

The risk of human- and mammal-infecting tick-borne viruses in northwest China and adjacent countries

Received: 16 December 2024

Accepted: 18 November 2025

Published online: 02 December 2025

 Check for updates

Hong Zhou ^{1,2,3,15} ✉, Hengcong Liu ^{4,15}, Yun-Xiao Wang ^{4,15}, Shanshan Zhao ^{5,6,15}, Moujian Guo ⁷, Mingxue Cui ^{2,3}, Jianjun Chen ⁸, Juan Li ^{1,2,3}, Suwen Wang ^{5,6}, Xudong Zhang ⁹, Jiaqi Qin ^{1,2}, Ziman Lv ^{5,6}, Zhenghai Ma ¹⁰, Ruiling Zhang ¹, Bin Zhang ¹¹, Wen Zheng ⁴, Jingkai Ji ¹², Tianze Wang ⁴, Edward C. Holmes ¹³, Yuan-Zhi Wang ^{5,6} ✉ & Weifeng Shi ^{4,7,14} ✉

The prevalence, evolution, and disease risk of pathogenic tick-borne viruses (TBVs) remains poorly understood in northwest China and adjacent countries, which are endemic for several TBVs of public health importance. Herein, we perform meta-transcriptomic sequencing of >9600 ticks collected across this vast geographic area to identify 92 RNA viruses and assemble 1567 viral genomes from 28 different tick species, including ten human- and mammal-infecting TBVs. Tacheng tick virus 1 (TcTV-1), Tacheng tick virus 2 (TcTV-2), Tamdy virus (TAMV), and Crimean-Congo hemorrhagic fever virus (CCHFV) are the most common human-infecting TBVs detected. We also report several pathogenic TBVs not previously identified in China (Burana virus and Bhanja virus), Kazakhstan (TcTV-1), and northwest China (Wad Medani virus and Alongshan virus). We predict a significant increase in the number of high-risk regions for the four major tick vectors of pathogenic TBVs, and the distribution of TAMV, TcTV-1, TcTV-2, and CCHFV, with 65.4% of the counties in this region identified as high risk for CCHFV. Therefore, the real disease burden caused by TBVs in northwest China and adjacent countries may be more underestimated than appreciated, and we call for strengthened field surveys and epidemiological surveillance in this vast region.

Ticks are second only to mosquitoes as vectors of human infectious diseases^{1,2}. There are more than 900 recognised species of ticks worldwide³ that transmit over 100 viral, bacterial and parasite pathogens to humans and other mammals⁴, including tick-borne viruses (TBVs) such as Crimean-Congo haemorrhagic fever virus (CCHFV)⁵, Severe Fever with Thrombocytopenia Syndrome Virus (SFTSV)⁶ and Alongshan virus (ALSV)⁷. Largely due to the increased threat posed by arthropod-borne viral diseases^{8,9} and viromic surveys facilitated by meta-transcriptomic sequencing, our knowledge of the genetic diversity of TBVs has significantly increased in the past two decades^{10–13}.

In total, 124 tick species and over 100 tick-borne agents have been described in China⁴. Most regions in northwest China have a high diversity of tick species, including the Xinjiang Uygur Autonomous Region (XUAR), Gansu province, Qinghai province, the Ningxia Hui Autonomous Region and the Inner Mongolia Autonomous Region (IMAR). Indeed, more than one-third of the tick species described in China can be found in the XUAR¹⁴. Importantly, several human- and mammal-infecting TBVs have also been identified in northwest China, and related human cases have been documented, including those associated with the well-studied CCHFV¹⁵, as well as the

A full list of affiliations appears at the end of the paper. ✉ e-mail: zhouh@sdfmu.edu.cn; wangyuanzhi621@126.com; shiwf@ioz.ac.cn

newly-discovered Tacheng tick virus 1 (TcTV-1)¹⁶ and Tacheng tick virus 2 (TcTV-2)¹⁷. In addition, Asian countries adjacent to northwest China have also faced a serious public health threat from TBVs. Aside from the epidemics caused by CCHFV in Pakistan, tick-borne encephalitis has been documented in both Kyrgyzstan and Mongolia. Also of note are the zoonotic TBVs Burana virus (BURV), Dhori virus (DHOV) and Tamdy virus (TAMV) that have been identified in ticks in Kyrgyzstan^{18,19}, and Uzbekistan²⁰ and confirmed to be infectious to humans or mammals^{21–25}. Considering the high diversity of both tick species and TBVs and the paucity of systematic sampling, as well as the large population engaged in livestock husbandry in northwest China and adjacent countries, a fuller understanding of the hidden diversity and distribution of ticks and TBVs, particularly to delineate the disease risk and burden of zoonotic TBVs in this vast area, is clearly warranted.

Despite these public health risks, our knowledge of the diversity, geographical distribution, and transmission risk of TBVs in northwest China and adjacent countries is limited. Herein, we collected >100,000 ticks and performed meta-transcriptomic sequencing of >9600 ticks to characterise their viromes, with the specific aim of documenting the presence of potentially pathogenic TBVs. We also performed a comprehensive literature search and built a database of tick species and TBVs in this region, then used this information to perform ecological modelling to map the presence of major tick species and four important TBVs.

Results

Distribution of the ticks and tick species diversity

We collected 100,406 ticks from 53 counties in XUAR ($n=85,960$), two counties in IMAR ($n=1602$), two counties in Qinghai ($n=475$), three counties in Gansu ($n=55$), China and 22 counties in Kazakhstan ($n=12,314$), from 2012 to 2024 (Fig. 1). Based on the morphological characterisations and the *COI* gene comparison, the ticks were classified into 28 tick species from 7 genera: 7 species in the genus *Dermacentor* ($n=50,120$, 49.92%), 5 species in *Hyalomma* ($n=37,666$, 37.52%), 3 species in *Rhipicephalus* ($n=7289$, 7.26%), 6 species in *Haemaphysalis* ($n=2876$, 2.86%), 2 species in *Argas* ($n=2153$, 2.14%), 4 species in *Ixodes* ($n=205$, 0.20%) and 1 species in *Ornithodoros* ($n=97$, 0.10%) (Fig. 1, Supplementary Data 1 and 2). A total of 55,345 blood-feeding ticks from the genera *Dermacentor*, *Haemaphysalis*, *Hyalomma*, *Rhipicephalus* and *Argas* were collected from domestic animals: sheep ($n=26,965$), cattle ($n=17,516$), camels ($n=5266$), horses ($n=3994$), dogs and cats ($n=1100$) and poultry ($n=504$). A total of 2824 blood-feeding ticks from seven genera were collected from wild animals: small rodents ($n=1151$) including *Rhombomys opimus* ($n=850$), *Spermophilus undulatus* ($n=209$), *Spermophilus dauricus* ($n=39$), *Rattus pyctoris* ($n=22$), *Niviventer confucianus* ($n=17$), *Apodemus agrarius* ($n=9$), *Apodemus draco* ($n=4$), *Meriones meridianus* ($n=1$) and other wild animals *Vormela peregusna* ($n=584$), *Lepus yarkandensis* ($n=523$), *Ochotona curzoniae* ($n=386$), *Pipistrellus pipistrellus* ($n=67$), *Vulpes vulpes* ($n=62$), *Hemichinus auritus* ($n=28$), *Meles leucurus* ($n=15$), *Lynx lynx* ($n=5$) and *Anourosorex squamipes* ($n=3$) (Supplementary Data 1). In addition, one tick (*Hy. asiaticum*) was collected from a patient²⁶.

The literature-based data provided essential supplementary information for the tick species included in the ecological modelling—*D. marginatus*, *D. nuttalli*, *D. silvarum* and *Hy. asiaticum*. This integration enabled a robust and precise modelling analysis, particularly for *D. silvarum*, as ecological modelling for this species would not have been feasible without the literature-based data (Supplementary Table 1). In addition, the diversity of tick species and their spatial distribution increased significantly when our field survey data were combined with literature-based data (Supplementary Data 3). Overall, 105 tick species from 11 genera were identified in 404 counties from the studied regions in China and neighbouring countries. The most widely distributed tick species was *Hy. anatolicum* found in 131 counties, followed by *Hy. asiaticum* in 122 counties, and *D. nuttalli* in 120 counties

(Supplementary Data 4). For 34 counties from Pakistan and 29 counties from China (XUAR, $n=25$; Gansu, $n=3$; IMAR, $n=1$), each reported >10 tick species, representing tick diversity hotspots. Notably, over 20 tick species were recorded in the Charsadda and Peshawar regions of Pakistan, while Wen county in Gansu and Kashgar city in XUAR reported 15–19 tick species (Supplementary Fig. 1).

Identification, composition and prevalence of RNA viruses in ticks

A total of 9610 samples of 20 tick species from 30 counties were pooled into 515 sequencing libraries, with 1–80 ticks per pool (Supplementary Data 1 and 5). Meta-transcriptomic sequencing generated a total of ~3.42 Tb raw data, ranging from 15,360 to 86,640,235 reads per library (Supplementary Data 5).

Accordingly, analyses of the assembled viral contigs from each library revealed 92 RNA virus species belonging to 19 defined virus families and unclassified families (Fig. 2 and Supplementary Data 6 and 7), from which 1567 full-length or near-complete genomes were obtained (Supplementary Data 6). The genome sizes of the RNA viruses identified in this study ranged from 2.3 Kb for Korla tick mitovirus to 19.8 Kb for Bachu tick virus 2, with most falling within 4–12 Kb. Among these, 56 viruses exhibited over 90% amino acid (aa) similarity in the RNA-dependent RNA polymerase (RdRp) protein or typical conserved domain to known viruses, and fell into 12 families: *Chuviridae*, *Flaviviridae*, *Hepeviridae*, *Iflaviridae*, *Lispiviridae*, *Nairoviridae*, *Orthomyxoviridae*, *Partitiviridae*, *Peribunyaviridae*, *Phenuiviridae*, *Rhabdoviridae* and *Totiviridae*. The remaining 36 viruses were more divergent, exhibiting 29.17–88.49% aa identities to known viruses, thereby meeting the criteria (i.e. <80% nucleotide identity across the complete genome or <90% aa identity of the RdRp domain with known viruses) to be classified as novel virus species (Table 1 and Supplementary Data 7).

The most common virus detected was Bole tick virus 1 (BLTV1, family *Phenuiviridae*), found in 222 pools (43.1% of all pools) from four tick species (*D. marginatus*, *Hy. anatolicum*, *Hy. asiaticum* and *R. turanicus*) collected in 15 counties. Taishun tick virus (TSTV) (*Rhabdoviridae*) was detected in 184 pools (35.7%) from seven tick species (*D. marginatus*, *Ha. punctata*, *Hy. anatolicum*, *Hy. asiaticum*, *Hy. detritum*, *Hy. scupense* and *R. turanicus*) collected in 18 counties, followed by Bole tick virus 4 (BLTV4) ($n=162$, 31.4%), Bole tick virus 3 (BLTV3) ($n=144$, 28.0%), Bole tick virus 2 (BLTV2) ($n=126$, 24.5%), Tick phlebovirus ($n=109$, 21.2%) and Brown dog tick phlebovirus 1 ($n=104$, 20.2%) (Supplementary Data 6). Other viral families, such as the *Dicistroviridae* ($n=6$, 1.16%), *Iflaviridae* ($n=9$, 1.74%), *Orthomyxoviridae* ($n=8$, 1.55%) and *Sedoreoviridae* ($n=1$, 0.19%) were less frequently detected. Several nairoviruses, including BURV, CCHFV, Nairoviridae sp., TAMV, Wusu tick nairovirus, novel Awat tick nairovirus (ATNV) and TcTV-1, were detected in the tick genera *Dermacentor*, *Haemaphysalis* and *Hyalomma*, but not in *Rhipicephalus* and *Argas* (Fig. 2).

