Nature & Anthropology

SCIEPublish

Review

Understanding Evolutionary Origin of Human Disease-Risk Through Ancient Human Genome

San-Ming Wang *

Faculty of Health Sciences, University of Macau, Macau, China

* Corresponding author. E-mail: sanmingwang@um.edu.mo (S.-M.W.)

Received: 22 May 2026; Revised: 14 September 2026; Accepted: 26 October 2026; Available online: 11 November 2026

ABSTRACT: Genetic variation is the driving force of evolution. A part of the variation is deleterious to the function of the affected genes, causing a high risk of developing various types of diseases in humans. Knowing the evolutionary origin of deleterious variation in modern humans is fundamental to protecting human health. Despite rich knowledge of the roles of deleterious variation in human diseases, however, its origin remains largely elusive due largely to the lack of variation information in ancient humans. Rapid progress in extracting DNA from Carbon-14-dated ancient human fossils and coupling it with high-throughput DNA sequencing has generated massive ancient human genomic DNA sequences, which may provide an opportunity to be used as a reference to determine the evolutionary origin of deleterious variation in modern humans. We explored this possibility by using the cancer-risk deleterious variation in human DNA damage repair (DDR) genes as a model. By systematic comparison of the variation data in hundreds of DDR genes between modern humans and ancient humans, we observed that 1066 (6.2%) of the 20,272 deleterious variants in 91 DDR genes of modern humans were also present in 73 DDR genes of ancient humans dated within the past 10,000 years, peaking between 7000 and 4000 years before present. We propose that the short history of deleterious variants in ancient humans can be largely attributed to human population expansion after the agricultural revolution. Our study demonstrates that the ancient human genome data can be a powerful means to bridge the gap between evolution and medicine, allowing deep understanding of the evolutionary origin of disease risk in modern humans. Our study also highlights that cancer risk by germline deleterious variation in modern humans can be a byproduct of human evolution.

Keywords: Origin; Deleterious variation; Disease risk; Modern humans; Ancient humans; Cancer; Haplotyping; C-14 dating; DNA damage repair; Predisposition

Deleterious Variation and Human Disease-Risk

Genetic variation is the driving force for evolution [1]. Two types of genetic variation can be involved in human diseases: the germline variation and somatic variation. Taking cancer-causing genetic variation as the example. The germline variation-related cancer accounts for about 20% of all cancer cases, whereas the somatic variation-related cancer accounts for 80%. The cancer-related germline variation present mostly in the DDR genes of hundreds is the genetic predisposition for high cancer risk. As germline variation is

© 2026 The authors. This is an open access article under the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/).

heritable, its origin can be traced along evolutionary history. Cancer risk-related germline variation has been extensively studied through decades' effects and is widely applied in clinical prevention, treatment, and prognosis of the cancer. In contrast, cancer-related somatic variation is caused by environmental factors. It occurs randomly, is not inheritable, and its origin cannot be traced along evolutionary history. This review focuses on cancer-related germline variation.

Germline variation includes benign variation and deleterious variation. While benign variation can be beneficial or neutral, deleterious variation damages the function of the affected genes, causing increased risk of developing various types of diseases, including cancer ("deleterious" is a biological term referring to the genetic variants that reduce fitness under purifying selection [2], whereas "pathogenic variation" is a medical term referring to the genetic variants that contribute to disease risk with clinical implications [3]. Although the two definitions are closely related, our study used "deleterious" instead of "pathogenic", as our study focused more on the evolutionary origin of DDR variation rather than its clinical aspects in ancient human populations). About 20% of human cancers are related to the deleterious variation in functionally important genes, mostly in the hundreds of DNA damage-repairing genes [4]. As a genetic predisposition, the deleterious variation is transmitted generation by generation in the carriers' families, causing high cancer risk in the variation carriers. This is best represented by the deleterious variation in *BRCA1*, which repairs double-strand DNA breaks, a serious type of DNA damage [5,6]. The damaged BRCA1 loses its normal function of repairing DNA breaks, leading to genome instability and a high risk of developing breast and ovarian cancer [7–9].

1. Current Status of Studying the Evolutionary Origin of Germline Deleterious Variation

"Where do we come from?" is an essential question in philosophy. The same question applies to the deleterious variation in modern humans, as the knowledge can deepen our understanding of disease risk in modern humans. For example, the founder deleterious variants *BRCA1* c.68_69delAG, *BRCA1* c.5266dupC, and *BRCA2* c.5946delT in the Ashkenazi Jewish population are estimated to have arisen 750–1500 [10], 1800 [11], and 580 years before present (BP) [12]. The information establishes a foundation to effectively prevent breast and ovarian cancer in the Ashkenazi Jewish women through population screening to identify the variant carriers before cancer development.

Evolutionary knowledge is believed to provide a key to understanding and curing human diseases [13]. Regardless of the rich knowledge of the involvement of deleterious variation in human disease risk, there is limited knowledge between evolution and human disease risk, as reflected by the fact that the evolutionary origin for most of the disease-risk variation remains unknown [14–18]. For example, realizing the functional importance of DDR gene variation in human cancer risk, many efforts have been made since the 1990s to study the evolutionary origin of the variation in DDR genes, e.g., the origin of *BRCA1/BRCA2* variation [14–18]. However, the accumulated information remains very limited. Since the 1990s, the evolutionary origins have only been determined for a small set of DDR variants present in a few DDR genes: only 68 variants in 7 DDR genes of *BRCA1*, *BRCA2*, *MLH1*, *MSH2*, *MSH6*, *MUTYH*, and *TP53* have been studied so far, of which only 44 were deleterious [19], whereas the others were variants of uncertain significance (VUS) or benign variation. Most of these studies were focused on evolutionary conservation between species within primates. Factors contributing to the deep gap between evolution and medicine include the very different aims that evolutionary study focuses more on population-wide, long-term fitness and reproduction, whereas medical study focuses more on individual-level, immediate curing and life-saving; poor communication between evolution and medicine; and lack of variation data from ancient humans as the reference to trace the origin of the disease risk-variation in modern humans.

2. Haplotyping Method for Dating Germline Deleterious Variation

Historically, determination of the origin of genetic variation relies on the use of the traditional haplotyping method [20]. The method was designed by the concept that genetic variation is conserved as a linked block in a genome. Following DNA recombination per generation, the size of the ancestor block shared between generations will become smaller. The size differences can be used to date genetic variation in the block. In practice, it uses flanking genetic markers to define a block, then uses Markov Chain Monte Carlo simulation and Bayesian approaches to test differences between the variant-containing block and the control block. The information is then used to calculate the age of the variation within the block. However, the accuracy of the haplotyping method can be affected by multiple factors, including the type of evolution selection imposed on the targeted genes [21], the size of the tested population [22], and the number of DNA recombination sites along given regions in chromosomes [23]. Its technical complexity has also restricted the use of haplotyping methods to date a few highly selected founder variants in high-impact disease-risk genes, such as the founder variant *BRCA1* c.68_69delAG in the Ashkenazi Jewish population [24]. These factors restrict its use for large-scale dating deleterious variation in humans.

3. C-14 Method for Dating Ancient Biological Fossils

The Carbon-14 (C-14) dating method is widely applied to date ancient biological fossils [25]. C-14 on earth is generated by cosmic rays. Through metabolism, C-14 integrates into living beings and reaches an equilibrium state between the environment and the body. Once life ceases, C-14 intake stops, while the C-14 already integrated into the body begins to decay. By measuring the remaining C-14 radioactivity in the dead body, the death time, *i.e.*, the age of the body, can be determined. With its half-life of 5730 years, C-14 is particularly suitable to date the human fossils formed after the last human out-of-Africa migration occurred 60,000–70,000 years ago till nowadays, including the human fossils formed within the past 11,000 years since the last ice age stopped with the following agricultural revolution and human population expansion. As C-14 decay is a basic physics feature not influenced by environmental factors, a C-14-dated result is considered more accurate than a haplotyping-dated result in reflecting the true biological age of genetic variation [26].

4. DNA Extraction from Dated Ancient Human Fossils and Sequencing

The C-14 dating method has been mainly used in archaeological studies for dating ancient biological fossils [27]. Since the mid-1990s, the technology of extracting DNA from ancient human fossils has gradually developed into maturation [28]. Sequences generated from the DNA extracted from C-14-dated ancient human fossils expand the power of C-14 dating from dating ancient fossils to dating ancient human genomes at single-base nucleotide resolution. Tens of thousands of ancient human genomes with different ages have been decoded, providing a potential source to trace the origin of genetic variation in modern humans.

