Molecular tracing of Japan-indigenous hepatitis E viruses

Yasuhiro Tanaka,1 Kazuaki Takahashi,2 Etsuro Orito,1 Yoshiyasu Karino,3 Jong-Hon Kang,4 Kazuyuki Suzuki,5 Atsushi Matsui,6 Akiko Hori,7 Hiroyuki Matsuda,8 Hiroshi Sakugawa,9 Yasuhiro Asahina,10 Tsuneo Kitamura,11 Masashi Mizokami1 and Shunji Mishiro2

The ancestor(s) of apparently Japan-indigenous strains of Hepatitis E virus (HEV) was probably of foreign origin, but it remains unclear when and from where it made inroads. In this study, 24 genotype 3 and 24 genotype 4 HEV strains recovered in Japan each showed a significant cluster, clearly distinct from those of foreign strains, in the phylogenetic tree constructed from an 821 nt RNA polymerase gene fragment. The evolutionary rate, approximately 0.86 x 10^-3 nucleotide substitutions per site per year, enabled tracing of the demographic history of HEV and suggested that the ancestors of Japan-indigenous HEV had made inroads around 1900, when several kinds of Yorkshire pig were imported from the UK to Japan. Interestingly, the evolutionary growth of genotype 3 in Japan has been slow since the 1920s, whereas genotype 4 has spread rapidly since the 1980s. In conclusion, these data suggest that the indigenization and spread of HEV in Japan were associated with the popularization of eating pork.
Japan. In this study, we first estimated the evolutionary rate of HEV by using Japan-indigenous genotype 3 and genotype 4 strains, which were phylogenetically distinct from the other strains in foreign countries. Then, based on this evolutionary rate, we traced the demographic history of HEV in Japan.

For linear-regression analyses within significant clusters, two independent datasets were applied: one was a Hyogo cluster (genotype 3) with JMO-Hyo03L, JTH-Hyo03L, JSO-Hyo03L, JYO-Hyo03L, JDEER-Hyo03L (these five isolates were obtained in April 2003) and JBOAR1-Hyo04 (April 2004) (Takahashi et al., 2004a), and another was a Sapporo cluster (genotype 4) with JSM-Sap95 (March 1995), JKK-Sap00 (November 2000), JYWSap02 (August 2002) and JTS-Sap02 (September 2002) (Takahashi et al., 2004b). GenBank accession numbers for these strains are given in Fig. 1. To elucidate the epidemiological history of the HEV population in Japan, 48 known and newly sequenced HEV strains (n = 24 for each of genotype 3 and 4) were used for molecular-evolutionary analyses. The nucleotide sequences of 28 strains for the molecular-clock analyses were determined in this study (the other 20 sequences dealt with in this paper were available from GenBank).

Nucleic acids were extracted from serum samples (50 μl) by using a commercial Smitest EX-R & D kit (Genome Science) and precipitated in a 2 ml tube. The pellet was air-dried for 15 min and then suspended in 10 μl autoclaved distilled water containing 10 U RNase inhibitor ml⁻¹ (TaKaRa

---

**Fig. 1.** Phylogenetic tree of the partial RNA polymerase region of the HEV genome. Twenty-four genotype 3 and 24 genotype 4 strains in Japan showed each significant cluster to have a high bootstrap value and to be distinct from other reference sequences (USA, Canada and Japanese minor strains in genotype 3; Chinese strains in genotype 4). Genetic distances have been transformed into a time scale of years by using estimates of the molecular clock (0.84 ± 0.04 nucleotide substitutions per site per year). Ten strains in bold are used for linear regression in Fig. 2. Strain names are followed by prefecture or city names in Japan: Hyo, Hyogo; Tok, Tokyo; Sai, Saitama; Sap, Sapporo; Iwa, Iwate; Kan, Kanagawa; Oki, Okinawa; Aom, Aomori; Nar, Nara; Tot, Tottori; Nii, Niigata; Toc, Tochigi; Toy, Toyama. Asterisks indicate strains that were newly sequenced in this study.
A sequence spanning 821 nt in the RNA-dependent RNA polymerase region (corresponding to nt 3961–4781 of the prototype Burmese HEV strain; GenBank accession no. M73218), including the GDD motif, was amplified by PCR in three overlapping regions with 20-mer primers deduced from known HEV sequences. Reverse transcription was carried out in the presence of Platinum Taq DNA Polymerase High Fidelity (Invitrogen). The final products were sequenced in an ABI 377 DNA sequencer (PE Biosystems) with an ABI Prism BigDye kit (Applied Biosystems). The sequences determined were utilized to confirm HEV genotypes and to construct phylogenetic trees. The reliability of the phylogenetic tree was assessed by bootstrap-resampling tests.

