

E5 SuperGroup A AA-Aln

most-likely	MINIPV...FVFTG.....TLLFLLCL..ALCVCLVSP.LLL.....SVSVYALVLLLLVL..LLW	47
HPV42	TV.-GLQYC.DST-CGTT.....GQK--L--FIVVGA--VC-.WIS-QNYPYPVWA-CL-SY-T----S-	62
HPV61	T--N-GRACV-DRWMVY.....D-PAAYTF--..FLSF-TKRIFGV....FLCHICRI-VHIHAS.VVV	57
HPV26	HTY---II..IP-I.....VYV-V--VCVCLL-CLL-.---IY-F-AS----F..CF-	50
HPV51	LYRHI-TI..A--I.....I---V---CVC-VL-CLL-.---QY-F-AA---I-.CF-	50
HPV30	LY-NKRH...ILLRLLYIATIVLFSVSVFV---CFVC-----YV-.---A-LFSTC---II..-F-	61
HPV66	SPY-AT...ID-CV.....ICV-A--FCVC---HFV-.---A-LFTSC-I-II..-F-	49
HPV18	-L.SLI....-L-C.....CV-M..YV-.H-.--P.....CMC-YAWV--F..VYI	40
HPV45	-L.SL....-LLC.....-SV---YV-.N-.--VQ.....-Y-C-FAW---F..FI	40
HPV39	--LL....-LVW.....-GV-I..YI-.N-.--P.....-H-C-Y-WII-F..VFI	40
HPV70	--AL....-LVW.....-AV---YI-.S-.--P.....-HLC-YMW--LF..VFI	40
HPV59	--TL....-CC.....VCV---N-.--Q.....-YMC-YTW--F..VYI	40
HPV16	-T-LDT...AST-L.....LAC---FCVL-----LIRP-----T-TSLII-----	50
HPV35h	--DLT...ASS-V.....L-C---FCVL--L--LVRS-----L-SALI-----I--	49
HPV31	--ELNIS...T-SI.....V-C---FCVL-F---IRP-V-----TL---IV..I--	50
HPV52	-EVV---QIA.AAT-T.....-ILPVVI..-FA--IL-IV-IILI.SDFV-YTSV--T-L-YL---	40
HPV33	--.F....-LC.....-I-F...C-S.L-LR-.I.....-I-T--WL-V-----	40
HPV58	--.L-I....-VC.....-I-F...C-.IFLR-.V.....-I-I--WL-V-----	41
HPV6bE5a	-EVV--QI..AAG.....-ST-I-PVII-FV--F--II-IVWI.SEFI-YTSV--T-L-YL---	58
HPV11E5a	-EVV--QIA.AAT-T.....-ILPVVI..-FA--IL-IV-IILI.SDFV-YTSV--T-L-YL---	58
HPV44	-EH--I...DATI-AT.....S-S-LPVVI..-F--F--IV-IICI.SDFI-YTSI--T-L-YL---	58
HPV55	-EH--I...DGTI-AT.....S-S-LPVV..-F--F--IV-IIYI.SDFI-YTSI--T-L-YM---	58
HPV13	-EF---...D-S-QAT.....SKS-LP-VI..-T--V--IITI-CI.SEFL-YTNV--T-I-YV---	58
PCPV1	-ELQV-P...VD--AKA.....TSQS-LP-LI..-T--FL-III-IFV.SEFLLYSSV--T-L-YL---	60
HPV34	--.FI....-IL.....-VFGF..YM-.LS-.S.-AV.....-C---WL---III.TFL	41
RhPV1E5a	TAT-LL....LVVF.....L-----G..VF-.A--ES-----CF---S---M...-F-	46
most-likely	.VVLTSPLQFFLLYLLFFYIPMLLIHFHAI..LLTQQ.	81
HPV42	.LQVLYFDY-F-C-IILG--SV-LTLL-.H-AI-. .	95
HPV61	.YGI-VAPPANHCCSYVNSV-CVCCS-A-. .CIPL..	89
HPV26	F--S--YITTYIV-ICL-F--ACFL--Y.T..VMVIAT	85
HPV51	F--A--Q-TT-FV--I---L-C--L-LY.T..F..LLQ	84
HPV30	F--AS-YITA-TIF-----L-VYA-V..W-INTQ	97
HPV66	F--A--FFDT-I-F-----T-C-YC--L..W-INHL	85
HPV18	..I---ATA-TV-VFC-LL----L-I-----SL-. .	73
HPV45	..I---TA-AV-ICCYLL--FVL-M..A..-H-I-. .	73
HPV39	.LIR-T--EV-FV-----VL--W-L-RL-M..DM.I..	72
HPV70	..H-T--M-CI-----IL--WFL-IL.S..VYA... .	71
HPV59	..I--SYEC---I---I--L--LYA---.S.I-. .	73
HPV16	.ITAA-AFRC-IV-II-V---LF---T--R..F..IT.	83
HPV35h	..TVAT--RC-CCF-C-L---GM-NA-Q..Y-AV-. .	83
HPV31	..IA---RC-CI-VV-I---LFV--T--S..F..S-Q	84
HPV52	..SIG--FKV-F-----L-F--FC--C-Q..Y-A-LQ	75
HPV33	..FVG--KI-FC---L-L--MC-N--Q..HM---E	75
HPV58	..SVG-A-RI-FC--I-L---MC-N--Q..Y----D	76
HPV6bE5a	.LL--T-----T--VC-C-A-Y--YY..VT---	91
HPV11E5a	.LL--T-----T-CVC-F-AFY--IY..VQ---	91
HPV44	.LL--A---Y--T-CVCFE-AWY-----VH---E	92
HPV55	.LL--TA---YV-T-CVCFE-AWYL--Y..VH.T-E	91
HPV13	.LL--T---Y--T-SLCFL-A-CV-QY..-Q.T-E	91
PCPV1	.LL--P-----T-SLCFL-AFC--QY..-Q---Q	94
HPV34	.H-SQ-L-KVYI--VCV-----A-V-Y--T..-QVT..	74
RhPV1E5a	..SIVN-FAA-G-C-FC-LT-L---L--LSVVYSRMM	83

