




# An overview of sequencing technology platforms applied to HTLV-1 studies: a systematic review

Felipe de Oliveira Andrade<sup>1</sup> · Marina Silveira Cucco<sup>2,3</sup> · Melina Mosquera Navarro Borba<sup>2</sup> · Reinaldo Conceição Neto<sup>4</sup> · Luana Leandro Gois<sup>2,5,6</sup> · Filipe Ferreira de Almeida Rego<sup>6</sup> · Luciane Amorim Santos<sup>2,3,5,6</sup> · Fernanda Khouri Barreto<sup>1</sup> 

Received: 14 March 2021 / Accepted: 20 June 2021 / Published online: 20 August 2021  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2021

## Abstract

Human T-lymphotropic virus type 1 (HTLV-1) was the first human retrovirus described. The viral factors involved in the different clinical manifestations of infected individuals are still unknown, and in this sense, sequencing technologies can support viral genome studies, contributing to a better understanding of infection outcome. Currently, several sequencing technologies are available with different approaches. To understand the methodological advances in the HTLV-1 field, it is necessary to organize a synthesis by a rigorous review. This systematic literature review describes different technologies used to generate HTLV-1 sequences. The review follows the PRISMA guidelines, and the search for articles was performed in PubMed, Lilacs, Embase, and SciELO databases. From the 574 articles found in search, 62 were selected. The articles showed that, even with the emergence of new sequencing technologies, the traditional Sanger method continues to be the most commonly used methodology for generating HTLV-1 genome sequences. There are many questions that remain unanswered in the field of HTLV-1 research, and this reflects on the small number of studies using next-generation sequencing technologies, which could help address these gaps. The data compiled and analyzed here can help research on HTLV-1, assisting in the choice of sequencing technologies.

## Introduction

It is estimated that 5–10 million people worldwide are infected with human T-lymphotropic virus type 1 (HTLV-1) [1–4]. Infected individuals can develop HTLV-1-associated

pathologies such as adult T-cell leukemia/lymphoma (ATLL in 2–5% of patients), HTLV-1-associated myelopathy/tropical spastic paraparesis (HAM/TSP in 0.25–3.8% of patients), HTLV-1-associated infectious dermatitis (IDH), and other inflammatory diseases such as uveitis and pneumonitis, or they can be classified as asymptomatic carriers [5–7]. The factors involved in the development of a particular clinical manifestation have not yet been elucidated, and HTLV-1-infected individuals remain without specific treatment [8–10].

The HTLV-1 genome structure is composed of two flanking regions, known as long terminal repeats (5' and 3' LTR), and the structural genes gag, pol, and env. There is also a non-structural region, pX, adjacent to the 3' LTR that encodes the regulatory and accessory proteins Tax, Rex, and HBZ [11]. Molecular characterization of the viral genome, based on sequencing combined with bioinformatics analysis, provides information on genomic regions such as viral integration sites and allows identification of mutations and epigenetic changes [12]. This information is important for the development of HTLV-1 specific vaccines and therapies.

Although HTLV-1 was the first human retrovirus described, the number of HTLV-1 sequences that have been

Handling Editor: Zhongjie Shi.

✉ Fernanda Khouri Barreto  
fernanda.khouri@hotmail.com

<sup>1</sup> Instituto Multidisciplinar em Saúde, Universidade Federal da Bahia, Campus Anísio Teixeira, Rua Hormindo Barros, 58, Bairro Candeias, Vitória da Conquista 45.029-094, Brazil

<sup>2</sup> Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil

<sup>3</sup> Faculdade de Medicina da Bahia, Programa de Pós-graduação em Ciências da Saúde, Universidade Federal da Bahia, Salvador, Brazil

<sup>4</sup> Faculdade Cruzeiro do Sul, Vitória da Conquista, Brazil

<sup>5</sup> Escola Bahiana de Medicina e Saúde Pública, Salvador, Brazil

<sup>6</sup> Universidade Católica do Salvador, Salvador, Brazil

determined is considerably smaller than for other important retroviruses, such as human immunodeficiency virus 1 (HIV-1). In March of 2021, there were 1,048,465 published HIV-1 sequences, while for HTLV-1 there were only 9,980 sequences available in the GenBank database. To perform some specific studies of virus modifications that could be associated to different manifestations in human hosts, it would be necessary to have clinical and epidemiological information about the patients. However, most studies do not give all the information necessary to connect viral mutations with the clinical status of the patient. Even with some sequences already published, an investment in the generation of more HTLV-1 sequences would allow the identification of new mutations that affect infection, which might be helpful for developing new diagnostic strategies.

In 1975, Sanger presented the first DNA sequencing technique, which was widely adopted and is still being used today. This technique is based on the use of modified chain terminators, which are dideoxynucleotides (ddNTPs) [13]. Sequencing techniques later evolved further, resulting in the emergence of next-generation sequencing (NGS), starting with second-generation technology. This technology brought new methodologies for determining nucleotide sequences with greater efficiency and speed, using systems such as 454 from Roche Applied Science, Solexa from Illumina, and Ion Torrent, which expanded the ways of sequencing genetic material [14, 15]. The main examples of second-generation technology are pyrosequencing and sequencing by synthesis (SBS). In this generation, the DNA polymerase acts in conjunction with a chemiluminescent enzyme, which, when complementing a template of a DNA strand, emits chemiluminescent signals, allowing the determination of the sequence [16].

Recently, a third generation has emerged, represented by nanopore sequencing (Oxford Nanopore Technologies) and Pacific Biosciences (PacBio) methodologies [15]. Unlike other sequencing technologies, these methods can be used to sequence unique DNA molecules and to produce longer read lengths in a shorter time than was possible in the previous generations [17]. The nanopore method stands out not only for generating long nucleotide chains through larger devices such as GridION and PromethION but also through small portable devices such as MinION and Flongle. This technique is based on the passage of genetic material through a nanopore membrane, which detects the electrical signals emitted during the passage of each nucleotide [18].

It should be noted that, in recent years, there has been significant technological diversification in genome sequencing, with more efficient, cheaper, and faster devices. Investigating which sequencing technology is most used to determine HTLV-1 genome sequences allows us to understand the limitations and possibilities of research carried out on the viral genome. This may help to fill the gaps in our knowledge

about this virus, such as the factors involved in the development of HTLV-1-associated diseases. Considering the importance of the technological choice for sequencing, in this article, we review the different technologies used to generate HTLV-1 sequences and the contributions of these techniques to new investigations of this retrovirus.

## Materials and methods

This study consists of a systematic literature review carried out in accordance with the guidelines of Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). A systematic search was conducted for studies in which performed partial or total sequencing of the HTLV-1 genome was performed. The articles were searched in the PubMed, Lilacs, SciELO, and EMBASE databases in May 2021. The search algorithm used was composed of subjects from the DeCS/MeSH database and additional keywords: ("*Human T lymphotropic virus 1*" OR "*HTLV-1*") AND "*sequence\**" AND ("*molecular sequence data*" OR "*sequencing*"). Through the search algorithm, all titles were cross-checked to identify possible duplicate studies.

For the selection of articles, the following inclusion criteria were applied to select studies: (i) the articles were in Portuguese, English or Spanish; (ii) they were original studies, and (iii) the articles presented complete or partial genome sequences of HTLV-1. Articles published since 2000 were included. The exclusion criteria were as follows: (i) studies not specifying the sequencing method, (ii) studies that did not generate HTLV-1 sequences or did not specify the number of sequences generated, (iii) animal studies, and (iv) studies in which genome sequencing was performed using a cell line. The HTLV-1 subtypes were not considered as a criterion for selection of articles.

The articles found on the platforms were initially filtered and selected by reading the title and abstract. Subsequently, a new selection was made by reading the full text. After reading and analyzing the selected articles, the data were collected and included in this review. The search for published studies was performed independently by two authors (F.O.A. and M.S.C.), and disagreements about all outcomes were resolved by consensus among all authors.

