***Tupaia* (tree shrew) adenovirus DNA: sequence of the left-hand fragment corresponding to the transforming early region of human adenoviruses**

U. Brinckmann, G. Darai¹ and R.M. Flügel*

Institute for Virus Research, German Cancer Research Center, and
¹Institute for Medical Virology of the University, 6900 Heidelberg, FRG

Communicated by L. Philipson
Received on 26 July 1983

The nucleotide sequence of the left-hand region of the adenovirus DNA of the phylogenetically interesting tree shrew, (*Tupaia belangeri*) has been determined. Transcription signals, initiation codons, splice sites and termination codons were assigned on the basis of its relatedness to the EIA region of human adenoviruses 5, 7 and 12. A consensus sequence for encapsidation of adenoviral DNA that is based on the established packaging region of adenovirus 16 is proposed. The sequenced region includes the gene for a 18 000 dalton polypeptide that corresponds to the EIA proteins of the transforming region of the human serotypes. The deduced amino acid sequence of the 18 000 dalton TAV polypeptide contains a highly conserved central domain that is homologous to the middle of the 30 000 dalton polypeptides of human adenovirus indicating the significance of the central amino acid residues for the biological function of the EIA proteins.

Key words: adenovirus DNA sequence/*Tupaia*/packaging sequence

Introduction

The structure and function of certain DNA regions within the left-hand region of adenoviruses has been of particular interest. The region EIA of the human serotypes 2, 5, 7 and 12 has been studied and characterised in great detail (see Philipson, 1979; Tooze, 1981). Thus, the origins of transcription for the early region EIA have been precisely mapped near coordinate 1.1–1.5 for different human serotypes. Furthermore, the EIA region and its gene products have been shown to be required for the expression of the remaining early viral genes (Pettersson *et al.*, 1976; Berk *et al.*, 1979; Jones and Shenk, 1979). Moreover, through the use of transfection experiments with defined adenoviral DNA fragments it has been established that this region of the genome and its gene products are required for cell transformation *in vitro* (Van der Eb *et al.*, 1979; Gallimore *et al.*, 1974; Shiroki *et al.*, 1977). In addition, certain cellular proteins, e.g., the heat-shock proteins are induced by the EIA gene products (Nevins, 1982).

There is evidence that the EIA 13S mRNA and/or its gene product of adenovirus 5 play a regulatory role in gene expression (Persson *et al.*, 1981; Ricciardi *et al.*, 1981). In this context, it is noteworthy that the comparison of the deduced amino acid sequences of EIA polypeptides between the different human adenoviruses 2, 5, 7 and 12 revealed strong regions of homology (Van Ormondt *et al.*, 1980). One way to define the domains of the EIA proteins which are necessary for gene expression and transforming ability is to perform comparative sequence studies of the EIA region of different

adenoviruses. Therefore, it was of interest to characterise the corresponding left-hand DNA elements of another transforming adenovirus which is phylogenetically separate from the human serotypes studied so far. We recently isolated and characterised an adenovirus (TAV) from *Tupaia belangeri* (the tree shrew) (Darai *et al.*, 1980; Matz *et al.*, 1980; Faissner *et al.*, 1980; Brinckmann *et al.*, 1983).

The inverted terminal repetition of (ITR) TAV DNA have been reported elsewhere (Brinckmann *et al.*, 1983). Here we report nucleotide sequences from the DNA region of TAV which corresponds to the EIA region of human adenoviruses which enables us to predict a reduced size for the adenoviral encapsidation sequence and to focus attention on the central domain of the EIA proteins.

