








Article

Molecular Epidemiology in Amerindians of the Brazilian Amazon Reveals New Genetic Variants in DNA Repair Genes

Amanda de Nazaré Cohen-Paes ^{1,*}, Angélica Leite de Alcântara ¹, Fabiano Cordeiro Moreira ¹, Marianne Rodrigues Fernandes ¹, Karla Beatriz Cardias Cereja Pantoja ¹, Darlen Cardoso de Carvalho ¹, João Farias Guerreiro ², Ândrea Ribeiro-dos-Santos ¹, Sidney Emanuel Batista dos Santos ¹, Paulo Pimentel de Assumpção ¹ and Ney Pereira Carneiro dos Santos ^{1,*}

¹ Oncology Research Nucleus, Universidade Federal do Pará, Belém 66073-000, Pará, Brazil

² Human and Medical Genetics Laboratory, Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém 66075-110, Pará, Brazil

* Correspondence: acohencaastro@gmail.com (A.d.N.C.-P.); npcasantosufpa@gmail.com (N.P.C.d.S.)

Abstract: Native American populations from the Brazilian Amazon have a low genetic diversity and a different genetic profile when compared to people from other continents. Despite this, few studies have been conducted in this group, and there is no description of their genetic data in the various currently existent international databases. The characterization of the genomic profile of a population not only has an impact in studies of population genetics, but also helps to advance diagnostic and therapeutic response studies, leading to the optimization of clinical applicability. Genetic variations in DNA repair genes have been associated with the modulation of susceptibility to various pathologies, as well as in their prognosis and therapy. This is the first study to investigate DNA repair genes in Amerindians from the Brazilian Amazon region. We investigated 13 important DNA repair genes in the exome of 63 Native Americans, comparing our results with those found in 5 continental populations, whose data are available in the Genome Aggregation Database. Our results showed that 57 variants already described in literature were differentially distributed in the Amerindian populations in relation to the continental populations, 7 of which have significant clinical relevance. In addition, 9 new variants were described, suggesting that they are unique to these populations. Our study reinforces the understanding that the Amazonian Native American population presents a unique genetic profile, and our findings may collaborate with the creation of public policies that optimize the quality of life of these groups as well as the Brazilian population, which presents a high degree of interethnic mixing with Amerindian groups.



Citation: Cohen-Paes, A.d.N.; de Alcântara, A.L.; Moreira, F.C.; Fernandes, M.R.; Pantoja, K.B.C.C.; Carvalho, D.C.d.; Guerreiro, J.F.; Ribeiro-dos-Santos, Â.; Santos, S.E.B.d.; Assumpção, P.P.d.; et al. Molecular Epidemiology in Amerindians of the Brazilian Amazon Reveals New Genetic Variants in DNA Repair Genes. *Genes* **2022**, *13*, 1869. <https://doi.org/10.3390/genes13101869>

Academic Editor: David Caramelli

Received: 5 September 2022

Accepted: 20 September 2022

Published: 15 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. 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/>).

Keywords: DNA repair; Native American; Brazil; genetic variants; continental populations

1. Introduction

Genome integrity depends not only on faithful replication, but essentially on proper repair of the data in DNA throughout its synthesis and processing. It is estimated that approximately 10,000 DNA bases are chemically modified every day in each cell [1,2]. These modifications can be spontaneous or induced by chemicals or radiation, causing genomic mutations such as nucleotide substitutions, amplifications, deletions, rearrangements, or chromosomal loss [3,4]. In order to avoid these mutations, cells have evolved different DNA repair mechanisms to respond to specific damage and prevent mutagenesis [5–7]. There are at least four DNA repair mechanisms that have been well-described in humans: base excision repair (BER), nucleotide excision repair (NER), homologous recombination repair (HRR), and the mismatch/mismatch repair (MMR) pathway [8–11].

Given the importance of their functions, several studies have associated genes of DNA repair pathways and susceptibility to diseases such as xeroderma pigmentosum [12,13], cockayne syndrome and trichothiodystrophy [13], and cancer [14–17]. Additionally, when

it comes to cancer, polymorphisms in these genes have been shown to be linked not only to its development, but also to the mechanism of resistance to pharmacological treatment [16,18–21].

Despite the amount of information available in literature regarding imbalances in gene expression of DNA repair genes, there are no studies reporting the frequency of mutations in these genes in Amerindian populations, especially in the Amazon. However, studies of precision medicine have demonstrated the importance of screening not only genetically homogeneous populations, such as the European one, but also Amerindian people distributed worldwide. It has been shown that populations with a high prevalence of Amerindian genomic ancestry have an increased predisposition to develop acute lymphoblastic leukemia, stomach cancer, and tuberculosis [22–25].

Thus, genomic studies on DNA repair genes in Amerindian populations may provide information on their molecular profile and may potentially discover new variants. The findings from molecular epidemiology studies may also promote the investigation of clinical implications regarding the molecular markers described, allowing them to be used as diagnostic and treatment tools for the Amerindian population, as well as for the admixed populations with marked Amerindian ancestry, such as the Brazilian one [26]. Finally, findings from this type of data are the basis for inferences about human evolutionary history [27]. The aim of this study is to characterize the molecular profile of genes present in DNA damage repair pathways in Amerindian populations from the Brazilian Amazon, and to compare these findings with data from continental populations described in the Genome Aggregation Database (gnomAD).

2. Materials and Methods

2.1. Study Population and Ethics

The study was approved by the National Research Ethics Committee (CONEP; available at: <http://conselho.saude.gov.br/comissoes-cns/conep/>) and by the Research Ethics Committee of the Tropical Medicine Center of the Federal University of Pará (CAE: 20654313.6.0000.5172). All individuals and community leaders signed an informed consent form.

The study population consisted of 63 Amerindians from the Brazilian Amazon. Amerindians represent 12 different Amazonian ethnic groups: Asurini do Xingu, Arara/Arara do Iriri, Araweté, Asurini do Tocantins, Awa-Guajá, Kayapó/Xikrin, Zo'é, Wajãpi, Karipuna, Phurere, Munduruku, and Yudjá/Juruna. These individuals were grouped into a single sample named NAT (Native American population).

The Amerindian population data were compared with representatives of five continental populations obtained from the Genome Aggregation Database (available at: <https://gnomad.broadinstitute.org/>; accessed on 15 February 2022), a public catalog of human variation and genotype data. This sample included 8128 individuals from Africa (AFR), 56,885 from Europe (EUR non-finish), 17,296 from the Americas (AMR), 9197 from East Asia (EAS), and 15,308 from South Asia (SAS).

2.2. Extraction of the DNA and Preparation of the Exome Library

DNA was extracted from a peripheral blood sample using the phenol-chloroform method described by Sambrook et al. [28]. To quantify the genetic material, the Nanodrop-8000 spectrophotometer (Thermo Fisher Scientific Inc., Wilmington, DE, USA) was used and its integrity was evaluated by 2% agarose gel electrophoresis.