E5 SuperGroups B-E AA-Aln

SuperC.con	M?hp?llLFLGL??gvQLlLl?FlLFFFlv?WD?fgCrC?lp?	34
	<p>hydrophilic domain which hydrophobic region can induce DNA synthesis mediates membrane in quiescent cells for BPV1 -> localization in BPV1 <- -> <- Cysteines which may be involved in dimer formation for BPV1 * *</p>	
GroupC1.con	MPNLWFLFLGLVAAQLLLLLLFLFLVYWDHFECSTGLPF	44
BPV1	-----	44
BPV2	-----	44
GroupC2.con	M?hpgLlLFLGL?fgvQLlLlvFlLFFFfVwWDqfGCrCe????	38
EEPv	-TY.-----T--L--M-----L-----NMQL	43
RPV	-N---F-----T-A-----L-----DGFIL	44
DPV	-N---F-----V-----I-----K--NFHM	44
OvPV1	-Q--F-----LW-----MF-----IF--KY----DKLPV	44
OvPV2	-Q--F-----LW-----MF-----F--KY----KLPV	44
SuperD.con	QTVPIFQRRPGPAEAGAVVFAIAITVTITIAIEGAALQWTRHAATESRETPGRKTKGHPRTGTMPWNTDSAN	73
BPV4	-----	73
	<p>L2 start for CRPV, HPV41 -> <- E2 end for CRPV <- E2 end for HPV41</p>	
SuperE.con	????????????sa????????????ahtqk????????q?Llgk??n?l????fi?hakl?????l????	22
HPV41	MKQKEK-SSNLSRFLKTLGCF-----SCDLCIK-C----GL-A-ILNNC-R---QRGAIVHPLM-	66
CRPV	MGFSDVYACNPF--AFVTQRFFVPINL-----.....VSW-H-HENAG-HHKT--Q----LAIAQ-TYRI	66
GroupE1.con	????????????????????????????????????YV???L?KI??????QT?A?L??KIKL??QQLL???	17
HPV1aR	MEVDFAWTCVHSPCIYSPPTPTYLEACKHCNKCIA--ENA-PP--YTPHAKY--P-H-TF----SI----IKY	73
HPV63E--NDE-LH--FILLVRL--I-P-IY----NK----TRF	39
E5b.con	M?MLTC??NDGDTWL?LWL??AF?V??LGLLL?HYRAV?G??TKC?KC????????DYV?M????GDY?YM	43
HPV6b-E5b	. -MLTCQFNDGDTWLGLWLLCAFIvGMLG--LMHYRAVQGDkHTKC-KCNKHNCND.DYVTMHYTTDGDYIY	70
HPV11-E5b	-V----HL-----F---FT--V-AV-----L-----H-TEK---A--KSNRNTV---Y-SHGDN---V--	73
SuperD.con	S	74
BPV4	-	74
SuperE.con	n????k????????????SLEDLESVVPEVLGVD	40
HPV41	-AMS-LHLLIVY	78
CRPV	-LKT-QLQIKFCSMAALVF-----	101
GroupE1.con	?NM?VWE??WEVW?L?????????VI??SVRVVGLLELLVQLQ	44
HPV1aR	C--A---FF----A-EQPEALEEEL--LP-----	117
HPV63	Y--G---YS----V-VLARVGVAGM--HL	68
E5b.con	N	44
HPV6b-E5b	MN	72
HPV11-E5b	-	74

E5 BLOCKS

ESBLOCKS: In light of the enormous diversity of E5 sequences, the MOTIF Algorithm was applied only to supergroup A sequences. Within supergroup A, two BLOCKS were identified, the second of which was supported by the Gibbs sampler algorithm. E5 similarities have been found to viral proteins such as HTLV p12 (see Halpern and McCance, Part III, this compendium); in a BLAST search using the supergroup A Cobbler sequence (aka HPV66, see below), weak matches were found to cellular proteins, partly due to the low-information content (high representation of leucine) of the E5 sequences.

