INTRODUCTION

Varicella-zoster virus (VZV) is a member of the Herpesviridae family that causes varicella (chickenpox) and zoster (shingles). The severity of the disease depends on the individuals' age upon acquiring infection, and the significantly young and old vulnerable individuals are associated with an increased risk of experiencing the disease (1,2).

The VZV genome is considered as the most conserved genome and contains approximately 70 open reading frames (ORFs). The genomic variation within the genome is approximately 0.1% (3,4). Based on single-nucleotide polymorphism (SNP) analysis, several geographical distinct genotypes have been reported previously (5). Comparison of the deoxyribonucleic acid (DNA) sequences of the Oka vaccine strain (vOka) and the parental Oka strain (pOka) showed that there were only 42 nucleotide differences along the whole genome, and 15 were located in the transactivator ORF 62 (6,7).

The varicella vaccine has been used in China for more than 30 years, and the vaccine strains from VZV-infected individuals are distributed epidemiologically (8). Considering the advancements in molecular genotyping methods, the molecular epidemiology of VZV in different geographical regions can be monitored. However, VZV isolates or samples from China have rare genotypes; specifically, the detailed geographical differences across the country have not been investigated. In the current study, we aimed to compare and analyze the genetic variations and sequences in Shanghai (eastern China) and Urumqi (western China, close to mid-Asia). Our results demonstrated that the 41 Shanghai strains were significantly highly homogenous, and the 41 Urumqi strains were highly variable from the wild-type strains.

MATERIALS AND METHODS

Patient and clinical enrollment: The study cohort included 88 patients clinically diagnosed with varicella and zoster infection. A total of 41 samples were collected in Tongren Hospital (Shanghai, China) and 47 from Karamay People’s Hospital (Karamay, China). Samples were collected as swabs of vesicle fluid in viral transport medium from skin lesions and transported to the laboratory within 4 h on wet ice (9). Virus was recovered by the inoculation of fluid from a single vesicle on a human embryonic lung fibroblast (HEL) monolayer and was confirmed by immunofluorescence with an anti-VZV gE monoclonal antibody (Biodesign International, ME, USA) (10). The study was approved by the Research Ethics Committee of The Tongren Hospital, Shanghai Jiao Tong University School of Medicine, and all patients provided written informed consent for inclusion in the study.
reaction (PCR): DNA was extracted using the PureLink Genomic DNA Mini kit (Life Technologies, Carlsbad, CA, USA) following the manufacturer’s instructions. All patients were referred to the dermatology department and were enrolled in the local hospitals. HELFs infected with the vOka, pOka, and an American wild-type strain (macrolide-lincosamide-streptogramin) served as the control. The VZV gene ORFs 1, 6, 12, 16, 17, 21, 22, 35, 37, 38, 50, 54, 55, 56, 60, and 66 were amplified by PCR (11). Primer sequences are listed in Table 1.

DNA sequencing and phylogenetic analysis: The electrophoresis products were purified using an agarose purification kit (Takara Biotechnology, Shiga, Japan) and were sequenced using a sequencing kit (Applied Biosystems, Foster city, CA, USA). An ABI Prism 7700 DNA Analyzer (Perkin Elmer Applied Biosystems, Foster city, CA, USA) was used to perform all the sequencing (12). The sequencing was compared with the published strains from PubMed using MEGA 4.02. SNP profiles were established to construct the phylogenetic tree. Evolutionary analysis was conducted using MEGA7 (13).

**RESULTS**

**VZV Oka stain SNP analysis:** ORF38 (Pst I, SNP69349), ORF54 (Bgl I, SNP95241), and ORF62 (Sna I, SNP106262) were the SNPs that were initially
used to distinguish between strains. VZV strains including \textit{Pst I- Bgl I+ Sma I-} and \textit{Pst I- Bgl I+ Sma I+} were previously reported in China. Here, we amplified ORF38, ORF54, and ORF62 and found that all the strains in the study were \textit{Pst I+ Bgl I+ Sma I-} (Table 2).

**VZV clade distribution in Shanghai and Urumqi:**

We used the VZV clade analysis (Fig. 1) to compare the sequence of ORFs 1, 6, 12, 16, 17, 21, 22, 35, 37, 38, 50, 54, 55, 56, 60, and 66 in all 88 samples (Table 2) (14). The majority of all samples (81.8%, 72/88) belonged to Clade 2. All Shanghai samples belonged to Clade 2, and 70.2% (33/47) of Urumqi samples belonged to Clade 2 (Fig. 2). We found a consecutive nucleotide (CGG) insertion in ORF1, which was previously reported in Chinese strains, confirming the existence of the consecutive nucleotide insertion in Clade 2 ORF60. Clade 5 comprised 23.4% (11/47) of all the samples. Interestingly, a total of 3 Urumqi samples (6.4%) could not be categorized in the existing clades (Fig. 2). Considering the geographical location of Urumqi, the strains possibly underwent evolution; hence, the prevalence of European (Clade 1), mid-Asian (Clade 5), and eastern Asian (Clade 2) clades caused high-frequency gene recombination.

