

# Computational Analysis of Human Adenovirus Type 22 Provides Evidence for Recombination among Species D Human Adenoviruses in the Penton Base Gene<sup>∇</sup>

Christopher M. Robinson,<sup>1</sup> Jaya Rajaiya,<sup>1</sup> Michael P. Walsh,<sup>2</sup> Donald Seto,<sup>2</sup> David W. Dyer,<sup>3</sup> Morris S. Jones,<sup>4</sup> and James Chodosh<sup>1\*</sup>

*Department of Ophthalmology, Howe Laboratory, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston, Massachusetts<sup>1</sup>; Department of Bioinformatics and Computational Biology, George Mason University, Manassas, Virginia<sup>2</sup>; Department of Microbiology and Immunology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma<sup>3</sup>; and Clinical Investigation Facility, David Grant USAF Medical Center, Travis, California<sup>4</sup>*

Received 17 April 2009/Accepted 18 June 2009

**Recombination in human adenoviruses (HAdV) may confer virulence upon an otherwise nonvirulent strain. The genome sequence of species D HAdV type 22 (HAdV-D22) revealed evidence for recombination with HAdV-D19 and HAdV-D37 within the capsid penton base gene. Bootscan analysis demonstrated that recombination sites within the penton base gene frame the coding sequences for the two external hypervariable loops in the protein. A similar pattern of recombination was evident within other HAdV-D types but not other HAdV species. Further study of recombination among HAdVs is needed to better predict possible recombination events among wild-type viruses and adenoviral gene therapy vectors.**

Adenoviruses were first isolated in 1953 (18, 31) and currently cause an array of human diseases. These include respiratory, gastrointestinal, and ocular surface infections, opportunistic infections in immune-deficient individuals, and possibly, obesity (9, 13, 14, 20, 40). Adenoviruses have also recently been used as gene therapy vectors (19, 35). Thus, while adenoviruses continue to cause significant morbidity and mortality in the human population, their existence also provides a potential benefit for the treatment of patients with an even broader range of ailments.

Since the characterization of the first human adenovirus (HAdV), 53 types have been identified and subsequently classified into seven species (A to G) on the basis of serology, restriction endonuclease digestion patterns, and to a lesser degree, genotyping. Recently, high-throughput sequencing technology has made whole-genome sequencing both rapid and affordable (27). However, the genomic sequences of 23 out of 53 HAdV types remain to be determined, with most of those within species D.

Species D HAdV type 22 (HAdV-D22) was originally isolated from a child in 1956 (3). While it is not clear what role HAdV-D22 plays in human disease, one report revealed a possible tropism for the eye (32). Recently, recombination with HAdV-D22 has been identified as the source of a novel HAdV, HAdV-D53, causing an outbreak of keratoconjunctivitis in Germany (37). HAdV-D22 recombination was also identified in a possible variant of HAdV-D53 that was isolated from a patient in Japan (2). Therefore, HAdV-D22 has shown the propensity to recombine with other viruses, with clinically im-

portant consequences. The emergence of new pathogenic HAdV genotypes, along with continued interest in HAdVs as vectors for human gene therapy, make adenovirus recombination a critically important issue.

HAdV-D22 was acquired from the American Type Culture Collection (Manassas, VA). The complete genome of the prototype strain AV-2711 (ATCC VR-1100) was sequenced on an Applied Biosystems (Foster City, CA) 3730 XL DNA sequencer in the Laboratory for Genomics and Bioinformatics at the University of Oklahoma Health Sciences Center using a previously described protocol (29). The sequence was validated by sequencing on an ABI SOLiD DNA sequencer. Sequences from both methodologies were 100% identical and provided 7,727-fold coverage for the genome.

The mVISTA Limited Area Global Alignment of Nucleotides (LAGAN) tool (<http://genome.lbl.gov/vista/index.shtml>) was used to align and compare the whole HAdV genomes (6) of HAdV-D22 and each of the other nine completely sequenced HAdV-Ds. Analysis revealed sequence diversity in the penton base, hexon, E3, and fiber open reading frames (Fig. 1). Surprisingly, comparison of HAdV-D22 to HAdV-D19 strain C (30) and HAdV-D37 (29) revealed considerable sequence conservation in the penton base gene (Fig. 1 and 2A and B).

Based on the sequence conservation seen in the penton base, Simplot 3.5.1 (<http://sray.med.som.jhmi.edu/SCROftware/simplot/>) and Recombination Detection Program (RDP) version 3.34 (<http://darwin.uvigo.es/rdp/rdp.html>) were used to identify possible recombination sites (21, 24). Bootscan analysis identified two possible recombination loci in the HAdV-D22 penton base, the first in HAdV-D37, encompassing nucleotides 400 to 600, and the second in HAdV-D19 at nucleotides 750 to 1350 (Fig. 3A). In silico amino acid analysis showed that these two probable recombination areas code for the two variable loops in the penton base protein (Fig. 3A), located on the exterior of the viral capsid (16, 44).

