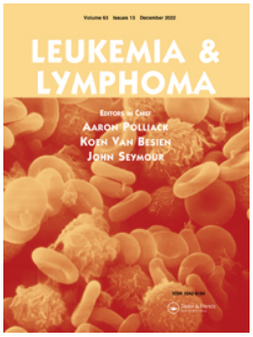
Exploration of whole genome and transcriptome sequencing data lacks evidence for oncogenic viral elements to drive the pathogenesis of T-cell prolymphocytic leukemia

Paurnima Patil, Umut H. Toprak, Julian Seufert, Till Braun, Stephan H. Bernhart, Laura Wiehle, Annika Müller, Matthias Schlesner, Marco Herling, Peter Lichter, Stephan Stilgenbauer, Reiner Siebert, Marc Zapatka

Angaben zur Veröffentlichung / Publication details:

Patil, Paurnima, Umut H. Toprak, Julian Seufert, Till Braun, Stephan H. Bernhart, Laura Wiehle, Annika Müller, et al. 2022. "Exploration of whole genome and transcriptome sequencing data lacks evidence for oncogenic viral elements to drive the pathogenesis of T-cell prolymphocytic leukemia." *Leukemia and Lymphoma* 63 (13): 3253–56. https://doi.org/10.1080/10428194.2022.2116933.



Leukemia & Lymphoma

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ilal20

Exploration of whole genome and transcriptome sequencing data lacks evidence for oncogenic viral elements to drive the pathogenesis of T-cell prolymphocytic leukemia

Paurnima Patil, Umut H. Toprak, Julian Seufert, Till Braun, Stephan H. Bernhart, Laura Wiehle, Annika Müller, Matthias Schlesner, Marco Herling, Peter Lichter, Stephan Stilgenbauer, Reiner Siebert & Marc Zapatka

To cite this article: Paurnima Patil, Umut H. Toprak, Julian Seufert, Till Braun, Stephan H. Bernhart, Laura Wiehle, Annika Müller, Matthias Schlesner, Marco Herling, Peter Lichter, Stephan Stilgenbauer, Reiner Siebert & Marc Zapatka (2022) Exploration of whole genome and transcriptome sequencing data lacks evidence for oncogenic viral elements to drive the pathogenesis of T-cell prolymphocytic leukemia, Leukemia & Lymphoma, 63:13, 3253-3256, DOI: 10.1080/10428194.2022.2116933

To link to this article: https://doi.org/10.1080/10428194.2022.2116933

LEUKEMIA & LYMPHOMA
2022, VOL. 63, NO. 13, 3253–3256
https://doi.org/10.1080/10428194.2022.2116933

Taylor & Francis
Taylor & Francis Group

LETTER TO THE EDITOR

Exploration of whole genome and transcriptome sequencing data lacks evidence for oncogenic viral elements to drive the pathogenesis of T-cell prolymphocytic leukemia

Paurnima Patil[a], Umut H. Toprak[b,c,d,e], Julian Seufert[c,d], Till Braun[f], Stephan H. Bernhart[g], Laura Wiehle[a], Annika Müller[h], Matthias Schlesner[b,i], Marco Herling[f,j], Peter Lichter[k,l], Stephan Stilgenbauer[h], Reiner Siebert[a] and Marc Zapatka[k]

[a]Institute of Human Genetics, Ulm University and Ulm University Medical Center, Ulm, Germany; [b]Bioinformatics and Omics Data Analytics, German Cancer Research Center (DKFZ), Heidelberg, Germany; [c]Faculty of Biosciences, Heidelberg University, Heidelberg, Germany; [d]Division Neuroblastoma Genomics, German Cancer Research Center (DKFZ), Heidelberg, Germany; [e]Hopp-Children's Cancer Center at the NCT Heidelberg (KiTZ), Heidelberg, Germany; [f]Department I of Internal Medicine, Center for Integrated Oncology (CIO), Aachen-Bonn-Cologne-Duesseldorf, Excellence Cluster for Cellular Stress Response and Aging-Associated Diseases (CECAD), Center for Molecular Medicine Cologne (CMMC), University of Cologne (UoC), Cologne, Germany; [g]Interdisciplinary Center for Bioinformatics, Transcriptome Bioinformatics, University of Leipzig, Germany; [h]Department of Internal Medicine III, University of Ulm, Ulm, Germany; [i]Biomedical Informatics, Data Mining and Data Analytics, Faculty of Applied Informatics and Medical Faculty, Augsburg University, Augsburg, Germany; [j]Clinic of Hematology, Cellular Therapy and Hemostaseology, University of Leipzig, Leipzig, Germany; [k]Division of Molecular Genetics, German Cancer Research Center (DKFZ), Heidelberg, Germany; [l]German Cancer Research Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany

