272
Current Topics in Microbiology and Immunology

Springer-Verlag Berlin Heidelberg GmbH

W. Doerfler P. Böhm (Eds.)

Adenoviruses: Model and Vectors in Virus-Host Interactions

Virion-Structure, Viral Replication and Host-Cell Interactions

With 60 Figures and 8 Tables

Springer

Prof. Dr. WALTER DOERFLER
Institut für Klinische und Molekulare Virologie
Universität Erlangen-Nürnberg
91054 Erlangen, Germany
e-mail: doerfler@viro.med.uni-erlangen.de

PETRA BÖHM
Institut für Genetik
Universität zu Köln
Weyertal 121
50931 Köln, Germany
e-mail: pboehm@scan.genetik.uni-koeln.de

Cover Illustration by M. Hösel (this volume):

Overexpression of the E4-ORF3 gene of Ad12 (C) or Ad2 (D) leads to the reorganization of the PML protein domains in BHK21 hamster cells. In untransfected cells (A) or in cells transfected with the pIRES2-EGFP vector (B), PML protein domains retain their spherical structures. All sucessfully transfected cells are green due to the expression of EGFP (enhanced expression of the green fluorescent protein). The PML protein is detectable by staining with primary anti-PML protein antibodies and secondary Cy3-anti mouse IgG. Bars designate magnifications. This figure was reproduced from Hösel et al. (2001a) see Chapter 14.

ISSN 0070-217X
ISBN 978-3-642-05517-1 ISBN 978-3-662-05597-7 (eBook)
DOI 10.1007/978-3-662-05597-7

Library of Congress Catalog Card Number 72-152360

Originally published by Springer-Verlag Berlin Heidelberg New York in 2003
Softcover reprint of the hardcover 1st edition 2003

Library of Congress Catalog Card Number 15-12910

Cover Design: Design & Production GmbH, Heidelberg
Typesetting: Fotosatz-Service Köhler GmbH, Würzburg
Production Editor: Christiane Messerschmidt, Rheinau
Printed on acid-free paper 27/3020 5 4 3 2 1 0

Preface

After volumes 109, 110 ("The Molecular Biology of Adenoviruses"), and 199 I–III ("The Molecular Repertoire of Adenoviruses"), volumes 272 and 273 of the *Current Topics in Microbiology and Immunology* in 2003 are the third series devoted to the analyses of molecular mechanisms elicited by adenovirus infections. Adenovirus research continues at a rigorous pace.

It is true, for many researchers in the biomedical sciences, adenoviruses are nowadays best known as possible vectors for human somatic gene therapy. However, our goal in editing the current volumes has been to present an overview on basic research dealing with the molecular biology of this group of viruses. Of course, much of the essential information on the adenoviral genomes, their replication and gene expression as well as information on other biologically relevant problems, such as adenoviral oncogenesis, virus–host interactions, and the immune response against adenoviruses, had already reached a highly sophisticated level when we edited the previous adenovirus volume in 1995. Nevertheless, many of the most exciting problems in adenovirology and on the host responses to the infection with these viruses both at the cellular and organismic levels are still unresolved but presently actively worked on. Possible applications of adenoviruses such as gene transfer vectors in human somatic gene therapy will vitally depend on advances in basic virology. Perhaps we can expect progress in foreign DNA transfer into mammalian cells and organisms only when we have derived all the necessary information from molecular adenovirus research. In this way, we might learn how the virus transports its DNA into an environment for which the virus has become specialized in the course of perhaps millions of years. Adenoviruses are experts when it comes to invading the human organism. This fact should also caution us when developing these viruses into vectors for human gene therapy.

Although predictions on future trends and realities in biomedical research can be risky at best, we venture to speculate that it will most likely not be the infectious virus itself that will eventually be directly applied in

human gene therapy. The virus as a pathogen with its high degree of expertise in maneuvering inside the human organism cannot be tamed sufficiently, even by the most ambitious gene technologists, to become an innocuous transfer vector.

Individual contributions to the two current volumes on adenoviruses encompass the entire gamut of molecular virology and of problems related to virus–host interactions, viral oncology and host defenses against viral infections. The chapters in this volume have been assigned to sections on virion structure, viral replication, and host–cell interactions; problems discussed in the next volume will relate to oncogenesis, immune responses, and gene therapy.

We hope that these books will again serve as a resource of information on current adenovirus research and help motivate novel and productive research on adenoviruses as efficient tools in mammalian molecular biology.