\* Corresponding author. Mailing address: Massachusetts Eye and Ear Infirmary, 243 Charles Street, Boston, MA 02114. Phone: (617) 573-6398. Fax: (617) 573-4324. E-mail: james\_chodosh@meei.harvard.edu.

<sup>∇</sup> Published ahead of print on 24 June 2009.

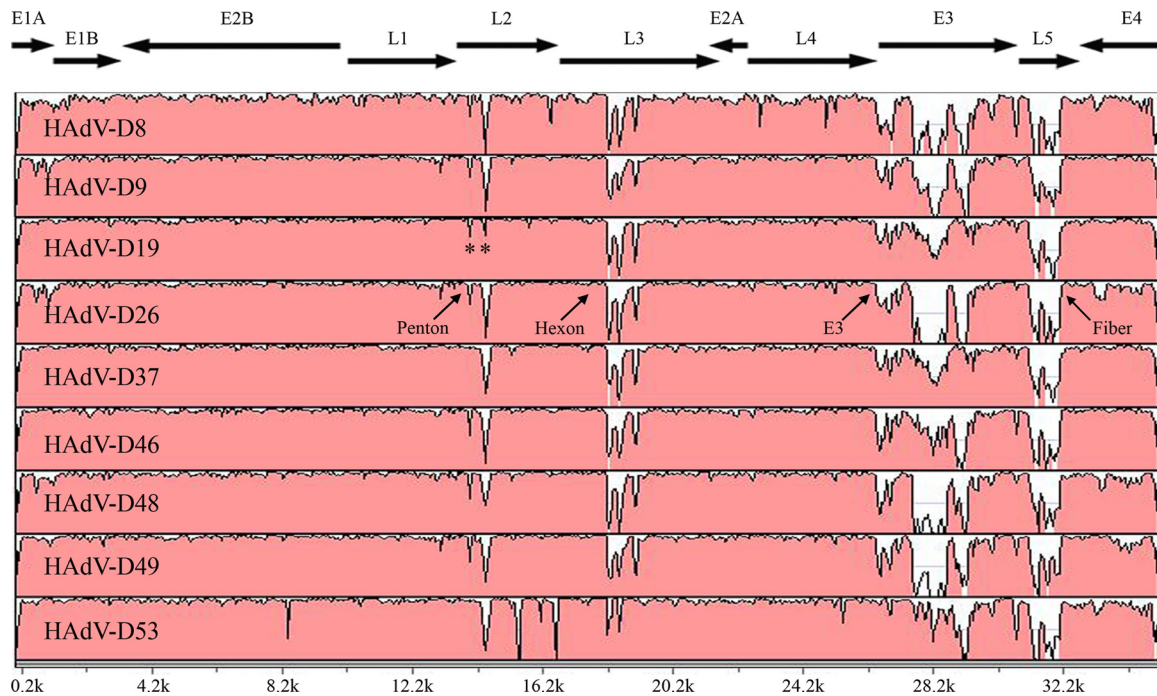


FIG. 1. Global pairwise sequence alignment of HAdV-D22 with nine other HAdV-D types. The percent sequence conservation is reflected in the height of each data point along the y axis. A conserved sequence in the penton base open reading frame is designated by an asterisk. GenBank accession numbers are as follows: HAdV-D8, AB448768; HAdV-D9, AJ854486; HAdV-D19, EF121005; HAdV-D26, EF153474; HAdV-D37, DQ900900; HAdV-D46, AY875648; HAdV-D48, EF153473; HAdV-D49, DQ393829; and HAdV-D53, FJ169625.

We extended our investigation to determine if these recombination sites were common to penton base genes of other HAdV genomes. We found evidence for recombination between HAdV-D9 and HAdV-D26 in both the nucleotide 400

to 600 and 750 to 1350 regions (Fig. 3B), between HAdV-D49 and HAdV-D48 in the nucleotide 400 to 600 region, and between HAdV-D49 and HAdV-D8 in the nucleotide 750 to 1350 region (Fig. 3C). A similar pattern was observed in one or both of

