Scrub typhus, an acute febrile illness that is widespread in the Asia-Pacific region, is caused by the bacterium *Orientia tsutsugamushi*, which displays high levels of antigenic variation. We conducted an investigation to identify the circulating genotypes of *O. tsutsugamushi* in 3 scrub typhus–endemic geographic regions of India: South India, Northern India, and Northeast India. Eschar samples collected during September 2010–August 2012 from patients with scrub typhus were subjected to 56-kDa type-specific PCR and sequencing to identify their genotypes. Kato-like strains predominated (61.5%), especially in the South and Northeast, followed by Karp-like strains (27.7%) and Gilliam and Ikeda strains (2.3% each). Neimeng-65 genotype strains were also observed in the Northeast. Clarifying the genotypic diversity of *O. tsutsugamushi* in India enhances knowledge of the regional diversity among circulating strains and provides potential resources for future region-specific diagnostic studies and vaccine development.

Scrub typhus is a vector-borne, acute febrile illness caused by *Orientia tsutsugamushi*, an obligate intracellular, gram-negative bacterium. Scrub typhus is widespread in the Asia-Pacific region, known as the “tsutsugamushi triangle.” Mite larvae, or chiggers, of the genus *Leptotrombidium* transmit the causative bacteria to humans through their bite. The infection is maintained in nature through transovarial transmission in the vector and a reservoir in small mammals (1,2). Clinical signs and symptoms of scrub typhus in humans are largely nonspecific, and if infection is not treated promptly and appropriately, it carries a high mortality rate (3).

The *Orientia* genome has a high degree of plasticity and is considered to be the most highly repetitive bacterial genome sequenced (4). This diversity is a result of high numbers of intragenomic deletions, duplications, and rearrangements with transposable and conjugative elements. These recombinations and rearrangements are unlikely to occur in dead-end hosts, but the details of this process are unclear (5).

Clarifying the epidemiology and genetic diversity of *O. tsutsugamushi* strains is essential to the development of rapid diagnostics and vaccines in disease-endemic areas. These efforts would also help in the early recognition and treatment of the disease. Currently, the most widely used method for strain classification is sequence analysis of the 56-kDa type-specific antigen (TSA), an immunodominant outer membrane protein unique to *O. tsutsugamushi*. With an open reading frame (ORF) of ≈1,600 bp, the 56-kDa TSA contains 516–541 amino acids and is involved in host cell invasion through the binding of fibronectin (6). Four hypervariable domains in this region, variable domains (VD) I–IV, are responsible for the large degree of antigenic variation in this gene. The direct interaction with the host, uniqueness to *O. tsutsugamushi*, and high level of variability make this protein an attractive target for studying the genetic variation among strains. This region is also highly immunogenic, making it a potential candidate as a vaccine target.

The process of conventional serotyping was a complex procedure, requiring reference serum samples and antigens, and is of limited use today. Greater diversity among the strains has been revealed by using molecular genotyping methods. Antigenic variations in *O. tsutsugamushi* from patients and rodents in different scrub typhus–endemic regions have been reported by testing using the 56-kDa TSA, which has led to identification of several new subtypes (1), such as Japanese Gilliam, Japanese Karp, Kawasaki, Kuroki, and Shimokoshi, in addition to the previously described prototypes Karp, Kato, and Gilliam (7,8).
Given the broad endemicity of scrub typhus in the Asia-Pacific region and variations in clinical manifestations that may be attributable to strain variation, thorough investigation into the regional distribution of genotypes is warranted. This study was conducted to identify the circulating 56-kDa antigen genotypes in 3 scrub typhus–endemic geographic regions of India: South India, Northern India, and Northeast India.

**Methods**

**Study Population**

All patients who sought care for fever and suspected scrub typhus at the 3 recruiting centers (Christian Medical College in Vellore, Indira Gandhi Medical College in Shimla, and the Dr. H. Gordon Roberts Hospital in Shillong) during September 2010–August 2012 were evaluated. A detailed clinical assessment of each patient for the signs and symptoms of scrub typhus, including a careful search for an eschar and basic laboratory studies, was documented by using a standardized form. Additional investigations, including blood cultures, quantitative buffy coat (testing for malarial parasites), abdominal ultrasound, and serologic testing for the causative agents of leptospirosis and dengue, were performed to rule out other common infections. Detection of *O. tsutsugamushi* was performed by using the Scrub Typhus Detect IgM ELISA (InBios International, Inc., Seattle, WA, USA) according to the manufacturer’s instructions; an optical density (OD) >0.5 was considered positive. De-roofed eschar samples were collected and stored in absolute alcohol at −70°C until DNA extraction. Patients with clinical illness compatible with scrub typhus, including an eschar and positive results for serum samples testing by IgM ELISA, were included in the study. The study was approved by the Institutional Review Board and Ethics Committee of Christian Medical College, and informed consent was obtained from all patients.

