Three novel, multiple recombinant types of species of human mastadenovirus D (HAdV-D 73, 74 & 75) isolated from diarrhoeal faeces of immunocompromised patients

Elias Hage,¹ ² Akshay Dhingra,¹ ² Uwe G. Liebert,³ Sandra Bergs,³ Tina Ganzenmueller¹ ² and Albert Heim¹ ²,*

Abstract

Species D is the largest of the seven species of human mastadenoviruses (HAdV), but few of its multiple types are associated with severe disease, e.g. epidemic keratoconjunctivitis. Many other types are hardly ever associated with significant diseases in immunocompetent patients, but have been isolated from the diarrhoeal faeces of terminal AIDS patients suggesting their role as opportunistic pathogens. Three novel HAdV-D strains were isolated from the faeces of three immunocompromised adult patients (clinical diagnoses: lymphoma, myelodysplastic syndrome and AIDS CDC3B, respectively). These strains were not typeable by imputed serology of the hexon and fibre gene and therefore complete genomic sequences were generated by next-generation sequencing (NGS). All three strains were multiple recombinants and fulfilled the criteria for designation as types 73, 74 and 75 with the penton/hexon/fibre genotype codes P67H45F27, P70H74F51 and P75H26F29, respectively. A novel genomic backbone and also a novel hexon neutralization epitope sequence were discovered in type 74, and a novel penton sequence in type 75. At the complete genome level, types 73, 74 and 75 were closely related neither to each other nor to type 70, which was previously isolated in the same region. However, these four HAdV-D types were closely related to each other in single genes and gene regions, e.g. penton, E1 and E4 due to recombination events in their phylogeny. In conclusion, regional co-circulation of opportunistic HAdV-D types facilitated co- and super-infections, which are essential for homologous recombination, and thus resulted in the evolution of novel genotypes by lateral gene transfer.

INTRODUCTION

Human adenoviruses (HAdVs) are associated with a wide spectrum of diseases depending on the organotropism of the HAdV type and the age and immunocompetence of the host. More than 70 different HAdV types have been described and are grouped into seven species (HAdV-A to HAdV-G). Respiratory tract infections are caused by various types of species HAdV-B and HAdV-C, gastroenteritis most frequently by types 40 and 41 of species HAdV-F and epidemic keratoconjunctivitis (EKC) almost exclusively by a few types of species HAdV-D [1–3].

HAdV can also cause life-threatening disseminated disease in allogeneic haematopoietic stem cell transplantation (HSCT) recipients and other severely immunocompromised patients. Types 1, 2 and 5 of species HAdV-C clearly predominate in disseminated disease, but this is also caused by type 31 of species HAdV-A in paediatric HSCT and several types (e.g. 11, 35) of species HAdV-B in adult SCT patients [4–6]. Only occasionally have types of species HAdV-D been involved in the aetiology of disseminated disease in adult SCT patients and other severely immunocompromised patients [6]. However, many HAdV-D types have been isolated from the stool of immunocompromised patients suffering from diarrhoea, especially from terminally ill AIDS patients [4], suggesting their role as opportunistic pathogens. Recently, the prototype of HAdV-D70 was isolated from the diarrhoeal faeces of an adult HSCT recipient [7]. By contrast, a small group of HAdV-D types (8, 37, 53, 54, 64 – previously 19a) are associated with the severe eye infection, epidemic keratoconjunctivitis [4].

Both recombination between species D types and selection of novel neutralization epitopes (‘immune escape’) may contribute to the huge diversity of HAdV-D types [8–10]. Classically, HAdV isolates were typed by neutralization testing (main neutralization epitope ε on the hexon) [11]. More recently, nucleic acid sequencing of the ε determinant on the hexon gene was proposed as an alternative procedure.
RESULTS

Type 73 (P67H45F27)

Type 73 was isolated from diarrhoeal faeces of a lymphoma patient treated with chemotherapy. Imputed serology yielded contradictory results for the hexon and fibre, which were unequivocally related to types 45 and type 27, respectively.

