Characterization and Molecular Cloning of a Human Parvovirus Genome

Susan F. Cotmore and Peter Tattersall

Abstract. *The genome of the small human virus serologically associated with erythrocyte aplasia and erythema infectiosum (fifth disease) is shown to be a linear, nonpermuted, single-stranded DNA molecule with self-priming hairpin termini, properties which are characteristic of the genomes of the family Parvoviridae. This human parvovirus chromosome was molecularly cloned into bacterial plasmid vectors and the cloned DNA was used to explore its relatedness to other mammalian parvovirus serotypes by DNA:DNA hybridization. It is not related to the human adeno-associated viruses but does show a distant evolutionary relationship to genomes of the helper-independent parvoviruses of rodents. This strongly suggests that it is an autonomous parvovirus, and as such is the first example of a member of this group of common animal pathogens to cause disease in man.*

The parvoviruses of vertebrates fall into two distinct groups, the adeno-associated viruses (AAV), now renamed the Dependoviruses to indicate their requirement for a helper adenovirus or herpesvirus, and the larger subgroup of autonomously replicating viruses (*1*). There are three serotypes of the AAV group now known to circulate in human populations in association with their helper viruses, but these do not appear to be responsible for any human disease (*2*). Although the parvoviruses are widely distributed throughout animal species, there have been no substantiated reports of a member of the autonomously replicating subgroup infecting the human population. In animals the autonomous parvoviruses are often extremely pathogenic. Because they require, for their own replication, a host function transiently expressed during S phase of growth and are unable to induce resting cells to enter the mitotic cycle, their lytic growth is confined to proliferating cell populations. Thus they commonly infect the fetus, causing resorption, abortion, or birth defects, and those which are pathogenic for adult animals attack proliferating cells of the gut epithelium and lymphoid system (*3*). With a genome size of about 5 kilobases (kb), their limited coding capacity renders them highly dependent on the synthetic machinery and programming of the host cell for their own reproduction. It is now also clear that at least one of these required host factors is developmentally regulated, and that this underlies the strain-dependent tissue specificity that some of these viruses exhibit (*4*).

This extreme dependence on host cell differentiation and physiological state probably explains the failure to isolate and propagate putative autonomous parvoviruses involved in human disease. Several candidate human viruses have been described, including the agent known as B19 or serum parvovirus–like virus, identified by Cossart and co-workers (*5*). After its initial isolation in 1975, the agent was shown to be widespread in the United Kingdom population, but on only one occasion was it associated with illness (*6*). In 1981 interest in this agent was revived by the finding that it was responsible for aplastic crisis in children with homozygous sickle cell disease or a variety of other hereditary hemolytic anemias (*6*). Large-scale seroepidemiologic studies of this agent have been hampered by scarcity of antigen, since it has not been possible to propagate it in animals or cell culture. Nevertheless, in 1983 this same virus was found to be the etiologic agent of erythema infectiosum, or fifth disease, a highly infectious rubelliform rash of childhood (*7*). Shortly thereafter, Summers *et al.* (*8*) demonstrated that particles found in B19 viremic plasma contained a single-stranded DNA of approximately 5.5 kb, suggesting the presence of a member of the Parvoviridae.

Isolation and Characterization of B19 DNA

A sample of viremic plasma (coded Wi) from an asymptomatic blood donor was obtained from B. Cohen of the U.K. Public Health service, and virions were partially purified from this by the rapid step-gradient technique described in Fig. 1. Because of their complex terminal secondary structure, the strand length of parvoviral DNA is best analyzed under completely denaturing conditions. Concentrated B19 virions were therefore lysed in alkali, subjected to electrophoresis in alkaline agarose, and compared with the DNA of minute virus of mice (MVM) (Fig. 1a). A single species of DNA molecule with a size of 5.45 kb was associated with the B19 virion preparation. To show that the B19 genome was indeed single-stranded in the virion, we used a rapid, low salt, non-denaturing procedure for its isolation and analyzed the DNA on a neutral agarose gel (Fig. 1b). Intact virions migrated in such a gel slightly slower than the single-stranded DNA released from them by this procedure (Fig. 1b, lanes 2 and 3). When placed under optimum reannealing conditions (T_m −20°C with a 40 percent guanine plus cytosine content being assumed), these single strands efficiently annealed to form duplex DNA migrating at approximately 5.4 kb, confirming the interpretation of Summers *et al.* (*8*), who proposed that B19 packages both DNA strands in approximately equal numbers into separate virions. It is interesting that the duplex DNA formed under these conditions migrated as a doublet, perhaps reflecting alternative annealing arrangements of the terminal palindromes present in these molecules. In order to confirm that this DNA is indeed the genome of the virus identified serologically as B19 (serum parvovirus–like virus), we conducted the studies depicted in Fig. 1c, using various sera specific for this antigen. Virions containing the 5.45-kb DNA species were efficiently and specifically immunoprecipitated by diagnostic antibodies, including a monoclonal antibody directed against the B19 particle (Fig. 1c, lane 6) (*9*). As this DNA species was also resistant to digestion with deoxyribonuclease in nondisrupted preparations, we conclude that the DNA is completely contained within capsids which themselves carry the B19 antigen.