**DNA Amplification and Sequence Analysis**

DNA was extracted from the homogenized eschar samples by using the QIAamp DNA Mini Kit (QIAGEN GmbH, Hilden, Germany) according to the manufacturer’s instructions. A standard PCR, targeting the 56-kDa protein, was performed as described by using the primers OtsuF: 5′-ATTGCTAGTGCAATGTCTG-3′ and OtsuR: 5′-GGATTATAGGGCTGAG-3′ (Sigma Aldrich, Bangalore, India) (9). This region encompasses ∼410 bp and contains the VD I–III hypervariable regions. The PCR products were purified by using the QIAquick PCR Purification Kit (QIAGEN) according to the manufacturer’s instructions; products were then subjected to a sequencing reaction using the BigDye Terminator Mix (Applied Biosystems, Foster City, CA, USA) and subsequent automatic sequencing using the ABI 310 Genetic Analyzer (Applied Biosystems). PCR and sequencing reactions were carried out in the Infectious Diseases Research laboratories at Christian Medical College in Vellore.

The sequences obtained were identified by comparison with sequences available in GenBank by using BLAST (http://blast.ncbi.nlm.nih.gov). All sequences obtained were submitted to GenBank (accession nos. KC153061–KC153085 and KF777306–KF777328 for samples from Vellore; KF777265–KF777290, KF777292, KF777294–KF777305 for samples from Shimla; and KF777329–KF777357 for samples from Shillong). Phylogenetic and molecular evolutionary analyses were conducted by using MEGA5 (10). The study sequences and reference sequences downloaded from GenBank were aligned by using ClustalW (http://www.clustal.org) and trimmed to the appropriate size. A phylogenetic tree with 1,000 bootstrap replications was constructed by using neighbor-joining methods with distances calculated by the maximum composite likelihood. Statistical analysis was performed by using SPSS software version 16.0 for Windows (SPSS IBM, Armonk, NY, USA). Descriptive data are given as mean (SD) or median (range).

**Results**

A total of 263 eschar samples were obtained from patients with confirmed scrub typhus from the 3 study sites: 95 from Vellore in South India, 72 from Shimla in Northern India, and 96 from Shillong in Northeast India. Of the patients from whom these samples were collected, 115 (43.7%) were male and 148 (56.3%) female. Mean patient age was 40 ± 12 years.

Of the 263 samples, 209 (79.5%) were successfully amplified by using 56-kDa conventional PCR. Of the 209 amplified samples, 130 (62.2%) yielded good sequencing reads (58 from Vellore, 42 from Shimla, and 30 from Shillong). These reads were then used for further data analysis.

**Sequence Analysis**

Sequence analysis revealed that Kato-like strains predominated; 80 (61.5%) of the 130 total samples analyzed yielded Kato-like strains, followed by Karp-like strains with 36 (27.7%) samples and Gilliam and Ikeda strains with 3 (2.3%) samples each. Four (3.1%) samples from Shillong were similar to Neimeng-65, and 4 (3.1%) from Shimla were similar to IHS-II. The Karp-like amplicons showed a nucleotide similarity of 88% to 95% to the Karp reference strain (GenBank accession no. AY956315), and the nucleotide similarity for Kato-like strains to the US Centers for Disease Control and Prevention (CDC) Kato reference strain (GenBank accession no. AY836148) ranged from 90% to 94%. Most sequences showed 90%–96% similarity to Hualien strains from Taiwan.
The distributions of genotypes by center are displayed in the Table. Among the 58 samples analyzed from Vellore, 45 (77.5%) were Kato-like, and most of those (35, 77.7%) showed 92%–96% sequence similarity to the Hualien-3 strain from Taiwan. Two Kato-like amplicons were 94% similar to another strain from Taiwan (GenBank accession no. KM06077b), and 4 others were 92%–95% similar to strain TA678 from South Korea. The remaining showed similarities ranging from 92% to 95% to various strains such as MZ02, KM06, and HL03 from Taiwan. Eleven (18.9%) samples from Vellore yielded Karp-like amplicons, and these showed 95%–98% sequence similarities to the clone T0122244_KH from Cambodia and Vietnam. One Gilliam strain found in Vellore showed 99% similarity to the BA344_1 Gilliam strain from Thailand (GenBank accession no. JN587265). One Ikeda strain was also found in Vellore. Phylogenetic analysis of the Vellore strains with reference strains showed 2 distinct clusters (Figure 1, panel A); one of the clusters included 11 strains from Vellore along with the Karp reference strains Taiwan CDC Karp, Kp-1, and Kp-2a.