By using NGS and de novo assembly, a single contig representing the genomic sequence of HAdV-D73 was constructed (average coverage 27 968-fold). The whole genome of HAdV-D73 was 35 190 bp in length with a GC content of 57.05 %. The genome was predicted to encode for 40 open reading frames (ORFs) with an organization similar to other members of species HAdV-D. The complete genomic sequence of HAdV-D73 was deposited in GenBank as KY618676.

The complete genome sequence was found to be most closely related to types 44 and 48 (96.3 % identity, Fig. 1a). Bootscan analysis indicated a multiple recombinant genome for the hexon and fibre, which were unequivocally related to types 45 and type 27, respectively (Fig. 2) with the major capsid proteins related to different other HAdV-D types, penton base to type 67, hexon (including the major neutralization determinant $\epsilon$) to type 45 and fibre to type 27 (Fig. 1b–d), with the fibre sequence significantly evolved since recombination with type 27. The more conserved parts of the genome were closely related with a multitude of HAdV species D types but could not be clearly assigned to a certain type (Fig. 2). The loop 1 sequence of the major neutralization epitope $\epsilon$ on the hexon was almost identical to type 45 (sequence divergence on the nucleotide level 0.36 %, on the protein level 0.0 %) but loop 2 was significantly divergent from type 45 (on nucleotide level 5.7 %, on protein level 7.52 %) and more closely related to types 26 and 75 (see below).

Type 74 (P70H74F51)

Type 74 was isolated from diarrhoeal faeces of a patient treated with allogeneic HSCT for a myelodysplastic syndrome. Its hexon sequence could not be associated with any previously described HAdV serotype by imputed serology.

A single contig representing the genome of HAdV-D74 was constructed using de novo assembly (average coverage 23 896-fold). The whole genome of HAdV-D74 was 35 155 bp in length with a GC content of 56.98 %. The genome was predicted to encode for 40 open reading frames (ORFs) with an organization similar to other members of species HAdV-D. The complete genomic sequence of HAdV-D74 was deposited in GenBank as KY618677.

Type 74 had a novel neutralization epitope sequence (Fig. 1c, hexon loop 1: identity to the most closely related type 13 on the nucleotide level only 71.2 %, on the protein level 68.2 %; hexon loop 2: 85.5 % identity on the nucleotide level, 91.2 % on the protein level). These deduced amino acid sequence identities surpassed the set identity thresholds (95.8 % for loop 1 and 98.8 % for loop 2) of imputed serology for identity with a previously identified serotype. Moreover, the complete genomic sequence of type 74 had only 94.4 % identity to the most closely related type 38 (Fig. 1a). Bootscan analysis indicated multiple recombination events in the phylogeny of type 74 (Fig. 3) with only the penton base sequence closely related to type 70 (Fig. 1b), whereas the fibre originated from type 51 but had evolved significantly since recombination (Fig. 1d). Moreover, recombination events were detected in the phylogeny of the early gene regions: for example, the E1 region was closely related to types 15 and 75 (see below) (Fig. 4a, b). By contrast the E3 region was found to have a novel sequence not closely related to any previous HAdV-D type sequence (Fig. 4c), whereas E4 was closely related to types 71 and 58 (Fig. 4d).

Type 75 (P75H26F29)

Type 75 was isolated from the faeces of an AIDS patient (CDC IIIB). Imputed serology yielded contradictory results for the hexon and fibre, which were unequivocally but distantly related to types 26 and 29, respectively.

A single contig representing the genome of HAdV-D75 was constructed using de novo assembly (average coverage 27 301-fold). The whole genome of HAdV-D75 was 35 104 bp in length with a GC content of 56.97 %. The genome was predicted to encode for 40 open reading frames (ORFs) with an organization similar to other members of species HAdV-D. The complete genomic sequence of HAdV-D75 was deposited in GenBank as KY618678.