ARTICLE HISTORY Received 4 May 2022; revised 12 July 2022; accepted 13 August 2022

According to data from the World Health Organization, a total of 9.9% of human cancers are linked to infections caused by viral pathogens [1,2]. A recent comprehensive analysis of the landscape of viral associations in whole-genome sequencing (WGS) data from more than 2600 tumor samples by the PCAWG (Pan-Cancer Analysis of Whole Genomes) working group on pathogens revealed that this number is considerably higher, as it reports an association in 16% [3]. The International Agency for Research on Cancer (IARC) [4] has reported that the four most prominent viral causes of cancer are human papilloma virus (HPV), hepatitis B virus (HBV), hepatitis C virus (HCV), and Epstein Barr virus (EBV) (official name: human herpes virus 4, HHV4, genus name: lymphocryptovirus) [5–8]. EBV was isolated from a cell line derived from a Burkitt lymphoma (BL) [9] and was the first oncogenic virus identified.

Different oncogenic viruses have been shown to drive or at least to be associated with subtypes of lymphomas or lymphatic leukemias [1]. EBV is considered to be associated not only with BL, but also with a range of other B-cell and T-cell lymphomas [10]. Human herpes virus 8 (HHV8, also named Kaposi-sarcoma-associated herpes virus, KSHV) is supposed to contribute to the pathogenesis of primary effusion lymphomas [11]. Regarding T-cell lymphomagenesis, a proven viral pathogen besides EBV is the human T-lymphotropic virus (HTLV). The HTLV family is a group of human retroviruses which are linked to the pathogenesis of adult T-cell leukemia/lymphoma [12].

Depending on the type of virus, its nucleic acids can be present in infected cells as RNA or DNA which could be integrated in the human host genome. Also germline transmission of viral sequences has been associated with lymphomagenesis [11]. Moreover, there are a number of viruses which are present in the blood cell compartments in the absence of neoplastic diseases, especially members of the *Herpesviridae* and *Anelloviridae* families [13].

T-cell prolymphocytic leukemia (T-PLL) is an aggressive T-cell malignancy [14]. The genetic hallmarks are the inv(14)/t(14;14) or t(X;14) which occur at a thymic differentiation stage and activate the *TCL1A* and *MTCP1* oncogenes [15,16]. While the mutational landscape of T-PLL has been well characterized, to this point no infectious or otherwise causative exogenous agents have been described to play an oncogenic role in T-PLL. With regard to a viral contribution to T-PLL pathogenesis, it has been reported that patients with T-PLL are seronegative for HTLV-I, and also T-PLL cells do not contain monoclonally integrated HTLV-I provirus [17]. Instead, EBV antigens, EBNA3C and LMP1 were found in T-PLL cells which

CONTACT Marc Zapatka m.zapatka@dkfz.de German Cancer Research Center, Molecular Genetics, Computational Cancer Genomics, Heidelberg D-69120, Germany

Supplemental data for this article is available online at https://doi.org/10.1080/10428194.2022.2116933

