The Role of ERCC1, BACH1 and NR5A2 Single-Nucleotide Polymorphisms with Pancreatic Cancer Risk and Global Epidemiology: Literature Systematic Review

https://doi.org/10.32635/2176-9745.RBC.2026v72n1.5363

O Papel dos Polimorfismos de Nucleotídeo Único nos Genes ERCC1, BACH1 e NR5A2 na Suscetibilidade ao Câncer de Pâncreas e na Epidemiologia Global: Revisão Sistemática da Literatura

El Papel de los Polimorfismos de Núcleotido Único en los Genes ERCC1, BACH1 y NR5A2 en la Susceptibilidad al Cáncer de Páncreas y la Epidemiología Global: Revisión Sistemática de la Literatura

Ana Clara Patrício Barbosa[1]; Camila Carreira Brito[2]; Giovanna Gilioli da Costa Nunes[3]; Natasha Monte da Silva[4]; João Victor Souza de Barros[5]; Lethycia Santana de Araujo Castro[6]; João Vitor Smith da Silva[7]; Jerffson Carvalho de Souza[8]; Laura Patrícia Albarello Gellen[9]; Ney Pereira Carneiro dos Santos[10]

ABSTRACT

Introduction: Pancreatic cancer is one of the most aggressive malignancies, ranking as the twelfth most common cancer and the sixth leading cause of cancer-related deaths worldwide. Its etiology is multifactorial, involving both genetic and non-genetic factors. **Objective:** Investigate the relationship between epidemiological data on global incidence and mortality of pancreatic cancer and specific genetic variations known as single nucleotide polymorphisms (SNPs). **Method:** A total of 253 SNPs were analyzed, identified through genome-wide association studies (GWAS). Allele frequencies were obtained from global population databases. Pearson's correlation analysis was utilized to evaluate the relationships between SNPs frequencies, incidence, and mortality rates. **Results:** The results identified 16 significant SNPs (p<0.05), among which rs2816938, rs372883, and rs2236575 were highly significant (p<0.001). These variants were primarily associated with higher mortality rates, particularly in European and American populations, while African and Southeast Asian populations showed lower SNPs frequencies and lower mortality rates. The findings suggest that genetic predisposition plays a crucial role in pancreatic cancer susceptibility and progression. **Conclusion:** The correlation between SNPs frequencies and epidemiological data reinforces the influence of population-specific genetic risk factors, highlighting the importance of personalized approaches in screening, prevention, and treatment strategies.

Key words: Pancreatic Neoplasms/etiology; Epidemiology/statistics & numerical data; Incidence; Mortality/trends; Genetic Variation.

RESUMO

Introdução: O câncer de pâncreas é uma das neoplasias malignas mais agressivas, ocupando a décima segunda posição entre os tipos de câncer mais comuns, e a sexta principal causa de mortes relacionadas ao câncer em todo o mundo. Sua etiologia é multifatorial, envolvendo fatores genéticos e não genéticos. **Objetivo:** Investigar a relação entre dados epidemiológicos sobre a incidência e mortalidade global do câncer de pâncreas e variações genéticas específicas conhecidas como polimorfismos de nucleotídeo único (SNP). **Método:** Analisaram-se 253 SNP, identificados por meio de estudos de associação genômica ampla (GWAS). As frequências alélicas foram obtidas em bancos de dados populacionais globais. Utilizando análise de correlação de Pearson, avaliaram-se as relações entre as frequências dos SNP, a incidência e as taxas de mortalidade. **Resultados:** Os resultados identificaram 16 SNP significativos (p<0,05), entre os quais rs2816938, rs372883 e rs2236575 se destacaram com alta significância (p<0,001). Essas variantes estiveram principalmente associadas a maiores taxas de mortalidade, especialmente em populações europeias e americanas, enquanto populações africanas e do Sudeste Asiático apresentaram frequências mais baixas dos SNP e menores taxas de mortalidade. Os achados sugerem que a predisposição genética desempenha um papel crucial na suscetibilidade e progressão do câncer de pâncreas. **Conclusão:** A correlação entre as frequências dos SNP e os dados epidemiológicos reforça a influência de fatores genéticos específicos de cada população, ressaltando a importância de abordagens personalizadas nas estratégias de rastreamento, prevenção e tratamento.

Palavras-chave: Neoplasias Pancreáticas/etiologia; Epidemiologia/estatística & dados numéricos; Incidência; Mortalidade/tendências; Variação Genética.

RESUMEN

Introducción: El cáncer de páncreas es una de las neoplasias malignas más agresivas, ocupando el duodécimo lugar entre los cánceres más comunes, y la sexta causa principal de muertes relacionadas con el cáncer a nivel mundial. Su etiología es multifactorial, involucrando factores genéticos y no genéticos. **Objetivo:** Investigar la relación entre los datos epidemiológicos sobre la incidencia y mortalidad global del cáncer de páncreas y variaciones genéticas específicas conocidas como polimorfismos de un solo nucleótido (SNP). **Método:** Se analizaron 253 SNP, identificados a través de estudios de asociación del genoma completo (GWAS). Las frecuencias alélicas se obtuvieron de bases de datos de poblaciones globales. Utilizando el análisis de correlación de Pearson, se evaluaron las relaciones entre las frecuencias de los SNP, la incidencia y las tasas de mortalidad. **Resultados:** Los resultados identificaron 16 SNP significativos (p<0,05), entre los cuales rs2816938, rs372883 y rs2236575 fueron altamente significativos (p<0,001). Estas variantes se asociaron principalmente con tasas de mortalidad más elevadas, particularmente en poblaciones europeas y americanas, mientras que las poblaciones africanas y del Sudeste Asiático mostraron frecuencias de SNP más bajas y tasas de mortalidad más reducidas. Los hallazgos sugieren que la predisposición genética desempeña un papel crucial en la susceptibilidad y progresión del cáncer de páncreas. **Conclusión:** La correlación entre las frecuencias de SNP y los datos epidemiológicos refuerza la influencia de los factores de riesgo genéticos específicos de cada población, destacando la importancia de enfoques personalizados en las estrategias de detección, prevención y tratamiento.

Palabras clave: Neoplasias Pancreáticas/etiología; Epidemiología/estadística & datos numéricos; Incidencia; Mortalidad/tendencias; Variación Genética.