Libraries were prepared using the Nextera Rapid Capture Exome (Illumina[®], San Diego, CA, USA) and SureSelect Human All Exon V6 (Agilent technologies, Santa Clara, CA, USA) kits, following the manufacturer's recommendations. The sequencing reactions were performed on the NextSeq 500[®] platform (Illumina[®], San Diego, CA, USA) using the NextSeq 500 High-Output v2 Kit 300 cycle kit (Illumina[®], San Diego, CA, USA).

2.3. Bioinformatic Analysis and Statistical Analyses

Reads in FASTQ format were analyzed for quality (FastQC v.0.11 <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>, accessed on 25 June 2021) and filtered to eliminate low-quality reads (fastx_tools v.0.13 http://hannonlab.cshl.edu/fastx_toolkit/; accessed on 25 June 2021). Then, the sequences were aligned with the reference genome (GRCh38) using the BWA v.0.7 tool (<http://bio-bwa.sourceforge.net/>; accessed on 20 March 2022).

After alignment, the generated file was indexed and sorted (SAMtools v.1.2—<http://sourceforge.net/projects/samtools/>; accessed on 20 March 2022). The alignment had to be processed to remove duplicate readings (Picard Tools v.1.129—<http://broadinstitute.github.io/picard/>; accessed on 20 March 2022), mapping quality recalibration and local realignment (GATK v.3.2—<https://www.broadinstitute.org/gatk/>). Finally, the result was processed in search of variants (GATK v.3.2, United States; <https://gatk.broadinstitute.org/hc/en-us>) of the reference genome. The allelic variants were annotated in the ViVa[®] (Viewer of Variants, Natal, RN, Brazil) software. Markers in PCLO gene were also selected in coverage, where the readout should be high-coverage, with a minimum of 10 reads (fastx_tools v.0.13—http://hannonlab.cshl.edu/fastx_toolkit/; accessed on 20 March 2022).

All statistical analyses were performed using the R Studio v.3.5.1 program (R Foundation for Statistical Computing, Vienna, Austria), including the discriminant analysis of principal components (DAPC). Significant differences in allele frequencies between populations were analyzed by Fisher's exact test. The false discovery rate (FDR) proposed by Benjamini and Hochberg [29] was used to correct the multiple analyses. Results were considered statistically significant when the p-value was less or equal than 0.05 ($p \leq 0.05$).

2.4. Selection of the Genetic Variations

We analyzed 13 gene components of 4 different DNA repair pathways in humans (base excision repair (BER), nucleotide excision repair (NER), homologous recombination repair and mismatch repair). A total of 432 variants were found, to which the following selection criteria were applied: (I) the read should be high-coverage, with a minimum of 10 reads (fastx_tools v.0.13—http://hannonlab.cshl.edu/fastx_toolkit/; accessed on 25 March 2022); (II) the predicted impact should be “modifier”, “moderate”, or “high” according to the software SnpEff (<https://pcingola.github.io/SnpEff/>; accessed on 25 March 2022); (c) the difference in allelic frequency of the variants between NAT populations and continental populations should be significant (p -value ≤ 0.05).

3. Results

After applying our selection criteria to the data from the complete exome of the 63 Amerindians investigated, 55 genetic variants were analyzed, which were distributed in 13 genes composing the DNA repair pathways in humans: DNA2, NEIL1, NEIL2, NEIL3, TOP3A, XPC, XRCC1, XRCC3, ERCC1, ERCC2/XPD, ERCC5, MSH3, and MSH4. Table 1 presents the genetic variants and their respective specifics: chromosomal location, genomic position, the wild allele and the mutant allele, the detailed genomic region, gene, and SNPId. We also find the allele frequency data for the world populations investigated here, and the respective allele frequency calculated for variants in the Amerindian population (NAT).

Thirteen of the investigated variants were differentially distributed in the NAT population when compared to the five continental populations investigated, among them: rs10823209 (DNA2); rs7689099 (NEIL3); rs2294913 (TOP3A); rs3212038, rs1799796, rs861531, and rs861537 (XRCC3); rs6151734 and rs1105524 (MSH3); rs3765682 (MSH4); rs907187, rs2293464, and rs1805404 (PARP1). Overall, for the other SNVs presented, the Amazonian indigenous population differed from two or three of the investigated world populations.

Table 1. Allele frequencies of the variants investigated in the Amerindian individuals (NAT) and continental populations (African (AFR), American population (AMR), East Asian (EAS), European (EUR), and South Asian (SAS)) described in the gnomAD database.

