

# Virome characterization in different types of non-Hodgkin lymphoma

Juliana D. Siqueira<sup>1\*</sup>, Marcelo A. Soares<sup>1,2</sup>, Esmeralda A. J. M. Soares<sup>1</sup>

<sup>1</sup>Programa de Oncovirologia, Instituto Nacional de Câncer, Rio de Janeiro, RJ, Brazil; <sup>2</sup>Departamento de Genética, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil; \*Postdoctoral researcher at Instituto Nacional de Câncer

## INTRODUCTION

Non-Hodgkin's lymphoma (NHL) is a heterogeneous group of lymphoid tissue neoplasms, with varied clinical, genetic and morphological characteristics. This lymphoma is the most common form of hematological malignancy worldwide. In addition to behavioral and environmental risk factors, some infectious agents cause or are related to the risk of developing NHL. These include the Epstein Barr virus (EBV), human herpes virus 8 (HHV-8), human T-cell lymphotropic virus type 1 (HTLV-1), human immunodeficiency virus (HIV), hepatitis C virus (HCV), and *Helicobacter pylori*.

The areas with the highest incidence of NHL are North America, Europe, Oceania and some countries of the African continent (Fig. 1). More than 10,000 new cases of NHL are estimated for Brazil in 2018.

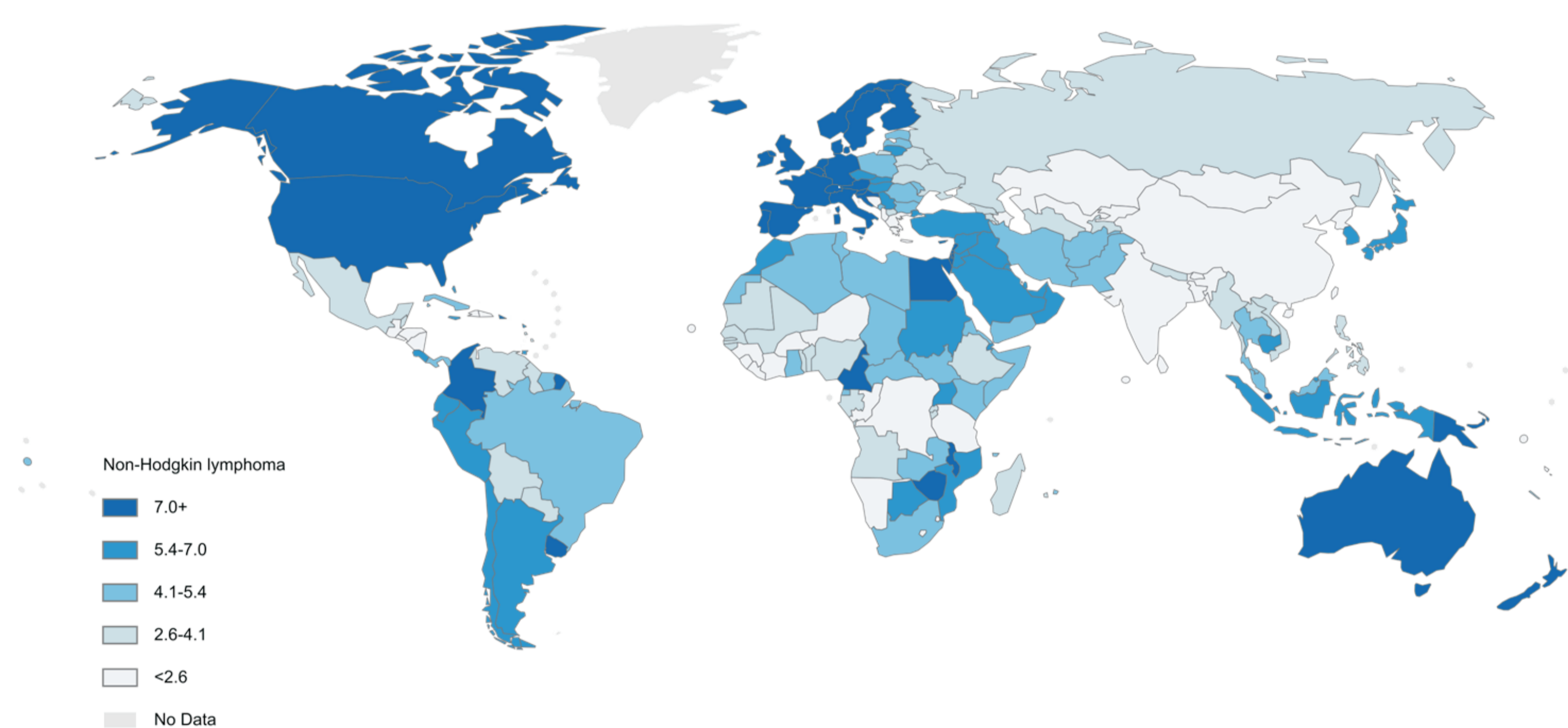


Figure 1. Estimated non-Hodgkin lymphoma incidence worldwide in 2012. Age-standardized rate per 100,000 (extract from Ferlay J et al., 2013).

Recently, a novel species of parvovirus was identified in Mycosis fungoides, a cutaneous T-cell NHL. Despite the low prevalence (5.4%), the infection was confirmed by *in situ* hybridization. However, the influence of this virus on cancer development and its prevalence has not yet been elucidated.