[1-3,5-8]Universidade Federal do Pará (UFPA), Faculdade de Medicina, Núcleo de Pesquisas em Oncologia. Belém (PA), Brasil. E-mails: ana.patricio.barbosa@ics.ufpa.br; camilalmr077@gmail.com; giovanna.nunes@ics.ufpa.br; joao.souza.barros@ics.ufpa.br; lethycia.castro@ics.ufpa.br; joao.smith@ics.ufpa.br; jerffson.souza@ics.ufpa.br. Orcid iD: https://orcid.org/0009-0002-2806-2003; Orcid iD: https://orcid.org/0009-0006-1150-7490; Orcid iD: https://orcid.org/0009-0002-9005-0281; Orcid iD: https://orcid.org/0009-0001-5618-7578; Orcid iD: https://orcid.org/0009-0008-3690-928X; Orcid iD: https://orcid.org/0009-0006-4971-0558; Orcid iD: https://orcid.org/0009-0004-1858-753X

[4,9,10]UFPA, Núcleo de Pesquisas em Oncologia. Belém (PA), Brasil. E-mails: ntshmonte@gmail.com; laura.patricia.agellen@hotmail.com; npcsantos.ufpa@gmail.com. Orcid iD: https://orcid.org/0000-0003-4061-5045; Orcid iD: https://orcid.org/0000-0001-8620-5987; Orcid iD: https://orcid.org/0000-0001-8087-7037

Corresponding authors: Giovanna Gilioli da Costa Nunes and Ney Pereira Carneiro dos Santos. Avenida Nazaré, 444 – Nazaré. Belém (PA), Brasil. CEP 66035-135. E-mails: giovanna.nunes@ics.com.br; npcsantos.ufpa@gmail.com

INTRODUCTION

Pancreatic cancer (PC) is the twelfth most common cancer and the sixth leading cause of cancer-related mortality worldwide. In 2022, it was estimated that approximately 510 thousand new cases and 467,000 deaths from pancreatic cancer occurred globally. The etiology of pancreatic cancer is multifactorial and highly complex, determined by a combination of genetic and non-genetic factors that include lifestyle aspects as smoking, obesity and heavy alcohol consumption[1].

The incidence of PC varies considerably across different ethnic groups and geographic locations. There are higher incidences in Western Asian men and in South America Uruguayan women[1]. In the United States, there is a slight variability among different ethnic groups, with the lowest incidence and mortality in Asian-American White men and the highest in Afro-American men, whose difference can be attributed to the multiple pancreatic cancer genetic risk loci across ethnic groups[2]. New epidemiological studies indicate that pancreatic cancer has high heritability and suggest a relationship between genetic factors and the onset of the disease. A meta-analysis revealed that approximately 16% of pancreatic cancer individuals had a family history of PC[3].

Genome-Wide Association Studies (GWAS) are an essential resource for the epidemiological and physical health outcomes, since its aim is to provide an estimated association between complex traits and genetic variants, which involves cascades of biological processes that influence and react to environmental exposures[4]. Single Nucleotide Polymorphisms (SNPs) are a type of genetic variation that is intended as a biological marker for various health implications as the likelihood of developing certain conditions[5]. Consequently, SNPs analysis plays a critical role in precision medicine, supporting the development of more specialized treatment regimens and personalized health-care strategies that consider the specific genetic makeup of an individual. SNPs are important in population genetics, as they can provide insights into evolutionary history, genetic variations and diversity in different human populations[6].

Over the past few years, GWAS have transformed cancer genetics. This kind of study has refined the understanding of the genetic mechanisms underlying pancreatic cancer etiology. Subsequently, higher PC incidence and mortality rates are observed in populations of European descent, as the majority of large-scale GWAS have been conducted within this demographic, further elucidating population-specific genetic risk factors[7].

This article aims to correlate epidemiological data on the incidence and mortality of pancreatic cancer worldwide with frequencies of important SNPs in GWAS associated with the susceptibility and severity of this neoplasm in different populations.

METHOD

This review was conducted in accordance with the guidelines outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)[8]. The review protocol was prospectively registered in the International Prospective Register of Systematic Reviews (PROSPERO)[9] under the identifier CRD420251067640.

An extensive literature search was undertaken to identify eligible studies published up to August 25, 2024 across PubMed, Scopus, and Web of Science databases. The full search strategy, including the use of Medical Subject Headings (MeSH) is provided in Table S1 of the Supplementary Material. To ensure the inclusion of the most recent studies, automated alerts were configured in each database to flag newly published articles that met the predefined search parameters. The inclusion criteria were GWAS, cohort studies, case-control studies, and meta-analyses that correlated SNPs with risk, incidence, mortality, prognosis, or susceptibility to pancreatic cancer published in English, Portuguese, or Spanish. The exclusion criteria were studies that did not specifically investigate pancreatic cancer or evaluated other cancer types, as well as studies lacking sufficient data for genetic association analysis. Retrieved records were imported into EndNote® reference management software (version X7, Thomson Reuters, Philadelphia, USA), and duplicate entries were eliminated through both automated detection and manual review. Title and abstract screening was conducted independently by two reviewers and discrepancies were resolved by consensus and, when necessary, in consultation with the senior author.

The data extraction process was conducted independently by four researchers. The selection of SNPs was based on a systematic review of peer-reviewed studies reporting genetic associations with pancreatic cancer. SNPs were included if they demonstrated statistically significant associations with susceptibility, incidence, mortality, risk, or prognosis of PC, irrespective of disease subtype.

To ensure the robustness of the findings, SNPs without significant associations or those lacking available allele frequency data in established genetic databases were excluded. The selection process prioritized variants with replicated evidence of association in independent populations, minimizing the risk of spurious findings due to population stratification or genotyping errors. SNPs screening process is illustrated in Figure 1, showing all the selection and exclusion steps.