After reading the selected articles, the following content was extracted from each one: (1) basic information (title, authors, year, objectives), (2) study design, (3) sequencing technology and method, (4) subjects (sample origin and region of the HTLV-1 genome sequenced), and (5) number of sequences generated. The data collected from the articles were tabulated using Microsoft Excel. The figures generated in this work were produced using the programs Adobe Photoshop and Microsoft PowerPoint (2019 versions).

This study was registered with the International Prospective Register of Systematic Reviews (PROSPERO) under the number CRD42020218387.

## Results

The search for studies identified a total of 574 articles, of which 350 were available in PubMed, 213 in EMBASE, one in SciELO, and 10 in Lilacs. Of these, 65 were excluded due to duplication, 404 were excluded after selection by title and abstract, and 43 after reading the full text. Ultimately, 62 articles were included in the systematic review (Fig. 1).

The articles indicated the use of three sequencing methodologies: Sanger, Illumina, and Ion Torrent. Sanger sequencing, which is the first generation of sequencing, was the most frequently used technique. Even after the emergence of NGS methodologies in 2004, it was observed that most published HTLV-1 studies continued to use the Sanger method preferentially. Among the 62 articles used in this review, 59 used the Sanger method, and, of these, 40 were carried out after 2004 (Table 1).

In most of the studies, a partial HTLV-1 genome sequence was determined. Of the 59 articles that used the Sanger method [19–77], 56 reported partial genome sequences, and three reported complete sequencing of the HTLV-1 genome. Of the four articles that used NGS [12, 75, 78, 79], two reported partial genome sequences and two reported complete genome sequences.

Another important aspect of these articles was the difference in the number of sequences generated for each region of the HTLV-1 genome: 1258 sequences of LTR, 89 sequences of gag, 124 of pol, 777 of env, and 1420 of the pX region (Fig. 2). It is important to highlight that there are four different overlapping open reading frames (ORFs) in the pX region that encode regulatory proteins and the HTLV-1 bzip domain gene (hbx), which is transcribed in the antisense direction from a promoter present in the 3'LTR. The number of sequences generated for each ORF and hbx are as follows: ORF I, 311; ORF-II, 54; ORF-III, 54; ORF-IV, 1153; hbx, 10. In addition, 14 partial genome sequences with the precise regions not described were found, and 228 complete HTLV-1 genome sequences were reported.

Brazil is the country with the largest number of sequences generated, distributed through 26 sequencing studies, followed by Japan, with seven. Colombia and France had four studies each, and Argentina and Chile each had three. Two studies each were performed in Gabon and Spain, and in Cuba, India, Israel, Italy, Mozambique, the UK, Portugal, and Russia, only one study was performed. Finally, there were also six articles that did not provide information about the origins of the sequences (Fig. 3).

It is important to note that, despite being a state-of-the-art technique employed for genome sequencing of other retroviruses such as HIV [80], no published study using third-generation sequencing for HTLV-1 was found. In addition, most of the studies refer to information generated more than 10 years ago, in which more than 3,000 sequences were generated, while the most recent studies generated only approximately 800 sequences (Fig. 4). Another interesting topic to be highlighted is the lack of clinical information about the patients included in the studies. Of the 62 articles included, 60% did not report the clinical status of the studied population.

## Discussion

In the 41 years since the discovery of HTLV-1, no effective therapeutic treatments or vaccines have been developed, and it is still not clear what determines different infection outcomes. During this period, diverse sequencing technologies have become available. The central aim of this systematic review was to summarize the different technologies used in the HTLV-1 field in order to guide the decision-making processes on the generation of new HTLV-1 genome sequences.

The Sanger method was the most commonly used for generating HTLV-1 sequences, followed by Illumina and Ion Torrent. All of these techniques have advantages and disadvantages. The characteristics of HTLV-1, as well as the specific aspects of each method must be taken into consideration.

One important aspect in HTLV-1 infection is that, after infection, the virus integrates into the host cell DNA as a provirus. Unlike HIV, in HTLV-1 infection, the circulating viral RNA is not easily detected in the plasma or serum, and additional techniques are usually needed prior to sequencing, such as PBMC separation and nested PCR [81, 82]. In this sense, the HTLV-1 sample extraction and preparation steps are an important point to consider during the choice of the sequencing platform to use.

Among the sequencing technology platforms, considering its low error rate, Sanger sequencing is considered the gold standard, despite being first-generation and having a high cost. Furthermore, it is possible to assess the sequencing quality based on other parameters, such as sequence length, sequencing depth, and GC content. One study reported that sequencing quality is more stable and GC depth distribution is better with Ion Torrent than with HiSeq 2000 [83]. Importantly, even when the goal is to sequence larger regions and/or the complete proviral genome, technologies such as Illumina and Ion Torrent produce small sequence reads. This read size, as well as the polymerase chain reaction (PCR) step, can impair the understanding of an essential aspect of HTLV-1 infection: clonality. While in patients with ATLL

