



# Biological Linkage of Ancient and Modern Populations of Northwestern Pakistan through their Mitochondrial Profiles

Nasir Ali<sup>1,2,6</sup>, Muhammad Ilyas<sup>2,3</sup>, Nazia Akbar<sup>1,2\*</sup>, Gohar Rahman<sup>2</sup>, Shakirullah Khan<sup>4</sup>, Mujaddad Ur Rehman<sup>6</sup>, Abdul Samad<sup>5</sup>, Habib Ahmad<sup>1,2</sup> and Bibi Nazia Murtaza<sup>7\*</sup>

<sup>1</sup>Department of Genetics, Hazara University, Mansehra, Pakistan

<sup>2</sup>Centre for Human Genetics, Hazara University, Mansehra, Pakistan

<sup>3</sup>Centre for Omic Sciences, Islamia College University, Peshawar, Pakistan

<sup>4</sup>Department of Archeology, Hazara University, Mansehra, Pakistan

<sup>5</sup>Directorate of Archaeology and Museums Khyber Pakhtunkhwa, Peshawar, Pakistan

<sup>6</sup>Department of Microbiology, Abbottabad University of Science and Technology, Abbottabad, Pakistan

<sup>7</sup>Department of Zoology, Abbottabad University of Science and Technology, Abbottabad, Pakistan

## ABSTRACT

Molecular anthropology is a new feature of archeology that can provide additional useful information to illustrate our past. Ancient DNA basically had small pieces of DNA that has been around for a long time. This DNA can be used as a sample to study different parts of a population that lived a long time ago. In order to study the continuity in maternal lineage between ancient and modern people of the north-western province, polymorphisms in the HVSI region of mitochondrial DNA sequences of 56 ancient samples (1000-3000 old) and 303 modern samples from three different regions (Chitral, Swat and Hazara) were analyzed for hyper variable sequences of mitochondrial DNA for the determination of their maternal affinities. Our results revealed 25% unique haplotypes in ancient samples whereas 28% unique haplotypes have been observed in the modern samples. The most frequent haplotype found was H2a (23% in ancient and 10% in modern samples). Our results show that with the exception of ancient Dir, there are genetic similarities among the ancient lines between Hazara, Chitral and Swat. It revealed that most of the ancient and modern samples belong to the H subclades. Current studies have revealed a significant genetic link between ancient and modern peoples of northwestern Pakistan. Our findings regarding maternal lineages of the people of Khyber Pakhtunkhwa, presumed in mtDNA, has their origin from the West Eurasian and South Asian stocks. These results we have offer a reference to better understand the biological linkage of ancient and modern inhabitants living in Khyber Pakhtunkhwa, Pakistan.

## Article Information

Received 18 May 2022

Revised 26 June 2022

Accepted 05 August 2022

Available online 30 August 2022

(early access)

Published 17 September 2023

## Authors' Contribution

MI and HA designed the experiments and supervised the research. NA and GR collected the bone specimens from archeological sites. NA, GR, BNM and MI performed bioinformatics analyses. SK and AS participated in the analyses and provided consultation. NA, MI, NA and AS wrote the paper. All authors have read and approved the manuscript.

## Key words

Ancient DNA, Pakistan, Pakhtuns, Gandhara, Genetics

## INTRODUCTION

Pakistan is a country with many different cultures and traditions. It has seen many systems in its history (one of the oldest). All the corpses, which live in this land and

contribute to its culture, are spread all over the country. Pakistan has many archaeological sites such as Taxila, Mahenjodaro, Gandhara, Harappa and Mehargar. These places are called the archaeological sites of Pakistan as well as the ancient learning centers of Pakistan dating back from Mughal Empire to Paleolithic period (Ahmed, 2014). The most famous archeological site is the Soanian culture which originated in the Soan valley near modern Islamabad. The Son Valley culture is considered to be the most popular Palaeolithic culture in Central Asia (Vishnyatsky, 1999). The most important Neolithic sites is Mehrgarh of Balochistan that dates back to 7000BC and 2000BC (Jarige, 2008). Mohenjo-daro and Harappa sites of Indus valley civilization that are the most famous and best-known sites (c2500 - 1900BCE) (Posehall, 2002).

