

ESCOLA BAHIANA DE MEDICINA E SAÚDE PÚBLICA
CURSO BIOMEDICINA

PEDRO GONDIM DA SILVEIRA

CENÁRIO CLÍNICO, MOLECULAR E EPIDEMIOLÓGICO DA
INFECÇÃO POR HTLV-1 A PARTIR DA MINERAÇÃO DE DADOS

SALVADOR – BA

2020

PEDRO GONDIM DA SILVEIRA

**CENÁRIO CLÍNICO, MOLECULAR E EPIDEMIOLÓGICO DA
INFECÇÃO POR HTLV-1 A PARTIR DA MINERAÇÃO DE DADOS**

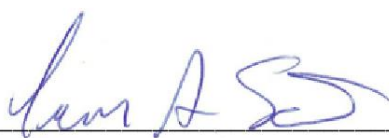
Trabalho de Conclusão de Curso apresentado à
Escola Bahiana de Medicina e Saúde Pública,
como parte dos requisitos para obtenção do
título de Bacharel em Biomedicina.

PEDRO GONDIM DA SILVEIRA

**CENÁRIO CLÍNICO, MOLECULAR E EPIDEMIOLÓGICO DA
INFECÇÃO POR HTLV-1 A PARTIR DA MINERAÇÃO DE DADOS**

Este Trabalho de Conclusão de Curso foi julgado adequado à obtenção do grau de Bacharel em Biomedicina e aprovada em sua forma final pelo Curso de Biomedicina da Escola Bahiana de Medicina e Saúde Pública.

Salvador – BA, 26 de junho de 2020.



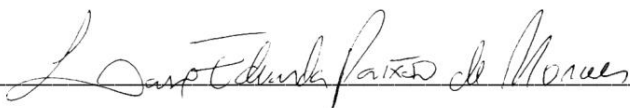
Profa. Dra. Luciane Amorim Santos

Escola Bahiana de Medicina e Saúde Pública



Profa. Dra. Thessika Hialla Almeida Araújo

Escola Bahiana de Medicina e Saúde Pública



Profa. Me. Laise Eduarda Paixão de Moraes

Fundação Oswaldo Cruz, Instituto Gonçalo Moniz – BA

AGRADECIMENTOS

Aos meus pais por fornecerem todo o apoio emocional e financeiro durante toda minha vida.

Além disso, sempre foram exemplos de pessoas humanas, me ensinaram o caminho correto a se trilhar, respeitando o próximo para poder ser respeitado.

A minha tia avó, por ter cuidado de mim como se fosse seu próprio filho.

A minha orientadora pelo apoio em todo os trabalhos em que fizemos, sempre educada e disposta a ensinar.

Aos meus verdadeiros amigos pela companhia nos momentos alegres e, sobretudo, nos momentos mais difíceis.

SUMÁRIO

1. Página de Título.....6

2. Resumo.....7

3. Introdução.....9

4. Metodologia.....10

5. Resultados.....11

6. Discussão.....13

7. Conflitos de interesse.....15

8. Referências.....16

9. Anexos.....19

10. Revista.....22

CENÁRIO CLÍNICO, MOLECULAR E EPIDEMIOLÓGICO DA INFECÇÃO POR HTLV-1 A PARTIR DA MINERAÇÃO DE DADOS

Pedro Gondim da Silveira^[1], Thessika Hialla Almeida Araújo^[1]

[1] Escola Bahiana de Medicina e Saúde Pública, Salvador, Bahia, Brasil

RESUMO

Introdução

O Human T-cell lymphotropic virus type 1 (HTLV-1) foi o primeiro retrovírus humano descrito associados a diversas doenças. Estima-se que 5 a 10 milhões de pessoas estejam infectadas pelo HTLV-1 em todo o mundo. o presente trabalho tem como objetivo descrever os cenários clínicos, moleculares e epidemiológicos da infecção pelo HTLV-1 utilizando o método de busca ativa em banco de dados biológicos (primários e secundários) e artigos.

Metodologia

A coleta de dados foi realizada em bancos de dados primário e secundário (HTLV-1 Molecular Epidemiology Database e Genbank) e artigos científicos de forma manual ou utilizando softwares específicos. Posteriormente, os dados foram tabulados e analisados no Excel (Microsoft, United States).

Resultados

Um total de 6840 sequências de genoma completo e parciais foram coletadas dos bancos de dados utilizados. O total de 6282(91,8%) das sequências tinham informação sobre a região geográfica e apenas 42% das sequências apresentou alguma informação sobre o perfil clínico. A maior parte das sequências (80%) se encaixa ao subtipo Cosmopolita (HTLV-1a) quando relacionada à região geográfica. Os dados étnicos, idade e sexo se mantiveram escassos.

Conclusão

Embora as tecnologias atuais tenham contribuído no desenvolvimento desse trabalho, permitindo entender um pouco do cenário atual da infecção pelo HTLV-1, a escassez dos dados de algumas variáveis ainda é um grande limitador. Os autores devem fornecer a quantidade máxima de informações possível, pois os dados moleculares podem ser favoráveis para muitas inferências diferentes sobre a infecção pelo HTLV-1 e, sobretudo úteis para o planejamento de políticas públicas de saúde.

Palavras-chave: Mineração de dados, HTLV-1, Banco de dados.

ABSTRACT

Introduction

Human T-cell lymphotropic virus type 1 (HTLV-1) was the first described human retrovirus associated with numerous diseases. It estimates that 5 to 10 million people are infected with HTLV-1 worldwide. The present work aims to describe the clinical, molecular and epidemiological scenarios of HTLV-1 infection using the active search method in biological databases (primary and secondary) and articles.

Methodology

Data collection was performed in primary and secondary databases (HTLV-1 Molecular Epidemiology Database and Genbank) and scientific articles manually or using specific software. Subsequently, the data were tabulated and analyzed in Excel (Microsoft, United States).

Results

A total of 6840 complete and partial genome sequences were collected from the databases used. A total of 6282 (91.8%) of the sequences had information about the geographic region and only 42% of the sequences had some information about the clinical profile. Most of the sequences (80%) fit the Cosmopolitan subtype (HTLV-1a) when related to the geographic region. Ethnic data, age and sex remained scarce.

Conclusion

Although current technologies have contributed to the development of this work, allowing us to understand a little about the current scenario of HTLV-1 infection, the scarcity of data for some variables is still a major limitation. Authors should provide as much information as possible, as molecular data can be favorable for many different inferences about HTLV-1 infection and, above all, useful for planning public health policies.