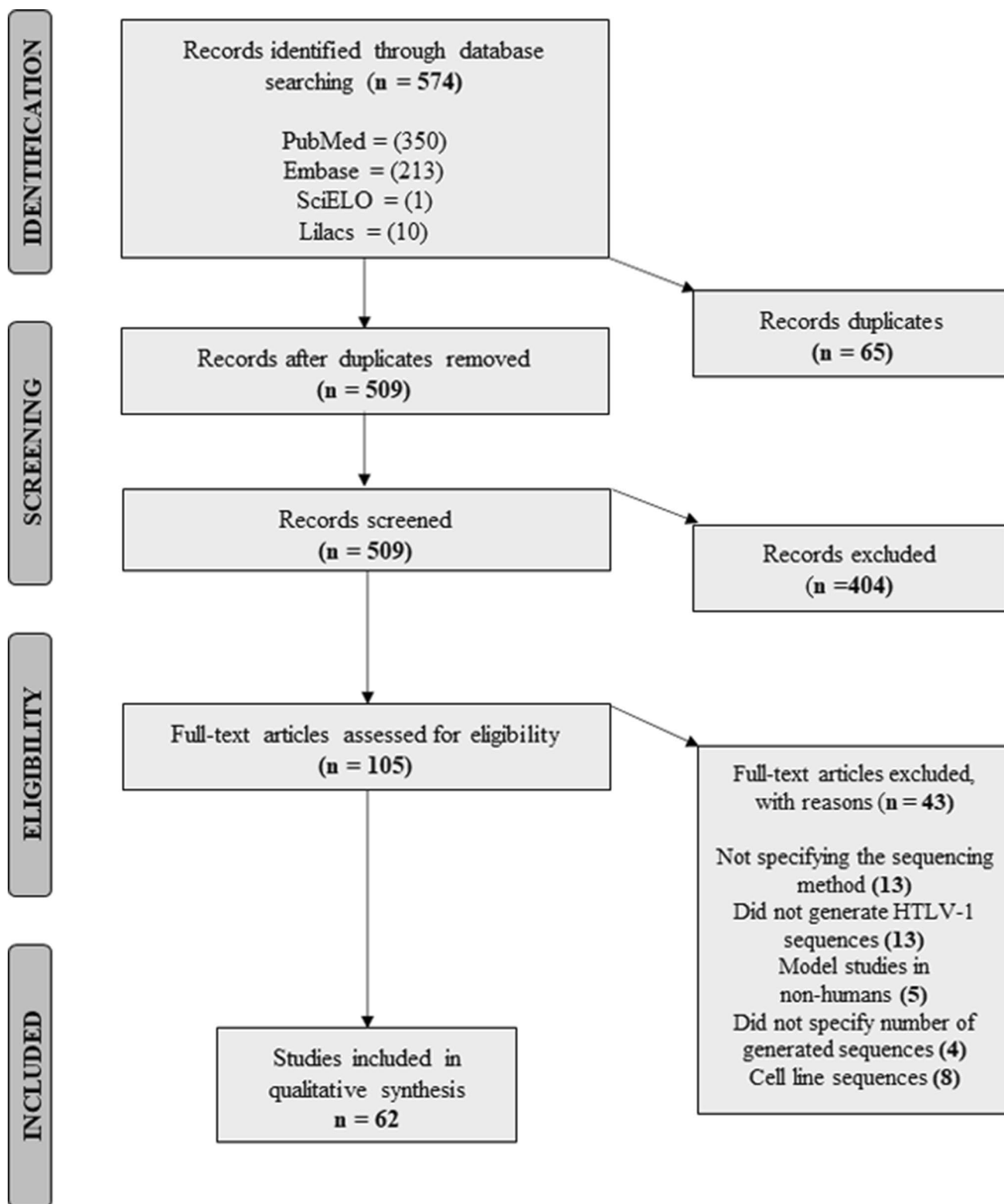


Fig. 1 Flow diagram for the systematic selection of studies to review

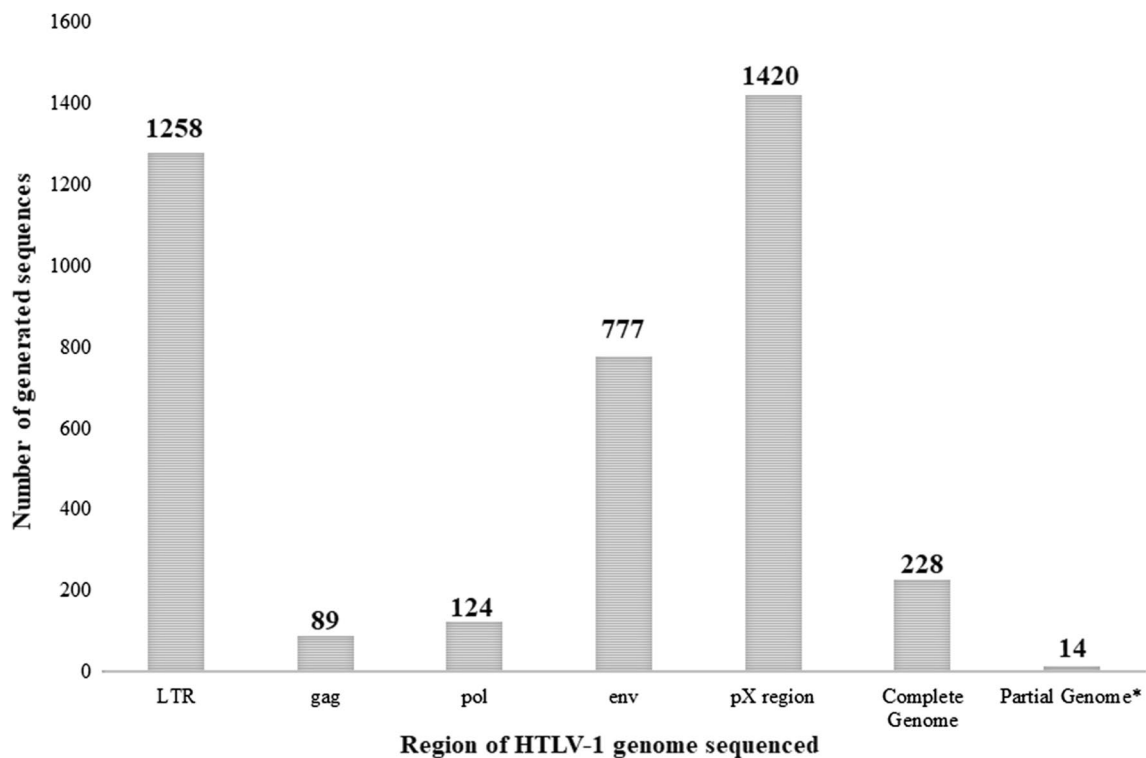
**Table 1** Summary of information collected from the 62 selected studies, including sample origin, sequencing methodology, equipment, and number of sequences generated

Reference	Sample origin	Sequencing methodology	Equipment	No. of sequences generated
[19]	NA <sup>b</sup>	Sanger	ABI® 377A DNA sequencer	218
[20]	NA	Sanger	Hitachi Fluorescent DNA sequencer SQ-5500	39
[21]	NA	Sanger	Hitachi Fluorescent DNA sequencer SQ-5500	64
[22]	France and UK	Sanger	ABI 373 automatic DNA sequencer	17
[23]	Chile	Sanger	Automated DNA sequencer	37
[24]	Brazil	Sanger	Fmol DNA sequencing system (Promega)	2
[25]	France	Sanger	ABI 377A DNA sequencer	208
[26]	Japan	Sanger	ABI 373 automatic DNA sequencer	178
[27]	NA	Sanger	SQ5500 automated sequencer	138
[28]	Colombia	Sanger	NA	12
[29]	Italy	Sanger	ABI PRISM automatic sequencer	6
[30]	Chile	Sanger	NA	50
[31]	Chile	Sanger	NA	128
[32]	Spain	Sanger	ABI 310 genetic analyzer	4
[33]	Colombia	Sanger	NA	12
[34]	France	Sanger	Applied Biosystems 377 DNA sequencer	1
[35]	Japan	Sanger	ABI 377 DNA sequencer	231
[36]	Colombia	Sanger	ABI Prism serie 3700	11
[37]	Brazil	Sanger	ABI 373 DNA Sequencer	3
[38]	Russia	Sanger	ABI 377 automatic DNA sequencer	8
[39]	India	Sanger	ABI automated DNA sequencer	7
[40]	Brazil	Sanger	ABI 377 Automated DNA Sequencer	26
[41]	France, Gabon and Iran	Sanger	ABI Prism 377 and Ceq2000 sequencer	65
[42]	Argentina	Sanger	ABI model 377 automated DNA sequencer	12
[43]	Brazil	Sanger	ABI 377 Sequencer	134
[44]	Israel	Sanger	ABI automated sequencer	1
[45]	Brazil	Sanger	ABI Prism 377 DNA Sequencer	2
[46]	Brazil	Sanger	ABI 373 DNA Sequencer	5
[47]	NA	Sanger	ABI 310 sequencer	334
[48]	NA	Sanger	ABI 310 autosequencer	316
[49]	Japan	Sanger	ABI 377 DNA Sequence	445
[50]	Brazil	Sanger	ABI 3100 genetic analyzer	46
[51]	Argentina	Sanger	ABI Prism 3100 Genetic Analyzer	44
[52]	Gabon	Sanger	Automatic sequencing system (Euro Sequence Gene Services)	34
[53]	Brazil	Sanger	ABI 3100 genetic analyzer	8
[54]	Brazil	Sanger	ABI 3100 genetic analyzer	5
[55]	Argentina	Sanger	ABI PRISM 377 Automated DNA sequencer	114
[56]	Japan	Sanger	ABI 3730 Sequencer	19
[57]	Brazil	Sanger	ABI PRISM 310 Genetic Analyzer	25
[58]	Brazil	Sanger	ABI 3100 genetic analyzer	8
[59]	Mozambique	Sanger	ABI 3730 Automated DNA Sequencer	25
[60]	Colombia	Sanger	ABI PRISM 310 sequencer	30
[61]	Portugal and Spain	Sanger	Automated DNA sequencing	47
[62]	Brazil	Sanger	NA	1
[63]	Brazil	Sanger	ABI 3130 genetic analyzer	13
[64]	Brazil	Sanger	ABI 3100 genetic analyzer	32
[65]	Brazil	Sanger	ABI 3100 genetic analyzer	146

