We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

6,600 Open access books available
177,000 International authors and editors
195M Downloads

154 Countries delivered to
TOP 1% Our authors are among the most cited scientists
12.2% Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Chapter

The Diversity of Parvovirus Telomeres

Marianne Laugel, Emilie Lecomte, Eduard Ayuso, Oumeya Adjali, Mathieu Mével and Magalie Penaud-Budloo

Abstract

Paroviridae are small viruses composed of a 4–6 kb linear single-stranded DNA protected by an icosahedral capsid. The viral genes coding non-structural (NS), capsid, and accessory proteins are flanked by intriguing sequences, namely the telomeres. Telomeres are essential for parvovirus genome replication, encapsidation, and integration. Similar (homotelomeric) or different (heterotelomeric) at the two ends, they all contain imperfect palindromes that fold into hairpin structures. Up to 550 nucleotides in length, they harbor a wide variety of motifs and structures known to be recognized by host cell factors. Our study aims to comprehensively analyze parvovirus ends to better understand the role of these particular sequences in the virus life cycle. Forty Paroviridae terminal repeats (TR) were publicly available in databases. The folding and specific DNA secondary structures, such as G4 and triplex, were systematically analyzed. A principal component analysis was carried out from the prediction data to determine variables signing parvovirus groups. A special focus will be put on adeno-associated virus (AAV) inverted terminal repeats (ITR), a member of the genus Dependoparvovirus used as vectors for gene therapy. This chapter highlights the diversity of the Paroviridae telomeres regarding shape and secondary structures, providing information that could be relevant for virus-host interactions studies.

Keywords: parvovirus, telomeres, DNA folding, DNA secondary structure, adeno-associated virus

1. Introduction

Many linear DNA viruses possess terminal repeats (TRs) known to be critical for viral genome stability and propagation [1]. A parallel can be drawn with human chromosome telomeres that are composed of GC-rich repeat sequences of 5–10 nucleotides. In cells, telomeres are critical to maintain the linear structure of the chromosomes. They can adopt specific secondary structures, such as G-quadruplexes (G4), providing structural characteristics for protein binding and genomic stability. In addition, Cellular telomeres play a role in transcription regulation, chromatin compaction, subcellular localization, and chromosome segregation.

Similarly, Paroviridae TRs have been demonstrated to be essential to several steps of the virus cycle. They vary in shape and size from approximatively 100 to 550
nucleotides [2]. Due to the presence of palindromic repeats, TRs can fold into T-, I-, J-, Y-, U- shape or simple hairpin-like structures. Up to now, the global shape has been named without any consensus, for example, Y-shape also being called “rabbit ears.”

Viral genomes in some genera (Ambi- and Itera-densovirus; Ave-, Dependo-and Erythro-parvovirus) are homotelomic meaning that both termini are similar but inverted, whereas, in other genera (Brevi- and Hepan-densovirus; Amdo-, Boca- and Proto-parvovirus), the 5’ and 3’ ends of the linear genome differ and therefore are called heterotelomic. The strand polarity packed in viral capsids may be related to the left and right TRs dissimilarity. Indeed, most of the heterotelomic parvoviruses encapsidate only one strand polarity, mainly negative. This preference may be due to inefficient nicking during replication or incomplete packaging signal at one TR [2]; for example, the minute virus of mice (MVM), a virus of the subfamily Parovirinae and genus Protoparvovirus, harbors a Y-shape left end and a longer U-shape structure on its right end. After replication and ori resolution, the single-stranded DNA of minus polarity is preferentially displaced from the left TR and encapsidated. For parvovirus with both polarities, the proportion can range from 1 to 50% and may be influenced by the host cell in which the virus is produced [3].

Parvoviral TRs are involved in many steps of the virus life cycle. They contain most of the cis-acting information required for genome replication and encapsidation, including tetranucleotide repeats that serve as binding sites for NS1 (Rep) oligomer, a resolution site necessary for the completion of the DNA strand copy, and a packaging signal. Recognized as DNA double-strand breaks (DSB) in the host cell, TR can trigger a DNA damage response (DDR), leading to the circularization and concatemerization of the viral genomes either by non-homologous end-joining (NHEJ) or homologous recombination (HR) [4]. Finally, transcription regulation elements are contained in the genome ends. For example, the MVM TRs contain both symmetric and asymmetric binding sites for transcription factors that modulate expression from the adjacent P4 promoter [5] and the Acheta Domestica Densovirus TRs contain a TATA box used for transcription initiation of NS gene on one side and VP on the other side [6].

