

OPEN

Epidemiology, evolution and transmission of human metapneumovirus in Guangzhou China, 2013–2017

Lina Yi^{1,2}, Lirong Zou¹, Jingju Peng³, Jianxiang Yu¹, Yingchao Song¹, Lijun Liang¹, Qianfang Guo¹, Min Kang¹, Changwen Ke¹, Tie Song¹, Jing Lu^{1,2,3} & Jie Wu¹

Human metapneumovirus (hMPV), first identified in 2001, is a major viral respiratory pathogen that worldwide reported. Fundamental questions concerning the dynamics of viral evolution and transmission at both regional and global scales remain unanswered. In this study, we obtained 32 G gene and 51 F gene sequences of hMPV in Guangzhou, China in 2013–2017. Temporal and spatial phylogenetic analyses were undertaken by incorporating publicly available hMPV G gene (978) and F gene (767) sequences. The phylogenetic results show different global distribution patterns of hMPV before 1990, 1990–2005, and 2006–2017. A sharply increasing hMPV positive rate (11%) was detected in Guangzhou 2017, mainly caused by the B1 lineage of hMPV. A close phylogenetic relation was observed between hMPV strains from China and Japan, suggesting frequent hMPV transmissions between these regions. These results provide new insights into hMPV evolution, transmission, and spatial distribution and highlight Asia as a new epicenter for viral transmission and novel variant seeding after the year 2005. Conducting molecular surveillance of hMPV in Asian countries is critical for understanding the global circulation of hMPV and future vaccine design.

Human metapneumovirus (hMPV), first identified in 2001, is a major viral respiratory pathogen that worldwide reported¹. It is estimated that 4–16% of acute respiratory tract infections are caused by hMPV^{2,3}. Retrospective serologic studies demonstrated 100% seroprevalence of hMPV antibodies among children ≤5 years of age⁴. Moreover, reinfection can occur throughout life^{5,6}, further contributing to the prevalence of hMPV. Clinical features of hMPV infection are similar to those associated with human respiratory syncytial virus infection, ranging from mild respiratory illnesses to severe bronchiolitis and pneumonia^{4,7}. Children under five years, aged adults, and patients with underlying diseases are more likely to develop severe diseases⁸.

hMPV is a single stranded, negative sense and enveloped RNA virus. The viral RNA is approximately 13 kb in length, containing eight genes coding for nine proteins⁹. The fusion (F) and attachment (G) proteins are two major hMPV surface glycoproteins that have played important roles in viral replication and host immune response¹⁰ and have been largely used to study hMPV genetic variation^{11–13}. Based on the genetic variability of the F and G genes, hMPV is classified into two main genetic groups, A and B, which are further divided into five lineages known as A1, A2a, A2b, B1, and B2^{14,15}. There may no direct association between the genotypes and the clinical courses¹⁶. However, antigenic variance is observed among different genetic lineages of hMPV in animal models and human infection cases^{10,17}.

The genetic evolution and transmission of hMPV are critical to epidemic control but have not been thoroughly studied as yet. Recently, infections of hMPV were increasingly reported in Asian countries^{6,18–24}, with novel variants of hMPV emerging^{23,24}, highlighting the risks of hMPV epidemics in these regions. Here, we conducted a five-year study on hMPV in Guangzhou, the most populous city in China, from January 2013 to December 2017.

¹Guangdong Provincial Center for Disease Control and Prevention, No. 160, Qunxian Road, Panyu District, Guangzhou, China. ²Guangdong Provincial Institute of Public Health, No. 160, Qunxian Road, Panyu District, Guangzhou, China. ³Southern Medical University, No. 1838, Shatai Road, Baiyun District, Guangzhou, People's Republic of China. Lina Yi and Lirong Zou contributed equally. Jing Lu and Jie Wu jointly supervised this work. Correspondence and requests for materials should be addressed to J.L. (email: Jimlu0331@gmail.com) or J.W. (email: wujie55@hotmail.com)

Through phylogeographic analyses, the spatial and evolution dynamics of hMPV were inferred by integrating the publicly available data and the newly generated sequences in this study.

Materials and Methods

Ethics statement. This study was approved by the Ethics Review Committee of the Guangdong Provincial Center for Disease Control and Prevention. We confirm that all methods were performed in accordance with the relevant guidelines and regulations. Respiratory samples were collected for public health purposes. All patients or their guardian(s) were informed about the study before providing written consent, and data were anonymized for analysis.

Clinical sample collection. Guangzhou in southern China is the country's largest city with a population of over 13 million in 2016. Observation was conducted in a sentinel hospital in Guangzhou through a Guangdong provincial respiratory surveillance system as previously described^{25,26}. Patients suspected of having acute respiratory tract infections were enrolled according to these criteria: acute fever ($T \geq 38^\circ\text{C}$) and/or abnormal leukocyte count with any one respiratory symptom (such as sore throat, cough, expectoration, or dyspnoea/tachypnoea). A total of 1,460 patients were enrolled from January 2013 to December 2017. Nasopharyngeal swabs (NPSs) were collected within 24 h after admission. Specimens were stored in 3 mL viral transport medium at -70°C until analysis.

Detection of human metapneumovirus. Total viral nucleic acids were extracted from 140 μl specimen and eluted with 60 μl water by using a QIAamp Viral RNA Mini kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. hMPV infection was detected by using real-time RT-PCR with a QIAGEN OneStep RT-PCR Kit. The hMPV RT-PCR primer set was used as previously described²⁷: hMPV-F: 5'-CGTCAGCTTCAGTCAATTCAACAGA-3', hMPV-R: 5'-ATTARGTCCAADGATATTGCTGGTGT-3' and hMPV-Probe: FAM-CTGCATTGTCTGAAAAYTGCCGCACAACATT-BHQ1. The thermocycling protocol was reverse transcription at 45°C for 10 min, followed by 10 min at 95°C , 40 cycles of 95°C 15 s, and 55°C 60 s.