We wish to express our gratitude to our colleagues, the authors of these chapters, for taking the time and effort to share the results of their exciting research and write these chapters. We are also indebted to Ms. Clauss at Springer Verlag in Heidelberg for her help in editing the manuscripts.

Köln, November 2002

Walter Doerfler
Petra Böhm

List of Contents

List of Contributors

(Their addresses can be found at the beginning of their respective chapters.)

1
Virion and Structure

CHAPTER 1

Molecular Evolution of Adenoviruses

M. BENKÖ, B. HARRACH

Abstract. New advances in the field of genetic characterization of adenoviruses originating from different animal species are summarized. Variations seen in the host range and specificity, pathogenicity, genomic arrangement or gene complement are much wider than expected based on previous studies of human adenoviruses. Several exceptional adenoviruses from the two traditional conventional genera are now removed, and proposed to form at least two new genera. The eventual host origin of the new genera, however, is not clarified. Novel results from the genomic and phylogenetic analyses of adenoviruses originating from lower vertebrate species (including reptiles, amphibians and fish) seem to imply that

M. BENKÖ, B. HARRACH
Veterinary Medical Research Institute of the Hungarian Academy of Sciences,
P.O. Box 18, 1581 Budapest, Hungary
e-mail: benko@novell.vmri.hu

probably five major clusters of adenoviruses exist corresponding to the five major classes of Vertebrata. Adenoviruses, which are now suspected to have common origin with enterobacterium phages from the family *Tectiviridae*, are perhaps very ancient indeed, and may have undergone a co-evolution with vertebrate hosts.

1 Introduction

Although adenoviruses have been isolated from every class of vertebrates including fishes, amphibians and reptiles (RUSSELL and BENKŐ 1999), according to the International Committee on Taxonomy of Viruses-approved official taxonomy (BENKÖ et al. 2000) the family *Adenoviridae* contains only two genera (*Mastadenovirus* and *Aviadenovirus*). For decades, the two genera seemed to be sufficient for the proper classification of the majority of the known mammalian and avian adenovirus types, nevertheless the system was compromised by several puzzling exceptions. Namely, half of the ten recognized bovine types did not fit properly into the genus *Mastadenovirus*. These viruses, compared to human adenoviruses (HAdVs) and other mammalian isolates showed unusual biological properties such as restricted growth, peculiar inclusion body morphology, elevated heat resistance and lack of the common genus specific antigen. Similarly, certain highly pathogenic, non-conventional avian adenovirus isolates were also recognized, which did not possess the genus specific antigen. Therefore, the genus *Aviadenovirus* had to be divided into three groups, and close relatedness was seen within the individual groups only.

To date, the full genomic sequence of one or more representatives of each conventional and atypical mammalian and avian adenovirus group is known. Analysis of the genomic organizations combined with results of phylogenetic calculations implies that at least two (probably three) new genera should be created within the family *Adenoviridae*. We have presumed that DNA sequence data of adenoviruses of lower vertebrate host species (frog, snake, fish) might shed light on possible reasons for the occurrence of the exceptional adenoviruses in mammals and birds. Analysis of the frog adenovirus genome revealed that one atypical avian adenovirus is indeed a close relative of the frog isolate (DAVISON et al. 2000). For a better understanding of adenoviral phylogeny, we have initiated molecular biological examinations of adenoviruses of lower vertebrate species. Presently, two adenovirus isolates, one from a corn snake

and another from a white sturgeon, are subjects of genome mapping and DNA sequencing studies in our laboratory. The accumulating DNA sequence and phylogenetic data provide evidence for the existence of five distinct groups, which should be classified into five different genera within the family *Adenoviridae*. The transition of these new concepts into scientific public opinion and recognition of the necessity of radical reform in adenoviral taxonomy, however, appear to be very slow and hampered likely by old stereotypes. This reluctance to accept the new taxonomic proposals is especially inconceivable because according to the original concept of adenovirus classification – in which the genera were determined and named by the classes of the host species (mammals and birds) – the viruses of lower vertebrate classes should also have been assigned to separate genera if differences in the biological and serological properties had warranted it.

