The E2 Proteins

Alison McBride[a] **and Gerald Myers**[b]

[a] *Laboratory of Viral Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892*
[b] *Theoretical Biology and Biophysics, MS K710, Los Alamos National Laboratory, Los Alamos, NM 87545*

Revised from an earlier version published in August 1996

Introduction

The papillomavirus E2 proteins regulate viral transcription and replication and therefore play a central role in the viral life-cycle. In BPV1, the full-length E2 gene product is a transcriptional transactivator that activates transcription from several viral promoters by binding to E2 binding sites located within enhancer elements in the LCR (reviewed in [68,75]). BPV1 also encodes two shorter forms of the E2 protein that antagonize the function of the full-length transactivator. The E2 repressors inhibit transcription by binding to and blocking the E2 binding sites and/or by forming heterodimers with the E2 transactivator. The E2 proteins of the human papillomaviruses can also function as transcriptional transactivators and the E2 proteins of oncogenic human papillomaviruses are able to activate E2-responsive promoters more efficiently than those of non-oncogenic viruses [47]. However most, though not all, studies find that the full-length E2 proteins of mucosal type viruses repress the activity of the E6 gene promoter [9,12,29,45,82,100].

Viral DNA replication requires the full-length E2 transactivator, the viral E1 protein and the replication origin. The replication origin contains an E1 binding site flanked on either side by E2 binding sites. The E1 protein has several replication-associated activities such as origin-specific binding and helicase activities and forms a complex with the E2 transactivator. The E2 protein probably plays an auxiliary role in replication by enhancing and regulating the functions of the E1 protein.

Plasmids containing the minimal replication origin can replicate transiently in cells expressing the E1 and E2 proteins but with time the replicated DNA is lost. Long term, stable maintenance of such plasmids requires additional E2 binding sites and expression of the E1 and E2 proteins [78]. The E2 transactivator protein and BPV-1 viral genomes are associated with mitotic chromosomes in dividing cells [88]. These studies suggest that the E2 protein may play a role in plasmid copy number control and viral genome segregation.

The E2 proteins may also play a role in packaging the viral genomes in virion particles. Viral DNA appears to be packaged much more efficiently in the presence of the E2 protein, in both insect and mammalian cell lines [86,114].

Thus, the E2 proteins are multifunctional and important for several steps of the viral life cycle. Most of the knowledge about the structure and function of the E2 proteins has been obtained with the BPV1 E2 proteins and this review will concentrate on these proteins, unless otherwise stated. However, in general, the structure and functions of the E2 proteins seem to be comparable in all E2 proteins that have been examined to date. This review will summarize studies that have mapped functions to the various regions of the E2 proteins

A. E2 gene products

The E2 proteins have been best characterized for BPV1. Three BPV1 E2 proteins have been identified and mapped to the E2 ORF [43,48]. The largest 48kD protein, expressed from the entire ORF, is a transcriptional transactivator and is required for viral DNA replication [90,103]. This protein has been designated E2-TA. Two smaller proteins, encoded by the 3′ half of the ORF, function as transcriptional repressors [21,50]: E2-TR is a 30kD protein expressed from the P3080 promoter and initiated at an internal initiation codon at residue 162; E8/E2 is a 28kD protein encoded by a message with a 1234∧3225 splice that encodes 11 amino acids from the E8 ORF joined to the C-terminal 205 amino acids of E2. HPV cDNAs that are capable of encoding C-terminal regions of HPV E2 proteins have been identified [4,19,30] but, as yet, there is no direct genetic or biochemical evidence that the human papillomaviruses encode truncated E2 repressor proteins.

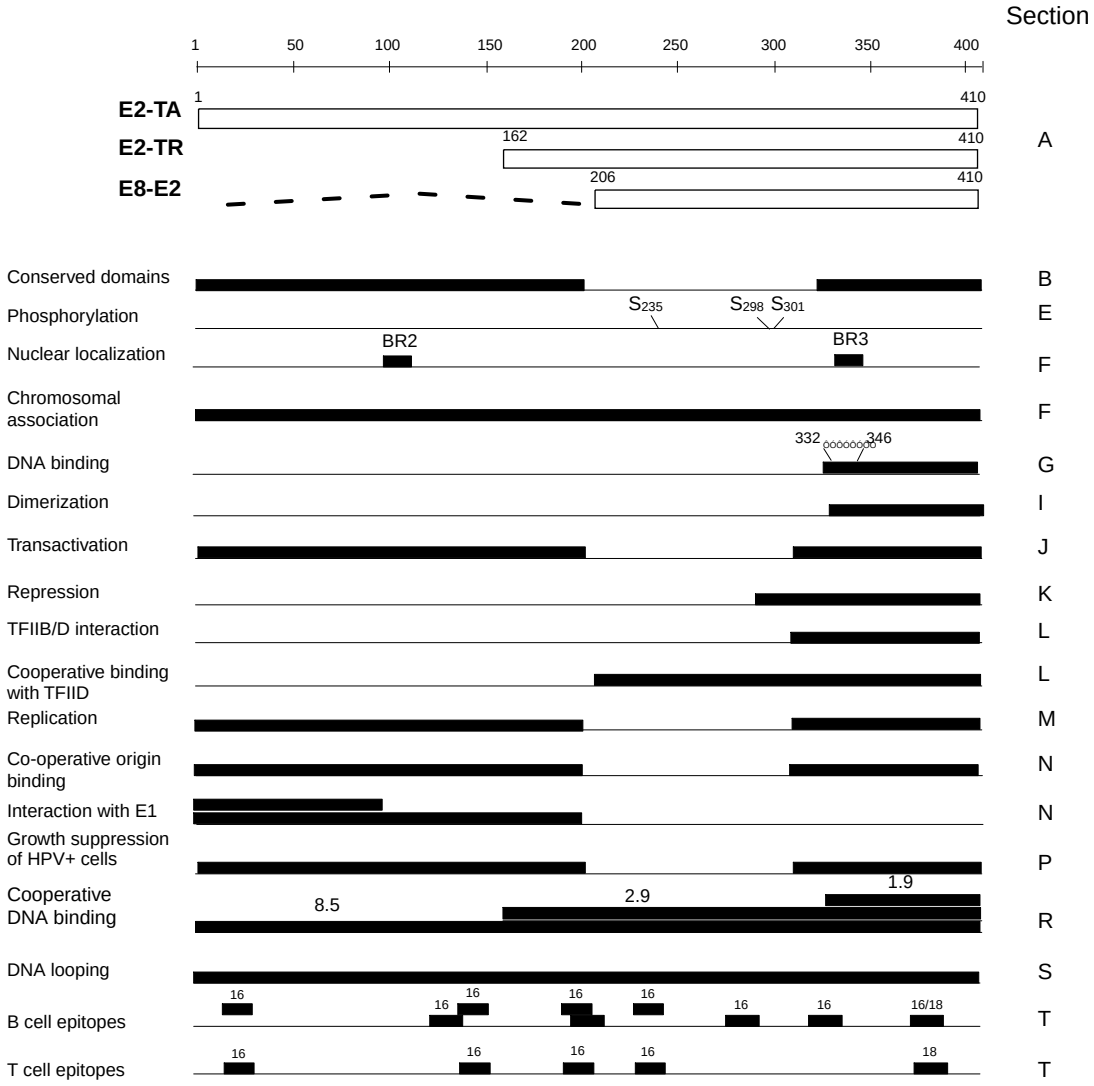


Figure 1. The structure of the three BPV1 E2 proteins are shown at the top of the figure. Below, the functions that have been mapped to different regions of the proteins are indicated. Refer to the section indicated to the right for more details. An alignment of E2 amino acid sequences is presented in Appendix A.

B. Conserved domains

Analysis of the predicted amino acid sequence of all papillomavirus E2 proteins shows that there are two conserved domains (see figure 1 and appendix A). An N-terminal domain of about 200 amino acids and a C-terminal domain of about 90 amino acids are separated by a non-conserved region of variable length that has been designated the hinge region. Notably, the hinge region overlaps the E4 open reading frame which is quite divergent among the papillomaviruses. The conserved E2 domains are approximately 35% similar among the papillomaviruses.

C. Protein structure

The E2-TA polypeptide consists of a C-terminal DNA binding domain linked to an N-terminal transactivation domain by a non-conserved hinge region. The E2 protein forms dimers that are mediated through the DNA binding domain and, as described below, the structure of the C-terminal domain has been solved.

The transactivation domain is approximately 200 amino acids and, unlike many other transactivation domains, appears to have a very constrained structure that is easily disrupted by deletion or certain non-conservative point mutations. The amino acid sequence of almost all of the papillomavirus E2 proteins is predicted to form two α-helices in the N-terminal region of the transactivation domain ([35] and Appendix B). However, as yet, there is no experimental evidence that such secondary structures exist in the transactivation domain.

The hinge region of the E2 proteins varies both in length and in amino acid composition among the E2 proteins. It has been postulated that this region forms a flexible link between the two domains and a study of the HPV16 E2 protein confirmed that the hinge is an unstructured region [34]. Antibodies were generated against overlapping peptides covering the entire E2 protein and it was found that only antibodies against the hinge region can recognized the native, undenatured E2 protein.

Figure 2. X-ray crystal structure of BPV-1 DNA binding domain (326–410) bound to DNA [40].

D. Protein turnover and Cell cycle Expression

The relative ratios of the three BPV1 E2 proteins have been measured in virally-transformed C127 cells as 1:10:3 for E2-TA/E2-TR/E8-E2 [43]. Within these cells E2-TA has a half-life of approximately 40 minutes and E2-TR and E8-E2 have half-lives of 10 and 15 minutes, respectively [43]. The ratio of the three BPV1 E2 proteins changes throughout the cell cycle with the ratio of E2 transactivator to repressors being highest at S phase and lowest at G1 [112].

E. Phosphorylation

BPV1, CRPV, HPV11 and HPV16 E2 proteins have been shown to be phosphorylated [5,13,64,69, 85] but the phosphorylation sites have only been identified in BPV1 E2 [52,64]. BPV1 E2 contains both phosphoserine and phosphothreonine [64]. Two major serine phosphorylation sites at positions

298 and 301 and a minor site at serine 235 have been mapped [52,64]. Mutation of E2 serine 301 to alanine results in a virus that replicates to a much greater copy number than wildtype BPV1 [66]. Viral genomes with an additional mutation at position 235 are defective in transformation and plasmid retention [52]. However, the region of the BPV1 E2 protein containing these phosphorylation sites is not conserved among the other papillomavirus E2 proteins.

F. E2 localization

All three BPV1 E2 proteins are located in the nucleus but a greater percentage of the full-length E2-TA protein is associated with insoluble chromatin and nuclear matrix components [43]. Two putative nuclear localization signals (NLS) have been identified in the BPV1 E2 proteins. A C-terminal peptide (BR3, residues 339–352, KCYRFRVKKNHRHR) which contains the DNA recognition helix of the DNA binding domain functions as a NLS both in the DNA binding domain and in heterologous proteins [87]. Deletion or mutation of a second signal in the transactivation domain (BR2, residues 107 to 115, KRCFKKGAR) results in a cytoplasmic E2 protein even though the C-terminal NLS is present. Therefore, it has been postulated that C-terminal NLSs may be masked in the E2-TA protein [87]. A recent study has shown that point mutations in the BR2 region (K111A, K112A) cause the protein to aggregate and be retained in the cytoplasm [1]. The E2-TA protein, but not the shorter repressor proteins, is found associated with mitotic chromosomes in dividing cells and this property may be important for viral genome segregation [88].

High levels of E2 expression are found primarily in the stratum spinosum of infected wart tissue which coincides with the region in which viral genome amplification occurs [17]. This may indicate that high levels of E2 are important for the switch to vegetative viral DNA replication.

G. DNA binding

The C-terminal domain of E2 (residues 326–410) binds specifically to DNA as a dimer (reviewed in [68]. The X-ray crystal structure of the C-terminal 85 amino acids of E2 bound to DNA was the first example of an anti-parallel β-barrel DNA binding structure [40]. As shown in figure 2, a dimer of the E2 DNA binding domain forms an eight-stranded anti-parallel β-barrel made up of four strands from each subunit. A pair of α-helices symmetrically positioned on the outside of the barrel contain the amino acids residues that are required for specific DNA interaction. Figure 3 shows the amino acid sequence of this recognition helix aligned with the homologous region from other papilloma virus E2 proteins. Also indicated on the figure are specific mutations that have been generated in this region of the E2 protein and their effect on DNA binding. Further studies have shown that sequences flanking the DNA binding domain (from the hinge region) contribute to DNA binding and are an integral part of the DNA binding domain [76]. The three dimensional structure of the HPV31 DNA binding domain in solution has also been determined by nuclear magnetic resonance [61]. The overall protein fold is very similar to the crystal structure of the BPV-1 domain but the DNA recognition helix appears to be flexible, as has been observed in a number of other DNA binding proteins. The DNA binding domain of the Epstein Barr virus EBNA1 protein has a very similar structure to the E2 DNA binding domain despite no sequence similarity [11].

The DNA recognition helix contains a highly conserved cysteine residue at position 340 that is very sensitive to oxidation [67]. This residue makes direct contacts with DNA yet E2 proteins with certain substitutions of this residue (glycine, serine, alanine) are still able to bind DNA. However, these proteins are not able to activate transcription efficiently in mammalian cells [37,67]. Similar reactive cysteines are found in the basic DNA-recognition regions of other proteins such as fos, jun and NFκB.

H. The E2 DNA Binding Site

The papillomavirus E2 proteins bind to DNA with high affinities which have been measured in the range of 4.5×10^{-9}M to 2×10^{-11}M [3,58,72,74,84,97].The consensus E2 recognition sequence

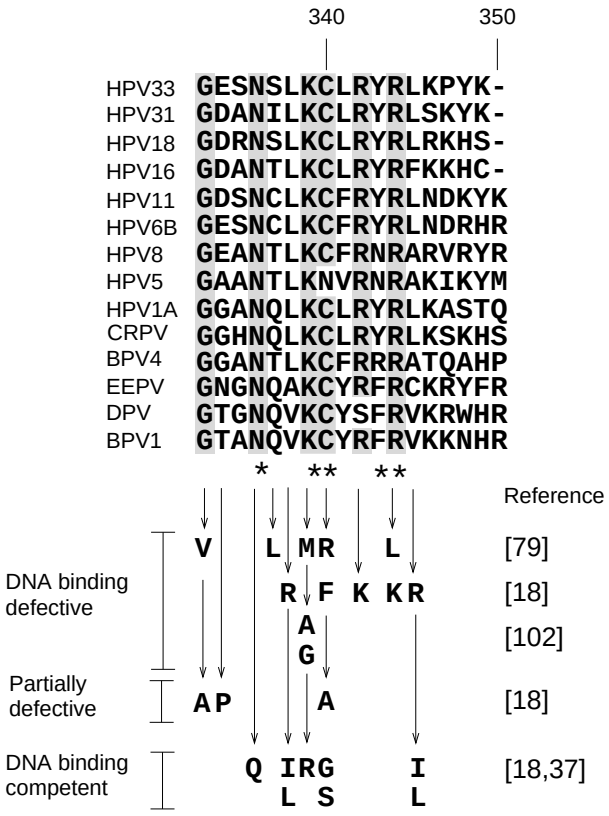


Figure 3 Alignment of the amino acid sequence of the DNA recognition helix of several papillomavirus E2 proteins. Residues in BPV1 E2 that directly contact DNA are indicated with an asterisk. Amino acid substitutions in this region of BPV1 E2 and their effect on DNA binding are shown below.

is ACCGN$_4$ CGGT [39] but the internal nucleotides and flanking sequences can greatly influence the affinity of E2 binding over several orders of magnitude [3,58,84,96,99]. There are 17 E2 binding sites in the BPV-1 genome and their affinities for the E2 proteins are over a 300-fold range (see Figure 4) [58]. Binding of the E2 protein to its consensus site induces a significant DNA bend [6,73,99] and E2 binding can be inhibited by CpG methylation of the ACCGN$_4$ CGGT motif [98].

In several mucosal type papillomaviruses, the full-length E2 protein appears to repress the promoter located upstream from the E6 gene. This probably occurs when the E2 proteins bind to E2 DNA binding sites that overlap binding sites for the cellular SP1 and TFIID transcription factors. In many cases it appears that these promoter proximal E2 binding sites have a lower affinity for the E2 protein than those sites located further upstream from the promoter start site [45,84,91]. This has led to a model in which low levels of E2 bind to the higher affinity upstream E2 sites and activate transcription, but at high levels of E2 protein the lower affinity proximal E2 sites are occupied leading to transcriptional repression. Figure 4 shows the relative affinity of the E2 binding sites in the non-coding region of several papillomaviruses.

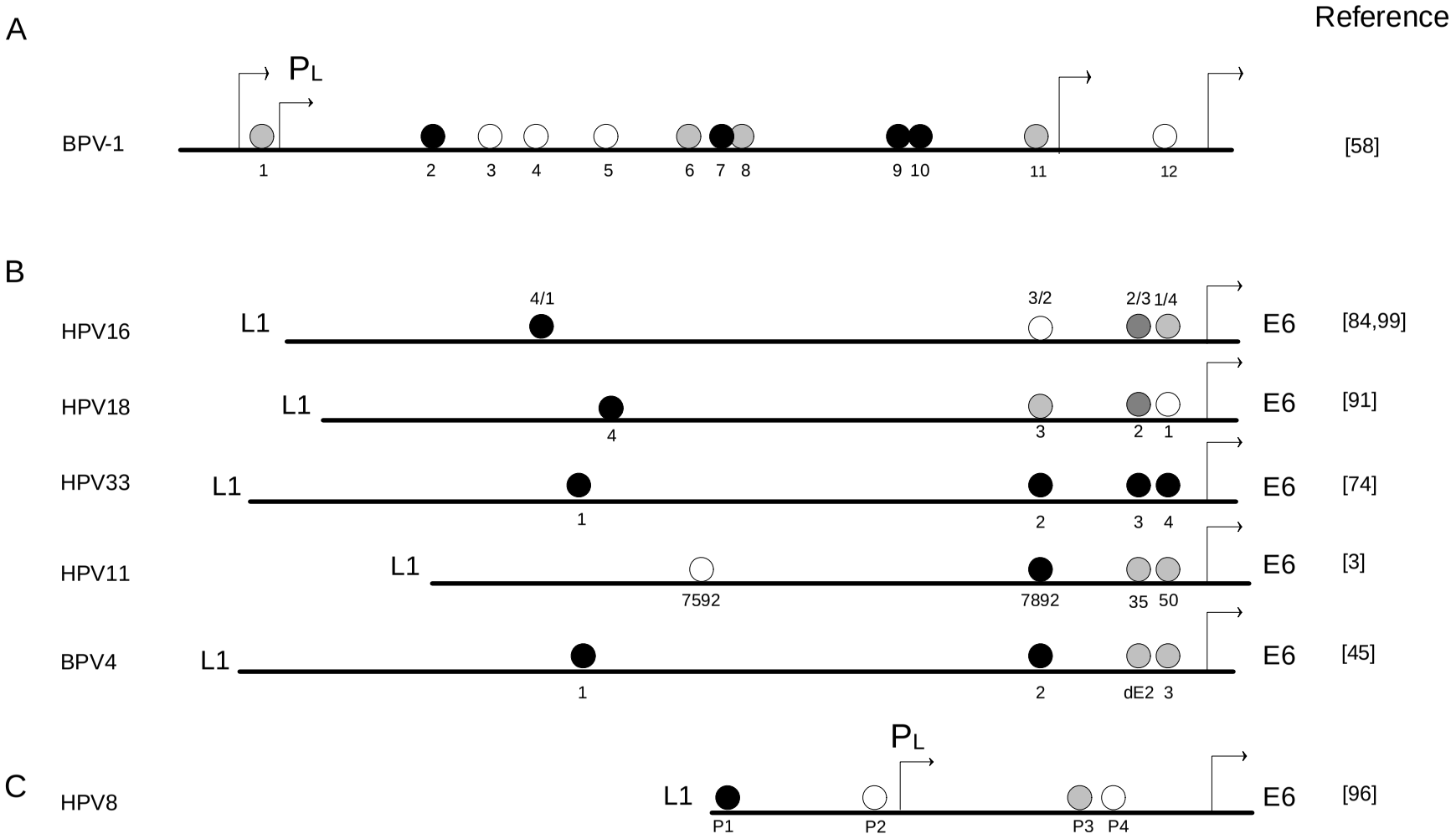


Figure 4. The location of E2 binding sites in the non-coding region or LCR of several papillomaviruses. The relative affinity or stability of E2 DNA-protein complexes is indicated as follows: open circle, low affinity or stability; gray circle, intermediate affinity or stability; black circle, high affinity or stability. The binding sites are labeled as described in the cited references. A. In BPV1, only the P_{7185} promoter is down-regulated by full-length E2 [93,105]. B. In the mucosal papillomaviruses shown, the major early promoter can be repressed by full-length E2 [29,45,82,101]. C. In the cutaneous papillomavirus, HPV8, only the late promoter is repressed by full-length E2 [96].

I. Dimerization

The DNA binding domain of the E2 proteins forms a stable dimer even in the absence of DNA [22,65]. The DNA binding and dimerization properties of this domain cannot be separated by deletion analysis; all deletions that have been tested eliminate both properties of the C-terminal domain. This is because dimerization involves an extensive subunit interface consisting of inter-backbone hydrogen bonds between the β-strands, interaction between side-chains of the β-strands in each subunit and an extensive hydrophobic core. This hydrophobic core contains a highly conserved tryptophan residue at position 360 which has been designated the tryptophan bridge ([22,79] see figure 2). The indole rings of W360 from each subunit are in Van der Waals contact which allows them to be crosslinked by UV irradiation [79]. Mutated E2 proteins containing hydrophobic residues at this position are functional but substitution of W360 by polar residues disrupts dimerization. Non-conservative mutations in other parts of the C-terminal domain can also eliminate both DNA binding and dimerization by disrupting protein structure [18]. Folding studies with the HPV16 E2 DNA binding domain have shown that the dimeric domain folds through a non-native monomeric intermediate [71].

J. Transactivation

When joined to a DNA binding domain, the N-terminal 194 amino acids of BPV-1 E2 are able to activate transcription from an E2-responsive promoter. The hinge region of the E2 proteins can be deleted with minimal effects on transactivation. Unlike many transactivation domains, the E2 N-terminal domain seems to have a very constrained structure as almost any deletion that has been made within this domain inactivates all E2-TA functions, presumably by disrupting protein conformation [65,108]. Even certain point mutations (e.g. BPV-1 P106G, G106A) may disrupt protein structure and therefore caution must be used to interpret mutational analyses of the E2 transactivation domain. It is likely, however, that mutations that eliminate one E2 function but not another do not extensively disrupt protein structure.

Recently several groups have undertaken systematic mutational analyses of the transactivation domains of the BPV-1, HPV16 and HPV11 E2 proteins. Several approaches were used and in most cases care was taken to try to avoid mutations that would disrupt protein structure. In one BPV-1 E2 study, conservative substitutions were generated, where possible, for each amino acid that is highly conserved among papillomavirus E2 proteins [16]. A second approach, used for both HPV16 and BPV-1 E2 proteins, was to change conserved charged residues to alanine residues as this can remove an essential side chain with minimal effects on protein structure [1,32,83]. A third study used a yeast screen to select for BPV-1 E2 mutants that were no longer able to activate transcription [15]. And finally, a study of the HPV11 E2 protein used a combination of the first and second approaches and generated both conservative substitutions and alanine residues for each conserved charged residue [107]. The latter study is quite informative as it shows that, in some cases, different substitutions of the same residue can give rise to dissimilar phenotypes. For example, HPV11 E2 proteins R7K and D96E are not defective for transactivation but proteins with alanine at these position have greatly reduced activity. Conversely, substitution of glutamic acid at position 39 with aspartic acid greatly reduces E2 activity but substitution with alanine has no effect. Table 1 shows an abridged summary of mutations in the E2 transactivation domain. This table lists only amino acid residues in which at least one mutation (among the different PV E2 proteins) has a significant effect on transactivation (10% or less of wild type activity). In the studies referenced many more mutations were analyzed and found to have more minimal effects on transactivation. Some proteins were also found to be relatively unstable and the original studies should be consulted for more details. Residues that are highly conserved among the papillomaviruses are shown in bold letters.

As can be seen in figure 1, no short linear sequence in the N-terminal domain seems to be important for transactivation and mutations that eliminate or greatly reduce this function appear to be scattered throughout the domain (with the caveat that some may disrupt overall domain structure). In the study of BPV E2 in which conservative changes were made in highly conserved residues, three mutations (W33F, E39D and K111R) were found to inactivate the transcriptional activation function. (Two other mutations that inactivated E2 function, P106G and G156A, were thought to disrupt protein structure.)

Table 1 Transactivation function of mutated E2 proteins

Residue	BPV1	HPV16	HPV11	Reference
R7	A, 35;		A, 10; K, 200;	[32, 107]
Q15	H, 3;			[15]
I30	A, 5;			[32]
L31	P, 5;			[38]
Y32	H, 3;			[38]
W33	F, 0; K, 0;	A, 30;		[16,32,83]
R37	K, 140; A, 25;	A, 0;	A, 5; K, 7;	[1, 16, 83, 107]
E39	D, 0; G, 3; A, 65; A, 70;	A, 110;	A, 150; D, 15	[1, 15, 16, 32, 83, 107]
P60	G, 10; A, 30;			[16, 32]
Q66	R, 0;			[38]
I73	L, 20;N, 0; A, 0;	A, 0;		[16, 32, 83]
E74	A, 0; A, 0;		A, 15; D, 5;	[1, 32, 107]
L82	A, 0;			[32]
F87	S, 3;			[15]
E90	A, 45; A, 50;	A, 95;	A, 0; D, 5;	[1, 32, 83, 107]
W92	F, 70; R, 2; A, 0; R, 0;	A, 5;		[15, 16, 32, 83]
S/T93	P, 5;	A, 80;		[38, 83]
D96			A, 5; E, 40;	[107]
W99	C, 3;			[15]
P106	G, 0; S, 0; A, 55;			[16, 32, 38]
K111	R, 0; A, 0;		A, 0; R, 0;	[1, 16, 107]
K112	R, 90; A, 0;A, 0;	A, 20;	R, 40;	[1, 16, 32, 83, 107]
F121	A, 10	A, 80;		[32, 83]
Y131	A, 5;			[32]
Y/W134		A, 10;		[83]
Y138	H, 9;	A, 45;		[15, 83]
W145	R, 1;			[15]
G156	A, 0;			[16]
Y159	F, 110;A, 0;			[16, 32]
156-159(GLYY)		dl, low;		[77]
Y169	A, 0;			[32]
E176/D174	A, 0; A, 50;G, 40;	A, 80;		[1, 32, 38, 83]
S181	F, 10;			[38]
181-182	PRSR insertion, ts;			[28]
R208	G, 10;			[38]

In each column the amino acid substitution is followed by its approximate activity expressed as a % of wild type activity. ts, temperature sensitive; dl, deletion. Highly conserved residues are shown in bold.

However, a W33A mutation in HPV16 E2 gave low but detectable activity suggesting that perhaps the phenylalanine side chain in the BPV E2 mutation interfered with transactivation. Similarly, while aspartic acid or glycine substitutions of residue 39 greatly reduced E2 activity in BPV and HPV11 E2 proteins, all three E2 proteins with alanine residues at this position were not defective. In agreement, however, BPV and HPV11 E2 proteins with K111A or K111R mutations were completely defective. The highly conserved arginine residue at position 37 has also been mutated in all three E2 proteins; R37K was partially defective in HPV11 E2 but not in BPV E2 and R37A resulted in low activity of

both BPV1 and HPV11 E2 proteins but no activity in HPV16 E2. The effect of other mutations of the E2 proteins can be seen in table 1. Clearly, more investigation is required to determine whether the differences observed in these studies are due to the different papillomavirus E2 proteins, the different amino acid substitutions or differences in assay conditions. Mutations throughout the transactivation domain were isolated using a yeast screening technique to isolate randomly generated transactivation-defective E2 proteins [15]. This study identified a pattern of bulky hydrophobic (BH) residues in two regions of the E2 protein (residues 26 to 47 and 87 to 107) similar to those that have been found to be important for transcriptional activation in VP16, GAL4 and RTA; mutation of several of these BH residues in E2 inactivated the transcriptional activation function [15]. Other notable mutations in the transactivation domain are a deletion of residues 156 to 159 which eliminates E2 activity but probably disrupts protein conformation; a G156A substitution is also thought to disrupt BPV E2 structure. A four amino acid insertion (PRSR) between residues 181 and 182 of the BPV E2 protein is temperature sensitive; E2 proteins containing this mutation are able to activate transcription at 32EC but not at 39EC [28].

K. Transcriptional Repression

The BPV1 E2-TR and E8/E2 repressors contain a small portion of the transactivation domain, the hinge region and the C-terminal DNA binding/dimerization domain (see figure 1). Transcriptional repression by these truncated E2 proteins is thought to be due to competitive DNA binding to the E2 binding sites in the viral enhancer elements and to heterodimer formation among transactivator and repressor species (reviewed in [68]. A C-terminal 121 amino acid polypeptide containing the DNA binding/dimerization domain is sufficient for repression of E2-mediated transactivation [63].

A different type of transcriptional repression can be mediated by the full-length transactivator proteins and this depends on the position of E2 binding sites with respect to proximal promoter elements. In many human papillomaviruses two E2 binding sites are positioned between a conserved SP1 site and the TATA box of the major E6/E7 promoter. Binding of E2 to these sites is thought to inhibit binding of these cellular factors resulting in repression of basal promoter activity [9,29,82,97,100]. A C-terminal domain of the E2 proteins is sufficient for this repression in transient assays [29,100].

