Journal of General Virology (1998), **79**, 2499–2505. Printed in Great Britain

Four geographically distinct genotypes of JC virus are prevalent in China and Mongolia: implications for the racial composition of modern China

Jing Guo,[1] Chie Sugimoto,[1] Tadaichi Kitamura,[2] Hideki Ebihara,[3] Atsushi Kato,[2] Zheng Guo,[4] Jing Liu,[5] Shu Ping Zheng,[6] Yue Ling Wang,[7] Yen Qun Na,[8] Makoto Suzuki,[2] Fumiaki Taguchi[3] and Yoshiaki Yogo[1]

[1] Department of Viral Infection, The Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan

[2] Department of Urology, Faculty of Medicine, The University of Tokyo, Tokyo, Japan

[3] Department of Microbiology, School of Allied Health Sciences, Kitasato University, Kanagawa, Japan

[4,7] Department of Pediatrics[4] and Department of Cardioangiology[7], The First Affiliated Hospital Hubei Medical University, Wuhan, People's Republic of China

[5] Department of Obstetrics and Gynaecology, The Maternity and Child Health Hospital in Hubei, Wuhan, People's Republic of China

[6] Department of Pediatrics, The Fifth Hospital of Wuhan, Wuhan, People's Republic of China

[8] Department of Urology, The First Hospital, Beijing Medical College, Beijing, People's Republic of China

JC polyomavirus (JCV) is ubiquitous in humans, persisting in renal tissue and excreting progeny in urine. It has been shown that the genotyping of urinary JCV offers a novel means of tracing human migrations. This approach was used to elucidate the racial composition of modern China. JCV isolates in the Old World were previously classified into nine distinct genotypes. One of them (B1) has a wide domain, encompassing part of Europe and the entirety of Asia. By constructing a neighbour-joining phylogenetic tree, all B1 isolates detected so far were classified into four distinct groups (B1-a to -d), each occupying unique domains in the world. According to this revised classification system of JCV DNAs, four genotypes (CY, SC, B1-a and -b) were found to be prevalent in China and Mongolia (Mongolia was studied instead of Inner Mongolia, which is part of China). There was a remarkable variation in the incidence of genotypes among the sites of sample collection. CY was more frequently detected in Northern China, SC was predominant in Southern China and B1-b was detected only in Mongolia. B1-a was spread throughout China. These data were statistically analysed and the observed regional differences in the incidence of genotypes were found to be significant. It is likely that these differences in JCV distribution in China reflect the intermingling of different population groups that constitute modern China.

Introduction

JC polyomavirus (JCV) was first isolated in 1971 from the brain of a patient with progressive multifocal leukoencephalopathy (PML) (Padgett *et al.*, 1971). This virus is ubiquitous among humans, however, infecting children asymptomatically and then persisting in renal tissue (Padgett & Walker, 1973, 1976; Chesters *et al.*, 1983; Tominaga *et al.*, 1992; Kitamura *et al.*, 1997). Most adults excrete JCV DNA in the urine (Kitamura *et al.*, 1990, 1994), from where it can readily be isolated by molecular cloning or the PCR method (Yogo *et al.*, 1990; Flægstad *et al.*, 1991).

JCV is transmitted frequently from parents to children (Kunitake *et al.*, 1995), but is rarely transmitted between human populations (Kato *et al.*, 1997). JCV DNAs in the Old World are classified into several distinct genotypes, of which the

Author for correspondence: Yoshiaki Yogo.
Fax +81 3 5449 5409. e-mail yogo@ims.u-tokyo.ac.jp

The DNA sequence data reported here have been deposited in the GSDB, DDBJ, EMBL and NCBI nucleotide sequence databases with accession numbers C22715–C22762.

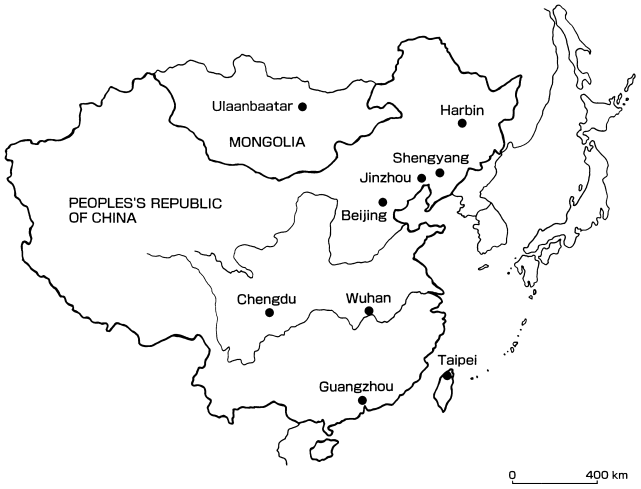

Fig. 1. Eight sites in China and Mongolia at which urine samples were collected.