Chromosome	Position	Wild	Variant	Region Detailed	Gene	SNPID	AFR	AMR	EAS	EUR	SAS	NAT
chr10	68470579	A	C	Intron	<i>DNA2</i>	rs10823209	0.1702	0.1304	0.2576	0.1397	0.1443	0.0527
chr10	68446276	TA	T	Intron	<i>DNA2</i>	rs34922453	0.5121	0.1948	0.0608	0.1473	0.0572	0.0139
chr10	68414950	T	A	3'-UTR	<i>DNA2</i>	rs1801041	0.2288	0.3172	0.2342	0.3119	0.000	0.0455
chr15	75351373	T	C	Intron	<i>NEIL1</i>	rs11634109	0.09616	0.1283	0.02490	0.2639	0.06537	0.0834
chr4	177335759	C	G	Nonsynonymous coding	<i>NEIL3</i>	rs7689099	0.05816	0.1333	0.02583	0.1165	0.0900	0.000
chr4	177353596	C	T	Nonsynonymous coding	<i>NEIL3</i>	rs13112358	0.6919	0.7025	0.3976	0.7599	0.5836	0.3645
chr4	177353681	A	C	Nonsynonymous coding	<i>NEIL3</i>	rs13112390	0.7222	0.7650	0.4594	0.7902	0.6484	0.5509
chr17	18280697	G	T	Intron	<i>TOP3A</i>	rs6502644	0.3374	0.2471	0.07407	0.3158	0.1864	0.0429
chr17	18305230	G	A	Intron	<i>TOP3A</i>	rs7207123	0.6050	0.2903	0.07764	0.3230	0.1978	0.0572
chr17	18285570	C	T	Intron	<i>TOP3A</i>	rs2294914	0.09835	0.1614	0.1618	0.1037	0.1845	0.2715
chr17	18299682	C	T	Intron	<i>TOP3A</i>	rs2294913	0.1523	0.2384	0.3018	0.2434	0.1797	0.0334
chr17	18290934	C	T	Nonsynonymous coding	<i>TOP3A</i>	rs28671051	0.1744	0.03313	0.003654	0.004159	0.01035	0.000
chr17	18290697	C	A	Intron	<i>TOP3A</i>	rs3817992	0.1525	0.2388	0.3031	0.2435	0.1799	0.381
chr3	14178523	G	C	Nonsynonymous coding	<i>XPC</i>	rs1870134	0.005	0.09175	0.2159	0.005	0.03741	0.2937
chr14	103711849	A	G	Intron	<i>XRCC3</i>	rs3212038	0.1582	0.2901	0.3624	0.3361	0.2017	0.091
chr14	103699590	T	C	Intron	<i>XRCC3</i>	rs1799796	0.1579	0.2891	0.3584	0.3333	0.1999	0.5477
chr14	103706470	C	A	Intron	<i>XRCC3</i>	rs861531	0.2652	0.2664	0.08507	0.3754	0.2405	0.0136
chr14	103700738	C	T	Intron	<i>XRCC3</i>	rs861537	0.4698	0.5578	0.4226	0.7082	0.4332	0.2805
chr19	45408744	A	C	Intragenic	<i>ERCC1</i>	rs735482	0.2856	0.2733	0.4344	0.1405	0.2635	0.3047
chr19	45409148	G	A	Intragenic	<i>ERCC1</i>	rs2336219	0.2779	0.2728	0.4343	0.1404	0.2636	0.3065
chr19	45419065	G	T	Intron	<i>ERCC1</i>	rs3212961	0.2362	0.2584	0.4443	0.1286	0.2521	0.336
chr19	45409085	A	G	Intragenic	<i>ERCC1</i>	rs762562	0.2855	0.2730	0.4345	0.1403	0.2635	0.3047
chr19	45420395	A	G	Protein structural locus	<i>ERCC1</i>	rs11615	0.8862	0.7043	0.7344	0.3743	0.5293	0.8985
chr19	45368886	T	C	Intron	<i>ERCC2/XPD</i>	rs2298860	0.0008012	0.1631	0.008170	0.0002323	0.0005553	0.000
chr19	45351661	T	G	Nonsynonymous coding	<i>ERCC2/XPD</i>	rs13181	0.2251	0.1993	0.07552	0.3795	0.3715	0.1429
chr19	45364001	C	T	Nonsynonymous coding	<i>ERCC2/XPD</i>	rs1799793	0.1058	0.1956	0.04341	0.3545	0.3499	0.1191
chr19	45365033	T	G	Intron	<i>ERCC2/XPD</i>	rs1799785	0.0017	0.0934	0.1567	0.0008	0.0018	0.000
chr13	102858951	T	C	Intron	<i>ERCC5</i>	rs4150299	0.5741	0.3061	0.2633	0.5930	0.5020	0.2339
chr13	102872412	C	T	Intron	<i>ERCC5</i>	rs4150360	0.2319	0.3543	0.2088	0.1848	0.4366	0.1954
chr5	80654689	C	T	5'-UTR	<i>MSH3</i>	rs1105525	0.04785	0.1127	0.06772	0.1783	0.2405	0.000
chr5	80728800	T	G	Intron	<i>MSH3</i>	rs6151734	0.0005	0.01465	0.001345	0.0000	0.0008	0.0577
chr5	80672452	A	G	Intron	<i>MSH3</i>	rs1677653	0.2746	0.2230	0.02350	0.2631	0.2780	0.0136
chr5	80672439	T	A	Intron	<i>MSH3</i>	rs1650648	0.2740	0.2225	0.02310	0.2631	0.2785	0.0395

Table 1. Cont.

Chromosome	Position	Wild	Variant	Region Detailed	Gene	SNPID	AFR	AMR	EAS	EUR	SAS	NAT
chr5	80654693	A	G	5'-UTR	<i>MSH3</i>	rs1105524	0.9187	0.6310	0.3927	0.6668	0.7514	0.4063
chr5	80854162	A	G	Nonsynonymous coding	<i>MSH3</i>	rs184967	0.8920	0.8667	0.9987	0.8429	0.8849	1
chr1	75912678	T	A	Intron	<i>MSH4</i>	rs28693610	0.1073	0.2917	0.5569	0.3903	0.5051	0.2059
chr1	75867478	T	G	Intron	<i>MSH4</i>	rs5745433	0.1600	0.2276	0.01961	0.2521	0.1856	0.0364
chr1	75803775	G	A	Nonsynonymous coding	<i>MSH4</i>	rs5745325	0.2978	0.2335	0.02654	0.2839	0.2133	0.0397
chr1	75867487	T	C	Intron	<i>MSH4</i>	rs3765682	0.0276	0.1524	0.2678	0.07137	0.05466	0.5167
chr1	226407946	C	G	5'UTR	<i>PARP1</i>	rs907187	0.05392	0.3036	0.4301	0.1631	0.1135	0.6855
chr1	226402008	C	T	Intron	<i>PARP1</i>	rs2666428	0.6124	0.4619	0.8060	0.3293	0.3771	0.6572
chr1	226392392	C	T	Intron	<i>PARP1</i>	rs2280712	0.05191	0.1076	0.3364	0.1682	0.1792	0.0136
chr1	226367601	A	G	Nonsynonymous coding	<i>PARP1</i>	rs1136410	0.05391	0.3020	0.4289	0.1593	0.1121	0.6588
chr1	226388595	G	A	Intron	<i>PARP1</i>	rs2293464	0.3651	0.1432	0.3374	0.1691	0.1771	0.0807
chr1	226388569	A	G	Intron	<i>PARP1</i>	rs2255403	0.1527	0.3101	0.4676	0.1596	0.1968	0.7032
chr1	226402132	T	C	Intron	<i>PARP1</i>	rs1805407	0.3885	0.1454	0.3373	0.1692	0.1798	0.0715
chr1	226385701	T	C	Intron	<i>PARP1</i>	rs1805408	0.3891	0.1453	0.3383	0.1691	0.1805	0.0794
chr1	226402257	G	A	Protein structural locus	<i>PARP1</i>	rs1805404	0.1248	0.4249	0.4472	0.1663	0.1030	0.000
chr1	226379307	C	G	Intron	<i>PARP1</i>	rs732284	0.3362	0.1388	0.3385	0.1689	0.1767	0.1016
chr14	20356854	C	T	Intron	<i>PARP2</i>	rs878157	0.09806	0.2069	0.2778	0.2703	0.3581	0.0143
chr14	20345513	A	G	Intron	<i>PARP2</i>	rs3093890	0.07983	0.1972	0.1907	0.2536	0.3426	0.0143
chr14	20357164	T	G	Intron	<i>PARP2</i>	rs200223594	0.1486	0.2147	0.2761	0.2765	0.3761	0.1084
chr14	20350953	T	A	Intron	<i>PARP2</i>	rs3093904	0.07983	0.1978	0.1918	0.2534	0.3440	0.0143
chr14	20357162	T	TTTCAC	Intron	<i>PARP2</i>	rs10625811	0.1490	0.2150	0.2767	0.2768	0.3762	0.113
chr14	20346854	C	T	Intron	<i>PARP2</i>	rs1713430	0.7030	0.4084	0.3690	0.3587	0.5194	0.4181

Figure 1 shows a discriminant analysis of principal components (DAPC) scatter-plot of the six populations analyzed. The DAPC analysis allows for the visualization of well-defined clusters according to the similarity or difference of the frequency distributions of each investigated SNV. Figure 1 demonstrates a significant distance between the Amerindian population (NAT) and the African population.

