AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY 128:185–193 (2005)

Genetic Diversity of JC Virus in the Saami and the Finns: Implications for Their Population History

Hiroshi Ikegaya,[1]* Huai-Ying Zheng,[2] Pekka J. Saukko,[1] Leena Varesmaa-Korhonen,[3] Tapani Hovi,[4] Timo Vesikari,[5] Hideki Suganami,[6] Tomokazu Takasaka,[2] Chie Sugimoto,[7] Yasuo Ohasi,[8] Tadaichi Kitamura,[2] and Yoshiaki Yogo[2]

[1]*Department of Forensic Medicine, University of Turku, Turku, Finland 20520*
[2]*Department of Urology, Faculty of Medicine, University of Tokyo, Tokyo, Japan 113-0033*
[3]*Ivalo Health Center, Ivalo, Finland 99800*
[4]*Department of Microbiology, National Public Health Institute, Mannerheimintie, Helsinki, Finland 00300*
[5]*Department of Virology and Vaccine Research, University of Tampere Medical School, Tampere, Finland 33014*
[6]*Graduate Course of Biomedical Statistics, Tokyo University of Science, Tokyo, Japan162-8602*
[7]*Tsukuba Primate Center for Medical Science, National Institute of Infectious Diseases, Tsukuba, Japan 305-0843*
[8]*Department of Biostatistics/Epidemiology and Preventive Health Science, Faculty of Medicine, University of Tokyo, Tokyo, Japan 113-0033*

KEY WORDS JC virus; Saami; Finnish; genetic diversity; human migration

ABSTRACT The JC virus (JCV) genotyping method was used to gain insights into the population history of the Saami and the Finns, both speaking Finno-Ugric languages and living in close geographic proximity. Urine samples from Saami and Finns, collected in northern and southern Finland, respectively, were used to amplify a 610-bp JCV-DNA region containing abundant type-specific mutations. Based on restriction site polymorphisms in the amplified fragments, we classified JCV isolates into one of the three superclusters of JCV, type A, B, or C. All 15 Saami isolates analyzed and 41 of 43 Finnish isolates analyzed were classified as type A, the European type, and two samples from Finns were classified as type B, the African/Asian type. We then amplified and sequenced a 583-bp JCV-DNA region from the type A isolates of Saami and Finns. According to type-determining nucleotides within the region, we classified type A isolates into EU-a1, -a2, or -b. Most type A isolates from Saami were classified as EU-a1, while type A isolates from Finns were distributed among EU-a1, EU-a2, and EU-b. This trend in the JCV-genotype distribution was statistically significant. On a phylogenetic tree based on complete sequences, most of the type A isolates from Saami were clustered in a single clade within EU-a1, while those from Finns were distributed throughout EU-a1, EU-a2, and EU-b. These findings are discussed in the context of the population history of the Saami and the Finns. This study provides new complete JCV DNA sequences derived from populations of anthropological interest. Am J Phys Anthropol 128:185–193, 2005.

Although most Central and Southern Europeans belong to the Indo-European language family, some Northern Europeans belong to the Finno-Ugric language family, and hence their ethnic origin has interested many geneticists. By using the JC virus (JCV) genotyping method, we investigated the population history of the Saami and the Finns, belonging to the Finno-Ugric language family and living in close geographic proximity.

JCV is a member of the *Polyomaviridae* family. Its genome is a single molecule of covalently closed, circular double-stranded DNA about 5,100 bp in length (Cole and Conzen, 2001). JCV is ubiquitous among humans, infecting children asymptomatically, and then persisting in renal tissue (Chesters et al., 1983; Kitamura et al., 1997; Padgett and Walker, 1973; Tominaga et al., 1992; Walker and Padgett, 1983). In adults, renal JCV replicates and progeny viruses are excreted in urine (Kitamura et al., 1990, 1994; Agostini et al., 1996). The main mode of transmission of JCV is from parents to children through long-term cohabitation (Kunitake et al., 1995; Kato et al., 1997; Suzuki et al., 2002; Zheng et al., 2004). Kunitake et al. (1995) estimated that only 50% of virus transmission occurs from parents to children. Nevertheless, the strong ethnic distribution of JCV (see below) can be explained, as transmission outside the family is often mediated by close contact of children with other family members (e.g., grandparents) or

Grant sponsor: Ministry of Education, Culture, Sports, Science and Technology, Japan; Grant sponsor: Ministry of Health, Labor and Welfare, Japan.

*Correspondence to: Hiroshi Ikegaya, M.D., Ph.D., Department of Forensic Medicine, Graduate School of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan.
E-mail: ikegaya-tky@umin.ac.jp

Received 17 June 2004; accepted 12 August 2004

DOI 10.1002/ajpa.20189
Published online 18 March 2005 in Wiley InterScience (www.interscience.wiley.com).

with individuals from the same community (e.g., babysitters) (Kunitake et al., 1995).

JCV strains throughout the world are classified into more than 10 geographically distinct genotypes (Sugimoto et al., 1997; Guo et al., 1998; Agostini et al., 2001a). The evolutionary relationships among JCV genotypes were examined by Sugimoto et al. (2002a,b), using the whole-genome approach with which a highly reliable phylogeny of JCV isolates can be reconstructed (Hatwell and Sharp, 2000). The results of these studies showed that ancestral JCV first divides into three superclusters, types A, B, and C, which subsequently split into the various genotypes of JCV distributed throughout the world. A split in type A generated three genotypes, EU-a, -b, and -c. EU-a and -b mainly contain European and Mediterranean isolates, and EU-c contains northeastern Siberian isolates. The split in type B generated Af2 (the major African subtype), B1-c (a minor European subtype), and various Asian subtypes (e.g., B1-a, -b, and -d, B2, CY, MY, and SC). Type C generated a single subtype (Af1), consisting of isolates derived from western Africa. Pavesi (2003) and Chima et al. (1998) supported the African origin of JCV by identifying Af1 (or type 6 according to the nomenclature by Agostini et al., 2001a) as the putative ancestral genotype.

The unique mode of JCV transmission, the geographic distribution patterns of various genotypes of JCV, and the mode of JCV evolution together showed JCV to be a marker of human populations (Sugimoto et al., 1997; Agostini et al., 1997; Disotell, 2003). Indeed, the JCV genotyping method was used to elucidate the origins of various ethnic groups worldwide (Kitamura et al., 1998; Sugimoto et al., 2002b; Agostini et al., 1997, 2001b; Guo et al., 1998; Jobes et al., 2001; Fernandez-Cobo et al., 2002; Yanagihara et al., 2002; Saruwatari et al., 2002; Miranda et al., 2002; Zheng et al., 2003).

The Saami and the Finns speak Finno-Ugric languages and live in close geographic proximity. Their population history was inferred based on studies using human mtDNA and the Y chromosome as markers. As described below (see Discussion), different views were suggested by these studies: those using mtDNA supported the view that both populations were formed by European populations, whereas those using the Y chromosome suggested a contribution by Asians in the formation of the Saami and the Finns. In this study, a complementary approach based on JCV was used to solve this issue.

MATERIALS AND METHODS

Urine samples

Urine samples from Finns were collected with informed consent from 98 general patients aged 40 years or older in Helsinki and Tampere. Urine samples from Saami were collected from general patients, aged 40 years or older, at the Ivalo Health Center. The places of residence of the Saami urine donors were Inari (n = 17), Utsjoki (n = 14), Petsamo (n = 10), Pello (n = 1), Enontekiö (n = 4), Tampere (n = 1), Rovaniemi (n = 2), and Karasjoki (n = 1). To our knowledge, no Saami donors had any ancestry from other ethnic groups. They gave informed consent prior to their inclusion in the study. Urine samples were stored in 10 mM EDTA (pH 8.0), and sent to the Laboratory of Viral Infection (Department of Virology, Institute of Medical Science, University of Tokyo, Japan), where DNA was extracted as described previously (Kitamura et al., 1990). This study was approved by the Ethics Committee of the Hospital District of Varsinais-Suomi.

Polymerase chain reaction

From extracted DNA, the 610-bp IG region and 583-bp ET region of the viral genome were amplified by polymerase chain reaction (PCR), using primer pairs P1/P2 and E3/E4, respectively, under conditions described previously (Kunitake et al., 1995). The sequences of P1 and P2 were described previously (Kunitake et al., 1995). E3 (5′-*GGCTGC*-AGGCCACTCATACACCCAAAGTA-3′) consisted of nucleotides (nt) 1255–1276 of the JCV genome (Frisque et al., 1984), and a 5′-terminal heptanucleotide (shown in italics), which was added to generate a *Pst*I site, and E4 (5′-*GGAAGCT*TGCACTGGCTT-CCCTGCACCAT-3′) consisted of nt 1881–1860, and a 5′-terminal heptanucleotide (shown in italics), which was added to generate a *Hin*dIII site. The IG region was previously established as a region of the JCV genome containing abundant type-determining sites (Ault and Stoner, 1992). The ET region encompasses the 3′-terminal region of the VP2/3 gene and the 5′-terminal region of the VP1 gene, and was established in this study as a region containing multiple nucleotide changes unique to each subtype within type A.

Restriction fragment length polymorphism analysis

Restriction fragment length polymorphism (RFLP) analysis analysis was performed for amplified IG fragments using four restriction enzymes (*Dde*I, *Hinc*II, *Nlu*I, and *Pvu*II), as described (Kitamura et al., 1998).

Molecular cloning

PCR-amplified fragments were digested with a combination of *Hin*dIII and *Pst*I. The recovered fragments were ligated, and were *Hin*dIII- phosphatase-treated and *Pst*I-digested, alkaline with pBluescript II SK (+) (Stratagene, La Jolla, CA), and used to transform competent *Escherichia coli* HB101 cells (Takara Shuzo Co. Ltd., Kyoto, Japan).

Plasmids containing the IG or ET region were prepared using a Qiagen Plasmid Mini kit (Qiagen GmbH, Hilden, Germany).

Entire JCV DNAs were cloned into pUC19 at a unique *Bam*HI site, as described previously (Yogo et al., 1991). The resultant complete JCV DNA clones were prepared using a QIAGEN Plasmid Maxi kit (QIAGEN). Purified plasmids were sequenced as described previously (Sugimoto et al., 2002a).

DNA sequencing

Purified recombinant plasmids containing partial or complete genomes were sequenced with an autosequencer (ABI Prism 373S DNA Analyzer, Applied Biosystems, Foster City, CA). Primers for sequencing partial genomes (IG and ET regions) were T3 and T7 promoters within the vector, and those for sequencing complete genomes were described previously (Sugimoto et al., 2002a).

Phylogenetic analysis

The noncoding regulatory region of the JCV genome was excluded from phylogenetic analysis. This was because an accurate multisequence alignment is hampered by the presence of extensive sequence rearrangements usually occurring in the regulatory regions of JCV isolates from the brains of patients with progressive multifocal leukoencephalopathy (Yogo and Sugimoto, 2001). DNA sequences were aligned using CLUSTAL W (Thompson et al., 1994), with a gap-opening penalty of 15.00 and gap-extension penalty of 6.66. To evaluate the phylogenetic relationships among DNA sequences, we constructed neighbor-joining (NJ) trees (Saitou and Nei, 1987), using the CLUSTAL W program. Divergences were estimated by the two-parameter method (Kimura, 1980), using CLUSTAL W. Phylogenetic trees were visualized using TREEVIEW (Page, 1996). To assess the confidence of branching patterns of NJ trees, 1,000 bootstrap replicates were made (Felsenstein, 1985). Bootstrap probabilities larger than 70% were considered significant (Hills and Bull, 1993).

Statistical analysis

The differences in genotype distribution between the Saami and Finns (2 × 3 contingency table) were analyzed using the overall likelihood ratio chi-square test. When there was a statistically significant difference, in order to explore the details of the difference, the chi-square value was divided into two portions (Agresti, 2002), i.e., the difference in intratype A subtypes and the difference in intra-EU-a subgroups between the two ethnic groups. All analyses were performed using SAS software version 8.2 (SAS Institute, Inc., Cary, NC).