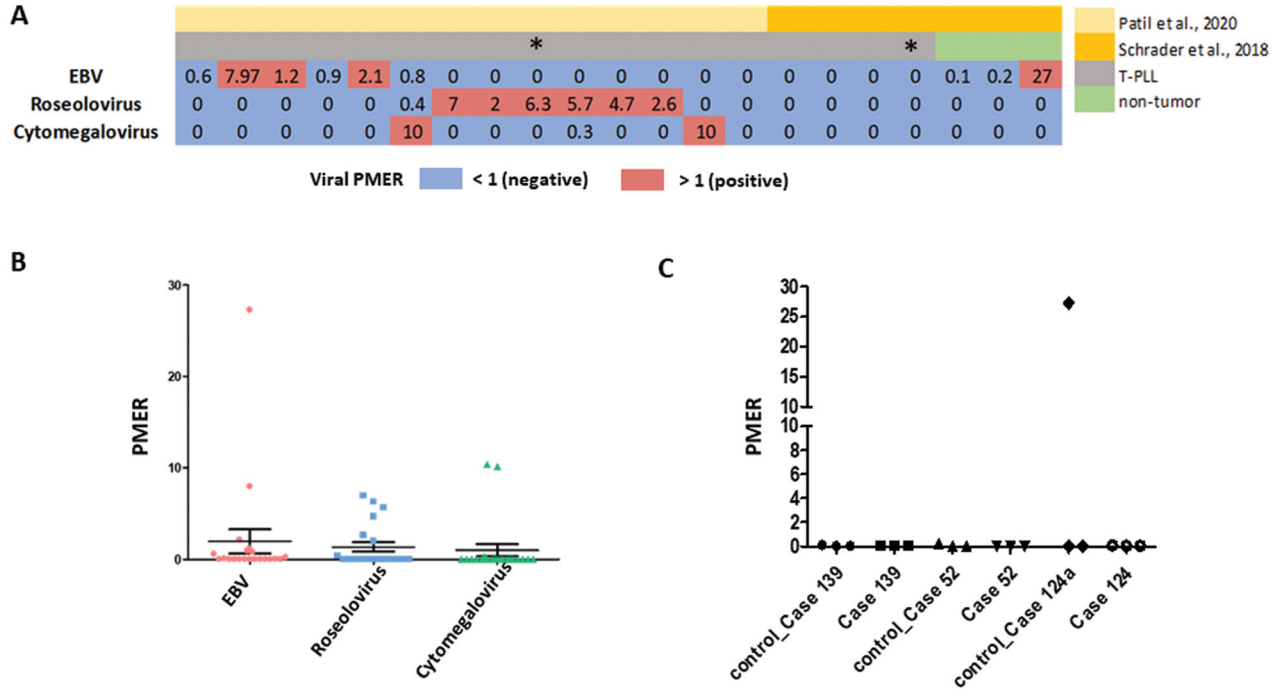


Figure 1. Viral reads detected in T-PLL samples and non-tumor samples in the WGS data. (A) The heatmap depicts viral reads per million extracted reads (PMER) per sample (in columns) for three viruses: EBV, roseolovirus and cytomegalovirus (in rows) based on the WGS data of T-PLL cases ($n = 16$) and SUP-T11 (the case 124 and case 124a which is the case analysis twice is marked with an asterisk) and non-tumor samples. The samples published in Patil et al. [16] and in Schrader et al. [19] are depicted in the top bar. The heatmap shows PMER for viruses detected per sample with low to high read values. (B) The dot plot shows the viral load in PMER on *x*-axis for all samples (T-PLL and non-tumor). (C) Bar plot shows the viral reads detected three T-PLL samples and their respective non-tumor controls. The *y*-axis depicts the PMER for viruses detected per sample.

suggests that EBV infection might affect tumor growth in T-PLL [18]. However, a specific role of EBV in the oncogenesis of T-PLL is yet to be explored. A previous study on whole genome and exome sequencing data of T-PLL cases revealed no evidence of viral sequences [19].

Herein, we aimed at an un-biased investigation of a potential viral contribution to the pathogenesis of T-PLL. To this end, we mined previously published WGS of 16 T-PLL cases and the SUP-T11 cell line (total $n = 17$) which harbored a *TCRAD::TCL1A* ($n = 16$) or *TCRAD::MTCP1* ($n = 1$) juxtaposition and of non-tumor samples from 3/16 T-PLL patients [16,19]. In addition, we included previously published RNA-sequencing (RNA-seq) data of 10/16 T-PLL cases and of SUP-T11 in the current analysis [16]. It needs to be noted that the samples analyzed using WGS and RNA-seq were obtained at the same time point.