regional distribution patterns are compatible with the migration and expansion of *Homo sapiens sapiens* (Sugimoto *et al*., 1997). On the basis of these findings, we proposed that JCV genotyping is a novel means of tracing human migrations (Sugimoto *et al*., 1997). Indeed, Agostini *et al*. (1997) detected Asian genotypes of JCV in samples from Native Americans and a Pacific Island population, which is consistent with the Asian origin of these native populations. In the present study, this approach was used to elucidate the racial composition of modern China.

We previously showed that three JCV genotypes (CY, SC and B1) are spread in China and that a single genotype (B1) predominates in Mongolia (Sugimoto *et al*., 1997). Genotype B1 has a wide domain, extending from Europe through the Near East to the eastern edge of Asia (Sugimoto *et al*., 1997). By a phylogenetic method, we further divided this genotype into four distinct groups (B1-a to -d), each occupying a unique domain. According to this revision of our JCV classification system, four JCV genotypes (CY, SC, B1-a and -b) were identified in China and Mongolia. The incidence of these JCV genotypes varied significantly depending upon the regions. We discuss the implications of these findings for the racial component of modern China.

Methods

■ **JCV isolates.** Most of the JCV isolates studied here were previously obtained from urine specimens collected from immunocompetent individuals in various geographical regions (Guo *et al*., 1996; Sugimoto *et al*., 1997). Five isolates (#202–#205 and GS/B) were obtained from the brain of PML patients (Ault & Stoner, 1992; Loeber & Dörries, 1988), and three were from the urine of patients with human immunodeficiency virus infection (#223) or multiple sclerosis (#227, #230) (Agostini *et al*., 1998*b*). The sites in China and Mongolia at which urine samples were collected are shown in Fig. 1.

■ **Phylogenetic analysis.** The reported 610 bp long IG sequences (Table 1) were used for the phylogenetic analysis. (The IG region was previously established as a region of the JCV genome that contains abundant type-determining sites; Ault & Stoner, 1992.) A neighbour-joining (NJ) phylogenetic tree (Saitou & Nei, 1987) was constructed using the CLUSTAL W program (Thompson *et al*., 1994). Divergences were estimated by the two-parameter method (Kimura, 1980). A phylogenetic tree was visualized using the TREEVIEW 1.4 program (Page, 1996). The bootstrap test was applied to estimate the confidence of the branching patterns of the NJ tree (Felsenstein, 1985).

Table 1. Origin of IG sequences used for phylogenetic analysis

JCV isolates	Country (city)	Reference
N3, N4, N6, N7	Netherlands (Deventer)	Sugimoto *et al*. (1997)
GS/B	Germany	Loeber & Dörries (1988)
GR-2, -5 to -8, -10, -11, -14, -16	Greece (Athens)	
MU-5	Mauritius (Port Louis)	
TU-5, -9, -11	Turkey (Ankara)	
SA-3, -5, -9, -11 to -13	Saudi Arabia (Riyadh)	
IN-12	India (Varanasi)	
SL-1 to -5	Sri Lanka (Colombo)	
MO-2 to -5, -7 to -10	Mongolia (Ulaanbaatar)	
MN-2, -4, -8, -9, -13, -15	Myanmar (Yangon)	Sugimoto *et al*. (1997)
ML-1	Malaysia (Masai)	
PH-2, -5, -7, -8	Philippines (Pamalican Is.)	
SJ-5	China (Shenyang/Jinzhou)	
CB-2, -9	China (Beijing)	
CD-1, -10	China (Chengdu)	
GZ-1, -2, -6 to -8	China (Guangzhou)	
C2, C7	China (Taiwan)	
#202 to #205	USA	Ault & Stoner (1992)
#223, #227, #230	USA	Agostini *et al*. (1998*b*)

■ **Statistical analysis.** Fisher's exact test was performed using a network algorithm described by Mehta & Patel (1983).