A reconstructed tree was built on the RNA polymerase region by using a heuristic maximum-likelihood (ML) topology search with stepwise addition and nearest neighbour-interchange algorithms. Tree likelihood scores were calculated by using the HKY85 model (Hasegawa et al., 1985) with the molecular clock enforced, using PAUP version 4.0b8. Using the estimated topology, all possible root positions were evaluated under a single-rate dated-tips (SRDT) model with the computer software TipDate v1.2 and the root that yielded the highest likelihood was adopted (Rambaut, 2000). The program provided an ML estimate of the rate and also the associated date of the most recent common ancestor of the sequences, using a model that assumed a constant rate of nucleotide substitution. The molecular clock was tested by a likelihood-ratio test between the SRDT model and a general unconstrained branch-length model [different-rate (DR) model].

For estimates of demographic history, a non-parametric function $N(t)$, also known as a skyline plot, was obtained by transforming the coalescent intervals of an observed genealogy into a piecewise plot that represented an effective population size through time (Pybus et al., 2001; Pybus & Rambaut, 2002). A parametric ML was estimated by several models with the computer software GENIE v3.5 to build a statistical framework for inferring the demographic history of a population on phylogenies reconstructed from sampled DNA sequences (Pybus & Rambaut, 2002). This model assumes a continuous epidemic process in which the viral transmission parameters remain constant through time. Model fitting was evaluated by likelihood-ratio tests of the parametric ML estimates (Lemey et al., 2003; Pybus et al., 2003; Tanaka et al., 2005). Approximate 95 % confidence intervals for the parameters were estimated by using the likelihood-ratio test statistics.

A phylogenetic tree in the partial RNA polymerase region of the HEV genome is represented in Fig. 1. A functional gene, such as the RNA polymerase gene, is suitable for molecular-evolutionary analyses based on the neutral theory, because the substitution of functional genes is based on the neutral theory. The 24 genotype 3 and 24 genotype 4 strains in Japan showed a significant cluster with a high bootstrap value, which was the major Japanese cluster distinct from other strains found in foreign countries by molecular-evolutionary analyses. Such a significant cluster is suitable for the following coalescent analysis. Additionally, the tree topology based on the RNA polymerase region, including functional genes, was quite similar to that based on complete genomes (data not shown).

To determine the evolutionary rate of HEV, the 48 Japan-indigenous HEV strains (Fig. 1) were subjected to further molecular-evolutionary analyses. The molecular-evolutionary rate was estimated by two independent methods. In brief, linear-regression analyses using highly similar strains, i.e. six genotype 3 strains in Hyogo and four genotype 4 strains in Sapporo, indicated that a molecular-evolutionary rate was $(0.81 \pm 0.88) \times 10^{-3}$ nucleotide substitutions per site per year (Fig. 2). Second, TipDate (v1.2) was used to compare the DR model with the single-rate (SR) and SRDT models. The SRDT model provided an adequate fit to the data ($P>0.05$; see Supplementary Table S1, available in JGV Online). Based on the SRDT model, the mean rate of nucleotide substitutions was estimated to be $(0.81 \pm 0.94) \times 10^{-3}$ nucleotide substitutions per site per year.

