Table 1 Classification of presented viruses

Flavivirus[4]
Dengue virus,[1–4] ***Japanese encephalitis virus*** (West Nile virus, Kunjin virus, Murray Valley encephalitis virus, Japanese encephalitis virus, St. Louis encephalitis virus), ***Yellow fever virus*** (yellow fever virus), ***Tick-borne encephalitis virus*** (Powassan virus, Langat, louping ill virus, Central European encephalitis virus, Russian spring-summer encephalitis virus)
Rabies[19,21]
Lyssavirus (rabies virus, Lagos bat virus, Mokola virus, Duvenhage virus, European bat lyssavirus 1, European bat lyssavirus 2, Australian bat virus)
Hantavirus[24]
Hantaan group (Hantaarn virus, Seoul virus, Dobrava virus, Saaremaavirus)
Puumala group (Puumala virus, Prospect Hill virus, Tula virus, Khabarovsk virus, Topograv virus, Isla Vista virus, Bloodland Lake virus)
Sin Nombre group (Sin Nombre virus, New York virus, Bayou virus, Black Creek Canal virus, Laguna Negra virus, Andes virus, Rio Mamore virus, Rio Segundo virus, El Moro Canyon virus)
Subfamily Paramyxovirunae[27]
***Morbillivirus*–Henipavirus** (Nipah virus, Hendravirus)
Subfamily Pneumovirinae[30]
Metapneumovirus (avian pneumovirus, human metapneumovirus)
Coronavirus[35]
group I (human coronavirus 229E, canine coronavirus, feline infectious peritonitis virus, porcine transmissible gastroenteritis virus, porcine epidemic diarrhea virus)
group II (human coronavirus OC43, bovine coronavirus, porcine hemagglutinating encephalomyelitis virus, rat sialodacryoadenitis virus, mouse hepatitis virus)
group III (turkey coronavirus, avian infectious bronchitis virus)
group IV (severe acute respiratory syndrome coronavirus)

Bold italics represents genus. Subfamily **Paramyxovirinae** and Subfamily **Pneumovirinae** belong to Family **Paramyxoviridae**. See references for more information on these viruses and for phylogenetic trees.

viruses, and other structural and non-structural proteins might also affect the virulence.[7] Host immune response also affects the virulence of these viruses. At present, there is no clear explanation for the dengue and dengue hemorrhagic fever.

Japanese encephalitis virus

Japanese encephalitis virus belongs to *Flaviviridae* and is the prototype of Japanese encephalitis antigenic complex.[4,9] Humans are infected through the virus-infected *Culex* mosquito bites. One of 300–1000 infected persons can develop encephalitis. Nearly 20% of encephalitis cases die and approximately 50% recover with sequel. Approximately 50 000 encephalitis cases and 10 000 death cases are annually recorded. Life cycle of *Culex*-pig/bird-*Culex*-human is well known but there is no evidence of human to human transmission. Pigs and birds are effective viral amplification hosts, serving as natural hosts for infection. The gene of Japanese encephalitis virus is approximately 11 kb positive single-stranded RNA. The composition of structural and non-structural proteins are the same as in dengue virus. The serotype is one, but between four and six genotypes exists in the envelope gene. Geographic variation has also been reported.[10,11] Although the induction of mutation in envelope gene affect the difference in neurovirulence, other structural and non-structural genes also affect neurovirulence. Phylogenetic tree of nucleic acids are mainly envelope gene. Antigen of antigen-antibody reaction is also an envelope antigen.

Yellow fever virus

Yellow fever virus belongs to *Flavivirus*.[4,8] Hemorrhagic fever type has a high mortality rate. *Aedes* mosquitoes act as vectors in Africa and *Haemagogus* mosquitoes are vectors in South America. Humans are natural hosts in Africa and the virus is maintained through the cycle of *Aedes*-human-*Aedes* transmission. In contrast, monkeys are natural hosts in South America and the virus is maintained through the cycle of *Haemagogus*-monkey-*Haemagogus*. Humans are mainly infected in forest areas, by mosquito bites. The virus spread to the United States with the transportation of Africans to the United States.[1] At present, yellow fever is not prevalent in Asia. Its structural and non-structural compositions are same as dengue virus. Two genotypes are found from envelope gene analysis: Group I was isolated in east and central Africa, Group IIA in west Africa and Group IIB in America.[12,13] Although the envelope gene makes the differences of infectivity and pathogenicity, other structure and non-structure genes also affect infectivity and pathogenicity. Envelope gene and proteins are used for diagnosis and molecular epidemiology.[8]

West Nile virus

The virus survives through the cycle of bird and *Culex* mosquito. Usually there are no symptoms in humans, however this virus can cause encephalitis, especially in the elderly.

The virus can also cause death in humans, although this is rare and deaths are more common in birds and horses. West Nile virus has been found mainly in Africa, and often in the Middle East, South Europe and West Asia. However, it was reported in the United States in 1999.[14] At the time of writing there are no reports of West Nile fever disease in Asia except India. Considering the spread of West Nile virus in the United States, outbreaks in Asian countries are highly likely to occur in the near future.

The virus has been detected in *Culex* mosquito, birds and animals in urban areas in the United States. Viremia is found in the early stages of the infection, and the virus is transmitted through the placenta and blood transfusion also. Two genotypes are known: Genotype I is found in Europe and the United States and Genotype I and II are found in Africa.[15] Genotype I viruses can be further subdivided based on the origins of the viruses; European and the United States, Australia and India. Kunjin virus is common in Australia (Table 1).[16]

St. Louis encephalitis virus is common in the United States and belongs to genus *Japanese encephalitis virus*.[4] The virus is carried through the cycle of bird and *Culex* mosquito. Human and cattle are often infected.

Zoonotic virus

Being the largest continent, most Asian countries can be connected via land to European and African countries, and animals can move easily from one country to another. In this section, rabies virus and hantavirus are described. Lassa virus, Ebola virus and Marburg virus are not included because the diseases are not found or are quite rare in Asia.[1,17,18]

Rabies virus

Rabies virus belongs to *Rabdoviridae* and is approximately 12 kb negative single stranded RNA virus. The virus causes zoonosis among human, animals and birds.[19] The virus in the salivary glands of animals invades from the bitten site. Other routes such as corneal transplant, eye injury, and respiratory infection have also been reported.[20] The virus infects muscle cells and moves to peripheral nerves, and eventually moves and grows in central nervous cell. It is fatal when it infects the brain. In Japan, rabies infections were first reported in humans in 1954, and in animals in 1856.[21] Wild foxes, wolves, skunks, raccoons and bats are natural hosts in 'forest type' rabies and it has spread worldwide with the exception of a few countries including Japan. Dogs are a common cause of the spread of 'urban type' rabies in Asia, Africa and middle/south America.

Pre- or postexposure prophylactic vaccinations are useful for the prevention of virus invasion to the central nervous system. The virus consists of nucleoprotein (N protein), non-structure protein (NS protein), membrane protein (M protein), glycoprotein (G protein) and large protein (L protein). G protein is closely related to infection (neurotoxin) and is a target for infection control. Recombinant vaccine has been developed using the G gene. Negri body in neuronal cells is found in the hippocampus of infected animals. The Negri body is composed of virus nucleoprotein. The nucleoprotein gene is conserved well and isuseful for molecular epidemiological analyses of the infected area. Genotype 1 (serotype 1) of classical rabies in Asia, Europe, Africa and the United States develop in separate clusters. The cluster is also differentiated within Asia. Sequence also has further small variations in the Philippines between Luzon Island and Mindanao Island.[22]

Rabies virus was previously thought to have only one serotype, but a new type which shows the symptom of classical rabies has been found in bats. One example is Australian bat virus of genotype 7, and a case of human death involving this genotype has been reported.[7]

Hantavirus

Hantavirus belongs to genus Hanta of *Bunyaviridae*. It is divided into two groups; the first group includes the virus which spreads in Eurasian Continent and causes hemorrhagic fever with renal syndrome. The other group is widespread in South and North America and causes hantavirus pulmonary syndrome. There are 24 species (serotypes/genotypes); four species cause hemorrhagic fever with renal syndrome and six species cause hantavirus pulmonary syndrome.[23] The virus is secreted through urine, stool and saliva from rodents. Humans who aspirate or contact the virus can have the disease. In infected rodents, infection persists without symptoms. Hantavirus causes zoonosis except for one species. Bunyavirus consists of three negative stranded RNA (L, M and S). Each segment codes L protein (polymerase), envelope protein (G1 and G2) and Nucleoprotein (NP). The molecular epidemiological analysis is mainly done in nucleotides of M gene. Three different clusters were discovered: (1) Hantaan virus, Dodrava virus, Seoul virus and Thailand virus from Murinae rodents; (2) Puumala virus from Arvicolinae rodents which cause hemorrhagic fever with renal syndrome; and (3) Sin Nombre virus, New York virus, Bayou virus, Black Creek Canal virus, Andes virus, Lagna Negra virus from Sigmodontinae rodents which causes hantavirus pulmonary syndrome.[24,25]

New viruses

Outbreaks of Nipha virus and SARS coronavirus occurred in 1998–1999 and 2002–03, respectively. These viruses were found mainly in Asia, although they were not common in

children. Methapneumovirus of unknown origin was reported as the cause of respiratory infection in children in 2001. These emerging viruses warn us of the challenges humanity faces in the new century.

Nipah virus

Nipah virus is a negative single stranded RNA virus and belongs to *Paramyxoviridae*. An outbreak of Nipah virus was recorded in 1998–1999 in Malaysia.[26] This virus was transmitted from pig to humans and it caused encephalitis in approximately 40% of infected persons. Pigs were infected from the urine of fruit bats living in the forest of Nipah. Pig to pig and pig to human transmission occurs through saliva and urine. Human to human transmission is rarely recognized. Respiratory syndrome is common in pigs but convulsions, abnormal movement and miscarriage are also recognized. The mortality rate for this virus is low. The virus infects dogs, cats and horses with the rise of the antibody level. The virus reached Singapore via pigs imported from Malaysia. More than 10 humans were infected and one person died. Hendra virus is similar to Nipah virus and was discovered in horses and spread in Australia in 1994.[27] It was transmitted to humans and caused respiratory symptoms. Currently classification of Nipah virus and Megapneumovirus is recommended as belonging to the genus Henipavirus, because the molecular size of these viruses are approximately 18 kb and different from other viruses of Paramyxoviridae (15–16 kb).[27-29]

Metapneumovirus

A new virus, metapneumovirus, was discovered in 2001 in a child with respiratory symptoms.[30] The same group had been recognized in animals and this was the first report in humans. This virus was included in mononegavirus (negative single strand RNA virus). The virus belongs to genus *Metapneumovirus*, subfamily Pneumovirinae in family Paramyxoviridae, but does not belong to genus *Pneumovirs* which includes respiratory syncytium virus.[30] The virus has characters of *Paramyxoviridae*, that is, glycoprotein and fusion protein. However there are no characteristics of hemagglutinin and pneuraminidase. Respiratory syncytium virus has been known to have one serotype (genotype) but recently it was divided into subgenus (suserotype) A and B.[31] There are differences within one subtype across geographic locations and time. More detailed analysis is required for the complete understanding of metapneumovirus.

Almost all children aged <5 years old have antibodies against metapneumovirus.[30,32] This was found in countries worldwide, including Japan.[33]

Severe acute respiratory syndrome coronavirus

Coronavirus is approximately 27–31 kb plus single stranded RNA virus. There are genus *coronavirivirus* and *torovirus*. Human coronavirus was cultured in human tracheal cells in 1965. Coronavirus was discovered in vertebrates and birds. Membrane protein, envelope protein, spike and glycoprotein with hemagglutinin and esterase function, and nucleoprotein have been observed. Coronavirus causes mainly respiratory infection, however it can also cause gastrointestinal infection and rarely neuronal infection. Outbreaks of gastroenteritis have also been reported. Molecular epidemiological studies, understanding of variation of serotype and re-infection are not well established for this virus. Human coronavirus spreads mostly in children in winter. Infection routes are droplet and/or airborne infection and fecus-oral infection. Human coronavirus 229E and OC43 strains were known long before the recent severe acute respiratory syndrome (SARS) coronavirus outbreak in China.[34,35]

Severe acute respiratory syndrome first occurred in Kuangtung in Guangdong province in China. Three-hundred-and-five people had respiratory disease and five persons died at that time. Following that, a SARS outbreak occurred in hotels, hospitals and condominiums in Hong Kong in March, 2003. Finally it spread to 29 countries including Vietnam, Taiwan, Singapore and Canada, and was spread mainly via travellers. The disease spread in hospitals where health workers and relatives of SARS patients were in close with patients. Fortunately, SARS disappeared in June 2003. More than 8000 people were infected and more than 800 people died of SARS in the 2003 outbreak. Considering the course of this disease, the disease started from the breeder of palm civet in Guangdong province and it is a zoonosis. The transmission route is usually droplet and/or air infection. However, fecus-oral infection and contact infection also occur. Effective management of a new respiratory infection is not easy. Research on diagnosis, treatment and prophylaxis are under way in order to counter the reappearance of SARS in the winter of 2004.[36]

There are four serotypes in Coronaviridae, and they are described in Table 1. There is some variation in the SARS coronavirus, although the variation is minor and at this time there is no need for subtyping.[34,35]

Conclusion

Changes in viral gene and host defense mechanisms are important factors in the spread of infection. Changes in habit and environment also affect the spread of infection. Molecular approach nowadays gives us quick and detailed analysis. In this paper, we have briefly described the following viruses at the molecular level: flaviviruses (dengue virus, Japanese

encephalitis virus, yellow fever virus, West Nile virus); zoonotic viruses (rabies virus, hantavirus); and new viruses (Nipah virus, metapneumovirus, severe acute respiratory syndrome coronavirus).

