microscopical focal destruction (tunnels) in exhumed human bones. Med Sci Law 21:243–265.

Hare PE. 1988. Organic geochemistry of bone and its relation to the survival of bone in the natural environment. In: Behrensmeyer AK, Hill A, editors. Fossils in the making: Vertebrate taphonomy and paleoecology. Chicago: University of Chicago Press.

Janaway RC. 1997. The decay of buried human remains and their associated materials. In: Hunter J, Roberts C, Martin A, editors. Studies in crime: An introduction to forensic archaeology. London: Routledge. p 58–85.

Jans MME, Nielsen-Marsh CM, Smith CI, Collins MJ, Kars H. 2004. Characterisation of microbial attack on archaeological bone. J Archaeol Sci 31:87–95.

Kovarik C, Stewart D, Cockerell C. 2005. Gross and histologic postmortem changes of the skin. Am J Forensic Med Pathol 26:305–308.

Lee Goff M. 2009. Early postmortem changes and stages of decomposition in exposed cadavers. Exp Appl Acarol 49:21–36.

Luna LG. 1968. AFIP: Manual of histological staining methods, 3rd ed. New York, NY: McGraw-Hill Book Company.

Maat GJR, Baig MS. 1990. Fossilized human blood cells from a Hellenistic settlement. Int J Anthropol 5:277–280.

Maat GJR. 1991. Ultrastructure of normal and pathological fossilized red blood cells compared with pseudopathological biological structures. Int J Osteoarchaeol 1:209–214.

Marchiafava V, Bonucci E, Ascenzi A. 1974. Fungal osteoclasia: A model of dead bone resorption. Calcif Tissue Res 14:195–210.

O'Brien TG, Kuehner AC. 2007. Waxing grave about adipocere: Soft tissue change in an aquatic context. J Forensic Sci 52: 294–301.

Penttilä A, Laiho K. 1981. Autolytic changes in blood cells of human cadavers II. Morphological studies. Forensic Sci Int 17:121–132.

Quatrehomme G, Işcan MY. 1997. Postmortem skeletal lesions. Forensic Sci Int 89:155–165.

Schultz M. 2001. Paleohistopathology of bone: A new approach to the study of ancient diseases. Am J Phys Anthropol 106–147.

Schultz M. 2006. Microscopic investigation of excavated skeletal remains: A contribution to paleopathology and forensic medicine. In: Haglund WD, Sorg MH, editors. Forensic taphonomy. The postmortem fate of human remains. Boca Raton, FL: CRC Press. p 200–222.

Schultz M. 2012. Light microscopic analysis of macerated pathologically changed bones. Bone histology. An anthropological perspective. Boca Raton, FL: CRC Press. p 253–296.

Schweitzer MH, Wittmeyer JL, Horner JR. 2007. Soft tissue and cellular preservation in vertebrate skeletal elements from the cretaceous to the present. Proc Biol Sci 274:183–197.

Setzer TJ, Sundell IB, Dibbley SK, Les C. 2013. Technical note: A histological technique for detecting the cryptic preservation of erythrocytes and soft tissue in ancient human skeletonized remains. Am J Phys Anthropol 152:566–568.

Stout SD, Teitelbaum SL. 1976. Histological analysis of undecalcified thin sections of archaeological bone. Am J Phys Anthropol 44:263–269.

Strasser S, Zink A, Kada G, Hinterdorfer P, Peschel O, Heckl WM, Nerlich AG, Thalhammer S. 2007. Age determination of blood spots in forensic medicine by force spectroscopy. Forensic Sci Int 170:8–14.

Thiele J, Kvasnicka HM, Zankovich R, Diehl V. 2000. Relevance of bone marrow features in the differential diagnosis between essential thrombocythemia and early stage idiopathic myelofibrosis. Haematologica 85:1126–1134.

Thiele J, Kvasnicka HM, Zankovich R, Diehl V. 2001. The value of bone marrow histology in differentiating between early stage polycythemia vera and secondary (reactive) polycythemias. Haematologica 86:368–374.

van den Berg H, Kluin PM, Vossen JM. 1990. Early reconstitution of haematopoiesis after allogeneic bone marrow transplantation: A prospective histopathological study of bone marrow biopsy specimens. J Clin Pathol 43:365–369.

van der Merwe AE, Maat GJR, Steyn M. 2010. Ossified haematomas and infectious bone changes on the anterior tibia: Histomorphological features as an aid for accurate diagnosis. Int J Osteoarchaeol 20:227–239.

von Hunnius TE, Roberts CA, Boylston A, Saunders SR. 2006. Histological identification of syphilis in pre-Columbian England. Am J Phys Anthropol 129:559–566.

Voss SC, Forbes SL, Dadour IR. 2008. Decomposition and insect succession on cadavers inside a vehicle environment. Forensic Sci Med Pathol 4:22–32.

Wain J, Pham VB, Ha V, Nguyen NM, To SD, Walsh AL, Parry CM, Hasserjian RP, HoHo VA, Tran TH, Farrar J, White NJ, Day NP. 2001. Quantitation of bacteria in bone marrow from patients with typhoid fever: Relationship between counts and clinical features. J Clin Microbiol 39:1571–1576.

Wilkins BS, Erber WN, Bareford D, Buck G, Wheatley K, East CL, Paul B, Harrison CN, Green AR, Campbell PJ. 2008. Bone marrow pathology in essential thrombocythemia: Interobserver reliability and utility for identifying disease subtypes. Blood 111:60–70.

Wu Y, Hu Y, Cai J, Ma S, Wang X, Chen Y, Pan Y. 2009. Timedependent surface adhesive force and morphology of RBC measured by AFM. Micron 40:359–364.

Paolo Fattorini[1]
Carlo Previderé[2]
Ilaria Carboni[3]*
Giorgio Marrubini[4]
Solange Sorçaburu-Cigliero[1]
Pierangela Grignani[2]
Barbara Bertoglio[2]
Paolo Vatta[5]
Ugo Ricci[3]

[1]Department of Medicine,
 Surgery and Health, University
 of Trieste, Trieste, Italy
[2]Department of Public Health,
 Experimental and Forensic
 Medicine, Section of Legal
 Medicine and Forensic Sciences,
 University of Pavia, Pavia, Italy
[3]SOD Diagnostica Genetica,
 A.O-U Careggi, Firenze, Italy
[4]Department of Drug Sciences,
 University of Pavia, Pavia, Italy
[5]Scuola Internazionale Superiore
 di Studi Avanzati (SISSA),
 Functional and Structural
 Genomics sector, Trieste, Italy

Research Article

Performance of the ForenSeq™ DNA Signature Prep kit on highly degraded samples

Next generation sequencing (NGS) is the emerging technology in forensic genomics laboratories. It offers higher resolution to address most problems of human identification, greater efficiency and potential ability to interrogate very challenging forensic casework samples. In this study, a trial set of DNA samples was artificially degraded by progressive aqueous hydrolysis, and analyzed together with the corresponding unmodified DNA sample and control sample 2800 M, to test the performance and reliability of the ForenSeq™ DNA Signature Prep kit using the MiSeq Sequencer (Illumina). The results of replicate tests performed on the unmodified sample (1.0 ng) and on scalar dilutions (1.0, 0.5 and 0.1 ng) of the reference sample 2800 M showed the robustness and the reliability of the NGS approach even from sub-optimal amounts of high quality DNA. The degraded samples showed a very limited number of reads/sample, from 2.9–10.2 folds lower than the ones reported for the less concentrated 2800 M DNA dilution (0.1 ng). In addition, it was impossible to assign up to 78.2% of the genotypes in the degraded samples as the software identified the corresponding loci as "low coverage" ($< 50x$). Amplification artifacts such as allelic imbalances, allele drop outs and a single allele drop in were also scored in the degraded samples. However, the ForenSeq™ DNA Sequencing kit, on the Illumina MiSeq, was able to generate data which led to the correct typing of 5.1–44.8% and 10.9–58.7% of 58 of the STRs and 92 SNPs, respectively. In all trial samples, the SNP markers showed higher chances to be typed correctly compared to the STRs. This NGS approach showed very promising results in terms of ability to recover genetic information from heavily degraded DNA samples for which the conventional PCR/CE approach gave no results. The frequency of genetic mistyping was very low, reaching the value of 1.4% for only one of the degraded samples. However, these results suggest that further validation studies and a definition of interpretation criteria for NGS data are needed before implementation of this technique in forensic genetics.