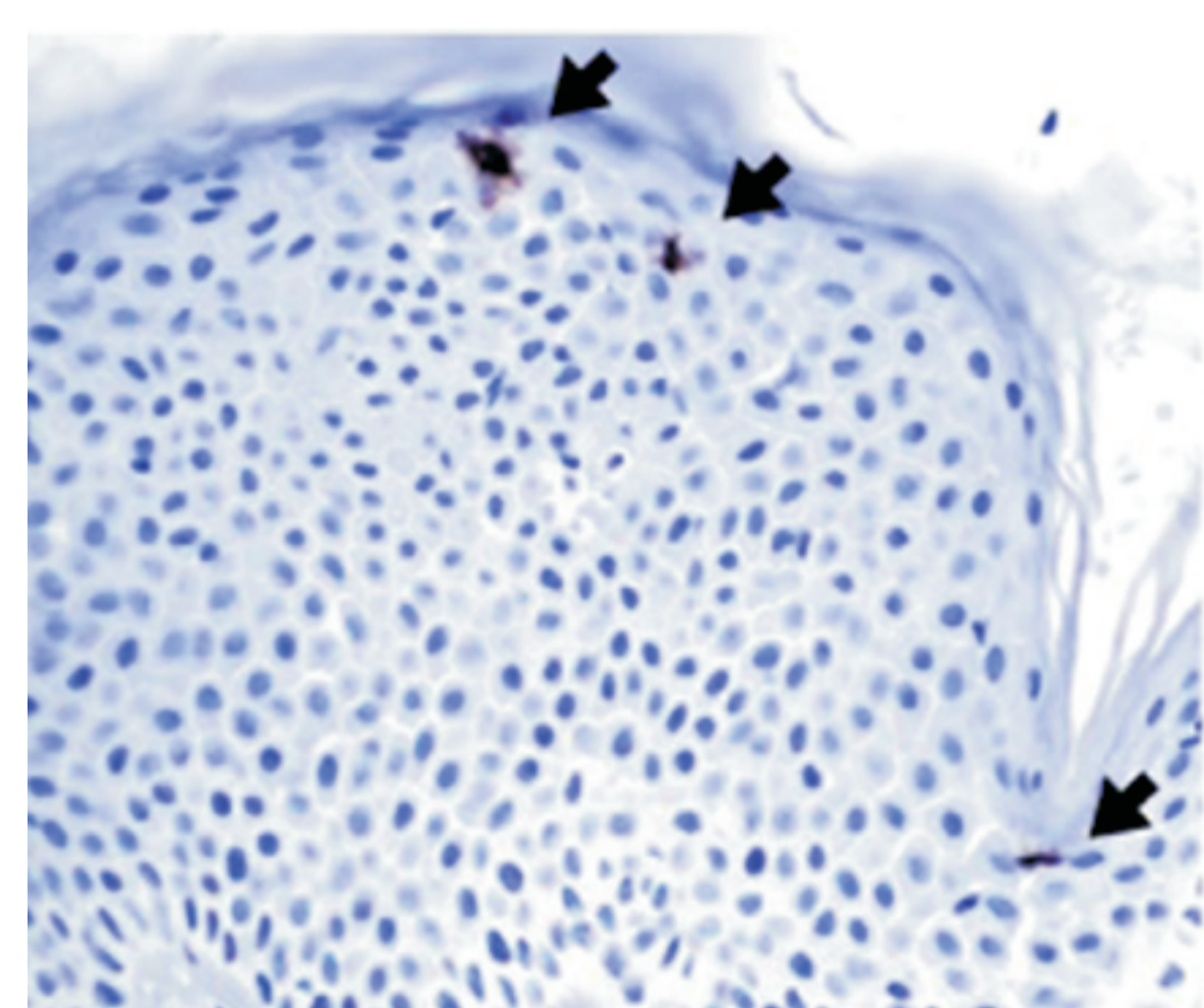


Figure 2. Cutavirus DNA detectable within 3 cells. Arrows indicate *in situ* hybridization cutavirus probe signal (extracted from Phan et al., 2016).

Although factors related to altered immune function and several infectious agents appear to play a role in the etiology of NHL, these factors do not explain the majority of the cases. This suggests that others factors or infectious agents could be involved in the NHL development. Identifying the cancer-infectious agent relationship is important for the development of cancer prevention and early diagnosis. Therefore, the aim of this study is to characterize, through metagenomics analysis, the viral composition of the NHL tumor tissue in the different types of NHL.

## METHODS

This study has been approved by the Ethics Research Committee of Instituto Nacional de Câncer (INCA).