References

1 Mackenzie JS, Chua KB, Daniels PW *et al.* Emerging viral diseases of Southeast Asia and Western Pacific. *Emerg. Infect. Dis.* 2001; 7 (Suppl.): 497–504.

2 Storch GA. Diagnostic Virology. In: *Fields Virology* (eds Knipe, DM, Howley, PM). Lippincott, Williams & Wilkins, Philadelphia 2001; 493–532.

3 Shimada K. Gene diagnosis of infectious diseases. *Nippon Rinsho* 1992; **50** (Suppl.): 1–6 (in Japanese).

4 Burke DS, Monath TP. Flaviviruses. In: *Fields Virology* (eds Knipe, DM, Howley, PM). Lippincott, Williams & Wilkins, Philadelphia 2001; 1043–125.

5 Kuno G, Chang GJ, Tsuchiya KR, Karabatsos N, Cropp CB. Phylogeny of the genus flavivirus. *J. Virol.* 1998; **72**: 73–83.

6 Wang E, Ni H, Xu R *et al.* Evolutionary relationship of endemic/epidemic and sylvatic dengue viruses. *J. Virol.* 2000; 74: 3227–34.

7 Kurane I. Dengue virus. *Nippon Rinsho* 2003; **61** (Suppl. 3): 486–90 (in Japanese).

8 Monath TP, Tsai TF. Flaviviruses. In: *Clinical Virology* (eds Richman, DD *et al.*). ASM Press, Washington DC 2002; 1097–151.

9 Kurane I. Japanese encephalitis virus. *Nippon Rinsho* 2003; **61** (Suppl. 3): 491–3 (in Japanese).

10 Solomon T, Ni H, Beasley DW, Ekkelenkamp M, Cardosa MJ, Barrett AD. Origin and evolution of Japanese encephalitis virus in southeast Asia. *J. Virol.* 2003; 77: 3091–8.

11 Holbrook MR, Barrett ADT. Molecular epidemiology of Japanese encephalitis virus. *Curr. Top. Microbiol. Immunol.* 2002; **267**: 75–90.

12 Kurane I. Yellow fever. *Nippon Rinsho* 2003; **61** (Suppl. 3): 494–6 (in Japanese).

13 Chang GJ, Cropp BC, Kinney RM, Trent DW, Gubler DJ. Nucleotide sequence variation of the envelope variation of the envelope protein gene indentifies two distinct genotypes of yellow fever virus. *J. Virol.* 1995; **69**: 5773–80.

14 Lanciotti RS, Roehrig JT, Deubel V *et al.* Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern United States. *Science* 1999; **286**: 2333–7.

15 Lanciotti RS, Ebel GD, Deubel V *et al.* Complete genome sequences and phylogenetic analysis of West Nile virus strain isolated from the United States, Europe and the Middle East. *Virology* 2002; **298**: 96–105.

16 Scherret JH, Mackenzie JS, Hall RA, Deubel V, Gould EA. Phylogeny and molecular epidemiology of West Nile and Kunjin viruses. *Curr. Top. Microbiol. Immunol.* 2002; **267**: 373–90.

17 Buchmeier MJ, Bowen MD, Peters CJ. *Arenaviridae*: The viruses and their replication. In: *Fields Virology* (eds Knipe, DM, Howley, PM). Lippincott, Williams & Wilkins, Philadelphia 2001; 1635–68.

18 Sanchez A, Khan AS, Zaki SR, Nabel GJ, Ksiazek TG, Filoviridae. Marburg and Ebola viruses. In: *Fields Virology* (eds Knipe, DM, Howley, PM). Lippincott, Williams & Wilkins, Philadelphia 2001; 1279–304.

19 Mattos CA, Mattos CC, Rupprecht CE. Rhabdoviruses. In: *Fields Virology* (eds Knipe, DM, Howley, PM), Lippincott, Williams & Wilkins, Philadelphia 2001; 1245–77.

20 Holzer MP, Salomon KD. Bobcat bite injury of the eye and ocular adnexa. *Arch. Oph.* 2003; **121**: 918–9.

21 Mannen K. Rhabdovirus – The possibility of rabies reemergence in rabies free. *Japan-Uirusu* 2002; **52**: 21–5.

22 Nishizono A, Mannen K, Elio-Villa LP *et al.* Genetic analysis of rabies virus isolates in the Philippines. *Microbiol. Immunol.* 2002; **46**: 413–7.

23 Nichol ST. Bunyaviruses. In: *Fields Virology* (eds Knipe, DM, Howley, PM), Lippincott, Williams & Wilkins, Philadelphia 2001; 1603–33.

24 Plyusnin A. Genetics of hantaviruses: implications to taxonomy. *Arch. Virol.* 2002; **147**: 665–82.

25 Arikawa J. Hantavirus. *Nippon Rinsho* 2003; **61**: 533–8 (in Japanese).

26 Chua KB, Bellini WJ, Rota PA *et al.* Nipah virus: a recent emergent deadly paramyxovirus. *Science* 2000; **288**: 1432–5.

27 Wang L, Harcourt BH, Yu M *et al.* Biology of Hendra and Nipah viruses. *Microbes Infect.* 2001; **3**: 279–87.

28 Kai C. Nipah virus infection. *Nippon Rinsho* 2003; **61** (Suppl. 2): 292–5.

29 Franke J, Essbauer S, Ahne W, Blahak S. Identification and molecular characterization of 18 paramyxoviruses isolated from snakes. *Virus Res.* 2001; **80**: 67–74.

30 van den Hoogen BG, de Jong JC, Groen J *et al.* A Newly Discovered Human Pneumovirus Isolated from Young Children with Respiratory Tract Disease. *Nat. Med.* 2001, 719–24.

31 Collins PL, Chanock RM, Murphy BR. Respiratory syncytial virus. In: *Fields Virology* (eds Knipe, DM, Howley, PM). Lippincott, Williams & Wilkins, Philadelphia 2001; 1443–85.

32 Kahn JS. Human metapneumovirus: a newly emerging respiratory pathogen. *Curr. Opin. Infect. Dis.* 2003; **16**: 255–8.

33 Suzuki A, Watanabe O, Nishimura H. Detection of human metapneumovirus from wheezing children in Japan. *Kansenshogaku Zasshi* 2003; 77: 467–8.

34 Rota P, Oberste S, Monroe S *et al.* Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science 2003*; **300**: 1394–9.

35 Ksiazek T, Erdman D, Cynthia S *et al.* A novel coronavirus associated with severe acute respiratory syndrome. *N. Engl. J. Med.* 2003; **348**: 1953–65.

36 Pearson H, Clarke T, Abbott A, Knight J, Cyranoski D. SARS-What have we learned? *Nature* 2003; **424**: 121–6.

JOURNAL OF CLINICAL MICROBIOLOGY, Mar. 2004, p. 1305–1307 Vol. 42, No. 3
0095-1137/04/$08.00+0 DOI: 10.1128/JCM.42.3.1305–1307.2004

Genetic Diversity of Norovirus and Sapovirus in Hospitalized Infants with Sporadic Cases of Acute Gastroenteritis in Chiang Mai, Thailand

Grant S. Hansman,[1,2] Kazuhiko Katayama,[2] Niwat Maneekarn,[3] Supatra Peerakome,[3] Pattara Khamrin,[3] Supin Tonusin,[4] Shoko Okitsu,[1] Osamu Nishio,[5] Naokazu Takeda,[2] and Hiroshi Ushijima[1]*

Department of Developmental Medical Sciences, Graduate School of Medicine, University of Tokyo,[1] and Department of Virology II[2] and Infectious Disease Surveillance Center,[5] National Institute of Infectious Diseases, Tokyo, Japan, and Department of Microbiology, Faculty of Medicine, Chiang Mai University,[3] and Department of Pediatrics, McCormic Hospital,[4] Chiang Mai, Thailand

Received 14 March 2003/Returned for modification 14 July 2003/Accepted 24 November 2003

Stool specimens from hospitalized infants with sporadic gastroenteritis in Chiang Mai, Thailand, between July 2000 and July 2001 were examined for norovirus and sapovirus by reverse transcription-PCR and sequence analysis. These viruses were identified in 13 of 105 (12%) specimens. One strain was found to be a recombinant norovirus.

Norovirus and sapovirus are two of the four genera of the family *Caliciviridae* and are well-characterized agents of human acute gastroenteritis (6–9). These viruses can be transmitted by a variety of routes, including food (15) and water (16). Three genogroups (GI, GII, and GIII) for norovirus and four genogroups (GI, GII, GIII, and GIV) for sapovirus are thought to exist, though only norovirus GI and GII and sapovirus GI, GII, and GIV are known to infect humans (11, 18). Numerous molecular epidemiological studies have revealed a global distribution of these viruses (2–4, 17, 19). However, very few molecular epidemiological studies have been conducted in Asian countries other than Japan. In this study, we detected norovirus and sapovirus in stool specimens from hospitalized infants with gastroenteritis in Thailand and partially sequenced the capsid gene to determine genogroups and genotypes.

One hundred five stool specimens collected from hospitalized infants (ranging from 1 month to 5 years of age) with acute sporadic gastroenteritis in Chiang Mai, Thailand, between July 2000 and July 2001 were examined for norovirus and sapovirus by reverse transcription-PCR. This included 52 specimens from McCormic Hospital, 21 specimens from Chiang Mai University Hospital, 23 specimens from Nakornping Hospital, and nine specimens from Sanpatong Hospital. RNA was extracted with the QIAamp viral RNA minivacuum protocol (Qiagen) according to the manufacturer's instructions. Reverse transcription was carried out in a final volume of 20 μl with 10 μl of RNA in 50 pmol of random hexamer (Takara), 1× Superscript II reverse transcription buffer (Invitrogen), 10 mM dithiothreitol (Invitrogen), 0.4 mM each of the four deoxynucleoside triphosphates (Roche), 1 U of RNase inhibitor (Toyobo), and 10 U of Superscript II reverse transcriptase (Invitrogen). Reverse transcription was performed at 42°C for 1 h, followed by deactivation of reverse transcriptase at 72°C for 15 min.

* Corresponding author. Mailing address: Department of Developmental Medical Sciences, Graduate School of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. Phone: (813) 5841-3590. Fax: (813) 5841-3629. E-mail: ushijima@m.u-tokyo.ac.jp.

The norovirus PCR primers were selected from three reports that described detection of a broad range of strains (10, 13, 14). For norovirus GI we used primers COG1F (sense) and G1SKR (antisense). For norovirus GII we used primers G2F3 (sense) and G2SKR (antisense). For sapovirus, we used novel capsid gene region primers (corresponding to nucleotides 5083 to 5516 of Manchester virus; GenBank accession number X86560), the SV5317 primer (sense; 5′-CTC GCC ACC TAC RAW GCB TGG TT-3′), and the SV5749 primer (antisense; 5′-CGG RCY TCA AAV STA CCB CCC CA-3′ [where R is A or G; W is A or T; S is C or G; Y is C or T; V is A, C, or G; and B is C, G, or T]). PCR was carried out with 5 μl of cDNA in a PCR mixture containing 33 pmol of each primer, 1× *Taq* DNA polymerase buffer B (Promega), 0.2 mM each of the four deoxynucleoside triphosphates, 2.5 U of *Taq* polymerase (Promega), and up to 50 μl of distilled water. PCR was performed at 94°C for 5 min, followed by 35 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 60 s, and a final extension at 72°C for 5 min. Reverse transcription-PCR products were sequenced and used for genetic classification. Partial and complete nucleotide sequencing and phylogenetic analysis were performed as previously described (11). The nucleotide sequences determined in this study have been deposited in GenBank under accession numbers AY237410 to AY237423.

Norovirus and sapovirus were detected in 12% (13 of 105) of stool specimens from infants admitted to three of the four hospitals in the Chiang Mai region. The age at infection ranged from 4 months to 5 years. All but one of the infants infected with sapovirus were 12 months of age or younger, the youngest infant being 4 months of age. Norovirus also mostly infected infants 12 months of age or younger. One infant was infected with both norovirus and sapovirus strains. Figure 1a shows the phylogenetic tree of the nine norovirus capsid sequences isolated together with reference sequences. The Thai sequences belonged to three distinct norovirus GI genotypes and three norovirus GII genotypes. One norovirus GII sequence (isolate Mc37) that did not cluster with any of the published genotypes was characterized further by complete genome sequencing.