Additional supporting information may be found in the online version of this article at the publisher's web-site

1 Introduction

Human identification in forensics is conventionally performed by PCR analysis of a selected core of autosomal STR markers [1, 2]. In addition to this standard approach, and depending on the particular casework, lineage markers (i.e., Y-chromosome and mitochondrial DNA markers), X-chromosome STRs and autosomal SNPs can provide useful tools to complement genetic identifications, to determine exclusions and to generate investigative leads based on phenotypic or ethnicity-predicting DNA markers [3]. All these DNA typing systems can be multiplexed together to maximize the discriminating power required to address

Correspondence: Dr. Paolo Fattorini, Department of Medicine, Surgery and Health, University of Trieste, Italy, Strada per Fiume 447, 34149, Italy
E-mail: fattorin@univ.trieste.it
Fax: +39-04-03996265

---

*Current address: Institute of Public Health, Section of Legal Medicine, Catholic University, Rome, Italy

Colour Online: See the article online to view Fig. 2 in colour.

most problems of human identification. The final step is then represented by capillary electrophoresis (CE), the primary methodology to separate and detect alleles and base sequences.

During the last ten years, a new technology has been developed, based on the use of high-throughput platforms [4]. The Next Generation Sequencing (NGS) technology, in fact, offers the opportunity to sequence entire genomes or hundreds of selected genes/loci simultaneously in a very efficient way [5–8]. In the last few years, new NGS platforms, based on different sequencing chemistries, were developed to an extraordinary high-throughput level leading to the reduction of the DNA sequencing cost per sample. For example, in ancient DNA (*aDNA*) studies, a discipline with many overlapping issues with forensic genetics [9], the NGS approach, either in the form of direct (shot-gun) or targeted enrichment or sequence capture sequencing, now represents the "gold standard" in technology [10, 11]. More recently, the first kits designed for forensic identification purposes have been developed and customized [12].

Reliability of the results obtained from degraded/damaged DNA samples is a critical concern in forensic genetics as well as in *aDNA* analyses [9]. In fact, in *post mortem* tissues, the degree of DNA degradation increases at a rate, which mostly depends on environmental conditions [13]. It is well known that DNA damage hinders polymerization; consequently, the main feature of degraded samples is that they are quite refractory to PCR amplification [13, 14]. It is therefore a common finding to achieve partial or no profiles at all from these kinds of damaged/degraded DNA samples which can be characterized, in addition, by the presence of PCR artifacts potentially affecting the correct assignment of the genotypes. To overcome this concern, several analytical strategies have been developed such as the increase of the number of PCR cycles [14, 15], the reduction of the size of the amplicons [16, 17], the post-PCR cleanup of the amplified products [18], and, more recently, the repair of the DNA template in a pre-PCR step [19]. Currently, NGS represents the technology with the highest sensitivity; for this reason, there is a strong interest in understanding its ability to recover as much genetic information as possible from highly degraded forensic DNA samples for which the conventional PCR/CE approach failed [12]. This potential clearly emerged from the first study on this topic which was performed using the HID-Ion AmpliSeq™ Identity Panel kit on sonicated DNA samples [20].

In this paper, a trial set of DNA samples was artificially enriched in Apurinic-Apirymidinic sites (A-P sites) [21], the most frequent lesion found in aged and probably even in forensic DNA samples [22]. This set of damaged DNA samples was used to evaluate the performance and reliability of the ForenSeq™ DNA Signature Prep kit (in its beta version) using the Illumina MiSeq NGS sequencer. This kit allows the simultaneous typing of 29 autosomal STR markers, 25 Y-STR markers, 9 X-STR and 95 identity informative SNP markers when using primers mix A [23].

The aim of this study is to evaluate the ability of this new technology to produce reliable genetic results from difficult samples, in comparison with the conventional PCR/CE approach.

2 Materials and methods

2.1 Samples

Sample FM was obtained from the buffy coat derived from 500 mL of a male donor's blood who provided informed consent in agreement with the Ethical Committee of the A.U-O of Trieste (Italy). Briefly, after phenolic extraction and ethanol precipitation, the sample was resuspended in bi-distilled water, quantified by NanoDrop ND-1000 (Thermo Fisher Scientific Inc., MA, USA), aliquoted in Eppendorf tubes and stored at −20°C until use.

To enrich the samples in A-P sites, a method previously published was followed [24]. Briefly, 200 μL aliquots, each one containing about 13 μg of DNA, were incubated in a water bath at 70°C for 6.0, 8.0 and 10.0 hours (namely sample FM-6, FM-8 and FM-10, respectively). After incubation, each sample was immediately centrifuged in TE (10 mM Tris pH 7.4, 0.1 mM Na$_2$EDTA pH 8.0)-equilibrated Ultracel 3 K Amicon Ultra columns (Millipore, MA, USA) for 30 min at 12 000 rpm. The retained samples were washed once with TE, centrifuged at 12 000 rpm for 20 min and finally recovered following the manufacturer's recommendations. The samples were then adjusted with TE to a final volume of 300 μL and stored at −20°C until further use. A negative control of sample preparation (NCSP), represented by TE filtered through Ultracel 3 K column, was also included.

To assess the molecular weight (MW) of the samples, aliquots of each sample were run through 1.2% agarose gels stained with Ethidium Bromide (0.5 μg/mL). λ/HindIII and Easy Ladder (purchased from Bioline, UK) were run simultaneously as MW markers.

2.2 DNA quantification

One-microliter aliquots from each sample were quantified using NanoDrop spectrophotometry. In addition, qPCR-based methods were used to test the amplifiability of the samples. Quantifiler Human DNA Quantification kit and Quantifiler Human DNA Trio Quantification kit (Thermo Fisher Scientific) were used to test 1.0 μL aliquots of each sample in triplicate assays, following the manufacturer's recommendations, on a 7500 Real Time PCR System (Thermo Fisher Scientific). Calibration was performed in triplicate. No template control and NCSP were included. The degradation index (D.I.) for the Quantifiler Trio was calculated by the ratio between the SA (small amplicon) and LA (large amplicon) quantification data. The UV/qPCR ratios were calculated as previously published [24].

DNA control sample 2800 M was provided with the ForenSeq™ DNA Signature kit at a nominal concentration of 10 ng/µL.

2.3 NGS experiment

2.3.1 Library preparation and sequencing

The *beta*-version of the ForenSeq™ DNA Sequencing kit (Illumina) was used to prepare the libraries. As recommended in the ForenSeq™ DNA Signature Prep Guide [23], one nanogram of the unmodified control sample FM DNA, together with three scalar amounts (1.0, 0.5 and 0.1 ng) of sample 2800 M were used as positive controls.

The maximum input DNA volume (5.0 µL) was utilized for the degraded samples FM-6, FM-8 and FM-10. The resulting DNA amounts were 1.68, 0.840 and 0.055 ng for samples FM-6, FM-8 and FM-10, respectively, as assessed by the *Quantifiler Trio* SA probe. Similarly, two dilutions of sample FM-6, namely FM-6 (1:1) and FM-6 (1:3) containing 0.840 and 0.420 ng, respectively, were also prepared and included in the experiment. In total, nine samples were analyzed in duplicate tests. A negative PCR amplification control (no template—NTC) and the NCSP were processed alongside the steps of the experiment, according to the manufacturer's recommendation [23].

Briefly, the DNA samples were amplified and tagged through 18 cycles in a GeneAmp® PCR System 9700 thermal cycler, according to the protocol described in [23]. The DNA primer mix A (DPMA) selected for this experiment allowed the simultaneous amplification of 62 STR markers (28 autosomal, 25 Y-chromosome, 9 X-chromosome) plus Amelogenin and 95 identity informative SNPs. The target enrichment procedure was completed by adding index adapters to the tagged DNA templates and setting up 15 additional PCR cycles. Each library was then purified and normalized by the bead-based recommended procedure. Five microliter aliquots of each one of the twenty normalized libraries were pooled together. Seven microliters of the pooled libraries were diluted with the hybridization buffer (HT1), added to the Human Sequencing Control (HSC) and denatured following the recommended procedures.

