CancerHERVdb: Human Endogenous Retrovirus (HERV) Expression Database for Human Cancer Accelerates Studies of the Retrovirome and Predictions for HERV-Based Therapies

Erik Stricker,[a,b] Erin C. Peckham-Gregory,[b] Michael E. Scheurer[b]

aBaylor College of Medicine, Department of Molecular Virology and Microbiology, Houston, Texas, USA
bBaylor College of Medicine, Department of Pediatrics, Houston, Texas, USA

ABSTRACT In this study, we sought to create a database summarizing the expression of human endogenous retroviruses (HERVs) in various human cancers. HERVs are suitable therapeutic targets due to their abundance in the human genome, overexpression in various malignancies, and involvement in various cancer pathways. We identified articles on HERVs from PubMed and then prescreened and automatically categorized them using the portable document format (PDF) data extractor (PDE) R package. We discovered 196 primary research articles with HERV expression data from cancer tissues or cancer cell lines. HERV RNA and protein expression was reported in brain, breast, cervical, colorectal, endocrine, gastrointestinal, kidney/renal/pelvis, liver, lung, genital, oral cavity, pharynx, ovary, pancreas, prostate, skin, testicular, urinary/bladder, and uterus cancers, leukemias, lymphomas, and myelomas. Additionally, we discovered reports of HERV RNA-only overexpression in soft tissue cancers including heart, thyroid, bone, and joint cancers. The CancerHERVdb database is hosted in the form of interactive visualizations of the expression data and a summary data table at https://erikstricker.shinyapps.io/cancerHERVdb/. The user can filter the findings according to cancer type, HERV family, HERV gene, or a combination thereof and easily export the results with the corresponding reference list. In our report, we provide examples of potential uses of the CancerHERVdb, such as identification of cancers suitable for off-target treatment with the multiple sclerosis-associated retrovirus (MSRV)-Env-targeting antibody GNbAC1 (now named temelimab) currently in phase 2b clinical trials for multiple sclerosis or the discovery of cancers overexpressing HERV-H long terminal repeat-associating 2 (HHLA2), a newly emerging immune checkpoint. In summary, the CancerHERVdb allows cross-study comparisons, encourages data exploration, and informs about potential off-target effects of HERV-targeting treatments.

IMPORTANCE Human endogenous retroviruses (HERVs), which in the past have inserted themselves in various regions of the human genome, are to various degrees activated in virtually every cancer type. While a centralized naming system and resources summarizing HERV levels in cancers are lacking, the CancerHERVdb database provides a consolidated resource for cross-study comparisons, data exploration, and targeted searches of HERV activation. The user can access data extracted from hundreds of articles spanning 25 human cancer categories. Therefore, the CancerHERVdb database can aid in the identification of prognostic and risk markers, drivers of cancer, tumor-specific targets, multicancer spanning signals, and targets for immune therapies. Consequently, the CancerHERVdb database is of direct relevance for clinical as well as basic research.

KEYWORDS HERV, cancer, CancerHERVdb, database, expression, human endogenous retrovirus

Human endogenous retroviruses (HERVs) are more abundant in the human genome than protein-coding, microRNA (miRNA), and other regulatory noncoding RNA (ncRNA) sequences taken together (1). Over 8% of the human genome encodes long terminal repeat (LTR)-containing elements such as human endogenous retroviruses (HERVs) (1, 2). While the role and function of the majority of HERVs are still unknown, increased expression of explicit HERV gene products has become a hallmark for several cancers (3), e.g., HERV-K in melanoma (4, 5), HERV-E in clear cell renal cell carcinoma (6), and HERV-H in colorectal cancer (7), or has been shown to have potential as a driver for oncogenesis, e.g., HERV-K np9 (8, 9).

Simple retroviruses such as HERVs of the gammaretroviral family carry *gag*, *pol*, and *env* genes flanked by LTRs, while complex retroviruses such as HERVs of the betaretroviral (e.g., HERV-K) or spumaviral (e.g., HERV-L) families carry additional nonstructural genes (10). The number and composition of accessory genes can vary between different HERVs and can include functions such as transcriptional control (e.g., HERV-L *tas/bel1*), transport regulation (e.g., HERV-K *rec*), defense against antiviral factors (e.g., HERV-L *bet*), and regulation of host cell pathways (e.g., HERV-K *np9*) (8, 10). HERVs have been shown to have inserted themselves into several cancer-related pathways using their viral proteins and altering transcriptional control mechanisms. Accordingly, we set out to create a database summarizing HERV RNA and protein expression in different cancers.

Our goal was to assess the presence of HERV-carried gene expression in known cancers and outline HERVs which are potentially suitable for targeted cancer therapies. Possible uses of HERVs for therapeutic purposes include HERV epitopes for vaccination, which have been shown to be safe and able to generate tumor-specific immune cells (11–14). In addition, the advent of CAR-T cells (15) and checkpoint blockade inhibitors (16) might offer additional options to enhance tumor-specific immunogenicity (17). Furthermore, direct pharmacological targeting of HERVs expressed in tumors or HERV-derived restriction factors such as suppressyn (HUGO Gene Nomenclature Committee [HGCN]: ERVH48-1), a HERV-F-derived inhibitor of syncytin-1 (HGCN: ERVW-1)-mediated fusion, might be another avenue for therapeutic approaches.

RESULTS

PDE outperforms PubMed search term optimization and reduces manual evaluation. As of 5 May 2022, a title and abstract search in PubMed indicated 3,960 scientific articles published on endogenous retroviruses (ERVs), including human ERVs. Of the identified papers, 1,104 articles were not available as full-text articles and 121 articles were not in English, resulting in 2,735 full-text articles available for our analysis (a modified PRISMA [Preferred Reporting Items for Systematic Reviews and Meta-Analyses] flow diagram is shown in Fig. 1). To categorize the articles into cancer-related and noncancer articles, we used our recently published portable document format (PDF) data extractor (PDE) software package (18), effectively reducing the number of articles to be evaluated manually from 2,735 to 1,080, i.e., by 60%. Based on the results, we detected 27 duplicates. Combining the corresponding filter in PubMed with the PDE results, we identified 155 cancer-related HERV review articles. Through assessment of the reviews, we discovered 60 additional HERV-specific articles that were not a result of the PubMed search because of a lack of HERV-specific keywords in the title and abstract. The remaining 1,020 possibly cancer-related HERV articles were grouped into 26 cancer categories, including nonhuman cancers, by the PDE analyzer. Through manual evaluation, we determined that 419 articles were on HERVs and human cancers. Manually evaluated articles included 18 files that were secured and 24 files that were nonreadable by the PDE analyzer. We decided to exclude articles on cancers associated with other viral infections using the PDE and recommend the recently published review on other viruses triggering HERVs by Chen et al. (19) for a detailed summary. Of the 419 articles with qualitative cancer data, 196 articles included quantitative data on HERV expression in cancer tissues or cancer cell lines.