PCR amplification of the full-length F and G protein genes. For hMPV positive samples, the viral RNA was reverse transcribed into cDNA using a SuperScript™ IV First-Strand Synthesis System (Invitrogen, USA). Then the full length of F and G genes was amplified using traditional nested PCR with primers newly designed in this study. Primers used for G gene amplification were: hMPV-G-F: 5'-GAGAACATTCGRRC RATWGAYATG -3', hMPV-G-R1: 5'-AGATARACATTRACAGTRGAYTCA-3', and hMPV-G-R2: 5'-CAAYAACA GGGTTTTCYAYAGCA-3'. The primer sets for F gene were: outer primers set hMPV-F-F1-5'-CAATGCA GGTATAACACCAGCAATATC-3' and hMPV-F-R1 5'-GCAACAATTGAACTGATCTTCAGGAAAC-3', and inner primers set hMPV-F-F2-5'-ACATGCCAACATCTGCAGGACAAATAAAC-3' and hMPV-F-R2 5'-ACATG CTGTTACCTTCAACTTTGC-3'. The PCR products were purified and sequenced using an ABI3730xl DNA Analyzer at IGE Biotech Co., Ltd. (Guangzhou, China). In total, 51 and 32 samples (out of the total 160 RT-PCR positive samples) F and G genes were sent for sequencing, respectively. Sequences generated in this study have been submitted to GenBank (accession numbers MK450165-MK450246).

Sequence alignment and maximum-likelihood phylogenetic analysis. The hMPV sequences generated in this study were combined with all publicly available hMPV G and F genes sequences with known sampling dates and locations in GenBank. Partial sequences (<80%) of the G and F genes were excluded, and identical sequences collected in the same sampling location on the same date were removed to improve computation time. In total, 978 G gene sequences (568 genogroup-A and 410 genogroup-B) and 767 F gene sequences (365 genogroup-A and 402 genogroup-B) were included in the phylogenetic analysis (Table S1) with sequences generated in this study. Multiple sequence alignment was performed using ClustalW, and alignments were minimally edited by hand using Aliview²⁸. Maximum-likelihood (ML) trees were estimated for G and F genes in RaxML²⁹. Temporal accumulation of genetic divergence was assessed from maximum-likelihood, mid-point-rooted phylogenies using the linear regression approach implemented in TempEst (formerly Path-O-Gen)³⁰.

Phylogeographic analysis. Compared to the F gene, more hMPV G gene sequences from different regions were available in the public database. Therefore, we performed the phylogeographic analysis on genogroup-A and genogroup-B G gene sequences separately to provide an outline of hMPV distribution and transmission. As previously described^{25,31,32}, Bayesian Markov chain Monte Carlo (MCMC) phylogenetic inference was performed using BEAST under a SRD06 nucleotide substitution model and a GMRF Bayesian skyride coalescent model³³. Preliminary analysis indicated high values for the coefficient of variation parameter of the molecular clock model, therefore an uncorrelated lognormal (UCLD) relaxed clock model was used in the final analysis to accommodate variation in substitution rates among branches³⁴. Three independent MCMC runs of 1×10^8 steps were computed and 10–20% burn-in was discarded from each, resulting in a total of 2.0×10^8 total steps. Model parameters and trees were sampled every 10,000 MCMC steps. Convergence and behavior of MCMC chains was inspected using Tracer v1.6 (<http://tree.bio.ed.ac.uk>). A subset of 1,000 trees was randomly drawn from the combined posterior distribution of trees and used as an empirical distribution for subsequent phylogeographic analysis³⁵.

Results

Epidemiology of hMPV in southern china. A total of 1,460 ARI patients (927 males and 533 females) were enrolled in Guangzhou from January 2013 to December 2017 (Fig. 1). HMPV was detected in 76 samples (5.2%), including 50 males and 26 females (1.92:1). No significant gender difference was observed ($p = 0.669$). The age of enrolled patients ranged from 1 month to 88 years, with the majority (82.7%) being children no more than 5 years (median age: 11 months, IQR, 0.3–3 years). HMPV positive cases were aged from 1 month to 81 years (median age: 1 year, IQR, 0.4–2.5 years). Half (52.6%) of infected individuals were children no more than 1 year. The distribution of HMPV prevalence among age groups was significantly different ($P = 0.016$), and children aged 1 to 2 years

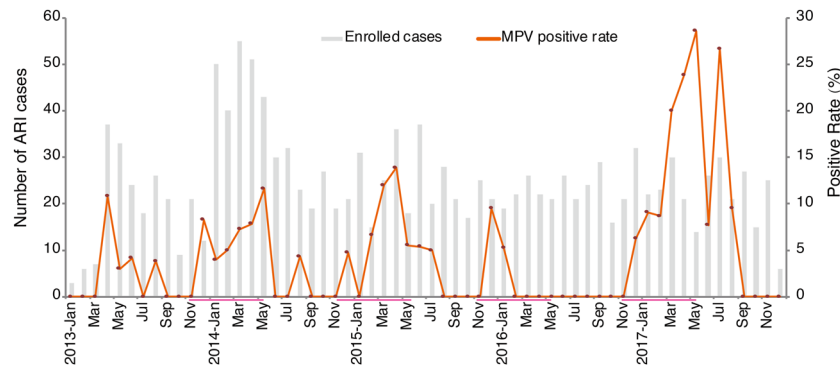


Figure 1. Month distribution of acute respiratory infection cases and hMPV positive rates in Guangzhou, 2013–2017. Duration from November to the next May is highlight to show the seasonality of hMPV epidemic.

had the highest prevalence (9.1%, 15/164). The clinical presentations of hMPV infected cases included fever (69, 90.8%), cough (69, 90.8%), sore throat (38, 50.0%), expectoration (16, 90.8%), shortness of breath (10, 13.2%) and nasal discharge (6, 7.9%). The seasonal distribution showed that hMPV circulated predominantly in the period from December of the first year to May of the next (Fig. 1). In the years 2013–2015, the hMPV detection rate was approximately even, ranging from 3.7% to 5.1%, similar with the rate detected in central and eastern China. An apparent low frequency of circulation of hMPV in the year 2016 was observed, with a low positive rate of 1.1%. However, in the year 2017, hMPV showed a remarkable increase, viruses being detected in 31 ARI cases (11.9%).

