Human Papillomavirus Type 16 Variant Analysis of E6, E7, and L1 Genes and Long Control Region in Identification of Cervical Carcinomas in Patients in Northeast China†

Qinglong Shang,1# Yan Wang,1# Yong Fang,1# Lanlan Wei,1 Sijia Chen,1 Yuhui Sun,2 Baoxin Li,3 Fengmin Zhang,1 and Hongxi Gu1*

Department of Microbiology, Harbin Medical University, Heilongjiang Key Laboratory of Infection and Immunity, Heilongjiang Province, and Pathogenic-Biological Key Laboratory, Heilongjiang Higher Education Institutions, Harbin 150081, People’s Republic of China1; and Department of Gynecology, First Affiliated Hospital, Harbin Medical University, Harbin 150001, People’s Republic of China2; and Department of Pharmacology, Harbin Medical University, and Bio-pharmacological Key Laboratory of Heilongjiang Province, Harbin 150081, People’s Republic of China3

Received 1 November 2010/Returned for modification 31 January 2011/Accepted 3 May 2011

Human papillomavirus type 16 (HPV 16) plays a cardinal role in the pathogenesis of cervical cancer. HPV 16 has intratypic variants which show different geographical distributions and different oncogenic potentials. To analyze the presence of sequence variations of HPV 16 variants in northeast China, 71 cervical carcinomas were identified by HPV typing. HPV 16-positive specimens were analyzed by PCR-directed sequencing in the E6, E7, and L1 genes and the LCR (long control region). The variation data were compared with those of neighboring districts. In this hospital-based study, HPV 16 was the most common type (73.24%). In HPV 16-positive specimens, 67.31% belonged to the European (E) lineage, while 32.69% were Asian (As) variants. The Asian-American (AA), African-1 (Af-1), African-2 (Af-2), and northern American (NA) lineages were not detected. The most frequently observed variation sites were T178G (32.69%) in E6; A647G (34.62%), G666A (38.46%), and T846C (32.69%) in E7; C6826T (36.17%) and G7060A (61.70%) in L1; and G7521A (98.08%) in the LCR. The most prevalent amino acid variations were D25E in E6 and N29S in E7. In addition, 28 novel variations of HPV 16 were reported. Some covariations between different genes were obtained. In this study, HPV 16 variants belonged to the European lineage and the Asian lineage. Compared with neighboring districts, the distribution of HPV 16 variants in northeast China had a typical pattern. As the first report on HPV 16 variants in northeast China, it should be helpful for designing a HPV vaccine and HPV vaccination program in China.

Human papillomavirus type 16 (HPV 16) is the primary etiology of cervical cancer, which is the second most common type of cancer in women worldwide (29). HPV 16 variants, which vary by ≤2% from HPV 16 prototype nucleotide sequences, have been identified as the following six phylogenetic branches: European (E), Asian (As), Asian-American (AA), African-1 (Af-1), African-2 (Af-2), and northern American (NA) variants (18, 52). Several researchers had reported correlations between specific HPV 16 variants and persistent viral infection, followed by the development of malignant lesions (3, 4, 16, 37, 49, 50). Non-European variants were found to be associated with an excess risk of cervical cancer (37). These variants had been found to show different geographic distributions, while some sequence variations had oncogenic potentials. In HPV 16 variants, the L83V mutation in E6 in the Swedish and Italian populations and D25E in E6 in the Japanese population were reported to be associated with the progression of cervical carcinoma (27, 53, 54). The HPV 16 Asian variant was the major causative agent associated with cervical cancer in Japan and northeast Thailand (10).

China has one of the highest incidence rates of cervical cancer, and approximately 13,200 new cases are reported every year (33). Recent data showed that the mortality of cervical cancer was 2.55/100,000 people in the China mainland. The highest mortality existed in northwest China (10.69/100,000 people in Xinjiang and 9.36/100,000 in Gansu) and central China (4.98/100,000 people in Hunan and 4.90/100,000 in Jiangxi). In southwest China, the mortality of cervical cancer was 1.53/100,000 people in Sichuan. In northeast China, the mortality of cervical cancer was 2.12/100,000 people in Heilongjiang, 1.97/100,000 in Jilin, and 1.35/100,000 in Liaoning (57).