Results

Sub-classification of genotype B1

We previously grouped 60 JCV isolates as genotype B1 (Sugimoto *et al.*, 1997). In addition, we found that one JCV isolate from Germany (GS/B) (Loeber & Dörries, 1988) and seven from the United States (#202–#205, #223, #227 and #230) (Ault & Stoner, 1992; Agostini *et al.*, 1998*b*) belong to B1 (data not shown). For these 68 B1 isolates, an unrooted NJ phylogenetic tree was constructed using reported IG sequences. According to the constructed tree (Fig. 2), the 68 isolates diverged into four clusters, which we tentatively designated genotypes B1-a to -d. From the tree (Fig. 2), the

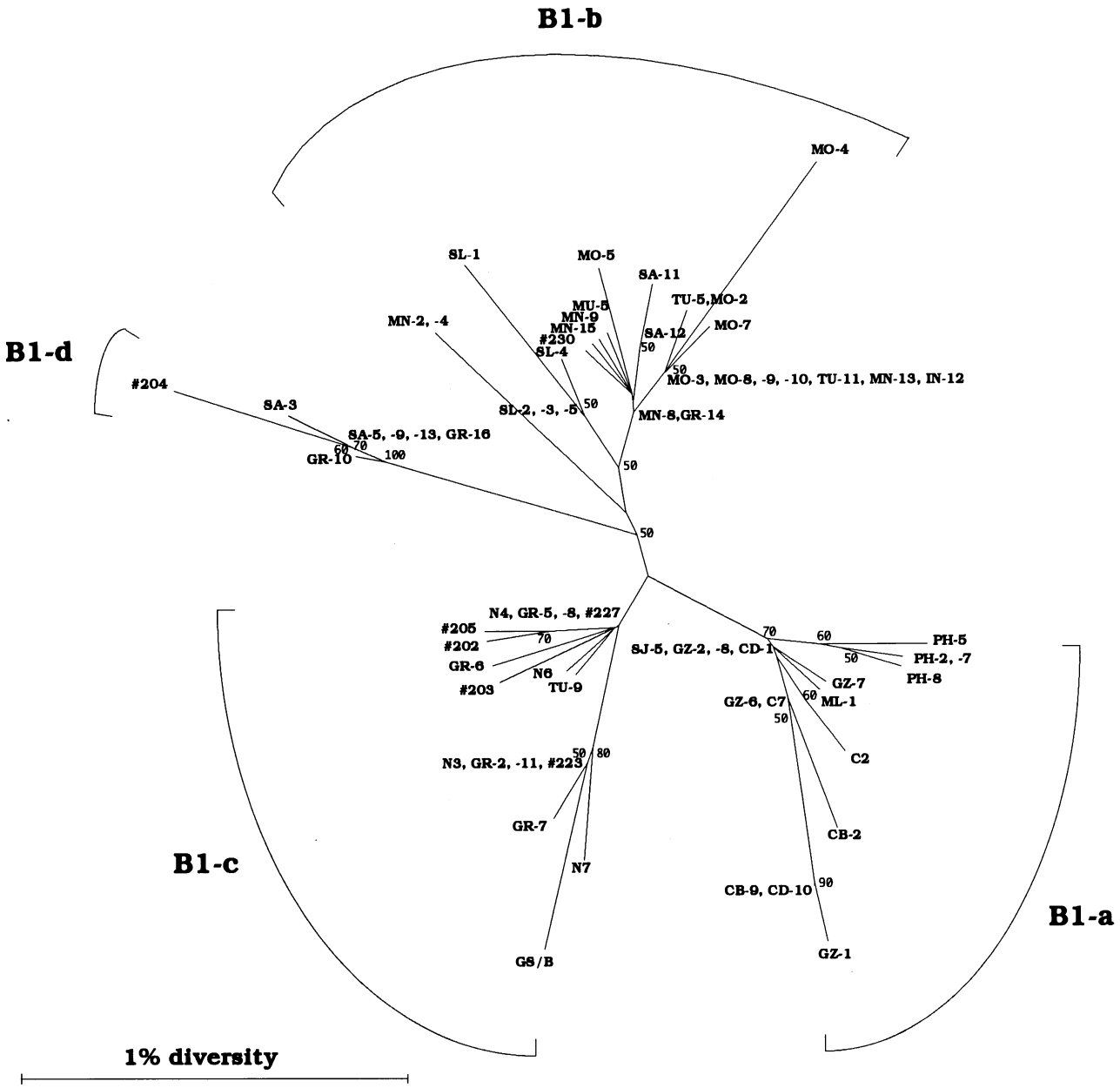

Fig. 2. Unrooted NJ tree constructed to correlate various B1 isolates in the world. The NJ tree was constructed from reported IG sequences (see Table 1) using the CLUSTAL W program (Thompson *et al.*, 1994). The phylogenetic tree was visualized by the TREEVIEW 1.4 program (Page, 1996). The symbol for each sequence is shown in Table 1. The bootstrap confidence levels obtained by 100 replicates are shown (only significant values are indicated). Genotypes (B1-a to -d) for major clusters are shown.

Table 2. Worldwide distribution of genotypes B1-a to -d

Country	No. of JCV isolates*				
	B1-a	B1-b	B1-c	B1-d	Total
Netherlands	0	0	4	0	4
Germany	0	0	1	0	1
Greece	0	1	6	2	9
Mauritius	0	1	0	0	1
Turkey	0	2	1	0	3
Saudi Arabia	0	2	0	4	6
India	0	1	0	0	1
Sri Lanka	0	5	0	0	5
Mongolia	0	8	0	0	8
Myanmar	0	6	0	0	6
Malaysia	1	0	0	0	1
Philippines	4	0	0	0	4
China	12	0	0	0	12
USA	0	1	5	1	7
Total	17	27	17	7	68

* Estimated from Fig. 2.

numbers of isolates belonging to each genotype were estimated for each geographical region (Table 2). The following correlation is evident between these genotypes and geographical regions.

B1-a included only isolates from East Asia (Malaysia, Philippines and China). Therefore, B1-a represents one of the major East Asian JCV genotypes (the other major East Asian genotypes are CY, SC and MY; Sugimoto *et al.*, 1997).