Genetic divergence and lineage classification of hMPV. To understand the lineage distribution and the cause(s) for the increase of the hMPV epidemic in 2017, 32 G genes and 51 F genes from 2013 through 2017 were sequenced. The evolution of hMPV at the global level was investigated by integrating all other publicly available G gene and F gene sequences (Materials and Methods). Due to the great genetic distance observed between two genogroups of hMPV (nucleotide identity of G gene between A and B groups is less than 57%), sequences from A and B genogroups were analyzed separately (Figs 2–4).

The hMPV A genogroup can be divided into four lineages (A1, A2a, A2b, A2c) based on the phylogeny topology (Figs 2a and 4a) and genetic diversity (Fig. 2d). The heat map of identity matrix clearly showed A1 strains were divergent from all A2 lineages (identity less than 85.2%), but similar genetic distances or identities were observed among A2a, A2b, and A2c (Fig. 2d, Table 1). Therefore, A2c was named as a novel A2b lineage in a previous study²³ but marked as a novel lineage in this study due to the comparable genetic identities between each pair of the A2 sub-lineages (Fig. 2d, Table 1). For instances, the pairwise nucleotide identities between strains from A2b and A2c lineages were range from 87.5–93.3% which was similar with that observed between strains from A2a and A2b (86.0–93.3%). Within lineage, the divergent patterns showed by the G genes existed according to the time, seemingly having little correlation with the isolation date (Fig. 2d).

Global distribution of hMPV lineages. The phylogeographic analysis of the G gene showed a spatial pattern of hMPV lineages. Strains near the root of genogroup A primarily circulated in America (Fig. 2a). Although an American origin for the genogroup A of hMPV seems the most probable (posterior probability = 1.0), this might also be caused by sampling biases. Between 1990 and 2005, hMPV infections began to be detected in other continents. Notably, regional clusters of hMPV were observed in lineages A1 and A2a (Fig. 2a), suggesting the local circulation of these hMPV viruses. However, these circulations were always temporal as the regional phylogenetic clusters in A1 and A2a lineages were only persistent during 2–3 years (Fig. 2a).

In contrast, a cluster in A2b, early detected in America, dominantly circulated in Asian countries after 2005 when the hMPV surveillances started in these regions^{22,36,37}. This cluster was persistent in Asia for more than 10 years, and the transmissions of this cluster of strains from Asia to other continents were observed. More obviously, the young lineage A2c most likely originated around 2001 (1997–2003, 95%HPD) and was well established in Asian countries. Current surveillance data showed A2c was now the major epidemic lineage in genogroup A. A similar pattern was observed for the F gene (Fig. 4). The A2c lineage of hMPV was endemic in Asia as the F gene phylogeny and introductions of A2c strains from Asia into other continents were observed after 2010 (Fig. 4a). In genogroup B, the sequences near the root of the phylogenetic tree were also mainly from America (Figs 3a and 4b). Lineages B1 and B2 of hMPV were established in Asian countries after 2010, and thereafter hMPV transmissions between Asia and Europe were frequently observed (Fig. 5).

Local transmission of hMPV. The continental-specific distribution was observed for hMPV genetic lineages (Fig. 5). Then, the next intriguing question became how hMPV circulated at a country level. In particular, the increasing epidemic of hMPV in Guangzhou in 2017 was caused by the local strains circulating in previous seasons or the strains being imported from another region. The phylogenetic analyses of hMPV strains isolated in Guangzhou 2013–2017 showed the predominant lineage was varied in different years (Figs 2a and 3a). Based on the phylogeny of F genes, 26, 1, 15, and 10 strains were classified into A2c, A2b, B1, and B2 lineages, respectively (Fig. 4). In 2017, we identified one hMPV A2c strain with 111nt-dup in the G gene (P17072_GD-CHN_110ntDup_2017, Fig. 2b). This strain is closely related with the 111nt-dup hMPV novel variant most