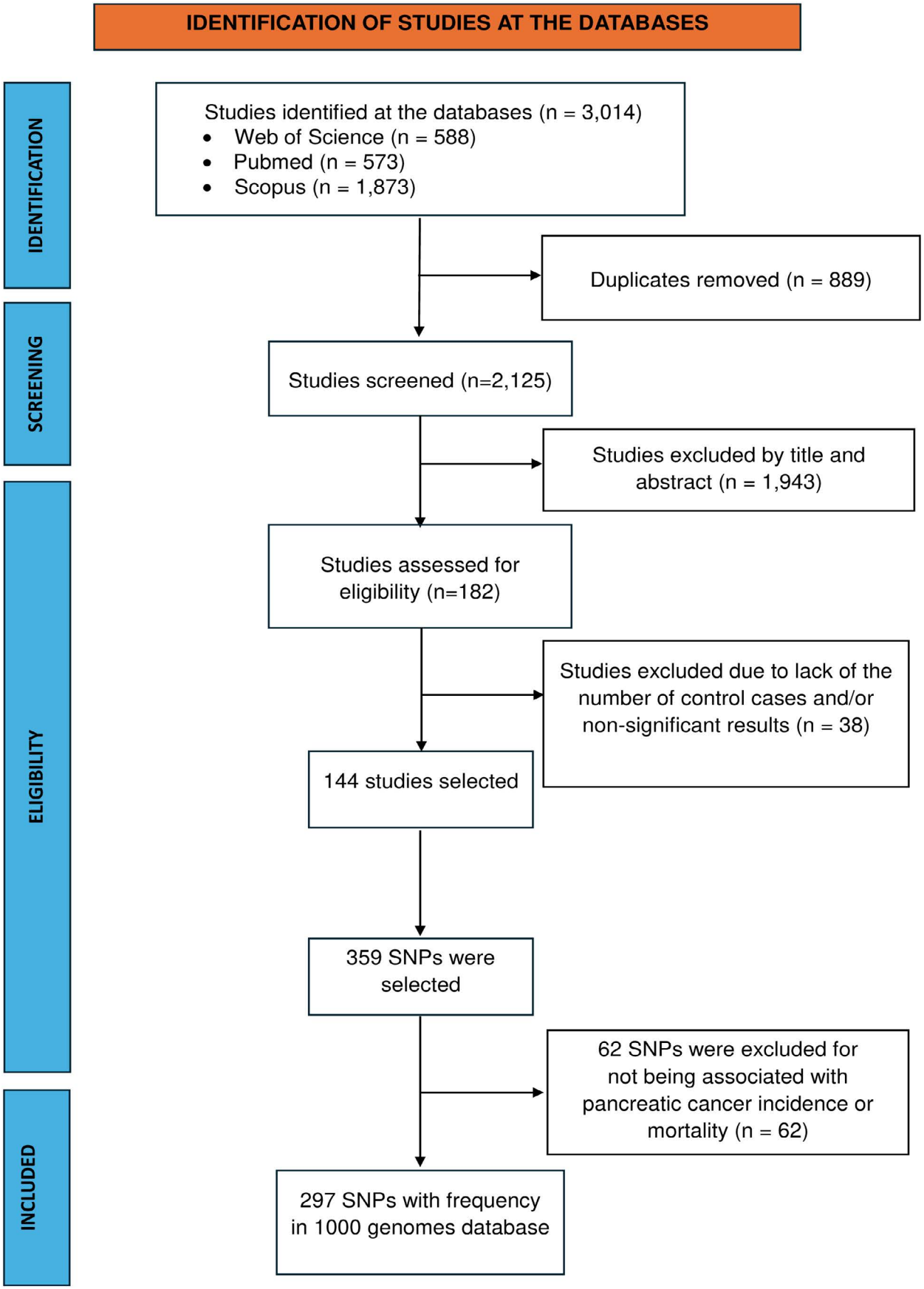

Figure 1. Study selection flowchart
Source: Adapted from PRISMA[8].

The methodological quality and potential risk of bias of the included studies were assessed using the Newcastle-Ottawa Scale (NOS)[10]. Two independent reviewer teams conducted the evaluation, and any discrepancies were resolved after discussions with a third reviewer. The NOS examines three key domains: selection (maximum of 4 stars), comparability (up to 2 stars), and exposure (up to 3 stars), allowing a maximum total score of 9 points per study.

In addition, the Strengthening the Reporting of Genetic Association Studies (STREGA)[11] guidelines were employed to assess the quality of reporting. These recommendations are structured into five core domains: genotyping methods and associated errors, population stratification, haplotype variation, Hardy–Weinberg equilibrium, and replication. The first domain encompasses five distinct criteria: genotyping platform, error and call rates, batch genotyping procedures, genotyping center or laboratory, and the number of successfully genotyped individuals – totaling nine evaluative items. To quantify reporting quality, one point was attributed to each item, resulting in a score ranging from 0 to 9, with higher values reflecting superior reporting rigor. This assessment was independently conducted by three reviewers, discrepancies were resolved in consultation with the senior author.

Pancreatic cancer incidence and mortality data were retrieved from the World Health Organization (WHO) via the Global Cancer Observatory platform[12]. Allele frequency information for genetic variants across continental populations was obtained from the 1000 Genomes Project, focusing on five major ancestry-based groups: Europe (EUR), Africa (AFR), East Asia (EAS), South Asia (SAS), and the Americas (AMR). WHO epidemiological data were compared with genomic data from Phase 3 of the 1000 Genomes Project[13]. Pursuant to the classification used by the 1000 Genomes Project Consortium, populations are clustered based on the predominant ancestral component. Accordingly, the analysis sought to correlate global epidemiological metrics with genetic variant frequencies in EUR, AFR, EAS, SAS, and AMR populations.

Pearson's[14] correlation coefficient was employed to evaluate the relationship between incidence/mortality rates of pancreatic cancer and the allele frequencies of selected variants. The analyses were conducted using the 'cor.test'[15] function from the 'stats' package in the R programming language. Following Bonferroni[16] correction for multiple comparisons, the correlation coefficient (r), coefficient of determination (r^2), *p*-value, and corresponding 95% confidence intervals were reported. All graphical representations were generated using the 'ggplot2'[17] package. Statistical significance was defined as a *p*-value ≤ 0.05.

RESULTS

The Age-Standardized Incidence Rate (ASIR) for pancreatic cancer was analyzed, which corresponds to the frequency at which new cases of the disease are diagnosed within a specific period in different regions or over time. Additionally, the Age-Standardized Mortality Rate (ASMR) for pancreatic cancer was examined, which is calculated based on the number of deaths attributed to the disease within a given period, adjusting for age distribution in a population.