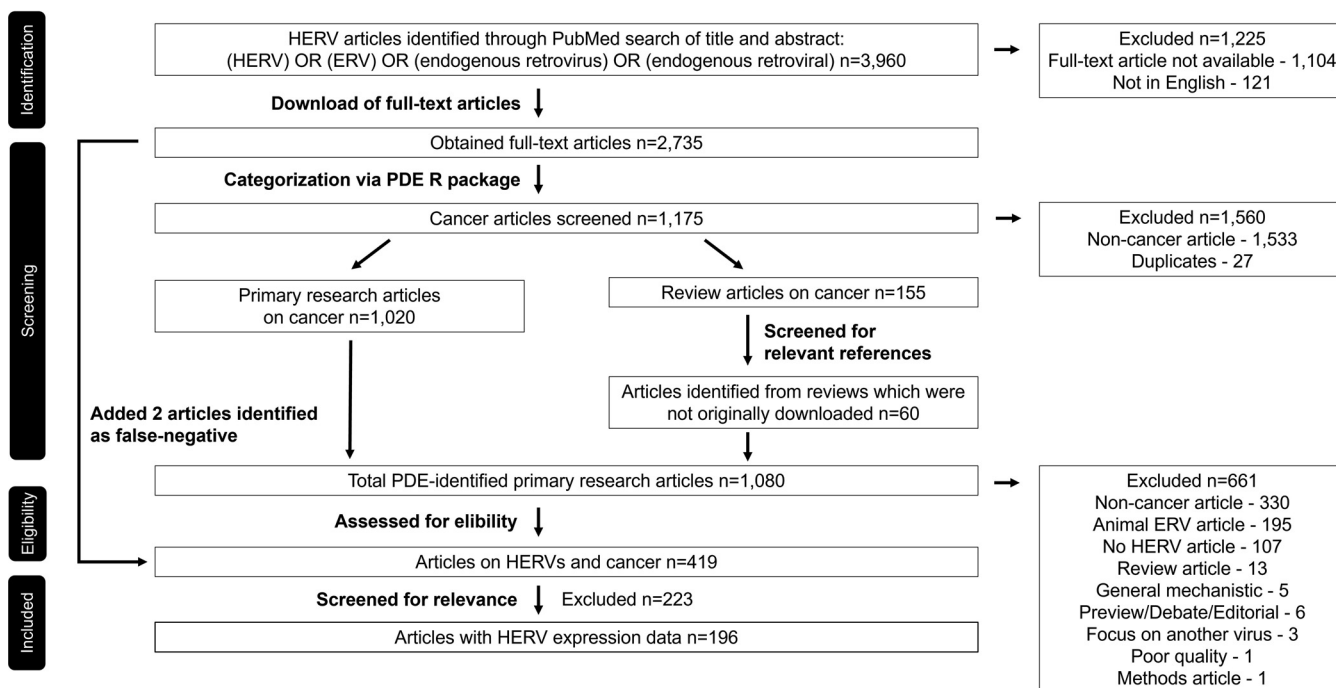Using the PDE reader tool in combination with assessment of the full-text articles,

FIG 1 PRISMA flowchart on how articles were identified, obtained, and assessed. Initially, HERV-related articles were identified through a PubMed search, and full-text articles were obtained using the PubMed-Batch-Download software developed by Bill Greenwald and Texas Medical Center (TMC) library access, filtered using the PDE R package, and in a final step manually evaluated.

we determined the efficacy of the PDE analyzer and identified 443/1,020 (43.4%) false positives, 2/2,735 (0.07%) false negatives, 195/1,080 (18.1%) animal ERV papers, and, of the cancer-related HERV articles, 62/581 (10.7%) articles assigned to the wrong category. Upon further examination, we found that the PDE produced false-negative calls based on the missing integrity of one PDF file (20) and the low ratio of cancer keywords because of the use of only one cancer cell line in a multiple sclerosis-focused paper (21). In total, we determined a specificity of 83.3% and sensitivity of 99.7% for the systematic use of the PDE analyzer. Interestingly, when the search term "cancer" was added to the PubMed search, a total of 1,041 articles were indicated of which the PDE analyzer correctly identified 116 (11.1%) as actually non-cancer related (see Table S1 in the supplemental material). Furthermore, evaluation of HERV-related full-text articles using the PDE analyzer revealed 77 articles not found with the search term "cancer." The results were even more discordant when using the MeSH term "neoplasms."

ERVs are studied predominantly in animal models. Analyses show nonhuman cancers as the most prevalent topic in the ERV literature, which was expected as ERVs were first described in animals and findings from animal models and animal ERVs (especially porcine and murine ERVs) are a predominant result of using the keywords "ERV" and "endogenous retrovirus" compared to "HERV" or "human endogenous retrovirus" (Fig. 2). Furthermore, evaluation of the HERV literature indicates that breast cancer is the most commonly studied human cancer, followed by testicular cancer, skin cancer, lymphoma, colorectal cancer, and brain cancer. We found lymphoma, skin cancer, and leukemia to be the main malignancies studied using animal ERV models. No articles on endocrine cancers, beyond thyroid and pancreas, or on brain cancers were found.

When evaluating publication date distributions, it was interesting to observe five eras of HERV cancer research. While initial reports of HERVs in cancers were published in 1978, we observed a first peak of HERV publication in the 1990s with a predominant focus on testicular and germ cell cancers (Fig. 3). In the 2000s, the emphasis of HERV research shifted slightly to breast cancers. In the years 2008 to 2013, reports of HERVs in skin cancers dominated the literature, followed by a prolific era of HERV and cancer
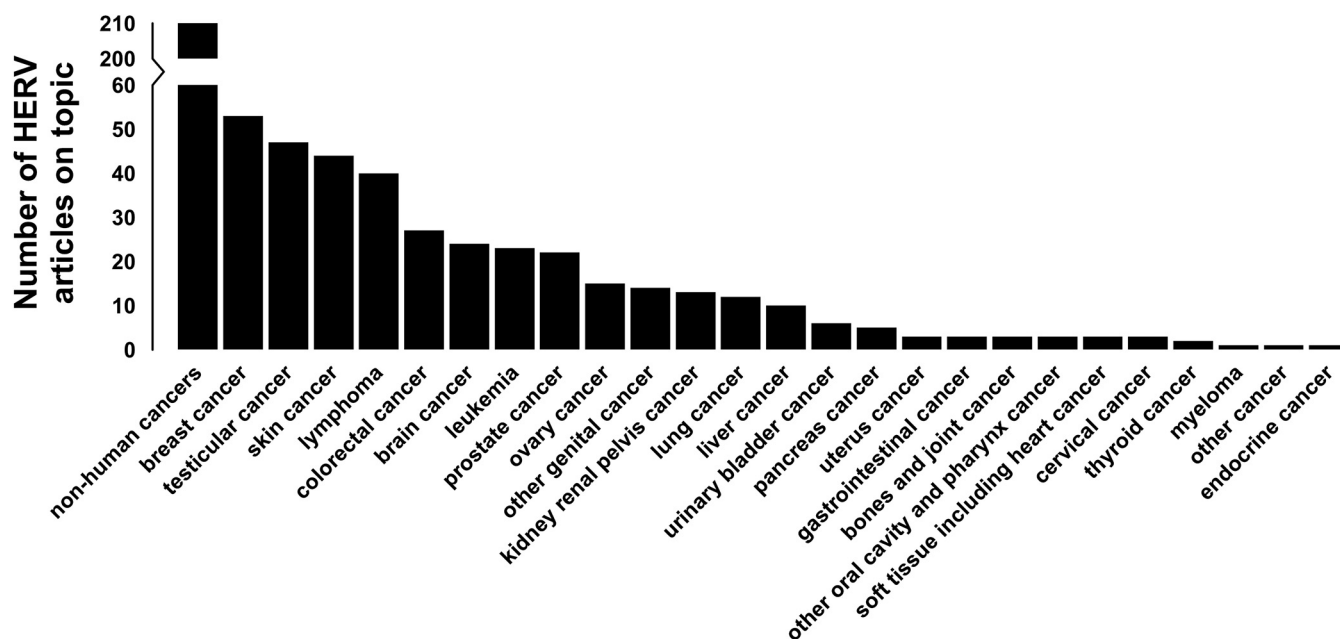
FIG 2 Distribution of cancer-related HERV papers by cancer type. A total of 2,735 full-text HERV-related papers were searched for cancer-related keywords. Articles with ≥0.2% of all words being general cancer keywords were considered cancer related. The resulting 655 cancer-related primary research HERV articles were probed for the cancer type that was mentioned most often and divided into one of 26 cancer categories, including nonhuman cancers and a "multiple cancer" category. Nonhuman cancers were included for completeness.

publications. Research between 2014 and 2019 covered a variety of cancers with lymphomas, breast cancers, and colorectal cancers displaying peak numbers. Starting in that time, we observed a rising number of brain cancer articles which coincides with an exponential increase in neurobiology-focused HERV articles currently dominating the field. Lastly, publications reporting HERVs in leukemia have remained nearly constant since 1993.

HERV expression was detected in a variety of cancers and for a large range of HERV families. Through the manual review of the full-text articles, we determined HERV expression data, scientific findings, and impact. We found HERV transcripts, HERV proteins, HERV-directed antibodies, and HERV-specific T cells to be overexpressed in various cancers (Fig. 4; Table S1). Our literature analysis showed that multiple viral gene products were elevated in various cancer cell lines but lacked data from matching malignant tissues (e.g., HERV-W *pol*, HERV-P *gag*, and HERV-T *pol*), while others were reported to be significantly upregulated in both cell lines and cancer tissues (e.g., HERV-K [HML-2] *gag*, HERV-K *env*, and HERV-H *env*). The HERV-K (HML-2) family is among the most studied families and, together with HERV-W Env (mostly syncytin-1; HCGN: ERVW-1) and HERV-FRD Env (syncytin-2; HCGN: ERVFRD-1), one of the earliest families with existing antibodies for viral protein detection. Meanwhile, HERV antibodies for HERV-K (HML-4) reverse transcriptase (RT), HERV-R Env, HERV-H Gag and LTR products, HERV-E Env, HERV-S, HERV-V1 Env, HERV-Fc Env, and HERV-MER61 have been created. In addition to commonly studied HERV families, our analyses distilled understudied HERV families and rare cancers such as HERV-Pb env RNA in brain cancer tissues (22), HERV-V1 Env protein expression in ovary cancer tumor tissues (23), and HERV-R env RNA expression in primary thyroid tumor tissue samples (24).