B1-b was composed of one major and two minor clusters. The major cluster contained 75% of the B1-b isolates examined, including isolates from Greece, Turkey, Saudi Arabia, India and Mongolia. In addition, one isolate (#230) from the United States was found in the major cluster of B1-b. One minor cluster of B1-b contained all five isolates from Sri Lanka, and the other minor B1-b contained two from Myanmar.

B1-c included isolates from Europe (Netherlands, Germany and Greece) and the United States. It appears that B1-c represents the minor JCV genotype prevalent in Caucasians (the major JCV genotype in Caucasians is EU; Sugimoto *et al.*, 1997; Agostini *et al.*, 1998*a*). Agostini *et al.* (1998*b*) reported that an Asian JCV genotype (type 2B according to their classification system) exists in Europe and the United States. This genotype corresponds to B1-c, since it included three isolates (#223, #227 and GS/B) that we assigned as B1-c.

B1-d included isolates from Greece and Saudi Arabia. One isolate (#204) from the United States was found in B1-d.

JCV genotypes prevalent in China

We used the revised system described above to classify the JCV isolates which we have obtained from various sites of China and Mongolia (Fig. 1) (Guo *et al.*, 1996; Sugimoto *et al.*, 1997). The numbers of JCV isolates in individual genotypes are shown on the bottom line of Table 3. [We found that a sequenced Taiwan JCV, Taiwan-3 (Ou *et al.*, 1997), belongs to genotype SC. However, since Taiwan-3 was obtained from an immunocompromised patient, this isolate was not used to elucidate the distribution of JCV genotypes in China.]

It is clear that four major genotypes (CY, SC, B1-a and -b) are prevalent in China and Mongolia. These accounted for 97% of the JCVs examined. Two isolates from Harbin and Chengdu (HB-2 and CD-7, respectively) which were previously included in genotype SC (Sugimoto *et al.*, 1997) were grouped here as genotype X, since it appeared that these isolates together with a few isolates from Thailand constitute an independent cluster (Sugimoto *et al.*, 1997).

Table 3. Distribution of various JCV genotypes in China and Mongolia

B1-a and -b were assigned according to the NJ tree shown in Fig. 2, and genotypes CY and SC were assigned as described by Sugimoto *et al.* (1997). X was assigned to two isolates, HB-2 and CD-7 (see text).