As the most prevalent genotype, the prevalences of HPV 16 among different geographical regions in China were similar (7, 42, 45, 56). However, the distribution of HPV 16 variants in China was studied less. The Asian lineage was reported in southwest China (31.0% in Sichuan) and southern China (50.6% in Hong Kong). The European prototype was reported in southwest China (23.0% in Sichuan), central China (15.52% in Hubei), and southern China (30.0% in Hong Kong) (5, 6, 35). No data on HPV 16 variants and sequence variations were reported in northern China and northeast China.

To characterize the prevalence of HPV 16 variants and nu-
clotid polymorphisms in northeast China, we have investigated the HPV 16 E6, E7, and L1 genes and the long control region (LCR) in cervical carcinomas. The results were compared with data reported from other neighboring areas.

MATERIALS AND METHODS

Preparation of clinical specimens. Recruitment of study subjects was conducted from June 2005 through December 2006. The subjects were women with newly diagnosed invasive cervical carcinoma (ICC) that was histological confirmed at Tumor Hospital of Harbin Medical University in Heilongjiang province. The inpatients came from Harbin, Jilin, and the north Inner Mongolia region, which lies in the north district of northeast China. For our database, 71 patients were identified as having squamous carcinomas, except sample 13 (adenocarcinoma) and sample 53 (adenosquamous carcinoma). Cervical samples were obtained from women undergoing surgery. Following the cervical punch biopsy, the biopsy tissue was subjected to histological processing and examined by two independent pathologists. Multiple aliquots were cut and stored at −70°C. The study protocol was approved by the institutional ethical committee. Written informed consent was obtained from each study subject.

The efficiency of PCR-amplified DNA was checked by PCR amplification of the 8-globin gene (forward primer, 5'-CAAC-CCCTGTC-CCCTGTC-3' and reverse primer, 5'-AAGCCAAGCCACACGTTTCC-3'). Amplification without a DNA template was used to monitor contamination in both HPV and β-globin reactions. PCR products were sequenced using a DNA Engine Peltier thermal cycler (Bio-Rad). A 260-bp amplicon was determined with 2% agarose electrophoresis (26).

Detection and typing of HPV. DNA was extracted by a standard proteinase K digestion and phenol–chloroform extraction method (40). The DNA samples were determined by spectrophotometric quantitation and 0.7% agarose electrophoresis. All specimens were first subjected to PCR amplification with HPV consensus primers GPs+ (5'-TTCTTCTGTGGTAGTACACRTGCATAGC-3') and GPs+ (5'-GAAAATAAAGCTGAAATATTCATTTG-3'). HPV 16- and 18-positive DNA was identified by specific primers (5'-AAGGGCAGTGAACCGAAGGT-3' and 5'-AACTTCATCCACGTTCACC-3'). Amplification without a DNA template was used to monitor contamination in both HPV and β-globin reactions. PCR products were sequenced on a DNA Engine Peltier thermal cycler (Bio-Rad). A 267-bp amplicon was determined with 2% agarose electrophoresis (26).

Sequence analysis of the E6, E7, and L1 genes and the LCR of HPV 16 by PCR-directed sequencing. HPV 16 L1-E7 fragments and LCR-E7 fragments were amplified by high-fidelity PCR and sequenced. The L1-E7 fragment, with an amplicon size of 3.3 kb, was amplified with primers 5'-TTCTGTCCCAGTGGCAGG-3' and 5'-TTCTGTCCAGCTCG-3'. Amplification without a DNA template was assessed by PCR amplification of the β-globin gene (forward primer, 5'-CAAC-CCCTGTC-CCCTGTC-3' and reverse primer, 5'-AAGCCAAGCCACACGTTTCC-3'). Amplification without a DNA template was used to monitor contamination in both HPV and β-globin reactions. PCR products were sequenced using a DNA Engine Peltier thermal cycler (Bio-Rad). A 267-bp amplicon was determined with 2% agarose electrophoresis (26).

Sequence analysis of the E6, E7, and L1 genes and the LCR of HPV 16 by PCR-directed sequencing. HPV 16 L1-E7 fragments and LCR-E7 fragments were amplified by high-fidelity PCR and sequenced. The L1-E7 fragment, with an amplicon size of 3.3 kb, was amplified with primers 5'-TTCTGTCCCAGTGGCAGG-3' and 5'-TTCTGTCCAGCTCG-3'. Amplification without a DNA template was assessed by PCR amplification of the β-globin gene (forward primer, 5'-CAAC-CCCTGTC-CCCTGTC-3' and reverse primer, 5'-AAGCCAAGCCACACGTTTCC-3'). Amplification without a DNA template was used to monitor contamination in both HPV and β-globin reactions. PCR products were sequenced using a DNA Engine Peltier thermal cycler (Bio-Rad). A 267-bp amplicon was determined with 2% agarose electrophoresis (26).