The TRs secondary structures, motifs composition, and their role in the virus-cell cycle have been under-examined. In this study, DNA secondary structures of the Paroviridae TRs, including non-canonical secondary structures, have been predicted. We have shown a high diversity of parvovirus telomeres characteristics even within a genus. This chapter may provide significant knowledge for Paroviridae classification and interaction with host cells.

2. The intriguing shape diversity of parvovirus telomeres

2.1 Size, GC content, and shape of parvovirus ends

Although Paroviridae genomes have been extensively studied, in particular for phylogenetic and evolutionary analyses, the sequence and characteristics of their telomeres are not clearly described. Therefore, we analyzed Paroviridae TRs sequences publicly available in the NCBI GenBank database. Paroviridae complete genomes were downloaded. At least one representative virus per genus was selected based on their notoriety and information available on the internal committee on the taxonomy of viruses (ICTV) website. TRs were annotated following GenBank annotation or information available in the literature (Table S1). TR sequences of homotelomic genomes were then verified by aligning the 5’ and 3’ ends. Sequences
differing in length and showing no homology (with no common palindromic regions between 5' and 3' ends) were discarded from the data set. Finally, the presence of palindromic regions was verified by RNAfold (method described later) [7]. A total of 40 Parvoviridae 5' and the 3' TRs sequences were extracted for further analysis. Among those are 17 Densovirinae, 22 Parvovirinae, and 1 unclassified Parvoviridae telomeres.

First, the length of each TR was determined and listed in Table S1. Interestingly, TR length varies within a single genus (Table 1), for example, going from 122 nucleotides for the PcDV to 550 nucleotides for the GmDV in the same genus Ambidensovirus. Second, the percentage of GC was calculated for each parvovirus TR. Figure 1 highlights the GC content diversity of TRs between paroviruses. The minimum and maximum GC content was observed for the left TR of AalDV2 and AAV2 with 32.4% and 69.7%, respectively. Within the genus Ambidensovirus, the percentage of GC ranges from 35% to 58.5% with the lowest GC content being attributed to the 5' TR of the PcDV (Figure 1b). Comparatively, telomeres of the human chromosomes contain 50–55% of G and C bases whereas the whole human genome contains 40.9% GC on average [8].

To visualize the general shape and the secondary structures of the viral TRs, the folding of each parvovirus TR was predicted by RNAfold program using parameters of the Turner model for single-stranded RNA and DNA and the Matthews model for double-stranded DNA [7]. Additionally, mFold program was used in the DNA mode to corroborate the predictions [9]. The most thermodynamically stable structures, or minimum free energy (MFE) structures, obtained on the RNAfold web server were used to propose a classification of the TR (Figure 2). Four groups were constituted according to the number of hairpin loops at their extremity and named H1 (previously named U- and I-shapes in the literature), H2 (corresponding to J-, Y-, T-shapes), H3, and H4 (Figure 2, Table S1). This classification based on the number of terminal hairpin loops after folding and on additional structural characteristics may be more informative and precise than the global shape. Moreover, this nomenclature is applicable to all parvovirus TR. Interestingly, TR sequences and shapes differ within a genus. For example, among Ambidensovirus, CpDV and DicDV 5' TRs are both classified in the H1 group although they only share 43% of sequence homology. In the genus Bocaparvovirus, HBoV1 and BpV1 left ends are 62% homologous in sequence but form a terminal H1 and H2 shape, respectively. Phylogenetic and evolution analyses of Paroviridae have been constructed on the basis of NS1 proteins homology. Telomeres have never been considered as a classification criterion.