L. Interaction with Cellular Proteins

The full-length BPV1 E2-TA protein has been shown to interact with the cellular replication protein RPA [56], a novel cellular protein AMF-1 [14] and the cellular transcription factor SP1 [59]. The C-terminal 127 amino acids of E2 or fusion proteins containing the C-terminal 160 amino acids of E2 are unable to interact with Sp1, implying that the N-terminal domain is required for this interaction [59].

The BPV1 E2 proteins interact with and cooperatively bind to DNA with the cellular basal transcription factors, TFIID and TFIIB [80,92]. The DNA binding domain of E2 (residues 310–410) is sufficient for protein-protein interaction with these factors [80]. However, the E2 hinge region is required in addition to the DNA binding domain (residues 204–410) for cooperative binding to DNA with TFIID or TFIIB [80,92].

M. Replication

Almost all studies of papillomavirus replication have examined E2 functions that are important for transient DNA replication and the plasmid maintenance function of the virus rather than vegetative viral DNA replication. However, using a cell culture system in which viral genomes are amplified in a subpopulation of cells, Alderborn et al. demonstrated that a temperature sensitive BPV1 E2 protein (see tables 1 and 2) that is defective for transactivation and plasmid maintenance replication at the non-permissive temperature was able to amplify large amounts of the viral genome in division arrested cells [2].

Most other studies have examined the ability of E2 to replicate plasmids in a transiently transfected cells. The full-length E2 protein is necessary for viral DNA replication [103]. The E2 protein binds cooperatively to the replication origin with the E1 protein [62,70,89,110], interacts with at least one cellular replication protein, RPA [56] and alleviates nucleosomal-mediated repression of replication [57]. The HPV E2 proteins seem to play a similar role in replication and, in fact, certain combinations of BPV1 and HPV E1 and E2 proteins are capable of initiating replication from various papillomavirus origins [20,23,36]. This indicates that the replication functions of these proteins are quite well conserved. Efficient replication by the BPV1 proteins in vivo requires the E2 protein. However, E2 is not necessary for replication of naked DNA templates in vitro, although it can enhance replication at low concentrations of the E1 protein [110,111]. In addition, it has been shown that the E1 protein of HPV1a is sufficient for replication in vivo [36]. Therefore, it appears that E2 plays an auxiliary role in replication and that E1 is the principal replication protein.

The transactivation domain of E2 is absolutely required for DNA replication [108]. In several studies, amino acid substitutions have been generated in the transactivation domains of the BPV1, HPV11 and HPV16 E2 proteins to determine which regions of this domain are important for the replication function(s) [1,15,16,32,38,83,107]. In the interest of space, only those mutated proteins that have 10% or less of wild type replication activity in at least one papillomavirus E2 protein are represented here. As shown in table 2, mutations throughout the domain affect the replication properties of the E2 protein but mutations in the same residue do not always give similar phenotypes in the different E2 proteins. Clearly, more investigation is required to determine whether the differences observed in these studies are due to the different papillomavirus E2 proteins, the different amino acid substitutions or differences in assay conditions. In many cases proteins defective for replication are also defective for transactivation but this is not always the case and separation of these properties are discussed in section N below. Two of the studies [16,83], also examined the ability of these mutated E2 proteins to interact with the E1 protein. In the BPV1 study, proteins that were defective for replication could still interact with the E1 protein (W33F, E39D, K111R). However, in the HPV16 study many of the replication-defective proteins with alanine substitutions could not interact efficiently with the E1 protein.

No particular amino acid sequence of the hinge region of the E2 protein is required for DNA replication. However, some nonspecific sequence is required between the two conserved domains to maintain the replication function. Two proteins with large deletions of the hinge region ($E2_{\Delta 220 \beta 309}$ and $E2_{\Delta 213 \beta 309}$) are unable to promote replication yet they can activate transcription more efficiently than the wildtype E2 protein [108]. It is possible that the two domains are too closely linked which sterically hinders one or more replication function.

In most cases, an intact E2 DNA binding domain is required for BPV1 DNA replication [104,108] and an E2 DNA binding site is required in the replication origin [102]. However, several E2 proteins have been identified that are defective in DNA binding but can support DNA replication to some extent. In one study, two out of ten E2 proteins with deletions in the DNA binding domain were able to support DNA replication at low levels [108]. In another study, DNA binding defective E2 proteins with point mutations in the DNA binding domain were able to promote replication in vitro [60]. BPV1 E2 proteins with mutations in the redox sensitive cysteine residue at position 340 in the DNA binding domain are able to support replication but are unable to activate transcription [37]. However, it is likely that the DNA binding and dimerization properties of the E2 protein are important for its replication function in the complete viral life cycle; the position of the E2 binding sites with respect to the E1 binding site is conserved among papillomavirus origins which indicates that they have an important function. It is possible that the mutated proteins have acquired some property that allows them to compensate for the absence of DNA binding, such as increased stability or increased interaction with the E1 protein. In addition, these experiments have been carried out using in vitro or transient assays in which the E1 and E2 proteins are most likely expressed at quite high levels and the replicon DNA is probably not assembled completely in chromatin and therefore all functions of the E2 proteins may not be required.

It is not clear whether the E2 repressors can repress transcription. Disruption of E2-TR expression in BPV1 results in a virus that replicates at much higher copy number than wild type [49,81] however, this effect could be indirect and due to lack of transcriptional repression. Notably, the E2 sites flanking the origin in BPV1 (sites 11 and 12) have a relatively weak affinity for the E2 protein.

Table 2 Replication function of mutated E2 proteins

Residue	BPV1	HPV16	HPV11	Reference
Q15	H, very low;			[38]
I30	A, 10;			[32]
W33	F, 5; K, 3;	A, 20;		[16, 32, 83]
E39	D, 5; G, 100; A, 125; A;15;	A, 10;	A, 15; D, 10;	[1, 16, 32, 38, 83, 107]
Q66	R, 5;			[38]
L82	A, 5;			[32]
W92	F, 45; R, 0; R, 5; A, 5;	A, 20;		[16, 32, 38, 83]
W99	C, 5;			[38]
P106	G, 0; S, 0; A, 100;			[16, 32, 38]
K111	R, 5; A, 0;		A, 5; R, 5;	[1, 16, 107]
K112	R, 60; A, 0; A, 5;	A, 65;	R, 100;	[1, 16, 32, 83, 107]
F121	A, 5;	A, 30;		[32, 83]
Y131	A, 10;			[32]
W145	R, 5;			[38]
G156	A, 0;			[16]
Y169	A, 0;			[32]
E/D176	G, 100; A, 0; A, 60;	A, 95; (D174)		[1, 32, 38, 83]
181-182	PRSR insertion, ts;			[28]

In each column the amino acid substitution is followed by its activity expressed as a % of wild type activity. ts, temperature sensitive. Highly conserved residues are shown in bold.

Binding of E2 to these sites is greatly increased in the presence of the E1 protein. This may be important for viral replication to ensure that only the E1-E2 complex, and not the E2-TR repressor protein, can bind to the origin region with high affinity.

N. Interaction and cooperative binding to the replication origin with the E1 protein.

The viral DNA replication origin contains an E1 binding site flanked by E2 binding sites [104]. The E1 and E2 proteins interact to form a protein complex and bind cooperatively to the origin of replication [10,62,70,89,110]. Formation of the replication preinitiation complex requires specific protein-protein and protein-DNA interactions between the E1 and E2 proteins and their respective DNA binding motifs. A biochemical study found that the entire transactivation domain and the DNA binding domain of BPV1 E2 are required for enhancement of E1 binding to the origin region (the hinge is not necessary). One study found that the entire transactivation domain is necessary and probably sufficient for interaction with E1 and the DNA binding domain is required for binding the complex to the origin DNA [109]. Another study found that the DNA binding domain of E2 also interacted with E1 when the E1 and E2 binding sites were adjacent to each other [8]. The E2-TR protein is unable to interact with or cooperatively bind to the origin with the E1 protein [70,109]. Other studies have shown that the N-terminal 190 amino acids of HPV16 E2 can interact with E1 [113] and that the N-terminal 140 amino acids of HPV16 E2 were unable to interact with E1 [94]. This suggests that the entire E2 transactivation domain is required for interaction with E1 in vitro. However, the first 91 amino acids of BPV1 E2 were able to interact with E1 in the yeast two hybrid system [7]. Two classes of mutations in the transactivation domain have been identified that interfere with the E1-E2 interaction. Mutations in the first class may identify regions of the E1 protein that are involved in protein-protein interaction because, although they are defective in E1 binding, they are transactivation competent [16,83]. These mutations

identify a region of E2 in the vicinity of residues 20, 33 and 39 and a second region containing residues 178 and 188 as being important for E1 interaction (see table 3). Notably, an antibody against HPV16 E2 residues 18 to 41 is able to block the E1-E2 interaction, supporting the fact that this region is important for complex formation [42]. The second class of E2 mutations that eliminate E1-E2 interaction are more difficult to interpret (see Table 3). These mutations are also defective in transcriptional activation and/or replication and the overall conformation of the N-terminal domain may be disrupted. Notably, these mutations are either deletions or mutations of or to prolines or glycines.

Table 3 E1 binding properties of mutated E2 proteins

Mutated E2 proteins that are defective in E1 interaction but functional in other E2 properties	Reference	Mutated E2 proteins that are defective both in E1 interaction and in other E2 properties	Reference
BPV1 E20D	[16]	HPV16 Δ23–26 HPV16	[77]
HPV16 W33A	[83]	L26P	[77]
HPV16 E39A	[83]	BPV1 P106G	[16]
HPV16 Y178A	[83]	BPV1 G106A	[16]
BPV1 V188L	[16]	HPV16 Δ156–159	[77, 94]

O. Separation of transactivation and replication properties

The E2-TA protein is important for transcriptional transactivation and DNA replication and several E2 mutants have been isolated that separate these properties. E2 proteins with deletions of the entire hinge region are able to activate transcription at wildtype levels yet are unable to support replication [108]. Conversely, a subset of proteins with deletions in the DNA binding domain of E2 are unable to activate transcription, yet can support DNA replication [108]. Other mutated E2 proteins with amino acid substitutions in the DNA binding domain (R344L, C340F) are also unable to activate transcription but can enhance in vitro replication [60]. Proteins with mutations in the redox-sensitive C340 are defective in transactivation in mammalian cells yet they can support viral DNA replication [37]. Several point mutations in the N-terminal domain also separate the transactivation and replication properties of E2 and these are listed in Table 4. Only mutations that are quite defective in one function (5% or less activity) and show reasonably high levels of activity in the other function are represented here. The cited studies should be referred to for additional mutated proteins that have are low in one activity and high in the others. This summary shows that the requirements for replication seem to be much less stringent than those for activation of transcription. Two regions appear to be important for transactivation but not replication; mutations in residues R37, I73 and/or E74 separate these functions in all three E2 proteins (BPV-1 E2 R37A also shows differential activity but as it retains 20% transactivation activity it is not shown in Table 4). Only two categories of mutants have been identified that cannot support DNA replication but can activate transcription. Proteins in one category have deletions of the entire hinge region and it has been postulated that this may cause some steric hindrance that interferes with one or more of the functions required for replication [108]. An HPV16 E2 protein with an E39A substitution is also defective for DNA replication and not transactivation but this is probably due to the inability of this protein to interact with the E1 protein [83]. (The same mutation in BPV-1 and HPV11 E2 proteins also reduces replication activity to approximately 15% but does not greatly affect transactivation.) These results probably reflect the fact that E2 plays a primary role in transcriptional activation but only an auxiliary role in DNA replication.

P. Growth suppression by the papillomavirus E2 proteins

The BPV1 E2-TA protein is able to suppress the growth of, and induce apoptosis in, HeLa cells, a line that is derived from an HPV-containing cervical carcinoma [101] [24,31,44]. HeLa cells are dependent on expression of the endogenous HPV18 E6 and E7 proteins for continued cell growth and

Table 4 Separation of transactivation and replication properties of the E2 proteins

E2 proteins that can support replication but not transactivation	Reference	E2 proteins that can activate transcription but not support DNA replication	Reference
BPV1		**BPV1**	
L31P, Y32H, E39G, F87S, S93P,	[38]	$E2_{\Delta 220-309}$, $E2_{\Delta 212-309}$	[108]
E74A	[1]	**HPV16**	
I73N, I73A, E74A,	[32]	E39A	[83]
$E2_{1-210}$, $E2_{1-376}$	[108]		
C340S, C340G	[37]		
HPV16			
R37A, I73A, W92A	[83]		
HPV11			
R37A, E74D, E90A, E90D, D96A	[107]		

Only those mutated proteins that are almost completely inactive in one assay (5% or less than wild type activity) are shown.

it is thought likely that E2 suppresses growth, at least in part, by repressing transcription of the E6/E7 P105 promoter. An intact transactivation and DNA binding domain are required for growth suppression even though the E2-TR repressor is able to repress expression from the P105 promoter in transient assays [31]. One explanation for this difference is that the E2 transactivation domain may be required to alleviate nucleosomal repression of the integrated HPV18 genome [31]. E2 may also inhibit cell growth by other mechanisms as one study has found that E2 expressed from a recombinant virus can also inhibit growth of an HPV-negative cervical carcinoma line [44]. This was not found when E2-containing HPV-negative cells were isolated by drug selection, perhaps because of the different E2 expression levels [31].

Overexpression of HPV31 E2 in human keratinocytes induces an S-phase arrest [33]. Cells undergo multiple rounds of DNA replication without undergoing mitosis. Clearly, this could be important for vegetative replication by allowing sustained synthesis of viral DNA.

Q. Alleviation of Nucleosomal Repression

The E2-TA protein can antagonize nucleosomal repression of BPV1 DNA replication in vitro [57]. This function depends on the presence of E2 DNA binding sites in the origin and therefore probably requires the E2 DNA binding domain. It is not known if the E2 transactivation domain is required for this function but in general the transactivation functions of cellular factors are required to relieve nucleosomal repression. Binding of the E2 protein in vivo under conditions in which E2 dimers activate transcription synergistically also results in a pronounced change in chromatin structure [51]

R. Cooperative DNA binding

BPV1 E2-TA binds cooperatively to two adjacent DNA binding sites with a cooperativity parameter of 8.5. The 86 amino acid DNA binding domain and the E2-TR protein exhibit much less cooperativity (factors of 1.9 and 2.9, respectively) which implies that the N-terminal domain of E2 is important for this function [72]. However, this cooperativity of DNA binding is not sufficient to explain the great synergy of transcription obtained with one versus two DNA binding sites.

S. Looping

The BPV1 E2-TA protein can form stable loops between widely spaced DNA sites that are visible by electron microscopy [46]. The shorter E2-TR and E8-E2 proteins are unable to form such loops implying that the transactivation domain is required for this function.

T. B and T cell epitopes

A number of linear B cell epitopes have been mapped in the HPV16 E2 protein. The E2 open reading frame (ORF) was synthesized as a set of overlapping 20-residue peptides which were tested for reactivity with HPV16-infected patient sera [25]. The E2 ORF is the most reactive of all the HPV16 E2 ORFs and four of the most reactive peptides (E2:9, residues 121–140; E2:13, residues 181–200; E2:17, residues 241–260; E2:19 residues 271–290) are shown in figure 1. In some instances, specific epitopes are recognized preferentially by either IgG, IgA or IgM antibodies. Figure 1 also shows another major IgG and IgA reactive epitope from HPV16 and HPV18 E2, designated p245 (residues 328–345) [26]. There is an association between serum antibodies against some of these epitopes and HPV-associated lesions and carcinomas [27,41,53,55,95,106]. A T-helper cell epitope overlaps the p245 B-cell epitope in HPV18 E2 but not in HPV16 E2 [54]. As summarized in figure 1, four T-helper cell epitopes have also been identified in HPV16 E2 (residues 11–25, 141–155, 191–205 and 231–245) and shown to overlap with additional IgG specific B-cell epitopes [27].

References

[1] Abroi, A., R. Kurg and M. Ustav. (1996) Transcriptional and replicational activation functions in the BPV1 E2 protein are encoded by different structural determinants. *J.Virol*. **70**:6169–6179.

[2] Alderborn, A., N. Jareborg and S. Burnett. (1992) Evidence that the transcriptional trans-activating function of the bovine papillomavirus type 1 E2 gene is not required for viral DNA amplification in division-arrested cells. *J.Gen.Virol*. **73**:2639–2651.

[3] Alexander, K.A. and W.C. Phelps. (1996) A fluorescence anisotropy study of DNA binding by HPV-11 E2C protein: a hierarchy of E2-binding sites. *Biochemistry* **35**:9864–9872.

[4] Baker, C. and C. Calef. 1995) Maps of papillomavirus mRNA transcripts, In G. Myers, H.-U. Bernard, C. Baker, A. Halpern, H. Delius, and J. Icenogel (eds.), *Human Papillomaviruses 1995*. Los Alamos National Laboratory, p. III-3–III-20

[5] Barbosa, M.S. and F.O. Wettstein. (1988) E2 of cottontail rabbit papillomavirus is a nuclear phosphoprotein translated from an mRNA encoding multiple open reading frames. *J.Virol*. **62**:3242–3249.

[6] Bedrosian, C.L. and D. Bastia. (1990) The DNA-binding domain of HPV-16 E2 protein interaction with the viral enhancer: Protein-induced DNA bending and role of the nonconserved core sequence in binding site affinity. *Virology* **174**:557–575.

[7] Benson, J.D. and P.M. Howley. (1995) Amino-terminal domains of the bovine papillomavirus type 1 E1 and E2 proteins participate in complex formation. *J.Virol*. **69**:4364–4372.

[8] Berg, M. and A. Stenlund. (1997) Functional interactions between papillomavirus E1 and E2 proteins. *J.Virol*. **71**:3853–3863.

[9] Bernard, B.A., C. Bailly, M.-C. Lenoir, M. Darmon, F. Thierry and M. Yaniv. (1989) The human papillomavirus type 18 (HPV18) E2 gene product is a repressor of the HPV18 regulatory region in human keratinocytes. *J.Virol*. **63**:4317–4324.

[10] Blitz, I.L. and L.A. Laimins. (1991) The 68-kilodalton E1 protein of bovine papillomavirus is a DNA binding phosphoprotein which associates with the E2 transcriptional activator in vitro. *J.Virol*. **65**:649–656.

[11] Bochkarev, A., J.A. Barwell, R.A. Pfuetzner, W.J. Furey, A.M. Edwards and L. Frappier. (1995) Crystal structure of the DNA-binding domain of the Epstein-Barr virus origin-binding protein EBNA 1. *Cell* **83**:39–46.

[12] Bouvard, V., A. Storey, D. Pim and L. Banks. (1995) Characterization of the human papillomavirus E2 protein: evidence of trans-activation and trans-repression in cervical keratinocytes. EMBO J. **13**:5451–5459.

[13] Bream, G.L., C.A. Ohmstede and W.C. Phelps. (1993) Characterization of human papillomavirus type 11 E1 and E2 proteins expressed in insect cells. *J.Virol*. **67**:2655–2663.

[14] Breiding, D., F. Sverdrup, M.J. Grossel, N. Moscufo, W. Boonchai and E.J. Androphy. (1997) Isolation of a BPV1 E2 transactivation domain binding factor required for both transcriptional activation and DNA replication. *Molecular and Cellular Biology*, in press.

[15] Breiding, D.E., M.J. Grossel and E.J. Androphy. (1996) Genetic analysis of the bovine papillomavirus E2 transcriptional activation domain. *Virology* **221**:34–43.

[16] Brokaw, J.L., M. Blanco and A.A. McBride. (1996) Amino acids critical for the functions of the BPV-1 E2 transactivator. *J.Virol*. **70**:23–29.

[17] Burnett, S., A.-C. Ström, N. Jareborg, A. Alderborn, J. Dillner, J. Moreno-Lopez, U. Pettersson and U. Kiessling. (1990) Induction of bovine papillomavirus E2 gene expression and early region transcription by cell growth arrest: Correlation with viral DNA amplification and evidence for differential promoter induction. *J.Virol*. **64**:5529–5541.

[18] Carruth, M. and A.A. McBride. (1997) unpublished data

[19] Chiang, C.-M., T.R. Broker and L.T. Chow. (1991) An E1M∧E2C fusion protein encoded by human papillomavirus type 11 is a sequence-specific transcription repressor. *J.Virol*. **65**:3317–3329.

[20] Chiang, C.M., M. Ustav, A. Stenlund, T.F. Ho, T.R. Broker and L.T. Chow. (1992) Viral E1 and E2 proteins support replication of homologous and heterologous papillomaviral origins. *Proc Natl.Acad.Sci.USA* **89**:5799–5803.

[21] Choe, J., P. Vaillancourt, A. Stenlund and M. Botchan. (1989) Bovine papillomavirus type 1 encodes two forms of a transcriptional repressor: Structural and functional analysis of new viral cDNAs. *J.Virol*. **63**:1743–1755.

[22] Corina, K., S.R. Grossman, J. Barsoum, S.S. Prakash, E.J. Androphy and R.B. Pepinsky. (1993) The tryptophan bridge is a critical feature of the papillomavirus E2 DNA binding domain. *Virology* 391–396.

[23] Del Vecchio, A.M., H. Romanczuk, P.M. Howley and C.C. Baker. (1992) Transient replication of human papillomavirus DNAs. *J.Virol*. **66**:5949–5958.

[24] Desaintes, C., C. Demeret, S. Goyat, M. Yaniv and F. Thierry. (1997) Expression of the papillomavirus E2 protein in HeLa cells leads to apoptosis. *EMBO J*. **16**:504–514.

[25] Dillner, J. (1990) Mapping of linear epitopes of human papillomavirus type 16: The E1, E2, E4, E5, E6 and E7 open reading frames. *Int.J.Cancer* **46**:703–711.

[26] Dillner, J., L. Dillner, J. Robb, J. Willems, I. Jones, W. Lancaster, R. Smith and R. Lerner. (1989) A synthetic peptide defines a serologic IgA response to a human papillomavirus-encoded nuclear antigen expressed in virus-carrying cervical neoplasia. *Proc.Natl.Acad.Sci.USA* **86**:3838–3841.

[27] Dillner, L., A. Zellbi, E. Avall-Lundqvist, P. Heino, C. Eklund, C.A. Pettersson, O. Forslund, B.G. Hansson, M. Grandien and P. Bistoletti. (1995) Association of serum antibodies against defined epitopes of human papillomavirus L1, E2, and E7 antigens and of HPV DNA with incident cervical cancer. Cancer Detect.Prev. **19**:381–393.

[28] DiMaio, D. and J. Settleman. (1988) Bovine papillomavirus mutant temperature sensitive for transformation, replication and transactivation. *EMBO J*. **7**:1197–1204.

[29] Dong, G., T.R. Broker and L.T. Chow. (1994) Human papillomavirus type 11 E2 proteins repress the homologous E6 promoter by interfering with the binding of host transcription factors to adjacent elements. *J.Virol*. **68**:1115–1127.

[30] Doorbar, J., A. Parton, K. Hartley, L. Banks, T. Crook, M. Stanley and L. Crawford. (1990) Detection of novel splicing patterns in a HPV16-containing keratinocyte cell line. *Virology* **178**:254–262.

[31] Dowhanick, J.J., A.A. McBride and P.M. Howley. (1995) Suppression of cellular proliferation by the papillomavirus E2 protein. *J.Virol*. **69**:7791–7799.

[32] Ferguson, M.F. and M.R. Botchan. (1996) Genetic analysis of the activation domain of bovine papillomavirus protein E2:its role in transcription and replication. *J.Virol*. **70**:4193–4199.

[33] Frattini, M.G., S.D. Hurst, H.B. Lim, S. Swaminathan and L.A. Laimins. (1997) Abrogation of a mitotic checkpoint by E2 proteins from oncogenic human papillomaviruses correlates with increased turnover of the p53 tumor suppressor protein. *EMBO J*. **16**:318–331.

[34] Gauthier, J.M., J. Dillner and M. Yaniv. (1991) Structural analysis of the human papillomavirus type 16–E2 transactivator with antipeptide antibodies reveals a high mobility region linking the transactivation and the DNA-binding domains. *Nucleic Acids.Res*. **19**:7073–7079.

[35] Giri, I. and M. Yaniv. (1988) Structural and mutational analysis of E2 trans-activating proteins of papillomaviruses reveals three distinct functional domains. *EMBO J*. **7**:2823–2829.

[36] Gopalakrishnan, V. and S.A. Khan. (1995) E1 protein of human papillomavirus type 1a is sufficient for initiation of viral DNA replication. *Proc.Natl.Acad.Sci.USA* **91**:9597–9601.

[37] Grossel, M.J., J. Barsoum, S.S. Prakash and E.J. Androphy. (1996) The BPV-1 E2 DNA-contact helix cysteine is required for transcriptional activation but not replication in mammalian cells. *Virology* **217**:301–310.

[38] Grossel, M.J., F. Sverdrup, D.E. Breiding and E.J. Androphy. 1996. Transcriptional activation function is not required for stimulation of DNA replication by BPV-1 E2. *J.Virol*. (In Press)

[39] Hawley-Nelson, P., E.J. Androphy, D.R. Lowy and J.T. Schiller. (1988) The specific DNA recognition sequence of the bovine papillomavirus E2 protein is an E2-dependent enhancer. *EMBO J*. **7**:525–531.

[40] Hegde, R.S., S.R. Grossman, L.A. Laimins and P.B. Sigler. (1992) Crystal structure at 1.7 A of the bovine papillomavirus-1 E2 DNA- binding domain bound to its DNA target. *Nature* **359**:505–512.

[41] Heino, P., C. Eklund, V. Fredriksson-Shanazarian, S. Goldman, J.T. Schiller and J. Dillner. (1995) Association of serum immunoglobulin G antibodies against human papillomavirus type 16 capsids with anal epidermoid carcinoma [see comments]. *J.Natl.Cancer Inst*. **87**:437–440.

[42] Hibma, M.J., K. Raj, S.Y. JEly, J. Stanley and L. Crawford. (1995) The interaction between human papillomavirus type 16 E1 and E2 proteins is blocked by an antibody to the N-terminal region of E2. *Eur.J.Biochem*. 229 :517–525.

[43] Hubbert, N.L., J.T. Schiller, D.R. Lowy and E.J. Androphy. (1988) Bovine papilloma virus-transformed cells contain multiple E2 proteins. *Proc.Natl.Acad.Sci.USA* 85 :5864–5868.

[44] Hwang, E.S., D.J.2. Riese, J. Settleman, L.A. Nilson, J. Honig, S. Flynn and D. DiMaio. (1993) Inhibition of cervical carcinoma cell line proliferation by the introduction of a bovine papillomavirus regulatory gene. *J.Virol*. **67**:3720–3729.

[45] Jackson, M.E. and M.S. Campo. (1995) Both viral E2 protein and the cellular factor PEBP2 regulate transcription via E2 consensus sites within the bovine papillomavirus type 4 long control region. *J.Virol*. **69**:6038–6046.

[46] Knight, J.D., R. Li and M. Botchan. (1991) The activation domain of the bovine papillomavirus E2 protein mediates association of DNA-bound dimers to form DNA loops. *Proc.Natl.Acad.Sci.USA* 88 :3204–3208.

[47] Kovelman, R., G.K. Bilter, E. Glezer, A.Y. Tsou and M.S. Barbosa. (1996) Enhanced transcriptional activation by E2 proteins from the oncogenic human papillomaviruses. *J.Virol*. **70**:7549–7560.

[48] Lambert, P.F., N.L. Hubbert, P.M. Howley and J.T. Schiller. (1989) Genetic assignment of multiple E2 gene products in bovine papillomavirus-transformed cells. *J.Virol*. **63**:3151–3154.

[49] Lambert, P.F., B.C. Monk and P.M. Howley. (1990) Phenotypic analysis of bovine papillomavirus type 1 E2 repressor mutants. *J.Virol*. **64**:950–956.

[50] Lambert, P.F., B.A. Spalholz and P.M. Howley. (1987) A transcriptional repressor encoded by BPV-1 shares a common carboxy-terminal domain with the E2 transactivator. *Cell* **50**:69–78.

[51] Lefebvre, O., G. Steger and M. Yaniv. (1997) Synergistic transcriptional-activation by the papillomavirus E2 protein occurs after DNA binding and correlates with a change in chromatin structure. *J.Mol.Biol*. **266**:465–478.

[52] Lehman, C.W., D.S. King and M.R. Botchan. (1997) A papillomavirus E2 phosphorylation mutant exhibits normal transient replication and transcription but is defective in transformation and plasmid retention. *J.Virol*. **71**:3652–3665.

[53] Lehtinen, M., A. Leminen, J. Paavonen, P. Lehtovirta, H. Hyoty, E. Vesterinen and J. Dillner. (1995) Predominance of serum antibodies to synthetic peptide stemming from HPV 18 open reading frame E2 in cervical adenocarcinoma. *J.Clin.Pathol*. **45**:494–497.

[54] Lehtinen, M., G. Stellato, H. Hyoty, P. Nieminen, E. Vesterinen and J. Paavonen. (1992) A T-helper cell epitope overlaps a major B-cell epitope in human papillomavirus type 18 E2 protein. *APMIS* **100**:1022–1026.

[55] Lenner, P., J. Dillner, F. Wiklund, G. Hallmans and U. Stendahl. (1995) Serum antibody responses against human papillomavirus in relation to tumor characteristics, response to treatment, and survival in carcinoma of the uterine cervix. Cancer Immunol.Immunother. **40**:201–205.

