Whole-genome Sequencing for Tracing the Transmission Link between Two ARD Outbreaks Caused by a Novel HAdV Serotype 7 Variant, China

Shaofu Qiu1,*, Peng Li1,*, Hongbo Liu1,*, Yong Wang1,*, Nan Liu1,*, Chengyi Li1, Shenlong Li1, Ming Li2, Zhengjie Jiang3, Huandong Sun3, Ying Li2, Jing Xie2, Chaojie Yang3, Jian Wang3, Hao Li2, Shengjie Yi1, Zhihao Wu1, Leili Jia1, Ligui Wang3, Rongzhang Hao3, Yansong Sun1, Liuyu Huang1, Hui Ma4, Zhengquan Yuan1 & Hongbin Song1

From December 2012 to February 2013, two outbreaks of acute respiratory disease caused by HAdV-7 were reported in China. We investigated possible transmission links between these two seemingly unrelated outbreaks by integration of epidemiological and whole-genome sequencing (WGS) data. WGS analyses showed that the HAdV-7 isolates from the two outbreaks were genetically indistinguishable; however, a 12 bp deletion in the virus-associated RNA gene distinguished the outbreak isolates from other HAdV-7 isolates. Outbreak HAdV-7 isolates demonstrated increased viral replication compared to non-outbreak associated HAdV-7 isolate. Epidemiological data supported that the first outbreak was caused by introduction of the novel HAdV-7 virus by an infected recruit upon arrival at the training base. Nosocomial transmission by close contacts was the most likely source leading to onset of the second HAdV-7 outbreak, establishing the apparent transmission link between the outbreaks. Our findings imply that in-hospital contact investigations should be encouraged to reduce or interrupt further spread of infectious agents when treating outbreak cases, and WGS can provide useful information guiding infection-control interventions.

Human adenoviruses (HAdVs) are non-enveloped, icosahedral, double-stranded DNA viruses that are classified into seven species (A–G) based on their nucleotide homology, haemagglutination, biochemical and biological properties; at least 51 serotypes are recognized1–3. HAdV can cause a wide range of clinical syndromes, including respiratory tract illnesses, conjunctivitis, cystitis, gastroenteritis, neurological diseases; in some cases infection directly results in death1,4. HAdV is reported to be one of the primary causes of acute respiratory disease (ARD) worldwide, which is responsible for 5–10% of lower respiratory tract infections in children throughout the world5,6. Serotype 7 (HAdV-7) accounts for nearly 20% of all HAdV infections reported to the World Health Organization4,6, and is more frequently associated with severe illnesses, particularly in children aged <7 years and persons with underlying health conditions7,8.
Moreover, HAdV-7 is among the most prevalent serotypes associated with ARD outbreaks in institutional, hospital and community settings. Whole-genome sequencing (WGS) has recently been applied to rapidly discriminate outbreaks caused by bacteria such as *Escherichia coli* O104: H4, methicillin-resistant *Staphylococcus aureus* and *Mycobacterium tuberculosis*, carbapenem-resistant *Klebsiella pneumoniae* and *Legionella pneumophila*, but has been used with limited study size or scope in describing transmission links during the course of outbreaks caused by viruses. From December 2012 to February 2013, two ARD outbreaks caused by HAdV-7 occurred at two training bases located in different cities (separated by approximately 170 km) in Hubei Province, China. A total of 946 persons were symptomatic as part of these two outbreaks. An investigation of these two HAdV-7 outbreaks to determine whether a potential transmission link could be identified is described using WGS coupled with clinical and epidemiological investigations.

**Results**

**Overview of the outbreaks.** Base A (Xiangyang) houses 1290 recruits, of whom 858 (66.5%) fell ill with ARD from December 10, 2012 to January 28, 2013 (Fig. 1). Among the 858 cases, 321 (37.4%) were hospitalized, 108 (12.6%) developed acute pneumonia and 37 were considered severe cases. Base B (Jingmen) houses 865 recruits; 88 (10.2%) were diagnosed with ARD from February 11 to 27, 2013. Among the 88 cases, 42 (47.7%) were hospitalized, 16 (18.2%) with pneumonia and 4 severe cases were recorded. Under normal circumstances, these training facilities are not directly connected; this is primarily due to the geographic separation (Figure S1) and self-contained management systems. The initial hypothesis was that these outbreaks were unrelated.

The mean ages of the cases at the onset of illness were 19.7 and 20.1 years, respectively (Table S5). A small number of cases from both outbreaks had underlying medical conditions; the most common condition was allergy to antibiotics (6.1% & 8.0%). Analysis of clinical characteristics of the hospitalized
cases from the two outbreaks showed very similar clinical features (Table 1). A total of 36.2% and 38.1% of the hospitalized cases from the outbreaks showed acute pneumonia. Among the pneumonia cases, more than 50% had bilateral pneumonia, and 34.3% and 25.0% of them developed into severe cases, respectively. Hospitalized cases from the outbreaks also showed very similar clinical laboratory and radiologic findings, and abnormal findings such as abnormalities on chest radiograph and CT were commonly observed among the hospitalized cases from the outbreaks (Table 2). These abnormal manifestations may be indicators of multiorgan involvement among the cases17.