Keywords: Data mining, HTLV-1, Database.

INTRODUÇÃO:

O Human T-cell lymphotropic virus type 1 (HTLV-1) foi o primeiro retrovírus humano descrito¹. Desde a sua identificação até os estudos mais atuais, tem sido relacionado à inúmeras manifestações clínicas tais como: Mielopatia associada ao HTLV-1/paraparesia espástica tropical (HAM/TSP)², leucemia/linfoma de células T do adulto (ATLL)³, uveíte associada ao HTLV-1⁴, dermatite infecciosa⁵, além de doenças reumáticas como síndrome de Sjögren e artrite reumatoide^{6,7}.

Estima-se que aproximadamente 5 a 10 milhões de pessoas estejam infectadas pelo HTLV-1 em todo o mundo⁸. Dados epidemiológicos mostram que a infecção por este vírus tem distribuição mundial⁹, no entanto, algumas áreas são endêmicas para esta infecção: sudoeste do Japão^{10,11}, África sub-Saariana¹², regiões do Caribe¹³, áreas localizadas no Irã e Melanésia¹⁴ e Brasil¹⁵.

O real cenário da prevalência da infecção por HTLV-1 pode está subestimado, levando-se em consideração que a maioria dos estudos epidemiológicos são realizados em populações específicas, principalmente em doadores de sangue, gestantes e indivíduos hospitalizados⁸.

Independentemente disto, a epidemiologia do HTLV-1 ainda apresenta muitos desafios. Estudos moleculares contribuíram e ainda contribuem para a aquisição de conhecimentos sobre a distribuição do vírus e as suas características moleculares, a partir da geração de sequências nucleotídicas, e o gerenciamento e mineração apropriado destes dados, a partir bancos de dados primários e secundários, podem fornecer informações consistentes adicionais sobre a infecção por HTLV-1, abrindo perspectivas de novos conhecimentos biomédicos e de saúde pública¹⁶.

Em resposta à necessidade de obter mais informações sobre as sequências de HTLV-1 já geradas e disponíveis, além de dados que possam apoiar a compreensão da patogênese viral, a via de transmissão, polimorfismos, relações genótipo-fenótipo, distribuição geográfica e evolução viral, assim, o presente trabalho tem como objetivo descrever os cenários clínicos, moleculares e epidemiológicos da infecção pelo HTLV-1 utilizando o método de busca ativa em banco de dados biológicos (primários e secundários) e artigos.

METODOLOGIA:

Bancos de Dados de Epidemiologia Molecular do HTLV-1

Inicialmente, foi feita uma lista com as variáveis (idade, sexo, perfil clínico, genótipo, região geográfica, etnia, identificação de isolado e região gênica) mais frequentes no banco de dados de Epidemiologia Molecular do HTLV-1¹⁷. Posteriormente, essas variáveis foram combinadas utilizando o algoritmo de busca implementado no banco de dados, que também permite o download e organização de planilhas formato de arquivo *xls*.

Genbank

Em seguida, buscamos os dados das sequências mais atualizadas no banco de nucleotídeos Genbank¹⁸ até o dia 23 janeiro de 2020. Para isto, foi realizada a pesquisa em modo avançado usando o descritor “HTLV-1[Organism]” e posterior *download* dos números de acessos em formato *FASTA*. A extração dos metadados presentes no arquivo foi obtida através do software gbmunge (<https://github.com/sdwfrost/gbmunge>)¹⁹, capaz de levantar dados relativos à data (publicação e coleta), país de origem, tamanho da sequência, e número de acesso, gerando assim, uma segunda planilha no formato de arquivo *xls*.

Por fim, combinamos as duas planilhas e excluimos as duplicações encontradas entre dos bancos de dados. As informações ausentes tais como: nome do isolado, sexo, idade, etnia, estado clínico, subtipo, subgrupo e região gênica) foram garimpadas de forma manual em artigos. Além disto, os genomas completos e as sequências LTR (*Long Terminal Repeat*) foram submetidas à subtipagem na plataforma Genome Detective²⁰. Após a etapa inicial de coleta e organização dos dados, foram criadas planilhas utilizando o Excel (Microsoft, United States) e subsequentemente, foram feitas as análises descritivas e as tabelas relacionadas.

RESULTADOS:

Atualmente, a soma dos números de sequências presentes nos dois bancos de dados analisados representa 6.840, dos quais 6.711 (98,1%) possuem informação sobre identificação do isolado e 129 (1,9%) sem esta informação, entretanto, também foram incluídas nas análises.

Região genômica do HTLV-1 – Através da busca ativa foi encontrado um total de 6.786 (99%) de sequências virais com informações de região genômica (Fig. I). Sendo a região da proteína PX a mais frequente, seguido dos LTR (long terminal repeat) e região do envelope viral (ENV) respectivamente. Houve um total de 274 sequências com mais de uma região genômica descrita nos bancos de dados.

Região geográfica dos isolados de HTLV-1 – Entre as 6.840 sequências, 6.282 (91,84%) possuem caracterização geográfica. Na distribuição continental (Fig. II), a América do Sul teve maior número de sequências (54,08%, 3.397/6.282), seguido da África (18,97%, 1.192/6.282), Ásia (16,28%, 1.023/6.282), América Central (3,90%, 245/6.282), Oceania (2,74%, 172/6.282), América do Norte (2,47%, 155/6.282) e Europa (1,56%, 98/6.282). É importante destacar que, em relação à América do Sul, o Brasil é o local com maior representação das sequências virais disponíveis (53,66%, 1.823/3.397), seguido da Argentina (9,53%, 324/3.397).

Região geográfica versus genótipo - A associação dos dados entre região geográfica e genótipo mostrou um total de 2.805 sequências virais (41%). O resultado indicou que a predominância mundial foi do subtipo Cosmopolita “a” (81,92%, 2.298/2.805), seguido de subtipo “b” (15,4%, 432/2.805), “c” (3,6%, 101/2.805), “d” (0,64%, 18/2.805), “e” (0,07%, 2/2.805), “f” (0,46%, 13/2.805), “g” (0,39%, 11/2.805); apresentados na Fig. III). Em relação ao subtipo Cosmopolita “a”, houve uma frequência maior do subgrupo “A” transcontinental (72,15%, 1658/2.298), seguido de subgrupo “B” (5,5%, 127/2.298), “D” (4,35%, 100/2.298), “C” (1%, 23/2.298) e “E” (0,17%, 4/2.298).