[56] Li, R. and M.R. Botchan. (1993) The acidic transcriptional activation domains of VP16 and p53 bind the cellular replication protein A and stimulate in vitro BPV-1 DNA replication. *Cell* **73**:1207–1221.

[57] Li, R. and M.R. Botchan. (1994) Acidic transcription factors alleviate nucleosome-mediated repression of DNA replication of bovine papillomavirus type 1. *Proc.Natl.Acad.Sci.USA* **91**:7051–7055.

[58] Li, R., J. Knight, G. Bream, A. Stenlund and M. Botchan. (1989) Specific recognition nucleotides and their DNA context determine the affinity of E2 protein for 17 binding sites in the BPV- 1 genome. *Genes Dev*. **3**:510–526.

[59] Li, R., J.D. Knight, S.P. Jackson, R. Tjian and M.R. Botchan. (1991) Direct interaction between Sp1 and the BPV enhancer E2 protein mediates synergistic activation of transcription. *Cell* **65**:493–505.

[60] Li, R., L. Yang , E. Fouts and M.R. Botchan. (1993) Site-specific DNA-binding proteins important for replication and transcription have multiple activities. *Cold Spring Harbor Symp.Quant.Biol*. **58**:403–413.

[61] Liang, H., A.M. Petros, R.P. Meadows, H.S. Yoon, D.A. Egan, K. Walter, T.F. Holzman, T. Robins and S.W. Fesik. (1996) Solution structure of the DNA-binding domain of a human papillomavirus E2 protein - evidence for flexible DNA-binding regions. *Biochemistry* **35**:2095–2103.

[62] Lusky, M., J. Hurwitz and Y.-S. Seo. (1993) Cooperative assembly of the bovine papilloma virus E1 and E2 proteins on the replication origin requires an intact E2 binding site. *J.Biol.Chem*. **268**:15795–15803.

[63] McBride, A.A. (1996) unpublished observations

[64] McBride, A.A., J.B. Bolen and P.M. Howley. (1989) Phosphorylation sites of the E2 transcriptional regulatory proteins of bovine papillomavirus type 1. *J.Virol*. **63**:5076–5085.

[65] McBride, A.A., J.C. Byrne and P.M. Howley. (1989) E2 polypeptides encoded by bovine papillomavirus type 1 form dimers through the common carboxyl-terminal domain: Transactivation is mediated by the conserved amino-terminal domain. *Proc.Natl.Acad.Sci.USA* **86**:510–514.

[66] McBride, A.A. and P.M. Howley. (1991) Bovine papillomavirus with a mutation in the E2 serine 301 phosphorylation site replicates at a high copy number. *J.Virol*. **65**:6528–6534.

[67] McBride, A.A., R.D. Klausner and P.M. Howley. (1992) Conserved cysteine residue in the DNA-binding domain of the bovine papillomavirus type 1 E2 protein confers redox regulation of the DNA-binding activity in vitro. *Proc Natl.Acad.Sci.USA* **89**:7531–7535.

[68] McBride, A.A., H.R. Romanczuk and P.M. Howley. (1991) The Papillomavirus E2 Regulatory Proteins. *J.Biol.Chem*. **266**:18411–18414.

[69] Meneguzzi, G., R. Lathe, M.P. Kieny and N. Vogt. (1989) The E2 trans-activating protein of bovine papillomavirus type 1 (BPV1) is serine-phosphorylated in vivo. *Oncogene* **4**:1285–1290.

[70] Mohr, I.J., R. Clark, S. Sun, E.J. Androphy, P. MacPherson and M.R. Botchan. (1990) Targeting the E1 replication protein to the papillomavirus origin of replication by complex formation with the E2 transactivator. *Science* **250**:1694–1699.

[71] Mok, Y.K., G.G. de Prat, P.J. Butler and M. Bycroft. (1996) Equilibrium dissociation and unfolding of the dimeric human papillomavirus strain-16 E2 DNA-binding domain. *Protein Sci*. **5**:310–319.

[72] Monini, P., S.R. Grossman, B. Pepinsky, E.J. Androphy and L.A. Laimins. (1991) Cooperative binding of the E2 protein of bovine papillomavirus to adjacent E2-responsive sequences. *J.Virol*. **65**:2124–2130.

[73] Moskaluk, C. and D. Bastia. (1988) DNA bending is induced in an enhancer by the DNA-binding domain of the bovine papillomavirus E2 protein. *Proc.Natl.Acad.Sci.USA* **85**:1826–1830.

[74] Muller, F., T. Giroglou and M. Sapp. (1997) Characterization of the DNA-binding activity of the E1 and E2 proteins and the E1/E2 complex of human papillomavirus type 33. *J.Gen.Virol*. **78**:911–915.

[75] O'Connor, M., S.-Y. Chan and H.-U. Bernard. 1995) Transcription factor binding sites in the long control region of the genital HPVs, In G. Myers, H. Bernard, C. Baker, A. Halpern, H. Delius, and J. Icenogel (eds.), *Human Papillomaviruses 1995*. Los Alamos National Laboratory, p. III-21–III-40

[76] Pepinsky, R.B., S.S. Prakash, K. Corina, M.J. Grossel, J. Barsoum and E.J. Androphy. (1997) Sequences flanking the core DNA-binding domain of bovine papillomavirus type 1 E2 contribute to DNA-binding function. *J.Virol*. **71**:828–831.

[77] Piccini, A., A. Storey, P. Massima, L. Banks, P. Massimi and V. Bouvard. (1995) Mutations in the human papillomavirus type 16 E2 protein identify multiple regions of the protein involved in binding to E1. *J.Gen.Virol*. **76**:2909–2913.

[78] Piirsoo, M., E. Ustav, T. Mandel, A. Stenlund and M. Ustav. (1996) Cis and trans requirements for stable episomal maintenance of the BPV-1 replicator. *EMBO J*. **15**:1–11.

[79] Prakash, S.S., S.R. Grossman, R.B. Pepinsky, L.A. Laimins and E.J. Androphy. (1992) Amino acids necessary for DNA contact and dimerization imply novel motifs in the papillomavirus E2 transactivator. *Genes Dev*. **6**:105–116.

[80] Rank, N.M. and P.F. Lambert. (1995) Bovine papillomavirus type 1 E2 transcriptional regulators directly bind two cellular transcription factors, TFIID and TFIIB. *J.Virol*. **69**:6323–6334.

[81] Riese, D.J., II, J. Settleman, K. Neary and D. DiMaio. (1990) Bovine papillomavirus E2 repressor mutant displays a high- copy-number phenotype and enhanced transforming activity. *J.Virol*. **64**:944–949.

[82] Romanczuk, H., F. Thierry and P.M. Howley. (1990) Mutational analysis of cis elements involved in E2 modulation of human papillomavirus type 16 P97 and type 18 P105 promoters. *J.Virol*. **64**:2849–2859.

[83] Sakai, H., T. Yasugi, J.D. Benson, J.J. Dowhanick and P.M. Howley. (1996) Targeted mutagenesis of the human papillomavirus type 16 E2 transactivation domain reveals separable transcriptional activation and DNA replication functions. *J.Virol*. **70**:1602–1611.

[84] Sanders, C.M. and N.J. Maitland. (1994) Kinetic and equilibrium binding studies of the human papillomavirus type-16 transcription regulatory protein E2 interacting with core enhancer elements. *Nucleic Acids.Res*. **22**:4890–4897.

[85] Sanders, C.M., P.L. Stern and N.J. Maitland. (1995) Characterization of human papillomavirus type 16 E2 protein and subdomains expressed in insect cells. *Virology* **211**:418–433.

[86] Schiller, J., personal communication

[87] Skiadopoulos, M.H. and A.A. McBride. (1996) The BPV-1 E2 transactivator and E2–TR repressor use different nuclear localization signals. *J.Virol*. **70**:1117–1124.

[88] Skiadopoulos, M.H. and A.A. McBride. Unpublished data.

[89] Spalholz, B.A., A.A. McBride, T. Sarafi and J. Quintero. (1993) Binding of bovine papillomavirus E1 to the origin is not sufficient for DNA replication. *Virology* **193**:201–212.

[90] Spalholz, B.A., Y.C. Yang and P.M. Howley. (1985) Transactivation of a bovine papilloma virus transcriptional regulatory element by the E2 gene product. *Cell* **42**:183–191.

[91] Steger, G. and S. Corbach. (1997) Dose-dependent regulation of the early promoter of human papillomavirus type 18 by the viral E2 protein. *J.Virol*. **71**:50–58.

[92] Steger, G., J. Ham, O. Lefebvre and M. Yaniv. (1995) The bovine papillomavirus 1 E2 protein contains two activation domains: one that interacts with TBP and another that functions after TBP binding. *EMBO J*. **14**:329–340.

[93] Stenlund, A. and M.R. Botchan. (1990) The E2 trans-activator can act as a repressor by interfering with a cellular transcription factor. *Genes Dev*. **4**:123–136.

[94] Storey, A., A. Piccini, P. Massima, V. Bouvard and L. Banks. (1995) Mutations in the human papillomavirus type 16 E2 protein identify a region of the protein involved in binding to E1 protein. *J.Gen.Virol*. **76**:819–826.

[95] Strickler, H.D., J. Dillner, M.H. Schiffman, C. Eklund, A.G. Glass, C. Greer, D.R. Scott, M.E. Sherman, R.J. Kurman and M. Manos. (1994) A seroepidemiologic study of HPV infection and incident cervical squamous intraepithelial lesions. *Virol.Immunol*. **7**:169–177.

[96] Stubenrauch, F. and H. Pfister. (1994) Low-affinity E2-binding site mediates downmodulation of E2 transactivation of the human papillomavirus type 8 late promoter. *J.Virol*. **68**:6959–6966.

[97] Tan, S.H., L.E. Leong, P.A. Walker and H.U. Bernard. (1994) The human papillomavirus type 16 E2 transcription factor binds with low cooperativity to two flanking sites and represses the E6 promoter through displacement of Sp1 and TFIID. *J.Virol*. **68**:6411–6420.

[98] Thain, A., O. Jenkins, A.R. Clarke and K. Gaston. (1996) CpG methylation directly inhibits binding of the human papillomavirus type 16 E2 protein to specific DNA sequences. *J.Virol*. **70**:7233–7235.

[99] Thain, A., K. Webster, D. Emery, A.R. Clarke and K. Gaston. (1997) DNA binding and bending by the human papillomavirus type 16 E2 protein. Recognition of an extended binding site. *J.Biol.Chem*. **272**:8236–8242.

[100] Thierry, F. and P.M. Howley. (1991) Functional analysis of E2-mediated repression of the HPV18 P105 promoter. New Biol. **3**:90–100.

[101] Thierry, F. and M. Yaniv. (1987) The BPV1–E2 trans-acting protein can be either an activator or a repressor of the HPV18 regulatory region. *EMBO J*. **6**:3391–3397.

[102] Ustav, E., M. Ustav, P. Szymanski and A. Stenlund. (1993) The bovine papillomavirus origin of replication requires a binding site for the E2 transcriptional activator. *Proc Natl.Acad.Sci.USA* **90**:898–902.

[103] Ustav, M. and A. Stenlund. (1991) Transient replication of BPV-1 requires two viral polypeptides encoded by the E1 and E2 open reading frames. *EMBO J*. **10**:449–457.

[104] Ustav, M., E. Ustav, P. Szymanski and A. Stenlund. (1991) Identification of the origin of replication of bovine papillomavirus and characterization of the viral origin recognition factor E1. *EMBO J*. **10**:4321–4329.

[105] Vande Pol, S.B. and P.M. Howley. (1990) A bovine papillomavirus constitutive enhancer is negatively regulated by the E2 repressor through competitive binding for a cellular factor. *J.Virol*. **64**:5420–5429.

[106] Wikstrom, A., C. Eklund, G. von Krogh, P. Lidbrink and J. Dillner. (1992) Levels of immunoglobulin G antibodies against defined epitopes of the L1 and L2 capsid proteins of human papillomavirus type 6 are elevated in men with a history of condylomata acuminata. *J.Clin.Microbiol*. **30**:1795–1800.

[107] Winokur, P., personal communication

[108] Winokur, P.L. and A.A. McBride. (1992) Separation of the transcriptional activation and replication functions of the bovine papillomavirus-1 E2 protein. *EMBO J*. **11**:4111–4118.

[109] Winokur, P.L. and A.A. McBride. (1996) The transactivation and DNA binding domains of the BPV-1 E2 protein have different roles in co-operative origin binding with the E1 protein. *Virology* **221**:44–53.

[110] Yang, L., R. Li, I.J. Mohr, R. Clark and M.R. Botchan. (1991) Activation of BPV-1 replication in vitro by the transcription factor E2. *Nature* **353**:628–632.

[111] Yang, L., I. Mohr, E. Fouts, D.A. Lim, M. Nohaile and M. Botchan. (1993) The E1 protein of bovine papillomavirus 1 is an ATP-dependent DNA helicase. *Proc.Natl.Acad.Sci.USA* **90**:5086–5090.

[112] Yang, L., I. Mohr, R. Li, T. Nottoli, S. Sun and M. Botchan. (1991) Transcription factor E2 regulates BPV-1 DNA replication in vitro by direct protein-protein interaction. *Cold Spring Harbor Symp.Quant.Biol*. **56**:335–346.

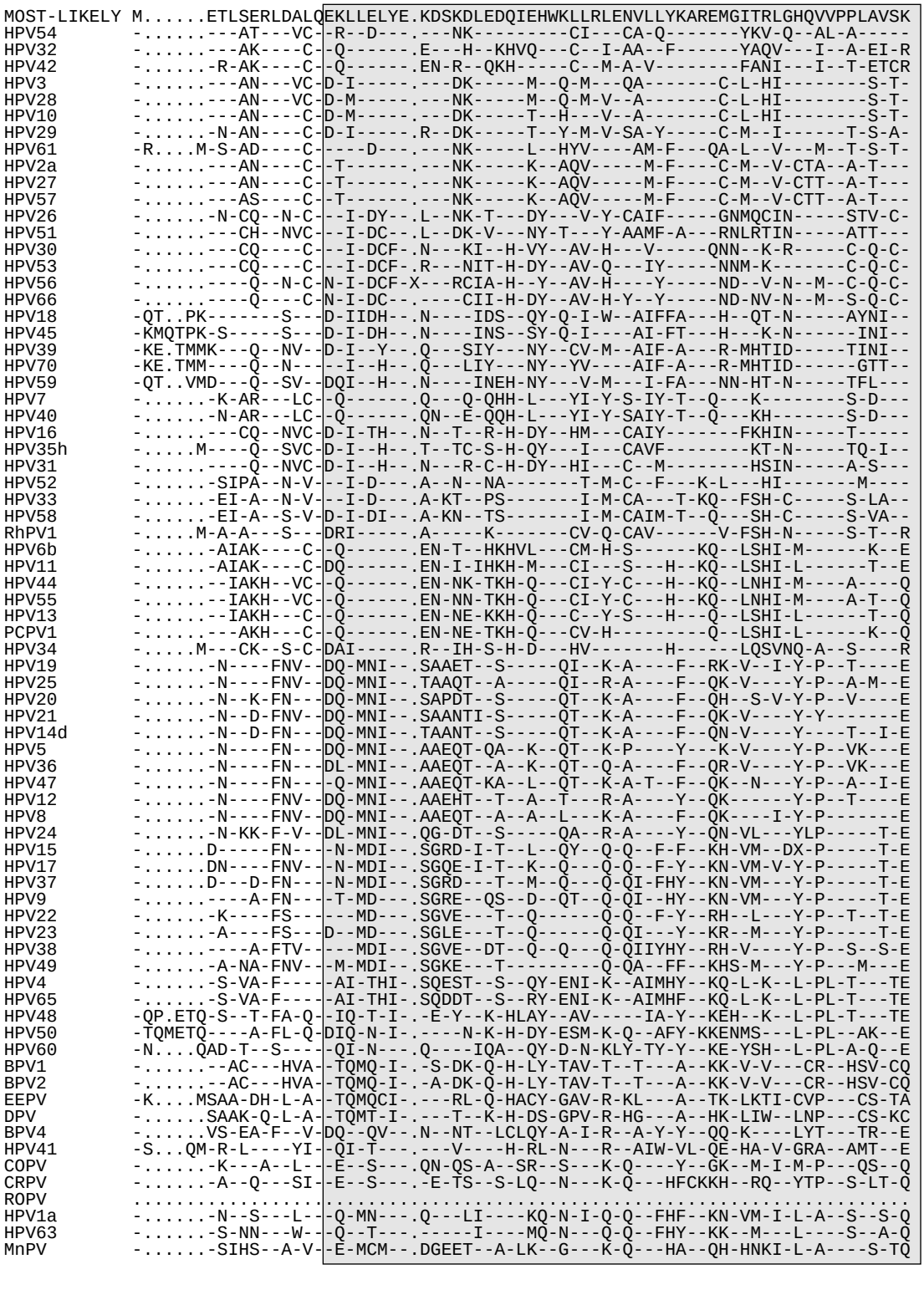
[113] Yasugi, T., J.D. Benson, H. Sakai, M. Vidal and P.M. Howley. (1997) Mapping and characterization of the interaction domains of human papillomavirus type 16 E1 and E2 proteins. *J.Virol*. **71**:891–899.

[114] Zhou, J., personal communication

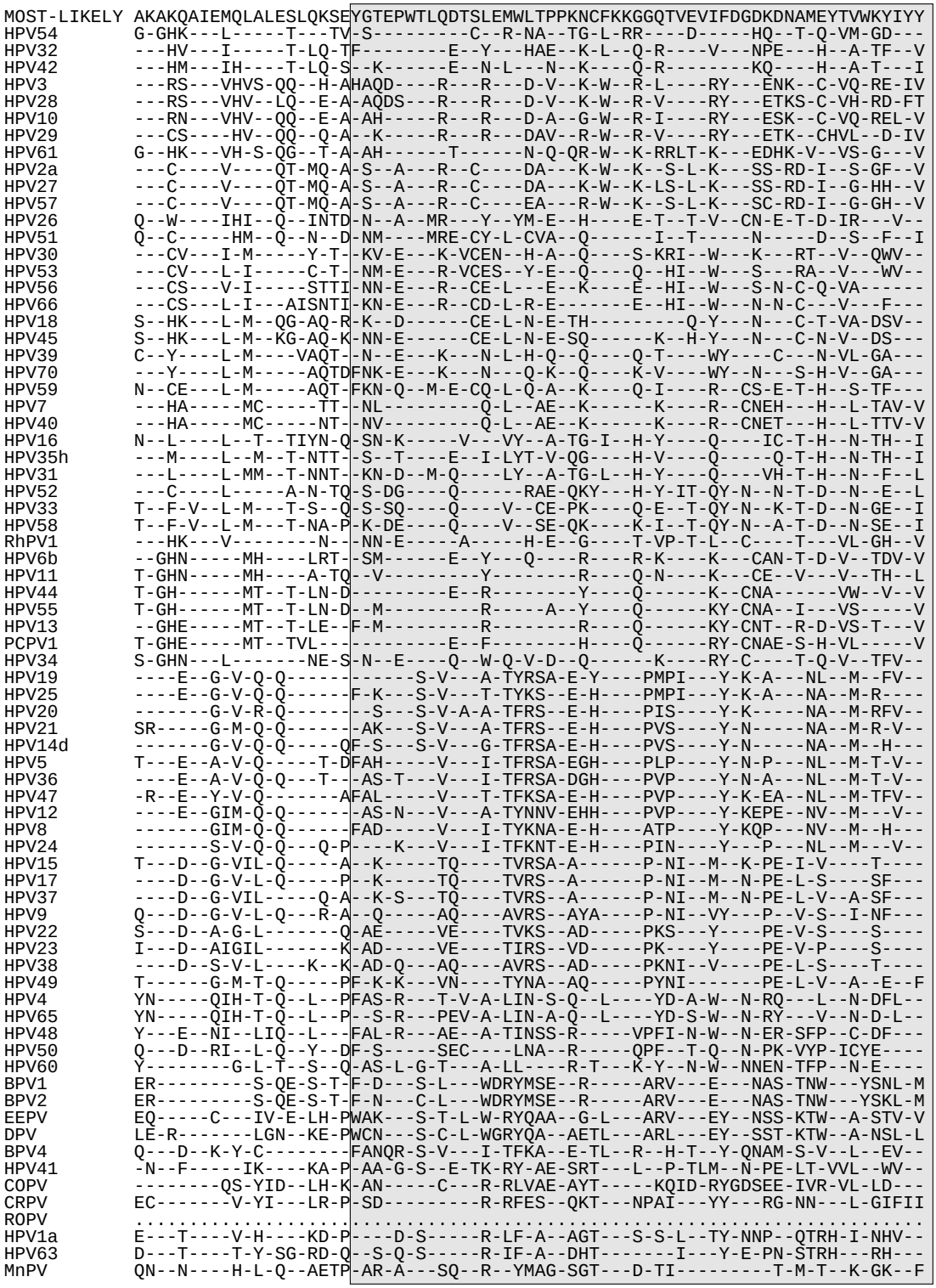
Appendix A: E2 Amino Acid Sequence Alignment

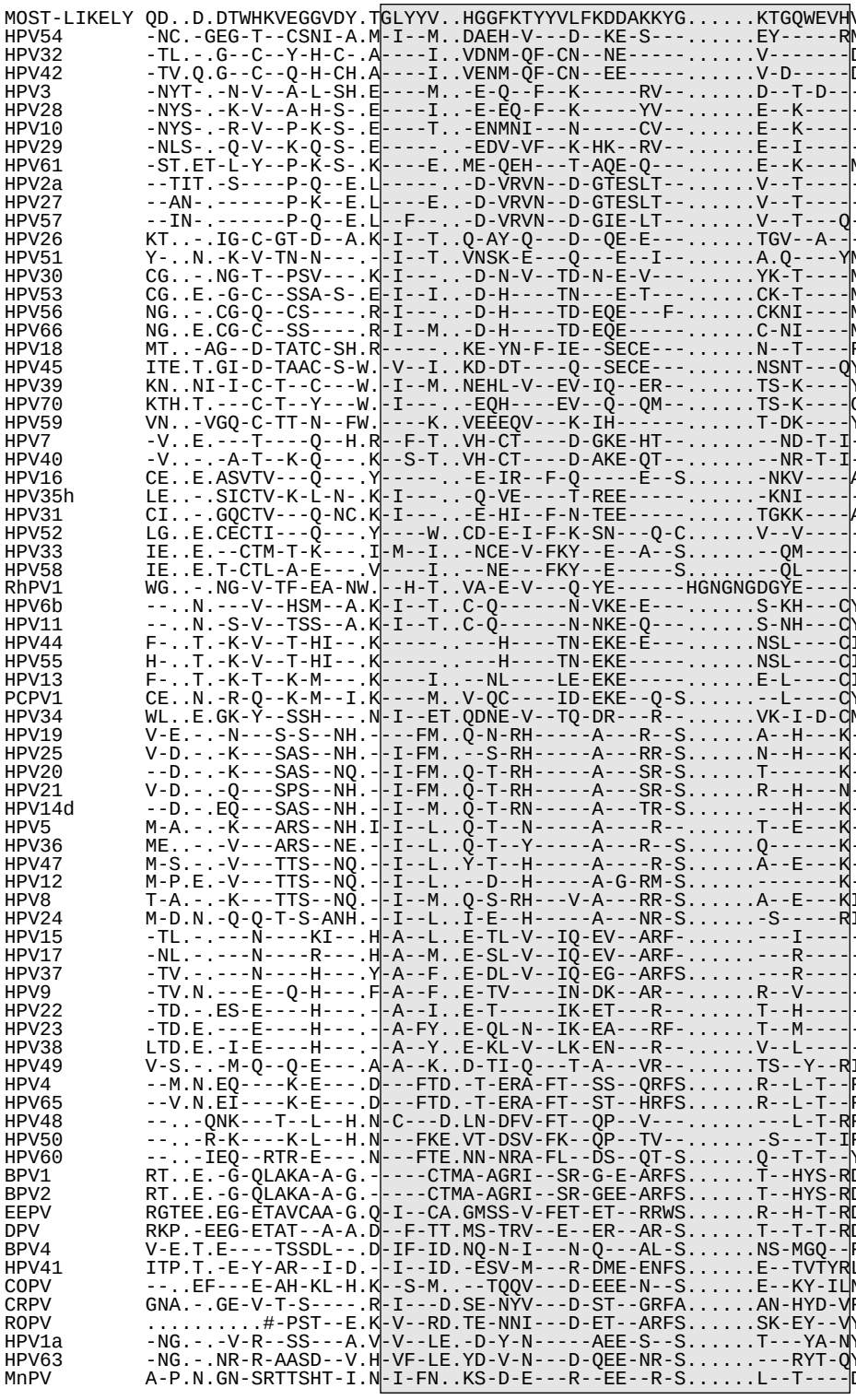
The following alignment of E2 amino acid sequences was generated using a Hidden Markov Method (HMM) of analysis, as described in Farmer and Myers, Papillomavirus Alignments and Structure Predictions, in *Human Papillomaviruses 1996*. The alignment differs from what is provided in Part II insofar as A and B supergroup sequences have been aligned together by the algorithm (whereas in Part II they are separately aligned). This alignment should be viewed as merely an hypothesis; it becomes the basis for the structural prediction output in Appendix B.

E2 BLOCKS: The MOTIF algorithm recognizes four BLOCKS in a complete E2 protein sequence alignment, of which the first is approximately confirmed by the Gibbs Sampler algorithm and the fourth is exactly confirmed. The latter BLOCK (residues 306 to 324 in the most likely sequence) is encoded by the DNA recognition helix, which is usually denoted the second conserved domain. BLOCKS 1–3 constitute the first conserved domain of E2. BLAST analysis of the MOTIF generated cobbler sequence (aka HPV13, shown below) gave strong scores and probabilities to other E2 sequences, as expected; among non-E2 sequences, no significant matches were found.