ASIR of pancreatic cancer was 7.8 per 100,000 individuals in the European population, followed by the American population (6.4), the Pacific region (5.1), the African population (2.1), and Southeast Asia (1.3). Similarly, ASMR for pancreatic cancer was also the highest in the European population, at 7.1 per 100,000, followed by the American population (5.4), the Pacific region (4.4), the African population (2.0), and Southeast Asia (1.2). These pancreatic cancer incidence and mortality data grouped by population are shown below in Figure 2.

A total of 253 SNPs were selected from the literature as described in Table 1. Among these single nucleotide polymorphisms, 16 SNPs (rs59519100, rs1785932, rs789744, rs117648907, rs2816938, rs7675998, rs2504938, rs2504956, rs372883, rs10500715, rs3792267, rs2236575, rs872106, rs1045485, rs2268578, rs3792267) were found to be significant according to Pearson's correlation analysis ($p<0.05$).

Among these SNPs, three highly significant variants were identified, with a *p*-value lower than 0.001 (rs2816938, rs372883, rs2236575). Of the three highly significant SNPs, all are associated with mortality, with one of them also showing an association with incidence.

It was also observed that among the 16 significant SNPs, 11 showed higher incidence rates in the European continent (rs1785932, rs789744, rs117648907, rs2816938, rs7675998, rs2504938, rs372883, rs10500715, rs2236575, rs872106, rs1045485), 2 in the Asian population (rs59519100, rs2268578), and three in the American continent (rs3792267, rs2504956, rs3792267).

Of the 15 significant variants related to mortality, three extremely significant variants have been identified: rs2816938, rs372883, and rs2236575. These variants are presented in Figure 3, a graph that correlates the allele frequency of each one in the four populations analyzed in relation to incidence and mortality rates.

Analyzing the correlation among these variants and the epidemiological data for each specific population group, a recurring pattern across all correlations was observed, as they show a directly proportional relationship with the mortality rate. This means that the higher the prevalence of this polymorphism in a population, higher is the observed mortality rate for pancreatic cancer.

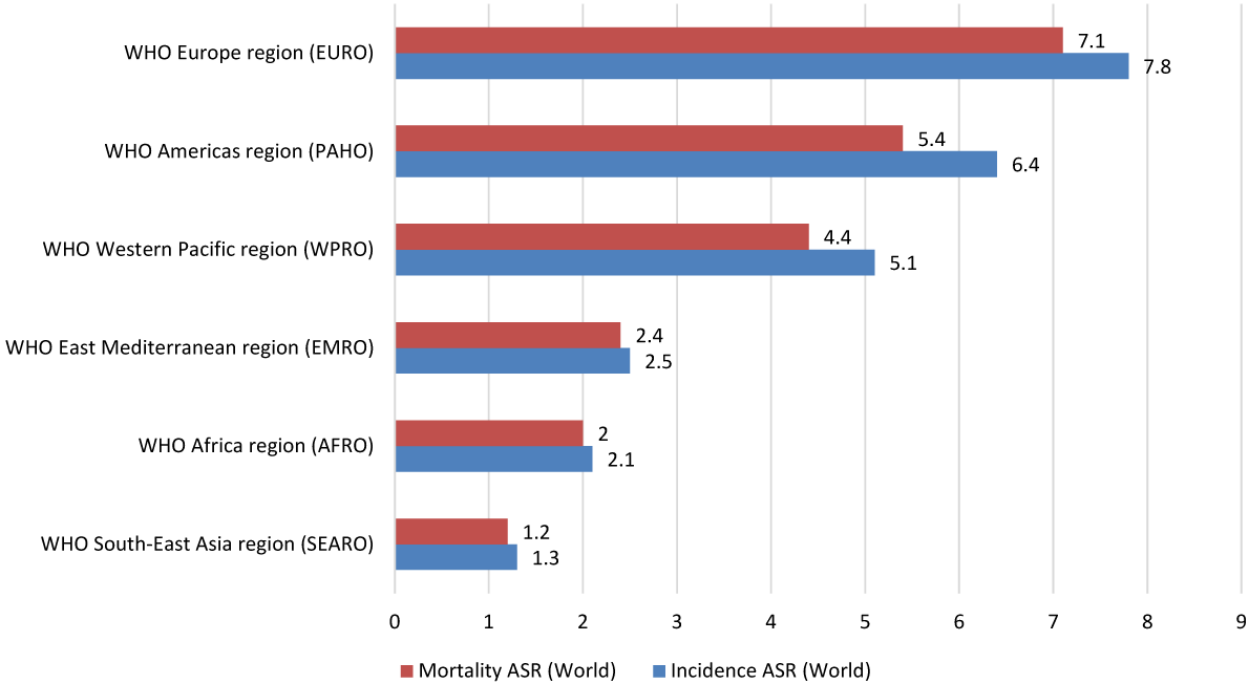

Figure 2. Incidence and mortality rates per 100,000 habitants from the WHO International Agency for Research on Cancer (IARC) in different populations