5. Ancient Human Genome Data as a Potential Source to Identify Origin of Germline Deleterious Variation in Modern Humans

The rapid development of ancient DNA genome sciences is generating massive ancient human genome data. For the first time, it is possible to directly use the DNA variation data in ancient humans as the reference to study genetic variation in modern humans. Indeed, efforts have been made to address several fundamental biological questions related to metabolism, pathogen exposure, and environmental adaptation, such as the origin of microbe infection in humans [29], human adaptation to extreme environmental conditions like high altitude [30], the evolution of HLA and ABO blood types [31], the genetic basis of pigmentation [32], and the development of behavior [33] and intelligence [34]. However, few efforts have been made to trace the evolutionary origin of the genetic variation contributing to human disease risk.

Notably, no efforts have been made to understand the evolutionary origin for the genetic variation contributing to human cancer risk.

Several important factors need to be considered when using ancient human genome data to study genetic variation in modern humans. Since separating from chimpanzees, the human lineage has evolved over six million years to Homo sapiens today. To trace the origin of human disease risk, it would be ideal to have ancient human variation data across the entire period of human evolution as the reference. However, this is currently impossible to reach due to the lack of DNA samples from the fossils covering the entire period of human evolution. However, modern humans are the direct descendants of the latest out-of-Africa human migration, which occurred 60,000–70,000 years ago. Further, it is well documented that the genetic variation in modern humans occurred largely after this period, especially in the later stage of the last 10,000 years [35,36]. We reasoned that the availability of C-14-dated ancient human genome data, which were derived mostly from the ancient human fossils aged within the last 10,000 years, may provide an immediate reference to study the origin of human disease risk. Compared to the study of genetic variation in modern humans, studying genetic variation analysis in ancient DNA samples also faces a reliability issue, as the quality of ancient DNA and the derived sequences can be affected by multiple factors, such as DNA degradation, uneven genome coverage, sequencing artifacts, and biased sample availability. However, these issues were commonly present in the early phase of ancient DNA extraction and sequencing. Following extensive efforts made afterwards in ancient DNA extraction, sequencing, and variant identification, the situation has been significantly improved [37]. Therefore, the reliability issue is of less concern nowadays. It is also important to note that our study imposed a critical quality control condition to ensure the reliability of the ancient genetic variants, that is, the use of modern human variants as the bait to identify the counterpart variants in ancient humans. Imposing this condition basically prevents the inclusion of artifactual variants present in ancient genome sequences.

6. Proof-of-Principle Studies: The Origin of Germline Deleterious Variation in Human DDR Genes

We proposed that rich ancient human genome data could be valuable to bridge the gap between evolution and medicine for better understanding the origin of human disease risk. In the past years, my laboratory has been systematically exploring such a possibility. We used cancer risk variation in the DNA damage repair (DDR) system as a model in our study. The DDR system preserved in vertebrates is essential in maintaining genome stability [38–40]. In humans, the DDR system consists of 9 DDR pathways with at least 169 DDR genes. Each pathway repairs a given type of DNA damage: the base excision repair (BER) pathway repairs small DNA lesions, the direct reversal (DR) pathway repairs ubiquitous alkylation damage, the Fanconi anemia (FA) pathway repairs strand cross-link damages, the homologous recombination (HR) and non-homologous end joining (NHEJ) pathways repair double-strand DNA breaks, the mismatch repair (MMR) pathway repairs DNA mismatch errors, and the nucleotide excision repair (NER) pathway repairs helix-distorting DNA damage. Importantly, the DDR genes themselves are prone to germline variation. Part of the variation is deleterious in disrupting the function of the affected genes, causing genome instability, a precondition for oncogenesis. The DDR deleterious variation is well-determined as the genetic predisposition for cancer risk in modern humans [41,42].

We first tested the deleterious variation in *BRCA1* and *BRCA2* (*BRCA*), two of the most important DDR genes related to breast and ovarian cancer [8]. Before the study, however, we need to know if human *BRCA* deleterious variation could originate from non-human species through evolutionary conservation. If this were the case, it would be meaningless to study their human origins by exploring the ancient human genome data. In the study, we used the 6624 *BRCA* deleterious variants from the ClinVar database [43] to represent their presence in modern humans. Using a phylogenetic approach, we searched for the presence of human *BRCA* deleterious variants in 100 vertebrate species in eight clades of primates, Euarchontoglires, Laurasiatheria, Afrotheria, Mammalia, Aves, Sarcopterygii, and fish. The results showed that the human

BRCA deleterious variants were nearly absent in the species of primates closest to humans, including chimps, gorillas, orangutans, gibbons, rhesus monkeys, crab-eating macaques, and baboons. A group of human *BRCA* deleterious variants was present in non-primate species, mostly in the clades of Sarcopterygii and Fish, which are evolutionarily more distant from humans than primates. Based on the results, we concluded that the human *BRCA* deleterious variation was not inherited from non-humans (Figure 1) [44].

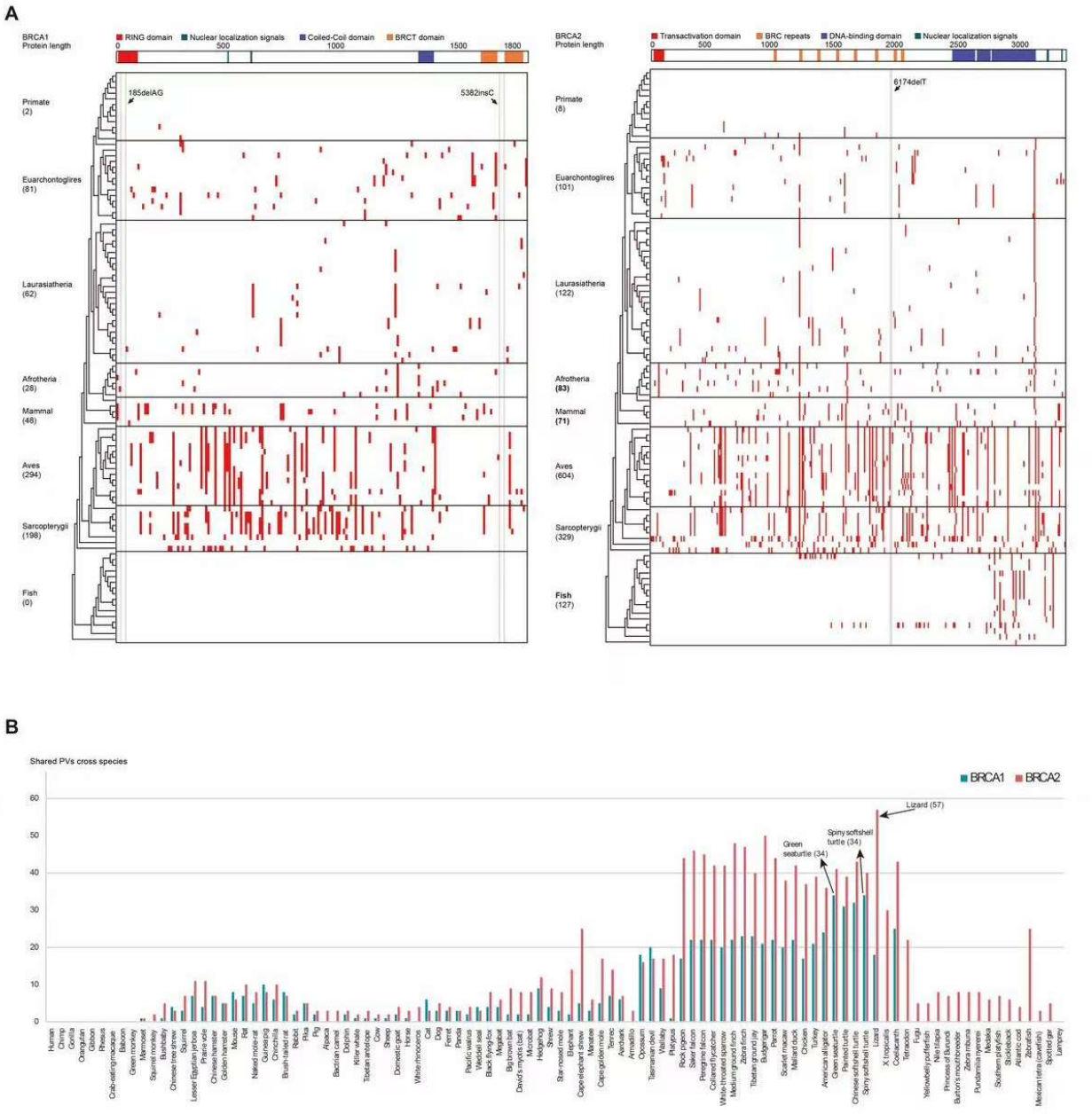


Figure 1. Phylogenetic analysis for the relationship of *BRCA* deleterious variation between human and non-human species. (**A**) Distribution of human *BRCA1* and *BRCA2* deleterious variants in non-human species. Left: *BRCA1*; right: *BRCA2*; (**B**) Combined distribution of modern-human *BRCA* deleterious variants in non-human species. The figure shows no direct evolutionary connection of *BRCA* deleterious variation between modern humans and non-human species. (Li J, et al. Breast Commun 2026;2:12 | https://dx.doi.org/10.21037/bc-25-4, accessed on 22 May 2026).