**Epidemiologic investigations of the outbreaks.** Based on the similar demographic characteristics, clinical presentations, laboratory and radiographic findings, possible epidemiological associations that might link the outbreaks was explored. Preliminary investigations identified a 22-year-old male recruit (Index case A) who had ARD symptoms before reporting to Base A (Xiangyang) on December 6, 2012 (Fig. 2). He arrived at the training base on December 10, 2012. During the next one month, he was treated in the clinic of Base A by empirical administration of roxithromycin, lincomycin, ceftriaxone and ribavirin, however his condition did not resolve and by January 10, 2013 had worsened to persistent high fever (39–39.9 °C) and shortness of breath. The case was admitted to Xiangyang affiliated hospital on January 11, 2013 where he was diagnosed as a severe pneumonia case with multiorgan damage. He had an elevated WBC count (12.5 × 10⁹ cells/liter), creatine kinase (56667 U/liter), lactate dehydrogenase (348 U/liter), aspartate aminotransferase (687 U/liter), alanine aminotransferase (303 U/liter) and C-reactive protein (28.16 mg/liter), and decreased level of potassium (3.39 mmol/liter). He was treated with additional antimicrobial and antiviral agents including azithromycin, levofloxacin, cefepime and ribavirin, and methylprednisolone was intravenously administered (range, 40–320 mg per day). Liver-protective and myocardial nutritional drugs were also prescribed. After 25 days of hospitalization the case made a full recovery. A serum sample collected on January 10, 2013 was positive for HAdV IgA, and a HAdV-7 isolate was identified from a nasopharyngeal sample collected on January 17, 2013 by PCR and viral culture, suggesting that the case had prolonged shedding of adenovirus (Fig. 2). From December to January, this case was not quarantined at Base A, and he participated in normal training and educational activities including the training course opening ceremony on December 20, 2012. During this period, 9 recruits living in the same dormitory (9/12) with the index case successively showed ARD symptoms; HAdV-7 was detected in 7 (77.8%) by PCR analysis of nasopharyngeal samples.

During our investigation, an observation that cases unable to be diagnosed and/or treated in the medical clinics associated with Base A and B were usually transferred to Xiangyang affiliated hospital was made. Importantly, while the outbreak was occurring at Base A, during which time >300 cases were sent to the Xiangyang affiliated hospital in one month, a 19-year-old male recruit from base B (Jingmen) (Index case C) had been sent for treatment to Xiangyang affiliated hospital for lumbago and backpain on February 3, 2013 (Fig. 2). Physicians did not strictly comply with infection control measures established by the hospital; primarily because of poor awareness of infection control for the inpatients. While Index case C was receiving treatment, he reported having two, half an hour conversations with an acquaintance from Base A (Index case B) who was hospitalized for ARD (Fig. 2). After returning to Base B, Index Case C participated in routine training exercises but developed ARD on February 11, 2013. Infection with HAdV-7 was confirmed by PCR and subsequent DNA sequence analysis; five roommates of Index case C subsequently showed ARD symptoms.

**Laboratory investigations and genome sequencing.** A total of 377 HAdV-7 isolates were identified by PCR and viral culture; 323 were from outbreak cases including the three index cases, 48 were from close contacts to these cases, and 6 were from non-recruit local ARD patients. The positive rates of HAdV-7 for the outbreak cases, close contacts and non-recruit ARD cases were 79.2%, 30.0% and 2.6%, respectively (Figure S2). Among the non-recruit cases, HAdV-7 was identified from four cases from Xiangyang City and two from Jingmen City. Phylogenetic analysis based on whole genes extracted from sequenced genomes showed that the HAdV-7 isolates from the outbreak cases, close contacts and non-recruit cases were indistinguishable in the hexon and fiber genes.