Sequence analysis of the E6, E7, and L1 genes and the LCR of HPV 16 by PCR-directed sequencing. HPV 16 L1-E7 fragments and LCR-E7 fragments were amplified by high-fidelity PCR and sequenced. The L1-E7 fragment, with an amplicon size of 3.3 kb, was amplified with primers 5'-TTCTGTCCCAGTGGCAGG-3' and 5'-TTCTGTCCAGCTCG-3'. Amplification without a DNA template was assessed by PCR amplification of the β-globin gene (forward primer, 5'-CAAC-CCCTGTC-CCCTGTC-3' and reverse primer, 5'-AAGCCAAGCCACACGTTTCC-3'). Amplification without a DNA template was used to monitor contamination in both HPV and β-globin reactions. PCR products were sequenced using a DNA Engine Peltier thermal cycler (Bio-Rad). A 267-bp amplicon was determined with 2% agarose electrophoresis (26).

Variant identification. HPV 16 variants were classified into phylogenetic classes and subclasses on the basis of their sequence variation at the E6 open reading frame (ORF), which is the most informative for this purpose. The E6 sequences were then compared with the reference sequences using the HPV 16R database to classify HPV 16 variant classes. The European (E) lineage has approximately complete homology with the HPV 16R database, but with no signature patterns of other classes identified. Changes in the following lineages were noted: Asian (As), T178G, A14I, G132C, C143G, G145T, T286A, A289G, and C335T; Af-2, T1109C, G132T, C143G, G145T, T286A, A289G, C335T, and A403G; NA-1, G145T, T286A, A289G, C335T, and T350G; and AA, G145T, T286A, A289G, C335T, and A532G (18, 46, 51, 52). Sequence variations in E6, E7, L1, and LCR in this study were compared with other data reported from neighboring geographical districts, which originated from cervical carcinomas in hospital-based studies.

Statistical analysis. Statistical analysis was performed using the chi-square test to determine the differences between northeast China and other geographical regions regarding HPV 16 variants and the European lineage distribution of cancer patients. P values of less than 0.05 were considered statistically significant.

RESULTS

HPV typing. All 71 cervical carcinomas were β-globin gene positive, indicating the presence of adequate cells in DNA samples. The HPV infection rate of cervical carcinomas was 98.59% (70 of 71). HPV 16 was the most common type (73.24%, 52 of 71), followed by HPV 18 (11.27%, 8 of 71).

HPV 16 variant identification. In HPV 16-positive specimens, 67.31% (55 of 82) belonged to the European lineage, while 32.69% (17 of 52) belonged to the Asian lineage. The AA, A1-I, A1-2, and NA lineages were not detected. The HPV 16 European lineage contained the European prototype (54.29%, 19 of 35) and European variants (45.71%, 16 of 35) (see Table 3).

E6 gene sequence variations. In HPV 16-positive samples, the nucleotide variation rate of HPV 16 E6 was 63.46% (33 of 52) (Table 1), whereas the amino acid variation rate was 57.78% (30 of 52). Nucleotide variation positions of E6 included three silent mutations (G94A, A131C, and T241G) and six missense mutation positions (G176A, T178G, T276G, T341G, T350G, and A442C), which led to amino acid variations (D25N, D25E, N58S, C80G, L83V, and E113D). The most frequent sequence variation site was nt 36 (36.54%), L83V (5.77%), D25N (5.77%), and N58S (5.77%). The rank orders of incidence of the HPV 16 prototype and E6 variants were as follows: D25E (42.31%), the prototype (34.54%), L83V (5.77%), D25N (5.77%), and N58S (3.85%). Only one novel mutation (T341G) in E6 was found for the first time, which resulted in amino acid substitution C80G (Table 2).