<table>
<thead>
<tr>
<th>Subfamily</th>
<th>Genus</th>
<th>5' TR length</th>
<th>3' TR length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parvovirinae</td>
<td>Bocaparvovirus</td>
<td>[140–161]</td>
<td>[161–200]</td>
</tr>
<tr>
<td></td>
<td>Dependoparvovirus</td>
<td>[141–455]</td>
<td>[141–455]</td>
</tr>
<tr>
<td></td>
<td>Erythroparvovirus</td>
<td>[94–383]</td>
<td>[94–383]</td>
</tr>
<tr>
<td>Densovirinae</td>
<td>Ambidensovirus</td>
<td>[122–550]</td>
<td>[122–550]</td>
</tr>
<tr>
<td></td>
<td>Brevidensovirus</td>
<td>[98–182]</td>
<td>[134–165]</td>
</tr>
<tr>
<td></td>
<td>Iteradensovirus</td>
<td>[101–271]</td>
<td>[101–271]</td>
</tr>
</tbody>
</table>

Table 1. Minimal and maximal length of five-prime and three-prime terminal repeats within genera of the Paroviridae family.
2.2 Comprehensive analysis of DNA secondary structural elements

The global analysis of the parvovirus TR has highlighted their broad diversity, even within the same genus. To study the TR divergence, an in-depth prediction of the secondary structures followed by a principal component analysis (PCA) have been realized. Secondary structure elements (Figure 2) and non-B form DNA structures were included as variables in the PCA.

Non-canonical specific structures are susceptible to be recognized by cellular proteins and thus to be essential in the virus-host interactions. For example, a recent study reported that special structures in DNA, such as quadruplex structures, can preferentially bind to IFI16 and trigger more potent type I IFN responses than those produced by the same sequence in dsDNA [10]. Such structures are intrinsic in many viral genomes,
such as those of EBV and HPV [11]. Rich in GC, viral telomeres may also contain non-B DNA structures, such as G-quadruplexes (G4) or triplexes. Therefore, putative G4 and triplexes were determined in all the parvoviral TR. G4 have been non-canonical DNA secondary structures formed by G-rich sequences (Figure 3a). Present in human telomeres, they are suggested to participate in chromosome stability maintenance [12]. G4 have also been shown to be present and play major roles in almost all virus families [13]. G4 have also been described in some parvovirus telomeres [14] but has not been systematically predicted in all parvovirus ends. G4 were predicted using the online tool QGRS-mapper [15] using the search parameters QGRS max length 45, min G-group size 2, loop size from 0 to 12. These criteria are deliberately drastic to increase the stringency and relevance of the G4 prediction. Three values were collected—the raw number of predictive G4, with and without overlaps, and the QGRS max-score rewarding the G4 that are more likely to form. The erythroparvovirus B19V contains four G4 without overlaps which represent the maximum number of these non-B motifs for parvovirus TR. Including overlaps, CeDV and SifDV TRs harbor the highest number of putative tetraplex DNA structures with 296 G4. Of note, the two brevidensoviruses lack any predictive G4 in their ends. No correlation exist between the length of TR and the number of predictive G4 (data not shown).

In parallel, triplexes are important non-B form DNA structures for protein recognition, such as for the binding of p53 factor [16]. Triplex can form at homopurine:homopyrimidine sequences with mirror symmetry (Figure 3b). The triplex package of the R program was used to predict the existence of intramolecular triplex DNA structures in parvovirus TR [17]. Only two triplexes were found, both in the bocaparvovirus BAAV ends.

Finally, a PCA was performed for the forty left TRs and with the following variables—length, GC content, shape, max G-score, raw number of G4 with overlaps, and secondary structures elements (hairpins loops, interior loops, junction loops, and stems) collected from RNAfold analysis. The R package FactoMineR was used [18]. The main PCA variables are the stems and loops (Figure 4a). The “hairpin loops” criteria is one of the most important element allowing division of parvoviruses into groups, hence the relevance of our proposed classification in shapes H1 to H4. Clustering was subsequently realized on the three most informative dimensions.
corresponding to more than 70% of the cumulative variance. Five clusters were obtained (Figure 4b).

Cluster 1, composed of individuals such as the HBoV1 and AAV2, one of the most famous paroviruses in the gene therapy field, is characterized by a high value for the variable “GC content.” Parovirus B19 belongs to cluster 2, a group characterized by high values for the G4 scores and TR length. Cluster 3 mainly depends on the shape class. Individuals in cluster 4 hold a similar number of multiloops and hairpin loops. Finally, viruses in cluster 5 share many DNA structure common features (stems, interior loops, hairpin loops, multiloops, and length). Clusters do not perfectly correlate with phylogenetic classification (Figure 4c), however, we observed that cluster 1 is only composed of Dependoparvovirus and Bocaparvovirus and cluster 5 contains two Ambidensovirus, GmDV and PsiDV. Interestingly, the latter highly differs from other groups.