Kato-like sequences also predominated among the 30 analyzed sequences from Shillong; 17 (56.6%) samples belonged to this strain. These genotypes, however, were generally more similar (94%–95%) to the Hualien-13 strain from Taiwan. Some were also highly similar (98%) to the clone T1116018_KH from Cambodia and Vietnam, and 1 was 99% similar to the reference strain Taiwan CDC Kato. Seven samples (23.3%) from Shillong were Karp-like strains, 2 (6.6%) were Gilliam strains, and 4 (13.3%) were Neimang-65 strains. Two major clusters were observed on phylogenetic analysis, with a few of the strains clustering with the Hualien-13 strains and the remaining strains scattered in a second, larger cluster with several subgroups (Figure 1, panel B).

Of the 42 sequences analyzed from Shimla, 18 (42.8%) were Kato-like, and most showed 94%–96% similarity to the Hualien-3 strain. The samples from Shimla also included 18 (42.8%) Karp-like samples, which were most similar to the MZ01 strain obtained from chiggers from Taiwan. Two (4.7%) samples were found to be of the Ikeda strain and 4 (9.5%) of the IHS-II cluster. Two distinct clusters with subgroups were observed on phylogenetic analysis (Figure 1, panel C).

A comparison of sequences obtained in this study revealed that the samples from the 3 centers clustered separately on the neighbor-joining tree. The Vellore strains formed a distant clade, whereas the Shillong and Shimla strains clustered more closely together (Figure 2, http://wwwnc.cdc.gov/EID/article/21/1/14-0580-F2.htm).

**Discussion**

Scrub typhus has been recognized as a major cause of undifferentiated acute febrile illness in India (11). Previous studies have reported the prevalence of scrub typhus in Vellore and Shimla (12–15), whereas reports from Shillong are few and fairly recent (16,17). Scrub typhus has been reported from states throughout India: Kerala, Karnataka, Andhra Pradesh, and Tamil Nadu in South India; Bihar, Maharashtra, Jammu Kashmir, Himachal Pradesh, Uttarakhand, and Rajasthan in Northern India; and Meghalaya, Sikkim, and West Bengal in Northeast India (9,12,18–24). Despite the broad effect of scrub typhus in India, little genotyping has been performed, with results available only from Himachal Pradesh (9). Similarly, although scrub typhus has been reported from neighboring countries, such as Pakistan, Myanmar, and Nepal, genotype data are lacking from these regions, making comparisons difficult (25,26).

In this study, we identified the circulating genotypes of *O. tsutsugamushi* in 3 scrub typhus–endemic regions in India by analyzing a variable portion of the 56-kDa antigen gene. Overall, 80 (61.5%) of 130 samples were Kato-like strains, which have previously been found by serologic testing and genotyping in other locations, primarily Japan and Cambodia, but not as a predominant strain (1,4,27,28). The strains we identified were largely similar to the Hualien-3 reference sample, which was originally reported from Taiwan (6). However, some Kato-like strains showed closer similarity to the TA678 strain from South Korea. In Shillong, the Kato-like strains were most similar to another Taiwan strain, Hualien-13. Other Kato-like strains that bore a closer resemblance to the clone T1116018_KH from Cambodia and Vietnam were also found in Shillong.

The highest proportion of Kato-like strains (77.5%) was found at Vellore in South India, with progressively smaller proportions of this strain seen as locations progressed to the north. The highest proportion of Kato-like strains in a single region (21.5%) was previously reported from Cambodia (4); South India’s climate is more similar to Cambodia’s climate than to Northern India’s climate, so temperature or other environmental factors may play a role in the selecting for this strain. However, most of the Kato-like strains we found, particularly from Vellore, resembled strains from Taiwan rather than Cambodia, which could argue against a climate-based selection process. Kato-like strains in Shillong, which is the geographically closest of our sites to Cambodia, resemble the strains found in Southeast Asia.

**Table.** Geographic distribution of *Orientia tsutsugamushi* genotypes in 3 regions of India, September 2010–August 2012

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Vellore, n = 58</th>
<th>Shimla, n = 42</th>
<th>Shillong, n = 30</th>
<th>Total, n = 130</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kato-like</td>
<td>45 (77.5)</td>
<td>18 (42.8)</td>
<td>17 (56.6)</td>
<td>80</td>
</tr>
<tr>
<td>Karp-like</td>
<td>11 (18.9)</td>
<td>18 (42.8)</td>
<td>7 (23.3)</td>
<td>36</td>
</tr>
<tr>
<td>Gilliam</td>
<td>1 (1.7)</td>
<td>0 (0.0)</td>
<td>2 (6.6)</td>
<td>3</td>
</tr>
<tr>
<td>Ikeda</td>
<td>1 (1.7)</td>
<td>2 (4.7)</td>
<td>0 (0.0)</td>
<td>3</td>
</tr>
<tr>
<td>IHS-II</td>
<td>0 (0.0)</td>
<td>4 (9.5)</td>
<td>0 (0.0)</td>
<td>4</td>
</tr>
</tbody>
</table>
Diversity of O. tsutsugamushi, India