E7 gene sequence variations. Forty-four (84.62%) patients showed six mutation sites for the E7 gene, of which two, A645C and A647G, were missense mutations: asparagine—serine (N29S) and leucine—phenylalanine (L28F). The remaining four, G666A (Fig. 1b), T760C, T843C, and T846C, led to silent mutations. The amino acid variation rate was 36.54% (19 in

### TABLE 1. Cases exhibiting the HPV 16 prototype or variant in E6, E7, the LCR, and L1

<table>
<thead>
<tr>
<th>Gene/region</th>
<th>Prototype</th>
<th>Variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>E6</td>
<td>19 (36.54)</td>
<td>33 (63.46)</td>
</tr>
<tr>
<td>E7</td>
<td>8 (15.38)</td>
<td>44 (84.62)</td>
</tr>
<tr>
<td>LCR&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0 (0.00)</td>
<td>52 (100.00)</td>
</tr>
<tr>
<td>L1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2 (4.26)</td>
<td>45 (95.74)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Variant analysis for the partial LCR sequence was done in samples 48, 49, 50, 51, and 52 (Fig. 2).

<sup>b</sup> Variant analysis for L1 was done using only 47 cases.
The most frequently observed mutations were G666A (38.46%, 20 of 52), A647G (34.62%, 18 of 52), and T846C (32.69%, 17 of 52), while the most frequent amino acid mutation was N29S (34.62%).

L1 gene sequence variations. The HPV 16 L1 genes (nt 5559 to 7154) from 47 HPV 16-positive cervical cancers patients were determined. Of all 19 nucleotide variation positions, six were missense mutations (A5883C, G6085A, A6164G, A6166C, A6427G, A6505A), and seven were silent mutations (A5836A, A5902C, A5902G, A6166C, A6166G, A6934A, A6934C). The remaining six mutations were transitions (G6085A, A6164C, A6166A, A6427G, A6505G, A6505A). The amino acid changes induced by these mutations are shown in Table 2. The most common amino acid changes were Serine to Threonine (315) and Threonine to Valine (3), with 17 and 9 mutations, respectively. The least common changes were Lysine to Serine (3) and Arginine to Thrigine (2), with 1 and 2 mutations, respectively.

TABLE 2. Novel variants of HPV 16 E6 and L1 and the LCR according to nucleotide position, variant, altered amino acid, and mutation

<table>
<thead>
<tr>
<th>Gene/region</th>
<th>Nt position</th>
<th>Amino acid</th>
<th>Codon no.</th>
<th>Type of mutation</th>
<th>No. of mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>E6</td>
<td>341</td>
<td>T</td>
<td>G</td>
<td>Cysteine Glycine</td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>5569</td>
<td>G</td>
<td>A</td>
<td>Valine Valine</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>5596</td>
<td>A</td>
<td>T</td>
<td>Threonine Threonine</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>5836</td>
<td>A</td>
<td>G</td>
<td>Serine Serine</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>5902</td>
<td>C</td>
<td>A</td>
<td>Threonine Threonine</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td>6085</td>
<td>G</td>
<td>A</td>
<td>Methionine Isoleucine</td>
<td>175</td>
</tr>
<tr>
<td></td>
<td>6164</td>
<td>A</td>
<td>G</td>
<td>Threonine Alanine</td>
<td>202</td>
</tr>
<tr>
<td></td>
<td>6166</td>
<td>C</td>
<td>A</td>
<td>Threonine Threonine</td>
<td>202</td>
</tr>
<tr>
<td></td>
<td>6427</td>
<td>G</td>
<td>A</td>
<td>Arginine Arginine</td>
<td>289</td>
</tr>
<tr>
<td></td>
<td>6505</td>
<td>A</td>
<td>G</td>
<td>Serine Serine</td>
<td>315</td>
</tr>
<tr>
<td></td>
<td>6658</td>
<td>A</td>
<td>T</td>
<td>Threonine Threonine</td>
<td>366</td>
</tr>
<tr>
<td></td>
<td>6667</td>
<td>A</td>
<td>T</td>
<td>Serine Serine</td>
<td>369</td>
</tr>
<tr>
<td></td>
<td>6938</td>
<td>C</td>
<td>G</td>
<td>Proline Alanine</td>
<td>460</td>
</tr>
<tr>
<td></td>
<td>7142</td>
<td>A</td>
<td>C</td>
<td>Lysine Glutamine</td>
<td>528</td>
</tr>
<tr>
<td>LCR</td>
<td>7168</td>
<td>A</td>
<td>G</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>7174</td>
<td>A</td>
<td>C</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>7233</td>
<td>A</td>
<td>G</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>7419</td>
<td>A</td>
<td>G</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>7429</td>
<td>A</td>
<td>T</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>7660</td>
<td>A</td>
<td>G</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>7667</td>
<td>A</td>
<td>C</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>7700</td>
<td>T</td>
<td>G</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>7781</td>
<td>T</td>
<td>A</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>7855</td>
<td>G</td>
<td>A</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>7859</td>
<td>A</td>
<td>C</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>G</td>
<td>A</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>G</td>
<td>A</td>
<td></td>
<td>17</td>
</tr>
</tbody>
</table>