The CancerHERVdb database provides open access to HERV expression data in cancers. The results of this study are available in the form of a publicly available website, CancerHERVdb, found at the following address: https://erikstricker.shinyapps.io/cancerHERVdb/. The landing page of the CancerHERVdb site provides the user with an interactive version of Fig. 4, which is newly generated based on the latest data table carrying HERV expression data in cancers each time the website loaded (Fig. 5A). The data table is updated quarterly based on curated user submissions though the website and recent publications. For the identification of new HERV-related cancer publications, the PubCrawler, a
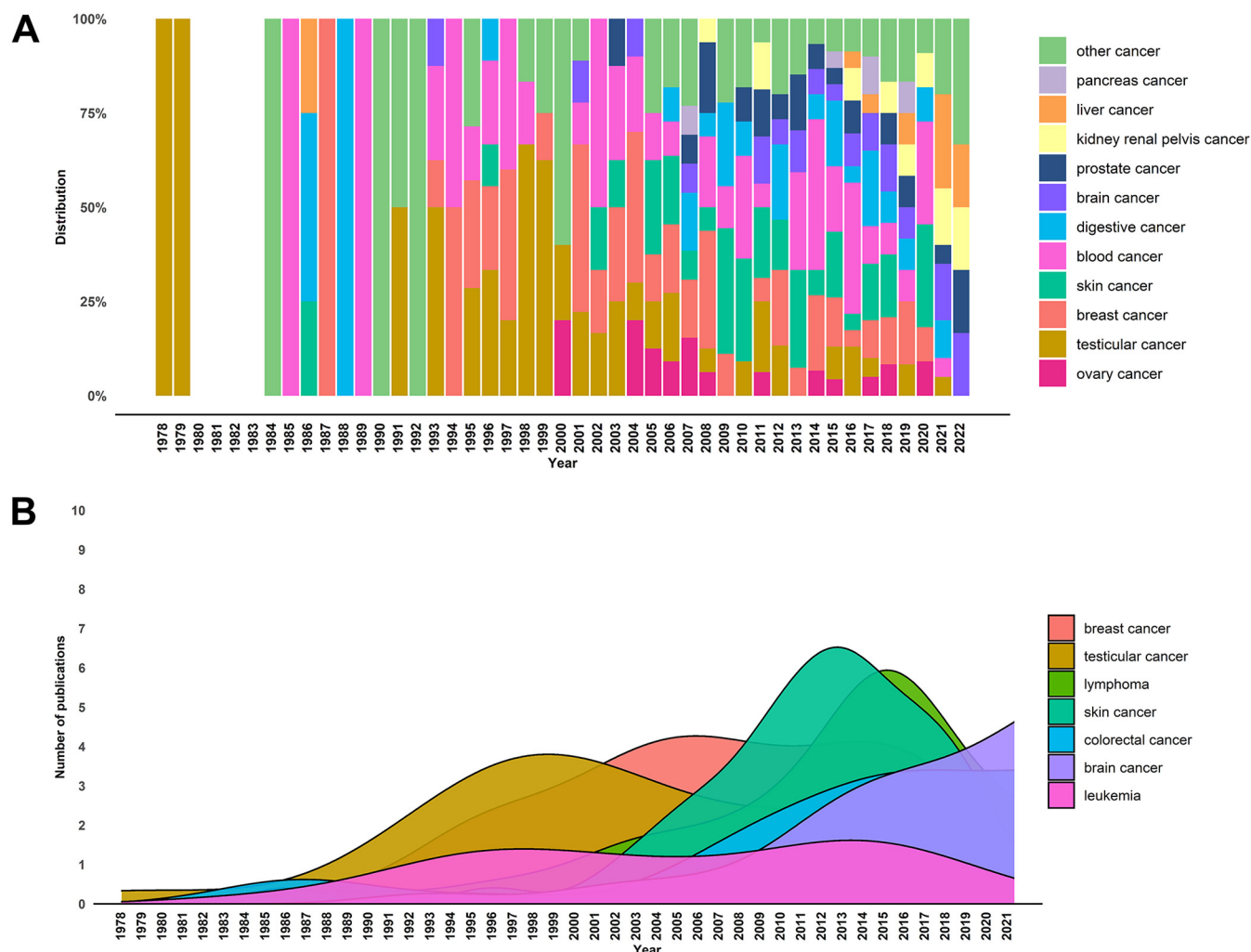
FIG 3 Distribution of cancer-related HERV papers by year. Articles with a single cancer focus were sorted by year of publication and used for visualization. (A) For distribution ratios, gastrointestinal cancer and colorectal cancer were summarized to digestive cancers, leukemia and lymphoma to blood cancer, and lung cancer, uterus cancer, thyroid cancer, myeloma, bone and joint cancer, cervical cancer, soft tissue cancers including heart cancer, urinary bladder cancer, other genital cancer, and other oral cavity and pharynx cancer to other cancer. (B) The seven most studied cancers in HERV-focused articles were evaluated for their publication distribution. Publication density was plotted against the year of publication.

free web-based program which sends daily notifications for newly published articles in user-specified areas of interest (25), will be used with the same PubMed search words but without publication date restrictions. Authors of novel publications with HERV expression data in cancers will be contacted per email and encouraged to submit their findings through the CancerHERVdb website's submission form. The date of the latest update is continuously displayed at the bottom of the website. The underlying data table can be browsed as well as downloaded on the webpage under the "Table Browser" tab (Fig. 5B), offering maximum access to the data. Under the "Custom Search" tab, the user can obtain detailed information on any cancer subtype, HERV family, HERV gene, or a combination thereof. Searches for cancer + HERV family, HERV family + HERV gene, and cancer + HERV family + HERV gene will list all articles reported to correspond to HERV expression with their results (Fig. 5C). The displayed results can be downloaded by the user as a table or Word document providing a reference list with PubMed web links. The search can also be initiated by clicking on data points displayed in any of the graphs within the CancerHERVdb site. For all graphs, on mouse hover, the website displays more information on the respective data point. Lastly, the user can visualize data by cancer type, by HERV family, or by HERV gene (Fig. 5D). A search by cancer type, for example, will present the user
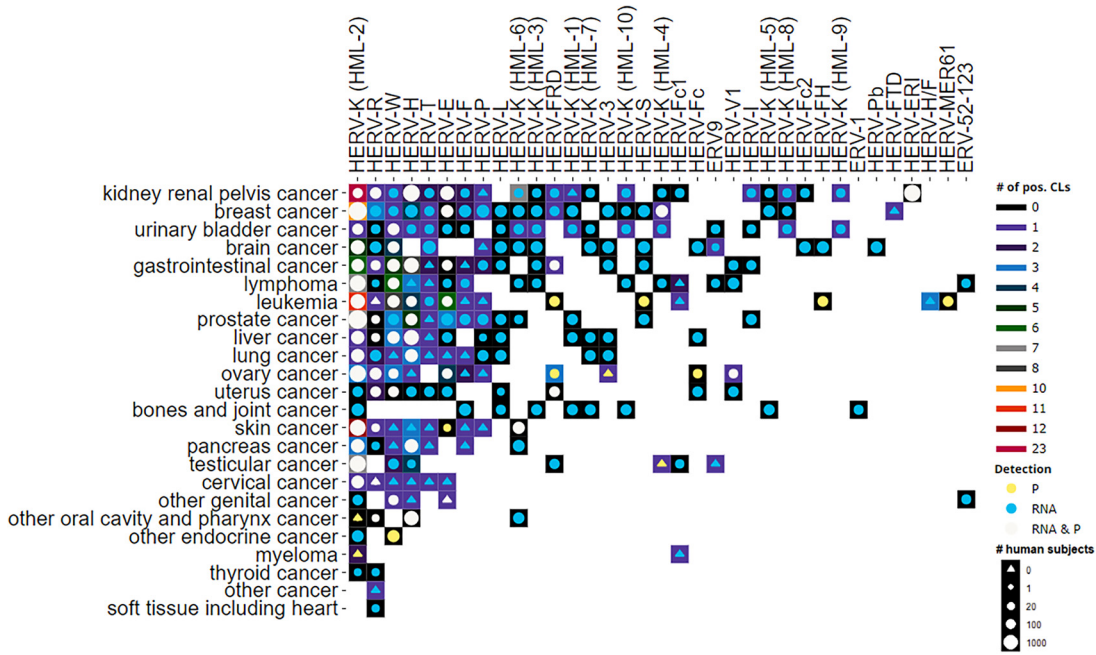
FIG 4 HERV expression by HERV family and cancer type. Background color represents number of different cell lines with expression of HERV (# of pos. CLs). Detection of HERV expression was classified as either protein (P), which includes anti-HERV antibody, or anti-HERV T-cell evidence or RNA, and is depicted by circle or triangle color (Detection). The size of the circle signifies number of cancer patients evaluated, with triangles denoting HERVs evaluated only in cell lines (# human subjects).