Geographical region	No. of isolates examined	No. of isolates belonging to genotype				
		CY	SC	B1-a	B1-b	X
Harbin	6	5	0	0	0	1
Shengyang/Jinzhou*	7	5	1	1	0	0
Beijing	10	8	0	2	0	0
Wuhan	10	4	6	0	0	0
Chengdu	10	0	7	2	0	1
Guangzhou	13	1	7	5	0	0
Taipei	9	1	6	2	0	0
Ulaanbaatar	12	3	1	0	8	0
Total (%)	77	27 (35)	28 (36)	12 (16)	8 (10)	2 (3)

* Since Shengyang and Jinzhou are closely located, the isolates from these regions were combined.

Table 4. Incidence of various genotypes in North China, South China and Mongolia

Table 2 was revised for this statistical analysis. Incidence higher than that in the other regions is underlined.

Geographical region	No. of isolates examined	No. of isolates (%) belonging to genotype			
		CY	SC	B1-a	B1-b
North China*	23	18 (<u>78</u>)	1 (4)	3 (13)	0 (0)
South China†	42	6 (14)	26 (<u>62</u>)	9 (<u>21</u>)	0 (0)
Mongolia‡	12	3 (25)	1 (8)	0 (0)	8 (<u>67</u>)

* Harbin, Shengyang, Jinzhou and Beijing.
† Wuhan, Chengdu, Guangzhou and Taipei.
‡ Ulaanbaatar.

Table 5. Statistical analysis of the differences in genotype incidence between geographical regions

Data shown in Table 4 were analysed using Fisher's exact test (Mehta & Patel, 1983).

Genotype	Test	*P* value
CY	North China vs South China	0·01
	North China vs Mongolia	0·01
SC	South China vs North China	0·01
	South China vs Mongolia	0·01
B1-a	South China vs North China	–*
	South China vs Mongolia	0·1
B1-b	Mongolia vs North China	0·01
	Mongolia vs South China	0·01

* Not significant.

Each genotype appeared to have a unique geographical distribution (Table 3). CY was more frequently detected in Northern China, including Harbin, Shengyang/Jinzhou and Beijing. In contrast, SC was detected mainly in Southern China, including Chengdu, Guangzhou and Taipei. B1-b was found only in Ulaanbaatar, Mongolia. B1-a was spread in both Northern and Southern China, with the highest incidence in Guangzhou, but this genotype was not found in Ulaanbaatar.

For the statistical examination of the observed regional differences in genotype incidence, the sites of urine collection were grouped into three regions, North China, South China and Mongolia. North China included Harbin, Shengyang, Jinzhou and Beijing; South China included Wuhan, Chengdu, Guangzhou and Taipei; and Mongolia included only Ulaanbaatar. The incidence of CY, SC, B1-a and -b in these regions is shown in Table 4. For all genotypes, the genotype incidence in one region (underlined in Table 4) was higher than that in the other regions. These trends were examined using Fisher's exact test (Mehta & Patel, 1983), and the following statements were confirmed (Table 5). (i) CY was more prevalent in North China than in the other regions. (ii) SC was more prevalent in South China than in the other regions. (iii) B1-b was more prevalent in Mongolia than in the other regions. (iv) B1-a was apparently more prevalent in South China than in Mongolia, but its incidence did not significantly differ between North and South China.

Discussion

Using the revised JCV classification system, which differentiates genotype B1 into B1-a to -d, we identified four JCV genotypes (CY, SC, B1-a and -b) in China and Mongolia (Mongolia was studied instead of Inner Mongolia, which is part of China). Furthermore, we found that there is a remarkable variation in the incidence of JCV genotypes among the sites of urine collection. The observed regional differences in the incidence of genotypes were statistically significant. As shown previously (Kato *et al.*, 1997; Sugimoto *et al.*, 1997; Agostini *et al.*, 1997), there is a close correlation between JCV genotypes and human populations. The above-noted findings on the distribution of JCV genotypes in China, therefore, have several implications for the origin of modern Chinese.

(i) A JCV genotype named CY was more frequently detected in Northern China, whereas another genotype, SC, was predominant in Southern China. This finding is consistent with the genetic distinction between Northern and Southern Chinese (Cavalli-Sforza *et al.*, 1994 and references cited therein).

(ii) The finding that the third major JCV genotype (B1-a) was spread throughout China suggests that the third lineage, represented by B1-a, occurs in the modern Chinese population. This lineage has not been identified by studies on human genes (Cavalli-Sforza *et al.*, 1994).

(iii) It is thought that more than 90% of modern Chinese belong to a single group called Han (Beijing Normal University, 1983). The present detection of three major JCV genotypes in China suggests that the so-called Han group is a population formed by the intermixture of several genetically distinct groups.