* M, missense; S, silence; Ts, transition; Tv, transversion.
A6180C, C6938G, and A7142C) which led to K108T, M175I, T202A, N207T, P460A, and K528Q, respectively. The remaining 13 were silent mutations. There were three common nucleotide variations at nt 6667, nt 6826, and nt 7060, which were silent variations (Fig. 1c). The most frequent amino acid mutation was N207T (6.38%, 3 out of 47). In 47 HPV 16-positive cervical carcinomas, the nucleotide variation rate in the HPV 16 L1 region was 95.74% (45 of 47), whereas the amino acid variation rate was 14.89% (7 of 47). There were 14 novel nucleotide mutations in the L1 gene (Table 2).

LCR sequence variations.
Forty-seven whole LCR fragments and five partial LCR fragments were sequenced. A total of 32 nucleotide variation positions were detected (Fig. 2). The most commonly observed LCR variation was the transition replacement G7521A (98.08%), which is shown in Fig. 1d. We reported a total of 13 novel nucleotide variation sites in LCR, which were observed here for the first time (Table 2). The nucleotide variation rate in the LCR was 100% in 47 HPV 16-positive cervical carcinomas.

Covariations.
In the HPV 16 Asian branch, covariations of T178G, A647G, T846C, A7175C, T7177C, T7201C, C7270T, A7287C, A7289C, A7730C, G7842A, and C24T were detected, resulting in covariations of predicted amino acids D25E and N29S (32.69%, 17/52) (Fig. 2). Covariations of G666A (E7), C6826T (L1), and A7168G (LCR) were detected (36.17%, 17 of 47) in the E branch in this study.

Variant comparison with neighboring districts.
HPV 16 variant distribution in this study was different from that in central and southern China, Russia, Indonesia, and north India, whereas it was not different from that in Sichuan (south-east China). The nucleotide positions of detected mutations are written vertically across the top and are indicated by the corresponding nucleotide letters. The absence of the samples are indicated on the left. HPV 16 reference: Ep, European prototype; As, Asian.
Central and southern China—
Central China—Hubei (4) 5 58 23 (39.66) 35 (60.34) 0 0 0 0 9 (39.13) 14 (60.87)
Southwest China—Sichuan (3) 41 21 16 (76.19) 5 (23.81) 0 0 0 0 4 (25.00) 12 (75.00)
Southwest China—Sichuan (2) 35 113 78 (69.03) 35 (30.97) 0 0 0 0 26 (33.33) 52 (66.67)
Northeast China (1) 52 35 (67.31) 17 (32.69) 0 0 0 0 19 (54.29) 16 (45.71)

those in Sichuan (southwest China) and Japan and different
ysis, the variant distribution in northeast China was the same as
andering countries (Russia and Japan) (52). Hong Kong, Mongolia,
 China, and the second group contained central China and
China (Hubei, Jiangxi, and Guangdong). The differ-
the Chinese population, as shown in Table 3. There were two
pecific groups with HPV 16 variant distribution. The first
t contained southwest China (Sichuan) and northeast China,
dergontowhereasothanothersincentralandsouthernChina
 lower than those in central and southern China. Other muta-
tion frequencies varied among different districts (Table 3).

**DISCUSSION**

Located in the center of northeast Asia, northeast China is
close to north China and neighbors Korea, Russia, Mongolia,
and Japan. In history, more communication lies between
China, Mongolia, Korea, Russia, and Japan. Northeast China
has a population of around 110 million people, who immi-
grated mainly from north China (Shandong and Hebei). How-
ever, no HPV 16 variant data from north China were provided,
where most of the population originated. Since HPV 16 vari-
ants are associated with geographic diversity and oncogenic
risk, we analyzed the data with China origin (southwest China,
central China, and southern China) and those from neighbor-
king countries (Russia and Japan) (52). Hong Kong, Mongolia,
and Korea were excluded from the analysis since no individual
data on cervical carcinoma samples were reported (6, 8, 9).