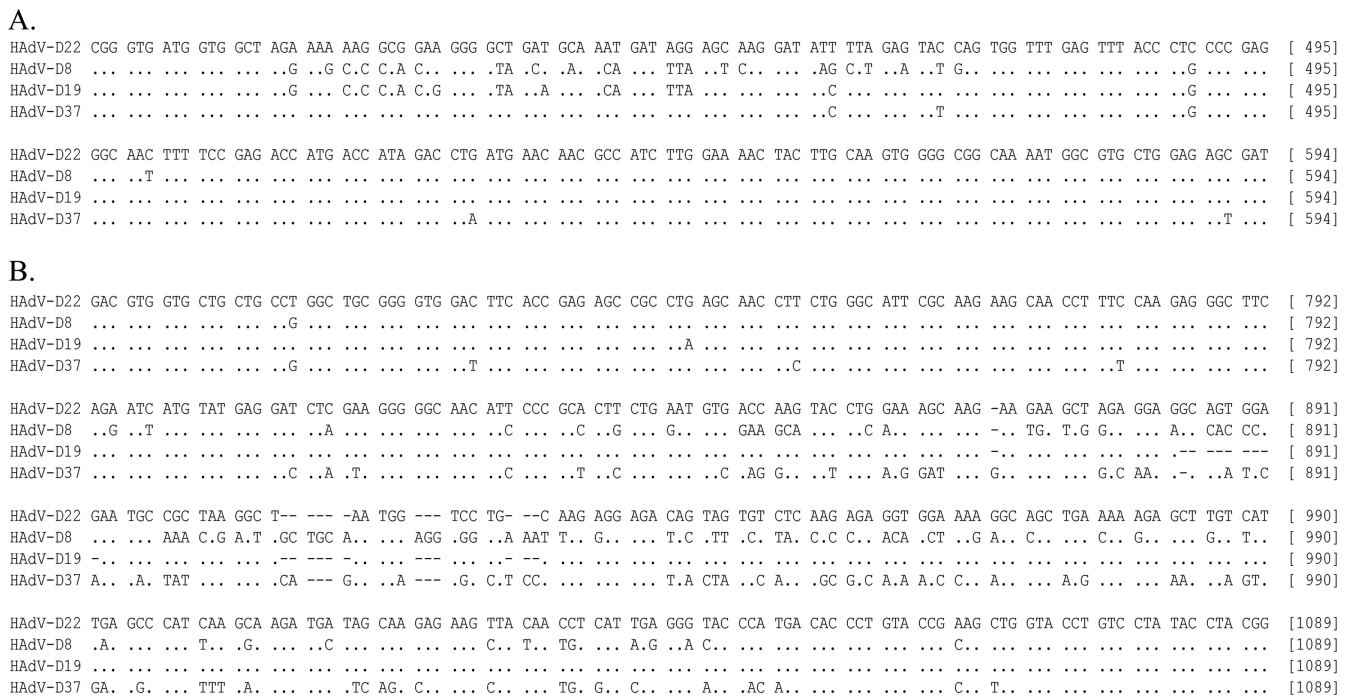
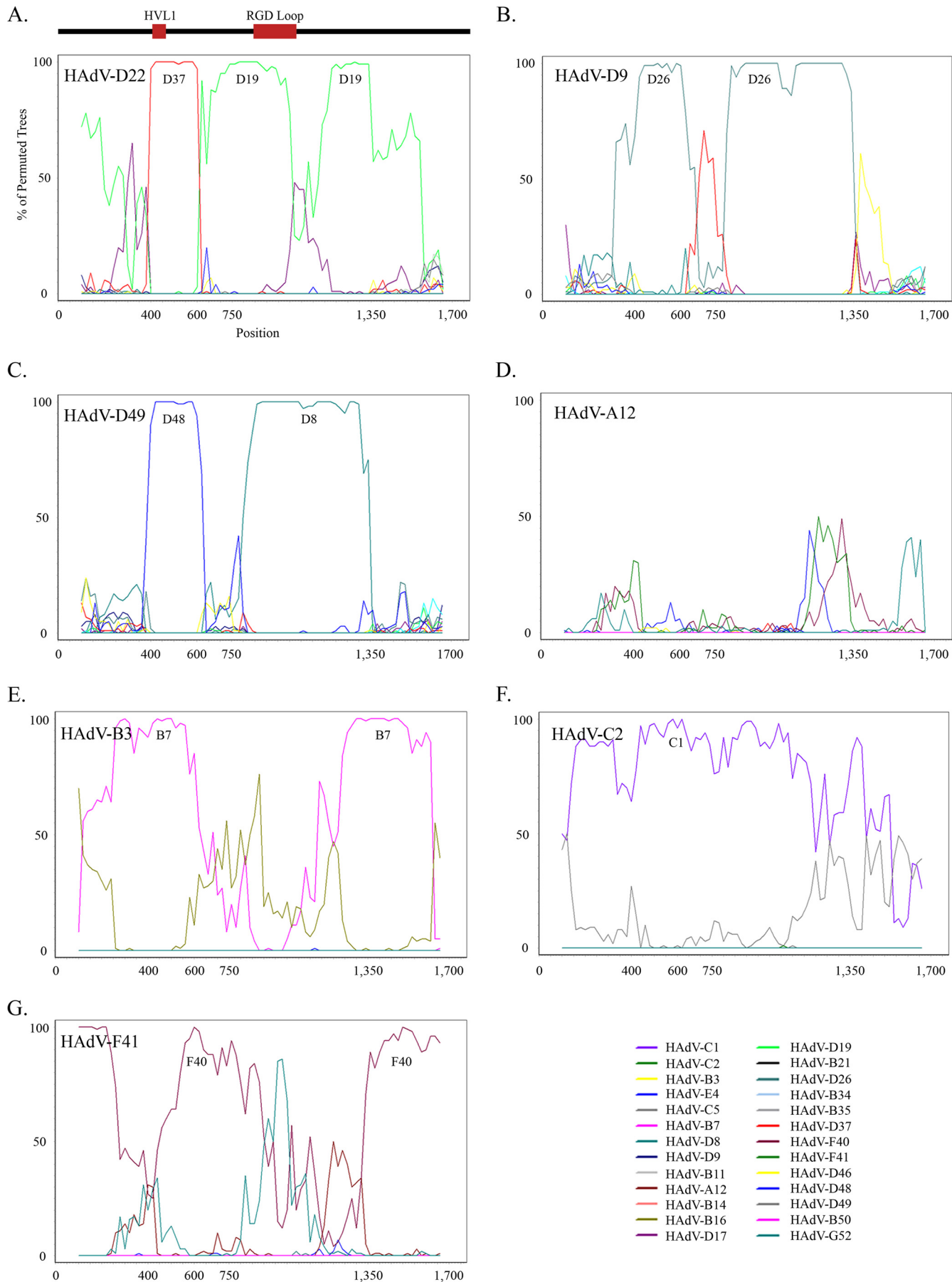