TABLE 1. Finnish and Saami JCV isolates whose complete DNA sequences were determined in this study

Isolate	Ethnic origin	Geographic origin	Genotype[1]	Accession no.[2]
Sam - 3	Saami	Utsjoki	EU - a1	AB127342
Sam - 8	Saami	Inari	EU - a2	AB127343
Sam - 9	Saami	Tampere[3]	EU - a1	AB127344
Sam - 12	Saami	Rovaniemi	EU - a1	AB127345
Sam -15	Saami	Petsamo	EU - a1	AB127346
FL - 3	Finn	Tampere	EU - a1	AB127347
FL - 4	Finn	Tampere	EU - a1	AB127348
FL - 7	Finn	Tampere	EU - a2	AB127349
FL - 8	Fin	Tampere	EU - b	AB127350
FL - 11	Finn	Tampere	EU - a2	AB127351
FL - 13	Finn	Helsinki	EU - a2	AB127353

[1] Determined according to Figure 1.
[2] GSDB, DDBJ, EMBL, and NCBI nucleotide sequence database.
[3] Parents of urine donor lived in Inari.

RESULTS

Superclusters of JCV in Saami and Finnish samples

We detected JCV DNA from Saami and Finn urine samples using a PCR that amplified the 610-bp IG region. The detection rates were 15/50 (30%) and 43/98 (44%) in the urine samples from the Saami and Finns, respectively. We classified the JCV isolates in these samples into superclusters of JCV by RFLP analysis of amplified IG fragments (Kitamura et al., 1998). All 15 isolates from the Saami and most isolates from the Finns were classified as type A (also designated as EU; Kitamura et al., 1998). Two isolates from Finns, however, were classified as type B. Phylogenetic analysis of these type B isolates revealed that they belonged to B1-c and -b (data not shown). B1-c is a minor genotype of JCV in Europe, while B1-b is widespread in West and Central Asia (Sugimoto et al., 1997).

Phylogenetic analysis using the whole-genome approach

We established five and six JCV-DNA clones containing complete type A sequences from the samples of Saami and Finns, respectively (Table 1). The complete type A JCV DNA sequences determined in this and previous studies (Table 2), together with a complete GH-1 sequence as the outgroup, were aligned and used to construct an NJ tree. Strain GH-1 (Sugimoto et al., 2002a) was chosen as the outgroup because it belongs to the predicted ancestral genotype Af1 (Pavesi, 2003). According to the resultant tree (Fig. 1), type A generated three clusters, EU-a to -c, with high bootstrap probability (99–100%). The first split in type A generated EU-c and non-EU-c, and the subsequent split in the latter generated EU-a and -b. In addition, we found that EU-a split into two major clusters, EU-a1 and -a2, with high bootstrap probability (88% or 96%).

TABLE 2. Origins of JCV isolates whose complete DNA sequences were used in phylogenetic analysis (see Fig. 1) [1]

Genotype	Isolate	Geographic origin	Accession no.[2]	Reference[3]
EU - a1	SW - 3	Sweden	AB048575	a
EU - a1	G2	Germany	AB038251	b
EU - a1	N5	Netherlands	AB074588	c
EU - a1	IT - 2	Italy	AB074582	c
EU - a1	# 124	USA	AF015526	e
EU - a1	Mad - 1	Madison, WI	J02226	d
EU - a1 / Jpn	AT - 2	Northeast Japan	AB048569	a
EU - a1 / Jpn	AT - 4	Northeast Japan	AB048570	a
EU - a1 / Jpn	HR - 5	Northeast Japan	AB048572	a
EU - a1 / Jpn	HR - 13	Northeast Japan	AB048571	a
EU - a1 / Jpn	SD - 9	Northeast Japan	AB048573	a
EU - a1 / Arc	ES - 1	Arctic Canada	AB074578	c
EU - a1 / Arc	ES - 3	Arctic Canada	AB074579	c
EU - a1 / Arc	KO - 2	Northeast Siberia	AB074585	c
EU - a1 / Arc	KO - 3	Northeast Siberia	AB074586	c
EU - a1 / Arc	KO - 5	Northeast Siberia	AB074587	c
EU - a2	G4	Germany	AB074580	c
EU - a2	G5	Germany	AB074581	c
EU - a2	N2	Netherlands	AB048574	a
EU - a2	UK - 2	England	AB048576	a
EU - a2	IT - 3	Italy	AB074583	c
EU - a2	IT - 5	Italy	AB048568	a
EU - a2	IT - 8	Italy	AB074584	c
EU - a2	SP - 7	Spain	AB074591	c
EU - a2	# 123	USA	AF015527	e
EU - b	N25	Netherlands	AB048565	a
EU - b	UK - 1	England	AB048567	a
EU - b	SP - 1	Spain	AB048566	a
EU - b	GR - 3	Greece	AB048563	a
EU - b	MR - 7	Morocco	AB048564	a
EU - b	# 402	USA	AF015528	e
EU - c	AM - 5	Northeast Siberia	AB074576	c
EU - c	AM - 7	Northeast Siberia	AB074577	c
EU - c	AM - 18	Northeast Siberia	AB074575	c
EU - c	SI - 1	Northeast Siberia	AB074589	c
EU - c	SI - 7	Northeast Siberia	AB074590	c

[1] Isolates for which complete sequences were determined in this study are shown in Table 1.
[2] GSDB, DDBJ, EMBL, and NCBI accession numbers.
[3] a, Sugimoto et al., 2002a; b, Kato et al., 2000; c, Sugimoto et al., 2002b; d, Frisque et al., 1984; e, Agostini et al., 1998.

All five Saami isolates were found in the EU-a1 cluster. In contrast, Finnish isolates were spread into various clusters, two in EU-a1, three in EU-a2, and one in EU-b. Four of the five Saami isolates, together with a single Danish isolate (N5), formed a distinct cluster (the Sam cluster) with a high bootstrap probability (96%). The urine donors, from whom the four Sam isolates were obtained, were unrelated and lived in different regions (Table 1).

Subclassification of type A isolates

By examining the aligned type A sequences used for the phylogenetic tree (Fig. 1), we identified a 583-bp region (ET region) of the genome where nucleotide variation was most extensive among intratype A genotypes (i.e., EU-a1, -a2, -b, and -c) (Table 3). The number of nucleotide differences varied from 3 (between EU-a1 and -a2) to 7 (between EU-a2 and -c). We attempted to PCR-amplify this region from Saami and Finn urine samples (excluding those in which complete JCV DNA sequences were analyzed), and obtained amplified fragments from 10 and 21 urine samples in the Saami and Finns, respectively. Sequencing of the amplified fragments unequivocally classified each isolate into EU-a1, -a2, or -b (no isolate was classified as EU-c).

Based on the type A subclassification data obtained by phylogenetic analysis (Fig. 1) and by nucleotide variations in the ET region, we estimated the frequency of each intratype A genotype in the Saami and Finns (Table 4), and the findings can be summarized as follows: 1) EU-b was detected in the Finns at a low rate (11%), but no EU-b was detected in the Saami. The observed difference in the detection of EU-b between the Saami and Finns, however, was not statistically significant. 2) In the Finns, both EU-a1 and -a2 were detected at similar rates, while in the Saami, EU-a1 was more frequently detected than EU-a2.

The difference in genotype distribution was analyzed using the overall likelihood ratio chi-square

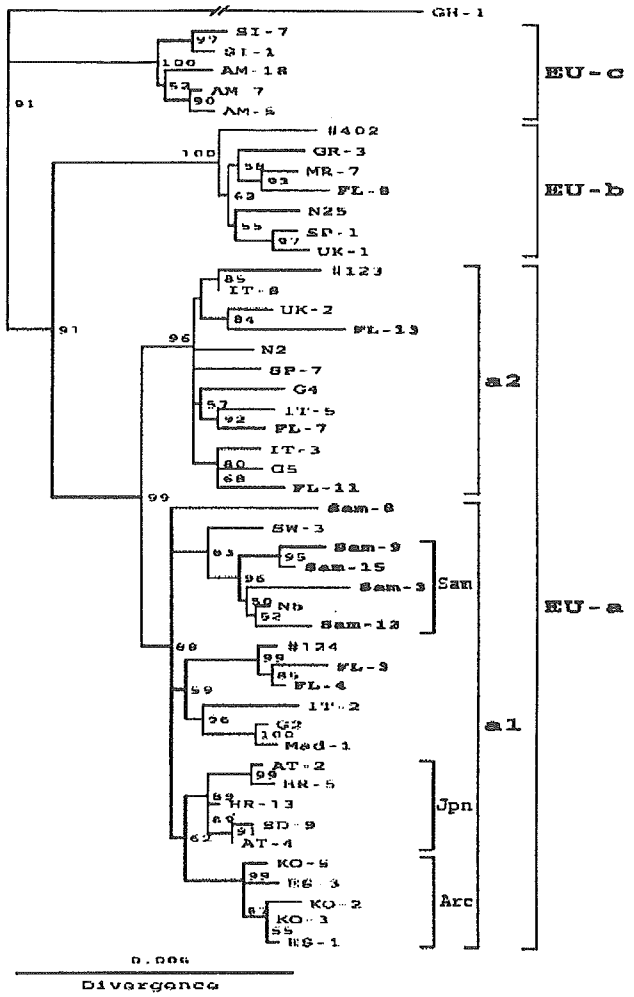

Fig. 1. NJ phylogenetic tree relating 47 complete type A JCV DNA sequences. Phylogenetic tree was constructed from complete sequences, excluding regulatory sequences, using NJ method. An isolate (GH-1) belonging to type C was used as outgroup. Symbols for sequences are shown in Tables 1 and 2. Numbers at nodes in tree indicate bootstrap probabilities (percent) obtained by 1,000 replicates (only those ≥50% are shown). Intratype A subtypes (EU-a, -b, and -c) and intra-EU-a subgroups (EU-a1 and -a2) are indicated. In addition, some ethnically distinct clusters within EU-a1 (EU-a/Arc (Arc), EU-a/Jpn (Jpn), and Sam are indicated.

test. There was a significant difference in genotype distribution between the Saami and Finns (χ^2 = 12.67, df = 2, P = 0.0018). Therefore, the chi-square statistics (12.67) were divided into two portions, 2.78 and 9.89. Although there was no significant difference in intratype A subtypes between the Saami and Finns (χ^2 = 2.78, df = 1, P = 0.0956), a statistically significant difference in intra-EU-a subgroups between the two ethnic groups was detected (χ^2 = 9.89, df = 1, P = 0.0017). The results of these analyses showed that the differences between Saami and Finns were mainly characterized by different proportions of EU-a1 and EU-a2.

DISCUSSION

Inference based on JCV genotypes

Using the JCV marker, this study provides insights into the population history of the Saami and the Finns, both speaking Finno-Ugric languages and living in close geographic proximity. We found that all or most of the JCV isolates detected in both populations belonged to type A, a supercluster mainly including European and Mediterranean isolates (Sugimoto et al., 2002a,b). This finding suggests that, like other Europeans, the Saami and Finns are of European descent. We detected no or few isolates in the Saami and the Finns that belonged to the Asian genotypes of JCV, and it therefore appears that the contribution by Asians to the formation of the Saami and the Finns was not significant.

According to the phylogenetic analysis of five type 1 isolates (three from the United States and two from Hungary), Agostini et al. (2001b) showed that type 1 divided into two groups, types 1A and 1B (type 1 corresponds to EU-a). In this study, based on complete JCV DNA sequences of a larger number (35 in total) of type A isolates, mostly derived from Europeans (including the Saami and Finns), we confirmed that EU-a divided into two groups, EU-a1 and -a2, with high bootstrap probability (EU-a1 and -a2 corresponded to types 1A and 1B, respectively). Although the above phylogenetic analysis was conducted using only the NJ method, it should be noted that if fully determined JCV DNA sequences are used as the data set, all aspects of JCV evolution observed in the NJ are reproduced in the phylogenetic analysis with the maximum likelihood method (a typical analysis based multiple tree algorithms) (Sugimoto et al., 2002a,b; Zheng et al., 2003; Yogo et al., 2003).

According to this updated classification of type A JCVs, we obtained the following findings: 1) EU-a1 and -a2 comprised most strains isolated from the Saami and the Finns; 2) rare EU-b isolates occurred in the Finns, but not in the Saami; 3) EU-c was detected in neither the Finns nor the Saami; 4) EU-a1 and -a2 occurred to an almost equal extent in the Finns, while in the Saami, EU-a1 was more prevalent than EU-a2; and 5) in a phylogenetic tree relating fully sequenced type A isolates, the type A isolates from Saami were all found within EU-a1, while those from Finns were widespread in EU-a1, -a2, and -b. In addition, we detected a few isolates classified as type B (B1-b and B1-c) only in the Finnish samples. From these findings, we conclude that the genetic diversity of JCV is more marked in the Finns than in the Saami. We therefore inferred that the Saami (at least the Saami population in northern Finland) were formed by a single founding population and that they did not significantly admix with other populations in the course of their subsequent history. On the other hand, the genetic diversity of JCV in the Finns may represent an admix-

TABLE 3. Nucleotide variations among various subgroups within type A

Genotype	Nucleotides at indicated positions[1]								
	1282	1308	1361	1651	1689	1753	1818	1843	1850
EU - a1	G	C	G	A	A	A	G	G	A
EU - a2	G	C	G	A	G	A	G	T	G
EU - b	G	C	A	T	A	A	C	G	A/G
EU - c	A	T	G	A	A	T	C	G	A

[1]Nucleotide numbers are those of Mad - 1 (Frisque et al., 1984), starting from midpoint of origin of replication and proceeding clockwise.