The knowledge of molecular anthropology gives an

\* Corresponding author: nazia.murtaza@gmail.com, naziasaadiq@gmail.com 0030-9923/2023/0005-2451 \$ 9.00/0



Copyright 2023 by the authors. Licensee Zoological Society of Pakistan.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

anthropologist and archaeologist a new perspective to better understand what happened in the past and explain it with the possibilities for the future (Iwamura *et al.*, 2004). Over the past few decades, geneticists and molecular geneticists have examined DNA mutations in citizens from this region to document genetic linkages and history. This assertion is slowly evolving from the study of human evolution to advanced DNA testing to address issues of population lamination and genetic vulnerability or disease resistance and genome-wide Association research. It has already been reported that mtDNA research of this region has a European maternal lineage (Akbar *et al.*, 2016). Multiple mtDNA genotypes exist in humans (Stewart *et al.*, 2021). We intend to review the mtDNA gene expression of human bones from archaeological sites and museums and compare it with modern ones of the same region. It is concluded that there is a significant genetic link between the ancient and modern peoples of northwestern Pakistan. In this study genealogical representation of ancient and modern human populations are identified that is there is no genetic differentiation between the studied area except for Dir but targeting only mtDNA HV1 and HVII regions.

## MATERIALS AND METHODS

### *Ethical approval and permissions*

Ethical support was obtained from Hazara University Bioethical Institutional Committee, prior to the start of the study. Proper permission was obtained from the Directorate of Archeology and Museums of Khyber Pakhtunkhwa for collecting and using ancient bone specimens for research purposes. Furthermore, a special consent form was signed from all the living participants who provided their biological material in the form of buccal swabs. The participants were ensured about the privacy of the data.

### *Sample collection*

A trained team consists of archeologists, geneticists and curator was sent to target areas (Chitral, Swat and Hazara) for collecting ancient bone specimens (Fig. 1A). In the presence of archeologists and site excavation experts from Department of Archeology, Hazara University, Mansehra ancient human bone samples were collected from the sites and skeletons from museums that were preserved previously (Fig. 1B). All the archeological sites have already been excavated under the supervision of Directorate of Archeology and Museums Khyber Pakhtunkhwa. Our team of experts visited the sites and collected the bones under a comprehensive precautionary procedure to avoid contamination by using standard operating procedures for bone successful recovery. A standard method was adopted to dig and excavation of

the graves. The human bones and teeth, human skeletal remains and other remains found were collected and stored in properly labeled sterile bags. The samples were saved in the form of photos, documents in written form and annotated before further use.