Perfil clínico - A investigação sobre o perfil clínico mostrou que um total de 2.873 sequências armazenadas apresentavam informações de estado clínico associado ao HTLV-1. A predominância foi de indivíduos assintomáticos (46%, 1338/2.873), seguido de HAM/TSP

(41%, 1168/2.873) e ATLL (10%, 292/2.873). Também foram relatadas doenças, como: dermatites, cânceres, paralisias, síndrome de Sjögren, diabetes, tuberculose, micose fúngica, estrongiloidíase, síndromes oculares e doenças cardíacas gerais, as quais em conjunto, apresentaram um total de 3% (75/2.873) das sequências virais, representados pela Fig. IV.

Sexo, idade e etnias - Utilizando a variável sexo, 14,5% (993/6840) dessas sequências possuíam informações, sendo que houve uma predominância do sexo feminino (57%, 532/993). Para a variável idade, 8,7% (594/6840) das sequências possuíam informações. A média de idade foi 44,8 anos, variando de 3 a 100 anos, quando em valor crescente a mediana é de 45 anos de idade e distância interquartil de 25 anos.

Os dados relativos à etnia são muito escassos 3,2% (221/6840), sendo que pretos foram mais frequentes (46,1%, 102/221), seguido de indígenas (37,5%, 83/221), caucasianos (9,5%, 21/221), asiáticos (6,3%, 14/221) e raças mistas (0,6%, 1/221).

A análise conjunta das variáveis perfil clínico, sexo, subtipos e região geográfica pode ser visualizada na Tabela I. Foi encontrada uma maior prevalência de indivíduos assintomáticos, do sexo feminino e sequências do subtipo Cosmopolita “a”. A região com mais informações sobre sequências foi a América do Sul (42,6%), seguido da África (37%.

Visto que, a maior porcentagem de dados moleculares e epidemiológicos, relacionados a infecção pelo HTLV-1, foram encontrados relacionados com a América do Sul, quando observamos informações de perfil clínico versus genótipo viral, notamos que há 244 sequências virais. Todas as 244 sequências foram identificadas, nas anotações dos bancos de dados, como subtipo “a”, e quando avaliamos a classificação dos subgrupos, encontramos uma maior predominância do subgrupo “A” (90,57%, 221/244), sendo que 61,06% (149/244) são sequências oriundas de indivíduos assintomáticos; 29,5% (72/244) de HAM/TSP; 5,32% (13/244) de ATLL e 4,1% (10/244) de outras manifestações clínicas. Para os demais subgrupos encontrados, 8,6% (21/244) do subgrupo “b” 90,47% (19/21) de indivíduos assintomáticos e

9,52% (2/21) de ATLL), e para finalizar, 0,82% (2/244) para os subgrupos “C” e “E”, ambas caracterizadas pelo perfil clínico HAM/TSP.

DISCUSSÃO

A coleta e análise das sequências selecionadas nos bancos de dados, demonstraram um aumento de 3 (três) vezes no número total de sequências, quando comparado à um estudo realizado em 2014, com o mesmo propósito de inferir o cenário global da infecção do HTLV-1 através de *data mining*²¹. Houve um aumento considerável durante o curto período de 6 anos, todavia, ainda continua deficiente a quantidade de informações das variáveis que avaliam o cenário epidemiológico do HTLV-1, mesmo após a implementação das políticas de submissão BioProject e BioSample, em 2011²².

A alta representatividade da região pX está relacionada com os estudos da patogênese viral, enquanto as regiões LTR e env do HTLV-1, por apresentarem uma maior variabilidade genética em relação às outras regiões genômicas, são utilizadas na caracterização genética dos subtipos virais²³. Mais da metade das sequências achadas nos bancos de dados são do gene LTR, env ou genoma completo, o que possibilita a genotipagem a partir de algoritmos bem definidos e já descritos na literatura²⁰.

O perfil clínico apresentou uma mudança considerável, além do número absoluto total de sequências virais com informação de perfil clínico terem aumentado, assintomáticos estão mais frequentes nos achados, seguidos de HAM/TSP e ATLL respectivamente. Entretanto, as variáveis como: sexo, idade, etnia, subtipo e subgrupos continuam com uma grande carência de dados, a mesma dificuldade encontrada nos estudos anteriores^{17, 21}. Essa observação ratifica a necessidade de os autores fornecerem a quantidade máxima de informações possíveis, pois os dados moleculares, clínicos e epidemiológicos podem ser úteis para muitas inferências em relação à infecção pelo HTLV-1 e, portanto, úteis para incentivar as políticas de prevenção, sobretudo facilitando a busca através de softwares.

A infecção pelo HTLV-1 é predominantemente cosmopolita transcontinental e heterogênea, sendo endêmica sobretudo na América do Sul, na África, Ásia e Oceania^{24, 25}. Foi possível observar esta distribuição heterogênea nos bancos selecionados para esse estudo em cada região geográfica. Entretanto houve aumento no número de sequências da região da África, ultrapassando a Ásia em comparação com estudos de 2014²¹, assim como a região da América Central que ultrapassou a região endêmica da Oceania. A alta prevalência de infecção pelo HTLV-1 na América do Sul, África e Ásia²⁴, condiz com a grande quantidade de sequências reunidas no banco de dados.

Os estudos das sequências relacionadas a América do Sul, demonstraram ser mais da metade do total. O Brasil é o país com maior quantidade de sequências desse continente (53,66%) e, portanto, possui uma maior quantidade de dados epidemiológicos e moleculares da infecção por HTLV-1. Por conta disso, torna-se necessária maiores estudos nos outros países dessa região.

Embora a infecção por HTLV-I ocorra em todo o mundo, a distribuição varia de acordo com a localização geográfica, fatores étnicos e raciais ou em grupos populacionais mais expostos aos fatores de risco²⁶. Apesar disto, não foi possível apontar inferências sobre sexo, idade e etnia devido à escassez de dados. Todavia, estudos étnicos epidemiológicos, demonstraram a associação da infecção do HTLV-1 e o desenvolvimento de ATLL ou HAM/TSP²⁷. Portanto, essas informações devem ser investigadas em casos de infecção por HTLV-1, sobretudo estudos do background genético de indivíduos infectados analisando os determinantes polimórficos de alelos de antígenos leucocitários humanos e sua resposta imune ao HTLV-1.

