Epidermodysplasia verruciformis associated human papillomaviruses present a subgenus-specific organization of the regulatory genome region

Armin Ensser and Herbert Pfister
Institut für Klinische und Molekulare Virologie, Universität Erlangen-Nürnberg, Loschgestraße 7, 8520 Erlangen, FRG

Received March 6, 1990; Revised and Accepted May 30, 1990 EMBL accession nos X52059–X52062 (incl.)

ABSTRACT

The regulatory regions of the human papillomaviruses (HPVs) 9, 17, 20, and 36 were mapped, sequenced, and aligned. They revealed an arrangement of putative protein binding sites specific for the epidermodysplasia verruciformis associated HPVs. Three subgroups could be differentiated.

INTRODUCTION

At least 18 human papillomavirus (HPV) types seem to be specifically associated with epidermodysplasia verruciformis (ev), a lifelong skin disease characterized by the development of multiple flat warts and macular skin lesions (1,2). In skin cancers, which occur in 30% of the patients, HPVs 5 and 8 are most prevalent; HPV 14, 17 and 20 have been found once each.

The nucleotide sequences of HPVs 5, 8, 19, and 25 (3–5) revealed the typical genome organisation of papillomaviruses except for a surprisingly short noncoding region between open reading frames (ORF) L1 and E6. This genome segment contains important cis-acting regulatory elements of papillomaviruses (6) and is therefore called long control region (LCR). All ev-LCRs showed good homology in the regions of the putative binding sites of the viral, transcription-activating E2 protein, and particularly in two motifs with 33 and 29 bp, respectively, called M33 and M29 (5).

MATERIALS AND METHODS

The cloning of DNAs of HPV 9, 17a, 20, and 36 was described previously (7–10). Standard protocols were used for all restriction enzyme digestions, preparations of plasmid and phage DNAs, molecular cloning, radioactive labelling and blot hybridisation (11). ^{35}S-Dideoxy-plasmid-sequencing of both strands with commercially available kits (pUC-Sequencing-kit, Boehringer Mannheim and T7-Sequencing-kit, Pharmacia) followed the instructions of the manufacturers. For sequence analysis, the UWGCG-program package (version 5.3) was used (12).

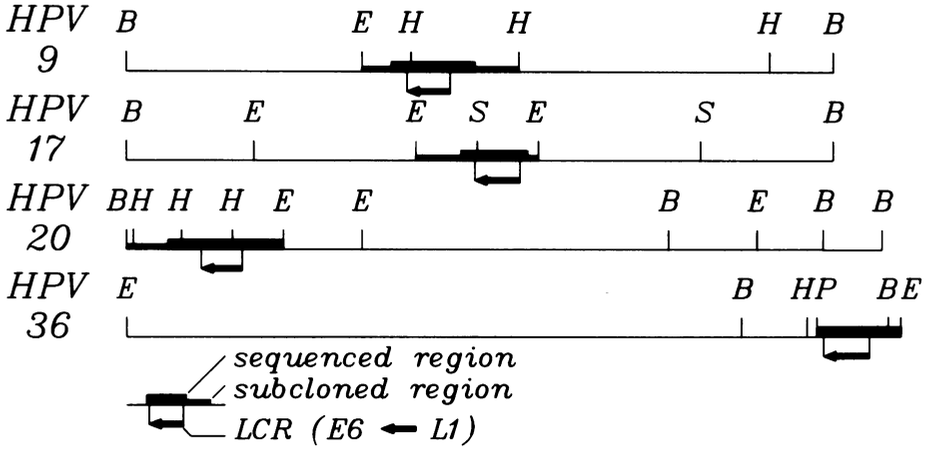

Fig.1. Physical maps of restriction enzymes cleavage sites in the genomes of HPVs 9, 17, 20, and 36, and localisation of the sequenced regions and the LCRs. B = Bam HI; E = Eco RI; H = Hind III; P = Pvu II; S = Sac I.