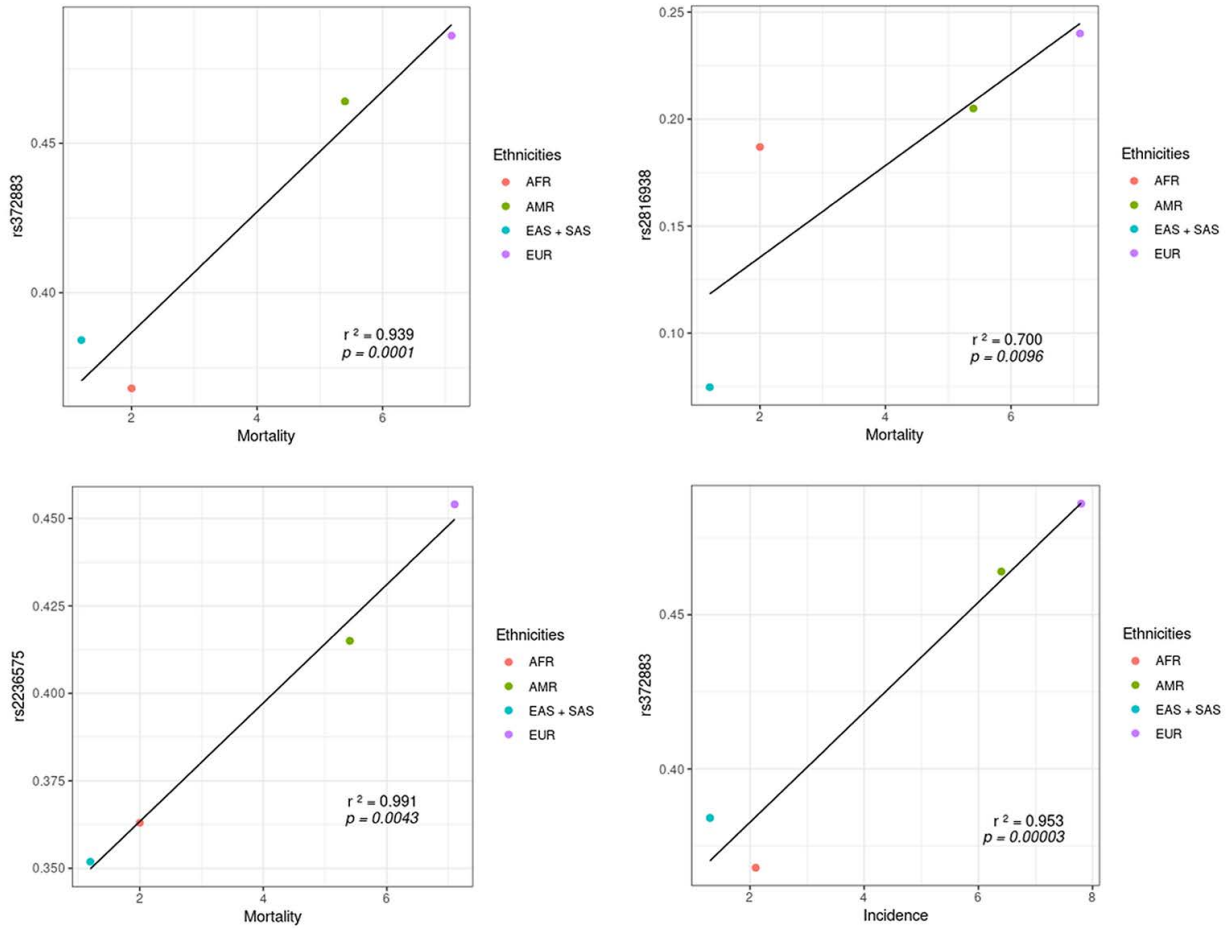

Figure 3. SNPs associated with pancreatic cancer incidence and mortality rates in different populations

Additionally, the American and European populations are at the positive extreme (higher variant frequency and higher mortality), while the Southeast Asian and African populations occupy the opposite extreme (lower variant frequency and lower mortality).

The rs372883 variant was found to be extremely significant regarding the incidence of pancreatic cancer (Figure 3) and directly proportional to allelic expression. In other words, the higher the allele frequency, greater is the susceptibility to the disease in a given population, reinforcing a positive relationship.

The African population presents the lowest values of the rs372883 SNP and one of the lowest incidences of the analyzed condition. Similarly, Asians also exhibit low SNPs values and low incidence. The proximity between these two groups suggests that they may share certain genetic variations that reduce the influence of the rs372883 SNP on disease incidence. The American population exhibits intermediate conditions, while the inhabitants of the European continent show the highest incidence values, suggesting that this SNP has a greater impact on this group.

Overall, among the SNPs analyzed that showed a highly significant and directly proportional relationship with the mortality rate, the rs372883 variant was the most outstanding (p = 0.0000736). The European continent stood out among ethnic groups, exhibiting the highest allelic expression rates of the significant variants in both incidence and mortality, whereas the African continent had the lowest incidence rates across all variants.

Among the studies evaluated, only 13 identified statistically significant SNPs and were consequently included in the risk of bias assessment using two independent evaluation tools. The methodological quality of these 13 studies, as assessed by the NOS[10], is presented in Table 1[18-30]. Overall, NOS scores ranged from four to nine stars, with the majority of studies receiving 8 or 9 stars. All studies were awarded a star in items 1 through 4 of the 'Selection' domain, as well as in items 1 and 2 of the 'Exposure' domain. The 'Comparability' domain, which evaluates whether potential confounding factors among cases and controls were adequately identified and controlled in the analysis, allows a maximum of two stars. Accordingly, all included studies received full points in item 1b of this domain. However, seven studies – namely, Couch et al.[22], Han et al.[23], Huang et al.[24], Zhang et al.[25], Pistoni et al.[28], Shan et al.[29], and Tian et al. – did not receive a star for item 3 in the 'Exposure' domain.

Table 2 presents the results of the reporting quality assessment according to the STREGA guidelines. Total scores ranged from four to nine, with Zhang et al.[25] and Wu et al.[27] achieving the highest overall scores.

DISCUSSION

Pancreatic cancer is one of the most aggressive and lethal cancers, influenced by both genetic and environmental factors. GWAS have identified SNPs associated with susceptibility to pancreatic cancer, particularly the variants rs2236575, rs372883, and

Table 1. Methodological quality of the studies based on NOS

Studies	Selection				Comparability		Exposure			Total Score
	Item 1	Item 2	Item 3	Item 4	Item 1a	Item 1b	Item 1	Item 2	Item 3	
Zhao et al.[18]	*	*	*	*	*	*	*	*	*	9
Campa et al.[19]	*	*	*	*	*	*	*	*	*	9
Mohelnikova-Duchonova et al.[20]	*	*	*	*	*	*	*	*	*	9
Wu et al.[21]	*	*	*	*	*	*	*	*	*	9
Couch et al.[22]	*	*	*	*	*	*	*	*		8
Han et al.[23]	*	*	*	*	*	*	*	*		8
Huang et al.[24]	*	*	*	*	*	*	*	*		8
Zhang et al.[25]	*	*	*	*	*	*	*	*		8
Fong et al.[26]	*	*	*	*	*	*	*	*	*	9
Wu et al.[27]	*	*	*	*	*	*	*	*	*	9
Pistoni et al[28]	*	*	*	*	*	*	*	*		8
Shan et al.[29]	*	*	*	*	*	*	*	*		8
Tian et al.[30]	*	*	*	*	*	*	*	*		8

Captions: *Selection – Item 1: Is the case definition adequate?; Item 2: Representativeness of the cases; Item 3: Selection of controls; Item 4: Definition of controls. Comparability – Item 1a and 1b: Comparability of cases and controls based on the design or analysis. Exposure – Item 1: Ascertainment of exposure; Item 2: Same method of ascertainment for cases and controls; Item 3: Non-response rate.