Results and Discussion

The sequencing strategy used is outlined in Figure 1. The sequence from nucleotide 1 to 765 was derived from DNA which was prepared from wild-type TAV. The sequence downstream from the *Bam*HI cleavage site was derived from the *Bam*HI fragment B cloned into the plasmid pAT153. The nucleotides spanning the *Bam*HI cleavage site were determined from wild-type virus DNA. In order to extend the previously determined sequence (Brinckmann *et al.*, 1983) of the ITRs of the left-hand end of TAV DNA, use was made of the *Bam*HI fragment C (1–767). The DNA fragment was purified by agarose gel electrophoresis and recleaved with either one of the following endonucleases: *Aos*II, *Ava*I, *Ava*II, *Alu*I, *Dde*I, *Hpa*II and *Hha*I. The generated fragments were labelled at their 3' ends with the Klenow fragment of the *Escherichia coli* DNA polymerase I and an appropriate [α -³²P]dNTP defined by the recognition sequence of the given endonuclease. In some cases the 5' termini of the DNA fragments were labelled with T4 polynucleotide kinase and [γ -³²P]ATP which were recut subsequently with another suitable restriction enzyme to separate both 5'-labelled fragments from each other. The labelled fragments were sequenced according to Maxam and Gilbert (1977). More than 80% of the sequence was determined from both strands or at least twice when the same strand was used.

The sequence shown in Figure 3 starts at nucleotide 1 and ends at position 1215. The nucleotide sequence confirms the presence and absence of all restriction sites mapped here (Figure 2) or as previously described (Matz *et al.*, 1980). The nucleotides from 450 to 861 of the established sequence represent an apparently uninterrupted translational reading frame assuming that the first AUG is used as initiation codon. The corresponding AUG initiation codon of a hypothetical TAV polypeptide at nucleotide 450 is preceded by a typical signal for initiation of transcription by RNA polymerase II (Corden *et al.*, 1980), TATATA, at nucleotides 352–357 (Figures 3 and 4). This promoter is itself preceded by a CAT-box sequence at 311–315, CACAC. Similar signal sequences have been implicated in promoter regulation of many eukaryotic and viral mRNAs (Corden *et al.*, 1980; McKnight *et al.*, 1981). Based on these assignments and on the strong homology with the EIA region of ad12, a potential cap site

*To whom reprint requests should be sent.

Fig. 1. Diagram of sequenced TAV DNA fragments and corresponding restriction endonuclease cleavage sites. The top line shows distance in bp from the left terminus of TAV DNA. The restriction enzymes have been abbreviated as follows: *Aos*II (A), *Bam*HI (B), *Ava*I (C), *Dde*I (D), *Ava*I (E), *Hha*I (F), *Alu*I (G) and *Hpa*II (H). Sequences covered in individual reactions and gels are shown as arrows pointing from the 5' to 3' and represent fragments that were labelled at the 3' end. Arrows with an asterisk mark fragments labelled at the 5' end with T4 polynucleotide kinase and [γ - 32 P]ATP.

Fig. 2. Physical map for a number of restriction sites in the left-end region of TAV DNA. Lower line shows distance in bp from the left terminus of TAV DNA.