We then tested whether *BRCA* deleterious variants in modern humans could be present in ancient humans. From public sources, we collected ancient human genome data sets derived from 2792 ancient human fossils dated between 45,000- and 300-years BP (before present). We searched the presence of modern-human *BRCA* deleterious variants in the ancient humans and identified 50 ancient variant carriers

dated within 10,000 years BP except one (BRCA1 c.181T>G) dated at 37,470 BP. The 50 ancient carriers carried 46 modern-human *BRCA* deleterious variants, of which 28 carriers carried 24 *BRCA1* variants and 22 carriers carried 22 *BRCA2* variants (Table 1). Of the 24 *BRCA1* variants, 14 (58.3%) were stop-gain, six (25%) were located at the BRCT domain, and 4 (c.3607C>T, c.4327C>T, c.5095C>T, c.5503C>T) were recurrent in more than one carrier; of the 22 *BRCA2* variants, 15 (68.2%) were stop-gain, and three were located at the nucleic acid binding domain. These ancient carriers lived in Africa, Europe, Asia, Oceania, and Central and North America (Figure 2). The oldest *BRCA1* carrier of *BRCA1* c.181T>G lived in Voronezh Oblast, Russia, dated 37,470 years BP [45]; the youngest *BRCA1* carrier of *BRCA1* c.2599C>T lived in Fujian, China, dated 307 years BP [46]; the oldest *BRCA2* carrier of *BRCA2* c.9573G>A lived in Bor District, Serbia, dated 7874 years BP [47]; and the youngest *BRCA2* carrier of *BRCA2* c.8009C>T lived in Shefa, Vanuatu, dated 350 years BP [48].

Table 1. *BRCA* deleterious variation shared between modern and ancient humans.

Age (BP)	Fossil Site	Variation		Type	Domain
		cDNA	Protein		
BRCA1					
37,470 ± 1210	Voronezh Oblast, Russia	c.181T>G	p.Cys61Gly	missense	Zinc finger
9871 ± 1650	Bor District, Serbia	**c.5095C>T**	**p.Arg1699Trp**	missense	BRCT
7796.5 ± 75.5	Shandong, China	c.2059C>T	p.Gln687Ter	stopgain	-
4971 ± 450	Catalonia, Spain	c.1A>G	p.Met1Val	startloss	-
4671 ± 50	Sistan and Baluchestan, Iran	**c.4327C>T**	**p.Arg1443Ter**	stopgain	-
4121 ± 100	Senec, Slovakia	**c.3607C>T**	**p.Arg1203Ter**	stopgain	-
3350 ± 200	Sardinia, Italy	c.1618G>T	p.Glu540Ter	stopgain	-
3096 ± 20	Hope Town, Bahamas	**c.5503C>T**	**p.Arg1835Ter**	stopgain	BRCT
2110 ± 30	Ballito Bay, South Africa	c.643C>T	p.Gln262Ter	stopgain	
2110 ± 30	Ballito Bay, South Africa	c.949C>T	p.Gln317Ter	stopgain	
2110 ± 30	Ballito Bay, South Africa	c.441+1G>A	-	splice site	
2031	Khovsgol, Mongolia	c.135-1G>A	-	splice site	-
1980 ± 20	Ballito Bay, South Africa	c.3967C>T	p.Gln1323Ter	stopgain	
1933	Rostov Oblast, Russia	c.5353C>T	p.Gln1785Ter	nonsense	BRCT
1695 ± 195	Chukotka, Russia	c.2761C>T	p.Gln921Ter	nonsense	-
1471 ± 50	Girona, Catalonia	c.4255G>T	p.Glu1419Ter	stopgain	-
1221 ± 100	Monsenor Nouel, Dominica	c.2309C>A	p.Ser770Ter	stopgain	-
921 ± 500	Monsenor Nouel, Dominica	**c.3607C>T**	**p.Arg1203Ter**	stopgain	-
921 ± 500	Monsenor Nouel, Dominica	c.5096G>A	p.Arg1699Gln	missense	BRCT
921 ± 500	Monsenor Nouel, Dominica	**c.4327C>T**	**p.Arg1443Ter**	stopgain	-
921 ± 500	Monsenor Nouel, Dominica	c.5216A>T	p.Asp1739Val	missense	BRCT
921 ± 500	Monsenor Nouel, Dominica	**c.5095C>T**	**p.Arg1699Gln**	missense	BRCT
721 ± 100	San Pedro de Macors, Dominica	c.2389G>T	p.Glu797Ter	stopgain	-
721 ± 100	San Pedro de Macors, Dominica	**c.5503C>T**	**p.Arg1835Ter**	stopgain	BRCT
480 ± 30	Eland Cave, South Africa	c.5074G>A	p.Asp1692Asn	nonsynonymous	BRCT
480 ± 30	Eland Cave, South Africa	c.3403C>T	p.Gln1135Ter	stopgain	
419 ± 94	Guangxi, China	c.4573C>T	p.Gln1525Ter	stopgain	SRD
307.5 ± 26.5	Fujian, China	c.2599C>T	p.Gln867Ter	stopgain	-
BRCA2					
7874 ± 72	Bor District, Serbia	c.9573G>A	p.Trp3191Ter	stopgain	NAB
7030 ± 50	Castile and Leon, Spain	c.9466C>T	p.Gln3156Ter	stopgain	NAB
4950 ± 150	Baden, Germany	c.6952C>T	p.Arg2318Ter	stopgain	-
4571 ± 350	Kladno, Central Bohemian Region	c.9382C>T	p.Arg3128Ter	stopgain	-
4384 ± 200	Scotland, UK	c.3922G>T	p.Glu1308Ter	stopgain	-

4250 ± 1200	Vrancea County, Romania	c.7806-2A>G	-	splicing	-
3858 ± 200	Budapest, Hungary	c.8695C>T	p.Gln2899Ter	stopgain	-
3546 ± 325	Atyrau region, Kazakhstan	c.244A>T	p.Lys82Ter	stopgain	-
2921 ± 100	NWFP, Pakistan	c.1189C>T	p.Gln397Ter	stopgain	-
2921 ± 100	NWFP, Pakistan	c.3469G>T	p.Glu1157Ter	stopgain	-
2821 ± 100	East Kazakhstan, Kazakhstan	c.1825C>T	p.Gln609Ter	stopgain	-
2821 ± 100	East Kazakhstan, Kazakhstan	c.2455C>T	p.Gln819Ter	stopgain	-
2330 ± 25	St. Helena, South Africa	c.2905C>T	p.Gln969Ter	stopgain	
2321 ± 100	NWFP, Pakistan	c.5200G>T	p.Glu1734Ter	stopgain	-
2131	Khovsgol, Mongolia	c.316+1G>A	-	splice site	-
2083	Omnogovi, Mongolia	c.7617+1G>A	-	splice site	-
1980 ± 20	Ballito Bay, South Africa	c.7480C>T	p.Arg2494Ter	stopgain	
1783	Issyk Kul, Kyrgyzstan	c.7007G>A	p.Arg2336His	missense	-
1697	South Kazakhstan, Kazakhstan	c.7977-1G>A	-	splice site	-
1100 ± 500	Santo Domingo, Dominica	c.171C>A	p.Tyr57Ter	stopgain	-
480 ± 30	Eland Cave, South Africa	c.475G>A	p.Val159Met	nonsynonymous	
350 ± 100	Shefa, Vanuatu	c.8009C>T	p.Ser2670Leu	missense	NAB

Note: Bold refers to the variants present in more than one ancient individual. (Li J, et al. Life Sci Alliance. 2022, 5, e202101263, https://doi.org/10.26508/lsa.202101263, accessed on 22 May 2026).