3. Focus on adeno-associated virus (AAV) inverted terminal repeats (ITR)

The use of vectors derived from the adeno-associated virus (AAV) for gene delivery encounters a growing success for the treatment of a variety of human diseases.

Figure 3.
Two non-B secondary structures: G-quadruplex (G4) and triplex. (a) 3D representation of a theoretical G4. (b) 3D representation of a G4 in the parovirus B19 five-prime telomere and correspondence to the G tetrads sequence (c). (d) Triplex theoretical 2D representation. (e) 3D representation of a triplex of the five-prime bovine AAV terminal repeat and (f) correspondence to the sequence. (b) and (e) figures were obtained using the Jmol software (Jmol: An open-source Java viewer for chemical structures in 3D. http://www.jmol.org/).
Nevertheless, the scientific community has recently faced tragic toxicity of AAV vectors administered intravenously at high doses in several clinical trials [20]. AAV vectors are generated by inserting a recombinant genome usually flanked by AAV-2 inverted terminal repeats (ITR) in an AAV capsid. The recent side-effects observed in human trials have raised the question of DNA sensing, in particular, ITR detection and subsequent cellular responses [21]. Considering the importance of providing new knowledge in this field, a special focus on AAV-2 ITR was included in our study.

The homotelomeric AAV2 possesses two identical ITR of 145 nucleotide-long. The first 125 bases contain three palindromic sequences allowing the ITR to form a T-shape structure composed of two small inverted repeat sequences (BB and CC) and a larger repeated sequence (AA) (Figure 5b). According to our analysis, AAV2 ITR belongs to group H2 (Figure 5c). A fourth proximal region called D remains single-stranded if not annealed to the opposite polarity strand or not in an intramolecular manner to the D region in 3’. Each ITR can be found in two alternative configurations termed “flip” and “flop” distinguishable by the BB’-CC’ orientation (Figure 5a), as a direct result of the replication mechanism. There are nomenclature inconsistencies in the literature. Here, ITR regions are named based on Lusby and Berns’ publication [22] and ordered as followed ABB’CC’A/D from 5’ to 3’.

For historical reasons and the sake of convenience, most of the AAV vectors contain the ITR of AAV serotype 2, the sole viral sequences required for the

Figure 4.
Principal component analysis (PCA) conducted on the five-prime telomere repeats data set. The PCA analysis was conducted using the R software and the Factoshiny package. (a) Contribution graph of each variable. (b) Clusterisation of the 40 parvovirus TR. (c) Correspondence of each parvovirus in the 5 clusters.

[19]. Nevertheless, the scientific community has recently faced tragic toxicity of AAV vectors administered intravenously at high doses in several clinical trials [20]. AAV vectors are generated by inserting a recombinant genome usually flanked by AAV-2 inverted terminal repeats (ITR) in an AAV capsid. The recent side-effects observed in human trials have raised the question of DNA sensing, in particular, ITR detection and subsequent cellular responses [21]. Considering the importance of providing new knowledge in this field, a special focus on AAV-2 ITR was included in our study.

The homotelomeric AAV2 possesses two identical ITR of 145 nucleotide-long. The first 125 bases contain three palindromic sequences allowing the ITR to form a T-shape structure composed of two small inverted repeat sequences (BB’ and CC’) and a larger repeated sequence (AA’) (Figure 5b). According to our analysis, AAV2 ITR belongs to group H2 (Figure 5c). A fourth proximal region called D remains single-stranded if not annealed to the opposite polarity strand or not in an intramolecular manner to the D’ region in 3’. Each ITR can be found in two alternative configurations termed “flip” and “flop” distinguishable by the BB’-CC’ orientation (Figure 5a), as a direct result of the replication mechanism. There are nomenclature inconsistencies in the literature. Here, ITR regions are named based on Lusby and Berns’ publication [22] and ordered as followed ABB’CC’A/D from 5’ to 3’.

