in Table 8. Eight of the 11 nucleotide substitutions accompanied no amino acid change, but 3 caused amino acid substitutions.

Rearrangements (deletions) in FA-1a and FC-1a were within regions encoding viral proteins, VP2/3 and LT, respectively (Tables 3 and 5). As the deletion in FA-1a caused no frame shift, FA-1a could produce a shorter VP2/3 lacking C-terminal 21 amino acids. On the other hand, the deletion in FC-1a caused a frame shift in the first exon of LT, and therefore FC-1a would have produced a truncated LT containing only 12 N-terminal amino acids.

Discussion

In this study, we sequenced many complete JCV DNA clones in both parents and children between whom JCV was very probably transmitted. We detected multiple JCV DNA sequences in 4 of the 5 parents and in 3 of the 6 children. It is unlikely that these variations in JCV DNA sequences were introduced during molecular cloning, because of the high fidelity of DNA synthesis in bacterial cells. Phylogenetic analysis suggested that multiple JCV DNA sequences detected in each family are related to each other. As superinfection by a different JCV strain is rather rare [21], it is likely that in each family, one of the detected sequences is the prototypical one from which the others were generated by nucleotide substitution or sequence rearrangement. Although we could not definitely determine which sequence is actually prototypical in each family, it is possible that the sequences at the nodes of family-specific clusters represented the prototypical ones (Figs. 1 and 2).

On the basis of the present data shown in Tables 3, 4, 6, and 7, we could identify JCV DNA sequences that were transmitted from parents to children in families A, B, D, and E. (i) Only a single JCV DNA sequence (FA-1) was detected in the parent of family A, and this sequence was mainly found in two children. Therefore, it was inferred that a JCV strain containing FA-1 was transmitted to 2 children of family A. (ii) Three sequences (FB-1, -2, and -3) were detected in the parent, while only a single sequence (FB-2) was found in the child. Therefore, it was inferred that FB-2 was transmitted from the parent to the child. (iii) Four sequences (FD-1, -1a, -2, and -3) were detected in the parent (Table 6), while only a single sequence (FD-1) was found in the child (Table 6). Therefore, it was inferred that FD-1 was transmitted from the parent to the child. (iv) FE-1 and -2 were detected in the parent of family E, while FE-1 and -3 detected in the child. Therefore, it was inferred that FE-1 was transmitted from the parent to the child.

In family C, however, we could not identify a common sequence in the parent and the child. Six sequences (FC-2, -3, -3a, -4, 4a, and -5) were detected in the parent of family C, while 2 sequences (FC-1 and -1a) were found in the child. According to phylogenetic analysis (Fig. 2), FC-1 was found at the node of the family-C cluster. Although FC-1 was not detected in the urine sample collected from the parent in this study, we may assume that the parent carried FC-1 when he cohabited with his child (this strain may have disappeared in the parent because

of frequent changes in the genome). If this assumption can be accepted, we may infer that FC-1 was transmitted to the child of family C.

To examine the possibility that JCV is transmitted preferentially within the family, Kunitake et al. [24] PCR-amplified the 610-bp IG region from urine specimens collected from all members of 7 families. JCV strains were identified by the nucleotide sequences of the amplified IG regions. From the obtained results, the authors concluded that JCV is transmitted frequently from parents to children. Nevertheless, since JCV strains with identical IG sequences often occur in the same geographical region [8, 14, 23, 27, 33, 34], the detection does not necessarily indicate that a JCV strain was transmitted from the parent to the child. In this study, we detected a completely identical JCV DNA sequence in both the parent and the child belonging to each of the four families. We thus provided support for the proposed mode of JCV transmission, i.e. parent-to-child transmission during cohabitation.

Acknowledgement

We are grateful to all urine donors. This study was supported in part by grants from the Ministry of Health, Labor and Welfare, Japan.

References

1. 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
2. Agostini HT, Ryschkewitsch CF, Brubaker GR, Shao J, Stoner GL (1997) Five complete genomes of JC virus Type 3 from Africans and African Americans. Arch Virol 142: 637–655
3. Agostini HT, Ryschkewitsch CF, Stoner GL (1998) JC virus Type 1 has multiple subtypes: three new complete genomes. J Gen Virol 79: 801–805
4. Agostini HT, Shishido-Hara Y, Baumhefner RW, Singer EJ, Ryschkewitsch CF, Stoner GL (1998) JC virus Type 2: definition of genotypes based on DNA sequence analysis of ten complete genomes. J Gen Virol 79: 1143–1151
5. Agostini HT, Ryschkewitsch CF, Stoner GL (1998) Complete genome of a JC virus subtype Type 6 from the brain of an African American with progressive multifocal leukoencephalopathy. J Hum Virol 1: 267–272
6. Agostini HT, Jobes DV, Stoner GL (2001) Molecular evolution and epidemiology of JC virus. In: Khalili K, Stoner GL (eds) Human polyomaviruses: molecular and clinical perspectives. John Wiley & Sons, New York, pp 491–526
7. 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
8. Chang D, Sugimoto C, Wang M, Tsai RT, Yogo Y (1999) JC virus genotypes in a Taiwan aboriginal tribe (Bunun): implications for its population history. Arch Virol 144: 1081–1090
9. 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

10. Daniel R, Shah K, Madden D, Stagno S (1981) Serological investigation of the possibility of congenital transmission of papovavirus JC. Infect Immunol 33: 319–321
11. Dörries K (2001) Latent and persistent polyomavirus infection. In: Khalili K, Stoner GL (eds) Human polyomaviruses: molecular and clinical perspectives. John Wiley & Sons, New York, pp 197–235
12. Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783–791
13. Frisque RJ, Bream GL, Cannella MT (1984) Human polyomavirus JC virus genome. J Virol 51: 458–469
14. Guo Z, Zheng SP, Sugimoto C, Wang YL, Zheng HY, Takasaka T, Kitamura T, Guo J, Yogo Y (2001) JC virus genotypes in northwestern China: Implications for its population history. Anthropol Sci 109: 203–212
15. Hills DM, Bull JJ (1993) An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst Biol 42: 182–192
16. Iida T, Kitamura T, Guo J, Taguchi F, Aso Y, Nagashima K, Yogo Y (1993) Origin of JC polyomavirus variants associated with progressive multifocal leukoencephalopathy. Proc Natl Acad Sci USA 90: 5062–5065
17. 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
18. Kato A, Sugimoto C, Zheng HY, 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
19. 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
20. 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
21. 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
22. 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
23. Kitamura T, Sugimoto C, Ebihara H, Kato A, Guo J, Taguchi F, Tominaga T, Ogawa Y, Ohta N, Kizu N, Imamura K, Funai H, Kurosawa T, Ichikawa S, Suzuki T, Chiba K, Nagashima K, Yasumoto S, Yogo Y (1998) Peopling of Japan as revealing by genotyping of urinary JC virus DNA. Anthropol Sci 106: 311–325
24. 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
25. Loeber G, Dörries K (1988) DNA rearrangements in organ-specific variants of polyomavirus JC strain GS. J Virol 62: 1730–1735
26. Major EO (2001) Human polyomavirus. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA, Roizman B, Straus SE (eds) Fields virology, 4th ed. Lippincott Williams & Wilkins, Philadelphia, Penn, pp 2175–2196
27. Miranda JJ, Sugimoto C, Paraguison R, Takasaka T, Zheng HY, 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
28. Nathanson N (1996) Epidemiology. In: Fields BN, Knipe DM, Howley PM (eds) Fields virology, vol 1, 3rd edn. Lippincott-Raven, New York, pp 251–271

29. Padgett BL, Walker DL, ZuRhein GM, Eckroade RJ, Dessel BH (1971) Cultivation of papova-like virus from human brain with progressive multifocal leucoencephalopathy. Lancet i: 1257–1260
30. 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
31. Page RDM (1996) TREEVIEW: An application to display phylogenetic trees on personal computers. Comp Appl Biosci 12: 357–358
32. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406–425
33. Saruwatari L, Sugimoto C, Kitamura T, Ohno N, Sakai E, Shresta P, Hoa BK, Phi PTP, An HPH, Tuyet NTA, Honjo T, Kobayashi N, Zheng HY, Takasaka T, Yogo Y (2002) Asian domains of four major genotypes of JC virus, Af2, B1-b, CY and SC. Arch Virol 147: 1–10
34. Sugimoto C, Kitamura T, Guo J, Al-Ahdal MN, Shchelkunov SN, Otova B, Ondrejka P, Chollet JY, Safi SEl, 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
35. Sugimoto C, Hasegawa M, Kato A, Zheng HY, Ebihara H, Taguchi F, Kitamura T, Yogo Y (2002) Evolution of human polyomavirus JC: implications for the population history of humans. J Mol Evol 54: 285–297
36. Suzuki M, Zheng HY, Takasaka T, Sugimoto C, Kitamura T, Beutler E, Yogo Y (2002) Asian genotypes of JC virus in Japanese-Americans suggest familial transmission. J Virol 76: 10074–10078
37. Taguchi F, Kajioka J, Miyamura T (1982) Prevalence rate and age of acquisition of antibodies against JC virus and BK virus in human sera. Microbiol Immunol 26: 1057–1064
38. 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
39. 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
40. Walker DL, Padgett BL (1983) The epidemiology of human polyomaviruses. Prog Clin Biol Res 105: 99–106
41. 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. J Virol 64: 3139–3143
42. 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
43. Yogo Y, Sugimoto C (2001) The archetype concept and regulatory region rearrangement. In: Khalili K, Stoner GL (ed) Human polyomaviruses: molecular and clinical perspectives. John Wiley & Sons, New York, pp 127–148
44. Zheng HY, 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

Author's address: Dr. Huai-Ying Zheng, Department of Urology, Faculty of Medicine, The University of Tokyo, Bunkyo-ku, 113-8655 Tokyo, Japan; e-mail: teizhy-tky@umin.ac.jp

Stability of JC Virus Coding Sequences in a Case of Progressive Multifocal Leukoencephalopathy in Which the Viral Control Region Was Rearranged Markedly

Huai-Ying Zheng, PhD; Yukiharu Yasuda, MD; Shunichi Kato, MD; Tadaichi Kitamura, MD; Yoshiaki Yogo, PhD

• ***Context.***—**It is generally accepted that JC virus variants in the brains of patients with progressive multifocal leukoencephalopathy are generated from archetypal strains through sequence rearrangement (deletion and duplication, or deletion alone) in the control region. This change is thought to occur during persistence of JC virus in patients.**

Objective.—**The present study was performed to ascertain whether amino acid substitution in the viral proteins is involved in the generation and propagation of JCV variants with rearranged control regions.**

Design.—**Many complete JC DNA clones were established from brain tissues (cerebellum, occipital lobe, and brainstem) autopsied in a case of progressive multifocal leukoencephalopathy in which multiple distinct control sequences were detected. Control and coding sequences were determined and compared among the JC DNA clones.**