TABLE 4. Subclassification of type A isolates detected from Saami and Finns

Ethnic group	No. of isolates analyzed	Frequency of JCV genotype (%)[1]			
		EU - a			EU - b
		Total	a1	a2	
Saami	15	15 (100)	13 (87)	2 (13)	0 (0)
Finns	27	24 (89)	9 (33)	15 (56)	3 (11)

[1]Classified based on phylogenetic analysis (Fig. 1) or nucleotide variations in ET region. No isolate was classified as EU-c.

ture of different populations carrying distinct JCV genotypes.

An interesting feature of the updated pattern for the evolution of type A JCV (Fig. 1) is that EU-a2 mainly contained isolates from Europeans, and that EU-a1 contained not only those from Europeans but also those from three distinct ethnic groups, northeastern Siberians/Inuits, northeastern Japanese, and Saami. Based on this finding, the following scenario is conceivable: An ancestral population carrying proto-EU-a split into two subpopulations, one carrying EU-a1 and the other carrying EU-a2. The subpopulation carrying EU-a1 would have generated various subgroups that migrated through Eurasia to the west and east, whereas the subpopulation carrying EU-a2 would have generated various subgroups that migrated only to Europe.

EU-b (or type 4 according to the alternative nomenclature; Agostini et al., 2001a) was detected in various areas of Europe (Agostini et al., 2001b; Dubois et al., 2001; Pagani et al., 2003). Interestingly, it appears that the frequency of EU-b significantly differs among geographic regions. Thus, EU-b occurred at higher frequency (>23%) in Southwest Europe (e.g., Spain, France, and Italy), and occurred at lower frequency (<18%) in Northeast Europe (e.g., Germany, Poland, and Hungary). In addition, EU-b occurred at notably high frequency in some ethnic groups (i.e., Basques, Sinti, and Roma; Agostini et al., 2001b). Our observation that EU-b did not occur in the Saami and occurred in the Finns at only low frequency is consistent with the regional distribution pattern of EU-b described above, suggesting that people carrying EU-b were distributed at higher rates in Southwest than Northeast Europe.

In the Finnish samples, the presence of an isolate belonging to B1-c (also named type 2B) deserves some additional attention. This genotype appears to have originated in Asia, as it was shown by phylogenetic analysis that B1-c belongs to the type B supercluster that contains the major African genotype (Af2 or type 3) and most of the Asian genotypes of JCV (Sugimoto et al., 2002a). B1-c occurs at higher frequency (about 30%) in the Netherlands, Greece (Guo et al., 1998) and Macedonia (unpublished findings), while it occurs at lower frequency (2–10%) in Germany, France, Italy, and Finland (Agostini et al., 2001b; Dubois et al., 2001; Pagani et al., 2003; this study). It has not been detected in East Europe, including Russia, Poland, Hungary, and the Czech Republic (Sugimoto et al., 1997; Agostini et al., 2001b). These regional distribution patterns of B1-c suggest ancient human dispersal carrying B1-c from the Near East to South Europe. Indeed, Pavesi (2003) proposed the hypothesis that the early peopling of Europe was mediated by two distinct human migrations, one carrying the European genotype (EU-a/b), and the other the Euro-Asiatic genotype (B1-c).

Compatibility with genetic findings

Geneticists have investigated the origin of the Saami and Finns, analyzing the mtDNA and Y-chromosomal diversity in these ethnic groups.

About 40% of the Saami carry mtDNA belong to haplogroup V (Torroni et al., 1998), and this high incidence is linked to very low sequence variation (Torroni et al., 1998, 2001) ("haplogroup" refers to the grouping of mtDNAs with mutational events defined by single-nucleotide polymorphisms, restriction enzyme recognition sites, or deletions at the same location). It was suggested that the Iberian peninsula is the most likely homeland of this haplogroup (Torroni et al., 1998). Moreover, many Saami mtDNAs (32–52%) harbor the "Saami" motif, defined by three nucleotide variants belonging to a subcluster (U5) of haplogroup U (Sajantila et al., 1995; Finnilä et al., 2000). U5 is widespread in Europe, albeit at lower frequency (Simoni et al., 2000). "Asian" mtDNA haplogroups (M and Z) occur only at low frequencies (Lahermo et al., 1996; Torroni et al., 1998; Meinilä et al., 2001). On the other hand, the Finns harbor all nine mtDNA haplogroups observed in other Europeans populations (Torroni et al., 1996, 1998). Nevertheless, U5, which was frequently observed in the Saami (see above), is more common in Finland than in other European populations

(Finnilä et al., 2000; Meinilä et al., 2001). Based on mtDNA, it therefore appears that both the Saami and Finns are mainly of European origin. In addition, it was suggested that the Saami (but not the Finns) were markedly influenced by the founder effect or the bottleneck phenomenon, and that Finns were more significantly admixed with the Saami than other Europeans speaking Indo-European languages.

According to a detailed study on Y-chromosomal diversity in Europe (Rosser et al., 2000), most of the ethnic groups speaking Finn-Ugric languages, including the Saami and Finns, harbor two classes of haplogroups, one that spread throughout Europe, and the other that spread throughout Northeastern Europe and Central/Northern Asia. The latter is represented by haplogroup 16, which carries a novel mutation (Tat C mutation) (Zerjal et al., 1997). Zerjal et al. (1997) suggested Central Asia as the geographic origin of this mutation, while Lahermo et al. (1999) suggested northern Eurasia. From the Y-chromosomal polymorphisms found in the Saami and Finns, it can be inferred that both ethnic groups were formed by multiple founding populations. However, as Y-chromosomal diversity is lower in Finns than in the other European populations, including the Saami, it appears that the Finns, rather than the Saami, were influenced by the founder effect or bottleneck phenomenon (Kittles et al., 1999). Explanations of the low diversity in the Finnish Y chromosome (in contrast to diverse Finnish mtDNA) have been offered in the literature (Kittles et al., 1999).

Recent studies showed that type A JCV occurs not only in European and Mediterranean areas but also in the Far East, including northern Japan and northeastern Siberia (Kitamura et al., 1998; Sugimoto et al., 2002a,b). This type A JCV evolved as distinct clusters (EU-a/Jpn and EU-a/Arc) within subtype EU-a or a novel subtype designated EU-c. According to the geographic distribution of type A JCV, it was previously suggested that small bands, derived from an ancestral population that generated modern Europeans, migrated through Eurasia from west to east, and colonized northeastern Siberia and northern Japan (Sugimoto et al., 2002a,b; Yogo et al., 2003). By analogy, it may be speculated that an ancestral population that generated modern Europeans also generated a population harboring the Tat C mutation; this population further split into two subpopulations, one moving to Northeastern Europe, the other moving to Northern Asia. If this speculation proves valid, then the Saami as well as the Finns carry two classes of Y-chromosome haplogroups, both of Caucasoid origin.

CONCLUSIONS

The population history of the Saami and Finns was previously inferred based on studies using mtDNA and the Y chromosome as markers. In this study, we used a new marker, the JCV genotype, to infer their population history. Studies based on the JCV genotype and mtDNA suggested that both populations are mostly of Caucasoid origin, but studies based on the Y chromosome suggested that Asians significantly contributed to the formation of the Saami and the Finns. Furthermore, there are other disagreements among studies based on the three markers as to the numbers of founding populations in the Saami and Finns. We believe that the population history of the Saami and Finns should be elucidated by more detailed studies using a multidisciplinary approach.

ACKNOWLEDGMENTS

We thank all sample donors. The helpful suggestions of the reviewers are gratefully acknowledged.

LITERATURE CITED

Agostini HT, Ryschkewitsch CF, Stoner GL. 1996. Genotype profile of human polyomavirus JC excreted in urine of immunocompetent individuals. J Clin Microbiol 34:159–164.

Agostini HT, Yanagihara R, Davis V, Ryschkewitsch CF, Stoner GL. 1997. Asian genotypes of JC virus in Native Americans and in a Pacific Island population: markers of viral evolution and human migration. Proc Natl Acad Sci USA 94:14542–14546.

Agostini HT, Ryschkewitsch CF, Stoner GL. 1998. JC virus type 1 has multiple subtypes: three new complete genomes. J Gen Virol 79:801–805.

Agostini HT, Jobes DV, Stoner GL. 2001a. Molecular evolution and epidemiology of JC virus. In: Khalili K, Stoner GL, editors. Human polyomaviruses: molecular and clinical perspectives. New York: John Wiley & Sons. p 491–526.

Agostini HT, Deckhut A, Jobes DV, Girones R, Schlunck G, Prost MG, Frias C, Perez-Trallero E, Ryschkewitsch CF, Stoner GL. 2001b. Genotypes of JC virus in East, Central and Southwest Europe. J Gen Virol 82:1221–1231.

Agresti A. 2002. Categorical data analysis, 2nd ed. New York: Wiley-Interscience. p 82–83.

Ault GS, Stoner GL. 1992. Two major types of JC virus defined in progressive multifocal leukoencephalopathy brain by early and late coding region DNA sequences. J Gen Virol 73:2669–2678.

Chesters PM, Heritage J, McCance DJ. 1983. Persistence of DNA sequences of BK virus and JC virus in normal human tissues and in diseased tissues. J Infect Dis 147:676–684.

Chima SC, Ryschkewitsch CF, Stoner GL. 1998. Molecular epidemiology of human polyomavirus JC in Biaka Pygmies and Bantu of Central Africa. Mem Inst Oswaldo Cruz 93: 615–623.

Cole CN, Conzen SD. 2001. *Polyomaviridae*: the viruses and their replication. In: Knipe DM, Howley PM, Griffin DE, Martin MA, Lamb RA, Roizman B, Straus SE, editors. Fields virology,4th ed. Philadelphia: Lippincott Williams & Wilkins. p 2141–2173.

Disotell TR. 2003. Discovering human history from stomach bacteria. Genome Biol 4:213.

Dubois V, Moret H, Lafon ME, Brodard V, Icart J, Ruffault A, Guist'hau O, Buffet-Janvresse C, Abbed K, Dussaix E, Ingrand D. 2001. JC virus genotypes in France: molecular epidemiology and potential significance for progressive multifocal leukoencephalopathy. J Infect Dis 183:213–217.

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

Fernandez-Cobo M, Agostini HT, Britez G, Ryschkewitsch CF, Stoner GL. 2002. Strains of JC virus in Amerind-speakers of North America (Salish) and South America (Guarani), Na-Dene-speakers of New Mexico (Navajo), and modern Japanese suggest links through an ancestral Asian population. Am J Phys Anthropol 118:154–168.

Finnilä S, Hassinen IE, Ala-Kokko L, Majamaa K. 2000. Phylogenetic network of the mtDNA haplogroup U in northern Finland based on sequence analysis of the complete coding region by conformation-sensitive gel electrophoresis. Am J Hum Genet 66:1017–1026.

Frisque RJ, Bream GL, Cannella MT. 1984. Human polyomavirus JC virus genome. J Virol 51:458–469.

Guo J, Sugimoto C, Kitamura T, Ebihara H, Kato A, Guo Z, Liu J, Zheng SP, Wang YL, Na YQ, Suzuki M, Taguchi F, Yogo Y. 1998. Four geographically distinct genotypes of JC virus are prevalent in China and Mongolia: implications for the racial composition of modern China. J Gen Virol 79:2499–2505.

Hatwell JN, Sharp PM. 2000. Evolution of human polyomavirus JC. J Gen Virol 81:1191—1200.

Hills DM, Bull JJ. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst Biol 42:182–192.

Jobes DV, Friedlaender JS, Mgone CS, Agostini HT, Koki G, Yanagihara R, Ng TCN, Chima SC, Ryschkewitsch CF, Stoner GL. 2001. New JC virus (JCV) genotypes from Papua New Guinea and Micronesia (type 8 and type 2E) and evolutionary analysis of 32 complete JCV genomes. Arch Virol 146:2097–2113.