Focusing on viral pathogens, we analyzed the raw reads in the whole genome and transcriptome sequencing by Kraken2 (version 2.0.8-beta) [20] (Supplementary Methods). We selected on average 10,831,317 (range 1,775,676–34,753,984) genomic raw and partially unmapped reads as previously described [3] which were analyzed in the 16 T-PLL cases, SUP-T11 and three non-tumor samples (Supplementary Table S2). Thereby, we detected viral raw reads for 25 different viruses with a range of 1–444 reads (Supplementary Table S2). Of these, per million extracted reads (PMER) >1 were detected for the genus Lymphocryptovirus (e.g. EBV) in 3/17 T-PLL and 1/3 non-tumor control samples, cytomegalovirus (CMV) in 2/17 cases and roseolovirus in 6/17 cases (Supplementary Table S3). The heatmap in Figure 1(A) depicts the PMER for each virus detected in each sample. In addition, the dot plots indicate the maximum PMER detected for EBV, roseolovirus, and CMV in the samples (Figure 1(B)). The WGS data of the three T-PLL cases (case 124a, case 139, and case 52) were previously analyzed using VirusFinder2.0 [19]. In line with this previous study, our current analysis using Kranken2 did not show presence of viral sequences with >1 PMER in these three cases, validating the previously published lack of viral sequences in these samples.

We found that EBV showed the highest reads counts (PMER = 27.3) in a non-tumor sample (control_case 124a) of a T-PLL case. Nevertheless, the tumor sample (124a) of the same case collected at the same time point was negative for EBV (PMER = 0) (Figure 1(C)) despite analyzing a comparable number of extracted genomic reads (Supplementary Table S2). Furthermore, a tumor sample collected nine months later was negative for EBV (case 124). In addition, we performed a polymerase chain reaction (PCR) based verification for both the tumor (case 124a) and the respective control (non-tumor) sample. Again, we confirmed that the tumor (case 124a) sample was negative for EBV, but the non-tumor control was

positive for EBV. As the tumor cells lacked detectable traces of EBV also by PCR, our results suggest that the patient was infected with EBV at some point, but the EBV infection was not evident in T-PLL cells. In the light of this case, it is important to note that not in all samples analyzed herein T-PLL cells were sorted to high purity (Supplementary Table S1). Thus, we cannot prove with absolute certainty that viral reads detected are derived from or even restricted to the tumor cells. Interestingly, roseolovirus was detected in T-PLL samples ($n = 5$) which were all sorted using flow cytometry and, in the SUP-T11 cell line. In contrast, none of the non-sorted samples ($n = 6$) were positive for reads from this virus suggesting positivity in the tumor and not the normal cells (Supplementary Table S3). In an analysis using SvABA [21], we could not identify any evidence for integration of viral sequences in to the human genome in any of the WGS sequenced cases (Supplementary methods).

Next, we performed the same analyses for detecting the viral reads in the RNA-seq data of 11/17 T-PLL samples (including SUP-T11) which were already analyzed using WGS data. Non-tumor control samples were not included in the transcriptome analysis. On average 15,355,751 (range 10,624,225–17,562,490) raw transcriptomic reads per sample were selected (Supplementary Table S4). Using this approach, we found that none of the 10 T-PLL cases and SUP-T11 cell line were positive for any viruses as the PMER threshold of >1 was not reached in any sample (Supplementary Table S5). Thus, our transcriptomic data show no evidence of virus-associated pathogenesis in T-PLL.

In the present study, we explored whole genome and transcriptome data for investigating a potential viral contribution to the pathogenesis of T-PLL in an unbiased approach. Among the previously known oncogenic viruses in leukemias and lymphomas [10,18], the most prominent viral sequences in WGS data were those of EBV in 3/17 (18%) T-PLL samples and one non-tumor control sample. In line, previous studies have provided evidence of EBV infection in T-PLL that might play a role in tumor development but the mechanism by which EBV might affect tumorigenesis in T-PLL is yet to be identified [17]. Here, we could not link EBV infection consistently with tumorigenesis in T-PLL as a relatively low percentage of samples was EBV positive. Among, the other prevalent viruses we detected in T-PLL is CMV which has known to affect immunosuppressed patients [22]. However, studies have also reported its role in cancer progression [23,24]. In addition, roseolovirus showed several hits in T-PLL. Roseolovirus which is a member of Herpesviridae family has been shown to be present in tumor cells but has not been reported to play any tumorigenic effect [3]. Furthermore, we could not identify actively transcribed viral genes for roseolovirus or CMV at the transcriptome level of the cases which showed highest number of reads in the WGS data. This is in accordance with previous reports describing a latent state of these viruses in blood mononuclear cells [25]. Although the PMERs of CMV and roseolovirus were above our threshold in a subset of cases, these viruses to the best of our knowledge have not been described as oncogenic drivers in lymphomas or lymphatic leukemias [26,27].