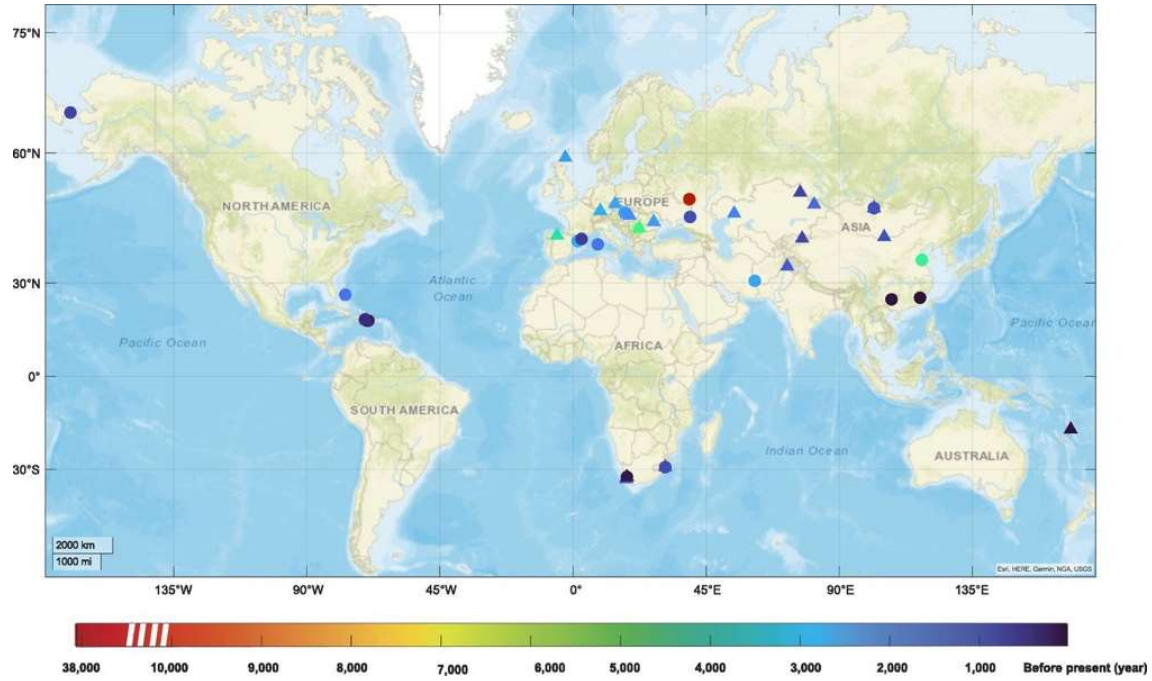


Figure 2. Geographic locations of ancient carriers with *BRCA* deleterious variants. Circle dot: ancient *BRCA1* variant carrier; triangle dot: ancient *BRCA2* variant carrier; irregular dot: overlapped *BRCA1* and *BRCA2* variant carriers (Li J, et al. Life Sci Alliance. 2022, 5, e202101263, https://doi.org/10.26508/lsa.202101263, accessed on 22 May 2026).

Based on the results, we concluded that *BRCA* deleterious variation in modern humans largely originated from humans themselves during recent human history within 10,000 years. To know if the results are only specific for *BRCA* deleterious variation, we applied the same approach to test the deleterious variants in several other DDR genes of *MUTYH*, *PALB2*, *TP53*, *MLH1*, *MSH2*, and *MSH6*. We obtained very similar results that most of the deleterious variation in these DDR genes arose in the recent human history of 10,000 years, as seen in *BRCA* [49–52]. The results raised an interesting question: if the short-arising history observed in these DDR genes might be universal for the deleterious variation in all DDR genes? To test this possibility, we expanded the study to all 169 DDR genes [19,53]. In the study, we

increased the ancient genome sequence data sets from the 2792 ancient humans used in the *BRCA* study to 5031 available dated between 45,045- and 100-years BP (Table S1) including several Neanderthal genome sequences available, and used the 20,272 deleterious variants in 91 DDR genes from the ClinVar database to represent the modern-human DDR deleterious variation. In the 5031 ancient humans, we identified 959 carriers dated mostly within the last 10,000 years, carrying 1266 modern-human deleterious variants in 73 DDR genes, of which 67 deleterious variants in 33 DDR genes were recurrent in more than one carrier (Figure 3A). The 959 ancient carriers lived in the Great Steppe and Asian continent. Comparing the variation rate at 1000-year intervals during the 10,000 years showed that the highest carrier rate of 0.18/1000 individuals was at the 6000–6999 and 4000–4999 intervals (Figure 3B, Table S2). The oldest carrier of *LIG4* c.833G>A was a Brazilian dated 9550 years BP; the youngest carrier of *MSH6* c.1407T>A was a Russian dated 1070 years BP. Of the DDR pathways affected by the deleterious variants, the homologous recombination (HR) pathway had the highest number of 636 deleterious variants in 21 (67.7%) of its 37 genes.

The same variant can have multiple carriers from the first (oldest) carrier to the last (youngest), distributed over a wide time range (Figure 3A). The significance of identifying the first carrier is that it narrows down the arising time of the variant by excluding any later carrier as the origin. Therefore, the chance that most of these first carriers represent the true origin of the variants should be high (Figure 3A). Figure 3B shows the DDR variation at 1000-year intervals in the 5031 ancient humans. A two-peak distribution was present, suggesting that the arising time of DDR variation is dynamic rather than simple accumulation. Among the possible factors can be the changes of population size, as exemplified by the vanished population following the collapse of the LIANGZHU civilization in Zhejiang, China, around 3000 BP due to climate change-caused sea level increase during the period [54].

Consistent with the presence of Neanderthal DNA in modern humans, we also identified 5 DDR deleterious variants in 5 Neanderthals dated between 80,000- and 50,000-years BP, including *ATM* c.5918+1G>A, *ATM* c.8584+1G>A, *ATM* c.7089+1G>A, *BRCA2* c.9076C>T, and *CHEK2* c.988C>T (Table 2).

Table 2. Deleterious variation of modern humans identified in Neanderthal.

Gene	HGVSc	Neanderthal	Age	Location
ATM	c.5918+1G>A	Chagyrskaya K	>50,000	Altai mountains, Russia
ATM	c.8584+1G>A	Chagyrskaya B	>50,000	Altai mountains, Russia
ATM	c.7089+1G>A	Chagyrskaya 8, Vi33.19	80,000 to 50,000	Altai mountains, Russia
BRCA2	c.9076C>T	Chagyrskaya F	80,000	Altai mountains, Russia
CHEK2	c.988C>T	Chagyrskaya B	>50,000	Altai mountains, Russia Vindija cave, Croatia

We also compared the quantitative distribution for the DDR deleterious variants shared between modern and ancient humans. Using the ratio between individual variant/total variants as the indicator, we observed that modern and ancient humans had a similar pattern of quantitative distribution for the shared deleterious variants, with the deleterious variants in *ATM*, *BRCA2*, and *BRCA1* at the highest (Figure 4). The deleterious variants in *TP53*, *MLH1*, and *MSH2* were exceptionally higher in ancient humans but lower in modern humans, possibly being suppressed in modern humans to prevent their deleterious effects.