The complete genome sequence was most closely related to type 51 (96.1 % identity, Fig. 1a). The loop 1 sequence of the major neutralization epitope $\epsilon$ on the hexon was not closely related to any previous (sero-)type (sequence identity 98.8 % on the nucleotide level, 90.3 % on the amino acid sequence level compared to type 26). However, loop 2 was closely related to type 26 with only 1 % sequence divergence.
Fig. 1. Phylogenetic analysis of the HAdV-D73, -D74 and -D75 nucleic acid sequences with all other prototype sequences of the species HAdV-D. (a) Complete genomic sequences; (b) penton base gene sequences; (c) hexon gene sequences; (d) fibre gene sequences. Sequences of species HAdV-B and HAdV-C are depicted as compressed clusters. Neighbour-joining trees were generated based on the Kimura 2-parameter model with MEGA 7 software. Bootstrap values below 80% are not robust and are therefore not depicted.
on the amino acid level (6.42% nucleic acid sequence divergence). Bootscan analysis did not confirm an intra-hexon recombination, probably because loop 2 is too short to yield a significant result. Other recombination events were supported by bootscanning (Fig. 5): the sequence of the E3 region was closely related to type 51 (Fig. 4c), whereas the fibre gene was related to type 29 and several other types (58, 63 and 70) clustering with the type 29 sequence (Fig. 1d). The penton base gene sequence of type 75 was found to be novel and not closely related to any previous types of species HAdV-D (Fig. 1b).

**DISCUSSION**

Three novel human mastadenovirus types were isolated from the diarrhoeal faeces of immunocompromised patients and designated as types 73 (P67H45F27), 74 (P70H74F51) and 75 (P75H26F29). Patients were from the same region and treated in the same hospital where type 70 was previously isolated [7]. However, there is no epidemiological evidence for an infection chain directly linking the immunocompromised patients from which types 70, 73, 74 and 75 were isolated. Type 70 was isolated in August 2013, but types 73–75 in January and February 2015. Moreover, their whole-genomic sequences did not cluster with each other. In contrast to the whole-genome level, these viruses were found to be closely related in specific open reading frames. For instance, type 75 was found to be related both to types 73 (in the E4 region) and 74 (almost identical E1A and E1B sequence). Moreover, type 74 was also related to the previously described type 70 [7] in the penton base gene and fibre gene. These results clearly show that types 70, 73, 74 and 75 have common ancestors at least for single ORFs, and this finding can only be explained by recombination events in their phylogeny.

Homologous recombination requires co-infection of host cells with two or more HAdV-D types. This is facilitated by persistent HAdV-D coinfections or a combination of a persistent infection with a HAdV-D type and an acute super-infection with another HAdV-D type, because coincidental acute double infection with two HAdV-D types seems to be rather improbable due to the short period of HAdV DNA replication in acute infections. Persistent infections have been reported in adult immunocompromised (e.g. AIDS) patients with long-term faecal HAdV-D shedding [15, 16], but so far not in immunocompetent individuals. However, transmission and evolution of HAdV-D types has to be considered in a broader context of community-acquired, asymptomatic (or at least subclinical) infections of non-