2.3.2 NGS run and data analysis

The sequencing run was performed on the MiSeq® system (Illumina) converted into the MiSeq FGx Forensic Genomic System version according to the manufacturer's protocol [23, 25]. Data analysis was initially performed using the provided *beta* version of the ForenSeq Universal Analysis Software. Quality metrics (intensity, i.e., number of reads and the corresponding alleles) for each run were then converted into a Microsoft Office Excel 2007 spreadsheet and analyzed with Excel and R, version 3.0.1 (2013-05-16, Copyright 2013 The R Foundation for Statistical Computing). The following analyses were then performed.

2.3.3 Molecular analysis

The mean number of reads/sample was calculated from the overall number of reads for the STR and SNP panels in the two experiments. For example, the calculation performed for the 1.0 ng dilution of sample 2800 M is the following: (511 850 + 246 002)/2 = 378 926 (see Supporting Information Table S1). To evaluate the intra-experiment variability (e.g., between the data recovered from the two replicates of each sample), the number of reads for each marker was transformed in two-axes plot diagrams. The concordance was then calculated by r^2.

The average depth of coverage (DoC)/locus was calculated by dividing the overall number of reads generated for a specific panel (in the two runs) by the number of markers of the same panel which gave at least 50 reads (that is the cut-off value for locus call). For example, the 0.1 ng dilution of sample 2800 M gave an overall number of 191 732 reads generated by 125 markers producing at least 50 reads. Therefore, the average DoC/locus was calculated as follows: 191 732/125 = 1534 (see Supporting Information Table S1). The same raw datasets were also used to calculate the median and the min/max values.

The relative DoC (*r*DoC) of each locus was calculated for samples FM and the 1.0 ng dilution of sample 2800 M by dividing the average number of reads of that specific marker by the average number of reads of the corresponding panel. For example, a total of 324 939 and 336 841 reads were generated for the STR panel in the two replicates of control sample FM, respectively (mean = 330 890). In those tests, 22 541 and 23 115 reads were assigned to locus D22S1045, respectively (mean = 22 828). Therefore, the *r*DoC of locus D22S1045 in sample FM was calculated as 22 828/330 890 = 0.0689.

For the STR panel, the percentage of stutter reads was calculated by dividing the number of reads identified as stutter products (e.g., -1 repeat) by the total number of reads of the corresponding STR panel.

2.3.4 Genetic analysis

The default settings of the ForenSeq Universal Analysis Software Guide [26] were applied to the data generated in the experiment. The software was programmed to provide analysis metrics, describing the quality of the data generated locus per locus.

The minimum DoC to assign a genotype depends on the NGS technology and on the end-point of the study [27–30]. The only pivot study performed with the ForenSeq™ DNA Sequencing kit reported a minimum DoC of 10x [31]. Our preliminary data analysis, performed with a cut off of 30x, evidenced mistyping error rates of up to 2.2% in these three trial samples [32]. For this reason, we tried to reduce these error rates by increasing the minimum coverage for "locus call" to 50x. Below this cut-off value, each specific locus was tagged by the software with the flag "low coverage".

The analytical (ΛT) and interpretation (IT) thresholds were set up as follows: for loci with DoC from 50x to 650x: 10 and 30 reads, respectively; for loci with DoC $\geq$ 651x: 1.5% and 4.5% of the total reads of the locus, respectively.

The heterozygote balance for each STR and SNP locus was calculated by dividing the number of reads of the lower covered allele by the higher one. Values $<$ 0.60 were tagged by the software with the flag "imbalance".

Stutter bands for the STR loci were identified by the software according to predefined % intensity values of the true allele, as reported in the ForenSeq Universal Analysis Software Guide [26]. These values show a high degree of variability, ranging from 7.5% for D4S441 and Penta E to 50% for DYS481. The software tagged the locus with a "stutter flag" when stutters were above the predefined % intensity values.

The correctness of the allelic calls was initially verified comparing the genotypes recovered from each replicate of sample 2800 M to the expected genotypes of this reference control sample reported in the ForenSeq™ DNA Signature Prep Guide [23]. The allelic calls of each replicate of samples FM-6, FM-8 and FM-10 were compared to the genotypes obtained from the reference sample FM. The comparison allowed an easy identification of the amplification artifacts in each replicate.

The interpretation guidelines for genotyping the NGS data obtained from the two replicate analyses of each sample are based on the "*consensus*" approach [33] and are described as follows, according to [24] with minor changes:

(i) no result (NR) for a given locus was defined when at least one of the two replicates gave either low coverage (e.g., $<$ 50x) or allelic reads below the interpretation threshold;
(ii) the genotype for a given locus was considered correct (C) if both tests gave the expected genotype;
(iii) when different genotypes are obtained at a given locus in the two replicate analyses, the result was considered

unreliable (U), as it was not possible to unambiguously assign a genotype for that given locus;
(iv) the genotype for a given locus was considered wrong (W) if both replicates showed the same incorrect genotype.

In order to identify sequence variants, STR alleles generated by NGS were verified with STRait Razor [34] when possible.

2.4 Conventional PCR/CE

The same maximum input DNA volume (5.0 μL) utilized for the NGS analysis of degraded samples FM-6, FM-8 and FM-10, was used for the conventional STR typing approach amplifying the 16 STR loci plus Amelogenin contained in the commercial kit PowerPlex® ESI 17 system (Promega, USA), in duplicate tests. Five-hundred picograms of the unmodified control samples FM and 2800 M were tested. Both NTC and NCSP were included. Thirty cycles of PCR were performed under standard conditions. The resulting amplified PCR products were then separated by capillary electrophoresis using a 3500 Genetic Analyser Sequencer (Thermo Fisher Scientific). The analytical (AT) and interpretation (IT) thresholds were set at 100 and 300 rfu respectively. The genetic typing of the samples was performed following the "*consensus*" approach described above.

3   Results and discussion

3.1 DNA degradation

In order to produce damaged DNA samples useful to test the reliability of the NGS approach, three trial DNAs (namely FM-6, FM-8 and FM-10) were artificially degraded by aqueous hydrolysis, according to a simple and well-defined

Table 1. Molecular features of the trial DNA samples. NanoDrop: amount of DNA as assessed by UV spectrophotometry at 260 nm (mean $\pm$ s.d.). qPCR-based quantification: amount of DNA (mean $\pm$ s.d.) as assessed by real-time PCR using the *Quantifiler* and *Quantifiler Trio* kits, respectively. The calibration data showed r^2 values $>$ 0.998 in the ranges of quantification (from 50.0 ng to 23.0 pg for *Quantifiler* and from 50.0 ng to 5.0 pg for *Quantifiler Trio*). hTERT: human telomerase reverse transcriptase. Y: Y-specific probe. SA: small autosomal probe. LA: large autosomal probe. In round brackets, the length of the molecular probes, in bp. In square brackets, the UV/qPCR ratios. IPC: C_t values of the Internal Positive Control

Sample	Incubation time at 70 °C (hours)	NanoDrop (ng/μL)	qPCR-based quantification (ng/μL)					
			Quantifiler		Quantifiler Trio			
			hTERT (62 bp)	IPC	Y (75 bp)	SA (80 bp)	LA (214 bp)	IPC
FM (ctrl)	0	32.9 $\pm$ 0.7	31.384 $\pm$ 2.721 [1.0]	25.2 $\pm$ 0.2	29.063 $\pm$ 2.454 [1.1]	32.946 $\pm$ 3.094 [1.0]	44.531 $\pm$ 4.251 [0.7]	27.3 $\pm$ 0.1
FM-6	6.0	38.0 $\pm$ 0.8	1.305 $\pm$ 0.158 [29.1]	25.3 $\pm$ 0.1	0.274 $\pm$ 0.021 [138.7]	0.336 $\pm$ 0.030 [113.1]	$<$ LOQ	26.9 $\pm$ 0.1
FM-8	8.0	32.9 $\pm$ 1.2	0.399 $\pm$ 0.037 [416.5]	25.2 $\pm$ 0.2	0.124 $\pm$ 0.011 [1,096.7]	0.168 $\pm$ 0.007 [1,134.4]	undetermined	26.8 $\pm$ 0.2
FM-10	10.0	38.8 $\pm$ 1.0	0.045 $\pm$ 0.010 [862.5]	25.4 $\pm$ 0.2	0.008 $\pm$ 0.001 [4,850.0]	0.011 $\pm$ 0.001 [3,227.0]	undetermined	26.7 $\pm$ 0.1

depurination protocol [24]. This procedure enriches the templates in A-P sites [21], promoting DNA degradation *via* β-elimination [35]. In addition, deamination of the nucleobases occurs in these samples even though at low level [36]. The main molecular features of these three trial samples and of the control sample FM are summarized in Table 1. Samples FM-6, FM-8 and FM-10 exhibited a mean molecular weight (MW) < 500 bp, with a degree of degradation proportional to the length of the treatment (see Supporting Information Fig. S1).