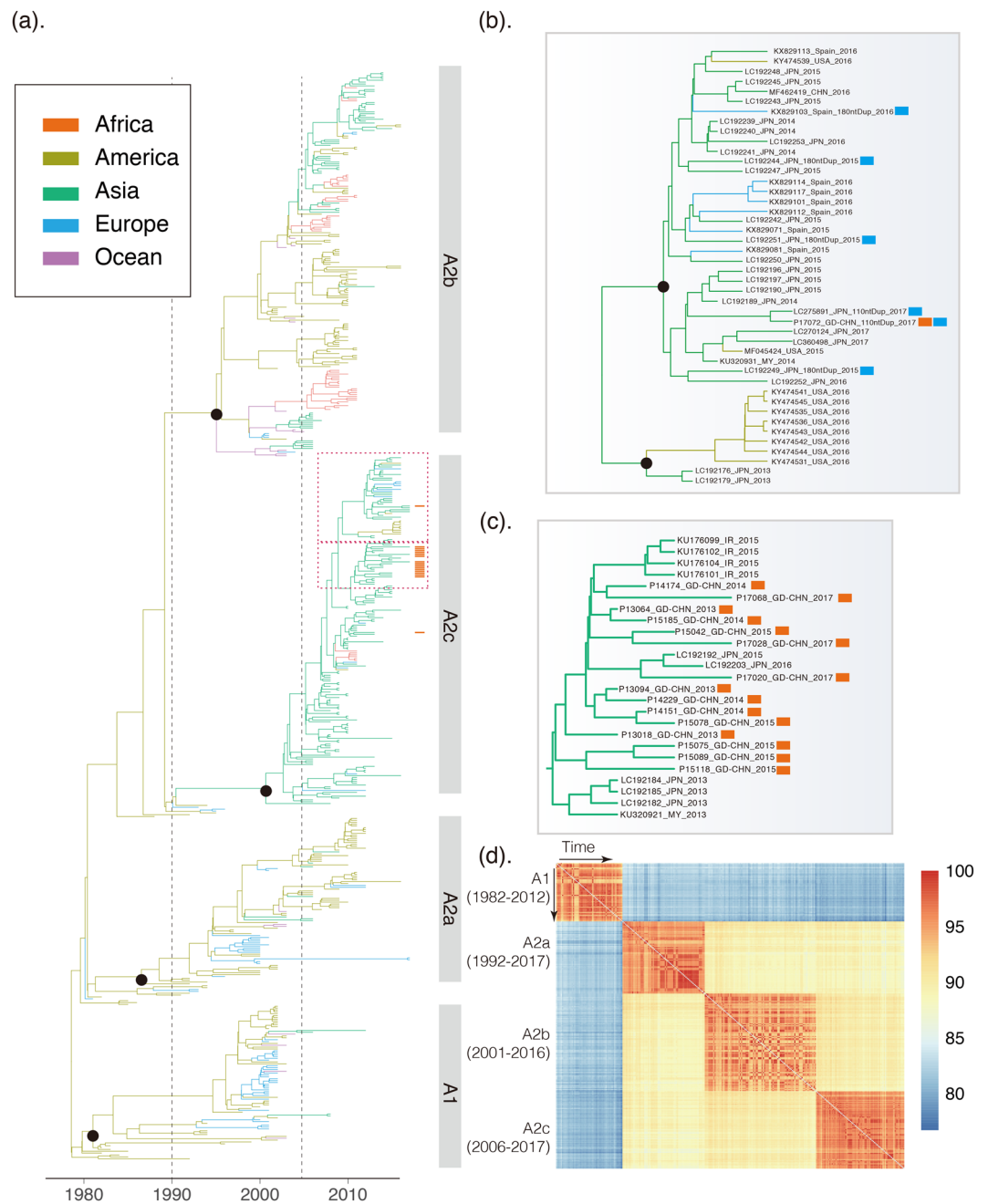


Figure 2. Genetic evolution and spatial spread of the genogroup A hMPV lineages. **(a)** Bayesian Maximum Clade Credibility (MCC) tree of G gene from genogroup A hMPV viruses. sequences reported in this study from Guangzhou are highlighted by orange bars. Branch colours represent the most probable ancestral locations of each branch. Four major lineages of hMPV are observed, denoted A1, A2a, A2b, and A2c. Black circles indicate posterior support >0.95 at the root node of each lineage. The clusters containing major circulating strains in Guangzhou are marked with box. These clusters are enlarged and names of strains are included in **(b,c)**. **(b)** The cluster of hMPV including the 180-nt duplication novel variant (marked with light blue bars) and 110-nt duplication variants (marked with an orange bars). Guangzhou hMPV sequences are also highlighted by red bars in **(b,c)**. **(d)** Heatmap of nucleotide identity matrix of hMPV strains. G genes are grouped by genetic lineages and arranged according to the isolation time.

recently identified in Japan in 2017²⁴. Notably, all these novel variants were scattered in a cluster of A2c lineage, indicating these variants were likely the result of multiple incidences of mutations rather than having evolved from a single mutation strain (Fig. 2a,b). In this scenario, the mutation leading to a duplication fragment in the G gene may easily have occurred for hMPV. In Guangzhou, hMPV infections were mainly caused by A2c, B1, and B2 lineage strains and none of the A1 and A2a lineage strains were identified in hMPV infection cases (Figs 2 and 4). Notably, in 2017, B1 lineage strains replaced the A2c and were likely the major causative strains leading to the increasing epidemiology of hMPV infections in Guangzhou (Figs 1, 3 and 4b). To trace the transmission