**Fig. 2.** Similarity plot (a) and bootscan analysis (b) of the HAdV-D73 whole-genomic sequence with other HAdV-D types. The bootscan plot demonstrates the phylogenetic relationship of HAdV-D73 to selected HAdV-D types calculated as bootstrap values (% permuted trees) of the sliding 1000 bp window, and the similarity plot indicates the genetic distances of the HAdV-D73 genome to selected HAdV-D type sequences in the sliding 1000 bp window. Approximate positions of the penton base, hexon, fibre genes and the E3 and E4 gene region are depicted. Prototype sequences selected for bootscanning appear in colour, with all other HAdV-D sequences depicted in grey (a). A window size of 1000 bp and a step size of 100 bp were set for bootscanning analysis using SimPlot software.
immunocompromised adult hosts, as an infection chain linking the patients harbouring types 70, 73, 74 and 75 could be excluded. These asymptomatic HAdV-D infections in immunocompetent individuals remain undiagnosed, but probably also result in HAdV-D persistence. However, little is known about these issues in the otherwise healthy adult population. In contrast, a recent study has described persistence of HAdV-DNA in lymphoid cells of the lamina propria of healthy children and active virus replication in epithelial cells of the gastrointestinal tract [17]. Species HAdV-C and HAdV-A predominated [17], similar to the HAdV species distribution found in acute respiratory and intestinal HAdV infections of children [18], and consequently also in severe infections of immunocompromised children [5]. Recombinant genomes were also frequently observed in clinical isolates of species HAdV-C [19], probably due to co- and super-infections with different types of species HAdV-C during HAdV-C persistence. A similar mechanism of HAdV-D DNA persistence in lymphoid cells of the lamina propria and virus replication in the epithelial cells of the gut of immunocompetent adults can be speculated, but data supporting this hypothesis are not currently available. For comparison, many of the HAdV-D types were originally isolated from adult AIDS and other adult immunocompromised patients [20–23], and many of these new types originated from multiple recombination events between HAdV-D types similar to types 73–75 [8, 9]. However, these previous studies did not show a phylogenetic relationship between HAdV-D isolates from the same region, whereas the present study shows phylogenetic relationships between novel HAdV-D types isolated in the same geographical region. This highlights a regional co-circulation of multiple HAdV-D types and recombination between these viruses driving the evolution of novel types by lateral gene transfer.

Although classical serological cross-neutralization experiments are not required for the description of new HAdV genotypes, hexon sequences may be used to predict neutralization properties. For example, the loop 1 sequence of the main neutralization epitope of type 73 was almost identical to type 45 and thus cross-neutralization is highly probable. Surprisingly, the loop 2 sequence of type 73 clustered differently with the sequences of type 26 and 75 (suggesting potential cross-neutralization) but an intra epitope recombinant site between loops 1 and 2 was not confirmed by bootscanning. In contrast type 74 has a novel hexon sequence not closely related to any previously described serotype and thus it seems probable that type 74 should also fulfil the criteria for a new serotype. The hexon neutralization epitope of type 75 was found to be phylogenetically

Fig. 3. Similarity plot (a) and bootscan analysis (b) of the HAdV-D74 whole-genomic sequence with other HAdV-D types. The bootscan plot demonstrates the phylogenetic relationship of HAdV-D74 to selected HAdV-D types calculated as bootstrap values (% permuted trees) of the sliding 1000 bp window, and the similarity plot indicates the genetic distances of the HAdV-D74 genome to selected HAdV-D type sequences in the sliding 1000 bp window. Approximate positions of the penton base, hexon, fibre genes and the E3 and E4 gene region are depicted. Prototype sequences selected for bootscanning appear in colour, with all other HAdV-D sequences depicted in grey (a). A window size of 1000 bp and a step size of 100 bp were set for bootscanning analysis using SimPlot software.
Fig. 4. Phylogenetic analysis of the HAdV-D73, -D74 and -D75 early gene region nucleic acid sequences with all other prototype sequences of the species HAdV-D. (a) E1A region; (b) E1B region; (c) E3 region; (d) E4 region. Neighbour-joining trees were generated based on the Kimura 2-parameter model with MEGA 7 software. Bootstrap values below 80% are not robust and are therefore not depicted.
related to type 26 with a loop 1 sequence that has significantly evolved (9.7% divergence on protein level), but a more conserved loop 2 sequence was found (only 1% sequence divergence to type 26 on protein level). The latter indicated a potential cross-neutralization with type 26 and 73 (see above).