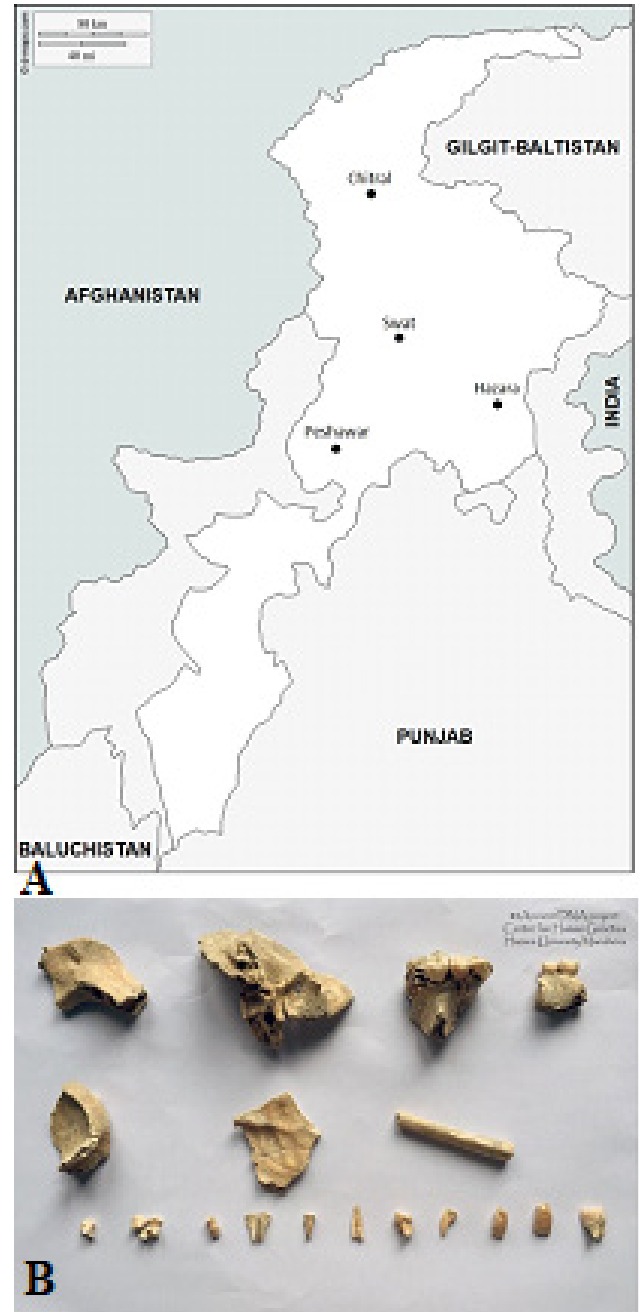


Fig. 1. The map of the Northern and western province of Pakistan named as Khyber Pakhtunkhwa with the regions shown from where the ancient human specimens (A) and modern samples (B) were collected.

Buccal swabs were taken from the modern individuals currently residing, in and surrounding areas of Swat, Hazara and Chitral, from where ancient human bone specimens were collected. The Swab sample were given a unique code.

#### *DNA Isolation from bone specimens*

The 2 mm outer surface of each of the three specimen was first wrinkled using a scalpel. Each side of the bone was UV irradiated for 10 min. The bone piece was powdered using UV sterilized mil crusher (mortar and pestle) in the presence of liquid nitrogen and transferred into a UV-free DNA bag. Twenty-five milligram bone powder was measured from each of the three samples and placed in a 1.5ml Eppendorf tube for DNA extraction. The procedure was performed in a clean room with a complete set of body suits worn solely for bone crushing. For extraction of ancient DNA from different bone pieces and teeth standard protocols were used as described in Raghavan *et al.* (2014) with slight modifications. Modern genomic DNA was extracted by following the protocol recommended by Akbar *et al.* (2015).

#### *mtDNA HVSI region analysis*

The ancient samples collected from Khyber Pakhtunkhwa were amplified and analyzed using three pairs of primers targeting mtDNA HVSI regions (Table I).

**Table I. List of primers used to amplify the mtDNA region of HVSI.**

Primers name	Primer sequence
IFb-16128	5' GGTACCATAAATACTTGACCACCT 3'
IR-16348	5' GACTGTAATGTGCTATGTACGGTAAA 3'
IF-16226	5' CTGTTCTTTTCATGGGGAAGC 3'
IF-16445	5' ATTGATTTACGGAGGATGG 3'
Hv1 59-50	5' CTCCACCATTAGCACCCAAAGCTAAG 3'
Hv1 59-50	5' GATATTGATTTACGGAGGATGGTGGTC 3'

For PCR amplification 2-3ng of ancient DNA with 25µl final volume of reaction mixture was processed using DreamTaq® Green ready-mix. The thermal cycler of final reaction mixture comprised the initial denaturation for 10min at 95°C following 40 cycles of denaturation at 94°C for 30 sec primers attachment at 60°C for 40 sec and cyclic elongation at 72°C for 1 min. The final elongation was set at 72°C for 5min and hold at 10∞. The same procedure was followed for modern DNA testing.

#### *Sanger sequencing and data analysis*

The PCR products were shipped to Macrogen, Korea

for nucleotide sequencing using ABI 3735 analyzer. The raw data received was processed using different publicly available softwares. The sequences obtained in FASTA format were BLASTed (The Basic Local Alignment Search Tool) to find regions of local similarity between sequences and to confirm that it is actual human mtDNA. Following the confirmation, the FASTA MitoMaster system (<https://www.mitomap.org>) was run to identify different variants in DNA sequences and biological implications of each variant. The haplogroup configuration was created using Haplogrep (<https://haplogrep.i-med.ac.at>) and was confirmed by MyTree. The genetic length of the genes between observant people was analyzed using the Arlequin app. The phylogenetic tree was developed using Molecular Evolutionary Genetics Analysis (MEGA 7).

