Hepatitis C virus genotypes and subtypes circulating in Mainland China

Ying Chen1,2,*, Changshun Yu1,2,*, Xueru Yin3, Xiaolei Guo1, Shangwei Wu1,2 and Jinlin Hou3

The hepatitis C virus (HCV) exhibits global genotypic diversity. HCV genotyping plays an important role in epidemiological studies and clinical management. Herein, we report the results of HCV genotype and subtype detection in a large number of clinical samples, as performed by an independent laboratory in China. In total, four HCV genotypes and 18 subtypes were identified among 32,030 patients from 29 provinces and municipalities in China. Five dominant subtypes were detected from 98.84% of the samples: 1b (n = 16,713, 52.18%), 2a (n = 9,188, 28.69%), 3b (n = 2,261, 7.06%), 6a (n = 2,052, 6.41%) and 3a (n = 1,479, 4.62%). Twelve rare subtypes were detected, of which four (that is, 6b, 6j, 6q and 6r) are reported for the first time in the Chinese population. Genotypes 4, 5 and 7 were not detected. Mixed infections of the dominant subtypes were found in a small portion of samples (n = 65, 0.203%), in the following combinations: 1b–2a, 1b–3b, 1b–6a, 3a–3b, 1b–3a and 2a–6a. No mixed infections with rare subtypes were found. Males, compared with females, showed higher HCV subtype diversity, a lower percentage of HCV1b and 2a and a higher percentage of rare subtypes and mixed infections. Our analyses revealed the comprehensive distribution patterns of HCV genotypes in the general population of Mainland China. HCV genotypic patterns were differentially distributed on the basis of geography, sex and age.

Keywords: genotyping; hepatitis C virus; subtyping

INTRODUCTION

More than 185 million people have been estimated to be infected with hepatitis C virus (HCV) worldwide, including 350,000 cases of death each year due to infection.1 The prevalence of HCV in China has been estimated to be 1%,2 thus making the Chinese HCV carrier population (14 million) one of the largest in the world. HCV is a positive-sense, single-stranded RNA (9.6 kb) virus of the Flaviviridae family that shows high genetic heterogeneity and can be classified into seven major genotypes (HCV 1–7) including 67 confirmed and 20 provisional subtypes, on the basis of phylogenetic analysis of genome sequences.3 Correct determination of HCV genotype is fundamental for evaluation of the epidemiological status of HCV infection and for administration of the most appropriate anti-viral regimen. Furthermore, HCV genotype and subtype are important indicators for antiviral therapeutic responses.4,5 The Chinese Society of Hepatology and the Chinese Society of Infectious Diseases recommend that HCV genotyping should be done prior to starting anti-viral therapy.6

HCV genotype distributions show different geographic patterns.4 HCV genotypes 1, 2 and 3 have a broad global distribution, whereas genotypes 4, 5 and 6 are seen only in certain areas. For example, genotype 4 is mainly found in the Middle East and North Africa, genotype 5 is seemingly confined to South Africa, genotype 6 is primarily distributed in southern China and South-East Asia,7 and genotype 7 has been reported only in central Africa.8 Four HCV genotypes (1, 2, 3 and 6) have been reported in China to date, among which subtypes 1b, 2a, 3b, 6a and 3a are the most prevalent subtypes.9 Several studies have previously been conducted in China but from relatively small populations and narrow geographical areas; thus, the research findings may not be reliable enough to draw solid conclusions.10,11 Therefore, we conducted an investigation of HCV genotype distributions in a variety of populations from different areas in Mainland China. All the tests included in this investigation were performed by a single, independent medical laboratory.

MATERIALS AND METHODS

Study population and sample collection

Between January 2013 and January 2017, a total of 56,627 consecutive serum samples were collected for HCV genotyping and subtyping by Kingmed Diagnostics (Guangzhou, Guangdong). Sex and age data corresponding to the samples were analyzed with the test results from this study. For patients with multiple specimens or orders submitted, only the first specimen was analyzed. HCV quantification was also performed at the same time for 734 of the samples.