Hy. asiaticum harboured the highest diversity of TBVs, with up to 38 virus species. Other tick species that carried more than 15 TBVs included *R. turanicus* ($n=30$), *D. marginatus* ($n=23$), *Hy. anatolicum* ($n=17$) and *A. persicus* ($n=15$). Viruses from the *Phenuiviridae* and *Rhabdoviridae* and several unclassified viruses, were the most widely distributed among different tick species, and showed relatively high abundance (Fig. 2). For example, the viral family *Phenuiviridae*, including BLTV1, Brown dog tick phlebovirus 1, Changping Tick Virus 1, *Dermacentor reticulatus* uukuvirus, Kharabali tick phlebovirus, Qinghai Lake uukuvirus, TcTV-2, Gonghe tick phlebovirus (GTPV), Tick phlebovirus and Xinjiang tick phlebovirus, was detected in 14 tick species (Supplementary Data 6).

Otyrar county in Turkistan (Kazakhstan) was the location with the highest variety of TBVs ($n=38$). Other counties that carried more than 15 TBVs included Bachu county ($n=21$), Shihezi city ($n=22$),

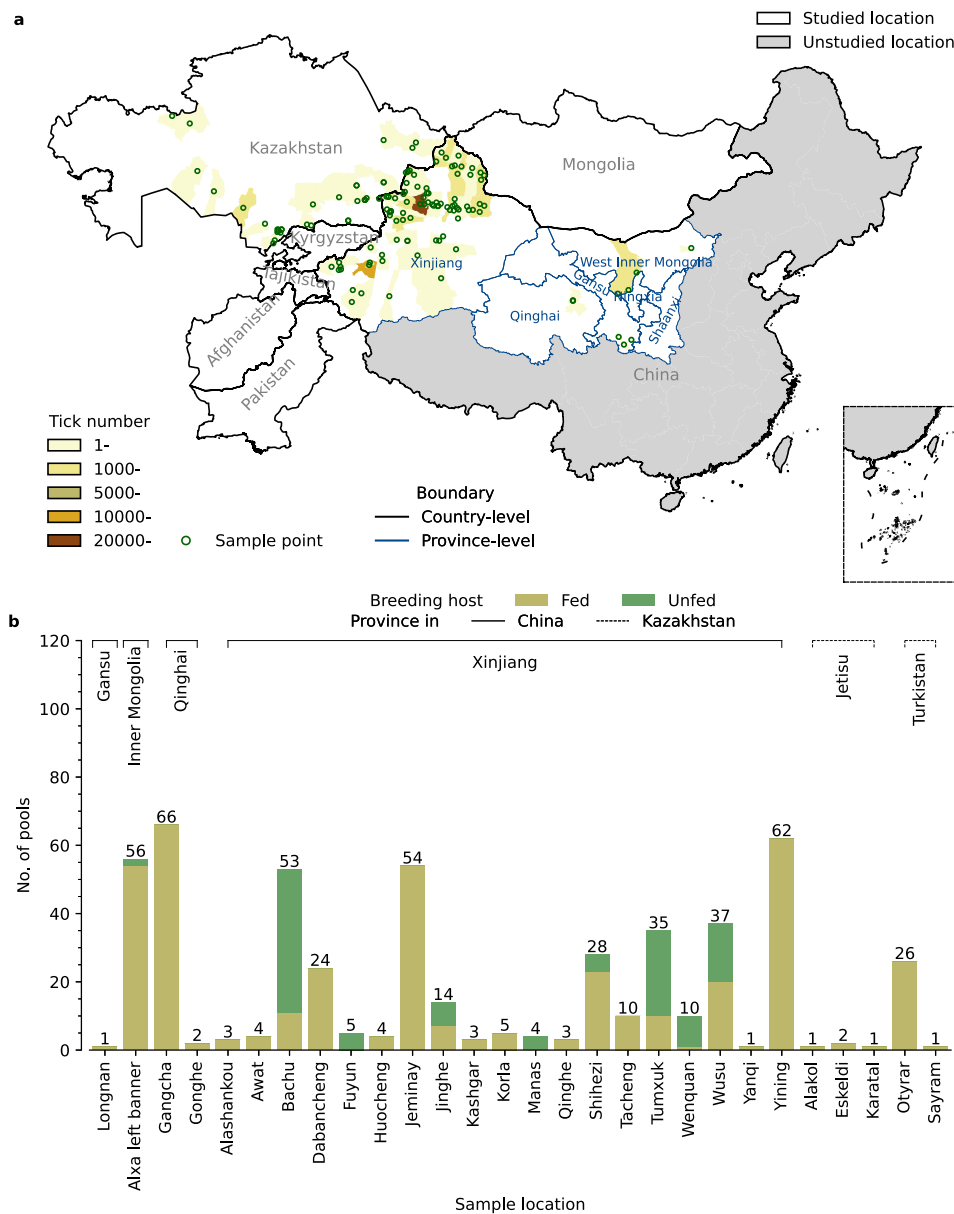


Fig. 1 | Overview of the ticks analysed in this study. a Geographic distribution of the ticks collected in northwest China and Kazakhstan. Shading colour represents the number of ticks collected at each location. Each sampling site is geo-referenced to the China and Kazakhstan maps based on its latitude and longitude. The Chinese

provinces unsampled are shown in grey. **b** Numbers of the pools for meta-transcriptomic sequencing per sampling location. Colours indicate the unfed/blood-feeding status of the ticks.

Tumxuk city ($n = 20$), Tacheng city ($n = 18$) and Yining city ($n = 15$) in XUAR, China, respectively. The literature-based data provided essential supplementary information for the four human-infecting TBVs included in the ecological modelling—CCHFV, TAMV, TcTV-1 and TcTV-2. Indeed, ecological modelling for CCHFV, TAMV and TcTV-2 would not have been feasible without the literature-based data (Supplementary Table 2). Furthermore, the diversity of TBVs and their spatial distribution increased significantly when our field survey data were combined with literature-based data (Supplementary Data 8). A total of 106 TBVs were identified across 161 counties in China and neighbouring countries. The Otyrar region in Kazakhstan recorded 38 TBVs, while Bachu county in XUAR, China, reported 24 TBVs (Supplementary Fig. 2). The most widely distributed TBV was CCHFV, found in 118 counties, followed by tick-borne encephalitis virus (TBEV) in 27 counties, and TSTV in 23 counties (Supplementary Data 9).

Composition and prevalence of human- or mammal-infecting TBVs

Among the 92 TBVs identified, 10 were associated with human and/or other mammalian infection—ALSV, BURV, CCHFV, DHOV, TAMV, TcTV-1, TcTV-2, Alxa tick phlebovirus (ATPV), Bhanja virus (BHAV) and Wad Medani virus (WMV)—with the recently described TBVs ATNV and GTPV also considered to pose a threat to public health (Table 1). Three of these viruses (ALSV, ATPV and TAMV) were detected in IMAR, GTPV was detected in Qinghai, TcTV-1 and TcTV-2 were detected in Kazakhstan, and 10 pathogenic TBVs (with the exceptions of ALSV and GTPV) were found in XUAR.

BURV, CCHFV, TAMV and TcTV-1 fell within the family *Nairoviridae*. Three of the 515 libraries (0.58%) tested positive for BURV, and three complete or near-complete genome sequences were generated. BURV was not previously identified in China. The three positive pools were from *Ha. punctata* ticks collected from Shihezi city ($n = 2$) and

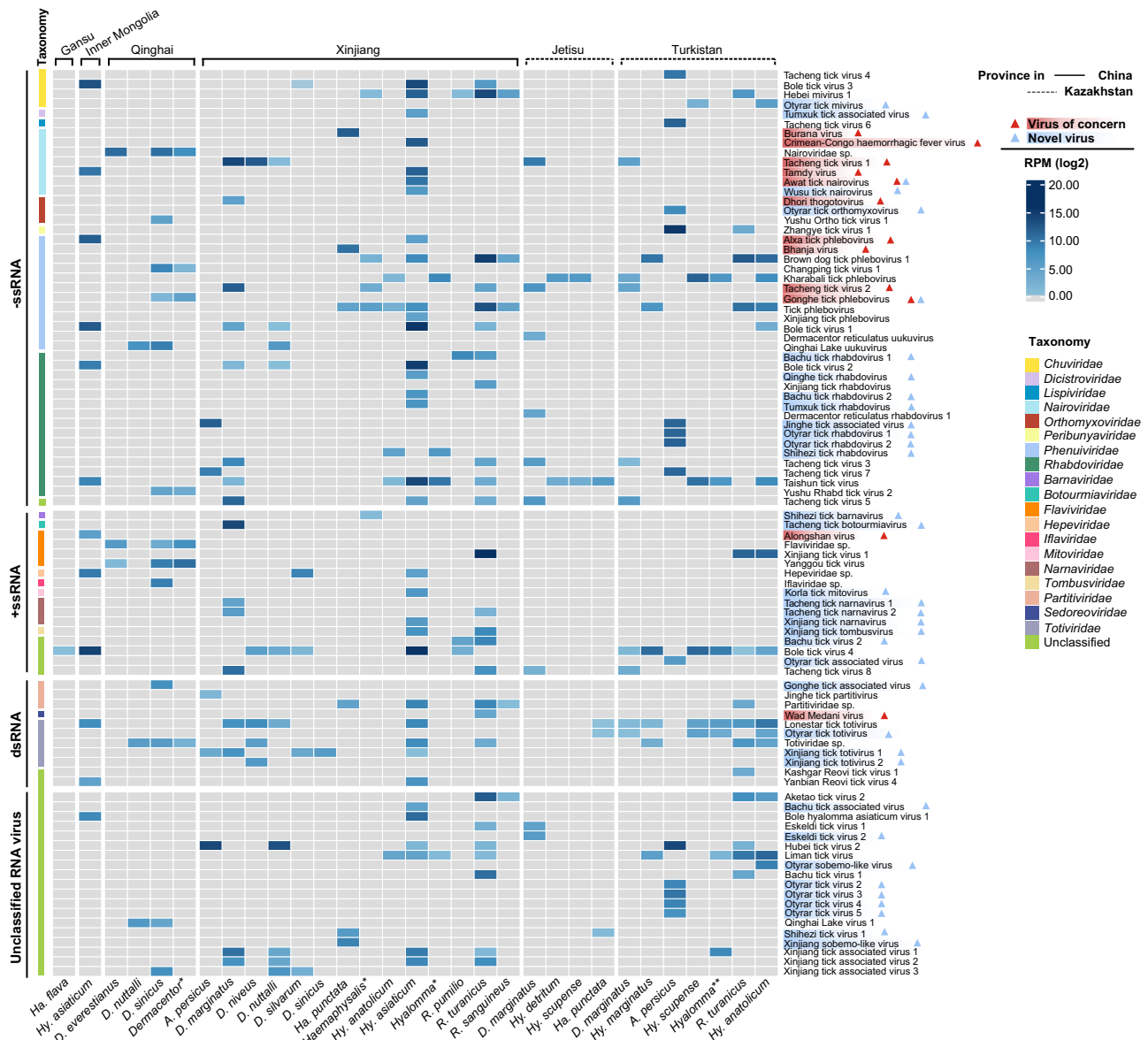


Fig. 2 | Distribution and abundance of the tick-associated viruses by tick species. The abundance of the tick-associated viruses is represented as a logarithm of the number of mapped viral reads per million total reads (RPM). Each column represents a different tick species from the geographic regions studied, while each row represents a virus species. The tick genus marked with an asterisk indicates that more than one species was mixed into a pool: *Derma-centor**—mixture of *D. sinicus* and *D. everestianus*, or *D. sinicus* and *D. silvatum*; *Haemaphysalis**—mixture of *Ha. sulcata* and *Ha. lagostrophis*; *Hyalomma**—mixture of *Hy. anatolicum* and *Hy.*

asiaticum; *Hyalomma**—mixture of *Hy. asiaticum* and *Hy. scupense*. Sampling sites are shown at the provincial level at the top. Tick taxonomy (species) and virus taxonomy (family) are shown at the bottom and left, respectively, and the colours of the strips on the left indicate the virus family. Red triangles indicate ‘viruses of concern’, defined as those that are pathogenic or potentially pathogenic to humans or mammals. Blue triangles indicate ‘novel viruses’, defined as those sharing <90% aa sequence identities with known RdRps.

