Emerging Microbes & Infections
2025, VOL. 14, 2494704 (12 pages)
https://doi.org/10.1080/22221751.2025.2494704

OPEN ACCESS Check for updates

A tombus-like virus in patients with lower respiratory tract infection: an observational study based on meta-transcriptomic sequencing

Run-Ze Ye[a,b*], Jia-Qi Zhao[d*], Hui Xie[c*], Lin Zhao[d*], Cheng Gong[c], Zhen-Fei Wang[d], Ning Yue[e,f], Luo-Yuan Xia[b], Ke Song[d], Bao Dong[b], Ning Wang[d], Wan-Ying Gao[d], Yu-Yu Li[g], Xiao-Ming Cui[b], Jiao-Jiao Pang[a], De-Dong Ma[h], Hao Wang[i], Jia-Fu Jiang[b,j], Yule Liu[e,f], Ye Feng[b], Na Jia[b,j], Wen-Qing Sun[k], Xiao-Peng Qi[l], Lu-Tao Du[m], Yuguo Chen[a*], Tao Jiang[b*], Fang Huang[c*] and Wu-Chun Cao[b,d,j*]

[a]Department of Emergency Medicine, Qilu Hospital of Shandong University, Jinan, People's Republic of China; [b]State Key Laboratory of Pathogen and Biosecurity, Academy of Military Medical Sciences, Beijing, People's Republic of China; [c]Beijing Center for Disease Prevention and Control; Beijing Academy for Preventive Medicine, Beijing, People's Republic of China; [d]Institute of EcoHealth, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, People's Republic of China; [e]MOE Key Laboratory of Bioinformatics, Center for Plant Biology, School of Life Sciences, Tsinghua University, Beijing, People's Republic of China; [f]Tsinghua-Peking Center for Life Sciences, Beijing, People's Republic of China; [g]Institute of Pathogen Biology, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, People's Republic of China; [h]Department of Pulmonary and Critical Care Medicine, Qilu Hospital, Shandong University, Jinan, People's Republic of China; [i]Department of Critical Care Medicine, Qilu Hospital of Shandong University, Jinan, People's Republic of China; [j]Research Unit of Discovery and Tracing of Natural Focus Diseases, Chinese Academy of Medical Sciences, Beijing, People's Republic of China; [k]Department of Intensive Care Unit, Public Health Clinical Center Affiliated to Shandong University, Jinan, People's Republic of China; [l]Key Laboratory for Experimental Teratology of the Ministry of Education, Advanced Medical Research Institute, Cheeloo College of Medicine, Shandong University, Jinan, People's Republic of China; [m]Department of Clinical Laboratory, Qilu Hospital of Shandong University, Shandong Provincial Key Laboratory of Innovation Technology in Laboratory Medicine, Jinan, People's Republic of China

ABSTRACT

The identification of a novel virus related to the family *Tombusviridae*, provisionally named human tombus-like virus (hTLV), is significant in the context of ongoing surveillance for respiratory pathogens. Meta-transcriptomic sequencing was utilized to detect respiratory pathogens in patients with lower respiratory tract infections (LRIs) in Jinan, China, from 2022 to 2023. The additional hTLV infections were identified through retrospective analysis of meta-transcriptome data collected in Beijing, China, from 2016 to 2019, prior to the COVID-19 outbreak. Phylogenetic analyzes indicated that hTLVs were clustered with a Jingmen tombus-like virus 2 but in a distinct clade. The hTLVs genomes consist of a single-stranded positive-sense RNA genomes of 4.7–4.8 kb in size, and contained four putative open reading frames (ORF1–4). The RNA-dependent RNA polymerase protein of hTLV shared significant sequence similarity containing three conserved motifs with 15, 24, and 15 amino acids, respectively. The hTLV genome included the canonical Gly_{376}-Asp_{377}-Asp_{378} (GDD) catalytic residues, which were a unifying feature of viruses in the family *Tombusviridae*. The main clinical manifestations of the 23 patients were fever, cough, expectoration and dyspnea, with varying degrees of lung infection or abnormalities in other laboratory indicators. Serological studies showed that fourfold rise in IgG titers in sera of a patient between acute and convalescent phase by ELISA. Identification of the pathogens for acute respiratory tract infections is essential for timely public health interventions and clinical management. The discovery of a novel virus, hTLV, in patients with LRIs highlights the continuous emergence of new respiratory pathogens in humans.

ARTICLE HISTORY Received 1 February 2025; Revised 11 April 2025; Accepted 13 April 2025

KEYWORDS Human tombus-like virus; Tombus-like virus; viral discovery; viral genome characteristics; phylogenetic analysis; China

Introduction

Lower respiratory tract infections (LRIs), defined as clinician-confirmed or radiologically confirmed pneumonia or bronchiolitis [1,2], are represent a leading infectious cause of morbidity and mortality, imposing a significant global disease burden [3]. Various bacteria, viruses, and fungi are responsible for LRIs, with viruses as major pathogens. Although intensive etiological investigations have been conducted, the causative agents remain unknown in a considerable proportion of LRIs. The global COVID-19 pandemic further emphasizes the urgent need for identifying

CONTACT Yuguo Chen chen919085@sdu.edu.cn Department of Emergency Medicine, Qilu Hospital of Shandong University, Jinan, Shandong 250012, People's Republic of China; Fang Huang hhffxdd@126.com Beijing Center for Disease Prevention and Control; Beijing Academy for Preventive Medicine, Beijing 100013, People's Republic of China; Wu-Chun Cao caowuchun@126.com, jiangtao@bmi.ac.cn State Key Laboratory of Pathogen and Biosecurity, Academy of Military Medical Sciences, Beijing 100071, People's Republic of China
*These authors Run-Ze Ye, Jia-Qi Zhao, Hui Xie and Lin Zhao contributed equally. The senior authors Yuguo Chen, Tao Jiang, Fang Huang and Wu-Chun Cao also contributed equally.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/22221751.2025.2494704.

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group, on behalf of Shanghai Shangyixun Cultural Communication Co., Ltd
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

the etiology of acute respiratory tract infections for early public health interventions and effective clinical management [4]. We have discovered a previously unrecognized virus belonging to the family *Tombusviridae* of order *Tolivirales* through meta-transcriptomic sequencing of sputum, bronchoalveolar lavage fluid (BALF) and oropharyngeal swab samples from 23 epidemiologically unrelated patients with LRIs based on an etiological surveillance project at Jinan, Shandong Province and a retrospective study at Beijing metropolis, China [5].