**Table 1** (continued)

Reference	Sample origin	Sequencing methodology	Equipment	No. of sequences generated
[66]	Brazil	Sanger	ABI 3100 genetic analyzer	18
[67]	Cuba	Sanger	Genome Lab Dye Terminator Cycle Sequence	12
[78]	Brazil	Illumina	Illumina MiSeq System	90
[68]	Brazil	Sanger	ABI 1373 Automated DNA Sequencer	14
[75]	Brazil	Sanger and Ion Torrent	ABI 3130xl Genetic Analyzer / Ion 314TM Chip 8-pack	22
[69]	Iran	Sanger	ABI 3730 Sequencer	2
[70]	Brazil	Sanger	ABI PRISM 310 Genetic Analyzer	2
[71]	Japan and Brazil	Sanger	ABI PRISM 3740 Genetic Analyzer	14
[72]	Brazil	Sanger	ABI PRISM 3100 Genetic Analyzer	24
[12]	Japan	Illumina	Illumina MiSeq or NextSeq	98
[73]	Iran	Sanger	NA	5
[79]	Brazil	Ion Torrent	Ion 318™ Chip PGM	31
[74]	Brazil	Sanger	ABI 3130 Genetic Analyzer	132
[76]	Brazil	Sanger	ABI 3100 Genetic Analyzer	21
[77]	Brazil and Japan	Sanger	ABI 1373 Sequencer	90

<sup>a</sup>ABI, Applied Biosystems; NA, not available; \*, sequence region not described. Total for each instrument used: ABI 377/377A DNA Sequencer, 12; ABI 3100 Genetic Analyzer, 10; ABI PRISM 310 Genetic Analyzer, 6; ABI 373 DNA sequencer, 4; Hitachi Fluorescent DNA sequencer SQ-5500, 3; ABI 3730 Sequencer, 3; ABI 3130/3130xl genetic analyzer, 3; ABI PRISM automatic sequencer, 3; Automated DNA sequencer, 2; ABI 1373 Automated DNA Sequencer, 2; Illumina MiSeq, 2; ABI Prism 3700, 1; Fmol DNA sequencing system (Promega), 1; Ion 318™ Chip PGM, 1; ABI PRISM 3740 Genetic Analyzer, 1; Illumina NextSeq, 1; Ceq2000 sequencer, 1; Automatic sequencing system (Euro Sequence Gene Services), 1; Genome Lab Dye Terminator Cycle Sequence 1

**Fig. 2** Numbers of sequences generated for each region of the HTLV-1 genome

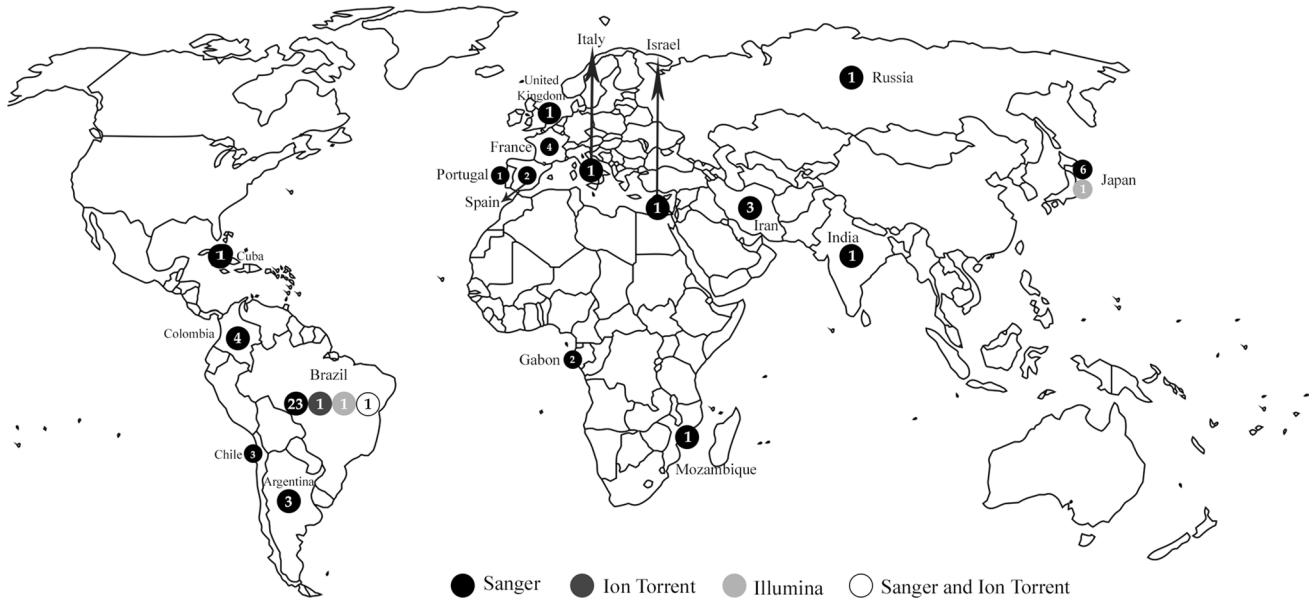


Fig. 3 Geographic origin of HTLV-1 sequences and methodology used for sequence generation