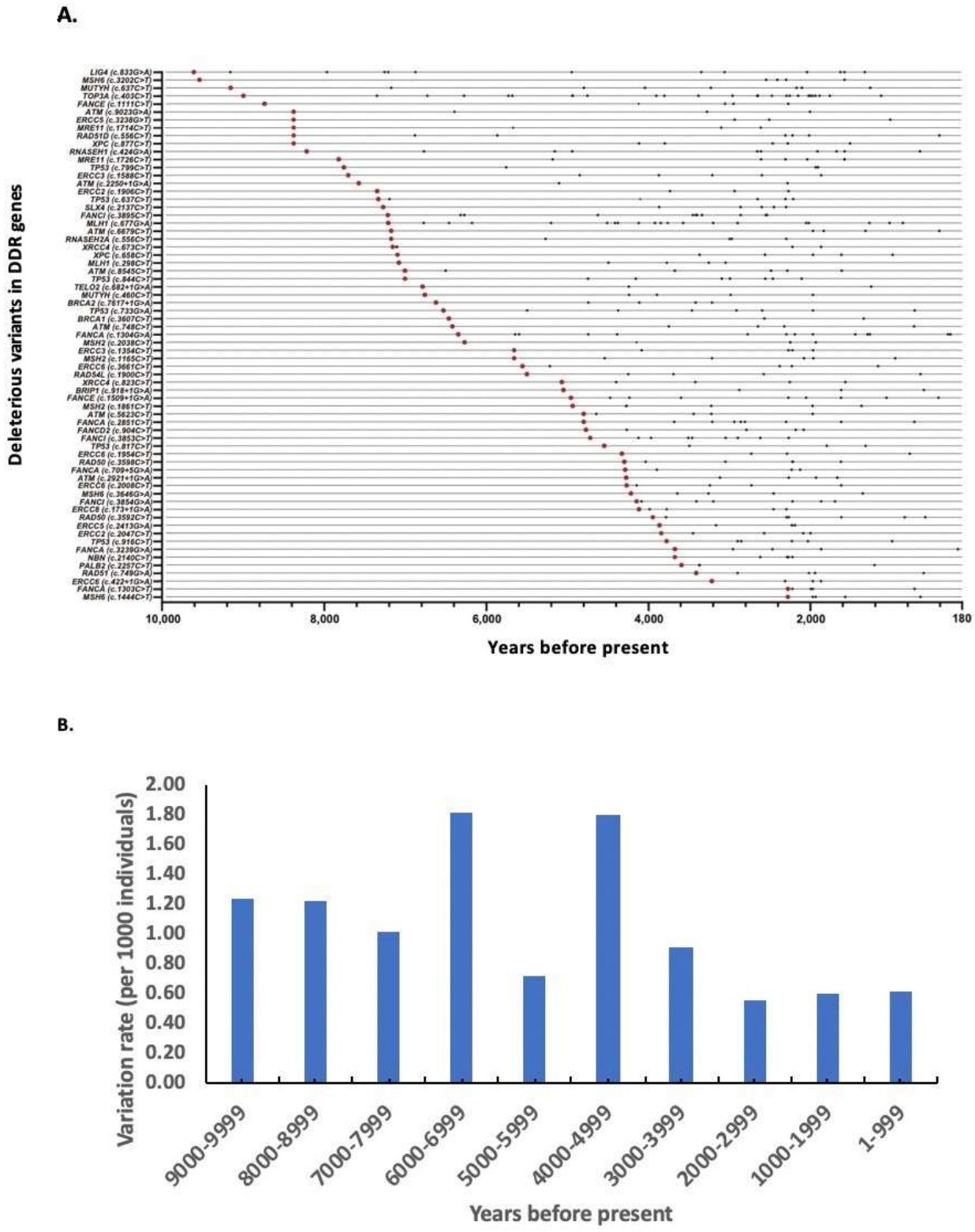


Figure 3. DDR deleterious variation shared between modern and ancient humans within the last 10,000 years. (**A**) Time distribution of the ancient carriers of DDR deleterious variations. It shows the timing of the carriers for the 67 recurrent deleterious variations in 33 DDR genes within the last 10,000 years. The red dot in each line refers to the first carrier for each variant; the black dot(s) in the same line refer to the later carrier(s) carrying the same variant. The line to the left of the red dot shows the absence of variant carriers in samples older than the first carrier. (**B**) Variation of DDR deleterious variation in ancient humans at 1000-year intervals during the past 10,000 years. The rate in each interval was normalized by dividing the number of ancient carriers at each interval with the total number of 5031 ancient individuals. The 6000–6999 and 4000–4999 intervals showed the highest variation among all intervals (Zhao B, et al., BMC Cancer 2024, 24, 415. DOI:10.1186/s12885-024-12160-6, accessed on 22 May 2026).

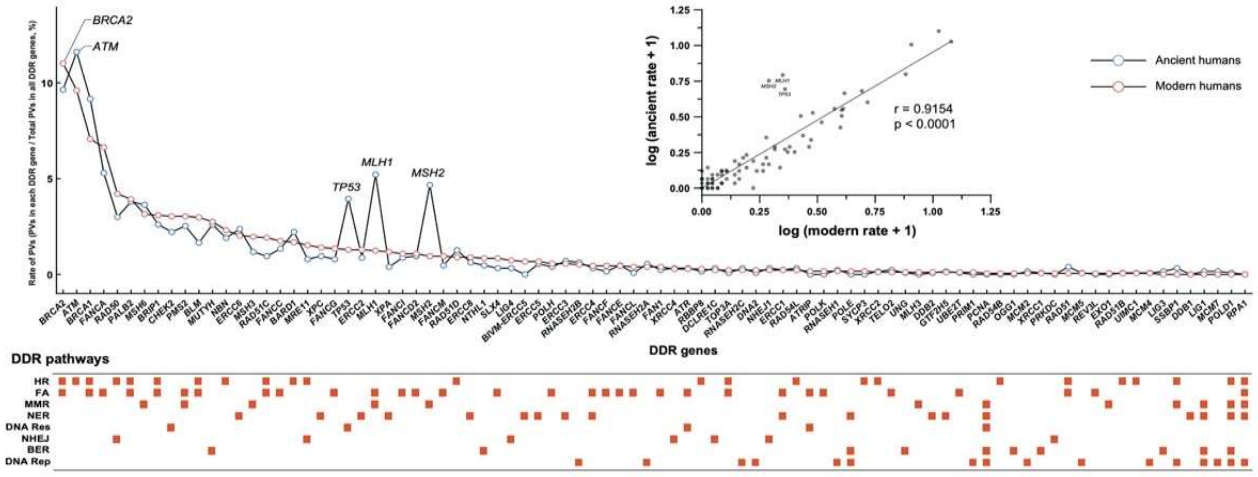


Figure 4. Quantitative comparison of DDR deleterious variation between ancient and modern humans. The comparison was made between the 1266 DDR deleterious variants identified in ancient humans in this study and the 1781 DDR deleterious variants identified in global modern humans by our previous study [53]. The rate of deleterious variants in each DDR gene in ancient or modern humans was calculated by the number of deleterious variants in each gene/total number of deleterious variants in all DDR genes. The red rectangle in the lower part refers to the deleterious variation-containing DDR genes in the affected DDR pathways. The data show that most of the DDR deleterious variation shared between ancient and modern humans had a similar pattern of quantitative distribution (Zhao B, et al. BMC Cancer 2024, 24, 415. DOI:10.1186/s12885-024-12160-6,accessed on 22 May 2026).

We further tested whether the DDR deleterious variation identified in ancient humans originated in Africans before the latest Out-of-Africa human migration or in descendant non-Africans after the latest Out-of-Africa human migration. High sharing among ancient humans, Africans, and non-Africans would imply that the variation arose before the Out-of-Africa migration, and high sharing between ancient humans and non-Africans but not Africans would suggest that the variation arose in non-Africans after the out-of-Africa migration. We extracted DDR deleterious variants from the genome data of 28,872 African individuals and 222,342 non-African individuals. We then compared the deleterious variation of ancient humans with those of Africans and non-Africans. Of the 397 shared ancient variants, 15 (3.8%) were shared with Africans only, 106 (26.7%) with both Africans and non-Africans, and 276 (69.5%) with non-Africans only. The non-proportional rates of population size (28,872 Africans /222,342 non-Africans =13%) and shared variants (15 shared with Africans only and 276 with non-Africans only = 5.4%) suggest that the DDR deleterious variation of ancient humans mostly originated in descendant non-Africans of the latest Out-of-Africa migrants (Figure 5) [55].

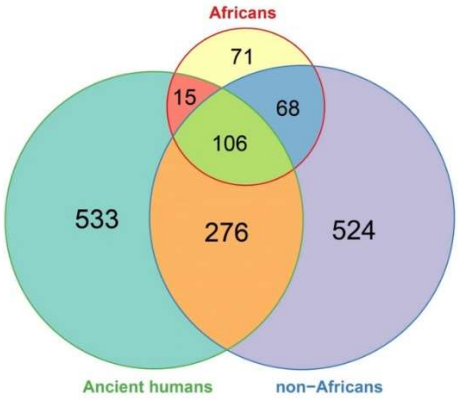


Figure 5. Comparison of DDR deleterious variation between ancient humans, Africans, and non-Africans. It shows that over two third of the ancient DDR PVs were shared only with non-Africans, highlighting their non-African origins (He J, et al. Front Genet. 2024, 15, 1408952. DOI:10.3389/fgene.2024.1408952, accessed on 22 May 2026).

We also tested the possibility that human admixture could transfer DDR deleterious variation between different ethnic human populations. Over five million Brazilian individuals carry Portuguese heritage formed during Portugal's global colonization era, constituting 2.3% of Brazil's population of 214 million. In the study, we observed in the Brazilian population the presence of three well-determined Portuguese-*BRCA* founder deleterious variants (*BRCA1* c.2037delinsCC with unknown arose time, *BRCA1* c.3331_3334del arose over 2000 years ago [56], and *BRCA2* c.156_157insAlu arose over 500 years ago [57]). The results indicate that human admixture can indeed impact human health, such as cancer risk, by transferring disease-risk genetic variation between different ethnic populations (Figure 6) [58].

A.

No.	Variants	Tested cases	# Carrier	Carrier rate (%)	Population Frequency	
					GMAF*	gnomAD**
BRCA1						
11	**c.3331_3334del*** **	**4,131**	**84**	**2.03**	-	-
3	c.211A>G	2,416	32	1.32	-	-
4	c.3477_3480del	158	1	0.63	-	0.00000
14	c.3477_3480del	158	1	0.63	-	0.00000
7	c.4165_4166del	998	6	0.60	-	-
9	c.470_471del	2,265	13	0.57	-	-
2	**c.2037delinsCC*** **	**1,261**	**5**	**0.40**	-	-
13	c.3257T>G	769	3	0.39	-	-
10	c.66dup	780	3	0.38	-	-
6	c.3817C>T	1,916	7	0.37	-	-
5	c.3627dup	649	2	0.31	-	0.00001
8	c.4484G>T	418	1	0.24	-	0.00000
1	c.1016dup	1,267	2	0.16	-	-
12	c.116G>A	1,267	1	0.08	-	0.00003
BRCA2						
3	c.1310_1313del	44	1	2.27	-	-
2	**c.156_157insAlu*** **	**4,471**	**30**	**0.67**	-	-
1	c.2T>G	2,727	14	0.51	0.00001	-
14	c.9382C>T	2,429	12	0.49	-	0.00001
12	c.9004G>A	1,067	4	0.37	-	-
7	c.6656C>G	2,523	8	0.32	-	-
13	c.9098_9099insA	649	2	0.31	-	-
10	c.8488-1G>A	2,528	7	0.28	-	-
5	c.4808del	2,528	6	0.24	-	-
8	c.7007G>A	1,067	2	0.19	-	-
9	c.7738C>T	1,916	3	0.16	-	-
6	c.6468_6469del	649	1	0.15	-	-
11	c.8695C>T	649	1	0.15	-	-
15	c.956dup	649	1	0.15	-	-
4	c.4380_4381del	1,267	1	0.08	-	-
	Total	7,711****	254			

*Global minor allele frequency
**Genome Aggregation Database
***Portuguese *BRCA* founder variants
**** Non-reduandant cases

B.