FIG. 2. Multisequence alignment of HAdV-D penton base genes. Nucleotide alignment of the bp 397 to 594 (A) and bp 694 to 1089 (B) regions of the penton base gene.



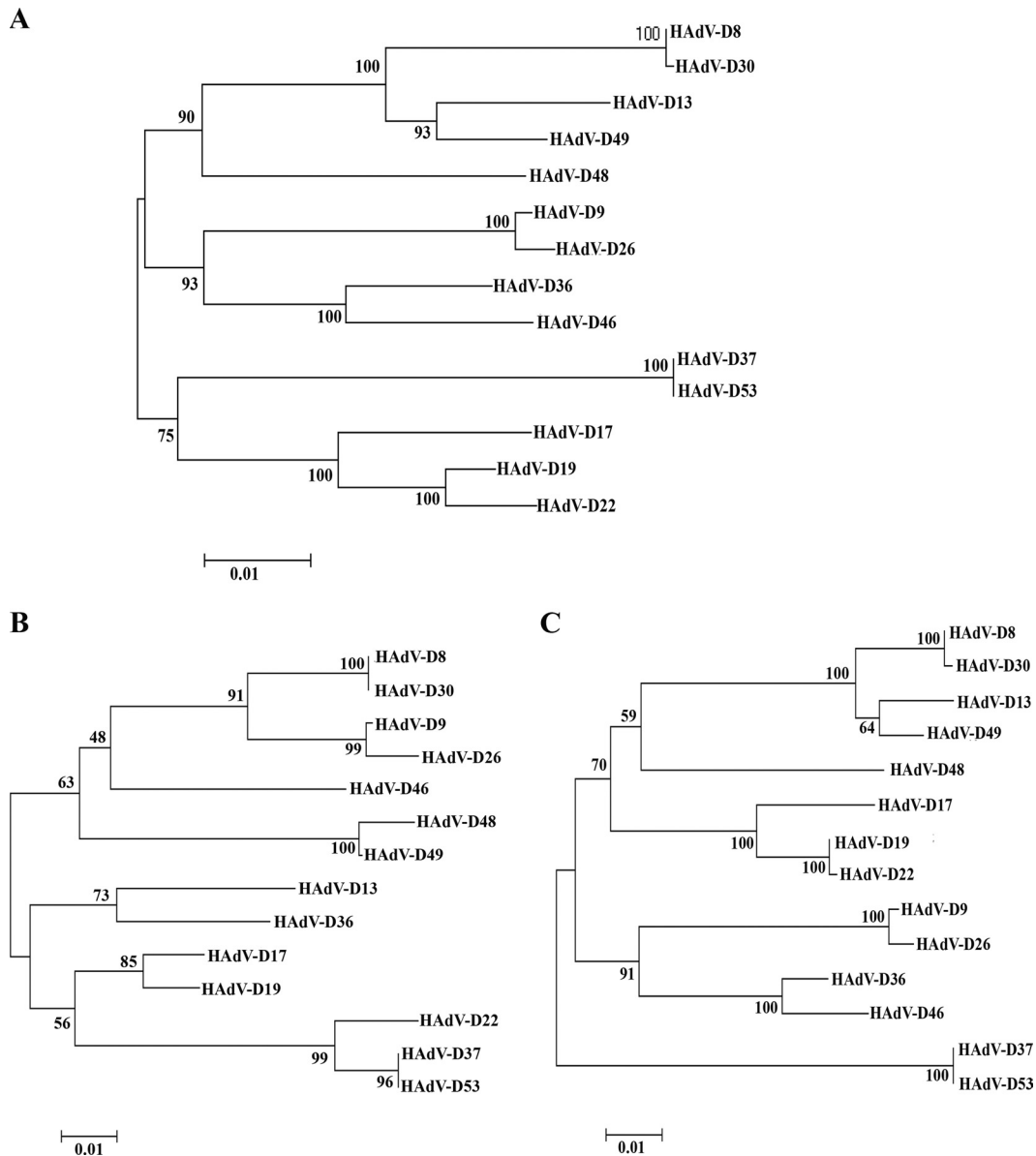


FIG. 4. Phylogenetic analysis of HAdV-D penton base genes. Bootstrap neighbor-joining tree designed using MEGA 4.0.2. The Gonnet protein weight matrix in ClustalX alignment was used, along with complete deletion options. Percent bootstrap confidence levels (1,000 replicates) are shown on the relevant branches. (A) Analysis of sequenced HAdV-D penton base genes. (B) Analysis of the nucleotide 400 to 600 region of sequenced HAdV-D penton base genes. (C) Analysis of the nucleotide 750 to 1350 region of sequenced HAdV-D penton base genes.

these nucleotide regions for HAdV-D8, -17, -19, -26, -37, and -48 (data not shown). Remarkably, this pattern of recombination in the penton base gene was unique to HAdV species D (Fig. 3D, E, F, and G).

Bootstrap-confirmed neighbor-joining phylogenetic trees of HAdV-D penton base genes were constructed with Molecular

Evolutionary Genetics Analysis (MEGA) 4.0.2 (<http://www.megasoftware.net/index.html>) to examine viral evolution in HAdV-D (34). Analysis of the entire penton gene revealed a close relationship of HAdV-D22 and HAdV-D19 strain C (Fig. 4A) (30). Additional phylogenetic trees were constructed to encompass the two proposed recombination sites within this