Results.—**Twenty-eight control-region and 20 coding sequences of JC virus were compared. Five rearranged control sequences were detected, but they could be classified into 3 groups that shared common structural features. Viral coding sequences were identical among clones with different control regions and among clones derived from different brain regions.**

Conclusion.—**In the present case, nucleotide substitution in the viral coding regions (and resultant amino acid change in the viral proteins) was involved neither in the genesis of rearranged JC virus variants nor in the proliferation of demyelinated lesions in the brain.**

(*Arch Pathol Lab Med.* 2004;128:275–278)

Progressive multifocal leukoencephalopathy (PML) is a fatal demyelinating disease of the central nervous system that affects individuals with decreased immunocompetence.[1] The causative agent is the human polyomavirus JC virus (JCV), first isolated in 1971 from the brain of a PML patient.[2] Although PML was once a rare disease, it is now a common opportunistic infection in patients with acquired immunodeficiency syndrome.[3]

Most individuals are asymptomatically infected with JCV during childhood.[4,5] The infecting JCV reaches the kidney, probably through viremia, and persists there throughout life.[6,7] In adults, the renal JCV replicates and excretes progeny in urine.[8–11] Renal JCV DNA carries the archetype control region (archetype CR),[7] whereas JCV DNA in the brain of PML patients contains various CRs (PML-type CRs) that harbor deletions and duplications, or deletions alone, in reference to the archetype.[12–15]

To explain the correlation between archetype and PML-type JCVs, Yogo and Sugimoto[15] proposed the archetype concept, which is formulated as follows: (1) JCVs with the archetype CR are circulating in the human population; (2) the archetype CR is highly conserved, in marked contrast to the hypervariable CRs (PML-type CRs) of JCVs in the brain of PML patients; (3) each PML-type CR is produced from the archetype by deletion and duplication, or by deletion alone; (4) the shift of the CR from archetype to PML type occurs during persistence in the host; and (5) PML-type JCVs never return to the human population.

The archetype concept assumes that no amino acid change in the viral proteins is involved in the generation of PML-type JCV. To examine this assumption, we studied a PML case in which multiple, distinct, PML-type JCVs were detected in autopsied brain tissues.[16] We established and sequenced many complete JCV DNA clones. Viral coding sequences were identical among clones with different control regions and among clones derived from different brain regions. We concluded that nucleotide substitution (and resultant amino acid change) was involved neither in the genesis of rearranged JCVs nor in the proliferation of demyelinated lesions in the brain.

MATERIALS AND METHODS

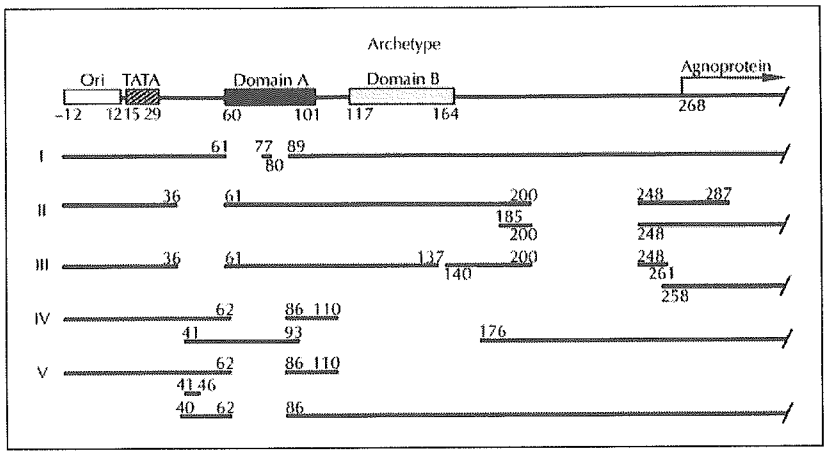
A detailed case report was published previously.[16] In brief, a 14-year-old boy with Wiskott-Aldrich syndrome suffered from progressive impairment of ocular movement and anarthria 6 months after allogenic bone marrow transplantation. T2-weighted magnetic resonance imaging showed high-signal areas in the right occipital lobe, cerebellum, and pons. JC viral DNA was detected in the cerebrospinal fluid by nested polymerase chain re-

Accepted for publication November 4, 2003.

From the Department of Urology, The University of Tokyo, Tokyo, Japan (Drs Zheng, Kitamura, and Yogo), and the Department of Pediatrics, Tokai University of School of Medicine, Isehara, Japan (Drs Yasuda and Kato). Dr Yasuda is now with Yasuda Pediatric Clinic, Machida, Japan.

The authors have no relevant financial interest in the products or companies described in this article.

Reprints: Yoshiaki Yogo, PhD, Department of Urology, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan (e-mail: yogo-tky@umin.ac.jp).



Diagrammatic representation of the detected JC virus (JCV) control region (CR) sequences. The structure of the archetypal CR is schematically shown at the top. The origin of replication (Ori), TATA sequence (TATA), and agnoprotein are indicated.[25] Domain A indicates a sequence duplicated in many progressive multifocal leukoencephalopathy (PML)–derived JCV isolates, and domain B indicates a sequence deleted in many PML-derived JCV isolates.[21] The JCV CRs detected in this study (I to V) are shown with deletions relative to the archetype described as gaps. Reading from left to right, when a repeat is encountered, the linear representation is displaced to the line below and to a position corresponding to the sequence of the archetype. Numbers below each box and the ends of lines are nucleotide numbers indicating end locations (nucleotide numbers are those of the archetype).[12]

Numbers of Complete JC Virus DNA Clones With Distinct Control Region (CR) Sequences Established From Various Regions of the Brain

Brain Region	Total No. of Clones*	No. of Clones With CR Sequences*†				
		I	II	III	IV	V
Cerebellum	14 (6)	12 (4)	1 (1)	0	0	1 (1)
Occipital lobe	7 (7)	2 (2)	3 (3)	2 (2)	0	0
Brainstem	8 (7)	5 (4)	0	0	3 (3)	0

* Numbers of clones whose complete DNA sequences were determined are indicated within parentheses.
† Schematic representations of CRs are shown in the Figure.

action. Several therapeutic approaches were not effective, and the patient died 8 months after bone marrow transplantation.

DNA was previously extracted from autopsied brain tissues (cerebellum, occipital lobe, and brainstem).[16] From these DNA samples, entire JCV DNA sequences were cloned into pUC19 at the unique *Bam*HI site, as described elsewhere.[17] The resultant recombinant plasmids containing complete JCV DNA sequences were prepared using a QIAGEN Plasmid Midi kit (QIAGEN GmbH, Hilden, Germany). Purified plasmids were sequenced as described by Sugimoto et al.[18] The CR sequences were aligned by eye in reference to the archetype sequence,[12] while entire DNA sequences, excluding the CR sequences, were aligned with the CLUSTAL W program (ftp://ftp-igbmc.u-strasbg.fr).[19]

RESULTS

Control Region Sequences

We established 14 clones from the cerebellum, 7 from occipital lobe, and 8 from brainstem (29 total). We first determined the CR sequences of these clones. We identified 5 rearranged CR sequences, designated CRs I to V. These sequences are diagrammatically represented in the Figure in reference to the archetype CR at the top. Deletions in rearranged sequences I through V are shown as gaps, with duplications depicted by parallel lines. Control regions I, II, and IV corresponded to TK-1a, -1c, and -1d, respectively, previously identified by polymerase chain reaction from the same brain tissue DNA.[16] Control regions III and V were not detected in the previous study.[16]

The structural features of the 5 CR sequences are summarized as follows: (1) CR I had only deletions, whereas the others (II to V) had both deletions and duplications. (2) Control regions II and III had 2 common deletions, spanning nucleotides (nt) 37 to 60 and nt 201 to 247. However, duplications in CRs II and III were unique, and CR III had a small deletion (nt 138–139) not present in CR II. These features suggest that the same intermediate carrying only the common deletions generated CRs II and III by different subsequent changes. (3) Control regions IV and V shared the same deletion (nt 63–85) and the same breakpoints (nt 41 and 110). However, these CRs also had a unique deletion (IV) and unique duplications (IV and V). A common intermediate probably generated IV and V, but it was not readily inferred how this intermediate produced CRs IV and V. All in all, it was concluded that at least 3 independent CR rearrangements occurred in a JCV strain (or strains) in the present PML case.

In the present case of PML, we previously amplified various rearranged CRs of JCV in the brain using polymerase chain reaction and found that each rearranged CR showed a unique distribution pattern in the brain.[16] These distribution patterns were confirmed in the present study, for which many complete JCV DNA clones were analyzed (see the Table). Thus, CR I (TK-1a) was widespread in the brain, with the highest incidence in the cerebellum; CR II (TK-1c) mainly occurred in the occipital lobe; and CR IV (TK-1d) mainly occurred in the brainstem. In addition, CR III, which was first detected in the present study but was structurally related to CR II (see above), was detected only in the occipital lobe, where CR II was mainly detected.

Coding Sequences

JC virus DNA clones were classified into groups according to origin and CR sequences (Table). We determined the complete coding sequences of representative or all clones belonging to these groups (Table). A single complete coding sequence of JCV was mainly detected in the brain in the current case, but a minor complete coding

sequence was detected in the 3 clones with CR IV. These minor clones carried T at a position (nt 824) within the *VP2* (a minor capsid protein) gene, whereas the major clones carried C at this position. A clone with CR V (CR structurally related to CR IV) carried C, rather than T, at nt 824. The C/T nucleotide change at position 824 caused an amino acid difference of alanine/valine (a change between amino acids with similar properties).

The result regarding the complete coding sequences described above had 2 implications. First, the complete coding sequence was identical regardless of the structures of CRs. The CR sequences were classified into 3 groups with distinct rearrangements. Thus, the identity of the complete coding sequence among clones with distinctly rearranged CRs suggested that in the present PML case, JCV DNA sequences with distinctly rearranged CRs were generated from the same archetype strain. Second, the complete coding sequence was identical among clones that carried the same CR, even if they were derived from different brain regions. For example, the same complete coding sequence (ie, the sequences harboring C at nt 824) was detected in all clones with CR I derived from all 3 brain regions. Likewise, this sequence was detected in all clones with CR II derived from 2 brain regions (cerebellum and occipital lobe).

COMMENT

It is now generally accepted that JCV variants in the brains of PML patients (PML-type JCVs) are generated from archetype strains through sequence rearrangement (deletion and duplication, or deletion alone) in the CR.[12–14,20,21] Yogo and Sugimoto[15] developed this view into the archetype concept. This concept assumes that no significant amino acid change in the viral proteins is involved in the generation of PML-type JCVs. We used a PML case in which multiple distinct PML-type JCVs were detected in the brain to examine this assumption.[16] We established and sequenced many complete JCV DNA clones from autopsied brain tissues and compared CR and coding sequences among the JCV DNA clones obtained. Five CR sequences were detected in total, but they could be classified into 3 groups that we believe evolved independently from the archetype. The same complete coding sequence was detected in most clones with various CRs, excluding CR IV. Although we did not analyze the archetype strain that would have generated the various PML-type JCVs, it can be assumed that no common mutation occurred during the genesis of PML-type JCVs.[22] Therefore, the detection of the same complete coding sequence in most JCV DNA clones, regardless of the structure of the CR, suggested that the detected complete coding sequence originally existed in the hypothetical archetype JCV and had been conserved during the generation of various PML-type JCVs.