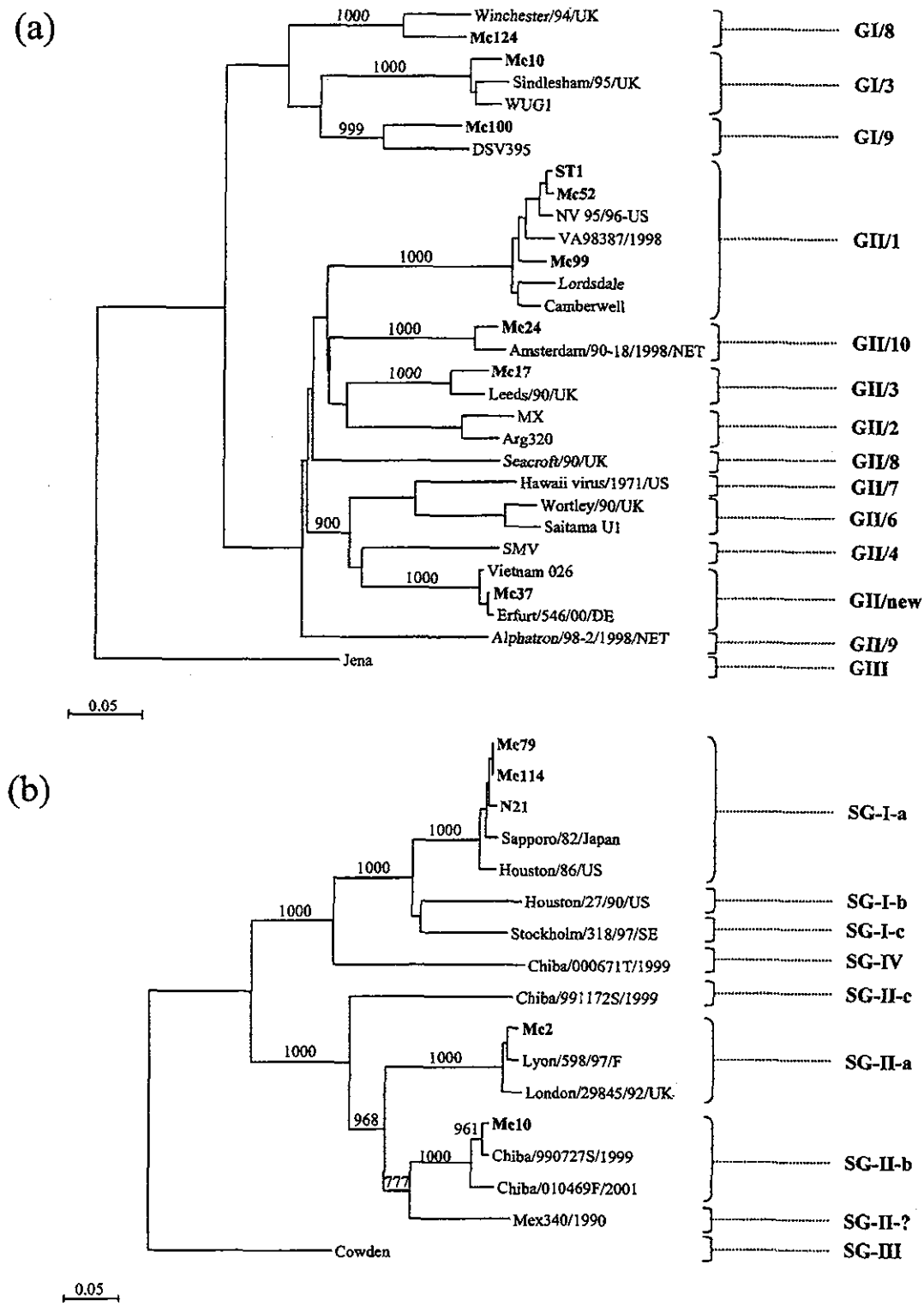

FIG. 1. Phylogenetic analysis of sequences isolated in Thailand. (a) Norovirus capsid sequences (264 bp). (b) Sapovirus capsid sequences (376 bp). The numbers on each branch indicate the bootstrap values for the genotype. *Thai sequences are represented in bold. For example, Mc2 is* the strain isolated at McCormic Hospital from patient 2. Mc, McCormic Hospital; St, Sanpatong Hospital; N, Nakornping Hospital. Norovirus sequences were classified according to the scheme of Katayama et al. (11), and sapovirus sequences were classified based on the scheme of Okada et al. (18). GenBank accession numbers for the reference strains are as follows: Vietnam 026, AF504671; Alphatron/98-2/1998/NET, AF195847; Amsterdam/98-18/1998/NET, AF195848; Arg320, AF190817; Camberwell, U46500; Chiba/000671T/1999, AJ412805; Chiba/990727S/1999, AJ412795; Chiba/991172S/1999, AJ412797; Chiba/010469F/2001, AJ412820; Cowden, AF182760; DSV395, U04469; Erfurt/546/00/DE, AF427118; Hawaii virus/1971/US, U07611; Houston/86/US, U95643; Houston/27/90/US, U95644; Jena, AJ011099; Leeds/90/UK, AJ277608; London/29845/92/UK, U95645; Lordsdale, X86557; Lyon/598/97/F, AJ271056; Mex340/1990, AF435812; MX, U22498; 408/97003012/1996/FL (NV 95/96-US), AF080558; Saitama U1, AB039775; Sapporo/82/Japan, U65427; Seacroft/90/UK, AJ277620; Sindlesham/95/UK, AJ277615; SMV, AY134748; Stockholm/318/97/SE, AF194182; VA98387/1998, AY038600; Winchester/94/UK, AJ277609; Wortley/90/UK, AJ277618; and WUG1, AB081723.

The genome of Mc37 comprised 7,541 nucleotides, excluding the poly(A) tail, and contained three open reading frames (ORFs). The ORF1 sequence showed 97.2% nucleotide identity to that of Saitama U1 virus (AB039775) but only 71.3 and 67.9% nucleotide identity in ORF2 and ORF3, respectively. Consequently, strain Mc37 likely represents a novel recombinant norovirus.

Figure 1b shows the phylogenetic tree of the five sapovirus sequences isolated together with reference sequences. The sapovirus primers detected both GI and GII sapovirus sequences. Three of the five sapovirus sequences belonged to one sapovirus GI cluster, SG-I-a. The two other sapovirus sequences, sapovirus isolates Mc2 and Mc10, belonged to two distinct sapovirus GII clusters, SG-II-a and SG-II-b, respectively. The sapovirus Mc2 sequence showed 78.5% nucleotide identity to the sapovirus Mc10 sequence. The sequences with the closest matches to the sapovirus Mc10 sequence were from two strains isolated in Japan, Chiba/010469F/2001 virus (AJ412820) and Chiba/990727S/1999 virus (AJ412795), showing 95 and 97% nucleotide identity, respectively. The next closest sequence in the GenBank database (Mex340/1990, AF435812) showed only 82% nucleotide identity.

Our results are consistent with those from similar studies. In a report from Spain, 14.19% of stool specimens were positive for norovirus and sapovirus (1), and the majority of strains belonged to norovirus GII (10.65%), followed by norovirus GI (2.26%) and sapovirus (1.29%). Also, in an Australian report, the overall annual minimum incidence rate in hospitalized children was 8.5% for norovirus and 0.6% for sapovirus infection (12). The majority of norovirus strains detected in this Australian report and another from Ireland (5) were of the Lordsdale virus cluster (GII/1). Our study identified several norovirus sequences in this cluster that closely matched a norovirus 95/96-US strain sequence. Recently, several reports have highlighted the importance of the 95/96-US strain's having global distribution and causing a significant number of outbreaks of gastroenteritis (2, 5, 17, 19). In conclusion, these data have described great genetic diversity among both norovirus and sapovirus strains cocirculating in the Chiang Mai region of Thailand and increased the evidence for the worldwide distribution of these viruses.

This work was supported by grants-in-aid from the Ministry of Education and the Ministry of Health and Welfare, Japan. We are grateful to the Japanese Monbusho for the Ph.D. scholarship provided to Grant Hansman.

We thank Peter White for assistance with the manuscript.

REFERENCES

1. **Buesa, J., B. Collado, P. Lopez-Andujar, R. Abu-Mallouh, J. Rodriguez Diaz, A. Garcia Diaz, J. Prat, S. Guix, T. Llovet, G. Prats, and A. Bosch.** 2002. Molecular epidemiology of caliciviruses causing outbreaks and sporadic cases of acute gastroenteritis in Spain. J. Clin. Microbiol. **40**:2854–2859.
2. **Fankhauser, R. L., J. S. Noel, S. S. Monroe, T. Ando, and R. I. Glass.** 1998. Molecular epidemiology of "Norwalk-like viruses" in outbreaks of gastroenteritis in the United States. J. Infect. Dis. **178**:1571–1578.
3. **Farkas, T., T. Berke, G. Reuter, G. Szucs, D. O. Matson, and X. Jiang.** 2002. Molecular detection and sequence analysis of human caliciviruses from acute gastroenteritis outbreaks in Hungary. J. Med. Virol. **67**:567–573.
4. **Farkas, T., X. Jiang, M. L. Guerrero, W. Zhong, N. Wilton, T. Berke, D. O. Matson, L. K. Pickering, and G. Ruiz-Palacios.** 2000. Prevalence and genetic diversity of human caliciviruses (HuCVs) in Mexican children. J. Med. Virol. **62**:217–223.
5. **Foley, B., J. O'Mahony, C. Hill, and J. G. Morgan.** 2001. Molecular detection and sequencing of "Norwalk-like viruses" in outbreaks and sporadic cases of gastroenteritis in Ireland. J. Med. Virol. **65**:388–394.
6. **Green, K. Y.** 1997. The role of human caliciviruses in epidemic gastroenteritis. Arch. Virol. Suppl. **13**:153–165.
7. **Green, K. Y., T. Ando, M. S. Balayan, T. Berke, I. N. Clarke, M. K. Estes, D. O. Matson, S. Nakata, J. D. Neill, M. J. Studdert, and H. J. Thiel.** 2000. Taxonomy of the caliciviruses. J. Infect. Dis. **181**(Suppl. 2):S322–S330.
8. **Jiang, X., C. Espul, W. M. Zhong, H. Cuello, and D. O. Matson.** 1999. Characterization of a novel human calicivirus that may be a naturally occurring recombinant. Arch. Virol. **144**:2377–2387.
9. **Jiang, X., J. Wang, and M. K. Estes.** 1995. Characterization of SRSVs using RT-PCR and a new antigen ELISA. Arch. Virol. **140**:363–374.
10. **Kageyama, T., S. Kojima, M. Shinohara, K. Uchida, S. Fukushi, F. B. Hoshino, N. Takeda, and K. Katayama.** 2003. Broadly reactive and highly sensitive assay for Norwalk-like viruses based on real-time quantitative reverse transcription-PCR. J. Clin. Microbiol. **41**:1548–1557.
11. **Katayama, K., H. Shirato-Horikoshi, S. Kojima, T. Kageyama, T. Oka, F. Hoshino, S. Fukushi, M. Shinohara, K. Uchida, Y. Suzuki, T. Gojobori, and N. Takeda.** 2002. Phylogenetic analysis of the complete genome of 18 Norwalk-like viruses. Virology **299**:225–239.
12. **Kirkwood, C. D., and R. F. Bishop.** 2001. Molecular detection of human calicivirus in young children hospitalized with acute gastroenteritis in Melbourne, Australia, during 1999. J. Clin. Microbiol. **39**:2722–2724.
13. **Kobayashi, S., K. Sakae, Y. Suzuki, H. Ishiko, K. Kamada, K. Suzuki, K. Natori, T. Miyamura, and N. Takeda.** 2000. Expression of recombinant capsid proteins of chitta virus, a genogroup II Norwalk virus, and development of an ELISA to detect the viral antigen. Microbiol. Immunol. **44**:687–693.
14. **Kojima, S., T. Kageyama, S. Fukushi, F. B. Hoshino, M. Shinohara, K. Uchida, K. Natori, N. Takeda, and K. Katayama.** 2002. Genogroup-specific PCR primers for detection of Norwalk-like viruses. J. Virol. Methods **100**: 107–114.
15. **Koopmans, M., J. Vinj, M. de Wit, I. Leenen, W. van der Poel, and Y. van Duynhoven.** 2000. Molecular epidemiology of human enteric caliciviruses in The Netherlands. J. Infect. Dis. **181**(Suppl. 2):S262–S269.
16. **Lawson, H. W., M. M. Braun, R. I. Glass, S. E. Stine, S. S. Monroe, H. K. Atrash, L. E. Lee, and S. J. Englender.** 1991. Waterborne outbreak of Norwalk virus gastroenteritis at a southwest US resort: role of geological formations in contamination of well water. Lancet **337**:1200–1204.
17. **Noel, J. S., R. L. Fankhauser, T. Ando, S. S. Monroe, and R. I. Glass.** 1999. Identification of a distinct common strain of "Norwalk-like viruses" having a global distribution. J. Infect. Dis. **179**:1334–1344.
18. **Okada, M., K. Shinozaki, T. Ogawa, and I. Kaiho.** 2002. Molecular epidemiology and phylogenetic analysis of Sapporo-like viruses. Arch. Virol. **147**: 1445–1451.
19. **White, P. A., G. S. Hansman, A. Li, J. Dable, M. Isaacs, M. Ferson, C. J. McIver, and W. D. Rawlinson.** 2002. Norwalk-like virus 95/96-US strain is a major cause of gastroenteritis outbreaks in Australia. J. Med. Virol. **68**:113–118.