TaqMan PCR HCV typing and Sanger sequencing confirmation

HCV RNA was extracted using a commercial extraction kit (QIAamp Viral RNA Mini Kit, QIAGEN, Shanghai, China) according to the manufacturer’s instructions. HCV genotyping was performed on all
samples by using a TIB HCV TaqMan PCR kit (Triplex International Bioscience, Fuzhou, China), which detects the HCV 1b, 2a, 6a, 3a and 3b subtypes, as approved by the China Food and Drug Administration, according to the manufacturers’ instructions. The low detection limit of the TIB HCV TaqMan PCR kit is ~ 1000 IU/mL. Samples that failed detection with the TIB HCV TaqMan kit were subsequently analyzed by Sanger sequencing as previously described. Briefly, two sets of universal primers targeting the core-envelope 1 (C/E1) region of the HCV genome were used for nested PCR amplification and sequencing analysis (Primer set 1: F1, 5′-GCAAC AGGGA ACCTT CCTGG TTGCT C-3′; R1, 5′-CGTAG GGGAC CAGTT CATCA TCAT-3′; Primer set 2: F2, 5′-AACCT TCCTG GTTGC TCTTT CTCTA T-3′, R2, 5′-GTTCA TCATC ATATC CCATG CCAT-3′). The product lengths of primer set 1 and set 2 were 495 and 474 bp, respectively. The subtypes were determined by phylogenetic comparison to reference sequences for HCVs, by using MEGA 4.0.14

HCV quantification
Along with HCV genotyping, an HCV quantification assay was also performed on a group of 760 samples with a COBAS AmpliPrep TaqMan kit. The low detection limit of this kit is 15 IU/mL.

Statistical analysis
The software R 3.2.1 was used to generate heat maps and for cluster analysis. The Simpson’s diversity index \( D = 1 - \sum (N_i/N)^2 \) was calculated to show the differences in distribution patterns between female and male HCV carriers.

RESULTS
Characteristics of HCV genotypes
During the study period, a total of 56 627 serum specimens from 1900 hospitals located in 29 provinces/municipalities, including Anhui, Beijing, Chongqing, Fujian, Gansu, Guangdong, Guangxi, Guizhou, Hainan, Hebei, Heilongjiang, Henan, Hubei, Hunan, Inner-Mongolia, Jiangsu, Jiangxi, Jilin, Liaoning, Ningxia, Qinghai, Shaanxi, Shandong, Shanghai, Shanxi, Sichuan, Tianjin, Yunnan, and Zhejiang, were obtained for HCV quantification testing by the KingMed Center for Clinical Laboratory. Among the serum specimens, 33 010 met the inclusion criteria and were included in this study, of which 32 030 samples (97.03%) were successfully genotyped/subtyped; however, 980 samples (2.97%) failed the genotyping. The distribution of the samples with positive genotyping results is shown in Figure 1 and is listed in detail in Supplementary Materials S1. Sequencing analysis was done for all rare subtypes. For quality control, sequencing analysis was also applied to confirm the results of the major subtypes, including 394 HCV1b, 35 HCV2a, 6 HCV3a, 53 HCV3b and 120 HCV6a cases. For all samples (major subtypes) that were confirmed by sequencing analysis, the results from the TaqMan analysis and sequencing analysis were consistent.

Ninety-eight of the samples that failed genotyping were also used for HCV quantification with a COBAS AmpliPrep TaqMan kit. Out of these 98 samples, HCV-RNA was not detectable in 91 samples and was detectable at only very low concentrations in 7, whereas the highest copy number obtained was 221 IU/mL. Six hundred and thirty-six samples that were positively genotyped were also positive in the HCV quantitative analysis, and the lowest copy number was 174 IU/mL. Four genotypes (1, 2, 3 and 6) and 17 subtypes were detected in total. HCV genotype 1 was the most common genotype \( (n = 16 799, 52.45%) \), followed by genotype 2 \( (n = 9 193, 28.70%) \), genotype 3 \( (n = 3 736, 11.66%) \) and genotype 6 \( (n = 2 332, 7.28%) \). Five major subtypes were detected among the 31 659 samples (98.84%): 1b \( (n = 16 713, 52.18%) \), 2a \( (n = 9 188, 28.69%) \), 3b \( (n = 2 261, 7.06%) \), 6a \( (n = 2 052, 6.41%) \) and 3a \( (n = 1 479, 4.62%) \). Mixed infections with the major subtypes were identified in 65 samples (1b–2a, 1b–3b,
was 0.6688 and 0.5892 (P < 0.001) of HCV2 and showed an increasing trend from north to south (Henan: 0.21%; Hubei: 4.42%; Hunan: 7.10E−234), respectively.