The authors are in the Department of Human Genetics, Yale University School of Medicine, New Haven, Connecticut 06510.

Properties of B19 Genomic DNA

A universal property of the parvovirus chromosome is that its terminal sequences are palindromic, which leads to the formation of hairpin duplexes at both ends of virion single-stranded DNA (*10, 11*). The hairpin at the 3′ end of each molecule can serve as a primer for a number of prokaryotic and eukaryotic DNA polymerases, allowing the synthesis of the complementary strand in vitro (*11*). To look for similar structures in B19 DNA, we examined the activity of the viral genome as a primer-template for *Escherichia coli* DNA polymerase I Klenow fragment. As shown in Fig. 2a, B19 DNA directed the synthesis of a single duplex DNA molecule with a molecular weight somewhat greater than 5 kb. When this molecule was cleaved with Bam HI, two equimolar doublet bands at 3.8 and 3.9 kb and 1.4 and 1.5 kb were detected (Fig. 2a, lane 2). However, if the Bam HI digest was denatured and quench-cooled prior to electrophoresis, the upper bands of each doublet ran as denatured DNA, while the lower bands remained unchanged in mobility (Fig. 2a, lane 3). This indicates that the lower band in each doublet represents a terminal fragment in which the template and product strands are covalently linked. These bands therefore contain the primer termini at the 3′ end of each original template strand. The upper band of each doublet represents the other end of such replication products, containing the 5′ end of each template strand, a fragment in which there is no covalent linkage between product and template strand.

This scheme of replication in vitro is outlined diagrammatically in Fig. 2c. The difference in size between the two bands in the doublet is a result of the ability of this Klenow enzyme preparation to strand-displace and copy the 5′ hairpin present on the template strand (*11*), as shown in Fig. 2c, step 2. Some Klenow preparations and most other DNA polymerases are unable to perform this type of strand-displacement synthesis and so cannot replicate this region of the parvovirus chromosome in vitro. The interpretation of these findings (Fig. 2c) is further supported by an experiment in which these same Klenow-replicated and "extended" products of B19 DNA were digested with Pst I, which cleaved the genome twice. This yielded terminal doublet bands of 3.0 and 3.1 kb and 1.4 and 1.5 kb, each of which is represented in half-molar yield, and a unique internal fragment of 0.8 kb present in molar yield (Fig. 2a, lane 4). As before, on denaturation the DNA present in the lower band of each doublet "snapped back" on quench-cooling, whereas all of the DNA present in the upper bands of each pair and all of the DNA in the internal fragment band was completely denatured.

Because *E. coli* DNA polymerase and the Klenow derivative have an editing function capable of removing unpaired bases in a 3′ to 5′ direction, the ability of B19 DNA to serve as a primer-template does not prove that the 3′ terminal nucleotides are base-paired in the hairpin. We therefore examined the ability of the genome to serve as a primer-template for avian myeloblastosis virus (AMV) reverse transcriptase, which does not possess such an editing function and requires a base-paired 3′ hydroxyl nucleotide as primer. As shown in Fig. 2b, this enzyme used B19 DNA very efficiently as a primer-template, once more synthesizing duplex DNA molecules of approximately 5.4 kb (lane 1), in which the product and template strands were covalently continuous (lane 2). Digestion of the reverse transcriptase product with Bam HI revealed only the two bands at 3.8 and 1.4 kb (lane 3) which comigrated with the "turn-around" fragments in Fig. 2a, lanes 2 and 3, since this enzyme cannot perform the strand-displacement copying of the 5′ terminal hairpins (Fig. 2c, step 2). Upon denaturation half of the DNA in each of these bands "snapped back," remaining at the same mobility (lane 4), and was resistant to digestion with the single-strand–specific mung bean nuclease, whereas the denatured form (shown as $^{\Delta}$A and $^{\Delta}$B in Fig. 2a, lane 3) were completely digested by this nuclease. Similarly, digestion of the reverse transcriptase product with Pst I now yielded three fragments (lane 5), the smallest of which was completely denaturable. Again, half of the DNA in each of the terminal fragment bands was completely denaturable and half "snapped back" on quench-cooling. This finding supports the contention that B19 packages both strands, and thus both primer 3′ terminal hairpins, in equal amounts. Further support for this comes from the quantitation of the remaining "snap back" bands in lanes 4 and 6, which indicates that they were present in equimolar amounts. Similar analyses of the duplexes synthesized in vitro were performed with the use of several different restriction endonucleases, yielding an uniquely oriented restriction map (Fig. 2d). All of these analyses were consistent with the proposal that the two individually packaged strands found in B19 virions are exact complements of each other.