For historical reasons and the sake of convenience, most of the AAV vectors contain the ITR of AAV serotype 2, the sole viral sequences required for the
replication and packaging of the recombinant genome in AAV capsids. Additional functions of AAV2 ITR have been described, such as a promoter activity [23], a role in the virus persistence either through genome integration [24] or recombination to form monomeric or concatemeric episomes [25].

Strikingly, the GC content of AAV2 ITR corresponds to the highest score of all studied parvoviral telomeres (69%). No predictive G4 was found using stringent parameters, unlike Satkunanathan et al. who described 18 QGRS inside the AAV2 ITR sequences [14]. In addition, no potential triplex structure was found. According to the PCA (Figure 4), AAV2 belongs to cluster 1 with several other Dependoparvoviruses.

Putative binding sites for human transcription factors (TF) have already been described in AAV2 ITR [26, 27]. We completed this work by analyzing human TF for AAV serotypes 1 to 7 ITRs using the Alggen-Promo tool with 0% sequence dissimilarity (Table 2) [28, 29]. Five human TF sites were found in AAV ITR: C/EBPbeta, Pax-5, YY1, AP-2alphaA, and GR-alpha. The C/EBPbeta was found in most of the
AAV serotypes and unlike the other found TF, is mainly involved in immune responses. In our study, the GR-alpha was only found in AAV5 ITR contains two predictive TF binding sites, one for C/EBPbeta and other for YY1 (Figure 5d). YY1 participates in the initial steps of replication by binding to the p5 promoter region of AAV [30, 31].

### 4. Conclusion

The genomic and structural diversity of parvovirus is today classified by phylogeny analysis showing an expected separation between parvoviruses and densoviruses, but its robustness is relative, suggesting that the introduction of new sequences could change our perception of their evolutionary history [32]. The diversity of sequences, structures, and genomic organizations of parvoviruses suggest evolutionary histories that are probably more complex than those illustrated by current phylogenies. These observations led us to analyze and characterize the intriguing terminal sequences present in all parvoviruses, namely the telomeres.

This chapter highlights the diversity of *Paroviridae* telomeres through a complete analysis of the terminal ends folding and secondary structures. Their length, GC content, and global shape vary even within a genus between phylogenetically closely related viruses. Evolution also led to heterotelomeric viruses with completely different left and right extremities. The diversity suggests high importance of these particular structures. Yet, factors involved in the TR selective pressure are unknown. Cotmore and Tattersall suggested a link between the resolution mechanism, the strand polarity, and the TR conformation [4, 33], while Tijssen et al. suggested that the significant differences in size and secondary structure of genome end between genera might reflect a dependence on specific cellular factors necessary for replication and encapsidation [34]. Consistently, we hypothesized that TR may have evolved according to the interactions with their replicase, helper virus co-factors, and/or cell
host proteins. Based on data integration of predictive DNA secondary structures in a PCA, new groups were made that were distinct from the ICTV phylogenetic classification conducted from the NS replicase sequences.

Additionally, the significance of specific secondary structures in the parvovirus life cycle and the relation with strand polarity of the packaged linear genome are interesting topics deserving further investigations. The MVM, canine parvovirus (CPV), BPV1 (Parvovirinae), and the AalDV2 (Densovirinae) encapsidate only or predominantly negative-strand polarity genomes and possess heterotelomeric TR [35, 36]. On the contrary, the homotelomeric AAV2 encapsidates both strands polarities at the same level. By having different shapes and different secondary structure elements, the TR directly impacts the polarity of the encapsidated strand.

Finally, a special emphasis was put on the ITRs of the adeno-associated virus serotype 2, taking into consideration its importance in the world of gene transfer using viral vectors. Particular motifs and secondary structures within AAV ITR may have a significant impact on gene transfer efficiency. Indeed, it has already been demonstrated that AAV2 ITRs are detected by cellular factors belonging to the NHEJ and HR-DNA damage pathways [37]. The viral telomeres may also be recognized by DNA sensors which subsequently could restrict AAV vectors transduction or activate innate immune responses [21]. Consistent with this hypothesis, a variety of cellular proteins have been shown to interact with AAV2 ITR, such as nucleophosmin (NPM1), a protein involved in ribosome biogenesis and nucleolus transport of basic proteins. Notably, NPM1 binds preferentially G4. The restriction factor FKBP52 in its phosphorylated form also binds to the ITR in the D region, inhibits the second strand synthesis, and consequently decreases transgene expression [38]. Thus, the involvement of ITR recognition by cellular factors is central to understand the extent of subsequent responses to the rAAV DNA that can negatively impact the therapeutic gene expression and cause potential safety concerns for the patients. Using drastic parameters, no putative G4 or triplex were found in AAV2 ITR contrary to a previous study [14]. The formation of these non-conventional DNA motifs highly depends on the adjacent sequences as well as pH and ion concentration conditions and thus requires to be confirmed experimentally.