There were 636 cases with both positive genotyping results and positive viral load results. The relationships between viral loads and different percentages of HCV6 and an increasing trend from north to south (Henan: 38.88%; Hubei: 22.09%; Hunan: 0.89%).

The prevalence of HCV genotype 6 gradually increased from northern to southern China (Figure 3B). HCV2 was the dominant genotype in the northeast (50.93%) but accounted for less than 10% in southwestern (5.23%) and southern China (8.23%). The three provinces in central China showed very different percentages (P < 0.001) of HCV2 and a decreasing trend from north to south (Henan: 38.88%; Hubei: 22.09%; Hunan: 0.89%).

The distribution of genotype 2 gradually decreased from northern to southern China (Figure 3B). The three provinces in central China showed very different percentages (P < 0.001) of HCV2 and a decreasing trend from north to south (Henan: 0.21%; Hubei: 4.42%; Hunan: 21.35%). The seven provinces in eastern China also showed very different percentages (P < 0.001) of HCV6 and an increasing trend from north to south (Shandong, Anhui and Jiangsu: lower than 2%; Shanghai: 8.09%; Fujian, Zhejiang and Jiangxi: higher than 16%).

From the southwest to the northeast, the distribution of HCV genotype 3 exhibited a gradually decreasing trend (Figure 3C). The percentage of HCV3 was 38.7% in the southwest, 13.69% in the south, 8.09%; Fujian, Zhejiang and Jiangxi: higher than 16%). From the southwest to the northeast, the distribution of HCV genotype 3 exhibited a gradually decreasing trend (Figure 3C). The percentage of HCV3 was 38.7% in the southwest, 13.69% in the south, 11.11% in the northwest, and 2.9% in the northeast.

In general, the distributions of HCV genotypes in the southern provinces of China displayed higher diversity than those in the northern provinces. In all of the northern provinces, HCV1b and 2a were the two predominant subtypes and accounted for 86% or more of the total (Supplementary Materials S1 and Figure 2), whereas in the southern provinces, HCV3b and 6a represented the most substantial proportions of the HCVs.

**Table 1 HCV genotyping and viral load**

| Subtypes (number) | <10⁴ | 10⁴–10⁵ | 10⁵–10⁶ | 10⁶–10⁷ | 10⁷–10⁸ | Geomean | Viral load (IU/mL) |
|------------------|------|---------|---------|---------|---------|---------|-------------------|
| HCV1b (n=381)    | 5.51%| 10.50%  | 30.18%  | 51.66%  | 2.36%   | 2.08E+06 |
| HCV2a (n=140)    | 5.71%| 18.57%  | 35.90%  | 38.57%  | 2.14%   | 1.88E+06 |
| HCV3a (n=25)     | 4.00%| 16.60%  | 24.00%  | 32.00%  | 4.00%   | 5.27E+06 |
| HCV3b (n=45)     | 0.08%| 22.22%  | 24.44%  | 33.11%  | 22.22%  | 5.30E+06 |
| HCV6a (n=41)     | 4.88%| 7.32%   | 24.99%  | 43.66%  | 19.31%  | 5.53E+06 |

We used red color in the table through Excel to show the relative viral load. The darker the red color, the higher the viral load. However, the viral load of minor subtypes were not compared because of the lack of samples.