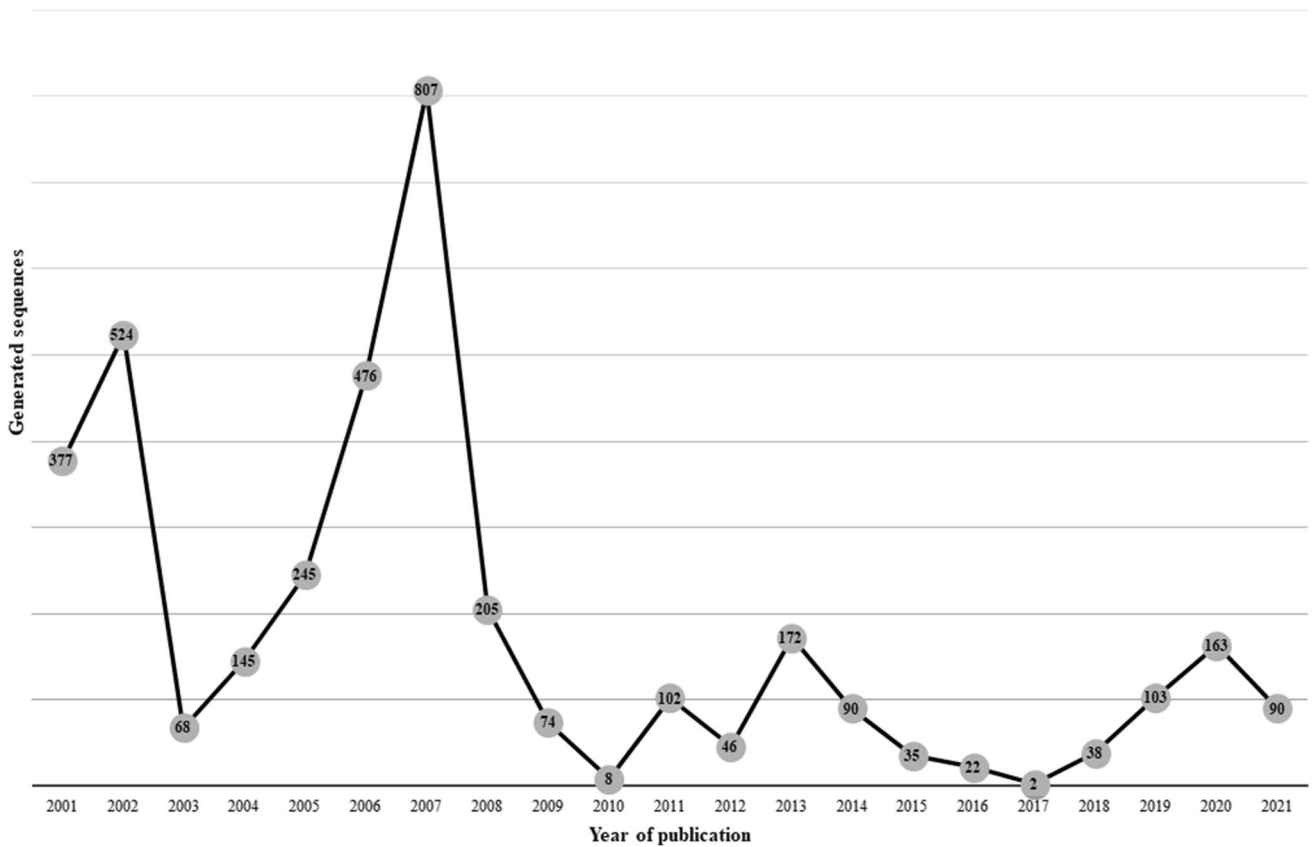


Fig. 4 Number of HTLV-1 sequences generated between 2000 and 2020

there is a monoclonal pattern, in patients with IDH or HAM/TSP, and in asymptomatic carriers, a polyclonal pattern is found [84, 85]. Therefore, the small size of the genome sequencing readout may make it difficult to identify viral quasispecies and may give an unrealistic biological picture.

It is important to highlight that sequencing of viral genomes is important for understanding the infection process [86]. Therefore, the use of few and old sequencing methods, despite the emergence of more innovative, faster, and often less expensive technologies, makes the goal of developing better alternatives for infection control and the understanding of viral pathogenesis increasingly distant. In addition, animal models are important in HTLV-1 research and have allowed significant advances in the understanding of viral infection and pathogenesis. Each animal model has its advantages. Rats are used in studies involving HAM/TSP, and non-human primates are used in studies analyzing the immune response and viral persistence [87].

The emergence of new sequencing protocols has led to a reduction in the time required and production costs [88]. Despite that, no article included in this study used more recent technologies, such as the third-generation sequencing. MinION and PacBio could be an interesting alternative, due to their shorter processing time, despite providing sequences with regular quality, when compared to older methods. These methods can be useful in HTLV-1 research, increasing the number of partial and/or complete sequences available on the platforms and contributing to a better understanding of the virus-host relationship.

In addition to the predominance of the older techniques, most of the studies focused on sequencing specific regions of the genome, with few studies generating complete genome sequences. The LTR and pX regions were the most frequently sequenced. This could be because of the importance of the LTR for the subtyping and the fact that pX encodes the HTLV-1 regulatory proteins. In this context, it is relevant to point out that complete genome sequencing is essential for the identification of gene functions and their involvement in disease as well as for vaccine development.

This systematic review demonstrated a deficit in the number of HTLV-1 sequences. However, this study has an important limitation, since sequences can be deposited in databases such as GenBank without being necessarily associated with a published article. However, our data corroborate an ongoing study carried out by our group that highlights the deficit of complete HTLV-1 genome sequences available in the GenBank database. In this study, we verified that only 242 complete HTLV-1 genome sequences were available in the GenBank database, and most of these sequences did not include clinical and epidemiological information about the patient.

On the other hand, the majority of studies provided geographical information about the samples sequenced. Most of

them were from endemic regions such as Japan and Brazil. Another country that deserves attention is Colombia. The Colombian island of Tumaco has a high population density and a very high prevalence of HAM/TSP, which is why this region is a focus of study of HTLV-1 [33]. Moreover, few articles from Africa were found, despite being the continent with the highest endemicity of HTLV-1 [1]. The European continent also contributes to the generation of HTLV-1 sequences, although relatively few articles describe the sequencing. Some studies did not report the origin of the sequence, which limits their epidemiological value. The sum of studies from each country does not correspond to the number of articles included, because some studies include samples from different countries, such as Bandeira et al., 2018 [71].

Interestingly, only 21 articles included in this review were published in the last 10 years, which is equivalent to almost 30% of the total number of studies, revealing that there is still low investment in research in the HTLV-1 field. The encouragement of more investments in HTLV-1 studies may contribute to an increased number of HTLV-1 sequences generated in different geographic regions, and this can assist in the understanding of the global and regional distribution of this virus [1].

There are gaps to be filled in relation to information on HTLV-1 infection. Although it was the first human retrovirus described and has been proven to be associated with the development of diseases, studies on the pathogenesis and treatment of this virus are not encouraged, and worse, investment in research is decreasing [89], demonstrating that HTLV-1 is still a neglected virus [90, 91]. Thus, more investment in HTLV-1 research and the implementation of worldwide prevention strategies will be the main motor for the eradication of these infections.

## Conclusion

The analysis of the articles selected for this systematic review showed that the number of studies sequencing the HTLV-1 genome is much lower than for other retroviruses, and most of these studies still opt for Sanger sequencing despite the emergence of new methodologies. This demonstrates a lack of investment in this field. It is important to note that Sanger sequencing has advantages over other methods. However, NGS techniques also have characteristics that may be important for answering questions that remain about HTLV-1 infection. Investments in HTLV-1 research are needed, mainly in the use of more current methodologies, since they are methodologies that have been developed through lessons learned and improved by the previous generation.



**Author contributions** FOA: methodology, formal analysis, investigation, writing. MSC: methodology, formal analysis, investigation, writing. MMNB: methodology, investigation and writing. RCN: methodology. LLG: conceptualization and writing—review and editing. FFAR: conceptualization and writing—review and editing. LAS: conceptualization and writing—review and editing. FKB: conceptualization, formal analysis, writing—review and editing and supervision. All authors read and approved the final manuscript.

**Funding** This research was funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), N. 421342/2018-8.

## Declarations

**Conflict of interest** The authors declare no conflict of interest. The funder had no role in the design of the study; in the collection, analysis, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