Kato A, Kitamura T, Sugimoto C, Ogawa Y, Nakazato K, Nagashima K, Hall WW, Kawabe K, Yogo Y. 1997. Lack of evidence for the transmission of JC polyomavirus between human populations. Arch Virol 142:875–882.

Kato A, Sugimoto C, Zheng H-Y, Kitamura T, Yogo Y. 2000. Lack of disease-specific amino acid changes in the viral proteins of JC virus isolates from the brain with progressive multifocal leukoencephalopathy. Arch Virol 145:2173–2182.

Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 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. J Infect Dis 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. J Clin Microbiol 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. J Clin Microbiol 35:1255–1257.

Kitamura T, Sugimoto C, Ebihara H, Kato A, Guo J, Taguchi F, Tominaga T, Ogawa Y, Ohta N, Kizu N, Imamura K, Funaki H, Kurosawa T, Ichikawa S, Suzuki T, Chiba K, Nagashima K, Yasumoto S, Yogo Y. 1998. Peopling of Japan as revealed by genotyping of urinary JC virus DNA. Anthropol Sci 106:311–325.

Kittles RA, Bergen AW, Urbanek M, Virkkunen M, Linnoila M, Goldman D, Long JC. 1999. Autosomal, mitochondrial, and Y chromosome DNA variation in Finland: evidence for a male-specific bottleneck. Am J Phys Anthropol 108:381–399.

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. J Clin Microbiol 33:1448–1451.

Lahermo P, Sajantila A, Sistonen P, Lukka M, Aula P, Peltonen L, Savontaus ML. 1996. The genetic relationship between the Finns and the Finnish Saami (Lapps): analysis of nuclear DNA and mtDNA. Am J Hum Genet 58:1309–1322.

Lahermo P, Savontaus ML, Sistonen P, Beres J, de Knijff P, Aula P, Sajantila A. 1999. Y chromosomal polymorphisms reveal founding lineages in the Finns and the Saami. Eur J Hum Genet 7:447–458.

Meinilä M, Finnilä S, Majamaa K. 2001. Evidence for mtDNA admixture between the Finns and the Saami. Hum Hered 52:160–170.

Miranda JJ, Sugimoto C, Paraguison R, Takasaka T, Zheng H-Y, Yogo Y. 2003. Genetic diversity of JC virus in the modern Filipino population: implications for the peopling of the Philippines. Am J Phys Anthropol 120:125–132.

Padgett BL, Walker DL. 1973. Prevalence of antibodies in human sera against JC virus, an isolate from a case of progressive multifocal leukoencephalopathy. J Infect Dis 127:467–470.

Pagani E, Delbue S, Mancuso R, Borghi E, Tarantini L, Ferrante P. 2003. Molecular analysis of JC virus genotypes circulating among the Italian healthy population. J Neurovirol 9:559–566.

Page RD. 1996. TREEVIEW: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12: 357–358.

Pavesi A. 2003. African origin of polyomavirus JC and implications for prehistoric human migrations. J Mol Evol 56:564–572.

Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, Amos W, Armenteros M, Arroyo E, Barbujani G, Beckman G, Beckman L, Bertranpetit J, Bosch E, Bradley DG, Brede G, Cooper G, Corte-Real HB, de Knijff P, Decorte R, Dubrova YE, Evgrafov O, Gilissen A, Glisic S, Golge M, Hill EW, Jeziorowska A, Kalaydjieva L, Kayser M, Kivisild T, Kravchenko SA, Krumina A, Kucinskas V, Lavinha J, Livshits LA, Malaspina P, Maria S, McElreavey K, Meitinger TA, Mikelsaar AV, Mitchell RJ, Nafa K, Nicholson J, Norby S, Pandya A, Parik J, Patsalis PC, Pereira L, Peterlin B, Pielberg G, Prata MJ, Previdere C, Roewer L, Rootsi S, Rubinsztein DC, Saillard J, Santos FR, Stefanescu G, Sykes BC, Tolun A, Villems R, Tyler-Smith C, Jobling MA. 2000. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. Am J Hum Genet 67:1526–1543.

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

Sajantila A, Lahermo P, Anttinen T, Lukka M, Sistonen P, Savontaus ML, Aula P, Beckman L, Tranebjaerg L, Gedde-Dahl T, Issel-Tarver L, DiRienzo A, Pääbo S. 1995. Genes and languages in Europe: an analysis of mitochondrial lineages. Genome Res 5:42–52.

Saruwatari L, Zheng H-Y, Takasaka T, Sugimoto C, Sakai E, Bo Bo, Nwe Nwe Aung, Kitamura T, Yogo Y, Ohno N. 2002. Peopling of Myanmar as demonstrated by genotyping of urinary JC virus DNA. Anthropol Sci 110:235–249.

Simoni L, Calafell F, Pettener D, Bertranpetit J, Barbujani G. 2000. Geographic patterns of mtDNA diversity in Europe. Am J Hum Genet 66:262–278.

Sugimoto C, Kitamura T, Guo J, Al-Ahdal MN, Shchelkunov SN, Otova B, Ondrejka P, Chollet J-Y, El-Safi S, Ettayebi M, Grésenguet G, Kocagöz T, Chaiyarasamee S, Thant KZ, 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. Proc Natl Acad Sci USA 94:9191–9196.

Sugimoto C, Hasegawa M, Kato A, Zheng H-Y, Ebihara H, Taguchi F, Kitamura T, Yogo Y. 2002a. Evolution of human polyomavirus JC: implications for the population history of humans. J Mol Evol 54:285–297.

Sugimoto C, Hasegawa M, Zheng H-Y, Demenev V, Sekino Y, Kojima K, Honjo T, Kida H, Hovi T, Vesikari T, Schalken JA, Tomita K, Mitsunobu Y, Ikegaya H, Kobayashi N, Kitamura T, Yogo Y. 2002b. JC virus strains indigenous to northeastern Siberians and Canadian Inuits are unique but evolutionally related to those distributed throughout Europe and Mediterranean areas. J Mol Evol 55:322–335.

Suzuki M, Zheng H-Y, Takasaka T, Sugimoto C, Kitamura T, Beutler E, Yogo Y. 2002. Asian genotypes of JC virus in Japanese-Americans suggests familial transmission. J Virol 76:10074–10078.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 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.

Torroni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, Scozzari R, Obinu D, Savontaus ML, Wallace DC. 1996. Classification of European mtDNAs from an analysis of three European populations. Genetics 144:1835–1850.

Torroni A, Bandelt HJ, D'Urbano L, Lahermo P, Moral P, Sellitto D, Rengo C, Forster P, Savontaus ML, Bonne-Tamir B, Scozzari R. 1998. mtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. Am J Hum Genet 62:1137–1152.

Torroni A, Bandelt HJ, Macaulay V, Richards M, Cruciani F, Rengo C, Martinez-Cabrera V, Villems R, Kivisild T, Metspalu E, Parik J, Tolk HV, Tambets K, Forster P, Karger B, Francalacci P, Rudan P, Janicijevic B, Rickards O, Savontaus ML, Huoponen K, Laitinen V, Koivumaki S, Sykes B, Hickey E, Novelletto A, Moral P, Sellitto D, Coppa A, Al-Zaheri N, Santachiara-Benerecetti AS, Semino O, Scozzari R. 2001. A signal, from human mtDNA, of postglacial recolonization in Europe. Am J Hum Genet 69:844–852.

Walker DL, Padgett BL. 1983. The epidemiology of human polyomaviruses. Prog Clin Biolo Res 105:99–106.

Yanagihara R, Nerurkar VR, Scheirich I, Agostini HT, Mgone CS, Cui X, Jobes DV, Cubitt CL, Ryschkewitsch CF, Hrdy DB, Friedlaender JS, Stoner GL. 2002. JC virus genotypes in the western Pacific suggest Asian mainland relationships and virus association with early population movements. Hum Biol 74:473–488.

Yogo Y, Sugimoto C. 2001. The archetype concept and regulatory region rearrangement. In: Khalili K, Stoner GL, editors. Human polyomaviruses: molecular and clinical perspectives. New York: John Wiley & Sons. p 127–148.

Yogo Y, Iida T, Taguchi F, Kitamura T, Aso Y. 1991. Typing of human polyomavirus JC virus on the basis of restriction fragment length polymorphisms. J Clin Microbiol 29:2130–2138.

Yogo Y, Zheng H-Y, Hasegawa M, Sugimoto C, Tanaka S, Honjo T, Kobayashi N, Ohta N, Kitamura T. 2003. Phylogenetic analysis of JC virus DNAs detected in Ainus: an attempt to elucidate the origin and diversity of the Ainu. Anthropol Sci (Jpn Ser) 111:19–34.

Zerjal T, Dashnyam B, Pandya A, Kayser M, Roewer L, Santos FR, Schiefenhovel W, Fretwell N, Jobling MA, Harihara S, Shimizu K, Semjidmaa D, Sajantila A, Salo P, Crawford MH, Ginter EK, Evgrafov OV, Tyler-Smith C. 1997. Genetic relationships of Asians and Northern Europeans, revealed by Y-chromosomal DNA analysis. Am J Hum Genet 60:1174–1183.

Zheng H-Y, Sugimoto C, Hasegawa M, Kobayashi N, Kanayama A, Rodas A, Mejia M, Nakamichi J, Guo J, Kitamura T, Yogo Y. 2003. Phylogenetic relationships among JC virus strains in Japanese/Koreans and Native Americans speaking Amerind or Na-Dene. J Mol Evol 56:18–27.

Zheng H-Y, Kitamura T, Takasaka T, Chen Q, Yogo Y. 2004. Unambiguous identification of JC polyomavirus strains transmitted from parents to children. Arch Virol 149:261–273.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Virological Methods 126 (2005) 37–43
Journal of Virological Methods
www.elsevier.com/locate/jviromet

JC virus genotyping using formalin-fixed, paraffin-embedded renal tissues

Hiroshi Ikegaya [a, *], Hirotaro Iwase [a, b], Huai-Ying Zheng [c, 1], Makoto Nakajima [a], Koichi Sakurada [a, d], Takehiko Takatori [a, d], Masashi Fukayama [e], Tadaichi Kitamura [f], Yoshiaki Yogo [c, 1]

[a] *Department of Forensic Medicine, Graduate School of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*
[b] *Department of Legal Medicine, Faculty of Medicine, Chiba University, Chiba, Japan*
[c] *Laboratory of Viral Infection, Department of Microbiology and Immunology, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan*
[d] *National Research Institute of Police Science, Chiba, Japan*
[e] *Department of Pathology, Graduate School of Medicine, University of Tokyo, Tokyo, Japan*
[f] *Department of Urology, Faculty of Medicine, University of Tokyo, Tokyo, Japan*

Received 12 October 2004; received in revised form 18 January 2005; accepted 26 January 2005
Available online 26 February 2005

Abstract

Recently genotyping of JC virus (JCV) DNA in renal tissue was reported to be useful to identify the geographic origin of unidentified cadavers. In the above study, autopsied tissue samples without storage or stored in a frozen state were used. This study examined JCV DNA sequence modifications caused by formalin-fixation, in an attempt to elucidate whether formalin-fixed, paraffin-embedded tissue samples can also be used to determine the genotypes of JCV DNA in the kidney. In four cases, a 610 bp typing region of the JCV genome was PCR-amplified from renal tissues stored for 1 year in three different states: frozen at −80 °C [Amaker, B.H., Chandler, F.W., Huey, L.O., Colwell, R.M., 1997. Molecular detection of JC virus in embalmed, formalin-fixed, paraffin-embedded brain tissue. J. Forensic Sci., 1157–1159], formalin-fixed, paraffin-embedded [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. J. Gen. Virol. 73, 2669–2678], and soaked in 5% formalin [Baksh, F.K., Finkelstein, S.D., Swalskey, P.A., Stoner, G.L., Ryschkewitsch, C.F., Randhawa, P.R., 2001. Molecular genotyping of BK and JC virus in human polyomavirus-associated interstitial nephritis after renal transplantation. Am. J. Kidney Dis. 38 (2), 354–365]. The amplified fragments were cloned, and the resultant clones were sequenced. In frozen samples, single sequences ('original' sequences) were detected in all cases. In formalin-fixed, paraffin-embedded samples, not only the original sequences but also those with 1–6 base substitutions were detected. From formalin-soaked samples, the original sequences and those with 1–5 and 10–13 substitutions were detected. The genotyping of JCV DNA was not hampered by the presence of 1–6 substitutions, but a shift in JCV genotypes was observed in sequences with 10–13 substitutions. Thus, it was concluded that the genotypes of JCV DNA in the kidney can be determined only with specimens stored in a frozen state or formalin-fixed for a short time.