**Figure 2** Hepatitis C virus (HCV) subtype distribution of males and females. Minor—minor subtypes; Mixed—mixed infections. The number of male carriers was much greater than females in this study, and male carriers showed a higher level of diversity. The Simpson’s diversity index of male and female carriers was 0.6688 and 0.5892 (P = 1.23E−234), respectively.
The distributions of subtypes of HCV genotypes 1, 2, and 3 are shown in Table 2. For HCV genotype 1, HCV1b was the dominant subtype, and HCV1a was found mainly in southern (1.97%), northwestern (0.98%), and southwestern China (0.82%); Hainan (8.70%; south China) and Chongqing (4.88%; southwest) had the highest prevalence of HCV1a in this study. For HCV genotype 2, HCV2a was the dominant subtype, and HCV2b was found mainly in Shanghai (5.88%) and Fujian (4.17%). For HCV genotype 3, subtype 3b was slightly more prevalent than 3a in the northeast, southwest, east, and north of China, whereas in the northwest, 3a was more prevalent than 3b in southern and central China, the prevalence of 3a and 3b was highly similar.

HCV genotype 6 strains were detected in 2332 samples (7.28%), with the most prevalent being 6a (2045/2332, 87.69%) followed by 6n (226/2332, 9.69%), 6u (36/2332, 1.54%), 6g (4 cases), 6w (2 cases), 6c (1 case), 6b (1 case), 6j (1 case), 6q (1 case), and 6r (1 case), and 5 strains that could not be subtyped (Supplementary Materials S1). In most provinces, 6a was the most common subtype of the HCV6 genotypes, whereas in Yunnan, 6n and 6a represented 63.6% and 23.6%, respectively. In Sichuan and Guizhou, two provinces adjacent to Yunnan, 6n was rare and accounted for only 8.5% and 4.0%, respectively. The proportion of 6n cases in eastern China was 10.7% (Table 3).

Mixed infections
Mixed infections with two genotypes or subtypes were identified in 65 samples from 13 provinces (Chongqing, Guangdong, Guangxi, Heilongjiang, Henan, Hunan, Inner-Mongolia, Jilin, Liaoning, Shaanxi, Sichuan, Tianjin, Yunnan), accounting for 0.203% of all the HCV-positive samples. The majority of combinations of HCV genotype-subtype in the mixed infections were 1b–2a (31 cases, 47.7%), followed by 1b–3b (20 cases, 32.3%), 1b–6a (6 cases, 9.2%), 3a–3b (4, 6.2%), 1b–3a (3 cases, 4.6%) and 2a–6a (1 case, 1.5%). Of the 65 mixed infections, 48 samples (73.84%) were from males, and the other 17 (26.16%) were from females (Supplementary Materials S1). Mixed infections with the rare subtypes were not observed in our study.

DISCUSSION
Chronic HCV infection is the leading cause of end-stage liver disease and is associated with a 1%-5% annual risk of developing hepatocellular carcinoma (HCC) after developing cirrhosis and with an incidence of approximately 10%-20% over 20-30 years of HCV infection. A previous study has shown that elimination of HCV can be achieved in the coming decades with concentrated strategies to improve diagnosis and treatment; however, there is not yet a comprehensive understanding of HCV, especially concerning genotypes with potential for use in the development of targeted and efficient strategies for HDV elimination.