Huocheng county ($n = 1$), XUAR (Table 1, and Supplementary Data 6). In addition, CCHFV was identified in *Hy. asiaticum* in 10 positive pools (1.94%): Korla city ($n = 5$), Aksu prefecture ($n = 2$), Kashgar city ($n = 2$) in southern XUAR and Altay prefecture ($n = 1$) in northern XUAR. Only 12 libraries (2.33%) tested positive for TAMV, and eight complete or near-complete genome sequences were generated from *Hy. asiaticum* collected in Alxa Left Banner ($n = 5$), Bachu county ($n = 1$), Dabancheng district ($n = 1$) and Korla city ($n = 1$). In addition, 21 libraries (4.08%) tested positive for TcTV-1, with 12 complete or near-complete genome sequences obtained. Positive samples were identified from the tick species *D. marginatus*, *D. niveus* and *D. nuttalli* collected from Tacheng city ($n = 9$), Jinghe county ($n = 6$), Wenquan county ($n = 1$), Yanqi Hui Autonomous county ($n = 1$) and Yining city ($n = 1$), China, as well as Eskeldi ($n = 1$), Otyrar ($n = 1$) and Sayram ($n = 1$), Kazakhstan,

respectively. This represents the definite identification of TcTV-1 in Kazakhstan. Additionally, a novel virus species from the genus *Ortho-nairivirus* was identified from four pools of *Hy. asiaticum* collected in Awat county, XUAR. This virus sequence shared the closest relationship (79.89% aa similarity in RdRp) with the TAMV strain LEIV-1308Uz, and was named ATNV (Table 1).

BHAV and TcTV-2 belong to the family *Phenuiviridae*. BHAV was first isolated from ticks (*Haemaphysalis intermedia*) in India in 1954, and a human laboratory BHAV infection was reported in the USA²⁷. We identified four BHAV from *Ha. punctata* infesting cattle and sheep collected from Shihezi city ($n = 3$) and Huocheng county ($n = 1$) (Table 1 and Supplementary Data 6). To our knowledge, BHAV was not identified in China and also in *Ha. punctata* ticks before this study. We identified 17 libraries (3.30%) that tested positive for TcTV-2, from

Table 1 | Human- and mammal-infecting viruses, and novel TBVs identified in this study

Virus	Positive rate (%)	Positive libraries	Complete genomes	Family	Genus	RdRp Length (aa)	Closest relative virus GenBank accession	Identity (%)
ALSV	0.39	2	2	Flaviviridae	unclassified	914	Alongshan virus UXX19107	99.11
ATNV	0.78	4	4	Nairoviridae	Orthonairovirus	3997	Tamdy virus AKC89322	79.89
BURV	0.58	3	3	Nairoviridae	Orthonairovirus	3972	Burana virus YP_009666118	99.62
CCHFV	1.94	10	5	Nairoviridae	Orthonairovirus	3945	Orthonairovirus haemorrhagiae ADD64466	98.99
TAMV	2.33	12	8	Nairoviridae	Orthonairovirus	3991	Orthonairovirus tomdense WNT98870	98.47
TcTV-1	4.08	21	12	Nairoviridae	Orthonairovirus	3962	Tacheng Tick Virus 1 YP_009304986	99.41
DHOV	0.39	2	1	Orthomyxoviridae	Thogotovirus	519	Thogotovirus dhorienae QB064970	98.65
ATPV	1.55	8	8	Phenuiviridae	Phlebovirus	2098	Alxa tick phlebovirus UXL90859	99.41
BHAV	0.78	4	4	Phenuiviridae	Bandavirus	2082	Bhanja virus AFO66272	99.23
GTPV	0.39	2	0	Phenuiviridae	Uukuvirus	2194	Tacheng tick virus 2 QIP68202.1	74.12
TcTV-2	3.30	17	17	Phenuiviridae	Uukuvirus	2189	Tacheng Tick Virus 2 QDF45276	99.45
WMV	0.19	1	0	Sedoreoviridae	Orbivirus	596	Wad Medani virus ATX74899.1	95.64

D. marginatus, *R. turanicus* and *Haemaphysalis* sp. Ticks collected from Tacheng city ($n = 9$), Yining city ($n = 4$), Shihezi city ($n = 1$) and Yanqi Hui Autonomous county ($n = 1$), China, as well as Eskeldi ($n = 1$) and Sayram ($n = 1$), Kazakhstan, respectively, from which 17 complete or near-complete genome sequences were obtained. A novel virus species from the genus *Uukuvirus* (*Phenuiviridae*) was identified from two pools of *Dermacentor* ticks collected in Gonghe county, Qinghai, China. It shared the closest (74.12% aa similarity in the RdRp) relationship with TcTV-2 strain TY1 isolated from a patient’s blood¹⁷, and was named GTPV (Table 1). In addition, another TBV belonging to the family *Phenuiviridae*–ATPV–was identified from one pool collected from Korla city and seven pools from Alxa Left Banner (1.55%), representing the identification of ATPV in XUAR, China. ATPV was only isolated from *Dermacentor* and *Hyalomma* ticks in Alxa Left Banner and Alxa Right Banner in 2022¹¹.

ALSV was first identified to be associated with human disease in northeast China⁷, and has now been found in mammalian hosts and ticks, including *Ixodes persulcatus* and *Ixodes ricinus* in multiple countries outside of Asia²⁸. Notably, two pools of *Hy. asiaticum* collected from a Bactrian camel from Alxa Left Banner tested positive for ALSV. Two near-complete genome sequences were generated (Table 1 and Supplementary Data 6), representing the identification of ALSV in western IMAR as well as in *Hy. asiaticum*. The infected camel may be a potential source of ALSV, and the ALSV detected might originate from *Hy. asiaticum* feeding on a viremic camel. Both the tick-transmitted DHOV and Thogoto virus (THOV) have been known to cause human infections, with clinical manifestations ranging from benign febrile symptoms to meningoencephalitis²⁵, and have been identified in both China and Kyrgyzstan¹⁹. Two pools of *D. marginatus* from Tacheng city and the Ili region tested positive for DHOV, and one set of near-complete genome sequences of PA, PB1, PB2, NP, M and GP were assembled (Table 1). Finally, one pool of *Hy. asiaticum* tested positive for WMV, which can be transmitted by a range of hard ticks and cause pathogenesis within mammalian cell lines and mice^{29,30}. The positive library was from Tumxuk city, and near-complete genome sequences of VP2-VP7 and NS1-NS3 were assembled (Table 1).

Evolution of human- or mammal-infecting TBVs

We next evaluated the evolutionary history of the 12 important TBVs (Fig. 3 and Supplementary Figs. 3–12). Phylogenetic analyses of 351 full-length CCHFV genome sequences globally (Supplementary Data 10) indicated that the NP gene of CCHFV could be placed into eight lineages: Asia 1-2, Africa 1-3 and Europe 1-3 (Fig. 3a and Supplementary Fig. 3). The NP gene of our five sequences fell within the Asia 2 lineage. JMNO1 and XMLLC14 formed a distinct sub-lineage with four sequences from Uzbekistan and Kazakhstan. The GP gene of CCHFV could be placed into 10 lineages, comprising Asia 1-4, Africa 1-3 and 3* and Europe 1-2 (Fig. 3a). Five sequences–XHYYL8 and KELML01–KELML04–fell within the Asia 4 lineage, while JMNO1, XHYYJ2 and XML09 fell within the Asia 2 lineage. XMLLC14 fell within the Africa 3 lineage (Fig. 3a). Eight lineages were also present in the phylogenetic tree of the RdRp gene (Fig. 3a). XHYYL8 and XHYYJ2 fell into the Asia 1 and Asia 2 lineages, respectively. JMNO1, XML09 and XMLLC14 fell into the Asia 3 lineage. Taken together, the newly identified CCHFV sequences belonged to four genotypes, including two known genotypes: Asia 2-Asia 2-Asia 2 (XHYYJ2), and Asia 2-Asia 2-Asia 3 (JMNO1, XML09), and two novel genotypes: Asia 2-Asia 4-Asia 1 (XHYYL8), and Asia 2-Africa 3-Asia 3 (XMLLC14) (Fig. 3b).

Phylogenetic analyses similarly revealed that the TAMV sequences identified here formed a distinct sister clade to the lineage of TAMV and Wanowrie virus previously reported (Fig. 3c and Supplementary Fig. 4). Phylogenetic analyses of the RdRp gene of the newly identified TcTV-1 viruses revealed two clusters: one in *D. marginatus* and *D. niveus*, as well as the reference strain QH1 associated with human infection, and another cluster comprising sequences in *D. marginatus*

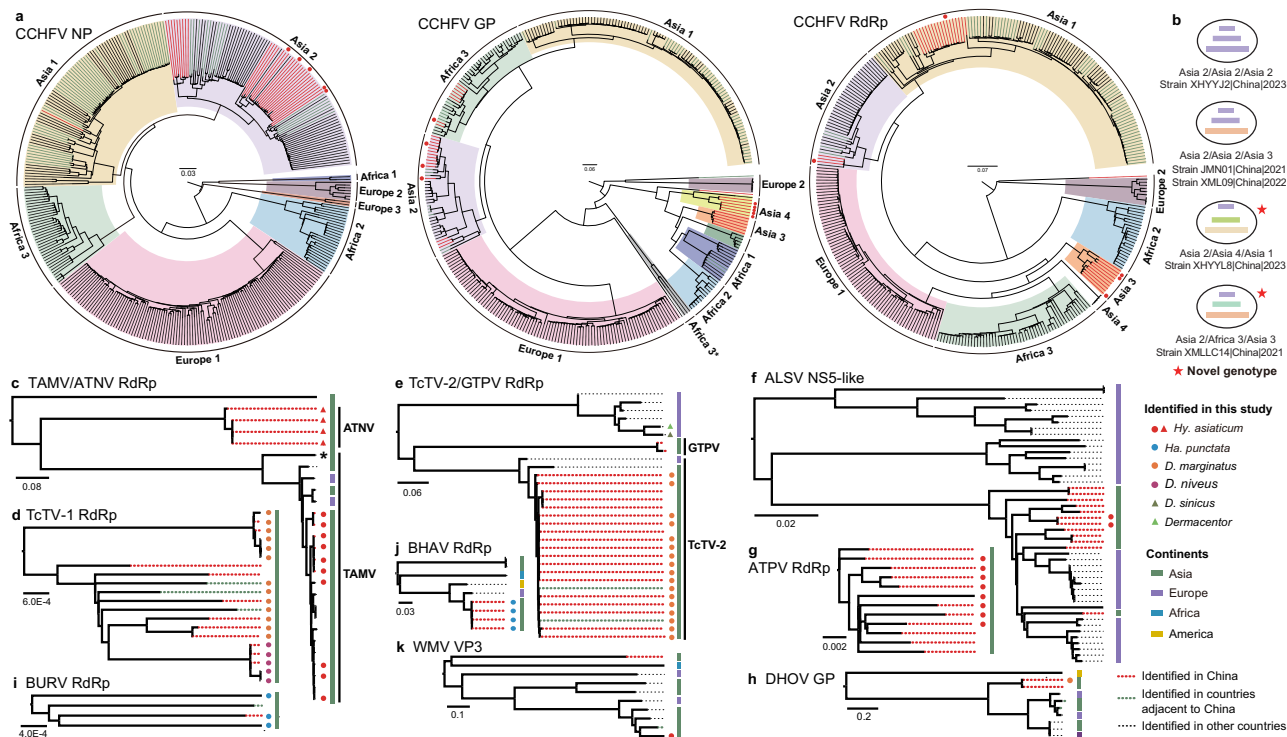


Fig. 3 | Phylogenetic analyses of pathogenic human-/mammal-infecting TBVs. **a** Phylogenetic trees of the CCHFV NP (left), GP (middle), and RdRp (right) genes. Genotypes are labeled and shaded in different colours. **b** Genotypes of the CCHFV sequences obtained in this study. The three gene segments (shown as horizontal bars starting from top to bottom of the ‘virion’) are NP, GP and RdRp. Different colours of the gene segments represent different virus lineages. The two novel genotypes are denoted with red stars. **c–k** Phylogenetic trees of the 11 important TBVs discovered in this study. In all trees, the viruses identified here are marked with solid circles (known viruses) and triangles (novel viruses), and the colours represent different tick species. The asterisk (*) in panel c represents the Wanowrie

virus. Branch colours represent the isolation country: China in red, adjacent countries in green, and other countries in black. The isolation continents of all the viruses are also highlighted with different colours. Trees are midpoint-rooted, and the scale bars represent the number of nucleotide substitutions per site. Bootstrap values >70% are shown for major nodes. ALSV Alongshan virus, ATNV Awat tick nairovirus, ATPV Alxa tick phlebovirus, BHAV Bhanja virus, BURV Burana virus, CCHFV Crimean-Congo haemorrhagic fever virus, DHOV Dhori virus, GTPV Gonghe tick phlebovirus, TAMV Tamdy virus, TcTV-1 Tacheng tick virus 1, TcTV-2 Tacheng tick virus 2, WMV Wad Medani virus. (See Supplementary Figs. 3–12 for each tree in detail).

collected from cattle in this study (Fig. 3d and Supplementary Fig. 5). Phylogenetic analysis also revealed that the TcTV-2 sequences identified here grouped with those from China, as well as those from Romania and Turkey. In addition, GTPV formed a distinct sister clade (Fig. 3e and Supplementary Fig. 6).