Keywords: JC virus; Genotyping; Renal tissue; Formalin; Paraffin; Base substitution

1. Introduction

Although the number of unidentified cadavers is increasing throughout the world (Cattaneo et al., 2000), there is no reliable means of tracing their origins. Recently it is reported that genotyping JC virus (JCV) DNA in the kidney or urine provides useful information as to the geographic origin of unidentified cadavers (Ikegaya et al., 2002). JCV asymptomatically infects most humans during childhood and persists in the kidneys throughout life (Chesters et al., 1983; Ikegaya et al., 2002, 2004; Kitamura et al., 1997; Padgett and Walker, 1973; Tominaga et al., 1992). JCV DNA can be classified into more than 10 genotypes that occupy unique domains in different parts of the world (Sugimoto et al., 1997; Yogo et al., 2004).

* Corresponding author. Tel.: +81 3 5841 3366. fax: +81 3 5841 3367.
E-mail address: ikegaya-tky@umin.ac.jp (H. Ikegaya).
[1] Department of Urology, Faculty of Medicine, The University of Tokyo, Tokyo, Japan.

0166-0934/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.jviromet.2005.01.019

In the study noted above, autopsied tissue samples without storage or stored frozen were used to PCR-amplify and sequence a 610 bp region of the JCV genome usually used for the genotyping (Ault and Stoner, 1992). However, there are a considerable number of unidentified cadavers for which only formalin-fixed, paraffin-embedded or formalin-soaked tissue samples are available. If JCV DNA sequences in tissue samples stored under such conditions could be accurately determined, the application of the JCV genotyping method would be expanded greatly.

Indeed, JCV DNA fragments have been amplified by PCR from formalin-fixed, paraffin-embedded tissue samples (Amaker et al., 1997; Telenti et al., 1990), and their nucleotide sequences have been determined to classify the detected JCV DNA into subtypes (Baksh et al., 2001; Stoner and Ryschkewitch, 2000). However, to our knowledge, no studies have been conducted to examine nucleotide modifications that might be introduced into JCV DNA in tissue samples during formalin-fixation before being embedded in paraffin. To clarify this issue, we examined the influence of formalin-fixation by comparing DNA sequences detected in renal tissue samples stored differently: (a) frozen at −80 °C for a year, (b) fixed in formalin for 2 weeks and then embedded in paraffin, and (c) soaked in formalin for a year.

2. Materials and methods

2.1. Tissue samples

Both kidneys were excised from eight Japanese cadavers with various causes of death (Table 1). All cadavers had no immunosuppressive diseases. The partially excised kidney blocks were stored for 1 year in three states: (a) frozen (i.e. frozen at −80 °C), (b) formalin-fixed, paraffin-embedded (i.e. embedded in paraffin at 20 °C after fixation in 5% formalin for 2 weeks), and (c) formalin-soaked (i.e. soaked in 5% formalin at 20 °C). Slices (10 mm × 10 mm × 2 mm) were excised from two different portions of the medulla of each kidney stored in each state. Paraffin-embedded samples were deparaffinized, and formalin-soaked samples were deformalinated, by the conventional methods. All samples were minced with scissors and digested with proteinase K in the presence of 0.5% sodium dodecyl sulfate at 56 °C for 16 h. The digest was extracted once with phenol and once with chloroform/isoamyl alcohol (24:1 by volume); DNA was recovered by ethanol precipitation and then dissolved in 200 μL water.

2.2. PCR

The 610 bp IG region of the JCV genome (Ault and Stoner, 1992) was PCR-amplified from the DNA extracted from renal tissue samples, using primers P1 and P2 (Kunitake et al., 1995). The total reaction volume of 50 μL contained 2.5 μL of sample DNA, 125 units of HotStar Taq DNA polymerase (QIAGEN GmbH, Hilden, Germany), 200 μM of each dNTP, 1.5 mM $MgCl_2$, 0.5 μM primers and a PCR Buffer supplied by the manufacturer. After activation at 95 °C for 15 min, amplification was performed for 50 cycles. The cycle profile was 94 °C for 1 min, 55 °C for 1 min, and 72 °C for 2 min. Both activation and amplification were carried out in a PC-800 (Astec, Fukuoka, Japan). The size of the amplicons was 666 bp. PCR was conducted twice for each DNA sample.

2.3. Cloning and sequencing

The amplified fragments were digested with a combination of *Hin*dIII and *Pst*I, which excises the IG region (Kunitake et al., 1995). The digested DNA was ligated to *Hin*dIII- and *Pst*I-digested, alkaline phosphatase-treated pBluescript II SK (+) (Stratagene, La Jo11a, USA), and used to transform *Escherichia coli* HB101-competent cells (Takara Shuzo Co. Ltd., Kyoto, Japan). Recombinant clones containing the IG region were selected by digestion with a combination of *Hin*dIII and *Pst*I followed by agarose gel electrophoresis. Purified recombinant plasmids were sequenced with an autosequencer (ABI Prism 3700 DNA Analyzer, Applied Biosystems, Foster City, CA, USA). Six clones from each DNA sample were sequenced.

2.4. Phylogenetic analysis

Neighbor-joining (NJ) phylogenetic trees (Saitou and Nei, 1987) were constructed using the CLUSTAL W Program Version 1.7 (Tompson and Higgins, 1994). Divergences were estimated by the two-parameter method (Kimura, 1980). Phylogenetic trees were visualized with TREEVIEW Program Ver-

Table 1
Eight cases studied

Case	Age (years)/gender	Post-mortem time (h)	Cause of death	Detection of JCV DNA[a]
1	65/M	24	Exsanguination	+
2	69/M	72	Frail chest	+
3	74/F	20	Pulmonary embolism	−
4	43/M	13	Exsanguination	−
5	64/M	24	General peritonitis	+
6	53/M	19	Burns	−
7	33/M	24	Traumatic shock	+
8	40/M	15	Strangulation	−

[a] See text.

sion 1.4 (Page, 1996). To assess the confidence of the NJ tree branching patterns, 1000 bootstrap replicates (Felsenstain, 1985) were performed with CLUSTAL W.

2.5. Statistical analysis

Statistical analysis was performed using the Bartlett test and the one-factor ANOVA test, with the significance level set at 5% using Microsoft Excel 2001 software.

3. Results

3.1. Detection of the IG region

As described in Section 2, we excised slices from two different portions of the medulla of each kidney. Using each slice, we performed PCR twice to amplify the IG region. Thus, we repeated PCR eight times using renal tissue samples derived from the same cadaver and stored in the same state. If amplified products of about 670 bp were obtained at least once, we regarded the renal tissue as positive for the JCV IG region. The specificity of the amplified fragments was confirmed by analyzing their sequences, as described below.

In four cases (cases 1, 2, 5, and 7), we could amplify the IG region from renal tissue samples stored in the frozen state. The tissue samples excised from these cadavers but stored in paraffin-embedded and formalin-soaked states also gave positive amplifications. In cases where the IG region was not detected in frozen samples, the IG region was detected in neither paraffin-embedded nor formalin-soaked samples. Positive amplifications were frequently obtained with only one of the duplicated DNA samples from stored in not only frozen but also paraffin-embedded and formalin-soaked tissues. Thus, as judged from the appearance of amplified fragments, renal tissue samples stored in three different states gave rise to similar amplification results.

3.2. Sequence comparisons of amplified IG fragments

We cloned the amplified fragments into a plasmid vector, and purified clones were sequenced. We sequenced six clones for each JCV-DNA positive tissue sample. In total, we identified 25 different IG sequences (SEQ-1 to -25) (Fig. 1). Table 2 shows the tissue samples in which these sequences were detected.

A single sequence was detected in the frozen tissue sample in each case. We designated this sequence as the original sequence that had not undergone base modifications. The original sequences in cases 1 and 2 were SEQ-1, and those in cases 5 and 7 were SEQ-11 and -20, respectively.

In paraffin-embedded tissue samples, not only the original but also sequences with a lower rate of base substitutions (i.e. 1–6 substitutions) were detected in all four cases. For ex-

Cases	Sequence	2151	2152	2158	2161	2168	2178	2188	2190	2210	2239	2247	2251	2264	2296	2316	2317	2341	2371	2395	2416	2431	2451	2475	2513	2514	2518	2523	2563	2593	2604	2606	2619	2642	2645	2663	2665	2680	2687	2699	2712	2723	No. of substitutions	Genotypes
1	SEQ1	A	T	A	T	A	T	A	T	A	T	G	A	A	C	T	C	C	G	T	T	A	A	A	T	T	A	A	A	A	C	T	G	G	A	A	T	T	A	T	A	C	-	MY
	SEQ2	A	T	A	T	A	T	G	T	A	T	G	A	A	C	T	C	C	G	T	T	A	A	A	T	T	A	A	A	A	C	T	G	G	A	A	T	T	A	T	A	C	1	MY
	SEQ3	A	T	A	T	A	T	A	T	A	T	G	A	A	C	T	C	C	G	T	T	A	A	A	T	T	A	A	A	A	C	T	G	G	G	A	T	T	A	T	A	C	1	MY
	SEQ4	A	T	G	T	A	T	A	C	A	G	G	G	C	A	C	T	A	A	A	T	C	A	A	A	T	T	A	A	A	A	C	C	G	A	G	A	T	T	A	T	G	T	13	CY
	SEQ5	A	T	A	C	A	C	A	T	A	G	G	G	C	A	C	T	A	C	A	T	C	A	A	A	T	T	A	A	A	A	C	C	G	A	G	A	T	T	A	C	G	T	13	CY
2	SEQ1	A	T	A	T	A	T	A	T	A	T	G	A	A	C	T	C	C	G	T	T	A	A	A	T	T	A	A	A	A	C	T	G	G	A	A	T	T	A	T	A	C	-	MY
	SEQ6	A	T	A	T	A	T	A	T	A	T	G	A	A	C	T	C	C	G	T	T	A	A	G	T	T	A	A	A	A	C	T	G	G	A	A	T	T	A	T	A	C	1	MY
	SEQ7	A	T	A	T	A	T	A	T	A	T	G	A	A	I	T	C	C	G	T	T	A	A	A	T	T	A	A	A	A	C	T	G	G	A	A	T	T	A	T	A	C	1	MY
	SEQ8	A	T	A	T	A	T	A	T	G	T	G	A	A	C	T	C	C	G	T	T	A	A	G	A	T	T	A	A	A	A	C	T	G	G	A	A	T	T	A	T	A	C	2	MY
	SEQ9	G	T	A	T	A	T	A	T	A	A	A	A	C	T	C	C	G	T	T	A	A	A	T	T	A	A	A	A	C	T	G	G	A	A	T	T	A	T	A	C	2	MY
	SEQ10	A	T	A	C	A	T	A	T	A	G	G	G	C	A	C	C	C	C	G	T	T	A	A	A	T	T	A	A	A	A	C	T	G	G	A	A	T	T	A	T	A	C	4	MY
5	SEQ11	A	T	A	T	A	T	A	T	A	T	G	A	A	C	T	C	C	G	T	T	A	A	A	T	T	A	A	A	A	C	T	T	G	A	A	T	T	A	T	A	C	-	MY
	SEQ12	A	T	A	T	A	T	A	T	A	T	G	A	A	C	T	C	C	G	G	C	T	A	A	A	T	T	A	A	A	A	C	T	T	G	A	A	T	T	A	T	A	C	1	MY
	SEQ13	A	T	A	T	A	T	A	T	G	T	G	A	A	C	T	C	C	G	T	T	A	G	A	T	T	A	A	A	A	C	T	T	G	A	A	A	T	T	A	T	A	C	2	MY
	SEQ14	A	T	G	T	A	T	A	T	A	T	G	A	G	I	T	C	C	G	T	T	A	A	A	A	C	T	A	A	A	A	C	T	T	G	A	A	T	T	A	T	A	C	5	MY
	SEQ15	A	T	G	T	A	T	A	T	A	T	G	A	G	I	T	C	C	G	T	T	A	A	A	A	C	T	A	A	G	A	C	T	T	G	A	A	T	T	A	T	A	C	5	MY
	SEQ16	A	T	G	T	A	T	A	T	A	T	G	A	G	I	T	C	C	G	T	T	A	A	A	A	C	T	A	A	G	A	C	T	T	G	A	A	T	T	A	T	-	C	6	MY
	SEQ17	A	T	G	T	A	T	A	T	A	T	G	A	G	I	T	C	C	G	T	T	A	A	A	A	C	T	A	A	G	A	C	T	T	G	A	A	T	T	G	T	A	C	6	MY
	SEQ18	A	T	A	C	A	T	A	T	A	G	G	G	C	A	C	T	A	C	A	T	C	A	A	A	T	T	A	A	A	A	C	C	T	A	G	A	T	T	A	T	G	T	10	CY
	SEQ19	A	T	A	C	A	T	A	T	A	G	G	G	C	A	C	T	A	C	A	T	C	G	A	A	T	T	A	A	A	A	C	C	T	A	G	A	T	T	A	T	G	T	12	CY
7	SEQ20	A	T	A	C	A	T	A	T	A	G	G	G	C	A	C	T	A	C	A	T	C	A	A	A	T	T	A	A	A	G	C	C	T	A	G	A	T	T	A	T	G	T	-	CY
	SEQ21	A	T	A	C	A	T	A	T	A	G	G	G	C	A	C	T	A	C	A	T	C	A	A	A	T	T	A	A	A	G	I	C	T	A	G	G	T	T	A	T	G	T	2	CY
	SEQ22	A	T	A	C	A	T	A	T	A	G	G	G	C	A	C	T	A	C	A	T	C	A	A	A	T	T	A	A	A	G	C	C	T	A	G	A	C	C	A	T	G	T	2	CY
	SEQ23	A	C	A	C	A	T	A	T	A	G	G	G	C	A	C	T	A	C	A	T	C	A	A	A	T	C	G	I	A	G	C	C	T	A	G	G	T	T	A	T	G	T	5	CY
	SEQ24	A	T	A	I	A	T	A	T	A	I	G	A	A	C	T	C	C	A	T	C	A	A	A	T	T	A	A	A	G	C	C	T	G	G	A	T	T	A	T	G	T	5	CY
	SEQ25	A	T	G	I	A	T	A	T	A	G	G	G	C	G	C	T	A	C	G	T	I	A	A	A	C	T	A	A	G	G	C	I	T	G	A	A	T	T	A	T	A	C	12	MY