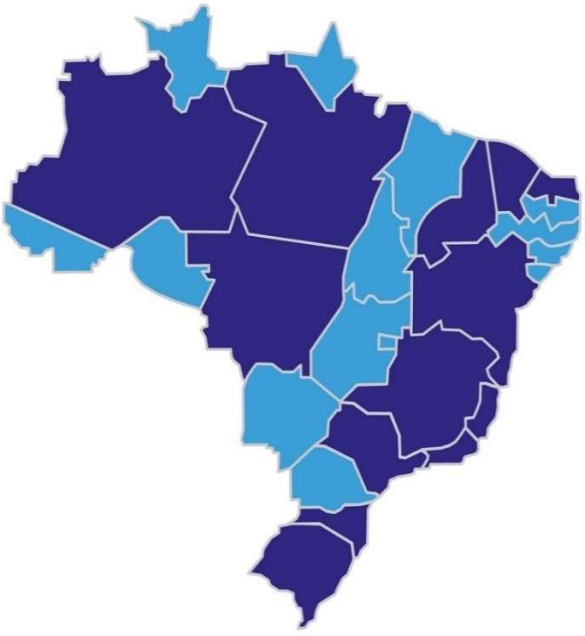


Figure 6. The presence of Portuguese *BRCA* founder deleterious variants in the Brazilian population. (**A**) Of the 29 of the 56 Portuguese-derived variants shared with Brazilians, the three founder variants (in bold) were highly present in the Brazilian

population; (**B**) Resident location of the Brazilian carriers of Portuguese *BRCA* variants in multiple states of Brazil (darker blue). (Andaluz S, et al. BMC Genomics. 2024, 25, 416. DOI:10.1186/s12864-024-10311-4, accessed on 22 May 2026).

To test the consistency of haplotyping-dated and C-14-dated results, we compared the age of deleterious variation dated by both methods. Of the ages for the 12 DDR deleterious variants dated by both methods (Table 3), 6 were highly consistent, but 6 were highly inconsistent. For example, *BRCA1* c.4485-1G>A was dated as 3800 years BP by haplotyping but 34,425 years BP by C-14. As C-14 dating is less affected by different factors than haplotyping, we consider that the age dated by the C-14 method should be more reliable than the age dated by the haplotyping method.

Table 3. Age comparison for the pathogenic variants determined by both haplotyping and C-14 dating.

Gene	cDNA	Age (Year Before Present)				
		Haplotyping Dating		C-14 Dating		
		Pubmed ID	Time	Pubmed ID	#Carrier	Time *
Similar Age						
MUTYH	c.1103G>A	23361220	8675	26595274; 29743352; 34937049; 34937049	4	8315
MUTYH	c.452A>G	23361220	7500	30988490; 31699931	2	7701
BRCA1	c.5503C>T	35377489	3975	33157037; 35366416	2	4575
BRCA1	c.3598C>T	35377489	3850	35561686; 29743352; 33579752; 33618348; 34937049; 34937049; 34937049	7	4250
TP53	c.1010G>A	26618902	2000	31147405; 31699931; 33618348; 34937049	4	2530
MLH1	c.306+5G>A	20858721	2200	29743352	1	3818
Very Different Age						
BRCA1	c.4485-1G>A	35377489	3800	29743352; 33122380	2	34425
BRCA1	c.4327C>T	15883839	2000	29743352; 31488661; 33157037; 33157037; 34937049	5	4600
MSH6	c.10C>T	25318681	543	29743352; 26062507	2	6713
BRCA2	c.7480C>T	11039581	400-200	29743352	1	6319
BRCA1	c.5153-1G>A	19912264	380	30417088	1	1724
MSH2	c.2152C>T	30968502	273	31699931	1	1946

* Note: The earliest time if present in multiple samples.

In summary, our study reveals that DDR deleterious variation in modern humans mostly originated in the recent human history of 10,000 years BP. The following factors can contribute to the result:

1. Evolution selection. The genes under positive selection can contain more genetic variation, as shown by the presence of abundant deleterious variants in *BRCA1* [19]. Furthermore, the short history of deleterious variation arising within the past 10,000 years may not allow evolution selection to function effectively by either eliminating or fixing the variation in the genome, leading to the accumulation of variation [59,60].
2. Beneficial effects. Certain deleterious variations can be developmentally beneficial by promoting fitness, although they can lead to high disease risk in individual carriers. *BRCA1* variation is an example. *BRCA1* variants are so highly abundant that there is one carrier among a few hundred individuals in modern humans. Besides the cancer risk, evidence shows that the *BRCA1* variant carriers have higher reproduction capacity than non-carriers, suggesting *BRCA1* variants may enhance human reproduction [61].
3. Size of genes. Genes with larger coding region often have more deleterious variations than those with smaller coding region. For example, BRCA1 has 1863 residues, BRCA2 has 3418 residues, and ATM has 3056 residues. All of them had highly abundant deleterious variants (Figure 4).
4. Cultural influence. Human consanguineous marriage can substantially alter the carrier rate of *BRCA* deleterious variation in the practicing population [62,63], and human admixture can transfer disease-risk deleterious variation between ethnic human populations [58].
5. Human evolution history. The short history of human DDR variation is coincident with several key events of human evolution history after the latest Out-of-Africa human migration. These include the ending of the last glacial period, the agricultural revolution and industrial revolution, human population expansion, *etc*. It is estimated that only a few thousand human individuals moved out of Africa during the latest Out-of-Africa human migration, and the size of the human population remained lower till the end of the last glacial period of 11,000 years ago [64,65]. The following agricultural revolution and industrial revolution later on led to the expansion of the human population till over 8 billion nowadays [66,67]. As the human genome has a fixed variation rate of 2.5×10^{-8} per nucleotide site per generation [68], the drastic increased population size provides a favorable explanation for the rapid arising of deleterious variation in modern humans within the past 10,000 years.

7. Limitation of Our Study

There are several limitations in our current study. Therefore, more evidence is required to support the conclusion that DDR deleterious variation in modern humans mostly originated in recent human history.

1. Limited time range for the ancient human samples. The current ancient human data mostly consist of individuals aged within the past 10,000 years. As such, the deleterious variation that originated earlier than this time range may be missed. The situation can be improved when more old ancient human samples can be identified and sequenced, as shown by a recent study that the number of sequenced ancient human cases has rapidly increased from 5031 used in our recent study to over 15,000 cases in 2026 [69].
2. The first carrier identified may not be the true first carrier. Such a situation may particularly affect the first carriers closer to 10,000 years BP, as not many ancient samples older than 10,000 years are currently available. This should be much less of concern for the first carriers distant from 10,000 years BP, as non-carrier samples older than these first carriers are widely present (Figure 3A).
3. Lower quality of ancient human genome sequences. Extensive technical improvement for ancient DNA extraction and sequencing has been made, resulting in decoding a tenth of thousands of high-quality ancient human genomes. The ancient human genome sequence data used in our study were dominantly generated in recent years with tight quality control, ensuring high reliability of the identified ancient variants. The use of modern-human variants to fish out their ancient counterparts greatly eliminates the contamination of artifactual variants.
4. Uneven age distribution of ancient human samples. Although the ancient human samples were aged mostly within the Holocene of 11,000 years BP, they were not equally distributed within this period.

5. Lack of experimental validation for the ancient variants. Ideally, validation should be performed for the ancient variants. However, the rarity of ancient DNA samples basically prevents attempts at validation study.
6. A variant shared between ancient and modern humans may not have a hereditary relationship. This is particularly true for the "hot-spot" variation loci that can generate spontaneous variation, as shown in certain *TP53* variants [70,71].

8. Conclusions

Our study provides proof-principle evidence that ancient human genome data are valuable for understanding the evolutionary basis of human disease risk. Based on the results from our study, we consider that cancer risk caused by germline deleterious variation in modern humans is largely a byproduct of human evolution. We propose that establishing and timely expanding the ancient human variation databases should greatly facilitate the efforts of determining the age and origin of deleterious variation contributing to human disease risk (https://genemutation.boboz.io/dbddr-ancienthumans/, accessed on 30 September 2026).

Supplementary Materials

The following supporting information can be found at: https://www.sciepublish.com/article/pii/1320, Table S1: List of 5031 ancient human samples used in the study; Table S2: List of 1266 DDR deleterious variants shared between modern and ancient humans.