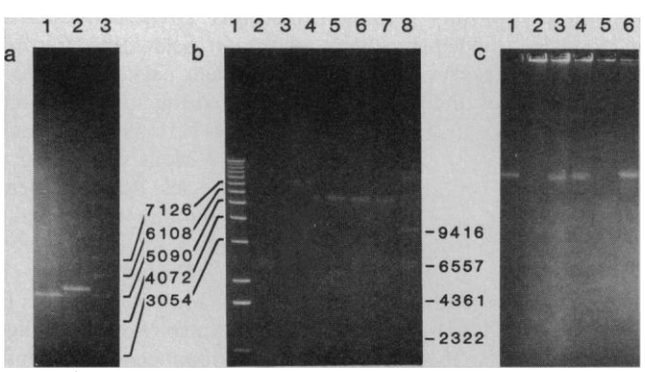


Fig. 1. Analysis of B19 genomic DNA from plasma concentrate. Plasma containing the B19 Wi isolate (*23*) was diluted with an equal volume of 100 m*M* tris and 10 m*M* EDTA (*p*H 8.7) and then the mixture was centrifuged at 1000 rev/min, for 20 minutes at 5°C in a Sorvall SS-34 rotor. The supernatant was layered on a step gradient composed of 1 ml of 10 percent glycerol, 1 ml of 30 percent glycerol, and 0.3 ml of 60 percent metrizamide, all in 50 m*M* tris and 0.5 m*M* EDTA, *p*H 8.7 (TE 8.7). Virions were recovered from the metrizamide layer after centrifugation at 3500 rev/min for 5 hours at 5°C in an SW41 rotor and dialyzed against TE 8.7. (a) Concentrated virions lysed in 0.03*N* NaOH for 15 minutes at room temperature were subjected to electrophoresis in a 0.8 percent alkaline agarose gel (lane 2), compared with MVM DNA (lane 1), and polymerized 1-kb unit markers (Bethesda Research Laboratories) (lane 3) after staining with ethidium bromide. (b) DNA was isolated from virions by digestion in 1 percent sodium dodecyl sulfate with Proteinase K (1 mg/ml) for 5 minutes at room temperature. After rapid deproteinization by centrifugation through a column of Sephacryl S-300 (Pharmacia) in 10 m*M* tris and 1 m*M* EDTA, *p*H 7.5, the DNA solution was made 50 m*M* NaCl and annealed at 57°C. At various times portions were analyzed by electrophoresis in a 1 percent neutral agarose gel. (Lane 1) One-kilobase ladder marker; (lane 2) intact B19 virions; (lane 3) isolated DNA, prior to annealing; (lane 4) annealed 1.5 hours; (lane 5) annealed 3 hours; (lane 6) annealed 4.5 hours; (lane 7) annealed 6 hours; (lane 8) heat-denatured Hind III digest of phage λ DNA. (c) Virion samples immunoprecipitated by the technique of Kessler (*24*) were subjected to electrophoresis on a 1 percent alkaline agarose gel. (Lane 1) Concentrated virus marker; (lane 2) immunoprecipitate, no serum; (lane 3) immunoprecipitate, B19-positive P serum; (lane 4) immunoprecipitate, B19-positive BD45642 serum; (lane 5) immunoprecipitate, B19-negative serum; (lane 6) immunoprecipitate, α-B19 monoclonal antibody (9).