To rapidly measure the degree of genetic relatedness among the HAdV-7 isolates from the outbreak and non-recruit cases, WGS was performed on ten HAdV-7 isolates including those from the index cases. Genome sequences were deposited in the GenBank database (accession nos. KJ019879-KJ019888). By aligning the sequenced genomes and using 0901 HZ as a reference genome (accession no. JF800905), the nine outbreak isolates, including those from the index cases (accession nos. KJ019879, KJ019880 and KJ019885), were 100% identical. Phylogenetic analysis of the WGS showed that the outbreak isolates formed an independent sublineage among the HAdV-7 strains (Fig. 3A). Comparative genomic and phylogenetic analyses showed that the outbreak isolates had genomes nearly identical with an isolate CDC228 (accession no. KJ019884) identified from a local patient, as well as other five strains (accession nos. JF800905, KC440171, JX625134, KF268314 and KF268316) circulating in China6,18,19. Further phylogenetic analysis of the individual whole genes revealed that the outbreak isolates clustered with the local isolate CDC228 as well as the above five strains, and shared 100% similarities in the hexon and fiber genes (Fig. 3B,C). However, the outbreak strains formed a separate subclade in the phylogenetic trees based on the penton, DNA polymerase and VA RNA genes (Figs 3D and 4A,C). Genome comparisons revealed that the outbreak isolates had only minor differences in the penton and DNA polymerase genes.
| Characteristic                              | Xiangyang outbreak (n = 321) | Jingmen outbreak (n = 42) |
|--------------------------------------------|------------------------------|---------------------------|
| **Symptoms or signs**                      |                              |                           |
| Fever                                      |                              |                           |
| Any-no. (%)                                | 319 (99.4)                   | 42 (100.0)                |
| Subgroup-no. (%)                           |                              |                           |
| 37.3–38.0 °C                               | 10 (3.1)                     | 0                         |
| 38.1–39.0 °C                               | 131 (40.8)                   | 20 (47.6)                 |
| > 39.0 °C                                  | 178 (55.5)                   | 22 (52.4)                 |
| Maximal temperature (°C), mean             | 39.2 ± 0.6                   | 39.4 ± 0.7                |
| Duration of fever, median (range), days    | 4.0 (1–20)                   | 4.0 (2–11)                |
| Cough                                      | 304 (94.7)                   | 40 (95.2)                 |
| Throat congestion                          | 296 (92.2)                   | 36 (85.7)                 |
| Sore throat                                | 246 (76.6)                   | 32 (76.2)                 |
| Sputum production -no./total no. (%)       | 234/321 (72.9)               | 34/42 (81.0)              |
| Yellow sputum                              | 145/234 (62.0)               | 24/34 (70.6)              |
| White sputum                               | 89/234 (38.0)                | 10/34 (29.4)              |
| Blood-streaked sputum                      | 28/234 (12.0)                | 6/34 (17.6)               |
| Swelling of tonsils                        | 211 (65.7)                   | 24 (57.1)                 |
| Fatigue                                    | 139 (43.3)                   | 21 (50.0)                 |
| Pharyngeal lymphoid follicular hyperplasia | 126 (39.3)                   | 15 (35.7)                 |
| Purulent exudate on the tonsil             | 102 (31.8)                   | 12 (28.6)                 |
| Rhinorrhea                                 | 79 (24.6)                    | 9 (21.4)                  |
| Myalgia                                    | 76 (23.7)                    | 8 (19.0)                  |
| Headache                                   | 64 (19.9)                    | 8 (19.0)                  |
| Dizziness                                  | 61 (19.0)                    | 5 (11.9)                  |
| Shortness of breath                        | 54 (16.8)                    | 8 (19.0)                  |
| Chest pain                                 | 51 (15.9)                    | 9 (21.4)                  |
| Diarrhea                                   | 30 (9.3)                     | 5 (11.9)                  |
| Nasal congestion                           | 29 (9.0)                     | 2 (4.8)                   |
| Chill                                      | 23 (7.2)                     | 2 (4.8)                   |
| Enlargement of lymph nodes                 | 20 (6.2)                     | 4 (9.5)                   |
| Abdominal pain                             | 19 (5.9)                     | 3 (7.1)                   |
| Bronchitis                                 | 18 (5.6)                     | 1 (2.4)                   |
| Nausea                                     | 12 (3.7)                     | 5 (11.9)                  |
| Vomiting                                   | 8 (2.5)                      | 1 (2.4)                   |
| Palpitation                                | 3 (0.9)                      | 1 (2.4)                   |
| Submandibular lymphadenitis                | 3 (0.9)                      | 0                         |
| Tinnitus, earache                          | 3 (0.9)                      | 0                         |
| Skin rash                                  | 2 (0.6)                      | 1 (2.4)                   |
| Convulsion, coma                           | 2 (0.6)                      | 0                         |
| Testicular pain                            | 2 (0.6)                      | 0                         |
| **Clinical outcome**                       |                              |                           |
| Hospitalization, no./total no. (%)         | 321/858 (37.4)               | 42/88 (47.7)              |
| Time from illness onset to hospitalization, median (range), days | 3.0 (1–36) | 3.0 (1–7) |
| Duration of hospitalization, median (range), days | 10.0 (6–25) | 9.0 (7–25) |
| Pneumonia, no./total no. (%)               | 108/321 (36.2)               | 16/42 (38.1)              |
| Admission to an intensive care unit        | 37/108 (34.3)                | 4/16 (25.0)               |
| Heart injury                               | 14/108 (13.0)                | 2/16 (12.5)               |
| Liver injury                               | 11/108 (10.2)                | 3/16 (18.8)               |

Continued
with CDC228 and the above reference strains (Table 3, Figs 3D and 4A), but a distinct point mutation was observed among them when compared with other HAdV-7 strains, resulting in the replacement of Leu381 with Ile (Table 3 and Fig. 4B). Additionally, all of the outbreak isolates showed an obvious difference with other HAdV-7 strains in the VA RNA genes (Fig. 4C), with a unique deletion of 12 base pair in the VA RNA II region (Table 3 and Fig. 4D). The universality of this unique deletion and the distinct point mutation were verified among an additional 46 HAdV-7 outbreak strains by PCR and DNA sequencing (data not shown).

**Growth kinetics of viral replication.** The in vitro growth kinetics of viral replication were examined using isolate XY1 as representative of the outbreak-associated HAdV-7 viruses. Similar growth kinetics were observed with the non-outbreak associated isolate CDC228 between 0–12h post-infection, while XY1 had a reproducibly higher viral titer than that of CDC228 between 12–48h post-infection, indicating enhanced viral replication of outbreak isolates during this time period (Fig. 5).