FIG. 3. Bootscan analysis of HAdV-D penton base genes. Comparison of HAdV-D22 (A) with completely sequenced HAdV types. Similar comparisons of the same region were performed with HAdV-D9 (B), HAdV-D49 (C), HAdV-A12 (D), HAdV-B3 (E), HAdV-C2 (F), or HAdV-F41 (G) as the reference type. HAdV-D53 was left out of the analysis due to the 100% identity of its penton base to that of HAdV-D37. The axes of all panels are as labeled in panel A. GenBank accession numbers are as follows: HAdV-A12, NC\_001460; HAdV-B3, DQ086466; HAdV-C2, AC\_000007; and HAdV-F41, DQ315364.



gene. Phylogenetic analysis of the nucleotide 400 to 600 region revealed a close identity among HAdV-D22, HAdV-D37, and HAdV-D53 (Fig. 4B). Analysis of the nucleotide 750 to 1350 region of the penton base gene revealed a close identity of HAdV-D22 and HAdV-D19 (Fig. 4C).

In summary, comparison of the HAdV-D22 genome to other sequenced HAdV-D genomes identified substantial sequence divergence in the penton, hexon, E3, and fiber open reading frames. Differences between HAdV-D genomes in these areas have previously been described by us in reports on HAdV-D19 strain C and HAdV-D37 genomes (29, 30). However, the sequence conservation among HAdV-D22, HAdV-D19, and HAdV-D37 in the penton base gene was unexpected and suggests recombination among these viruses. Bootscan analysis identified recombination events at two regions within the penton base gene, encompassing nucleotides 400 to 600 and 750 to 1350. Further analysis of other HAdV-D penton genes suggests that these areas represent hot spots for recombination.