The members of the family *Tombusviridae* have traditionally been known to infect plants. However, the recent phylogenetic analyzes based on viral genomes obtained from meta-transcriptomic sequencing revealed that a diversity of novel viruses in the family *Tombusviridae* was associated with invertebrates and mammals such as rabbits and calves [6–8], suggesting that the host distribution of *Tombusviridae* may be more diverse than previously recognized. To date, it remains unknown whether the members of the family *Tombusviridae* are able to cause infections or diseases in humans. Identification of novel viruses in clinical samples is challenging due to low viral loads and limitations of traditional diagnostic methods, which require prior knowledge of viral genomes or antigenic properties [9,10]. In contrast to traditional methods, meta-transcriptomic sequencing enables unbiased detection of both known and unknown RNA viruses. It facilitates the discovery and identification of viruses, especially in the investigation of unexplained disease outbreaks or emerging infectious diseases. This approach enables researchers to identify novel viral sequences, understand the genome structure of viruses, and reconstruct evolutionary relationships [11–14]. This approach has uncovered novel viruses, such as human bocavirus and human rhinovirus group C, expanding our understanding of virus diversity. Importantly, many newly identified viruses are associated with respiratory diseases that threaten individual and population-level health security [15,16]. In this study, we explored potential novel viruses in patients with lower respiratory tract infections through meta-transcriptomic sequencing, and reported the discovery, phylogenetic classification, genomic structure of a new member of the family *Tombusviridae* that we provisionally named human tombus-like virus (hTLV), and described the epidemiological and clinical characteristics of the patients with hTLV infections.

Methods

Study design and viral genome assembly

In a multicenter surveillance study in Jinan, Shandong Province, China, meta-transcriptome sequencing was conducted on the patients with acute respiratory tract infection after the outbreak of COVID-19 in 2022 – 2023.

All raw sequencing reads were aligned to the human reference genome (GRCh38) using hisat2 (version 2.2.1) to remove human related-reads [17]. According to the results of *de novo* assembly, we search for possible virus sequences by comparing against National Center for Biotechnology Information (NCBI) non-redundant nucleotide database through the blastn program (version 2.11.0) and against the non-redundant protein database through Diamond blastx (version 2.0.9) [18], with an e-value threshold of 1e-5 to retain high sensitivity at a low false-positive rate. To eliminate mis-assembly and false-positive virus, quality-controlled reads were mapped to the assembled complete genomes using Bowtie2 (version 2.3.5.1) [19]. To mitigate potential sequencing-derived contamination, we estimated the read ratio between the highest-abundance library and other lower-abundance libraries on the same sequencing chip. If the ratio was below the index-hopping threshold for the sequencing platform, the reads from the lower-abundance libraries were considered as potential cross-contamination during library preparation and were excluded from further analysis [20]. To mitigate the risk of index-hopping, we set a threshold of 0.1% as the maximum allowable rate based on previous studies [21,22].

RT–PCR amplification and verification of viral genome

To verify the viral genome sequences detected in meta-transcriptome sequencing output, we conducted RT–PCR assays for amplifying the viral sequences using a serial of overlap primers designed according to meta-transcriptome-assembly genome (MAG) sequences of hTLVs (Table S1). All available samples either positive or negative for hTLVs were amplified for the virus sequences and confirmed by Sanger sequencing. At the same time, to eliminate potential contamination, all buffers, reagents, and plasticware used for nucleic acid extraction, and amplification were subjected to UV irradiation. Furthermore, a blank control, sterile enzyme-free water was performed to assess the possibility of cross-contamination and potential reagent contamination. The Sanger sequencing results were subsequently used for mutual validation with the sequences assembled from meta-transcriptome sequencing output.

Retrospective screening of meta-transcriptome data

To explore whether hTLV is also present in other human populations, we conducted retrospective screening of meta-transcriptome sequencing data

from the study population in Beijing. In a multicenter surveillance study of the human seasonal coronavirus (HCoV) infections based on the Respiratory Pathogen Surveillance System (RPSS) at Beijing Center for Disease Prevention and Control from January 2016 to December 2019 at Beijing Metropolis, China [5], meta-transcriptome sequencing was conducted on a total of 321 respiratory samples (sputum, oropharyngeal swab, and BALF) positive for HCoV tested by multiplex real-time fluorescent reverse transcription polymerase chain reaction (RT–PCR) kit (Jiangsu Macro & Micro Test, Nantong, China), and 94 HCoV-negative samples. The quality-controlled reads were mapped to the assembled complete genomes using Bowtie2 (version 2.3.5.1) [19], and *de novo* assembly were further conducted.

Phylogenetic analysis

All available reference genomes of order *Tolivirales* as of Oct 20, 2022 were downloaded from the NCBI/GenBank database. All assembled viral genomes were aligned with global available sequences using MAFFT (version 7.487) [23]. We used function trimAl (version v1.4.rev15) [24] to remove ambiguously aligned regions. Phylogenetic trees were constructed using the maximum likelihood (ML) method with 1,000 bootstrap replicates in IQ-TREE (version 2.1.4) [25], employing the best-fit model SYM + I + R8 for complete genomes and model Q.pfam + F + R8 for RdRps according to Bayesian Information Criterion (BIC). We also constructed the phylogenetic trees using the software MrBayes (version 3.2.7) [26] for comparison. The phylogenetic trees were rooted at midpoint and visualized using FigTree software (version v1.4.4) (http://beast.community/figtree). We further constructed the phylogenetic analyzes based on RdRp domain using ML method and MrBayes.

Viral genomes annotation and mutation analysis

Open reading frames (ORF) of all viral genomes were defined, and annotated through a BLAST-like algorithm, by alignment with the full length of each annotation using Geneious Prime 2024.0.2 (https://www.geneious.com). The conserved domains were identified using RSP-blast (v2.6.0) compared with CDD database (v3.20) and verified on the HAMMER web server (https://www.ebi.ac.uk/Tools/hmmer/). The conserved motifs of viral RdRp were searched by using CDD Tools (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) [27]. The codon gene sequences of the RdRp protein were translated and extracted in Geneious Prime. Multiple sequence alignment of these codon gene sequences was accomplished using MAFFT (version 7.487).

Enzyme-linked immunosorbent assay

Serum samples were collected at the time of admission as acute serum samples, and serum samples taken 14 days after admission or before discharge are collected as convalescent serum samples. The enzyme-linked immunosorbent assay (ELISA) was developed to detect specific immunoglobulins G (IgG) and immunoglobulin M (IgM) against hTLV in the sera using the recombinant hTLV capsid protein (0.1 μg/mL) as a coating antigen. The sera of patients with hTLV infection in acute phase and convalescent phase were tested for the assessment of hTLV seroprevalence. The cut-off value was determined by comparing the mean ± 2 standard deviation (SD) of IgG/IgM ELISA by using sera from healthy young adults. Any sample identified as positive at the screening 1:80 dilution was tittered out. The titer of specific antibodies was the reciprocal of the highest doubling dilution in which has a reading about the cut-off.

Results

Identification of hTLV in Jinan city, China

In response to COVID-19 outbreak, we conducted an etiological surveillance study at two hospital in Jinan City, China from January 2022 to June 2023, and used New Coronavirus (2019-nCoV) Nucleic Acid Detection Kit (Wuhan Easy Diagnosis Biomedicine Co.,ltd.) by reverse transcription polymerase chain reaction (RT–PCR) to test for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). We performed meta-transcriptome sequencing on respiratory samples including oropharyngeal swab, sputum, and BALF, from 107 patients with acute respiratory infections who tested negative for SARS-CoV-2. From the meta-transcriptomic sequencing output, two complete meta-transcriptome-assembly genome (MAG) sequences and contigs of hTLVs were obtained in the samples from three patients with LRIs by using *de novo* assembly. We then designed specific primers according to assembled virus sequences (Figure S1) and performed RT–PCR followed by Sanger sequencing to verify accuracy of viral genomes identified using transcriptome-assembly (Figure S2). Two complete genomes of hTLVs were confirmed, and the other one partial sequence was obtained by specific RT–PCR and Sanger sequencing. Among the remaining 104 samples, five positive samples were detected by RT–PCR assay and confirmed by Sanger sequencing. Among them, we assembled one complete genome sequence. As a result, a total of three hTLV full genome sequences were obtained and deposited to GenBank with the accession number PQ434694 – PQ434696.