Molecular Cloning of the B19 Genome

The insertion of a linear single-stranded DNA genome with terminal hairpins into a circular double-stranded plasmid vector poses some technical problems which have been overcome in a number of ways in the cloning of other parvovirus genomes (*12, 13*). In cloning the genome of the murine parvovirus MVM, we had previously found that it was not possible to obtain both complete terminal sequences in the same clone (*13*). Rather than attempting to clone the entire genome of B19 in one operation, therefore, we cloned each terminus separately, exploiting the ability of some preparations of *E. coli* polymerase Klenow fragment to copy the 5′ terminal hairpin by strand-displacement (as shown in Fig. 2a). Synthesis of the extended 5′ palindrome, duplex form of B19 DNA from single-stranded DNA was accomplished as described in the legend to Fig. 2. The product of this reaction was ligated to Bam HI recombination linkers and cleaved with Bam HI. The resulting fragments were ligated into the Bam HI site of plasmid pAT153 (*14*), transformed into *E. coli* strain LE392, and cultured in the presence of ampicillin (*13*). Ampicillin-resistant colonies were isolated and their plasmids screened by the procedure of Barnes (*15*), yielding two plasmids containing B19-specific inserts, pYT101 and pYT102, corresponding in size and restriction pattern to the viral 1.5-kb and 3.9-kb Bam HI fragments, respectively.

Analysis both soon after cloning and after subsequent amplification in *E. coli* showed that each of these clones underwent a series of deletion events involving up to 200 base pairs at or near the viral terminus. This behavior has been noted previously for the right-hand end (viral 5′) palindrome of MVM DNA cloned in a bacterial plasmid (*13*), where selection for the deletion of the central portion, but by no means all, of the palindrome appears to occur. This is not, however, a general property of cloned parvoviral palindromes, since the left-hand end (viral 3′) palindrome of MVM and the terminally redundant palindromes of AAV are stable when cloned into plasmid vectors (*12, 13*). The reason for the instability of the MVM right-hand palindrome and the mode of its deletion in *E. coli* remain to be determined, but this property appears to be shared with both ends of the cloned B19 genome.

Relatedness of B19 to Other Mammalian Parvoviruses

We explored the relatedness of B19 to other distinct parvovirus serotypes by using a mixture of the two clones to probe a blot of several representative parvovirus DNA's separated by electrophoresis in an alkaline agarose gel (Fig. 3). Several identical blots were hybridized and washed under different stringency conditions essentially as described by Howley *et al.* (*16*). At relatively low stringency (Fig. 3c), cross-hybridization was detected between B19 and KRV, H-1, CPV, PPV, MVM, TVX, and LuIII, whereas no signal was detected with BPV or AAV-2. Under these conditions regions of identity containing up to 25 percent mismatch were detected. Raising the stringency to allow less than 15 percent mismatch (Fig. 3b) did not eliminate the cross-hybridization between B19 and KRV, H-1, MVM, or LuIII very significantly, but did reduce or abolish the signal with CPV, PPV, and TVX. Comparison of the signals obtained for heterologous as opposed to homologous hybridization (Fig. 3) indicated that the homology between B19 DNA and other parvovirus genomes is extremely limit-

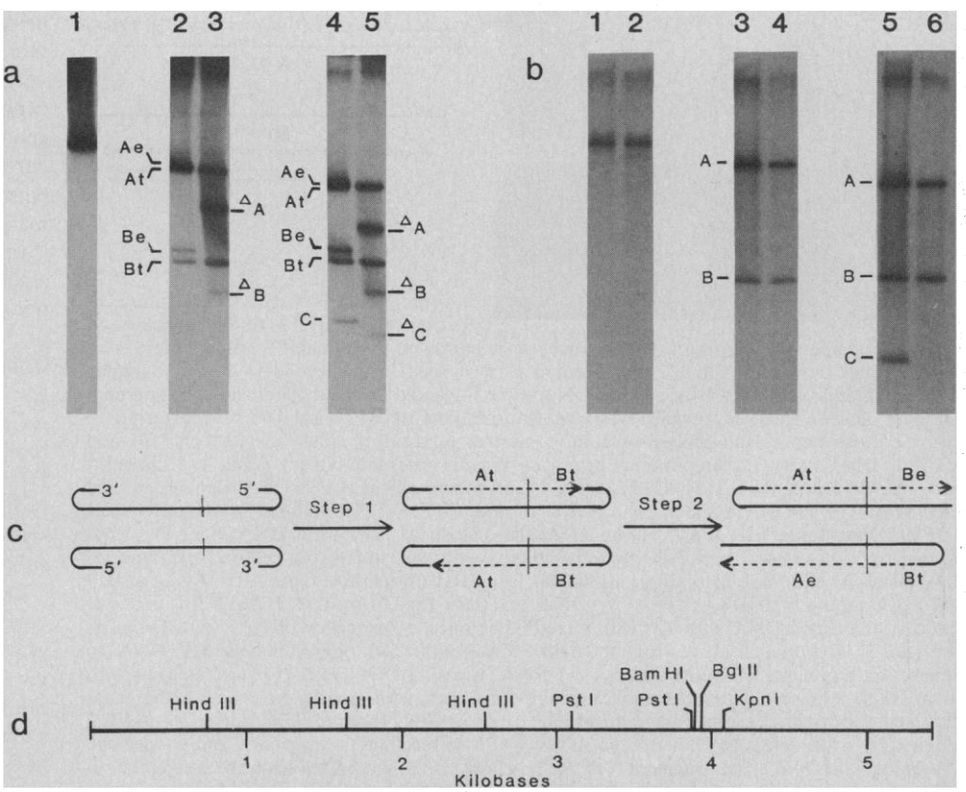
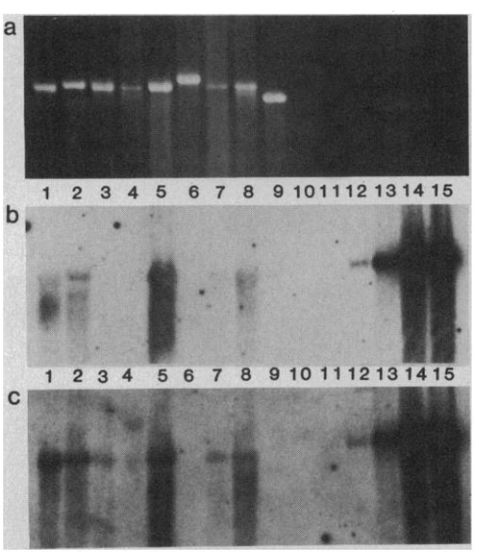


