Molecular Phylogenetic Analysis of *Orientia tsutsugamushi* Based on the *groES* and *groEL* Genes

Satoru Arai,¹ Kenji Tabara,² Norishige Yamamoto,³ Hiromi Fujita,⁴ Asao Itagaki,² Marina Kon,³ Hiroshi Satoh,¹ Kazuko Araki,¹ Keiko Tanaka-Taya,¹ Nobuhiro Takada,⁵ Yasuhiro Yoshikawa,⁶ Chiaki Ishihara,⁷ Nobuhiko Okabe,^{1,8} and Kazunori Oishi¹

Abstract

DNA sequences encoding the GroES and GroEL proteins of *Orientia tsutsugamushi* were amplified by the PCR and sequenced. Pairwise alignment of full-length *groES* and *groEL* gene sequences indicated high sequence similarity (90.4–100% and 90.3–100%) in *O. tsutsugamushi*, suggesting that these genes are good candidates for the molecular diagnosis and phylogenetic analysis of scrub typhus. Comparisons of the 56-kD type-specific antigen (TSA) protein gene and the *groES* and *groEL* genes showed that genotypes based on the 56-kD TSA gene were not related to a cluster containing the *groES* and *groEL* genes in a dendrogram, suggesting that a gene rearrangement may be associated with homologous recombination in mites.

Key Words: *Orientia tsutsugamushi*—*groES* and *groEL* genes—Phylogeny—Scrub typhus—Japan.

Introduction

THE BACTERIUM *Orientia tsutsugamushi* belongs to the family Rickettsiaceae and is transmitted by trombiculid mites such as *Leptotrombidium pallidum*, *Leptotrombidium akamushi*, and *Leptotrombidium scutellare*. It is the causative agent of scrub typhus, an emerging zoonosis, spreading throughout central, southern, and eastern Asia, and Oceania (Kelly et al. 2009). In Korea, 17,451 scrub typhus cases, *i.e.*, 7.2 cases per 100,000 individuals, were reported from 2001 to 2005 (Bang et al. 2008). In Japan, 313–791 cases have been reported annually since 1999, when nationwide patient surveillance began. Less than ten fatal cases of scrub typhus have been reported in Japan each year, but the public health impact of scrub typhus has remained in endemic areas. The number of people that participate in outdoor activities in forests, mountains, and similar areas is increasing in Japan, which increases the risk of transmission. Recently, several research groups conducted field surveys to determine the prevalence of *O. tsutsugamushi* among wild rodents and mites (Ogawa and Ono 2008). The serological diversity (Tamura et al. 1984)

and genotypic diversity (Tamura et al. 2001) of *O. tsutsugamushi* have also been examined. These studies suggest that many genotypes/serotypes of *O. tsutsugamushi* are distributed throughout Japan and that the risk of infection with scrub typhus remains high.

Phylogenetic and phylogeographic analyses of *O. tsutsugamushi* have typically been based on the gene encoding the 56-kD type-specific antigen (TSA) protein thus far. This protein is expressed on the outer membrane surface and is the primary immunogen in human infections that is responsible for eliciting neutralizing antibodies. In general, a 56-kD TSA-based indirect fluorescence antibody technique (IFAT) is used for the diagnosis and classification of serotypes. There is a high diversity of 56-kD TSA genotypes among *O. tsutsugamushi*, and consequently type-specific PCR based on the 56-kD TSA gene is conducted to identify different genotypes. However, genes encoding major immune antigens that are modulated by the immune response of the host, such as the 56-kD TSA gene, are usually not suitable for phylogenetic and phylogeographic analyses. Housekeeping genes, such as heat shock protein genes or the 16S ribosomal RNA (rRNA) gene,

¹Infectious Disease Surveillance Center, National Institute of Infectious Diseases, Tokyo, Japan.

²The Shimane Prefectural Institute of Public Health and Environmental Science, Matsue, Shimane, Japan.

³The Saitama Prefectural Institute of Public Health, Saitama, Japan.

⁴Mahara Institute of Medical Acarology, Anan, Tokushima, Japan.

⁵University of Fukui, Eiheiji-cho, Fukui, Japan.

⁶Chiba Institute of Science, Choshi, Chiba, Japan.

⁷Rakuno Gakuen University, Ebetsu, Hokkaido, Japan.