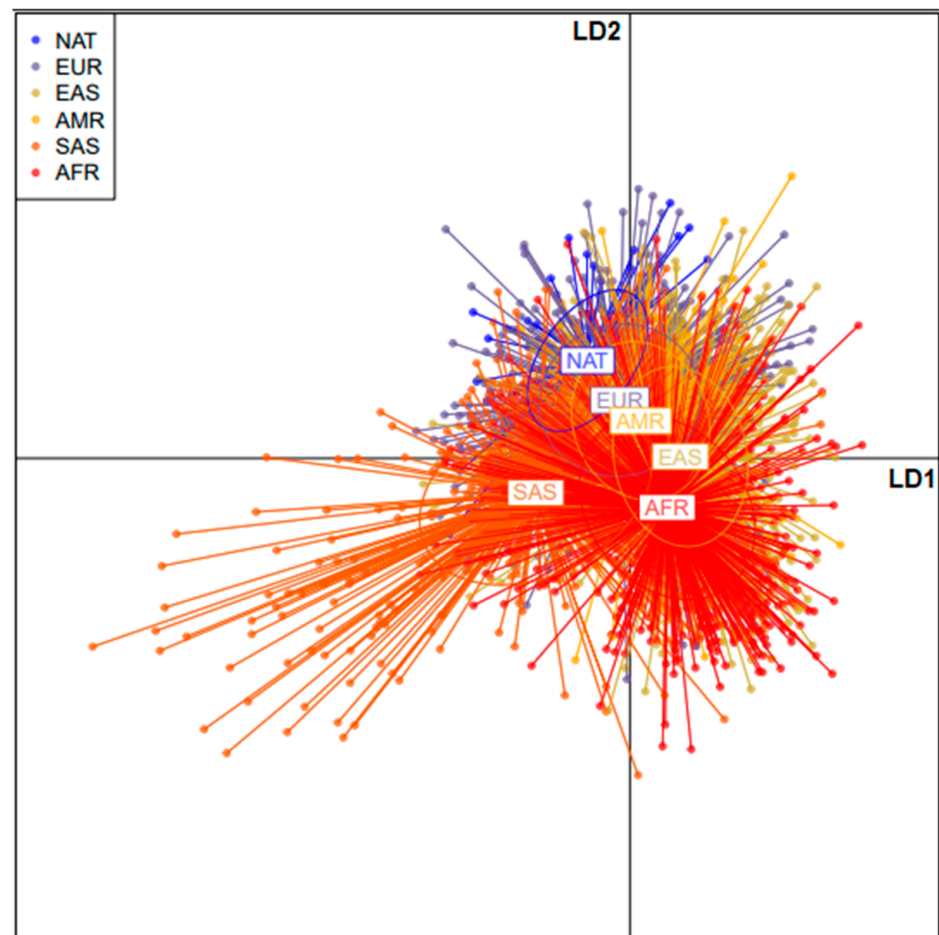


Figure 1. Discriminant analysis of principal components (DAPC) of the analyzed variants in the selected repair genes for Native American populations (NAT), African population (AFR), European population (EUR), American population (AMR), East Asian population (EAS), and South Asian population (SAS).

Table 2 shows the pairwise analysis between the allele frequencies in Amerindians and each of the five continental populations of the gnomAd database for each variant of the repair genes investigated, seven of which have a clinical impact for the development of some pathology or on the therapeutic efficacy of some pathway in question according to the ClinVar database (National Center for Biotechnology Information—NCBI), and three of which have a high impact according to the SNPeff software. For rs1799796 (XRCC3) and rs184967 (MSH3), the NAT population showed a significantly different distribution to the frequencies in relation to American and European populations, these SNVs being respectively associated with susceptibility to breast cancer and hereditary cancer predisposition syndrome.

Supplementary Table S1 shows the same analysis as Table 2 on the 48 markers that are statistically significant, but not clinically significant according to the classification available in the ClinVar database (NCBI), nor classified as high-impact.

Table 2. Comparison between the allelic frequency of the study Native American populations (NAT) and five continental populations described in the gnomAD database for each variant investigated and their respective clinical impacts according to ClinVar.

Gene	SNPID	Variation Type	Impact	Molecular Function	NAT vs. AFR	NAT vs. AMR	NAT vs. EAS	NAT vs. EUR	NAT vs. SAS	ClinVar
XPC	rs1870134	SNV	Moderate	NER ^a	0	0.0003	1	0	0	Xeroderma pigmentosum C Breast cancer susceptibility ND*
XRCC3	rs1799796	SNV	Modifier	HRR ^b	0	0.0019	0.2189	0.0433	0	ND*
TOP3A	rs76300532	SNV	High	HRR ^b	1	1	1	1	1	Efficacy/toxicity to carboplatin, cisplatin, oxaliplatin Xeroderma pigmentosum D; Non-small-cell lung cancer and osteosarcoma
ERCC1	rs11615	SNV	High	NER ^a	1	0.0146	0.0864	0	0	Xeroderma pigmentosum D; Non-small-cell lung cancer and osteosarcoma
ERCC2/ XPD	rs13181	SNV	Moderate	NER ^a	1	1	1	0.004	0.0065	Xeroderma pigmentosum D Xeroderma pigmentosum G
ERCC2/ XPD	rs1799785	SNV	Modifier	NER ^a	1	0.3452	0.0031	1	1	Hereditary cancer-predisposing syndrome ND*
ERCC5	rs4150360	SNV	Modifier	NER ^a	1	0.495	1	1	0.0034	ND*
MSH3	rs184967	SNP	Moderate	MMR ^c	0.1433	0.0234	1	0.0033	0.056	ND*
PARP1	rs1805404	SNV	High	BER ^d	0.0345	0	0	0.0012	0.1387	ND*

* ND = no data; ^a NER = nucleotide excision repair; ^b HRR = homologous recombination repair; ^c MMR = mismatch repair; ^d BER = base excision repair.

Finally, Table 3 describes the chromosomal/genomic position data, the wild and mutant alleles, the impact, type of genetic variation, protein exchange that the mutation may entail, detailed region, and repair pathway that these markers comprise in nine novel genetic variants found in the investigated Amerindian populations. These variants have never been described in the literature and are present in DNA2, PARP2, TOP3A, ERCC2, ERCC5, and MSH3 genes, of which one has a predicted high impact, one has a modifier impact, and the others have a moderate impact.

Table 3. New variants found in Native American population from Brazilian Amazon.