In this hospital-based study, the samples contained two HPV
16 lineages, the European lineage (67.31%) and the Asian
lineage (32.69%), which suggested that the European and
Asian lineages were the main lineages in northeast China.
In Table 3, the European and Asian lineages in these studies
were circulated, while the AA, Af-1, Af-2, and NA lineages
were reported less often. The European lineage was the main
lineage in Russia, north India, and Indonesia. By statistical
alysis, the variant distribution in northeast China was the same as
those in Sichuan (southwest China) and Japan and different
from those in central and southern China and neighboring
countries, except Japan.

Southwest China (Sichuan) and northeast China were areas
with a moderate mortality of invasive cervical cancer, while
central China (Hubei and Jiangxi) had a high mortality (57).
Northeast China and southwest China have the same HPV 16
variant distribution by statistical analysis, while Hubei, Jiangxi,
and Guangdong have higher Asian variant frequencies.
Interestingly, it was reported that HPV 16 Asian variants might
be linked to a high incidence of cervical cancer in China (48).
We wondered if the distribution character of HPV 16 variants,
especially Asian variants, had a relationship with the mortality
of invasive cervical cancer in these geographical areas. But this
consideration needs further investigation by population-based
study in more districts in China.

The European and Asian lineages were the main patterns in
the Chinese population, as shown in Table 3. There were two
specific groups with HPV 16 variant distribution. The first
group contained southwest China (Sichuan) and northeast
China, and the second group contained central China and
southern China (Hubei, Jiangxi, and Guangdong). The differ-
ence had no significance within each group. The two groups
were different from neighboring countries. In this analysis, the
Japan sample was different from that of the first group, al-
though it had no differences with the northeast China sample.
Data from east China (Zhejiang) belonged to the second group
(20). Data from Hong Kong (southern China) also conformed
to the second group, which belonged to a population-based
study (6). However, the distributions of the European proto-
type and European variants in northeast China were specific,
because the data were different from data of any other area by
statistical analysis. Identification of HPV 16 variants may be
helpful to evaluate the potential impact of HPV vaccine and
design diagnostics and therapeutic methods on cervical cancer
(15, 44). Although further confirmation of this study is still
needed, the results may be helpful for the study of the HPV
vaccination program in China.

Covariation of nucleotide sequences among these genes was

**TABLE 3. Comparison of HPV 16 variant distributions and sequence variations with neighboring districts and countries**

<table>
<thead>
<tr>
<th>Variant (sample no.)</th>
<th>Reference</th>
<th>No. of cases sequenced</th>
<th>No. (%) of cervical carcinoma cases&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>HPV 16 variant</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>E</td>
</tr>
<tr>
<td>Northeast China (1)</td>
<td>52</td>
<td>35 (67.31) 17 (32.69) 0 0 0 0 9 (39.13) 14 (60.87)</td>
<td></td>
</tr>
<tr>
<td>Southwest China—Sichuan (2)</td>
<td>35</td>
<td>113</td>
<td>78 (69.03) 35 (30.97) 0 0 0 0 19 (54.29) 16 (45.71)</td>
</tr>
<tr>
<td>Southwest China—Sichuan (3)</td>
<td>41</td>
<td>21</td>
<td>16 (76.19) 5 (23.81) 0 0 0 0 4 (25.00) 12 (75.00)</td>
</tr>
<tr>
<td>Central China—Hubei (4)</td>
<td>5</td>
<td>58</td>
<td>23 (39.66) 35 (60.34) 0 0 0 0 9 (39.13) 14 (60.87)</td>
</tr>
<tr>
<td>Central and southern China—Jiangxi, Guangdong (5)</td>
<td>47</td>
<td>55</td>
<td>19 (34.55) 36 (65.45) 0 0 0 0 5 (26.32) 14 (73.68)</td>
</tr>
<tr>
<td>Japan (6)</td>
<td>27</td>
<td>43</td>
<td>24 (55.81) 19 (44.19) 0 0 0 0 5 (20.83) 19 (79.17)</td>
</tr>
<tr>
<td>Northeast Thailand (7)</td>
<td>10</td>
<td>23</td>
<td>6 (26.09) 17 (73.91) 0 0 0 0 1 (4.17) 22 (95.83)</td>
</tr>
<tr>
<td>Russia (8)</td>
<td>19</td>
<td>32</td>
<td>30 (93.75) 0 0 0 0 2 (6.25) 28 (93.75)</td>
</tr>
<tr>
<td>Indonesia (9)</td>
<td>11</td>
<td>22</td>
<td>19 (86.36) 2 (9.09) 0 1 (4.55) 0 0 7 (36.84) 12 (63.16)</td>
</tr>
<tr>
<td>North India (10)</td>
<td>32</td>
<td>60</td>
<td>52 (86.67) 0 0 0 2 (3.33) 6 (10.00) 18 (30.62) 34 (69.38)</td>
</tr>
</tbody>
</table>