Novel HAdV types with E1 and E4 ORFs originating from different types can result in high virulence, such as in the case of the novel subtype HAdV-B21a [24], because the E3 ligase complexes formed by E4orf6 and E1B55k of different HAdV types can have an altered substrate specificity [25]. For example, the novel type 74 had its early gene regions 1 and 4 derived from different, previously described HAdV-D types (types 15 and 71, respectively). Moreover, type 74 has a novel E3 gene region sequence. The E3 region codes for multiple gene products that interfere with the immune response and thus may enhance virulence. However, there is no clinical or epidemiological evidence suggesting a high virulence of type 74. This virus was recovered from an immunosuppressed patient with diarrhoea (as were types 70, 73 and 75), but reports on type 74-associated diseases in the general population are lacking. Therefore, the high virulence of type 74 is merely hypothetical. On the other hand, adenovirus-associated diseases are not notifiable in most countries and thus an association of type 74 with clinically significant disease may have been missed. Nevertheless, lateral gene transfer of the E1, E3 and E4 genome regions by homologous recombination can be regarded as a mechanism that might enhance the virulence of novel HAdV genotypes and could generate a highly virulent, emerging HAdV type in the future.

In conclusion, all three novel HAdV-D types should be regarded as opportunistic pathogens that evolved by multiple recombination events, probably during persistent infections.

**METHODS**

**Isolation of the novel HAdV types**

Adenovirus isolation from stool samples was performed on human lung carcinoma cells A549 (ATCC, CCL-185) until 90% CPE formation. Briefly, about 0.1 g of stool was suspended in 1 ml 0.9% NaCl and sterile filtered (0.2 µm filter). Ninety per cent confluent cell cultures were washed with PBS, and subsequently stool suspension and fresh culture medium were added to the cells. Cells were checked every 3 to 4 days for CPE formation. Viral DNA was extracted from cell culture supernatant at 90% CPE using the Qiagen blood kit (Qiagen, Hilden, Germany).
**Imputed serology by partial sequencing**

Initial typing was performed by Sanger sequencing of parts of the neutralization epitope (loop 1 and loop 2) of the hexon gene (imputed serology) and the fibre knob gene as previously described [26].

**Next-generation sequencing and bioinformatics**

Library preparation was performed using the Nextera XT DNA sample preparation kit (Illumina, San Diego, CA) with 1 ng of DNA as input following manufacturers’ guidelines. After purification and quality control, libraries were sequenced on an Illumina MiSeq platform (2×300 bp paired-end run). The obtained sequence data was quality controlled and de novo assembled using CLC Genomics Workbench (version 9, Aarhus, Denmark). Genome annotations were transferred from the most closely related type and cross-checked by using an ORF finder followed by BLAST analysis using Geneious (version 10).

**Phylogenetic and recombination analysis**

Multiple sequencing alignments of the nucleotide and predicted amino acid sequences were constructed using the MAFFT online server (http://mafft.cbrc.jp/alignment/server/) implementing an iterative refinement method (FFT-NS-i). The alignment was visualized and checked for errors using AliView (version 1.18). Phylogenetic analysis was performed using MEGA (version 7) by implementing the neighbour-joining approach (Kimura 2-parameter model) with 1000 bootstrap replicates.

Similarity plots and recombination detection (bootscan approach) was performed using the SimPlot software version 3.5.1 [27] with a window size of 1000 bp and a step size of 200 bp.

### Funding information

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) Collaborative Research Centre (SFB) 900, core project Z1; and the Robert-Koch-Institut, FKZ 1369-380. E. H. and A. D. were supported by the Hannover Biomedical Research School (HBRS) and the Centre for Infection Biology (ZIB).

### Acknowledgements

We thank Gabriele Harste, Jenny Witthuhn, and Sandra Flucht (Institut für Virologie, Medizinische Hochschule Hannover) for excellent technical assistance.

### Conflicts of interest

The authors declare that there are no conflicts of interest.

### References


Five reasons to publish your next article with a Microbiology Society journal

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as ‘excellent’ or ‘very good’.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at microbiologyresearch.org.