

CHARACTERIZATION OF *CDK7* GENE VARIANTS ASSOCIATED WITH CANCER IN AMERINDIANS OF THE BRAZILIAN AMAZON

Amanda de Nazaré Cohen-Paes¹, Darlen Cardoso de Carvalho³, Elisa da Silva Meneses², Kezia Nascimento Góes¹, Débora Monteiro Carneiro do Vale¹, Diego Di Felipe Ávila Alcântara¹, Marianne Rodrigues Fernandes², João Farias Guerreiro², Rommel Mario Rodriguez Burbano^{2,3}, Ney Pereira Carneiro dos Santos²

¹ Laboratório de Biologia Molecular do Hospital Ophir Loyola, Ophir Loyola Hospital, Belém, Pará, Brazil

²Oncology Research Nucleus, Universidade Federal do Pará, Belém, Pará, Brazil

³Faculdade de Medicina Estácio de Castanhal, Castanhal, Pará, Brazil

Introduction: Cyclin-dependent kinase 7 (CDK7) is frequently overexpressed in various malignancies due to its crucial role in transcription regulation and cell cycle progression. Moreover, CDK7 has emerged as a promising therapeutic target in cancer treatment. Amerindian populations from the Brazilian Amazon possess a distinct genetic profile characterized by low genetic diversity compared to other continental populations. Investigating these populations not only enhances our understanding of human population genetics but also contributes to the development of more precise diagnostic and therapeutic strategies, improving clinical outcomes for underrepresented groups.

Objectives: This study aimed to characterize four CDK7 gene variants —rs2972388, rs12657418, rs2291519, and rs2972367— in an Amazonian Amerindian cohort and to compare their frequencies with those of global populations using data from the Genome Aggregation Database (gnomAD). **Methods:** The study was approved by the National Committee for Ethics in Research (CAAE: 20654313.6.0000.5172). The cohort consisted of 64 Amerindian individuals from the Brazilian Amazon (referred to as NAT). Genomic DNA was extracted from peripheral blood using the phenol-chloroform method. Whole-exome sequencing was conducted on the Illumina NextSeq500® platform. Variant frequencies in the NAT group were compared to those in five continental populations from gnomAD: African (AFR, n=8,128), European non-Finnish (EUR, n=56,885), American (AMR, n=17,296), East Asian (EAS, n=9,197), and South Asian (SAS, n=15,308). Statistical analyses were performed using R Studio version 3.5.1. **Results:** This is the first study to describe CDK7 variants in Amazonian Amerindians. Our data also details the chromosomal location, alleles, mutation type and effect, and frequency of each variant in both the NAT population and comparative continental groups. Notably, rs2972388 showed a significantly different distribution in NAT compared to AFR, being more frequent in the African population ($p<0.001$). However, its frequency in the Amerindian group (0.694) was the highest among all continental populations studied. This variant has been previously associated with increased risk for Differentiated Thyroid Cancer (DTC), which

accounts for over 80% of thyroid cancer cases and exhibits high incidence rates in Brazil.

Conclusion: Understanding the prevalence of CDK7 variants such as rs2972388 in Amerindian populations is essential for shaping public health strategies, including cancer screening and prevention, especially in genetically admixed populations like those in Brazil.

Keywords: CDK7; Amerindian populations; cancer susceptibility.