Acknowledgements

The authors would like to thank Judith Penzes for the initial discussion about phylogeny.

This work was supported by Nantes University Hospital.

Conflict of interest

The authors declare no conflict of interest.
## Appendix

<table>
<thead>
<tr>
<th>Subfamily</th>
<th>Genus</th>
<th>Virus name</th>
<th>Abbreviation</th>
<th>Accession number</th>
<th>5’ TR length</th>
<th>5’ TR shape</th>
<th>3’ TR length</th>
<th>3’ TR shape</th>
<th>Reference used for TR annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parvovirinae</strong></td>
<td>Aveparvovirus</td>
<td>Chicken parvovirus ABU-P1</td>
<td>ChiPV</td>
<td>GU214704.1</td>
<td>206</td>
<td>H4</td>
<td>206</td>
<td>H4</td>
<td>[39]</td>
</tr>
<tr>
<td></td>
<td>Bovaparvovirus</td>
<td>Bovine parvovirus-1 strain Abinantia</td>
<td>BPV1 HBoV1</td>
<td>NC038895</td>
<td>161</td>
<td>H2</td>
<td>161</td>
<td>H1</td>
<td>[40]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Human bocavirus 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>adeno-associated virus 2</td>
<td>AAV-2</td>
<td>NC_001401</td>
<td>145</td>
<td>H2</td>
<td>145</td>
<td>H2</td>
<td>[41]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>adeno-associated virus 5</td>
<td>AAV-5</td>
<td>NC_006152.1</td>
<td>167</td>
<td>H2</td>
<td>167</td>
<td>H2</td>
<td>[42]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adeno-associated virus-Go.1 (caprine)</td>
<td>AAV-Go1</td>
<td>DQ335246</td>
<td>167</td>
<td>H2</td>
<td>167</td>
<td>H2</td>
<td>[45]</td>
</tr>
<tr>
<td></td>
<td>Avian adeno-associated virus ATCC VR-865</td>
<td>AAAV</td>
<td></td>
<td>NC004828</td>
<td>142</td>
<td>H2</td>
<td>142</td>
<td>H2</td>
<td>[46]</td>
</tr>
<tr>
<td></td>
<td>Avian adeno-associated virus strain DA-1</td>
<td>AAV-DA1</td>
<td></td>
<td>NC006263</td>
<td>142</td>
<td>H2</td>
<td>143</td>
<td>H2</td>
<td>[47]</td>
</tr>
<tr>
<td></td>
<td>Avian adeno-associated virus strain YZ-1</td>
<td>AAV-YZ1</td>
<td></td>
<td>GQ368252</td>
<td>141</td>
<td>H2</td>
<td>141</td>
<td>H2</td>
<td>[48]</td>
</tr>
<tr>
<td></td>
<td>Bearded dragon parovirus</td>
<td>BDPV</td>
<td></td>
<td>NC027429</td>
<td>257</td>
<td>H2</td>
<td>257</td>
<td>H2</td>
<td>[49]</td>
</tr>
<tr>
<td></td>
<td>Bovine AAV</td>
<td>BAAV</td>
<td></td>
<td>NC005889</td>
<td>172</td>
<td>H2</td>
<td>172</td>
<td>H2</td>
<td>[50]</td>
</tr>
<tr>
<td></td>
<td>Duck parovirus strain FJM3</td>
<td>DuPV-FJM3</td>
<td></td>
<td>KR075690</td>
<td>359</td>
<td>H1</td>
<td>359</td>
<td>H1</td>
<td>[51]</td>
</tr>
<tr>
<td>Subfamily</td>
<td>Genus</td>
<td>Virus name</td>
<td>Abbreviation</td>
<td>Accession number</td>
<td>5’ TR length</td>
<td>5’ TR shape</td>
<td>3’ TR length</td>
<td>3’ TR shape</td>
<td>Reference used for TR annotation</td>
</tr>
<tr>
<td>-----------</td>
<td>-------</td>
<td>-----------------------------------</td>
<td>--------------</td>
<td>------------------</td>
<td>--------------</td>
<td>-------------</td>
<td>--------------</td>
<td>-------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Duck parvovirus strain M8</td>
<td>DuPV-M8</td>
<td>KR029614</td>
<td>387</td>
<td>H1</td>
<td>387</td>
<td>H1</td>
<td>[52]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Duck parvovirus strain NMZJD110</td>
<td>DuPV-NMZJ</td>
<td>KR075691.1</td>
<td>415</td>
<td>H1</td>
<td>415</td>
<td>H1</td>
<td>[52]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Goose parvovirus</td>
<td>GPV</td>
<td>U25749</td>
<td>444</td>
<td>H1</td>
<td>444</td>
<td>H1</td>
<td>[52]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Muscovy duck parvovirus FM</td>
<td>DPV</td>
<td>NC_006147.2</td>
<td>457</td>
<td>H1</td>
<td>455</td>
<td>H1</td>
<td>[51]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Muscovy duck parvovirus YY</td>
<td>MudPV-YY</td>
<td>KU844281</td>
<td>452</td>
<td>H1</td>
<td>452</td>
<td>H1</td>
<td>[51]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Serpentine adeno-associated virus</td>
<td>SAAV</td>
<td>NC006148</td>
<td>154</td>
<td>H2</td>
<td>154</td>
<td>H2</td>
<td>[53]</td>
</tr>
<tr>
<td>Erythroparvovirus</td>
<td></td>
<td>B19 virus isolate J35 Simian adeno-associated virus</td>