Table 2. The quality of reporting using the STREGA guideline

Studies	Genotyping methods and platforms	Error rates and call rates	Genotyping in batches	Laboratory/ center where the genotyping was done	Number of successfully genotyped individuals	Description of modeling population stratification	Description of modeling haplotype variation	Hardy-Weinberg equilibrium was considered	Statement of whether the study is the first report of a genetic association, a replication effort, or both	Scores
Zhao et al.[18]	No	No	No	No	Yes	Yes	No	Yes	No	2
Campa et al.[19]	No	No	No	No	Yes	Yes	No	Yes	No	2
Mohelnikova-Duchonova et al.[20]	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	7
Wu et al.[21]	Yes	No	No	No	Yes	Yes	No	Yes	Yes	5
Couch et al.[22]	Yes	Yes	No	Yes	Yes	Yes	No	Yes	No	6
Han et al.[23]	Yes	Yes	No	No	Yes	No	No	No	Yes	4
Huang et al.[24]	Yes	No	No	Yes	Yes	No	No	Yes	Yes	4
Zhang et al.[25]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	8
Fong et al.[26]	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes	7
Wu et al.[27]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	8
Pistoni et al.[28]	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	6
Shan et al.[29]	Yes	No	No	Yes	Yes	Yes	No	Yes	No	5
Tian et al.[30]	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	6

rs2816938, linked to the ERCC1, BACH1, and NR5A2 genes, respectively. These genes play essential roles in biological mechanisms that influence carcinogenesis, such as DNA damage response, epithelial-mesenchymal transition (EMT), and metabolic regulation[25].

The ERCC1 gene (Excision Repair Cross-Complementation Group 1) is involved in nucleotide excision repair (NER), a critical process for maintaining genomic integrity[31], which consists of several essential steps, including DNA damage recognition, unwinding of the double helix in the affected region, making cuts on the DNA strands before and after the damaged site, and subsequently repairing the generated gap[32]. The variant rs2236575 affects the efficiency of the ERCC1-XPF complex, compromising the removal of DNA damage induced by mutagenic agents. This leads to the accumulation of mutations that can promote genomic instability, a key factor in the development of pancreatic neoplasms[24]. The incidence of this variant is higher in European populations, where elevated rates of pancreatic cancer incidence and mortality are also observed. Thus, failure in the DNA damage response directly contributes to tumor initiation and progression[33].

The BACH1 gene (BTB and CNC Homology 1) regulates the EMT, a process crucial for cancer progression and metastasis. The variant rs372883 has been identified as a critical factor in regulating EMT, acting in the repression of epithelial genes as FOXA1 and activation of mesenchymal genes as SNAI2. This alteration leads to loss of cell adhesion, increased mobility, and tumor invasion[30]. The presence of rs372883 is more common in European and American populations, where more aggressive disease is observed. Studies indicate that high expression of BACH1 is also associated with a more glycolytic metabolism, favoring tumor growth in hypoxic environments and contributing to the progression of pancreatic ductal adenocarcinoma (PDAC)[27].

Interestingly, although the presence of the variant rs2816938 is higher in European and American populations, mortality is higher in Asian and African populations, suggesting that other environmental or genetic factors may modulate its protective effect. Ethnic distribution data reveal that African populations have the lowest mortality and the lowest frequency of the SNP rs372883, while European populations have the highest mortality associated with the high frequency of this genetic marker. American and Asian populations fall in intermediate positions on this spectrum, suggesting a possible interaction between genetic factors and predisposition to specific health conditions[26].

In the Brazilian context, the applicability of SNPs in assessing pancreatic cancer susceptibility must be considered within the framework of the country's unique genetic composition. Recognized as the most ethnically

admixed nation globally, Brazil harbors a genomic background shaped by contributions from Indigenous peoples, Europeans, Africans, and Asians[34]. This extensive admixture emphasizes the imperative to broaden access to genetic testing within the Brazilian National Health System (SUS), where such services remain scarce. The need is further underscored by the fact that most SNPs significantly associated with pancreatic cancer incidence and mortality have been identified primarily in European populations. Addressing this gap is therefore not only a matter of scientific relevance but also a pressing public health priority for Brazil.

The analysis of the graph correlating the allele frequency of SNP rs372883 with pancreatic cancer incidence reveals a strong positive relationship (r^2 = 0.953; p = 0.00003), indicating that 95.3% of the variation in disease incidence across populations can be explained by the prevalence of this polymorphism[20]. The European population has the highest allele frequency (0.486) and incidence (7.8 cases per 100,000), followed by the American population (AMR: 0.464; 6.4 cases per 100,000). In contrast, African (AFR: 0.368; 2.1 cases per 100,000) and South Asian (SAS: 0.308; 1.3 cases per 100,000) populations correlate with lower neoplasia rates, reinforcing the ethnic-epidemiological gradient[27].

Despite the East Asian population (EAS: 0.458) exhibiting an intermediate allele frequency, the incidence in the region remains low (1.3 cases per 100,000), suggesting that other protective genetic or environmental factors may mitigate the risk associated with rs372883 in this population[23]. The graphical distribution shows that ancestral genetic factors contribute to the disparity in disease incidence. The observed pattern suggests that rs372883, linked to BACH1 regulation, may influence susceptibility to pancreatic cancer, partially explaining the higher mortality in populations with a high allele frequency (EUR and AMR) as proposed by Wu C, et al.[27].

The gene NR5A2 (Nuclear Receptor Subfamily 5 Group A Member 2) plays a fundamental role in maintaining the differentiation of acinar cells, which are responsible for the production of digestive enzymes[35]. Experimental studies have demonstrated that loss of NR5A2 expression in acinar cells accelerates acinar-to-ductal metaplasia (ADM), particularly in the presence of oncogenic mutations in the KRAS gene[35]. ADM is considered one of the early steps in the formation of precursor lesions of PDAC[5].