The propensity for recombination among different HAdV-Ds in the penton base gene can be understood in the context of penton base function. The adenovirus penton base acts as the ligand for a secondary attachment event that is critical to host cell internalization (38) and thus is critical to infection. The penton base protein contains two hypervariable loops located on the surface of the viral capsid (16, 44). One loop contains the host cell integrin binding Arg-Gly-Asp (RGD) motif mediating attachment to host cell integrins (8, 38). The RGD motif is conserved in almost every HAdV penton base protein, with the exception of HAdV-F40 and HAdV-F41, which do not use host cell integrins for internalization (1, 11). The second exterior loop, known as the variable loop (HVL1), has no known function. Both of these regions are highly variable among different HAdV types. We identified recombination around both of these loops for multiple viruses within HAdV-D (Fig. 3A).

In recent epidemiological studies, patients were identified with coincident clinical infections with two or more adenovirus types (17, 36). Simultaneous infection by more than a single HAdV type is possible because of conserved host receptor affinity, common tissue tropisms, and an absence of immunity across serotypes. Our data, along with evidence from previous studies (22, 23, 37, 43), identified recombination events among viruses with similar tissue tropisms, providing evidence that the restriction of tissue tropism might determine in part the observed recombination within adenoviruses of the same species. Recently, a novel HAdV, HAdV-D53, was isolated from an outbreak of keratoconjunctivitis. Subsequent analysis revealed recombination among HAdV-D22, HAdV-D37, HAdV-D8, and a previously unknown adenoviral sequence, suggesting the potential for the emergence of new pathogens, with important ramifications for human disease.

Previous work provides evidence for recombination in HAdVs (10, 12, 22, 25, 26, 37, 39, 41–43), but the mechanisms of recombination have yet to be identified. Recombination may result from selective pressure from the host immune system relative to surface capsid proteins, a nucleotide motif that directs cellular recombination machinery to the local sites on the viral DNA, or a combination of both. Two eukaryotic recombinases, RAD51 and Dmc1, both homologues of the bacterial recombinase RecA, act in host cell DNA recombina-

tion (5, 33). RAD51 mediates recombination during mitosis, while Dmc1 acts during meiosis. RAD51 is of potential interest because it colocalizes with promyelocytic leukemia nuclear bodies in the nucleus (4). Adenoviral proteins E1A and E4 Orf3 have been shown to interact with promyelocytic leukemia nuclear bodies, which play an important role in adenoviral replication (7, 15). An interaction between RAD51 and adenoviral DNA has not been studied. A proposed motif consisting of CCNCCNTNNCCNC was recently identified as being associated with loci of recombination and genome instability in humans (28). Although not present in the viruses we studied, sequencing of other viruses within HAdV-D may yet reveal a consensus site for recombination. The elucidation of recombination mechanisms for HAdVs should allow a better understanding of adenoviral evolution.

In conclusion, our analysis of the penton base gene of HAdV-D identified a potential paradigm for adenovirus recombination and the emergence of pathogenic strains. An in depth understanding of adenovirus recombination and evolution is critical to ensure the safety of adenoviral gene therapy.

**Nucleotide sequence accession number.** The HAdV-D22 genome sequence obtained in this study has been deposited in GenBank under accession number FJ404771.