**Discussion**

In China, HAdVs are not notifiable infectious pathogens, and large-scale epidemiological data from HAdV infections nationwide is currently not available. According to literature published from 1983 to 2013, at least 30 HAdV outbreaks have been reported in China, most of these have occurred since the year 20006,20–31. To date, at least ten HAdV-7 outbreaks in China have been described, 5 were reported before 199332–35, one in 2009, one in 201119, one in 201228 and two described in this study. Interestingly, the HAdV-7 outbreaks occurring before 1993 were mainly associated with pharyngoconjunctival fever symptoms in children; no deaths were reported22–35. However, recently reported HAdV-7 outbreaks were all associated with symptoms of ARD, sometimes leading to severe or fatal illnesses in otherwise healthy persons living in China6,28. In this study the two HAdV-7 outbreaks mainly infected the healthy young adults, and serious or even life-threatening infections were often observed; fortunately no deaths were recorded. Several distinct clinical characteristics were observed from cases comprising the two outbreaks reported in this study. Most of the hospitalized patients with sputum production had yellow sputum (62.0% & 70.6%), and leukocytosis (WBC count, > 10.0 x 10⁹ cells/liter) was more frequently seen in hospitalized patients (Table 2). These characteristics were notable differences from those described in the outbreaks associated with other HAdVs or other respiratory viruses36–38. Moreover, the high attack rate (66.5%) observed during the Base A (Xiangyang) outbreak places this outbreak as one of the largest HAdV-7 outbreaks documented worldwide2. These observations imply that virulence associated with HAdV is changing in China; this may be due to the emergence of more virulent new genotypes within existing serotypes, host factors, or a combination of these possibilities.

Combining the clinical observations, laboratory studies, epidemiological investigations and WGS analyses, the first case (Index case A) was identified as a recruit whose home was Hunan not Hubei Province. Distinct genetic differences were documented between the genomes of outbreak-associated and circulating, non-outbreak associated HAdV-7 isolates. WGS analysis strongly supported the hypothesis that the outbreak was caused by an imported HAdV-7 isolate. We were not able to investigate the prevalence of HAdV-7 in other regions of China, especially that in Hunan Province, therefore we were unable to determine the exact source of the HAdV-7 outbreak isolate. A significant contributor to the dispersal of the virus and extension of the outbreak was the lack of awareness of infectious disease control and the absence of properly trained medical personnel at training base A. Index case A, who had ARD symptoms in early December 2012, was not quarantined during the time of his treatment at Base A. This patient was possibly a super-spreader as he probably had prolonged shedding of adenovirus. Prolonged shedding of adenovirus and its hardy nature are therefore conducive to the dissemination of adenovirus to others especially the people in close contact6. To effectively prevent and control the occurrence of infectious disease outbreaks in the schools and training institutions especially when freshmen or new recruits enrolled, strengthening of pathogen surveillance and quarantine inspection to rapidly find and timely investigate potential transmission is critical.

This study has clearly demonstrated supplementing traditional outbreak investigations with WGS analysis provides a powerful discriminatory method for identifying transmission links between apparently unrelated outbreaks. In this case, WGS clearly established the role of nosocomial infection in the

| Characteristic                | Xiangyang outbreak (n = 321) | Jingmen outbreak (n = 42) |
|------------------------------|-------------------------------|----------------------------|
| Gallbladder injury            | 3/108 (2.8)                  | 1/16 (6.3)                 |
| Meningitis irritation         | 2/108 (1.9)                  | 0                          |
| Secondary infections          | 10/108 (9.3)                 | 2/16 (12.5)                |

Table 1. Clinical characteristics of the hospitalized patients from the two outbreaks*. *In this study continuous variables were summarized as means (± SD) or medians (range).
transmission and introduction of HAdV-7 into a geographically unlinked location. Without WGS analysis, it is likely that nosocomial transmission of HAdV-7 among patients in the hospital would not have been appreciated, leading to the false hypothesis that the two outbreaks were independent. Further, the relationship between the outbreak and circulating non-outbreak-associated isolates in the greater community became clear using WGS analysis. In the absence of WGS, endemic circulating HAdV isolates most likely would have been mistakenly identified as outbreak isolates because both isolates were indistinguishable using traditional genotyping based on the sequences of hexon and fiber genes. WGS revealed that HAdV-7 isolates from the two outbreaks were genetically indistinguishable, while could be distinguished by small but significant changes, with a non-synonymous substitution in DNA polymerase variable.