In the phylogenetic trees of ALSV (Fig. 3f and Supplementary Fig. 7), ATPV (Fig. 3g and Supplementary Fig. 8), and DHOV (Fig. 3h and Supplementary Fig. 9), most of the newly identified sequences were grouped within previously described viruses from China. The NSS-like protein of ALSV detected in *Hy. asiaticum* in this study shared 99.11% amino acid identity with the ALSV strain NE-TH4 found in *Ixodes persulcatus*, and clustered closely with NE-TH4 and other strains also from *I. persulcatus* identified in China (Fig. 3f and Supplementary Fig. 7). Hence, this may simply reflect feeding on a viremic camel host. In the phylogenetic trees of BURV, the three novel Chinese sequences clustered with a Kyrgyzstan sequence (NC_0434397-NC_043439), the only one available in GenBank (Fig. 3i and Supplementary Fig. 10). Similarly, in the phylogenetic trees of BHAV (Fig. 3j and Supplementary Fig. 11), the viruses newly described formed a separate and novel clade, although this may reflect a lack of reference viruses. However, the novel Chinese WMV sequences clustered with previously described viruses from Tajikistan, Turkmenistan and Azerbaijan in the phylogenetic trees of VP3 and other genes (Fig. 3k and Supplementary Fig. 12). Generally, phylogenetic analyses of the 12 viruses estimated using maximum likelihood (ML) and Bayesian methods revealed consistent results (Supplementary Figs. 3–12).

Risk map for tick species

Risk maps were projected for four tick species that serve as primary vectors for four human- and mammal-infecting TBVs with sufficient data for modelling (Supplementary Table 3). In the main analysis, the BRT model used all tick data (1950–2024) as dependent variables, and risk factors comprised climatic, land type and livestock as independent variables (Supplementary Table 4). Since over 90% of the tick sampling records in our comprehensive data set spanned from 1980 to 2024 (Supplementary Fig. 13), climatic data between 1980 and 2024 were used to derive bioclimatic predictors. After screening the multicollinearity among the 19 candidate predictors, four predictors—bio5 (max temperature of warmest month), bio8 (mean temperature of wettest quarter), bio10 (mean temperature of warmest quarter) and bio18 (precipitation of warmest quarter)—were included in the BRT model as representative climatic risk factors (Supplementary Figs. 14 and 15 and Supplementary Tables 5 and 6). After correcting sampling bias by a logistic model (Supplementary Table 7), the models were configured with the following parameters: for *D. marginatus* and *D. nuttalli*, the models used a tree complexity of 7, a learning rate of 0.001 and a bag fraction of 0.7; for *D. silvarum*, a tree complexity of 5, a learning rate of 0.001 and a bag fraction of 0.7 were applied; and for *Hy. asiaticum*, a tree complexity of 5, a learning rate of 0.01 and a bag fraction of 0.8 were applied (Supplementary Table 8). The estimated average testing area under the curve (AUC) ranged from 0.80 to 0.87, sensitivity ranged from 0.78 to 0.87 and specificity ranged from 0.70 to 0.84 (Supplementary Table 9), indicating modest support for the projections. Importantly, the model predicted that high-risk counties

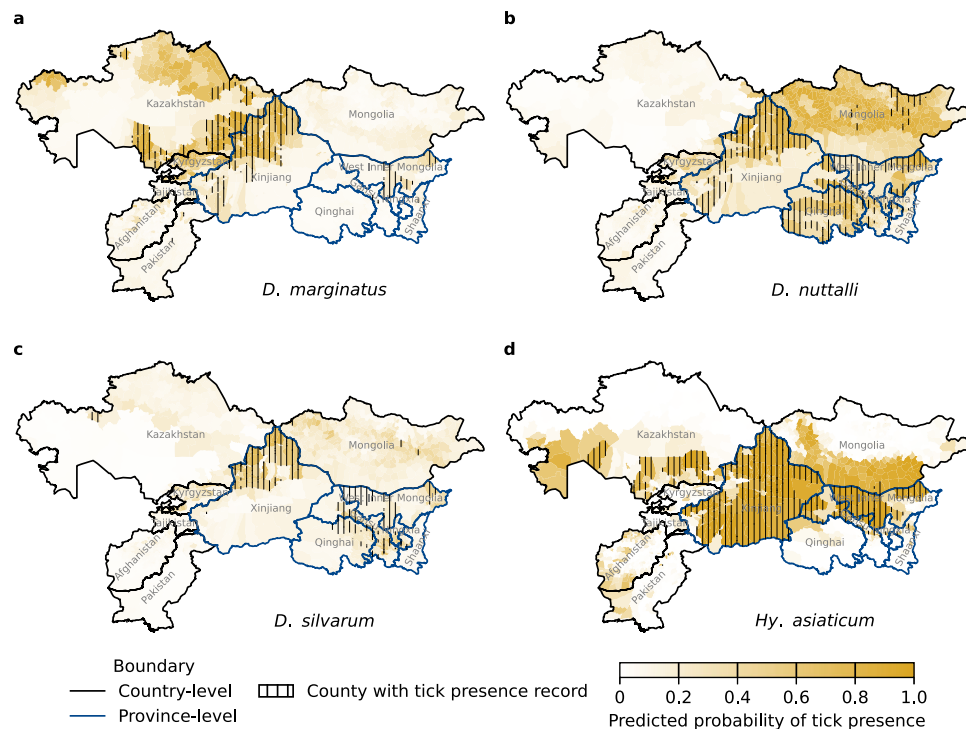


Fig. 4 | Modelling the distribution of the four major tick species in northwest China and adjacent countries. a–d Model-predicted probabilities of the presence of *D. marginatus*, *D. nuttalli*, *D. silvarum* and *Hy. asiaticum*. Counties with the

observed tick presence are highlighted using vertical lines. Geographic range of the map: 18°–57° N, 45°–140° E.

for tick species were more extensive than currently observed (Fig. 4). The number of counties with *D. silvarum* increased by 1327.3% ((942–66)/66 × 100%), rising from 66 counties (observed) to 942 counties (predicted). This was followed by *D. marginatus* (1016.0%, from 81 to 904), *D. nuttalli* (871.7%, from 120 to 1166) and *Hy. asiaticum* (665.6%, from 122 to 934) (Supplementary Table 10). All four tick species showed a wide distribution across northwest China and neighbouring countries. Specifically, *D. marginatus* demonstrated a significant distribution potential in Kazakhstan and Kyrgyzstan (Fig. 4a). *D. nuttalli* was projected to have high distribution potential across the whole of Mongolia (Fig. 4b), and a similar distribution pattern was also observed in *D. silvarum*, although with a probability of less than 20% (Fig. 4c). *Hy. asiaticum* displayed a distinct distribution pattern, primarily within latitudes 30°–50° N (Fig. 4d). Accordingly, the modelled geographic area and population size of the four tick species increased by 213.5% to 733.3%, and 584.8% to 1058.3%, respectively (Supplementary Table 10). The relative contribution of risk factors varied among tick species (Supplementary Table 11). The highest contributing risk factors were the density of horses for *D. marginatus* (11.1%), bio8 for *D. nuttalli* and *D. silvarum* (10.3% and 11.8%, respectively), and the density of cattle for *Hy. asiaticum* (8.5%). The partial dependence plots of risk factors with over 5% relative contribution showed that different risk factors had different impacts on the occurrence risk of ticks (Supplementary Figs. 16–19). Two additional analyses, focusing on climatic and tick distribution data for periods before 2010 and after 2010, were conducted. For all four tick species, the modelled high-risk counties, geographic area and affected population sizes were broader than the observed counts, supporting the robustness of our modelling analyses (Supplementary Figs. 20 and 21 and Tables 12–14).

Risk map for TBVs

We also created risk maps for four human- and mammal-infecting TBVs with abundant data for ecological modelling. In the main analysis, the

BRT model used all TBV data (1950–2024) as dependent variables, and risk factors used in the modelling of ticks, along with two demographic factors, as independent variables (Supplementary Table 4). The model-predicted distribution of ticks was used to define the control counties for the corresponding TBVs, and therefore, it was not included in the model as an independent variable. The models were configured with the following parameters: for CCHFV and TAMV, the models used a tree complexity of 3, a learning rate of 0.01 and a bag fraction of 0.7; for TcTV-1, a tree complexity of 7, a learning rate of 0.001 and a bag fraction of 0.7 was applied; and for TcTV-2, a tree complexity of 3, a learning rate of 0.005, and a bag fraction of 0.76 was applied (Supplementary Table 15). The estimated average testing AUC ranged from 0.92 to 0.98, sensitivity ranged from 0.92 to 1.0 and specificity ranged from 0.89 to 0.97, indicating highly accurate predictions (Supplementary Table 16). The modelling results showed that high-risk counties with TBVs were much more extensive than currently observed (Fig. 5). The number of counties with TAMV increased by 2844.4% ((265–9)/9 × 100%), rising from 9 counties (observed) to 265 counties (predicted). This was followed by TcTV-2 (1,100.0%, from 15 to 180), CCHFV (733.9%, from 118 to 984) and TcTV-1 (311.1%, from 9 to 37) (Supplementary Table 17). Remarkably, 984 counties (65.4% of 1505) were identified as high risk for CCHFV (Fig. 5a). CCHFV exhibited 96.3% distribution potential in XUAR (104 of 108 counties predicted as high-risk) and 68.2% (15/22) in Ningxia, 64.7% (33/51) in IMAR, 30.0% (27/90) in Gansu, 18.69% (20/107) in Shaanxi and 15.56% (7/45) in Qinghai, China. The distribution potential of CCHFV was also extremely high in Afghanistan (91.8%, 301/328), Kyrgyzstan (88.6%, 39/44), Pakistan (85.8%, 121/141) and Tajikistan (79.3%, 46/58). In contrast, although lower in Kazakhstan (59.77%, 104/174) and Mongolia (49.55%, 167/337) compared to other countries, the CCHFV risks remained high in these regions relative to those of other viruses. TAMV had a similar but lower distribution potential compared to CCHFV and was predicted to affect 265 counties (17.6% of 1505) (Fig. 5b). TcTV-1 and TcTV-2 had a similar pattern, with high distribution potential in northern