Acknowledgments

The author thanks all laboratory members for their contribution to the study and the Information and Communication Technology Office of the University of Macau for providing the high-performance computing resources for the study.

Ethics Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

All C-14-dated ancient human genome data and deleterious variation data used in the study were obtained from public data sources freely available, as indicated in the text citation.

Funding

This research was funded by the Macau Science and Technology Development Fund (085/2017/A2, 0077/2019/AMJ, and 0032/2022/A1).

Declaration of Competing Interests

The author declares no competing interests in the study.

References

1. Hershberg R. Mutation—The Engine of Evolution: Studying Mutation and Its Role in the Evolution of Bacteria. *Cold Spring Harb. Perspect. Biol.* **2015**, *7*, a018077. DOI:10.1101/cshperspect.a018077

2. MacArthur DG, Manolio TA, Dimmock DP, Rehm HL, Shendure J, Abecasis GR, et al. Guidelines for investigating causality of sequence variants in human disease. *Nature* **2014**, *508*, 469–476. DOI:10.1038/nature13127
3. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **2015**, *17*, 405–424. DOI:10.1038/gim.2015.30
4. Lovely C, Rampersad M, Fernandes Y, Eberhart J. Gene–environment interactions in development and disease. *WIREs Dev. Biol.* **2017**, *6*, e247. DOI:10.1002/wdev.247
5. Li X, Heyer WD. Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res.* **2008**, *18*, 99–113. DOI:10.1038/cr.2008.1
6. Stewart MD, Merino Vega D, Arend RC, Baden JF, Barbash O, Beaubier N, et al. Homologous Recombination Deficiency: Concepts, Definitions, and Assays. *Oncol.* **2022**, *27*, 167–174. DOI:10.1093/oncolo/oyab053
7. Jin J, Cao J, Li B, Li T, Zhang J, Cao J, et al. Landscape of DNA damage response gene alterations in breast cancer: A comprehensive investigation. *Cancer* **2023**, *129*, 845–859. DOI:10.1002/cncr.34618
8. Welcsh PL. BRCA1 and BRCA2 and the genetics of breast and ovarian cancer. *Hum. Mol. Genet.* **2001**, *10*, 705–713. DOI:10.1093/hmg/10.7.705
9. Forrest IS, Chaudhary K, Vy HMT, Petrazzini BO, Bafna S, Jordan DM, et al. Population-Based Penetrance of Deleterious Clinical Variants. *JAMA* **2022**, *327*, 350. DOI:10.1001/jama.2021.23686
10. Mathieson I, Lazaridis I, Rohland N, Mallick S, Patterson N, Roodenberg SA, et al. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **2015**, *528*, 499–503. DOI:10.1038/nature16152
11. Hamel N, Feng BJ, Foretova L, Stoppa-Lyonnet D, Narod SA, Imyanitov E, et al. On the origin and diffusion of BRCA1 c.5266dupC (5382insC) in European populations. *Eur. J. Hum. Genet.* **2011**, *19*, 300–306. DOI:10.1038/ejhg.2010.203
12. Neuhausen SL, Godwin AK, Gershoni-Baruch R, Schubert E, Garber J, Stoppa-Lyonnet D, et al. Haplotype and Phenotype Analysis of Nine Recurrent BRCA2 Mutations in 111 Families: Results of an International Study. *Am. J. Hum. Genet.* **1998**, *62*, 1381–1388. DOI:10.1086/301885
13. Caulin AF, Maley CC. Peto's Paradox: Evolution's prescription for cancer prevention. *Trends Ecol. Evol.* **2011**, *26*, 175–182. DOI:10.1016/j.tree.2011.01.002
14. American Association for Cancer Research. BRCA Exchange Launches. *Cancer Discov.* **2019**, *9*, 311–312. DOI:10.1158/2159-8290.CD-NB2019-008
15. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, et al. ClinVar: Public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* **2016**, *44*, D862–D868. DOI:10.1093/nar/gkv1222
16. Bodey GP. Evaluation of new agents in breast carcinoma and other chemotherapy-sensitive tumors. *Eur. J. Cancer Clin. Oncol.* **1985**, *21*, 273–276. DOI:10.1016/0277-5379(85)90124-5
17. Loeb LA. Human Cancers Express a Mutator Phenotype: Hypothesis, Origin, and Consequences. *Cancer Res.* **2016**, *76*, 2057–2059. DOI:10.1158/0008-5472.CAN-16-0794
18. Keinan A, Clark AG. Recent Explosive Human Population Growth Has Resulted in an Excess of Rare Genetic Variants. *Science* **2012**, *336*, 740–743. DOI:10.1126/science.1217283
19. Zhao B, Li J, Sinha S, Qin Z, Kou SH, Xiao F, et al. Pathogenic variants in human DNA damage repair genes mostly arose in recent human history. *BMC Cancer* **2024**, *24*, 415. DOI:10.1186/s12885-024-12160-6
20. Slatkin M, Rannala B. Estimating Allele Age. *Annu. Rev. Genom. Hum. Genet.* **2000**, *1*, 225–249. DOI:10.1146/annurev.genom.1.1.225
21. Li J, Zhao B, Qin Z, Kou SH, Chian JS, Xiao F, et al. Pathogenic variation in human DNA damage repair genes was originated from the evolutionary process of modern humans. *Genes Dis.* **2026**, *13*, 101916. DOI:10.1016/j.gendis.2025.101916
22. Kong A, Thorleifsson G, Gudbjartsson DF, Masson G, Sigurdsson A, Jonasdottir A, et al. Fine-scale recombination rate differences between sexes, populations and individuals. *Nature* **2010**, *467*, 1099–1103. DOI:10.1038/nature09525
23. Mackiewicz D, de Oliveira PMC, Moss de Oliveira S, Cebrat S. Distribution of Recombination Hotspots in the Human Genome—A Comparison of Computer Simulations with Real Data. *PLoS ONE* **2013**, *8*, e65272. DOI:10.1371/journal.pone.0065272
24. Laitman Y, Feng BJ, Zamir IM, Weitzel JN, Duncan P, Port D, et al. Haplotype analysis of the 185delAG BRCA1 mutation in ethnically diverse populations. *Eur. J. Hum. Genet.* **2013**, *21*, 212–216. DOI:10.1038/ejhg.2012.124
25. Libby WF. Atmospheric Helium Three and Radiocarbon from Cosmic Radiation. *Phys. Rev.* **1946**, *69*, 671–672. DOI:10.1103/PhysRev.69.671.2
26. Hajdas I, Ascough P, Garnett MH, Fallon SJ, Pearson CL, Quarta G, et al. Radiocarbon dating. *Nat. Rev. Methods Primers* **2021**, *1*, 62. DOI:10.1038/s43586-021-00058-7