In JCV isolates with CR IV, we detected a single nucleotide change in the *VP2* (a minor capsid protein) gene. Because this nucleotide substitution caused only a change between amino acids (alanine and valine) with similar properties, the structure and function of *VP2* would not have been influenced by this mutation. All in all, we can conclude that at least in the current case, amino acid substitution in the viral proteins played no important role in the genesis of JCVs with various rearranged CRs.

In the present PML case, the same rearranged CRs occurred in multiple brain regions.[16] For instance, CR I occurred in 3 regions (cerebellum, occipital lobe, and brainstem), and CR II occurred in 2 regions (cerebellum and occipital lobe). We believe this observation reflects the proliferation of demyelinated lesions in the brain.[16] We detected the same complete coding sequence in JCV DNA clones derived from different brain regions. We concluded that nucleotide substitutions and resultant amino acid changes rarely occurred in the proliferating demyelinated lesions in the brain.

JC viruses that persist in renal tissue and that are excreted in the urine carry the archetype CR and represent JCVs circulating in human populations.[15] Zheng et al[23] recently investigated how frequently renal/urinary JCVs undergo nucleotide substitution in their coding regions. In brief, they established 5 to 9 complete JCV DNA clones (61 in total) from the urine of 11 individuals (parents and children) belonging to 5 families. The complete sequences of these clones were determined and compared in each family. In the viral coding sequences, 1 or a few nucleotide substitutions per individual were detected in 5 individuals, but none were detected in 6 individuals. It is not easy to compare mutation rates in the coding regions between archetype and PML-type JCVs on the basis of the findings in the present and previous study,[23] as only a single PML case was analyzed in the present study. Nevertheless, it is likely that mutation in the JCV genome is not very frequent in PML patients, although sequence rearrangements in the CR more frequently occur in immunosuppressed patients than in immunocompetent patients.[24]

References

1. Walker DL. Progressive multifocal leukoencephalopathy. In: Vinken PJ, Bruyn GW, Klawans HL, Koetsier JC, eds. *Demyelinating Diseases*. Amsterdam, The Netherlands: Elsevier; 1985:503–524. *Handbook of Clinical Neurology*; Vol 3(47).
2. Padgett BL, Walker DL, ZuRhein GM, Eckroade RJ, Dessel BH. Cultivation of papova-like virus from human brain with progressive multifocal leucoencephalopathy. *Lancet*. 1971;1(7712):1257–1260.
3. Berger JR, Kaszovitz B, Post MJ, Dickinson G. Progressive multifocal leukoencephalopathy associated with human immunodeficiency virus infection: a review of the literature with a report of sixteen cases. *Ann Intern Med*. 1987;107:78–87.
4. Padgett BL, Walker DL. Prevalence of antibodies in human sera against JC virus: an isolate from a case of progressive multifocal leukoencephalopathy. *J Infect Dis*. 1973;127:467–470.
5. Padgett BL, Walker DL. New human papovaviruses. *Prog Med Virol*. 1976;22:1–35.
6. Chesters PM, Heritage J, McCance DJ. Persistence of DNA sequences of BK virus and JC virus in normal human tissues and in diseased tissues. *J Infect Dis*. 1983;147:676–684.
7. Tominaga T, Yogo Y, Kitamura T, Aso Y. Persistence of archetypal JC virus DNA in normal renal tissue derived from tumor-bearing patients. *Virology*. 1992;186:736–741.
8. Kitamura T, Aso Y, Kuniyoshi N, Hara K, Yogo Y. High incidence of urinary JC virus excretion in nonimmunosuppressed older patients. *J Infect Dis*. 1990;161:1128–1133.
9. Kitamura T, Kunitake T, Guo J, Tominaga T, Kawabe K, Yogo Y. Transmission of the human polyomavirus JC virus occurs both within the family and outside the family. *J Clin Microbiol*. 1994;32:2359–2363.
10. Kitamura T, Sugimoto C, Kato A, et al. Persistent JC virus (JCV) infection is demonstrated by continuous shedding of the same JCV strains. *J Clin Microbiol*. 1997;35:1255–1257.
11. Agostini HT, Ryschkewitsch CF, Stoner GL. Genotype profile of human polyomavirus JC excreted in urine of immunocompetent individuals. *J Clin Microbiol*. 1996;34:159–164.
12. Yogo Y, Kitamura T, Sugimoto C, et al. Isolation of a possible archetypal JC virus DNA sequence from nonimmunocompromised individuals. *J Virol*. 1990;64:3139–3143.
13. Ault GS, Stoner GL. Two major types of JC virus defined in progressive multifocal leukoencephalopathy brain by early and late coding region DNA sequences. *J Gen Virol*. 1992;73:2669–2678.
14. Agostini HT, Ryschkewitsch CF, Singer EJ, Stoner GL. JC virus regulatory region rearrangements and genotypes in progressive multifocal leukoencephalopathy: two independent aspects of virus variation. *J Gen Virol*. 1997;78:659–664.
15. Yogo Y, Sugimoto C. The archetype concept and regulatory region rear-

rangement. In: Khalili K, Stoner GL, eds. *Human Polyomaviruses: Molecular and Clinical Perspectives*. New York, NY: John Wiley & Sons; 2001:127–148.

16. Yasuda Y, Yabe H, Inoue H, et al. Comparison of PCR-amplified JC virus control region sequences from multiple brain regions in PML. *Neurology*. 2003; 61:1617–1619.

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

18. Sugimoto C, Hasegawa M, Kato A, et al. Evolution of human polyomavirus JC: implications for the population history of humans. *J Mol Evol*. 2002;54:285–297.

19. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 1994;22:4673–4680.

20. Flaegstad T, Sundsfjord A, Arthur RR, Pedersen M, Traavik T, Subramani S. Amplification and sequencing of the control regions of BK and JC virus from human urine by polymerase chain reaction. *Virology*. 1991;180:553–560.

21. Iida T, Kitamura T, Guo J, et al. Origin of JC polyomavirus variants associated with progressive multifocal leukoencephalopathy. *Proc Natl Acad Sci U S A*. 1993;90:5062–5065.

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

23. Zheng HY, Kitamura T, Takasaka T, Chen Q, Yogo Y. Unambiguous identification of JC virus strains transmitted from parents to children. *Arch Virol*. 2004. In press.

24. Kitamura T, Satoh K, Tominaga T, et al. Alteration in the JC polyomavirus genome is enhanced in immunosuppressed renal transplant patients. *Virology*. 1994;198:341–345.

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

Available online at www.sciencedirect.com
SCIENCE DIRECT
Microbes and Infection 6 (2004) 596–603
Microbes and Infection
www.elsevier.com/locate/micinf

Original article

Regional distribution of two related Northeast Asian genotypes of JC virus, CY-a and -b: implications for the dispersal of Northeast Asians

Huai-Ying Zheng [a,*], Pengyun Zhao [a], Hideki Suganami [b], Yasuo Ohasi [c], Hiroshi Ikegaya [d], Jung-Chul Kim [e], Chie Sugimoto [f], Tomokazu Takasaka [a], Tadaichi Kitamura [a], Yoshiaki Yogo [a]

[a] *Department of Urology, Faculty of Medicine, The University of Tokyo, Hongo 7-3-1, Tokyo 113-8655, Japan*
[b] *Graduate Course of Biomedical Statistics, Tokyo University of Science, Tokyo 162-8601, Japan*
[c] *Department of Biostatistics/Epidemiology and Preventive Health Science, Faculty of Medicine, The University of Tokyo, Tokyo 113-8655, Japan*
[d] *Department of Forensic Medicine, The Graduate School of Medicine, The University of Tokyo, Tokyo 113-0033, Japan*
[e] *Department of Immunology, School of Medicine, Kyungpook National University, Taegu, South Korea*
[f] *Tsukuba Primate Center for Medical Science, National Institute of Infectious Diseases, Tsukuba 305-0843, Japan*

Received 25 September 2003; accepted 20 February 2004

Available online 09 April 2004

Abstract

JC virus (JCV) is a useful marker to trace human dispersal. Two genotypes of JCV (MY and CY) are mainly distributed in Northeast Asia. The population history of people carrying MY has been studied in some detail but that of people carrying CY remains poorly understood. To gain insights into the population history of Northeast Asians carrying CY we analyzed the genetic variation in CY isolates. We constructed a neighbor-joining phylogenetic tree from 28 complete CY DNA sequences: on the resultant tree the CY DNA sequences diverged into two clades, designated CY-a and -b, each clustered with a high bootstrap probability. The split into CY-a and -b was estimated to have occurred about 10 000 years ago, based on K_s values (synonymous substitutions per synonymous site) and the suggested rate of synonymous nucleotide substitutions. Comparison of the 28 complete CY sequences revealed six nucleotide mismatches between CY-a and -b, one of which showed a restriction fragment length polymorphism (RFLP). We then PCR-amplified a region of the genome containing this polymorphic site from many CY isolates in various Northeast Asian populations and classified the isolates into CY-a or -b according to the RFLP analysis. CY-a was more abundant than CY-b in various Chinese and Japanese populations but CY-b was more abundant than CY-a in South Koreans. On the basis of the present findings we inferred the population history in East Asians carrying CY.

Keywords: JC virus; Phylogenetic analysis; Northeast Asians; Human migrations

1. Introduction

JC virus (JCV) is a member of the Polyomaviridae family. Its genome is a single molecule of covalently closed, circular double-stranded DNA about 5100 bp in length [1]. JCV is ubiquitous among humans, infecting children asymptomatically, then persisting in renal tissue [2–6]. In adults, renal JCV replicates and progeny viruses are excreted in urine [7–9]. Although JCV transmission is categorized as horizontal (i.e. transmission among individuals after birth) [10], JCV is frequently transmitted from parents to children during long-term cohabitation [11–13], and it is rarely transmitted between human populations [14].

JCV isolates worldwide belong to a single serotype [15], but they can be classified into more than 10 groups (designated as genotypes) according to nucleotide variations in their genomes [16]. (Two systems have been used to classify JCV isolates; nomenclature correspondence between the systems is shown elsewhere [16].) Each of these genotypes occupies a unique domain in the world. (1) The European genotypes EU-a and -b are spread throughout Europe and Mediterranean areas [17,18]. (2) A genotype (B1-c) related to the Asian genotypes occurs in some regions of Europe (e.g. the Netherlands, Greece) [17,18]. (3) The African genotype Af2 is spread not only throughout Africa but also in West and South Asia [17,19,20]. (4) A minor African genotype

Abbreviations: AIK, Akaike Information Criterion; JCV, JC virus; LT, large T antigen; NJ, neighbor-joining; nt, nucleotide; RFLP, restriction fragment length polymorphism.