Chromosome (chr)	Position	Wild	Variant	Impact	Variation Type	Change Protein	Region Detailed	Gene	Molecular Function
chr10	68422310	T	A	Moderate	SNV	p.Glu871Val	Nonsynonymous coding	DNA2	BER
chr10	68422732	C	A	Moderate	SNV	p.Gln789His	Nonsynonymous coding	DNA2	BER
chr14	20346875	T	A	Moderate	SNV	p.Cys109Ser	Nonsynonymous coding	PARP2	BER
chr17	18271782	C	T	Modifier	SNV	c.*3020G>A	3'-UTR	TOP3A	HRR
chr17	18314776	CA	C	High	INDEL	p.Met1fs	Frame shift	TOP3A	HRR
chr19	45351734	G	A	Modifier	SNV	c.2191-13C>T	Intron	ERCC2	NER
chr19	45364177	CA	C	Modifier	INDEL	c.815+57delT	Intron	ERCC2	NER
chr13	102865876	G	A	Moderate	SNV	p.Glu722Lys	Nonsynonymous coding	ERCC5	NER
chr5	80654908	G	GCAG...	Moderate	INDEL	p.Ala68_Pro69i	Codon insertion	MSH3	MMR

For all missense variants, we used two other impact prediction tools (SIFT and PolyPhen). Table 4 describes this data for each possible gene transcript in the investigated genomic region.

Table 4. Risk prediction of the new missense variants characterized in the study.

Chromosome (chr)	Position	Gene	Wild	Variant	Sequence	SIFT	PolyPhen		
chr10	68422310	DNA2	T	A	ENSG00000138346	0.03	Deleterious	0.506	Possibly damaging
chr10	68422310	DNA2	T	A	ENSG00000138346	0.08	Tolerated	0.877	Possibly damaging
chr10	68422310	DNA2	T	A	ENSG00000138346	0.05	Tolerated	0.909	Possibly damaging
chr10	68422310	DNA2	T	A	ENSG00000138346	0.35	Tolerated	0.694	Possibly damaging

Table 4. Cont.

Chromosome (chr)	Position	Gene	Wild	Variant	Sequence	SIFT		PolyPhen	
chr13	102865876	ERCC5	G	A	ENSG00000134899	0.61	Tolerated	0.184	Benign
chr13	102865876	ERCC5	G	A	ENSG00000134899	0.6	Tolerated	0.024	Benign
chr13	102865876	ERCC5	G	A	ENSG00000134899	0.64	Tolerated	0.024	Benign
chr14	20346875	PARP2	T	A	ENSG00000129484	0.46	Tolerated	0	Benign
chr14	20346875	PARP2	T	A	ENSG00000129484	0.42	Tolerated	0.003	Benign
chr14	20346875	PARP2	T	A	ENSG00000129484	0.4	Tolerated	0	Benign
chr14	20346875	PARP2	T	A	ENSG00000129484	0.38	Tolerated	0.003	Benign

4. Discussion

The study of genetic variations in repair genes has aided the understanding of various pathological mechanisms, describing risks associated with both individual and population genotypes based on the exposure of cells to xenobiotics that lead to genomic instability [10,30–37].

The past process of global colonization is known to have played an important role in defining current patterns of genetic diversity, and it partially explains geographic variation in susceptibility to certain complex diseases [38,39]. Thus, predisposition to certain illnesses may have an intrinsic relationship with genomic ancestry [40–42]. Recent investigations have shown that Brazilian Amerindian populations have a unique genetic profile, yet it remains undescribed in major population databases [23,43,44].

Therefore, we analyzed the complete exome of 13 DNA repair genes never previously investigated in Amerindian populations, representative of tribes from the Brazilian Amazon. Our description of frequencies and our DAPC analysis demonstrated that the African, South Asiatic, and Amazonian Native American populations were positioned at opposite extremes, being the most genetically distinct regarding the investigated variants, corroborating with findings on the history of human populations [45]. Our results also showed that at least seven of the investigated variants had a high clinical impact, both in disease predisposition and in modulating therapeutic response. The markers rs1799796 of the XRCC3 gene and rs184967 of the MSH3 gene were shown to be statistically significant regarding their distribution in the NAT population when compared to the AMR and EUR ones. The XRCC3 gene encodes a protein involved in the HRR pathway and in the repair of strand breaks caused by X-rays [11,46]. Based on the function of XRCC3, mutations in this gene are related to the development of various neoplastic types, such as osteosarcoma [47], bladder cancer [48], and thyroid cancer [9], among others [11]. The rs1799796 investigated here is primarily associated with susceptibility to breast cancer, an association reinforced by the investigation of Niu and colleagues (2021), who performed a meta-analysis of 13 major studies previously published in literature [49–51].

The MSH3 gene is part the MMR system, whose function is to repair DNA after cross-linking chains and recognize and correct base–base mismatches and insertion/deletion loops generated during DNA replication and homologous recombination [11,52]. The combination of genetic variants in this gene can disrupt cellular responses to DNA damage, directly influencing an individual's sensitivity to carcinogens, so that mutations in the MSH3 gene can modulate everything from carcinogenesis and metastatic progression to therapeutic response [17,53]. The rs184967 (c.2846A>G; Gln940Arg) was associated with increased risk of proximal colon cancer ($p = 0.005$), as well as a worse progression-free survival [54]. This polymorphism is also associated with familial breast cancer [55,56] and colon and rectal cancer [54].

Finally, we found nine new variants, among which three are INDEL-like, and one—which causes a reading matrix change—has an estimated high clinical impact. This mutation is located in the TOP3A gene, responsible for homologous recombination-mediated repair of double-stranded DNA during DNA synthesis, an essential component in the mitochondrial DNA replication process [57]. The remaining novel variants found are distributed among four important DNA repair pathways in humans (Table 3). This data reinforces the importance of studying Amerindian populations, to discover the clinical impact of these mutations, since they are found in critical genes for maintaining genomic integrity. The advantages of studying populations with low genetic diversity such as ours

for screening complex diseases are the high degree of linkage disequilibrium, reduced haplotype complexity, and greater potential for identifying rare variants [58,59]. Kuhn and collaborators in 2012 conducted a study analyzing the genomic profile of the Xavante tribe with other populations, including Brazilian populations, and demonstrated that the indigenous population investigated remained genetically isolated, potentially providing a unique opportunity for hereditary disease-mapping studies [22].

This is the first study to investigate DNA repair genes in Amerindian populations from the Brazilian Amazon, which are genetically unique and not yet described in any of the available databases on human genetic variability. Knowledge of different patterns in human genetic diversity is important in many areas of medical genetics, and it can be used as a tool to maximize understanding regarding susceptibility, diagnosis, prognosis, and therapeutic management for Native American populations, as well as for populations with high Amerindian ancestry, such as the Brazilian one.

5. Conclusions

Our study reinforces the understanding that the Amazonian Native American population presents a unique genetic profile. The characterization of repair genes in this population is an important tool for future studies regarding their association with complex diseases in these populations and also in ones with a high degree of admixture with these groups. Furthermore, our data may contribute to the creation of public policies that optimize the quality of life of Amerindian populations investigated here.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes13101869/s1>, Supplementary Table S1 shows analysis on the 49 markers that are statistically significant, but not clinically significant according to the classification available in the ClinVar database (NCBI), nor classified as high-impact.