#### REFERENCES

- Albinsson, B., and A. H. Kidd. 1999. Adenovirus type 41 lacks an RGD alpha(v)-integrin binding motif on the penton base and undergoes delayed uptake in A549 cells. *Virus Res.* **64**:125–136.
- Aoki, K., H. Ishiko, T. Konno, Y. Shimada, A. Hayashi, H. Kaneko, T. Ohguchi, Y. Tagawa, S. Ohno, and S. Yamazaki. 2008. Epidemic keratoconjunctivitis due to the novel hexon-chimeric-intermediate 22/37/H8 human adenovirus. *J. Clin. Microbiol.* **46**:3259–3269.
- Bell, S. D., Jr., T. R. Rota, and D. E. McComb. 1960. Adenoviruses isolated from Saudi Arabia. III. Six new serotypes. *Am. J. Trop. Med. Hyg.* **9**:523–526.
- Bischof, O., S. H. Kim, J. Irving, S. Beresten, N. A. Ellis, and J. Campisi. 2001. Regulation and localization of the Bloom syndrome protein in response to DNA damage. *J. Cell Biol.* **153**:367–380.
- Bishop, D. K., D. Park, L. Xu, and N. Kleckner. 1992. DMCI: a meiosis-specific yeast homolog of *E. coli* recA required for recombination, synaptonemal complex formation, and cell cycle progression. *Cell* **69**:439–456.
- Brudno, M., C. B. Do, G. M. Cooper, M. F. Kim, E. Davydov, E. D. Green, A. Sidow, and S. Batzoglou. 2003. LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. *Genome Res.* **13**:721–731.
- Carvalho, T., J. S. Seeler, K. Ohman, P. Jordan, U. Pettersson, G. Akusjarvi, M. Carmo-Fonseca, and A. Dejean. 1995. Targeting of adenovirus E1A and E4-ORF3 proteins to nuclear matrix-associated PML bodies. *J. Cell Biol.* **131**:45–56.
- Cheresh, D. A., and R. C. Spiro. 1987. Biosynthetic and functional properties of an Arg-Gly-Asp-directed receptor involved in human melanoma cell attachment to vitronectin, fibrinogen, and von Willebrand factor. *J. Biol. Chem.* **262**:17703–17711.
- Chu, W., and D. Pavan-Langston. 1979. Ocular surface manifestations of the major viruses. *Int. Ophthalmol. Clin.* **19**:135–167.
- Crawford-Mikszta, L. K., and D. P. Schnurr. 1996. Adenovirus serotype evolution is driven by illegitimate recombination in the hypervariable regions of the hexon protein. *Virology* **224**:357–367.
- Davison, A. J., E. A. Telford, M. S. Watson, K. McBride, and V. Mautner. 1993. The DNA sequence of adenovirus type 40. *J. Mol. Biol.* **234**:1308–1316.
- de Jong, P. J., G. Valderrama, I. Spigland, and M. S. Horwitz. 1983. Adenovirus isolates from urine of patients with acquired immunodeficiency syndrome. *Lancet* **i**:1293–1296.
- Dhurandhar, N. V., P. Kulkarni, S. M. Ajinkya, and A. Sherikar. 1992. Effect of adenovirus infection on adiposity in chicken. *Vet. Microbiol.* **31**:101–107.
- Dingle, J. H., and A. D. Langmuir. 1968. Epidemiology of acute, respiratory disease in military recruits. *Am. Rev. Respir. Dis.* **97**(Suppl):1–65.
- Doucas, V., A. M. Ishov, A. Romo, H. Juguilon, M. D. Weitzman, R. M. Evans, and G. G. Maul. 1996. Adenovirus replication is coupled with the dynamic properties of the PML nuclear structure. *Genes Dev.* **10**:196–207.
- Fuschiotti, P., G. Schoehn, P. Fender, C. M. Fabry, E. A. Hewat, J. Chroboczek, R. W. Ruigrok, and J. F. Conway. 2006. Structure of the dodecahedral penton particle from human adenovirus type 3. *J. Mol. Biol.* **356**:510–520.
- Gray, G. C., T. McCarthy, M. G. Lebeck, D. P. Schnurr, K. L. Russell,