The vast majority of our present knowledge about the adenoviral virion structure, genomic organization and replication strategies comes from studies performed on different human serotypes, most frequently on type 2 or 5 (HAdV-2, HAdV-5) that belong to the former subgenus C recently established as species *Human adenovirus C* (Benkö et al. 2000). Members of the other HAdV species (A–F) might show significantly different properties in their replication ability, cell, tissue or organ tropism, receptor usage, pathogenicity, etc. Several slight differences in the genomic organization have also been recognized, nevertheless the six human adenovirus species containing more than 50 serotypes constitute a relatively uniform cluster. Primate, non-human adenoviruses comprising at least 26 different types are obviously closely related to HAdVs. However, adenoviruses originating from non-primate mammalian host species exhibit much greater heterogeneity, as illustrated by the excellent example of bovine adenoviruses (BAdVs).

In the next few pages summarizing our and others' recent results, we try to present lines of evidence that justify the establishment of the proposed new genera and re-classification of several adenovirus types. Based on the emerging picture, a hypothesis on the possible molecular evolution of adenoviruses will be described with special emphasis given to different evolutionary events that might have led to the somewhat unclear present-day situation, which is not easy to interpret. After a short overview on the exceptional adenoviruses, the four different types of genomic organization observed in adenoviruses studied so far will be discussed according to the main transcription units. Then the results of former and more

recent phylogenetic studies will be presented and contrasted with each other. The occurrence and distribution of different adenoviruses from the different host species will be shown with the possible explanations. Finally, an even less firmly based – but certainly fascinating – theory on the origin of adenoviruses will be presented taking us back in time hundreds of millions of years.

2
Unusual Adenoviruses

In the genus *Mastadenovirus*, five bovine isolates (serotypes 4, 5, 6, 7, and 8) distinguished as subgroup 2 BAdVs (Bartha 1969) were recorded as exceptions. By restriction enzyme analysis (Benkö et al. 1988), the genome of these viruses was found to be smaller than the genome of the typical (so-called subgroup 1) BAdVs (types 1, 2, 3, and 9), and no cross-reaction could be detected between the members of the two subgroups in Southern blot hybridizations (Benkö et al. 1990). BAdV-4 has been chosen for further study, and in its very first DNA sequences determined, an interesting codon usage heavily biased towards AT rich triplets was recognized (Benkö and Harrach 1994). Additionally, a strange isolate referred to as strain OAV287 originating from diseased lambs in Australia has also been described (Boyle et al. 1994). Because of the partial cross-neutralization with BAdV-7, no official ovine adenovirus type number has been assigned to it. Nevertheless, OAV287 became the subject of detailed studies as a potential gene delivery vector (Vrati et al. 1995; Khatri et al. 1997; Xu et al. 1997), and was the very first fully sequenced representative of these 'atypical mastadenoviruses' recently proposed to form a third genus (Benkö and Harrach 1998). Most of the characteristic features in the genome organization of this putative new genus were originally derived from the unique genome arrangement found in OAV287 (Both 2002).

In the genus *Aviadenovirus*, only the members of group I (comprising the largest number of virus types isolated from different poultry or wild bird species) were similar to each other and shared common antigens, whereas groups II and III each represented by only one serotype seemed to be very different (Hess 2000; McFerran and Smyth 2000). Group II was assigned to allocate different virus isolates related to three distinguished disease entities, namely the hemorrhagic enteritis of turkey (THE), the marble spleen disease (MSD) of pheasant, and the splenomegaly of

chicken. The isolates originating from the different diseases, however, were indistinguishable by serum neutralization. The official taxonomic name of these viruses is turkey adenovirus type 3, while two earlier turkey isolates (types 1 and 2) are conventional and belong to group I. For ease of reference, the abbreviation THEV will be used in the following. In the first published DNA sequence of THEV including the penton base and core protein genes, a high genomic AT content was also noted (JUCKER et al. 1996). Group III contained one virus, which caused severe disease in layer hen flocks all over the world commencing in 1976; its colloquial name is 'egg drop syndrome '76' or, in short, EDS virus. With retrospective studies, the presence of antibodies to EDS virus was detected in the sera collected before 1976 from a number of bird species including duck, which was supposed to be the reservoir of the virus. Therefore, the official name of the EDS virus is duck adenovirus type 1, but it will be referred to here as EDS virus.