Retrospective identification of hTLV in meta-transcriptome sequencing data from Beijing, China

To further investigate the infection of hTLV in patients with LRIs, we conducted a retrospective screening of meta-transcriptome sequencing for 321 respiratory samples (oropharyngeal swab, sputum, and BALF) positive for HCoVs using multiple real time RT–PCR tests in a multicenter surveillance study on HCoVs based on RPSS at Beijing Metropolis, China [28]. From the meta-transcriptomic sequencing output, ten complete and three nearly complete MAG sequences of hTLVs were obtained in the samples from patients with LRIs by using *de novo* assembly and verified by performing RT–PCR followed by Sanger sequencing. Among them, two complete genome sequences were obtained by specific RT–PCR and Sanger sequencing. As a result, twelve complete and one partial (with 969-bp absence at 3`-end) genomes of hTLVs were validated from MAG sequences. Furthermore, two additional complete hTLV genome sequences were identified by RT–PCR amplification and Sanger sequencing in sputum samples, one of which from HCoV-positive patient and the other from one of 94 HCoV-negative patients. The 15 hTLV genome sequences were deposited to GenBank under the accession number OQ147328 – OQ147340 and OQ077518 – OQ077519 (Table S1). Totally, 23 patients were confirmed to be infected by hTLV, of which 17 nearly complete genomes and one partial genome were available for further analyzes (Figure S2).

Phylogenetic analysis of hTLV

The 18 hTLV genome sequences shared nucleotide (nt) identities of 80.03% – 99.94% with each other (Figure S3), and of only 34.45% – 36.24% with the most closely related viral genome sequence of an invertebrate-associated Jingmen tombus-like virus 2 (GenBank accession no. NC_033696) [6]. The phylogenetic analysis based on the whole genome sequences revealed that these viral genomes fell in the family *Tombusviridae* of order *Tolivirales* (Figure 1A). Subsequently, we constructed the phylogenetic tree based on the amino acid (aa) of the most conserved RdRp protein from the 18 hTLVs and all viral RdRp in the order *Tolivirales*. The phylogenetic trees base on RdRp protein generated by either maximum likelihood or MrBayes analysis showed consistent topology with that constructed using completed genomic nt sequences (Figure 1B). The phylogenies revealed that the viruses identified in this study were distinct from all known viruses, and clustered with a Jingmen tombus-like virus 2 but in a separate clade, which represented a new member between families *Tombusviridae* and *Sinhaliviridae*. The newly discovered hTLV had the RdRp of aa identity of 91.47% – 100% with each other (Figure S4), but of only 26.13% – 26.85% and 19.82% – 20.89% with available viral sequences in the closest families *Tombusviridae* and *Sinhaliviridae*, respectively. At present, there is no species or even genus demarcation of the family *Tombusviridae* according to the current scheme of virus classification by the International Committee on Taxonomy of Viruses (ICTV) [29]. According to the commonly accepted species demarcation threshold (80% identity of the genome-wide nucleotide sequence, and 90% aa identity of RdRp protein) [6,30], the virus identified from patients with LRIs in this study should be a new member related to the family *Tombusviridae*, considering the so low identity of the genome as well as the RdRp protein sequences with other related viruses and the distance in the phylogenetic trees. As mentioned above, we provisionally named it human tombus-like viruses (hTLV).

Genome structure of hTLV

Genome sequence analyzes showed that the hTLVs had single, positive-strand RNA genomes of 4.7–4.8 kb in size, which was slightly smaller than that of the Lake Sinai Virus TO (GenBank accession no. NC_035116) or Towan virus (GenBank accession no. KX160090). The hTLV genome contained four putative open reading frames (ORF1–4) deduced directly from the nucleotide sequence, flanked by relatively short 5' and 3' nontranslated RNA (NTR) segments of 114 and 747 nucleotides (Figure 2). The coding region was about 3953 nt long, with the ORF1–4 composed of 410, 540, 306–307, and 111 amino acids, respectively. Analysis and comparisons with gene expressions of the members in the family *Tombusviridae* suggested that the hypothetical protein 2 CDS resulting from ORF2 might be translated from different ORFs. The larger ORF2 arose from the internal initiation codon at end of ORF1 with a 151-nt overlap, similar to that of the Jingmen tombus-like virus 2 or Lake Sinai Virus. The hTLV might also use – 1 frameshifting to generate a smaller ORF2 of 491 aa (Figure 2). Both different hypothetical protein 2 CDS contained the highly conserved motif that encodes putative viral RdRp of 213 aa in the family *Tombusviridae*. In addition, the ORF3 encoded a putative coat protein shared with aa similarity of only 29.57–30.49% to that of the distantly related Towan virus (GenBank accession no. AOG30802). We further compared the sequences of ORF3 and ORF4 with typical movement proteins (MP) from the family *Tombusviridae* [31] and found that both ORF3 and ORF4 of hTLV differed from known MP, providing the evidence that hTLV should not be a plant-associated virus (Figure S5).