SHORT REPORT

Characterisation of hexon and fibre genes of a novel strain of adenovirus involved in epidemic keratoconjunctivitis

A K Adhikary, T Inada, J Numaga, E Suzuki, H Ushijima, U Banik, A Mukouyama, S Matsuno, N Okabe

J Clin Pathol 2004;57:95–97

Aims: To characterise a novel strain (M86) of adenovirus (Ad) involved in epidemic keratoconjunctivitis (EKC).
Methods/Results: The virus strain was neutralised by antisera to both Ad35 and Ad11. Restriction endonuclease analysis of genomic DNA showed 98% and 88% homology with Ad11 and Ad35, respectively. The deduced amino acid sequence of the hypervariable regions of (HVRs) of the hexon gene showed a higher homology with Ad35 (94.4%) than with Ad11 (83.7%). However, it was 100% homologous to Ad35 in HVRs 1, 2, 3, and 6 and to Ad11 in HVRs 4 and 6. In the fibre knob, the isolate was more homologous to Ad11 (99.4%) than to Ad35 (29.1%).
Conclusion: This novel strain of adenovirus showed similarities with both Ad11 and Ad35. The isolation of a novel strain like Ad35+11 is important because of its association with EKC.

Adenovirus type 11 (Ad11) and adenovirus type 35 (Ad35) belong to subgenus B2, and cause opportunistic infections mainly among the immunocompromised patients.[1] Ad35 was isolated for the first time from a renal transplant recipient with interstitial pneumonia,[2] whereas Ad11 was isolated from a faecal specimen of a child with poliomyelitis.[3] Here, we report the isolation of a novel strain of adenovirus from a 25 year old otherwise healthy male patient with severe clinical manifestations of epidemic keratoconjunctivitis (EKC) in southern Japan.

METHODS AND RESULTS

Immunochromatography confirmed the causative agent as adenovirus (strain M86).[4] Conjunctival scrapings were isolated in A549 cells and the viral titre was also determined in a microtitre plate containing a confluent monolayer of A549 cells. Aliquots (25 µl) of 100 tissue culture infectious doses of virus ($100TCID_{50}$) were incubated with 25 µl of serially diluted type specific antisera at 37°C for 60 minutes and then inoculated into A549 cells. Viral growth was inhibited by two different type specific antisera, anti-Ad11 and anti-Ad35, at a 256-fold dilution, and the strain was identified as Ad35+11. Although Ad11 infrequently causes keratoconjunctivitis, Ad35 or a novel strain like M86 (Ad35+11) has never been reported as an ocular pathogen.[5] Therefore, this strain was subjected to a detailed study at the molecular level.

Viral DNA extraction and restriction endonuclease analysis of M86 with BamHI, BglII, BstEII, EcoRI, HindIII, PstI, SalI, SmaI, XbaI, and XhoI (Boehringer Mannheim, Mannheim, Germany) were carried out to investigate homology with the serologically related prototypes (Ad35 and Ad11), as described previously.[6] Genomic homology between M86, Ad35, and Ad11 was calculated from published restriction patterns of Ad11, Ad35, and M86, using the percentage of pair wise co-migrating restriction fragments of a pair divided by the total number of bands in the pair. The isolate (M86) showed 98% and 88% homology with Ad11 and Ad35, respectively (fig 1). Higher homology of the new strain in restriction endonuclease analysis with Ad35 and Ad11 provide supportive evidence that the new strain might have evolved from the recombination of these two parent viruses.[7]

The fibre knob enables the virus to attach to the cellular receptor and, together with the hexon protein, defines the serological specificity of the adenoviruses. Therefore, the hexon gene and the fibre gene were analysed to compare the immunological data with the molecular biological results, in addition to looking for any possible variation that might be related to ocular pathogenicity. The hypervariable regions

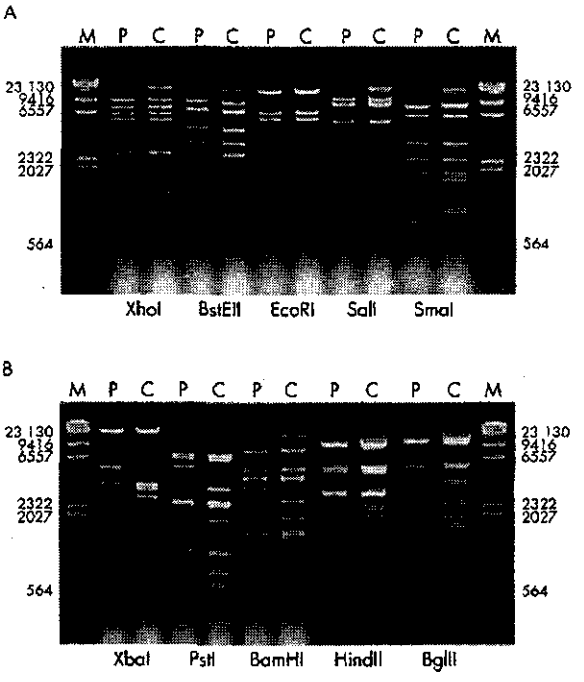

Figure 1 Restriction patterns obtained after cleavage of Ad35p (P) and M86 (C). (A) Restriction endonucleases XhoI, BstEII, EcoRI, SalI, and SmaI; (B) XbaI, PstI, BamHI, HindIII, and BglII. The samples were electrophoresed on a 1.2% agarose gel. A HindIII digest of λ DNA (lane M) was run as a molecular weight standard.

Abbreviations: AA, amino acid; Ad, adenovirus; EKC, epidemic keratoconjunctivitis; HVR, hypervariable region

Table 1 Hexon gene amino acid homologies (%) of M86 with other members of subgenus B and subgenus D

Ad	HVR1	HVR2	HVR3	HVR4	HVR5	HVR6	HVR7	Overall
Ad3	11.7	36.8	33.3	28.5	12.5	36.3	20.0	56.5
Ad7	11.7	41.1	33.3	38.0	31.2	50.0	16.2	61.2
Ad11	45.4	50.0	86.6	100.0	87.5	100.0	85.0	86.7
Ad34	48.9	50.0	100.0	28.5	62.5	50.0	15.3	70.7
Ad35	100.0	100.0	100.0	42.8	87.5	100.0	95.0	94.4
Ad8	12.7	9.5	12.1	42.2	12.5	41.6	10.0	51.3
Ad19	10.8	22.2	18.1	28.5	25.0	66.6	10.0	53.2
Ad37	10.8	41.1	18.1	25.0	10.6	66.6	10.0	52.7

Per cent of homology between the indicated hypervariable regions (HVRs) of M86 and the HVRs of adenovirus (Ad) subgenus B and subgenus D.

(HVRs) of the hexon gene and all regions of the fibre gene were sequenced by overlapping primers from genomic DNA by direct cycle sequencing. Multiple sets of primers for the hexon and fibre genes were selected based on alignment of the hexon gene (x76549 (Ad3), x76551 (Ad7), AB018424 (Ad11), AB018425 (Ad14), x74662 (Ad16), AY008279 (Ad21), AB018246 (Ad34), and AB018427 (Ad35)) and fibre gene sequences (m12411 (Ad3), m23696 (Ad7), L08231 (Ad11), AB065116 (Ad14), u06106 (Ad16), u06107 (Ad21), u10271 (Ad34), and u10272 (Ad35)) available from GeneBank. The sequences were determined by a Genetic Analyser 310 (Applied Biosystem, Foster City, California, USA). DNASIS software (Hitachi Software Ltd, Tokyo, Japan) was used for sequence alignment and analysis. The amino acid (AA) sequences of these residues were deduced. The AA sequences of M86 were compared with the available sequences of Ad3, Ad7, Ad11, Ad34, Ad35, Ad8, Ad19a, and Ad37 involved in keratoconjunctivitis. The nucleotide sequence data reported in our paper will appear in the DDBJ/GeneBank nucleotide sequence database with the accession numbers AB098564 (hexon gene) and AB098563 (fibre gene).

Table 2 Fibre gene nucleotide and amino acid homologies (%) of M86 with other members of adenovirus (Ad) subgenus B and subgenus D

	Ad11	Ad34	Ad35	Ad3	Ad7	Ad8	Ad19/37
Tail							
DNA	100.0	99.2	99.2	80.6	97.6	48.8	48.0
Protein	100.0	97.6	97.6	86.0	97.6	37.2	30.0
Shaft							
DNA	98.1	71.7	71.7	55.2	90.1	32.9	31.8
Protein	98.9	63.7	63.7	52.2	91.2	16.4	15.3
Knob							
DNA	99.8	45.0	45.4	50.6	93.7	27.9	32.2
Protein	99.4	28.9	29.1	28.8	91.6	8.3	10.6
Overall							
DNA	99.3	59.7	59.9	56.3	93.2	30.4	34.1
Protein	99.3	47.8	47.8	41.8	92.3	12.6	12.6

Per cent of homology between different fibre regions of M86 and those of adenovirus (Ad) subgenus B and subgenus D.

In HVRs of the hexon, M86 showed an overall 94.4% AA homology with Ad35 and 86.7% with Ad11 (table 1). However, it was 100% homologous to Ad35 in HVRs 1, 2, 3, and 6 and to Ad11 in HVRs 4 and 6 (fig 2). The fibre knob showed high AA homology (99.4%) with Ad11, but only 29.1% homology with Ad35 (table 2).

DISCUSSION

Members of subgenus D adenoviruses (Ad8, Ad19, and Ad37) are the common agent of EKC. Occasionally, the members of subgenus B (Ad3 and Ad7) and subgenus E (Ad4) are also related to EKC. The tropism of adenoviruses for conjunctival or corneal cells depends on the presence of certain amino acids in the knob, which attaches the virus to the specific cellular receptor.[5] The fact that the fibre knob of M86 has 99.4% homology with Ad11, with only a single AA difference, but only 29.1% homology with Ad35, means that it is able to attach to conjunctival and corneal cells (fig 3).

```
                          L1<                                  HVR 1
Ad35     1   SGTAYNSLAP KGAPNASQWI AKGVPTAAAA GNGEEEHETE EKTATYTFAN  50
M86      1   SGTAYNSLAP KGAPNASQWI AKGVPTAAAA GNGEEEHETE EKTATYTFAN  50
Ad11     1   SGTAYNSLAP KGAPNTSQWI AEGVKNTTGE EHVTEE---E TNTTTYTFGN  50
                        >    <      HVR 2     >        < HVR 3
Ad35    51   APVKAEAQIT KEGLPIGLEI SAENESKPIY ADKLYQPEPQ VGDETWTDLD 100
M86     51   APVKAEAQIT KEGLPIGLEI SAENESKPIY ADKLYQPEPQ VGDETWTDLD 100
Ad11    51   APVKAEAEIT KEGLPVGLEV S-DEESKPIY ADKTYQPEPQ LGDETWTDLD 100
               >                          <    HVR 4          >
Ad35   101   GKTEEYGGRA LKPTTNMKPC YGSYAKPTNL KGGQAKPKNS EPSSEKIEYD 150
M86    101   GKTEEYGGRA LKPDTKMKPC YGSFAKPTNV KGGQAKQKTT EQPNQKVEYD 150
Ad11   101   GKTEKYGGRA LKPDTSMKPC YGSFAKPTNV KGGQAKQKTT EQPNQKVEYD 150
                    <    HVR 5     >          <     HVR 6
Ad35   151   IDMEFFDNSS QRTNFSPKIV MYAENVGLET PDTHVVYKPG TEDTSSEANL 200
M86    151   IDMEFFDAAS QRTNFSPKIV MYAENVGLET PDTHVVYKPG TEDTSSEANL 200
Ad11   151   IDMEFFDAAS QKTNLSPKIV MYAENVNLET PDTHVVYKPG TEDTSSEANL 200
              >
Ad35   201   GQQSMPNRPN YIGFRDNFIG LMYYNSTGNM GVLAGQASQL NAVVDLQDRN 250
M86    201   GQQSMPNRPN YIGFRDNFIG LMYYNSTGNM GVLAGQASQL NAVVDLQDRN 250
Ad11   201   GQQSMPNRPN YIGFRDNFIG LMYYNSTGNM GVLAGQASQL NAVVDLQDRN 250

Ad35   251   TELSYQLLLD SLGDRTRYFS MWNQAVDSYD PDVRVIENHG VEDELPNYCF 300
M86    251   TELSYQLLLD SLGDRTRYFS MWNQAVDSYD PDVRVIENHG VEDELPNYCF 300
Ad11   251   TELSYQLLLD SLGDRTRYFS MWNQAVDSYD PDVRVIENHG VEDELPNYCF 300
              L2<                 HVR 7                  >
Ad35   301   PLDGIGVPTT SYKSIVPNGE DNNNWKEPEV NGTSEIGQGN LFAMEINLQA 350
M86    301   PLNGIGVPTT SYKSIVPNGE DNNNWKEPEV NGTSEIGQGN LSAMEINLQA 350
Ad11   301   PLDGIGVPTT SYKSIVPNGD NAPNWKEPEV NGTSEIGQGN LFAMEINLQA 350

Ad35   351   NLWRSFLY.. .......... .......... .......... .......... 400
M86    351   NLWRSFLY.. .......... .......... .......... .......... 400
Ad11   351   NLWRSFLY.. .......... .......... .......... .......... 400
```