To resolve the problems of taxonomy, subgroup 2 BAdVs and OAV287, as well as the group II and III avian adenoviruses were recently proposed to be re-classified (BENKÖ and HARRACH 1998). As a first step, however, only their removal from the respective genera could be achieved (BENKÖ et al. 2000). Two pending proposals concern the establishment of two new genera. The traditional (*Mastadenovirus* and *Aviadenovirus*) genera contain viruses isolated from host species of the respective class of vertebrates (mammals and birds). The phylogenetic distance and differences between their genomic organization are accepted to reflect the evolutionary distance between the hosts, i. e., the two vertebrate classes. In addition, there are two new genera proposed to contain viruses that are phylogenetically closely related but of miscellaneous (mammalian, avian, amphibian, etc.) origin. The genus *Atadenovirus* was proposed for the allocation of the atypical BAdVs (serotypes 4–8) together with the former avian group III EDS virus. The other new proposed genus *Siadenovirus* contains only two putative members (DAVISON et al. 2000), but both of them, the former group II aviadenovirus THEV and frog adenovirus, have been fully sequenced. Because of the unclear host origin, the proposed names of the two new genera refer to other common properties of the candidate members. The genomes of atadenoviruses studied so far have very high (> 60%) AT content, while siadenoviruses contain a seemingly unique gene, a sialidase gene homolog at the left-hand side of their genomes (DAVISON and HARRACH 2002). In the following comparison of the different genomes, these putative genus names will be used.

3
Four Major Types of Genome Organization

Initially, based on partial DNA sequences from animal (bovine, canine, murine, etc.) mastadenoviruses, the adenoviral genome in general was thought to be very conservative and similar to that of HAdV-2. Variations were seen especially in the E1 and E3 regions, but such differences were frequently encountered also among the different HAdV species. Due to the increase of interest in non-human adenoviral expression vectors and gene delivery systems (Löser et al. 1999), the study of animal adenoviruses intensified considerably resulting in many new partial or complete DNA sequences during the past decade. By now, as many as 16 full genomic sequences are available from animals compared to the previously overwhelming number (five) of HAdVs.

Analysis of the animal adenovirus sequences revealed several unexpected features including unusual genome organizations. The genome of the CELO (chicken embryo lethal orphan) virus, the prototype of conventional group I avian adenoviruses officially called fowl adenovirus type 1 (FAdV-1) was found to be considerably longer and organized somewhat differently to that of the mastadenoviruses (Chiocca et al. 1996). In contrary, the genome size of OAV287 (Vrati et al. 1996b) is much shorter than that of HAdVs and is similar to that estimated for subgroup 2 BAdVs (Benkö et al. 1988). Moreover, OAV287 also has a special genome arrangement, which was analyzed and characterized in detail by Khatri and Both (1998). The full genome sequence of the EDS virus (Hess et al. 1997) and THEV (Pitcovsky et al. 1998) was also determined, and they both were very different from that demonstrated for FAdV-1. Additional full genomic sequences from typical mammalian and avian adenoviruses including mouse adenovirus type 1 (Meissner et al. 2000), BAdV-3 (Reddy et al. 1998b), porcine adenovirus type 3 (Reddy et al. 1998a) and porcine adenovirus type 5 (Nagy et al. 2001), FAdV-9 (Ojkic and Nagy 2000), as well as the most recently completed genome sequence of BAdV-4 (Benkö et al. 2001) all strengthened the observation that basically four major genome organization patterns exist (Davison et al. 2000).

In addition to the genome arrangements of the mastadenoviruses and the aviadenoviruses (typical members of the two traditional genera), two novel genomic organizations were recognized. In accordance with the results of the phylogenetic analyses discussed later, the genome organization of EDS virus, the atypical bovine and some other ruminant adeno-