MOST-LIKELY M.....ETLSERLDALQEKLLELYE.KDSKDLEDQIEHWKLLRLENVLLYKAREMGITRLGHQVVPPLAVSK 66
HPV54 -.....---AT---VC--R--D---.---NK---------CI---CA-Q-------YKV-Q--AL-A----- 66
HPV32 -.....---AK----C--Q------.E---H--KHVQ---C--I-AA--F------YAQV---I--A-EI-R 66
HPV42 -.....-R-AK----C--Q------.EN-R--QKH-----C--M-A-V-------FANI---I--T-ETCR 66
HPV3 -.....---AN---VC-D-I-----.---DK-----M--Q-M---QA-------C-L-HI--------S-T- 66
HPV28 -.....---AN---VC-D-M-----.---NK-----M--Q-M-V--A-------C-L-HI--------S-T- 66
HPV10 -.....---AN----C-D-M-----.---DK-----T--H---V--A-------C-L-HI--------S-T- 66
HPV29 -.....-N-AN----C-D-I-----.R--DK-----T--Y-M-V-SA-Y-----C-M--I------T-S-A- 66
HPV61 -R....M-S-AD----C-----D---.---NK-----L--HYV----AM-F---QA-L--V---M--T-S-T- 68
HPV2a -.....---AN----C--T------.---NK-----K--AQV-----M-F----C-M--V-CTA--A-T--- 66
HPV27 -.....---AN----C--T------.---NK-----K--AQV-----M-F----C-M--V-CTT--A-T--- 66
HPV57 -.....---AS----C--T------.---NK-----K--AQV-----M-F----C-M--V-CTT--A-T--- 66
HPV26 -.....-N-CQ--N-C---I-DY--.L--NK-T---DY---V-Y-CAIF-----GNMQCIN-----STV-C- 66
HPV51 -.....---CH--NVC---I-DC--.L--DK-V---NY-T---Y-AAMF-A---RNLRTIN-----ATT--- 66
HPV30 -.....---CQ----C---I-DCF-.N---KI--H-VY--AV-H---V-----QNN--K-R-----C-Q-C- 66
HPV53 -.....---CQ----C---I-DCF-.R---NIT-H-DY--AV-Q---IY-----NNM-K-------C-Q-C- 66
HPV56 -.....----Q--N-C-N-I-DCF-X---RCIA-H--Y--AV-H----Y-----ND--V-N--M--C-Q-C- 67
HPV66 -.....----Q----C-N-I-DC--.----CII-H-DY--AV-H-Y--Y-----ND-NV-N--M--S-Q-C- 66
HPV18 -QT..PK-------S---D-IIDH--.N----IDS--QY-Q-I-W--AIFFA---H--QT-N-----AYNI-- 70
HPV45 -KMQTPK-S-----S---D-I-DH--.N----INS--SY-Q-I----AI-FT---H---K-N------INI-- 72
HPV39 -KE.TMMK---Q--NV--D-I--Y--.Q---SIY---NY--CV-M--AIF-A---R-MHTID-----TINI-- 71
HPV70 -KE.TMM----Q--N---I--H--.Q---LIY---NY--YV----AIF-A---R-MHTID------GTT-- 71
HPV59 -QT..VMD---Q--SV--DQI--H--.N----INEH-NY---V-M---I-FA---NN-HT-N-----TFL--- 70
HPV7 -.....-K-AR---LC--Q------.Q---Q-QHH-L---YI-Y-S-IY-T--Q---K--------S-D--- 66
HPV40 -.....-N-AR---LC--Q------.QN--E-QQH-L---YI-Y-SAIY-T--Q---KH-------S-D--- 66
HPV16 -.....---CQ--NVC-D-I-TH--.N--T--R-H-DY--HM---CAIY-------FKHIN-----T----- 66
HPV35h -.....M----Q--SVC-D-I--H--.T--TC-S-H-QY---I---CAVF--------KT-N-----TQ-I-- 67
HPV31 -.....----Q--NVC-D-I--H--.N---R-C-H-DY--HI---C--M--------HSIN-----A-S--- 66
HPV52 -.....-SIPA--N-V---I-D---.A--N--NA-------T-M-C--F---K-L---HI-------M---- 66
HPV33 -.....-EI-A--N-V---I-D---.A-KT--PS-------I-M-CA---T-KQ--FSH-C-----S-LA-- 66
HPV58 -.....-EI-A--S-V-D-I-DI--.A-KN--TS-------I-M-CAIM-T--Q---SH-C-----S-VA-- 66
RhPV1 -.....M-A-A---S---DRI-----.A-----K-------CV-Q-CAV------V-FSH-N-----S-T--R 67
HPV6b -.....-AIAK----C--Q------.EN-T--HKHVL---CM-H-S------KQ--LSHI-M------K--E 66
HPV11 -.....-AIAK----C-DQ------.EN-I-IHKH-M---CI---S---H--KQ--LSHI-L------T--E 66
HPV44 -.....--IAKH--VC--Q------.EN-NK-TKH-Q---CI-Y-C---H--KQ--LNHI-M----A----Q 66
HPV55 -.....--IAKH--VC--Q------.EN-NN-TKH-Q---CI-Y-C---H--KQ--LNHI-M----A-T--Q 66
HPV13 -.....--IAKH---C--Q------.EN-NE-KKH-Q---C--Y-S---H---Q--LSHI-L------T--Q 66
PCPV1 -.....---AKH---C--Q------.EN-NE-TKH-Q---CV-H---------Q--LSHI-L------K--Q 66
HPV34 -.....M---CK--S-C-DAI-----.R--IH-S-H-D---HV-------H------LQSVNQ-A--S----R 67
HPV19 -.....-N----FNV--DQ-MNI--.SAAET--S-----QI--K-A----F--RK-V--I-Y-P--T----E 66
HPV25 -.....-N----FNV--DQ-MNI--.TAAQT--A-----QI--R-A----F--QK-V----Y-P--A-M--E 66
HPV20 -.....-N--K-FN---DQ-MNI--.SAPDT--S-----QT--K-A----F--QH--S-V-Y-P--V----E 66
HPV21 -.....-N--D-FNV--DQ-MNI--.SAANTI-S-----QT--K-A----F--QK-V----Y-Y-------E 66
HPV14d -.....-N--D-FN---DQ-MNI--.TAANT--S-----QT--K-A----F--QN-V----Y----T--I-E 66
HPV5 -.....-N----FN---DQ-MNI--.AAEQT-QA--K--QT--K-P----Y---K-V----Y-P--VK---E 66
HPV36 -.....-N----FN---DL-MNI--.AAEQT--A--K--QT--Q-A----F--QR-V----Y-P--VK---E 66
HPV47 -.....-N----FN---Q-MNI--.AAEQT-KA--L--QT--K-A-T--F--QK--N---Y-P--A--I-E 66
HPV12 -.....-N----FNV--DQ-MNI--.AAEHT--T--A--T---R-A----Y--QK------Y-P--T----E 66
HPV8 -.....-N----FNV--DQ-MNI--.AAEQT--A--A--L---K-A----F--QK----I-Y-P------E 66
HPV24 -.....-N-KK-F-V--DL-MNI--.QG-DT--S-----QA--R-A----Y--QN-VL---YLP-----T-E 66
HPV15 -.....D-----FN---N-MDI--.SGRD-I-T--L--QY--Q-Q--F-F--KH-VM--DX-P-----T-E 66
HPV17 -.....DN----FNV--N-MDI--.SGQE-I-T--K--Q---Q-Q--F-Y--KN-VM-V-Y-P-----T-E 66
HPV37 -.....D---D-FN---N-MDI--.SGRD---T--M--Q---Q-QI-FHY--KN-VM---Y-P-----T-E 66
HPV9 -.....----A-FN---T-MD---.SGRE--QS--D--QT--Q-QI--HY--KN-VM---Y-P-----T-E 66
HPV22 -.....-K----FS---MD---.SGVE---T--Q------Q-Q--F-Y--RH--L---Y-P--T--T-E 66
HPV23 -.....-A----FS---D--MD---.SGLE---T--Q------Q-QI---Y--KR--M---Y-P-----T-E 66
HPV38 -.....----A-FTV--MDI--.SGVE--DT--Q--Q---Q-QIIYHY--RH-V----Y-P--S--S-E 66
HPV49 -.....-A-NA-FNV--M-MDI--.SGKE---T---------Q-QA--FF--KHS-M---Y-P---M---E 66
HPV4 -.....-S-VA-F----AI-THI-.SQEST--S--QY-ENI-K--AIMHY--KQ-L-K--L-PL-T---TE 66
HPV65 -.....-S-VA-F----AI-THI-.SQDDT--S--RY-ENI-K--AIMHF--KQ-L-K--L-PL-T---TE 66
HPV48 -QP.ETQ-S--T-FA-Q--IQ-T-I-.-E-Y--K-HLAY--AV-----IA-Y--KEH--K--L-PL-T---TE 71
HPV50 -TQMETQ----A-FL-Q-DIQ-N-I-.----N-K-H-DY-ESM-K-Q--AFY-KKENMS---L-PL--AK--E 72
HPV60 -N....QAD-T--S----QI-N---.Q----IQA--QY-D-N-KLY-TY-Y--KE-YSH--L-PL-A-Q--E 68
BPV1 -.....--AC---HVA--TQMQ-I-.-S-DK-Q-H-LY-TAV-T--T---A--KK-V-V---CR--HSV-CQ 66
BPV2 -.....--AC---HVA--TQMQ-I-.-A-DK-Q-H-LY-TAV-T--T---A--KK-V-V---CR--HSV-CQ 66
EEPV -K....MSAA-DH-L-A--TQMQCI-.---RL-Q-HACY-GAV-R-KL---A--TK-LKTI-CVP---CS-TA 68
DPV -.....SAAK-Q-L-A--TQMT-I-.---T--K-H-DS-GPV-R-HG---A--HK-LIW--LNP---CS-KC 66
BPV4 -.....VS-EA-F--V-DQ--QV--.N--NT--LCLQY-A-I-R--A-Y-Y--QQ-K----LYT---TR--E 66
HPV41 -S...QM-R-L----YI--QI-T---.---V----H-RL-N---R--AIW-VL-QE-HA-V-GRA--AMT--E 69
COPV -.....-K---A--L---E--S----.QN-QS-A--SR--S---K-Q----Y--GK--M-I-M-P---QS--Q 66
CRPV -.....-A--Q---SI--E--S---.-E-TS--S-LQ--N---K-Q---HFCKKH--RQ--YTP--S-LT-Q 66
ROPV 0
HPV1a -.....-N--S---L---Q-MN---.Q---LI----KQ-N-I-Q-Q--FHF--KN-VM-I-L-A--S--S-Q 66
HPV63 -.....-S-NN---W---Q--T---.-----I----MQ-N---Q-Q--FHY--KK--M---L----S--A-Q 66
MnPV -.....-SIHS--A-V--E-MCM--.DGEET--A-LK--G---K-Q---HA--QH-HNKI-L-A----S-TQ 66

E2 Appendix A





```
MOST-LIKELY QD..D.DTWHKVEGGVDY.TGLYYV..HGGFKTYYVLFKDDAKKYG......KTGQWEVHVGGEVI.......    193
HPV54       -NC.-GEG-T--CSNI-A.M-I--M..DAEH-V---D--KE-S---.....EY-----RM-SSI-.......    195
HPV32       -TL.-.G--C--Y-H-C-.A----I..VDNM-QF-CN--NE----.....V-------D-TQ--.......    194
HPV42       -TV.Q.G--C--Q-H-CH.A----I..VENM-QF-CN--EE----.....V-D-----D-NQ--.......    194
HPV3        -NYT-.-N-V--A-L-SH.E----M..-E-Q--F--K-----RV--.....D--T-D---K-.....HH.    197
HPV28       -NYS-.-K-V--A-H-S-.E----I..-E-EQ-F--K-----YV--.....E--K----K-.....HH.    197
HPV10       -NYS-.-R-V--P-K-S-.E----T..-ENMNI---N-----CV--.....E--K----K-.....HH.    197
HPV29       -NLS-.-Q-V--K-Q-S-.E-----..-EDV-VF--K-HK--RV--.....E--I----K-.....HH.    197
HPV61       -ST.ET-L-Y--P-K-S-.K----E..ME-QEH---T-AQE-Q---.....E--K----M-NT--Y..EPC.    201
HPV2a       --TIT.-S----P-Q--E.L-----..-D-VRVN--D-GTESLT--.....V--T----A-T-....HH.    197
HPV27       --AN-.------P-K--E.L----E..-D-VRVN--D-GTESLT--.....V--T----A-R-....HH.    197
HPV57       --IN-.------P-Q--E.L--F---..-D-VRVN--D-GIE-LT--.....V--T---Q---R-....YH.    197
HPV26       KT..-.IG-C-GT-D--A.K-I--T..Q-AY-Q---D--QE-E---.....TGV--A---C-Q-.......    193
HPV51       Y-..N.-K-V-TN-N---.--I--T..VNSK-E---Q---E--I---.....A.Q----YMY-T-.......    192
HPV30       CG..-.NG-T--PSV---.K-I---..-D-N-V--TD-N-E-V---.....YK-T----M-N-S-.......    193
HPV53       CG..E.-G-C--SSA-S-.E-I--I..-D-H----TN---E-T---.....CK-T----M-KQS-.......    193
HPV56       NG..-.CG-Q--CS----.R-I---..-D-H----TD-EQE---F-.....CKNI----MEN-S-.......    194
HPV66       NG..E.CG-C--SS----.R-I--M..-D-H----TD-EQE---------C-NI----MET-S-.......    193
HPV18       MT..-AG--D-TATC-SH.R------..KE-YN-F-IE--SECE---.....N--T----F-NN-....DCN    201
HPV45       ITE.T.GI-D-TAAC-S-W.-V--I..KD-DT----Q--SECE---.....NSNT---QY--N-....DCN    203
HPV39       KN..NI-I-C-T--C---W.-I--M..NEHL-V--EV-IQ--ER---.....TS-K----YN-NI-....HCP    202
HPV70       KTH.T.---C-T--Y---W.-I---..-EQH----EV--Q--QM---.....TS-K----CN-NI-....HCP    202
HPV59       VN..-VGQ-C-TT-N--FW.----K..VEEEQV---K-IH---------.T-DK----YN-K-....DCY    201
HPV7        -V..E.---T----Q--H.R--F-T..VH-CT----D-GKE-HT---------ND-T-I--SR-.......    193
HPV40       -V..-.-A-T--K-Q----.K--S-T..VH-CT----D-AKE-QT---------NR-T-I--SH-.......    193
HPV16       CE..E.ASVTV---Q---.Y------..-E-IR--F-Q-----E--S.....-NKV----A--Q-.......    193
HPV35h      LE..-.SICTV-K-L-N-.K-I---..-Q-VE----T-REE---------KNI----Q-.......    194
HPV31       CI..-.GQCTV---Q-NC.K-I---..-E-HI--F-N-TEE---------TGKK----A--Q-.......    193
HPV52       LG..E.CECTI---Q---.Y----W..CD-E-I-F-K-SN---Q-C.....V--V----Q-.......    193
HPV33       IE..E.--CTM-T-K---.I-M--I..-NCE-V-FKY--E--A--S.....--QM----Q-.......    193
HPV58       IE..E.T-CTL-A-E---.V----I..--NE---FKY--E-----S.....-QL----SR-.......    193
RhPV1       WG..-.NG-V-TF-EA-NW.--H-T..VA-E-V---Q-YE------HGNGNGDGYE----T-M.....H.    201
HPV6b       --..N.---V--HSM--A.K-I--T..C-Q------N-VKE-E---.....S-KH---CY-ST-.......    193
HPV11       --..N.-S-V--TSS--A.K-I--T..C-Q------N-NKE-Q---.....S-NH---CY-ST-.......    193
HPV44       F-..T.-K-V--T-HI---.K-----.---H----TN-EKE-E---.....NSL----CI-SSI-.......    193
HPV55       H-..T.-K-V--T-HI---.K-----.---H----TN-EKE---------NSL----CI-SS-.......    193
HPV13       F-..T.-K-T--K-M---.K----I..--NL----LE-EKE---------E-L----CI-ST-.......    193
PCPV1       CE..N.-R-Q--K-M--I.K----M..V-QC----ID-EKE--Q-S.....--L----CYDSK-.......    193
HPV34       WL..E.GK-Y--SSH---.N-I--ET.QDNE-V--TQ-DR---R-.....VK-I-D-CM--K-.......    195
HPV19       V-E.-.-N---S-S--NH.----FM..Q-N-RH-----A---R--S.....A--H---K-NK-TVF..TPVT    199
HPV25       V-D.-.-K---SAS--NH.--I-FM..--S-RH-----A---RR-S.....N--H---K-NKDTVF..TPVT    199
HPV20       --D.-.-K---SAS--NQ.--I-FM..Q-T-RH-----A---SR-S.....T------K-NK-TVF..APVT    199
HPV21       V-D.-.-Q---SPS--NH.--I-FM..Q-T-RH-----A---SR-S.....R--H---N-NK-TVF..APVT    199
HPV14d      --D.-.EQ---SAS--NH.--I--M..Q-T-RN-----A---TR-S.....---H---K-NK-TVF..TPVT    199
HPV5        M-A.-.-K---ARS--NH.I-I--L..Q-T--N-----A----R---.....T--E---K-NK-TVF..APVT    199
HPV36       ME..-.-V---ARS--NE.--I--L..Q-T--Y-----A---R--S.....Q------K-NK-TVF..APVT    198
HPV47       M-S.-.-V---TTS--NQ.--I--L..Y-T--H-----A----R-S.....A--E---K-NK-TVF..TPVT    199
HPV12       M-P.E.-V---TTS--NQ.--I--L..--D--H-----A-G-RM-S.....-------K-NK-TVF..APVT    199
HPV8        T-A.-.-K---TTS--NQ.--I--M..Q-S-RH---V-A---RR-S.....A--E---KINKDTVF..APVT    199
HPV24       M-D.N.-Q-Q-T-S-ANH.--I--L..I-E--H-----A---NR-S.....-S-----RINK-TVF..APVT    199
HPV15       -TL.-.---N----KI---.H-A--L..E-TL-V--IQ-EV--ARF-.....---I----NEDT-F..APVT    199
HPV17       -NL.-.---N----R---.H-A--M..E-SL-V--IQ-EV--ARF-.....---R----NEDT-F..APVT    199
HPV37       -TV.-.---N----H---.Y-A--F..E-DL-V--IQ-EG--ARFS.....---R----NKDT-F..APVT    199
HPV9        -TV.N.---E--Q-H---.F-A--F..E-TV----IN-DK--AR-.....R--V----NKDIVF..APVT    199
HPV22       -TD.-.ES-E----H---.--A--I..E-T-----IK-ET---R-.....T--H----NKDTVF..TPVT    199
HPV23       -TD.E.---E----H---.--A-FY..E-QL-N--IK-EA---RF-.....T--M----NKDTVF..TPVT    199
HPV38       LTD.E.-I-E----H---.--A--Y..E-KL-V--LK-EN---R-.....V--L----NKDTVF..TPVT    199
HPV49       V-S.-.-M-Q--Q-E---.A-A--K..D-TI-Q---T-A---VR-.....TS--Y--RINN-TVF..APVT    199
HPV4        --M.N.EQ----K-E---.D---FTD.-T-ERA-FT--SS--QRFS.....R--L-T--FKTQ-.......    195
HPV65       --V.N.EI----K-E---.D---FTD.-T-ERA-FT--ST--HRFS.....R--L-T--FKTQ-.......    195
HPV48       --..-QNK---T--L--H.N-C---D.LN-DFV-FT--QP--V---.....---L-T-RFKNKT-.......    200
HPV50       --..-R-K----K-L--H.N---FKE.VT-DSV-FK--QP--TV--.....-S---T-IFKNKT-.......    201
HPV60       --..-IEQ--RTR-E----.N---FTE.NN-NRA-FL--DS--QT-S.....Q--T-T--YKNQI-.......    197
BPV1        RT..E.-G-QLAKA-A-G.--I----CTMA-AGRI--SR-G-E-ARFS.....T--HYS-RDQDR-Y...AGV.    198
BPV2        RT..E.-G-QLAKA-A-G.--I----CTMA-AGRI--SR-GEE-ARFS.....T--HYS-RDQDR-Y...AGV.    198
EEPV        RGTEE.EG-ETAVCAA-G.Q-I--CA.GMSS-V-FET-ET--RRWS.....R--H-T-RDN.D----.YHST    201
DPV         RKP.-EEG-ETAT--A-A.D--F-TT.MS-TRV--E--ER--AR-S.....T--T-T-RDNDRTY....HS.    198
BPV4        V-E.T.E----TSSDL--.D-IF-ID.NQ-N-I---N-Q---AL-S.....NS-MGQ--FESK-L.......    195
HPV41       ITP.T.-E-Y-AR--I-D.--I--ID.-ESV-M---R-DME-ENFS.....E--TVTYRL-SALV.......    198
COPV        --..EF---E-AH-KL-H.K--S-M..--TQQV---D-EEE-N--S.....E--KY-ILNQPTT-....PTT    197
CRPV        GNA.-.GE-V-T-S----.R-I---D.SE-NYV---D-ST--GRFA.....AN-HYD-VFQNMRL.......    195
ROPV        .........#-PST--E.K-V--RD.TE-NNI---D-ET--ARFS.....SK-EY--VYKSQKL.......     47
HPV1a       -NG.-.-V-R--SS---A.V-V--LE.-D-Y-N-----AEE-S--S.....T---YA-NYR-KRF...TNVM    199
HPV63       -NG.-.NR-R-AASD--V.H-VF-LE.YD-V-N---D-QEE-NR-S.....---RYT-QYE-KRF...TNV.    198
MnPV        A-P.N.GN-SRTTSHT-I.N-I-FN..KS-D-E---R--EE--R-S.....L--T----D-L-THSLLIPVT    201
```

E2 Appendix A

```
MOST-LIKELY .SSPPSVS...........STGQAEV....STSE..TTSTLS.............................    217
HPV54       .F--A---.............--EE-LSI...-STG..-AEHTRP.............................    221
HPV32       .V--A-I-.............--TTT-AEV..-S-G..L-ELVQT............................    221
HPV42       .V--API-.............--T...S....TDA-..IP--G-T............................    216
HPV3        .D-FDP--.............--R..-I....PAPG..PLYA.C............................    218
HPV28       .HAFDP--.............--R..-I....PAAG..PLC-GD.............................    219
HPV10       .DAFDP--.............--R..-I....--PG..PVC-.-............................    218
HPV29       .NAFDP--.............--Q..-.....PATG..PLYASH............................    219
HPV61       .A-VS-TQ...........DAV.R-.....--A-..-AGH-RD.............................    225
HPV2a       .T-.A---.............--Q.-SA....-DD-..PL-PIRT...........................    220
HPV27       .T-.A---.............--Q.-TA....-DD-..PL-PIRL...........................    220
HPV57       .T-.A---.............--Q.-AT....-DDD..-L-P-RS...........................    220
HPV26       .CC-EF--.............--CSSNQI....--AK..-AEPV-...........................    218
HPV51       .TC-EY--.............--CSDAL.....--TT..-VEQ--...........................    216
HPV30       .YC-D---.............--LRSN-.....-PV-..-VVEYNT..........................    218
HPV53       .YC-D---.............--FRSN-.....-SV-..-VNEYYS..........................    218
HPV56       .YC-D---.............--CRYN-.....-PV-..-VNEYNT..........................    219
HPV66       .YC-D---.............--CRYN-....PPV-..-VNEYNN..........................    218
HPV18       .D-MC-T-...........DDT...-.....-ATQ..LVKQ-Q.............................    222
HPV45       .D-MC-T-...........DDT...-.....-ATQ..IVRQ-Q.............................    224
HPV39       .D-MC-T-...........DGS...-.....P-T-..L-TE--.............................    223
HPV70       .D-MY-T-...........DDT...-.....P-T-..L-AE-Q.............................    223
HPV59       .D-MC-T-...........DEQ...-.....--AG..SSEQ--Y............................    223
HPV7        .C-.--TV...........EGL.PI-....APVD..IRHPAA..............................    215
HPV40       .C--STIE...........EH-LPI-E...TADARPP-TATDT.............................    221
HPV16       .LC-T--F...........-SN..-.-....-SP-..IIRQ.HL............................    215
HPV35h      .VC-E--F...........-ST..-L.....-A-..IATQ-H..............................    216
HPV31       .VF-E--F...........-SD..-I....-FAG..IVTK-PT.............................    216
HPV52       .VC-A---...........-NE...-.....--T-..-AVH-C.............................    214
HPV33       .VC-T-I-...........-NQ...I.....--T-..-ADI.QT............................    214
HPV58       .VC-T-IP...........-DQ...I.....--T-..-ADP.K.............................    213
RhPV1       .Y-.D---...........-ATHCDKL...P-V-..IV-G-Q..............................    225
HPV6b       .C--A---...........--T.Q-.....-IP-..S-TYTP..............................    216
HPV11       .C--A---...........--V.R-.....-IA-..P-TYTP..............................    216
HPV44       .C--A-I-...........--V.QD-....-IAG..PA-HS-..............................    216
HPV55       .C--A-I-...........--V.QD-....-IAG..PA-HT-..............................    216
HPV13       .C--A---...........--V.Q-.....-IAG..PA-Y.-T.............................    216
PCPV1       .C--A---...........--V.Q-.....-IAG..P--H.-T.............................    216
HPV34       .CF.AP-F...........-PC...-.....--P-..IVRP-HT............................    217
HPV19       .--T-PD-...........PG--RDPN...TS-KTP--T-D-ASRLSPTASREQSQ................    240
HPV25       .--T-PE-...........PG---DSN...TS-KTP--D-A-..RLSPTGSGERSQ................    238
HPV20       .--T-PD-...........PG---DSN...AS-QTPA-T-D-TTRQSPRKQSQQ..................    238
HPV21       .--T-PD-...........PG---DSN...TS-TTPA-T-D-TSRLSSTRKQSQQ.................    239
HPV14d      .--T-PE-...........PG---DSN...TS-KTP--A-D-TSRLSPADSRKQSQ................    240
HPV5        .--T-PG-...........PG---DTN...T-PATP----TAVDSTSRQLTTSKQP................    240
HPV36       .--T-PG-...........PG---DTN...AS-KTS--T-.ATVDSTTKQLTTSEQP...............    239
HPV47       .--T-PG-...........PG--TDPD...TS-KTP--T-AATDTSPRRQSINKQSQQTETKRRGYGRRPSS    256
HPV12       .--T-PG-...........PGQRDPD....A--KTPA--SD-TTRSSDKQSQQA..................    237
HPV8        .--T-PG-...........PP---DTD...TAAKTP---AD-TSRQQRSPAKQPQ.................    239
HPV24       .--T-PD-...........PG-SR-L....PG-T..AN-KA-SPTQQPQQACSDETTKRK............    241
HPV15       .--.SPAA...........GE-ATSID...-AP-SPANRQ--STSVSSRKRTPP..................    237
HPV17       S---AAGE...........G-DASPIN...AA-RSSPARG--ATSVSTRTTQR...................    238
HPV37       S---AAGE...........G-DG-ASVH..TV-GSPLARGF-TTSVSTRKRT....................    238
HPV9        .--S-PTG...........DG-ETSKHT.L-R-GSP---R-PATTVPTGGSR...................    238
HPV22       .--T-P-G...........VAS...QN...-AP-..PA--SDSPQRSSQV.....................    229
HPV23       .--T-P-G...........DASNNA-......P-AS---LS-PQRSPSTN.....................    231
HPV38       .--T-P-G...........DST.DSA....-RAALPEP--SVSPERPPSQT.....................    233
HPV49       .--T-PSTGLRESSNASPVHD-V.D-T....P--TTA--T-F-TTTATATATGAPELSSKTGTR.........    257
HPV4        .--.-I--.............--YSSSF....D-E-QQLPGP.-T..........................    220
HPV65       .--.SV--.............--NTPSF....DFE-QQLPGP.-T..........................    220
HPV48       .-A.SVT-.............-SR..NTN...PS---.SRVG--T...........................    223
HPV50       .H-.SVT-.............-SRS-FG....PAD-QPGP--SY............................    226
HPV60       .-A.-VT-.............-SK-SSDD...Y--K..AGQQPH............................    221
BPV1        .--TS-DF...........RDRPDG-W...VA--GPEGDPAG..............................    225
BPV2        .--TS-DF...........RDRPDG-.....A--GPEGDPAG..............................    224
EEPV        .FGA-PH-...........RNDRDCIEGFW-DAG..ERRG.-..............................    228
DPV         .H-A--H-RETIE...GLWN-G-.R-R......GR.P-N-PDR.............................    229
BPV4        ...-SVT-.............-LR...-G....--GGQRG-Q-GDH..........................    219
HPV41       .NV-EP-T...........V-D..SS.....--R-..R-PKVL..............................    220
COPV        .-A.AGT-...........GPELPGH....-A-G..SGAC.-..............................    219
CRPV        .--.SVT-.............-PQ.PL-....-AP-.D-VPEEA.............................    218
ROPV        .-VSSVT-.............--PLRPIAL.GN-LDNA-A-PAP.............................     76
HPV1a       .--TS-PR...........AA-APA-H...-DYP..-L-ESDTAQQST........................    230
HPV63       .M-.-VN-...........-PLRTSGS...P-DTNPA-QGQ-T..............................    225
MnPV        .--T-QTG...........FPRGDP-R...LHGN..--TG-PIPLRNSSSNQILL.................    239
```

```
MOST-LIKELY .........................TTTTTATQA...RS...RGTSEP.TRSRPR..SRSRSRSR.....   248
HPV54       .....................ANS..-PR-DNS....TK...A...I-.CTPP-................   241
HPV32       ....................TDLYN--P-P-T....IT...-SNCD-.DGTDGILYKDPTPTTP.....   257
HPV42       ..................KLVQQVC.--NPLH-T....T-IDNHHADCT.DGTAY...NVPIQT-P.....   255
HPV3        ..........................--QAP--....AQ...V-A---.GPEQK....-Q-LETVYG...   246
HPV28       ..........................--KAS-E....T-...V-AT-..GPQQK....-Q-LETLN....   246
HPV10       ........................N--PAS--....AQ...V-A---.GPEQK....-Q-LEAV.....   245
HPV29       ........................N--RSP--....AP...L-PE-..GQE-K....-R-LEAVGPGPQ   251
HPV61       ........................N--Q-T-TP...TCV..GP-QTSTSVQT-.................   249
HPV2a       ......................AVSPVPAPV-AS-E.STG...A-RAA-P-QALC...-AQAPT-P.....   257
HPV27       ......................AVSPVPAPASAAS...AR...T--AP-.-NLLC...TAPAPP-P.....   254
HPV57       ......................AAAAV-A-ATAT....TA...V...P-.-LQDS....AQAPS-P.....   249
HPV26       .....................NA.--Q--EAY....VP...V--K-..-EAPY.................   240
HPV51       ........................N-P--NPLT...TC...V-AK-AQ-QQ-K.................   239
HPV30       .......................YN-YQ-P-TS...TP...V-AN-A.AS-AR.................   241
HPV53       ........................HK-P-TST.....P...V--Y-A.SS-LR.................   238
HPV56       ........................HK---T-S....T-...V-NQDA.AV-HR.................   240
HPV66       ........................HR---TAS....TF...V-AQDA.AV-HR.................   239
HPV18       ........................H-PSPYSST...V-...V--AKT.YGQTS.................   244
HPV45       ........................HAS-STPKT...A-...V--PK-.HIQT-.................   246
HPV39       ........................N--A-HST-...TT...PC-QK..-IPP-.................   244
HPV70       ........................H--PAH-A-...TT...PC-KK..-K-A-.................   244
HPV59       ......................PSA..-PPE--Y....LG...P...QTWN-QTK.............   244
HPV7        ....................-DA-D--KV...HD...APYAL-.ASTTK...VYND-HAP.....   245
HPV40       ......................PDA..PA-AT-E....TV...G...P..AQAP-.............   240
HPV16       ......................ANH..PAA-H-K-...VA...L--E-..-QTTI.............   238
HPV35h      ....................AYN--E-H-K-...C-...V--T-..-QKTN...............   239
HPV31       ......................ANN..---SNSKT...CA...L----GVR-ATT.............   241
HPV52       ......................-E-SK-S-...V-...V-AKD..-HLQ-.................   234
HPV33       ......................DND.NRPPQ-AAK...-R...-PADTT.DTAQ-.............   239
HPV58       ....................--EAT.....N...N...-S.-QGTK.....................   227
RhPV1       ......................HINPSPPP....AN...PSAK-..NVWSS.................   245
HPV6b       ....................AQ-S-L....V-...SS-K-D.AVQT-.....................   235
HPV11       ....................AQ---PT....V-...AC-T-D.GV-A-.....................   236
HPV44       ......................SSTT--LAQ-SST...LP...I--A-D.CVDA-.............   242
HPV55       ......................SSTT--LAQ-SP-...LP...TC---E.RVDP-.............   242
HPV13       ........................--S-Q-ST-...V-...CSA--E.CVQA-.............   238
PCPV1       ......................--LAQ--C-...V-...SIAT-D.SVQA-.............   238
HPV34       ....................SNS.....SNA-D...AG...V...P..--K-H.............   234
HPV19       ............QTNTKGRRYERRPSSR-RRQ-Q-R....QK...-SR-K..SK--S-.-----L-SN....   286
HPV25       ............QTSTKGRRYERRPSS.R-RRQQA-----.-QR..-SR-KSRS---SQ.-----I-------   286
HPV20       ..........TNTKGRRYGRRPSSR-RR--Q-RQ...-R...-SR-KSKSK--S-.-----H-------   285
HPV21       ..........TNTKGRRYGRRPSSR-RR--Q-HQ...-R...-SR-K..S---S-.-----L-------   284
HPV14d      ...........QANTKGRRYGRRPSSR-RR--E-RQ...-R...-SR-K..S---S-.-----L-------   286
HPV5        ..........QQTETRGRRYGRRPSS.KSRRSQ--....QR...-SR-R..H---S-.-----K-QTHTT.   289
HPV36       ..........QQTETKGRKYGRRPSS.R-RRPQAK....QR...-SR-RH.RS--S-.----Q---HT....   286
HPV47       RTRRPQTHQRRSRSRSRSRSSSQTHSS.-----T-Y....--...-S--LNK--A-S-.---T---T.....   314
HPV12       ............DPRRKGYGRRPSS.R-RRQE-----.QR...-SR-R..Y--QS....N-----Q.....   277
HPV8        ............QTETKGRRYGRRPSS.R-RPQKE-....R...-SR-R..H-T-S-.---L--V-AVGS.   286
HPV24       .......RYGRRESSPTDSRCRRRSSS.RQKKQGRR-...--...-TR-R..CS-TQ...T----T-.....   289
HPV15       ..............RTEARRYNRKESSP-----RR-K...-Q...GQRQ-..DTA-R.....T--G-QE...   282
HPV17       ..............TSPRRYRRKASS.P-A--TRHK...-QD.I-...R..S--TS-..G-QAISRG.....   279
HPV37       ..............PPRRYRRKASS.P---A-R-K...-Q...G...-..DTATR-.....T--GKQATSR   281
HPV9        ..............TSSRRYQRKASS.P--RKKR-....-QG..E-EG-GEGEETNY..R-Q----KGRTET   287
HPV22       ................THRYGRKASSP-I--IRR-K...-RE..-QRQ-TP--R-K...T------T.....   272
HPV23       ................RRYGRKASS.P-A--RR-K...-Q...G..K-TL--R-K...T----------.....   269
HPV38       ..............ARRYGRKASS.PS--SRR-....-K...GQRETTG-QR-RK..----------.....   273
HPV49       .....KGRYGRKDSSPTAASNSRKEVSRRRSRSR-RT...-R...-EA-T..S--QKA..----------.....   310
HPV4        ......................SYSEV-EQASPT....-R...-KPRKSDAT-TT...-PETEGV-.....   254
HPV65       ......................PTY..-EL-QASP...CG....-K-RE.SQPTSTT.-PET-GL-.....   254
HPV48       ......................SSS....SESPR.....-R...PSI--NSNTES-...TS-T--L.....   252
HPV50       ....................DKSQQERS....G-...GQPKALQDTEP-....T-T-TV-.....   255
HPV60       ..................FFASSSS.P---DGG-....T-...Q...-GVSS-TT...-P-AV-L.....   254
BPV1        ......................KEA.EPAQPVSSL...LG...SPACG-IRAGLGW..V-DGP--H.....   260
BPV2        ......................KEA.EPAQPVSSL...LG...SPACV-IRAGLGW..V-DGP-PH.....   259
EEPV        ................RGSDTTDRA.LPYPA-R-SPIC-P...VR-G-..N---AV..H-QAPY-.....   269
DPV         ......................AVL.H-PPGGNTVH..GP...VRAC-..N-G-SI..N-PTPY-.....   262
BPV4        ....................ARGRS.RPSPRSSRD...AR...GRQQRAQSS--S-.-----------..   255
HPV41       ......................RPQGSRRR....-N...EE-G--VAPAPK.................   242
COPV        ......................L-PRKGPS....-R...P-RRSS.RFP-R.................   240
CRPV        ....................PDSAVPAAQKKT.....G...P...KT.--TLG....-R----PG....   247
ROPV        ....................AVS.AGPAHHP-....TP...V.K-LSGPV--Y...G-R----PGVG..   111
HPV1a       ..................SIDYTEL.PGQGETS-V...-Q...-QQKT-.V-R--Y..G-R----P.....   268
HPV63       ......................QTA.RKAE-KGS....-H...HPK-PAVRKR--Y..G-R----P.....   259
MnPV        ............REGRGDYPDGARRET.RRYYQGPTP...TP...-SL-P-IY-PP-...-YEE--R-R....   286
```