with a stacked bar graph containing information on positivity rates in tumor tissues tested in addition to a graph displaying publication numbers by year.

The CancerHERVdb website can be used to discover the most suitable cancers for HERV-targeting therapeutics. The most substantial application of the presented synthesis of cancers with HERV expression is the exploration of HERV-based treatments. The multiple sclerosis-associated retrovirus (MSRV)-Env-targeting antibody GNbAC1 (now named temelimab) is currently in phase 2b clinical trials for the treatment of multiple sclerosis (MS) and displays promising results without noticeable adverse effects (26, 27). In patients, teme-limab has been shown to bind HERV-W Env, which has been associated with the development of certain autoimmunity-mediated diseases such as MS, type 1 diabetes mellitus, and chronic inflammatory demyelinating polyneuropathy (CIDP), although displaying specific cross-reactivity to ERVW-1 only at high concentrations (26). A search on the CancerHERVdb website indicates HERV-W Env protein expression in 136/161 (84.5%) colorectal cancer samples (28, 29), 84/103 (81.55%) liver cancer samples (30), and 62/82 (75.6%) urothelial cell carcinoma samples (31). Furthermore, a search showed HERV-W Env protein or RNA detected in leukemia, endocrine cancer, uterus cancer, testicular cancer, placental cancer, ovary cancer, renal cell carcinoma, and astrocytoma samples. In addition, skin cancer, prostate cancer, pancreas cancer, lung cancer, cervical cancer, and breast cancer cell lines were previously reported to display HERV-W Env protein or RNA expression. Alternatively, the restriction factor suppressyn (HGCN: ERVH48-1), a HERV-F-derived inhibitor of ERVW-1-mediated fusion, might be an effective therapeutic in these cancers.

While immune checkpoint blockade inhibitors are battling therapeutic resistance in the clinic, HERV-H LTR-associating protein 2 (HHLA2) is considered the next immune checkpoint for antitumor therapy (32). A search for HERV-H LTR expression on the CancerHERVdb website indicated HHLA2 protein expression in 164/228 (71.9%) pancreatic ductal adenocarcinoma samples (33, 34), 138/201 (68.66%) oral squamous cell carcinoma samples (35), 82/134 (61.1%) gastrointestinal cancer samples (36), 119/218 (54.59%) intrahepatic cholangiocarcinoma samples (37), 103/202 (50.99%) hepatocellular carcinoma samples (38, 39), 59/126 (46.83%) various lung cancer samples (40), and 198/490 (40.4%) clear cell renal cell carcinoma samples (41, 42). In many cases, our
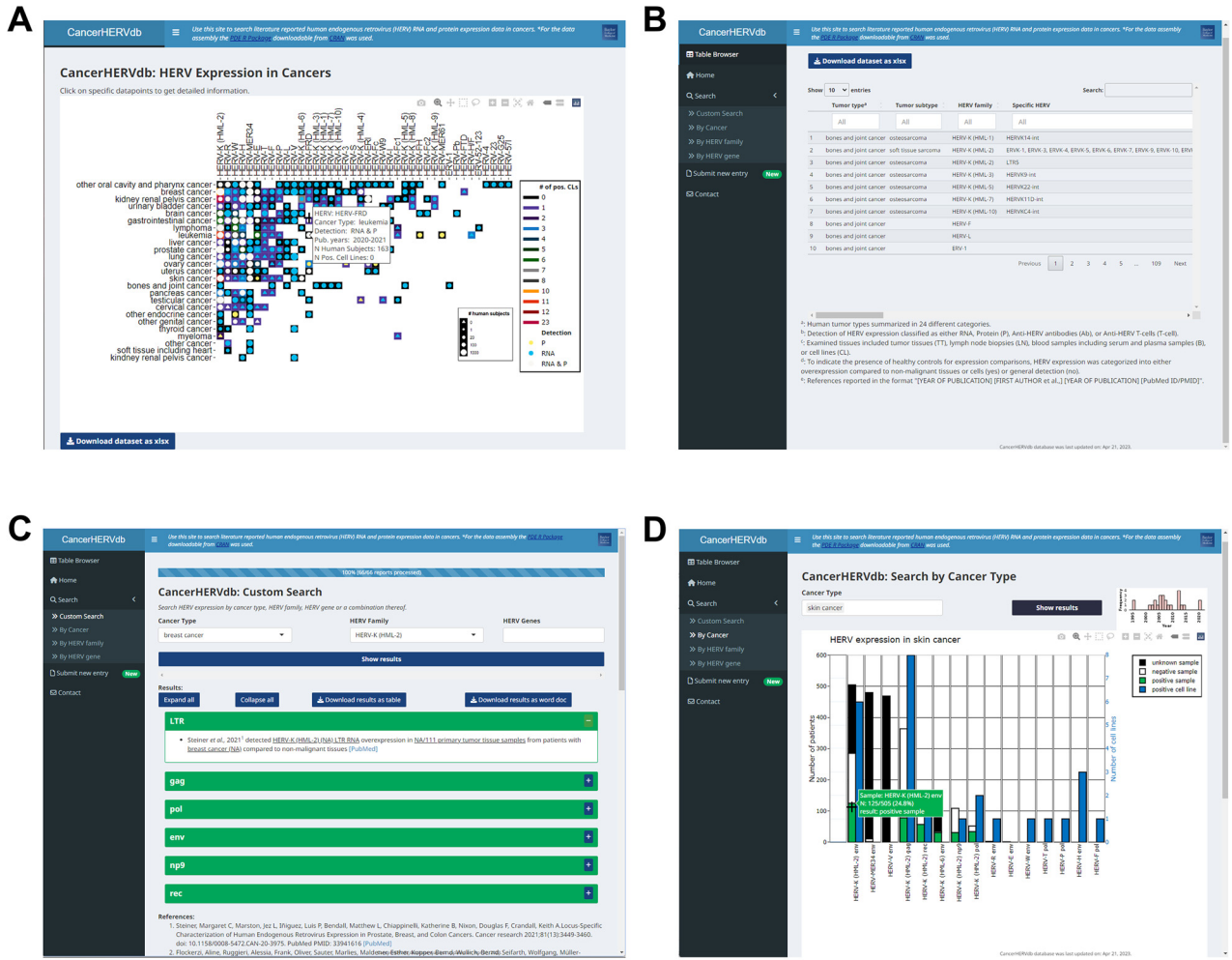
FIG 5 CancerHERVdb website screenshots. (A) Home: landing page displaying an interactive version of Fig. 4, which displays data point information on mouse hover and initiates a custom search on mouse click. (B) Table Browser: data table fully browsable and freely downloadable. (C) Custom Search: interface allowing the user to search for HERV expression data by cancer type, HERV family, HERV gene, or a combination thereof. (D) Search by Cancer Type: specific search interface providing summary data for a single or multiple cancer types. Search by HERV Family and Search by HERV Gene have similar interfaces and result displays. The bar graph is interactive, displaying detailed information on mouse hover and initiating a custom search on mouse click.

analysis revealed verification of the expression data from two or more independent groups. HHLA2-based therapeutics are certainly in the foreseeable future since Bhatt et al. were already successful in developing HHLA2-targeting antibodies that specifically block its immunoinhibitory activity in mouse models (43).