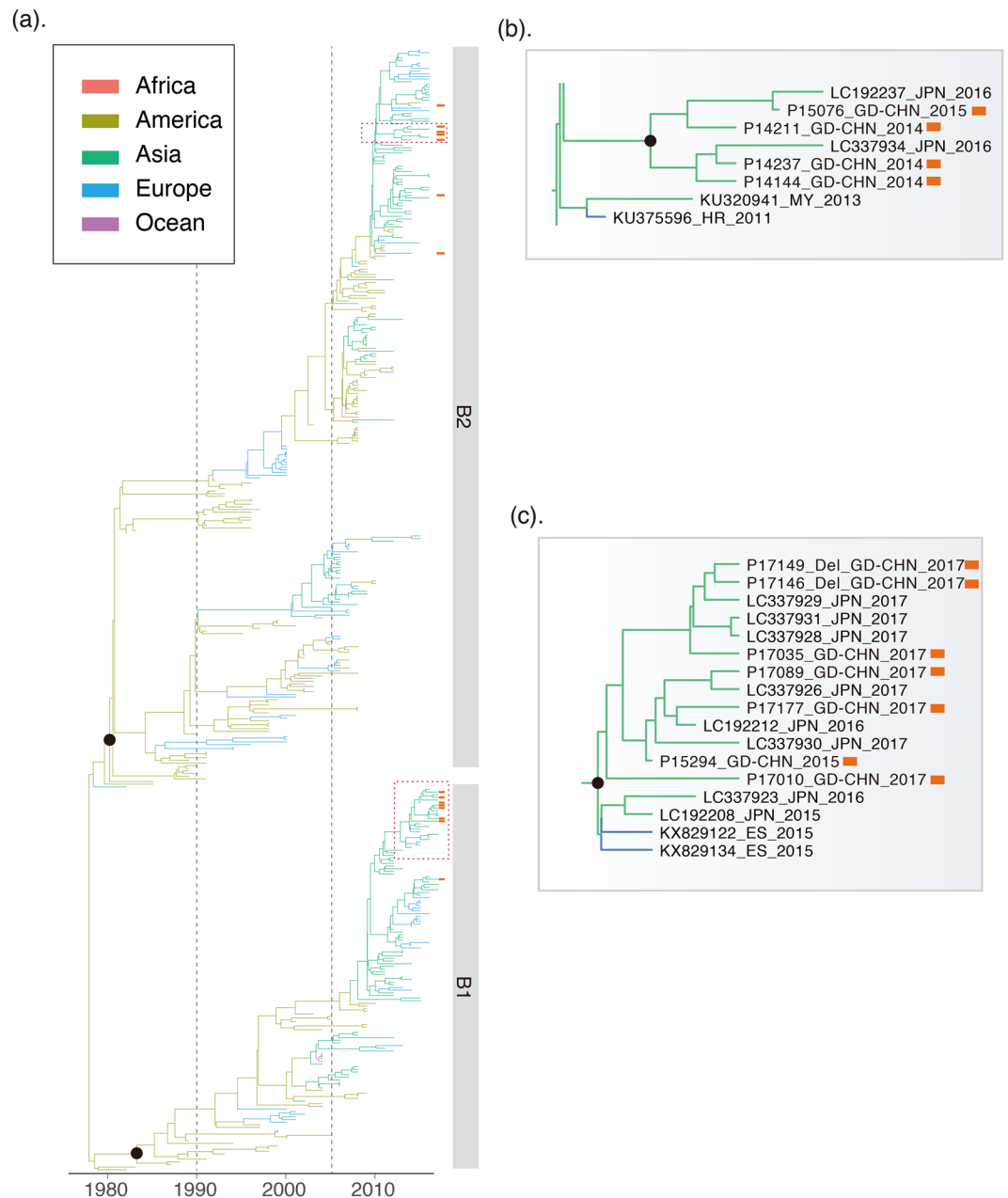


Figure 3. Genetic evolution and spatial spread of the genogroup B hMPV lineages. (a) Bayesian Maximum Clade Credibility (MCC) tree of G gene from genogroup B hMPV viruses. sequences reported in this study from Guangzhou are highlighted by red bars. Branch colours represent the most probable ancestral locations of each branch. Four major lineages of hMPV are observed, denoted B1 and B2. Black circles indicate posterior support >0.95 at the root node of each lineage. The clusters containing major circulating strains in Guangzhou are marked with box. These clusters are enlarged and names of strains are included in (b,c).

of hMPV, we reconstructed the potential origins of Guangdong hMPV strains based on the phylogeny of the G gene. The hMPV strains isolated in Guangzhou in different years were always closely related (Figs 2b,c and 3b,c). However, the contemporary strains from Japan were always clustered together with strains of Guangzhou, highlighting the frequent transmissions among cities in China and Japan. The B1 strains causing the increasing hMPV epidemic in Guangzhou 2017 were more likely imported from Japan since closely related sequences were not found in Guangzhou but in Japan in the previous season (Fig. 3c).

Discussion

Despite the first isolation in 2001, serological data suggested hMPV has been circulating globally for more than 50 years¹. The great genetic distance between two genogroups (A and B) of hMPV indicates the origin of hMPV should be much earlier. To date, the molecular evolution, spatial distribution, and transmission of hMPV have not been adequately studied. Molecular surveillance will also be helpful to understand the transmission pathways