⁸Kawasaki City Institute for Public Health, Kawasaki, Kanagawa, Japan.

are feasible candidates for such analyses because they are unaffected by the host immune response. The *groEL* gene, which encodes the 60-kD heat shock protein GroEL, is more diverse than the 16S rRNA gene. The sequences of the *groES* and *groEL* genes might be reconstructed more accurately with gene tree reconstructions, similar to 16S rRNA-based phylogeny. Thus, a comparison of the 56-kD TSA gene and the *groEL* or *groES* genes would help understand the molecular epidemiology of *O. tsutsugamushi*.

This study compared the entire *groES* and *groEL* gene sequences of six reference samples, *i.e.*, the Kaisei, Sato, Kato, Kuroki, Matsuzawa, and Shimokoshi strains, and 15 field strains of *O. tsutsugamushi*. Phylogenetic comparisons between the 56-kD TSA gene and the *groES* or *groEL* genes suggest that gene rearrangement may be associated with homologous recombination in mites.

Materials and Methods

Genomic DNA was extracted using a PureLink genomic DNA mini kit (Invitrogen, San Diego, CA) from six *O. tsutsugamushi* reference strains isolated from patients in Japan, *i.e.*, Kaisei, Sato, Kato, Kuroki, Matsuzawa, and Shimokoshi, and 15 field isolates from rodents in Japan, *i.e.*, UAP1, UAP2, UAP4, UAP6, UAP7, FAR1, FAR2, HSB1, HSB2, HSB3, KNP1, KNP2, CMM1, O2, and O3 from infected L929 cells (Fujita et al. 2000, Tamura et al. 2001). DNA samples from two wild

rodents in Japan (SH216 and 05-5) and two patients in Shimane, Japan (TD-7 and TD-17) were used for amplification (Tabara et al. 2012). Additionally, three DNA samples, SH205, SH234, and SH245, from rodents that were captured in Japan were also used for amplification. The gene encoding the *O. tsutsugamushi* 56-kD TSA protein was amplified from DNA samples by PCR using the primer set described by Furuya et al. (1993). To identify long sequences, 1722–1880 bp fragments were obtained using the primers Ot56-M329F (5'-CTAAGGTTAAYTGTCRAGGTGC-3') and Ot56-Rev1512R (5'-TCCACATACACACCTTCAGCAGCA-3'). In addition to the 56-kD TSA gene, the *groES* and *groEL* genes were also amplified. PCR was performed using oligonucleotide primers, which were newly designed based on the Ikeda strain and other *O. tsutsugamushi* strains. The primers groEL-14F (5'-TTGTACATRGCGATCAATGTCGT-3') and groEL-1667R (5'-TAGAAATCCATTCCGCCCATAC-3') were used for the first-round reaction, and the primers groEL-14F and groEL-772R (5'-GAGCTTCTCCGTCTA CATCATCAG-3') were used for the second-round reactions to detect the *groEL* gene. In addition, the primers groESL-M594F (5'-GGCTATARGCAATGATATAGCTAAACCAG-3') and groEL-2258R (5'-GCGAACTATTAGTGCTAG CAACTAC-3') were newly designed to identify the *groESL* gene operon of *O. tsutsugamushi*. First- and second-round PCR were performed using 20-μL reaction mixtures containing 250 μL of deoxyribonucleotide triphosphates