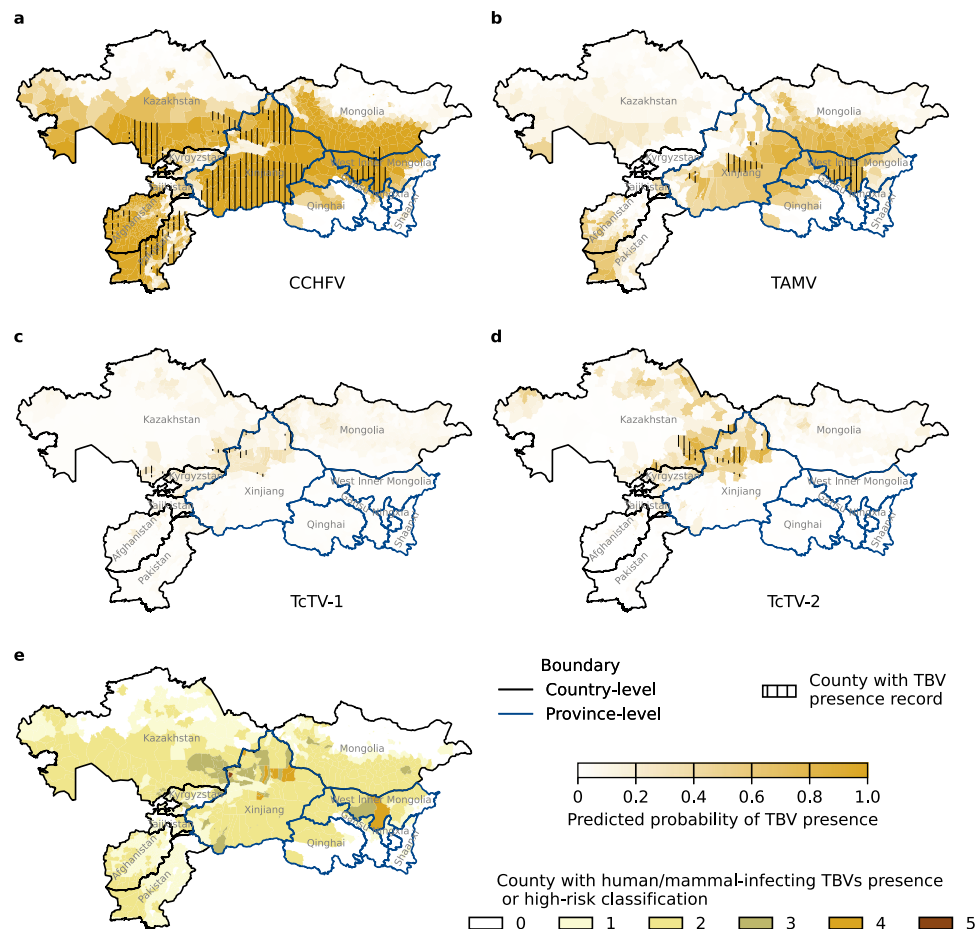


Fig. 5 | Modelling the distribution of the major zoonotic TBVs in northwest China and adjacent countries. a–d Model-predicted probabilities of the presence of CCHFV, TAMV, TcTV-1 and TcTV-2, along with locations reporting their presence (no matter they were reported as cases or detected in ticks). Counties with the

observed TBV presence are highlighted using vertical lines. **e** Distribution of the TBVs in counties with recorded and model-predicted presence. Geographic range of the map: 18°–57° N, 45°–140° E. CCHFV Crimean-Congo haemorrhagic fever virus, TAMV Tamdy virus, TcTV-1 Tacheng tick virus 1, TcTV-2 Tacheng tick virus 2.

XUAR, China, and adjacent areas of Kazakhstan (Fig. 5c, d). Accordingly, the predicted geographic area and affected population size of the four TBVs increased by 320.4% to 1906.9%, and 297.9% to 1778.6%, respectively (Supplementary Table 17). Overall, Alxa Left Banner in IMAR and northern XUAR in China were identified as high-risk regions, reporting at least four of the 12 high-risk TBVs, either through recorded presence or classification as high-risk (Fig. 5e). The distribution patterns of these TBVs were heavily shaped by the distribution of their primary vectors (e.g. *Hy. asiaticum* for CCHFV and TAMV). The relative contribution of risk factors varied among TBVs (Supplementary Table 18). For CCHFV and TAMV, which have a single tick vector, the built-up area and density of cattle were the most significant contributing factors, with a maximum contribution of up to 26.8%. In contrast, for TcTV-1 and TcTV-2, which are associated with multiple tick vectors, horse density contributed 28.8% and 13.7%, respectively. Similarly, the partial dependence plots showed that different risk factors had different impacts on the occurrence risk of TBVs (Supplementary Figs. 22–25). For example, basically, the occurrence risk of CCHFV, TAMV, TcTV-1 and TcTV-2 increased with the decrease in precipitation of the warmest quarter (bio18). However, when the mean temperature of the wettest quarter (bio8) exceeded 20 °C, the risk of TAMV occurrence increased until reaching -23 °C. In the additional analyses for CCHFV before 2010 and all four TBVs after 2010, the modelled high-risk counties, geographic area and affected population size were generally consistent and also broader than the observed,

supporting the robustness of our modelling analyses (Supplementary Figs. 26 and 27 and Tables 19 and 20).

Discussion

This is a systematic surveillance and assessment of tick species and TBV diversity in northwest China and adjacent countries. Compared to previous studies, sample collection was more systematic and performed in 82 counties covering northwest China (XUAR, IMAR, Qinghai and Gansu) and Kazakhstan, with the dominant tick vector species for known zoonotic TBVs in this region chosen for meta-transcriptomic sequencing. In addition, we performed a systematic literature search using the terms ‘tick’, ‘tick-borne disease’, ‘tick-borne virus’, ‘diversity’, ‘risk factor’, ‘northwest China’ and ‘Central Asia’, in various combinations, to search studies published between January 1950 and October 2024 in six major reference databases. From this literature search, we identified 105 tick species from 11 genera in 404 counties from the studied regions in China and neighbouring countries. Thus, the diversity of tick species and TBVs, as well as their spatial distribution, increased when combining our field surveys with the literature-based data. Finally, we identified 92 RNA viruses and assembled 1567 viral genomes from our collection of 9610 samples of 20 tick species from 30 counties, including 10 viruses that infect or have the potential to infect humans and/or other mammals, as well as two novel viruses closely related to known pathogenic TBVs. Notably, BHAV and BURV were not

previously reported in China, ALSV and WMV were not previously reported in northwest China, and TcTV-1 was never previously reported in Kazakhstan. CCHFV was predicted to have the most extensive geographic distribution, approximating the distribution of its principal vector and reservoir ticks of the genus *Hyalomma*^{5,31}. The 10 CCHFV genome sequences generated here belonged to four genotypes, including two novel genotypes defined for the first time.

For decades, the geographic distribution of BHAV was thought to cover southern and Central Asia, Africa, and southern (partially also central) Europe^{32–34}, with some genomic evidence for BURV in Kyrgyzstan¹⁸. These viruses have been detected in *Ha. intermedia* and *Rhipicephalus decoloratus*, as well as in several cases of patients with febrile illness. In *Ha. punctata*, we identified three BHAV sequences from Shihezi city and one from Huocheng county in XUAR, China. Moreover, we identified two BURV sequences from the same pools from which two BHAV were detected, indicating an expansion of both viruses into China. We also observed an expansion of the newly documented pathogenic TBV, TcTV-1, into Kazakhstan, Central Asia. These viruses should be included in the diagnostic assessment of symptomatic cases associated with tick bites and vector surveillance efforts. Of note, 43.1% of the pools tested positive for BLTV1, 28.0% were positive for BLTV3, and 31.4% were positive for BLTV4. A wide distribution range of these viruses may pose potential public health risks that warrant further investigation and monitoring.

Our modelling predicts that the number of high-risk counties for the four tick species will be 6.6-fold larger than currently observed values. The relative contribution of risk factors varied among tick species, and their impacts on the occurrence risk of ticks also varied. Additionally, we identified several tick diversity hotspots in China and Pakistan, highlighting the need for increased surveillance in these regions as they may serve as reservoirs for novel TBVs.

The model-predicted number of high-risk counties for the four human/mammal-infecting TBVs were also larger than observed, consistent with previous studies^{4,35}. CCHFV was predicted to have the widest distribution potential, affecting >95% of counties in XUAR, China, as well as >75% of the counties in a few adjacent countries, excluding Kazakhstan and Mongolia. Similarly, the relative contribution of risk factors and their impacts on the occurrence risk of TBVs also varied, suggesting the complicated biological and ecological factors jointly shaping the diversity of TBVs in northwest China and adjacent countries.

Our study has some limitations. No *Ha. longicornis* and only 14 *I. persulcatus* ticks were collected in this study, which might account for the absence of SFTSV and TBEV. The absence of the tick species is more likely due to habitat selection, seasonal mismatch, insufficient sampling intensity and current niche shrinkage. *I. persulcatus* prefers coniferous and broad-leaved mixed forests with humidity >80%^{36–38}, while the sampling sites of this study were mainly shrub grasslands and farmland edges, which were not suitable for the survival of *I. persulcatus*. *Ha. longicornis* was found in regions between 18°–53° latitude in the northern hemisphere. The critical equilibrium relative humidity of *Ha. longicornis* is about 85%, and if humidity drops below this threshold, ticks will lose their water continuously^{39,40}. This might explain why *Ha. longicornis* is mainly distributed in coastal areas, and Changji Hui Autonomous Prefecture is the only recorded location of this tick in XUAR, China⁴¹. Moreover, the grazing area of local cattle and sheep, and the community structure of wild small mammals have changed, which may further shrink the suitable habitat range of the two tick species. In addition, the grid data for livestock was sourced from 2015, and land cover from 2021, while the tick species and TBV data were collected over a broad time period. This may introduce an unknown effect on the model-predicted results. The low detection rate of TBVs in field-collected ticks, coupled with the high underreporting rate of tick-borne disease cases, poses challenges in establishing accurate

criteria for defining a control county. The stricter selection criteria in this study resulted in fewer counties being defined as controls, which further contributes to an overestimation of high-risk counties for TBVs. However, from a disease prevention perspective, overestimation is preferable to underestimation. For instance, based on our literature review, TBEV has been reported in 27 counties, although it was not detected in this study. In addition to sampling deficiency, misidentification of other flaviviruses as TBEV by qPCR or serological investigation might account for the inconsistent positive rate. In 2022, a significant serological cross-reaction between the tick-borne flavivirus Karshi Virus and TBEV was documented, implying that the reported seroprevalence of TBEV in humans and animals in northwest China may be overly high⁴². In-depth molecular detection and epidemiological investigations should therefore be performed to confirm the TBEV prevalence and incidence.

Overall, our results provide evidence that at least 10 pathogenic TBVs circulate in northwest China and adjacent countries, including some that were previously undetected. The high-risk regions of these TBVs might also be much more extensive than currently known, and most of the counties in the region were predicted to be at high risk for CCHFV. In sum, this study provides valuable resources for the prevention and control of tick-borne disease in this region and also globally.

Methods

Sample collection

From 2012 to 2024, a total of 100,406 ticks were collected from 82 counties in XUAR ($n = 85,960$), Qinghai ($n = 475$), Gansu ($n = 55$), and IMAR ($n = 1602$), China, and Kazakhstan ($n = 12,314$) (Supplementary Data 1). In total, there were 53 counties covering northern and southern XUAR, and 22 counties in Kazakhstan, with a nationwide distribution. The collection sites included different ecological environments, such as desert, shrubland, steppe and farmland. The latitude and longitude of each collection site were recorded. We collected 41,847 unfed ticks by flagging, and 58,559 blood-feeding ticks were removed from different domestic or wild animals, such as camels, cattle, horses, sheep, dogs, poultry, mice, bats, hedgehogs, badgers, foxes, rabbits and mustelids.