* Corresponding author. Tel.: +81-3-5800-8662; fax: +81-3-5800-8917.
E-mail address: teizhy-tky@umin.ac.jp (H.-Y. Zheng).

doi:10.1016/j.micinf.2004.02.013

(Af1) occurs in Central and West Africa [17,20]. (5) Various genotypes (e.g. B1-a, -b, -d, B2, CY, MY, and SC) are spread in Asia, with their domains partially overlapping [17,21–26]. (6) A genotype (EU-c) related to the European genotypes (EU-a and -b) occurs in Northeast Siberia [27]. (7) Three genotypes (type 8A, type 8B, and type 2E) not found on the Asian continent but related to the Asian genotypes are spread in the western Pacific [28,29]. (8) A European genotype (EU-a) occurs in native Americans belonging to Inuits [27], while an Asian genotype (MY) occurs in native Americans belonging to Na-Denes and Amerinds [30–32].

Sugimoto et al. [27,33] examined the evolutionary relationships among JCV genotypes by using the whole-genome approach with which a highly reliable phylogeny of JCV isolates can be reconstructed [34,35]. Using the same method, Jobes et al. [28] and Yanagihara et al. [29] examined the evolutionary relationships among Oceanian and Asian JCV genotypes. The results of these studies showed that an ancestral JCV first divided into three superclusters, type-A, -B, and -C, which subsequently split into various genotypes of JCV distributed throughout the world. A split in type-A generated three genotypes, EU-a, -b, and -c. The split in type-B generated Af2, B1-c, and various Asian and Oceanian genotypes. Type-C generated a single subtype (Af1). Pavesi [36] recently suggested that type-C could represent the ancestral type, and that type-A could reflect its direct origin from the ancestral lineage.

The unique mode of JCV transmission, the geographic distribution patterns of various JCV genotypes, and the mode of JCV evolution altogether indicate that JCV evolved with human populations, and therefore JCV should serve as a novel marker of human population history [17,30,37,38]. Indeed, the JCV genotype profiles were often as expected from historical facts or common views among anthropologists [12,21–23,25,26,28–32]. However, it has also been reported that some aspects of variation in JCV differ with our knowledge of human population history [39].

Two genotypes of JCV (MY and CY), both generated by splits in type-B [33], are mainly distributed in Northeast Asia [17]. These genotypes have unique patterns of geographic distribution. Thus, MY is restricted to Japan and South Korea, whereas CY is widespread in Northeast Asia, including Japan, South Korea, and China. Interestingly, MY is also prevalent in various Native American populations throughout the Americas, excluding those in the Arctic [27,30–32]. It was shown by Zheng et al. [32] that splits in the MY genotype generated seven intra-MY subgroups, two (MY-a and -b) containing only Japanese/Korean isolates and five (MY-c to -g) containing only Native American isolates. Based on these findings, it was inferred that 10,000–30,000 years ago, a human population carrying proto-MY, probably living in Northeast Siberia or Beringia, diverged into several subpopulations, carrying MY-a to -g, and that the subpopulations carrying MY-a and -b migrated to the Japanese archipelago and the Korean peninsula and those carrying MY-c to -g to the Americas [32].

However, the population history of Northeast Asians carrying CY remains to be elucidated. The objective of the present study was to gain insights into the dispersal of the CY-carrying group. To achieve this end, we utilized genetic variations in CY isolates. Suzuki et al. [12] constructed an NJ phylogenetic tree from complete JCV DNA sequences worldwide, including those from Japanese Americans. On the resultant tree, genotype CY split into two clades with a high bootstrap probability. However, one of these clades contained only American isolates (one Japanese-American and one Southern-Californian isolate). We therefore examined whether Asian CY isolates are also included in this clade. We confirmed that CY isolates worldwide are classified into two clades (designated as CY-a and -b), both containing a substantial number of Asian isolates. Comparison of the complete CY-a and -b sequences revealed six nucleotide mismatches, one of which showed a restriction fragment length polymorphism (RFLP). Using this RFLP, we revealed the ratios of CY-a and -b in various Northeast Asian populations. Furthermore, we estimated the time scale for the split of CY into CY-a and -b based on K_s values (synonymous substitutions per synonymous site) and the suggested rate of synonymous nucleotide substitutions. On the basis of the findings obtained, we inferred the population history of East Asians carrying CY.

2. Materials and methods

2.1. Urine samples

The sites of urine collection are shown in Fig. 1. Urine specimens from Changchun, Seoul, and Taegu were collected in this study, while those from X'ian, Lanzhou, Urumuqi, Kyoto, Yaizu, and Okinawa were collected in previous studies [14,22,23]. Urine donors were native inhabitants of each site (healthy volunteers or general patients) aged 50 years or older. Urine samples were collected in 10 mM EDTA (pH 8.0), and sent to the Laboratory of Viral Infection, Department of Microbiology and Immunology, Institute of Medical Science, University of Tokyo, where DNA was extracted as described previously [7].

2.2. Selection of isolates belonging to CY

Isolates belonging to genotype CY were selected using phylogenetic and RFLP methods. The phylogenetic method involved PCR-amplification of the 610-bp IG region [40], sequencing of the amplified fragments, and construction of a neighbor-joining (NJ) tree [41] (see [17] for details). The PFLP method was carried out as described by Kitamura et al. [23] with the following modification. Isolates belonging to group X were selected according to the RFLP procedure as described [23]. As it was possible that group X contained isolates belonging to not only CY but also B1-b, we amplified non-coding regulatory regions from isolates assigned as

Fig. 1. A map showing nine sites of urine collection.

group X, and subjected them to RFLP analysis using *Afl*III. CY was cleaved, but B1-b was not cleaved, with this enzyme between nucleotide (nt) 222 and 223 (nucleotide numbers were those of CY [42], GSDB, DDBJ, EMBL and NCBI accession number: AB038249).

2.3. Sub-classification of CY into CY-a and -b

From DNA extracted from urine specimens containing CY isolates, a 454-bp region of the viral genome, spanning nt 4607 to 5060, was amplified by PCR using primer pairs C1 and C2 under the conditions described [11]. C1 (5′-GCAATCAAAGCAATAGCAATCTATCC-3′) spanned nt 4581–4606, and C2 (5′-GGGCGTGGAGGCTTTTTAGG-AGGCCA-3′) spanned nt 5086–5061. PCR mixtures were extracted with phenol and filtered through spin-columns containing Sephadex G-25. Purified PCR mixtures were digested at 37 °C for 1 h with 10 units of *Aat*I (Toyobo Biochemicals, Osaka, Japan). The digest was resolved by electrophoresis on a 1.8% agarose gel stained with ethidium bromide.

2.4. DNA sequencing

Complete JCV DNA clones previously established from East Asians [43,44] were sequenced with an autosequencer (ABI PRISM 373S DNA Analyzer, Applied Biosystems, Foster City, USA) as described previously [33].

JCV isolates whose complete DNA sequences were used for the phylogenetic analysis are shown in Table 1.

2.5. Phylogenetic analysis

The non-coding regulatory region of the JCV genome was excluded from phylogenetic analysis, as this region is hypervariable especially in JCV isolates derived from the brains of PML patients [45]. Rates of synonymous substitution were estimated using the Diverge program in the GCG Wisconsin package. DNA sequences were aligned using CLUSTAL W [46] 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 NJ trees using the CLUSTAL W program. Divergences were estimated by the two-parameter method [47] using CLUSTAL W. Phylogenetic trees were visualized using TREEVIEW [48]. To assess the confidence of branching patterns of the NJ trees, 1000 bootstrap replicates were performed [49].

2.6. Statistical analysis

Based on the geographical differences in genotypic incidence, an appropriate grouping of five regions in Northeast Asia was statistically explored. There were 52 all possible groupings, for example, AAAAA, AA BBB, A BB CC, ..., and A B C D E, in which each capital letter represented the different groups, and the logistic model was applied to each grouping. We decided that the grouping that provided the best fit would be the most appropriate grouping. If the number of group is increased, the goodness of fit is improved superficially, therefore, the AIC (Akaike Information Criterion, −2-log likelihood + 2 × group number) was selected as the criterion of the goodness of fit. The AIC evaluates the model fitness, taking account of the number of unknown parameters (model complexity) and the model with the smallest AIC was considered to be the best model [50]. Then the likelihood ratio test between the full (non-grouping) model (A B C D E) and selected model was used to evaluate the fitness of the selected model [51]. The 95% confidence interval for CY-a proportion of the selected model was calculated. All analyses were carried out using SAS software ver. 8.2 (SAS Institute Inc., Cary, North Carolina USA).

Table 1
JCV isolates whose complete DNA sequences were used for the phylogenetic analysis (Fig. 2)

Isolate	Genotype	Geographic origin	Accession no. [a]	Reference	
				For isolate	For sequence
CY	CY-a	Tokyo, Japan	AB038249	[60]	[42]
MS	CY-a	Tokyo, Japan	AB081654	[60]	[12]
SI	CY-a	Tokyo, Japan	AB081617	[60]	[12]
Tky-2a	CY-a	Tokyo, Japan	AB038255	[61]	[42]
UA	CY-a	Tokyo, Japan	AB081618	[60]	[12]
FO [b]	CY-b	Yaizu, Japan	AB118231	[43]	– [c]
NK [b]	CY-a	Yaizu, Japan	AB118232	[43]	–
NY [b]	CY-b	Yaizu, Japan	AB118233	[43]	–
ES [b]	CY-b	Yaizu, Japan	AB118234	[43]	–
SS [b]	CY-b	Yaizu, Japan	AB118235	[43]	–
SK-2	CY-b	Seoul, South Korea	AB118651	[44]	–
SK-3	CY-b	Seoul, South Korea	AB118652	[44]	–
SK-5	CY-b	Seoul, South Korea	AB118653	[44]	–
CB-1	CY-a	Beijing, China	AB118654	[44]	–
CB-3	CY-a	Beijing, China	AB048560	[44]	[33]
CB-4	CY-b	Beijing, China	AB118655	[44]	–
CW-1	CY-a	Wuhan, China	AB118656	[44]	–
CW-3	CY-a	Wuhan, China	AB118657	[44]	–
CW-5	CY-b	Wuhan, China	AB118658	[44]	–
CW-6	CY-a	Wuhan, China	AB118659	[44]	–
MO-1	CY-a	Ulaanbaatar, Mongolia	AB048561	[44]	[33]
MO-6	CY-a	Ulaanbaatar, Mongolia	AB048562	[44]	[33]
J2-10 [d]	CY-a	Los Angeles, USA	AB081600	[12]	[12]
J3-7 [e]	CY-b	Los Angeles, USA	AB081605	[12]	[12]
J3-9 [e]	CY-a	Los Angeles, USA	AB081607	[12]	[12]
J3-11 [e]	CY-a	Los Angeles, USA	AB081602	[12]	[12]
J3-13a [e]	CY-a	Los Angeles, USA	AB081603	[12]	[12]
LA-17 [f]	CY-b	La Jolla, USA	AB081610	[12]	[12]

[a] GSDB, DDBJ, EMBL and NCBI accession numbers.
[b] Isolated from renal transplant patients.
[c] This study.
[d] Isolated from a second-generation Japanese American.
[e] Isolated from a third-generation Japanese American.
[f] Isolated from a general southern Californian.