Fig. 2. Synthesis of B19 DNA in vitro. DNA isolated from concentrated virus by sedimentation in alkaline sucrose (*12*) was used to prime synthesis of complementary strands in the presence of ^{32}P-labeled deoxycytidine triphosphate by different DNA polymerases. The reaction product and its restriction endonuclease fragments (*25*) were analyzed by electrophoresis in 1.4 percent neutral agarose gels and by autoradiography, and molecular weights were calculated from the relative migration of ethidium bromide–stained marker fragments in the same gel. (a) Product synthesized by *E. coli* DNA polymerase I Klenow fragment. (Lane 1) Undigested; (lane 2) digested with Bam HI; (lane 3) as in lane 1, heat-denatured (90°C, 2 minutes), then quench-cooled; (lane 4) digested with Pst I; (lane 5) as in lane 4, heat-denatured, then quench-cooled. Suffixes e and t denote the "extended" and "turn-around" forms of terminal restriction fragments; Δ denotes the position of denatured restriction fragments. (b) Product synthesized by AMV reverse transcriptase. (Lane 1) Undigested; (lane 2) as in lane 1, heat-denatured, quench-cooled, digested with mung bean nuclease; (lane 3) digested with Bam HI; (lane 4) as in lane 3, heat-denatured, quench-cooled, digested with mung bean nuclease; (lane 5) digested with Pst I; (lane 6) as in lane 5, heat-denatured, quench-cooled, digested with mung bean nuclease. (c) Cartoon of DNA synthesis on B19 single-stranded DNA. The solid line indicates the template (virion) strand with its terminal hairpin duplexes. The dotted line represents the complementary strand synthesized by DNA polymerase and the arrow indicates the direction of synthesis (5′ → 3′). The vertical slash shows the position of the Bam HI site, and A_e, A_t, B_e, and B_t indicate Bam HI fragments. (d) Partial restriction map of the B19 (Wi isolate) genome. Endonucleases found not to cut within the genome were Cla I, Eco RI, Eco RV, Pvu I, Sac I, and Sal I.

ed. The strongest cross-hybridization, obtained with MVM, is only in the order of 0.1 percent of the homologous reaction. Given the genome size of about 5 kb, this indicates that B19 probably shares with the rodent parvoviruses a short sequence (or sequences) with close homology, suggesting a distant evolutionary relationship.

The B19 clones described here were derived from virus present in plasma obtained in the early 1970's from an asymptomatic British blood donor (*5*). They hybridize with 100 percent efficiency to DNA present in plasma obtained in 1982 from a patient in aplastic crisis from Augusta, Georgia [shown as B19 (Au) in Fig. 3, lane 15] demonstrating their efficacy in the detection of B19 viremia.