| Variable                      | Value                                      |
|-------------------------------|--------------------------------------------|
| WBC count                     |                                            |
| Median (range), 10^9 cells/liter | 10.4 (3.1–25) 10.5 (3.4–15.4)             |
| >10 x 10^9 cells/L, no. (%)    | 156 (52.3) 21 (50.0)                      |
| <4 x 10^9 cells/L, no. (%)     | 23 (7.7) 3 (7.1)                          |
| RBC count, median (range), 10^{12} cells/liter | 4.46 (3.45–6.0) 4.31 (3.47–4.97) |
| Platelet count, median (range), 10^12 cells/liter | 213.2 (129–414) 223.5 (137–414) |
| Hemoglobin, median (range), g/liter | 133 (98–312) 134 (102–156)               |
| Glucose, >6.2 mmol/liter, no. (%) | 76 (25.5) 9 (21.4)                         |
| Creatine Kinase, >200 U/liter, no. (%) | 97 (32.6) 18 (42.9)                    |
| Creatine kinase MB fraction, >25 U/liter, no. (%) | 85 (28.5) 18 (42.9)                    |
| Lactate dehydrogenase, >250 U/liter, no. (%) | 39 (13.1) 5 (11.9)                      |
| Aspartate aminotransferase, >40 U/liter, no. (%) | 43 (14.4) 6 (14.3)                      |
| Alanine aminotransferase, >40 U/liter, no. (%) | 41 (13.8) 6 (14.3)                      |
| Total bilirubin, >17.1 μmol/liter, no. (%) | 27 (9.1) 4 (9.5)                        |
| Uric acid, >416 umol/liter, no. (%) | 17 (5.7) 2 (4.8)                         |
| Creatinine, >133 μmol/liter, no. (%) | 13 (4.4) 2 (4.4)                         |
| Potassium                     |                                            |
| median (range), mmol/liter     | 3.76 (3.0–5.61) 3.75 (3.18–4.57)        |
| <3.5 μmol/liter, no./total no. (%) | 32 (10.7) 8 (19.0)                  |
| Sodium, median (range), mmol/liter | 138 (132–143) 138 (131–142)            |
| Erythrocyte sedimentation rate, >15 mm/hr, no. (%) | 79 (26.5) 20 (47.6)                     |
| C-reactive protein, >10 mg/liter, no. (%) | 117 (39.3) 15 (35.7)                     |
| Abnormalities on electrocardiogram, no./total no. (%) | 101/298 (33.9) 16/42 (38.1)             |
| Sinus arrhythmia               |                                            |
|                                | 60/101 (59.4) 11/16 (68.8)                |
| High left ventricular voltage  |                                            |
|                                | 27/101 (26.7) 5/16 (31.3)                 |
| Bundle branch block            |                                            |
|                                | 3/101 (3.0) 1/16 (6.3)                    |
| T-wave change                  |                                            |
|                                | 7/101 (6.9) 0                             |
| Ventricular premature contraction |                                            |
|                                | 5/101 (5.0) 1/16 (6.3)                    |
| Abnormalities on chest radiograph and CT, no./total no. (%) | 108/298 (36.2) 16/42 (38.1)             |
| Local patchy shadowing         |                                            |
|                                | 101/108 (93.5) 15/16 (93.8)               |
| Ground-glass opacities         |                                            |
|                                | 3/108 (2.8) 0                             |
| Pulmonary nodule               |                                            |
|                                | 6/108 (5.6) 1/16 (6.3)                    |
| Mediastinal lymphadenecescisis |                                            |
|                                | 5/108 (4.6) 2/16 (12.5)                   |
| Consolidation                  |                                            |
|                                | 4/108 (3.7) 1/16 (6.3)                    |
| Involvement of left lung       |                                            |
|                                | 11/108 (10.2) 1/16 (6.3)                  |
| Involvement of right lung      |                                            |
|                                | 41/108 (38.0) 4/16 (25.0)                 |
| Involvement of both lungs      |                                            |
|                                | 56/108 (51.9) 11/16 (68.8)                |

Table 2. Laboratory and radiographic findings of the hospitalized patients from the two outbreaks†.
†The clinical laboratory data on admission were available for 298 patients from the Xiangyang outbreak.
gene and a unique 12 bp deletion in the VA RNA II region when compared to endemic HAdV-7 isolates in China. This finding, along with the epidemiological picture, identified the most likely transmission link between the two outbreaks.

Epidemiological investigation revealed that Xiangyang affiliated hospital played a critical role in the continuation and further transmission of the outbreak. This study identified areas of patient management and infection control that were subsequently strengthened at the hospital as well as the medical clinics at Base A and B, including contact tracing investigations and infection-control interventions to reduce the risk of further transmission, infection and associated morbidity. As a result, rapid and effective control of both outbreaks occurred. Ironically, hospitals remain a central problem of inadvertent infections, since they provide a suitable ecosystem in which a variety of pathogens can emerge and persist, possibly resulting in serious life-threatening infections and significant hospital-associated outbreaks. Pathogen transmission in hospitals and the transmission from hospitals to the community must be recognized as a major driving force for infectious disease epidemics or outbreaks, and more attention should be paid to prevent their potential as a public health threat.

Another major finding of this study is the unexpected identification of a novel HAdV-7 variant by WGS. Genetic variation can greatly drive the emergence of new pathogens or novel variants, resulting in multidrug resistance, and increased adaptation, transmissibility and virulence. New pathogens or novel variants have in more recent years been more frequently detected, seriously threatening public health on both local and global scales. Recent examples illustrating this point are provided by the 2009 influenza pandemic caused by the 2009 H1N1 virus, the worldwide concern about antimicrobial resistance posed by NDM-1 producing bacteria, and the large outbreak of haemolytic uraemic syndrome caused by a strain of E. coli O104:H4. In particular, China is one of the largest reservoirs for the emergence of new pathogens or novel variants, such as the SARS coronavirus in 2003, a novel bunyavirus in 2011, and a novel avian-origin influenza A (H7N9) virus in 2013. In recent years, novel species or variants of HAdVs have been commonly reported, with enhanced virulence and ability of cross-species transmission. Growth kinetics revealed that the outbreak-associated HAdV-7 isolate seemingly had higher viral replication than the local HAdV-7 isolate. It would be interesting to determine what impact the 12 bp deletion in VA RNA II gene has on viral pathogenicity. Continuous surveillance is urgently required to explore the spatiotemporal variation and evolution of HAdVs.