3. Results

3.1. Sub-classification of CY into CY-a and -b

We determined 14 complete JCV (CY) DNA sequences from Northeast Asia (six from China, five from Japan, and three from South Korea), and constructed an NJ phylogenetic tree from these sequences together with 14 complete CY sequences reported previously (Table 1) [12,33,42] (the complete sequence of isolate ET-3 [33], an isolate belonging to Af2, was included as an outgroup). According to the resultant tree (Fig. 2), CY diverged into two clades, designated as CY-a and -b, with high bootstrap probabilities (≥ 90%). CY-a as well as -b included many Northeast Asian isolates, suggesting that both clades originated in Northeast Asia.

Using the suggested rate of synonymous nucleotide substitutions (4×10^{-7} substitutions per synonymous substitution site per year) [33,35], we attempted to estimate the time scale of the divergence of CY JCVs. We calculated the average K_s values (synonymous substitutions per synonymous site) between isolates belonging to CY-a and -b for each of the three genes, VP1, VP2, and large T antigen (LT) (Table 2). As the mean of these K_s values, calculated with weight for the length of each gene, we obtained 0.009 synonymous substitutions per synonymous site for CY-a vs. CY-b. According to the suggested rate of synonymous nucleotide substitutions (4×10^{-7}/synonymous site/year) [33,35], we estimated that the split between CY-a and -b occurred 11,000 years ago.

3.2. Sites of the genome showing nucleotide variations between CY-a and -b

We aligned the 28 complete CY DNA sequences to find sites showing nucleotide differences between CY-a and -b. We identified nucleotide variations at six positions (Table 3), one within the non-coding region between the VP1 and LT genes, two within the agnogene, and three within the LT gene. The three nucleotide-changes within the LT gene caused no amino acid changes, while both nucleotide-changes within the agnogene caused amino acid changes (Table 3) (the effect of these amino acid changes on viral

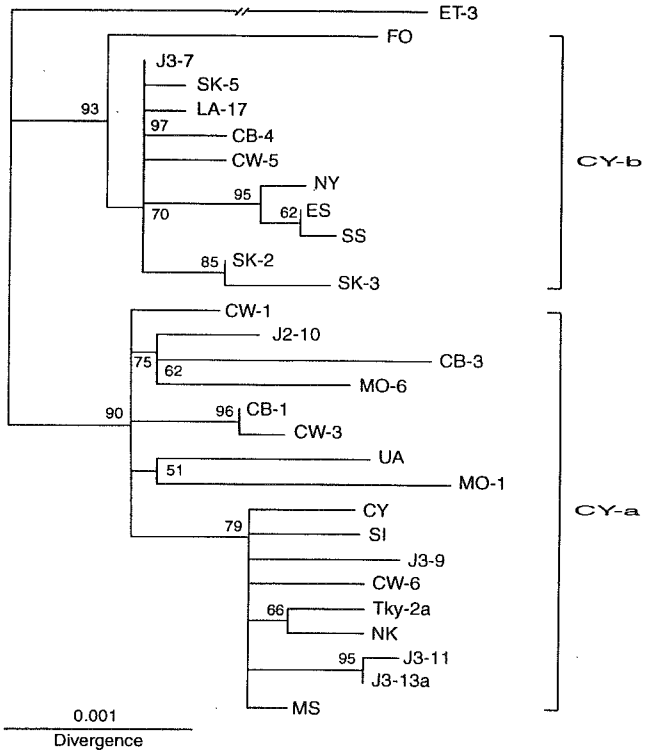


Fig. 2. NJ phylogenetic tree relating 28 complete JCV DNA sequences belonging to CY. An NJ phylogenetic tree was constructed from 28 complete CY sequences reported in this and previous studies. The phylogenetic tree was visualized using TREEVIEW. The tree was rooted using a genotype Af2 isolate (ET-3) [33] as the outgroup. The symbols for isolates are shown in Table 1. The numbers at nodes in the tree indicate the bootstrap confidence levels (percent) obtained with 1000 replications (only values ≥50% are shown). Intra-CY subgroups are indicated.

growth remains to be clarified). One of the nucleotide changes within the LT gene (nt 4950) showed an RFLP between CY-a and -b, involving *Aat*I; CY-a was cleaved with *Aat*I at this site, while CY-b was not.

3.3. Distribution of CY-a and -b in various Northeast Asian populations

We selected nine CY-rich Northeast-Asian sites, including four sites in China, two in South Korea, and three in Japan (Fig. 1). CY isolates at these sites were classified into CY-a and -b based on the RFLP analysis using *Aat*I. The distribution of CY-a and -b in each Northeast Asian population thus obtained is shown in Table 4, and can be summarized as follows. (1) CY-a was more frequently detected than CY-b at all the sites examined in China. The percentage of CY-a ranged from 56% to 73%. (2) CY-a was also more frequently detected than CY-b at the three sites in Japan (Kyoto, Yaizu, and Okinawa). The percentage of CY-a was consistently high, ranging from 81% to 87%. (3) In contrast, CY-b was more frequently detected than CY-a at the two sites in South Korea, Seoul and Taegu.

Table 2
Synonymous nucleotide substitutions between CY-a and -b

Gene	Synonymous nucleotide substitutions between CY-a and -b [a]
VP1	0.007
VP2	0.004
LT	0.013
Mean [b]	0.009

[a] For three genes, the estimated numbers of synonymous substitutions per synonymous site are shown as averages of pair-wise comparisons between five CY-a (CY, CB-3, CW-1, MS, and UA) and four CY-b (CW-5, J3-7, NY and SK-2).
[b] Calculated with weight for the length of each gene.

Table 3
Nucleotide differences in the complete sequences between CY-a and -b

Nucleotide position [a]	Gene	Nucleotide	
		CY-a	CY-b
nt 442	Agnogene	G (E) [b]	C (Q)
nt 449	Agnogene	G (R)	C (T)
nt 2509	– [c]	G	A
nt 3737	LT	C	T
nt 4019	LT	C	T
nt 4950	LT	G [d]	A

[a] Nucleotide numbers are those of an archetype isolate CY [42], starting from the midpoint of the origin of replication and proceeding clockwise.
[b] Amino acid differences are shown in parentheses by one letter.
[c] Noncoding region between the VP1 and LT genes.
[d] Constitutes a cleavage site for a restriction enzyme (*Aat*I).

Table 4
Distribution of CY-a and -b in various Northeast Asian regions

Geographic region	Genome incidence for CY [a]	No. of CY isolates analyzed	No. (%) of isolates classified as	
			CY-a	CY-b
Changchun	0.61	33	24 (73)	9 (27)
X'ian	0.67 [b]	16	9 (56)	7 (44)
Lanzhou	0.63 [b]	10	7 (70)	3 (30)
Urumuqi	0.41 [b]	9	6 (67)	3 (33)
Seoul	0.73	18	1 (6)	17 (94)
Taegu	0.47	12	4 (33)	8 (67)
Kyoto	0.81 [c]	21	17 (81)	4 (19)
Yaizu	0.76 [c]	16	13 (81)	3 (19)
Okinawa	0.91 [c]	69	60 (87)	9 (13)

[a] (No. of isolates belonging to CY)/(total no. of isolates).
[b] Cited from [22].
[c] Cited from [23].

3.4. Statistical analysis of regional differences in genotype incidence

For the statistical examination of the observed regional differences in genotype incidence (see above), the sites of urine collection were grouped into five regions, Northeast China, Northwest China, South Korea, Hondo Japan, and Okinawa (Table 5). Northeast China included only Changchun; Northwest China included X'ian, Lanzhou, and Urumuqi; South Korea included Seoul and Taegu; Hondo Japan included Kyoto and Yaizu; and Okinawa included only Okinawa. Genotype incidences for CY-a and -b in these regions are shown in Table 5. The incidence was higher for CY-a than CY-b in Northeast China, Northwest China, Hondo Japan, and Okinawa. In contrast, it is higher for CY-b than for CY-a in South Korea.

Table 5
Incidence of CY-a and -b in five Northeast Asian areas [a]

Geographic areas	No. of CY isolates analyzed	No. (%) of isolates classified as	
		CY-a	CY-b
Northeast China[b]	33	24 (73)	9 (27)
Northwest China[c]	35	22 (63)	13 (37)
South Korea[d]	30	5 (17)	25 (83)
Hondo Japan[e]	37	30 (81)	7 (19)
Okinawa	69	60 (87)	9 (13)

[a] Table 4 was revised for statistic analysis.
[b] Changchun.
[c] X'ian, Lanzhou, and Urumuqi.
[d] Seoul and Taegu.
[e] Kyoto and Yaizu.

These regional differences in genotypic incidence were statistically analyzed using logistic models. An appropriate model was selected, based on the AIC. The minimum AIC was 208.606 among the possible 52 models, and the selected model consisted of three groups (Northeast China and Northwest China), (Hondo Japan and Okinawa) and (South Korea). The model fitness was worse when South Korea and any other region were combined into one group. According to the result of the likelihood ratio test, there was no statistically significant difference between the most complicated model (containing five groups) and the selected model. This means that the selected model was not oversimplified. Moreover, there was a significant difference between the selected model and the simpler model, which grouped South Korea against the four other regions ($P < 0.001$). The proportion of CY-a and 95% confidence intervals for the grouped regions were as follows: the Northeast China and Northwest China group, 67.6% (56.0, 78.0); the Hondo Japan and Okinawa group, 84.9% (77.3, 90.9); and South Korea, 16.7% (6.3, 32.5). The proportions in the group combining Northeast China and Northwest China and in the Hondo Japan and Okinawa group were similar. It was found that the proportion in South Korea was considerably different from these two groups.

4. Discussion

Two major genotypes of JCV (MY and CY), both generated by splits in type-B [33], are distributed in Northeast Asia [17]. Interestingly, MY is also spread among Native Americans excluding Aleut-Eskimos [27], although CY is restricted to East Asia [17]. Zheng et al. [32] revealed that MY isolates worldwide split into several subgroups, those containing only Japanese and Korean isolates and those containing only Native American isolates. This finding along with timings of MY divisions [32] allowed the inference regarding the population history of people carrying MY [32] (see Section 1).