Figure 1. Phylogenetic distribution of Orientia tsutsugamushi isolates from scrub typhus patients in Vellore (A), Shillong (B), and Shimla (C), India, September 2010–August 2012. Isolates were identified on the basis of the 56-kDa TSA gene. Evolutionary history was inferred by using the neighbor-joining method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1,000 replicates) is shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. Evolutionary distances were computed by using the maximum composite likelihood method. Scale bars indicate base substitutions per site. Sequences identified in this study were deposited in GenBank under accession nos. KC153061–KC153085 and KF777306–KF777328 (OTV), KF777329–KF777357 (OTSG), and KF777265–KF777290, KF777292, and KF777294–KF777305 (OTSh). OTV, O. tsutsugamushi from Vellore; OTSG, O. tsutsugamushi from Shillong; OTSh, O. tsutsugamushi from Shimla.
Karp-like strains, although less common than Kato-like in our study, are more widespread globally, reported by serologic testing and genotyping in Southeast Asia (Thailand, Vietnam, Cambodia, the Philippines, and Malaysia), Japan, southern China, Taiwan, Mongolia, Oceania, and Australia; these strains also are cited as the most common circulating strain in several Southeast Asia countries and in Japan (1,3,27–30). By serologic testing, these strains have been previously reported in India and in Pakistan just to the north (1). In our study, Karp-like strains were at their highest proportion at Shimla in the northern part of the country (equal to the amount of Kato-like strains), with progressively smaller proportions of this grouping successively farther south. Most Karp-like strains from Vellore showed high sequence similarity to strains previously found in Cambodia and Vietnam; strains in 5 samples showed a high degree of similarity to the Ld-1a strain (GenBank accession no. JN415541), which was reported from trombiculid mites in Thailand. Unfortunately, sequencing data were not sufficient to assess for similarities in the Karp-like samples from Shillong or Shimla.

Our study also identified Gilliam-like strains in Vellore and Shillong; all of these sample strains were similar to the BA244_1 strain previously found in mites from Thailand. Gilliam strains have been reported in a variety of locations, including Japan, Korea, China, Taiwan, Thailand, Russia, Tajikistan, and Northern India (1,29,31–35). All of our samples, including those from Shillong in Northeast India, were most similar to genotypes previously seen in Thailand, not to the Gilliam-like strain from Myanmar (GenBank accession no. DQ286233) that was previously reported in Northern India. In addition, we isolated Neimeng-65 strains in Shillong. This finding is not surprising in Northeastern India, considering the proximity of the area to the border of China, where Neimeng-65 was originally reported in rodents from Inner Mongolia and Xinjian (GenBank accession no. DQ514319). We also found Ike-da-like strains from Vellore and Shimla; these strains were previously reported primarily from Japan (27).

Initial molecular descriptions of *O. tsutsugamushi* in India were reported in Shimla, where 2 new genotypes, IHS-I (GenBank accession no. DQ286233DQ530440) and IHS-II (GenBank accession no. DQ286233DQ530441), were identified based on partial 56k-Da TSA sequences (9). Strains in 4 samples, all from Shimla, in this study showed the highest nucleotide similarity to the IHS-II strain. Seerangayee strains and the Kuroki strain of the Boryong cluster, previously reported in India (1,34), were not identified in our study. Other groupings, such as Kawasaki, TA673, TA716, and Japanese Gilliam, were also not observed.

Our study has limitations. Not all scrub typhus patients have an eschar, so we could not run PCR testing on samples from all patients. If some strains are more likely than others to produce an eschar, some strains may have been over- or underrepresented in our study. In addition, not all eschar samples were amenable to good sequencing, which could also have skewed our results.

In summary, we identified the circulating genotypes of *O. tsutsugamushi* in 3 scrub typhus–endemic regions of India. Kato-like strains were found to be predominant in the South and Northeast, whereas an equal prevalence of Karp-like and Kato-like strains was observed in Northern India. The prevalence of antigenic diversity can have wider implications in vaccine development; also, a potential association between strain variation and pathogenicity of scrub typhus has reported in mice studies (36,37). In addition, targeted serodiagnosis will require further knowledge of this variability. Thus, an accurate picture of the local antigenic diversity will be essential for the development of region-specific vaccines and diagnostics.

This work was supported by the Indian Council of Medical Research (no. 30/3/16/2008/ECD-II).

Dr. Varghese is professor of infectious diseases at Christian Medical College, Vellore, India. His research includes the epidemiology (including molecular epidemiology), pathogenesis, early recognition, and management of rickettsial infections, particularly scrub typhus.

### References