Another important feature is that all treated samples were quite refractory to polymerization as shown by the limited amount of DNA assessed by the four qPCR probes used in this study. This finding emerged clearly when the UV quantification data were compared to the amount of DNA assessed by each molecular probe [24], thus calculating each probe-related UV/qPCR ratio. In addition, it was impossible to calculate the degradation index for the *Quantifiler Trio* data as the LA gave no results for all the trial DNA samples. Negative control samples (NCSP and NTC) gave the expected results (*undetermined* for all the probes). The untreated sample FM showed UV/qPCR ratios close to one (0.739–1.132) and a degradation index of 0.74. These values are typical of high-molecular-weight DNA [24]. No inhibition was detected in all samples as shown by the C_t values of the IPC (Internal Positive Control) probes of the qPCR kits. As shown in Table 1, those values were in fact consistent with the average IPC C_t values of the quantification standards (25.0 ± 0.2 for *Quantifiler* and 27.2 ± 0.7 for *Quantifiler Trio*).

In conclusion, the three samples, whose molecular features are described above, can be used as models to mimic problematic forensic DNA samples.

3.2 NGS analysis

3.2.1 MiSeq run data

The MiSeq sequencing run showed a mean cluster density of 1112 K/mm² and produced about 13.2 M reads. Of these, about 92.8% passed the quality filters. An overall value of 0.196% of the molecules had a cluster phasing while about 0.088% had a cluster pre-phasing. Therefore, all the parameters of this run completely fulfilled the recommended parameters of MiSeq platforms [26, 31].

3.2.2 Control DNAs (samples 2800 M and FM) and negative controls

The beta version of the ForenSeq™ DNA Signature kit contained four STRs (namely DXS10148, DXS8377, DYS456 and SE33) which are now not present in the current commercial version of the ForenSeq™ DNA Signature kit. In this paper, the genetic results for these four markers were not considered.

3.2.2.1 Control sample 2800 M

The NGS results obtained for the duplicate tests of scalar amounts (1.0, 0.5 and 0.1 ng) of control sample 2800 M are reported in Supporting Information Table S1. The mean number of reads/sample was 378 926, 301 731 and 138 926 for DNA input of 1.0, 0.5 and 0.1 ng, respectively. These values exceeded by 37.7–46.5% those reported elsewhere for the same amounts of this control sample [31]. The most likely explanation for this finding is that in our experiment 20 samples per run were pooled, whereas Churchill et al. [31] pooled 32 samples per run. In addition, it was observed that about 67.1–69.0% of the reads were generated for the STR panel and no more than 2.2–4.5% of the reads of the STR panel were identified as stutter products (–1 repeat).

Reads were generated for all the markers of both panels, in each one of the three 2800 M DNA dilutions. The 0.1 ng DNA dilution showed the highest variability between the two replicates, as indicated by lower r^2 values denoted on the concordance column. The average DoC across STR and SNP panels for the three 2800 M DNA dilutions are shown in Supporting Information Figs. S2 and S3, respectively. No correlation between the average DoC found for each marker and the length of the corresponding amplicon was found for the two panels of markers. This difference in the average DoC within and between STR and SNP markers has already been described [31]. According to previous reports it could be ascribed mainly to a different efficiency of the target enrichment procedure caused by the peculiar melting temperature of each one of the multiplex-couples of primers used in the experiments [14, 31, 37].

Out of the 58 STR loci (plus Amelogenin) considered in this study, low coverage of the locus (e.g., DoC < 50x) was found only for the 0.1 ng 2800 M DNA dilution, with a frequency of 0.009. The imbalance of heterozygous loci was observed mainly in the 0.1 and 0.5 ng dilutions of sample 2800 M (frequency of 0.482 and 0.089, respectively, versus 0.036 of the 1.0 ng dilution), as previously reported [31]. In addition, the most unbalanced allele was the longer one with a frequency of 0.764, a finding already observed in previous studies [38–40]. Two stutter reads above the defined parameters were found: the first one affected one of the two 1.0 ng replicates at locus DYS576 and the second one affected one of the two 0.1 ng replicates at locus D12S391. No other artifact was scored for the STR panel of markers. Sequence variance (e.g., repeat-motif variations) was observed at locus D9S1122 [31] in all the replicates.

Out of the 95 SNP markers considered here, locus rs2342747 gave the homozygous genotype A/A versus the expected A/G in all the tests performed. This locus has been already described to be prone to locus and allele drop outs [31]. Low coverage of the loci was found only in the 0.1 ng dilution with a frequency of 0.016. In addition, for the SNP panel of markers, imbalance was observed mainly in the 0.1 and 0.5 ng dilutions (frequency of 0.428 and 0.128, respectively, versus 0.071 of the 1.0 ng dilution). Reads below the IT were

scored in the 0.1 ng sample with a frequency of 0.028. No other drop outs were observed.

The results described here show the robustness and reliability of the NGS approach even from sub-optimal amount of high-quality DNA [12, 31]. However, it is likely that the low frequency of artifacts found in these dilutions of the control sample 2800 M is due to high number of reads/sample which were obtained, thanks to the reduced number of samples ($n = 20$) pooled in the library.

3.2.2.2 Control sample FM

As described in paragraph 2.3.3, the NGS analysis of 1.0 ng of the control sample FM generated a mean value of 472 502 reads, among which 70.0% were obtained for the STR panel with 1.3% of the reads of this panel identified as stutters (see Tables 2 and 3). The total number of the reads of sample FM exceeded by 19.8% the value found for the same input DNA (1.0 ng) of control sample 2800 M. A variation (in reads/sample) of up to 34.5% has been already described in a

Table 2. NGS analysis of the STR panels for samples FM, FM-6, FM-8 and FM-10. Reads: overall number of reads obtained in the two tests. In bold characters, the reads assigned to markers with ≥ 50x coverage. In brackets, the overall number of markers with ≥ 50x coverage in the two tests. Concordance: r^2 values of the X-Y plots representing the concordance between the two replicate tests. DoC: depth of coverage/locus (reads/locus) calculated on markers with ≥ 50 reads. The mean values (± standard deviation) are reported in italic while the median values are in bold characters. The min/max values are in brackets. Low coverage: frequency of markers showing < 50 reads not considering the locus DXS8377 (*) (see also paragraph 3.2.2). Imbalance: frequency of heterozygous imbalanced loci. In brackets, the ratio between the reads of the two alleles (mean ± standard deviation). Drop out: frequency of drop outs for heterozygous markers. I.T.: frequency of reads below the interpretation threshold. Stutter: percentage of the reads attributed to stutter products out of the total number of reads. In brackets, the frequency of stutter products whose number of reads exceeded the defined thresholds (see Materials and Methods for details). Drop in: frequency of drop ins

Sample	Reads	Concordance	DoC	Low coverage	Imbalance	Drop out	I.T.	Stutter	Drop in
FM (ctrl)	661 780 **661 707** (124/126)	0.965	*5 336 ± 5 415* **3817** (143; 32 339)	0*	0.057 (0.562 ± 0.026)	0	0	1.27 % (0)	0
FM-6	131 734 **131 482** (69/126)	0.915	*1 905 ± 2 873* **725** (51; 12 703)	0.436	0.323 (0.429 ± 0.112)	0.117	0.029	1.91 % (0)	0.014
FM-8	44 571 **44 492** (44/126)	0.807	*1 011 ± 1 328* **468** (54; 5 765)	0.634	0.440 (0.361 ± 0.112)	0.320	0	3.73 % (0.045)	0
FM-10	18 804 **18 734** (28/126)	0.627	*669 ± 802* **337** (56; 3 126)	0.761	0.277 (0.302 ± 0.120)	0.444	0.111	0.31 % (0.027)	0