E2 Appendix A

```
MOST-LIKELY ...PRKR.....................................................................   252
HPV54       .....---....................................................................   244
HPV32       .....-......................................................................   261
HPV42       .....-......................................................................   259
HPV3        ..EQQQQ.....................................................................   251
HPV28       ...WEQQ.....................................................................   250
HPV10       ...DGQH.....................................................................   249
HPV29       QQQQQQH.....................................................................   258
HPV61       ...-H--.....................................................................   253
HPV2a       ...-A--.....................................................................   261
HPV27       ...-A--.....................................................................   258
HPV57       ...-P--.....................................................................   253
HPV26       ...-G--.....................................................................   244
HPV51       .....-......................................................................   240
HPV30       ...-G--.....................................................................   245
HPV53       ...-G--.....................................................................   242
HPV56       ...-G--.....................................................................   244
HPV66       ...-G--.....................................................................   243
HPV18       ...AAT-.....................................................................   248
HPV45       ...AT--.....................................................................   250
HPV39       ...S---.....................................................................   248
HPV70       ...SC-C.....................................................................   248
HPV59       ...TG--.....................................................................   248
HPV7        .....-......................................................................   249
HPV40       .....---....................................................................   243
HPV16       .....Q-.....................................................................   240
HPV35h      ....H--.....................................................................   242
HPV31       ...ST--.....................................................................   245
HPV52       ....-Q--....................................................................   238
HPV33       ...LT-L.....................................................................   243
HPV58       .....-......................................................................   228
RhPV1       ...-A--.....................................................................   249
HPV6b       .....-......................................................................   239
HPV11       .....-......................................................................   240
HPV44       ...-C--.....................................................................   246
HPV55       ...-C--.....................................................................   246
HPV13       ...-C--.....................................................................   242
PCPV1       ...-Y--.....................................................................   242
HPV34       .....-......................................................................   235
HPV19       ...R-S-SKSRRKASTTR..........................................................   301
HPV25       ...S-S-SRSESQSSKRRSRSRSRR...................................................   308
HPV20       ...S-S-SESPRRRSRYRSRSGSR....................................................   306
HPV21       ...S-S-SRSYSRSRSQSSDQPQYRFRSGG..............................................   311
HPV14d      ...S-SQSSERRSRYRSRSR........................................................   303
HPV5        RSTT-S-STSLTKTRALTSRSRS.....................................................   312
HPV36       ...-TT-SATTRSRSPSLAKTGVQ....................................................   307
HPV47       ...STTS.....................................................................   318
HPV12       ...SQT-ALGATSVSRSSRSPSVTQIRNR...............................................   303
HPV8        TTVS-S-SSSLTKAVRPRSRSRSRG...................................................   311
HPV24       ...R-S-.....................................................................   293
HPV15       .ISRGGN.....................................................................   288
HPV17       ...GER-.....................................................................   283
HPV37       GGDR-R-.....................................................................   288
HPV9        ERGGER-.....................................................................   294
HPV22       ...EQRG.....................................................................   276
HPV23       ...EQRG.....................................................................   273
HPV38       ...TNRG.....................................................................   277
HPV49       ...STS-.....................................................................   314
HPV4        ...L-R-.....................................................................   258
HPV65       ...V-RG.....................................................................   258
HPV48       ...RER-.....................................................................   256
HPV50       ...L-RG.....................................................................   259
HPV60       ...R-R-.....................................................................   258
BPV1        ...-YNF.....................................................................   264
BPV2        ...-YHF.....................................................................   263
EEPV        ...APSS.....................................................................   273
DPV         ...TPQS.....................................................................   266
BPV4        ...-T-G.....................................................................   259
HPV41       .....-......................................................................   243
COPV        ...SGG-.....................................................................   244
CRPV        ...VQR-.....................................................................   251
ROPV        .FD-ARS.....................................................................   117
HPV1a       ...RGGG.....................................................................   272
HPV63       ...RDTT.....................................................................   263
MnPV        ..KL-R-QDGRVKYPASPYRTKPPGETSSDDEDEGRGGHEPRPQRRLPRGLRDRGERAPERRRPPVQEGEEDV    357
```

```
MOST-LIKELY .........................PRGRGRSPS...TTSSPSSQRS.....................RG   273
HPV54       .........................A-V....Y-...-DQQ-H-T.-.....................D.   259
HPV32       .........................Y-QSL.Q-P......TKHL-HY.....................GVT  279
HPV42       .........................Y-QC-Q------...QHLQH-NPSI...................PSIP 282
HPV3        .........................QQQQQQQQQ...H-QT-AP-TT.....................E-A  273
HPV28       .........................Q-Q...Y-Q......T--T-TT.....................E-A  266
HPV10       .........................QQQ-QG-KD.........-T-KA...................AE-A 267
HPV29       .........................QQQQQ.QQT...P-HT--T-AC.....................A-T  279
HPV61       .........................Q-L...HRD...REQQ-DTTQK.....................DN   271
HPV2a       .........................Q-VIV..GQ...QHPR-D-T-T.....................V-   280
HPV27       .........................Q-VIV..GQ...QHLR-D-T-T.....................V-   277
HPV57       .........................Q-VIV..GQ...QWQQ-D-T-K.....................VR   272
HPV26       .........................R-LS-..-D...--VTTVTTVTT.................AATQP  267
HPV51       .........................Q-LTE..-D.........--TI-....................PL.  254
HPV30       .........................--TTE..-D.......-TDTT-Q....................SA.  261
HPV53       .........................--TTE.PD-...-D-TTQ-TTT.....................AR   262
HPV56       .........................--L-E.-EF...DS-RE-HAKC...................VTTH 266
HPV66       .........................--ASE.-EP...DS-RE-YAHC.....................V.   262
HPV18       .........................-GHC-.LAE.........K-HC.....................G.   261
HPV45       .........................--QC-.LTE...........-HH.....................G-.  263
HPV39       .........................--QCA.VTE.........PTEP.....................D.   261
HPV70       .........................GVS...R-.............E.....................TD   257
HPV59       .........................--QC-.YTQ........HP-ST.....................SV.  263
HPV7        .........................R-D....GD...LSI-AVDGC-.....................GRK. 267
HPV40       .........................R-N....GH...LPITTTVGK-.....................L-   260
HPV16       .........................--S...E-D...-GNPCHTTKL.....................LH   258
HPV35h      .........................L--.....G...-ELPYNPTKR.....................VR   258
HPV31       .........................--T...E-E...HRNTHHPNKL.....................LR   263
HPV52       .........................R-PDV..TD...SRNTKYPNNL.....................L--  258
HPV33       .........................FCA.....D............P.....................A.   249
HPV58       .........................R-LDL.PD-...RDNTQY-TKYTD.................CAVD. 252
RhPV1       .........................V-RSDSGGD............P.....................V-.  261
HPV6b       .........................A--VQQ--C.........NALC.....................VAH  255
HPV11       .........................A--..............TNNTLC.....................VAN  254
HPV44       .........................---...P-T...N-NNARNTVC.....................V-N  265
HPV55       .........................---...PTT...N-NNARDTV-.....................V-H  265
HPV13       .........................Q--PS-PIG...NPQNTQ-IVC.....................VTD  264
PCPV1       .........................L--PSHCAR...KLQNT-NIVC.....................ATD  264
HPV34       .........................QCD....-D............E.....................G.   242
HPV19       ..................GRGRGSPTATSDQSS------...A---TT-L--RGSSRVG......RSRGGRSRV-  347
HPV25       ..........KTSATRGRGPGSPTTTTSD-AAR.------....-AT----QRSR............SRAGSSR 355
HPV20       ............GRVALRAITTTTTTTTR-AG-G--T...S---TT----RQ..............LRGGGR 349
HPV21       .....QVSLITTATTTTTTATNYSTRGSG--SS.-T-...SST-KRPR-P.....................--  354
HPV14d      ..............SRQKEVSRITTTT-----GS-...S---KR---A.....................--  336
HPV5        ............RGRSPTTCRRGGGRS--R-S------..-S--CTT---QRARAESST..TRGARGSRGS--  367
HPV36       ............RVSTRSRSRSTSRRGGR-R-S------..-S--TTTTNKRSRVRAETT.GSRGARGGRGA--  364
HPV47       .........................R--GR.GS-...-RQRSR-PSTYTSKRSREGNTRGRGRGRQGRA-  359
HPV12       .......RSRSQSRGRGGRGSSTDTTTKR-RSR.-S-...SNTRKR---G.................ERGR 346
HPV8        ...........RATATSRRRAGRGSPRRR-STS------..-NTFKR---GG................GR--  352
HPV24       .........................STS--NRRC...RGDT-RG--G.....................VS   314
HPV15       .........................Q-R-R--RE...-SI--AWG-G.....................GRSR 311
HPV17       .........................Q-R-E--Y-...RD--R-PN-G.....................RG-  305
HPV37       .........................R-E-SY-RD...-S---DRG-G.....................GR   309
HPV9        .........................R---SS-AD...S-TPTDRR-G.....................--   315
HPV22       .........................G-ATR--L-...RE-AE-PR-G.....................GRG- 299
HPV23       .........................G-ETQ--S-...RGA-K-PR-G.....................GRS- 296
HPV38       .........................G-DTR--S-...RG--V-PT-G.....................-R   298
HPV49       .........................GSRGSGGSV...---RD--PKR.....................TR--  337
HPV4        .........................R-EGKSG-G...SGET-RKR-R.....................GG-  280
HPV65       .........................R-Q-KSG-G...PGET--KR-R.....................GGGR 281
HPV48       .........................R-EPRE-GT...-DTT-RRRGT.....................KRKL 279
HPV50       .........................R-E-E.HH-...YRHRK-QSEL.....................GAD  280
HPV60       .........................SNEQQ-EL-...SRE--RTK-RRV.................PDEVDR 285
BPV1        .........................-A-S-G-IL...RS--TPV-GTV.....................PVDL 288
BPV2        .........................-A-S-G-LL...RSA-TPV-GPV.....................PVDL 287
EEPV        .........................-GSSV.G-D...SP-ES-R-VP.....................LVLL 295
DPV         .........................--SGV.G-D...---PLP-PVPQNPR.............CVSLPD- 295
BPV4        .........................-HSS-..RD...-RLPSPGRPP.....................GG-R 280
HPV41       .........................R--...AYG...RR---KA--R.....................TA   261
COPV        .........................G-LGR.GG-...GELP-QP-P-S.....................SWSP 267
CRPV        .........................-AKQRKQAA...PDEAD-AAGD.....................I-P  273
ROPV        .........................R-QGKHPTDFNAN-I-AD-TD-.....................TD   141
HPV1a       .........................R-EGE.-TP.............-.....................-T   283
HPV63       .........................L-RGE.GE-...ARA-AG-GER.....................VAF  284
MnPV        DGVGALLDDLKLYQEPPGDPVEDSDSPGS-LTP.A-P...DL-RYD-T-L.....................QVDA 407
```

E2 Appendix A

```
MOST-LIKELY GGPVDSGRRRSRSSSTS.SSNKR.....RLGRLLDEARDP  307
HPV54       ..--GCDND-HI-.DDN.NK-QG.....-H..TSSGD.TT  287
HPV32       NV---P-SQ-.VT.-DN.NN-Q-.....-N..PC.GNQTT  308
HPV42       SAS--P-LCGV-TN-EN..C---.....-N..HC.GSQAT  312
HPV3        RQ-L-TD-T-D-D.T-C..PHPI.....GH..RS-PD.CV  302
HPV28       SQ-L-VT-TSDCD.T-C..PYTV.....GH..PS-PD.CA  295
HPV10       --Q---D-T-LCD.TR-..AHPV.....-H..PS-PD.CA  296
HPV29       ------N-T-DCD.----.Q-PY.....-H..PS-PD.CA  308
HPV61       HKR---TDQW.INGHRN.-TETG.....DN...C-SY.SS  300
HPV2a       E-E-ECYNK--I-.DSN.RTDP-.....WG..HG-TD.SV  310
HPV27       E-Q-ECYDK--I-.N-N.-TAP-.....WD..HG-TD.TV  307
HPV57       E-Q-ECQND--IRNPD-..TDP-.....-G..HS-LD.AV  302
HPV26       -QS--YTNNNLH-T-GG..HHPG.....-D..TS.SDQTV  297
HPV51       ..S--NTNNQIHCG-G-..T-TG.....GH..QS.ATQTA  282
HPV30       ..ARE-HAN-VNTNN-N..NRQC.....LGGATCYNTEVDGGYKTT  298
HPV53       ESYAECVA-N.TD.N-N..N-T-.....KHLPGGASCNNTEIDSGYKTA  302
HPV56       THIS-TDNTD---R-IN.NN-HP.....GDKTT  293
HPV66       ..TT-TDISN.NAN-R-PRI-TQ.....SH...C.GDKTT  290
HPV18       ..--NPLLGA.AT.P-G..N---.....-K..-CSGN.TT  287
HPV45       ..VNTHVHNPLLC-----.N---.....-K..VCSGN.TT  291
HPV39       -VSL-HLNNPLH-N--G..H-T-.....-Y..-SCGN.TT  291
HPV70       -VF--.....LVT.-KG..C---.....-H..QCCGD.TT  281
HPV59       ..S--YCDNPVVRLHPG..N-P-.....-H..IPCSN.TT  291
HPV7        ..Y--T-N-A.--PDIE.-N--I.....-N...SGGGHST  295
HPV40       -EY--TAD-T.-TPDPE.-N-GH.....-N...CGGGSST  290
HPV16       RDS--..SAPILT.AFN.--H-G.....-I..NCNSN.TT  286
HPV35h      LSA---VD-GVY-T-DC..T--D.....-C..GSCST.TT  288
HPV31       -DS---VNCGVI-.AAA.CT-QT.....-A..VSCP-.TT  293
HPV52       QQS---TT-GLVT.A-E.CT--G.....-V...AHTTCTA  288
HPV33       ...L-..N-TA-T.A-N.CT--Q.....-T..VCSSN.VA  274
HPV58       ..SRPR-GGL.H-.T-N.CTY-G.....-N..VC.SSKVS  279
RhPV1       ..AL-GKS-SVLCG-AH.NNATG.....SSGD.....SD.YT  289
HPV6b       I------NHNLITNNHD..QHQ-.....-N..NSNSS.AT  285
HPV11       IRS---TINNIVTDNYN..KHQ-.....-N..NCHS-.AT  284
HPV44       SDS---TNNNILPN-YN..---G.....-D..NNYCT.AT  295
HPV55       SDS---TNNNIYPN-YN..---G.....-D..NNFCT.AT  295
HPV13       YDTL--ANNNINVNHYN..N--G.....-D..NSYC-.AT  294
PCPV1       R-TL--ENNI.NNNNYN.NN-QQ.....-N..NSNSS.GT  294
HPV34       ..-L-FVHNLQPT.TD-..-TQC.....T-.....HN.VA  266
HPV19       RSRGRGK-S-.E-P-PT..NT--SRRQSGSS.RLHGVSAD......AVGTSVHTVSGRNTG------E--L--  410
HPV25       --RGRG----H-L.-E-.PTS--SRRESGSV.RLHGVSAD......AVGTSVHTVSSRHTG------E--L--  419
HPV20       --SRQRA-G-.------PTPS--SRGESESV.RQHGISPS......DVGTAVYTVSSRHTG---------L--  414
HPV21       -AIGG-SG-GR------P-PS--SRGKSESV.RQRGISPD......DVGKSLQSVSTRNTG---------L--  420
HPV14d      R-RGG.S-G-.------PT-S--SRRESESS.RQRGISPS......DVGKSLQSVSSRNTG---------L--  400
HPV5        -SRGGR----G----S-SPAH--SRGGSAKL.RGVSPGE........VGGSLRSVSSKHTG------E-----  431
HPV36       -SGGGRR-G--S-.---.PAH--SREHSVRS.RGVSPDQ........VGKSLRSVSSKHTG------E--L--  426
HPV47       SSGGREQ---R--F---PD-S--VRRESPKY.RGVSPSE........VGKQLRSVGAKHSG------E-----  423
HPV12       -ERGGR-K--D------.PTP--SRAGSRSS.RQRGVSPE.......QVGRSLQSVSSKHRG------E--L--  411
HPV8        R-RGSR---E.------PTPT--SRGESSRL.RGVSPSE........VGRSVQSVSAKHTG---------I--  415
HPV24       TSSRGR--GSR----S-.-PTP-TKASQRGC.DTRSVRDSGIS.PGDVGRKLQTVSGRNSG------E--L--  384
HPV15       R--TTRSQSK-L-R-R-.R-KS-SRGSSPR..GGISPAD........VGSSVRSLGRKHTG--E---E-----  373
HPV17       SSGGPTT-SQ---L-R-.R-RS-SRSRGSSAGGGVAPEQ........VGKSVRSVGRNPGG--T---E-----  369
HPV37       SRGGPET-SQ---L-R-.R-RS-SRGSS.SR.GGVAPDA........VGKSVRTVGRDHSG--K---------  371
HPV9        --RGPTT-SQ---R-R-.H-RS-SRGGTASR.VGVSPDE........VGTRVRSVGAGHHG--A---A--K--  378
HPV22       ---LTRS-S----RTRE.-VDGG..........GVAPDE........VGATLRSIGRQHSG--AQ---A-K--  353
HPV23       ---LTRS-S----.PE-..VTGG..........GVAPSE........VGASLRSVSRHSSG--AQ---A-K--  348
HPV38       R-GG--R--G.PV.TR-.R-RSLSRASSAG..GGISPDK........VGTAVRSVGRQSGG--T---AD-A--  358
HPV49       R-RGGRS--S.PT.P--..TS--ERRRSRSR.GGEPVSGGVGISPDKVGSRVQTVSGRHLG------E--S--  405
HPV4        R-GGETELGSAP-PAEV.G-RH-.....QVERQGLS---L-QA-----  322
HPV65       --GETRLESA.P-PGEV.GIRH-.....TVERQGLS---Q-QA-----  322
HPV48       -SDSAPTPSEVG-R--T.LARHG.....YS-----QE-----  315
HPV50       SA-TPEEVG-.---.H-V..AAHG.....LS--R--QE-----  313
HPV60       QSA-G-APTAEEVG-RH.R-LP-SGIS.....--A--QG-----  323
BPV1        ASRQEEEEQS.PD.--E..EEPV.....TLPRRTTNDGFH--.K-GGS  326
BPV2        APRQEEEENQ-PD.--E..EEPV.....TVPRHTSDAD-FH-LK-GQS  327
EEPV        P--S-PAPPS.PD-TDV.IAEGDKEPE.....-FSI-SKPG.GQ  331
DPV         F-RGEEDNPP-PDQHDV.IP-PQPKEP.....-F.S-FGSSGGL  332
BPV4        R-TPERE-CP.GT.P-P..PTPD.....Q........VGGRSSTPKRQASS--AQ-I-A-Y--  326
HPV41       AS--SR-NGG-SD.F---.GESD.....EGHRVRH-AL-KK.T-GVA  299
COPV        PS-QQV-SKHQLR.T--..-AGG.....------Q--Y--  299
CRPV        PA-E-V---TTTVGR-P.PGRN-.....--RE-IT--S--  307
ROPV        FS-AFRPPTP-EVGRRN.TTAP-.....ESARGLGG-VRQ-IS-----  183
HPV1a       P-S-P-A-DV.G-IH-T..PQ-G.....HSS--R---Q--W--  318
HPV63       IS-G-.....VGT-TR-..PP-G.....GQS--R--IQ-----  315
MnPV        ESSPPRTP-PAPTLVAE.CTPG-.....PSPQT.....GSGQQALGEPPSRPS-G..HCRDP-TA  459
```

```
MOST-LIKELY P.....VIHLKGDANTLKCFRYRLKKKYKGLFKNISTTWHWVGGD...GT..ERLGI..VTITFTSETQRQDF    368
HPV54       -.....IV-F--EP-------Q-IQ-.--H--EQA-S----ACVP...--T.KNR--..--L-YS-VE---Q-    348
HPV32       -.....----Q--P-C---L-W-----NCSH--TQV-S---LTEK-...Y-RDSKD---.I--HYYN-E--DK-    371
HPV42       -.....----Q--P-C---L-F-----RNCSH--TQV-S---LTEN-...C-RDTKT---.I--HYYD-A--NL-    375
HPV3        -.....----R--P-C-------N-GKNK-YSRT-S--R-S.CE...SE..NQCAY..---WY--YG--EA-    362
HPV28       -.....-V-----P-C-------H-GKRK-YCKT-S--R-S.CE...SE..NQAAF..---WY--YS--NE-    355
HPV10       -.....----R--P-S-------HHGKRK-YSRS-S--R-S.CE...SE..NQAAF..--LWY--D---TE-    356
HPV29       -.....----R--P-S-------QNGK---YCKA-S--R-SCEP...EN..QS.AF..---WY--V---AE-    368
HPV61       -.....-------P-K-------QHSVPE--DKA-S----A--Q...S-..T-AAF..--LWYVNVE--KQ-    361
HPV2a       -.....----R----C-------VQ-HKDV-YARV-S----A--N...-D..KT.AF..--LWY--VE--TE-    370
HPV27       -.....----R----C-------VQ-HKDK-YDRV-S----A--K...CD..KT.AF..--VWY--VE--KE-    367
HPV57       -.....----Q-E--C-------VQ-HKDV--VKA-S----AC-N...-D..KT.AF..--LWYK-QE--AE-    362
HPV26       F.....IV-----T-S---L---F--.H---YC-V-S----TSN-...TN..QQ.---.-----N-I---NN-    356
HPV51       F.....IV-----T-C-------FT-.H---Y--V-S-----T......SN..TKT--.---V-D-AH--ET-    339
HPV30       -.....-V----EP-R---L---CQ-.H-H--V---S-Y--T.NT...H-..-Y.SY..I-VVYKD----AN-    356
HPV53       -.....-V-I--E--R---L---FQ-.H-Q--VTV-S-Y--TNVN...CA..VNNSY..I-VVYKD-----K-    362
HPV56       -.....-V----EP-R---C---FQ-.--T--VDVTS-Y--TST-...NK..NY.S-..I--IYKD----NS-    352
HPV66       -.....------E--R---C---FQ-.--T--TDVT--Y--TST-...NK..DS.S-..I--LYKD----DT-    349
HPV18       -.....I------R-S---L----R-.HSDHYRD--S----T.-A...-N..-KT---.L-V-YH-----TK-    346
HPV45       -.....I------K-S---L----R-.-ADHYSE--S----T..G...CN..KNT---.L-V-YN--V--NT-    349
HPV39       -.....I------K-G---L----Q-.-DT--E---C----IR-K...---.KNA---.L-V-YAT-S---K-    351
HPV70       -.....IV-----K-G---L---R-.FNS-YE---C----I--K...-S..KHT---.L-V-Y-T-A---K-    341
HPV59       -.....I------K-G---L---R-.VHW--E---S----T-NR...-S..AKT---.L-L-Y------NE-    351
HPV7        -.....I-Q-E----C-------T-.VSH-YT-S----R-TTES...R-..NKNA-..I-L-YS-VH--SQ-    355
HPV40       -.....I-Q-E--C-------G-.VSH--C-S----R-TTES...R-..-KNA-..I-L-YS-VQ--S--    350
HPV16       -.....IV-----------L---F--.HCT-YTAV-S----T-HN...VK..HKSA-.---L-YD--W--DQ-    346
HPV35h      -.....IV-----------L---G-.--A-YQDA-S--R-TCTN...DK..KQIA-.---L-Y-T-Y--DK-    348
HPV31       -.....I--------I---L---S-.--Q-YEQV-S----TCT-...-K..HKNA-.---L-YI-TS--D--    353
HPV52       -.....I------P-S---L---V-T.H-S-YVQ--S----TSNE...C-N.NK---.----YSD-----Q-    349
HPV33       -.....IV----ES-S---L---P.--E-YSSM-S----TSDN...KN..SKN--.---V--VT-Q-Q-M-    334
HPV58       -.....IV-----P-S---L---P.F-D-YC-M-S----TSD-...KG..DKV--.---V-Y-T-----L-    339
RhPV1       -.....IV----ES-C---L-F--G-.H-H-YI---S--R-A.NH...AS..-K.A-.--V--AN-L---Q-    347
HPV6b       -.....IVQFQ-ES-C-------NDRHRH--DL--S----ASSK...AP..HKHA-.---V-YD--E---Q-    346
HPV11       -.....IVQ-Q--S-C-------ND---H--ELA-S----ASPE...AP..HKNA-.---L-YS--E---Q-    345
HPV44       -.....-VQ-Q----C---L---HA---T--VAA-S--R-TCS-...TS..SN.AL.--L-YVD-Q--Q-    355
HPV55       -.....-VQ-Q--P-C---L---HA-H-T--VAA-S--R-TCS-...TS..SKHAL.--L-YVN-E--EQ-    356
HPV13       -.....IVQ-Q--S-C-------HE---D--LLA-S----TAPN...NS..QKHAL.--L-YVN-Q-----    355
PCPV1       -.....IVQ-Q--S-N-------HD---H--MLA-S----TASS...NS..TKNA-.---L-YVN-Q-----    355
HPV34       -.....IV-----K-S---L---MH-G-SH--N-VT-----T.NN...TN..SKC-V..I-FM-S-TS-QKQ-    326
HPV19       -.....--LVR-EP---RS--N-A-HM-R---SSF--A-S--A--...-I..----RTRML-S-V-FN--KH-    473
HPV25       -.....--LVR------RS--N-A-HM-T---SSF--A-S--A--...-I..----RSRML-S-I-NS--KH-    482
HPV20       -.....--LVR-EP-------N-A-QR-T--Y-SF--A-S--A--...--..----RSRML-S-I-FS--K--    477
HPV21       -.....--LVR-EP-------N-A-L--A--Y-AF--A-S--A--...--..----RSRML-S-F-FE--K--    483
HPV14d      -.....--LVR--P---R---N-A-Q-FT--YRAF--A-S--A--...--..----RSRML-S-F-FN--R--    463
HPV5        -.....--IV--A-----NV-N-A-I--M---RSF----S--A--...--..----RPRML-S-S-Y---R--    494
HPV36       -.....--LVR-E--------N-A-I--M--YRSF----S--A--...--..----RPRML-S-S-YN--R--    489
HPV47       -.....--LVR----------N-ARN--R---RSF---FS--A--...SI..----RSRML-S-SCL---R--    486
HPV12       -.....--IC--G--------N-ARH--T----AF----S--A--...S-..----RPRML-S---TN--K--    474
HPV8        -.....--LVR-E--------N-ARVR-R----YF----S--A--...S-..----RSRML-L---AG--K--    478
HPV24       -.....--L-R-G--------N-A-LR-R-HY-AF--S-S--AA-...--..----RSRLLVS---FK--SG-    447
HPV15       -.....--L-R----K-----F-A----QD-V-YY----S----T...SN..D-I-RSRLLLA-S-N-E-EL-    436
HPV17       -.....--L-R-E--K-------A--R-GS-V-YY----S---AN...TN..D-I-RSRMLLA-NTYDE-EL-    432
HPV37       -.....--V-R----K---Y---A---HGN-V-YY----S----S...TN..D-I-RSRMLLA-Q-N-E-EL-    434
HPV9        -.....LML-R----V---Y-F-ER--KR--V-YY----S---E-...SC..D-V-RARMILA-DTYEH--Q-    441
HPV22       -.....--L-R-A------Y---FR--HA-S-QF-----S----H...T-..D-I-RSRIL-S-HTDRE-EKC    416
HPV23       -.....--L-R-G------Y---FR--HA-K-YYV----S-I--H...S-..D-V-RARML-A-H-NHE-EKC    411
HPV38       -.....--L-R--------Y---FR--HA-G-RFV----S-I-DA...SN..D-I-RSRMLLA-Y--S--EK-    421
HPV49       -.....--L-R--P-I---Y---D--RKL--V-HY----S---V-...-N..--I-RSRMLLS---NST-SQY    468
HPV4        -.....M-L---T--S---W---KVNSNCCN-LFM--V-N--.----.CS..HNHSR..ML-A-D-TD--DA-    382
HPV65       -.....M-L---T--S---W---KQNSSNCG-LFM--V-N--.----.VS..-NHSR..ML-A-K-PG--DS-    382
HPV48       -.....LVLFT-QQ-N---W-N-CTT--AS--LCF-SV-K-L-PN...SD..GGAAK..-LVA-K-DA---V-    376
HPV50       -.....-LIIT-QQ-N---W---FSQ--AD-YECC-SA-K-L-PK...SE..GYR-DAKLL-A-KNPE--LS-    376
HPV60       -.....ILLI--L--S---W---.---.-TRY--CM--VFR--DI-...VP..-S.SRHKLLVV-NDT---DV-    384
BPV1        C.....FALIS-T--QV--Y-F-V--NHRHRYE-CT---FT-ADN...-A...-Q-QAQIL---G-PS-----    389
BPV2        C.....FALIS-S--QV--Y-F-V--NHRHRYE-CT--SFT-ADN...-A...-Q-QAQIL---G-PG-----    390
EEPV        -.....CLI-S-NG-QA--Y-F-C-RYFREHYQH-T---WT--ER...-S..--H-DAC-LV--KDSS--GV-    394
DPV         -.....CLLIS-TG-QV--YSF-V-RWHRDKYHHCT---WA--EQ...-S..--P-DAT-IV--KDQS--SM-    395
BPV4        -.....-LL-Q-A--------R-ATQAHPHK-LCM--S-T--SKT...SP..LKS-H.RML-A-SNSE--NC-    388
HPV41       -AEGHYLVGA--PV-S-R-L--Kw-N--S-DIMYLG--FT-TES-.....--.--C-SGRFFCA-SN--K-EK-    367
COPV        -.....-LV-A--P-S---I---I---SH-HR--YLGA----K-TS-GDGASK..HDR-SARMLLA-L-DQ--E--    365
CRPV        -.....--C---GH-Q---L---I---S-HSS--DC-----S--DTT...S-..C---SGRML-K-ADSE--DK-    370
ROPV        -.....--C---GN-Q---L---I---A-HRT--DC-----S--DNS...S-..C-V-SGR-L-K-KD-A--EKV    246
HPV1a       -.....-VCV--G--Q---L---I---ASTQVD-DS-------TDRK...N-..--I-SARMLVK-ID-A--EK-    381
HPV63       -.....I-C---GP-Q---L---I-ASNSSD-ES--------HNK...C-..D-V-HARMLVR-I-TE--DR-    378
MnPV        C.....LLII--SS-QV--L-F-I---SWHHS--SY-----Q--PSV...-S..N-I-RSRILVMCEDSA-MDR-    522
```