A phylogenetic tree demonstrating VZV subclades, variants, and unclassified strains is illustrated in Fig. 3. The phylogenetic analysis illustrated that all the 41 Shanghai strains were highly homologous. Isolated

<table>
<thead>
<tr>
<th>Location</th>
<th>Number</th>
<th>\textit{Pst I-}</th>
<th>\textit{Bgl I+}</th>
<th>\textit{Sma I+}</th>
<th>Clade 2</th>
<th>Clade 5</th>
<th>Unclassified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shanghai</td>
<td>41</td>
<td>0</td>
<td>41</td>
<td>0</td>
<td>41</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Urumqi</td>
<td>47</td>
<td>0</td>
<td>47</td>
<td>0</td>
<td>31</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>88</td>
<td>0</td>
<td>88</td>
<td>0</td>
<td>72</td>
<td>11</td>
<td>3</td>
</tr>
</tbody>
</table>

![Fig. 1. VZV clades analysis that was utilized in the study.](image)

![Fig. 2. (Color online) Analysis of the strain-associated SNPs of the Shanghai and Urumqi samples. S-all represented SNPs in Shanghai samples. Variations existed in Urumqi samples but not identified in Shanghai samples. U-others presented SNPs in Clades 2 that were not included.](image)
Geographical Separation and VZV Complexity

VZV is a highly infectious disease, causing significant outbreaks of varicella in various populations. The present study described the largest study on VZV genotyping in China, and the major finding of the study was that the geographical region contributes significantly to VZV genomic variations.

Novel CGG insertions were present in the VZV isolates. Additionally, new nucleotide substitutions were observed in Chinese VZV isolates. According to the worldwide distribution of VZV clades, Clades 1 and 3 dominate in European and American countries, whereas Clades 4 and 5 are mainly observed in the immigrants of African origins. Consistent with the result of a previous study, we found that Clade 2 is the dominant clade in Chinese populations (15). Clade 5 in Urumqi strains indicated the genetic recombination events during VZV evolution from the history of immigration (Fig. 2) (10).

Phylogenetic analysis showed that all the 41 Shanghai strains included in this study were highly homologous (Fig. 3) (16). This was not surprising because the strains were isolated in Shanghai, where 99.9% of the residents are native Chinese. These results were consistent with the results of the previous observations (17). However, it was slightly different from the predominant genotype in suburban Shanghai in the previous report (18).

Genetic variations are observed in the strains from Urumqi, where the populations are mixed with immigrants, including Caucasian and Mongolian. These results confirmed that VZV strain distribution is associated with immigration patterns (16,19). On the contrary, phylogenetic analysis using the full nucleotide sequences identified five distinct clades, consistent with the result of the previous findings in other countries (20). Concatenated amino acid sequences and coding nucleotide sequences can be utilized in future phylogenetic analysis. The undetermined clades in Urumqi suggest that new clade or subclade could be extended to western China, which is geographically close to Europe and mid-Asia.

Phylogenetic analyses using the sequences of the individual ORFs suggested that the 12 ORFs might be important in distinguishing vaccine strains from clinical strains. ORF0, also known as ORFS/L, is considered to be essential for VZV growth and encodes a membrane protein with 129 amino acid residues, which is essential for vesicular trafficking and altering cell adhesion molecules in infected cells (21). Genetic variations observed in the current study will serve as the molecular markers in future epidemiological studies. Four or five clusters are differentiated using the SNPs and insertions, which are within the residing local geographical areas.

In conclusion, the genetic variations in VZV strains significantly highlight the ongoing status of VZV infection, and the genetic variations in VZV strains during continuing disease epidemics should be highly considered for future vaccine strategies and clinical management.
Acknowledgments This work was supported by the Changning District Committee of Science and Technology (CNKW2016Y04), the Shanghai Jiaotong University "Medical Professionals Cross Fund" (YG2017QN59) and Science and Technology Innovation Project of Shanghai Jiao Tong University School of Medicine (16XJ21011).

Conflict of interest None to declare.

References