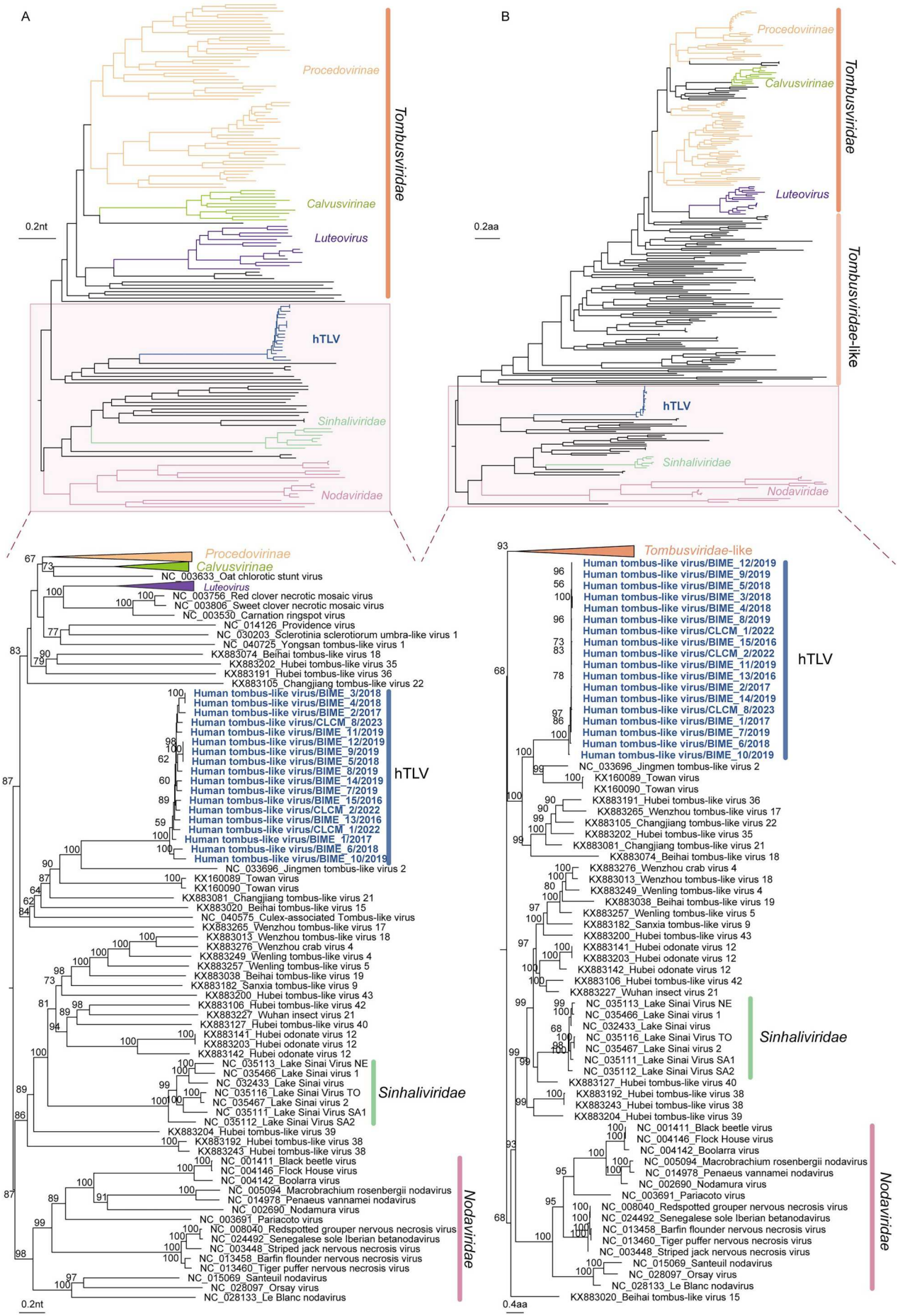

Figure 1. Phylogenetic tree of hTLV using maximum likelihood method. (A) Phylogenetic trees of hTLV based on the complete genome. (B) Phylogenetic trees of hTLV based on the RdRP protein. The viral genome sequences newly obtained in this study are marked in dark bule. Bootstrap values ≥60% are present on the branch nodes.

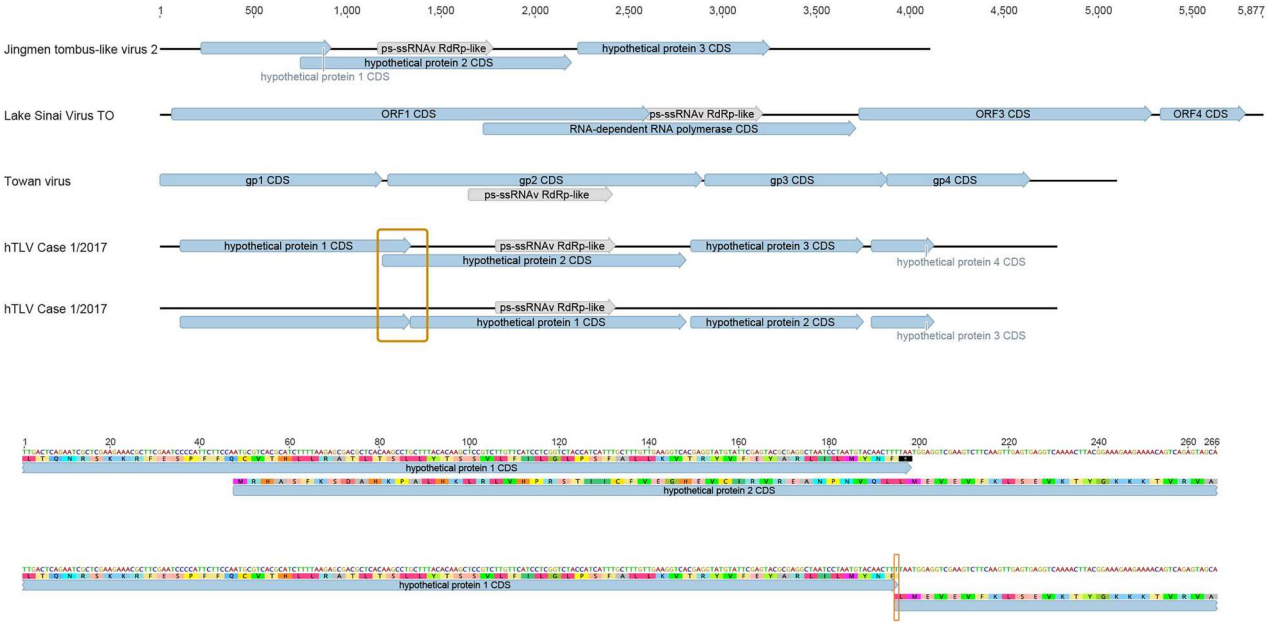

Figure 2. Structure of hTLV compared with related viruses. The putative read-through stop codon was in the orange box.

Variation in core conserved domain of rdrp

We used RSP-blast to compare with CDD database, and found that hTLV ORF2 encoded a putative RdRp of 639 nt distantly related to those of species in the family *Tombusviridae* with aa similarity of 43.20–44.66%. Sequence comparison of the predicted aa sequences of hTLV RdRps indicated that they shared significant sequence similarity characterized by three conserved motif A, B, and C containing 15, 24, and 15 aa, respectively (Figure 3). All motifs located in the palm subdomains of RdRp, which might be important for nucleotide incorporation fidelity of RdRp in the Sequence Cluster cd23174 (https://www.ncbi.nlm.nih.gov/Structure/cdd/). Notably, motif A contained a highly conserved aspartic acid residue (Asp_{641}) in the RdRp of all hTLVs. The unifying feature of the family *Tombusviridae* showed that each member species possessed a conserved polymerase containing the canonical Gly_{376}-Asp_{377}-Asp_{378} (GDD) catalytic residues in motif C. The GDD catalytic residues has been confirmed to be important for the genome replication of the viruses in the family *Tombusviridae* via mutational analysis [32].