In conclusion, WGS has potential for use in aiding outbreak investigations given its enhanced discriminatory capacity, including identifying the known and emerging pathogens, analyzing the variation...
Figure 3. Phylogenetic analysis of the outbreak isolates based on WGS and individual genes. Panels (A–D) represent the phylogenetic trees of the whole-genome, and individual entire hexon, fiber and penton genes, respectively. Outbreak strains are highlighted in circle (●), and the local strain CDC 228 from Xiangyang City is indicated in triangle (▲). Bootstrap analysis was performed with 1,000 replicates. Representative HAdV strains with GenBank accession numbers are included in the phylogenetic trees, and they are also shown in Table S2.
and evolution, determining the origins and transmission dynamics of the outbreaks, and informing appropriate infection-control interventions. In this study the definite transmission link between two ARD outbreaks caused by a novel HAdV-7 variant by integration of genomic and epidemiological data was demonstrated. HAdV outbreaks have been increasingly reported in recent years in China, concomitant with the variation and evolution of HAdVs, more attention should be given to the potential threat on public health, and strengthen the nationwide and systematic surveillance of HAdVs.

Figure 4. Phylogenetic analysis and pairwise alignment of the outbreak isolates based on DNA polymerase and VA RNA. Panels (A,C) represent the phylogenetic trees of DNA polymerase and VA RNA genes, respectively. Panels (B,D) represent the pairwise alignment analyses of the outbreak strains based on the DNA polymerase amino acid sequences and VA RNA genes, respectively.
Methods

Outbreak investigations. From December 10, 2012 to January 28, 2013, an outbreak of ARD occurred in a recruit training base (Base A) located in Xiangyang City, Hubei Province. A second ARD outbreak was reported in a different training base (Base B), located in Jingmen City, Hubei Province from February 11 to 27, 2013 (Figure S1). Two field investigation teams, comprised of health officials, epidemiologists, technicians and medical staff, were dispatched to investigate each outbreak. Demographic and epidemiologic data were collected through interviews and field observations using a standardized survey. Medical records of hospitalized patients were abstracted by physicians who had been taking care of cases; these were copied and provided to the investigation teams. A probable or confirmed case with HAdV-7 infection were defined as previously described49,50. A close contact was defined as a person that has closely contacted with the suspected or confirmed cases within 8 days; close contacts without ARD symptoms were quarantined for medical observation for 8 days. We defined a severe case as a patient who was admitted to an ICU for serious complications, such as pneumonia, heart and liver failure, shock, and sepsis27,51. Mild cases were reviewed by trained physicians, quarantined and treated in the training base; moderate or severe cases were transferred to Xiangyang affiliated hospital in the Xiangyang City, Hubei Province.

Laboratory investigation. A total of 796 nasopharyngeal specimens were collected during the outbreak period; 328 from Base A (Xiangyang City) and 80 from Base B (Jingmen City) outbreak patients, 160 from the close contacts including 89 from Base A and 71 from Base B, and 228 from non-recruit patients diagnosed with ARD in other hospitals in the local consisting of 128 from Xiangyang and 100 from Jingmen. Blood and sputum samples from severe cases were also collected and analyzed for possible bacterial infections using biochemical test kits (API, bioMerieux, Marcy letoile, France) and 16S rRNA sequencing (Table S1). Additionally, 135 sera samples were obtained from the new recruits of training base A, which were collected for medical examination of the new recruits by the Xiangyang affiliated hospital.

Table 3. Comparative genomic analysis of the outbreak strains with the reference strains and the local strain CDC228§. “–” means no change or not applicable. §The reference genomes include the HAdV-7 strains DG01_2011 (KC440171) and 0901HZ/ShX/2009 (JF800905). Outbreak strains mean the nine HAdV-7 outbreak strains. *Synonymous mutation.
hospital when they reported on December 10, 2012. Specimens were sent to the Institute of Disease Control and Prevention, Academy of Military Medical Sciences (Beijing) central laboratory for identification and confirmation. For the nasopharyngeal samples collected in the early outbreak, multiplex PCR assays, Seeplex® RV12 ACE and Seeplex® Pneumabacter detection kits (Seegene, Seoul, Korea) were used to screen the suspected pathogens including Adenovirus, Coronavirus, Parainfluenza virus, Influenza A virus, Influenza B virus, Respiratory syncytial virus A, Bordetella pertussis, Legionella pneumophila, Chlamydophila pneumoniae, Streptococcus pneumoniae, Haemophilus influenzae, and Mycoplasma pneumonia. A commercial HAdV real-time PCR kit (BioPerfectus Technologies, Jiangsu, China) was used to detect virus from all of the nasopharyngeal samples. ELISA classic adenovirus kits (Institute Virion/Serion GmbH, Würzburg, Germany) were used to detect IgA and IgG antibodies against all serotypes of HAdV from sera samples. Viral culture was performed using Hep-2 cells6,27. Specimens positive for HAdVs were subtyped after direct sequencing the PCR amplified products of partial segments of viral hexon, fiber and virus-associated RNA (V A RNA) genes using primers either described previously or designed as a part of this study (Table S1). Viral growth kinetics were based on a study performed with outbreak-associated representative strain XY1, isolated from the Index case A from Base A, which was measured and compared with the HAdV-7 strain CDC228, recovered from a non-outbreak associated case, as described previously52,53.