1 TATCATCAAT TATATATCTG AGCATTTTGA CGTAATGAGG TGACAGCTAA
51 GTTCGAAGTG TCAATCTGTT GTGGAGGTAT GTTTGTGTAG GTTTGAGG
101 GTTTTATTTT TGGTGGGTTT GCGATTTTGG GCGGTTTCCG GAGAACGTTG
151 AGTATATGAA GTTACAGGCG CAAATCTGGA TATTCTGGC GGTTCAGGTTG
201 TATCTTTTGA CCGATTTTATG AGAGGTGTGT TAATTTTAT GTTATTTGTG
251 TATTTTATTT TCTGATAT ATATATGATG TTTCTGTGAA GTTCTATGAT
301 GTTCTGATAT TACAGAGTTA ATATGATGAT TTAACGAGCG CCGTTCTCTT
351 TATATATCTG GTTACAGGTTT TATCTGTGTT GAGCTTTTCT TTAACAGCTG
401 TGAGCTGTTG TTTAGAGGCT TATCTGTCTT GTTGGGTTCT TAAAGAGGAG
451 GTTCGAAGTG TCAATCTGTT GTTACAGGTT GTTTTGTAGA GTTTTGTAT
501 TATGATGTCT GTTCTGTGGA GTTCTGATCT TCTCTGGAG CCGTTGTCTT
551 TATGATGTCT GTTCTGTGGA ATATTTTGGT GAGCTGTGAT CCGTTGTCTT
601 AGGCTGTGGA TCGGTTATTT CCGTATGTTT TGTGTGAGGA GGTTCAGGTT
651 GTTTCAGGTT TCGTCTTTGA AATCAATGAA GAGGTTTAAG CCGTTGTCTT
701 TCGAGAGTCT TCGAGAGGAG CAGGACAAGA ATGAGATCTA TGTAAAGCA
751 AGGCTGAGAT GAGTGGGAT CCGTCTATCT TCGTCTTTCT ATGCTATAT
801 CCGTAAAGCG CCGTCTTTGT TTAAGTAAAG TATGTTTCTT GCAATTTTCT
851 GTTGTGTCTT TATGCTTTCT ATCATGATTT TATGTTTCTT AGGCTGTG
901 AGGAGCTGGT AAGACCAAGA GCAACACGAA GGGGTGGCTG AAAATTTACT
951 GAAACGTCAA GCGGACAGTT CCGTTGTGCT ATCTGTGCA AAGGCTTTT
1001 GTTGTATGAG ACTTTTAAAT GCGATTTTGG CAATACAGGG TGCTAAATC
1051 TGTGTGTGTT GTTGTCTCTA ATCAATCTCT GGTGGTGTCT AGGTTCTTCT
1101 GCGCTAAGAT TACTAAGTTA ATTTATTTCA TAAAGTTTCT TAATCTGTAT
1150 CAATTTTATG ACTCGCTTCT AGGCGTTTAC TTGAAGGCTT TGCGTAGTGG
1201 TGATTTTCTT AAAAT

Fig. 3. Nucleotide sequence of fragment *Bam*HI C and part of *Bam*HI B of TAV DNA. Promoter signals are indicated by dotted lines. Initiator and terminator codons of the 18-kd polypeptide are marked by open and closed circles. The proposed encapsidation sequence is boxed. Arrows mark inverted repeats. The end of the inverted terminal repetition is marked by an asterisk, the poly(A) addition signal is underlined once.

Fig. 4. The coding regions of early region of TAV DNA. The top line represents the left-end region of TAV DNA in bp. The next three lines refer to the three possible reading frames. The heavy portions in these frames are open for translation; vertical lines indicate termination codons. The lower interrupted lines represent the hypothetical TAV mRNAs with the splice sites based on comparison to human adenoviruses. The box designated A+T represents the A+T-rich region, P = promoter.

for an early TAV mRNA can be presumed close to nucleotide 381 (Figures 3 and 4). The sequence AA(C)TAAA which precedes most poly(A) addition sites is located at nucleotide 1129. In addition, there are direct repeats in the 5'-non-coding region e.g., GTTTA at 819 and 858 which correspond to the so-called -100 box (Dierks *et al.*, 1983). Nearly all distances between characteristic and regulatory sites in the 5'-non-coding EIA region of the TAV genome are approximately the same as those found in the EIA region of the extensively studied ad5, 7 and 12. The relative non-variance in length between regulation signals holds also true for the hypothetical splice donor and acceptor sites in TAV DNA which have been deduced by comparison with the corresponding regions and sites in the human adenoviruses. As shown in Figure 4 the different splice donor sites and the common splice acceptor site of the TAV EIA region show a striking homology, namely 74% identical nucleotides, to EIA of ad7 DNA. There is a 70% A+T-rich region in TAV DNA between nucleotides 827 and 892 (Figures 3 and 4). A corresponding A+T-rich region has been shown to occur and to be spliced out in the mRNAs of human ad2, 5, 7 and 12 (see Tooze, 1981). The presence and length of this A+T-rich sequence in TAV DNA lends further support to the first hypothetical splicing pattern shown in Figure 4 (827–892) linking together reading frame 3 and 1 and thereby permits the deduction of an amino acid sequence for the largest TAV polypeptide. However, there are other splice donor sites between the initiator at 450 and the probable splice donor site at nucleotide 827 which would obey the GU-AG rule, e.g., those at 496, 628 and 642. S1 nuclease RNA mapping experiments are necessary to prove whether they can serve as splice signals. The nucleotide homology between TAV and human adenovirus DNA in the EIA region ranges from 46.3% for ad5 (562 identical letters) and 47.0% for ad7 (572 identical letters) to 50.6% for ad12 (615 identical letters). The high homology at the nucleotide level in the vicinity of the splice sites supports the notion that the viruses are related and is further strengthened by the deduced amino acid sequence of the 18-kd TAV polypeptide which is compared with the deduced largest EIA polypeptide of ad5 shown in Figure 5B.