<sup>a</sup> A dash indicates that no data were provided. Chi-square ($\chi^2$) tests were applied in this study. Regarding HPV 16 variant distribution, the differences had no significance between samples 1 and 2, 1 and 3, 2 and 3, 1 and 2 + 3, 4 and 5, 1 and 6 ($P$ value > 0.05), while statistical significance existed between samples 1 and 4/5/7/8/9/10, 1 and 4 + 5, 2 + 3 and 6/7/8/9/10, and 4 + 5 and 6/7/8/9/10 ($P$ value < 0.05). Regarding the distribution of the HPV 16 European prototype and European variant, statistical significance existed between 1 and 2/3/4/5/6/7/8/9/10.
analyzed, since the HPV 16 E6, E7, and L1 genes and the LCR were sequenced at one DNA molecule by primer walking. Among the Asian lineage, covariations of predicted amino acids D25E and N29S might be a typical pattern. Covariations of G666A (E7) and C6826T (L1), reported as Javanese variants, were found in 72.73% of Indonesian samples (Fig. 2) (11). Validation of Javanese variants needed more confirmation, because G176A, T178A, and T241G were detected in this variant apart from A276G.

Amino acid mutations in HPV 16 variants should be considered in the vaccine design and evasion of the natural immune response (12, 17, 47). As the target of humoral and cellular immune responses, the amino acid changes of HPV 16 E6 may have effects (2, 13, 39). D25E and L83V are associated with the elevated risk of cervical carcinomas and vary geographically due to genetic differences between populations (1, 3, 18, 22, 25, 26, 32, 34). The D25E and L83V distribution was lower in the Asian population and higher in the American and European populations (5, 6, 47). The L83V distribution was lower in the Asian population and higher in the American and European populations (5, 6, 47). The L83V distribution was lower in the American and higher in the European and American populations. In Table 3, L83V mutations in northeast China accounted for 5.77%, close to those in central China and southern China, in contrast to that in southwest China (27.61% in Sichuan), 32.56% in Japan, and 84.38% in Russia (Moscow) (5, 6, 8, 9, 19, 47, 52). G94A was also reported in southwest China and Mongolia (8, 35). Sequence variation at E7 displays geographic dependence (14, 36). The E7 N29S replacement may be associated with a higher oncogenic risk (12, 23, 38).

In this study, N29S accounted for 34.62% of mutations, which varied from 14.29% (southwest China) to 70.21% (central and southern China). G666A varied between northeast China, southwest China, central and southern China, and Indonesia (11, 41, 47).

Compared with E6 gene variations, the HPV 16 L1 gene product is more conservative in terms of its amino acid sequence. HPV 16 L1 nucleotide variations were found in 95.74% of specimens, while amino acid mutations were located only in 14.89% of specimens. The amino acid changes of K108T, A175L, I202A, N207T, P460A, and K528Q were reported previously. The silent variation C6826T has been reported from southeast Asia (52). To our knowledge, 14 novel mutations in this study were not reported previously.

The LCR is the most variable segment of the HPV 16 genome in different populations (21, 22, 24, 43, 52). The LCR is the binding site of various cellular and viral transcription factors. In this study, the nucleotide variation rate in the LCR was 100%. A total of 5 of 32 mutation positions (15.63%) impacted cellular transcription factor binding sites: C7310T lies in AP1, C7395T lies in GRE, T7441G lies in YY1, T7714G lies in NF1, and G7842A and G7842T lie in Oct-1. Over half of the isolates (32/52) showed the above-mentioned mutation sites. The exact oncogenic potentials of these nucleotide mutations still need further investigation.