DISCUSSION

The CancerHERVdb database centralized findings on HERV transcription and translation in a variety of cancers. While it was expected that breast, skin, and colorectal cancers would be frequent foci of HERV-associated cancer articles, lung, bladder, and prostate cancers were considerably underrepresented in the HERV literature. Contrary to the assumption that these cancers display low expression of HERVs, a quick search on the CancerHERVdb website reveals 59/116 (46.8%) lung cancer tissue samples with detection of HERV-H LTR expression, 244/1,357 (18%) prostate cancer samples positive for HERV-K *gag* expression, and 72.7 to 100% of bladder cancer tissues positive for HERV-W *env*, HERV-W *pol*, HERV-T *pol*, HERV-Rb *pol*, HERV-K *pol*, HERV-E *gag*, and HERV-E *pol* expression. Interestingly, only a few lung and bladder cancer cell lines have been reported to exhibit HERV expression. Since lung cancers are frequently assessed through imaging technology and pathological examinations and not molecular

methods, only recent studies have started to evaluate HERVs as novel biomarkers (44). In contrast, bladder cancer has been and remains a substantially understudied and underfunded cancer which is also represented in the available data on HERVs in bladder cancers. Publication date distribution analysis showed that early HERV research focused on testicular cancer followed by breast cancer. This was likely a result of the discovery of HERV virus-like particles in testicular cancer cell lines Tera-1 by Bronson et al. in 1979 (45) and GH by Boller et al. in 1993 (46), followed by reports of HERV particles in the breast cancer cell line T47D by Seifarth et al. (1995) (47) and Patience et al. (1996) (48). HERV research peaked in the 2010s as next-generation sequencing methods improved in accuracy and affordability. Current cancer-HERV research indicates a surge in brain cancer studies, which coincides with an increase of studies revealing the relevance of HERVs in neurodegenerative diseases such as multiple sclerosis (MS), amyotrophic lateral sclerosis, Alzheimer's disease, and schizophrenia (49). In addition, clinical trials of GNbAC1, a monoclonal antibody specifically targeting HERV-W Env, for the treatment of multiple sclerosis are promising and will also have a positive impact on cancer research and potential treatments (27).

Our assessment showed that nonhuman cancers dominate the ERV cancer literature. This is not surprising as HERVs were originally discovered as endogenous mouse mammary tumor virus (MMTV)-related sequences (50), and animal models such as pig (51), mouse (52), and koala (53) are frequently researched for the study of infectious endogenous retroviruses. Discoveries of various HERV-K distributions in human populations and frequent observations of insertions and oncogene activation through ERVs in mice (52) and chickens (54, 55) led to the hypothesis of HERV retrotransposition as a cancer driver (56–58). However, no *de novo* germ line or somatic insertions of HERVs could be identified in human tissues or cancers despite the sequencing of more than 3,500 human cancer genomes over the last 20 years (57, 59, 60). The theory that the almost-5-times more abundant long interspersed nuclear elements (LINEs) are better substrates for nonallelic recombination in humans (61–64) is supported by the observation that reverse transcriptase (RT) inhibitors display effects as treatment in some human cancers (65).

In our CancerHERVdb database, we included HERV RNA, HERV protein, anti-HERV antibody, and anti-HERV T-cell expression to provide a thorough overview on potential interference of HERV protein as well as RNA products with cancer pathways. While we discourage the deduction of HERV protein expression from RNA levels, HERVs have been reported to form long noncoding RNAs (lncRNAs) (66–68) as well as antisense transcripts inhibiting transcription (69). We also decided not to report negative results for HERV expression in cancers, as the large variety within HERV families discourages deducing an absence of HERV expression from the lack of detection by a single probe or antibody. In addition, our restrictions in gaining access to all HERV primary literature in the PubMed database allow the potential for missing data. In the future, we intend to review the titles and abstracts of articles without full-text availability for HERV expression data and encourage the authors to submit their data together with a full-text PDF copy of their study to us. As of now, however, a gap in the CancerHERVdb database cannot be equated with an absence of HERV expression.

Even though publications on HERVs and cancer appear to be slowly declining, we predict the overcoming of two obstacles to be catalysts for another surge: (i) the increase in accuracy and scalability of long-read sequencing methods and (ii) the streamlining of findings on HERVs in databases. With the creation of the CancerHERVdb database website, we hope to contribute to overcoming the second obstacle by providing easier access to literature reporting HERV expression data in cancers. The CancerHERVdb website allows a general overview of which HERVs are expressed in a specific cancer in addition to expression information on specific HERV families or genes. This can be useful for the exploration of potential off-target effects of drugs or development of multicancer treatments. Furthermore, the CancerHERVdb website gives easy access to information on rare and understudied cancers and HERV families, such as oral cavity and pharynx cancer or the HERV-Pb family.

To ensure a thorough and complete methodology, we used a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) approach for our evaluation of the relationship between HERVs and cancer (outcome) without focusing on any specific intervention. We demonstrated the use of the PDE R package as an accurate and efficient prescreening alternative to the restrictive manual title and abstract-only assessment used with software like DistillerSR, RevMan, or Covidence. To our knowledge, PDE is the only software allowing high-throughput detection of a list of filter and search words in a large number of PDF files. We proved the usefulness of full-text searches for the detection of side notes, methodologies, and minor results generally not mentioned in the abstract, adding another layer of sensitivity to the literature evaluation. It should be noted that the PDE R package did not completely replace our manual review of all articles but rather reduced the number of manually assessed articles at an early stage and highlighted potentially relevant paragraphs. The PDE is also the only free software which enables the high-throughput export of tables from PDF files.

We recognize that the PDE was highly dependent on the availability and integrity of the PDF full-text articles. Nevertheless, documentation files were created for secured, nonreadable, or image-only PDF files or tables. In addition, PDE-driven analysis can always be supplemented by a title and abstract screen of non-readily available articles. At this point, the PDE does not distinguish between the article body and the reference section, leading to a notable fraction of the filter words being detected via the reference section. While this still indicates the general topic of the article, the risk of false positives arising from reference inclusion is evident, and consequently, features intended to be implemented in future updates of the PDE package include the voluntary exclusion of the reference section for search word detection. Lastly, the sensitivity and specificity of the PDE-driven prescreen are predominantly influenced by the selection of the filter words and filter word threshold. While abbreviation detection (e.g., detection of "ALL" in context when using "acute lymphoblastic leukemia" as a filter word) aids in processing comprehensiveness, a thorough filter word library is essential. Even though there always remains a risk of underdetection, our analyses continuously showed that relevant articles contained a filter word ratio significantly above the average. In conclusion, we showed that the PDE R package can enhance the literature search for review articles, gene or disease curation, risk factor analysis, and general literature reviews.

The CancerHERVdb website should be used with the awareness of certain limitations. First, the database includes results from studies with varying quality without providing a summarizing quality score. While a ranking system, for example, with a reducibility score of one to five stars assigned to each report, would be of certain advantage, we did not have access to an automated scoring algorithm and opted against assigning a subjective score. The accuracy and quality of HERV detection techniques are rapidly improving, which would suggest a dynamic rather than a fixed score. Therefore, the data from the CancerHERVdb database might be useful to train a scoring algorithm, although this exceeded the scope of this study. Nonetheless, we provide the user with information which allows a relative comparison of studies. For examples, detection type, patient and cell line numbers tested, publication number, and date ranges are available for each data point on mouse hover on the landing page of the CancerHERVdb website (Fig. 5A), while detailed information on detected HERV loci is accessible in the summary data table (Fig. 5B) or through a search on the website (Fig. 5C). Second, several factors are contributing to an increased loading time on the CancerHERVdb website: (i) all graphical displays are generated dynamically from the underlying data table to provide the most up-to-date visualizations, (ii) the CancerHERVdb is coded in R shiny and has undergone only basic processing optimization, and (iii) the CancerHERVdb website is hosted on a free server, restricting computing power. Lastly, the CancerHERVdb database allows no relative comparisons of expression levels between studies and healthy control databases. Nevertheless, to best evaluate the suitability of certain HERVs as therapeutic targets or biomarkers, information on expression levels relative to healthy control tissues is certainly of advantage. Therefore, in the summary data table and

search reports on the website, we indicated studies where nonmalignant tissues or cells were used to describe relative overexpression and contrasted them with reports with general detection of a specific HERV expression. However, as demonstrated in a study assessing the comparability of locus-specific HERV transcriptome data sets by Hamann et al. (70), variable read depths and batch effects make comparison between studies challenging. This is especially true for less abundant HERV elements, which require similar read depth and the application of computational batch effect reduction to distinguish nondetection from nonpresence. In addition, depending on the specific primers and antibodies used for expression analyses, only certain subpopulations of HERVs are detected. Most studies used a preselected combination of primers and antibodies and thus do not capture the large range of HERV sequence variability. Therefore, seemingly contradictory absences of detection reported by different authors might be explained in this way. In addition, the data displayed on the CancerHERVdb website are restricted only to HERV family and gene resolution; nonetheless, specific HERV loci, where available, can be found in the table browser.