De novo genome sequencing and phylogenetic analysis. Ten representative HAdV-7 isolates, including one from a non-recruit ARD case in Xiangyang City, 4 from Base A (Xiangyang) outbreak cases and 5 from Base B (Jingmen) outbreak cases (Table S2), were selected and submitted for de novo WGS using a MiSeq system (Illumina, San Diego, CA, USA). Virus DNAs were extracted from the nasopharyngeal specimens using the QIAamp MinElute Virus Spin Kit (Qiagen) according to the manufacturer's instruction and randomly fragmented to an average size of 350 bp by using a Covaris S200 system (Covaris Inc). Sequencing libraries were then prepared by using the TruSeq standard DNA preparation protocol (Illumina, San Diego, CA, USA). Paired-end 2 × 150 bp sequencing was performed on a Illumina MiSeq system. After filtering low quality bases and trimming of the adapter, De novo assemblies were performed using Velvet 1.2.08 and VelvetOptimiser 2.24 with k values of between 21 and 151 for searching maximal N50 and the largest contig value. Sanger sequencing was used to close gaps and verify ambiguous sequences with primers listed in Table S3, designed by aligning with a reference genome 0901HZ (JF800905). Information on the number of reads sequenced, the number of viral reads and the depth of coverage of the virus was provided in Table S4. Genome annotations were completed based on the annotation of 0901HZ. Comparative genomic analysis was performed to find nucleotide acid changes and amino acid substitutions between the outbreak strains and the local or other HAdV-7 strains circulating in China. The 12-bp deletion in the VA RNA gene and the point mutation in DNA polymerase gene were verified by screening an additional 46 HAdV-7 outbreak strains by PCR and DNA sequencing with primers listed in Table S1. Multiple-sequence alignments were conducted using Clustalw2. Thirty-seven additional sequences of HAdVs were retrieved from GenBank (Table S2), and
neighbor-joining phylogenetic trees with 1000 bootstrap replicates were constructed based on WGS and individual gene sequences by using the MEGA software version 6.06.

All experimental protocols were approved by the Medical Ethics Committee of the Academy of Military Medical Sciences. The methods were carried out in accordance with the approved guidelines. Written informed consents for the use of their clinical samples were obtained from the corresponding patients.