Table 3. NGS analysis of the SNP panels obtained from samples FM, FM-6, FM-8 and FM-10. Reads: overall number of reads obtained in the two tests. In bold characters, the reads assigned to markers with ≥ 50x coverage. In brackets, the overall number of markers with ≥ 50x coverage in the two tests. Concordance: r^2 values of the X-Y plots representing the concordance between the two replicate tests. DoC: depth of coverage/locus (reads/locus) calculated on markers with ≥ 50 reads. The mean values (± standard deviation) are reported in italic while the median values are in bold characters. The min/max values are in brackets. Low coverage: frequency of markers showing < 50 reads, not considering the *rs2342747 and 430046* markers (*) (see paragraph 3.2.2.2). Imbalance: frequency of heterozygous imbalanced loci. In brackets, the ratio between the reads of the two alleles (mean ± standard deviation). Drop out: frequency of drop outs for heterozygous markers. I.T.: frequency of reads below the interpretation threshold

Sample	Reads	Concordance	DoC	Low coverage	Imbalance	Drop out	I.T.
FM (ctrl)	283 224 **283 155** (186/190)	0.904	*1 522 ± 1 044* **1 205** (128; 5 784)	0*	0.045 (0.477 ± 0.077)	0	0
FM-6	128 789 **128 129** (137/190)	0.977	*935 ± 1 865* **303** (54; 11 630)	0.257	0.326 (0.407 ± 0.135)	0.102	0.073
FM-8	66 647 **66 026** (92/190)	0.958	*718 ± 1 740* **216** (53; 11 518)	0.494	0.242 (0.357 ± 0.156)	0.242	0.054
FM-10	30 015 **29 322** (42/190)	0.949	*698 ± 1 400* **167** (63; 7 290)	0.757	0.166 (0.357 ± 0.018)	0.333	0.050

previous paper [31], for the same amount (1.0 ng) of two different DNA control samples (cell lines 2800 M and 9947A). Therefore, it is likely that the inter-sample variation found here could be ascribed to poor precision in the quantification of the samples as well as to intrinsic chemical features of the templates [31, 35]. Reads were generated for all markers of both panels, with low variability between the two replicates ($r^2 = 0.965$ and 0.904 for STR and SNP markers, respectively).

Very similar variations in DoC for STR markers were found between sample FM and the control DNA 2800 M. This emerged by the comparison of the relative DoC (*r*DoC) of each locus in the two samples (see Supporting Information Fig. S4). The mean ratio value ± s.d. was, in fact, 1.015 ± 0.412 (median value = 0.991; min = 0.078; max = 3.370). It is interesting to note that the markers characterized by the lowest ratios (DY456 and DXS8377) were removed from the commercial version of the kit, while marker DYS392, showing a ratio of 3.370, has been already described [31] as the most variable in terms of DoC among different samples. The frequency of imbalance of heterozygote STR loci was 0.057, for sample FM, similar to the one found for control sample 2800 M (0.036). Marker DYF387S1 showed a tri-allelic pattern (37/38/39) which is a frequent finding for this duplicated marker [41].

Also, very similar variations in DoC were found for SNP loci between samples FM and 2800 M (see Supporting Information Fig. S5). The mean ratio value ± s.d. was, in fact, 1.056 ± 0.274 (median value = 1.045; min = 0.005; max = 1.632). Also for this panel of markers, it is interesting to note that the markers characterized by the lowest ratios (*rs2342747* and *rs430046*) have been already described [31] as those most affected by artifacts. In addition, these two loci showed low coverage in all the replicates. For this reason, they were not further considered in this study. Allelic imbalance of SNP loci showed a frequency of 0.045 similar to the one (0.071) calculated for the same amount of DNA of sample 2800 M.

In conclusion, control DNA 2800 M and the high-molecular-weight DNA sample FM showed comparable features in terms of NGS typing results for both panels of STR and SNP markers.

3.2.2.3 Negative controls

No reads were scored for the STR panel in the two negative controls. A total of 14 and 18 reads were scored in the NTC and, respectively, the NCSP for the SNP panel. Below the AT, no more than 2 reads/locus were found.

3.2.3 Samples FM-6, FM-8 and FM-10

Quantification of degraded samples is a critical concern in forensic and ancient DNA analysis [13, 22]. Molecular quantification of such samples provides, in fact, only an assessment of the DNA amount "relative" to the qPCR probe selected [24]. In addition, in this study, the degradation index of samples FM-6, FM-8 and FM-10 could not been calculated

thus supporting the evidence of a severe modification of the templates. For these reasons, we did not follow the recommendation of using 1.0 ng of genomic DNA as reported in the ForenSeq™ DNA Sequencing kit protocol [23] but rather decided to employ the maximum allowed DNA volume for the NGS reaction, corresponding to 5.0 μL (equivalent to 1.64–0.055 ng of DNA, as assessed by the SA probe).

The NGS results achieved for these degraded samples are shown in Tables 2 and 3, for STR and SNP markers, respectively. A significant decrease in the number of reads was evidenced in these damaged templates in comparison to the unmodified sample FM. This reduction was 35.2 and 9.4 folds lower compared to the control FM, for the STR and the SNP panels, respectively, and 10.2 and 2.9 folds lower than the 2800 M 0.1 ng dilution, for the STR and the SNP panels, respectively. Finally, the number of markers which produced reads in both replicates was dramatically reduced.

As expected, the mean DoC/locus decreased severely, especially for the STR loci, in relation to the increase in degradation of the templates. The most likely explanation is that the target enrichment procedure (performed by PCR) was inhibited by the molecular damage of the samples [14, 24, 42, 43]. This general trend is clearly described in Supporting Information Figs. S6 and S7 where the number of reads obtained for the three degraded FM samples and the unmodified FM control were reported and compared for each marker of the STR and SNP panels. An interesting finding was that NGS produced a higher number of reads in sample FM-6 than in control FM for three STR (namely, *DYS385A-B, DYS505* and *TPOX*) and ten SNP markers (*rs8037429, rs1413212, rs2269355, rs2040411, rs1493232, rs576261, rs1058083, rs1028528, rs221956* and *rs1490413*) out of the total number of 155 markers. A possible explanation for this result can be the DNA amount of sample FM-6 employed for the tests (1.68 ng as assessed by the SA). Nevertheless, other different mechanisms could be involved. Even in sample FM-10 (where 0.055 ng of DNA was used), in fact, markers *rs2269355* and *rs576261* gave higher DoCs than in control sample FM.

A possible explanation for this result is that the primer annealing sequences for the amplification of these loci could be less damaged than those of the other markers, being poorly represented in purinic compounds. For this reason, an unknown high number of intact annealing sequences, preferentially amplified in a general situation of damaged sequences, could be the explanation for the increase of the DoC in these markers. Even the nucleotide position of the purinic lesion within the primer sequence could affect the PCR efficiency. However, more studies are necessary to elucidate this finding.

As reported in Tables 2 and 3, the analysis of STR and SNP data showed a high percentage of loci which gave none or less than 50 reads in the three trial samples (up to 75.7–76.1% in sample FM-10) and a positive trend between the extent of DNA degradation and the % of low-covered loci. Although this setting of the cut-off value is arbitrary, and leads to a loss of genetic information of up to 5.7% (when compared to the cut-off of 30 x) [32], we believe that a conservative approach

focused on the reliability of the results is what is needed in forensic genetics when validation studies are underway to better understand the potential of an innovative technique.

Last, the percentage of low-covered loci was higher in the STR panels that in the SNP ones, for samples FM-6 and FM-8 (43.6% versus 25.7%, in FM-6 and 63.4% versus 49.4%, in FM-8).

The allelic imbalance varied from 0.277 to 0.440 and from 0.166 to 0.326, for STR and SNP markers, respectively. In 70.3% of the STR markers, the allele with higher molecular weight was less amplified.

Allelic reads below the 30 reads interpretation threshold (IT) were scored for loci with a DoC between 50 and 650 reads. In detail, three STR and seventeen SNP markers showed allelic reads below the IT in the trial DNA samples.

Allelic drop outs were seen for both STR and SNP markers. As expected, the frequency of these artifacts increased proportionally to the extent of degradation in the trial samples, reaching the highest frequency in the most degraded sample FM-10 (up to 44.4% and 33.3% for STRs and SNPs, respectively). In addition, the longer allele of the heterozygous genotype dropped out in 60.8% of the cases, for STR markers. In Fig. 1, the number of reads of the surviving alleles for the dropped out genotypes are shown, subdivided in four intervals. Interestingly, allelic drop out occurred even for STR markers characterized by a DoC ≥ 601x, reaching the considerable coverage of 1445, for allele 7 of the TH01 locus, in sample FM-8. For SNP markers this misleading result was scored only for loci with DoCs ≤ 244x. These results highlight that a potentially misleading amplification artifact such as an allelic dropout could be a crucial concern even for NGS technology especially when the analyses are performed on highly degraded samples.