FIG. 1. Phylogenetic trees generated using the Bayesian Markov chain Monte Carlo estimation methods (MCMC) with the GTR+I+Γ model of evolution, which was estimated from the data based on the alignment of the coding regions of the 271- to 285-nucleotide *groES* (ES), 1651- to 1668-nucleotide *groEL* (EL), and 420- to 1481-nucleotide 56-kD type-specific antigen (TSA) (56K) genes of the following *Orientia tsutsugamushi* strains: Kaisei (56-kD TSA, JX235719; groESL, JX188391) from patient, Sato (56-kD TSA, JX235718; groESL, JX188392) from patient, Kato (56-kD TSA, M63382; groESL, JX188393) from patient, Kuroki (56-kD TSA, M63380; groESL, JX188394) from patient, Matsuzawa (56-kD TSA, AF173043; groESL, KC688324) from patient, Shimokoshi (56-kD TSA, M63381; groESL, JX188395) from patient, UAP1 (56-kD TSA, AF302991; groESL, JX188396) from *Apodemus speciosus*, UAP2 (56-kD TSA, AF302992; groESL, KC688329) from *A. speciosus*, UAP4 (56-kD TSA, AF302993; groESL, JX188397) from *A. speciosus*, USP6 (56-kD TSA, AF302994; groESL, KC688325) from *A. speciosus*, UAP7 (56-kD TSA, AF302995; groESL, JX188398) from *A. speciosus*, FAR1 (56-kD TSA, AF302989; groESL, JX188399) from *A. speciosus*, FAR2 (56-kD TSA, AF302990; groESL, KC688326) from *A. speciosus*, HSB1 (56-kD TSA, AF302983; groESL, JX188400) from *A. speciosus*, HSB2 (56-kD TSA, AF302984; groESL, JX188401) from *A. speciosus*, HSB3 (56-kD TSA, AF302985; groESL, KC688327) from *A. speciosus*, CMM1 (56-kD TSA, AF302986; groESL, KC688328) from *A. speciosus*, KNP1 (56-kD TSA, AF302987; groESL, KC688320), KNP2 (56-kD TSA, AF302988; groESL, KC688321), O2 (56-kD TSA, KC688322; groESL, KC688330) from *A. speciosus*, O3 (56-kD TSA, KC688323; groESL, KC688331) from *Apodemus argenteus*, SH205 (56-kD TSA, KC693730; groESL, KC688332) from *A. speciosus*, SH234 (56-kD TSA, KC693731; groESL, KC688333) from *A. speciosus*, SH245 (56-kD TSA, KC693732; groESL, KC688334) from *A. speciosus*, TD-7 (56-kD TSA, JX188388; groESL, JX235722) from patient, TD-17 (56-kD TSA, JX188387; groESL, JX235720) from patient, 05-5 (56-kD TSA, JX188390; groESL, JX235721) from rodent, and SH216 (56-kD TSA, JX188389; groESL, JX188402) from *Mus musculus*. The phylogenetic positions of these strains are also shown relative to the following representative *O. tsutsugamushi* strains: Gilliam (groESL, AY191585) from patient, Taiwan (56-kD TSA, DQ485289) from patient, Ikeda (56-kD TSA, AF173033; groESL, NC010793) from patient, Yunchon (56-kD TSA, U19903) from patient, Karp (56-kD TSA, M33004; groESL, M31887) from patient, Kawasaki (56-kD TSA, M63383; groESL, AY191587) from patient, Boryong (56-kD TSA, AM494475; groESL, NC009488) from patient, Nishino (56-kD TSA, AF173048) from patient, Young-worl (56-kD TSA, AF43141; groESL, AY191588) from patient, 402I (56-kD TSA, AF173047) from patient, Darnley (56-kD TSA, AY860955) from patient, FPW1038 (56-kD TSA, EF213087; groESL, EF551288) from patient, FPW2016 (56-kD TSA, EF213085; groESL, EF551289) from patient, Fuji (56-kD TSA, AF210834) from *Leptotrombidium fuji*, IHS I (56-kD TSA, DQ530440) from patient, Hwasung (groESL, QY191589) from patient, LF-1 (56-kD TSA, AF173050) from *Leptotrombidium fletcheri*, LX-1 (56-kD TSA, AF173042) from *Leptotrombidium mite*, TA678 (56-kD TSA, U19904) from *Rattus rattus*, TA716 (56-kD TSA, U19905) from *Menetes berdmorei*, TSA763 (56-kD TSA, RTU80636) from *Rattus rajah*, TW52-1 (56-kD TSA, AY222630) from *Rattus norvegicus*, UT302 (56-kD TSA, EF213095) from patient, and UT418 (56-kD TSA, EF213087; groESL, EF551310) from patient, with *Rickettsia helvetica* (DQ442911), *Rickettsia prowazekii* (Y15783), and *Rickettsia typhi* (AF462073) as outgroups. Genotypes based on the 56-kD TSA gene are shown on the right of 56K. The numbers at each node are the posterior node probabilities (>0.70) based on 150,000 trees. Two replicate MCMC runs consisting of six chains of 10 million generations were sampled every 100 generations with a burn-in of 25,000 (25%). The scale bar indicates the nucleotide substitutions per site. The country of origin is shown in parentheses.

(dNTPs), 1 U of Expand Long template DNA polymerase (Roche Applied Science, Penzberg, Upper Bavaria, Germany), and 0.25 μ M of each primer.