Our findings do neither rule out nor support the hypothesis that the viruses detected in the WGS data of the T-PLL samples, or any other viruses might have been active at an earlier time point during tumor development and therefore are not expressed at the time of sampling. We did not detect viral sequence integration in the genome of T-PLL patients which renders the possibility of viruses to act as an initiating factor in T-PLL by driving or inactivating tumor-associated host genes unlikely. As recently many mature lymphatic neoplasms have been shown to have a long evolutionary history with initiating events occurring early in life [28], it might be challenging to sample T-PLL at initiating phases of the disease to address this question. Thus, such viral 'hit-and-run hypothesis' needs likely to be addressed in targeted experiments in model systems like mouse models for T-PLL.

Our analyses of WGS and RNA-seq datasets show that oncogenic viruses are present in relatively few T-PLL cases. Although our study shows no strong evidence of tumorigenic viral involvement in T-PLL pathogenesis, we cannot exclude a hit-and-run mechanism in which viruses were active at an initiating phase of T-PLL pathogenesis but then get lost during disease evolution, as it has been proposed for some EBV-positive lymphomas [29].

Acknowledgements

The authors thank the members of the cancer genetics laboratories of the Institutes of Human Genetics in Ulm and Kiel for expert support and discussion. We gratefully acknowledge contributing centers for enrolling all patients. The author thank the Omics IT and Data Management Core Facility of the German Cancer Research Center (DKFZ) for excellent technical support.

Author contributions

MH provided the WGS data of three T-PLL patient samples and reviewed the manuscript. UHT, JS, and MS performed bioinformatic analyses of WGS data. SHB did the bioinformatic analysis of mRNA data. TB coordinated the EBV PCR verification experiments. LW and PL reviewed the paper. MZ performed the bioinformatic analysis for detection of the viral pathogens, interpreted the data and reviewed the paper. PP: performed data analysis, interpreted the data, and wrote the manuscript. RS designed and developed the study, provided intellectual input, and reviewed the paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by grants of the German Research Foundation (DFG) in the framework of the Collaborative Research Centre SFB 1074 (Projects B1 and B9N/B9). MH was supported by the EU Transcan-2 Consortium ERANET-PLL and by the ERAPerMed Consortium JAKSTAT-TARGET.

ORCID

Stephan H. Bernhart http://orcid.org/0000-0002-5928-9449
Laura Wiehle http://orcid.org/0000-0001-6081-3925
Matthias Schlesner http://orcid.org/0000-0002-5896-4086
Reiner Siebert http://orcid.org/0000-0002-6492-5323
Marc Zapatka http://orcid.org/0000-0001-8287-5967