Stutter products above the defined setting parameters do not constitute an important artifact, as they were found only in three cases. However, this spurious PCR product could originate potentially misleading results. In fact, the NGS results of

Figure 1. Number of surviving alleles recorded in the trial samples for the STR and SNP panels. X-axis: number of reads subdivided in four categories; Y-axis: number of observations.

locus D17S10301 for sample FM-8 showed a stutter product (allele 11) whose number of reads (326) was even higher than the one of the expected real genotype 12/12 (75 reads). Again, the NGS results of locus TH01 (expected real genotype 7/9.3) for sample FM-10 showed instead allele 6 with a 35x coverage (and then above the interpretation threshold), the surviving allele 7 with a 265x coverage and allele 9.3 dropped out.

A single PCR artifact identifiable as "drop in" was observed in the damaged trial samples. This occurred at locus D22S1045 of sample FM-6 which even showed a genotype characterized by alleles 16 (35x) and 17 (57x) instead of the expected true genotype 11/12.

In conclusion, these results clearly point out the need for replicates, at least in two different experiments carried out under the same analytical conditions, to prevent the risk of mistyping, especially for poorly covered loci.

3.2.4 Dilutions of sample FM-6

The NGS analyses of two dilutions of 5.0 μL each of sample FM-6, namely FM-6 (1:1) and FM-6 (1:3), are reported in Supporting Information Tables S2 and S3 (for STR and SNP markers, respectively). A trivial result was that the total mean number of reads/sample and the mean DoC/locus were always lower than the ones described for sample FM-6. The frequency of loci with DoC < 50x increased dramatically reaching the values of 0.587 and 0.594 for STRs and SNPs, respectively, for the FM-6 dilution 1:3.

Several allelic reads below the IT were scored. This last finding was observed for loci with a low depth of coverage (mean 99x ± 38x for STRs and 65x ± 19x for SNPs). In two cases, spurious reads substituted the dropped out allele of a heterozygous locus (D2S441 and D3S1358). Other spurious reads were sporadically generated for both STR and SNP markers. Allelic drop outs were observed for both STR and SNP markers. The DoC of the surviving allele was ≤234x and ≤74x, for STR and SNP markers, respectively.

An interesting amplification artifact was also found at locus D7S820 (expected genotype 8/8) of dilution FM-6 (1:1). In detail, the NGS approach displayed 93 reads of the expected (GATA)$_8$ repeat sequence and 38 reads of a mutated (AATA)(GATA)$_7$ one. The molecular origin of this repeat sequence variation is only presumptive but it is well known that the presence of A-P sites can promote preferential and unspecific incorporation of Adenine [43]. In addition, another possible explanation for this finding is that deamination of the Cytosine to Uracil can lead to G→A transitions [44] as the one observed for the first base of the tetrameric repeat.

In conclusion, even the NGS data results obtained from two dilutions of the less damaged trial sample FM-6 are consistent in identifying the risk of mistyping when the molecular data of poorly covered loci are used for personal identification purposes, in particular if a single NGS experiment is performed.
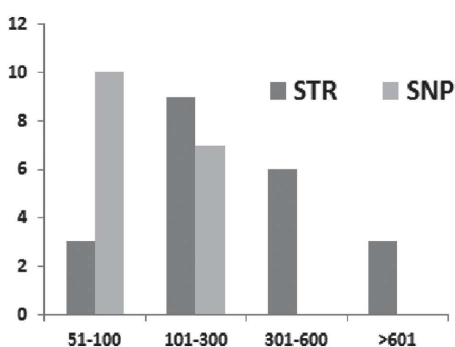
3.2.5 Genetic typing of the samples

The ability of the NGS approach in recovering reliable genetic results from artificially degraded samples was checked by comparing the genetic typing obtained for samples FM-6, FM-8, FM-10 with the corresponding reference DNA FM. Fifty-eight STRs and ninety-three SNPs were the bulk set of markers for the comparisons.

The interpretation guidelines for genotyping the NGS data obtained from the two replicate analyses of each sample are based on the "*consensus*" approach [24, 33] and have been already described in paragraph 2.3.4. The concordance of results among the replicates of the degraded samples and the final comparison with the expected genotypes of reference sample FM allowed us to assign each STR and SNP genotype to one of the following four categories: correct (C), no result (NR), unreliable (U) and wrong (W) result. Figures 2A and 2B show the distribution of the frequencies of each category, for STR and SNP markers respectively, obtained from the analysis of the three trial samples.

The frequency of correctly typed loci was inversely related (r^2 from 0.971 to 0.999) to the degree of chemical damage of the templates (see Supporting Information Fig. S8). As expected, the less damaged sample FM-6 showed the highest frequency of correctly typed loci (44.8 and 58.7% for the STR

Figure 2. 2A and 2B. Results of the genetic typing for trial samples, according to the interpretation guidelines described in paragraph 2.3.4 (C: correct typing; NR: no result; U: unreliable typing; W: wrong typing). A: STR panel; B: SNP panel. X-axis: samples FM-6, FM-8 and FM-10; Y-axis: frequency of each genotyping category.

and the SNP panels, respectively) while the most degraded sample FM-10 revealed the lowest number of markers successfully typed (5.1 and 10.9%, for the STR and the SNP panels, respectively). All these data together indicate that the SNPs panel is two times more efficient than the STR panel in providing correct genotypes, at least for the extremely degraded sample FM-10.
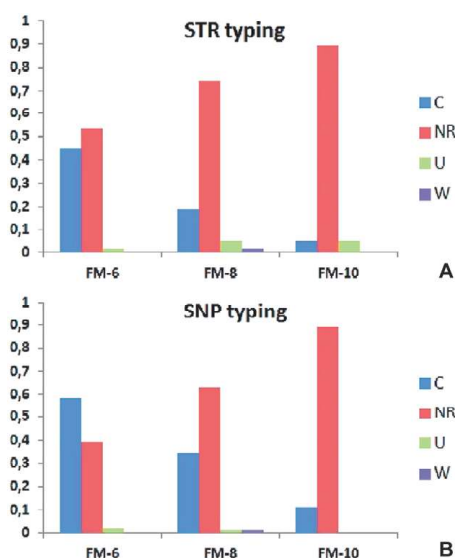
In Figs. 3 and 4 are represented the STR and SNP markers which showed correct results for the degraded samples FM-6, FM-8, FM-10, and their relative amplicon sizes. No correlation between the amplified markers and the corresponding amplicon size was found, at least for STR loci. Markers DYS505 and DYS385A-B were correctly typed for the most degraded sample FM-10, in fact, even though they are characterized by quite different amplicon sizes (174 and 335 bp, respectively). Differently, for the most degraded sample FM-10, the correctly typed SNPs seem to cluster on the left side of Fig. 4, characterized by loci with short amplicon sizes.

The STR and SNP markers correctly typed in sample FM-10 also gave the expected correct genotypes in the less damaged samples FM-6 and FM-8, as shown in Figs. 3 and 4.
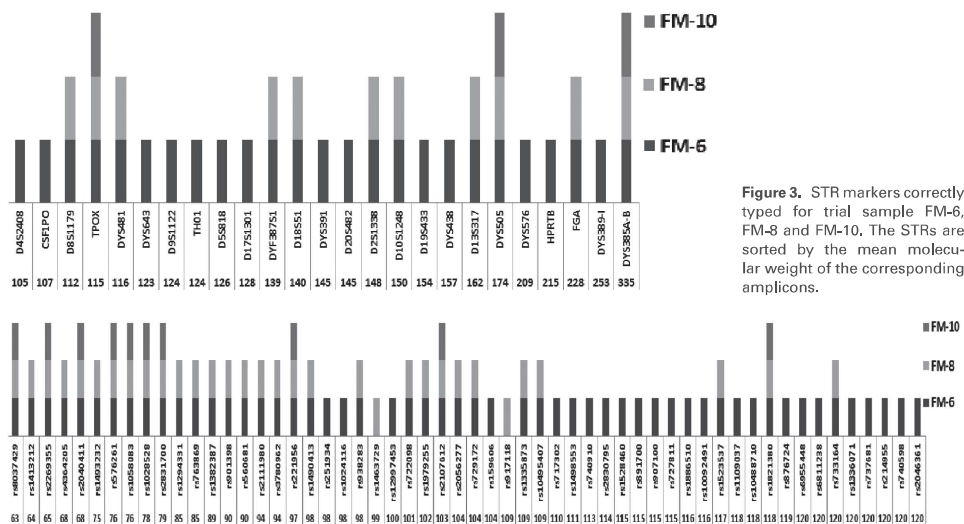
A positive trend between the number of loci which provided "no result" and the extent of degradation of the samples was also observed. The frequency of the loci which gave "unreliable results", that is different typing results in the two replicates, was higher for the STR panel, reaching the value of 5.2% for sample FM-10 (for the SNP panel no more than 2.2% of the markers gave this outcome) (see Fig. 2A and B).