The patients infected with hTLV

As mentioned above, sixteen of the 23 patients infected with hTLV were detected from meta-transcriptome sequencing output, and subsequently confirmed by RT–PCR amplification followed by Sanger sequencing. The other seven were identified by direct RT–PCR amplification and Sanger sequencing (Table 1). The hTLV genome sequences were mostly obtained from sputum (12), oropharyngeal swab (10), and BALF (1) samples. The interval from the onset of illness to sample collection ranged from 1 d to 29 days with a median of seven days. Out of the 23 patients, one occurred in 2023, seven in 2022, seven in 2019, four in 2018, two in 2017, and two in 2016. Median age of the 23 patients was 63 years (range 52–83 years), and 17 (73.91%) were male (Table 1). Through RT–PCR testing for various respiratory viruses and bacteria in all samples (Table S2), we found that six patients (CLCM-1, CLCM-2, CLCM-4, and CLCM-6, CLCM-7, and BIME-15) were solely infected with hTLV, with no other respiratory pathogens detected. Additionally, patients CLCM-3 and CLCM-5 had co-infections with H1N1 and a patient (CLCM-8) was infected with Influenza B virus, respectively. Among them, patient CLCM-3 was also infected with *Streptococcus pneumoniae*. Fourteen patients (BIME-1 to BIME-14) were co-infected with various human coronaviruses, including HCoV-229E, HCoV-OC43, and HCoV-NL63.

The main clinical manifestations of the 23 patients were fever (temperature >37.0°C; 100%), cough (78.26%), expectoration (86.96%) and dyspnea (34.78%). The respiratory rates ranged from 16 to 35 per minute with a media of 20, and seven patients had the abnormal respiratory rate faster than 20 per minute. Their heart rates ranged from 61 to 135 with a media of 94 per minute, and six patients had faster heart rates (>100 per minute). Eleven and eight patients developed lung rales or respiratory failure, ten of whom were transferred to the intensive care unit (ICU). Fourteen (60.87%) of 23 patients were recorded to have at least one underline chronic illnesses, such as malignancy, coronary heart disease, diabetes mellitus, chronic obstructive pulmonary disease, stroke, cirrhosis, bronchitis, and tuberculosis. Out of 21 patients examined by chest computed

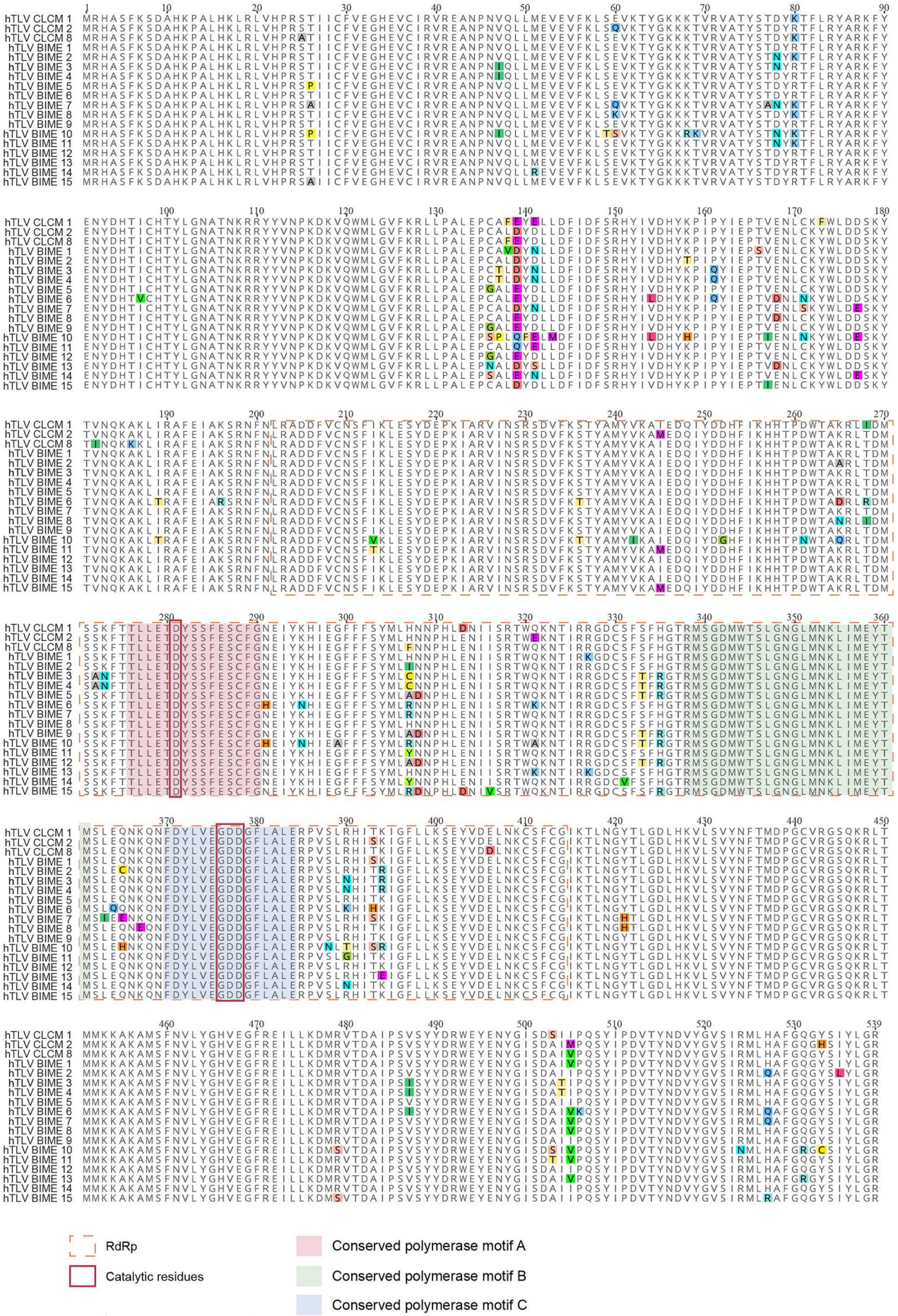

Figure 3. Amino acid mutations in hTLV RNA-dependent RNA polymerase proteins in this study. The RdRp domain was shown in light red box, Catalytic residues were shown in the dark red box, and the Conserved polymerase motifs A, B, and C were in pink, green, and blue shades, respectively.

Table 1. The characteristics of 23 patients infected with hTLV.