Figure 2 Comparison of predicted amino acid sequences of seven hypervariable regions (HVRs) of M86 with that of adenovirus type 11 (Ad11) and Ad35. The sequences of loop 1 (L1) and loop 2 (L2) were aligned to obtain maximal homology. Deduced amino acid sequences of Ad11 and Ad35 were obtained from GeneBank (accession numbers ABO18424 (Ad11) and ABO18427 (Ad35)).

```
M86     1   WTGVNPTEAN CQIMNSSESN DCKLILTLVK TGALVTAFVY VIGVSNNFNM  50
Ad11    1   WTGVNPTEAN CQIMNSSESN DCKLILTLVK TGALVTAFVY VIGVSNNFNM  50
Ad35    1   WTGINPP-PN CQIVENTNTN DGKLTLVLVK NGGLVNGYVS LVGVSDTVNQ  50

M86    51   LTTHRNINFT AELFFDSTGN LLTRLSSLKT PLNHKSGQNM ATGAITNAKG 100
Ad11   51   LTTHRNINFT AELFFDSTGN LLTRLSSLKT PLNHKSGQNM ATGAITNAKG 100
Ad35   51   MFTQKTANIQ LRLYFDSSGN LLTEESDLKI PLKNKSSTA- TSETVASSKA 100

M86   101   FMPSTTAYPF NDNSREKENY IYGTCYYAAS -DRTAFPIDI SVMLNRRAIN 150
Ad11  101   FMPSTTAYPF NDNSREKENY IYGTCYYTAS -DRTAFPIDI SVMLNRRAIN 150
Ad35  101   FMPSTTAYPF NTTTRDSENY IHGICYYMTS YDRSLFPLNI SIMLNSRMIS 150

M86   151   DETSYCIRIT WSWNTGDAPE VQTSATTLVT SPFTFYYIRE DD..... 200
Ad11  151   DETSYCIRIT WSWNTGDAPE VQTSATTLVT SPFTFYYIRE DD..... 200
Ad35  151   SNVAYAIQFE WNLNASESPE --SNIATLTT SPFFFSYITE DD..... 200
```

Figure 3 Comparison of predicted fibre knob sequences of M86, adenovirus type 11p (Ad11p), and Ad35p. The sequences were aligned to obtain maximal homology. The deletions are represented by dashes. Deduced amino acid sequences of Ad11p and Ad35p were obtained from *GeneBank* (accession numbers L08231 (Ad11p) and U10272 (Ad35p)).

Neutralisation of the infectivity of adenoviruses is primarily carried out by antibodies against the hexon protein. Antigenic determinants (epitopes) located in two or more of the seven HVRs in loop 1 and loop 2 of the hexon react with neutralising antibodies.[9] These HVRs are highly conserved within the serotype.[10] However, the position of the epitope in the HVRs and number of amino acids forming them are yet to be determined. The construction of a chimaera in the HVRs of the hexon could change the antigenic specificity of the virus, enabling it to escape type specific neutralisation.[11] M86 was 100% homologous to Ad35 in HVRs 1, 2, 3, and 6 and to Ad11 in HVRs 4 and 6. This sequence variation reflects the preceding mutation or recombination events involving the HVRs of the hexon, which is expressed by a mixed antigenic character in the neutralisation test. This novel arrangement in the HVRs might enable the virus to circumvent existing immunity.

> "The construction of a chimaera in the hypervariable regions of the hexon could change the antigenic specificity of the virus, enabling it to escape type specific neutralisation"

The isolation of a strain like M86 as a new aetiological agent of EKC is medically and epidemiologically important because it shows that recombination or mutation involving the HVRs of the hexon gene can enable the non-ocular adenoviruses to become ocular pathogens. It is also possible that such strains can circumvent existing immunity and might be responsible for outbreaks of EKC in the future; this may be especially important in developing countries, where the detection of adenoviruses in the clinical setting is not available. It is also important to accumulate data on the HVRs of the hexon gene and fibre gene sequences of the EKC strains to predict their possible role in keratoconjunctivitis.

Take home messages

- We have characterised a novel strain (M86) of adenovirus (Ad) involved in epidemic keratoconjunctivitis
- The virus strain showed homology with both Ad35 and Ad11 using both immunological and molecular biological techniques
- It high homology with the fibre knob sequences of Ad11 may provide it with the ability to invade ocular cells
- The mixed antigenic characteristics of virus strains such as this may enable them to circumvent existing immunity

ACKNOWLEDGEMENTS

Supported by grant-in-aid for Scientific Research from the Japan Society for the Promotion of Science (JSPS), Japan.

Authors' affiliations

A K Adhikary, T Inada, U Banik, A Mukouyama, S Matsuno, N Okabe, Infectious Disease Surveillance Centre, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan
J Numaga, Department of Ophthalmology, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
E Suzuki, H Ushijima, Department of Developmental Medical Sciences, Graduate School of Medicine, The University of Tokyo

Correspondence to: Dr A K Adhikary, Mailing, Infectious Disease Surveillance Centre, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan; arun@nih.go.jp

Accepted for publication 10 July 2003

REFERENCES

1 **Hierholzer** JC. Adenoviruses in the immunocompromised host. *J Clin Microbiol* 1992;5:262–74.
2 **Stalder** H, Hierholzer JC, Oxman MN. New human adenovirus (candidate adenovirus type 35) causing fatal disseminated infection in a renal transplant recipient. *J Clin Microbiol* 1977;6:257–65.
3 **Kibrick S**, Melendez L, Ender JF. Clinical association of enteric viruses with particular reference to agents exhibiting properties of ECHO group. *Ann NY Acad Sci* 1957;67:311–29.
4 **Uchio E**, Aoki K, Saitoh W, *et al*. Rapid diagnosis of adenoviral conjunctivitis on conjunctival swab by 10-minute immunochromatography. *Ophthalmology* 1996;104:1294–9.
5 **Kitamura N**. Genome analysis of adenovirus type 7 and adenovirus type 11. *Jpn J Opthalmol* 2001;45:22–30.
6 **Adhikary AK**, Numaga J, Kaburaki T, *et al*. Genetic characterisation of adenovirus type 8 isolated in Hiroshima city over a 15 year period. *J Clin Pathol* 2003;56:120–5.
7 **Sambrook** J, Sleigh M, Engler JA, *et al*. The evolution of the adenoviral genome. *Ann NY Acad Sci* 1980;354:426–52.
8 **Elz B**, Adrian T, Pring-Akerblom P. Recombinant fiber proteins of human adenoviruses Ad9, Ad15 and Ad19: localization of the haemagglutination properties and the type specific determinant. *Res Virol* 1997;148:5–10.
9 **Crawford-Miksza L**, Schnurr DP. Analysis of adenovirus hexon proteins revealed the location of structure of seven hypervariable regions containing serotype-specific residues. *J Virol* 1996;70:1835–40.
10 **Li QG**, Wadell G. Genetic variability of hexon loops 1 and 2 between seven genome types of adenovirus 7. *Arch Virol* 1999;144:1739–49.
11 **Roy S**, Shirley PS, McClelland A, *et al*. Circumvention of immunity to the adenovirus major coat protein hexon. *J Virol* 1998;72:6875–9.

Journal of General Virology (2004), 85, 283–292 DOI 10.1099/vir.0.19633-0

Genotype C of hepatitis B virus can be classified into at least two subgroups

Tran Thien-Tuan Huy,[1,2,3] Hiroshi Ushijima,[2] Vo Xuan Quang,[3] Khin Maung Win,[4] Pairoj Luengrojanakul,[5] Kaoru Kikuchi,[6] Tetsutaro Sata[1] and Kenji Abe[1]

[1]Department of Pathology, National Institute of Infectious Diseases, Toyama 1-23-1, Shinjuku-ku, Tokyo 162-8640, Japan

[2]Department of Developmental Medical Sciences, Graduate School of Medicine, University of Tokyo, Tokyo, Japan

[3]Department of Gastroentero-Hepatology, Cho Ray Hospital, Ho Chi Minh City, Vietnam

[4]Department of Hepatology, Yangon General Hospital, Yangon, Myanmar

[5]Department of Gastroenterology, Mahidol University Siriraj Hospital, Bangkok, Thailand

[6]Gastroenterology Section, Okinawa Chubu Hospital, Okinawa, Japan

Correspondence
Kenji Abe
kenjiabe@nih.go.jp

Received 8 September 2003
Accepted 30 September 2003

A genomic characterization of hepatitis B virus (HBV) was done for 56 pre-S1/pre-S2 genes and 10 full-length HBV genotype C isolates from five Asian countries. Phylogenetic analysis of the pre-S1/pre-S2 genes revealed two major groups within genotype C: one for isolates from southeast Asia including Vietnam, Myanmar and Thailand (named HBV/C1) and the other for isolates from Far East Asia including Japan, Korea and China (named HBV/C2). This finding was confirmed by phylogenetic analysis based on the full-length sequence of 32 HBV genotype C isolates, including 22 from database entries. Two isolates from Okinawa, the island off the southern end of Japan, formed a different branch. Specific amino acid sequence changes were identified in the large S protein (amino acids 51, 54, 60, 62 and 73) and P protein (amino acids 231, 233, 236, 248, 252 and 304). Our results indicate that genotype C of HBV can be classified into at least two subgroups.

INTRODUCTION

Hepatitis B virus (HBV) infection is a global health problem, with more than 350 million people chronically infected worldwide (Lee, 1997). The infection is associated with a wide spectrum of clinical symptoms, ranging from acute or fulminant hepatitis to various forms of chronic liver disease, such as chronic hepatitis, cirrhosis and hepatocellular carcinoma. HBV has been classified into genotypes A–G, with an intergenotypic diversity of at least 8 % in the full genome sequence (Okamoto *et al.*, 1988; Norder *et al.*, 1994). HBV genotypes have a distinct geographical distribution and correlate with severity of liver disease (Kidd-Ljunggren *et al.*, 2002). Genotypes B and C are prevalent in Asia, and genotype C causes more serious liver disease than genotype B (Shiina *et al.*, 1991; Orito *et al.*, 2001). Moreover, HBV strains in the same genotype may differ in their capacity to induce clinical liver disease. Subgroup Ba of genotype B, which is recombinant with genotype C, is found predominantly in southeast Asian countries and appears to have more detrimental effects than subgroup Bj (Sugauchi *et al.*, 2002a). Recently, a novel genotype C variant has been found in Australian aborigines (Sugauchi *et al.*, 2001). Therefore, it is possible that virological differences in HBV genotype C exist in Asian countries resulting in different clinical outcomes for patients. To analyse this further, we carried out genomic characterization of HBV genotype C isolates and found that they could be classified into at least two subgroups.

METHODS

Source of sera, genotyping and serotyping of HBV. HBV DNA-positive sera were obtained from 56 patients in five different Asian countries: 17 from Vietnam, 21 from Myanmar, four from Thailand, nine from China and five from Japan (three from Tokyo and two from Okinawa). All sera were found to be HBV genotype C by PCR genotyping using type-specific primers as reported previously (Naito *et al.*, 2001). The HBV serotype was inferred from the

The DDBJ/GenBank/EMBL accession nos. of sequences reported in this study are HBV-VT101, AB112063; HBV-VT103, AB111946; HBV-VT140, AB112065; HBV-MY624, AB112066; HBV-MY656, AB112348; HBV-MY683, AB112408; HBV-TH123, AB112471; HBV-TH124, AB112472; HBV-047, AB115417; and HBV-055, AB115418.

sequence of the S gene (Magnius & Norder, 1995). The serum samples were kept at −40 °C or below until used. Informed consent for participation in this study was obtained from each individual.

Extraction of DNA. Viral DNA was extracted from 100 µl serum using a DNA/RNA extraction Kit (SepaGene RV-R; Sanko Junyaku Co.). The resulting pellet was eluted in 50 µl RNase-free water and kept at −20 °C until used.

Amplification of HBV DNA. The region of the HBV genotype C sequence covering the pre-S region (522 bases from the beginning of pre-S1 to the end of pre-S2) was amplified by heminested PCR in all 56 isolates. The full-length nucleotide sequences of 10 HBV isolates belonging to genotype C were also determined. HBV DNA was amplified by PCR with LA *Taq* (TaKaRa Shuzo) or AmpliTaq Gold DNA polymerase (Applied Biosystems). The sequences of oligonucleotide primers and their combinations used in this study are listed in Table 1. To obtain the entire sequence, the first round of PCR was carried out for 40 cycles (98 °C for 10 s, 50 °C for 20 s and 72 °C for 2·5 min) followed by extension at 72 °C for 10 min. The second round was carried out for 35 cycles (94 °C for 1 min, 55 °C for 1 min and 72 °C for 1 min) followed by extension at 72 °C for 7 min in order to amplify the five overlapping fragments that covered the full genome. The PCR products were separated by 1 % agarose gel electrophoresis and purified using a QIA quick gel extraction kit (Qiagen) for sequence analysis.

DNA sequencing. Purified PCR products were subjected to direct sequencing using the ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems). Sequences of amplified DNA were determined using an automated DNA sequencer ABI 377 (Perkin Elmer).