There are a number of minor aboriginal groups in China (Beijing Normal University, 1983). Minor JCV genotypes (X in Harbin and Chengdu) may have originated from one of these groups. However, it appears that we missed many other minor genotypes, because most of the aboriginal groups are probably located in regions not examined in this study. We emphasize that the present data on the distribution of JCV genotypes in China should provide the basis of future studies on the origins of the minor aboriginal groups in China. For example, if a significant number of individuals of an aboriginal group carry a new JCV genotype that can be distinguished from those identified in this study, it would substantially support the notion that this group is unique not only in culture but also in origin.

The present data regarding the prevalent JCV genotypes in China also have implications for the origins of neighbouring human populations. Some examples are described below.

(i) Mongolians were unique in that they contain genotype B1-b, which did not occur in Chinese. However, they carried two minor JCV genotypes, CY and SC, that are prevalent in China. From these data, it appears that the Mongolian population originally represented a single ethnic group, but was later partly mixed with other ethnic groups in China.

(ii) Two JCV genotypes, CY and MY, are predominant in Japan (Sugimoto *et al*., 1997), suggesting that the modern Japanese population was formed by the intermixing of two peoples. The finding of CY as a major JCV genotype in China implies that one of the founding groups of China also contributed to the founding of modern Japanese. However, since MY, the other major genotype in Japan, has not been detected in China, the racial affinity between modern Japanese and Chinese is partial.

(iii) The findings that SC is the JCV genotype predominant in Southern China (the present study) and that this genotype also predominates in Southeast Asia including Malaysia, Thailand and Indonesia (Sugimoto *et al*., 1997) suggest a close genetic relationship between Southern Chinese and Southeastern Asians. Thus, the regional distribution of SC is in good agreement with a genetic study of Southeastern Asians (Cavalli-Sforza *et al*., 1994).

We are grateful to Dr Yoshiyuki Nagai for helpful discussions and to Drs Chikuma Hamada and Akifumi Yafune for their help in statistical analysis. We thank Dr Gerald L. Stoner for sending a copy of the manuscript by Agostini *et al*. (1998*b*) prior to publication. This study was supported in part by grants from the Ministry of Education, Science, Sports and Culture, Japan and from the Ministry of Welfare, Japan. C.S. was supported by JSPS Research Fellowships for Younger Scientists.

References

Agostini, H. T., Yanagihara, R., Davis, V., Ryschkewitsch, C. F. & Stoner, G. L. (1997). Asian genotypes of JC virus in Native Americans and in a Pacific Island population: markers of viral evolution and human migration. *Proceedings of the National Academy of Sciences, USA* **94**, 14542–14546.

Agostini, H. T., Ryschkewitsch, C. F. & Stoner, G. L. (1998*a*). JC virus Type 1 has multiple subtypes: three new complete genomes. *Journal of General Virology* **79**, 801–805.

Agostini, H. T., Shishido-Hara, Y., Baumhefner, R. W., Singer, E. J., Ryschkewitsch, C. F. & Stoner, G. L. (1998*b*). JC virus Type 2: definition of subtypes based on DNA sequence analysis of ten complete genomes. *Journal of General Virology* **79**, 1143–1151.

Ault, G. S. & Stoner, G. L. (1992). Two major types of JC virus defined in progressive multifocal leukoencephalopathy brain by early and late coding region DNA sequences. *Journal of General Virology* **73**, 2669–2678.

Beijing Normal University (1983). The investigation and statistics of the population of China. The third census in China. In *A Dictionary on the China National Conditions*, pp. 109–111. The Publishing House of Beijing Normal University.

Cavalli-Sforza, L. L., Menozzi, P. & Piazza, A. (1994). The history and geography of human genes. Princeton: Princeton University Press.

Chesters, P. M., Heritage, J. & McCance, D. J. (1983). Persistence of DNA sequences of BK virus and JC virus in normal human tissues and in diseased tissues. *Journal of Infectious Diseases* **147**, 676–684.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791.

Flægstad, T., Sundsfjord, A., Arthur, R. R., Pedersen, M., Traavik, T. & Subramani, S. (1991). Amplification and sequencing of the control regions of BK and JC virus from human urine by polymerase chain reaction. *Virology* **180**, 553–560.