Case No.	CLCM-1	CLCM-2	CLCM-3	CLCM-4	CLCM-5	CLCM-6	CLCM-7	CLCM-8	BIME-1	BIME-2	BIME-3	BIME-4	BIME-5	BIME-6	BIME-7	BIME-8	BIME-9	BIME-10	BIME-11	BIME-12	BIME-13	BIME-14	BIME-15
Identification method	MTA	MTA	MTA	RT-PCR	RT-PCR	RT-PCR	RT-PCR	RT-PCR	MTA	MTA	MTA	MTA	MTA	MTA	MTA	MTA	MTA	MTA	MTA	MTA	MTA	RT-PCR	RT-PCR
Sample	OS	OS	OS	OS	OS	OS	OS	OS	Sputum	OS	Sputum	Sputum	OS	Sputum	Sputum	Sputum	BALF	Sputum	Sputum	Sputum	Sputum	Sputum	Sputum
IOS (days)	6	2	3	10	29	29	4	10	12	NA	4	10	1	9	NA	9	3	15	4	NA	2	3	8
Year of onset	2022	2022	2022	2022	2022	2022	2022	2023	2017	2017	2018	2018	2018	2018	2019	2019	2019	2019	2019	2019	2016	2019	2016
Age (year)	57	78	65	68	55	64	64	81	64	53	83	61	60	63	67	63	53	52	59	60	79	63	68
Gender	Male	Male	Female	Male	Male	Female	Male	Male	Male	Male	Male	Male	Female	Male	Male	Male	Male	Female	Male	Female	Male	Female	Male
Clinical manifestations																							
Temperature (°C)	41	38.3	37.3	37.3	38.5	40.4	39.5	38.6	37.5	40	37	38.5	38	37.2	39	38	39	37.3	38	NA	37.8	38.4	38.1
Cough	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NA	Yes	Yes	Yes
Expectoration	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dyspnea	No	No	No	Yes	No	No	No	No	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes	No	Yes	NA	No	No	No
Respiratory rate	24	18	16	18	20	16	20	20	23	16	35	27	20	20	18	22	21	20	20	NA	28	20	20
Heart rate	108	61	79	64	80	96	100	64	82	115	135	72	89	96	105	92	106	100	109	NA	98	80	87
Lung rales	Yes	Yes	No	Yes	Yes	No	Yes	No	Yes	No	Yes	No	No	No	No	Yes	Yes	No	No	NA	Yes	No	Yes
Respiratory failure	NA	NA	No	NA	NA	NA	NA	NA	Yes	No	Yes	Yes	No	Yes	No	No	Yes	No	Yes	NA	Yes	No	Yes
ICU	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	No	No	No	No	No	Yes	No	No	NA	Yes	No	No
Underline illnesses	HD, HTN	HD, HTN, DM, TB	CHD	HTN	None	HTN, stroke	DM	HTN, COPD	Malignancy	None	CHD	None	None	TB	Bronchitis	TB, HD	Cirrhosis	None	Stroke, malignancy	None	None	None	None
Laboratory tests																							
CT or X-ray	+	+	-	+	+	+	+	+	+	+	NA	+	-	+	+	+	+	+	-	NA	+	-	+
WBC	16.21	6.16	7.33	6.18	8.75	6.95	9.79	8.09	5.3	29.5	7.2	15.9	9.88	6.25	14.52	6.5	15.73	4.1	5.98	NA	15.3	6.1	5.4
RBC	3.6	3.6	4.36	4.58	4.05	4.03	4.66	3.88	4.5	4.4	3.8	4.6	4.68	3.82	3.85	4.58	4.03	2.97	5.4	NA	1.9	4.25	2.7
ALT/AST	+	-	-	-	-	NA	+	-	-	NA	NA	+	NA	-	-	-	+	-	-	NA	-	-	-
PaO_2 (mmHg)	72.5	87.1	NA	NA	NA	86.8	NA		68	NA	113	63	NA	82	72	46	41	NA	77	NA	82	NA	NA
SaO_2 (%)	94.9	NA	NA	NA	NA	NA	NA		92.5	NA	99	94	NA	98	95.5	82.2	NA	NA	97	NA	98	NA	NA
$PaCO_2$ (mmHg)	34.3	29.9	NA	NA	NA	35.5	NA		42	NA	39	32	NA	37	27	64	37	NA	45	NA	42	NA	NA

MTA, Meta-transcriptome assembly; OS, Oropharyngeal swab; BALF, Bronchoalveolar lavage fluid; IOS, Interval between onset and sample collection; HCoV, Human seasonal coronavirus; HTN, Hypertension; CHD, Coronary heart disease; TB, Tuberculosis; HD, Heart disease; DM, Diabetes mellitus; CT or X-ray, Chest computed tomography or X-ray chest radiography; WBC, White blood cell count; RBC, Red blood cell count; ALT/AST, Alanine aminotransferase or aspartate aminotransferase; PaO_2, Alveolar oxygen partial pressure; SaO_2, Arterial oxyhemoglobin saturation; $PaCO_2$, Pressure of carbon dioxide.

tomography (CT) or X-ray radiography, 17 (80.95%) displayed abnormal results. Eight (36.36%) of 22 patients had abnormally higher white blood cell (WBC) counts, indicating possible coinfections with bacteria. The red blood cell (RBC) count was lower in twelve patients. The hepatic aminotransferase (alanine aminotransferase or aspartate aminotransferase) values were abnormally elevated in four patients. Among the 12 patients who underwent blood gas analysis, case BIME-3 treated in ICU had the highest alveolar oxygen partial pressure (PaO_2), possibly due to oxygen therapy. Seven patients had abnormally lower PaO_2, and four of whom showed an arterial oxyhemoglobin saturation of below 95%. The pressure of carbon dioxide ($PaCO_2$) was in the normal range of 35–45 mmHg among seven patients. Four patients had the $PaCO_2$ below 35 mmHg, and one patient (case BIME-8) had a higher $PaCO_2$ of 64 mmHg (Table 1). A comparative analysis of clinical symptoms between six patients with hTLV monoinfection and 17 co-infected patients based on the clinical symptoms available revealed no statistically differences in symptoms (all $P > 0.05$) (Table S3).

Serological results of hTLV infections in the patients

To study the prevalence of newly identified hTLV virus in the human population, we tested sera from PCR-confirm patients by ELISA with recombinant hTLV capsid as a coating antigen. From the analysis of mean ± 2SD of the 1:80 dilution, cut-off value of 0.658 was selected.

The paired serum sample was obtained from one patient during the acute and convalescent phase of infection. Acute serum samples were collected on the 6^{th} day after the onset of illness, and convalescent serum samples were collected on the 42^{nd} day after the acute phase samples were collected. The fourfold or greater rise in IgG titers between acute and convalescent-phase samples was determined in this patient. A single acute-phase sample obtained on the 29^{th} day after the onset from the patient, tested positive for IgG and IgM.

Discussion

Here we report the characterization of hTLV, a new member in the family *Tombusviridae*, identified via meta-transcriptome sequencing and confirmed by RT–PCR amplification followed by Sanger sequencing from 23 patients with LRIs in Shandong Province and Beijing, China. All the cases were older than 52 years of age with a media of 63 years, and required hospitalization. Both phylogenetic analysis and genomic structure support the argument that hTLV should be a new member of the family *Tombusviridae*, distinct from all known tombusvirus species. The genome sequences of hTLVs only have approximately 34.45% – 36.24% nucleotide identities to the most closely related Jingmen tombus-like virus 2. In the phylogenetic trees, the hTLVs of this study formed a separate clade (Figure 1). Typically, tombusviruses possess 3–5 kb sized genomes with 3–5 ORFs but lack a large open reading frame that spans the entire genome [33]. The hTLVs have genomes of 4.7–4.8 kb in size with four ORFs. Sequence alignment of the ORF3 and ORF4 of hTLV showed no significant similarity to the movement protein (MP) in several typical viruses of the family *Tombusviridae* [31]. It is suggested that hTLV may represent a novel viral species in animals, sharing some genomic structural features with the family *Tombusviridae* members. Tombusviruses also usually contain a read-through stop codon that can be suppressed or encode a frame-shift mechanism for translation of the RdRp. The read-through stop codon has not been recognized in the hTLV genome organization, and requires further investigations on encoded protein size to clarify this phenomenon. The conserved domain found in other tombusviruses [34] is also identified in hTLVs.