The dominant tick vector species for known zoonotic TBVs collected from understudied counties were selected for meta-transcriptomic sequencing. Samples collected in the early years of the study (2012–2015) were not well-stored and therefore were not chosen for sequencing. From 2016 to 2024, a total of 7486 tick samples chosen for meta-transcriptomic sequencing were collected from 46 counties in northern XUAR and seven counties in southern XUAR, including the dominant tick species of *D. marginatus*, *D. nuttalli*, *D. silvarum*, *Hy. asiaticum*, *Hy. anatolicum*. In addition, densely populated counties with different ecological environments were also included in the study. Because Gansu, Qinghai and IMAR have been less well studied, all the ticks from Gansu ($n = 55$) and Qinghai ($n = 402$), and the majority (1068 out of 1602) from the IMAR were chosen for sequencing. Ticks of the same species and collection location from IMAR were randomly chosen for sequencing. In addition, 599 ticks of 13 species from five counties of Kazakhstan were also selected for meta-transcriptomic sequencing. These samples included the dominant tick species of *D. marginatus* and *Hy. asiaticum* in Kazakhstan^{43,44}.

In sum, a total of 9610 ticks containing 3493 unfed and 6117 feeding ticks from 30 counties were used for meta-transcriptomic sequencing. All samples were captured alive and then transported from the field to the laboratory, and stored at -80°C until further processing.

Sample processing and meta-transcriptomic sequencing

All ticks were pretreated as described previously⁴⁵. They were first washed in 1 mL of 70% ethanol, followed by the $1 \times$ PBS solution. The

species and developmental stage of each tick were tentatively identified using the dissecting microscope (Olympus, Japan). Ticks were divided into pools on the basis of species, sampling sites and time, and unfed/blood-feeding status. The pools of unfed ticks were suspended in 1 × PBS and homogenised with steel beads (2 mm) employing a tissue crusher (Jingxin, Suzhou, China) at 4 °C and 70 Hz for 120 s. The pools of blood-feeding ticks were pestled in a mortar under liquid nitrogen. Total RNA of each pool was extracted using RNAiso Plus (Takara, Japan) from the ground products and purified according to the manufacturer's instructions. RNA quantity and quality were checked, and used for RNA library construction. Ribosomal RNA (rRNA) was removed using the MGIEasy rRNA Depletion Kit (Human-Mouse-Rat) (MGI, Shenzhen, China). A total of 515 RNA sequencing libraries were then constructed using the MGIEasy mRNA library Prep Kit (MGI, Shenzhen, China). Paired-end (100 bp) sequencing of each RNA library was performed on the MGISEQ-2000RS platform in our laboratory.

Data analysis and confirmation of tick species

Raw sequencing reads were first subjected to quality control, using Fastp v.0.20.0 and Trimmomatic v.0.36⁴⁶ to remove adaptor and low-quality reads. To further confirm the tick species, clean reads were mapped onto a custom tick reference database using Bowtie2 v2.4.1, employing the end-to-end mode and searching for at most two distinct alignments with 10 parallel search threads⁴⁷. The reference database only included representative mitochondrial cytochrome c oxidase I (*COI*) gene sequences from 109 tick species, including 5 tick species of the genus *Argas*, 13 tick species of the genus *Dermacentor*, 25 tick species of the genus *Haemaphysalis*, 15 tick species of the genus *Hyalomma*, 34 tick species of the genus *Ixodes* and 17 tick species of the genus *Rhipicephalus* (Supplementary Data 11). The tick species was then determined based on the maximum number of mapped reads and the coverage of the assembled *COI* gene sequences. Based on the SAM files from Bowtie2, SAMtools (v1.10)⁴⁸ was used to calculate the coverage depth of the *COI* gene, and Geneious Prime (v2023.2.1) was used to check the mapping results manually. These tools were run with default settings unless otherwise specified.

Sequence assembly and virus discovery

Clean sequence reads generated after quality control were processed for de novo assembly using Trinity v2.5.1 with default settings⁴⁹. The assembled contigs were annotated by BLASTN and Diamond BLASTX⁵⁰ against the GenBank non-redundant nucleotide (nt) and non-redundant protein (nr) databases (downloaded on 2021/02/22), with e-value thresholds of 1×10^{-10} and 1×10^{-5} , respectively. The contigs annotated as the same tick-associated virus in individual libraries were first merged using the 'De Novo Assemble' tool embedded in Geneious Prime with the sensitivity parameter set to 'Highest Sensitivity/Slow'. Subsequently, reads were directly mapped to the sequence of a re-merged close relative using Bowtie2 employing the end-to-end mode and 15 parallel search threads. A consensus genome sequence was obtained from the mapped reads as the final sequence, with thresholds of the consensus $\geq 90\%$ at all sites and the number of degenerate bases $\leq 0.2\%$ of the total genome length. Ambiguous bases were determined by a consensus threshold setting (90%) for IUPAC ambiguity codes in Geneious. The relative abundance of each virus was estimated using the number of reads per million (RPM), calculated as 'mapped viral reads/raw reads*one million'.

The assembled contigs were annotated by BLASTN and Diamond BLASTX against the GenBank non-redundant nt and nr databases, which included all prokaryotic and eukaryotic species. All contigs annotated as prokaryotes and eukaryotes were removed, and contigs annotated as viruses were retained for further analyses. In addition, we removed any possible contaminating viral sequences as previously reported⁵¹. Finally, if the read abundance of a virus was less than 0.1% of the highest value for that virus among other libraries, the virus in

question was discarded to exclude potential false-positives, as the index-hopping rate is usually between 0.01% and 0.1%⁵².

Assembled viral contigs exceeding RPM ≥ 1 and 1000 bases in length were retained for further classification. Open reading frames (ORFs) of these viruses were predicted using the online tool ORFfinder (<https://www.ncbi.nlm.nih.gov/orffinder/>), and functional predictions of conserved domains were performed against the Conserved Domain Database (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>). Amino acid sequences with positive hits to the RdRp domain were retained and compared against the non-redundant protein database at NCBI using the BLASTP programme (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome). A commonly applied rule for the delineation of a new virus species is that it exhibits $<80\%$ nucleotide identity across the complete genome or $<90\%$ amino acid identity in the RdRp to known viruses^{53–57}. For known viruses identified in this study, the virus strains were named as the best-matched virus from the BLASTP comparison, followed by the sequencing pool ID. For putative novel viruses, they were named using the sampling location and the best match viral family from the BLASTX search if the amino acid identity in the RdRp was between 40% and 90%. In cases where RdRp amino acid identity was $<40\%$, viruses were named using the sampling location and the term 'tick-associated virus'.

Phylogenetic analysis

For the viruses of potential health concern, all complete or near-complete sequences available in the NCBI database were included as references for phylogenetic analysis. For CCHFV, sequences with 100% sequence identity were removed to minimise redundancy. Nucleotide sequences for each group were aligned using MAFFT v7.490⁵⁸, employing the E-INS-i algorithm in Geneious. Phylogenetic trees were then inferred using the ML method in IQ-TREE v1.6.1241, with the best-fit substitution model and 1000 bootstrap replicates⁵⁹. Phylogenetic analysis was also performed using a Bayesian method as specified in MrBayes⁶⁰, employing the GTR nucleotide substitution model. Ten million steps were run, with trees and parameters sampled every 10,000 steps. The first 10% of the samples were discarded as burn-in.

Literature review

A literature review on the spatial distribution of ticks and TBVs was conducted, and the resulting database was combined with our field surveys to achieve accurate and reliable model-predicted risk maps. The search terms included ('tick' or 'tick-borne') and ('China' or 'Gansu' or 'Inner Mongolia' or 'Ningxia' or 'Qinghai' or 'Shaanxi' or 'Xinjiang' or 'Afghanistan' or 'Kazakhstan' or 'Kyrgyzstan' or 'Mongolia' or 'Pakistan' or 'Tajikistan'). For the TBVs of interest, their names were added to the search terms using the operator 'or'. Studies published between January 1950 and October 2024 were searched in six major electronic databases: PubMed, Web of Science, Embase, China Wan Fang, China National Knowledge Infrastructure and the Chinese Scientific Journal Database (Supplementary Table 21). The studies identified were initially screened by titles and abstracts, and the full texts of those relevant to the topic were downloaded. Only studies that provided clear county-level locations (for provinces in China) or district-level locations (for neighbouring countries) of tick species or TBVs were included (Supplementary Fig. 28). Data were recorded using a standardised form that captured the study date, study location, coordinates, tick species, hosts and detection results for TBVs. In cases where more than one tick species was isolated from the same host in a single study, a separate record was created for each tick species. Similarly, if the same tick species was isolated from different hosts, a record was created for each host. If the same tick species was isolated from the same host in different years, a record was created for each year. Additionally, if multiple TBVs were found in the same tick species, a record was created for each TBV. However, if a tick species or TBV was reported more

than once in the same county or district within the same year, only one record was created in our database. Any ambiguous information reported in the studies was discussed among team members, and such records were excluded if the ambiguity could not be clarified. Each record in the database was reviewed independently by two team members. The included studies were provided in the Supplementary List of Included Studies.

Risk factors

Based on previous studies^{4,35,61–63}, both climatic and environmental risk factors significantly influence the distribution of tick species. Given that ticks are blood-feeding vectors, the distribution of host animals also plays a critical role in shaping tick distribution. As with TBVs, whose distribution is largely determined by the distribution of their major vectors, the aforementioned risk factors were also incorporated into the model for projecting TBV distribution. Additionally, because high-risk populations for TBVs are primarily farmers and individuals with a high probability of contact with ticks, two demographic risk factors, including population density and the proportion of the rural population among the total population, were included in the model for TBV distribution projection. These risk factors were used as independent variables in the BRT model.

Climatic data from 1980 and 2024 across six provinces in China (Gansu, Inner Mongolia, Ningxia, Qinghai, Shaanxi and Xinjiang) and six neighbouring countries (Afghanistan, Kazakhstan, Kyrgyzstan, Mongolia, Pakistan and Tajikistan) were sourced from the European Centre for Medium-Range Weather Forecasts⁶⁴. This data set provides high spatiotemporal resolution gridded climatic data from 1940 onwards, including daily mean temperature, maximum temperature, minimum temperature, dewpoint temperature and total precipitation (Supplementary Figs. 29–33). County-level averaged values of these variables were subsequently calculated. Given that bioclimatic predictors recommended by the U.S. Geological Survey better capture the seasonal trends than traditional meteorological variables^{65,66}, and have been widely used in ecological studies^{4,67}, 19 bioclimatic predictors were calculated, with their yearly averages used as risk factors in the ecological model. To avoid overfitting and enhance model interpretability, multicollinearity among these 19 candidate predictors was screened. A clustering analysis was performed based on their pairwise correlations. Specifically, a binary distance matrix was constructed, where the distance between any pair of variables was set to 0 if the absolute value of their Pearson correlation coefficient exceeded 0.8, and 1 otherwise^{67–70}. The optimal number of clusters was determined using the Krzanowski and Lai index (sum of these two indexes) via the hierarchical clustering method due to a small data size^{58–61}. From each cluster, only one predictor, the one with the lowest mean absolute value of correlation coefficients with all variables outside its cluster, was selected for modelling⁷¹. County-level averaged values of relative humidity were calculated using dewpoint and mean temperature, and these values were selected for inclusion in the modelling.

Global grid data on elevation at a spatial resolution of 30 m were downloaded from the Land Processes Distributed Active Archive Center⁷², and the spatial distribution of elevation is shown in Supplementary Fig. 34. We rescaled the elevation to a spatial resolution of 3000 m and then estimated the mean elevation at the county level for provinces in China and at the district level for neighbouring countries. Global grid data for land cover at a spatial resolution of 30 m were downloaded from the ESA WorldCover 2021 archive⁷³, and the spatial distribution of 11 land cover types is shown in Supplementary Fig. 35. We rescaled the land cover data to a spatial resolution of 3000 m and then estimated the proportion of each land cover type at the county level for provinces in China and at the district level for neighbouring countries.