This study revealed that five subtypes (1b, 2a, 3b, 6a and 3a) comprise 98.67% of the HCV infections in Mainland China, and the data published by others have also shown similar results. Regional differences in the distribution of HCV genotypes were also demonstrated. HCV1b and 2a comprise the greatest proportion of HCV.
Subtypes in northern China, whereas the subtype composition showed more diversity in southern China. Subtype 1b is dominant in most areas of China. Subtype 2a is common in northern China and dominant in the northeast provinces (that is, Heilongjiang, Jilin and Liaoning), but is much less common in southern China. Subtype 3b is dominant in Shanghai (34.62%) and Yunnan (44.95%). These results may be explained by the unique locations and diversification of the population. Shanghai is located on the north–south border and is well known as a free-trade zone, which results in a complex transient population. Yunnan borders Vietnam, Laos and Burma, and its ethnic minority groups might lead to remarkable diversity. Southeast Asia is one of the most important drug-producing bases worldwide, and southwestern China is an important transit area for drug smuggling. Thus, the proportion of drug abusers in the south is much higher than in the north. On the basis of previous studies, intravenous drug abuse associated with genotype 3 has become the major risk factor for HCV infection, thus potentially explaining why 3b is the predominant subtype in Yunnan. Although HCV genotype 6 has been widely reported in Asia, the prevalence has not been well studied in China. A previous study has demonstrated that HCV genotype 6 has significant genetic diversity and probably contains at least 22 (6a–6v) subtypes. Our results indicated that HCV genotype 6 is common in the southern parts of Mainland China (Hainan, Guangxi, Guangdong, Guizhou, Chongqing, Hunan, Sichuan and Yunnan) and is relatively rare in north China. The current study demonstrated that HCV6a is the dominant genotype 6 subtype in China; however, in Yunnan Province, the prevalence of HCV subtype 6n is higher than 6a, in contrast with the general distribution of this subtype. Dashan Hu et al have reported that subtype 1b is dominant in the Guangdong Province from 2004–2011, followed by HCV2a; our data show that

| Area          | HCV1  | HCV2  | HCV3  |
|---------------|-------|-------|-------|
|               | 1a    | 1b    | 2a    | 2b    | 1b    | 2b    | 3a   | 3b   | n=   |
| Central       |       |       |       |       |       |       |      |      |      |
| Henan         | 0.0011| 0.9998| 4609  | 0.9992| 0.0008| 2487  |       | 0.4578| 0.5422| 166  |
| Hubei         | 0.0003| 0.9997| 3779  | 0.9992| 0.0008| 2423  | 0.6667| 0.3333| 6     |
| Hunan         | 0.0061| 0.9939| 653   | 1.0000| 0.0000| 9     | 0.4328| 0.5672| 134   |
| East          |       |       |       |       |       |       |      |      |      |
| Anhui         | 0.0000| 1.0000| 425   | 1.0000| 0.0000| 99    | 0.4545| 0.5455| 22    |
| Fujian        | 0.0000| 1.0000| 52    | 0.9583| 0.0417| 24    | 0.4000| 0.6000| 10    |
| Jiangsu       | 0.0023| 0.9777| 430   | 1.0000| 0.0000| 39    | 0.4000| 0.6000| 15    |
| Jiangxi       | 0.0000| 1.0000| 43    | NA    | NA    | 0     | 0.3333| 0.6667| 3     |
| Shandong      | 0.0000| 1.0000| 53    | 1.0000| 0.0000| 311   | 0.3000| 0.7000| 30    |
| Shanghai      | 0.0128| 0.9872| 78    | 0.9412| 0.0588| 17    | 0.2969| 0.7031| 64    |
| Zhejiang      | 0.0000| 1.0000| 46    | 1.0000| 0.0000| 6     | 0.2000| 0.8000| 10    |
| North         |       |       |       |       |       |       |      |      |      |
| Hebei         | 0.0000| 1.0000| 1089  | 1.0000| 0.0000| 681   | 0.5000| 0.5000| 6     |
| Inner-Mongolia | 0.0000| 1.0000| 241   | 1.0000| 0.0000| 205   | 0.5454| 0.4545| 11    |
| Shanxi        | 0.0000| 1.0000| 653   | 1.0000| 0.0000| 231   | 0.4211| 0.5789| 19    |
| Tianjin       | 0.0000| 1.0000| 53    | 1.0000| 0.0000| 14    | 0.3148| 0.6852| 54    |
| Northeast     |       |       |       |       |       |       |      |      |      |
| Heilongjiang  | 0.0000| 1.0000| 653   | 1.0000| 0.0000| 744   | 0.3889| 0.6111| 36    |
| Jilin         | 0.0016| 0.9894| 621   | 0.9899| 0.0111| 874   | 0.3333| 0.6667| 12    |
| Liao ning     | 0.0000| 1.0000| 1104  | 1.0000| 0.0000| 1066  | 0.2019| 0.7981| 104   |
| Northwest     |       |       |       |       |       |       |      |      |      |
| Gansu         | 0.0000| 1.0000| 458   | 1.0000| 0.0000| 486   | 0.7487| 0.2539| 386   |
| Ningxia       | 0.0000| 1.0000| 16    | 1.0000| 0.0000| 6     | 0.5758| 0.4242| 33    |
| Qinghai       | 0.0000| 1.0000| 10    | 1.0000| 0.0000| 10    | 0.5000| 0.5000| 6     |
| Shaanxi       | 0.0143| 0.9857| 1049  | 1.0000| 0.0000| 1287  | 0.0000| 1.0000| 6     |
| South         |       |       |       |       |       |       |      |      |      |
| Guangdong     | 0.0150| 0.9850| 1730  | 1.0000| 0.0000| 326   | 0.5749| 0.4251| 487   |
| Guangxi       | 0.0252| 0.9748| 357   | 1.0000| 0.0000| 14    | 0.4706| 0.5294| 51    |
| Hainan        | 0.0870| 0.9130| 92    | 1.0000| 0.0000| 15    | 0.2885| 0.7115| 52    |
| Southwest     |       |       |       |       |       |       |      |      |      |
| Chongqing     | 0.0488| 0.9512| 205   | 1.0000| 0.0000| 96    | 0.4207| 0.5793| 164   |
| Guizhou       | 0.0175| 0.9825| 229   | 1.0000| 0.0000| 15    | 0.4050| 0.6000| 200   |
| Sichuan       | 0.0028| 0.9972| 1404  | 1.0000| 0.0000| 136   | 0.3541| 0.6459| 418   |
| Yunnan        | 0.0033| 0.9967| 612   | 1.0000| 0.0000| 44    | 0.2567| 0.7447| 1371  |
HCV6a has already replaced 2a as the second most common subtype with a rate of 24.16%. That the transmission of HCV type 6 primarily by injectable drug use may explain the HCV6 epidemics that have spread throughout southern China. In addition, transient populations, new modes of transmission and increasing global travel may change the proportion of HCV genotypes in Mainland China.