Author Contributions: All authors contributed significantly to the preparation of the manuscript. Experimental study design, article writing, and statistical analysis: A.d.N.C.-P.; English proofreading, review, and correction: A.L.d.A., K.B.C.C.P., M.R.F. and D.C.d.C.; statistical and bioinformatics analysis: F.C.M.; funding acquisition: J.F.G. and Â.R.-d.-S.; resources: S.E.B.d.S. and P.P.d.A.; coordination: N.P.C.d.S.; All authors have read and agreed to the published version of the manuscript.

Funding: We acknowledge funding from UFPA (Universidade Federal do Pará), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), and PROPESP—UFPA (Pró-Reitoria de Pesquisa e Pós-Graduação da Universidade Federal do Pará). Dr Ney Santos is supported by CNPq/Produtividade (CNPQ 309999/2021-9). This work is part of the Rede de Pesquisa em Genômica Populacional Humana (Biocomputacional-Protocol No. 3381/2013/CAPES), FAPESPA (Amazon Foundation for the Support of Studies and Research) and Pro-rector of Research and Post-graduation (Propesp) of the Federal University of Pará.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Brazilian National Commission for Ethics in Research (CONEP) identified by No 1062/2006 and 123/98.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The authors confirm that the data supporting the findings of this study are available within the article and its Supplementary Material. Raw data of the studied genes are available from the corresponding author, upon reasonable request.

Acknowledgments: We acknowledge the Universidade Federal do Pará (UFPA); Oncology Research Center (NPO/UFPA); Graduate Program in Oncology and Medical Sciences (PPGOCM/UFPA); Human and Medical Genetics Laboratory (LGHM/UFPA).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **1993**, *362*, 709–715. [[CrossRef](#)]
2. Albelazi, M.S.; Martin, P.R.; Mohammed, S.; Mutti, L.; Elder, R.H. The Biochemical Role of the Human NEIL1 and NEIL3 DNA Glycosylases on Model DNA Replication Forks. *Genes* **2019**, *10*, 315. [[CrossRef](#)]
3. Zheng, L.; Meng, Y.; Campbell, J.L.; Shen, B. Multiple roles of DNA2 nuclease/helicase in DNA metabolism, genome stability and human diseases. *Nucleic Acids Res.* **2020**, *48*, 16–35. [[CrossRef](#)]
4. Oliveira, R.T.G.D.; França, I.G.; LR Junior, H.; Riello, G.B.; Borges, D.D.P.; Cavalcante, G.M.; Magalhães, S.M.; Pinheiro, R.F.C. 9253-6T > c REV3L: A novel marker of poor prognosis in Myelodysplastic syndrome. *Hematol. Transfus. Cell Ther.* **2021**, *43*, 377–381. [[CrossRef](#)] [[PubMed](#)]
5. Maynard, S.; Schurman, S.H.; Harboe, C.; de Souza-Pinto, N.C.; Bohr, V.A. Base excision repair of oxidative DNA damage and association with cancer and aging. *Carcinogenesis* **2008**, *30*, 2–10. [[CrossRef](#)]
6. Wallace, S.S.; Murphy, D.L.; Sweasy, J.B. Base excision repair and cancer. *Cancer Lett.* **2012**, *327*, 73–89. [[CrossRef](#)]
7. Kelley, M.R.; Logsdon, D.; Fishel, M.L. Targeting DNA repair pathways for cancer treatment: What's new? *Future Oncol.* **2014**, *10*, 1215–1237. [[CrossRef](#)]
8. Carter, R.J.; Parsons, J.L. Base Excision Repair, a Pathway Regulated by Posttranslational Modifications. *Mol. Cell. Biol.* **2016**, *36*, 1426–1437. [[CrossRef](#)]
9. Sarwar, R.; Mahjabeen, I.; Bashir, K.; Saeed, S.; Kayani, M.A. Haplotype Based Analysis of XRCC3 Gene Polymorphisms in Thyroid Cancer. *Cell. Physiol. Biochem.* **2017**, *42*, 22–33. [[CrossRef](#)] [[PubMed](#)]
10. Alimu, N.; Qukuerhan, A.; Wang, S.; Abdurehim, Y.; Kuyaxi, P.; Zhang, B.; Yasheng, Y. The association between XRCC1 polymorphism and laryngeal cancer susceptibility in different ethnic groups in Xinjiang, China. *Int. J. Clin. Exp. Pathol.* **2018**, *11*, 4595–4604. [[PubMed](#)]
11. Liu, W.; Ma, S.; Liang, L.; Kou, Z.; Zhang, H.; Yang, J. The association between XRCC3 rs1799794 polymorphism and cancer risk: A meta-analysis of 34 case-control studies. *BMC Med. Genom.* **2021**, *14*, 117. [[CrossRef](#)] [[PubMed](#)]
12. Broughton, B.C.; Berneburg, M.; Fawcett, H.; Taylor, E.M.; Arlett, C.F.; Nardo, T.; Stefanini, M.; Menefee, E.; Price, V.H.; Queille, S.; et al. Two individuals with features of both xeroderma pigmentosum and trichothiodystrophy highlight the complexity of the clinical outcomes of mutations in the XPD gene. *Hum. Mol. Genet.* **2001**, *10*, 2539–2547. [[CrossRef](#)] [[PubMed](#)]
13. Cleaver, J.E.; Lam, E.T.; Revet, I. Disorders of nucleotide excision repair: The genetic and molecular basis of heterogeneity. *Nat. Rev. Genet.* **2009**, *10*, 756–768. [[CrossRef](#)]
14. Vogelsang, M.; Wang, Y.; Veber, N.; Mwapagha, L.M.; Parker, M.I. The Cumulative Effects of Polymorphisms in the DNA Mismatch Repair Genes and Tobacco Smoking in Oesophageal Cancer Risk. *PLoS ONE* **2012**, *7*, e36962.
15. Miao, H.-K.; Chen, L.-P.; Cai, D.-P.; Kong, W.-J.; Xiao, L.; Lin, J. MSH3 rs26279 polymorphism increases cancer risk: A meta-analysis. *Int. J. Clin. Exp. Pathol.* **2015**, *8*, 11060–11067. [[PubMed](#)]
16. Li, Q.; Ma, R.; Zhang, M. XRCC1 rs1799782 (C194T) polymorphism correlated with tumor metastasis and molecular subtypes in breast cancer. *OTT* **2018**, *11*, 8435–8444. [[CrossRef](#)] [[PubMed](#)]
17. Caja, F.; Vodickova, L.; Kral, J.; Vymetalkova, V.; Naccarati, A.; Vodicka, P. DNA Mismatch Repair Gene Variants in Sporadic Solid Cancers. *Int. J. Mol. Sci.* **2020**, *21*, E5561. [[CrossRef](#)]
18. Kirschner, K.; Melton, D.W. Multiple roles of the ERCC1-XPF endonuclease in DNA repair and resistance to anticancer drugs. *Anticancer Res.* **2010**, *30*, 3223–3232.
19. Jiricny, J. The multifaceted mismatch-repair system. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 335–346. [[CrossRef](#)]
20. Polosina, Y.Y.; Cupples, C.G. MutL: Conducting the cell's response to mismatched and misaligned DNA. *BioEssays* **2010**, *32*, 51–59. [[CrossRef](#)]
21. Song, X.; Wang, S.; Hong, X.; Li, X.; Zhao, X.; Huai, C.; Chen, H.; Gao, Z.; Qian, J.; Wang, J.; et al. Single nucleotide polymorphisms of nucleotide excision repair pathway are significantly associated with outcomes of platinum-based chemotherapy in lung cancer. *Sci. Rep.* **2017**, *7*, 11785. [[CrossRef](#)]
22. Kuhn, P.C.; Horimoto, A.R.V.R.; Sanches, J.M.; Filho, J.P.B.V.; Franco, L.; Fabbro, A.D.; Franco, L.J.; Pereira, A.C.; Moises, R.S. Genome-Wide Analysis in Brazilian Xavante Indians Reveals Low Degree of Admixture. *PLoS ONE* **2012**, *7*, e42702. [[CrossRef](#)] [[PubMed](#)]
23. de Carvalho, D.C.; Wanderley, A.V.; Dos Santos, A.M.R.; Moreira, F.C.; de Sá, R.B.A.; Fernandes, M.R.; Modesto, A.A.C.; de Souza, T.P.; Cohen-Paes, A.; Leitão, L.P.C.; et al. Characterization of pharmacogenetic markers related to Acute Lymphoblastic Leukemia toxicity in Amazonian native Americans population. *Sci. Rep.* **2020**, *10*, 10292. [[CrossRef](#)] [[PubMed](#)]
24. Leal, D.F.D.V.B.; Da Silva, M.N.S.; Fernandes, D.C.R.D.O.; Rodrigues, J.C.G.; Barros, M.C.D.C.; Pinto, P.D.D.C.; Pastana, L.F.; Da Silva, C.A.; Fernandes, M.R.; De Assumpção, P.P.; et al. Amerindian genetic ancestry as a risk factor for tuberculosis in an amazonian population. *PLoS ONE* **2020**, *15*, e0236033. [[CrossRef](#)]
25. de Castro, A.N.C.L.; Fernandes, M.R.; de Carvalho, D.C.; de Souza, T.P.; Rodrigues, J.C.G.; Andrade, R.B.; Dos Santos, N.P.C. Polymorphisms of xenobiotic-metabolizing and transporter genes, and the risk of gastric and colorectal cancer in an admixed population from the Brazilian Amazon. *Am. J. Transl. Res.* **2020**, *12*, 6626–6636. [[PubMed](#)]
26. Cohen-Paes, A.D.N.; de Carvalho, D.C.; Pastana, L.F.; Dobbin, E.A.F.; Moreira, F.C.; de Souza, T.P.; Dos Santos, N.P.C. Characterization of PCLO Gene in Amazonian Native American Populations. *Genes* **2022**, *13*, 499. [[CrossRef](#)] [[PubMed](#)]