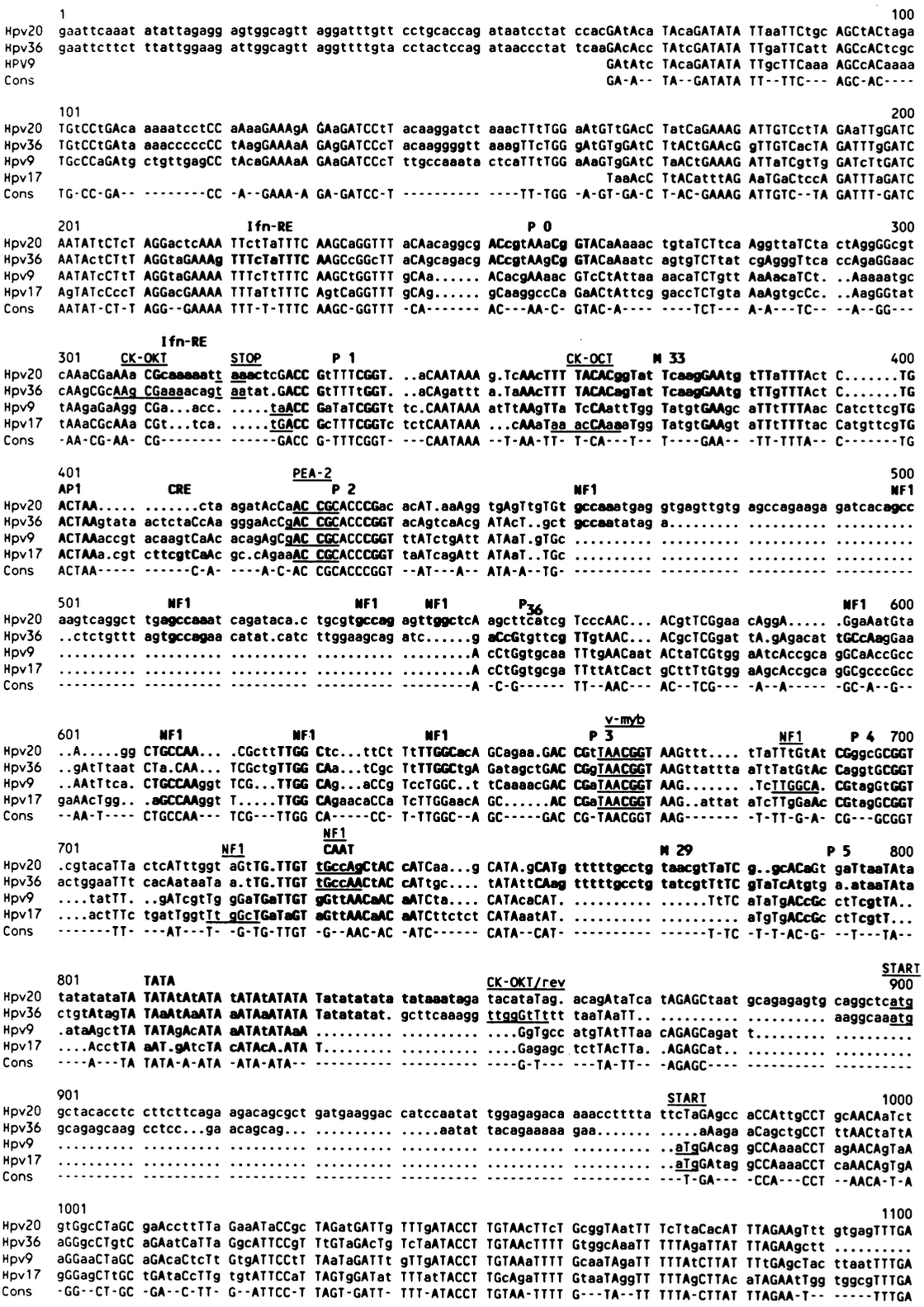

Fig.2. Alignment of the newly determined sequences of HPVs 20 and 36 and HPVs 9 and 17. Putative functionally important elements and protein binding sites are indicated in bold letters or underlined. Ifn-RE=Interferon responsive element. CK-OKT=Cytokeratin-octamer (19). P0−5,P_{36}=E2-protein binding sites. STOP=Translation stop codon of ORF L1. CRE=cAMP responsive element. PEA-2=Polyomavirus enhancer element. v-myb= Binding site for v-myb protein. NF1=Nuclear Factor 1 binding site. M33,M29=conserved motifs of 33 and 29bp respectively. CAAT=putative CAAT-box. TATA=stretch of A and T nucleotides. START=Translation start codon of ORF E6.