From an epidemiological perspective, genetic variants in NR5A2 have been associated with modulation of pancreatic cancer risk across different populations. Polymorphisms such as rs3790844 and rs2816938 have shown an inverse relationship with disease incidence, suggesting a protective effect[27,36]. These SNPs may influence the expression and transcriptional activity of NR5A2, preserving the structural and functional integrity of pancreatic tissue and reducing susceptibility to chronic inflammatory processes – a known risk factor for PDAC[33]. Conversely, reduced NR5A2 expression may create a pro-inflammatory microenvironment, favoring tumor initiation and progression[24,30]. Therefore, considering the high lethality of pancreatic cancer and its low early detection rate[1], the study of polymorphisms in NR5A2 may have relevant implications for personalized screening, risk biomarker identification, and the development of novel targeted therapeutic approaches.

The analysis of the distribution of these polymorphisms among different populations reveals distinct patterns of susceptibility and prognosis for pancreatic cancer. While variants such as rs2236575 and rs372883 increase the risk and aggressiveness of the disease through mechanisms related to genomic instability and EMT, the variant rs2816938 may play a protective role by preserving cellular homeostasis.

The identification of these genetic variants offers a promising field for advances in precision medicine. In this context, the global burden of the disease, combined with its high lethality, reinforces the importance of more targeted preventive strategies[1]. Identified SNPs can be incorporated into prediction models for personalized screening in high-risk individuals, especially in populations with a family history or exposure to relevant environmental factors[3,4]. Furthermore, certain variants have shown potential impact on therapeutic response, paving the way for targeted therapies and individualized selection of chemotherapy regimens[24,32]. Thus, the integration of these genetic markers into clinical programs can optimize early diagnosis, improve prognostic stratification, and maximize therapeutic efficacy.

These findings reinforce the influence of genetic factors on the incidence and tumor progression, highlighting the need for personalized approaches in the prevention and treatment of pancreatic cancer[29]. This study does not primarily focus on discussing population diversity and allele frequency distributions; therefore, it is limited with respect to a formal quantitative analysis of the heterogeneity of the populations evaluated and the variables reported. All analyses were conducted using secondary data extracted from previously published studies, which implies dependence on the methodological quality and standardization adopted by each author. Despite these limitations, the investigation of genetic variants associated with the incidence and mortality of

pancreatic cancer is a highly relevant area for clinical medicine. Therefore, the necessity of further studies is reinforced in order to expand and consolidate the knowledge in this field.

CONCLUSION

The present article developed an epidemiological analysis of genetic variants associated with the incidence and mortality rates of pancreatic cancer across different continents and ethnicities worldwide, highlighting the significant SNPs for a possible predisposition to this type of cancer. Based on the studies analyzed in this article, different GWAS studies were correlated, demonstrating that the SNPs (rs2816938, rs2236575 and rs372883) were significantly and positively associated with the incidence and mortality of pancreatic cancer in different ethnic groups. In this way, the discovery of these risk loci contributes to improving studies on prevention based on incidence, prognosis, toxicity, and targeted treatments for patients, enhancing the understanding of the genetic basis of pancreatic cancer.

ACKNOWLEDGEMENTS

This is an original work developed by the "*Faculdade de Medicina da Universidade Federal do Pará*".

CONTRIBUTIONS

All the authors contributed substantially to the study design, acquisition, analysis and interpretation of the data, writing and critical review. They approved the final version to be published.

DECLARATION OF CONFLICT OF INTERESTS

There is no conflict of interests to declare.

DATA AVAILABILITY STATEMENT

All content underlying the text is contained in the manuscript.

FUNDING SOURCES

"*CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior*".

REFERENCES

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74(3):229-63. doi: https://doi.org/10.3322/caac.21834
2. SEER: Surveillance, Epidemiology, and End Results [Internet]. Bethesda: National Cancer Institute; 2024. [cited 2024 Oct 26]. Available from: https://seer.cancer.gov/statfacts/html/pancreas.html
3. Yang Y, Shi J, Huang J, et al. Case-control trials on risk factors for pancreatic cancer: a systematic review and meta-analysis. Iran J Public Health. 2023;52(8):1578-88. doi: https://doi.org/10.18502/ijph.v52i8.13397
4. Abdellaoui A, Dolan CV, Verweij KJH, et al. Gene-environment correlations across geographic regions affect genome-wide association studies. Nat Genet. 2022;54(9):1345-54. doi: https://doi.org/10.1038/s41588-022-01158-0
5. Ye ZM, Li LJ, Luo MB, et al. A systematic review and network meta-analysis of single nucleotide polymorphisms associated with pancreatic cancer risk. Aging (Albany NY). 2020;12(24):25256-74. doi: https://doi.org/10.18632/aging.104128
6. Aoki MN, Stein A, Carvalho J, et al. Susceptibility loci for pancreatic cancer in the Brazilian population. BMC Med Genomics. 2021;14(1):111. doi: https://doi.org/10.1186/s12920-021-00956-5
7. Campa D, Gentiluomo M, Stein A, et al. The PANcreatic Disease ReseArch (PANDoRA) consortium: ten years' experience of association studies to understand the genetic architecture of pancreatic cancer. Crit Rev Oncol Hematol. 2023;186:104020. doi: https://doi.org/10.1016/j.critrevonc.2023.104020
8. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. doi: https://doi.org/10.1136/bmj.n71
9. University of York. Centre for Reviews and Dissemination [Internet]. New York: University of York; 2019. PROSPERO - International prospective register of systematic reviews. 2023. [cited 2025 ago 31]. Available from: https://www.crd.york.ac.uk/PROSPERO/view/CRD420251067640
10. Wells GA, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses [Internet]. Ottawa: Ottawa Hospital Research Institute; 2009 [cited 2024 Oct 25]. Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp
11. Little J, Higgins JP, Ioannidis JP, et al. STrengthening the REporting of Genetic Association Studies (STREGA): an extension of the STROBE statement. PLoS Med. 2009;6(2):e22. doi: https://doi.org/10.1371/journal.pmed.1000022
12. GCO: Global Cancer Observatory [Internet]. Lyon: IARC; c1965-2022. [2024 Oct 2025]. Available from: https://gco.iarc.fr/

13. IGSR: The International Genome Sample Resource [Internet]. Cambridge: EMBL-EBI; ©2008-2025. [cited 2024 Oct 4]. Available from: https://www.internationalgenome.org/