The TAV polypeptide of 162 amino acids is surprisingly

Fig. 5. Amino acid sequence of the TAV largest polypeptide (18 kd) as deduced from the nucleotide sequence of TAV DNA (**A**) and the comparison of its corresponding domain to the largest EIA polypeptides of the human ad5 (**B**). Repeats are underlined. To maximise homology, dashes were introduced into the sequence at appropriate positions. Symbols are as follows: □ STPAG hydroxyl/small aliphatic, ⊕ NDEQ acid/acid amide, ~ ⊖ ⊖ hydrophilic, + HRK basic, . MILV aliphatic, ▽ FYW, aromatic. Upper row in **B** is TAV, lower in ad5.

short and it reveals the highest degree of homology of 41.4% when compared with the 31.9-kd polypeptide of ad5 (Dijkema *et al.*, 1980) (37% to the 28 kd of the ad7 and 35.8% to the 30 kd of the ad12 polypeptide). When the central part of the 18-kd TAV polypeptide (residues 86–132 in Figure 5) is inspected for identical amino acids, the homology reaches values of 63% for ad5, 61% for ad7, and 50% for ad12. Since TAV is a weakly transforming virus (Darai *et al.*, 1980) the degree of homology of the TAV 18-kd polypeptide to the EIA polypeptides does not seem to run parallel to the varying degree of transforming capacity from the subgroups A, B and C of the human adenoviruses 12, 7 and 5 (see Tooze, 1981). The high homology in the central domain of the adenoviral EIA polypeptides probably reflects a common functional property residing in this region required for inducing expression of viral genes during the early phase of infection. This conclusion is in agreement with the regulatory defects that result from a point mutation in the central part of the first exon of the ad5 hr 1 mutant as reported by Ricciardi *et al.* (1981).

At least three TAV polypeptides with a variable number of amino acids can be predicted from the splice pattern. The shortest TAV polypeptide is 62 amino acids long (6.9 kd) is formally analogous and homologous to the 6.3-kd ad7 polypeptide (Van Ormondt *et al.* 1980). Recent evidence reported by Svensson *et al.* (1983) shows that in the case of ad2, the carboxy-terminal part of the corresponding ad2 late polypeptide derives from a different reading frame and also results in a short polypeptide of only 6.1 kd. Here, only the largest TAV polypeptide will be considered in more detail, since it has a high degree of homology to those of ad12 and ad5 (Figure 5). On close inspection, the TAV 18-kd polypeptide which has an amino acid composition similar to the 30-kd polypeptide of ad12 can be divided into three structural domains analogous to SV40 T antigen (Clayton *et al.*, 1981). (i) The amino-terminal region of relatively low homology containing three tetrapeptide repeats -SPSF-(9-12,25-28) FPDC (57-60,84-97) and EEVE (63-66,77-80) (Figure 5A) which are not found in this part of the human EIA polypeptides. (ii) The central domain which extends from the