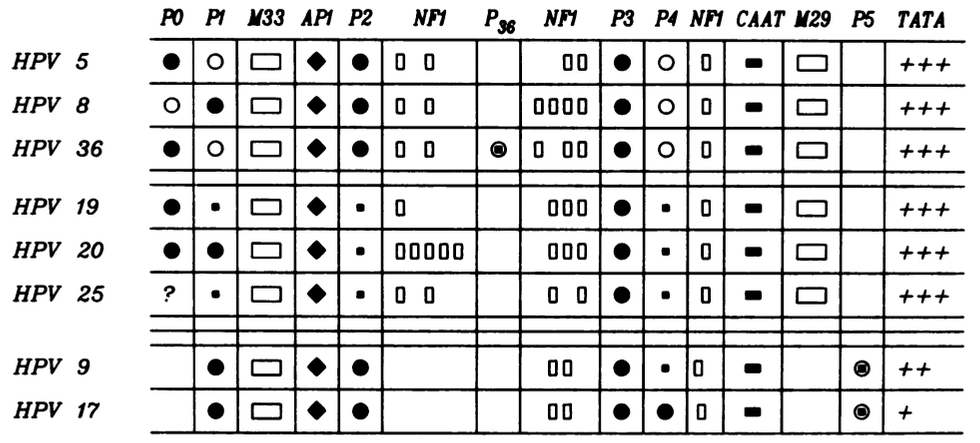

Fig.3. Schematic distribution pattern of conserved, putative regulatory elements in the LCR sequences of ev-associated HPV 5 (3), HPV 8 (4), HPVs 19 and 25 (5), and HPVs 9, 17, 20, 36 (this paper). P0–5, P_{36} = E2 protein binding sites: ● = $ACCGN_4CGGT$; ◉ = $ACCGN_4CGTT$; ○ = $ACCN_6GGT$; ▪ = degenerated palindrome. ◆ = AP1 binding site. □ = NF1 binding site. ▭ = M33, M29 conserved motifs of 33 and 29bp, respectively. ■ = CAAT box. TATA = TATA-stretch: +++ = ≈50bp; ++ = ≈20bp; + = TATA-like sequence in HPV 17. ? in HPV 25: DNA-sequence not available.

RESULTS AND DISCUSSION

The LCRs in the genomes of HPVs 9, 17, 20 and 36 were localized by Southern blot analysis using the homologous HPV 8 DNA as probe (Fig.1). The cross-hybridising fragments were subcloned into the Bluescribe vector (Vector Cloning Systems) and sequenced. The lengths of the LCRs from the translation stop codon of L1 to the start codon of E6 were 513 bp in HPV 20, 480 bp in HPV 36, 387 bp in HPV 9, and 379 bp in HPV 17. The aligned nucleotide sequences and the locations of conserved elements and putative recognition sites for DNA binding proteins (13) are shown in Fig.2.

A comparison with published LCR sequences of ev-associated HPVs confirmed the following, highly conserved landmarks (Fig.3): 1. Four palindromic E2 binding sites $ACCGN_4CGGT$ or slightly degenerated versions thereof (P1–4). Palindrome P3 is the most conserved among them and includes a putative binding site for the v-myb protein (14). 2. One binding site for the transcription factor jun/AP1 (13). 3. A cluster of direct or inverted repeats TGCCAA and related sequences, which were shown to bind the transcription factor NF1 (15). 4. A 'CAAT-motif', which is generally conserved among papillomaviruses (5).

HPVs 5, 8, 19, 20, 25, and 36 (group A) share the previously described conserved blocks with 33 and 29bp, respectively, an additional palindrome P0 within the 3'-part of ORF L1, and a striking stretch of about 50 A and T nucleotides. HPVs 5, 8, and 36, on the one hand, and HPVs 19, 20, and 25, on the other hand, form two slightly different subgroups A_1 and A_2. In A_2-viruses, palindromes P2 and P4 are incomplete, and 12bp are deleted between AP1 and P2; the AT rich stretch consists of monotonously alternating A and T, in contrast to subgroup A_1 viruses, which show two to four TAAA blocks followed by a shorter AT run. This AT stretch is without precedent among papillomaviruses, but similar motifs were identified within the autonomously replicating sequences (ARS) and promoter regions of yeast. Promoter activity was there shown to improve with increasing length of this element (16,17). Against the background of a rather uniform organization of group A virus LCRs, an apparently recent duplication can be noted in the case of HPV 20, which has a 21bp direct repeat with 3 mismatches between pos. 451 and 492 (Fig.2). The same duplication was observed in a second HPV 20 clone from another patient, which renders it unlikely to represent a cloning artefact. The sequence $ACCGN_4CGTT$, which was recently shown to bind E2 of bovine papillomavirus 1 (18), occurs once in HPV 36 (P_{36}).