In the present study, we attempted to elucidate the population history of Northeast Asians carrying CY, using the genetic variation in CY isolates. We found that Asian CY isolates diverged into two subgroups, designated as CY-a and -b, according to a phylogenetic tree constructed from complete JCV DNA sequences. These subgroups showed different distribution patterns in North Asia. Thus, CY-a was widespread in Northeast Asian populations, including Northeast and Northwest China, Hondo Japan, and Okinawa. Moreover, CY-a, rather than CY-b, was spread in South China where CY occurs as a minor genotype (unpublished). Such a widespread distribution of CY-a can well be explained by the expansion of a population carrying CY-a in Northeast Asia. On the other hand, the Korean peninsula was the only area where CY-b predominantly occurred. (The reason for the wide and restricted distribution of CY-a and -b, respectively, remains unclear). On the basis of these findings, the following inference may be made. A Northeast Asian population carrying proto-CY (a hypothetical ancestor that generated both CY-a and -b) diverged into two populations, one carrying CY-a and the other carrying CY-b. The former expanded throughout North China and to the western part of the Japanese archipelago; the latter expanded mainly along the Korean Peninsula.

It is generally accepted that modern Japanese were formed by two distinct ethnic groups, the Jomon people who colonized in Japan in the Neolithic period and the later "immigrants" from the Asian continent during Aeneolithic Yayoi and the prehistoric Kofun periods [52,53]. Based on time scales for the split of MY and CY [32, this study], it can be speculated that CY accompanied the "immigrants" from the Asian Continent, while MY was indigenous to the Jomon people. A specific question about the "immigrants" theory is which route the "immigrants" took when they migrated to the Japanese archipelago. Although some genetic studies suggested that the immigrants came through the Korean peninsula [54,55], this issue remains open. In the present study, we showed that CY divided into two subgroups, CY-a and -b, and that one (CY-a) of these subgroups mainly occurs in North China and Japan (both Hondo Japan and Okinawa) and the other (CY-b) predominantly occurs in South Korea. This finding suggests that the "immigrants" to Japan during the Yayoi and Kofun period migrated from North China (a region abundant in CY-a), rather than from the Korean peninsula (a region abundant in CY-b).

It should be noted that ethnic groups carrying JCV genotypes other than CY and MY partially contributed to the formation of modern Northeast Asians. SC (the predominant genotype of JCV in Southeast Asia, including South China) also occurs in North China at a lower frequency [17,22], suggesting that South Chinese partially contributed to the formation of North Chinese. B1-b (the predominant genotype of JCV in Central Asia [21,24]) occurs in Northwest China at a higher frequency [22], suggesting that Central Asians greatly contributed to the formation of Northwest Chinese. An ethnic group carrying B1-a (a minor genotype of JCV spread throughout China [17,21,22]) may have also contributed in part to the formation of North Chinese. A European-related ethnic group carrying EU-a/Jpn probably

contributed to the formation of Northeast Japanese and probably South Koreans [23,27,33,56].

Recently, genetic variation on the non-recombining portion of the Y chromosome has been extensively examined to investigate ancient human migrations and the population structures of human groups [57]. Karafet et al. [58] recently reported an extensive study of Y-chromosome variation in East Asians. The authors grouped Asian populations into Northeast, Southeast, and Central Asians. Based on the Y-chromosome variation among these groups, they suggested that most Northeast Asians are clearly distinguished from most Southeast Asians and that Northeast Asians have a closer genetic affinity with Central Asians than with Southeast Asians. The first suggestion appears consistent with the presence of different JCV genotype profiles in Northeast and Southeast Asia (CY and SC predominantly occur in Northeast and Southeast Asians, respectively [21]), although an alternative view was suggested by Chu et al. [59]. Nevertheless, the second suggestion does not necessarily agree with the JCV findings that Northeast, Southeast, and Central Asians carry unique major genotypes of JCV: CY (Northeast Asians), SC (Southeast Asians), and B1-b (Central Asians). This discrepancy may partly be due to the different grouping of populations. For example, Mongolians and Uzbeks who carry mainly B1-b genotype [21,24] were classified as Central Asians in JCV studies, while Mongolians and Uzbeks were classified as different groups (Northeast and Central Asians, respectively) in the Y-chromosome study noted above. Unfortunately, Y-chromosome variation has not been effectively used to elucidate the dispersals of Northeast Asians to the Korean peninsula and the Japanese archipelago.

Acknowledgements

We are grateful to all urine donors. This study was supported in part by grants from the Ministry of Health, Labor and Welfare, Japan.

References

[1] C.N. Cole, S.D. Conzen, fourth ed. in: D.M. Knipe, P.M. Howley, D.E. Griffin, M.A. Martin, R.A. Lamb, B. Roizman, S.E. Straus (Eds.), Fields Virology, Lippincott Williams & Wilkins, Philadelphia, 2001, pp. 2141–2174.

[2] B.L. Padgett, D.L. Walker, Prevalence of antibodies in human sera against JC virus, an isolate from a case of progressive multifocal leukoencephalopathy, J. Infect. Dis. 127 (1973) 467–470.

[3] D.L. Walker, B.L. Padgett, The epidemiology of human polyomaviruses, Prog. Clin. Biol. Res. 105 (1983) 99–106.

[4] P.M. Chesters, J. Heritage, D.J. McCance, Persistence of DNA sequences of BK virus and JC virus in normal human tissues and in diseased tissues, J. Infect. Dis. 147 (1983) 676–684.

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

[6] T. Kitamura, C. Sugimoto, A. Kato, H. Ebihara, M. Suzuki, F. Taguchi, K. Kawabe, Y. Yogo, Persistent JC virus (JCV) infection is demonstrated by continuous shedding of the same JCV strains, J. Clin. Microbiol. 35 (1997) 1255–1257.

[7] T. Kitamura, Y. Aso, N. Kuniyoshi, K. Hara, Y. Yogo, High incidence of urinary JC virus excretion in nonimmunosuppressed older patients, J. Infect. Dis. 161 (1990) 1128–1133.

[8] T. Kitamura, T. Kunitake, J. Guo, T. Tominaga, K. Kawabe, Y. Yogo, Transmission of the human polyomavirus JC virus occurs both within the family and outside the family, J. Clin. Microbiol. 32 (1994) 2359–2363.

[9] H.T. Agostini, C.F. Ryschkewitsch, G.L. Stoner, Genotype profile of human polyomavirus JC excreted in urine of immunocompetent individuals, J. Clin. Microbiol. 34 (1996) 159–164.

[10] R. Daniel, K. Shah, D. Madden, S. Stagno, Serological investigation of the possibility of congenital transmission of papovavirus JC, Infect. Immun. 33 (1981) 319–321.

[11] T. Kunitake, T. Kitamura, J. Guo, F. Taguchi, K. Kawabe, Y. Yogo, Parent-to-child transmission is relatively common in the spread of the human polyomavirus JC virus, J. Clin. Microbiol. 33 (1995) 1448–1451.

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

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

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

[15] E.O. Major, fourth ed, in: D.M. Knipe, P.M. Howley, D.E. Griffin, M.A. Martin, R.A. Lamb, B. Roizman, S.E. Straus (Eds.), Fields Virology, Lippincott Williams & Wilkins, Philadelphia PN, 2001, pp. 2175–2196.

[16] H.T. Agostini, D.V. Jobes, G.L. Stoner, in: K. Khalili, G.L. Stoner (Eds.), Human Polyomaviruses: Molecular and Clinical Perspectives, John Wiley & Sons, New York, 2001, pp. 491–526.

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

[18] H.T. Agostini, A. Deckhut, D.V. Jobes, R. Girones, G. Schlunck, M.G. Prost, C. Frias, E. Pérez-Trallero, C.F. Ryschkewitsch, G.L. Stoner, Genotypes of JC virus in East, Central and Southwest Europe, J. Gen. Virol. 82 (2001) 1221–1331.

[19] H.T. Agostini, C.F. Ryschkewitsch, G.R. Brubaker, J. Shao, G.L. Stoner, Five complete genomes of JC virus type 3 from Africans and African Americans, Arch. Virol. 142 (1997) 637–655.

[20] S.C. Chima, C.F. Ryschkewitsch, G.L. Stoner, Molecular epidemiology of human polyomavirus JC in the Biaka Pygmies and Bantu of Central Africa, Mem. Inst. Oswaldo. Cruz. 93 (1998) 615–623.

[21] J. Guo, C. Sugimoto, T. Kitamura, H. Ebihara, A. Kato, Z. Guo, J. Liu, S.P. Zheng, Y.L. Wang, Y.Q. Na, M. Suzuki, F. Taguchi, Y. Yogo, 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 (1998) 2499–2505.

[22] Z. Guo, S.P. Zheng, C. Sugimoto, Y.L. Wang, H.-Y. Zheng, T. Takasaka, T. Kitamura, J. Guo, Y. Yogo, JC virus genotypes in northwestern China: implications for its population history, Anthropol. Sci. 109 (2001) 203–212.

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

[24] L. Saruwatari, C. Sugimoto, T. Kitamura, N. Ohno, E. Sakai, P. Shresta, B.K. Hoa, P.T.P. Phi, H.P.H. An, N.T.A. Tuyet, T. Honjo, N. Kobayashi, H.-Y. Zheng, T. Takasaka, Y. Yogo, Asian domains of four major genotypes of JC virus, Af2, B1-b, CY and SC, Arch. Virol. 147 (2002) 1–10.

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

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

[27] C. Sugimoto, M. Hasegawa, H.-Y. Zheng, V. Demenev, Y. Sekino, K. Kojima, T. Honjo, H. Kida, T. Hovi, T. Vesikari, J.A. Schalken, K. Tomita, Y. Mitsunobu, H. Ikegaya, N. Kobayashi, T. Kitamura, Y. Yogo, 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 (2002) 322–335.

[28] D.V. Jobes, J.S. Friedlaender, C.S. Mgone, H.T. Agostini, G. Koki, R. Yanagihara, T.C.N. Ng, S.C. Chima, C.F. Ryschkewitsch, G.L. Stoner, 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 (2001) 2097–2113.

[29] R. Yanagihara, V.R. Nerurkar, I. Scheirich, H.T. Agostini, C.S. Mgone, X. Cui, D.V. Jobes, C.L. Cubitt, C.F. Ryschkewitsch, D.B. Hrdy, J.S. Friedlaender, G.L. Stoner, JC virus genotypes in the western Pacific suggest Asian mainland relationships and virus association with early population movements, Hum. Biol. 74 (2002) 473–488.

[30] H.T. Agostini, R. Yanagihara, V. Davis, C.F. Ryschkewitsch, G.L. Stoner, 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 (1997) 14542–14546.

[31] M. Fernandez-Cobo, H.T. Agostini, G. Britez, C.F. Ryschkewitsch, G.L. Stoner, 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 (2002) 154–168.

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

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

[34] D.V. Jobes, S.C. Chima, C.F. Ryschkewitsch, G.L. Stoner, Phylogenetic analysis of 22 complete genomes of the human polyomavirus JC virus, J. Gen. Virol. 79 (1998) 2491–2498.