Following the discovery of the novel virus, an important question is whether hTLV causes the infections of the patients with LRIs. To answer this question, we developed an ELISA to test specific antibodies against hTLV. Unfortunately, only one paired serum from acute and convalescent phase was available for a LRI patient from whom hTLV was detected, whether hTLV causes LRIs cannot be directly concluded. In addition, 14 of 18 hTLV genome sequences were initially detected from the meta-transcriptome sequencing output of respiratory samples collected from patients with HCoV infections, we cannot distinguish whether LRIs are caused by hTLV or by HCoV. However, six hTLV-infected patients (cases BIME-15, CLCM-1, CLCM-2, CLCM-4, CLCM-6, and CLCM-7) with respiratory failure were identified from the HCoV-negative group, and no other respiratory pathogen was detected, suggesting that hTLV should be infective and pathogenic to humans and a causal relationship between hTLV and LRI is likely. Our serological survey has provided further evidence that hTLV infection can cause a fourfold increase in IgG titers in an ILI patient.

Viruses in the family *Tombusviridae* have been previously known to infect plants with limits on relatively narrow natural host range [35]. However, plant pathogens viruses in the family *Tombusviridae* have been reported to cause human diseases [36]. The plant-derived pepper mild mottle virus, the member of *Tobamovirus*, has been reported to be associated with specific immune responses, and various clinical manifestations in humans [37]. Tobacco mosaic virus, another member in the family *Tombusviridae*,

has been identified in saliva of smokers [38], and subsequently proved to cause human infection by detection of specific antibodies against the virus [39]. Recently, a novel tombus-like virus, Statovirus, was found in human, macaque, mouse, and cow gastrointestinal tract samples [25], and was identified by metagenomic sequencing of human nasal – throat swabs from livestock farm workers experiencing acute respiratory diseasev [40]. As hTLVs identified in this study seem genetically most close to Jingmen tombus-like virus 2 from invertebrates and Towan viruses from orangutan, we speculate that the human virus might have originated from verity of animal hosts, which need further investigation.

As advances in metagenomic sequencing and analysis, whole-genome sequencing of either known or novel pathogens can now be done directly from clinical samples, helping to accurately elucidate the viral infection and characterization [41]. The discovery of hTLV genomes indicate that metagenomic sequencing is valuable approach to identifying various viruses including novel one from clinical samples, which can facilitate ecological and differential diagnosis of infections with known or unknown pathogens. Two more whole genomes of hTLVs were obtained by RT–PCR amplifications followed by Sager sequencing in two samples, from which whole genomes had not been assembled through metagenomics analysis. This finding suggests that RT–PCR currently remains more sensitive for the validation of the virus. However, this approach needs a priori knowledge of known pathogens and an extensive set of primers for amplification. The advantage of metagenomic approach over RT-PCRs lies in its capacity of identifying and assembling all known and novel viral genomes simultaneously, because metagenomic approaches do not target particular pathogens [42]. This untargeted manner makes it a promising application prospect of equally detecting expected pathogens as well as novel pathogens, such as SARS-CoV-2 [43]. The detection of hTLV in both pre-pandemic (Beijing 2016–2019) and post-pandemic (Jinan 2022–2023) surveillances, across patients with varying clinical presentations (Table 1), suggests this virus may have circulated but undetected in China.

The clinical manifestations of hTLV infection including fever, productive cough, and respiratory failure, share significant overlap with conventional respiratory viral infections caused by influenza viruses or coronaviruses (Table 1). It poses diagnostic challenges and underscores the need for differential characterization. Future nationwide surveillance incorporating seroepidemiology and metagenomic screening across diverse populations is needed to fully characterize the distribution and transmission dynamics of hTLV. Meanwhile, longitudinal cohort studies should establish whether hTLV exhibits seasonal co-circulation patterns with established pathogens, due to the differentiation is critical for developing targeted therapeutic strategies.

The discovery of a novel virus, hTLV, in patients with LRIs highlights the continuous emergence of new respiratory pathogens in humans, and warrants further investigations to characterize its pathogenicity and clinical importance. Metagenomic analysis is a promising and applicable approach to detection of respiratory viral pathogens in clinical samples independent of known viral genome sequence.

Ethical approval statement

The ethics approval for the protocol of this study was obtained from the Ethics Committee of Beijing Center for Disease Prevention and Control and Shandong University (LL202305002). Before enrollment, the purpose, procedures, potential healthy impact and benefits of this study were explained carefully to participants or their care providers, and written consents were obtained.

Disclosure statement

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

Funding

This study was supported by grants from the CAMS Innovation Fund for Medical Sciences [2019-I2M-5-049], the National Key Research and Development Program of China [2023YFC2305901], Open Project Program of the State Key Laboratory of Pathogen and Biosecurity [SKLPB2133], Young Talent of Lifting engineering for Science and Technology in Shandong, China [SDAST2024QTA079], Science and Technology Support Plan for Youth Innovation of Colleges and Universities of Shandong Province of China [2022KJ016], the Beijing Municipal Health Commission's Funds for the High-qualified Public Health Professionals Development Project [Leading Professionals-01-10], Cheeloo Young Scholar Program of Shandong University, and the Postdoctoral Fellowship Program and China Postdoctoral Science Foundation [BX20240215].

Data availability statement

All viral genomes obtained in this study have been deposited in NCBI GenBank under Accession numbers PQ434694 – PQ434696, OQ147328 – OQ147340 and OQ077518 – OQ077519 as listed in Table S1.

References

[1] GBD 2016 Lower Respiratory Infections Collaborators. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990-2016: a systematic analysis for the global burden of