O banco de dados de epidemiologia molecular do HTLV-1 e Genbank^{17,18}, e os softwares utilizados nesse estudo, permitiram uma busca e análise, com grande quantidade de dados de forma eficiente. Portanto, é muito importante fomentar estudos na área de bioinformática, em desenvolvimentos de softwares avançados e projetos de bancos de dados. Além disto, dados moleculares e epidemiológicos para sequências virais devem ser oferecidos com mais

frequência, já que essas informações podem ser usadas para o entendimento geral da infecção pelo HTLV-1, auxiliando no planejamento de políticas públicas de saúde²¹.

CONFLITO DE INTERESSES:

O autor declara que não houve conflitos de interesses.

REFERÊNCIAS:

1. Poiesz BJ, Ruscetti FW, Gazdar AF, Bunn PA, Minna JD, Gallo RC. 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*, 1980, 77: 7415-7419.
2. Gessain A, Vernant JC, Maurs L, Barin F, Gout O, Calender A, de Thé G. ANTIBODIES TO HUMAN T-LYMPHOTROPIC VIRUS TYPE-I IN PATIENTS WITH TROPICAL SPASTIC PARAPARESIS. *Lancet*, 1985, 326: 407-410.
3. Uchiyama T, Yodoi J, Sagawa K, Takatsuki K, Uchino H. Adult T-cell leukemia: clinical and hematologic features of 16 cases. *Blood*. 1977;50(3):481-492.
4. Mochizuki M, Watanabe T, Yamaguchi K, Takatsuki K, Yoshimura K, Shirao M, et al. HTLV-I Uveitis: a distinct clinical entity caused by htlv-i.: A Distinct Clinical Entity Caused by HTLV-I. *Jpn J Cancer Res*, 1992, 83: 236-239.
5. La Grenade L. HTLV-I-Associated Infective Dermatitis: past, present, and future.: Past, Present, and Future. *J Acquir Immune Defic Syndr Hum Retrovirol*, 1996, 13: 46-49.
6. Mccalum RM, Patel DD, Moore JO, Haynes BF. ARTHRITIS SYNDROMES ASSOCIATED WITH HUMAN T CELL LYMPHOTROPIC VIRUS TYPE I INFECTION. *Med Clin North Am*, 1997, 81: 261-276.
7. Nishioka K. HTLV-I Arthropathy and Sjögren Syndrome. *J Acquir Immune Defic Syndr Hum Retrovirol*, 1996, 13: 57-62.
8. Gessain A, Cassar O. Epidemiological Aspects and World Distribution of HTLV-1 Infection. *Front Microbiol*, 2012, 3: 388-388.
9. de Thé G, Kazanji M. An HTLV-I/II Vaccine: from animal models to clinical trials?. *J Acquir Immune Defic Syndr Hum Retrovirol*, 1996, 13: 191-198.
10. Yamaguchi K. Human T-lymphotropic virus type I in Japan. *Lancet*, 1994, 343: 213-216.

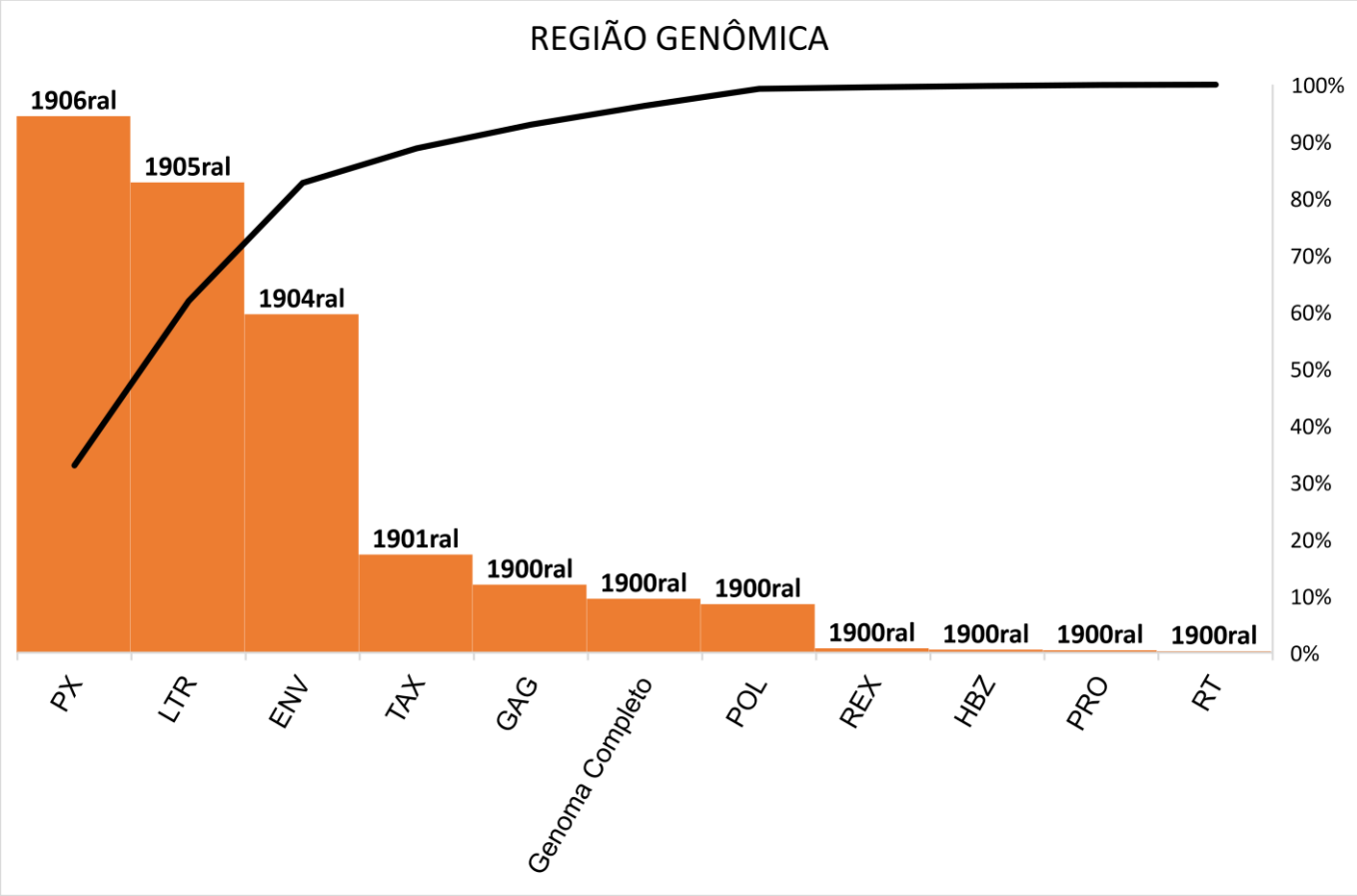
11. Mueller N, Okayama A; Stuver S; Tachibana N. Findings from the Miyazaki Cohort Study. *J Acquir Immune Defic Syndr Hum Retrovirol*, 1996, 13: 2-7.
12. Gessain A, Mahieux R, de Thé G. Genetic variability and molecular epidemiology of human and simian T cell leukemia/lymphoma virus type I. *J Acquir Immune Defic Syndr Hum Retrovirol*, 1996, 13: S132-145.
13. Oliveira PD, de Carvalho RF, Bittencourt AL. Adult T-cell leukemia/lymphoma in South and Central America and the Caribbean: systematic search and review.: systematic search and review. *Int J STD AIDS*, 2016, 28: 217-228.
14. Mueller N. The epidemiology of HTLV-I infection. *Cancer Causes Control*, 1991, 2: 37-52.
15. Catalan-Soares B, Carneiro-Proietti ABF, Proietti FA. Heterogeneous geographic distribution of human T-cell lymphotropic viruses I and II (HTLV-I/II): serological screening prevalence rates in blood donors from large urban areas in brazil.: serological screening prevalence rates in blood donors from large urban areas in Brazil. *Cad Saude Publica*, 2005, 21: 926-931.
16. Yoo I, Alafaireet P, Marinov M, Pena-Hernandez K, Gopidi R, Chang J, Hua L. Data Mining in Healthcare and Biomedicine: a survey of the literature.: A Survey of the Literature. *J Med Syst*, 2011, 36: 2431-2448.
17. Araujo THA, Souza-Brito LI, Libin P, Deforche K, Edwards D, de Albuquerque-Junior AE, et al. A Public HTLV-1 Molecular Epidemiology Database for Sequence Management and Data Mining. Belshaw R, editor. *PLoS ONE*. 2012;7: e42123.
18. Sayers EW, Cavanaugh M, Clark K, Ostell J, Pruitt KD, Karsch-Mizrachi I. GenBank. *Nucleic Acids Res*, 2018, 47: D94-D99.
19. Simon Frost. GitHub. Redmond, WA, USA; 2017 [updated 2018 Jul 6]. Available from: <https://github.com/sdwfrost/gbmunge>.
20. Cleemput S, Dumon W, Fonseca V, Karim WA, Giovanetti M, Alcantara LC, et al. Genome Detective Coronavirus Typing Tool for rapid identification and