<td>B19V</td>
<td>AY386330.1</td>
<td>383</td>
<td>H1</td>
<td>383</td>
<td>H1</td>
<td>[54]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Porcine parvovirus strain FMV10-1437266</td>
<td>PoPTV</td>
<td>NC022104.1</td>
<td>210</td>
<td>H2</td>
<td>210</td>
<td>H2</td>
<td>[56]</td>
</tr>
<tr>
<td>Unclassified</td>
<td></td>
<td>Porcine parvovirus strain FMV10-1437266</td>
<td>PoPTV</td>
<td>NC022104.1</td>
<td>210</td>
<td>H2</td>
<td>210</td>
<td>H2</td>
<td>[56]</td>
</tr>
<tr>
<td>Densovirinae</td>
<td>Ambidensovirus</td>
<td>Acheta domestica densovirus</td>
<td>AaDV</td>
<td>HQ827781</td>
<td>144</td>
<td>H1</td>
<td>144</td>
<td>H1</td>
<td>[57]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Blattella germanica densovirus</td>
<td>BgDV1</td>
<td>NC005041</td>
<td>217</td>
<td>H1</td>
<td>216</td>
<td>H1</td>
<td>[58]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Culex pipiens densovirus</td>
<td>CpDV</td>
<td>NC012685</td>
<td>285</td>
<td>H1</td>
<td>285</td>
<td>H1</td>
<td>[59]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diaphorina citri densovirus</td>
<td>DcDV</td>
<td>NC030296.1</td>
<td>210</td>
<td>H1</td>
<td>210</td>
<td>H1</td>
<td>[60]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Galleria mellonella densovirus</td>
<td>GmDV</td>
<td>NC_004286</td>
<td>550</td>
<td>H2</td>
<td>550</td>
<td>H2</td>
<td>[61]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Planococcus citri densovirus</td>
<td>PcDV</td>
<td>NC004289.1</td>
<td>122</td>
<td>H2</td>
<td>122</td>
<td>H2</td>
<td>[62]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pseudoplusia includens densovirus</td>
<td>PsiDV</td>
<td>NC019492.1</td>
<td>540</td>
<td>H2</td>
<td>540</td>
<td>H2</td>
<td>[63]</td>
</tr>
<tr>
<td></td>
<td>Brevidensovirinae</td>
<td>Aedes albopictus densovirus 2</td>
<td>AalDV2</td>
<td>NC004285.1</td>
<td>182</td>
<td>H2</td>
<td>134</td>
<td>H2</td>
<td>[64]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Anopheles gambiae densonucleosis virus</td>
<td>AgDV</td>
<td>NC_011317.1</td>
<td>98</td>
<td>H2</td>
<td>165</td>
<td>H2</td>
<td>[65]</td>
</tr>
<tr>
<td>Subfamily</td>
<td>Genus</td>
<td>Virus name</td>
<td>Abbreviation</td>
<td>Accession number</td>
<td>5’ TR length</td>
<td>5’ TR shape</td>
<td>3’ TR length</td>
<td>3’ TR shape</td>
<td>Reference used for TR annotation</td>
</tr>
<tr>
<td>---------------</td>
<td>------------------------</td>
<td>-------------------------------------------------</td>
<td>--------------</td>
<td>-----------------</td>
<td>--------------</td>
<td>-------------</td>
<td>--------------</td>
<td>-------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>I teradensovirus</td>
<td>Bombyx mori densovirus 1</td>
<td>BmDV1</td>
<td>NC003346.1</td>
<td>230</td>
<td>H1</td>
<td>230</td>
<td>H1</td>
<td>[66]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Casphalia extranea densovirus</td>
<td>CeDV</td>
<td>NC004288.1</td>
<td>230</td>
<td>H1</td>
<td>230</td>
<td>H1</td>
<td>[67]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Danaus plexippus plexippus iteravirus isolate Granby</td>
<td>DapDV</td>
<td>NC023842</td>
<td>239</td>
<td>H1</td>
<td>239</td>
<td>H1</td>
<td>[68]</td>
<td></td>
</tr>
<tr>
<td>Dendrolimus punctatus densovirus</td>
<td>DpDV</td>
<td>NC006555.1</td>
<td>200</td>
<td>H2</td>
<td>200</td>
<td>H2</td>
<td>[69]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Helicoverpa armigera densovirus</td>
<td>HaDV</td>
<td>NC015718</td>
<td>101</td>
<td>H4</td>
<td>101</td>
<td>H1</td>
<td>[70]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Papilio polyxenes densovirus</td>
<td>PpDV</td>
<td>NC018450.1</td>
<td>271</td>
<td>H1</td>
<td>271</td>
<td>H1</td>
<td>[71]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sibine fusca densovirus</td>
<td>SfDV</td>
<td>NC018399.1</td>
<td>230</td>
<td>H1</td>
<td>230</td>
<td>H1</td>
<td>[72]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unassigned</td>
<td>Acheta domesticus mini ambidensovirus isolate Kalamazoo</td>
<td>AdMADV</td>
<td>NC022564.1</td>
<td>199</td>
<td>H2</td>
<td>199</td>
<td>H2</td>
<td>[73]</td>
<td></td>
</tr>
</tbody>
</table>