aspartic acid residue 86 to the glutamic acid residue 134 (TAV polypeptide numbering system of Figure 5A) is characterised by its high homology to the EIA polypeptide of all three human adenoviruses, in particular to that of ad5 (Figure 5B), its high contents of acidic amino acids and several potential phosphorylation sites such as the serine residue 129 which is in an analogous environment to that of certain phosphorylated serines in the SV40 T antigen (Van Roy *et al.*, 1983). It has been reported that at least one human adenovirus EIA polypeptide is a phosphoprotein (Ricciardi *et al.*, 1981; Gaynor *et al.*, 1982). The equidistantly located four cysteine residues at positions 95, 98, 112 and 115 in TAV are conserved in the human serotypes. The secondary structure of the central domain with its high homology to the EIA human polypeptides was predicted using a computer program according to the method of Robson (1978). The first half of this region of strong homology (86–109) seems to be of a predominant turn/loop character, whereas the second part (110–132) is predicted to be of β -sheet structure. (iii) The third COOH-terminal domain is relatively short comprising ~25 amino acids with limited homologies to the EIA polypeptides of the human adenoviruses, suggesting that this part may not be essential for biological function.

Since it is relatively unusual that each of three different tetrapeptides occur twice in a short polypeptide of 162 amino acids, a computerised sequence alignment program for finding duplications within the TAV polypeptide was applied. A value of 33% homology within the TAV polypeptide sequence was found which might indicate evidence for a possible gene duplication event from an ancestral DNA sequence.

So far homologies and similarities of the sequenced TAV DNA region to EIA of ad12 have been emphasized including the length of the ITR of 166 (TAV) and 164 (ad12), and the relatively short distance between the end of the ITR and the promoter signals and the first initiation codons when compared with those of ad5 and 7. However, there are differences in the overall organisation in this part of the TAV genome, among them the even shorter distance from the TATATA-box and initiation codons to the end of the ITR than in the case of ad12, the relatively short 18-kd TAV polypeptide which does not contain glycosylation sites of the type NXS or NXT in contrast to the EIA polypeptides of the three human serotypes. In addition, another open reading frame starts at nucleotide 1007 with high homology to the EIB 20-kd polypeptide of human adenoviruses. Corresponding promoter signals for TATA, CAT - and - 100 boxes are located at position 878 (ATTTA), at 842 (CAAT), and at 819 (GT-TTA) that is repeated at 858. The distance between the stop codons of the human EIA polypeptides and the initiator AUG for the EIB protein is much longer in the human serotypes (151 - 172 nucleotides), whereas in TAV these sites are separated by three nucleotides only. Furthermore, the relatively low number of 449 nucleotides in the 5'-non-coding area of the TAV region points to an even more economical use of the information contents of TAV DNA, since this part of adenovirus genomes has been postulated to contain information required for encapsidation of viral DNA, as reported by Hammerskjöld and Winberg (1980).

The search for a common encapsidation sequence was based on a 100-bp long sequence established for human ad16 (Hammar skjöld and Winberg, 1980), a member of subgroup B to which ad7 also belongs. The corresponding 110-bp long sequence of ad7 was used to search for common features

within the first 500 bp of TAV and subsequently of ad5 and 12. Within all four adenoviral 5'-non-coding regions, 35 nucleotides were found to be identical and 16 were of the same type, i.e., either a purine (R) or a pyrimidine (Y) base residue at the same position. In all cases the identities (46.3%) are located within the 110 bp long sequence. A consensus sequence which possibly contains an essential part of the sequence necessary for encapsidation was found to occur within this 110-bp sequence of all adenoviral 5'-non-coding regions:

```
AD7   350 TATTTGCCGAGGGCCGAGTAGACTTTGA 377
AD5   341 TATTTGTCTAGGGCCGCGGGGACTTTGA 368
TAV   182 TATTTGGGCGGTCACGCTGTACTTTTGA 209
AD12  292 TATTTACCGAGGGCAGAGTGAAGTCTGA 319
      XX X X X      X      X      X X X X
```

Consensus TATTT(R)₃RG(N)₄G(N)₄RNYTYTGA

In addition, there are several inverted repeat sequences upstream of this consensus sequence in the case of ad7, TAV and ad12. Similarly located inverted repeats were reported for the bacteriophage λ encapsidation system (Miwa and Matsubara, 1982).