According to the conservative criteria described in interpretation guidelines for genotyping, only two loci gave incorrect ("wrong") results in both replicate tests, when compared to the expected genotypes, thus showing that a cut-off value of 50x eliminated genotyping errors in samples FM-6 and FM-10. The STR locus TH01 typing gave in both replicates of sample FM-8 only the reads of allele 7 (1,445x in the first test and 917x in the second one) instead of the expected genotype 7/9.3. The SNPs typing of locus rs722290 (expected genotype C/G) gave in both replicates of sample FM-8 only the reads of allele G (64x in the first test and 244x in the second one). This point can be extremely relevant when forensic casework will be routinely approached with the NGS technology.

The genetic typing of the two dilutions of sample FM-6, namely FM-6 (1:1) and FM-6 (1:3), showed that an increased dilution correlates negatively to the number of correctly typed loci. In addition, it was observed that the most diluted sample FM-6 (1:3) provided the highest frequencies of genotypes assigned to the category of "*no result*", according to the criteria described in paragraph 2.3.4 (see Supporting Information Fig. S9A and B). No incorrect genotype was scored while the frequency of "unreliable results" was ≤0.017. In Supporting Information Figs. S10 and S11 are represented the STR and SNP markers which showed correct results for the diluted samples FM-6 (1:1) and FM-6 (1:3). As already observed for the degraded samples FM-6, FM-8, FM-10, only a limited number of markers gave a good coverage thus allowing a correct typing in the diluted samples.

Figure 3. STR markers correctly typed for trial sample FM-6, FM-8 and FM-10. The STRs are sorted by the mean molecular weight of the corresponding amplicons.

Figure 4. SNP markers correctly typed for trial sample FM-6, FM-8 and FM-10. The SNPs are sorted by the molecular weight of the corresponding amplicons.

3.3 STR Analysis by CE

The same amounts of DNA utilized for the NGS analysis of degraded samples FM-6, FM-8 and FM-10, were used also for the conventional STR typing approach amplifying the 16 STR loci plus Amelogenin contained in the commercial kit PowerPlex® ESI 17 system, in duplicate tests. The results showed all the peculiar features of degraded/modified DNA samples, characterized by partial profiles with loss of high-molecular-weight markers and the presence of amplification artifacts such as allelic imbalances and drop outs (data not shown). The STR typing of sample FM-6 allowed the correct genetic typing of 8 out of 16 STRs (50%) plus Amelogenin, while the number of correctly scored STRs was 5 out of 16 (31.5%) plus Amelogenin, for sample FM-8. The conventional PCR-CE analysis provided reliable results from three low molecular weight loci (D3S1358, D16S539 and vWA) which gave no result by employing NGS-based tests. More interestingly, PCR-CE analysis gave the correct genotype (7/9.3) at locus sample TH01 of sample FM-8 while NGS provided the incorrect 7/7 genotype (see Supporting Information Table S4). No marker could be typed for sample FM-10. No amplicon was found in the negative controls (NTC and NCSP).

4   Concluding remarks

The results reported in this study are the first data gathered on the reliability of the NGS analysis using the ForenSeq™

DNA Sequencing kit on degraded samples. At present, only a single study [31] approached challenging DNA samples but this study reported the results focusing mainly on the concordance of the data with the conventional PCR/CE methods. In the present paper, a DNA sample (namely sample FM) was artificially enriched in A-P sites, thus producing three differently degraded trial DNA samples (FM-6, FM-8 and FM-10). Even if the limit of this assay is that the DNA samples were enriched in specific lesions [21, 24, 35, 36], despite the large number of different DNA damages which can be found in real case-work samples [13], we believe that this approach can confidently represent a model to study the performance and the limits of this new NGS technology in forensic genetics. The results of the three degraded samples were then compared with those of the undegraded sample FM and those of three dilutions (1.0, 0.5 and 0.1 ng) of reference sample 2800 M. Finally, also conventional PCR/CE analysis was performed on the three trial samples as well as on the undegraded control.

In order to increase the number of reads/sample, the libraries of a reduced number of samples (twenty) were pooled together in this study, a strategy suggested to maximize the depth of coverage when challenging samples are considered [23,37]. However, despite this approach, the number of reads obtained from the three trial DNA samples was 2.9–10.2 folds lower than the ones reported for the less concentrated control DNA 2800 M dilution (0.1 ng). This observation is consistent with findings already reported for other difficult samples [31]. In addition, it was impossible to assign up to 78.2% of the genotypes in the trial samples,

as the corresponding loci were flagged by the software as "low-covered" markers (< 50x). Both the decrease of the number of reads/sample and the increase of the percentage of low-covered loci were proportional to the extent of the degradation in each trial DNA sample. Other analytical parameters, such as the number of PCR cycles of the target enrichment procedure and library preparation should be evaluated in a future prospective in order to optimize the NGS approach to highly degraded samples [31]. In addition, since our data highlighted higher amplification efficiency for a specific set of STR and SNP markers in the trial samples, it could be interesting to investigate the amplification efficiency of these markers in other difficult DNA samples, potentially driving toward the selection of the best set of markers for forensic applications.

The ForenSeq™ DNA Sequencing kit run on the Illumina MiSeq was able to generate data which led to the correct typing of 44.8%, 18.9%, 5.1% of the STRs and 58.7%, 34.7%, 10.9% of the SNPs of trial samples FM-6, FM-8 and FM-10, respectively. These data support the common sense conclusion that the ability to recover reliable genotypes is inversely proportional to the extent of degradation of the samples. However, the SNP markers showed in all trial samples, that they had higher chances to be typed correctly compared to the STRs.

This NGS approach seems to offer very promising performances on heavily degraded DNA samples. In fact, while the conventional approach (PCR/CE) on the most degraded trial sample (sample FM-10) gave no results, three STRs (one autosomal and two Y-chromosome STRs) and ten autosomal SNPs gave reportable genotypes with this NGS assay (see Figs. 3 and 4). The 1/random match probability value (that is the LR) calculated for the autosomal loci [45] gave a result of 1.7×10^5, which can be considered an excellent achievement, for such a highly damaged DNA sample.

The NGS assay showed very reliable results. The frequency of the genetic mistyping was very low, reaching the value of 1.4% only for sample FM-8. It is interesting to note that, in two cases of erroneous typing, the same allele dropped out in the two replicates while the surviving alleles showed a considerable number of reads, far above the interpretation threshold value (1,445x and 917x for allele 7 of TH01 locus; 64x and 244x for allele G of rs722090 locus). This finding points out that even for NGS the risk of mistyping (that is, in these two cases, to type as homozygous a genotype instead of the correct heterozygous one) could not be eliminated by duplicate tests.

In addition, our results point out that validation studies and the definition of standard operating procedures for the interpretation of the NGS DNA typing results are needed in order to evaluate carefully the application of this technology in forensic human identification even if very promising results have been already achieved [12].

5 References

[1] Welch, L., Gill, P., Phillips, C., Ansell, R., Morling, N., Parson, W., Palo, J.U., Bastisch, I., *Forensic Sci. Int. Genet.* 2012, *6*, 819–826.

[2] Hares, D. R., *Forensic Sci. Int. Genet.* 2012, *6*, e52–e54.

[3] Roewer, L., *Investig. Genet.* 2013, *4*, 22–30.