References

1. Marinheiro, J. C. et al. A naturally occurring human adenovirus type 7 variant with a 1743 bp deletion in the E3 cassette. J Gen Virol 92, 2399–2404 (2011).
2. Yusof, M. A. et al. Human adenovirus type 7 outbreak in Police Training Center, Malaysia, 2011. Emerg Infect Dis 18, 852–854 (2012).
3. Seto, D., Chodosh, J., Brister, J. R. & Jones, M. S. Using the whole-genome sequence to characterize and name human adenoviruses. J Virol 85, 5701–5702 (2011).
4. Erdman, D. D. et al. Molecular epidemiology of adenovirus type 7 in the United States, 1966–2000. Emerg Infect Dis 8, 269–277 (2002).
5. Su, X. et al. Complete genome analysis of a novel E3-partial-deleted human adenovirus type 7 strain isolated in Southern China. Virol J 8, 91 (2011).
6. Tang, L., Wang, L., Tan, X. & Xu, W. Adenovirus serotype 7 associated with a severe lower respiratory tract disease outbreak in infants in Shaanxi Province, China. Virol J 8, 23 (2011).
7. Ghanaiem, H. et al. An outbreak of adenovirus type 7 in a residential facility for severely disabled children. Pediatr Infect Dis J 30, 948–952 (2011).
8. Lin, K. H. et al. A two decade survey of respiratory adenovirus in Taiwan: the reemergence of adenovirus types 7 and 4. J Med Virol 73, 274–279 (2004).
9. Rohde, H. et al. Open-source genomic analysis of Shiga-toxin-producing E. coli 0144:H4. N Engl J Med 365, 718–724 (2011).
10. Harris, S. R. et al. Whole-genome sequencing for analysis of an outbreak of meticillin-resistant Staphylococcus aureus: a descriptive study. Lancet Infect Dis 13, 130–136 (2012).
11. Koser, C. U. et al. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. N Engl J Med 366, 2267–2275 (2012).
12. GARDY, J. L. et al. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. N Engl J Med 364, 730–739 (2011).
13. Roeter, A. et al. Whole genome sequencing versus traditional genotyping for investigation of a Mycobacterium tuberculosis outbreak: a longitudinal molecular epidemiological study. PLoS Med 10, e1001387 (2013).
14. Walker, T. M. et al. Whole-genome sequencing to delineate Mycobacterium tuberculosis outbreaks: a retrospective observational study. Lancet Infect Dis 13, 137–146 (2012).
15. Snitkin, E. S. et al. Tracking a hospital outbreak of carbapenem-resistant Klebsiella pneumoniae with whole-genome sequencing. Sci Transl Med 4, 148ra116 (2012).
16. Reuter, S. et al. A pilot study of rapid whole-genome sequencing for the investigation of a Legionella outbreak. BMJ Open 3, e002173 (2013).
17. Tsou, T. P. et al. Community outbreak of adenovirus, Taiwan, 2011. Emerg Infect Dis 18, 1825–1832 (2012).
18. Tang, L., An, J., Yu, P. & Xu, W. Complete genome sequence of human adenovirus type 7 associated with fatal infant pneumonia. Genome Announc 1, e00182–12 (2013).
19. Zhao, S. et al. Re-emergent human adenovirus genome type 7d caused an acute respiratory disease outbreak in Southern China after a twenty-one year absence. Sci Rep 4, 7365 (2014).
20. Cai, X. & Wang, R. J. An outbreak of upper respiratory tract infection caused by adenovirus. Chin J Sch Health 32, 248–249 (2011).
21. Huang, G. et al. Outbreak of febrile respiratory illness associated with human adenovirus type 14p1 in Gansu Province, China. Influenza Other Respir Viruses 7, 1048–1054 (2013).
22. Huang, Y., Zhao, W., Chen, S. S., Zhang, S. M. & He, Y. X. The epidemiological investigation and treatment on an outbreak of pharyngoconjunctival fever caused by adenovirus. Modern Preventive Medicine 40, 1611–1615 (2013).
23. Jiang, L. et al. Serological study on the outbreak of acute respiratory infection caused by Adenovirus 3. Chin J Infect Dis 24, 3–6 (2006).
24. Li, Z. et al. Epidemiological study on an outbreak of respiratory infection due to adenovirus. Prac J Med & Pharm 28, 636–638 (2011).
25. Liu, W. et al. An Outbreak of Children Pharyngo-conjuctival Fever Induced by Type 3 Adenovirus. Disease Surveillance 20, 176–179 (2005).
26. Liu, X., Song, H., Zhang, L., Liu, Q. & Li, X. Epidemiological investigation of an outbreak of acute respiratory infection caused by adenovirus. J Prev Med Chin PLA 22, 199–200 (2004).
27. Xie, L. et al. Two adenovirus serotype 3 outbreaks associated with febrile respiratory disease and pharyngoconjunctival fever in children under 15 years of age in Hangzhou, China, during 2011. J Clin Microbiol 50, 1879–1888 (2012).
28. Yu, P. et al. Outbreak of acute respiratory disease caused by human adenovirus type 7 in a military training camp in Shaanxi, China. Microbiol Immunol 57, 553–560 (2013).
29. Zhang, Q., Seto, D., Cao, B., Zhao, S. & Wan, C. Genome sequence of human adenovirus type 55, a re-emergent acute respiratory disease pathogen in China. J Virol 86, 12441–12442 (2012).
30. Zhu, Z. et al. Outbreak of acute respiratory disease in China caused by B2 species of adenovirus type 11. J Clin Microbiol 47, 697–703 (2009).
31. Mi, Z. et al. Genomic analysis of HAdV-B14 isolate from the outbreak of febrile respiratory infection in China. Genomics 102, 448–453 (2013).
32. Hong, T. et al. Outbreak of acute pharyngoconjunctival fever caused by adenovirus type 7 in Beijing. Virologica Sinica 1, 27–30 (1986).
33. Peng, Z. W. et al. Two outbreaks of acute pharyngoconjunctival fever caused by adenovirus type 7. People's Military Surgeon 8, 24–26 (1983).
34. Zhou, H. M. & Hong, T. An outbreak of acute pharyngoconjunctival fever caused by adenovirus type 7. Acta Academiae Medicinae Xuzhou 5, 12–16 (1985).
35. Wang, Y. H. et al. Investigation on an outbreak of pharyngeal conjunctival fever caused by adenovirus type 7. Zhejiang Journal of Preventive Medicine 5, 17–19 (1993).
36. Cao, R. et al. Clinical features of the initial cases of 2009 pandemic influenza A (H1N1) virus infection in China. N Engl J Med 361, 2507–2517 (2009).
Acknowledgments

This work was supported by the Mega-projects of Science and Technology Research (nos. 2012ZX10004215, 2013ZX10004607 and 2013ZX10004218), the National Nature Science Foundation of China (nos. 31200942, 81371854, 81171554 and 81202252), and the Beijing Science and Technology Nova program (no. xx2013061). We thank John D. Klena (International Emerging Infections Program, US Centers for Disease Control and Prevention, Beijing, China; Global Disease Detection Branch, Division of Global Health Protection, Center for Global Health, Centers for Disease Control and Prevention, Atlanta, USA) for critical review of the manuscript.

Author Contributions

S.Q., Y.W., Y.S., L.H., H.M., Z.Y. and H.S. designed the study. S.Q., N.L., C.L., M.L., Y.L., Z.Y. and H.S. investigated the Xiangyang outbreak, and S.Q., N.L., S.L., L.H., Z.J., H.S., M.L., Y.L. and H.S. investigated the Jingmen outbreak. S.Q., N.L., C.L., M.L., Y.L., Z.Y. and H.S. analyzed data. S.Q., H.L., N.L., J.X., C.Y., J.W., H.L., S.Y., R.H., Z.W., L.J. and Y.W. performed the experiments. S.Q., P.L. and N.L. performed the genomic analyses. S.Q., P.L., N.L., Y.W., C.L., S.L., Z.Y. and H.S. wrote the first draft. All authors contributed to review and revision, and approved the final version.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Qiu, S. et al. Whole-genome Sequencing for Tracing the Transmission Link between Two ARD Outbreaks Caused by a Novel HAdV Serotype 7 Variant, China. Sci. Rep. 5, 13617; doi: 10.1038/srep13617 (2015).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/