In general, HCV genotyping relies on the amplification of a specific portion of the HCV genome and is followed by a detection step such as restriction fragment length polymorphism analysis, line probe reverse hybridization, TaqMan PCR or sequence analysis. Core/E1 and NS5B are the two most commonly used target regions. Qingxian Cai et al (2012) have compared the suitability of the core region and NS5B sequences as the target regions and have concluded that the core region is more suitable for clinical application and has greater amplification sensitivity. In addition, HCV genome sequencing improves understanding of resistance-associated variants, including those occurring naturally before treatment, those acquired by transmission at HCV infection, and those emerging after virologic failure.

In our practice, all HCV-positive samples (defined by virus load > 1000 IU/mL) of HCV subtypes 1b, 2a, 6a, 3a and 3b were successfully genotyped by TaqMan PCR. Validation was conducted by comparison of the TaqMan PCR results with sequence analysis data, and some of the validation data are currently published. For the other genotypes (that is, 6n, 1a, 6u, 6g, 6w, 6v, 2b, 6b, 6e, 6j, 6q and 6r), sequence analysis of the core/E1 region was used as the target, and all but three samples (3/243) were successfully subtyped. For the three samples that could not be subtyped, phylogenetic analysis showed that they could be grouped to HCV genotype 6 but differed from all the known HCV 6 subtypes (data not shown). These samples might belong to other rare HCV6 subtypes; therefore, further effort toward identification should be performed.