- A. E. Kajon, M. L. Landry, D. S. Leland, G. A. Storch, C. C. Ginocchio, C. C. Robinson, G. J. Demmler, M. A. Saubolle, S. C. Kehl, R. Selvarangan, M. B. Miller, J. D. Chappell, D. M. Zerr, D. L. Kiska, D. C. Halstead, A. W. Capuano, S. F. Setterquist, M. L. Chorzay, J. D. Dawson, and D. D. Erdman. 2007. Genotype prevalence and risk factors for severe clinical adenovirus infection, United States 2004–2006. *Clin. Infect. Dis.* **45**:1120–1131.
18. Hilleman, M. R., and J. H. Werner. 1954. Recovery of new agent from patients with acute respiratory illness. *Proc. Soc. Exp. Biol. Med.* **85**:183–188.
19. Jager, L., and A. Ehrhardt. 2007. Emerging adenoviral vectors for stable correction of genetic disorders. *Curr. Gene Ther.* **7**:272–283.
20. Kojaoghlani, T., P. Flomenberg, and M. S. Horwitz. 2003. The impact of adenovirus infection on the immunocompromised host. *Rev. Med. Virol.* **13**:155–171.
21. Lole, K. S., R. C. Bollinger, R. S. Paranjape, D. Gadkari, S. S. Kulkarni, N. G. Novak, R. Ingersoll, H. W. Sheppard, and S. C. Ray. 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J. Virol.* **73**:152–160.
22. Lukashev, A. N., O. E. Ivanova, T. P. Eremeeva, and R. D. Iggo. 2008. Evidence of frequent recombination among human adenoviruses. *J. Gen. Virol.* **89**:380–388.
23. Madisch, I., S. Hofmayer, C. Moritz, A. Grintzalis, J. Hainmueller, P. Pring-Akerblom, and A. Heim. 2007. Phylogenetic analysis and structural predictions of human adenovirus penton proteins as a basis for tissue-specific adenovirus vector design. *J. Virol.* **81**:8270–8281.
24. Martin, D. P., C. Williamson, and D. Posada. 2005. RDP2: recombination detection and analysis from sequence alignments. *Bioinformatics* **21**:260–262.
25. Mautner, V., and M. E. Boursnell. 1983. Recombination in adenovirus: DNA sequence analysis of crossover sites in intertypic recombinants. *Virology* **131**:1–10.
26. Mautner, V., and N. Mackay. 1984. Recombination in adenovirus: analysis of crossover sites in intertypic overlap recombinants. *Virology* **139**:43–52.
27. Morozova, O., and M. A. Marra. 2008. Applications of next-generation sequencing technologies in functional genomics. *Genomics* **92**:255–264.
28. Myers, S., C. Freeman, A. Auton, P. Donnelly, and G. McVean. 2008. A common sequence motif associated with recombination hot spots and genome instability in humans. *Nat. Genet.* **40**:1124–1129.
29. Robinson, C. M., F. Shariati, A. F. Gillaspay, D. W. Dyer, and J. Chodosh. 2008. Genomic and bioinformatics analysis of human adenovirus type 37: new insights into corneal tropism. *BMC Genomics* **9**:213.
30. Robinson, C. M., F. Shariati, J. Zaitshik, A. F. Gillaspay, D. W. Dyer, and J. Chodosh. 2009. Human adenovirus type 19: genomic and bioinformatics analysis of a keratoconjunctivitis isolate. *Virus Res.* **139**:122–126.
31. Rowe, W. P., R. J. Huebner, L. K. Gilmore, R. H. Parrott, and T. G. Ward. 1953. Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. *Proc. Soc. Exp. Biol. Med.* **84**:570–573.
32. Sever, J. L., and R. G. Traub. 1962. Conjunctivitis with follicles associated with adenovirus type 22. *N. Engl. J. Med.* **266**:1375–1376.
33. Shinohara, A., H. Ogawa, and T. Ogawa. 1992. Rad51 protein involved in repair and recombination in *S. cerevisiae* is a RecA-like protein. *Cell* **69**:457–470.
34. Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**:1596–1599.
35. Thirion, C., H. Lochmuller, Z. Ruzsics, M. Boelhaue, C. Konig, C. Thedieck, S. Kutik, C. Geiger, S. Kochanek, C. Volpers, and H. G. Burgert. 2006. Adenovirus vectors based on human adenovirus type 19a have high potential for human muscle-directed gene therapy. *Hum. Gene Ther.* **17**:193–205.
36. Vora, G. J., B. Lin, K. Gratwick, C. Meador, C. Hansen, C. Tibbetts, D. A. Stenger, M. Irvine, D. Seto, A. Purkayastha, N. E. Freed, M. G. Gibson, K. Russell, and D. Metzgar. 2006. Co-infections of adenovirus species in previously vaccinated patients. *Emerg. Infect. Dis.* **12**:921–930.
37. Walsh, M. P., A. Chintakuntlawar, C. M. Robinson, I. Madisch, B. Harrach, N. R. Hudson, D. Schnurr, A. Heim, J. Chodosh, D. Seto, and M. S. Jones. 2009. Evidence of molecular evolution driven by recombination events influencing tropism in a novel human adenovirus that causes epidemic keratoconjunctivitis. *PLoS ONE* **4**:e5635.
38. Wickham, T. J., P. Mathias, D. A. Chereshe, and G. R. Nemerow. 1993. Integrins alpha v beta 3 and alpha v beta 5 promote adenovirus internalization but not virus attachment. *Cell* **73**:309–319.
39. Williams, J., T. Grodzicker, P. Sharp, and J. Sambrook. 1975. Adenovirus recombination: physical mapping of crossover events. *Cell* **4**:113–119.
40. Wood, D. J. 1988. Adenovirus gastroenteritis. *Br. Med. J.* **296**:229–230.
41. Young, C. S., G. Cachianes, P. Munz, and S. Silverstein. 1984. Replication and recombination in adenovirus-infected cells are temporally and functionally related. *J. Virol.* **51**:571–577.
42. Young, C. S., and S. J. Silverstein. 1980. The kinetics of adenovirus recombination in homotypic and heterotypic genetic crosses. *Virology* **101**:503–515.
43. Zhu, Z., Y. Zhang, S. Xu, P. Yu, X. Tian, L. Wang, Z. Liu, L. Tang, N. Mao, Y. Ji, C. Li, Z. Yang, S. Wang, J. Wang, D. Li, and W. Xu. 2009. Outbreak of acute respiratory disease in China caused by B2 species of adenovirus type 11. *J. Clin. Microbiol.* **47**:697–703.
44. Zubieta, C., G. Schoehn, J. Chroboczek, and S. Cusack. 2005. The structure of the human adenovirus 2 penton. *Mol. Cell* **17**:121–135.