Similarly, we downloaded global livestock grid data at a spatial resolution of 10 km from the Food and Agriculture Organization of the

United Nations⁷⁴. The spatial distribution of the average number of cattle, goats, horses and sheep per square kilometre is shown in Supplementary Figs. 36–39. We then estimated their average values at the county level for provinces in China and at the district level for neighbouring countries.

In total, 21 risk factors at the county or district level were used to model the spatial distribution of tick species. Additionally, these 21 risk factors, along with population density and the proportion of the rural population relative to total population downloaded from the online website City Population⁷⁵, were utilised to model the spatial distribution of TBVs (Supplementary Table 4).

Case-control design

A case-control study design was employed to classify counties or districts into binary categories, with their binary status used as dependent variables in the BRT model^{4,67}. Counties with at least one record of any tick species were classified as surveyed, while those without records were considered unsurveyed. Within the surveyed counties, those with records of the analysed tick species were considered as ‘case’ counties (labelled 1), while those without records were classified as ‘control’ counties (labelled 0). Unsurveyed counties were excluded from the model building but included in risk mapping. To address potential sampling bias for tick species, a logistic regression model was developed. All 21 risk factors mentioned above were incorporated as independent variables, while the binary status of all counties served as the dependent variable. To avoid overfitting, a logistic regression model was initially run, and the significance level of all risk factors was calculated to identify key risk factors, using a threshold of 0.05. These key risk factors were then used in the final logistic regression model to calculate the model-predicted sampling probability for each county (the likelihood of being selected as a survey location). The reciprocals of the predicted sampling probabilities of all surveyed counties were rescaled to have a mean of one and were used as weights in the BRT model to counterbalance the sampling bias for tick species.

A similar case-control design was applied for TBVs. Counties where TBVs were detected in ticks or where human cases of the associated tick-borne disease were reported were classified as ‘case’ counties. To address the potential bias for TBVs, strict criteria were used to define a control county. Given that the distribution of TBVs is largely shaped by the distribution of their major tick vectors, model-predicted high-risk counties for these tick species were considered when defining control counties for TBVs. Specifically, for TBVs with a single major tick vector, a ‘control’ county must meet the following conditions: (1) it is included in the list of surveyed counties; (2) it is not classified as a ‘case’ county; and (3) it is not in the list of model-predicted high-risk counties for the tick vector. For TBVs with multiple major tick vectors, a county was classified as a ‘control’ if it was designated as a ‘control’ county for any tick vector, according to the criteria applied to viruses with a single tick vector.

Boosted regression tree model

A tree complexity of 5, a learning rate of 0.005 and a bag fraction of 75% have been widely used in studies due to their satisfactory performance^{4,35}. A grid search was conducted separately, considering tree complexity values of [3, 5, 7], learning rates of [0.001, 0.005, 0.010] and bag fractions of [0.7, 0.75, 0.8], using negative mean squared error to identify the optimal parameters for each tick species and TBV. These ranges were selected based on the aforementioned parameter set as a reference. Given the computational constraints—specifically, the substantial number of models required ([4 tick species + 4 TBVs] × 27 × 5)—it was not feasible to optimise all model parameters with smaller intervals. Following this, the BRT model was configured with the optimal tree complexity, learning rate and bag

fraction, along with a fixed number of trees set to 3000 and 'binary: logistic' objective using XGBoost. This model was subsequently employed to predict the probabilities of occurrence of ticks and TBVs across all counties within the studied area, as well as the relative contribution of risk factors. Note that a higher relative contribution for a certain risk factor only indicates a stronger influence on the response of the BRT model; the partial dependence plots were provided to visualise the effect of a variable on the fitted function after accounting for the average effects of all other variables in the model. A two-stage bootstrapping procedure was implemented to obtain a robust and parsimonious estimation. In each stage, the split-and-fit process was repeated several times. A training set comprising 75% of data points was randomly selected, while the remaining 25% served as a test set. In the first stage, the split-and-fit process was repeated 10 times to identify the key risk factors, defined as those with a mean relative contribution greater than 2%^{4,67}. In the second stage, only key risk factors were included, and the split-and-fit process was repeated 100 times. For each simulation, the predicted probabilities of occurrence across all counties for each tick species and TBV, along with the testing AUC, sensitivity and specificity, were recorded. The averaged predicted probabilities were used to generate the figure, while the mean and 95% confidence interval for the other three metrics were also reported. A dynamic cut-off value maximizing the sum of sensitivity and specificity along the receiver operating characteristic curve was chosen for each simulation to determine model-predicted high-risk counties for tick species. A constant cut-off probability of 50% was used to identify model-predicted high-risk counties for TBVs.

Robustness analysis of the modelling results

Climatic data spanning from 1980 to 2024, along with all tick species and TBVs distribution data, were utilised to develop the ecological model. To address potential bias arising from the variability of these variables, two additional analyses were conducted: one for the period before 2010 (using climatic data from 1980 to 2009 and distribution data recorded before 2009) and another for the period after 2010 (using climatic data from 2010 to 2024 and distribution data recorded after 2010). The daily mean temperature, maximum temperature, minimum temperature, dewpoint temperature and total precipitation exhibited temporal variations, resulting in distinct sets of the 19 bioclimatic predictors for the periods 1980–2024, 1980–2009 and 2010–2024. Consequently, the clusters identified and the selected representative bioclimatic predictors varied across analyses (Supplementary Table 12). Projections for the four tick species were conducted for both periods (Supplementary Table 13). However, due to data availability constraints, only CCHFV was modelled for the period before 2010, while all four TBVs were included for the period after 2010 (Supplementary Table 19).

Ethics statement

The study was approved by the ethics committee of the First Affiliated Hospital of Shihezi Medical University, including the procedures and protocols of specimen collection and processing. Informed consent was obtained from the patient involved (No. A2018-144-01). One tick (*Hy. asiaticum*) was collected from a patient, who was a woman living in Manas County, XUAR. On September 26, 2019, she noticed a tick on the right lower leg.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The sequence reads generated in this study have been deposited in the NCBI Sequence Read Archive (SRA) database under BioProject accession number PRJNA1190384 (SRA accession numbers:

SRR32560802- SRR32561316) [<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1190384>]. The assembled sequences are available at GenBank under accession numbers PV023382-PV023575. The study data is also available in an openly available repository, Figshare [<https://doi.org/10.6084/m9.figshare.28831238>]. All data required to develop the BRT model in this study are provided in the Supplementary Data or are accessible via the provided source links.

Code availability

The codes used in this study are archived on Zenodo with the <https://doi.org/10.5281/zenodo.17385901>. The corresponding GitHub repository can be found at <https://github.com/HengcongLiu/NorthWestBRTmodel>.

References

- Dantas-Torres, F., Chomel, B. B. & Otranto, D. Ticks and tick-borne diseases: a One Health perspective. *Trends Parasitol.* **28**, 437–446 (2012).
- Fang, L. Q. et al. Emerging tick-borne infections in mainland China: an increasing public health threat. *Lancet Infect. Dis.* **15**, 1467–1479 (2015).
- Shi, J., Hu, Z., Deng, F. & Shen, S. Tick-borne viruses. *Viol. Sin.* **33**, 21–43 (2018).
- Zhao, G. P. et al. Mapping ticks and tick-borne pathogens in China. *Nat. Commun.* **12**, 1075 (2021).
- Ergonul, O. Crimean-Congo haemorrhagic fever. *Lancet Infect. Dis.* **6**, 203–214 (2006).
- Yu, X. J. et al. Fever with thrombocytopenia associated with a novel bunyavirus in China. *N. Engl. J. Med.* **364**, 1523–1532 (2011).
- Wang, Z. D. et al. A new segmented virus associated with human febrile illness in China. *N. Engl. J. Med.* **380**, 2116–2125 (2019).
- Balakrishnan, V. S. WHO launches global initiative for arboviral diseases. *Lancet Microbe* **3**, e407 (2022).
- The World Health Organization Team. Pathogens prioritization: a scientific framework for epidemic and pandemic research preparedness. *R&D Blue Print*. 1–38 (2024).
- Harvey, E. et al. Extensive diversity of RNA viruses in Australian ticks. *J. Virol.* **93**, e01358-18 (2019).
- Kong, Y. et al. Metatranscriptomics reveals the diversity of the tick virome in Northwest China. *Microbiol. Spectr.* **10**, e0111522 (2022).
- Ni, X. B. et al. Metavirome of 31 tick species provides a compendium of 1,801 RNA virus genomes. *Nat. Microbiol.* **8**, 162–173 (2023).
- Zhang, Y. et al. Viromes and surveys of RNA viruses in camel-derived ticks revealing transmission patterns of novel tick-borne viral pathogens in Kenya. *Emerg. Microbes Infect.* **10**, 1975–1987 (2021).
- Sheng, J. et al. Tick distribution in border regions of Northwestern China. *Ticks Tick Borne Dis.* **10**, 665–669 (2019).
- Fereidouni, M. et al. Crimean-Congo hemorrhagic fever virus in Central, Eastern, and South-eastern Asia. *Viol. Sin.* **38**, 171–183 (2023).
- Liu, X. et al. A Tentative Tamdy Orthonairovirus related to febrile illness in Northwestern China. *Clin. Infect. Dis.* **70**, 2155–2160 (2020).
- Dong, Z. et al. Human Tacheng tick virus 2 infection, China, 2019. *Emerg. Infect. Dis.* **27**, 594–598 (2021).
- L'vov, D. K. et al. Taxonomic status of the Burana virus (BURV) (Bunyaviridae, Nairovirus, Tamdy group) isolated from the ticks *Haemaphysalis punctata* Canestrini et Fanzago, 1877 and *Haem. concinna* Koch, 1844 (Ixodidae, Haemaphysalinae) in Kyrgyzstan. *Vopr. Virusol.* **59**, 10–15 (2014).
- Fuchs, J. et al. Comparative study of ten Thogotovirus isolates and their distinct in vivo characteristics. *J. Virol.* **96**, e0155621 (2022).
- L'vov DK SG, et al. Virus "Tamdy"-a new arbovirus, isolated in the Uzbee S.S.R. and Turkmen S.S.R. from ticks *Hyalomma asiaticum*