Finally, a computer search for coding information was carried out for the left strand of TAV EIA region, since evidence for viral RNA, novel polypeptides, and open reading frames for the left-ward transcribed strand of ad2 was recently reported (Katze *et al.*, 1982; Gingeras *et al.*, 1982). However, open reading frames longer than one coding for 58 amino acids were not found in this region of TAV DNA neglecting possible spliced mRNAs.

More work is necessary to examine the role of the TAV EIA region in the process of transformation and in the regulation of viral genes and their products.

Materials and methods

Cells and virus

Tupaia adenovirus was propagated in cultures of *Tupaia* baby fibroblast cells and viral DNA extracted from purified virus as described previously (Darai *et al.*, 1980).

Restriction endonucleases and cleavage of DNA

Restriction enzymes were purchased from BRL and Boehringer and used according to the specifications from the manufacturers. The *AosII* enzyme was purchased from Vector Bioproducten B.V., Leiden, The Netherlands.

Cloning of the *Bam*HI fragment B and propagation of the recombinant

For molecular cloning of the *Bam*HI fragment B of TAV DNA (0.02–0.08 map unit, 1.36 kb) a recombinant DNA plasmid pTA-BE-A harbouring 13.7 kb of TAV genome (0.02–0.43 map unit, inserted into the *Eco*RI/*Bam*HI sites of the bacterial plasmid pAT153 (Twigg and Sheratt, 1980)) was used. pTA-BE-A was cleaved with *Bam*HI restriction endonuclease, the *Bam*HI B fragment of TAV was selected and inserted into the *Bam*HI site of pAT153. Thereby, a recombinant DNA plasmid (pTA-B-B) was constructed harbouring only the *Bam*HI B fragment of TAV DNA. *E. coli* C 600 was used for amplification of both selected recombinant DNA plasmids (Grosveld *et al.*, 1981). Extraction and purification of the plasmid DNA was done as described (Ish-Horowitz and Burke, 1981; Stenlund *et al.*, 1980).

Preparation of the *Bam*HI fragment C

Wild-type viral DNA was fully digested with the *Bam*HI enzyme, purified by electrophoresis on agarose gel, eluted from the gel, precipitated with ethanol and used as starting material for nucleotide sequencing of the left-hand end of the viral genome.

Sequencing procedure

Sequence analysis was performed according to the method of Maxam and Gilbert (1977). The sequence data were handled by using methods for protein analysis (Staden, 1978; Smith *et al.*, 1981; Kabsch and Sander, 1983; Lim, 1974; Garnier *et al.*, 1978).

Acknowledgements

We thank Helmut Bannert for outstanding assistance and Dr. Stüber, EMBL and Dr. James Fickett, Los Alamos, for providing computer programs. Dr. C. Sander, Max-Planck-Institut für Medizinische Forschung, Heidelberg, for discussion and help with the computerised protein analysis. This publication is dedicated to Friedrich Cramer on the occasion of his 60th birthday. This work is part 2 of U. Brinckmann's Ph.D. thesis at the TH Darmstadt.