## References

- Gessain A, Cassar O (2012) Epidemiological aspects and world distribution of HTLV-1 infection. *Front Microbiol* 3:388. <https://doi.org/10.3389/fmicb.2012.00388>
- Poiesz BJ, Ruscetti FW, Gazdar AF et al (1980) Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc Natl Acad Sci USA* 77:7415–7419. <https://doi.org/10.1073/pnas.77.12.7415>
- de-Thé G, Gessain A, Gazzolo L et al (1985) Comparative seroepidemiology of HTLV-I and HTLV-III in the French West Indies and some African countries. *Cancer Res* 45:4633s–4636s
- Hinuma Y, Nagata K, Hanaoka M et al (1981) Adult T-cell leukemia: antigen in an ATL cell line and detection of antibodies to the antigen in human sera. *Proc Natl Acad Sci USA* 78:6476–6480
- Gessain A, Barin F, Vernant JC et al (1985) Antibodies to human T-lymphotropic virus type-I in patients with tropical spastic paraparesis. *Lancet* 2:407–410. [https://doi.org/10.1016/s0140-6736\(85\)92734-5](https://doi.org/10.1016/s0140-6736(85)92734-5)
- La Grenade L, Manns A, Fletcher V et al (1998) Clinical, pathologic, and immunologic features of human T-lymphotropic virus type I-associated infective dermatitis in children. *Arch Dermatol* 134:439–444. <https://doi.org/10.1001/archderm.134.4.439>
- Yoshida M, Miyoshi I, Hinuma Y (1982) Isolation and characterization of retrovirus from cell lines of human adult T-cell leukemia and its implication in the disease. *Proc Natl Acad Sci USA* 79:2031–2035. <https://doi.org/10.1073/pnas.79.6.2031>
- Barreto FK, Khouri R, de Rego FFA et al (2016) Analyses of HTLV-1 sequences suggest interaction between ORF-I mutations and HAM/TSP outcome. *Infect Genet Evol* 45:420–425. <https://doi.org/10.1016/j.meegid.2016.08.020>
- Borba MMN, Farre L, Bittencourt AL et al (2019) Assessment of genetic diversity of HTLV-1 ORF-I sequences collected from patients with different clinical profiles. *AIDS Res Hum Retroviruses* 35:881–884. <https://doi.org/10.1089/AID.2019.0127>
- Cucco MS, de Moraes LEP, de Oliveira AF et al (2021) Molecular characterization of HTLV-1 genomic region hbz from patients with different clinical conditions. *J Med Virol*. <https://doi.org/10.1002/jmv.27005>
- Barreto FK, Araújo THA, Rego FF de A, Alcántara LCJ (2017) A fully annotated genome sequence of human T-cell lymphotropic virus type 1 (HTLV-1). *J Bioinf Com Sys Bio* 1(1):105
- Katsuya H, Islam S, Tan BJY et al (2019) The nature of the HTLV-1 provirus in naturally infected individuals analyzed by the viral DNA-capture-seq approach. *Cell Rep* 29:724–735.e4. <https://doi.org/10.1016/j.celrep.2019.09.016>
- Sanger F, Coulson AR (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol* 94:441–448. [https://doi.org/10.1016/0022-2836\(75\)90213-2](https://doi.org/10.1016/0022-2836(75)90213-2)
- Ansorge WJ (2009) Next-generation DNA sequencing techniques. *N Biotechnol* 25:195–203. <https://doi.org/10.1016/j.nbt.2008.12.009>
- Kchouk M, Gibrat J-F, Elloumi M (2017) Generations of sequencing technologies: from first to next generation. *Biol Med*. <https://doi.org/10.4172/0974-8369.1000395>
- Yoshinaga Y, Daum C, He G, O'Malley R (2018) Genome sequencing. *Methods Mol Biol* 1775:37–52. [https://doi.org/10.1007/978-1-4939-7804-5\\_4](https://doi.org/10.1007/978-1-4939-7804-5_4)
- Rhoads A, Au KF (2015) PacBio sequencing and its applications. *Genomics Proteomics Bioinformatics* 13:278–289. <https://doi.org/10.1016/j.gpb.2015.08.002>
- Jain M, Olsen HE, Paten B, Akeson M (2016) The oxford nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biol* 17:239. <https://doi.org/10.1186/s13059-016-1103-0>
- Leclercq I, Mortreux F, Cavrois M et al (2000) Host sequences flanking the human T-cell leukemia virus type 1 provirus in vivo. *J Virol* 74:2305–2312. <https://doi.org/10.1128/jvi.74.5.2305-2312.2000>
- Seki M, Higashiyama Y, Mizokami A et al (2000) Up-regulation of human T lymphotropic virus type 1 (HTLV-1) tax/rex mRNA in infected lung tissues. *Clin Exp Immunol* 120:488–498. <https://doi.org/10.1046/j.1365-2249.2000.01237.x>
- Nakane S, Shirabe S, Moriuchi R et al (2000) Comparative molecular analysis of HTLV-I proviral DNA in HTLV-I infected members of a family with a discordant HTLV-I-associated myelopathy in monozygotic twins. *J Neurovirol* 6:275–283. <https://doi.org/10.3109/13550280009030753>
- Morozov VA, Lagaye S, Taylor GP et al (2000) Chimeric matrix proteins encoded by defective proviruses with large internal deletions in human T-Cell leukemia virus type 1-infected humans. *J Virol* 74:3933–3940
- Sonoda S, Li HC, Cartier L et al (2000) Ancient HTLV type 1 provirus DNA of Andean mummy. *AIDS Res Hum Retroviruses* 16:1753–1756. <https://doi.org/10.1089/08892220050193263>
- Caterino-de-Araújo A, Favero A, de los Santos-Fortuna E et al (2000) HTLV-I/HTLV-II coinfection in an AIDS patient from São Paulo, Brazil. *AIDS Res Hum Retroviruses* 16:715–719. <https://doi.org/10.1089/088922200308710>
- Mortreux F, Leclercq I, Gabet AS et al (2001) Somatic mutation in human T-cell leukemia virus type 1 provirus and flanking cellular sequences during clonal expansion in vivo. *J Natl Cancer Inst* 93:367–377. <https://doi.org/10.1093/jnci/93.5.367>
- Furukawa Y, Kubota R, Tara M et al (2001) Existence of escape mutant in HTLV-I tax during the development of adult T-cell leukemia. *Blood* 97:987–993. <https://doi.org/10.1182/blood.v97.4.987>
- Okazaki S, Moriuchi R, Yosizuka N et al (2001) HTLV-1 proviruses encoding non-functional TAX in adult T-cell leukemia. *Virus Genes* 23:123–135. <https://doi.org/10.1023/a:1011840918149>
- Dominguez MC, Castillo A, Cabrera J et al (2002) Envelope sequence variation and phylogenetic relations of human T cell lymphotropic virus type 1 from endemic areas of Colombia. *AIDS Res Hum Retroviruses* 18:887–890. <https://doi.org/10.1089/08892220260190371>