Phylogenetic analysis. The 56 isolates of HBV genotype C from Asian countries recovered in this study were compared with isolates of genotype C (22 strains) and of other genotypes (11 strains) from the databases over the pre-S region. Forty-three isolates with full genome sequences (10 in this study and 33 from database entries) were also analysed. Nucleotide sequences were multiple-aligned using GENETYX for Windows version 6.0 software (Genetyx) and corrected manually by visual inspection. Genetic distances were calculated using the Kimura two-parameter method and phylogenetic trees were constructed by the neighbour-joining method (Saitou & Nei, 1987). To confirm the reliability of the pairwise comparison and phylogenetic tree analysis, bootstrap resampling and reconstruction were carried out 1000 times. Phylogenetic and molecular evolutionary analyses were done using MEGA version 2.1 (Kumar *et al.*, 2001).

RESULTS

Nucleotide sequences and phylogenetic analysis

By phylogenetic analysis of pre-S1/pre-S2 genes from 56 isolates obtained in this study and 33 from database entries, two major groups were found to be clustered within genotype C (Fig. 1a). Specifically, all 42 isolates from southeast Asia, including Thailand, Vietnam and Myanmar, were clustered in one branch and 34 isolates from Far East Asia, including Japan, China and Korea, formed another branch. In addition, there was a branch containing two isolates from Okinawa prefecture (the island off the

Table 1. PCR primers for HBV DNA used in this study

For the pre-S gene, P1/S2-2 were used for the first round of PCR and P1/S4R for the second round. For the complete sequence, five overlapping fragments were obtained by nested or heminested PCR using the following primer combinations for the second round of PCR: HBV-1/BG1R for fragment A, HBc1/PS8R for fragment B, P1/S1-2 for fragment C, S1-1/HBx2 for fragment D and HBx1/HBV-2 for fragment E. The first round of PCR was done with the primer combinations of HBV-1/S1-2 for fragments A to C and S1-1/HBV-2 for fragments D and E, respectively. Underlining of the primer sequence indicates the *Eco*RI site.

Primer	Nucleotide sequence (5′→3′)	Position*	Polarity
For pre-S gene			
P1	TCACCATATTCTTGGGAACAAGA	2817–2839	Sense
S2-2	GGCACTAGTAAACTGAGCCA	687–668	Antisense
S4R	AGAAGATGAGGCATAGCAGC	434–415	Antisense
For complete sequence			
HBV-1	CCGGAAGAATTCTTTTTCACCTCTGCCTAATCA	1821–1841	Sense
S1-2	CGAACCACTGAACAAATGGC	704–685	Antisense
BG1R	ATAGGGGCATTTGGTGGTCT	2316–2297	Antisense
HBc1	AGTGTGGATTCGCACTCCT	2269–2287	Sense
PS8R	ARGCCCTGAGCCTGAGGGCTC†	3098–3078	Antisense
P1	TCACCATATTCTTGGGAACAAGA	2817–2839	Sense
S1-1	TCGTGTTACAGGCGGGGTTT	192–211	Sense
HBV-2	CCGGAAGAATTCAAAAAGTTGCATGGTGCTGG	1825–1806	Antisense
HBx2	ACGTGCAGAGGTGAAGCGAAG	1604–1584	Antisense
HBx1	GTCCCCTTCTTCATCTGCCGT	1487–1507	Sense

*Nucleotide position based on the sequence of HBV genotype C (accession no. D50517).
†Degenerate nucleotide R=A/G.

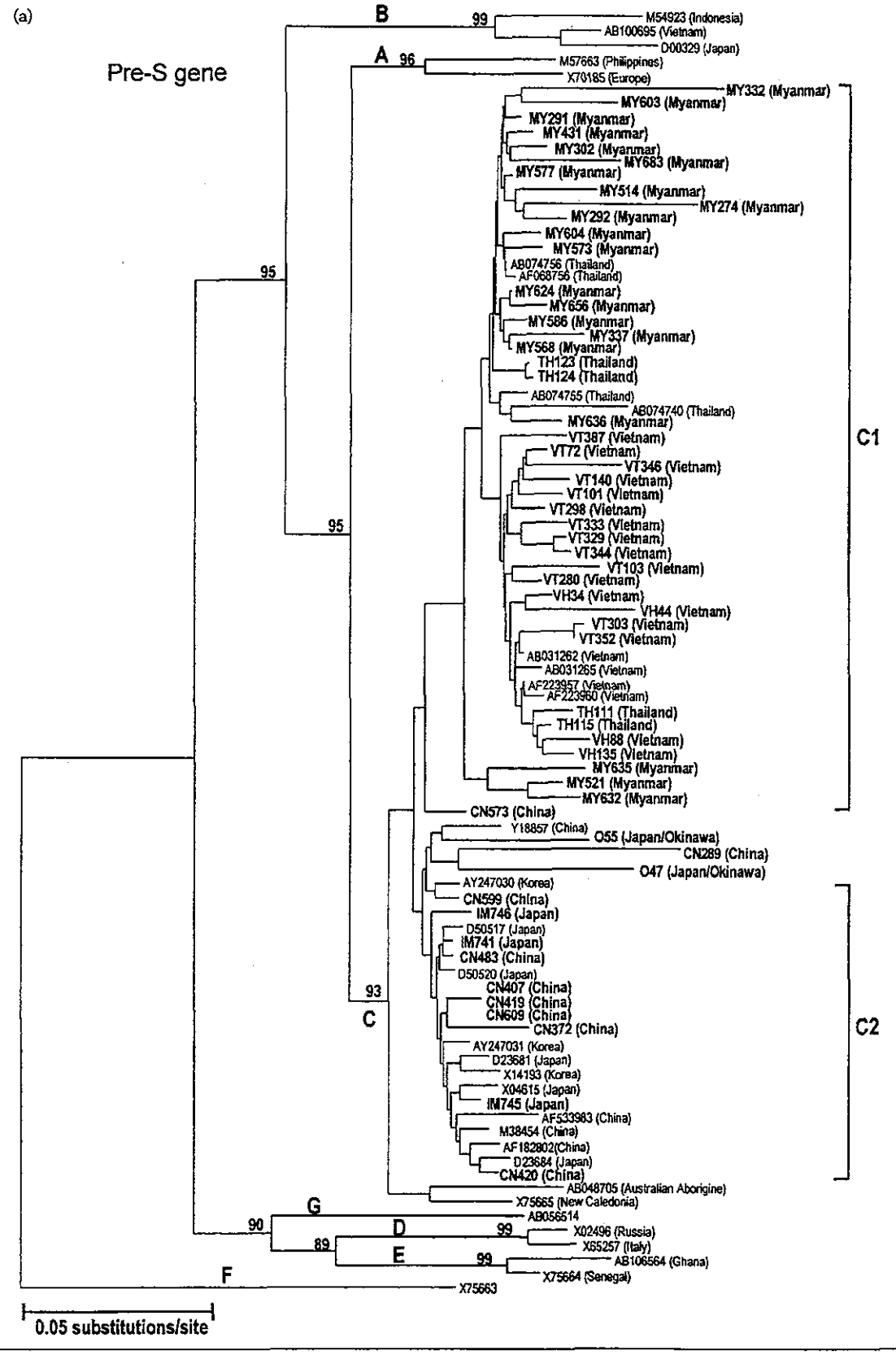

Fig. 1. For legend see page 287.

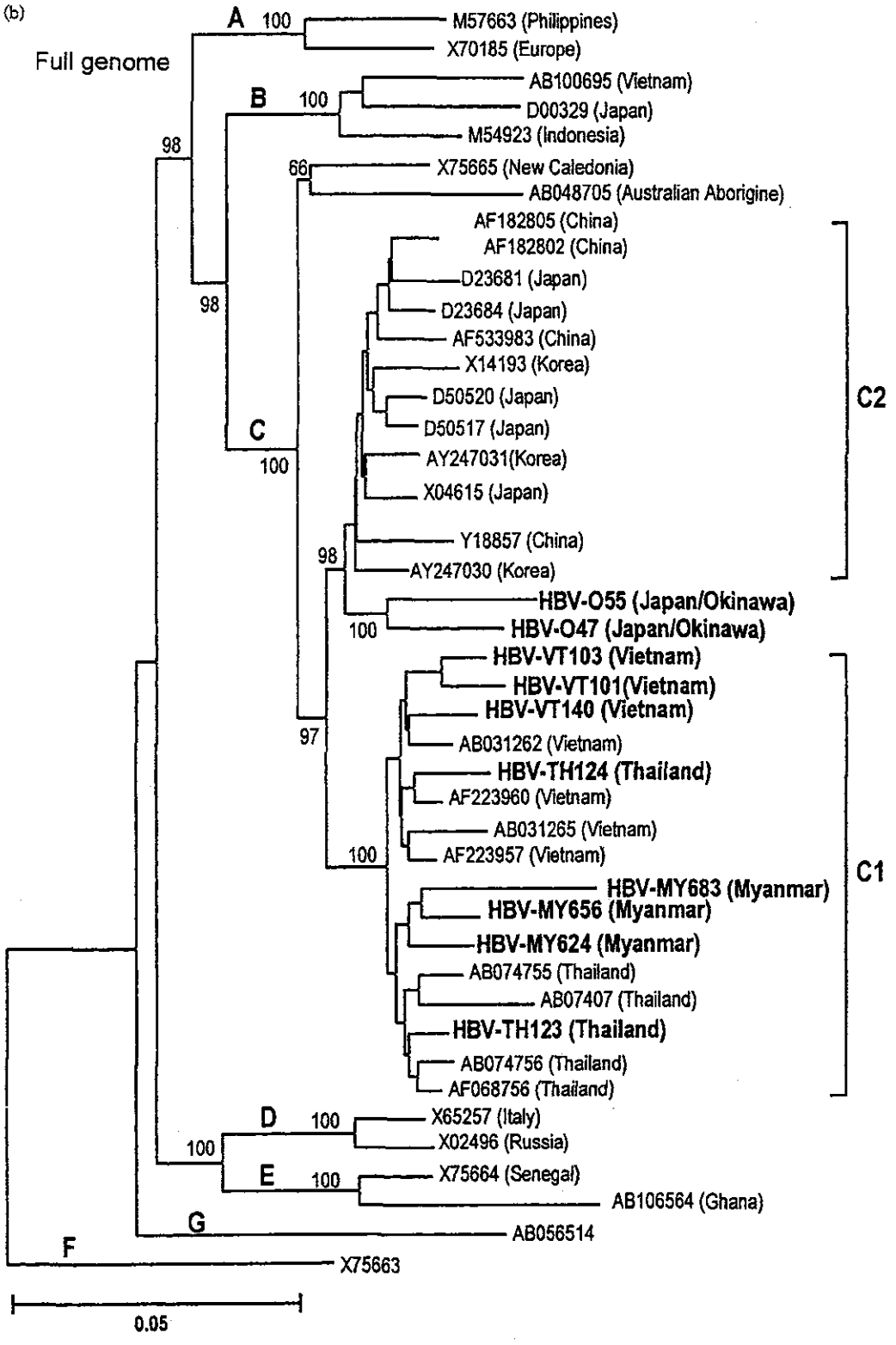

Fig. 1. *For legend see page 287.*

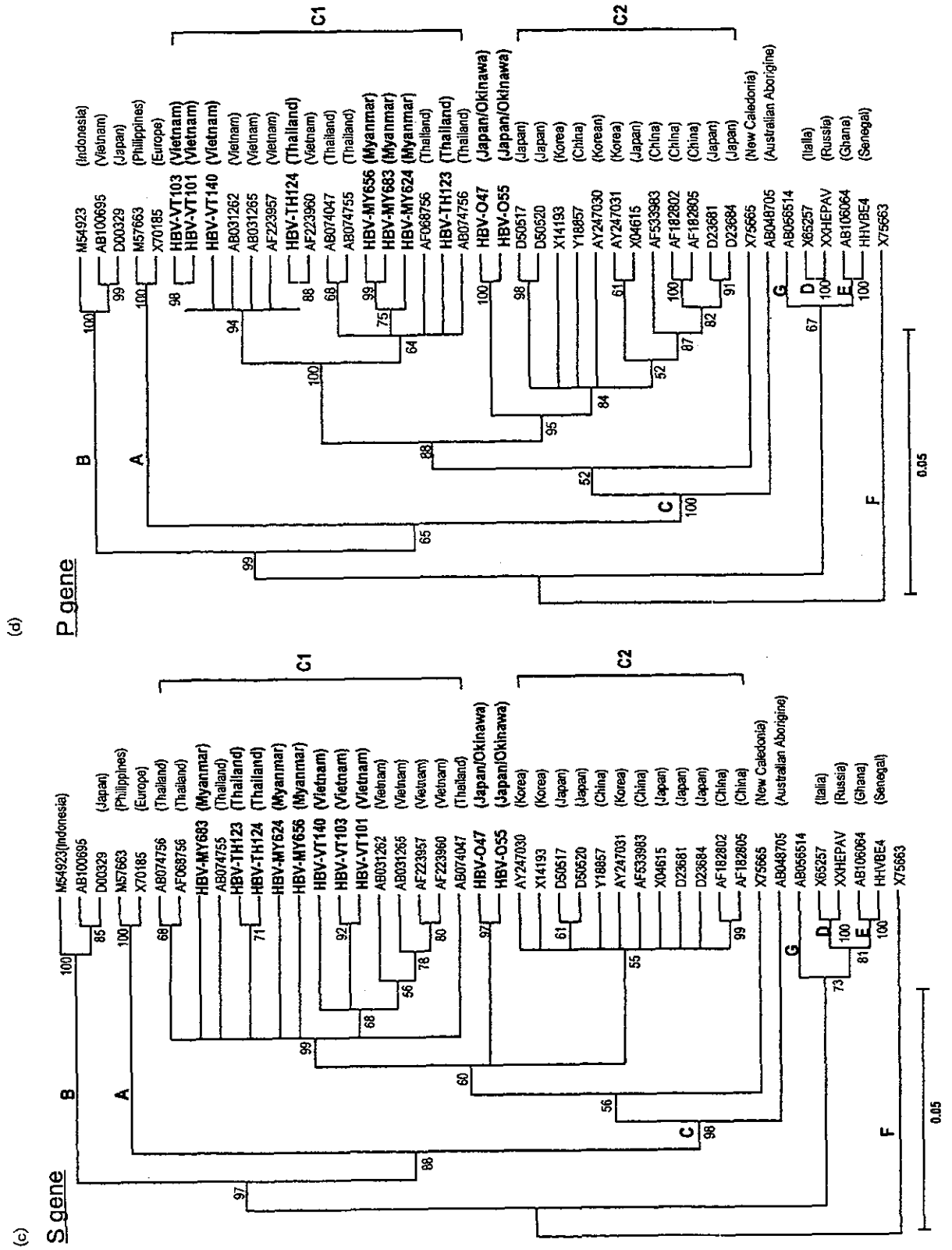

Fig. 1. Phylogenetic tree constructed on the pre-S gene (522 nucleotides) (a), the full genome (b), and the S (c) and P (d) genes, respectively. Bootstrap values are shown at the beginning of each main node. The length of the horizontal bar indicates the number of nucleotide substitutions per site. The origin of each strain is also shown in parentheses.