In this study, we investigated comprehensively sequence variations of the HPV 16 E6, E7, and L1 genes and the LCR based on long fragment sequencing. Five HPV 16-positive samples which lost the L1 ORF were sequenced on LCR-E7 fragments. Age was not considered in this study, for age was reported to have no relationship with the distribution of HPV 16 variants among women in the same geographic region (34). As a hospital-based sample study, our study was still limited by sample size. The conclusions in this study need further investigation by a large population-based study.

In the present study, it was demonstrated that HPV 16 variants in northeast China have typical distribution. HPV 16 variants belonged to the European lineage and the Asian lineage. D25E variation in E6 and N29S variation in E7 were the most prevalent variations. Twenty-eight novel variations of HPV 16 were reported. This is the first report on HPV 16 variant distribution and sequence variations in northeast China. Based on this study, we supposed that northeast China may be an area with moderate HPV 16 Asian variant prevalence. Further investigation should be emphasized on Asian variants in east Asia. We suggested that an HPV-preventive vaccine should be applied to control the HPV 16 infection in northeast China. These data lay the foundation for the HPV vaccination program in northeast China and are important for

---

**TABLE 3—Continued**

<table>
<thead>
<tr>
<th>Variation site</th>
<th>No. (%) of cervical carcinoma cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>D25E in E6</td>
<td>T359G (L83V) in E6</td>
</tr>
<tr>
<td>G94A in E6</td>
<td>T178A in E6</td>
</tr>
<tr>
<td>A276G (N958) in E6</td>
<td>C335T (H78Y) in E6</td>
</tr>
<tr>
<td>A442C (E113D) in E6</td>
<td>T183G (I27R) in E6</td>
</tr>
<tr>
<td>A647G (N298S) in E7</td>
<td>C749T (S63F) in E7</td>
</tr>
<tr>
<td>C669A in E7</td>
<td>C6826T (silence) in L1</td>
</tr>
<tr>
<td>G7521A in LCR</td>
<td>G7842A in LCR</td>
</tr>
<tr>
<td></td>
<td>No. (%) of cervical carcinoma cases</td>
</tr>
<tr>
<td>22 (42.31)</td>
<td>3 (5.77)</td>
</tr>
<tr>
<td>35 (30.97)</td>
<td>35 (30.97)</td>
</tr>
<tr>
<td>5 (23.81)</td>
<td>2 (9.52)</td>
</tr>
<tr>
<td>36 (62.07)</td>
<td>3 (5.17)</td>
</tr>
<tr>
<td>37 (67.27)</td>
<td>2 (3.64)</td>
</tr>
<tr>
<td>19 (44.19)</td>
<td>14 (32.56)</td>
</tr>
<tr>
<td>17 (73.9)</td>
<td>0</td>
</tr>
<tr>
<td>2 (9.09)</td>
<td>1 (4.55)</td>
</tr>
<tr>
<td>0</td>
<td>42 (70.00)</td>
</tr>
</tbody>
</table>

Continued...
diagnostic technique development and vaccine design for eradication of cervical cancers.

ACKNOWLEDGMENTS

This research was supported by grant 11551236 from the Heilongjiang Higher Education Institutions, funding (grants LC2009C26 and D201065) from the Natural Science Foundation of Heilongjiang Province, and grants 81000726 and 30901706 from the National Natural Science Foundation of China, People’s Republic of China. We have no conflicts of interest to declare.

REFERENCES


ERRATUM

Human Papillomavirus Type 16 Variant Analysis of E6, E7, and L1 Genes and Long Control Region in Cervical Carcinomas from Patients in Northeast China

Qinglong Shang, Yan Wang, Yong Fang, Lanlan Wei, Sijia Chen, Yuhui Sun, Baoxin Li, Fengmin Zhang, and Hongxi Gu

Department of Microbiology, Harbin Medical University, Heilongjiang Key Laboratory of Infection and Immunity, Heilongjiang Province, and Pathogenic-Biological Key Laboratory, Heilongjiang Higher Education Institutions, Harbin 150081, People’s Republic of China; and Department of Gynecology, First Affiliated Hospital, Harbin Medical University, Harbin 150001, People’s Republic of China; and Department of Pharmacology, Harbin Medical University, and Bio-pharmacological Key Laboratory of Heilongjiang Province, Harbin 150081, People’s Republic of China