29. Manca N, Perandin F, De Simone N et al (2002) Detection of HTLV-I tax-rex and pol gene sequences of thymus gland in a large group of patients with myasthenia gravis. *J Acquir Immune Defic Syndr* 29:300–306. <https://doi.org/10.1097/00126334-200203010-00012>
30. Ramirez E, Cartier L, Villota C, Fernandez J (2002) Genetic characterization and phylogeny of human T-cell lymphotropic virus type I from Chile. *Virus Res* 84:135–149. [https://doi.org/10.1016/S0168-1702\(02\)00005-9](https://doi.org/10.1016/S0168-1702(02)00005-9)
31. Ramirez E, Fernandez J, Cartier L et al (2003) Defective human T-cell lymphotropic virus type I (HTLV-I) provirus in seronegative tropical spastic paraparesis/HTLV-I-associated myelopathy (TSP/HAM) patients. *Virus Res* 91:231–239. [https://doi.org/10.1016/S0168-1702\(02\)00276-9](https://doi.org/10.1016/S0168-1702(02)00276-9)
32. Toro C, Rodés B, Poveda E, Soriano V (2003) Rapid development of subacute myelopathy in three organ transplant recipients after transmission of human T-cell lymphotropic virus type I from a single donor. *Transplantation* 75:102–104. <https://doi.org/10.1097/00007890-200301150-00019>
33. Balcázar N, Sánchez GI, García-Vallejo F (2003) Sequence and phylogenetic analysis of human T cell lymphotropic virus type I from Tumaco, Colombia. *Mem Inst Oswaldo Cruz* 98:641–648. <https://doi.org/10.1590/S0074-02762003000500010>
34. Leclercq I, Mortreux F, Rabaoui S et al (2003) Naturally occurring substitutions of the human T-cell leukemia virus type 1 3' LTR influence strand-transfer reaction. *J Virol Methods* 109:105–117. [https://doi.org/10.1016/S0166-0934\(03\)00052-1](https://doi.org/10.1016/S0166-0934(03)00052-1)
35. Furukawa Y, Usuku K, Izumo S, Osame M (2004) Human T cell lymphotropic virus type I (HTLV-I) p12I is dispensable for HTLV-I transmission and maintenance of infection in vivo. *AIDS Res Hum Retroviruses* 20:1092–1099. <https://doi.org/10.1089/aid.2004.20.1092>
36. Chávez M, Domínguez MC, Blank A et al (2004) Reconstrucción de la evolución molecular de la infección actual por el virus linfotrópico humano tipo I en Colombia. *Biomédica (Bogotá)*. <https://doi.org/10.7705/biomedica.v24i1.1245>
37. Vallinoto ACR, Muto NA, Pontes GS et al (2004) Serological and molecular evidence of HTLV-I infection among Japanese immigrants living in the Amazon region of Brazil. *Jpn J Infect Dis* 57:156–159
38. Morozov VA, Syrtsev AV, Ellerbrok H et al (2005) Mycosis fungoides in European Russia: no antibodies to human T cell leukemia virus type I structural proteins, but virus-like sequences in blood and saliva. *Intervirology* 48:362–371. <https://doi.org/10.1159/000086063>
39. Ohkura S, Yamashita M, Ishida T et al (2005) Phylogenetic heterogeneity of new HTLV type 1 isolates from southern India in subgroup A. *AIDS Res Hum Retroviruses* 21:325–330. <https://doi.org/10.1089/aid.2005.21.325>
40. Iñiguez AM, Otsuki K, Magalhães GP et al (2005) Genetic markers on the HTLV-1 p12I protein sequences from Brazilian HAM/TSP patients and asymptomatic HTLV-1 carrier isolates. *AIDS Res Hum Retroviruses* 21:580–582. <https://doi.org/10.1089/aid.2005.21.580>
41. Capdepon S, Londos-Gagliardi D, Joubert M et al (2005) New insights in HTLV-I phylogeny by sequencing and analyzing the entire envelope gene. *AIDS Res Hum Retroviruses* 21:28–42. <https://doi.org/10.1089/aid.2005.21.28>
42. Gastaldello R, Otsuki K, Barbas MG et al (2005) Molecular evidence of HTLV-1 intrafamilial transmission in a non-endemic area in Argentina. *J Med Virol* 76:386–390. <https://doi.org/10.1002/jmv.20370>
43. Kashima S, Alcantara LC, Takayanagui OM et al (2006) Distribution of human T cell lymphotropic virus type 1 (HTLV-1) subtypes in Brazil: genetic characterization of LTR and tax region. *AIDS Res Hum Retroviruses* 22:953–959. <https://doi.org/10.1089/aid.2006.22.953>
44. Shohat M, Shohat B, Mimouni D et al (2006) Human T-cell lymphotropic virus type 1 provirus and phylogenetic analysis in patients with mycosis fungoides and their family relatives. *Br J Dermatol* 155:372–378. <https://doi.org/10.1111/j.1365-2133.2006.07312.x>
45. Vallinoto ACR, Pontes GS, Muto NA et al (2006) Identification of human T-cell lymphotropic virus infection in a semi-isolated Afro-Brazilian quilombo located in the Marajó Island (Pará, Brazil). *Mem Inst Oswaldo Cruz* 101:103–105. <https://doi.org/10.1590/S0074-02762006000100020>
46. Souza LA, Lopes IGL, Maia EL et al (2006) Molecular characterization of HTLV-1 among patients with tropical spastic paraparesis/HTLV-1 associated myelopathy in Belém, Pará. *Rev Soc Bras Med Trop* 39:504–506. <https://doi.org/10.1590/S0037-86822006000500017>
47. Kim FJ, Lavanya M, Gessain A et al (2006) Intrahost variations in the envelope receptor-binding domain (RBD) of HTLV-1 and STLV-1 primary isolates. *Retrovirology* 3:29. <https://doi.org/10.1186/1742-4690-3-29>
48. Miyazaki M, Yasunaga J-I, Taniguchi Y et al (2007) Preferential selection of human T-cell leukemia virus type 1 provirus lacking the 5' long terminal repeat during oncogenesis. *J Virol* 81:5714–5723. <https://doi.org/10.1128/JVI.02511-06>
49. Kubota R, Hanada K, Furukawa Y et al (2007) Genetic stability of human T lymphotropic virus type I despite antiviral pressures by CTLs. *J Immunol* 178:5966–5972. <https://doi.org/10.4049/jimmunol.178.9.5966>
50. de Mota ACA, Van Dooren S, de Fernandes FMC et al (2007) The close relationship between South African and Latin American HTLV type 1 strains corroborated in a molecular epidemiological study of the HTLV type 1 isolates from a blood donor cohort. *AIDS Res Hum Retroviruses*. <https://doi.org/10.1089/aid.2006.0203>
51. Eirin ME, Dilernia DA, Berini CA et al (2008) Divergent strains of human T-lymphotropic virus type 1 (HTLV-1) within the Cosmopolitan subtype in Argentina. *AIDS Res Hum Retroviruses* 24:1237–1244. <https://doi.org/10.1089/aid.2008.0024>
52. Etenna SL-D, Caron M, Besson G et al (2008) New insights into prevalence, genetic diversity, and proviral load of human T-cell leukemia virus types 1 and 2 in pregnant women in Gabon in equatorial central Africa. *J Clin Microbiol* 46:3607–3614. <https://doi.org/10.1128/JCM.01249-08>
53. de Magalhães TS, Miranda ACAM, Alcantara LCJ et al (2008) Phylogenetic and molecular analysis of HTLV-1 isolates from a medium sized town in northern of Brazil: tracing a common origin of the virus from the most endemic city in the country. *J Med Virol* 80:2040–2045
54. de Rego FFA, Alcantara LCJ, de Moura Neto JP et al (2008) HTLV type 1 molecular study in Brazilian villages with African characteristics giving support to the post-Columbian introduction hypothesis. *AIDS Res Hum retroviruses* 24:673–677
55. Gastaldello R, Iñiguez AM, Otsuki K et al (2008) HTLV type 1 genetic types among native descendants in Argentina. *AIDS Res Hum Retroviruses* 24:1139–1146. <https://doi.org/10.1089/aid.2007.0299>
56. Eguchi K, Fujii H, Oshima K et al (2009) Human T-lymphotropic virus type 1 (HTLV-1) genetic typing in Kakeroma Island, an island at the crossroads of the ryukyans and Wajin in Japan, providing further insights into the origin of the virus in Japan. *J Med Virol* 81:1450–1456. <https://doi.org/10.1002/jmv.21540>
57. dos Santos EL, Tamegão-Lopes B, Machado LFA et al (2009) Molecular characterization of HTLV-1/2 among blood donors in Belém, State of Pará: first description of HTLV-2b subtype in the