**Fig. 2.** Linear-regression analyses within the partial RNA polymerase region for evolutionary rate of HEV. (a) The evolutionary rate of genotype 3 in the Hyogo cluster is estimated to be $0.88 \times 10^{-3}$ nucleotide substitutions per site per year; (b) the evolutionary rate of genotype 4 in the Sapporo cluster is estimated to be $0.81 \times 10^{-3}$ nucleotide substitutions per site per year.
substitutions per site per year, which was similar to the rate for Hepatitis C virus (Ina et al., 1994; Tanaka et al., 2002). When we used $0.84 \times 10^{-3}$ nucleotide substitutions per site per year, which was based on all 48 sequences (24 genotype 3 and 24 genotype 4), the time of the most recent common ancestor of Japan-indigenous genotype 3 was estimated to be in the 1900s (95% confidence interval, 1902–1917) and that of genotype 4 was approximately in the 1880s (1881–1898) (Fig. 1).

Based on the phylogenetic tree, the effective number of HEV infections through time, $N(t)$, was analysed by using a skyline plot for the Japan-indigenous HEV strains. The parameters for several models in GENIE v3.5 were examined (see Supplementary Table S2, available in JGV Online). Time $t$ was then transformed to year by using the constant rate ($0.84 \times 10^{-3}$ nucleotide substitutions per site per year), assuming the collecting time to be the present. Fig. 3 shows the skyline plots and population growth for the HEV strains, according to a specific demographic model in GENIE v3.5 with three parameters and a piecewise-expansion growth model, which was evaluated by likelihood-ratio testing (Ina et al., 1994; Lemey et al., 2003; Pybus et al., 2003; Tanaka et al., 2005). Our estimates of the effective numbers of HEV infections showed a transition from constant size to exponential growth in the 1920s (95% confidence interval, 1916–1930) among the genotype 3 population (Fig. 3a), whereas the rapid exponential growth among the genotype 4 population was dated in the 1980s (1978–1990) (Fig. 3b).

Because the natural course of HEV infection in human beings and animals is usually transient, not persistent as in the cases of hepatitis B and C viruses, it is almost impossible to estimate the molecular-evolutionary rate of HEV by using serial samples from an individual host. However, even though HEV does not persist in individual hosts, it could persist in the community by hopping from host to host successively. The first study attempting to estimate the number of synonymous mutations per synonymous site ($k_s$) of Hepatitis A virus (HAV) was reported by Sánchez et al. (2003). The estimated $k_s$ values from HAV strains isolated from a clam-associated outbreak varied from 0.038 for VP0 to 0.29 for VP1. Similarly, we estimated the evolutionary rate of HEV by using Japan-indigenous genotype 3 and genotype 4 strains isolated over time. The rate was estimated to be approximately $8 \times 10^{-3}$ nucleotide substitutions per site per year by two independent methods, which was around half of our previously estimated rate (Takahashi et al., 2004b). One of the reasons is that the molecular-evolutionary rate would depend on estimated genes; the previous report (Takahashi et al., 2004b) used complete sequences, whereas this study used only RNA polymerase sequences. Another reason is that the previous extrapolation of substitution rate on pairwise (direct) comparisons can give overestimates of the molecular clock and hence divergent times of HEV species, as reported previously (Ina et al., 1994). Based on the molecular clock, we traced the demographic history of HEV in Japan and the indigenization time was suggested to be similar (approx. 1900), but the spread time was quite different, between HEV genotypes 3 and 4 (1920s versus 1980s). Interestingly, in addition, the evolutionary growth of genotype 3 has been quite slow since the 1920s, whereas genotype 4 strains have spread rapidly in Sapporo since the 1980s.

Zoonosis has been implicated in HEV transmission. The first animal strain of HEV to be isolated and characterized was a swine HEV from a pig in the USA in 1997 (Meng et al., 1997). Since then, many swine HEV strains, which exhibit extensive genetic heterogeneity, have been identified worldwide and shown to be genetically related closely to strains of human HEV (Chandler et al., 1999; Hsieh et al., 1999; Huang et al., 2002; Okamoto et al., 2001; Wang et al., 2002). Recent findings suggested an interspecies HEV transmission between boar and deer in their wild life (Takahashi et al., 2004a) and that both animals might serve as an infection source for human beings. More recently, wild mongoose was newly added to the list of HEV-reservoir animals in Japan.