References

[1] Plummer M, de Martel C, Vignat J, et al. Global burden of cancers attributable to infections in 2012: a synthetic analysis. Lancet Glob Health. 2016;4(9):e609–e616.
[2] Parkin DM. The global health burden of infection-associated cancers in the year 2002. Int J Cancer. 2006;118(12):3030–3044.
[3] Zapatka M, Borozan I, Brewer DS, et al. The landscape of viral associations in human cancers. Nat Genet. 2020;52(3):320–330.
[4] Bouvard V, Baan R, Straif K, et al. A review of human carcinogens – part B: biological agents. Lancet Oncol. 2009;10(4):321–322.
[5] Muñoz N, Castellsagué X, de González AB, et al. Chapter 1: HPV in the etiology of human cancer. Vaccine. 2006;24:S1–S10.
[6] Bialecki ES, Di Bisceglie AM. Clinical presentation and natural course of hepatocellular carcinoma. Eur J Gastroenterol Hepatol. 2005;17(5):485–489.
[7] Hermine O, Lefrère F, Bronowicki J-P, et al. Regression of splenic lymphoma with villous lymphocytes after treatment of hepatitis C virus infection. N Engl J Med. 2002;347(2):89–94.
[8] Thompson MP, Kurzrock R. Epstein-Barr virus and cancer. Clin Cancer Res. 2004;10(3):803–821.
[9] Hausen HZ, Schulte-Holthausen H, Klein G, et al. Epstein–Barr virus in Burkitt's lymphoma and nasopharyngeal carcinoma: EBV DNA in biopsies of Burkitt tumours and anaplastic carcinomas of the nasopharynx. Nature. 1970;228(5276):1056–1058.
[10] Epstein MA, Achong BG, Barr YM. Virus particles in cultured lymphoblasts from Burkitt's lymphoma. Lancet. 1964;1(7335):702–703.
[11] Zhang E, Cotton VE, Hidalgo-Bravo A, et al. HHV-8-unrelated primary effusion-like lymphoma associated with clonal loss of inherited chromosomally-integrated human herpesvirus-6A from the telomere of chromosome 19q. Sci Rep. 2016;6:22730.
[12] Poiesz BJ, Poiesz MJ, Choi D. The human T-cell lymphoma/leukemia viruses. Cancer Invest. 2003;21(2):253–277.
[13] Moustafa A, Xie C, Kirkness E, et al. The blood DNA virome in 8,000 humans. PLoS Pathog. 2017;13(3):e1006292.
[14] Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. Blood. 2016;127(20):2375–2390.
[15] Virgilio L, Lazzeri C, Bichi R, et al. Deregulated expression of TCL1 causes T cell leukemia in mice. Proc Natl Acad Sci U S A. 1998;95(7):3885–3889.
[16] Patil P, Cieslak A, Bernhart SH, et al. Reconstruction of rearranged T-cell receptor loci by whole genome and transcriptome sequencing gives insights into the initial steps of T-cell prolymphocytic leukemia. Genes Chromosomes Cancer. 2020;59(4):261–267.
[17] Kojima K, Hara M, Sawada T, et al. Human T-lymphotropic virus type I provirus and T-cell prolymphocytic leukemia. Leuk Lymphoma. 2000;38(3–4):381–386.
[18] Lan K, Murakami M, Choudhuri T, et al. Detection of Epstein-Barr virus in T-cell prolymphocytic leukemia cells in vitro. J Clin Virol. 2008;43(3):260–265.
[19] Schrader A, Crispatzu G, Oberbeck S, et al. Actionable perturbations of damage responses by TCL1/ATM and epigenetic lesions form the basis of T-PLL. Nat Commun. 2018;9(1):697.
[20] Wood DE, Lu J, Langmead B. Improved metagenomic analysis with kraken 2. Genome Biol. 2019;20(1):257.
[21] Wala JA, Bandopadhayay P, Greenwald NF, et al. SvABA: genome-wide detection of structural variants and indels by local assembly. Genome Res. 2018;28(4):581–591.
[22] Kepler GM, Banks HT, Davidian M, et al. A model for HCMV infection in immunosuppressed patients. Math Comput Model. 2009;49(7–8):1653–1663.
[23] Herbein G, Kumar A. The oncogenic potential of human cytomegalovirus and breast cancer. Front Oncol. 2014;4:230.
[24] Herbein G. The human cytomegalovirus, from oncomodulation to oncogenesis. Viruses. 2018;10(8):408.
[25] Krug LT, Pellett PE. Roseolovirus molecular biology: recent advances. Curr Opin Virol. 2014;9:170–177.
[26] Ogata M. Human herpesvirus 6 in hematological malignancies. J Clin Exp Hematop. 2009;49(2):57–67.
[27] Wingard JR. Viral infections in leukemia and bone marrow transplant patients. Leuk Lymphoma. 1993;11(Suppl. 2):115–125.
[28] Maura F, Rustad EH, Boyle EM, et al. Reconstructing the evolutionary history of multiple myeloma. Best Pract Res Clin Haematol. 2020;33(1):101145.
[29] Mundo L, Del Porro L, Granai M, et al. Frequent traces of EBV infection in Hodgkin and non-Hodgkin lymphomas classified as EBV-negative by routine methods: expanding the landscape of EBV-related lymphomas. Mod Pathol. 2020;33(12):2407–2421.