Initial denaturation was conducted at 94°C for 2 min, followed by 30 cycles of denaturation at 94°C for 40 s, annealing at 55°C for 30 s, and extension at 72°C for 2 min using a GeneAmp PCR 9700 thermal cycler (Perkin-Elmer, Waltham, MA). PCR products were separated with MobiSpin S-400 spin columns (MoBiTec, Goettingen, Germany), and amplicons were sequenced directly using an ABI Prism 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA). Sequences were processed using Genetyx version 9.1 (Genetyx Corporation, Tokyo, Japan) and aligned with CLUSTALW and CLUSTALW2 (Thompson et al. 1994) using publicly available *O. tsutsugamushi* sequences. The phylogenetic analysis was conducted using Bayesian interference trees, which were generated with Bayesian Metropolis Hastings Markov Chain Monte Carlo (MCMC) tree-sampling methods implemented in MrBayes (Ronquist and Huelsenbeck 2003) using a GTR + I + Γ model of evolution, *i.e.*, the hierarchical likelihood-ratio test (hLRT) in MrModeltest2.3 (<http://www.abc.se/~nylander/mrmodeltest2/mrmodeltest2.html>) (Posada and Crandall 1998) with partitioning based on codon positions.

Results

Sequence and phylogenetic analysis of the *groESL* gene

The 271- to 285-bp and 1651- to 1668-bp *groES* and *groEL* genes of *O. tsutsugamushi* encoded predicted GroES and GroEL proteins with 90–95 and 550–555 amino acids, respectively. The interspecies variation among strains in the *groES* and *groEL* genes was negligible (0–5.6% at the nucleotide level and 0–4.0% at the amino acid level), with the exception of the Shimokoshi strain. Analysis of the *groESL* gene operon, which measures approximately 2.8-kbp and includes the *groES* and *groEL* genes of *O. tsutsugamushi*, revealed four known conserved motifs (two ribosomal binding sites, GGAGGA and AGGAG; one –35 region, TTGAAA; and one –10 region, TATGA) (Stover et al. 1990). A potential *E. coli* sigma 32 heat shock promoter (TTCACACTTGAA) was also found 5' upstream of the *groES* gene. This study confirmed the overall high sequence similarity of the *groESL* gene operon among strains.

The partially sequenced 56-kD TSA gene, measuring 420–1481 nucleotides in length, was amplified from the genomic DNA of the Sato, CMM1, KNP1, KNP2, SH234, 05-5, and TD-17, and Kaisei, UAP6, FAR2, HSB3, TD-7, and SH216 strains, which showed that the sequences were very similar to those of the Karp-related-type and Gilliam/Kawasaki-type *O. tsutsugamushi*, respectively (Fig. 1, 56K). Exhaustive phylogenetic analyses based on the nucleotide and deduced amino acid sequences of the *groES* and *groEL* genes, which were generated using Bayesian MCMC estimation methods, indicated that both genes were very similar in this species. Interestingly, the two phylogenetic trees were not entirely identical, although they strongly resembled each other. The *groES* and *groEL* genes were highly conserved among species, and the differences between strains were less than 9.7%. Thus, highly conserved genes such as *groES* or *groEL* genes could be suitable targets for the molecular diagnosis of scrub typhus.

Discussion

The present study identified the *groESL* gene operon of *O. tsutsugamushi* that includes the *groES* and *groEL* genes. The shared identities among strains were 90.4–100% and 90.3–100% for the *groES* and *groEL* genes, respectively. Recently, the *groEL* gene was used to identify *Streptococcus*, *Enterococcus*, *Staphylococcus*, *Bartonella*, *Mycobacterium*, and *Ehrlichia*. The present study demonstrated that the sequences of the heat shock protein genes *groES* and *groEL* are highly conserved among *O. tsutsugamushi* strains. Phylogenetic analyses indicated that the 56-kD TSA gene strongly resembled, but was not identical to, the *groEL* and *groES* genes, and these gene sequences are distinctly different from those in *Rickettsia*.

There are reports that indicate that the 56-kD TSA gene is polymorphic among species (Kollars et al. 2003), but gene-based molecular techniques have not been applied to investigate the relationships of the 56-kD TSA gene. Interestingly, phylogenetic analyses revealed no correlation between the 56-kD TSA gene and geographic location of *O. tsutsugamushi* strains and vector *Leptotrombidium*, although the gene is highly polymorphic among species (Fig. 1).