- asiaticum Schulee et Schlottke, 1929, and *Hyalomma plumbeum plumbeum* Panzer, 1796. *Arch. Virol.* **51**, 15–21 (1976).
21. Zhou, H. et al. Tamdy Virus in Ixodid ticks infesting Bactrian camels, Xinjiang, China, 2018. *Emerg. Infect. Dis.* **25**, 2136–2138 (2019).
 22. Moming, A. et al. Evidence of human exposure to Tamdy Virus, Northwest China. *Emerg. Infect. Dis.* **27**, 3166–3170 (2021).
 23. Cui, M. et al. Serological evidence of Bactrian camel infection with Tamdy Virus, Xinjiang, China. *Vector Borne Zoonotic Dis.* **24**, 842–845 (2024).
 24. Alkhovsky, S. V. et al. Complete genome coding sequences of Artashat, Burana, Caspiy, Chim, Geran, Tamdy, and Uzun-Agach viruses (Bunyvirales: Nairoviridae: Orthonairovirus). *Genome Announc* **5**, e01098–17 (2017).
 25. Moore, D. L. et al. Arthropod-borne viral infections of man in Nigeria, 1964–1970. *Ann. Trop. Med. Parasitol.* **69**, 49–64 (1975).
 26. Yang, M. et al. *Rickettsia aeschlimannii* Infection in a woman from Xinjiang, Northwestern China. *Vector Borne Zoonotic Dis.* **22**, 55–57 (2022).
 27. Matsuno, K. et al. Characterization of the Bhanja serogroup viruses (Bunyviridae): a novel species of the genus Phlebovirus and its relationship with other emerging tick-borne phleboviruses. *J. Virol.* **87**, 3719–3728 (2013).
 28. Gomer, A., Lang, A., Janshoff, S., Steinmann, J. & Steinmann, E. Epidemiology and global spread of emerging tick-borne Alongshan virus. *Emerg. Microbes Infect.* **13**, 2404271 (2024).
 29. Dedkov, V. G. et al. Isolation and characterization of Wad Medani virus obtained in the tuva Republic of Russia. *Ticks Tick Borne Dis.* **12**, 101612 (2021).
 30. Yadav, P. D. et al. Characterization of novel reoviruses Wad Medani Virus (Orbivirus) and Kundal Virus (Coltivirus) collected from *Hyalomma anatolicum* ticks in India during surveillance for Crimean Congo Hemorrhagic Fever. *J. Virol.* **93**, e00106–19 (2019).
 31. Semper, A. E. et al. Research and product development for Crimean-Congo haemorrhagic fever: priorities for 2024–30. *Lancet Infect. Dis.* **25**, e223–e234 (2024).
 32. Hubalek, Z. Biogeography of tick-borne bhanja virus (bunyviridae) in Europe. *Interdiscip. Perspect. Infect. Dis.* **2009**, 372691 (2009).
 33. Calisher, C. H. & Goodpasture, H. C. Human infection with Bhanja virus. *Am. J. Trop. Med. Hyg.* **24**, 1040–1042 (1975).
 34. Shah, K. V. & Work, T. H. Bhanja virus: a new arbovirus from ticks *Haemaphysalis intermedia* Warburton and Nuttall, 1909, in Orissa, India. *Indian J. Med. Res.* **57**, 793–798 (1969).
 35. Liu, K. et al. A national assessment of the epidemiology of severe fever with thrombocytopenia syndrome. *China Sci. Rep.* **5**, 9679 (2015).
 36. Tufts, D. M. et al. Distribution, host-seeking phenology, and host and habitat associations of *Haemaphysalis longicornis* Ticks, Staten Island, New York, USA. *Emerg. Infect. Dis.* **25**, 792–796 (2019).
 37. Herrmann, C., Voordouw, M. J. & Gern, L. *Ixodes ricinus* ticks infected with the causative agent of Lyme disease, *Borrelia burgdorferi sensu lato*, have higher energy reserves. *Int. J. Parasitol.* **43**, 477–483 (2013).
 38. Kholodilov, I. et al. Ixodid ticks and tick-borne encephalitis virus prevalence in the South Asian part of Russia (Republic of Tuva). *Ticks Tick Borne Dis.* **10**, 959–969 (2019).
 39. Jaenson, T. G. et al. First evidence of established populations of the taiga tick *Ixodes persulcatus* (Acari: Ixodidae) in Sweden. *Parasit. Vectors* **9**, 377 (2016).
 40. Shchuchinova, L. D., Kozlova, I. V. & Zlobin, V. I. Influence of altitude on tick-borne encephalitis infection risk in the natural foci of the Altai Republic, Southern Siberia. *Ticks Tick Borne Dis.* **6**, 322–329 (2015).
 41. Zhao, L. et al. Distribution of *Haemaphysalis longicornis* and associated pathogens: analysis of pooled data from a China field survey and global published data. *Lancet Planet. Health* **4**, e320–e329 (2020).
 42. Bai, Y. et al. Discovery of Tick-Borne Karshi virus implies misinterpretation of the Tick-Borne Encephalitis Virus Seroprevalence in Northwest China. *Front. Microbiol.* **13**, 872067 (2022).
 43. Perfilyeva, Y. V. et al. Tick-borne pathogens and their vectors in Kazakhstan—a review. *Ticks Tick Borne Dis.* **11**, 101498 (2020).
 44. Sang, C. et al. Tick distribution and detection of *Babesia* and *Theileria* species in Eastern and Southern Kazakhstan. *Ticks Tick Borne Dis.* **12**, 101817 (2021).
 45. Wang, Y. et al. Identification and phylogenetic analysis of Nairobi sheep disease virus from *Haemaphysalis longicornis* ticks in Shandong Province, China. *Ticks Tick Borne Dis.* **15**, 102375 (2024).
 46. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
 47. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
 48. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
 49. Grabherr, M. G. et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
 50. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).
 51. Li, N. et al. Nationwide genomic surveillance reveals the prevalence and evolution of honeybee viruses in China. *Microbiome* **11**, 6 (2023).
 52. Mahar, J. E., Shi, M., Hall, R. N., Strive, T. & Holmes, E. C. Comparative analysis of RNA virome composition in rabbits and associated ectoparasites. *J. Virol.* **94**, e02119–19 (2020).
 53. Shi, M. et al. The evolutionary history of vertebrate RNA viruses. *Nature* **556**, 197–202 (2018).
 54. Pettersson, J. H. et al. Circumpolar diversification of the *Ixodes uriae* tick virome. *PLoS Pathog.* **16**, e1008759 (2020).
 55. Wille, M. et al. Sustained RNA virome diversity in Antarctic penguins and their ticks. *ISME J.* **14**, 1768–1782 (2020).
 56. Xu, L. et al. Tick virome diversity in Hubei Province, China, and the influence of host ecology. *Virus Evol.* **7**, veab089 (2021).
 57. Babaian, A. & Edgar, R. Ribovirus classification by a polymerase barcode sequence. *PeerJ* **10**, e14055 (2022).
 58. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
 59. Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
 60. Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574 (2003).
 61. Gilbert, L., Aungier, J. & Tomkins, J. L. Climate of origin affects tick (*Ixodes ricinus*) host-seeking behavior in response to temperature: implications for resilience to climate change? *Ecol. Evol.* **4**, 1186–1198 (2014).
 62. Lambin, E. F., Tran, A., Vanwambeke, S. O., Linard, C. & Soti, V. Pathogenic landscapes: interactions between land, people, disease vectors, and their animal hosts. *Int. J. Health Geogr.* **9**, 54 (2010).
 63. Zhang, Y. K., Zhang, X. Y. & Liu, J. Z. Ticks (Acari: Ixodoidea) in China: Geographical distribution, host diversity, and specificity. *Arch. Insect Biochem. Physiol.* **102**, e21544 (2019).
 64. Climate Data Store. ERA5 post-processed daily statistics on single levels from 1940 to present. <https://cds.climate.copernicus.eu/datasets/derived-era5-single-levels-daily-statistics?tab=download> (Accessed 20 March 2025).

65. O'Donnell, M. S. & Ignizio, D. A. Bioclimatic predictors for supporting ecological applications in the conterminous United States. *Data Series No. 691, U.S. Geol. Surv.* (2012).
66. Amiri, M., Tarkesh, M., Jafari, R. & Jetschke, G. Bioclimatic variables from precipitation and temperature records vs. remote sensing based bioclimatic variables: Which side can perform better in species distribution modeling?. *Ecol. Inform.* **57**, 101060 (2020).
67. Chen, J. J. et al. Small mammals and associated infections in China: a systematic review and spatial modelling analysis. *Lancet Reg. Health West Pac.* **54**, 101264 (2025).
68. Miao, D. et al. Epidemiology and ecology of severe fever with Thrombocytopenia Syndrome in China, 2010–2018. *Clin. Infect. Dis.* **73**, e3851–e3858 (2021).
69. Wang, T. et al. Mapping the distributions of blood-sucking mites and mite-borne agents in China: a modeling study. *Infect. Dis. Poverty* **11**, 41 (2022).
70. Wang, T. et al. Mapping the distributions of mosquitoes and mosquito-borne arboviruses in China. *Viruses* **14**, 691 (2022).
71. Teng, A. Y. et al. Mapping the viruses belonging to the order Bunyavirales in China. *Infect. Dis. Poverty* **11**, 81 (2022).
72. Land Processes Distributed Active Archive Center (LP DAAC). ASTGTM v003. <https://opendap.cr.usgs.gov/opendap/hyrax/ASTER/ASTT/ASTGTM.003/2000.03.01/contents.html> (Accessed 20 October 2024).
73. European Space Agency. ESA WorldCover Project 2021. <https://worldcover2021.esa.int/> (Accessed 20 October 2024).
74. Food and Agriculture Organization of the United Nations. Livestock Systems-Global distributions. <https://www.fao.org/livestock-systems/global-distributions/en/> (Accessed 20 October 2024).
75. City Population. City population. <https://www.citypopulation.de/> (Accessed 20 October 2024).

Acknowledgements

The authors would like to thank Mengchan Hao, Yanhai Wang, Xiaoqing Zhang, Jiaying Wu, Sanling Fan, Juefu Hu and Lijia Jia for their contributions to sample collection, and Xiurong Wang and Yu Cao for their laboratory assistance. This study was supported by grants from the National Natural Science Foundation of China for Distinguished Young Scholars (32325003 to W.S.), National Natural Science Foundation of China (82202516 and 82472276 to H.Z., 82260399 to Y.-Z.W. and 31970174 to J.C.), Natural Science Key Project of Xinjiang Uygur Autonomous Region (2022B03014 to Y.-Z.W.), National Key Research and Development Program of China (2023YFC260550 to J.C., and 2022YFC2304004 to Y.-Z.W.), Taishan Scholars Programme of Shandong Province (tsqn202306264 to H.Z.), Joint Innovation Team for Clinical & Basic Research (202407 to W.S., H.Z., J.L. and J.J.) and AI for Science Program from Shanghai Municipal Education Commission (JWAIZD-3 to W.S.).

Author contributions

W.S., Y.-Z.W. and H.Z. conceived the study. H.Z., Y.-X.W., S.Z., M.G., M.C., S.W., X.Z. and J.Q. did the experiments. H.Z., Y.-X.W., S.Z., M.G., J.C., M.C., S.W., X.Z., J.Q., Z.L., Z.M., R.Z. and B.Z. collected and pre-treated the samples. W.S., H.L., H.Z., Y.-X.W., J.L., J.J., W.Z. and T.W. analysed the data. H.Z., H.L., Y.-X.W. and Y.-Z.W. accessed and verified the data. W.S., H.Z., H.L. and Y.-X.W. wrote the paper. Y.-Z.W. and E.C.H. revised the paper. All authors reviewed the final draft and agreed with its content and conclusions, had full access to all the data in the study, and had final responsibility for the decision to submit for publication.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-66873-8>.

Correspondence and requests for materials should be addressed to Hong Zhou, Yuan-Zhi Wang or Weifeng Shi.

Peer review information *Nature Communications* thanks Jiří Černý and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025

¹School of Clinical and Basic Medical Sciences, Shandong First Medical University & Shandong Academy of Medical Sciences, Ji'nan, China. ²Key Laboratory of Emerging Infectious Diseases in Universities of Shandong, Shandong First Medical University & Shandong Academy of Medical Sciences, Ji'nan, China. ³School of Public Health, Shandong First Medical University & Shandong Academy of Medical Sciences, Ji'nan, China. ⁴Shanghai Institute of Virology, Shanghai Jiao Tong University School of Medicine, Shanghai, China. ⁵Key Laboratory for Prevention and Control of Emerging Infectious Diseases and Public Health Security, the XPCC, School of Medicine, Shihezi University, Shihezi, China. ⁶NHC Key Laboratory of Prevention and Treatment of Central Asia High Incidence Diseases, School of Medicine, Shihezi University, Shihezi, China. ⁷Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China. ⁸CAS Key Laboratory of Special Pathogens, Wuhan Institute of Virology, Center for Biosafety Mega-Science, Chinese Academy of Sciences, Wuhan, China. ⁹Shandong Public Health Clinical Center, Shandong University, Ji'nan, China. ¹⁰College of Life Science and Technology, Xinjiang University, Urumqi, China. ¹¹College of Life Sciences and Technology, Inner Mongolia Normal University, Hohhot, China. ¹²School of Life Sciences, Shandong First Medical University & Shandong Academy of Medical Sciences, Taian, China. ¹³School of Medical Sciences, The University of Sydney, Sydney, Australia. ¹⁴School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China. ¹⁵These authors contributed equally: Hong Zhou, Hengcong Liu, Yun-Xiao Wang, Shanshan Zhao. ✉ e-mail: zhouh@sdfmu.edu.cn; wangyuanzhi621@126.com; shiwf@ioz.ac.cn