- characterization of novel coronavirus genomes. *Bioinformatics*. 2020, 36: 3552–3555.
21. Araujo THA, Barreto FK, Luiz Carlos Júnior A, Miranda ACAM. Inferences about the global scenario of human T-cell lymphotropic virus type 1 infection using data mining of viral sequences. *Memórias do Instituto Oswaldo Cruz*. 2014, 109: 448–451.
 22. Barrett,T., Clark,K., Gevorgyan,R., Gorelenkov,V., Gribov,E., Karsch-Mizrachi,I., et al. BioProject and BioSample databases at NCBI: facilitating capture and organization of metadata. *Nucleic Acids Res*, 2011, 40:D57–D63.
 23. Santos FLN, Lima F WM. Epidemiology, physiopathogenesis and laboratorial diagnosis of the HTLV-I infection. *J Bras Patol Med Lab*. 2005, 41:105-16.
 24. Gessain, A. Le rétrovirus humain oncogène HTLV-1: épidémiologie descriptive et moléculaire, origine, évolution et aspects diagnostiques et maladies associées. *Bull. Soc. Pathol. Exot*. 2011, 104:167.
 25. Proietti FA, Carneiro-Proietti ABF, Catalan-Soares BC, Murphy EL. Global epidemiology of HTLV-I infection and associated diseases. *Oncogene*. 2005, 24: 6058-68
 26. Blattner WA, Blayney DW, Robert-Guroff M, Sarngadharam MG, Kalyanaraman VS. Epidemiology of human T-cell leukemia/lymphoma virus. *Journal of Infectious Diseases*. 1983, 29:406-416.
 27. Sonoda S, Li HC, Tajima K. Ethnoepidemiology of HTLV-1 related diseases: ethnic determinants of HTLV-1 susceptibility and its worldwide dispersal. *Cancer Sci*. 2011,102: 295-301.