southern end of Japan) and two from China. Two database-derived isolates found in Australian Aborigines and New Caledonia formed an outgroup of the genotype. Phylogenetic analysis based on full genomic sequences of 10 HBV isolates from Vietnam (3), Myanmar (3), Thailand (2) and Okinawa (2) also confirmed the existence of two major subgroups within genotype C (Fig. 1b). Similar clusters were also identified by analysis of the large S and the P genes (Fig. 1c, d). From these results, we designated the two subgroups within the HBV genotype C subgroup C1 (HBV/C1) for isolates mainly distributed in southeast Asia and subgroup C2 (HBV/C2) for isolates in Far East Asia.

Nucleotide divergence between HBV/C1, HBV/C2 and other genotypes

To define the magnitude of intergenotypic and intragenotypic differences, pairwise analysis of nucleotide comparisons was performed over the complete HBV genome and the large S gene of 45 HBV strains representing all genotypes. Over the complete genome from genotype A to G, the mean percentages of nucleotide divergence between HBV/C1 and other genotypes were always higher than 8%: 9·5% (genotype A), 9·3% (genotype B), 10·7% (genotype D), 14·5% (genotype E), 14·1% (genotype F) and 12·8% (genotype G) (Table 2). In contrast, the nucleotide difference between HBV/C1 and HBV/C2 was 5·2% and the intragenotypic difference in each group was 3·1% and 3·9%, respectively. The above results, together with the nucleotide divergences in the large S gene, support the concept that HBV/C1 and HBV/C2 belong to genotype C and that they can be considered as separate subgroups within this genotype.

Amino acid changes in HBV/C1 and HBV/C2

By comparison of the deduced amino acid sequences among the complete HBV genomes, we identified a number of amino acid differences between HBV/C1 and HBV/C2 in the large S gene and P gene, but not in the X and core genes (Table 3). HBV/C1 isolates had the following consensus sequence of amino acids: 51Q, 54A, 60V, 62S, 73S/N, 125S and 227L in the large S gene and 136T/I, 142Q, 231M, 233S, 236P, 248S, 252Q, 304Q, 354Y and 358H in the P gene. Interestingly, two HBV isolates from Okinawa had almost identical amino acid sequences to those of HBV/C1, except for 51Q and 231M, which matched those of HBV/C2.

Serotypes and the nucleotide change at nt 1858 in HBV/C1 and HBV/C2

All of the HBV/C1 isolates recovered in this study belonged to serotype *adr* based on the amino acid sequences at positions 122 and 160 of the S protein. The T→C nucleotide change at position 1858 (C-1858) in codon 15 of the pre-core gene was found in 6 out of 16 isolates (37·5%) of HBV/C1, but in none of the HBV/C2 isolates.

DISCUSSION

HBV genotypes are distributed geographically, but their virulence and pathogenicity differ in each location (Lok, 2000; Kidd-Ljunggren *et al.*, 2002). In addition, subgroups of HBV genotype have also been reported in genotypes A (A′ and A-A′) (Bowyer *et al.*, 1997; Kramvis *et al.*, 2002) and B (Ba and Bj) (Sugauchi *et al.*, 2002a). The genotyping of HBV is important for clarifying the route of infection with and virulence of the virus. In particular, examination of sequence diversity among different isolates of the virus is important since variants may differ in their patterns of serological reactivity, replication of the virus, activity of liver disease, prognosis and response to treatment. In fact, patients infected with genotype C have a more aggressive clinical phenotype than those with genotype B (Orito *et al.*, 2001; Kao *et al.*, 2002). Interestingly, however, isolates within the same genotype can cause different clinical manifestations, e.g. between subgroups Ba and Bj (Sugauchi *et al.*, 2003). In the present study, we focused on genotype C because it is prevalent mainly in Asia and seems to contribute to progressive liver disease and poor clinical outcomes in infected patients. We found that genotype C detected in Asia could be classified into at least two subgroups, which we named HBV/C1 and HBV/C2. Notably, HBV/C1 was found only in southeast Asia including Vietnam, Myanmar and Thailand, while HBV/C2 was found in Far East Asia including Japan, Korea and China.

By means of phylogenetic analysis in the pre-S region containing the pre-S1 to the pre-S2 genes from 56 isolates in Asia, we identified two major subgroups within genotype C. In addition to these subgroups, there were additional small clusters consisting of two isolates from Okinawa and two isolates from China. It has been reported that HBV from Australian Aborigines showed 7·1% difference at the nucleotide level and belonged to a novel genotype C variant (Alestig *et al.*, 2001a). These above findings were also confirmed by phylogenetic analysis of the full genome sequences of 32 isolates. Based on analysis of the full genome sequences, two HBV isolates from Okinawa showed a closer relationship to C2 than to C1, but formed a different branch with a 100% bootstrap value. Therefore, HBV isolates from Okinawa could be considered as variants of genotype C, but do not belong to the C1 or C2 subgroups.

The time that has elapsed since the divergence of HBV/C1 can be estimated on the basis of the complete genome differences compared with HBV/C2. From the assumed rate of $4{\cdot}5 \times 10^{-5}$ mutations per site per year (Orito *et al.*, 1989), HBV/C1 would have diverged from genotype C about 160 years ago (data not shown). The geographical distribution of these subgroups could help us to understand how HBV is spreading in Asia.

The amino acid changes specific to HBV/C1 and HBV/C2 were concentrated in the pre-S1, pre-S2 and P regions, but not in the X and core regions. The pre-S1 region contains

Table 2. Mean percentage and range of nucleotide divergence over the complete genome and the large S gene among 43 isolates of HBV with different genotypes

Values above the diagonal correspond to pairwise comparisons of the complete genomes and those below the diagonal correspond to comparisons of the large S genes. Values in parentheses indicate the standard deviation (bootstrap value 1000; Kimura two-parameter method). Divergences in intragenotypes and/or intrasubgroups are shown in bold, the upper value corresponding to comparison of the complete genome and the lower to comparison of the large S genes.

Genotype	C1	A	B	C2	D	E	F	G
C1	**3·1 (3·0–3·2)** **2·9 (2·6–3·2)**	9·5 (9·0–10·0)	9·3 (8·8–9·8)	5·2 (4·9–5·5)	10·7 (10·2–11·2)	14·5 (13·9–15·1)	14·1 (13·5–14·7)	12·8 (12·2–13·4)
A	6·8 (6·2–7·4)	**4·9 (4·5–5·3)** **3·8 (3·2–4·4)**	10 (9·6–10·4)	9·3 (8·7–9·8)	10·1 (9·6–10·6)	13·5 (13·0–14·0)	14·4 (13·7–15·1)	11·5 (10·9–12·1)
B	8·7 (7·9–9·5)	8·7 (7·9–9·5)	**5·6 (5·2–6·0)** **2·8 (2·4–3·2)**	9·3 (8·9–9·7)	10·6 (10·1–11·1)	14·6 (13·9–15·3)	14·5 (13·9–15·1)	13·1 (12·5–13·7)
C2	4·0 (3·6–4·4)	6·8 (6·1–7·5)	8·5 (7·7–9·3)	**3·9 (3·7–4·1)** **2·8 (2·6–3·0)**	10·4 (9·9–10·9)	14·0 (13·4–14·6)	14·3 (13·7–14·9)	12·6 (12·0–13·2)
D	9·7 (8·8–10·6)	9·3 (8·4–10·2)	9·3 (8·4–10·2)	9·1 (8·3–9·9)	**1·2 (0·9–1·5)** **1·2 (0·9–1·5)**	11·1 (10·6–11·6)	14·2 (13·6–14·8)	11·7 (11·1–12·3)
E	9·9 (9·0–10·8)	10·1 (9·2–11·0)	10·6 (9·6–11·6)	9·4 (8·5–10·3)	7·1 (6·3–7·9)	**1·9 (1·5–2·3)** **1·9 (1·5–2·3)**	17·5 (16·9–18·1)	14·9 (14·3–15·5)
F	13·8 (12·7–14·9)	13·7 (12·6–14·8)	14·3 (13·2–15·4)	13·7 (12·5–14·9)	13·1 (12·0–14·2)	14·0 (12·8–15·2)	**4·6 (4·3–4·9)** **2·4 (2·8–3·2)**	14·9 (14·3–15·5)
G	9·8 (8·7–10·9)	8·5 (7·7–9·3)	11·4 (10·4–12·4)	9·5 (8·7–10·3)	8·0 (7·2–8·8)	8·4 (7·6–9·2)	13·2 (12·0–14·4)	**0·4 (0·3–0·5)** **0·6 (0·2–0·8)**

Table 3. Comparison of deduced amino acid sequences among the complete HBV genomes

	Genotype/isolate	S gene							P gene										Reference
		Pre-S1					Pre-S2		TP		Spacer						Pol		
	Amino acid...	51	54	60	62	73	125	227	136	142	231	233	236	248	252	304	354	358	
Consensus	*Genotype C*																		
	Subgroup C1 consensus	Q	A	V	S	S/N	S	L	T/I	Q	M	S	P	S	Q	Q	Y	H	
	Subgroup C2 consensus	H	E	A	A	G	T	S	A	K	L	R	S	P	R	H	H	N	
	Genotype A	H	A	V	A	G	T	S	V	Q	L	S	P	S	R	H	H	H	
	Genotype B	N	D	V	A	G	T	S	V	Q	L	G	Q	P/S	G	H	H	R	
	Genotype D	T	D	A	A	G	T	S	L	Q	L	R	Q	P	R	H	H	H	
	Genotype E	H	E	V	A	G	T	S	V	Q	L	G	Q	S	R	H	H	H	
	Genotype F	S	M	V	G	G	T	S	A	Q	L	N	Q	T	W	N	H	H	
	Genotype G	P	E	V	A	G	T	S	V	Q	L	R	Q	P	R	Y	H	H	
Genotype C isolates	**Subgroup C1**																		
	AB031262 (Vietnam)	Q	A	V	S	N	S	L	I	Q	M	S	P	S	Q	Q	Y	H	Yuasa *et al.* (2000)
	AB031265 (Vietnam)	Q	A	V	S	N	S	L	I	Q	M	S	P	S	Q	Q	Y	H	Yuasa *et al.* (2000)
	AF223957 (Vietnam)	Q	A	V	S	N	S	L	I	Q	M	S	P	S	Q	Q	Y	H	Alestig *et al.* (2001b)
	AF223960 (Vietnam)	Q	A	V	S	N	S	L	I	Q	M	C	P	S	Q	Q	Y	H	Alestig et al. (2001b)
	AF068756 (Thailand)	Q	A	V	S	S	S	L	T	Q	M	S	P	S	Q	Q	Y	H	P. Monkongdee and others (1998). Direct database submission
	AB074047 (Thailand)	Q	A	V	S	S	S	S	A	Q	M	S	P	S	Q	Q	H	H	Sugauchi *et al.* (2002b)
	AB074755 (Thailand)	Q	A	V	A	N	S	L	T	Q	M	S	P	S	Q	Q	Y	H	Sugauchi et al. (2002b)
	VT101 (Vietnam)	Q	A	V	S	N	S	L	I	Q	M	S	P	P	Q	Q	Y	H	This study
	VT103 (Vietnam)	H	A	V	S	G	S	L	I	Q	L	S	P	S	R	K	Y	H	This study
	VT140 (Vietnam)	Q	A	V	S	N	S	L	I	Q	M	S	P	P	Q	Q	Y	H	This study
	MY624 (Myanmar)	Q	A	V	S	S	S	L	T	Q	M	S	P	S	Q	Q	Y	H	This study
	MY656 (Myanmar)	Q	A	V	A	S	S	L	T	Q	M	S	P	S	Q	Q	Y	H	This study
	MY683 (Myanmar)	Q	A	V	S	S	S	L	T	Q	M	S	P	S	Q	Q	Y	R	This study
	TH123 (Thailand)	Q	A	V	A	S	S	L	A	Q	M	S	P	S	Q	Q	Y	H	This study
	TH124 (Thailand)	Q	A	V	S	S	S	L	T	Q	M	S	P	S	Q	Q	H	H	This study
	Subgroup C2																		
	D23681 (Japan)	H	E	A	A	G	T	S	A	K	L	R	S	P	R	H	H	N	Horikita *et al.* (1994)
	D50517 (Japan)	H	E	A	A	G	T	S	A	K	L	R	S	P	R	H	H	N	Asahina *et al.* (1996)
	D50520 (Japan)	H	E	A	A	G	T	S	A	K	L	R	S	P	R	H	H	N	Asahina et al. (1996)
	AY247030 (Korea)	H	E	A	A	G	T	S	V	K	L	R	S	P	R	H	H	N	J. W. Song and others (2003). Direct database submission
	AY247031 (Korea)	H	E	A	A	G	T	S	A	K	L	R	S	P	*R*	*H*	H	N	J. W. Song and others (2003). Direct database submission

Table 3. cont.