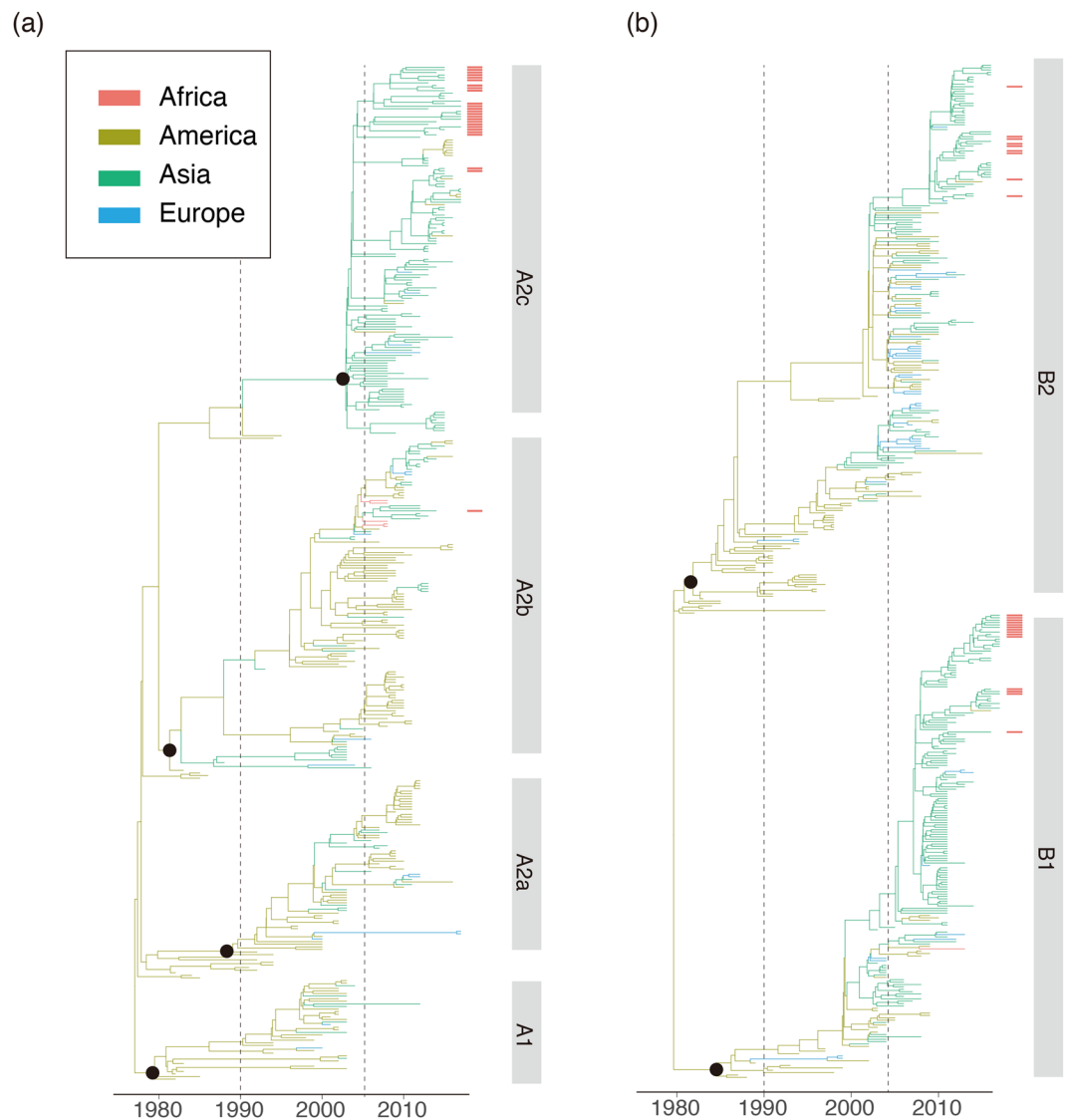


Figure 4. Bayesian Maximum Clade Credibility (MCC) tree of F gene from genogroup A (a) and genogroup B (b) hMPV viruses. sequences reported in this study from Guangzhou are highlighted by red bars. The color scheme of the phylogeny was the same as Figs 2 and 3.

Lineages	A1	A2a	A2b	A2c
A1	89.9–100 (81.8–100)	77.6–85.2 (61.2–72.9)	77.7–85.0 (59.3–72.4)	77.9–84.3 (59.8–72.9)
A2a	—	91.9–100 (86.4–100)	86.0–93.3 (75.7–87.4)	86.1–92.8 (75.7–86.9)
A2b	—	—	91.1–100 (83.2–100)	87.5–93.3 (76.2–87.4)
A2c	—	—	—	91.1–100 (83.2–100)

Table 1. Pairwise identities (%) of G genes from strains within and between hMPV lineages at both the nucleotide and amino acid levels (in bracket).

and to confirm the source of infection, when it is analyzed in conjunction with standard epidemiological investigation. In this study, we performed molecular surveillance on hMPV in Guangzhou, China in 2013–2017. By integrating all publicly available data, we described the spatial dynamics of different hMPV lineages. The establishments of hMPV were identified in Asian countries after 2005. Moreover, frequent transmissions of hMPV strains and the seeding of novel variants were observed in China and Japan, highlighting the risk of an hMPV epidemic in these regions.

The spatial transmission pattern of a virus is very important for epidemic control and potential vaccine design. Viruses like H3N2 seasonal influenza virus are highly transmissible, and a new adaptive variant can rapidly spread globally after its emergence^{38,39}. Thus, in the phylogeny of H3N2, it is rare to have a genetic cluster that is persistent and limited in a specific region. A sink-source model of H3N2 viral ecology is suggested in which

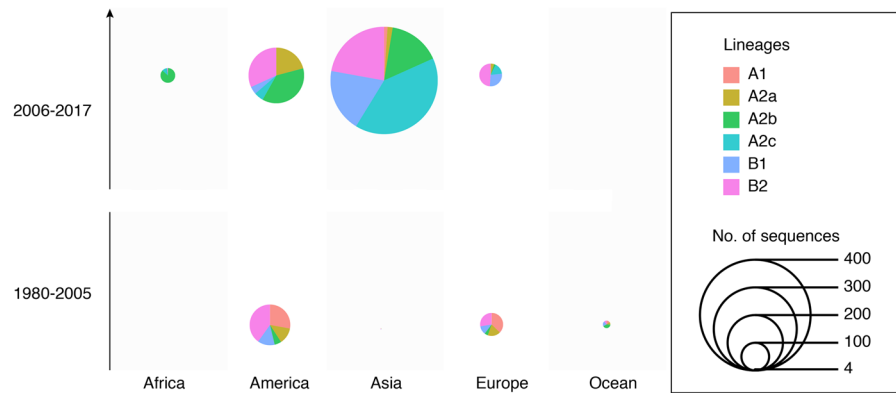


Figure 5. Continental distribution of hMPV genetic lineages in year 1980–2005 and 2006–2017. The number of lineages strains and spatial distribution are calculated according to the phylogeny of G genes in Figs 2 and 3.