[35] J.N. Hatwell, P.M. Sharp, Evolution of human polyomavirus JC, J. Gen. Virol. 81 (2000) 1191–1200.

[36] A. Pavesi, African origin of polyomavirus JC and implications for prehistoric human migrations, J. Mol. Evol. 56 (2003) 564–572.

[37] G.L. Stoner, D.V. Jobes, M. Fernandez Cobo, H.T. Agostini, S.C. Chima, C.F. Ryschkewitsch, JC virus as a marker of human migration to the Americas, Microbes Infect. 2 (2000) 1905–1911.

[38] T.R. Disotell, Discovering human history from stomach bacteria, Genome Biol. 4 (2003) 213.

[39] S. Wooding, Do human and JC virus genes show evidence of host–parasite codemography? Infect. Genet. Evol. 1 (2001) 3–12.

[40] G.S. Ault, G.L. Stoner, Two major types of JC virus defined in progressive multifocal leukoencephalopathy brain by early and late coding region DNA sequences, J. Gen. Virol. 73 (1992) 2669–2678.

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

[42] A. Kato, C. Sugimoto, H.-Y. Zheng, T. Kitamura, Y. Yogo, 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 (2000) 2173–2182.

[43] Y. Yogo, T. Kitamura, C. Sugimoto, K. Hara, T. Iida, F. Taguchi, A. Tajima, K. Kawabe, Y. Aso, Sequence rearrangement in JC virus DNAs molecularly cloned from immunosuppressed renal transplant patients, J. Virol. 65 (1991) 2422–2428.

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

[45] Y. Yogo, C. Sugimoto, in: K. Khalili, G.L. Stoner (Eds.), Human Polyomaviruses: Molecular and Clinical Perspectives, John Wiley & Sons, New York, 2001, pp. 127–148.

[46] J.D. Thompson, D.G. Higgins, T.J. Gibson, 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 (1994) 4673–4680.

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

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

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

[50] D.W. Hosmer, S. Lemeshow, second ed., Applied Logistic Regression, John Wiley & Sons, New York, 2000.

[51] P. McCullagh, J.A. Nelder, Generalized Linear Models, second ed, Chapman and Hall, London, 1989.

[52] Y. Dodo, H. Ishida, Population history of Japan as viewed from cranial nonmetric variation, Anthropol. Sci. 98 (1990) 269–287.

[53] K. Hanihara, Dual structure model for the population history of the Japanese, Japan Rev. 2 (1991) 1–33.

[54] K. Tokunaga, in: Y. Dodo (Ed.), Peopling the Japanese Islands (Series Mongoloid Peoples in Time and Space), University of Tokyo Press, Tokyo, 1995, pp. 193–210 (in Japanese).

[55] S. Horai, K. Murayama, K. Hayasaka, S. Matsubayashi, Y. Hattori, G. Fucharoen, S. Harihara, K.S. Park, K. Omoto, I.-H. Pan, mtDNA polymorphism in East Asian populations, with special reference to the peopling of Japan, Am. J. Hum. Genet. 59 (1996) 579–590.

[56] B.-H. Jeong, K.-H. Lee, E.-K. Choi, K. Kim, Y.-S. Kim, Genotyping of the JC virus in urine samples of healthy Korean individuals, J. Med. Virol. 72 (2004) 281–289.

[57] M.A. Jobling, C. Tyler-Smith, The human Y chromosome: an evolutionary marker comes of age, Nat. Rev. Genet. 4 (2003) 598–612.

[58] T. Karafet, L. Xu, R. Du, W. Wang, S. Feng, R.S. Wells, A.J. Redd, S.L. Zegura, M.F. Hammer, Paternal population history of East Asia: sources, patterns, and microevolutionary processes, Am. J. Hum. Genet. 69 (2001) 615–628.

[59] J.Y. Chu, W. Huang, S.Q. Kuang, J.M. Wang, J.J. Xu, Z.T. Chu, Z.Q. Yang, K.Q. Lin, P. Li, M. Wu, Z.C. Geng, C.C. Tan, R.F. Du, L. Jin, Genetic relationship of populations in China, Proc. Natl. Acad. Sci. USA 95 (1998) 11763–11768.

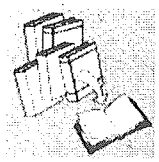
[60] Y. Yogo, T. Kitamura, C. Sugimoto, T. Ueki, Y. Aso, K. Hara, F. Taguchi, Isolation of a possible archetypal JC virus DNA sequence from nonimmunocompromised individuals, J. Virol. 64 (1990) 3139–3143.

[61] K. Kato, J. Guo, F. Taguchi, O. Daimaru, M. Tajima, H. Haibara, J. Matsuda, M. Sumiya, Y. Yogo, Phylogenetic comparison between archetypal and disease-associated JC virus isolates in Japan, Jpn. J. Med. Sci. Biol. 47 (1994) 167–178.

Rev. Med. Virol. 2004; 14: 179–191.
Published online in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/rmv.428

Reviews in Medical Virology

REVIEW



JC virus genotyping offers a new paradigm in the study of human populations

Yoshiaki Yogo[1]*, Chie Sugimoto[2], Huai-Ying Zheng[1], Hiroshi Ikegaya[3], Tomokazu Takasaka[1] and Tadaichi Kitamura[1]

[1]*Department of Urology, Faculty of Medicine, The University of Tokyo, Tokyo, Japan*
[2]*Tsukuba Primate Center for Medical Science, National Institute of Infectious Diseases, Tsukuba, Japan*
[3]*Department of Forensic Medicine, Graduate School of Medicine, University of Tokyo, Japan*

SUMMARY

A small DNA virus, named JC virus (JCV) and belonging to the *Polyomaviridae*, is attracting the attention of anthropologists worldwide, as JCV genotyping appears to be a novel means of elucidating human migrations and the origins of various ethnic groups. The basic properties of JCV, the regional distributions of JCV genotypes, and the phylogenetic relationships among various JCV genotypes are described. Then, a study is described in which the origin of the modern Japanese was extensively investigated using the JCV genotyping method. Based on JCV genotypes in neighboring areas, the origins of people who carried JCV genotypes to the Japanese Archipelago are discussed. Finally, the relationships between JCV genotypes and Y-chromosome haplogroups are examined, as genetic variation on the Y chromosome has recently been examined in detail to investigate ancient human migrations and the population structures of human groups. Copyright © 2004 John Wiley & Sons, Ltd.

Accepted: 21 January 2004

INTRODUCTION

A small DNA virus named JC virus (JCV) is attracting the attention of anthropologists worldwide, as it has been shown to serve as a novel marker of human populations. Indeed, a significant number of studies in which JCV was used as a marker with which to elucidate the origins of various ethnic populations have been published in the formal journals of the Anthropological Society of Nippon [1–5] and the American Association of Physical Anthropologists [6,7].

JCV is a member of the *Polyomaviridae* family. Its genome is a single molecule of covalently closed, circular double-stranded DNA about 5100 bp in length [8]. JCV is ubiquitous in the human population, infecting children asymptomatically, then persisting in renal tissue throughout life [9–13]. In most adults, renal JCV is not latent but replicates to excrete progeny into urine [14–16], from which JCV DNA can readily be recovered by means of molecular cloning or the polymerase chain reaction (PCR). JCV is usually transmitted from parents to children during cohabitation [17–19], but rarely transmitted among human populations, unless the populations are intermixed [20]. JCV strains in the world can be classified into more than ten major genotypes [21–23], with each genotype occupying a unique geographical domain, suggesting that the evolution of JCV occurred in association with human populations. Furthermore, JCV evolved two-orders of magnitude faster than the human chromosomes [24,25]. The high rate of evolution of JCV, together with its host-linked mode, appears to make JCV a highly sensitive and reliable marker in tracing the history of human populations.

This review provides basic information about JCV, essential to understanding why JCV can be used for the study of human populations. The regional distributions of JCV genotypes and the phylogenetic relationships among various JCV genotypes are reviewed. Next, a study [1] is introduced in which the origin of the modern Japanese

Corresponding author: Dr Y. Yogo, Department of Urology, Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, 113-8655 Tokyo, Japan. E-mail: yogo-tky@umin.ac.jp

Abbreviations used
IG region, VT-intergenic region; JCV, JC virus; PCR, polymerase chain reaction; PML, progressive multifocal leukoencephalopathy; RFLP, restriction fragment length polymorphism; Y haplogroup, Y-chromosome haplogroup.

was extensively investigated using the JCV genotyping method. Regarding the origin and expansion of modern humans, it is generally accepted that a single ancestral human population in Africa generated multiple major sub-populations, migrating to Europe and Asia, further to the Americas and Oceania [26–28]. Thus, based on the JCV genotypes in the world, an attempt is made to infer the origins of people who carried JCV genotypes to the Japanese Archipelago. Finally, the relationships are examined between JCV genotypes and Y-chromosome haplogroups (abbreviated as Y haplogroups), as genetic variation on the Y chromosome has recently been examined in detail to investigate ancient human migrations and the population structures of human groups [29].

BASIC INFORMATION ABOUT JCV

Discovery and taxonomy

JCV was first isolated in 1971 from the brain of a patient with progressive multifocal leukoencephalopathy (PML) [30] (PML is a fatal demyelinating disease of the central nervous system. It occurs in patients with an underlying disease which decreases their immunological capacity and in those who are immunosuppressed therapeutically or after organ transplantation [31]). The new virus was named after the patient from whom it was derived.

JCV is a small DNA virus, belonging to the *Polyomaviridae* family [8]. This family at present contains 13 members. All members have capsids of similar sizes formed by three capsid proteins (i.e. VP1, VP2 and VP3). All show a similar genomic organisation (Figure 1), with many highly conserved regions, demonstrating that they originated from a common ancestor. Furthermore, all have unique natural hosts. For instance, humans are the natural host of JCV, while the rhesus monkey is the natural host of simian virus 40, a virus closely related to JCV. Thus, it is likely that the *Polyomaviridae* coevolved with their hosts, and in the case of JCV, its origin should be traceable to that of modern humans.