References

- Berk, A.J., Lee, F., Harrison, T., Williams, J. and Sharp, P.A. (1979) *Cell*, **17**, 935-944.
- Brinckmann, U., Darai, G. and Flügel, R.M. (1983) *Gene*, **24**, 131-135.
- Clayton, C.E., Lovett, M. and Rigby, P.W. (1982) *J. Virol.*, **44**, 974-982.
- Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kédinger, C. and Chambon, P. (1980) *Science (Wash.)*, **209**, 1406-1414.
- Darai, G., Matz, B., Flügel, R.M., Grafe, A., Gelderblom, H. and Delius, H. (1980) *Virology*, **104**, 122-148.
- Dierks, P., Van Ooyen, A., Cochran, M.D., Dobkin, C., Reiser, J. and Weissmann, C. (1983) *Cell*, **32**, 695-706.
- Dijkema, R., Dekker, B.M.M., Van Ormondt, H., De Waard, A., Maat, J. and Boyer, H.W. (1980) *Gene*, **12**, 287-299.
- Faissner, A., Darai, G. and Flügel, R.M. (1980) *Intervirology*, **14**, 272-276.
- Gallimore, P.H., Sharp, P.A. and Sambrook, J. (1974) *J. Mol. Biol.*, **89**, 49-72.
- Garnier, J., Osguthorpe, D.J. and Robson, B. (1978) *J. Mol. Biol.*, **120**, 97-120.
- Gaynor, R.B., Tsukamoto, A., Montell, C. and Berk, A.J. (1982) *J. Virol.*, **44**, 276-285.
- Gingeras, T.R., Sciaky, D., Gelin, R.E., Dong, J.B., Yen, C.E., Kelly, M.M., Bullock, P.A., Parsons, B.L., O'Neill, K. and Roberts, R.J. (1982) *J. Biol. Chem.*, **257**, 13475-13491.
- Grosveld, F.G., Dahl, H.M., de Boer, E. and Flavell, R.A. (1981) *Gene*, **13**, 277-237.
- Hammarskjöld, M.L. and Winberg, C. (1980) *Cell*, **20**, 787-795.
- Ish-Horowitz, D. and Burke, J.F. (1981) *Nucleic Acids Res.*, **9**, 2984-2998.
- Jones, N. and Shenk, T. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 3665-3669.
- Kabsch, W. and Sander, C. (1983) *FEBS Lett.*, **155**, 179-191.
- Katze, M.C., Persson, H. and Philipson, L. (1982) *EMBO J.*, **1**, 783-789.
- Matz, B., Delius, H., Flügel, R.M. and Darai, G. (1980) *J. Gen. Virol.*, **51**, 421-423.
- Maxam, A. and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 560-564.
- McKnight, S.L., Gavis, E.R., Kingsbury, R. and Axel, R. (1981) *Cell*, **25**, 385-398.
- Miwa, T. and Matsubara, K. (1982) *Gene*, **20**, 267-279.
- Nevins, J.R. (1982) *Cell*, **29**, 913-919.
- Persson, H., Monstein, H.J., Akusjärvi, G. and Philipson, L. (1981) *Cell*, **23**, 485-486.
- Pettersson, U., Tibbetts, C. and Philipson, L. (1976) *J. Mol. Biol.*, **101**, 479-501.
- Philipson, L. (1979) *Adv. Virus Res.*, **25**, 357-405.
- Ricciardi, R., Jones, R., Cepko, C., Sharp, P. and Roberts, B. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 6121-6125.
- Shiroki, K., Segawa, K., and Shimojo, H. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 2274-2278.
- Smith, T.F., Waterman, M.S. and Fitch, W.M. (1981) *J. Mol. Evol.*, **18**, 38-46.
- Staden, R. (1978) *Nucleic Acids Res.*, **5**, 1013-1015.
- Stenlund, A., Perricaudet, M., Tiollais, P. and Pettersson, U. (1980) *Gene*, **10**, 47-52.
- Svensson, C., Pettersson, U. and Akusjärvi, G. (1983) *J. Mol. Biol.*, **165**, 475-499.
- Tooze, J., ed., (1981) *DNA Tumor Viruses*, 2nd edn. revised, Cold Spring Harbor Laboratory Press, NY, pp. 383-576e.
- Twigg, A. and Sheratt, D. (1980) *Nature*, **283**, 216-218.
- Van der Eb, A.J., Van Ormondt, H., Schrier, P.I., Lupker, J.H., Jochensen, H., van der Elsen, P.J., DeLeys, R.J., Maat, J., Van Beveren, C.P., Dijkema, R. and De Waard, A. (1979) *Cold Spring Harbor Symp. Quant. Biol.*, **44**, 383-399.
- Van Ormondt, H., Maat, J. and Dijkema, R. (1980) *Gene*, **12**, 63-76.
- Van Roy, F., Fransen, L. and Fiers, W. (1983) *J. Virol.*, **45**, 315-331.

Note added in proof

The left hand terminus of human ad16 has been sequenced. Our consensus sequence proposed for packaging adenovirus DNA is contained within the ad16 sequence at nucleotides 349–376 (M.L. Hammarskjöld, G. Winberg, personal communication).