## RESULTS

A total of 56 human bone including tooth types were collected from Chitral (n=3), Swat (n=43) and Hazara (n=10) of Khyber Pakhtunkhwa province of Pakistan. The artifacts collected from the same regions were used to estimate the possible date and age of the specimens (Lovejoy *et al.*, 1985; Hanihara and Suzuki, 1978).

#### *Mitochondrial haplotypes in ancient samples*

Table II shows the haplotypes and their frequencies for ancient and modern samples. The most frequent haplotype found in ancient samples was H2a which was about 23.2% in our samples, followed by H1a (10.7%). Total 25% of the samples were having unique haplotypes in which M30d was reported only in a Chitrali sample and is of purely South Asian origin. About ten of the Swati samples had unique haplotypes (A2z, H1ag1a, H1an2, H1bs, H26a1a1, L0d1c1a2, M7b1a, T1a, T2 and U6b3a) not found in other ancient samples. Similarly, H3a and I2b were only observed in samples collected from Hazara.

#### *Mitochondrial haplotypes in modern samples*

The haplotypes predicted using bioinformatics tools from analysis of 299 buccal swabs samples from Swat, Chitral, Hazara along with a neighboring area Dir. are shown in Table II. The most frequent haplotype identified was H2a (10%) followed by M3 (4.7%). More than 27% of haplotypes in modern samples were unique. The MDS (Multidimension scaling) was performed that shows the similarity and differences of the current day populations with that of ancient samples collected from the same area (Fig. 2) Also, the modern population and the ancient people, Ancient Khwazakhela and ancient Kabal are the nearest tehsils and are closer than others which suggests that the lineage of these individuals are different (Fig. 3).

**Table II. Summary of Haplotypes and their frequencies for ancient and modern samples.**

Region	No of haplotypes	Unique haplotypes	Frequent haplogroup	Frequency	Mega haplogroup	Origin
Ancient Chitral	3	3	M30d	33.3%	M	South Asia
Modern Chitral	22	17	H2a	38%	H	Southwest Asia
Ancient Swat	29	24	H2a	12%	H	Southwest Asia
Modern Swat	16	08	H2a	56%	H	Southwest Asia
Ancient Hazara	05	04	H2a	44%	H	Southwest Asia
Modern Hazara	56	32	U7 and M3	10.7%	U	West Asia
Dir Modern	49	34	J1b	20%	J	West Asia

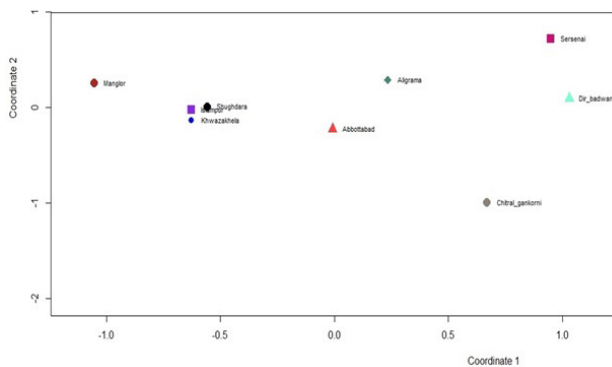


Fig. 2. Multidimensional scaling analysis for the ancient samples to further explain their genetic relationship among the ancients. Aligrama, Dir Badwan, Abbottabad and Sersnai are close to each other, except for Chitral Gankorni.

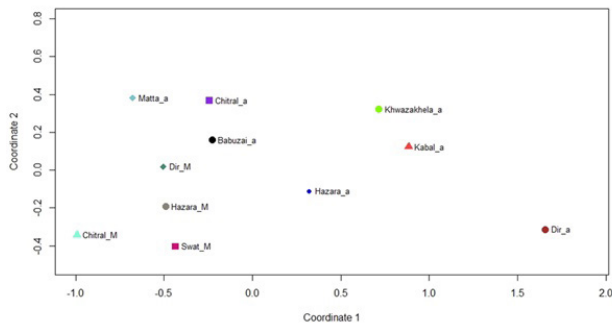


Fig. 3. Comparing the modern population and the ancient people, Ancient Khwazakhela and ancient Kabal are the nearest tehsils and are closer than others which suggests that the lineage of these individuals are different. Ancient Matta Ancient Babuzai and ancient Chitral are very close to each other in which Modern Hazara, Modern Swat and Modern Dir populations are concentrated. The current population of Hazara is the largest example of population migration from the Swat region where it represents a strong connection with the existing population of Swat and Dir.