27. Arnold JR, Libby WF. Age Determinations by Radiocarbon Content: Checks with Samples of Known Age. *Science* **1949**, *110*, 678–680. DOI:10.1126/science.110.2869.678
28. Pääbo S. Ancient DNA: Extraction, characterization, molecular cloning, and enzymatic amplification. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 1939–1943. DOI:10.1073/pnas.86.6.1939
29. Zeberg H, Pääbo S. The major genetic risk factor for severe COVID-19 is inherited from Neanderthals. *Nature* **2020**, *587*, 610–612. DOI:10.1038/s41586-020-2818-3
30. Huerta-Sánchez E, Jin X, Asan N, Bianba Z, Peter BM, Vinckenbosch N, et al. Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature* **2014**, *512*, 194–197. DOI:10.1038/nature13408
31. Irving-Pease EK, Refoyo-Martínez A, Barrie W, Ingason A, Pearson A, Fischer A, et al. The selection landscape and genetic legacy of ancient Eurasians. *Nature* **2024**, *625*, 312–320. DOI:10.1038/s41586-023-06705-1
32. Maróti Z, Nyerki E, Török T, Varga GI, Kalmár T. Robust imputation-based method for eye, hair, and skin colour prediction from low-coverage ancient DNA. *Sci. Rep.* **2026**, *16*, 7371. DOI:10.1038/s41598-026-38372-3
33. Pao GM, Zhu Q, Perez-Garcia CG, Chou SJ, Suh H, Gage FH, et al. Role of BRCA1 in brain development. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E1240–E1248. DOI:10.1073/pnas.1400783111
34. Dauyey K, Saitou N. Inferring intelligence of ancient people based on modern genomic studies. *J. Hum. Genet.* **2022**, *67*, 527–532. DOI:10.1038/s10038-022-01039-8
35. Ruan Y, Wang H, Chen B, Wen H, Wu CI. Mutations Beget More Mutations—Rapid Evolution of Mutation Rate in Response to the Risk of Runaway Accumulation. *Mol. Biol. Evol.* **2020**, *37*, 1007–1019. DOI:10.1093/molbev/msz283
36. Fu W, O'Connor TD, Jun G, Kang HM, Abecasis G, Leal SM, et al. Analysis of 6515 exomes reveals the recent origin of most human protein-coding variants. *Nature* **2013**, *493*, 216–220. DOI:10.1038/nature11690
37. Pääbo S, Poinar H, Serre D, Jaenicke-Després V, Hebler J, Rohland N, et al. Genetic Analyses from Ancient DNA. *Annu. Rev. Genet.* **2004**, *38*, 645–679. DOI:10.1146/annurev.genet.37.110801.143214
38. Chatterjee N, Walker GC. Mechanisms of DNA damage, repair, and mutagenesis. *Environ. Mol. Mutagen.* **2017**, *58*, 235–263. DOI:10.1002/em.22087
39. Jeggo PA, Pearl LH, Carr AM. DNA repair, genome stability and cancer: A historical perspective. *Nat. Rev. Cancer* **2016**, *16*, 35–42. DOI:10.1038/nrc.2015.4
40. Jackson SP, Bartek J. The DNA-damage response in human biology and disease. *Nature* **2009**, *461*, 1071–1078. DOI:10.1038/nature08467
41. Garber JE, Offit K. Hereditary Cancer Predisposition Syndromes. *J. Clin. Oncol.* **2005**, *23*, 276–292. DOI:10.1200/JCO.2005.10.042
42. Huang R, Zhou PK. DNA damage repair: Historical perspectives, mechanistic pathways and clinical translation for targeted cancer therapy. *Signal Transduct. Target. Ther.* **2021**, *6*, 254. DOI:10.1038/s41392-021-00648-7
43. ClinVar Database. Available online: https://www.ncbi.nlm.nih.gov/clinvar/ (accessed on 22 May 2026).
44. Li J, Zhao B, Huang T, Qin Z, Wang SM. Human *BRCA* pathogenic variants were originated during recent human history. *Life Sci. Alliance* **2022**, *5*, e202101263. DOI:10.26508/lsa.202101263
45. Seguin-Orlando A, Korneliussen TS, Sikora M, Malaspinas AS, Manica A, Moltke I, et al. Genomic structure in Europeans dating back at least 36,200 years. *Science* **2014**, *346*, 1113–1118. DOI:10.1126/science.aaa0114
46. Yang MA, Fan X, Sun B, Chen C, Lang J, Ko YC, et al. Ancient DNA indicates human population shifts and admixture in northern and southern China. *Science* **2020**, *369*, 282–288. DOI:10.1126/science.aba0909
47. Mathieson I, Alpaslan-Roodenberg S, Posth C, Szécsényi-Nagy A, Rohland N, Mallick S, et al. The genomic history of southeastern Europe. *Nature* **2018**, *555*, 197–203. DOI:10.1038/nature25778
48. Lipson M, Spriggs M, Valentin F, Bedford S, Shing R, Zinger W, et al. Three Phases of Ancient Migration Shaped the Ancestry of Human Populations in Vanuatu. *Curr. Biol.* **2020**, *30*, 4846–4856.e6. DOI:10.1016/j.cub.2020.09.035
49. Lei H, Li J, Zhao B, Kou SH, Xiao F, Chen T, et al. Evolutionary origin of germline pathogenic variants in human DNA mismatch repair genes. *Hum Genom.* **2024**, *18*, 5. DOI:10.1186/s40246-024-00573-0
50. Xiao F, Li J, Lagniton PNP, Kou SH, Lei H, Tam B, et al. Evolutionary Origin of MUTYH Germline Pathogenic Variations in Modern Humans. *Biomolecules* **2023**, *13*, 429. DOI:10.3390/biom13030429
51. Chian JS, Li J, Wang SM. Evolutionary Origin of Human PALB2 Germline Pathogenic Variants. *Int. J. Mol. Sci.* **2023**, *24*, 11343. DOI:10.3390/ijms241411343
52. Kou SH, Li J, Tam B, Lei H, Zhao B, Xiao F, et al. *TP53* germline pathogenic variants in modern humans were likely originated during recent human history. *NAR Cancer* **2023**, *5*, zcad025. DOI:10.1093/narcan/zcad025
53. Qin Z, Huang T, Guo M, Wang SM. Distinct landscapes of deleterious variants in DNA damage repair system in ethnic human populations. *Life Sci. Alliance* **2022**, *5*, e202101319. DOI:10.26508/lsa.202101319

54. Zhang H, Cheng H, Sinha A, Spötl C, Cai Y, Liu B, et al. Collapse of the Liangzhu and other Neolithic cultures in the lower Yangtze region in response to climate change. *Sci. Adv.* **2021**, *7*, eabi9275. DOI:10.1126/sciadv.abi9275
55. He J, Kou SH, Li J, Ding X, Wang SM. Pathogenic variants in human DNA damage repair genes mostly arose after the latest human out-of-Africa migration. *Front. Genet.* **2024**, *15*, 1408952. DOI:10.3389/fgene.2024.1408952
56. Tuazon AMDA, Lott P, Bohórquez M, Benavides J, Ramirez C, Criollo A, et al. Haplotype analysis of the internationally distributed BRCA1 c.3331_3334delCAAG founder mutation reveals a common ancestral origin in Iberia. *Breast Cancer Res.* **2020**, *22*, 108. DOI:10.1186/s13058-020-01341-3
57. Peixoto A, Santos C, Pinheiro M, Pinto P, Soares MJ, Rocha P, et al. International distribution and age estimation of the Portuguese BRCA2 c.156_157insAlu founder mutation. *Breast Cancer Res. Treat.* **2011**, *127*, 671–679. DOI:10.1007/s10549-010-1036-3
58. Andaluz S, Zhao B, Sinha S, Lagniton PNP, Costa DA, Ding X, et al. Using Portuguese BRCA pathogenic variation as a model to study the impact of human admixture on human health. *BMC Genom.* **2024**, *25*, 416. DOI:10.1186/s12864-024-10311-4
59. Simons YB, Turchin MC, Pritchard JK, Sella G. The deleterious mutation load is insensitive to recent population history. *Nat. Genet.* **2014**, *46*, 220–224. DOI:10.1038/ng.2896
60. Seo A, Steinberg-Shemer O, Unal S, Casadei S, Walsh T, Gumruk F, et al. Mechanism for survival of homozygous nonsense mutations in the tumor suppressor gene *BRCA1*. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 5241–5246. DOI:10.1073/pnas.1801796115
61. Smith KR, Hanson HA, Mineau GP, Buys SS. Effects of *BRCA1* and *BRCA2* mutations on female fertility. *Proc. R. Soc. B* **2012**, *279*, 1389–1395. DOI:10.1098/rspb.2011.1697
62. Belaiba F, Medimegh I, Bidet Y, Boussetta S, Baroudi O, Mezlini A, et al. BRCA1/BRCA2 Mutations Shaped by Ancient Consanguinity Practice in Southern Mediterranean Populations. *Asian Pac. J. Cancer Prev.* **2018**, *19*, 2963–2972. DOI:10.22034/APJCP.2018.19.10.2963
63. Bittles AH, Black ML. Consanguinity, human evolution, and complex diseases. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 1779–1786. DOI:10.1073/pnas.0906079106
64. Liu H, Prugnolle F, Manica A, Balloux F. A geographically explicit genetic model of worldwide human-settlement history. *Am. J. Hum. Genet.* **2006**, *79*, 230–237. DOI:10.1086/505436
65. Zhivotovsky LA, Rosenberg NA, Feldman MW. Features of Evolution and Expansion of Modern Humans, Inferred from Genomewide Microsatellite Markers. *Am. J. Hum. Genet.* **2003**, *72*, 1171–1186. DOI:10.1086/375120
66. Gignoux CR, Henn BM, Mountain JL. Rapid, global demographic expansions after the origins of agriculture. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 6044–6049. DOI:10.1073/pnas.0914274108
67. Henn BM, Cavalli-Sforza LL, Feldman MW. The great human expansion. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 17758–17764. DOI:10.1073/pnas.1212380109
68. Nachman MW, Crowell SL. Estimate of the Mutation Rate per Nucleotide in Humans. *Genetics* **2000**, *156*, 297–304. DOI:10.1093/genetics/156.1.297
69. Akbari A, Perry A, Barton AR, Kariminejad M, Gazal S, Li Z, et al. Ancient DNA reveals pervasive directional selection across West Eurasia. *Nature* **2026**, *654*, 419–428. DOI:10.1038/s41586-026-10358-1
70. Olivier M, Hollstein M, Hainaut P. TP53 Mutations in Human Cancers: Origins, Consequences, and Clinical Use. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, a001008. DOI:10.1101/cshperspect.a001008
71. Baugh EH, Ke H, Levine AJ, Bonneau RA, Chan CS. Why are there hotspot mutations in the TP53 gene in human cancers? *Cell Death Differ.* **2018**, *25*, 154–160. DOI:10.1038/cdd.2017.180