Phylogenetic analysis of the *groES*, *groEL*, and TSA genes detected no common clustering in the dendrograms of each phylogenetic tree (Fig. 1). Two and three clusters were detected in the *groES* and *groEL* gene-based dendrograms, respectively. The clustered strains did not belong to the same genotype according to analysis of the 56-kD TSA gene. A recent report identified genetic recombination between *O. tsutsugamushi* and its vector mite (Sonthayanon et al. 2010), which presaged the mismatches between the 56-kD TSA, *groES*, and *groEL* genes. Thus, the genomic DNA of *O. tsutsugamushi* is undoubtedly rearranged in different strains. More importantly, these findings show that the evolutionary history and transmission dynamics of *O. tsutsugamushi* are far richer and more complex than originally imagined. Given the sympatric and synchronistic coexistence of several *O. tsutsugamushi* strains and different species of vector mites in the field over evolutionary time, it seems plausible that ongoing exchanges within *O. tsutsugamushi* continue to drive its evolution. Phylogenetic analyses based on the *groESL* gene using more isolates will unveil further interesting evolutionary processes that have occurred in *O. tsutsugamushi*.

Acknowledgments

We thank Masashi Hamada for his technical assistance. This work was supported in part by a grant-in-aid from the Ministry of Health, Labor, and Welfare of Japan (Research on Emerging and Re-emerging Infectious Diseases, Health Science Research Grants), and by a grant-in aid from the Japan Society for the Promotion of Science 24405045 (Scientific Research grant B).

Author Disclosure Statement

No competing financial interests exist.

References

- Bang HA, Lee MJ, Lee WC. Comparative research on epidemiological aspects of tsutsugamushi disease (scrub typhus) between Korea and Japan. *Jpn J Infect Dis* 2008; 61:148–150.

- Fujita H, Watanabe Y, Hagiwara T, Takada N. Trombiculid mites and *Orientia tsutsugamushi* in field rodents in the northern part of Fukushima prefecture. *Ann Rep Ohara Hosp* 2000; 43:33–40.
- Furuya Y, Yoshida Y, Katayama T, Yamamoto S, et al. Serotype-specific amplification of *Rickettsia tsutsugamushi* DNA by nested polymerase chain reaction. *J Clin Microbiol* 1993; 31:1637–1640.
- Kelly DJ, Fuerst PA, Ching WM, Richards AL. Scrub typhus: The geographic distribution of phenotypic and genotypic variants of *Orientia tsutsugamushi*. *Clin Infect Dis* 2009; 48(Suppl 3):S203–S230.
- Kollars TM, Jr., Bodhidatta D, Phulsuksombati D, Tippayachai B, et al. Short report: Variation in the 56-kD type-specific antigen gene of *Orientia tsutsugamushi* isolated from patients in Thailand. *Am J Trop Med Hyg* 2003; 68:299–300.
- Ogawa M, Ono T. Epidemiological characteristics of tsutsugamushi disease in Oita Prefecture, Japan: yearly and monthly occurrences of its infections and serotypes of its causative agent, *Orientia tsutsugamushi*, during 1984–2005. *Microbiol Immunol* 2008; 52:135–143.
- Posada D, Crandall KA. MODELTEST: Testing the model of DNA substitution. *Bioinformatics* 1998;14:817–8.
- Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 2003; 19:1572–1574.
- Sonthayanon P, Peacock SJ, Chierakul W, Wuthiekanun V, et al. High rates of homologous recombination in the mite endosymbiont and opportunistic human pathogen *Orientia tsutsugamushi*. *PLoS Negl Trop Dis* 2010; 4:e752.
- Stover CK, Marana DP, Carter JM, Roe BA, et al. The 56-kilodalton major protein antigen of *Rickettsia tsutsugamushi*: Molecular cloning and sequence analysis of the sta56 gene and precise identification of a strain-specific epitope. *Infect Immun* 1990; 58:2076–2084.
- Tabara K, Kawabata H, Ando S, Arai S, et al. Epidemiological study of tsutsugamushi disease in Shimane Prefecture, Japan. *J Jpn Vet Med Assoc* 2012; 65:535–541.
- Tamura A, Takahashi K, Tsuruhara T, Urakami H, et al. Isolation of *Rickettsia tsutsugamushi* antigenically different from Kato, Karp, and Gilliam strains from patients. *Microbiol Immunol* 1984; 28:873–882.
- Tamura A, Yamamoto N, Koyama S, Makisaka Y, et al. Epidemiological survey of *Orientia tsutsugamushi* distribution in field rodents in Saitama Prefecture, Japan, and discovery of a new type. *Microbiol Immunol* 2001; 45:439–446.
- Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994; 22:4673–4680.

Address correspondence to:

Satoru Arai
Infectious Disease Surveillance Center
National Institute of Infectious Diseases
Toyama 1-23-1
Shinjuku, Tokyo 162-8640
Japan

E-mail: arais@nih.go.jp