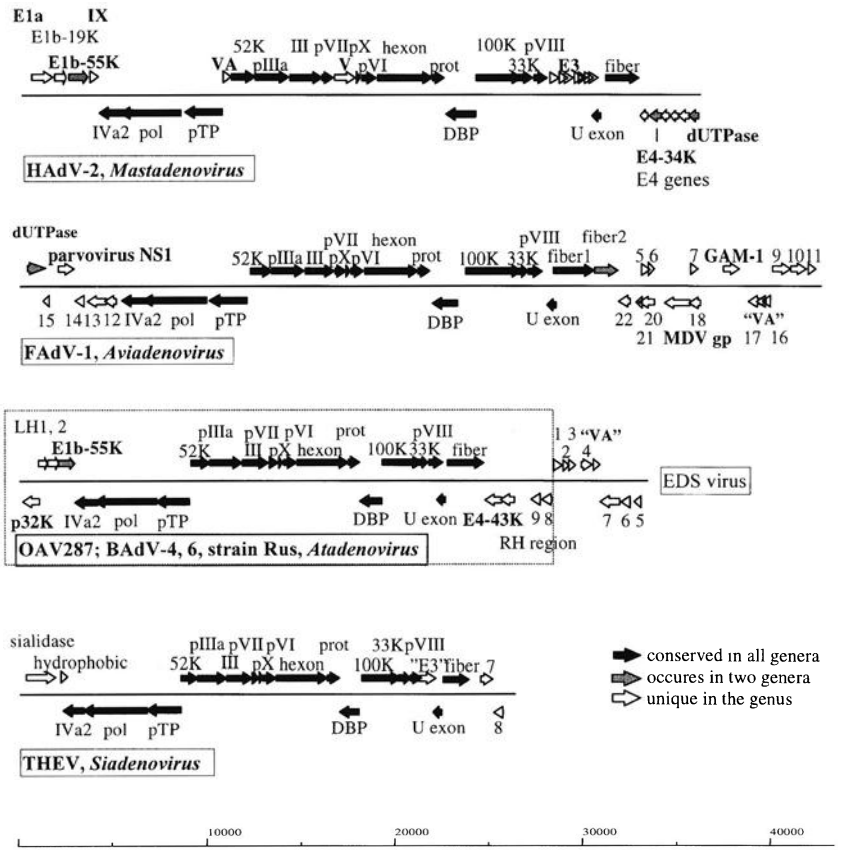

Fig. 1. Schematic presentation of the four different genome organization patterns

viruses was found to be very similar to that seen in OAV287, while the genome arrangement of an amphibian adenovirus isolate originating from leopard frog (*Rana pipiens*) turned out to be almost identical to that of the former group II 'avian' adenovirus THEV.

A schematic presentation of the four types of genome maps is shown in Fig. 1. Apparently, the highest degree of conservation in the adenovirus genome is in the middle part, which encodes the structural proteins and viral enzymes, whereas the two ends – containing mainly regulatory early regions – might exhibit larger variations, and these variations account for the remarkable differences seen in the overall genome sizes. Aviadenoviruses possess the largest genome, in spite of the fact that a couple of

structural proteins and early regions described in mastadenoviruses are missing from them (HARRACH 2002). The extension at the right hand part of the genome containing genes and transcription units unique for the genus *Aviadenovirus* add up to approximately 44–45 kilo base pairs (kb), whereas the smallest genome of about 26 kb is found in the two siadenoviruses. The genome size in mastadenoviruses ranges between 31 and 36 kb, while in atadenoviruses it is between 29 and 33 kb.

Traditionally, the GC content of the DNA was also considered as a characteristic marker of the genera and HAdV subgenera (the present species). With the analysis of the atypical adenoviruses, the formerly known range of GC content has widened and the lowest value is now 34%. The biased base composition of these viruses is accompanied by a preference for AT rich codons. The high AT content in the atadenoviruses and siadenoviruses results in gene overlaps and in a more compact genome. The actual cause of this phenomenon remains unknown.

In the following sections, the genomic organization of the four clusters will be summarized according to the transcription units conventionally classified into early (E) and late (L) regions. It should be emphasized that the presence or lack of the majority of the genes and genome parts in the different clusters seem to be consistent and characteristic of the four genera. Nevertheless, there also exist certain units, which show variations even within the individual genera.

3.1 Early Regions

The left-hand end of the genome contains different genes in the different clusters. In mastadenoviruses, after the left inverted terminal repeat (ITR), the E1A and E1B regions are found without exception. Candidate members of the proposed genera, as well as aviadenoviruses lack the E1A transcription unit. In aviadenoviruses, several open reading frames (ORFs) on both strands have been described at this location, but the function of most of them has not yet been elucidated (CHIOCCA et al. 1996). In atadenoviruses, at the position of E1A, the complementary strand encodes a gene of a novel structural protein p32K, which is transcribed leftward (VRATI et al. 1996b). A homolog of the p32K gene at the same position has been identified in every candidate atadenovirus examined so far including the corn snake isolate (FARKAS et al. 2002) and BAdV-4 through 8

(P. ÉLÖ, unpublished results), therefore it can be considered as a marker of atadenoviruses. In the two sequenced siadenoviruses, a putative sialidase gene was found on the *r* strand in the place of the E1 region. Beside mastadenoviruses, only the atadenoviruses seem to have a homolog of one of the E1B genes, namely the 55K protein, but this homology can be identified with special search options only, e.g., using the Conserved Domain database within GenBank.