new variants are seeded from the tropics (mainly East and Southeast Asia) to sink populations in temperate regions^{39,40}. In contrast, viruses like enterovirus A 71 (EVA71) are more endemic, and different genetic subtypes circulate and persist in different regions. For instance, the C4a sub-genotype of EVA71 mainly circulates in mainland China, while the C5 genotype of EVA71 is mainly detected in Vietnam⁴¹. The establishment and prevalence of the specific sub-genotype of a viral species could be determined by environment factors, population immunity, and/or the genetic background of hosts in this region. Here, we trace the global as well as local hMPV distribution and transmission by integrating publicly available sequences and newly generated sequences. We find that hMPV was endemic in America before 1990, indicating the possible origin of currently circulating hMPV lineages. Notably, this could be the case because hMPV was limited to circulating in America or could be caused by sampling bias as there were few samplings on other continents before 1990 (Fig. 5). Further retrospective study of hMPV in other continents would improve our understanding of the origin of hMPV. Between 1990 and 2000, transmissions from America to other continents were observed. However, most of these introductions were transient since the local phylogenetic clusters of hMPV were not sustained in the phylogeny, indicating hMPV was still not well established in the regions. After 2005, clusters of hMPV mainly composed by strains in A2b and A2c lineages became dominant in Asian countries. Thereafter, transmissions from Asia to other continents were frequently identified through the phylogenetic trees (Figs 2a and 4a). These results suggest the establishment of A2b and A2c lineages in Asia after 2005, and more Asian countries became epicenters for hMPV. In this scenario, it is reasonable to have novel variants of hMPV first identified in Asian regions. The 180-nt and 111-nt duplicates of hMPV in the G gene were first reported in Japan in 2016 and 2017, and these newly emerged variants caused an increasing epidemic in Japan^{23,24}. In Guangzhou, the duplicate variant was also detected and closely related with the strains identified in Japan with nucleotide identity at 99.3% (Fig. 2b). A2c was the predominant lineage identified in hMPV epidemics in Guangzhou, China 2013–2015. However, the increasing epidemic of hMPV infection in Guangzhou in 2017 was mainly caused by the B1 strains which were closely related viruses circulating in Japan in 2016 (Fig. 3). These results suggest the predominant hMPV lineage is varied among Asian countries in the same epidemic season, and rotation of these predominant lineages may be observed among these countries due to frequent viral transmissions.

Infections of hMPV were recently investigated in multiple regions of China^{20–22,37,42}. Different seasonal patterns were observed in different climatic zones. In central China, hMPV is predominantly detected in winter and spring²⁰. In northwest China, the annual peak of hMPV infection is observed in September through October¹⁹. However, in Guangzhou in southern China, the seasonal peaks of hMPV were detected mainly in March through May (Fig. 1). Although the temporal distribution and clinical symptoms of hMPV have been reported, the representative genetic sequences of hMPV in different regions of China are rarely investigated. Therefore, the geographic distributions and the transmission dynamics of hMPV lineages among different regions of China are still little understood. Our genetic description and analysis of hMPV viruses in southern China provide a baseline for future studies of the evolution and molecular epidemiology of this emerging virus in China.

In summary, our genetic and epidemiological analyses show that hMPV displayed different distribution and transmission patterns before 1990, 1990–2005, and 2006–2017. The firm establishment and continuous evolution of hMPV is observed in Asian countries after 2005, leading to the emergence of novel variants and the global transmission of these variants. In addition, close relations are observed among hMPV strains from China and Japan through the whole phylogenies. Therefore, careful monitoring of hMPV should be conducted throughout Asian countries to illustrate virus circulation in the area and sound an alarm for global hMPV surveillance.

References

1. Hoogen, B. Gvanden *et al.* A newly discovered human pneumovirus isolated from young children with respiratory tract disease. *Nature Medicine* **7**, 719–724 (2001).
2. Panda, S., Mohakud, N. K., Pena, L. & Kumar, S. Human metapneumovirus: review of an important respiratory pathogen. *International Journal of Infectious Diseases* **25**, 45–52 (2014).
3. Naz, R. *et al.* Etiology of acute viral respiratory infections common in Pakistan: A review. *Rev Med Virol* e2024, <https://doi.org/10.1002/rmv.2024> (2018).
4. Edwards, K. M. *et al.* Burden of human metapneumovirus infection in young children. *N. Engl. J. Med.* **368**, 633–643 (2013).