**Fig. 3.** ML estimates of $N(t)$ on the effective number of (a) HEV genotype 3 and (b) HEV genotype 4 infections in Japan. The parametric model is indicated by the black line and step-wise plots by the grey line, which represents corresponding non-parametric estimates of $N(t)$ (number as a function of time). Genetic distances have been transformed into a time scale of years by using estimates of the molecular clock in the partial RNA polymerase region of HEV.
(Nakamura et al., 2006). Notwithstanding the importance of these wild animals, pigs for food must be the major reservoirs of HEV: a recent Japanese study indicated that anti-HEV antibodies were detected in 1448 (58 %) of 2500 pigs from 2 to 6 months of age at 25 commercial swine farms in Japan (Takahashi et al., 2003). The importance of transmission of HEV from pigs to humans was further supported by a recent field study in Indonesia: Muslim people, for whom it is a taboo to eat or contact pigs, were significantly less frequently positive for anti-HEV than Hindu people (2-0 vs 20 %) (Surya et al., 2005).

Our molecular-evolutionary analyses suggested that HEV entered Japan around 1900. If we have traced the origin of Japan-indigenous HEV correctly back to about 100 years ago, what happened at that time in relevance to HEV’s indigenization? Several kinds of Yorkshire pig were imported for the first time in the history of Japan from the UK in 1900, by the Japanese government’s policy to introduce excellent domestic animals for food in Western countries to Japan, as a measure to nutritionally strengthen the people (especially soldiers) of this formerly vegetarian country. Since then, the Yorkshire pigs have been propagated in Japan and, in the 1930s, thousands of pigs were reported all over Japan (http://okayama.lin.go.jp/history/2-3-1-2.htm), suggesting that the domestic spread of HEV might have been associated with the popularization of pigs for food in Japan. Indeed, a previous phylogenetic analysis of 304 bp nucleotide sequence (ORF2) obtained from the two UK swine strains showed a close relationship with Japanese swine strains in genotype 3 (Banks et al., 2004), indicating that Japanese genotype 3 may have been imported from the UK. On the other hand, Japanese genotype 4 strains were related phylogenetically to Asian strains in Taiwan and China. As the HEV found in wild boars living in the Iriomote Island, near Taiwan, was of genotype 4 (unpublished results), the source of Japanese genotype 4 might be from Taiwan or the mainland of China. Note that a phylogenetic analysis showed that the Japanese swine and human HEV strains segregated into four clusters [three genotype 3 clusters (one major Japanese and two minor clusters) and one genotype 4 cluster], with the highest nucleotide identity being 94-4-100 % between swine and human strains in each cluster (Takahashi et al., 2003), suggesting that swine have served as one of the most important reservoirs for HEV to be transmitted to humans. The possible risk factor for transmission of HEV was to have eaten uncooked or undercooked pig liver and/or intestine 1-2 months before the onset of hepatitis E in Hokkaido, Japan (Mizuo et al., 2005). Such eating habits, which are particularly unique to those living in Hokkaido (Sapporo is one of the big cities there) in recent decades, might be one of the reasons that HEV has been widespread in this area since 1990, as supported by our molecular-evolutionary analyses in this study.

In conclusion, based on our present data, the indigenization and domestic spread of HEV in Japan are proposed to have been associated with the importation and popularization of pigs for food in Japan. However, there still remains a possibility of different scenarios. Another animal(s) might have carried the virus to Japan: for example, mongoose was imported from India to Japan in 1910 (Nakamura et al., 2006).

Acknowledgements

Contributions of authors are as follows: Y.T. performed molecular-clock analyses and wrote the manuscript; K.T. amplified and sequenced viral isolates; E.O. helped Y.T. with the molecular-clock analyses; Y.K., J.-H. K., K.S., A.M., A.H., H.M., H.S., Y.A. and T.K. provided HEV RNA-positive sera to K.T. for sequence determinations; M.M. supervised the molecular-clock analyses; and S.M. designed the study and helped Y.T. to write the manuscript. This work was supported in part by grants from the Ministry of Health, Labour and Welfare of Japan (200400676A) and from the United States–Japan Collaborative Medical Science Program (Hepatitis Panel). We greatly appreciate Dr Oliver G. Pybus (Department of Zoology, University of Oxford, Oxford, UK) for his enlightening advice on molecular-evolutionary analyses using GENIE v3.5.

References