The delayed early gene of protein IX, the only structural protein that is not transcribed from the major late promoter (MLP) was found only in mastadenoviruses. Recent studies revealed, that besides cementing the hexons and thus stabilizing the capsid, protein IX also has transcriptional activity (ROSA-CALATRAVA et al. 2001). The lack of protein IX (and V, contained in the late region) in the avi-, at- and siadenoviruses suggests that the virions of these animal adenoviruses are slightly different structurally. Perhaps other yet uncharacterized proteins (e.g., p32K in atadenoviruses) substitute protein IX (and/or V). Nevertheless, an interesting contradiction is that protein IX negative HAdV mutants show thermolabile virion phenotype (JONES and SHENK 1979), while atadenoviruses also lacking this protein were demonstrated to have increased heat resistance (BARTHA 1969). The other delayed early gene, IVa2, is present in every studied member of all four genera.

The conservation level of the E2 region on the *l* strand encoding proteins [DNA-binding-protein (DBP) precursor terminal protein (pTP) and DNA polymerase] essential for replication and maturation seems to be exceptional. Among the few genes identified thus far from the white sturgeon adenovirus genome, two E2 genes (pTP and DNA polymerase) were found, and their estimated relative map position corresponds to that in other genera (M. BENKÖ et al. 2002).

The E3 region apparently exists only in the members of the genus *Mastadenovirus* and no homologous genes were detected in any members of any other genera. The size and content of the E3 region is highly variable within the genus and differs even among HAdVs of different species. As this region is non-essential for in vitro virus replication, it can be deleted and replaced by foreign gene expression constructs. E3 is one of the most commonly used insertion sites in adenovirus gene delivery systems (RUSSELL 2000) and has been extensively studied in numerous animal adenoviruses too (e.g., RAVIPRAKASH et al. 1987; ESFORD and HAJ-AHMAD 1994; EVANS et al. 1998; IDAKAMANTI et al. 1999). In avi- and atadenoviruses, no ORFs are found on the *r* strand between the genes of

pVIII and the fiber. The putative E3 region (one ORF at the conventional location) of siadenoviruses has no homology with mastadenoviral E3 genes. However, a novel short ORF termed the U exon (Davison et al. 1993) was identified on the *l* strand between the genes of pVIII and the fiber in all four genera. This small coding region extends from an initiation codon to a splice donor site, and seems to code for the amino-terminal part of a protein; however, downstream exons have not been identified. U exon is present in each examined type of all four genera with the exception of mouse adenovirus type 1, porcine adenovirus type 5 and BAdV-10. The deduced amino acid sequences are conserved within, but not between the genera. A homolog of atadenoviral U exon was also identified in the snake adenovirus genome (Farkas et al. 2002).

From the E4 region, a single gene of the 34K protein of mastadenoviruses has its homolog in the atadenoviruses only. However, this short E4 region in the atadenoviruses is not at the right-hand end of the genome, but is followed by a unique 'E3' region (encoded on the *l* strand) that has no homology with the E3 regions of the mastadenoviruses or siadenoviruses, and was named E3 only because of its deletable nature (Xu et al. 1997). The corresponding right-hand end of the EDS virus, aviadenoviruses and siadenoviruses are partially characterized and numerous putative ORFs need confirmation and further studies (Chiocca et al. 1997; Payet et al. 1998; Davison et al. 2000).

3.2 Late Regions

As mentioned above and illustrated in Fig. 1, the late regions – occupying the middle part of the genomic DNA – seem to be much better conserved throughout the family. The single exception is protein V, which occurs only in the mastadenoviruses. Protein V is a structural core protein to which important role is attributed during the viral replication cycle. It remains associated with the viral DNA during its transport to the nucleus of infected cells (Matthews and Russell 1998a, b). Protein V is missing from the avi-, at- and siadenoviruses, and no other gene is found in its location as a replacement.

The presence and order of 12 genes (of the following proteins: 52K, pIIIa, III, pVII, pX, pVI, hexon, protease, 100K, 33K, pVIII, fiber) are fully conserved throughout the family, although the number of late mRNA classes in which clusters of genes are transcribed might be different, and