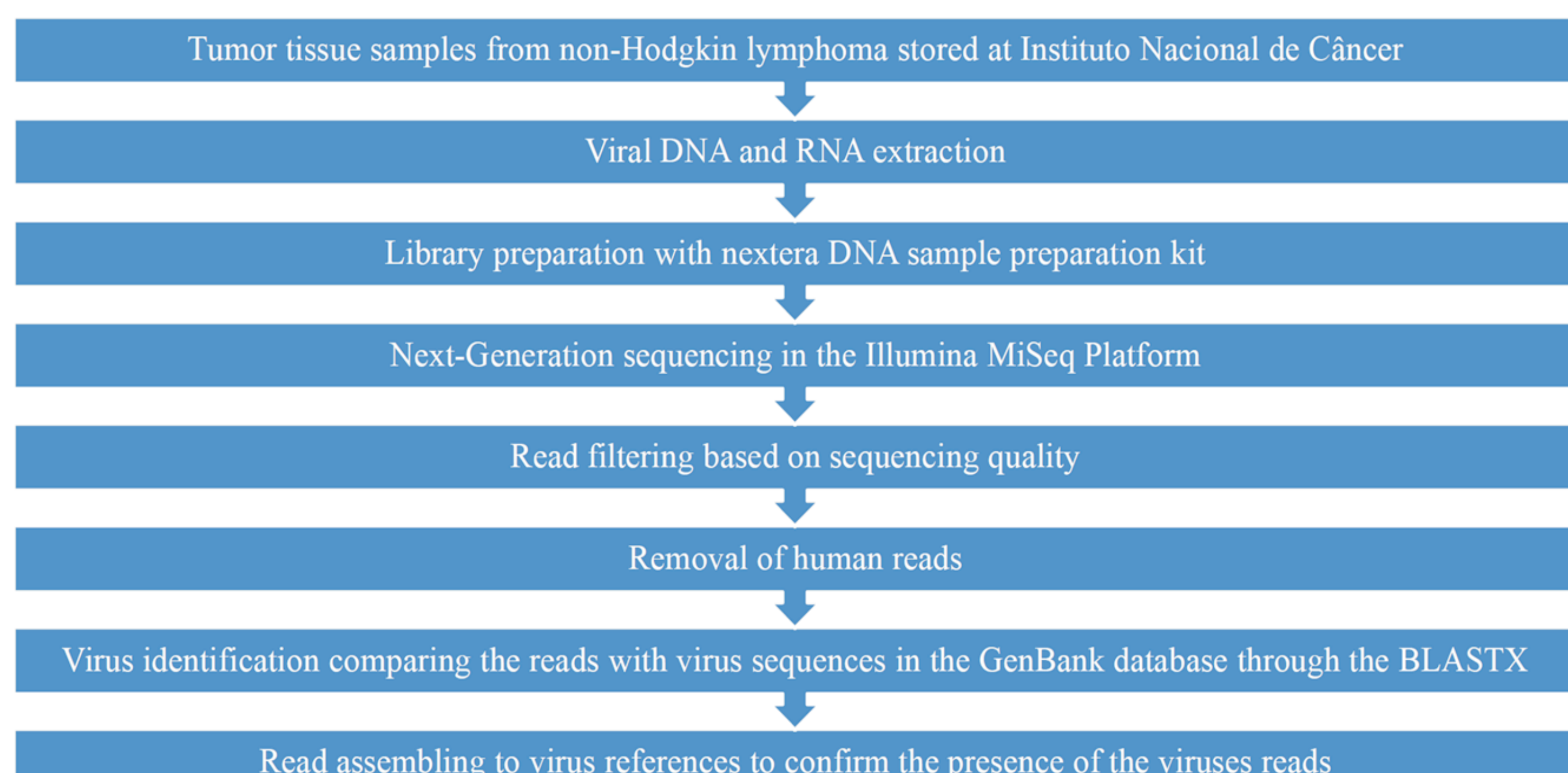


Figure 3. Flowchart of the methodology developed in the study.

Currently, 15 samples have been sequenced at least one time and presented over 400,000 reads (Table 1). A great percentage of reads obtained mapped to human genome and were discarded (Table 1).

Table 1. Distribution of the total and human reads obtained in the sequencing

Sample	Total number of reads	Human reads (%)
L01	404,100	262,698 (65)
L02	1,789,814	63,044 (4)
L03	1,999,132	1,022,992 (51)
L06	2,255,326	1,108,472 (49)
L07	2,171,246	951,007 (44)
L09	1,704,342	60,933 (4)
L10	1,928,742	886,906 (46)
L11	2,090,714	17,628 (1)
L12	1,929,148	581,289 (30)
L13	2,198,648	1,162,187 (53)
L15	1,326,104	764,987 (58)
L19	3,598,652	2,262,112 (63)
L20	454,990	209,992 (46)
L21	9,414,708	6,230,646 (66)
L24	6,414,128	5,300,109 (83)

Until now, 11 samples have been analyzed comparing by BLASTX the remaining reads to the Genbank non-redundant protein database. We found six viral families that infects vertebrates: *Anelloviridae*, *Flaviviridae*, *Hepadnaviridae*, *Herpesviridae*, *Papillomaviridae* and *Retroviridae* (Fig. 5).