The DNA homology between B19 and MVM was next used to try to determine the genetic layout of the B19 genome with respect to transcriptional orientation. The coding strategy of the MVM genome is known in some detail and the arrangement of its overlapping transcription units is depicted in Fig. 4. Transcription is unidirectional (*17*) and the genome, which has been sequenced (*18*), is arranged in two major coding blocks, with a nonstructural protein encoded in the left half and the capsid polypeptides in the right half (*19*). However, when each of the two B19 arms was used separately to probe blots of Eco RI–digested MVM DNA, identity was detected by both probes in all three MVM Eco RI fragments, although the 1.5-kb B19 probe gave a lower signal with the right-hand end of MVM. The consistent signal with both probes for the middle 50 percent and left-hand 20 percent of the MVM genome suggests that this identity does not reflect a single region of conserved coding sequence and thus cannot be used to orient the B19 genome. Again AAV DNA fragments were not detected with either B19 probe.



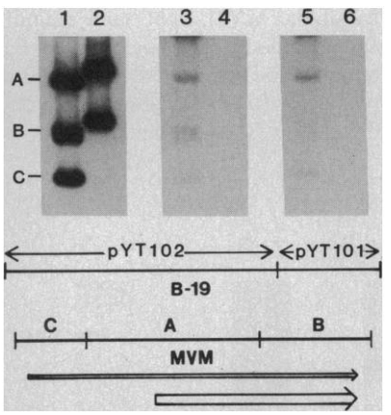


Fig. 3 (left). Cross-hybridization of cloned B19 DNA to the genomes of other mammalian parvoviruses. Parvoviruses were extracted from infected cell monolayers by freezing and thawing in TE 8.7 and purified as described in the legend to Fig. 1. Samples containing approximately 1 μg of viral DNA were subjected to electrophoresis in 1.4 percent alkaline agarose gels, in parallel with various dilutions of B19 viral DNA, transferred to nitrocellulose paper, and probed with an equimolar mixture of nick-translated pYT101 and pYT102 DNA's (*26*). (a) An ethidium bromide–stained representative gel. (Lane 1) Kilham rat virus (KRV) DNA; (lane 2) H-1 DNA; (lane 3) canine parvovirus (CPV) DNA; (lane 4) porcine parvovirus (PPV) DNA; (lane 5) minute virus of mice (MVM) DNA; (lane 6) bovine parvovirus (BPV) DNA; (lane 7) TVX parvovirus DNA; (lane 8) Lu III parvovirus DNA; (lane 9) adeno-associated virus (AAV) type 2 DNA; (lane 10) concentrated, uninfected cell extract; (lane 11) 0.00001 μg of B19 (Wi) DNA; (lane 12) 0.0001 μg of B19 (Wi) DNA; (lane 13) 0.001 μg of B19 (Wi) DNA; (lane 14) 0.01 μg of B19 (Wi) DNA; and (lane 15) 0.01 μg of B19 (Au) DNA. (b and c) Blots of the above DNA's probed with mixed B19 clones at stringency of T_m − 20°C (panel b) and T_m − 35°C (panel c). Fig. 4 (right). Cross-hybridization of cloned B19 DNA to restriction fragments of MVM and AAV-2 DNA. Cloned MVM DNA (*14*) was cleaved with Bam HI to release it from plasmid sequences and then with Eco RI to generate the three fragments shown in the lower diagram, which also shows the primary MVM transcripts. AAV-2 duplex DNA was cleaved with Eco RI to generate two terminal fragments (A and B) and one extremely small internal fragment (off gel). These DNA preparations were subjected to electrophoresis in 1.4 percent neutral gels, blotted, and probed with (lanes 1 and 2) an equimolar mixture of nick-translated MVM and AAV-2 DNA's; (lanes 3 and 4) nick-translated pYT102 DNA; and (lanes 5 and 6) nick-translated pYT101 DNA. The blot was processed at a stringency of T_m − 20°C. (Lanes 1, 3, and 5) MVM fragments; (lanes 2, 4, and 6) AAV-2 fragments. The exposure time for the autoradiographs in lanes 3 through 6 was 30 times longer than that for lanes 1 and 2.

Discussion

We have shown that the genome of B19 is a nonpermuted, single-stranded DNA molecule with terminal hairpin duplexes that is capable of priming synthesis of its complementary strand in vitro. These properties allow the virus to be unambiguously assigned to the Parvoviridae family (*10*). The lack of any detectable hybridization between the cloned form of the B19 genome and AAV-2 DNA suggests that B19 does not belong to the Dependovirus subgroup, since AAV genomes cross-hybridize extensively, even between serotypes specific for monkey or man (*2, 20*).