The current population of Hazara is the largest example of population migration from the Swat region where it represents a strong connection with the existing population of Swat and Dir. The phylogenetic tree was constructed to check the evolutionary relationship of ancient populations with the modern one (Fig. 4).

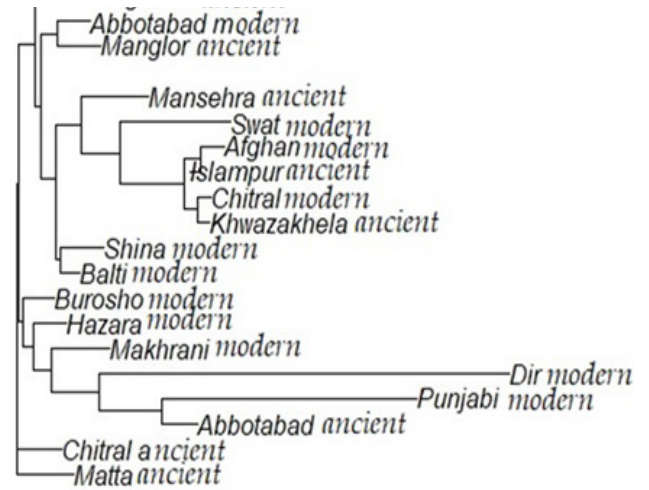


Fig. 4. The phylogenetic tree was constructed using MEGA7 that shows the evolutionary relationship of populations in current study to those reported previously from other regions of Pakistan.

**DISCUSSION**

The history of mankind can be explored with the help of Paleontology, Genetics and Archeology. Not only does this help to understand the history of mankind but it also helps to study contemporary human change. In the twentieth century genetics and paleontology were initiated together by many scientists to study the population assembly of the past, population changes and the relationship of these populations in terms of evolution (Jobling *et al.*, 2019). Mitochondrial DNA haplogroup classification is used

to study maternal lineage of ancient human populations (Kim *et al.*, 2022). In the current study, about three ancient samples and 63 modern ones from Chitral were used in this study. The most common haplotype is H2a with 38% frequency following J1d5 (11.1%) and J1d3a (9.5%). Numerous studies reported that haplogroup H may have originated in West Asia about 25,000 years ago and moved to Europe between 20,000 to 25,000 years ago (Richards *et al.*, 2000; Pereira *et al.*, 2005). The expansion of the subclades H1, H3 and haplogroup V reflects the second intra-European expansion from the Franco-Cantabrian region to the last glacial peak almost 13,000 years ago (Achilli *et al.*, 2004; Richards *et al.*, 2000). H2a, common in Europe, was first expanded in the northeast and south Caucasus that defines the western tip of Eurasia and later the southwestern part of Europe. Significant spread of H2a among the Chitrali people indicates that their maternal gene pool originated in western Eurasia. The addition on a positive scale suggests an admixture between Chitral and European peoples around 990 and 210 BCE as well as related to Alexander's intrusion in subcontinent India in 327-326 BCE (Hellenthal *et al.*, 2014). J1 is a subclade of JT from the mega haplogroup R that is the starting point for the subclade in Western Asia date back to 27000 years ago. The presence of the J1 haplogroup in our analysis is based on Quintana *et al.* (2004) where it is referred to as J1 is found between 9% of the Kalash minority of northwestern Pakistan. In Pakistan, where the West ethnic Eurasian population makes up 50%, the J1 rate is down by about 5%. The ancient Chitrali seem to have carried the same genetic signature from the Neolithic period from time to time, while Swat and Hazara carry large numbers of native maternal lineage such as H2a, Hv1 and Hv2. For example, 66% of Chitrali who were haplotyped in the first study were of haplogroup M.