27. Ribeiro-dos-Santos, A.M.; de Souza, J.E.S.; Almeida, R.; Alencar, D.O.; Barbosa, M.S.; Gusmao, L.; Santos, S. High-Throughput Sequencing of a South American Amerindian. *PLoS ONE* **2013**, *8*, e83340. [[CrossRef](#)] [[PubMed](#)]
28. Green, M.R.; Sambrook, J. *Molecular Cloning: A Laboratory Manual*, 4th ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 2012.
29. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B (Methodol.)* **1995**, *57*, 289–300. [[CrossRef](#)]
30. Gokkusu, C.; Cakmakoglu, B.; Dasdemir, S.; Tulubas, F.; Elitok, A.; Tamer, S.; Seekin, S.; Umman, B. Association Between Genetic Variants of DNA Repair Genes and Coronary Artery Disease. *Genet. Test. Mol. Biomark.* **2013**, *17*, 307–313. [[CrossRef](#)]
31. Czarny, P.; Kwiatkowski, D.; Toma, M.; Kubiak, J.; Sliwinska, A.; Talarowska, M.; Szmraj, J.; Maes, M.; Galecki, P.; Sliwinski, T. Impact of Single Nucleotide Polymorphisms of Base Excision Repair Genes on DNA Damage and Efficiency of DNA Repair in Recurrent Depression Disorder. *Mol. Neurobiol.* **2017**, *54*, 4150–4159. [[CrossRef](#)]
32. Feki-Tounsi, M.; Khlifi, R.; Louati, I.; Fourati, M.; Mhiri, M.-N.; Hamza-Chaffai, A.; Rebai, A. Polymorphisms in XRCC1, ERCC2, and ERCC3 DNA repair genes, CYP1A1 xenobiotic metabolism gene, and tobacco are associated with bladder cancer susceptibility in Tunisian population. *Environ. Sci. Pollut. Res.* **2017**, *24*, 22476–22484. [[CrossRef](#)] [[PubMed](#)]
33. Wu, H.; Li, S.; Hu, X.; Qin, W.; Wang, Y.; Sun, T.; Wu, Z.; Wang, X.; Lu, S.; Xu, D.; et al. Associations of mRNA expression of DNA repair genes and genetic polymorphisms with cancer risk: A bioinformatics analysis and meta-analysis. *J. Cancer* **2019**, *10*, 3593–3607. [[CrossRef](#)] [[PubMed](#)]
34. Yesil-Devecioglu, T.; Dayan, A.; Demirtunc, R.; Sardas, S. Role of DNA repair genes XRCC3 and XRCC1 in predisposition to type 2 diabetes mellitus and diabetic nephropathy. *Endocrinol. Diabetes Y Nutr.* **2019**, *66*, 90–98. [[CrossRef](#)] [[PubMed](#)]
35. Sharma, R.; Lewis, S.; Wlodarski, M.W. DNA Repair Syndromes and Cancer: Insights Into Genetics and Phenotype Patterns. *Front. Pediatr.* **2020**, *8*, 570084. [[CrossRef](#)] [[PubMed](#)]
36. Rybicka, M.; Wozniowiczka, A.; Sznarkowska, A.; Romanowski, T.; Stalke, P.; Dręzewski, M.; Verrier, E.; Baumert, T.; Bielawski, K. Liver Cirrhosis in Chronic Hepatitis B Patients Is Associated with Genetic Variations in DNA Repair Pathway Genes. *Cancers* **2020**, *12*, 3295. [[CrossRef](#)]
37. de Souza, M.R.; Rohr, P.; Kahl, V.F.S.; Kvitko, K.; Cappetta, M.; Lopes, W.M.; da Silva, J. The influence of polymorphisms of xenobiotic-metabolizing and DNA repair genes in DNA damage, telomere length and global DNA methylation evaluated in open-cast coal mining workers. *Ecotoxicol. Environ. Saf.* **2020**, *189*, 109975. [[CrossRef](#)]
38. Tishkoff, S.A.; Verrelli, B.C. Patterns of human genetic diversity: Implications for human evolutionary history and disease. *Annu. Rev. Genom. Hum. Genet.* **2003**, *4*, 293–340. [[CrossRef](#)]
39. Novembre, J.; Di Rienzo, A. Spatial patterns of variation due to natural selection in humans. *Nat. Rev. Genet.* **2009**, *10*, 745–755. [[CrossRef](#)]
40. Katkoori, V.R.; Jia, X.; Shanmugam, C.; Wan, W.; Meleth, S.; Bumpers, H.; Grizzle, W.E.; Manne, U. Prognostic Significance of p53 Codon 72 Polymorphism Differs with Race in Colorectal Adenocarcinoma. *Clin. Cancer Res.* **2009**, *15*, 2406–2416. [[CrossRef](#)]
41. Zhang, Q.; Ma, Y.-Y.; Wang, H.-J.; Shao, C.-M.; Zhang, J.; Ye, Z.-Y. Meta-analysis of the association between P53 codon 72 polymorphisms and gastric cancer. *J. Surg. Oncol.* **2013**, *107*, 360–366. [[CrossRef](#)]
42. Özdemir, B.C.; Dotto, G.-P. Racial Differences in Cancer Susceptibility and Survival: More Than the Color of the Skin? *Trends Cancer* **2017**, *3*, 181–197. [[CrossRef](#)] [[PubMed](#)]
43. Rodrigues, J.C.G.; De Souza, T.P.; Pastana, L.F.; Dos Santos, A.M.R.; Fernandes, M.R.; Pinto, P.; Wanderley, A.V.; De Souza, S.J.; Kroll, J.E.; Pereira, A.L.; et al. Identification of NUDT15 gene variants in Amazonian Amerindians and admixed individuals from northern Brazil. *PLoS ONE* **2020**, *15*, e0231651. [[CrossRef](#)] [[PubMed](#)]
44. Dobbin, E.A.F.; Medeiros, J.A.G.; Costa, M.S.C.R.; Rodrigues, J.C.G.; Guerreiro, J.F.; Kroll, J.E.; Santos, N.P.C.D. Identification of Variants (rs11571707, rs144848, and rs11571769) in the BRCA2 Gene Associated with Hereditary Breast Cancer in Indigenous Populations of the Brazilian Amazon. *Genes* **2021**, *12*, 142. [[CrossRef](#)] [[PubMed](#)]
45. Skoglund, P.; Reich, D. A genomic view of the peopling of the Americas. *Curr. Opin. Genet. Dev.* **2016**, *41*, 27–35. [[CrossRef](#)] [[PubMed](#)]
46. Khanna, K.K.; Jackson, S.P. DNA double-strand breaks: Signaling, repair and the cancer connection. *Nat. Genet.* **2001**, *27*, 247–254. [[CrossRef](#)] [[PubMed](#)]
47. Sobhan, M.R.; Forat-Yazdi, M.; Mazaheri, M.; Zare-Shehneh, M.; Neamatzadeh, H. Association between the DNA Repair Gene XRCC3 rs861539 Polymorphism and Risk of Osteosarcoma: A Systematic Review and Meta-Analysis. *APJCP* **2017**, *18*, 549–555.
48. Peng, Q.; Mo, C.; Tang, W.; Chen, Z.; Li, R.; Zhai, L.; Yang, S.; Wu, J.; Sui, J.; Li, S.; et al. DNA repair gene XRCC3 polymorphisms and bladder cancer risk: A meta-analysis. *Tumor Biol.* **2014**, *35*, 1933–1944. [[CrossRef](#)] [[PubMed](#)]
49. Ali, A.M.; AbdulKareem, H.; Al Anazi, M.; Parine, N.R.; Shaik, J.P.; Alamri, A.; Pathan, A.A.K.; Warsy, A. Polymorphisms in DNA Repair Gene XRCC3 and Susceptibility to Breast Cancer in Saudi Females. *BioMed Res. Int.* **2016**, *2016*, 8721052. [[CrossRef](#)]
50. Alkasaby, M.K.; Abd El-Fattah, A.I.; Ibrahim, I.H.; Abd El-Samie, H.S. Polymorphism of XRCC3 in Egyptian Breast Cancer Patients. *Pharm. Pers. Med.* **2020**, *13*, 273–282. [[CrossRef](#)]
51. Niu, H.; Yang, J.; Chen, X. Associations of rs1799794 and rs1799796 polymorphisms with risk of breast cancer: A meta-analysis. *J. Cancer Res. Ther.* **2021**, *17*, 1225–1233.
52. Yang, Z.; Liu, Z. Potential Functional Variants in DNA Repair Genes Are Associated with Efficacy and Toxicity of Radiotherapy in Patients with Non-Small-Cell Lung Cancer. *J. Oncol.* **2020**, *2020*, 3132786. [[CrossRef](#)] [[PubMed](#)]