ANEXOS

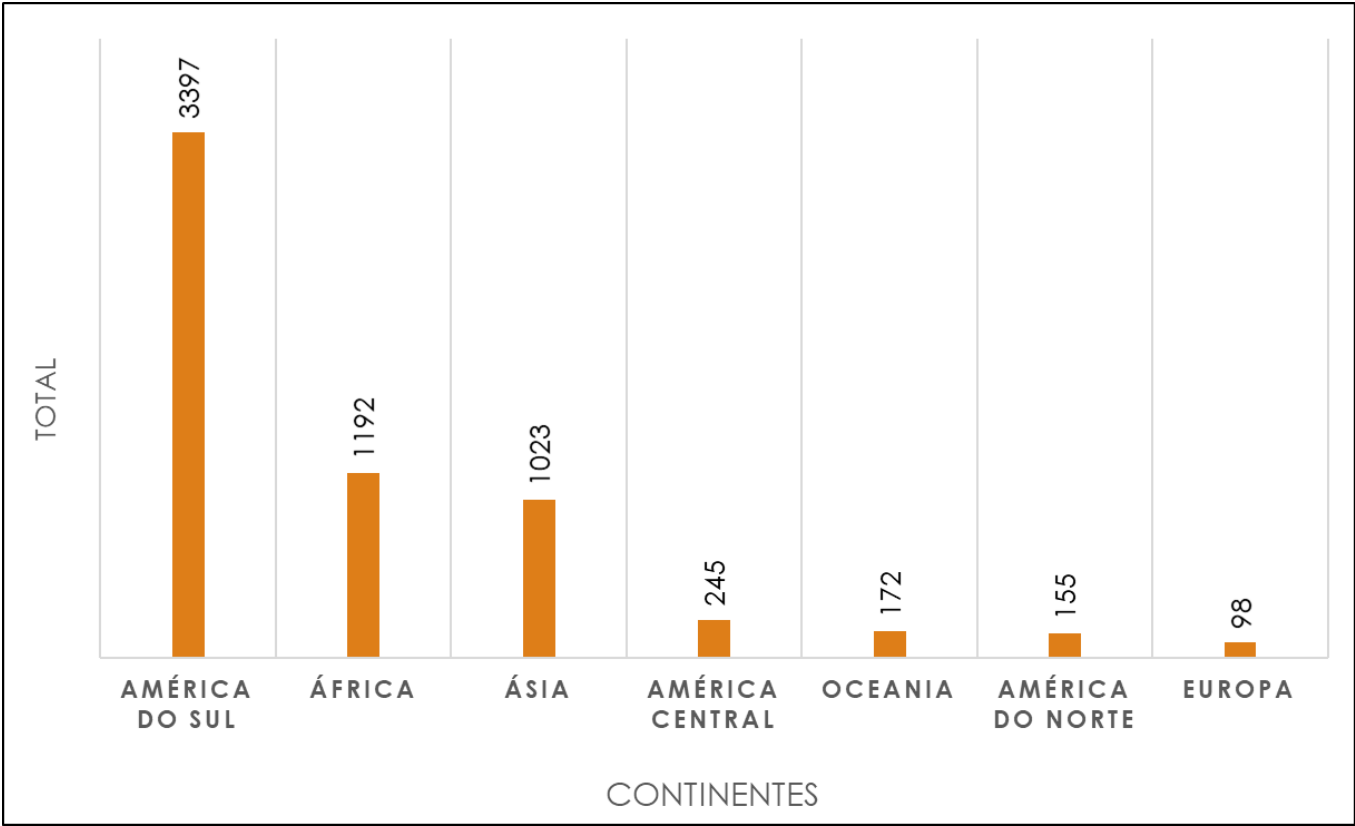
ANEXO A – Figura I

Fig. I: Distribuição das regiões genômicas em ordem decrescente de frequência das sequências achadas nos bancos de dados de Epidemiologia Molecular do HTLV-1 e Genbank.



ANEXO B – Figura II

Fig. II: Distribuição geográfica de 6282 isolados encontrados nos bancos de dados de Epidemiologia Molecular do HTLV-1 e Genbank.



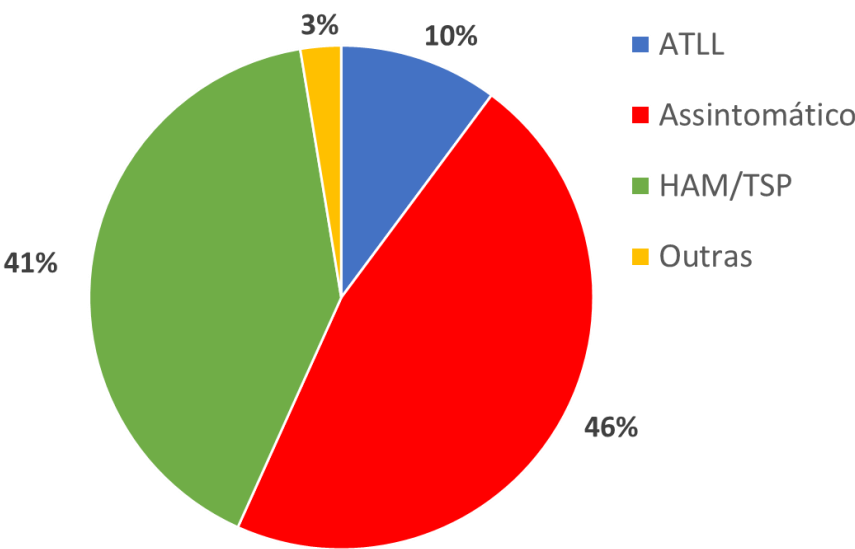
ANEXO C – Figura III

Fig. III: Distribuição geográfica do número absoluto de sequências do vírus linfotrópico de células T humanas tipo 1 (HTLV-1) armazenadas no Banco de Dados de Epidemiologia Molecular do HTLV-1 e Genbank, usando a variável subtipo.



ANEXO D – Figura IV

Fig. IV: Gráfico de porcentagem dos perfis clínicos achados nos bancos de dados de Epidemiologia Molecular do HTLV-1 e Genbank.



ANEXO E – Tabela

TABELA Distribuição do gênero, subtipo e origem geográfica entre os estados clínicos: paraparesia espástica tropical / mielopatia associada ao vírus linfotrópico de células T humanas tipo 1 (HTLV-1) (TSP / HAM), leucemia / linfoma de células T adultas (ATL), assintomáticos (ASS) e outras doença associada ao HTLV-1.

Estado Clínico	TSP/HAM n = 1168 n (%)	ATL n = 292 n (%)	ASS n = 1338 n (%)	Outros ^a n = 76 n (%)
Sequências com informação sobre sexo	98 (8.4)	47 (16.1)	148 (11.1)	26 (34.2)
Homem	37 (37.8)	32 (68.1)	61 (41.2)	13 (50)
Mulher	61 (62.2)	15 (31.9)	87 (58.8)	13(50)
Sequências com informação de Subtipo	177 (15.2)	102(34.9)	263 (19.7)	36 (47.4)
Subtipo				
“a”	173 (97.7)	94(92.2)	222 (84.4)	35 (97.2)
“b”	4 (2.3)	6(5.9)	38 (14.4)	NA
Outras	NA	2 (1.9)	3 (1.1)	1 (2.8)
Sequências com informação sobre origem geográfica	988 (84.6)	244 (83.6)	1223 (91.4)	74(97.4)

a: sequências de pacientes com Dermatites, Cânceres, Paralisias, Síndrome de Sjögren, Diabetes, Tuberculose, Micose fúngica, Estrongiloidíase, Síndrome ocular, e Doenças Cardíacas Gerais. NA: Informações não encontradas

Revista: Revista Da Sociedade Brasileira De Medicina Tropical/Journal Of The Brazilian Society Of Tropical Medicine

Regras para Submissão:

O manuscrito deve ser preparado usando *software* padrão de processamento de textos (Word) e salvo como arquivos .DOC ou .DOCX. A fonte preferencial é *Times New Roman* tamanho 12, com espaço duplo em todo o texto, título/legendas para as figuras, e referências, margens com pelos menos 3cm. O manuscrito deve ser dividido nas seguintes seções: Cartão de Apresentação (endereçada ao Editor-Chefe), Página de Título, incluindo Título do manuscrito, título corrente, Resumo, Palavras-chaves, Texto do Manuscrito, Declaração de Conflito de Interesses, Agradecimentos, Suporte Financeiro, Lista de Referências, legenda das Figuras. A Carta de Apresentação, Página de Título, Agradecimentos e Suporte Financeiro devem ser incluídos em documentos separados. Abreviações devem ser usadas com moderação.