Fig. 1. Nucleotide variations among IG sequences detected from renal tissue samples stored in various states. The nucleotides shown are those at positions where differences were found. Nucleotide numbering is that of strain Mad-1 (Frisque et al., 1984). Nucleotides different from those of the original sequence (shown at the top line in each case) are underlined.

Table 2
IG sequences detected in renal tissue samples stored in various states

Cases	IG sequence	No. of base substitutions	No. of clones detected from tissue samples stored in various states			Accession number[a]
			Frozen[b]	Paraffin[c]	Formalin[d]	
1	SEQ-1	–[e]	6	4	4	AB185175
	SEQ-2	1	0	1	0	AB185176
	SEQ-3	1	0	1	0	AB185177
	SEQ-4	13	0	0	1	AB185178
	SEQ-5	13	0	0	1	AB185179
2	SEQ-1	–	6	3	4	AB185175
	SEQ-6	1	0	0	1	AB185180
	SEQ-7	1	0	1	0	AB185181
	SEQ-8	2	0	0	1	AB185182
	SEQ-9	2	0	1	0	AB185183
	SEQ-10	4	0	1	0	AB185184
5	SEQ-11	–	6	2	0	AB185185
	SEQ-12	1	0	1	0	AB185186
	SEQ-13	2	0	0	2	AB185187
	SEQ-14	5	0	0	1	AB185188
	SEQ-15	5	0	1	0	AB185189
	SEQ-16	5	0	1	0	AB185190
	SEQ-17	6	0	1	0	AB185191
	SEQ-18	10	0	0	2	AB185192
	SEQ-19	12	0	0	1	AB185193
7	SEQ-20	–	6	4	3	AB185194
	SEQ-21	2	0	0	1	AB185195
	SEQ-22	2	0	0	1	AB185196
	SEQ-23	5	0	1	0	AB185197
	SEQ-24	5	0	1	0	AB185198
	SEQ-25	12	0	0	1	AB185199

[a] GSDB, DDBJ, EMBL and NCBI accession numbers.
[b] Stored at −80 °C for 1 year.
[c] Fixed in 5% formalin for 2 weeks and embedded in paraffin for 1 year.
[d] Stored in 5% formalin for 1 year.
[e] Indicates the original sequence.

Table 3
Base substitutions detected in paraffin-embedded and formalin-stored renal tissue samples

Substitution type	No. (%) of substitutions in paraffin-embedded samples[a]	No. (%) of substitutions in formalin-stored samples[b]
A → T	1 (2.9)	0 (0.0)
A → G	16 (44.4)	25 (32.9)
A → C	2 (5.6)	5 (6.6)
T → A	0 (0.0)	0 (0.0)
T → G	1 (2.8)	4 (5.3)
T → C	8 (22.2)	18 (23.7)
G → A	1 (2.8)	10 (13.2)
G → T	1 (2.8)	0 (0.0)
G → C	0 (0.0)	0 (0.0)
C → A	1 (2.8)	5 (6.6)
C → T	5 (13.8)	9 (11.8)
C → G	0 (0.0)	0 (0.0)
Total	36 (100.0)	76 (100.0)
Transitions	30 (83.3)	62 (81.6)
Transversions	6 (16.7)	14 (18.4)

Twenty-four clones were examined for each storage conditions.
[a] Fixed in 5% formalin for 2 weeks and embedded in paraffin for 1 year.
[b] Stored in 5% formalin for 1 year.

ample, the original sequence (SEQ-1) and sequences (SEQ-2 and -3) with a single substitution were found in case 1. The original sequence (SEQ-1) was identified in three clones, and sequences (SEQ-7, -9 and -10) with 1–4 substitution were found in case 2. The original sequence (SEQ-11) and sequences (SEQ-12, -15, -16 and -17) with 1–6 substitutions were found in case 5. The original sequence (SEQ-20) and sequences (SEQ-23 and -24) with five substitutions were found in case 7.

In formalin-soaked tissue samples, the original sequences were detected in three cases but not in one case. Sequences with a lower rate of base substitutions were detected in cases 2, 5 and 7 and those with a higher rate of substitutions (i.e. 10–13 substitutions), were detected in cases 1, 3 and 5. For example, the original sequence (SEQ-1) and sequences (SEQ-4 and -5) with 13 substitutions were found in case 1. The original sequence (SEQ-1) and sequences (SEQ-6 and -8) with 1 or 2 substitutions were found in case 2 (a higher rate of substitutions were not detected). The original sequence was not detected in case 5, but sequences (SEQ-13 and -14) with 2 or 5 substitutions and a sequence (SEQ-18) with 10 substitutions were found. The original sequence (SEQ-20), sequences (SEQ-21 and -22) with two substitutions and a se-

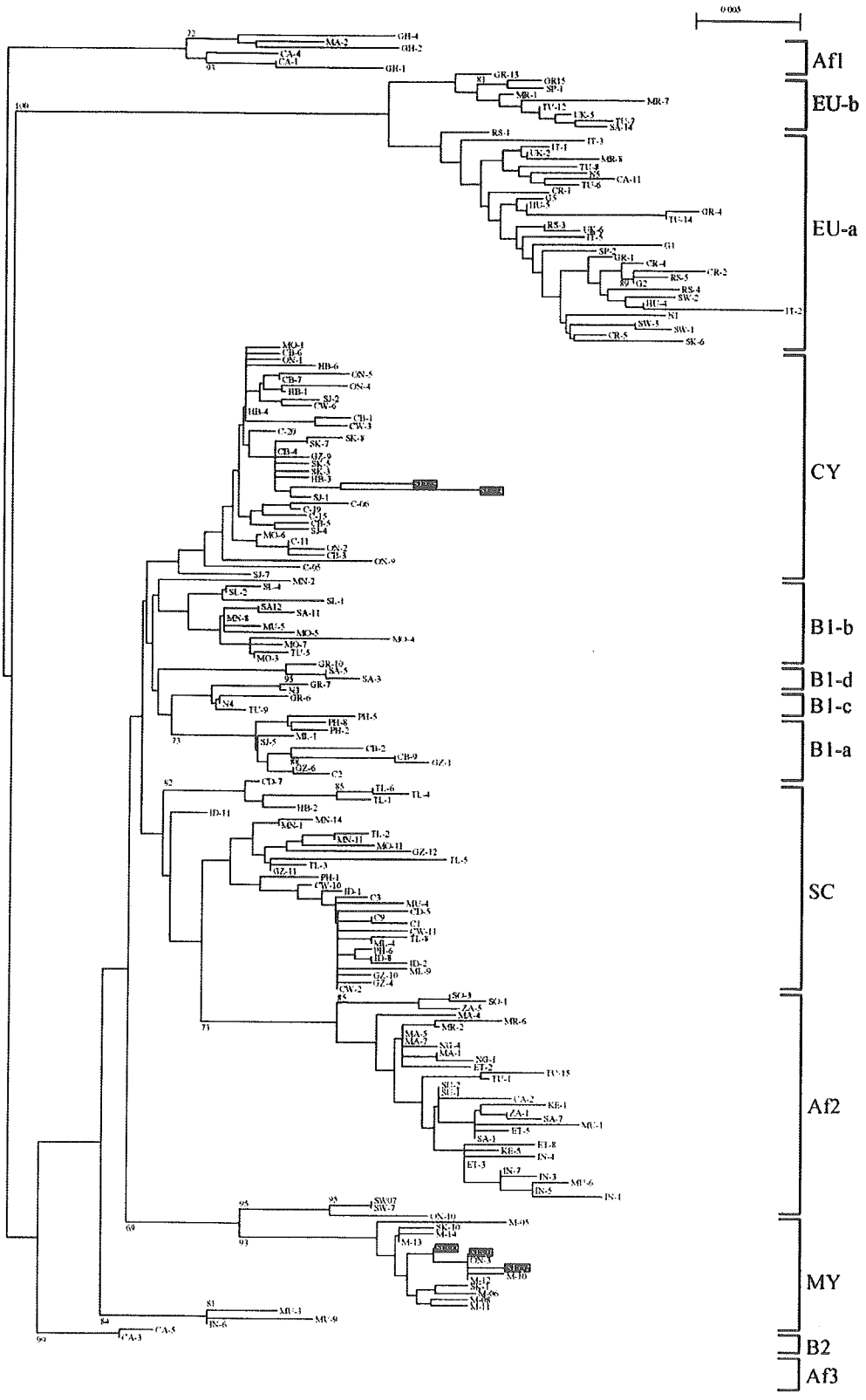

Fig. 2. An NJ phylogenetic tree showing the genotype shift detected in renal tissue samples stored in 5% formalin. An NJ phylogenetic tree was constructed from the five IG sequences detected in case 1 (see Fig. 1) and reference IG sequences reported previously (Sugimoto et al., 1997). The phylogenetic tree was visualized using TREEVIEW. The tree was rooted using genotype Af2 as the outgroup, since this genotype was suggested to indicate possible ancestral type (Pavesi, 2003). The symbols for reference sequences are described elsewhere (Sugimoto et al., 1997). The numbers at the tree nodes indicate the bootstrap confidence levels (percent) obtained with 1000 replications (only values ≥50% are shown). Genotypes are indicated on the right of the tree.

quence (SEQ-25) with 12 base substitutions were detected in case 7.

3.3. Base substitutions detected in paraffin-embedded and formalin-soaked tissue samples

The total number of base substitutions in paraffin-embedded samples was 36, while that in formalin-soaked samples was 76. This difference was statistically significant ($P < 0.05$), indicating that more base modifications occurred in formalin-soaked samples than in paraffin-embedded samples.

Table 3 lists all base substitutions detected in amplicons from paraffin-embedded and formalin-soaked tissue samples. In formalin-fixed, paraffin-embedded samples, base transitions ($n = 30$) were five times as frequent as transversions ($n = 6$), as in formalin-soaked samples.

3.4. Location of sequences with substitutions on phylogenetic trees

We examined whether the presence of lower and higher rates of base substitutions affected the genotypes of JCV. Five sequences (SEQ-1 to SEQ-5) detected in case 1 in paraffin-embedded and formalin-soaked tissue samples together with 197 IG sequences reported by Sugimoto et al. (1997) were used to construct an NJ phylogenetic tree. On the resultant trees (Fig. 2), the original sequence (SEQ-1) and sequences (SEQ-2 and -3) with a lower rate of base substitutions were found within the MY cluster, whereas sequences (SEQ-4 and -5) with a higher rate of base substitutions were located within the CY cluster. Similarly, we analyzed the distribution of sequences detected in cases 2, 5 and 7 on NJ phylogenetic trees (data not shown), and confirmed that the original sequences and those with a lower rate of base substitutions fell within the same genotypes (genotype MY in cases 2 and 5 and CY in case 7) (Fig. 2), and that sequences (SEQ-18, -19, -25) with a higher rate of substitutions were located within genotypes to which the original sequence did not belong (Fig. 2).