53. Vodicka, P.; Musak, L.; Frank, C.; Kazimirova, A.; Vymetalkova, V.; Barancokova, M.; Smolkova, B.; Dzupinkova, Z.; Jiraskova, K.; Vodenkova, S.; et al. Interactions of DNA repair gene variants modulate chromosomal aberrations in healthy subjects. *Carcinogenesis* **2015**, *36*, 1299–1306. [[CrossRef](#)] [[PubMed](#)]
54. Berndt, S.I.; Platz, E.A.; Fallin, M.D.; Thuita, L.W.; Hoffman, S.C.; Helzlsouer, K.J. Mismatch repair polymorphisms and the risk of colorectal cancer. *Int. J. Cancer* **2007**, *120*, 1548–1554. [[CrossRef](#)] [[PubMed](#)]
55. Hu, J.; Lee, E.; Allen, G. Abstract 10: DNA repair and nutritional risk models of triple-negative breast cancer. *Cancer Res.* **2014**, *74*, 10. [[CrossRef](#)]
56. Kappil, M.; Terry, M.B.; Delgado-Cruzata, L.; Liao, Y.; Santella, R.M. Mismatch Repair Polymorphisms as Markers of Breast Cancer Prevalence in the Breast Cancer Family Registry. *Anticancer Res.* **2016**, *36*, 4437–4441. [[CrossRef](#)] [[PubMed](#)]
57. Nicholls, T.J.; Nadalutti, C.A.; Motori, E.; Sommerville, E.W.; Gorman, G.S.; Basu, S.; Hoberg, E.; Turnbull, D.M.; Chinnery, P.F.; Larsson, N.-G.; et al. Topoisomerase 3 α Is Required for Decatenation and Segregation of Human mtDNA. *Mol. Cell* **2018**, *69*, 9–23.e6. [[CrossRef](#)]
58. Jorde, L.B.; Watkins, W.S.; Kere, J.; Nyman, D.; Eriksson, A.W. Gene mapping in isolated populations: New roles for old friends? *Hum. Hered.* **2000**, *50*, 57–65. [[CrossRef](#)]
59. Kristiansson, K.; Naukkarinen, J.; Peltonen, L. Isolated populations and complex disease gene identification. *Genome Biol.* **2008**, *9*, 109. [[CrossRef](#)] [[PubMed](#)]