Página de Título: deve incluir o nome dos autores na ordem direta e sem abreviações, juntamente com afiliações institucionais na seguinte ordem: Instituição dos autores, Departamento, Cidade, Estado e País. Para autores brasileiros, favor não traduzir os nomes das instituições. O endereço completo do autor para correspondência deve ser especificado, incluindo telefone e e-mail. A quantidade de autores e coautores por manuscrito deve ser limitada ao número real de autores que realmente contribuíram com o manuscrito. Exceto para estudos multicêntricos nacionais e internacionais, até vinte autores e coautores serão permitidos. Os nomes dos autores remanescentes serão publicados em notas de rodapé. Forneça o número do ORCID do autor correspondente e de todos os coautores.

Potenciais revisores: Os autores devem fornecer os nomes e informações de contato (e-mail e afiliação institucional) de três potenciais revisores imparciais de uma instituição diferente da dos autores.

Título: deve ser conciso, claro e o mais informativo possível. Não deve conter abreviações e não deve exceder a 250 caracteres.

Título Corrente: com no máximo 100 caracteres.

Resumo Estruturado: deve sumarizar os resultados obtidos e as principais conclusões de tal forma que um leitor, não familiarizado com o assunto tratado no texto, consiga entender as implicações do artigo. O resumo não deve exceder 250 palavras e deve ser estruturado com os seguintes tópicos: Introdução, Métodos, Resultados e Conclusões. O uso de abreviações deve ser evitado.

Palavras-chaves: 3 a 6 palavras chaves devem ser listados imediatamente abaixo do resumo estruturado (Exemplo: Tuberculose. Cuidados primários de saúde. Estrutura de serviços.). Por favor visite o website <https://www.ncbi.nlm.nih.gov/mesh>

Introdução: A introdução do artigo deve ser curta e destacar os propósitos para o qual o estudo foi realizado. Estudos prévios devem ser citados somente quando essencial.

Métodos: devem ser claros e suficientemente detalhados para que os leitores e revisores possam compreender precisamente o que foi feito e permitir que seja repetido por outros. Técnicas-padrões precisam apenas ser citadas.

Ética: em caso de pesquisas em seres humanos, os autores devem indicar se os procedimentos realizados estão em acordo com os padrões éticos do comitê de experimentação em seres humanos (institucional, regional ou nacional) e de acordo com a Declaração de Helsinki de 1964, revisada em 1975, 1983, 1989, 1996 e 2000. Para experimentação em animais, o autor deve indicar se seguiu um guia do conselho nacional de pesquisa em experimentação animal ou se qualquer lei sobre o cuidado e uso de animais em laboratório foi seguida. O número de aprovação deve ser enviado à Revista. No caso de pesquisa em seres humanos, os autores devem incluir na seção métodos

(subtítulo Considerações Éticas) uma declaração de que o estudo foi aprovado pelo Comitê de Ética Institucional.

Ensaio Clínico: No caso de Ensaio Clínicos, o manuscrito deve ser acompanhado pelo número e órgão de registro do ensaio clínico (Plataforma REBEC).

Resultados: devem ser um relato conciso e impessoal da nova informação (todos os achados relevantes positivos e negativos) revelados pelo estudo. Evitar repetir no texto os dados apresentados em tabelas e ilustrações e usar a grafia do verbo no passado.

Discussão: A discussão deve ser limitada à significância das novas informações e logicamente argumentada, considerando a relevância clínica, importância e limitações do estudo. Não incluir uma revisão geral sobre o assunto. Mantenha a discussão concisa e relevante. As principais conclusões devem ser apresentadas no último parágrafo.

Agradecimentos: devem ser curtos, concisos e restritos àqueles realmente necessários, e que não atendam aos critérios de coautoria. No caso de órgãos de fomento, não usar siglas.

Conflito de Interesse: todos os autores devem revelar qualquer tipo de conflito de interesse existente durante o desenvolvimento do estudo.

Suporte Financeiro: informar todos os tipos de fomento recebidos de agências de fomento ou demais órgãos ou instituições financiadoras da pesquisa.

Referências: Apenas as referências citadas no texto devem ser incluídas na lista ao final do manuscrito. Devem ser numeradas consecutivamente em ordem progressiva, usando números em arábico, na medida em que aparecem no texto. A lista de referência deve ser formatada de acordo com o estilo Vancouver. Listar todos os autores quando houver até seis. Para sete ou mais, listar os seis primeiros, seguido por "et al". Digitar a lista de referências com espaçamento duplo, em folha separada e no final do manuscrito. Referências de comunicações pessoais, dados não publicados ou manuscritos "em preparação" ou "submetidos para publicação", não devem constar na lista de referência.

Artigos aceitos para publicação devem ser listados como "*in press*" e a carta de aceitação deve ser fornecida. Esse material pode ser incorporado em local apropriado no texto, entre parênteses da seguinte forma: (AB Figueiredo: Comunicação Pessoal, 1980); (CD Dias, EF Oliveira: dados não publicados).

Citações no texto devem ser feitas pelo respectivo número das referências, acima da palavra correspondente, em ordem numérica crescente, separados por vírgula ou por hífen. Ex.: Mundo^{1,2}; Vida^{30,42,44-50}. As referências no fim do manuscrito devem estar de acordo com o sistema de requisitos uniformes utilizado para manuscritos enviados para periódicos biomédicos (Consulte: <http://www.nlm.nih.gov/citingmedicine>). Os títulos dos periódicos devem ser abreviados de acordo com o estilo usado no *Index Medicus* (Consulte: <https://www.ncbi.nlm.nih.gov/nlmcatalog/journals>).

A responsabilidade pelas citações bibliográficas contidas no texto e na lista de referências recai exclusivamente sobre os autores.