Guo, J., Kitamura, T., Ebihara, H., Sugimoto, C., Kunitake, T., Takehisa, J., Na, Y. Q., Al-Ahdal, M. N., Hallin, A., Kawabe, K., Taguchi, F. & Yogo, Y. (1996). Geographical distribution of the human polyomavirus JC virus types A and B and isolation of a new type from Ghana. *Journal of General Virology* **77**, 919–927.

Kato, A., Kitamura, T., Sugimoto, C., Ogawa, Y., Nakazato, K., Nagashima, K., Hall, W. W., Kawabe, K. & Yogo, Y. (1997). Lack of evidence for the transmission of JC polyomavirus between human populations. *Archives of Virology* **142**, 875–882.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* **16**, 111–120.

Kitamura, T., Aso, Y., Kuniyoshi, N., Hara, K. & Yogo, Y. (1990). High incidence of urinary JC virus excretion in nonimmunosuppressed older patients. *Journal of Infectious Diseases* **161**, 1128–1133.

Kitamura, T., Kunitake, T., Guo, J., Tominaga, T., Kawabe, K. & Yogo, Y. (1994). Transmission of the human polyomavirus JC virus occurs both within the family and outside the family. *Journal of Clinical Microbiology* **32**, 2359–2363.

Kitamura, T., Sugimoto, C., Kato, A., Ebihara, H., Suzuki, M., Taguchi, F., Kawabe, K. & Yogo, Y. (1997). Persistent JC virus (JCV) infection is demonstrated by continuous shedding of the same JCV strains. *Journal of Clinical Microbiology* **35**, 1255–1257.

Kunitake, T., Kitamura, T., Guo, J., Taguchi, F., Kawabe, K. & Yogo, Y. (1995). Parent-to-child transmission is relatively common in the spread of the human polyomavirus JC virus. *Journal of Clinical Microbiology* **33**, 1448–1451.

Loeber, G. & Dörries, K. (1988). DNA rearrangements in organ-specific variants of polyomavirus JC strain GS. *Journal of Virology* **62**, 1730–1735.

Mehta, C. R. & Patel, N. R. (1983). A network algorithm for performing Fisher's exact test in $r \times c$ contingency tables. *Journal of American Statisticians Association* **78**, 427–434.

Ou, W. C., Tsai, R. T., Wang, M., Fung, C. Y., Hseu, T. H. & Chang, D.

(1997). Genomic cloning and sequence analysis of Taiwan-3 human polyomavirus JC virus. *Journal of the Formosan Medical Association* **96**, 511–516.

Padgett, B. L. & Walker, D. L. (1973). Prevalence of antibodies in human sera against JC virus, an isolate from a case of progressive multifocal leukoencephalopathy. *Journal of Infectious Diseases* **127**, 467–470.

Padgett, B. L. & Walker, D. L. (1976). New human papovaviruses. *Progress in Medical Virology* **22**, 1–35.

Padgett, B. L., Walker, D. L., ZuRhein, G. M., Eckroade, R. J. & Dessel, B. H. (1971). Cultivation of papova-like virus from human brain with progressive multifocal leucoencephalopathy. *Lancet* i, 1257–1260.

Page, R. D. M. (1996). TREEVIEW: an application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* **12**, 357–358.

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

Sugimoto, C., Kitamura, T., Guo, J., Al-Ahdal, M. N., Shchelkunov, S. N., Otova, B., Ondrejka, P., Chollet, J.-Y., El-Safi, S., Ettayebi, M., Grésenguet, G., Kocagöz, T., Chaiyarasamee, S., Thant, K. Z., Thein, S., Moe, K., Kobayashi, N., Taguchi, F. & Yogo, Y. (1997). Typing of urinary JC virus DNA offers a novel means of tracing human migrations. *Proceedings of the National Academy of Sciences, USA* **94**, 9191–9196.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673–4680.

Tominaga, T., Yogo, Y., Kitamura, T. & Aso, Y. (1992). Persistence of archetypal JC virus DNA in normal renal tissue derived from tumor-bearing patients. *Virology* **186**, 736–741.

Yogo, Y., Kitamura, T., Sugimoto, C., Ueki, T., Aso, Y., Hara, K. & Taguchi, F. (1990). Isolation of a possible archetypal JC virus DNA sequence from nonimmunocompromised individuals. *Journal of Virology* **64**, 3139–3143.

Received 11 May 1998; Accepted 22 June 1998