disease study 2016. Lancet Infect Dis. 2018;18:1191–1210. doi:10.1016/S1473-3099(18)30310-4
[2] GBD 2016 Causes of Death Collaborators. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980-2016: a systematic analysis for the global burden of disease study 2016. Lancet. 2017;390:1151–1210. doi:10.1016/S0140-6736(17)32152-9
[3] GBD 2015 LRI Collaborators. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory tract infections in 195 countries: a systematic analysis for the global burden of disease study 2015. Lancet Infect Dis. 2017;17:1133–1161. doi:10.1016/S1473-3099(17)30396-1
[4] Boni MF, Lemey P, Jiang XW, et al. Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic. Nat Microbiol. 2020;5:1408–1417. doi:10.1038/s41564-020-0771-4
[5] Gong C, Zhang TG, Luo M, et al. Distribution of the atypical pathogens of community-acquired pneumonia to disease severity. J Thorac Dis. 2018;10:5991–6001. doi:10.21037/jtd.2018.10.50
[6] Shi M, Lin XD, Tian JH, et al. Redefining the invertebrate RNA virosphere. Nature. 2016;540:539–543. doi:10.1038/nature20167
[7] Mahar JE, Shi M, Hall RN, et al. Comparative analysis of RNA virome composition in rabbits and associated ectoparasites. J Virol. 2020;94:e02119. doi:10.1128/jvi.02119-19
[8] Hause BM, Nelson E, Christopher-Hennings J. Identification of a novel statovirus in a faecal sample from a calf with enteric disease. J Gen Virol. 2021;102(9). doi:10.1099/jgv.0.001655
[9] Liotti FM, Posteraro B, Mannu F, et al. Development of a multiplex PCR platform for the rapid detection of bacteria, antibiotic resistance, and candida in human blood samples. Front Cell Infect Microbiol. 2019;9:389. doi:10.3389/fcimb.2019.00389
[10] Dabisch-Ruthe M, Vollmer T, Adams O, et al. Comparison of three multiplex PCR assays for the detection of respiratory viral infections: evaluation of xTAG respiratory virus panel fast assay, RespiFinder 19 assay and RespiFinder SMART 22 assay. BMC Infect Dis. 2012;12:163. doi:10.1186/1471-2334-12-163
[11] Rosseel T, Ozhelvaci O, Freimanis G, et al. Evaluation of convenient pretreatment protocols for RNA virus metagenomics in serum and tissue samples. J Virol Methods. 2015;222:72–80. doi:10.1016/j.jviromet.2015.05.010
[12] Lewandowska DW, Zagordi O, Geissberger FD, et al. Optimization and validation of sample preparation for metagenomic sequencing of viruses in clinical samples. Microbiome. 2017;5(1):94. doi:10.1186/s40168-017-0317-z
[13] Tan W, Zhao X, Ma X, et al. A novel coronavirus genome identified in a cluster of pneumonia cases - wuhan, China 2019-2020. People's Republic of China CDC Wkly. 2020;2(4):61-62.
[14] Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med. 2020;382(8):727–733. doi:10.1056/NEJMoa2001017
[15] Jain S, Self WH, Wunderink RG, et al. Community-acquired pneumonia requiring hospitalization among U.S. adults. N Engl J Med. 2015;373(5):415–427. doi:10.1056/NEJMoa1500245
[16] Nair H, Simões EA, Rudan I, et al. Global and regional burden of hospital admissions for severe acute lower respiratory infections in young children in 2010: a systematic analysis. Lancet. 2013;381(9875):1380–1390. doi:10.1016/S0140-6736(12)61901-1
[17] Kim D, Paggi JM, Park C, et al. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019;37(8):907–915. doi:10.1038/s41587-019-0201-4
[18] Buchfink B, Reuter K, Drost HG. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat Methods. 2021;18:366–368. doi:10.1038/s41592-021-01101-x
[19] Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357–359. doi:10.1038/nmeth.1923
[20] Cobbin JC, Charon J, Harvey E, et al. Current challenges to virus discovery by meta-transcriptomics. Curr Opin Virol. 2021;51:48–55. doi:10.1016/j.coviro.2021.09.007
[21] Li Q, Zhao X, Zhang W, et al. Reliable multiplex sequencing with rare index mis-assignment on DNB-based NGS platform. BMC Genomics. 2019;20(1):215. doi:10.1186/s12864-019-5569-5
[22] MacConaill LE, Burns RT, Nag A, et al. Unique, dual-indexed sequencing adapters with UMIs effectively eliminate index cross-talk and significantly improve sensitivity of massively parallel sequencing. BMC Genomics. 2018;19(1):30. doi:10.1186/s12864-017-4428-5
[23] Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–780. doi:10.1093/molbev/mst010
[24] Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. Trimal: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25:1972–1973. doi:10.1093/bioinformatics/btp348
[25] Nguyen LT, Schmidt HA, von Haeseler A, et al. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32:268–274. doi:10.1093/molbev/msu300
[26] Ronquist F, Huelsenbeck JP. Mrbayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 2003;19:1572–1574. doi:10.1093/bioinformatics/btg180
[27] Lu S, Wang J, Chitsaz F, et al. CDD/SPARCLE: the conserved domain database in 2020. Nucleic Acids Res. 2020;48:D265–D268. doi:10.1093/nar/gkz991
[28] Ye RZ, Gong C, Cui XM, et al. Continuous evolution and emerging lineage of seasonal human coronaviruses: a multicenter surveillance and genomic epidemiology study. J Med Virol. 2023;95(6):e28861. doi:10.1002/jmv.28861
[29] Lefkowitz EJ, Dempsey DM, Hendrickson RC, et al. Virus taxonomy: the database of the international committee on taxonomy of viruses (ICTV). Nucleic Acids Res. 2018;46:D708–D717. doi:10.1093/nar/gkx932
[30] Shi W, Shi M, Que TC, et al. Trafficked malayan pangolins contain viral pathogens of humans. Nat Microbiol. 2022;7:1259–1269. doi:10.1038/s41564-022-01181-1

[31] Scholthof HB. Plant virus transport: motions of functional equivalence. Trends Plant Sci. 2005;10:376–382. doi:10.1016/j.tplants.2005.07.002
[32] Molinari P, Marusic C, Lucioli A, et al. Identification of artichoke mottled crinkle virus (AMCV) proteins required for virus replication: complementation of AMCV p33 and p92 replication-defective mutants. J Gen Virol. 1998;79(Pt3):639–647. doi:10.1099/0022-1317-79-3-639
[33] Janowski AB, Krishnamurthy SR, Lim ES, et al. Statoviruses, a novel taxon of RNA viruses present in the gastrointestinal tracts of diverse mammals. Virology. 2017;504:36–44. doi:10.1016/j.virol.2017.01.010
[34] Heinlein M. Plant virus replication and movement. Virology. 2015;479–480:657–71. doi:10.1016/j.virol.2015.01.025
[35] King AMQ, Adams MJ, Carstens EB, et al. Family - tombusviridae. San Diego: Elsevier; 2012:1111–38. doi:10.1016/B978-0-12-384684-6.00096-3
[36] Kim JS, Yoon SJ, Park YJ, et al. Crossing the kingdom border: human diseases caused by plant pathogens. Environ Microbiol. 2020;22:2485–2495. doi:10.1111/1462-2920.15028
[37] Colson P, Richet H, Desnues C, et al. Pepper mild mottle virus, a plant virus associated with specific immune responses, fever, abdominal pains, and pruritus in humans. PLoS One. 2010;5:e10041. doi:10.1371/journal.pone.0010041
[38] Balique F, Colson P, Raoult D. Tobacco mosaic virus in cigarettes and saliva of smokers. J Clin Virol. 2012;55:374–376. doi:10.1016/j.jcv.2012.08.012
[39] Liu R, Vaishnav RA, Roberts AM, et al. Humans have antibodies against a plant virus: evidence from tobacco mosaic virus. PLoS One. 2013;8:e60621. doi:10.1371/journal.pone.0060621
[40] Thi Kha Tu N, Thi Thu Hong N, Thi Han Ny N, et al. The virome of acute respiratory diseases in individuals at risk of zoonotic infections. Viruses. 2020;12:960. doi:10.3390/v12090960
[41] Ladner JT, Grubaugh ND, Pybus OG, et al. Precision epidemiology for infectious disease control. Nat Med. 2019;25:206–211. doi:10.1038/s41591-019-0345-2
[42] Chiu CY, Miller SA. Clinical metagenomics. Nat Rev Genet. 2019;20:341–355. doi:10.1038/s41576-019-0113-7
[43] Wu F, Zhao S, Yu B, et al. A new coronavirus associated with human respiratory disease in China. Nature. 2020;579:265–269. doi:10.1038/s41586-020-2008-3