4. Discussion

For the first time, the effect of formalin-fixation on the nucleotide sequence of JCV DNA in tissue samples was analyzed in this study. The JCV DNA sequence detected in tissue samples stored in a frozen state served as the source of unaffected JCV DNA sequences (designated original ones). Original and variant sequences were detected in tissue samples that underwent formalin-fixation. A comparison between original and variant sequences allowed us to detect base substitutions.

When JCV DNA with some bases modified with formalin is used as the template for PCR amplification, the modified bases were recognized as one of the four bases, sometimes as a base different from the original one, generating base substitutions that can be identified by sequencing. The present finding that base transitions much more frequently occurred than transversions in paraffin-embedded as well as formalin-soaked tissue samples suggests that Taq DNA polymerase can distinguish between modified purines and pyrimidines with some ambiguity.

It has been reported that as the time of exposure to formalin increased, the PCR-amplification of DNA was more severely damaged (Inoue et al., 1996; Rogers et al., 1990). In this study, the paraffin-embedded tissue samples were previously fixed with formalin for 2 weeks, while the formalin-soaked samples were fixed for 1 year. Thus, based on our results, we may state that the duration of fixation affected the appearance of base substitutions. More substitutions were found after long-term fixation than after short-term fixation. As a result, we detected IG sequences with a higher rate of base substitutions (i.e. 10–13 substitutions per IG sequence) only in tissue samples undergoing long-term fixation. We found a genotype shift in sequences with a higher rate of substitutions. In contrast, we did not detect a higher rate of substitutions in tissue samples undergoing short-term fixation, but instead a lower rate of substitutions (i.e. 1–6 substitutions per IG sequence) was detected. Sequences with a lower rate of substitutions did not cause a genotype shift.

In summary, it was found that a significant number of base substitutions occurred in viral DNA sequences amplified from formalin-fixed, paraffin-embedded renal tissues, and that the rate of base substitutions increased in tissues soaked in formalin for a prolonged time (1 year). Thus, it was concluded that the genotypes of JCV DNA in the kidney can be determined only with specimens stored in a frozen state or formalin-fixed for a short time (2 weeks). The present findings also suggest that special care should be paid to the diagnosis of viral diseases using formalin-fixed, paraffin-embedded tissues for which formalin-fixation time is unknown.

References

Amaker, B.H., Chandler, F.W., Huey, L.O., Colwell, R.M., 1997. Molecular detection of JC virus in embalmed, formalin-fixed, paraffin-embedded brain tissue. J. Forensic Sci., 1157–1159.

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. J. Gen. Virol. 73, 2669–2678.

Baksh, F.K., Finkelstein, S.D., Swalskey, P.A., Stoner, G.L., Ryschkewitsch, C.F., Randhawa, P.R., 2001. Molecular genotyping of BK and JC virus in human polyomavirus-associated interstitial nephritis after renal transplantation. Am. J. Kidney Dis. 38 (2), 354–365.

Cattaneo, C., Ritz-Timme, S., Shutz, H.W., Collins, M., Waite, E., Boormann, H., Grandi, M., Kaatsch, H.J., 2000. Unidentified cadavers and human remains in the EU: an unknown issue. Int. J. Legal Med. 113, N2–N3.

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. J. Infect. Dis. 147, 676–684.

Felsenstain, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Frisque, R.J., Bream, G.L., Cannella, M.T., 1984. Human polyomavirus JC virus genome. J. Virol. 51 (2), 458–469.

Ikegaya, H., Iwase, H., Sugimoto, C., Yogo, Y., 2002. JC virus genotyping offers a new means of tracing the origins of unidentified cadavers. Int. J. Legal Med. 116, 242–245.

Ikegaya, H., Iwase, H., Yogo, Y., 2004. Detection of identical JC virus DNA sequences in both human kidneys. Arch. Virol. 149, 1215–1220.

Inoue, T., Nabeshima, K., Kataoka, H., Kudo, M., 1996. Feasibility of archival non-buffered formalin-fixed and paraffin-embedded tissues for PCR amplification: an analysis of resected gastric carcinoma. Pathol. Int. 46, 997–1004.

Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. E 16, 111–120.

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. J. Clin. Microbiol. 35, 1255–1257.

Kunitake, T., Kitamura, T., Guo, J., Tominaga, T., Kawabe, K., Yogo, Y., 1995. Parent-to-child transmission is relatively common in the spread of the human polyomavirus JC virus. J. Clin. Microviol. 33, 1448–1451.

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. J. Infect. Dis. 127, 467–470.

Page, R.D., 1996. TREEVIEW: an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. 12, 357–358.

Pavesi, A., 2003. African origin of polyomavirus JC and implications for prehistoric human migrations. J. Mol. E 56, 564–572.

Rogers, B.B., Alpert, L.C., Hine, E.A.S., Buffone, G.J., 1990. Analysis of DNA in fresh and formalin fixed tissue by the polymerase chain reaction. Am. J. Pathol. 136 (3), 541–548.

Saitou, N., Nei, M., 1987. The neighbor joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. E 4, 406–425.

Stoner, G.L., Ryschkewitch, C.F., 2000. Reappraisal of progressive multifocal leuko-encephalopathy due to simian virus 40.1998. Acta Neuropathol. 96, 271–278.

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. Proc. Natl. Acad. Sci. U.S.A. 94, 9191–9196.

Telenti, A., Aksamit, A.J., Proper, J., Smith, T.F., 1990. Detection of JC virus DNA by polymerase chain reaction in patients with progressive multifocal leukoencephalopathy. J. Infect. Dis. 162, 858–861.

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.

Tompson, J.D., Higgins, D.G., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl. Acids Res. 22, 4673–4680.

Yogo, Y., Sugimoto, C., Zheng, H.-Y., Ikegaya, H., Takasaka, T., Kitamura, T., 2004. JC virus genotyping offers a new paradigm in the study of human populations. Rev. Med. Virol. 14, 179–191.

Journal of General Virology (2006), **87**, 303–306 DOI 10.1099/vir.0.81368-0

Short Communication

Stability of the BK polyomavirus genome in renal-transplant patients without nephropathy

Tomokazu Takasaka,[1] Nobuyuki Goya,[2] Hideki Ishida,[2] Kazunari Tanabe,[2] Hiroshi Toma,[2] Tomoaki Fujioka,[3] So Omori,[3] Huai-Ying Zheng,[1,4] Qin Chen,[1] Souichi Nukuzuma,[5] Tadaichi Kitamura[1] and Yoshiaki Yogo[1]

Correspondence
Yoshiaki Yogo
yogo-tky@umin.ac.jp

[1]Department of Urology, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan

[2]Department of Urology, Tokyo Women's Medical University, Tokyo, Japan

[3]Department of Urology, Iwate Medical University School of Medicine, Morioka, Japan

[4]Japanese Foundation for AIDS Prevention, Tokyo, Japan

[5]Kobe Institute of Health, Kobe, Japan

To clarify the stability of the BK polyomavirus (BKPyV) genome in renal transplant (RT) recipients, three to five complete BKPyV genomes from each of six RT recipients with surviving renal allografts were molecularly cloned. The complete sequences of these clones were determined and compared in each patient. No nucleotide difference was detected among clones in two patients, and a few nucleotide variations were found among those in four patients. In each of these patients a parental sequence (usually the major sequence), from which variant sequences (usually minor sequences) with nucleotide substitutions would have been generated, were identified. A comparison between the parental and variant sequences in each patient identified a single nucleotide substitution in each variant sequence. From these findings, it was concluded that the genome of BKPyV is stable in RT recipients without nephropathy, with only minor nucleotide substitutions in the coding region.

Received 22 July 2005
Accepted 4 November 2005

BK polyomavirus (BKPyV) is ubiquitous in human populations, infecting children and persisting in the kidney. The renal BKPyV is reactivated in immunocompromised patients, usually without obvious symptoms. A relatively small proportion of renal transplant (RT) recipients, however, develop polyomavirus-associated nephropathy (PVAN) (de Bruyn & Limaye, 2004). Although host factors (e.g. immunological conditions) are important in the pathogenesis of PVAN (de Bruyn & Limaye, 2004), the genetic change of BKPyV may also involve the progression of the disease. Indeed, several authors have reported sequence rearrangements in the transcriptional control region (TCR) of the BKPyV genome derived from diseased tissue (Chen *et al.*, 2001; Randhawa *et al.*, 2003). Furthermore, genetic variations in the BKPyV VP1 gene have been detected in bioptic renal specimens from patients with PVAN (Baksh *et al.*, 2001; Randhawa *et al.*, 2002). Nevertheless, the implications of these genetic changes for the pathogenesis of PVAN remain unclear.

Chen *et al.* (2004) recently investigated the genetic variability of BKPyV *in vivo*. They sequenced several full-length viral DNA clones obtained, using PCR, from the striated muscle and heart of a patient with BKPyV-associated capillary leak syndrome ($BKPyV_{CAP}$) and also from the urine of one human immunodeficiency virus type 2-positive subject ($BKPyV_{HI}$) and one healthy control subject ($BKPyV_{HC}$). They detected a high degree of variation (mean difference, 0·29 % per site) in the coding region among clones in $BKPyV_{CAP}$, and a lower, but still remarkable, degree of variation (mean difference, 0·1–0·2 % per site) among clones in $BKPyV_{HI}$ or $BKPyV_{HC}$. Non-synonymous nucleotide substitutions (i.e. those resulting in amino acid substitutions) were frequently observed in all subjects. In addition, the authors did not identify a potentially prototypal sequence that might have generated the variant sequences detected in each subject.

The findings reported by Chen *et al.* (2004) presented a sharp contrast to observations made with *JC polyomavirus* (JCPyV), a virus closely related to BKPyV. Using the standard method of cloning that allows one to obtain intact viral DNA molecules (Sambrook *et al.*, 1989), Zheng *et al.* (2004) established and sequenced five to nine complete JCPyV DNA clones in each of 11 healthy individuals (parents and children in five families), and compared the resultant sequences in each individual. Variations in the

The GenBank/EMBL/DDBJ accession numbers of the sequences reported in this paper are AB217917–AB217921.

Table 1. RT patients analysed in this study

All patients received a transplanted kidney from a living related donor.

Patient	Gender, age (years)	Time after transplant (months)	Serum creatinine (mg dl^{-1})*
1	F (51)	4	1·4
2	M (30)	13	1·5
3	M (42)	82, 83	1·4
4	F (36)	12	1·7
5	M (55)	3	1·1
6	M (49)	7	1·3

*To convert to μmol l^{-1} multiply by 88.

coding region were detected in six individuals, but not in five individuals. The detected variations were mostly single-nucleotide substitutions, and only three of 10 nt substitutions caused amino acid substitutions. Furthermore, the authors detected possible prototypal sequences at the nodes of family specific clusters of phylogenetic trees.

We examined the stability of the BKPyV genome in RT patients without PVAN, as a basis of future studies analysing possible genetic changes of BKPyV associated with the pathogenesis or progression of PVAV (see above). We established multiple full-sized BKPyV DNA clones from the urine of each of six RT patients with surviving renal allografts, by using the standard method of molecular cloning (Sambrook *et al.*, 1989). In each patient, three to five complete BKPyV DNA clones were sequenced and the resultant sequences were compared in each patient.

RT patients analysed in this study are shown in Table 1. Entire BKPyV DNAs were cloned into pUC19 at the unique *Bam*HI site by the standard method (Sambrook *et al.*, 1989) as described previously (Yogo *et al.*, 1991). The complete BKPyV DNA clones were prepared using a Qiagen Plasmid Maxi kit. Purified plasmids were sequenced as described

Table 2. Complete BKPyV DNA sequences detected in patients 1–6

Nucleotides (amino acids) at positions of the Dunlop genome (Seif *et al.*, 1979) are shown, when differences were detected among sequences from the same patients. The numbers of clones with individual sequences and the GenBank/EMBL/DDBJ accession numbers are also shown. Six clones (TW-1, TW-3, TW-4, TW-5, TW-8 and THK-9) will be reported elsewhere (T. Takasaka and others, unpublished data).