E2 Appendix A

```
MOST-LIKELY LNTVKIPKGVQVSLGYMD...SLG     389
HPV54       -V--R--PSISM---V-....--$     367
HPV32       -S---L-P-IKSCI---SMLQFM$     394
HPV42       -------S-IKSCI---SMLQFI$     398
HPV3        -S---V-P-I--I--H-SM..FT$     383
HPV28       -S---V-P-I--I--H-SM..FV$     376
HPV10       --V--V-P-I--I----S...IF$     376
HPV29       -AN----P-M-AI--H-S...VF$     388
HPV61       --R-T----I-ATA---SM..CI$     382
HPV2a       -TR-S----LIALP---SA..FV$     391
HPV27       -TR-N-----IALP---SA..FV$     388
HPV57       -TR-HL----KALP---SA..FV$     383
HPV26       -T-----QSITST--I-....--$     375
HPV51       IK-I-V-PS-TL---I-....T-$     358
HPV30       --V----PSIKIVM-H-TGV.DM$     378
HPV53       -DI----PS-SLV--H-TCV.DM$     384
HPV56       -SH----#S------Q-#.....      368
HPV66       --V----PS---I--Q-S...CP$     369
HPV18       ----A--DS--ILV---....TM$     365
HPV45       -DV-T--NS--I-V---....TI$     368
HPV39       -D-----SS-H--------...T-$     370
HPV70       -E--R--PS-H--V---....T-$     360
HPV59       -D-----NS--IHV---....-V$     370
HPV7        -AL-----TIKH---MLT...IM$     375
HPV40       -AI-----TIKH---MLT...LM$     370
HPV16       -SQ-----TIT--T-F-....-I$     365
HPV35h      -T-----NT-T--K---....-I$     367
HPV31       -------NT-S--T---....TI$     372
HPV52       -K-----NT---IQ-V-....--$     368
HPV33       -G-----PT--I-T-F-....T-$     353
HPV58       -------PT--I-T-V-....--$     358
RhPV1       -------ST-TL-Q-V-....TV$     366
HPV6b       -DV----PTISHK--F-SLH.L-$     368
HPV11       --S----PTIRHKV-F-SLH.L-$     367
HPV44       -----L-PK-TYKV---SLQ.L-$     377
HPV55       ----RL-PT-TYKV---SLQ.L-$     378
HPV13       -K-----PTITHK--F-SLQ.L-$     377
PCPV1       -------ATIKHT--F-SFQ.L-$     377
HPV34       -QCA---PTIS--S---....-I$     345
HPV19       DD--RY----DR-F-SF-...--$     493
HPV25       DDA-RY----DR-F-SF-...--$     502
HPV20       DE---Y----DR-F-SF-...--$     497
HPV21       DK---Y----DR-Y-SF-...--$     503
HPV14d      DQ---Y----DR-F-SF-...--$     483
HPV5        DEA-RY----DKAY-NL-...--$     514
HPV36       DDV-RY----EK-Y-NL-...--$     509
HPV47       DDA--Y----EW-Y-SL-...--$     506
HPV12       DE---Y----ETAY-NL-...--$     494
HPV8        DE---Y----DT-Y-NL-...--$     498
HPV24       -DL-RF----DW---SF-...K-$     467
HPV15       IKIM-L-P--DW----L-...D-$     456
HPV17       IQKM-L-P--DW---HL-...D-$     452
HPV37       -K-M-L-P--DW---HL-...E-$     454
HPV9        IR-M-L-PT-DW---NV-...D-$     461
HPV22       -QQM-L-L--EW-Y-QF-...D-$     436
HPV23       IQEM-L-L--DW-Y-QF-...D-$     431
HPV38       IQ-M-L-T--EW---QF-...D-$     441
HPV49       VKIM-L----EW-F-NF-...K-$     488
HPV4        VKHNLF--LCTYTY-SLN...--$     402
HPV65       VKHNLF--LCTYTY-SLN...--$     402
HPV48       ----H----TTIT--RL-...--$     396
HPV50       ----GL--NTTY-M-HL-...--$     396
HPV60       MKL-TL-R-CTYTF-TLN...--$     404
BPV1        -KH-PL-P-MNI-GFTASL..DF$     410
BPV2        -KH-PL-P-MNI-GFTASL..DF$     411
EEPV        -KR-PL-P-MRAQALT-IA..DF$     415
DPV         -QQ-PL-P-MSAHGVT-TV..DF$     416
BPV4        -AS-RL----SAVK-AL-...G-$     408
HPV41       -KS-----NIGLFRAHAE...K-$     387
COPV        MDR-TF--S-R-FR-GL-...E-$     385
CRPV        -SR-PL-STT--F--NFY...G-$     390
ROPV        -EE-P--RHM--FV-NFF...G-$     266
HPV1a       -ER-AL-RS-S-F--QFN...GS$     401
HPV63       -DK-VV--S-S-I--AF-...GS$     398
MnPV        -C-----A-MT-EQCS-A...-V$     542
```

```
      **COBBLER sequence from MOTIF**
>hpv_E2_ HPV13, with embedded consensus blocks
metiakhldacqEQLLDLYEKDSKDLEDHIQHWKLIRKENVLLYYAREKGITRLGHQPVP
PLAVSKakgheaiemqmtletllleseYGNEPWTLQDTSxEMWNTPPKNCFKKGGQTVEVM
YDGDKDNTMxYTMWKYIYYfdtdkwtkvkgmvdykGLYYTHDGxKVYYYVQFEEDAKKYGK
TGQWEVHigstvicspasvsstvqevsiagpasysttttstqastavscsaseecvqappc
krqrgpsrpignpqntqsivcvtdydtldsannninvnhynnnkgrdnsycaatPIILLK
GDANSLKCFRYRlhekykdlfllasstwhwtapnnsqkhalvtltyvneqqrqdflktvk
ipptithklgfmslqll
```

Appendix B: Secondary Structure Prediction from E2 Sequences

Protein sequences such as the E2 sequences that display less than 30% similarity might nevertheless be shown to have similar structures. In general, we tend to learn more about structure from dissimilar (but homologous) proteins than from highly similar proteins.

This appendix summarizes the secondary structure predictions over the E2 HMM-predicted sequence as determined by several different algorithms, Gibrat, Levin, DPM, and SOPMA. Two consensus structures are also reported, one based on the four different algorithms, the other (at the top of the printout) based on individual E4 sequences as analyzed by the SOPMA method. The derivation of an HMM model sequence ('most likely sequence') is discussed in appendix A and elsewhere in Part III (Farmer and Myers). The various methods for secondary structure prediction are also discussed elsewhere in Part III of this compendium.

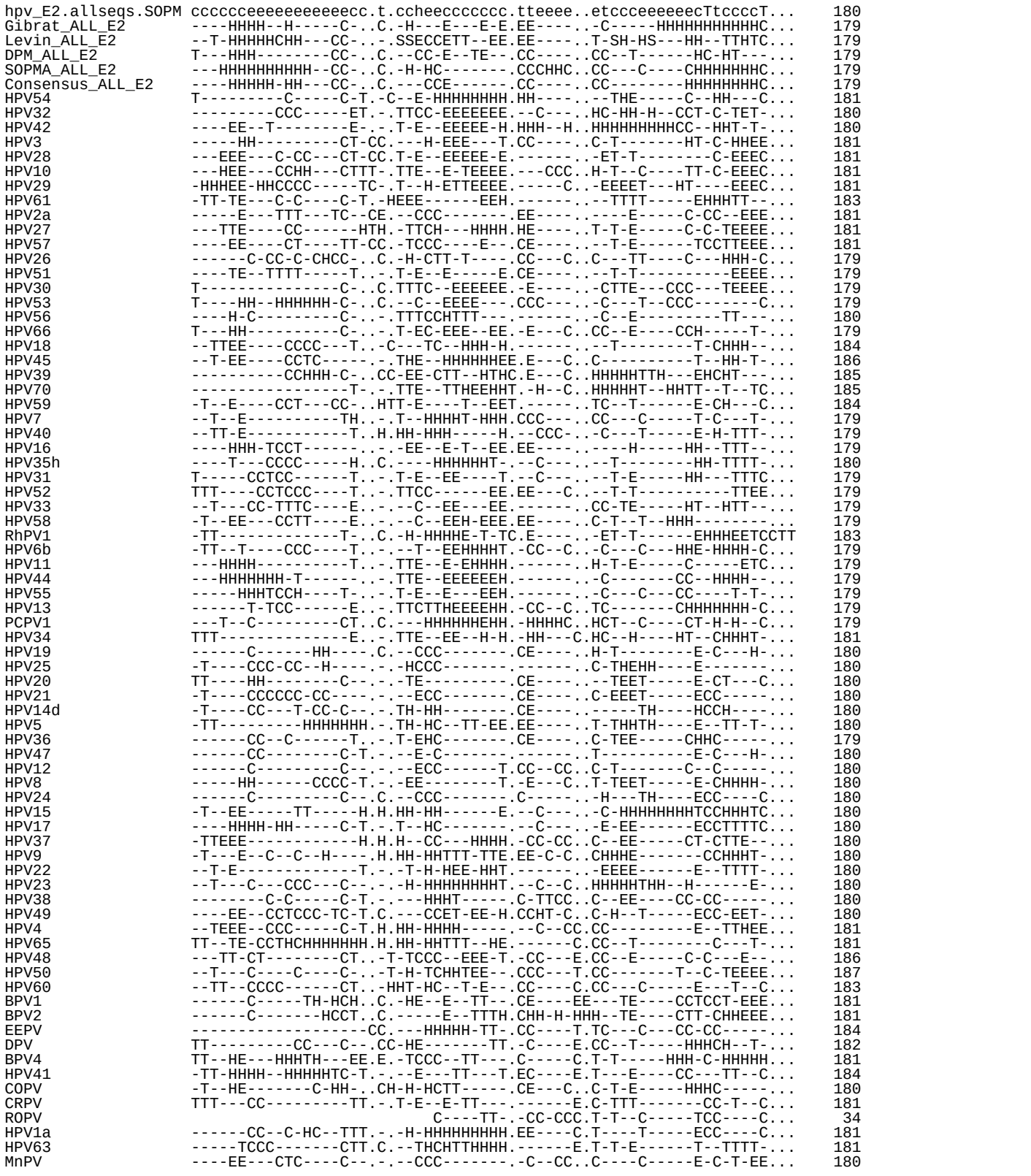
The structural code encompasses lower and upper case letters for alpha-helix (h, H), beta-sheet (e, E), turns (t, T), and random coil (c, C). States that are predictable with greater confidence are shown in upper case. The criteria for designating a state as upper or lower case are spelled out in the general discussion of this approach in Part III: states that are predicted in upper case letters have i) scores that are equal to or greater than the median average for scores assigned to that state over all positions and ii) scores that are in the upper quartile of difference from the second highest predicted state. Hence the absolute and the relative scores must meet stringent requirements to warrent upper case prediction.

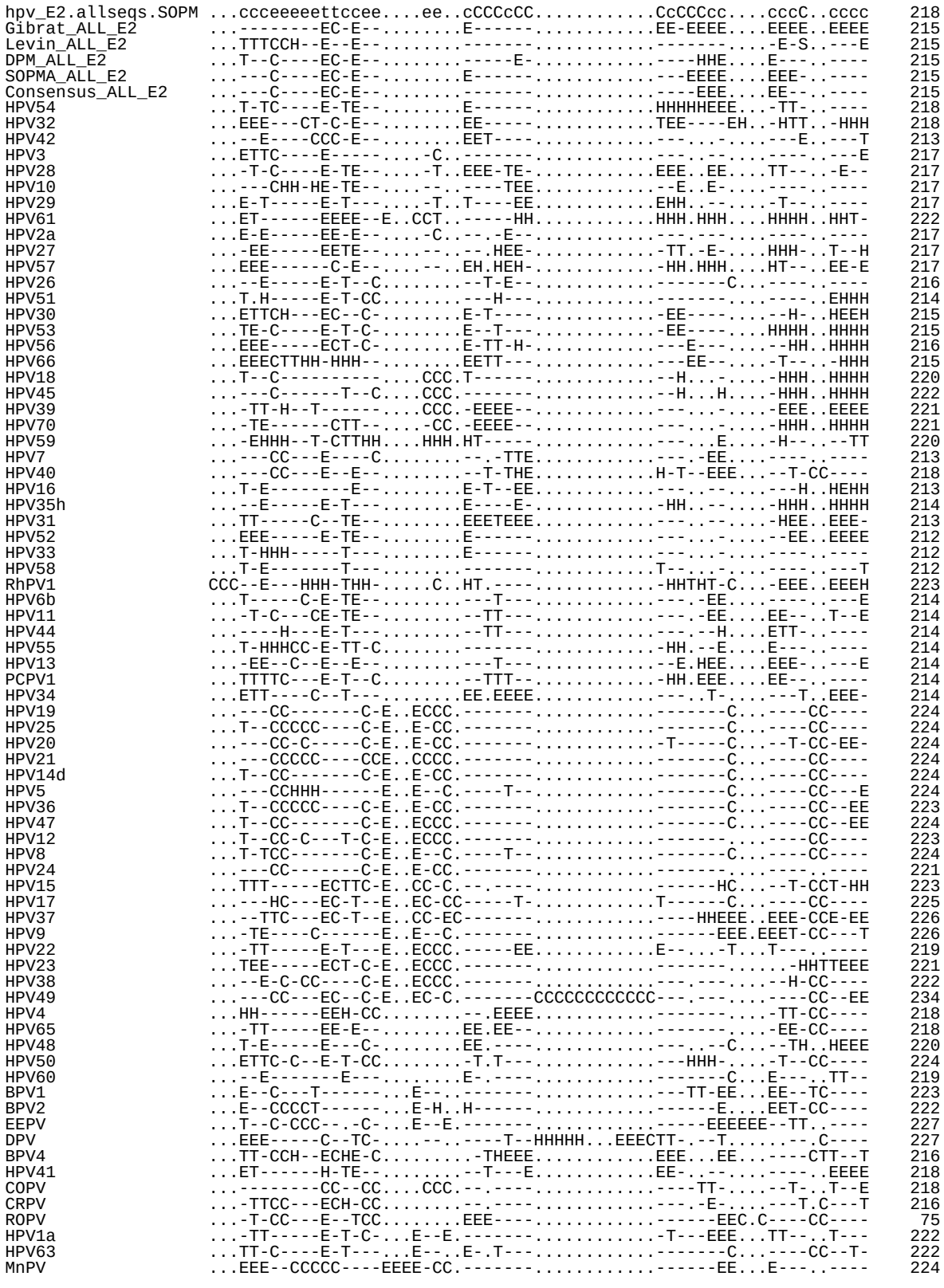
E2 Appendix B

```
hpv_E2.allseqs.SOPM h......hhhhhhhHHhHHHHHHHHh.hcchhhhhhhhhhhhhhchhhheeeeccccccceeeeccc     58
Gibrat_ALL_E2       -......--------------------.-HH------------H----HHHHHHHHHE--CC---     58
Levin_ALL_E2        -......--------------------.-HHT-----C----CCH-T-E--CCHTS---CCCTS--     58
DPM_ALL_E2          C......C------------------.CTTCC----E--E-E-H--EE---HHHHH-E-C-CEEE     58
SOPMA_ALL_E2        C......CC-----------------.-HH-------------H----HHHHHHH-EE-CCCEEE     58
Consensus_ALL_E2    C......C------------------.-HH-------------H----HHHHHHH--E-CCC---     58
HPV54               -......-------------------C.C-TT--T---------H----HHHHHHHHT-CHHHHH-     58
HPV32               -......--------------------.------------ETE-EE--E----TTTHH-CCCCTTE     58
HPV42               -......--------------------.-H-----------E-EE-EE---C----TECCC-EEE     58
HPV3                -.......----T-CCCC--C-----C.C--CCCCCCCEECEEEEE---HHHC-HH---C--CEEE     58
HPV28               -.........----CTT--C----------.C--T-CTC-----T--H--T-T-HH---------TEEE     58
HPV10               -......--------------------.-HH-------------H----HHH-T----E---TTEE     58
HPV29               -......---------C---------C.CT--------------H--TT----TT-TTEC--TTEE     58
HPV61               EH....H--------------------.C--CCC----EEETEETCCT------TTTT----T-E-     60
HPV2a               -.......------CC-C------EEC.C--CCCCCCCCCC-----------C---------C---     58
HPV27               -......--------------------.--H------------H----HH-CTT--T----T---     58
HPV57               -......--------------------.-TT-------C-----H----HH-C--H--E---C-EE     58
HPV26               -......CC----CCCCT------EE.C--TCCCCCCEECC--H----HHHHH-----CCCC-H-     58
HPV51               -........---C---CCC---EEEEEE.--TT-----EEETTETTCEEE---H-TH---H---TE-     58
HPV30               -......--------------------C.-TTTTCTTEEEEETCC-TCEE---C------HHHHTHH     58
HPV53               -......--------------------.-T-CCCTTC-------ETTEE-----TT---HHC--HH     58
HPV56               -......------------------CCT--C------------TTCEE------TT------HEH     59
HPV66               -......---------C----------.--T---EE-------CTTEEE---C--TT------EEE     58
HPV18               -CC..CH--------------------.THH---T-CETEEEEEEEET----HH-TT-E----EEE     62
HPV45               -HCCCCH--------------------.T-TTTCTTEEEEE--EH---E-H-H------C---EEE     64
HPV39               -HH.HHH----------T-------T.C-H----T--CCTTEEEETC-H--HHTTT-E---TTT-     63
HPV70               -HH.HHH--------------------.-H-----T-CCCTTEEEEEE----H-TTTT-----EEE     63
HPV59               -HH..HH---------C----------.THH---T---C-EEEEEECCE----H-T-----CC--E     62
HPV7                -.......--------CC---------.-HH--CTT-EETEEEEEE-EE----E--EEEHHHTEE-     58
HPV40               -......--------------------.-HH----T-EE-TEEEEEC-E----EEETEE----EEE     58
HPV16               -......-----CCC-C----------.T---TCETECCEETECH----HHHHHHHT-HT---TEE     58
HPV35h              E.....H--TTCCCCCC--------T.T--CECTTCCC-ETEEH----HHHHHHHTT-C---T--     59
HPV31               -.........-----CCC-C----------.C-H---TT-EE---TT-E--E----EEETT-CCC-TEE     58
HPV52               -........-C----------------.-H-T------------H--------TT--TE---TTEE     58
HPV33               -......--------------------.--H-CCCC-----CEEEEE--HH-C-TT---CTCHHH-     58
HPV58               -......---------------EE--.TTTTCCC-------EEEE---THHC--TTTEHHHHHH-     58
RhPV1               -.....H--------------------.-HH-----------H----H--TTTT---CCCCHEE     59
HPV6b               -......--------------------.-H-----CC---------C-E---TH---T-CCCCEE-     58
HPV11               -......--------------------.-HH----------C--H----HHHTT--TEE----EEE     58
HPV44               -.........-----CCCCC---------.--T--CEEEETCC-C-EE-EEHHHHH-HT--CHCTHEE     58
HPV55               -......--------------------.--TTCCEE---C--CEEEEEEHHHHHHHHH-C-C-EEE     58
HPV13               -......--------------------.-HH------------EEEEEHHTTE--TTECC-TEEE     58
PCPV1               -......--------------------.--T------------TTCCE---HH---T-CC--EE-     58
HPV34               -.....H------------------T.T-H-E--TTCC-----H--EEHHHTHHHETEC-CCTTE     59
HPV19               C......CC--CCCTCC----------.-HH------CCCEECC----E---H---------C---     58
HPV25               C......CC-TTCC--------CCCC.CHH----CCCCCCCCC-TC-E---H-----E--CC---     58
HPV20               -......--------------------.C--C-------C--------E---------E----E--     58
HPV21               C......CC-TCCC-----------.-HH--------TCCCC-T-----------C---EEE     58
HPV14d              -......CC---C-------------.C-H-----C--CCC-C----E----------C-CCEEE     58
HPV5                -......--------------------.-HH-----------CC-CCCE---------E----E--     58
HPV36               -.......-C-----------------.-HH----------CC--------H---------E--     58
HPV47               -.......-C-----------------.-HH--------E-CCC-TCCE----------C--C---     58
HPV12               C......CC---C-------------.C--CCC--CC-TCECC-C--E----------CC---     58
HPV8                -......--------------------.-HH----------EEC-T--E--------TE---C---     58
HPV24               -.......-C-TTCCCCC--EEEECCC.C--CC---C------C----E---------E-------     58
HPV15               -......C-----------------.T--CCCCEEEEECETT-TTEE----ETTTTE--CCT--     58
HPV17               -......CC-TCCC-----------.C--CC-EEECTTCTTC-C-EE----TT---E----E--     58
HPV37               -......CC---C-------------.--TCC------TTE----CTE---TTT---E-------     58
HPV9                -......--------------------.--------TTTCCCC-H----HHHH---T-E-------     58
HPV22               -......--------------------.CHH-------------H----H---TT-TTE--CC---     58
HPV23               -......--------------------.-HH-----------C----E----TTT--EC------     58
HPV38               C......CE-TTCEEEE----EEEEC.C----TCCCCTTCECEE----HHHH----TE-------     58
HPV49               -.......---TCCCCCEE--EEEECC.C--C-------CC-EC-TCCC----E----E-------     58
HPV4                -......--------------------.--HCC---------TC----E---HHHH---HTTT---     58
HPV65               -......-------------EEEECC.CT-C---T-------T-----HHHHHH----CHCC---     58
HPV48               CCT.TCC-------C-C--EEEEEEC.C-TTCCTTTEEEE--------HHHHHTTTEE---C---     63
HPV50               CCHHHHH---------------EEEC.C--TCCTCEEEEEC--H----HHHHHTTT-E-------     64
HPV60               CC....CCECC----------------.-HTC-----------H-EEE----ETTTT-C--C---     60
BPV1                -......----------T---------.E--CCCC--EEEEECC-----H---TT-TEE-----T-     58
BPV2                -......----------T---------.-HT-T----EEEEEC------HHHHHH-TEE-------     58
EEPV                EC....C--CCCC----CCCC-----C.C--CCCCTCEE------TT-E---C------C--CEE-     60
DPV                 C......-CC-------CCCEEEEE-.---TCCCTTECCCCCC-T-TTH-HH----EE-------     58
BPV4                -......--------------EEEEC.TTTTC-EEEEEECCEC-TTCC----EE-TTTTTT-EEE     58
HPV41               -H...HH--------------------.THHECTT---EEEE---CCCC----EETTT-CCCTTH-     61
COPV                -......C---------------EE-.T--C-----CTCCEEC-TTCE----E--T-E----E--     58
CRPV                -......--------------------.-H-----C-------T-CCTE----E--TT-C-TT---     58
HPV1a               -.......-C----------------C-.-H-CCCCCT-C-CC-------H------T-E----EE-     58
HPV63               -......--------------------.CHH--C----EEEECC-----H----TTT-E----EEE     58
MnPV                -......--------------------.-HH-----------CC-TTTE---CH-----CCCCH--     58
```