In conclusion, the CancerHERVdb database summarizes the literature on HERV RNA and protein expression in different cancers, allowing cross-study comparisons. In addition, we demonstrated the productive use of the PDE R package for automated literature processing and created a website called CancerHERVdb providing interactive access to our created database. We still believe that streamlining HERV sequences, conserved insertion sites, and nomenclature is of essence to further increase comparability of studies, and we plan to incorporate advances in any of these areas also in our CancerHERVdb website.

MATERIALS AND METHODS

Assembly of the full-text article library. We identified articles on ERVs published between 1 January 1957 and 5 May 2022 through PubMed using the following search terms in titles and abstracts: {(HERV[Title/Abstract]) OR (ERV[Title/Abstract]) OR (endogenous retrovirus[Title/Abstract]) OR (endogenous retroviral[Title/Abstract])} AND (1957/01/01:2022/05/05[pdat]). Using the corresponding filter in PubMed, we identified review articles, evaluated them for articles missed by the PubMed search, included corresponding articles in our systematic review, and excluded the review articles themselves using the PubMed filter mask from the bioinformatic analysis. We downloaded open access papers using the PubMed-Batch-Download software developed by Bill Greenwald (71), supplemented with a manual download through PubMed with Texas Medical Center (TMC) library access in portable document format (PDF). We obtained articles with restricted access through the Texas Medical Center Library using the OpenAthens plugin in EndNote. Only articles accessible and available in English were included.

For the categorization and evaluation of the articles on HERVs and cancer, we used our recently developed R package called PDF data extractor (PDE) available on CRAN (https://CRAN.R-project.org/package=PDE) (3, 18). We first searched all full-text HERV-related primary research articles for 97 general cancer keywords (see Table S2 in the supplemental material) and separated cancer-related from cancer-unrelated articles automatically using the filter word parameter of the PDE analyzer tool. Table S2 comprises the tsv file used for the analysis. For the PDE-facilitated filter categorization, we considered papers with ≥0.2% of all words being general cancer keywords to be cancer related. We evaluated all nonreadable files and later added the articles manually without the use of the PDE package. In the manual review, exclusion criteria comprised the absence of cancer data (noncancer articles), the focus on animal ERVs, animal cancers studied or the exclusive use of animal model systems (animal ERV article), the deficiency of HERV data (no HERV article), review articles, exclusive studies of pathways (general mechanistic), previews/debates/editorials, focus on non-HERV viruses (focus on other viruses), absence of primer sequences, probe or antibody descriptions (poor quality), and method articles. Using 366 cancer-type-specific keywords and the search word parameter of the PDE analyzer, we automatically sorted the resulting cancer-related HERV papers into 26 cancer categories, including nonhuman cancers, according to the cancer-type-specific keywords detected most often (Tables S2 and S3). We evaluated all downloaded full-text articles with the PDE reader to quickly identify false positives, i.e., noneligible articles. Lastly, we manually evaluated the identified full-text articles as well as all secured PDF files for HERV expression data, conclusions, and impact. For the qualitative synthesis, we weighted the scientific findings by number of articles reporting similar results and excluded findings that were based on inaccurately described or performed methods (e.g., lack of detail, no technical replicates, or inappropriate method for a certain conclusion). Articles on animal endogenous retroviruses were not considered for this study as it has been shown that data from animal systems are minimally applicable to human systems (72). In the same way, papers on nonhuman cancers were excluded from the full-text evaluation.

Extraction of HERV expression data. We evaluated all articles indicated by the PDE analyzer and manually verified them to be HERV and cancer related and containing data on HERV RNA, HERV protein, anti-HERV antibody, and/or anti-HERV T-cell expression. For each finding of expression data, we documented the following information: tumor category group (according to the 26 cancer categories based on the cancer-type-specific keywords), specific cancer type if available, HERV family, specific HERV loci if

available, detection (RNA or protein), gene products, number of tested and positive samples if available, number of tested and positive controls if available, cell lines examined if available, tissue examined, overexpression compared to nonmalignant tissues/cells or general expression, viral particles detected, PubMed identifier (ID), first author, and publication year (see Table S4 in the supplemental material). We assessed the quality of the expression data based on completeness of the methodology description. Where HCGN names for specific HERV loci were not available, chromosome bandings were used to match HERV sequences. In addition, HERV annotation tables included in the work of Subramanian et al. (73) and Broecker et al. (74) were used to assign unifying names to expressed HERV elements. Articles not providing the necessary data were excluded.

Visualization and database generation. We visualized the tabularized data using the ggplot2 (version 3.3.6) R package (75) in conjunction with the Cairo R package (version 1.5-15) (76). We used the dplyr (version 1.0.9) (77), stringr (version 1.4.0) (78), and R.utils (version 2.11.0) (79) packages to aid in data processing and analysis. To provide broad accessibility to the results, we generated interactive and dynamic visualizations of the results using the plotly R package (version 4.10.0) (80) and designed an R Shiny dashboard using the shiny (version 1.7.1) (81), shinyFiles (version 0.9.1) (82), shinydashboard (version 0.7.2) (83), g (version 2.0.3) (84), shinyjs (version 2.1.0) (85), shinymanager (version 1.0.400) (86), shinycssloaders (version 1.0.0) (87), mailR (version 0.8) (88), zip (version 2.2.0) (89), DT (version 0.22) (90), openxlsx (version 4.2.5) (91), and officer (version 0.4.2) (92) R packages. We incorporated dynamic citation retrieval for the R Shiny dashboard using the easyPubMed (version 2.13) (93) and rentrez (version 1.2.3) (94) R packages. The R Shiny dashboard app was published at https://erikstricker.shinyapps.io/cancerHERVdb/ through shinyapps.io by RStudio, and the underlying data table is updated quarterly based on submissions through the website and new publications. All scripts were executed on R version 4.2.0 (95).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.
SUPPLEMENTAL FILE 1, XLSX file, 0.3 MB.
SUPPLEMENTAL FILE 2, XLSX file, 0.02 MB.
SUPPLEMENTAL FILE 3, CSV file, 0.01 MB.
SUPPLEMENTAL FILE 4, XLSX file, 0.1 MB.

REFERENCES

1. Bannert N, Kurth R. 2004. Retroelements and the human genome: new perspectives on an old relation. Proc Natl Acad Sci U S A 101:14572–14579. https://doi.org/10.1073/pnas.0404838101.

2. Deininger PL, Batzer MA. 2002. Mammalian retroelements. Genome Res 12:1455–1465. https://doi.org/10.1101/gr.282402.

3. Stricker E, Peckham-Gregory EC, Scheurer ME. 2023. HERVs and cancer—a comprehensive review of the relationship of human endogenous retroviruses and human cancers. Biomedicines 11:936. https://doi.org/10.3390/biomedicines11030936.

4. Serafino A, Balestrieri E, Pierimarchi P, Matteucci C, Moroni G, Oricchio E, Rasi G, Mastino A, Spadafora C, Garaci E, Vallebona PS. 2009. The activation of human endogenous retrovirus K (HERV-K) is implicated in melanoma cell malignant transformation. Exp Cell Res 315:849–862. https://doi.org/10.1016/j.yexcr.2008.12.023.

5. Balestrieri E, Argaw-Denboba A, Gambacurta A, Cipriani C, Bei R, Serafino A, Sinibaldi-Vallebona P, Matteucci C. 2018. Human endogenous retrovirus K in the crosstalk between cancer cells microenvironment and plasticity: a new perspective for combination therapy. Front Microbiol 9:1448. https://doi.org/10.3389/fmicb.2018.01448.