When chronic hepatitis C is diagnosed and hepatic fibrosis is present, anti-viral therapies should be initiated immediately. Different treatment should be considered for the patients with different HCV genotypes; thus, the HCV genotype must be determined prior to anti-viral therapy. On the basis of the pivotal clinical trials for standard treatment, the sustainable virologic response (SVR) rate for genotype 1 is 91%–100%, and the SVR for genotype 4 patients is 100%. The SVR rate for genotype 2 is 95% and the SVR for

### Table 3 Distribution of HCV 6 subtypes

| Provinces   | 6a | 6n | 6u | 6 | 6g | 6w | 6v | 6xa | 6b | 6e | 6j | 6q | 6r |
|-------------|----|----|----|---|----|----|----|-----|----|----|----|----|----|
| **Central** | 100|     |    |   |    |    |    |     |    |    |    |    |    |
| Henan       | 100|     |    |   |    |    |    |     |    |    |    |    |    |
| Hubei       | 100|     |    |   |    |    |    |     |    |    |    |    |    |
| Hunan       | 100|     |    |   |    |    |    |     |    |    |    |    |    |
| **East**    | 86.7| 10.7|   | 1.3| 1.3|    |    |     |    |    |    |    |    |
| Anhui       | 100|     |    |   |    |    |    |     |    |    |    |    |    |
| Fujian      | 87.5|     |    | 6.3|    |    |    |     |    |    |    |    |    |
| Jiangsu     | 88.9|     |    | 11.1|    |    |    |     |    |    |    |    |    |
| Jiangxi     | 100|     |    |   |    |    |    |     |    |    |    |    |    |
| Shandong    | 87.5|     |    | 12.5|    |    |    |     |    |    |    |    |    |
| Shanghai    | 78.6|     |    | 21.4|    |    |    |     |    |    |    |    |    |
| Zhejiang    | 76.9|     |    | 23.1|    |    |    |     |    |    |    |    |    |
| **North**   | 94.4| 5.6|    |   |    |    |    |     |    |    |    |    |    |
| Hebei       | 100|     |    |   |    |    |    |     |    |    |    |    |    |
| Inner-Mongolia | 100|     |    |   |    |    |    |     |    |    |    |    |    |
| Shanxi      | 85.7|     |    | 14.3|    |    |    |     |    |    |    |    |    |
| Tianjin     | 100|     |    |   |    |    |    |     |    |    |    |    |    |
| **Northeast** | 92.7| 7.3|    |   |    |    |    |     |    |    |    |    |    |
| Heilongjiang| 100|     |    |   |    |    |    |     |    |    |    |    |    |
| Jilin       | 90.0|     |    | 10.0|    |    |    |     |    |    |    |    |    |
| Liaoning    | 92.6|     |    | 7.4|    |    |    |     |    |    |    |    |    |
| **Northwest** | 92.9| 3.6| 1.2| 1.2| 1.2|     |    |     |    |    |    |    |    |
| Gansu       | 50.0|     |    | 50.0|    |    |    |     |    |    |    |    |    |
| Ningxia     | 100|     |    |   |    |    |    |     |    |    |    |    |    |
| Shaanxi     | 93.6|     |    | 3.8| 1.3|   |    |     |    |    |    |    |    |
| **South**   | 99.1| 0.3|    | 0.2| 0.2| 0.1| 0.1| 0.1| 0.1| 0.1| 0.1|    |    |
| Guangdong   | 99.2|     |    | 0.2|    | 0.1| 0.1| 0.1| 0.1| 0.1|    |    |    |
| Guangxi     | 99.3|     |    | 0.7|    |    |    |     |    |    |    |    |    |
| Hainan      | 97.3|     |    | 2.7|    |    |    |     |    |    |    |    |    |
| **Southwest** | 63.6| 30.4| 4.8| 0.7| 0.3| 0.1|    |     |    |    |    |    |    |
| Chongqing   | 100|     |    |   |    |    |    |     |    |    |    |    |    |
| Guizhou     | 96.0|     |    | 4.0|    |    |    |     |    |    |    |    |    |
| Sichuan     | 89.8|     |    | 8.5| 1.7|   |    |     |    |    |    |    |    |
| Yunnan      | 23.6| 63.6| 10.1| 1.7| 0.7| 0.3|    |     |    |    |    |    |    |
| **Total**   | 87.99| 9.69| 1.46| 0.21| 0.13| 0.13| 0.09| 0.09| 0.04| 0.04| 0.04| 0.04| 0.04 | 2332
Supplementary Information for this article can be found on the Emerging Microbes & Infections website (http://www.nature.com/emi)