The phylogenetic classification used here refers to the most up-to-date from the International Committee on Taxonomy of Virus (ICTV) published in 2020. Abbreviations were taken from the literature; when not existing, they were created taking the first letters of the virus name. The TR shape were annotated according to our classification proposed in this chapter. The reference used for TR annotations does not always match with the first citation of the virus.

**Table S1.**
List of the forty Parvoviridae terminal repeats analyzed in the study.

DOI: http://dx.doi.org/10.5772/intechopen.102684
References


Lusby E, Fife KH. Nucleotide sequence of the inverted terminal repetition in adeno-associated virus DNA. Journal of Virology. 1980; 34:8


[34] Tijssen P, Pénzes JJ, Yu Q, Pham HT, Bergoin M. Diversity of small, single-stranded DNA viruses of invertebrates and their chaotic evolutionary past. Journal of Invertebrate Pathology. 2016;140:83-96


[47] Estevez C, Villegas P. Sequence Analysis, Viral Rescue from Infectious Clones and Generation of Recombinant Virions of the Avian Adeno-Associated Virus Telomeres
DOI: http://dx.doi.org/10.5772/intechopen.102684


[53] Farkas SL. A parvovirus isolated from royal python (Python Regius) is a member of the genus Dependovirus. Journal of General Virology. 2004;85(3):555-561. DOI: 10.1099/vir.0.19616-0