6. Cherkasova E, Malinzak E, Rao S, Takahashi Y, Senchenko VN, Kudryavtseva AV, Nickerson ML, Merino M, Hong JA, Schrump DS, Srinivasan R, Linehan WM, Tian X, Lerman MI, Childs RW. 2011. Inactivation of the von Hippel-Lindau tumor suppressor leads to selective expression of a human endogenous retrovirus in kidney cancer. Oncogene 30:4697–4706. https://doi.org/10.1038/onc.2011.179.

7. Boese A, Sauter M, Galli U, Best B, Herbst H, Mayer J, Kremmer E, Roemer K, Mueller-Lantzsch N. 2000. Human endogenous retrovirus protein cORF supports cell transformation and associates with the promyelocytic leukemia zinc finger protein. Oncogene 19:4328–4336. https://doi.org/10.1038/sj.onc.1203794.

8. Armbruester V, Sauter M, Roemer K, Best B, Hahn S, Nty A, Schmid A, Philipp S, Mueller A, Mueller-Lantzsch N. 2004. Np9 protein of human endogenous retrovirus K interacts with ligand of numb protein X. J Virol 78:10310–10319. https://doi.org/10.1128/JVI.78.19.10310-10319.2004.

9. Denne M, Sauter M, Armbruester V, Licht JD, Roemer K, Mueller-Lantzsch N. 2007. Physical and functional interactions of human endogenous retrovirus proteins Np9 and rec with the promyelocytic leukemia zinc finger protein. J Virol 81:5607–5616. https://doi.org/10.1128/JVI.02771-06.

10. Lindemann D, Steffen I, Pohlmann S. 2013. Cellular entry of retroviruses. Adv Exp Med Biol 790:128–149. https://doi.org/10.1007/978-1-4614-7651-1_7.

11. Mullins CS, Linnebacher M. 2012. Endogenous retrovirus sequences as a novel class of tumor-specific antigens: an example of HERV-H env encoding strong CTL epitopes. Cancer Immunol Immunother 61:1093–1100. https://doi.org/10.1007/s00262-011-1183-3.

12. Probst P, Stringhini M, Ritz D, Fugmann T, Neri D. 2019. Antibody-based delivery of TNF to the tumor neovasculature potentiates the therapeutic activity of a peptide anticancer vaccine. Clin Cancer Res 25:698–709. https://doi.org/10.1158/1078-0432.CCR-18-1728.

13. Sacha JB, Kim IJ, Chen L, Ullah JH, Goodwin DA, Simmons HA, Schenkman DI, von Pelchrzim F, Gifford RJ, Nimityongskul FA, Newman LP, Wildeboer S, Lappin PB, Hammond D, Castrovinci P, Piaskowski SM, Reed JS, Beheler KA, Tharmanathan T, Zhang N, Muscat-King S, Rieger M, Fernandes C, Rumpel K, Gardner JP, II, Gebhard DH, Janies J, Shoieb A, Pierce BG, Trajkovic D, Rakasz E, Rong S, McCluskie M, Christy C, Merson JR, Jones RB, Nixon DF, Ostrowski MA, Loudon PT, Pruimboom-Brees IM, Sheppard NC. 2012. Vaccination with cancer- and HIV infection-associated endogenous retrotransposable elements is safe and immunogenic. J Immunol 189:1467–1479. https://doi.org/10.4049/jimmunol.1200079.

14. Kraus B, Fischer K, Sliva K, Schnierle BS. 2014. Vaccination directed against the human endogenous retrovirus-K (HERV-K) gag protein slows HERV-K gag expressing cell growth in a murine model system. Virol J 11:58. https://doi.org/10.1186/1743-422X-11-58.

15. Krishnamurthy J, Rabinovich BA, Mi T, Switzer KC, Olivares S, Maiti SN, Plummer JB, Singh H, Kumaresan PR, Huls HM, Wang-Johanning F, Cooper LJ. 2015. Genetic engineering of T cells to target HERV-K, an ancient retrovirus on melanoma. Clin Cancer Res 21:3241–3251. https://doi.org/10.1158/1078-0432.CCR-14-3197.

16. Chiappinelli KB, Strissel PL, Desrichard A, Li H, Henke C, Akman B, Hein A, Rote NS, Cope LM, Snyder A, Makarov V, Budhu S, Slamon DJ, Wolchok JD, Pardoll DM, Beckmann MW, Zahnow CA, Merghoub T, Chan TA, Baylin SB, Strick R. 2015. Inhibiting DNA methylation causes an interferon response in cancer via dsRNA including endogenous retroviruses. Cell 162:974–986. https://doi.org/10.1016/j.cell.2015.07.011.

17. Bonaventura P, Alcazer V, Mutez V, Tonon L, Martin J, Chuvin N, Michel E, Boulos RE, Estornes Y, Valladeau-Guilemond J, Viari A, Wang Q, Caux C, Depil S. 2022. Identification of shared tumor epitopes from endogenous retroviruses inducing high-avidity cytotoxic T cells for cancer immunotherapy. Sci Adv 8:eabj3671. https://doi.org/10.1126/sciadv.abj3671.

18. Stricker E, Scheurer ME. 2021. PDF Data Extractor (PDE) - a free web application and R package allowing the extraction of tables from portable document format (PDF) files and high-throughput keyword searches of full-text articles. bioRxiv. https://doi.org/10.1101/2021.07.13.452159.

19. Chen J, Foroozesh M, Qin Z. 2019. Transactivation of human endogenous retroviruses by tumor viruses and their functions in virus-associated malignancies. Oncogenesis 8:6. https://doi.org/10.1038/s41389-018-0114-y.

20. Andersson AC, Merza M, Venables P, Ponten F, Sundstrom J, Cohen M, Larsson E. 1996. Elevated levels of the endogenous retrovirus ERV3 in human sebaceous glands. J Invest Dermatol 106:125–128. https://doi.org/10.1111/1523-1747.ep12329612.

21. Azebi S, Batsche E, Michel F, Kornobis E, Muchardt C. 2019. Expression of endogenous retroviruses reflects increased usage of atypical enhancers in T cells. EMBO J 38:e101107. https://doi.org/10.15252/embj.2018101107.

22. Buslei R, Strissel PL, Henke C, Schey R, Lang N, Ruebner M, Stolt CC, Fabry B, Buchfelder M, Strick R. 2015. Activation and regulation of endogenous retroviral genes in the human pituitary gland and related endocrine tumours. Neuropathol Appl Neurobiol 41:180–200. https://doi.org/10.1111/nan.12136.

23. Diaz-Carballo D, Saka S, Klein J, Rennkamp T, Acikelli AH, Malak S, Jastrow H, Wennemuth G, Tempfer C, Schmitz I, Tannapfel A, Strumberg D. 2018. A distinct oncogenerative multinucleated cancer cell serves as a source of stemness and tumor heterogeneity. Cancer Res 78:2318–2331. https://doi.org/10.1158/0008-5472.CAN-17-1861.

24. Kang YJ, Jo JO, Ock MS, Chang HK, Baek KW, Lee JR, Choi YH, Kim WJ, Leem SH, Kim HS, Cha HJ. 2014. Human ERV3-1 env protein expression in various human tissues and tumours. J Clin Pathol 67:86–90. https://doi.org/10.1136/jclinpath-2013-201841.

25. Hokamp K, Wolfe KH. 2004. PubCrawler: keeping up comfortably with PubMed and GenBank. Nucleic Acids Res 32:W16–W19. https://doi.org/10.1093/nar/gkh453.

26. Kornmann G, Curtin F. 2020. Temelimab, an IgG4 anti-human endogenous retrovirus monoclonal antibody: an early development safety review. Drug Saf 43:1287–1296. https://doi.org/10.1007/s40264-020-00988-3.

27. Hartung HP, Derfuss T, Cree BA, Sormani MP, Selmaj K, Stutters J, Prados F, MacManus D, Schneble HM, Lambert E, Porchet H, Glanzman R, Warne D, Curtin F, Kornmann G, Buffet B, Kremer D, Kury P, Leppert D, Ruckle T,

Barkhof F. 2022. Efficacy and safety of temelimab in multiple sclerosis: results of a randomized phase 2b and extension study. Mult Scler 28:429–440. https://doi.org/10.1177/13524585211024997.

28. Larsen JM, Christensen IJ, Nielsen HJ, Hansen U, Bjerregaard B, Talts JF, Larsson LI. 2009. Syncytin immunoreactivity in colorectal cancer: potential prognostic impact. Cancer Lett 280:44–49. https://doi.org/10.1016/j.canlet.2009.02.008.