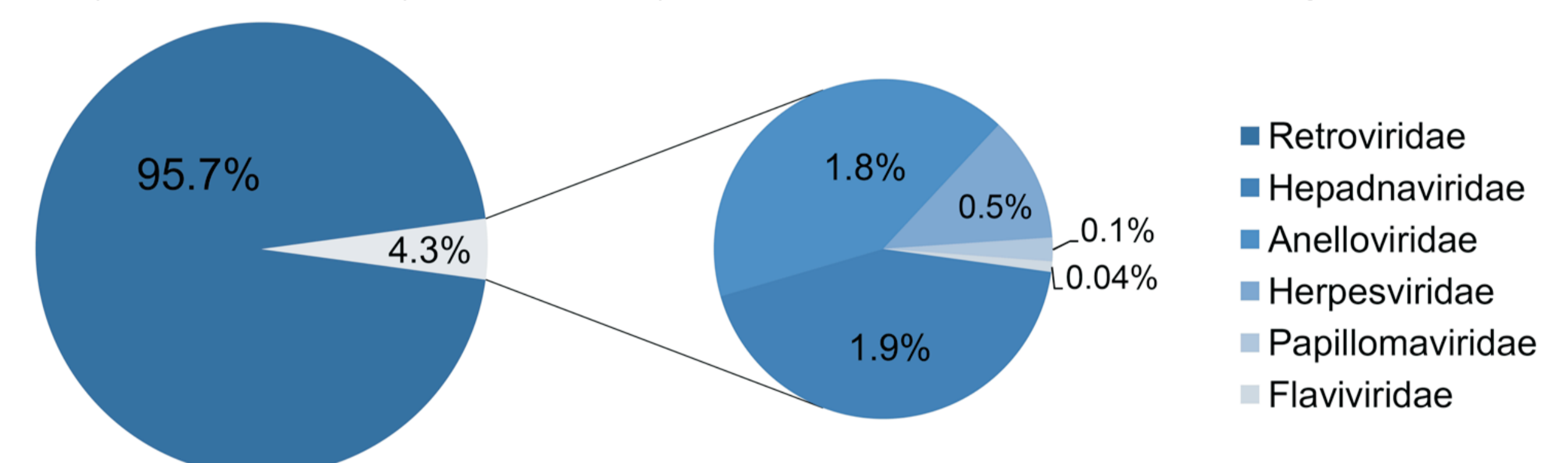


Figure 4. Distribution of sequence reads generated by viral metagenomics with translated protein sequence similarity to different viral families.

Reads from *Retroviridae* and *Hepadnaviridae* were identified as contaminant from samples from others projects sequenced in the same experiments and excluded from further analyses.

The reads from each sample were mapped to references from the viral families found and confirmed the presence of reads of EBV, human papillomavirus (HPV), human pegivirus (HPgV) and torque tenovirus (TTV) (Fig. 6).