Alguns exemplos de referências:

1. **Citação de Artigos em Geral:** Sobrenome, seguido das iniciais dos seis primeiros autores. Para sete ou mais autores, liste os seis primeiros, seguidos de "et al."), título completo do artigo (no idioma original), título abreviado do periódico (pode ser encontrado Em: <https://www.ncbi.nlm.nih.gov/nlmcatalog/journals>), ano de publicação, volume (número), páginas inicial e final abreviada.

Exemplo 1: Petitti DB, Crooks VC, Buckwalter JG, Chiu V. Blood pressure levels before dementia. Arch Neurol. 2005;62(1):112-6.

Exemplo 2: Freitas EC, Oliveira MF, Vasconcelos ASOB, Filho JDS, Viana CEM, Gomes KCMS, et al. Analysis of the seroprevalence of and factors associated with Chagas disease in an endemic area in northeastern Brazil. Rev Soc Bras Med Trop. 2016;50(1):115-21.

2. **Capítulo de livro:** Sobrenome, seguido das iniciais dos autores do capítulo, título completo do capítulo, editores, título do livro, Edição, local de publicação: editor, ano de publicação, páginas inicial e final do capítulo abreviada.

Exemplo: Blaxter PS, Farnsworth TP. Social health and class inequalities. In: Carter C, Peel JR, editors. Equalities and inequalities in health. 2nd ed. London: Academic Press; 1976. p. 165-78.

3. **Livro:** Sobrenome, seguido das iniciais dos autores do livro, título do livro, edição, local de publicação: editor, ano de publicação e número de páginas do livro.

Exemplo: Carlson BM. Human embryology and developmental biology. 4th ed. St. Louis: Mosby; 2009. 541 p.

4. **Sites:** Nome do autor/organização. Título da página [Internet]. Local de publicação: Nome do editor; Data ou ano de publicação [atualizado ano mês dia; Citado ano mês dia]. Disponível em: endereço.

Exemplo: Diabetes Australia. Diabetes globally [Internet]. Canberra ACT: Diabetes Australia; 2012 [updated 2012 June 15; cited 2012 Nov 5]. Available from: <http://www.diabetesaustralia.com.au/en/Understanding-Diabetes/DiabetesGlobally/>

5. **Dissertação/Tese:** A Revista da Sociedade Brasileira de Medicina Tropical não aceitará a citação de dissertação/mestrado, teses de doutorado ou similar.

6. World Health Organization (WHO). Chemotherapy of leprosy for control programmes. Technical Report Series 675. Geneva: WHO; 1982. 36 p.

7. Ministério da Saúde (MS). Secretaria de Vigilância em Saúde. Sistema Nacional de Vigilância em Saúde - Relatório de Situação: Mato Grosso do Sul. 5ª edição. Brasília: MS; 2011. 39 p.

Ilustrações: devem ser submetidas em arquivos separados, nomeados apenas com o número das figuras (exemplo: Figura 1; Figura 2). Todas as figuras devem ter numeração arábica, citadas no texto, pela primeira vez, em ordem numérica crescente. Autores podem disponibilizar figuras coloridas ou em preto e branco.

Título e Legendas: devem ser digitados com espaçamento duplo no final do manuscrito.

Dimensões: As dimensões das figuras não devem ultrapassar o limite de 18cm de largura por 23cm de altura. Veja abaixo a correta configuração para cada formato de figura:

- **Imagens/Fotografias:** devem ser obrigatoriamente submetidas em alta resolução no formato **TIFF**. Certifique-se que a mesma foi capturada na resolução mínima de 600 DPI, preferencialmente entre 900-1200dpi, preparadas utilizando programa de Editoração de Imagens (*Adobe Photoshop, Corel Photo Paint, etc*).
- **Gráficos:** Devem ser criados usando software estatístico e devem ser salvos/exportados com a extensão original (**.xls, .xlsx, .wmf, .eps ou .pdf**).
- **Mapas:** devem ser vetorizados (desenhados) profissionalmente, utilizando os softwares *Corel Draw* ou *Illustrator* em alta resolução.

Tabelas: devem ser digitadas com espaçamento simples, com título curto e descritivo (acima da tabela) e submetidas em arquivos separados. Legendas para cada tabela devem aparecer abaixo da mesma. O significado de todas as siglas e símbolos utilizados na tabela devem constar no rodapé da tabela. Todas as tabelas devem ter numeração arábica, citadas no texto, em ordem numérica crescente. Tabelas não devem ter linhas verticais, e linhas horizontais devem ser limitadas ao mínimo. Tabelas devem ter no máximo 18cm de largura por 23cm de altura, fonte *Arial*, tamanho 9.

Processo de submissão: Todos os manuscritos submetidos à Revista da Sociedade Brasileira de Medicina Tropical deverão utilizar apenas a via eletrônica. Todos os manuscritos deverão ser enviados via internet

para <http://mc04.manuscriptcentral.com/rsbmt-scielo>, seguindo as instruções no topo de cada tela. O processo de revisão pelos pares também será totalmente pela via eletrônica.

Sobre Reenvio e Revisões: a revista diferencia entre: a) manuscritos que foram rejeitados e b) manuscritos que serão reavaliados após a realização das correções que foram solicitadas aos autores.

Resubmissão: caso o autor receba uma carta informando que seu trabalho foi rejeitado e queira que os editores reconsiderem tal decisão, o autor poderá reenviá-lo. Neste caso, será gerado um novo número para o manuscrito.

Revisão: Se os revisores recomendarem rever seu manuscrito, ao devolvê-lo para uma segunda análise, por favor, encaminhe o manuscrito revisado e informe o mesmo número do manuscrito.

Após a Aceitação: Uma vez aceito para publicação, o processo de publicação inclui os passos abaixo:

a) Formulário de concessão de direitos autorais, fornecido pela secretaria da revista, deve retornar para a revista assinado pelos autores.

b) Provas: serão enviadas ao autor responsável, mencionado no endereço para correspondência, no formato PDF, para que o texto seja cuidadosamente conferido. Nesta etapa do processo de edição não serão permitidas mudanças na estrutura do manuscrito. Após os autores receberem as provas, deverão devolvê-las assim que possível.

c) Requerimentos para errata apenas serão aceitos no caso de falha cometida por parte do pessoal técnico da revista.

d) Os artigos aceitos serão disponibilizados na modalidade de publicação contínua na biblioteca SciELO.

Custos de Publicação: Não haverá custos de publicação.

A Revista da Sociedade Brasileira de Medicina Tropical não indica qualquer tipo de serviços de tradução.