That the B19 genome hybridizes with KRV, H-1, MVM, and LuIII DNA's indicates that B19 belongs to the autonomously replicating parvovirus subgroup. Its ability to hybridize equally well to all four of these viruses is not unexpected, since these four have been shown, by heteroduplex analysis, to be about 70 percent identical (*21*). It is the nature of the sequence shared between B19 and these viruses that is perhaps surprising. The cross-hybridization is extremely weak, indicating a very short region of homology. However, Fig. 4 shows that this homology is dispersed, both in MVM and in B19, and thus perhaps reflects the remnants of a series of repeated sequences in a common ancestral virus from which they separately diverged. It is interesting that the B19 genome is more closely related to the genomes of the autonomously replicating parvoviruses of rodents than to those of parvoviruses infecting domesticated animals. The relation between B19 and the "orphan" virus LuIII is also interesting since seroepidemiological studies have not yet identified the natural host species of LuIII, despite its original isolation from human cell cultures (*3*). Nevertheless, it is significant to the present analysis that, like B19, this autonomously replicating parvovirus encapsidates both DNA strands in separate virions (*22*), a property previously thought to be exclusive to members of the Dependovirus subgroup (*10*).

Definitive proof of the ability of B19 virus to replicate autonomously awaits a growth system in vitro in which helper viruses can be rigorously excluded. Studies in which the virus has been shown to inhibit the formation of colonies derived from both early (BFU-E) and late (CFU-E) erythroid progenitors, but not colonies derived from granulocyte-macrophage progenitors (CFU-C) have gone some way to solving this problem, as well as identifying as a target

cell for B19 a precursor lying between BFU-E and CFU-E in the erythropoietic pathway (*23*). This inhibition does not require coinfection with an exogenously added helper virus, but it remains to be determined whether this target cell supports a fully productive, lytic infection by B19. That B19 should show such an extreme target cell specificity is not surprising, since susceptibility to lytic parvovirus infection has been shown in several cases to be a function of host cell differentiation (*3, 4*). The identification of B19 as the causative agent of erythema infectiosum and its probable transmission by way of the upper respiratory tract suggest that this virus will also be found to replicate at other sites in the body. Now that cloned copies of the B19 genome are available many questions about the virus, its distribution, and its disease-producing potential become experimentally accessible.

References and Notes

1. P. Tattersall and D. C. Ward, in *Replication of Mammalian Parvoviruses*, D. C. Ward and P. Tattersall, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, 1978), pp. 3–12.
2. M. D. Hoggan, *Prog. Med. Virol.* **12**, 211 (1970); U. Bantel-Schaal and H. Zur Hausen, *Virology* **134**, 52 (1984).
3. G. Siegl, in *The Parvoviruses*, K. I. Berns, Ed. (Plenum, New York, 1984), pp. 297–362.
4. P. Tattersall and J. Bratton, *J. Virol.* **46**, 944 (1983); B. A. Spalholz and P. Tattersall, *ibid.*, p. 937.
5. Y. E. Cossart *et al.*, *Lancet* **1975-I**, 72 (1975).
6. N. Young and P. Mortimer, *Blood* **63**, 729 (1984).
7. M. J. Anderson *et al.*, *Lancet* **1983-I**, 1378 (1983).
8. J. Summers, S. E. Jones, M. J. Anderson, *J. Gen. Virol.* **64**, 2527 (1983).
9. B. J. Cohen, P. P. Mortimer, M. S. Pereira, *J. Hyg.* **91**, 113 (1983).
10. P. A. Bachmann *et al.*, *Intervirology* **11**, 248 (1979); G. Siegl *et al.*, *ibid.*, in press.
11. G. J. Bourguignon *et al.*, *J. Virol.* **20**, 290 (1976); P. Tattersall and S. F. Cotmore, unpublished.
12. R. J. Samulski *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2077 (1982); C. A. Laughlin *et al.*, *Gene* **23**, 65 (1983).
13. M. J. Merchlinsky *et al.*, *J. Virol.* **47**, 227 (1983).
14. A. J. Twigg and D. Sherratt, *Nature (London)* **283**, 216 (1980).
15. W. M. Barnes, *Science* **195**, 393 (1977).
16. P. M. Howley *et al.*, *J. Biol. Chem.* **254**, 4876 (1979).
17. D. Pintel *et al.*, *Nucleic Acids Res.* **11**, 1019 (1983).
18. C. R. Astell *et al.*, *ibid.*, p. 999.
19. S. F. Cotmore, L. J. Sturzenbecker, P. Tattersall, *Virology* **129**, 333 (1983).
20. J. A. Rose *et al.*, *J. Virol.* **2**, 999 (1968).
21. P. T. Banerjee *et al.*, *J. Mol. Biol.* **166**, 257 (1983).
22. D. E. Muller and G. Siegl, *J. Gen. Virol.* **64**, 1043 (1983); R. C. Bates *et al.*, *J. Virol.* **49**, 319 (1984).
23. P. P. Mortimer *et al.*, *Nature (London)* **302**, 426 (1983); N. Young *et al.*, *J. Clin. Invest.* **73**, 224 (1984); N. Young *et al.*, *ibid.*, in press.
24. S. W. Kessler, *J. Immunol.* **115**, 1617 (1975).
25. Restriction endonuclease and mung bean nuclease digestions were performed as suggested by the suppliers, New England Biolabs and P-L Biochemicals, respectively.
26. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1975); P. W. J. Rigby *et al.*, *ibid.* **113**, 237 (1977).