29. Diaz-Carballo D, Acikelli AH, Klein J, Jastrow H, Dammann P, Wyganowski T, Guemues C, Gustmann S, Bardenheuer W, Malak S, Tefett NS, Khosrawipour V, Giger-Pabst U, Tannapfel A, Strumberg D. 2015. Therapeutic potential of antiviral drugs targeting chemorefractory colorectal adenocarcinoma cells overexpressing endogenous retroviral elements. J Exp Clin Cancer Res 34:81. https://doi.org/10.1186/s13046-015-0199-5.

30. Zhou Y, Liu L, Liu Y, Zhou P, Yan Q, Yu H, Chen X, Zhu F. 2021. Implication of human endogenous retrovirus W family envelope in hepatocellular carcinoma promotes MEK/ERK-mediated metastatic invasiveness and doxorubicin resistance. Cell Death Discov 7:177. https://doi.org/10.1038/s41420-021-00562-5.

31. Yu H, Liu T, Zhao Z, Chen Y, Zeng J, Liu S, Zhu F. 2014. Mutations in 3′-long terminal repeat of HERV-W family in chromosome 7 upregulate syncytin-1 expression in urothelial cell carcinoma of the bladder through interacting with c-Myb. Oncogene 33:3947–3958. https://doi.org/10.1038/onc.2013.366.

32. Zhao R, Chinai JM, Buhl S, Scandiuzzi L, Ray A, Jeon H, Ohaegbulam KC, Ghosh K, Zhao A, Scharff MD, Zang X. 2013. HHLA2 is a member of the B7 family and inhibits human CD4 and CD8 T-cell function. Proc Natl Acad Sci U S A 110:9879–9884. https://doi.org/10.1073/pnas.1303524110.

33. Chen Q, Wang J, Chen W, Zhang Q, Wei T, Zhou Y, Xu X, Bai X, Liang T. 2019. B7-H5/CD28H is a co-stimulatory pathway and correlates with improved prognosis in pancreatic ductal adenocarcinoma. Cancer Sci 110:530–539. https://doi.org/10.1111/cas.13914.

34. Yan H, Qiu W, Koehne de Gonzalez AK, Wei JS, Tu M, Xi CH, Yang YR, Peng YP, Tsai WY, Remotti HE, Miao Y, Su GH. 2019. HHLA2 is a novel immune checkpoint protein in pancreatic ductal adenocarcinoma and predicts post-surgical survival. Cancer Lett 442:333–340. https://doi.org/10.1016/j.canlet.2018.11.007.

35. Xiao Y, Li H, Yang LL, Mao L, Wu CC, Zhang WF, Sun ZJ. 2019. The expression patterns and associated clinical parameters of human endogenous retrovirus-H long terminal repeat-associating protein 2 and transmembrane and immunoglobulin domain containing 2 in oral squamous cell carcinoma. Dis Markers 2019:5421985. https://doi.org/10.1155/2019/5421985.

36. Zhu Z, Dong W. 2018. Overexpression of HHLA2, a member of the B7 family, is associated with worse survival in human colorectal carcinoma. Onco Targets Ther 11:1563–1570. https://doi.org/10.2147/OTT.S160493.

37. Jing CY, Fu YP, Yi Y, Zhang MX, Zheng SS, Huang JL, Gan W, Xu X, Lin JJ, Zhang J, Qiu SJ, Zhang BH. 2019. HHLA2 in intrahepatic cholangiocarcinoma: an immune checkpoint with prognostic significance and wider expression compared with PD-L1. J Immunother Cancer 7:77. https://doi.org/10.1186/s40425-019-0554-8.

38. Luo M, Lin Y, Liang R, Li Y, Ge L. 2021. Clinical significance of the HHLA2 protein in hepatocellular carcinoma and the tumor microenvironment. J Inflamm Res 14:4217–4228. https://doi.org/10.2147/JIR.S324336.

39. Wang R, Guo H, Tang X, Zhang T, Liu Y, Zhang C, Yu H, Li Y. 2022. Interferon gamma-induced interferon regulatory factor 1 activates transcription of HHLA2 and induces immune escape of hepatocellular carcinoma cells. Inflammation 45:308–330. https://doi.org/10.1007/s10753-021-01547-3.

40. Farrag MS, Ibrahim EM, El-Hadidy TA, Akl MF, Elsergany AR, Abdelwahab HW. 2021. Human endogenous retrovirus-H long terminal repeat-associating protein 2 (HHLA2) is a novel immune checkpoint protein in lung cancer which predicts survival. Asian Pac J Cancer Prev 22:1883–1889. https://doi.org/10.31557/APJCP.2021.22.6.1883.

41. Chen L, Zhu D, Feng J, Zhou Y, Wang Q, Feng H, Zhang J, Jiang J. 2019. Overexpression of HHLA2 in human clear cell renal cell carcinoma is significantly associated with poor survival of the patients. Cancer Cell Int 19:101. https://doi.org/10.1186/s12935-019-0813-2.

42. Zhou QH, Li KW, Chen X, He HX, Peng SM, Peng SR, Wang Q, Li ZA, Tao YR, Cai WL, Liu RY, Huang H. 2020. HHLA2 and PD-L1 co-expression predicts poor prognosis in patients with clear cell renal cell carcinoma. J Immunother Cancer 8:e000157. https://doi.org/10.1136/jitc-2019-000157.

43. Bhatt RS, Berjis A, Konge JC, Mahoney KM, Klee AN, Freeman SS, Chen CH, Jegede OA, Catalano PJ, Pignon JC, Sticco-Ivins M, Zhu B, Hua P, Soden J, Zhu J, McDermott DF, Arulanandam AR, Signoretti S, Freeman GJ. 2021. KIR3DL3 is an inhibitory receptor for HHLA2 that mediates an alternative immunoinhibitory pathway to PD1. Cancer Immunol Res 9:156–169. https://doi.org/10.1158/2326-6066.CIR-20-0315.

apoptosis, an indicator of malignant transformation. Mob DNA 7:25. https://doi.org/10.1186/s13100-016-0081-9.

75. Wickham H. 2016. ggplot2: elegant graphics for data analysis. Springer-Verlag, New York, NY.

76. Urbanek S, Horner J. 2022. Cairo: R graphics device using Cairo graphics library for creating high-quality bitmap (PNG, JPEG, TIFF), vector (PDF, SVG, PostScript) and display (X11 and Win32) output.

77. Wickham H, François R, Henry L, Müller K. 2021. dplyr: a grammar of data manipulation, v1.0.7. https://CRAN.R-project.org/package=dplyr.

78. Wickham H. 2019. stringr: simple, consistent wrappers for common string operations.

79. Bengtsson H. 2021. R.utils: various programming utilities.

80. Sievert C. 2020. Interactive web-based data visualization with R, plotly, and shiny.

81. Chang W, Cheng J, Allaire JJ, Sievert C, Schloerke B, Xie Y, Allen J, McPherson J, Dipert A, Borges B. 2021. shiny: web application framework for R.

82. Pedersen TL, Nijs V, Schaffner T, Nantz E. 2020. shinyFiles: a server-side file system viewer for Shiny.

83. Chang W, Borges Ribeiro B. 2018. shinydashboard: create dashboards with 'Shiny'.

84. Granjon D. 2021. shinydashboardPlus: add more 'AdminLTE2' components to 'shinydashboard'.

85. Attali D. 2020. shinyjs: easily improve the user experience of your Shiny apps in seconds.

86. Thieurmel B, Perrier V. 2021. shinymanager: authentication management for 'Shiny' applications.

87. Sali A, Attali D. 2020. shinycssloaders: add loading animations to a 'shiny' output while it's recalculating.

88. Premraj R. 2021. mailR: a utility to send emails from R.

89. Csárdi G, Podgórski K, Geldreich R. 2020. zip: cross-platform 'zip' compression.

90. Xie Y, Cheng J, Tan X. 2022. DT: a wrapper of the JavaScript library 'DataTables'.

91. Schauberger P, Walker A. 2021. openxlsx: read, write and edit xlsx files.

92. Gohel D. 2022. officer: manipulation of Microsoft Word and PowerPoint documents.

93. Fantini D. 2019. easyPubMed: search and retrieve scientific publication records from PubMed.

94. Winter DJ. 2017. rentrez: an R package for the NCBI eUtils API. R J 9:520–526. https://doi.org/10.32614/RJ-2017-058.

95. R Core Team. 2020. R: a language and environment for statistical computing.