In HPV 9 and 17 (group B) no M29 and only a slimed down M33 could be detected. The AT stretch is shortened in HPV 9 to about 20bp and interrupted by a few other nucleotides. No classical TATA-box was found in HPV 17 in this region, but TATA-like sequences TTAAATGA and TACATACAATAT are located at pos. −36 and −22 relative to the putative E6-protein translation start codon. The group B virus sequences thus indicate that the characteristic TATA stretch and the M29 motif discussed before are not consistent features of ev-associated HPVs, and can therefore not account for their host-specificity. Despite good general homologies with group A viruses in the 3'-part of ORF L1, group B viruses display no P0 in this region, but they contain an $ACCGN_4CGTT$ following the CAAT-box (P5).

No obvious correlation exists between the LCR classification into subgroups A_1, A_2, and B and the assumed oncogenic potential of ev-HPVs. This analysis provides another example, however, that LCRs of papillomaviruses, which belong to the generally less conserved genome segments, show a subgenous-specific organization pattern in viruses with similar tropism and pathogenic properties as observed for viruses associated with anogenital lesions.

ACKNOWLEDGEMENTS

We are indebted to Dr. G. Orth who provided cloned DNAs of HPVs 9, 17a, and 36. We gratefully acknowledge helpful discussions with Dr. P.G. Fuchs. This work was supported by the Deutsche Forschungsgemeinschaft (PF 123/3–2).

REFERENCES

1. Orth, G. (1987) in Salzman, N.P., and Howley, P.M. (ed.), The Papovaviridae, Vol.2, Plenum Publishing Corp., New York 199–243.
2. De Villiers, E.-M. (1989) J.Virol. 63, 4898–4903.
3. Zachow, K.R., Ostrow, R.S., and Faras, A.J. (1987) Virology, 158, 251–254.
4. Fuchs, P.G., Iftner, T., Weninger, J., and Pfister, H. (1986) J.Virol., 58, 626–634.
5. Krubke, J., Kraus, J., Delius, H., Chow, L., Broker, T., Iftner, T., and Pfister, H. (1987) J.Gen.Virol., 68, 3091–3103.
6. Iftner, T. (1990) in Pfister, H. (ed.), Papillomaviruses and human cancer, CRC Press, Boca Raton (Fl.,USA), 181–202.
7. Kremsdorf, D., Favre, M., Jablonska, S., Obalek, S., Rueda, L.A., Lutzner, M.A., Blanchet-Bardon, C., van Vorst Vader, P.C., and Orth, G. (1984) J.Virol., 43, 1013–1018.
8. Gassenmaier, A., Lammel, M., and Pfister, H. (1984) J.Virol., 52, 1019–1023.
9. Kawashima, M., Favre, M., Jablonska, S., Obalek, S., and Orth, G. (1986) Virology, 154, 389–394.
10. Kremsdorf, D., Jablonska, S., Favre, M., and Orth, G. (1982) J.Virol., 43, 436–447.
11. Maniatis, T., Frisch, E.F., and Sambrook, J. (1982) Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
12. Devereux, J., Haeberli, P., and Smithies, O. (1983) Nucleic Acids Res., 12, 387–395.
13. Jones, N.C., Rigby, P.W.J., and Ziff, E.B. (1988) Genes & Development, 2, 267–281.
14. Biedenkapp, H., Borgmeyer, U., Sippel, A.E., and Klempnauer, K.-H. (1988) Nature, 335, 835–837.
15. Gloss, B., Yeo-Gloss, M., Meisterernst, M., Rogge, L., Winnacker, E.L. and Bernard, H.-U. (1989) Nucleic Acids Res., 17, 3519–3533.
16. Stinchcomb, D.T., Mann, C.,Selker, E., and Davis, R.W. (1981) ICN-UCLA Symp. Mol. Cell. Biol. 22, 473–488.
17. Struhl, K. (1985) Proc.Natl.Acad.Sci.USA, 82, 8419–8423.
18. Li, R., Knight, J., Bream, G., Stenlund,A., and Botchan, M. (1989) Genes & Development, 3, 510–526.
19. Blessing, M., Zentgraf, H., and Jocarno, J.L. (1987) EMBO J. 6, 567–575.