[4] Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y. J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Irzyk, G. P., Jando, S. C., Alenquer, M. L., Jarvie, T. P., Jirage, K. B., Kim, J. B., Knight, J. R., Lanza, J. R., Leamon, J. H., Lefkowitz, S. M., Lei, M., Li, J., Lohman, K. L., Lu, H., Makhijani, V. B., McDade, K. E., McKenna, M. P., Myers, E. W., Nickerson, E., Nobile, J. R., Plant, R., Puc, B. P., Ronan, M. T., Roth, G. T., Sarkis, G. J., Simons, J. F., Simpson, J. W., Srinivasan, M., Tartaro, K. R., Tomasz, A., Vogt, K. A., Volkmer, G. A., Wang, S. H., Wang, Y., Weiner, M. P., Yu, P., Begley, R. F., Rothberg, J. M., *Nature* 2005, *437*, 376–380.

[5] Metzker, M. L., *Nat. Rev. Genet.* 2010, *11*, 31–46.

[6] Korlach, J., Bjornson, K. P., Chaudhuri, B. P., Cicero, R. L., Flusberg, B. A., Gray, J. J., Holden, D., Saxena, R., Wegener, J., Turner, S. W., *Methods Enzymol.* 2010, *472*, 431–455.

[7] Lam, H. Y., Clark, M. J., Chen, R., Natsoulis, G., O'Huallachain, M., Dewey, F. E., Habegger, L., Ashley, E. A., Gerstein, M. B., Butte, A. J., Ji, H. P., Snyder, M., *Nat. Biotechnol.* 2012, *30*, 78–82.

[8] Haque, F., Li, J., Wu, H. C., Liang, H. J., Guo, P., *Nano Today* 2013, *8*, 56–74.

[9] Capelli, C., Tschentscher, F., Pascali, V. L., *Forensic Sci. Int.* 2003, *131*, 59–64.

[10] Orlando, L., Gilbert, M. T., Willerslev, E., *Nat. Rev. Genet.* 2015, *16*, 395–408.

[11] Schaefer, N. K., Shapiro, B., Green, R. E., *Mol. Ecol.* 2016, *25*, 2398–2412.

[12] Børsting, C., Morling, N., *Forensic Sci. Int. Genet.* 2015, *18*, 78–89.

[13] Alaeddini, R., Walsh, S. J., Abbas, A., *Forensic Sci. Int. Genet.* 2010, *4*, 148–157.

[14] Eckert, K. A., Kunkel, T. A., *PCR Methods Appl.* 1991, *1*, 17–24.

[15] Forster, L., Thomson, J., Kutranov, S., *Forensic Sci. Int. Genet.* 2008, *2*, 318–328.

[16] Butler, J. M., Shen, Y., McCord, B. R., *J. Forensic Sci.* 2003, *48*, 1054–1064.

[17] Coble, M. D., Butler, J. M., *J. Forensic Sci.* 2005, *50*, 43–53.

[18] Smith, P. J., Ballantyne, J., *J. Forensic Sci.* 2007, *52*, 820–829.

[19] Diegoli, T. M., Farr, M., Cromartie, C., Coble, M. D., Bille, T. W., *Forensic Sci. Int. Genet.* 2012, *6*, 498–503.

[20] Gettings, K. B., Kiesler, K. M., Vallone, P. M., *Forensic Sci. Int. Genet.* 2015, *19*, 1–9.

[21] Lindahl, T., Nyberg, B., *Biochemistry.* 1972, *11*, 3610–3618.

[22] Pääbo, S., Poinar, H., Serre, D., Jeenicke-Despres, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L., Hofreiter, M., *Ann. Rev. Genet.* 2004, *38*, 645–679.

[23] Illumina, ForenSeq™ DNA Signature Prep Guide, Part # 15049528 Rev. D, February 2015.

[24] Fattorini, P., Previderè, C., Sorçaburu Cigliero, S., Marrubini, G., Alù, M., Barbaro, A. M., Carnevali, E., Carracedo, A., Casarino, L., Consoloni, L., Corato, S., Domenici, R., Fabbri, M., Giardina, E., Grignani, P., Baldassarra, S. L., Moratti, M., Nicolin, V., Pelotti, S., Piccinini, A., Pitacco, P., Plizza, L., Resta, N., Ricci, U., Robino, C., Salvaderi, L., Scarnicci, F., Schneider, P. M., Seidita, G., Trizzino, L., Turchi, C., Turrina, S., Vatta, P., Vecchiotti, C., Verzeletti, A., De Stefano, F., *Electrophoresis.* 2014, *35*, 3134–3144.

[25] Illumina, CASAVA v1.8.2 User Guide, December 2011.

[26] Illumina, ForenSeq Universal Analysis Software Guide, Part # 15053876 Rev. B, February 2015.

[27] Ajay, S. S., Parker, S. C., Abaan, H. O., Fajardo, K. V., Margulies, E. H., *Genome Res.* 2011, *21*, 1498–1505.

[28] Eduardoff, M., Santos, C., de la Puerte, M., Gross, T. E., Fondevila M., Strobl, C., Sobrino, B., Ballard, D., Schneider, P. M., Carracedo, Á., Lareu, M. V., Parson, W., Phillips, C., *Forensic Sci. Int. Genet.* 2015, *17*, 110–121.

[29] Gettings, K. B., Kiesler, K. M., Faith, S. A., Montano, E., Baker C. H., Young, B. A., Guerrieri, R. A., Vallone, P. M., *Forensic Sci. Int. Genet.* 2016, *21*, 15–21.

[30] Goodwing, S., McPeron, J. D., McCombie, W. R., *Nature Rev. Genet.* 2016, *17*, 333–351

[31] Churchill, J. D., Schmedes, S. E., King, J. L., Budowle, B., *Forensic Sci. Int. Genet.* 2016, *20*, 20–29.

[32] Carboni, I., Fattorini, P., Previderè, C., Sorçaburu Ciglieri, S., Iozzi, S., Nutini, A. L., Contini, E., Pescucci, C., Torricelli, F., Ricci, U., *Forensic Sci. Int. Genet.* 2015, *Suppl. Series 5*, e83–e85.

[33] Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, S. Escaravage, N., Waits, L. P., Bouvet, J., *Nucleic Acids Res.* 1996, *24*, 3189–3194.

[34] Warshauer, D. H., King, J. L., Budowle, B., *Forensic Sci. Int. Genet.* 2015, *14*, 182–186.

[35] Lindahl, T., Andersson, A., *Biochemistry* 1972, *11*, 3618–3623.

[36] Lindahl, T., Nyberg, B., *Biochemistry* 1974, *13*, 3405–3410.

[37] Sims, D., Sudbery, I., Ilott, N. E., Heger, A., Ponting, A. P., *Nature Rev. Genet.* 2014, *15*, 121–132.

[38] Sommer, S., Courtiol, A., Mazzoni, C. J., *BMC Genomics* 2013, *14*, 542–552.

[39] Rockenbauer, E., Hansen, S., Mikkelsen, M., Børsting, C., Morling, N., *Forensic Sci. Int. Genet.* 2014, *8*, 68–72.

[40] Fordyce, S. L., Mogensen, H. S., Børsting, C., Lagace, R. E., Chang, C. W., Rajagopalan, N., Morling, N., *Forensic Sci. Int. Genet.* 2015, *14*, 132–140.

[41] Ballantyne, K. N., Goedbloed, M., Fang, R., Shaap, O., Lao, O. Wollstein, A., Choi, Y., van Duijn, K., Vermeulen, M., Brauer, S., Decorte, R., Poetsch, M., von Wurmb-Schwark, N., de Knijff, P., Labuda, D., Vézina, H., Knoblauch, H., Lessig, R., Roewer, L., Ploski, R., Dobosz, T., Henke, L., Henke, J., Furtado, M. R., Kayser, M., *Am. J. Hum. Genet.* 2010, *87*, 341–353.

[42] Brisco, M. J., Latham, S., Bartley, P. A., Morley, A. A., *BioTechniques* 2010, *49*, 893–897.

[43] Heyn, P., Stenzel, U., Briggs, A. W., Kircher, M., Hofreiter, M., Meyer, M., *Nucleic Acids Res.* 2010, *38*, e161.

[44] Hofreiter, M., Jaenicke, V., Serre, D., von Haeseler, A., Pääbo, S., *Nucleic Acids Res.* 2001, *29*, 4793–4799.

[45] The ALlele FREquency Database, https://alfred.med.yale.edu.

Acknowledgements