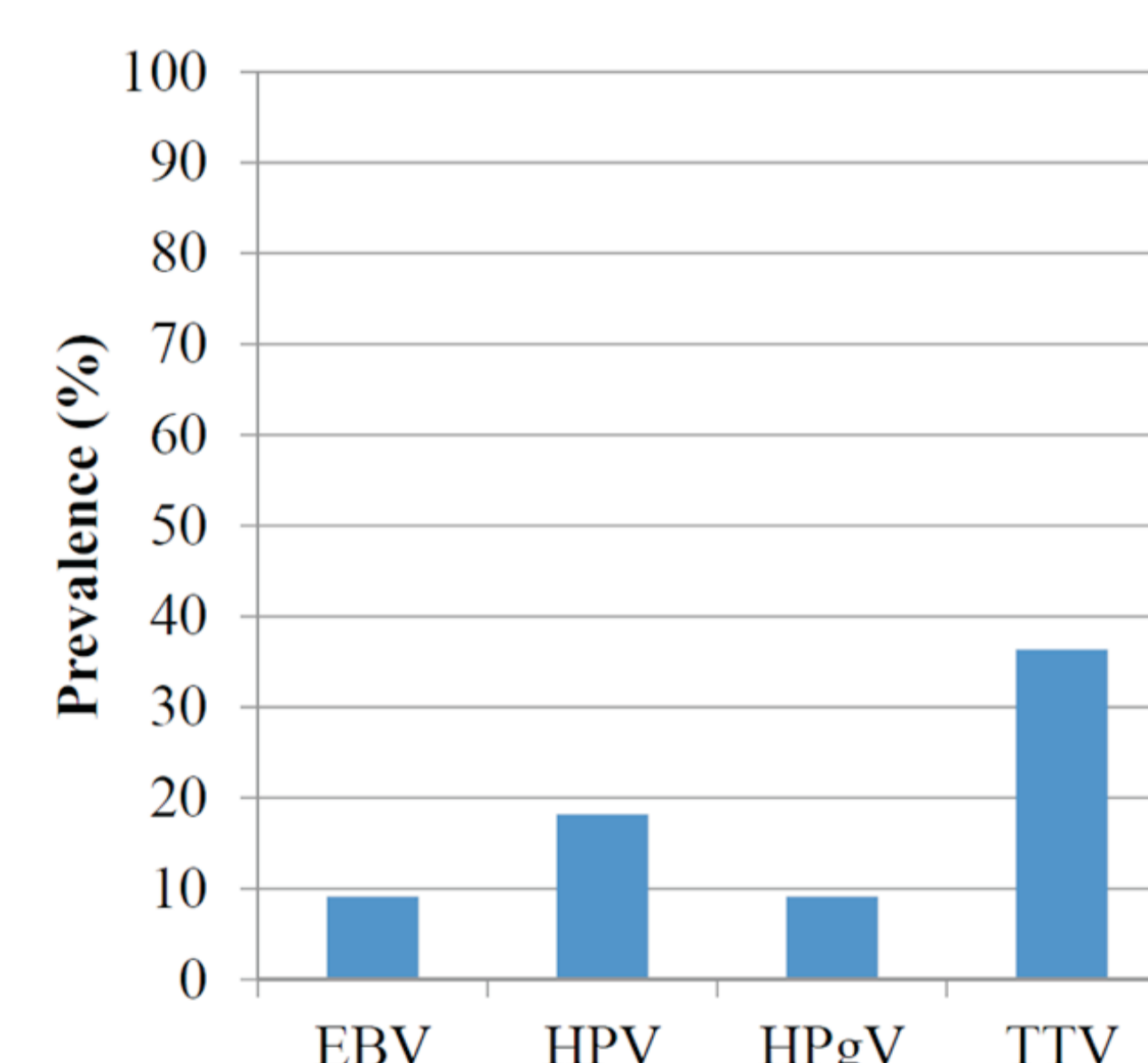


Figure 5. Prevalence of the different virus found in the samples by viral metagenomics. EBV: Epstein-Barr virus; HPV: human papillomavirus; HPgV: human pegivirus; TTV: torque tenovirus.

The EBV, a herpesvirus associated with different types of NHL, was found in one sample of diffuse large B-cell lymphoma.

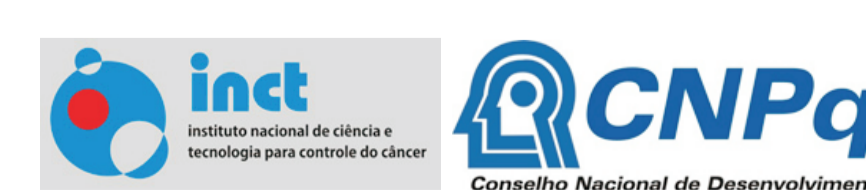
Anelloviruses were found in four samples of diffuse large B-cell lymphoma. These viruses have a ubiquitous presence without any clinical manifestation so far described in humans.

HPV is an epitheliotropic oncogenic virus and was found in two samples of follicular lymphoma. A recent study (Intaraphet et al., 2017) found an increase incidence of lymphoma in woman with HPV persistent infection, using conization as a surrogate marker of persistent HPV infection.

The flavivirus human pegivirus was found in one sample of diffuse large B-cell lymphoma. Some studies associate this latter virus with risk of lymphoma.

In summary, this study has described the viral diversity found in lymphomas, and the next steps will be evaluating the prevalence of each individual virus in the lymphoma cases.

### SUPPORT:



Projeto Gráfico: Setor de Edição e Informação Técnico-Científica / INCA