Swat valley is famous for its Gandhara culture and a number of invaders left their genetic signatures in this part of the world. The current people of Swat are migrated from the surrounding areas like Afghanistan and Central Asia long time ago. A number of archeological sites have been excavated in Swat by the Directorate of Archeology and Museums Khyber Pakhtunkhwa. With the help of trained team we collected and analysed some human bones from those sites and compared its maternal lineage with the residing people of Swat (Table II). It was reported that H2a is the most common haplotype in the ancient Swati samples followed by H32 with the frequency of 11.6% which is a Southwest Asian origin. On the other side, the modern samples had a significant number of individuals with unique haplotype with mostly originated from Europeans clad. These results are consistent with our archaeological evidence that people had high mobility

within the area in ancient and historical times. Haplotype C, found in modern population of Swat, originated in Central Asia about 30000 years ago and is a member of the M haplogroup. The Lineage C was not found here in Swat ancient samples. This would mean that Swat may have regional roots in ancient Neolithic Asia.

About 102 modern samples and 10 ancient samples from Hazara region were studied. Interestingly, all these ancient samples of Hazara region were found to be from Southwest Asian clad. The most common haplotype observed in modern Hazara is M3 and U7 followed by U7a, H2a and HV2. MtDNA haplogroup U7 belongs to western Eurasia is the ancestor of mega haplogroup R and thought to be originated around 40000 years near Black Sea. In South Asia, the U7 is found in 12% of Gujarat, the western state of India where the whole of India accounts for 2% and 5% in Pakistan. Of these 5%, it is distributed in the Hazara region near Pakistani Gujrat and Punjab. The majority of the haplotype U7 among the Abbottabad people can show the maternal genes flow from different ethnic groups living in the Indus Valley, while the high prevalence of haplotype HV2 among them shows a general impact on South Asia and the influence on their maternal gene pool.

Since this study uses only maternal sequences to confirm genetic relationships, further research on other markers within the genome could reveal different population dynamics during the study period. Increasing the number of people surveyed from these important populations could further unleash genetic links with other populations during this crucial period in European history.

It is concluded that Khyber Pakhtunkhwa's gene pool is thought to have been influenced by the flow of genes from neighbors and migration. The maternal lines in the modern genealogy are the signatures of the numbers populations that have lived in the area for centuries. Recently in the populations of Khyber Pakhtunkhwa, there is a dominating effect of Central Asian and West Eurasia.

## ACKNOWLEDGEMENTS

We are thankful to all the participants, museum curators and field experts for their valuable time. This paper is part of the PhD dissertation of Mr. Nasir Ali.

### *Funding*

Not applicable.

### *Availability of data and materials*

Not applicable.

### *Ethics approval*

Not applicable.

*Consent for publication*

Not applicable.

*Data availability statement*

The generated data can be obtained by contacting the corresponding author.

*Statement of conflict of interest*

The authors have declared no conflict of interest.