- Amazon region. *Rev Soc Bras Med Trop* 42:271–276. <https://doi.org/10.1590/S0037-86822009000300007>
58. de Rego FFA, Miranda AM, de Santos ES et al (2010) Seroprevalence and molecular epidemiology of HTLV-1 isolates from HIV-1 co-infected women in Feira de Santana, Bahia, Brazil. *AIDS Res Hum Retroviruses*. <https://doi.org/10.1089/aid.2009.0298>
  59. Vicente ACP, Gudo ES, Iñiguez AM et al (2011) Genetic characterization of human T-cell lymphotropic virus type 1 in Mozambique: transcontinental lineages drive the HTLV-1 endemic. *PLoS Negl Trop Dis*. <https://doi.org/10.1371/journal.pntd.0001038>
  60. Salcedo-Cifuentes M, Domínguez MC, García-Vallejo F (2011) Genome epidemiology and tropical spastic paraparesis associated with human T-cell lymphotropic virus type 1. *Rev Panam Salud Publica* 30:422–430
  61. Pádua E, Rodés B, Pérez-Piñar T et al (2011) Molecular characterization of human T cell leukemia virus type 1 subtypes in a group of infected individuals diagnosed in Portugal and Spain. *AIDS Res Hum Retroviruses* 27:317–322. <https://doi.org/10.1089/aid.2010.0195>
  62. Zanella L, Otsuki K, Marin MA et al (2012) Complete genome sequence of Central Africa human T-cell lymphotropic virus subtype 1b. *J Virol* 86:12451. <https://doi.org/10.1128/JVI.02258-12>
  63. Magri MC, de Brigido LFM, Rodrigues R et al (2012) Tax gene characterization of human T-Lymphotropic virus type 1 strains from Brazilian HIV-coinfected patients. *AIDS Res Hum Retroviruses* 28:1775–1778. <https://doi.org/10.1089/aid.2011.0389>
  64. Magri MC, de Macedo Brigido LF, Rodrigues R et al (2012) Phylogenetic and similarity analysis of HTLV-1 isolates from HIV-coinfected patients from the South and Southeast regions of Brazil. *AIDS Res Hum Retroviruses* 28:110–114. <https://doi.org/10.1089/aid.2011.0117>
  65. Miranda ACAM, Barreto FK, de Amarante MFC et al (2013) Molecular characterization of HTLV-1 gp46 glycoprotein from health carriers and HAM/TSP infected individuals. *Virology*. <https://doi.org/10.1186/1743-422X-10-75>
  66. Mota-Miranda ACA, Barreto FK, Baptista E et al (2013) Molecular study of HBZ and gp21 human T cell leukemia virus type 1 proteins isolated from different clinical profile infected individuals. *AIDS Res Hum Retroviruses* 29:1370–1372. <https://doi.org/10.1089/AID.2013.0015>
  67. Machado LY, Navea LM, Díaz HM et al (2013) Phylogenetic analysis of human T cell lymphotropic virus type 1 isolated from Cuban individuals. *AIDS Res Hum Retroviruses* 29:1168–1172. <https://doi.org/10.1089/AID.2012.0225>
  68. Bandeira LM, Uehara SNO, Asato MA et al (2015) High prevalence of HTLV-1 infection among Japanese immigrants in non-endemic area of Brazil. *PLoS Negl Trop Dis* 9:e0003691. <https://doi.org/10.1371/journal.pntd.0003691>
  69. Rafatpanah H, Torkamani M, Valizadeh N et al (2016) Prevalence and phylogenetic analysis of HTLV-1 in a segregated population in Iran. *J Med Virol* 88:1247–1253. <https://doi.org/10.1002/jmv.24448>
  70. de Aguiar SA, de Souza França SA, Santana BB et al (2017) Human T-lymphotropic virus 1aA circulation and risk factors for sexually transmitted infections in an Amazon geographic area with lowest human development index (Marajó Island, Northern Brazil). *BMC Infect Dis*. <https://doi.org/10.1186/s12879-017-2859-x>
  71. Bandeira LM, Uehara SNO, Puga MAM et al (2018) HTLV-1 intrafamilial transmission among Japanese immigrants in Brazil. *J Med Virol* 90:351–357. <https://doi.org/10.1002/jmv.24938>
  72. Ribeiro IP, Kozlowski AG, Dias de Matos MA et al (2018) HTLV-1 and -2 in a first-time blood donor population in Northeastern Brazil: prevalence, molecular characterization, and evidence of intrafamilial transmission. *J Med Virol* 90:1651–1657. <https://doi.org/10.1002/jmv.25231>
  73. Mirhosseini A, Mohareri M, Arab R et al (2019) Complete sequence of human T cell leukemia virus type 1 in ATLL patients from Northeast Iran, Mashhad revealed a prematurely terminated protease and an elongated pX open reading frame III. *Infect Genet Evol* 73:460–469. <https://doi.org/10.1016/j.meegid.2019.05.012>
  74. Campos KR, Caterino-de-Araujo A (2020) Provirus mutations of human T-lymphotropic virus 1 and 2 (HTLV-1 and HTLV-2) in HIV-1-coinfected individuals. *mSphere*. <https://doi.org/10.1128/mSphere.00923-20>
  75. de Rego FFA, de Oliveira T, Giovanetti M et al (2016) Deep sequencing analysis of human T cell lymphotropic virus type 1 long terminal repeat 5' region from patients with tropical spastic paraparesis/human T cell lymphotropic virus type 1-associated myelopathy and asymptomatic carriers. *AIDS Res Hum Retroviruses*. <https://doi.org/10.1089/aid.2015.0273>
  76. Aleluia MM, Mello MAG, Alcântara LCJ et al (2015) The origin of HTLV-1 in southern Bahia by phylogenetic, mtDNA and  $\beta$ -globin analysis. *Virology* 5:63–74. <https://doi.org/10.1016/j.virep.2015.05.002>
  77. Bandeira LM, Puga MAM, Weis-Torres SMS et al (2021) Human T-cell leukemia virus type 1 infection among Japanese immigrants and their descendants living in Southeast Brazil: a call for preventive and control responses. *PLoS Negl Trop Dis* 15:e0009066. <https://doi.org/10.1371/journal.pntd.0009066>
  78. Pessôa R, Watanabe JT, Nukui Y et al (2014) Molecular characterization of human T-cell lymphotropic virus type 1 full and partial genomes by illumina massively parallel sequencing technology. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0093374>
  79. Araújo THA, Barreto FK, Menezes ADL et al (2020) Complete genome sequence of human T-cell lymphotropic type 1 from patients with different clinical profiles, including infective dermatitis. *Infect Genet Evol* 79:104166. <https://doi.org/10.1016/j.meegid.2019.104166>
  80. Nguyen Quang N, Goudey S, Ségéral E et al (2020) Dynamic nanopore long-read sequencing analysis of HIV-1 splicing events during the early steps of infection. *Retrovirology*. <https://doi.org/10.1186/s12977-020-00533-1>
  81. Cabral F, Arruda LB, de Araújo ML et al (2012) Detection of human T-cell lymphotropic virus type 1 in plasma samples. *Virus Res* 163:87–90. <https://doi.org/10.1016/j.virusres.2011.08.014>
  82. Demontis MA, Sadiq MT, Golz S, Taylor GP (2015) HTLV-1 viral RNA is detected rarely in plasma of HTLV-1 infected subjects. *J Med Virol* 87:2130–2134. <https://doi.org/10.1002/jmv.24264>
  83. Liu L, Li Y, Li S et al (2012) Comparison of next-generation sequencing systems. *J Biomed Biotechnol*. <https://doi.org/10.1155/2012/251364>
  84. Bangham CRM, Cook LB, Melamed A (2014) HTLV-1 clonality in adult T-cell leukaemia and non-malignant HTLV-1 infection. *Semin Cancer Biol* 26:89–98. <https://doi.org/10.1016/j.semcancer.2013.11.003>
  85. Wattel E, Vartanian JP, Pannetier C, Wain-Hobson S (1995) Clonal expansion of human T-cell leukemia virus type I-infected cells in asymptomatic and symptomatic carriers without malignancy. *J Virol* 69:2863–2868
  86. Capobianchi MR, Giombini E, Rozera G (2013) Next-generation sequencing technology in clinical virology. *Clin Microbiol Infect* 19:15–22. <https://doi.org/10.1111/1469-0691.12056>
  87. Panfil AR, Al-Saleem JJ, Green PL (2013) Animal models utilized in HTLV-1 research. *Virology (Auckl)* 4:49–59. <https://doi.org/10.4137/VRT.S12140>
  88. Heather JM, Chain B (2016) The sequence of sequencers: the history of sequencing DNA. *Genomics* 107:1–8. <https://doi.org/10.1016/j.ygeno.2015.11.003>
  89. Martin F, Tagaya Y, Gallo R (2018) Time to eradicate HTLV-1: an open letter to WHO. *Lancet* 391:1893–1894. [https://doi.org/10.1016/S0140-6736\(18\)30974-7](https://doi.org/10.1016/S0140-6736(18)30974-7)

90. Cao F, Ji Y, Huang R et al (2000) Sequence note: nucleotide sequence analyses of partial envgp46 gene of human T-lymphotropic virus type I from inhabitants of Fujian Province in Southeast China. *AIDS Res Hum Retroviruses* 16:921–923. <https://doi.org/10.1089/08892220050042855>
91. Chou SK, Okayama A, Tachibana N et al (1995) Nucleotide sequence analysis of a full-length human T-cell leukemia virus type I from adult T-cell leukemia cells: a prematurely terminated

PX open reading frame II. *Int J Cancer* 60:701–706. <https://doi.org/10.1002/ijc.2910600522>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.