5. Aberle, S. W., Aberle, J. H., Sandhofer, M. J., Pracher, E. & Popow-Kraupp, T. Biennial spring activity of human metapneumovirus in Austria. *Pediatr. Infect. Dis. J.* **27**, 1065–1068 (2008).
6. Aberle, J. H., Aberle, S. W., Redlberger-Fritz, M., Sandhofer, M. J. & Popow-Kraupp, T. Human metapneumovirus subgroup changes and seasonality during epidemics. *Pediatr. Infect. Dis. J.* **29**, 1016–1018 (2010).
7. Martinello, R. A. *et al.* Human metapneumovirus and exacerbations of chronic obstructive pulmonary disease. *Journal of Infection* **53**, 248–254 (2006).
8. Schildgen, V. *et al.* Human Metapneumovirus: Lessons Learned over the First Decade †. *Clin Microbiol Rev* **24**, 734–754 (2011).
9. van den Hoogen, B. G., Bestebroer, T. M., Osterhaus, A. D. M. E. & Fouchier, R. A. M. Analysis of the genomic sequence of a human metapneumovirus. *Virology* **295**, 119–132 (2002).
10. van den Hoogen, B. G. *et al.* Antigenic and Genetic Variability of Human Metapneumoviruses. *Emerging Infectious Diseases* **10**, 658–666 (2004).
11. Chow, W. Z. *et al.* Genetic diversity, seasonality and transmission network of human metapneumovirus: identification of a unique sub-lineage of the fusion and attachment genes. *Scientific Reports* **6** (2016).
12. Papenburg, J. *et al.* Genetic diversity and molecular evolution of the major human metapneumovirus surface glycoproteins over a decade. *Journal of Clinical Virology* **58**, 541–547 (2013).
13. de Graaf, M., Osterhaus, A. D. M. E., Fouchier, R. A. M. & Holmes, E. C. Evolutionary dynamics of human and avian metapneumoviruses. *Journal of General Virology* **89**, 2933–2942 (2008).
14. Huck, B. *et al.* Novel human metapneumovirus sublineage. *Emerging Infect. Dis.* **12**, 147–150 (2006).
15. Biacchesi, S. *et al.* Genetic diversity between human metapneumovirus subgroups. *Virology* **315**, 1–9 (2003).
16. Park, J. Y. *et al.* Clinical and genetic features of human metapneumovirus infection in children. *Pediatr Int* **58**, 22–26 (2016).
17. van den Hoogen, B. G. *et al.* Experimental infection of macaques with human metapneumovirus induces transient protective immunity. *J. Gen. Virol.* **88**, 1251–1259 (2007).
18. Zhang, L. *et al.* Epidemiological and clinical features of human metapneumovirus in hospitalised paediatric patients with acute respiratory illness: a cross-sectional study in Southern China, from 2013 to 2016. *BMJ Open* **8**, e019308 (2018).
19. Li, X. *et al.* Viral etiologies and epidemiology of patients with acute respiratory infections based on sentinel hospitals in Gansu Province, Northwest China, 2011–2015. *J. Med. Virol.* **90**, 828–835 (2018).
20. Kong, W. *et al.* Circulation of human metapneumovirus among children with influenza-like illness in Wuhan, China. *J. Med. Virol.* **88**, 774–781 (2016).
21. Peiris, J. S. M. *et al.* Children with Respiratory Disease Associated with Metapneumovirus in Hong Kong. *Emerging Infectious Diseases* **9**, 628–633 (2003).
22. Zhang, C. *et al.* Detection and Genetic Diversity of Human Metapneumovirus in Hospitalized Children with Acute Respiratory Infections in Southwest China. *Journal of Clinical Microbiology* **50**, 2714–2719 (2012).
23. Saikusa, M. *et al.* 180-Nucleotide Duplication in the G Gene of Human metapneumovirus A2b Subgroup Strains Circulating in Yokohama City, Japan, since 2014. *Frontiers in Microbiology* **8** (2017).
24. Saikusa, M. *et al.* A novel 111-nucleotide duplication in the G gene of human metapneumovirus: HMPV evolution. *Microbiology and Immunology* **61**, 507–512 (2017).
25. Zou, L. *et al.* Evolution and Transmission of Respiratory Syncytial Group A (RSV-A) Viruses in Guangdong, China 2008–2015. *Frontiers in Microbiology* **07** (2016).
26. Ke, C. *et al.* Circulation of Reassortant Influenza A(H7N9) Viruses in Poultry and Humans, Guangdong Province, China, 2013. *Emerging Infectious Diseases* **20**, 2034–2040 (2014).
27. Yi, L. *et al.* A cluster of adenovirus type B55 infection in a neurosurgical inpatient department of a general hospital in Guangdong, China. *Influenza and Other Respiratory Viruses* **11**, 328–336 (2017).
28. Larsson, A. AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* **30**, 3276–3278 (2014).
29. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
30. Rambaut, A., Lam, T. T., Max Carvalho, L. & Pybus, O. G. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evolution* **2**, vew007 (2016).
31. Lu, J. *et al.* The Evolution and Transmission of Epidemic GII.17 Noroviruses. *Journal of Infectious Diseases* **214**, 556–564 (2016).
32. Lu, J. *et al.* Molecular Evolution, Diversity, and Adaptation of Influenza A(H7N9) Viruses in China. *Emerging Infectious Diseases* **24**, 1795–1805 (2018).
33. Gill, M. S. *et al.* Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. *Mol. Biol. Evol.* **30**, 713–724 (2013).
34. Drummond, A. J., Ho, S. Y. W., Phillips, M. J. & Rambaut, A. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**, e88 (2006).
35. Lemey, P. *et al.* Unifying viral genetics and human transportation data to predict the global transmission dynamics of human influenza H3N2. *PLoS Pathog.* **10**, e1003932 (2014).
36. Regev, L. *et al.* Characterization of Human Metapneumovirus Infections in Israel. *Journal of Clinical Microbiology* **44**, 1484–1489 (2006).
37. Li, J. *et al.* Prevalence of human metapneumovirus in adults with acute respiratory tract infection in Beijing, China. *J. Infect.* **64**, 96–103 (2012).
38. Russell, C. A. *et al.* The Global Circulation of Seasonal Influenza A (H3N2) Viruses. *Science* **320**, 340–346 (2008).
39. Bedford, T. *et al.* Global circulation patterns of seasonal influenza viruses vary with antigenic drift. *Nature* **523**, 217–220 (2015).
40. Rambaut, A. *et al.* The genomic and epidemiological dynamics of human influenza A virus. *Nature* **453**, 615–619 (2008).
41. Thao, N. T. T. *et al.* Evolution and Spatiotemporal Dynamics of Enterovirus A71 Subgenogroups in Vietnam. *The Journal of Infectious Diseases* **216**, 1371–1379 (2017).
42. Zhou, J.-Y. *et al.* Human bocavirus and human metapneumovirus in hospitalized children with lower respiratory tract illness in Changsha, China. *Influenza Other Respir Viruses* **12**, 279–286 (2018).

Acknowledgements

We gratefully acknowledge the efforts of the sentinel hospital in the investigation and sample collection. This work was supported by grants from Natural Science Foundation of Guangdong Province (2018A030313205), Science and Technology Planning Project of Guangdong (2018B020207006), and Science and Technology Planning Project of Guangzhou (201804010030).

Author Contributions

J.W., T.S. and L.Z. designed the study; L.Z., J.Y., Y.S., L.L., Q.G., M.K., C.K. and T.S. collected samples and prepared genome sequencing; J.L. and L.Y. analyzed data; J.L., L.Y. and L.Z. interpreted data; J.L. and L.Y. prepared the figures; and J.L., J.P., L.Y. and J.W. wrote the article.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-50340-8>.

Competing Interests: The authors declare no competing interests.

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 4.0 International License, which permits use, sharing, adaptation, 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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019