## REFERENCES

- Achilli, A., Rengo, C., Magri, C., Battaglia, V., Olivieri, A., Scozzari, R., Cruciani, F., Zeviani, M., Briem, E., Carelli, V., and Moral, P., 2004. The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. *Am. J. Hum. Genet.*, **75**: 910-918. <https://doi.org/10.1086/425590>
- Ahmed, M., 2014. *Ancient Pakistan an archaeological history: Volume III: Harappan Civilization-the Material Culture*. Amazon.
- Akbar, N., Ahmad, H., Nadeem, M.S., Ali, N., Saadiq, M., 2015. An efficient procedure for DNA isolation and profiling of the hyper variable MtDNA sequences. *J. Life Sci.*, **9**: 530-534. <https://doi.org/10.17265/1934-7391/2015.11.004>
- Akbar, N., Ahmad, H., Nadeem, M.S., Hemphill, B.E., Muhammad, K., Ahmad, W., and Ilyas, M., 2016. HVSI polymorphism indicates multiple origins of mtDNA in the Hazarewal population of Northern Pakistan. *Genet. Mol. Res.*, **15**: <https://doi.org/10.4238/gmr.15027167>
- Derenko, M., Malyarchuk, B., Bahmanimehr, A., Denisova, G., Perkova, M., Farjadian, S., and Yepiskoposyan, L., 2013. Complete mitochondrial DNA diversity in Iranians. *PLoS One*, **8**: e80673. <https://doi.org/10.1371/journal.pone.0080673>
- Hanihara, K., and Suzuki, T., 1978. Estimation of age from the pubic symphysis by means of multiple regression analysis. *Am. J. Physiol. Anthropol.*, **48**: 233-239. <https://doi.org/10.1002/ajpa.1330480218>
- Hellenthal, G., Busby, G.B., Band, G., Wilson, J.F., Capelli, C., Falush, D., and Myers, S., 2014. A genetic atlas of human admixture history. *Science*, **343**: 747-751. <https://doi.org/10.1126/science.1243518>
- Iwamura, E.S., Soares-Vieira, J.A., and Muñoz, D.R., 2004. Human identification and analysis of DNA in bones. *Rev. Hosp. Clin.*, **59**: 383-388. <https://doi.org/10.1590/S0041-87812004000600012>
- Jarrige, J.F., 2008. Mehrgarh neolithic. *Pragdhara*, **18**: 136-154.
- Jobling, M.A., Hurles, M., and Tyler-Smith, C., 2019. *Human evolutionary genetics: origins, peoples and disease*. Garland Sci., <https://doi.org/10.1201/9780203487211>
- Kim, K., Kim, D.H., and Kim, K.Y., 2022. Mitochondrial haplogroup classification of ancient DNA samples using haplotracker. *BioMed. Res. Int.*, <https://doi.org/10.1155/2022/5344418>
- Lovejoy, C.O., Meindl, R.S., Pryzbeck, T.R., and Mensforth, R.P., 1985. Chronological metamorphosis of the auricular surface of the ilium: a new method for the determination of adult skeletal age at death. *Am. J. Physiol. Anthropol.*, **68**: 15-28. <https://doi.org/10.1002/ajpa.1330680103>
- Pereira, L., Richards, M., Goios, A., Alonso, A., Albarrán, C., Garcia, O., Behar, D.M., Gölge, M., Hatina, J., Al-Gazali, L., and Bradley, D.G., 2005. High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium. *Genome Res.*, **15**: 19-24. <https://doi.org/10.1101/gr.3182305>
- Possehl, G.L., 2002. *The Indus civilization: A contemporary perspective*. Rowman Altamira.
- Quintana-Murci, L., Chaix, R., Wells, R.S., Behar, D.M., Sayar, H., Scozzari, R., Rengo, C., Al-Zahery, N., Semino, O., Santachiara-Benerecetti, A.S., and Coppa, A., 2004. Where west meets east: The complex mtDNA landscape of the southwest and Central Asian corridor. *Am. J. Hum. Genet.*, **74**: 827-845. <https://doi.org/10.1086/383236>
- Raghavan, M., DeGiorgio, M., Albrechtsen, A., Moltke, I., Skoglund, P., Korneliusen, T.S., Grønnow, B., Appelt, M., Gulløv, H.C., Friesen, T.M., and Fitzhugh, W., 2014. The genetic prehistory of the New World Arctic. *Science*, **345**. <https://doi.org/10.1126/science.1255832>
- Richards, A.J., Baguley, D.M., Yates, J.R., Lane, C., Nicol, M., Harper, P.S., Scott, J.D., Snead, M.P., 2000. Variation in the vitreous phenotype of Stickler syndrome can be caused by different amino acid substitutions in the X position of the type II collagen Gly-XY triple helix. *Am. J. Hum. Genet.*, **67**: 1083-1094. <https://doi.org/10.1086/321189>
- Stewart, J.B., and Chinnery, P.F., 2021. Extreme heterogeneity of human mitochondrial DNA from organelles to populations. *Nat. Rev. Genet.*, **22**: 106-118. <https://doi.org/10.1038/s41576-020-00284-x>
- Vishnyatsky, L.B., 1999. The paleolithic of central Asia. *J. World Prehist.*, **13**: 69-122. <https://doi.org/10.1023/A:1022538427684>