27. We thank B. J. Cohen for supplying B19 viremic plasma, P-serum, and monoclonal antibody. We also thank M. J. Anderson for B19-positive BD45642 serum and B19-negative serum, B. J. Carter for AAV-2 DNA, C. Joyce for Klenow enzyme preparations, G. Siegl, S. Rhode, S. M. Halling, and T. Molitor for providing virus seed stocks, J. Bratton for assistance with virus production, and D. DiMaio, A. L. Smith, and D. C. Ward for reading the manuscript. This work was supported by grant CA29303 from the National Institutes of Health and by a grant from the Cancer Research Institute.

29 May 1984; accepted 9 July 1984

RESEARCH ARTICLE

Molecular Characterization of Human T-Cell Leukemia (Lymphotropic) Virus Type III in the Acquired Immune Deficiency Syndrome

George M. Shaw, Beatrice H. Hahn, Suresh K. Arya
Jerome E. Groopman, Robert C. Gallo
Flossie Wong-Staal

The human retrovirus that has been isolated repeatedly from patients with the acquired immune deficiency syndrome (AIDS) and from persons at risk for the disease (*1–5*) shares many of the biological and physicochemical properties common to the human T-cell leukemia (lymphotropic) viruses (HTLV). These properties include tropism for T-lymphocytes, induction of multinucleated giant cells, a Mg^{2+} preferring reverse transcriptase of high molecular weight, a relatively small major core protein (molecular weight 24,000; p24), distant antigenic and nucleic acid homology, and a likely African origin (*1–4, 6–10*). Because of these similarities, and because of the uniform nomenclature adopted for the HTLV family of retroviruses (*11*), the AIDS associated virus was called HTLV-III.

Detailed characterization of HTLV-III and serologic testing of large numbers of patients with AIDS or AIDS-related complex (ARC) became possible when it was found that the virus could be transmitted to a human T-cell line, H9, that is largely resistant to the cytopathic effects of the virus but is a good virus producer (*1*). This cell line has served as the principal source of viral reagents for several seroepidemiological studies of HTLV-III in AIDS (*3–5, 12–16*), and as the source of virus in the present study of the molecular biology of the AIDS agent. In this article, we describe the molecular cloning of two full-length integrated proviral DNA forms of HTLV-III and an analysis of the HTLV-III genome in cell lines and fresh tissues from patients with AIDS or ARC.

Molecular Cloning of the HTLV-III Provirus

Sequences of HTLV-III were first detected in DNA of infected H9 cells (H9/HTLV-III) by Southern blot analysis with the use of a ^{32}P-labeled complementary DNA (cDNA) probe prepared from HTLV-III virions. There was no evidence of such sequences in uninfected H9 cells. The identity of these sequences in H9/HTLV-III was subsequently confirmed by using as a probe the cloned genome of HTLV-III derived from unintegrated linear viral DNA (*9*). Preliminary analyses of Southern digests of H9/HTLV-III DNA revealed that the virus was present in this cell line both as unintegrated DNA and as proviral DNA integrated into the cellular genome at multiple different sites. Since the HTLV-III provirus was found to lack Xba I restriction sites, a genomic library was constructed by using Xba I–digested

G. M. Shaw, B. H. Hahn, S. K. Arya, R. C. Gallo, and F. Wong-Staal are in the Laboratory of Tumor Cell Biology, Developmental Therapeutics Program, Division of Cancer Treatment, National Cancer Institute, Bethesda, Maryland 20205. J. E. Groopman is in the Division of Hematology and Oncology, Department of Internal Medicine, New England Deaconess Hospital, Boston, Massachusetts 02255.