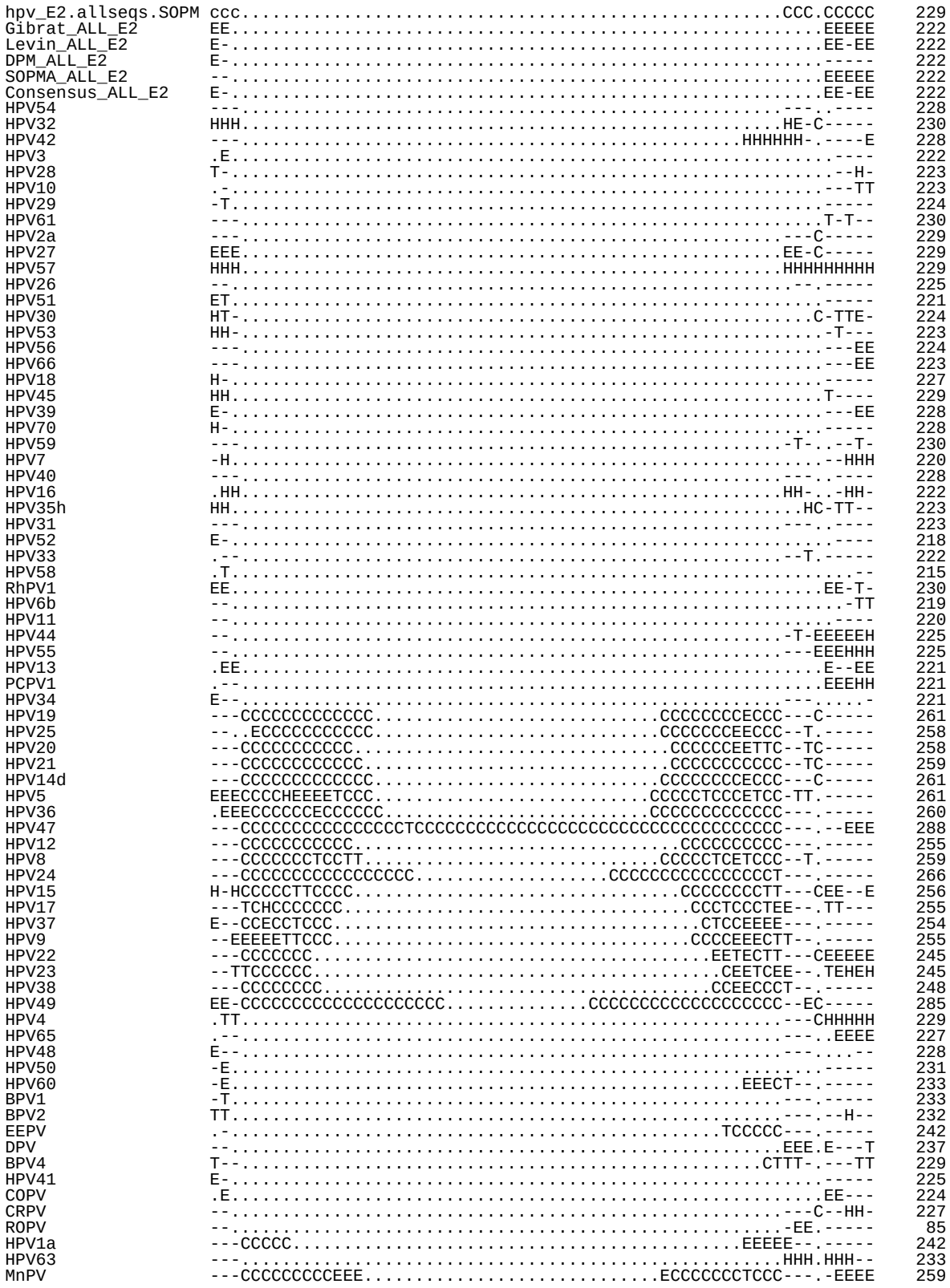
```
hpv_E2.allseqs.SOPM cccccchhhhhhHhhhhhhHhhhcccccccceeechchhhhhcccccccccccceEEEEcc  123
Gibrat_ALL_E2       --HHHHH-----------------HH---------HH-H---EEH-------------------  123
Levin_ALL_E2        ---E-H---C---------------HHT--S--C--STT-EECE---SSS--TTTT-C-------  123
DPM_ALL_E2          E----EE-------------------T------CCC-C-C---E--------TT-----------  123
SOPMA_ALL_E2        E---EHH-----------------HHHHHEE--CCHH-H-----HH--HHHHH------------  123
Consensus_ALL_E2    -----HH-----------------HH------CCH--H----E--------T------------  123
HPV54               -TTE-E-CTTTCC-----------HHHEEE---CCCEC---E-C-----------T-C------T  123
HPV32               EEE--H-CC---------------TH--EE-ETT-T-TH-----HH--HHEEE-TT-------E-  123
HPV42               EEEEHHHCCCCC-E----------HHHHH-----CHH-H----EE--T------TT-------ET  123
HPV3                E------CCCCC---EEEEECCCC---------CCC-C-CCC---H------T-T------CC--  123
HPV28               E----E-C--C----EE-------H--HH-TTTC-H-CH-----HH------TTTEE------TT  123
HPV10               E-T-EEECCCCCC-EEEEEE----HHHHH---TH-H--H---T----T-------EE------TT  123
HPV29               EEE-EE-CCCCCC------------HHT---T-H-HH-H-----HH--HHHHTTTTE------TT  123
HPV61               EEEEEEECTTCCC-EEEE---TCC-TT-E----H--E-H--------T---H--TT-------E-  125
HPV2a               ---EEE-CCCC----EEEEEECCEE--H---T-CCC-C-CEEEE-------H-HTT-------TT  123
HPV27               --HEEE-T----------------HHHHH-TT---------E-------------E-----CE-  123
HPV57               E-EEEEH--------EEEEE---EE-HH--TTTCCH-C---EEE---------TTT-------TT  123
HPV26               ----EE-CCCCCCCEEEEEEEC-EE--------CCC-C-CEEEEEE--------TT---------  123
HPV51               -----E-CCT------------CC-TTT----HH-H-CEEEEEEEE-TT-------E------ET  123
HPV30               HHHHHH-CCCCCEE----------H---EEET-HHHH-H-----T---------TT-------ET  123
HPV53               EHHHHHHCCC---------------HHH--HTHHHHH-H----CET----EE-T-T---------  123
HPV56               HHHHHH-CCC-------------C--EEE--HHHHHH-H--EEE---T----ET-T--------T  124
HPV66               -TTHHH-C----C----------HHHHH-HHHHHHH-H-----T---T-----HT-------E-  123
HPV18               EET----C----------------HH----THHHHHH-H---------T--EET-T---------  127
HPV45               E-T----CCCC------------CH-TT---THHHHH-H-----H------HHTTT-------ET  129
HPV39               -E----TTCC----------------------TH---HT-C----H---------TTE------T-  128
HPV70               -------CCC---------------HT---HTE----CH-----H--TTT----TTE------T-  128
HPV59               EEEEEE-CCT--------------HHHHHHHHHHHHH-H-----HH--H------TE--------  127
HPV7                -TT--HH---------------CC--------E---H------EHH-HHH---TT--C-----E-  123
HPV40               EET-EHH----------------CC-T----EEE-----C-C----H---H----TTT------E-  123
HPV16               EEEEEE-C-----------EE-EE---T--HTHC-H--H-----HH-TT--ETHTT-----C-TT  123
HPV35h              -----HH-----------------------TE---HT-C-E--E------E---EE-----CTT  124
HPV31               E-E--H-C----------------HHTT----H--C--H-----H--HHHHHHHT-----HHC-T  123
HPV52               E-EEEEECCC--------------HHTT---HE--HHTH------H-THHHH-TT---------T  123
HPV33               -HTHHH-CCCCC------------HT----HHHHHH--H----EE-----EE-TT----------  123
HPV58               -HHHHH-CCCEEEEE--------CT--T-------C--H-----H---HHHH-TTTE--------  123
RhPV1               -TT--HH----------------C-TTT---TH---E-HEEEE-H--TT----E-T-C-----E-  124
HPV6b               --TE-E-CCCCCC---------T----------CC---H----C----------TT------TT  123
HPV11               --T-EE-C-TCCC---EEEE-CCC----EEE-E----C-C--EEE---------T-T--------T  123
HPV44               -TTEEEECCTC-------------H------T----T--C--TC----EEEEE-TT--------T  123
HPV55               EET----CCCCTC-EEEEEE----H---E--TT--H---C----E----EEEETTT-------ET  123
HPV13               -----EEC-C--------------HH-HHHHEH--HETTCCEEE---------T---------ET  123
PCPV1               --T----CCTC---------EEEEEE---E-T-----TT-C---E--------TTT-------E-  123
HPV34               EEEEEE-CCTTCC----------C-T----HT-----T-C----HH-HTH---TT--------E-  124
HPV19               --EEEH---C------EEEE-CCC-----T---C--EC-CCC--H------------C-----E-  123
HPV25               --HEEH----CCC---EEEECCCC-----T---C--EE-CCCCC-------------C-----E-  123
HPV20               -EEEEHH--C--C---EEEE---C-----T---HH-E-HC---C-------------------E-  123
HPV21               ---------CCCC--EEEEE-CCC-----T---C--EC-CCCCC---------------------  123
HPV14d              --EEEH---CCCC--EEEEE-CCC---------C--EE-CCCC-HH--H----------------  123
HPV5                --HHHH----------EEE--CCC-HHT-T------EE-CCCTT-------E-----------E-  123
HPV36               ---EE-----------EEEECCCC------HH----EE-CCCTC-------E-------------  123
HPV47               --HHHHH--------EEEEE-CCC-----E---C--EE-CCCCC---------------------  123
HPV12               --EEEE---CCCC--EEEEE--CC-HHH---TTC--EE-CCCCC---------------------  123
HPV8                ---EEHH---------EEEE--CC---HTT---C--EE-CCCCC---T---H-----C-------  123
HPV24               ---------C-----EEEEECCCC----T-------EE-CCCCC---------------------  123
HPV15               ---------C-------EEE----HTHTT----C--H-H-----H------E-T---C-----E-  123
HPV17               -------C-CC---------------T-TT---------H-----H------EE---TC-----E-  123
HPV37               ---E--H---C----EEEEE--CC-T------E---EC-------------E-----C-----E-  123
HPV9                ---------------EEEEEE-------TT---CC-EC-----C-----EEE-----------E-  123
HPV22               --EEE-------------------TT-------C--E-----TC-------EE----C-----E-  123
HPV23               -------CCC--------------TT-------C--E--C---T-------EE--TT------T-  123
HPV38               ---HHHH-----CC--EEEE---C-------T-CHHHC-----C-------------C-----E-  123
HPV49               --------CC------EEEEECCC---------C--EC-C-CCC--------------------T  123
HPV4                --TEEHHC-TC----E---------H-----HHHHHHH-H----------H-HHTT---------E-  123
HPV65               --EEEHH--CC--CCEEEEEEC-EE------H-CCCH-H-------------HH----------  123
HPV48               -TEEEEEE------EE--EE----HHHHHHHHHHHHH--C-------TT-------E------E-  128
HPV50               -------CC---------------HT----T--CCHH-H----------------EE------E-  129
HPV60               --T---H--------------CC----HHHHHH-HHHTH---EEE------T----E------E-  125
BPV1                ---EEE-C---TT-----------H-HTTT---C--EE-C----H-------TTT-E-----T-T  123
BPV2                ---EEEE-----------------H--TTT---C--EE-CCCCC--------TTT-E--------  123
EEPV                -------CCCCCCCEEEEEEE--C---------CC-ECHCCCCC---------TT-E-----C--  125
DPV                 ---------C-C--CEEEEECT-C--T------C-HH--CCCCE---HHHHH----E--------  123
BPV4                E--------CC-----EEEE---C-HHHHTHHHH--E-TC--------T------T--------T  123
HPV41               --HEEEH---C-------------H---HHHTTHHHH-T-----HH-THHHHHHHT-CC----E-  126
COPV                --T---------------EEECEEE-TH--T--C--EC-CCC--HE---EEEE-TT------C--  123
CRPV                --HEH--C----------------------T-----EC-C----------------------ET  123
HPV1a               --HHHHH---------EEEE-CCC-----T---CCC-C-TCEEEE---TT----TT---------  123
HPV63               EHT----CC--------EE-CCTC----H----CCH---C-EEEE--------T-T-------ET  123
MnPV                -------CCCCCCCCEEEEEECCC------HHHHHC-C-CCEEEEE--------TT---------  123
```

E2 Appendix B




hpv_E2.allseqs.SOPM ...ccceeeeettccee....ee..cCCCcCC...........CcCCCcc....cccC..cccc 218
Gibrat_ALL_E2 ...---------EC-E--.....E------...........EE-EEEE....EEEE..EEEE 215
Levin_ALL_E2 ...TTTCCH--E--E--.....------..........----------....-E-S.----E 215
DPM_ALL_E2 ...T--C----EC-E--.....------E-...........----HHE....E---.---- 215
SOPMA_ALL_E2 ...---C----EC-E--.....E------...........---EEEE....EEE-..---- 215
Consensus_ALL_E2 ...---C----EC-E--.....------...........----EEE....EE--.---- 215
HPV54 ...T-TC----E-TE--.....E------...........HHHHHEEE...-TT-..---- 218
HPV32 ...EEE---CT-C-E--.....EE-----...........TEE----EH..-HTT..-HHH 218
HPV42 ...--E----CCC-E--.....EET----...........---.-.---.---E.---T 213
HPV3 ...ETTC----E-----...-C..------...........---..--.--.----.---E 217
HPV28 ...-T-C----E-TE--...-T..EEE-TE-...........EEE..EE....TT--..-E-- 217
HPV10 ...---CHH-HE-TE--...--..----TEE...........--E..E-...----..---- 217
HPV29 ...E-T-----E-T---...-T..T----EE...........EHH..-----.-T--.---- 217
HPV61 ...ET------EEEE--E..CCT..-----HH...........HHH.HHH....HHHH..HHT- 222
HPV2a ...E-E-----EE-E--...-C..--.-E--...........---.--.---.----.---- 217
HPV27 ...-EE-----EETE--...--..---.HEE-...........-TT.-E-....HHH-..T--H 217
HPV57 ...EEE------C-E--...--..EH.HEH-...........-HH.HHH....HT--..EE-E 217
HPV26 ...--E-----E-T--C.....--T-E--...........-------C...----.---- 216
HPV51 ...T.H-----E-T-CC.....---H---...........-------...-----..EHHH 214
HPV30 ...ETTCH---EC--C-.....E-T----...........-EE---...---H-..HEEH 215
HPV53 ...TE-C----E-T-C-.....E--T---...........-EE---...HHHH..HHHH 215
HPV56 ...EEE-----ECT-C-.....E-TT-H-...........---E---...--HH..HHHH 216
HPV66 ...EEECTTHH-HHH--.....EETT---...........---EE--...-T---.-HHH 215
HPV18 ...T--C----------....CCC.T------...........--H...-...-HHH..HHHH 220
HPV45 ...---C------T--C....CCC.------...........--H...H....-HHH..HHHH 222
HPV39 ...-TT-H--T------...CCC.-EEEE--...........---...-...-EEE..EEEE 221
HPV70 ...-TE------CTT-----..CC.-EEEE--...........---...-...-HHH..HHHH 221
HPV59 ...-EHHH--T-CTTHH....HHH.HT-----...........---...E....-H--..--TT 220
HPV7 ...---CC---E----C.....--.-TTE...........---.-EE...---------- 213
HPV40 ...---CC---E--E--.....--T-THE...........H-T--EEE...--T-CC---- 218
HPV16 ...T-E--------E--.....E-T--EE...........---.--.---H..HEHH 213
HPV35h ...--E-----E-T---.....E----E-...........-HH.--.--.-HHH..HHHH 214
HPV31 ...TT-----C--TE--.....EEETEEE...........---.--.---HEE..EEE- 213
HPV52 ...EEE-----E-TE--.....E------...........---.--.---EE..EEEE 212
HPV33 ...T-HHH-----T---.....E------...........---.--.---.----.---- 212
HPV58 ...T-E-------T---.....------...........T-...-...----..---T 212
RhPV1 CCC--E---HHH-THH-.....C..HT.----...........-HHTHT-C...-EEE..EEEH 223
HPV6b ...T-----C-E-TE--.....---T---...........---.-EE...----..---E 214
HPV11 ...-T-C---CE-TE--.....--TT---...........---.-EE....EE--..T--E 214
HPV44 ...----H---E-T---.....--TT---...........---.--H....ETT-..---- 214
HPV55 ...T-HHHCC-E-TT-C.....------...........-HH.--E....E---.---- 214
HPV13 ...-EE--C--E--E--.....---T---...........--E.HEE....EEE-..---E 214
PCPV1 ...TTTTC---E-T--C.....--TTT--...........-HH.EEE....EE--..---- 214
HPV34 ...ETT----C--T---.....EE.EEEE...........---..T-....---T..EEE- 214
HPV19 ...---CC-------C-E..ECCC.-------...........-------C...----CC---- 224
HPV25 ...T--CCCCC----C-E..E-CC.-------...........-------C...----CC---- 224
HPV20 ...---CC-C-----C-E..E-CC.-------...........-T-----C...--T-CC-EE- 224
HPV21 ...---CCCCC----CCE..CCCC.-------...........-------C...----CC---- 224
HPV14d ...T--CC-------C-E..E-CC.-------...........-------C...----CC---- 224
HPV5 ...---CCHHH------E..E--C.----T--...........-------C...----CC---E 224
HPV36 ...T--CCCCC----C-E..E-CC.-------...........-------C...----CC--EE 223
HPV47 ...T--CC-------C-E..ECCC.-------...........-------C...----CC--EE 224
HPV12 ...T--CC-C---T-C-E..ECCC.-------...........-------...----CC---- 223
HPV8 ...T-TCC-------C-E..E--C.----T--...........-------C...----CC---- 224
HPV24 ...---CC-------C-E..E-CC.-------...........-------...----.---- 221
HPV15 ...TTT-----ECTTC-E..CC-C.--.----...........------HC...--T-CCT-HH 223
HPV17 ...---HC---EC-T--E..EC-CC-----T-...........T------C...----CC---- 225
HPV37 ...--TTC---EC-T--E..CC-EC-------...........----HHEEE..EEE-CCE-EE 226
HPV9 ...-TE----C------E..E--C.-------...........------EEE.EEET-CC---T 226
HPV22 ...-TT-----E-T---E..ECCC.-----EE...........E-----T...T---.---- 219
HPV23 ...TEE-----ECT-C-E..ECCC.-------...........-------.....-HHTTEEE 221
HPV38 ...--E-C-CC----C-E..ECCC.-------...........---.---....--H-CC---- 222
HPV49 ...---CC---EC--C-E..EC-C.-------CCCCCCCCCCCC---.---....----CC--EE 234
HPV4 ...HH------EEH-CC.....--.EEEE...........--------....-TT-CC---- 218
HPV65 ...-TT-----EE-E--.....EE.EE--...........-------....-EE-CC---- 218
HPV48 ...T-E-----E---C-.....EE.----...........---..--C...--TH..HEEE 220
HPV50 ...ETTC-C--E-T-CC.....-T.T---...........---HHH-.....-T--CC---- 224
HPV60 ...--E-------E---.....E-.----...........------C...E---..TT-- 219
BPV1 ...E--C---T-------...E--..------...........---TT-EE...EE--TC---- 223
BPV2 ...E--CCCCT-------...E-H..H------...........------E....EET-CC---- 222
EEPV ...T--C-CCC--.-C-...E--E.------...........-----EEEEEE--TT..---- 227
DPV ...EEE-----C--TC-....--.----T--HHHHH...EEECTT-.--T.....-.C---- 227
BPV4 ...TT-CCH--ECHE-C.....-THEEE...........EEE...EE...----CTT--T 216
HPV41 ...ET------H-TE----------T---E...........EE-.---.....-----..EEEE 218
COPV ...--------CC--CC....CCC.--.----...........----TT-.....--T-..T--E 218
CRPV ...-TTCC---ECH-CC.....--.----...........----.-E-...---T.C---T 216
ROPV ...-T-CC---E--TCC.....EEE----...........------EEC.C----CC---- 75
HPV1a ...-TT-----E-T-C-...E--E.------...........-T---EEE...TT--..T--- 222
HPV63 ...TT-C----E-T---...E---.E-.T---...........-------C...----CC--T- 222
MnPV ...EEE--CCCCC----EEEE-CC.------...........------EE...E---..---- 224

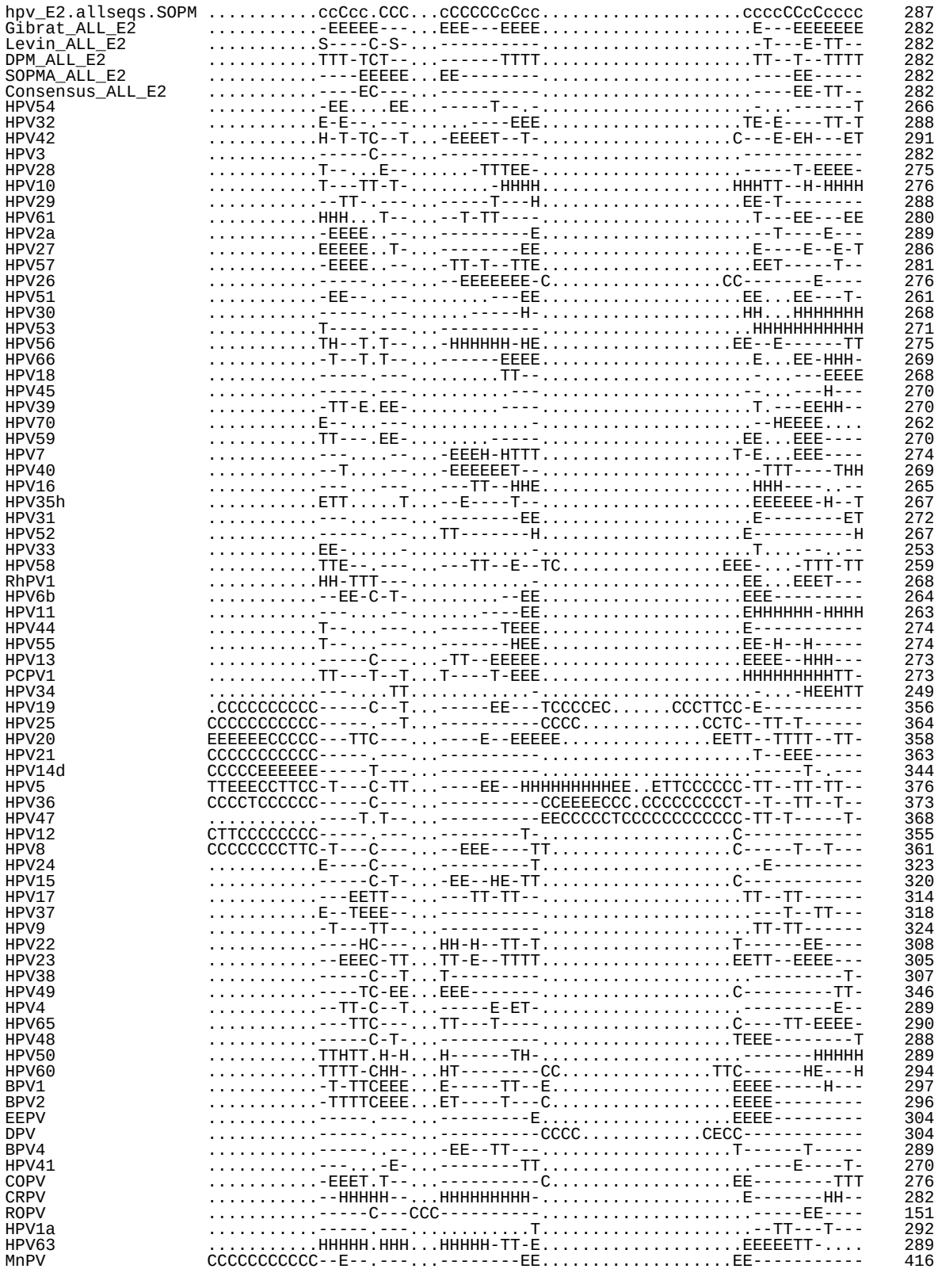
E2 Appendix B



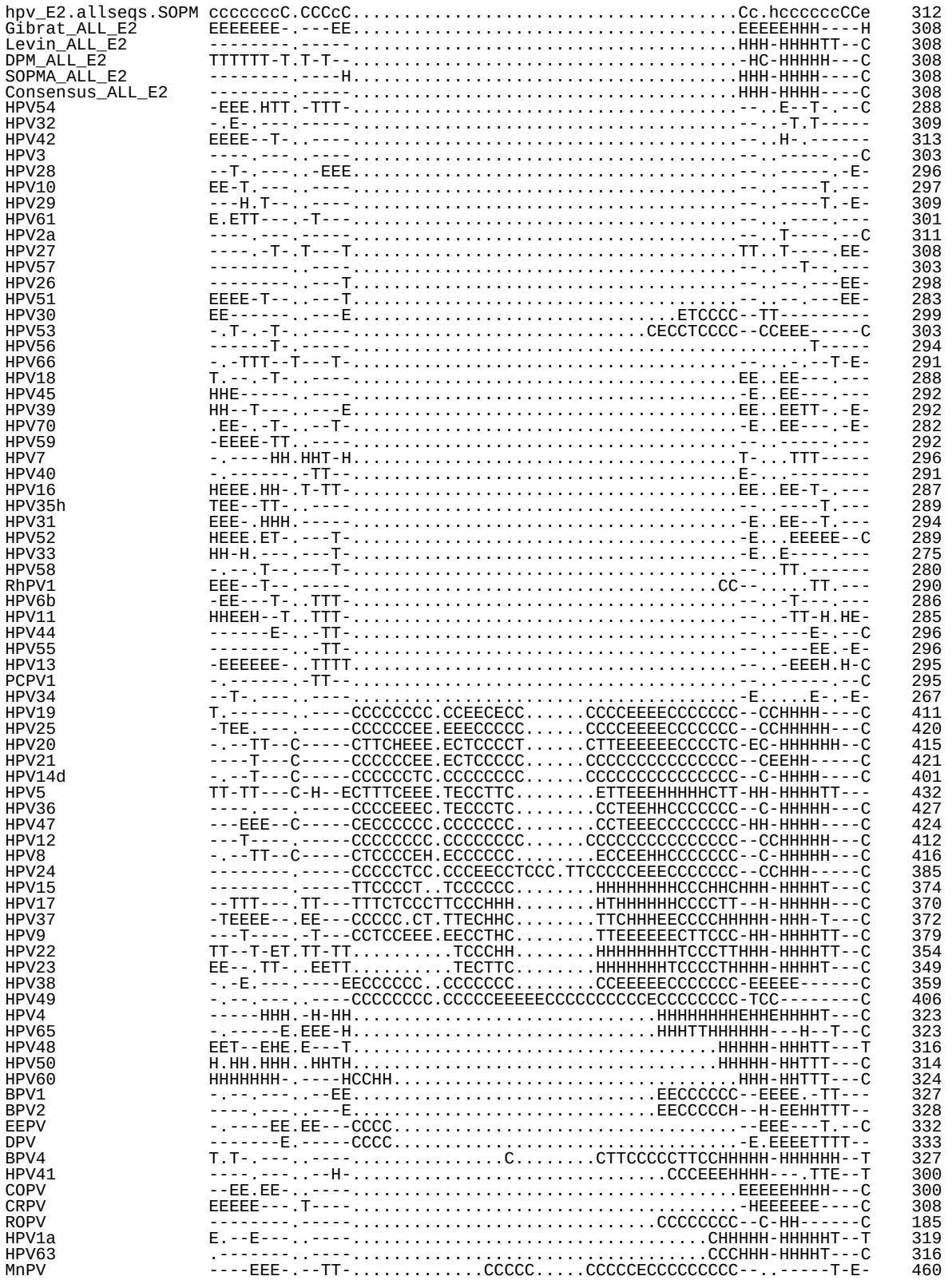
```
hpv_E2.allseqs.SOPM CCcc...cc...CCCCCC.CCCCC....CCCccCC........ccCC.................   257
Gibrat_ALL_E2       EEEE...EE...-------.-EE-EE..EEEE----.........--E-.................   252
Levin_ALL_E2        ----...-...------.-----C..C-------.........----.................   252
DPM_ALL_E2          --HH...-T...-T--T-.-----C..TTTTT---.........---T.................   252
SOPMA_ALL_E2        E---...-...------.-----C..C-------.........----.................   252
Consensus_ALL_E2    ----...-...------.-----C..C-------.........----.................   252
HPV54               ---...-...-...-..-.....................---.....................   244
HPV32               ---...-E...-------.T---EEEEC----T--.........---E.................   261
HPV42               EEE....EECCC--HH--.---EE...E--E----.........---H.................   259
HPV3                ---...-...-----..-----.....--HEEEEEC.....C--TT.................   251
HPV28               --H....HH...E--T-..----T....HHHHHHHH.......HHTT.................   250
HPV10               ---....HH...HHHH-...--HHH....HHHHHHH.......HTHT.................   249
HPV29               ---...-...------..-HHHH....-HEEEEECCCCCCCC----.....................   258
HPV61               --EE...EEC..------CEEE--.....................T--H.................   253
HPV2a               ----C.C--...-----C---E-...C-------.........----.....................   261
HPV27               ---H...-...------.--EEE...C-------.........----.....................   258
HPV57               HHH....E-...-...-.---HH....H------.........----.....................   253
HPV26               --E....E-...-----..-----.....................----.....................   244
HPV51               ------.EE...----HHHHHTH-.....................H.....................   240
HPV30               TT--...-E...E------.HHHH-.....................TT--.................   245
HPV53               ---....-E...EEEEE-.-----.....................----.................   242
HPV56               EE-...---...TT--HH.HEH--.....................TT--.................   244
HPV66               EEE....EE...E---HH.HHT--.....................-T--.................   243
HPV18               ---E...EE...-----E.E---T.....................T---.................   248
HPV45               -TT-...-...------.-----.....................----.................   250
HPV39               ---...-...-...-..-----.....................----.................   248
HPV70               --HH...-...-...-..----T.....................---E.................   248
HPV59               -EE...---...T...T-C-TTT-.....................----.................   248
HPV7                -HEH...---...--HH-H.HH--E...EE---T--.........----.................   249
HPV40               ---....E-...-...-.-----.....................---.................   243
HPV16               HHEE...EE...----H..---EE.....................--.................   240
HPV35h              ---E...EE...ET---..-----.....................H-E.................   242
HPV31               ---...-...E-EE--C--E--.....................----.................   245
HPV52               ---E...EE...EE---..-----.....................T---.................   238
HPV33               HHHH...---...------.-----.....................HHHE.................   243
HPV58               HHH.....-...-...-.---HH.....................H.....................   228
RhPV1               ---...-...----H..HH---.....................--HH.................   249
HPV6b               -EE....E-...--T---.-H---.....................----.................   239
HPV11               --E....E-...-----T.T----.....................T---.................   240
HPV44               TTT-...-...----HH.HH---.....................----.................   246
HPV55               --TT...-T...-------.-----.....................----.................   246
HPV13               -EEE...EE...E-HHH-.T----.....................----.................   242
PCPV1               HHHH...H-...-HHHH-.--T--.....................--HH.................   242
HPV34               -HHH...HT...-...-..-----.....................-...................   235
HPV19               ---...-...-----..-----C..C---E---C......----CCCCCCCCCCC.......   301
HPV25               ----...--C..------C-----C..C-------.........----CCCCCCCCCCCCCCCCCC   308
HPV20               ---...-...------C-----C..C-------.........--H-CCCCCCCEEEECCCCCC.   306
HPV21               ---...-...-----..-----C..C-------.........----CCEEECCCCCCCCCCCEE   306
HPV14d              ---...-...-----..-----C..C---E---.........----CCCCCCCECCCCC.....   303
HPV5                ---...-...----T..T----C..C----TT-EEEE.CCC--T-CCEEEEEEEEEECCCC..   312
HPV36               ---...-...------.-----C..C-------C.......----CCCCCCCCCTCCCEEE.   307
HPV47               EE-...-...------C-----C..C-----E-........E-E-.....................   318
HPV12               ---...-...-----..-----...---------.....----EEECEEECCCCCCCEEEE   299
HPV8                ---...-...-----..-----C..C-EEEEEEECC.EEE----CCHEEEECCCCCCCCCTT   311
HPV24               ---...-...-----..-------...C------.........----.....................   293
HPV15               -H-T...TT...T-HHH..HHHHH...C-H-TT--HE....EET---.....................   288
HPV17               ---...--C.C-...-..--TTTT..C-EEEEE-.....----.....................   283
HPV37               -H-T...TT...T...T..HEEEHC..CE--T---CHEECCCC----.....................   288
HPV9                ---....--C..--T---CT-HHHH..C-------CCCCTCCC----.....................   294
HPV22               EHHH...TTT..TT-TH-HH-H--...HEHHH--H........TTTT.....................   276
HPV23               HHTT...TT...T..HHEEHHTHH...HHHHHH-H........TTTT.....................   273
HPV38               ---....-T...---T--C-----C..C------.........----.....................   277
HPV49               ---...-...-----..-----C..C-------.........----.....................   314
HPV4                ---....HH...------C------...C---TE-E........EEE-.....................   258
HPV65               E---...-...----T-.-----CC.CTT---EE........EEE-.....................   258
HPV48               -TT...-...------C-------...C----HH.........-HT-.....................   256
HPV50               --T...-...---T--H-----...---EEEE........E-TT.....................   259
HPV60               T--...-...-...-C-------...C--HHEE.........HHH-.....................   258
BPV1                -HEE...ET...-T----EEEEEEE..E-TTTT--.........-EEE.....................   264
BPV2                HHHE...E-...------E-EEEEE..E-------.........-EEE.....................   263
EEPV                ----CCC--...-----..-----E..C------.........----.....................   273
DPV                 ----C..EE...HH---..-----C..C------.........----.....................   266
BPV4                ---...-...----TTT-----C..C------.........-T--.....................   259
HPV41               ---...-...HHH---C-----.....................T.....................   243
COPV                --T...-...TT---T.T--E-.....................TT--.................   244
CRPV                ---...-...-...--.EEEEE...---T--TT.....----.....................   251
ROPV                ---...-...-.----C------...C---T---CCE...CC----.....................   118
HPV1a               --E-...-...-T----.-----T..T-------.........-TT-.....................   272
HPV63               T--...-...------HC----TT..T-------.........---E.....................   263
MnPV                ----...-...------E------...C--H----C......C----CCCCCECCCCCCCCCTC    309
```