Ubiquitous in the human population

Soon after the isolation of JCV [30], it was found that this virus has hemagglutinating activity. Padgett and Walker [9] collected sera from various age groups in Madison, the United States, and

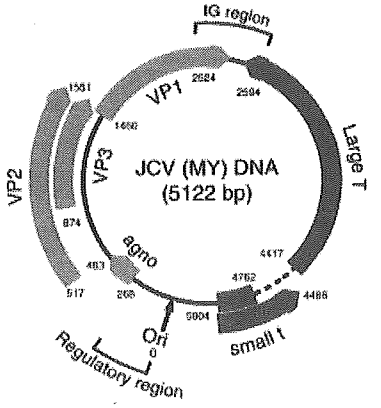


Figure 1. JCV genome. The circular double stranded DNA of an archetype strain (MY) is shown. The origin of replication (Ori), the regions encoding six viral proteins, the noncoding regulatory region and the 610-bp IG region frequently used for genotyping are indicated. From the specific sites near the origin of replication, the genome is transcribed anticlockwise and clockwise to generate early and late mRNAs, respectively. The early mRNAs are translated into two regulatory proteins, the large T and small t, while the late mRNAs are translated into the agnoprotein and three capsid proteins (VP1, VP2 and VP3). Marked changes occur in the regulatory regions of JCV isolates from the brain of PML patients (see text). The IG region spanning nucleotides 2122 to 2731 has abundant nucleotide substitutions and therefore has been used to type JCV isolates using the RFLP and the phylogenetic method

subjected them to hemagglutination inhibition tests to detect serum antibodies against JCV. Surprisingly, they found that even though PML is uncommon, antibodies against the virus are very common in the human population. The presence of antibodies (as evidence of past and present infection) indicated that infection begins in early childhood and that by middle age, 75% of individuals have been infected. In addition, no antibodies against JCV were detected in sera derived from various animals (five primates were included) [9], indicating that JCV has a very narrow host range. Similar sero-epidemiological studies have been carried out in various countries, and the results of these studies, along with those of Padgett and Walker [9], have established that JCV is ubiquitous in the human population [10].

Urinary excretion

In the early 1980s, studies using classic virological methods (cytology, electron microscopy and immunostaining) indicated that JCV is excreted in the urine of immunocompromised patients (e.g. organ transplant patients and pregnant women) [32]. Using a more sensitive method (i.e.

Southern hybridisation), Kitamura *et al.* [14] found that JCV is frequently excreted in the urine of immunocompetent patients without an obvious correlation with any diseases and that the detection rate increases with age. In most human populations, the detection rates (as assayed using PCR) in individuals aged 40 years or more were more than 50% [3,7,15,16,33,34], although lower detection rates were reported for some ethnic tribes [35–37]. Thus, urine offers a convenient source of JCV DNA for epidemiological studies of JCV.

It is emphasised that urinary JCV represents the virus circulating in the human population. This view was developed from the comparison of regulatory DNA sequences between JCVs (urine-derived JCVs) from the urine of non-immunocompromised individuals and those (PML-derived JCVs) from the brain of PML patients. Urine-derived JCV DNAs carry the archetype regulatory sequence from which various regulatory sequences of PML-derived JCVs were generated by sequence rearrangements in the course of persistence in patients [38]. Thus, urine-derived JCVs represent wild-type viruses, while PML-derived JCVs are mutant viruses that never return to the human population. For details, refer to a review [39].

Unique mode of transmission

Viruses are transmitted by two modes: horizontal and vertical transmission. The former denotes postnatal transmission between individuals, while the latter denotes transmission between parent and offspring. Vertical transmission may occur during gestation, by passage of the virus through the placenta, perinatally during birth, or from the mother through milk [40]. Serological studies [9,41,42] have shown that JCV is transmitted horizontally. Nevertheless, studies using urinary JCV DNA have provided insights into the mode of JCV transmission. (i) Kunitake *et al.* [17] PCR-amplified a 610 bp JCV DNA region (VT-intergenic region, abbreviated as the IG region) from urine specimens collected from all members of seven families. (The IG region is a region of the JCV DNA that contains relatively abundant sites for typing JCV DNAs [43].) JCV strains were identified based on the nucleotide sequences of the amplified IG regions. Strains detected in half of the JCV-positive children were identified in their parents. Furthermore, the same IG sequences were detected in the offspring as well as in the fathers or mothers at almost the same rates, suggesting that JCV transmission occurs both maternally and paternally. (ii) Kato *et al.* [20] studied whether JCV have been transmitted from the American population to the Japanese population both existing on a small island, Okinawa, Japan. No American JCV genotypes were detected in the Japanese population. (iii) Suzuki *et al.* [18] collected urine samples in Los Angeles from second and third generation Japanese-Americans whose parents and grandparents were all Japanese. From these urine samples, two JCV genotypes that predominantly occur in homeland Japanese [1] were mainly detected in each generation of the Japanese-Americans. Given these findings, it is likely that JCV is transmitted mainly from parents to children during long-term cohabitation. This mode of transmission has probably generated the close correlation, described below, between JCV genotypes and human populations.

GENOTYPES AND HUMAN POPULATIONS

Before the isolation of JCV DNAs with the archetype regulatory region [38], JCV isolates (all from the brains of PML patients) were classified as two types, types I and II, according to their sequence rearrangements in the regulatory regions [44]. Yogo *et al.* [45] proposed a new system for the classification of JCV based on the nucleotide differences distributed throughout the viral genome. Full-length JCV DNA clones were isolated at various sites in Asia, Africa and Europe, and subjected to restriction fragment length polymorphism (RFLP) analysis using nine restriction enzymes [45,46]. The results of these studies revealed that irrespective of their regulatory region rearrangements, JCV DNAs in the world can be classified into three groups (genotypes), types A, B and C. Type A is distributed throughout Europe, type B is distributed throughout Asia and Africa with a minor subtype found in Europe, and type C is located in West Africa.

To classify JCV isolates in the world into smaller groups, Sugimoto *et al.* [21] developed a method based on the DNA sequence variation in the 610 bp IG region [43]. They PCR-amplified and sequenced the IG regions of JCV DNA clones isolated either previously or directly from urine samples additionally collected. More than 200 IG sequences obtained from almost all representative

geographical sites in the Old World were used to construct a phylogenetic tree based on the neighbor-joining method [47]. The resultant tree identified 12 genotypes of JCV (EU, Af1, Af2, Af3, B1-a, B1-b, B1-c, B1–d, B2, CY, MY and SC), whose geographical distribution profiles in the Old World were unveiled [21,22].

A few years later, Stoner and his colleagues [48–50] detected three distinct JCV genotypes (2E, 8A and 8B) in the western Pacific. JCV genotype patterns in many additional sites have since been clarified [1–3,6,7,34–37,51–56; Sugimoto *et al.*, unpublished; Zheng *et al.*, unpublished; Takasaka *et al.*, unpublished]. Furthermore, as described below, some genotypes were subdivided (e.g. EU was divided into three genotypes, EU-a, -b and -c [36]).

Based on all these findings, the geographical distributions of the 18 genotypes of JCV were mapped (see Figure 2). (It should be noted that recently imported genotypes of JCV are not shown in Figure 2.) Features of the JCV genotype distribution in the world can be summarised as follows. (i) Each genotype has a unique domain. Thus, EU-a and -b are mainly spread in Europe and Mediterranean areas, and a minor genotype (B1-c) also occurs in the region; Af2 is spread throughout Africa, and in West and South Asia; Af1 and Af3 are localised to parts of Africa (West and Central Africa, respectively); and many genotypes (B1-a, -b, -d, B2, CY, MY and SC) are distributed in East Asia, with their domains partially overlapping. (ii) Multiple genotypes occur in areas where multiple human populations have intermixed. For example, European (EU-a, -b) and African (Af2) genotypes occur in North Africa facing the Mediterranean Sea; two European (EU-a/-b), an African (Af2) and an Asian (B1-b) genotype occur in West Asia, such as Turkey and Saudi Arabia; and an African (Af2) and two Asian genotypes (B2 and SC) occur in Mauritius (islands in the Indian Ocean). (iii) Closely related genotypes sometimes occur in remote areas. For instance, EU-a occurs in not only Europe and Mediterranean areas but also in Northeast Japan, South Korea, and Northeast Siberia and Arctic areas of

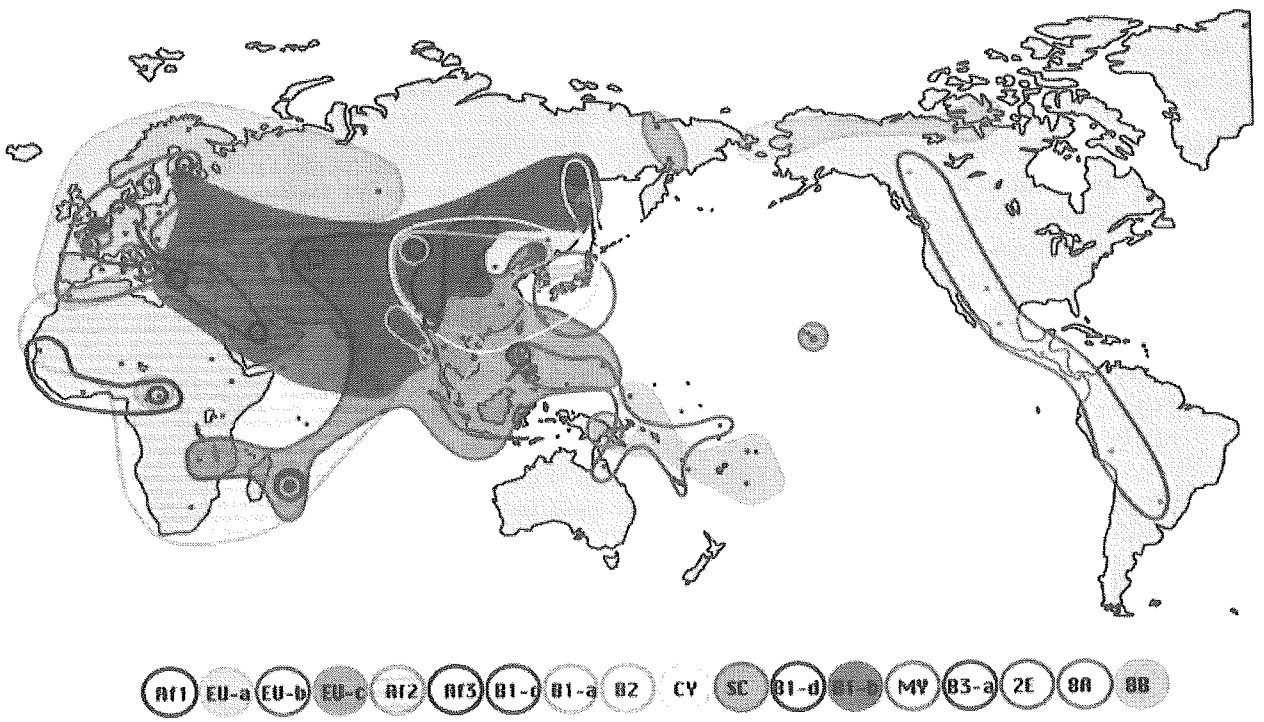


Figure 2. World map showing geographical distributions of 18 JCV genotypes in the world. Dots indicate sites from which urine samples were collected. Areas including where the same genotype was detected are shown as indicated at the bottom of the figure. This map was prepared based on published and unpublished data [1–3,6,7,21,22,34–37,48–56; Sugimoto *et al.*, unpublished; Zheng *et al.*, unpublished; Takasaka *et al.*, unpublished]