14. Henley SS, Golden RM, Kashner TM. Statistical modeling methods: challenges and strategies. Biostat Epidemiol. 2020;4(1):105-39. doi: https://doi.org/10.1080/24709360.2019.1618653

15. Cor.test: test of association/correlation [Internet]. Vienna: R Foundation for Statistical Computing; 2025. [cited 2024 Oct 4]. Available from: https://stat.ethz.ch/R-manual/R-devel/library/stats/html/cor.test.html

16. Bland JM. Multiple significance tests: the Bonferroni method. BMJ. 1995;310(6973):170. doi: https://doi.org/10.1136/bmj.310.6973.170

17. Wickham H. ggplot2: create elegant data visualisations using the grammar of graphics [Internet]. 3.5.1 ed. Houston: Posit Software, PBC; 2025 [cited 2024 Oct 4]. Available from: https://ggplot2.tidyverse.org

18. Zhao L, Liu H, Luo S, et al. Associations between genetic variants of KIF5B, FMN1, and MGAT3 in the cadherin pathway and pancreatic cancer risk. Cancer Med. 2020;9(24):9620-31. doi: https://doi.org/10.1002/cam4.3603

19. Campa D, Matarazzi M, Greenhalf W, et al. Genetic determinants of telomere length and risk of pancreatic cancer: a PANDoRA study. Inter J. Cancer. 2019;144(6):1275-83. doi: https://doi.org/10.1002/ijc.31928

20. Mohelnikova-Duchonova B, Krajsova I, Holubec L, et al. SLC22A3 polymorphisms do not modify pancreatic cancer risk, but may influence overall patient survival. Sci Rep. 2017;7(1):9426. doi: https://doi.org/10.1038/s41598-017-10049-y

21. Wu Chen, Kraft P, Stolzenberg-Solomon R, et al. Genome-wide association study of survival in patients with pancreatic adenocarcinoma. Gut. 2014;63(1):152-60. doi: https://doi.org/10.1136/gutjnl-2012-303477

22. Couch FJ, Wang X, McWilliams RR, et al. Association of breast cancer susceptibility variants with risk of pancreatic cancer. Cancer Epidemiol Biomarkers Prev. 2009;18(11):3044-8. doi: https://doi.org/10.1158/1055-9965.EPI-09-0306

23. Han Y, Jung KJ, Kim U, et al. Non-invasive biomarkers for early diagnosis of pancreatic cancer risk: metabolite genomewide association study based on the KCPS-II cohort. J Transl Med. 2023;21(1):878. doi: https://doi.org/10.1186/s12967-023-04670-x

24. Huang X, Zheng J, Li J, et al. Functional role of BTB and CNC homology 1 gene in pancreatic cancer and its association with survival in patients treated with gemcitabine. Theranostics. 2018;8(12):3366-79. doi: https://doi.org/10.7150/thno.23978

25. Zhang M, Wang Z, Obazee O, et al. Three new pancreatic cancer susceptibility signals identified on chromosomes 1q32.1, 5p15.33 and 8q24.21. Oncotarget. 2016;7(41):66328-43. doi: https://doi.org/10.18632/oncotarget.11041

26. Fong PY, Lee SH, Chan WC, et al. Association of diabetes susceptibility gene Calpain-10 with pancreatic cancer among smokers. J Gastrointest Cancer. 2010;41(3):203-8. doi: https://doi.org/10.1007/s12029-010-9130-7

27. Wu C, Miao X, Huang L, et al. Genome-wide association study identifies five loci associated with susceptibility to pancreatic cancer in Chinese populations. Nat Genet. 2011;44(1):62-6. doi: https://doi.org/10.1038/ng.1020

28. Pistoni L, Gentiluomo M, Lu Y, et al. Associations between pancreatic expression quantitative traits and risk of pancreatic ductal adenocarcinoma. Carcinogenesis. 2021;42(8):1037-45. doi: https://doi.org/10.1093/carcin/bgab057

29. Shan YS, Hsu HP, Lai MD, et al. Validation of genome-wide association study-identified single nucleotide polymorphisms in a case-control study of pancreatic cancer from Taiwan. J Biomed Sci. 2011;18:64. doi: https://doi.org/10.1186/1423-0127-18-64

30. Tian J, Chen C, Rao M, et al. Aberrant RNA splicing is a primary link between genetic variation and pancreatic cancer risk. Cancer Res. 2022;82(11):2084-96. doi: https://doi.org/10.1158/0008-5472.can-21-4367

31. He MG, Zheng K, Tan D, et al. Association between ERCC1 and ERCC2 gene polymorphisms and susceptibility to pancreatic cancer. Genet Mol Res. 2016;15(1). doi: https://doi.org/10.4238/gmr.15017879

32. McWilliams RR, Bamlet WR, Andrade M, et al. Nucleotide excision repair pathway polymorphisms and pancreatic cancer risk: evidence for role of MMS19L. Cancer Epidemiol Biomarkers Prev. 2009;18(4):1295-302. doi: https://doi.org/10.1158/1055-9965.epi-08-1109

33. Dai P, Li J, Li W, et al. Genetic polymorphisms and pancreatic cancer risk: a PRISMA-compliant systematic review and meta-analysis. Medicine (Baltimore). 2019;98(32):e16541. doi: https://doi.org/10.1097/md.0000000000016541

34. Nunes K, Pereira R, Santos C, et al. Admixture's impact on Brazilian population evolution and health. Science. 2025;388:eadl3564. doi: https://doi.org/10.1126/science.adl3564

35. Ueno M, Hosono S, Nakata B, et al. Genome-wide association study-identified SNPs (rs3790844, rs3790843) in the NR5A2 gene and risk of pancreatic cancer in Japanese. Sci Rep. 2015;5:17018. doi: https://doi.org/10.1038/srep17018

36. López de Maturana E, Rodríguez JA, Alonso L, et al. A multilayered post-GWAS assessment on genetic susceptibility to pancreatic cancer. Genome Med. 2021;13(1):15. doi: https://doi.org/10.1186/s13073-020-00816-4

Recebido em 15/7/2025
Aprovado em 27/8/2025

Associate editor: Fernando Lopes Tavares de Lima. Orcid iD: https://orcid.org/0000-0002-8618-7608
Scientific-editor: Anke Bergmann. Orcid iD: https://orcid.org/0000-0002-1972-8777