Genotype/isolate	S gene							P gene							Reference
	Pre-S1				Pre-S2		TP		Spacer			Pol			
	Amino acid...	51	54	60	62	73	125	227	136	142	231	233	236	248	252	304	354	358	
X14193 (Korea)		R	E	A	A	G	T	S	A	K	L	R	S	P	R	H	H	N	Rho *et al.* (1989)
AF182802 (China)		H	E	A	A	G	T	S	A	K	L	R	S	P	R	Y	H	N	Lin *et al.* (2001)
AF182805 (China)		H	E	A	A	G	T	S	A	K	L	R	S	P	R	H	H	N	Lin et al. (2001)
AF533983 (China)		H	E	A	A	G	T	S	V	K	L	R	S	P	R	H	H	N	Q. M. Dong and others (2002). Direct database submission
Genotype C isolates	C variants																		
	O47 (Okinawa)		Q	E	A	A	G	T	S	A	K	M	R	S	P	R	H	H	N	This study
	O55 (Okinawa)		Q	E	V	A	G	T	S	A	K	M	R	S	P	R	H	H	N	This study
	AB04875 (Australian aborigines)		H	E	A	A	G	T	S	A	K	L	R	S	P	R	H	H	N	Sugauchi *et al.* (2001)
	X75665 (New Caledonia)		H	E	A	A	G	T	S	T	K	L	R	S	P	R	H	H	N	Norder *et al.* (1994)

the HBV receptor for entering hepatocytes (Neurath *et al.*, 1986) and also has sites for transcriptional factors (Melegari *et al.*, 1994). It has been reported that mutations in the CCAAT motif located in the pre-S1 gene result in retention of the S protein and lead to the more aggressive form of HBV-related liver disease (Bock *et al.*, 1999). Therefore, the relationship between HBV/C1 and HBV/C2 and their virulence in chronic liver diseases including hepatocellular carcinoma is of great interest, since the prevalence of hepatocellular carcinoma associated with HBV infection is extremely high in Asia compared with other regions. Of note, genotype C correlates well with the occurrence of hepatocellular carcinoma in Japanese patients, but not in Taiwanese patients younger than 50 years of age (Kao *et al.*, 2000).

Although it is now well established that nt 1858 is critical for the emergence of a pre-core stop-codon mutant at codon 28, the role of the C-1858 variant on the course of HBV infection is still unclear. It has been reported that there was high prevalence of C-1858 strains of genotype C in southeast Asia (Lindh *et al.*, 1997). Furthermore, Alestig *et al.* (2001b) reported that C-1858 strains of HBV found in southeast Asia showed a common phylogenetic origin and represented one of the subgroups of HBV genotype C. However, our study has shown that 10 out of 16 (62·5%) strains of HBV/C1 detected in southeast Asia had T-1858, whereas the remaining six isolates had C-1858. Furthermore, no cases with the pre-core stop codon mutation in HBV/C1 were seen (data not shown). Therefore, we believe that the C-1858 phylogeny entity in southeast Asia does not give a representative view of HBV prevailing in southeast Asia.

Several recent studies have shown the existence of recombination between different HBV genotypes (Morozov *et al.*, 2000; Cui *et al.*, 2002; Sugauchi *et al.*, 2002a). We also looked for the possibility of such recombination in the HBV/C1 and HBV/C2 subgroups, but no evidence for such an event was found using SimPlot analysis (data not shown).

In conclusion, we have presented evidence for the existence of at least two subgroups within genotype C of HBV, designated HBV/C1 and HBV/C2. HBV/C1 prevails in the southeastern part of Asia including Vietnam, Myanmar and Thailand and HBV/C2 in the northeastern part of Asia including Japan, China and Korea. Conserved amino acid sequences between each subgroup were identified. Future studies are needed to determine whether these subtypes correlate with the progression of liver disease including hepatocellular carcinoma, which could influence treatment.

ACKNOWLEDGEMENTS

This study was supported in part by the Scholarship for Foreign Nationals in Japan of the Ichiro Kanehara Foundation, Grant-in-Aid for Science Research of the Ministry of Education, Science and Culture of Japan and the Ministry of Health, Labor and Welfare of Japan.

REFERENCES

Alestig, E., Hannoun, C., Horal, P. & Lindh, M. (2001a). Hepatitis B virus genotypes in Mongols and Australian Aborigines. *Arch Virol* **146**, 2321–2329.

Alestig, E., Hannoun, C., Horal, P. & Lindh, M. (2001b). Phylogenetic origin of hepatitis B virus strains with precore C-1858 variant. *J Clin Microbiol* **39**, 3200–3203.

Asahina, Y., Enomoto, N., Ogura, Y., Kurosaki, M., Sakuma, I., Izumi, N., Marumo, F. & Sato, C. (1996). Sequential changes in full-length genomes of hepatitis B virus accompanying acute exacerbation of chronic hepatitis B. *J Hepatol* **25**, 787–794.

Bock, C. T., Tillmann, H. L., Manns, M. P. & Trautwein, C. (1999). The pre-S region determines the intracellular localization and appearance of hepatitis B virus. *Hepatology* **30**, 517–525.

Bowyer, S. M., van Staden, L., Kew, M. C. & Sim, J. G. (1997). A unique segment of the hepatitis B virus group A genotype identified in isolates from South Africa. *J Gen Virol* **78**, 1719–1729.

Cui, C., Shi, J., Hui, L., Xi, H., Zhuoma, Quni, Tsedan & Hu, G. (2002). The dominant hepatitis B virus genotype identified in Tibet is a C/D hybrid. *J Gen Virol* **83**, 2773–2777.

Horikita, M., Itoh, S., Yamamoto, K., Shibayama, T., Tsuda, F. & Okamoto, H. (1994). Differences in the entire nucleotide sequence between hepatitis B virus genomes from carriers positive for antibody to hepatitis B e antigen with and without active disease. *J Med Virol* **44**, 96–103.

Kao, J. H., Chen, P. J., Lai, M. Y. & Chen, D. S. (2000). Hepatitis B genotypes correlate with clinical outcomes in patients with chronic hepatitis B. *Gastroenterology* **118**, 554–559.

Kao, J. H., Chen, P. J., Lai, M. Y. & Chen, D. S. (2002). Genotypes and clinical phenotypes of hepatitis B virus in patients with chronic hepatitis B virus infection. *J Clin Microbiol* **40**, 1207–1209.

Kidd-Ljunggren, K., Miyakawa, Y. & Kidd, A. H. (2002). Genetic variability in hepatitis B viruses. *J Gen Virol* **83**, 1267–1280.

Kramvis, A., Weitzmann, L., Owiredu, W. K. & Kew, M. C. (2002). Analysis of the complete genome of subgroup A′ hepatitis B virus isolates from South Africa. *J Gen Virol* **83**, 835–839.

Kumar, S., Tamura, K., Jakobsen, I. B. & Nei, M. (2001). MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* **17**, 1244–1245.

Lee, W. M. (1997). Hepatitis B virus infection. *N Engl J Med* **337**, 1733–1745.

Lin, X., Qian, G. S., Lu, P. X., Wu, L. & Wen, Y. M. (2001). Full-length genomic analysis of hepatitis B virus isolates in a patient progressing from hepatitis to hepatocellular carcinoma. *J Med Virol* **64**, 299–304.

Lindh, M., Andersson, A. S. & Gusdal, A. (1997). Genotypes, nt 1858 variants, and geographic origin of hepatitis B virus – large-scale analysis using a new genotyping method. *J Infect Dis* **175**, 1285–1293.

Lok, A. S. (2000). Hepatitis B infection: pathogenesis and management. *J Hepatol* **32**, 89–97.

Magnius, L. O. & Norder, H. (1995). Subtypes, genotypes and molecular epidemiology of the hepatitis B virus as reflected by sequence variability of the S-gene. *Intervirology* **38**, 24–34.

Melegari, M., Bruno, S. & Wands, J. R. (1994). Properties of hepatitis B virus pre-S1 deletion mutants. *Virology* **199**, 292–300.

Morozov, V., Pisareva, M. & Groudinin, M. (2000). Homologous recombination between different genotypes of hepatitis B virus. *Gene* **260**, 55–65.

Naito, H., Hayashi, S. & Abe, K. (2001). Rapid and specific genotyping system for hepatitis B virus corresponding to six major genotypes by PCR using type-specific primers. *J Clin Microbiol* **39**, 362–364.

Neurath, A. R., Kent, S. B., Strick, N. & Parker, K. (1986). Identification and chemical synthesis of a host cell receptor binding site on hepatitis B virus. *Cell* **46**, 429–436.

Norder, H., Courouce, A. M. & Magnius, L. O. (1994). Complete genomes, phylogenetic relatedness, and structural proteins of six strains of the hepatitis B virus, four of which represent two new genotypes. *Virology* **198**, 489–503.

Okamoto, H., Tsuda, F., Sakugawa, H., Sastrosoewignjo, R. I., Imai, M., Miyakawa, Y. & Mayumi, M. (1988). Typing hepatitis B virus by homology in nucleotide sequence: comparison of surface antigen subtypes. *J Gen Virol* **69**, 2575–2583.

Orito, E., Mizokami, M., Ina, Y., Moriyama, E. N., Kameshima, N., Yamamoto, M. & Gojobori, T. (1989). Host-independent evolution and a genetic classification of the hepadnavirus family based on nucleotide sequences. *Proc Natl Acad Sci U S A* **86**, 7059–7062.

Orito, E., Mizokami, M., Sakugawa, H., Michitaka, K., Ishikawa, K., Ichida, T., Okanoue, T., Yotsuyanagi, H. & Iino, S. (2001). A case–control study for clinical and molecular biological differences between hepatitis B viruses of genotypes B and C. Japan HBV Genotype Research Group. *Hepatology* **33**, 218–223.

Rho, H. M., Kim, K., Hyun, S. W. & Kim, Y. S. (1989). The nucleotide sequence and reading frames of a mutant hepatitis B virus subtype adr. *Nucleic Acids Res* **17**, 2124.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.

Shiina, S., Fujino, H., Uta, Y. & 7 other authors (1991). Relationship of HBsAg subtypes with HBeAg/anti-HBe status and chronic liver disease. Part I: analysis of 1744 HBsAg carriers. *Am J Gastroenterol* **86**, 866–871.

Sugauchi, F., Mizokami, M., Orito, E. & 7 other authors (2001). A novel variant genotype C of hepatitis B virus identified in isolates from Australian Aborigines: complete genome sequence and phylogenetic relatedness. *J Gen Virol* **82**, 883–892.

Sugauchi, F., Orito, E., Ichida, T. & 9 other authors (2002a). Hepatitis B virus of genotype B with or without recombination with genotype C over the precore region plus the core gene. *J Virol* **76**, 5985–5992.

Sugauchi, F., Chutaputti, A., Orito, E., Kato, H., Suzuki, S., Ueda, R. & Mizokami, M. (2002b). Hepatitis B virus genotypes and clinical manifestation among hepatitis B carriers in Thailand. *J Gastroenterol Hepatol* **17**, 671–676.

Sugauchi, F., Orito, E., Ichida, T. & 10 other authors (2003). Epidemiologic and virologic characteristics of hepatitis B virus genotype B having the recombination with genotype C. *Gastroenterology* **124**, 925–932.

Yuasa, R., Takahashi, K., Dien, B. V., Binh, N. H., Morishita, T., Sato, K., Yamamoto, N., Isomura, S., Yoshioka, K., Ishikawa, T., Mishiro, S. & Kakumu, S. (2000). Properties of hepatitis B virus genome recovered from Vietnamese patients with fulminant hepatitis in comparison with those of acute hepatitis. *J Med Virol* **61**, 23–28.