Sequence	Subtype*	nt 1744†	nt 1807†	No. clones	Accession no.‡
Patient 1					
TW-1§	Ic	G (E)	G (E)	2	AB211381
TW-1a	Ic	G (E)	C (Q)	1	AB217917
TW-1b	Ic	A (K)	G (E)	1	AB217918
		nt 410‖			
Patient 2					
TW-3	IV	T (L)	–	1	AB211391
TW-3a§	IV	G (R)	–	4	AB217919
Patient 3					
TW-4		–	–	3	AB211382
Patient 4					
TW-5	Ic	–	–	3	AB211383
		nt 1156¶			
Patient 5					
TW-8§	Ic	C (S)	–	3	AB211385
TW-8a	Ic	A (stop codon)	–	1	AB217920
		nt 1127¶			
Patient 6					
THK-9§	Ic	T (V)	–	2	AB211379
THK-9a	Ic	G (V)	–	1	AB217921

*Previously classified using phylogenetic analysis (T. Takasaka and others, unpublished data).
†Located in the VP1 gene.
‡GenBank/EMBL/DDBJ accession numbers.
§Ancestral sequence (see text).
‖Located in the agnogene.
¶Located in the VP2/3 gene.

elsewhere (T. Takasaka and others, unpublished data). The determined and reference sequences were aligned using the CLUSTAL W program (Thompson *et al.*, 1994). Translation of nucleotide sequences into amino acid sequences was performed with GENETYX-MAC version 11.10 (Genetyx). A neighbour-joining (NJ) phylogenetic tree (Saitou & Nei, 1987) was constructed using the CLUSTAL W program (Thompson *et al.*, 1994). Divergences were estimated with the two-parameter method (Kimura, 1980). The phylogenetic tree was visualized using TREEVIEW (Page, 1996). The confidence of branching patterns of the NJ trees was assessed based on 1000 bootstrap replicates (Felsenstein, 1985).

When we recently established full-sized BKPyV DNA clones from the urine of RT patients, we obtained multiple clones in six patients, but sequenced only a single clone for each patient (T. Takasaka and others, unpublished data). In the present study, we determined the clones left unsequenced. A comparison of the sequences determined in this and the other study (T. Takasaka and others, unpublished data) identified one to three unique sequences in each patient (Table 2). (i) Three sequences, TW-1, TW-1a and TW-1b, were detected in patient 1. In reference to TW-1, TW-1a and TW-1b carried single nucleotide substitutions at nt 1807 and 1744, respectively, within the VP1 gene [nucleotide numbers are those of the strain Dunlop (Seif *et al.*, 1979)]. Both nucleotide substitutions resulted in amino acid substitutions within a predicted outer loop (the BC loop) of the VP1 protein (Chang *et al.*, 1996) (Table 2). (ii) Two sequences, TW-3 and TW-3a, were detected in patient 2. In reference to TW-3, TW-3a carried a single nucleotide substitution at nt 410 within the agnogene. This nucleotide substitution resulted in an amino acid substitution. (iii) A single sequence, TW-4, was detected in patient 3. (iv) Similarly, a single sequence, TW-5, was detected in patient 4. (v) Two sequences, TW-8 and TW-8a, were detected in patient 5. In reference to TW-8, TW-8a carried a single nucleotide substitution at nt 1156 within the VP2/3 gene. This nucleotide substitution inserted a stop codon. (vi) Finally, two sequences, THK-9 and THK-9a, were detected in patient 6. In reference to THK-9, THK-9a carried a single nucleotide substitution at nt 1127 within the VP2/3 gene. This nucleotide substitution did not result in any amino acid substitution.

The TCR sequences of representative BKPyV clones in patients 1–6 were previously clarified, and all had the archetypal configuration (Takasaka *et al.*, 2004). In the

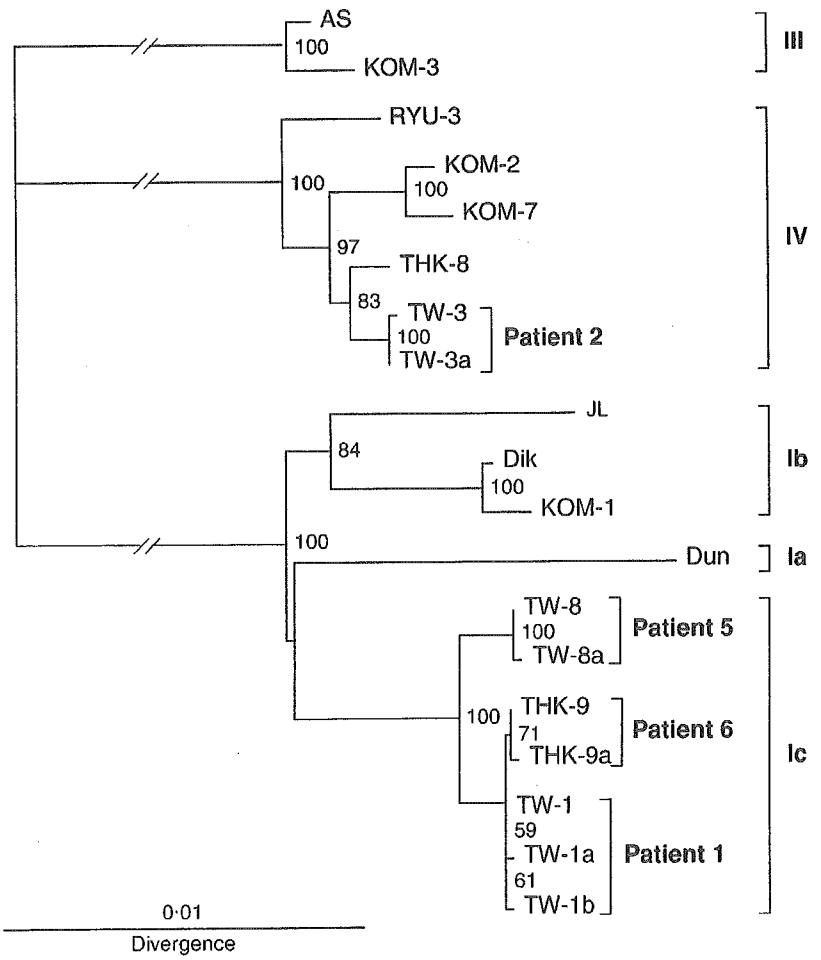

Fig. 1. NJ phylogenetic tree relating complete BKPyV DNA sequences detected in patients 1, 2, 5 and 6. An NJ phylogenetic tree was constructed from nine complete BKPyV sequences detected in patients 1, 2, 5 and 6, and 10 complete BKPyV sequences detected in unrelated individuals and belonging to subtype Ia, Ib, III or IV (Seif *et al.*, 1979; Tavis *et al.*, 1989; T. Takasaka and others, unpublished data) (the non-coding regulatory region of the BKPyV genome was excluded from this phylogenetic analysis). The phylogenetic tree was visualized using TREEVIEW. The tree was rooted using subtype III isolates (i.e. AS and KOM-3) as the outgroup. The numbers at nodes in the tree indicate the bootstrap confidence levels (percentage) obtained with 1000 replications (only values $\geqslant$50% are shown). Subtypes Ia, Ib, Ic, III and IV and sequences detected in patients 1, 2, 5 and 6 are indicated.

present study, we did not detect any difference in the TCR among clones derived from each patient.

To elucidate the evolutionary relationships among several unique sequences detected in patients 1, 2, 5 and 6, we constructed an NJ phylogenetic tree from the BKPyV DNA sequences detected in these patients together with reference sequences reported previously (Seif *et al.*, 1979; Tavis *et al.*, 1989; T. Takasaka and others, unpublished data). On the resultant tree (Fig. 1), the BKPyV DNA sequences in patients 1, 2, 5 and 6 formed individual clusters. We detected a sequence (TW-1, TW-3a, TW-8 or THK-9) at the node of each cluster, probably representing the prototypal sequence that generated variant sequences in each patient. It may be worth noting that the prototypal sequences were usually the major ones (Table 2).

In this study, we investigated the stability of the BKPyV genome in RT recipients without PVAN. We found that the genome of BKPyV is rather stable in these patients, with only minor nucleotide substitutions in the coding region. This finding forms the basis of further study on the genetic changes possibly involved in the progression of PVAN.

The finding noted immediately above is contradictory to a high intra-strain genetic diversity in BKPyV suggested by Chen *et al.* (2004). This discrepancy may be related to the difference in the methods used to obtain a full-sized genome. We used standard molecular cloning (Sambrook *et al.*, 1989), whereas Chen *et al.* (2004) used PCR amplification. Standard molecular cloning warrants the isolation of intact complete viral genomes, while PCR amplification inevitably involves replication errors, even though the frequency of errors may be reduced by using a thermostable DNA polymerase with proofreading activity. The frequently detected variations in BKPyV (Chen *et al.*, 2004) (see above) could have been introduced by the authors during PCR. Nevertheless, it remains to be elucidated whether the BKPyV genome undergoes a high variability in a specific disease, i.e. $BKPyV_{CAP}$.

In this study, we detected four non-synonymous nucleotide substitutions and one synonymous substitution. Of the four non-synonymous substitutions, three resulted in amino acid changes in VP1 and the agnoprotein, and one generated an incomplete VP2/3 protein due to the insertion of a stop codon. While viruses with incomplete VP2/3 proteins may not be infectious, it remains to be elucidated whether the amino acid changes in VP1 and the agnoprotein cause alterations in the properties of BKPyV.

Acknowledgements

This study was supported in part by grants from the Ministry of Health, Labor and Welfare, Japan.

References

Baksh, F. K., Finkelstein, S. D., Swalsky, P. A., Stoner, G. L., Ryschkewitsch, C. F. & Randhawa, P. (2001). Molecular genotyping of BK and JC viruses in human polyomavirus-associated interstitial nephritis after renal transplantation. *Am J Kidney Dis* **38**, 354–365.

Chang, D., Liou, Z.-M., Ou, W.-C., Wang, K.-Z., Wang, M., Fung, C.-Y. & Tsai, R.-T. (1996). Production of the antigen and the antibody of the JC virus major capsid protein VP1. *J Virol Methods* **59**, 177–187.

Chen, C.-H., Wen, M.-C., Wang, M., Lian, J.-D., Wu, M.-J., Cheng, C.-H., Shu, K.-H. & Chang, D. (2001). A regulatory region rearranged BK virus is associated with tubulointerstitial nephritis in a rejected renal allograft. *J Med Virol* **64**, 82–88.

Chen, Y., Sharp, P. M., Fowkes, M., Kocher, O., Joseph, J. T. & Koralnik, I. J. (2004). Analysis of 15 novel full-length BK virus sequences from three individuals: evidence of a high intra-strain genetic diversity. *J Gen Virol* **85**, 2651–2663.

de Bruyn, G. & Limaye, A. P. (2004). BK virus-associated nephropathy in kidney transplant recipients. *Rev Med Virol* **14**, 193–205.

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

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

Page, R. D. M. (1996). TREEVIEW: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* **12**, 357–358.

Randhawa, P. S., Khaleel-Ur-Rehman, K., Swalsky, P. A., Vats, A., Scantlebury, V., Shapiro, R. & Finkelstein, S. (2002). DNA sequencing of viral capsid protein VP-1 region in patients with BK virus interstitial nephritis. *Transplantation* **73**, 1090–1094.

Randhawa, P., Zygmunt, D., Shapiro, R., Vats, A., Weck, K., Swalsky, P. & Finkelstein, S. (2003). Viral regulatory region sequence variations in kidney tissue obtained from patients with BK virus nephropathy. *Kidney Int* **64**, 743–747.

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

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning, a Laboratory Manual*, 2nd edn. Cold Spring Harbor: Cold Spring Harbor Laboratory.

Seif, I., Khoury, G. & Dhar, R. (1979). The genome of human papovavirus BKV. *Cell* **18**, 963–977.

Takasaka, T., Goya, N., Tokumoto, T. & 14 other authors (2004). Subtypes of BK virus prevalent in Japan and variation in their transcriptional control region. *J Gen Virol* **85**, 2821–2827.

Tavis, J. E., Walker, D. L., Gardner, S. D. & Frisque, R. J. (1989). Nucleotide sequence of the human polyomavirus AS virus, an antigenic variant of BK virus. *J Virol* **63**, 901–911.

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 Res* **22**, 4673–4680.

Yogo, Y., Iida, T., Taguchi, F., Kitamura, T. & Aso, Y. (1991). Typing of human polyomavirus JC virus on the basis of restriction fragment length polymorphisms. *J Clin Microbiol* **29**, 2130–2138.

Zheng, H. Y., Kitamura, T., Takasaka, T., Chen, Q. & Yogo, Y. (2004). Unambiguous identification of JC polyomavirus strains transmitted from parents to children. *Arch Virol* **149**, 261–273.