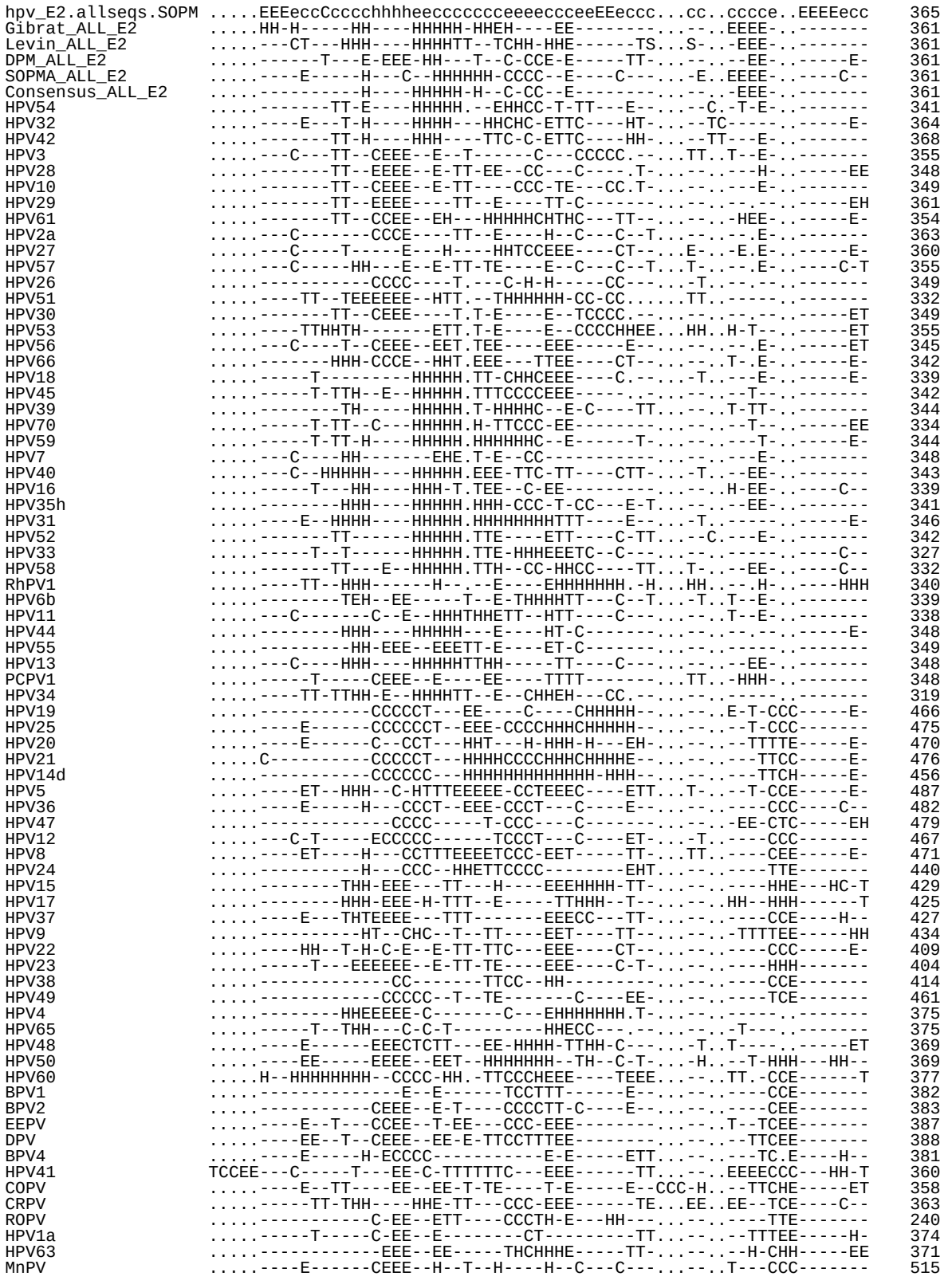
E2 Appendix B

```
hpv_E2.allseqs.SOPM ..............................................................  257
Gibrat_ALL_E2       ..............................................................  252
Levin_ALL_E2        ..............................................................  252
DPM_ALL_E2          ..............................................................  252
SOPMA_ALL_E2        ..............................................................  252
Consensus_ALL_E2    ..............................................................  252
HPV54               ..............................................................  244
HPV32               ..............................................................  261
HPV42               ..............................................................  259
HPV3                ..............................................................  251
HPV28               ..............................................................  250
HPV10               ..............................................................  249
HPV29               ..............................................................  258
HPV61               ..............................................................  253
HPV2a               ..............................................................  261
HPV27               ..............................................................  258
HPV57               ..............................................................  253
HPV26               ..............................................................  244
HPV51               ..............................................................  240
HPV30               ..............................................................  245
HPV53               ..............................................................  242
HPV56               ..............................................................  244
HPV66               ..............................................................  243
HPV18               ..............................................................  248
HPV45               ..............................................................  250
HPV39               ..............................................................  248
HPV70               ..............................................................  248
HPV59               ..............................................................  248
HPV7                ..............................................................  249
HPV40               ..............................................................  243
HPV16               ..............................................................  240
HPV35h              ..............................................................  242
HPV31               ..............................................................  245
HPV52               ..............................................................  238
HPV33               ..............................................................  243
HPV58               ..............................................................  228
RhPV1               ..............................................................  249
HPV6b               ..............................................................  239
HPV11               ..............................................................  240
HPV44               ..............................................................  246
HPV55               ..............................................................  246
HPV13               ..............................................................  242
PCPV1               ..............................................................  242
HPV34               ..............................................................  235
HPV19               ..............................................................  301
HPV25               ..........................................................CCCCCCC  315
HPV20               ..........................................................CEEEE  311
HPV21               ECTTC.....................................................CEEEEECCCCCC  323
HPV14d              ..............................................................C  304
HPV5                ...........................................................TCCC  316
HPV36               ..........................................................ECCCC  312
HPV47               ..............................................................  318
HPV12               ECCC.....................................................CCCCCCCCCC  313
HPV8                ..........................................................CEEEEE  317
HPV24               ..............................................................  293
HPV15               ..............................................................  288
HPV17               ..............................................................  283
HPV37               ..............................................................  288
HPV9                ..............................................................  294
HPV22               ..............................................................  276
HPV23               ..............................................................  273
HPV38               ..............................................................  277
HPV49               ..............................................................  314
HPV4                ..............................................................  258
HPV65               ..............................................................  258
HPV48               ..............................................................  256
HPV50               ..............................................................  259
HPV60               ..............................................................  258
BPV1                ..............................................................  264
BPV2                ..............................................................  263
EEPV                ..............................................................  273
DPV                 ..............................................................  266
BPV4                ..............................................................  259
HPV41               ..............................................................  243
COPV                ..............................................................  244
CRPV                ..............................................................  251
ROPV                ..............................................................  118
HPV1a               ..............................................................  272
HPV63               ..............................................................  263
MnPV                CCCCCCCCCCCCCCCCCCCCCCCTCCCCTCCCCCCCCCCCCCCCCCHHHHHHHHHCCCEEECCCT  374
```


hpv_E2.allseqs.SOPM
Gibrat_ALL_E2
Levin_ALL_E2
DPM_ALL_E2
SOPMA_ALL_E2
Consensus_ALL_E2
HPV54
HPV32
HPV42
HPV3
HPV28
HPV10
HPV29
HPV61
HPV2a
HPV27
HPV57
HPV26
HPV51
HPV30
HPV53
HPV56
HPV66
HPV18
HPV45
HPV39
HPV70
HPV59
HPV7
HPV40
HPV16
HPV35h
HPV31
HPV52
HPV33
HPV58
RhPV1
HPV6b
HPV11
HPV44
HPV55
HPV13
PCPV1
HPV34
HPV19
HPV25
HPV20
HPV21
HPV14d
HPV5
HPV36
HPV47
HPV12
HPV8
HPV24
HPV15
HPV17
HPV37
HPV9
HPV22
HPV23
HPV38
HPV49
HPV4
HPV65
HPV48
HPV50
HPV60
BPV1
BPV2
EEPV
DPV
BPV4
HPV41
COPV
CRPV
ROPV
HPV1a
HPV63
MnPV

E2 Appendix B



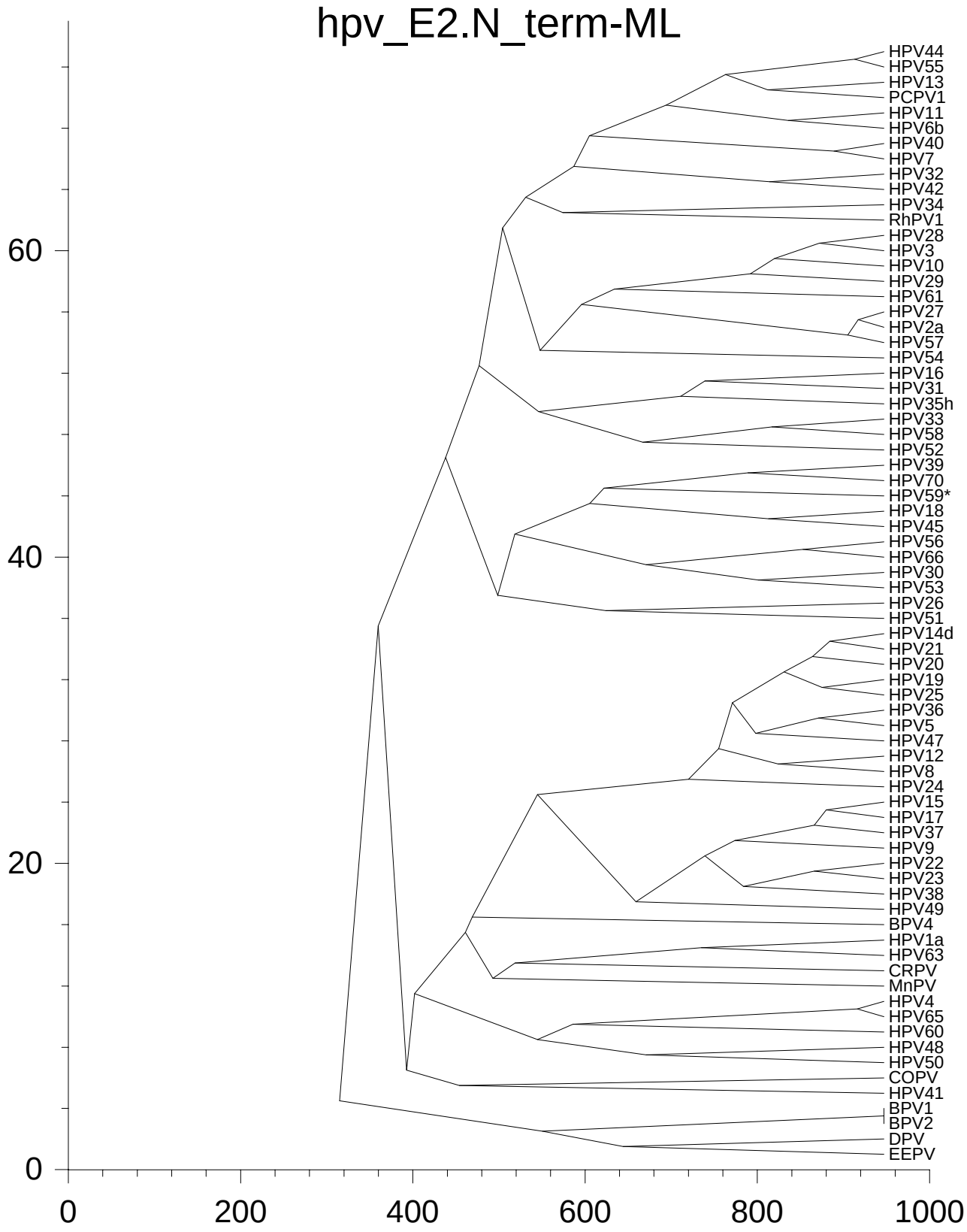

hpv_E2.allseqs.SOPM
Gibrat_ALL_E2
Levin_ALL_E2
DPM_ALL_E2
SOPMA_ALL_E2
Consensus_ALL_E2
HPV54
HPV32
HPV42
HPV3
HPV28
HPV10
HPV29
HPV61
HPV2a
HPV27
HPV57
HPV26
HPV51
HPV30
HPV53
HPV56
HPV66
HPV18
HPV45
HPV39
HPV70
HPV59
HPV7
HPV40
HPV16
HPV35h
HPV31
HPV52
HPV33
HPV58
RhPV1
HPV6b
HPV11
HPV44
HPV55
HPV13
PCPV1
HPV34
HPV19
HPV25
HPV20
HPV21
HPV14d
HPV5
HPV36
HPV47
HPV12
HPV8
HPV24
HPV15
HPV17
HPV37
HPV9
HPV22
HPV23
HPV38
HPV49
HPV4
HPV65
HPV48
HPV50
HPV60
BPV1
BPV2
EEPV
DPV
BPV4
HPV41
COPV
CRPV
ROPV
HPV1a
HPV63
MnPV

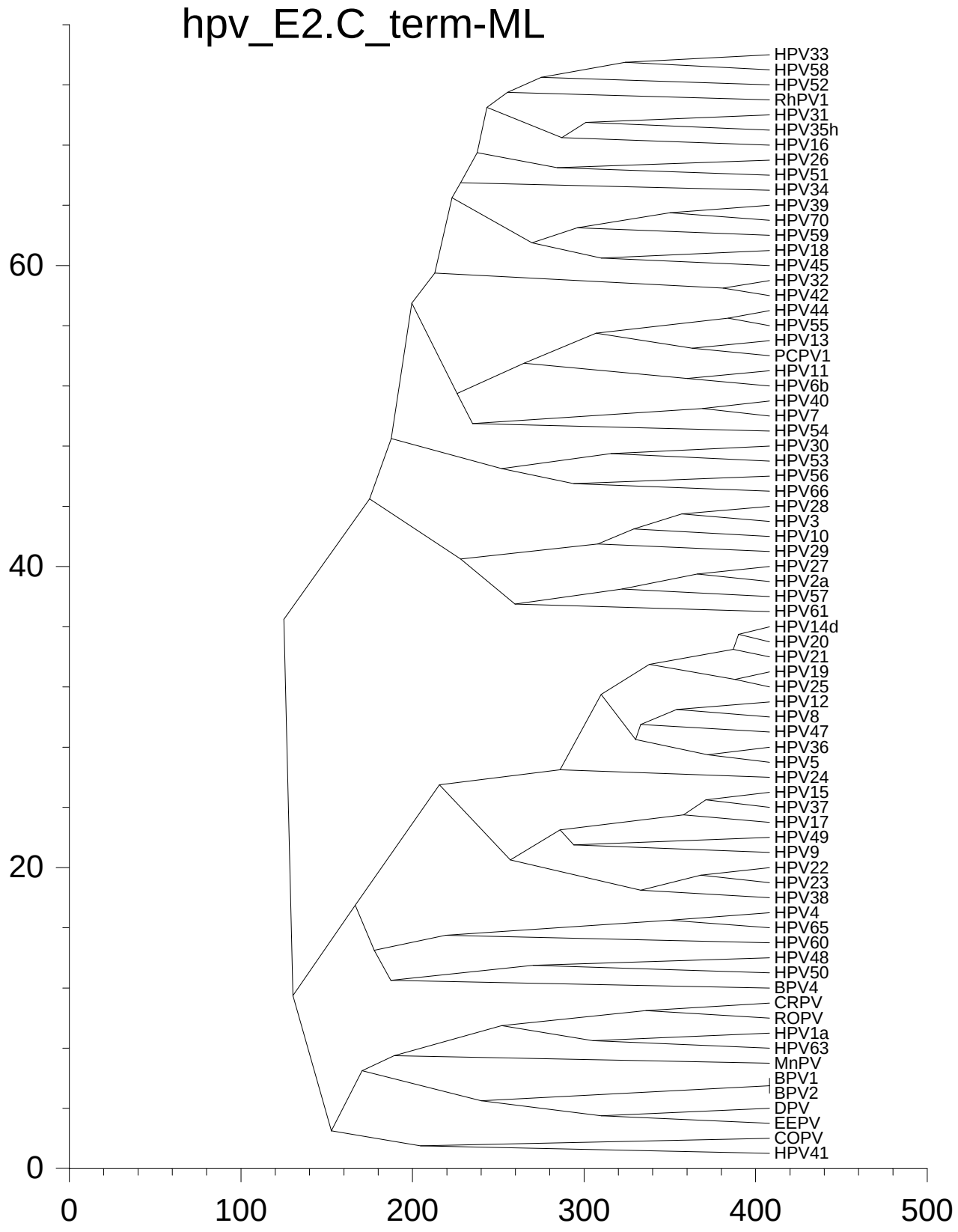
E2 Appendix B

```
hpv_E2.allseqs.SOPM cccccceeeeeccccccceeeccc...chc 393
Gibrat_ALL_E2       -HHHHHHHHHHE-----EE--CEE-...EC- 389
Levin_ALL_E2        -----HHHHH----S--EE--T--T...TTS 389
DPM_ALL_E2          H--HH---C--EE---EEE--C---...-C- 389
SOPMA_ALL_E2        ---EE-HHHHHE------E---EE-...-C- 389
Consensus_ALL_E2    ---HH-HHHHHE-----EE--C---...-C- 389
HPV54               --TT-E-----E--TEEEE---EE....-E- 368
HPV32               T--TTE----------------EEETCCT-T 395
HPV42               HHHHHHHHHH---TT----CHHHHHHCCTT- 399
HPV3                --------------TT------E--C..-C- 384
HPV28               T--T-------E--T--EE---EE-E..EEE 377
HPV10               ----TE-----E-----TE---EEE...EET 377
HPV29               T--T----HTC---THHHHHHTTEE...EEE 389
HPV61               ---THHHHH--E-----EHHT-EE-H..H-T 383
HPV2a               --TT----------T----CCC---C..H-H 392
HPV27               TT-THH-----E----EEECCTE--H..EEE 389
HPV57               TT---H--------T----CTT-E-T..HCE 384
HPV26               ---T-------E-------C-C--....EEE 376
HPV51               HH-TTH--------T--HHCCCEH....H-E 359
HPV30               -----E-----E------E---E-TTC.-TT 379
HPV53               -HTTH-C-H--------HE---HEEHC.-CE 385
HPV56               T--T----HHHE----EEE-HTT-T       370
HPV66               T-TTT---------T---E----EE...-CH 370
HPV18               -----------E-----EH-HHHH....-E- 366
HPV45               -H-TT------E-----TE-CT-E....EET 369
HPV39               -TT-HH--------TTEEETTT--....-EE 371
HPV70               TT-TTHH-------TT--E-----....EEE 361
HPV59               ---TTE-----------TE-H-EE....EEE 371
HPV7                ----H------E-----HHC--EEE...EET 376
HPV40               -TH--------E-----E----EEE...EET 371
HPV16               TT-TT--HHHH---TEEEE---EE....H-T 366
HPV35h              -T-TT------E--T-EEEC----....-EE 368
HPV31               ----T-HHC--E----T-ET--E-....-EE 373
HPV52               ---THE-----E----TEE---EE....EE- 369
HPV33               ----HH-------------C--EE....EE- 354
HPV58               --HHHE-----------EECTT--....--- 359
RhPV1               HH-HTHH----EETTEEE-C--EE....EEE 367
HPV6b               TT--HH----T--------C--EE-EC.TTE 369
HPV11               ----H----H-------T-C--EEEEC.TTE 368
HPV44               T--HHH----T-------E--C---EE.EEE 378
HPV55               ---HHH----------EEE--TT--EE.EEE 379
HPV13               ---HHH-----------HH---EEEEE.EET 378
PCPV1               ---TT----H---T--HH-HH-EETTT.TCT 378
HPV34               -----HHC--TT------ECCC--....-ET 346
HPV19               ------CCC-----TT---C-H--H...TC- 494
HPV25               ------CCCHH---TT-----H--H...TC- 503
HPV20               ------HHH--E--TT---HHHHHH...T-T 498
HPV21               -------CC-C---TT---CCC-HH...H-T 504
HPV14d              ------HHHHC---TT---CHH--H...TCT 484
HPV5                ------HHHHHH--TT--HHHHHHH...H-H 515
HPV36               ------CC---E--T----CCC---...T-- 510
HPV47               ------CCHHH---TT--HHHHH--...-C- 507
HPV12               ------CCH-----TT--HHHC--H...--H 495
HPV8                T-----CC---E--TT---CTC---...H-- 499
HPV24               -----T-----E--TT---CHHHHH...HC- 468
HPV15               TTHHHH-----E--TT--HHHHHH-...H-H 457
HPV17               --TTEE---TTT--T----CHC---...H-- 453
HPV37               TT-HHE--------T---HHHHHH-...H-H 455
HPV9                H-TT-E---------T--HHHH-HT...--E 462
HPV22               TT----HHHHHHHHEE-EECTTT--...H-H 437
HPV23               TTTHHHHHHHHH-E----E-TTTTT...-TH 432
HPV38               T----------E------E-HC-H-...TCE 442
HPV49               -----------E--T---HHHH---...-C- 489
HPV4                T----E--CCC--TTT------E-T...-CE 403
HPV65               TT----T-CCTE-T-EEE---C---...-C- 403
HPV48               T--TEE-----E---T--E--C--T...--- 397
HPV50               -T-HHHHHHH----T--E----E-H...H-- 397
HPV60               ---TT------E--T---E--C-E-...-C- 405
BPV1                --------------TT---CC-EEHH..--H 411
BPV2                TT-------------T---CC-EE-C..--E 412
EEPV                TT---E----C--------CHHEEEH..HC- 416
DPV                 TT--EE-------------CT-EEEE..ECE 417
BPV4                TTTTT------E--T--EHHTHHHH...H-H 409
HPV41               HHTHHHHHHHHH--TTHTHHCCHHH...HTT 388
COPV                TTT--------E---H--EC-C---...-CH 386
CRPV                ---THH----C---T--HE--HEH-...-C- 391
ROPV                -T-------CC------E---T-T-...-CE 268
HPV1a               H--T-HHH---E--TT-EE--C---...TT- 402
HPV63               TT-T-HHH---E-----EE--TT-E...-CT 399
MnPV                ------C--C-E--T---E--C---...-C- 543
```

Appendix C: Phenograms based on E2 Amino Acid Sequences

Phenetic analysis is a form of cluster analysis that attempts to capture the relatedness of sequences irrespective of evolutionary pathways—that is to say the simple similarity of sequences. In the following two phenograms, E2 amino acid sequences over the NH_2 -terminal (about 200 aa) and COOH-terminal (about 90 aa) regions are compared using the PIMA program as described by Korber et al. (*J. Virol*. **68**:6730–674, 1994). The intervening hinge region, which is highly diverse, has been excluded from these analyses. The PIMA approach employs a hierarchical scoring scheme that allows for conserved substitutions in addition to identities. The abscissa records the similarity scores, whereas the ordinate merely records the number of sequences being compared. Sequences 44 and 55 are closely related (in both stretches of the E2 protein); in contrast, many of the sequences differ by as much as 70% using this scoring method—they are connected by nodes that are a small fraction of the score possessed by identical sequences, for example BPV-1 and BPV-2. The amino acid sequences cluster in these analyses as one would predict on the basis of phylogenetic classification—the groups of the A and B supergroups stay together in both fragments. One exception to this pattern is the RhPV1 sequence, which clusters with A9 PV sequences in general and in the C-terminal fragment but not in the N-terminal fragment. Otherwise, there appear to be no unexpected similarities nor unexpected dissimilarities as are seen with E4 sequences (Doorbar and Myers, Part III, appendix C).


hpv_E2.N_term-ML
HPV44
HPV55
HPV13
PCPV1
HPV11
HPV6b
HPV40
HPV7
HPV32
HPV42
HPV34
RhPV1
HPV28
HPV3
HPV10
HPV29
HPV61
HPV27
HPV2a
HPV57
HPV54
HPV16
HPV31
HPV35h
HPV33
HPV58
HPV52
HPV39
HPV70
HPV59*
HPV18
HPV45
HPV56
HPV66
HPV30
HPV53
HPV26
HPV51
HPV14d
HPV21
HPV20
HPV19
HPV25
HPV36
HPV5
HPV47
HPV12
HPV8
HPV24
HPV15
HPV17
HPV37
HPV9
HPV22
HPV23
HPV38
HPV49
BPV4
HPV1a
HPV63
CRPV
MnPV
HPV4
HPV65
HPV60
HPV48
HPV50
COPV
HPV41
BPV1
BPV2
DPV
EEPV
0
20
40
60
0
200
400
600
800
1000


hpv_E2.C_term-ML
HPV33
HPV58
HPV52
RhPV1
HPV31
HPV35h
HPV16
HPV26
HPV51
HPV34
HPV39
HPV70
HPV59
HPV18
HPV45
HPV32
HPV42
HPV44
HPV55
HPV13
PCPV1
HPV11
HPV6b
HPV40
HPV7
HPV54
HPV30
HPV53
HPV56
HPV66
HPV28
HPV3
HPV10
HPV29
HPV27
HPV2a
HPV57
HPV61
HPV14d
HPV20
HPV21
HPV19
HPV25
HPV12
HPV8
HPV47
HPV36
HPV5
HPV24
HPV15
HPV37
HPV17
HPV49
HPV9
HPV22
HPV23
HPV38
HPV4
HPV65
HPV60
HPV48
HPV50
BPV4
CRPV
ROPV
HPV1a
HPV63
MnPV
BPV1
BPV2
DPV
EEPV
COPV
HPV41
0
20
40
60
0
100
200
300
400
500