CONSENSUS-E5.superA	m????????????????f?lcf???l?vC?????l????S?????sv????l1l?l1lW?????p1?	21
HPV51	LYRHIVTIAVFII..LL..-V--LCVC-VL-CLLPL-..L.-Q.....Y-FAAA---I-CF-FVVATSQL	58
HPV18	-L.....S.LI..-LF--CVCMY--CHVPL-..P-.....VCMCAYA-VLVFVYIV	41
HPV16	-T.NLD..TASTT..LLAC-L---CVL-C--LLIRP-..LL-V.....-TYTSLII-V----ITAASAF.	57
HPV31	-I.ELN..I.STVS.IVLC-L---CVL-F--LVIRP-..VL-V.....-YATL---IVI--VIATS--.	57
HPV52	-L...G..L.....-VF--ILLMVF-AVLRP-..LL-I.....-YAQV-V-V----VSIGS-F.	47
HPV58	-I..LP..I.....-VV--ILF-CL-IFLRP-..VL-I.....-IYAWL-V-V----VSVGSA-.	48
HPV6b-E5a	-E.VVPVQIAAGT..TST..I-PVIAFV--FVSII-IVWI-EFIVYT--LVLV---Y----LLLT--.	65
HPV11-E5a	-E.VVPVQIAAAT..TTT.LI-PVIAFA--ILSIV-IILI-DFVVYT--LVLV---Y----LLLT--.	65
PCPV1	-E.LQVVPVDF TAKATSQ.SL-PLLIA-T--FLSIIILIFV-EFLLYS--LVLV---Y----LLLT--.	67
CONSENSUS-E5.superA	???????F?y1?f?y?P?f?ih?????l??qg	33
HPV51TT-FV--I-F-L-CLLL-.LYTF-L.L-	84
HPV18	VITSPATA-TV-VFCFL-MLLL-.IHAI-S.L-	73
HPV16RC-IV-II-V-I-L-L--THARF-I.T.	83
HPV31RC-CI-VV-I-I-L-V--THASF-S.--	84
HPV52KV-FL--L-L-F-M-C--CHAQY-AQL-	75
HPV58RI-FC--I-L-I-MMC-NFHAQY-TQ-D	76
HPV6b-E5aQF-LLT-LVC-C-ALY--.YYIVTT.--	91
HPV11-E5aQF-LLT-CVC-F-A-Y--.IYIVQT.--	91
PCPV1QF-LLT-SLCFL-A-C--.QYILQTQ--	94

COBBLER sequence from MOTIF

```
>hvp_E5. HPV66, with embedded consensus blocks
spyiatidfcvicvFILCFCICLCVCIVxLLlsaslftscLLLCLLLWLLVatsffdtf
ilfl1lffiyiptlciychalwlinhl
```

* coordinate 3850 in HPV16R
 <- E2 end in HPV16R

most-likely	ATGATAaAtATACTTGTA.....TTTGTAAATTAGT.....	30
HPV42	-CAG-T...--GGAC-TCAATACTGT...GACTCC-CA-CGTGTGGGACAACC.....G	49
HPV61	--CC-----C-AT-C--G-CGAGCTTGTGTG---AT-GAT-GATGGTGTAC.....	51
HPV26	CAC-CCT-----C---ATCATT.....A--CCTT---TA.....	36
HPV51	T--TAT-GACATA----ACCATT.....GCA--GT---T.....	36
HPV30	T-ATAC--C-ACAAACGCCAT.....A--T-GC-GCTACGCTGTTATATATGCAACCATTGTTTTAT	64
HPV66	TC-CC-T---TGCAACC.....A--ATT--T--GTA.....	33
HPV18	---T--...TC---A-T.....--T--T--T-C.....	27
HPV45	---C--...TCTT-A--G.....--T--T-GT-C.....	27
HPV39	-----T--T-G---.....--T--GG-GT-G.....	27
HPV70	----T...GC-T-G--G.....--T--GG-GT-G.....	27
HPV59	-----...CC-----.....--GTG-T--.....	27
HPV16	---C---C-TGA-ACT.....GCATCC-CA-CATTA.....	33
HPV35h	-----GAC---AC-.....GC-TCC-G--C-GTG.....	30
HPV31	----TG-AC--AA-A-TTCT.....ACA---GC-T.....	33
HPV52	---T-.....GGAT--.....--T--T--.....	24
HPV33	-----...T---T.....--T--AT--.....	24
HPV58	-----...T---C-A-T.....--T--TG--T.....	27
HPV6bE5a	---GA-GTGG-G-C---CAAATA.....GC--C-GGA-CA.....	36
HPV11E5a	---GAGGTAG-G-C---CAAATTGCT...GCA-C--CA-C-ACA.....	42
HPV44	---GA-C-C---C-A--.....GA--CT-C--TAGGGGCAACC.....	39
HPV55	---GA-C---C-A--.....GA--GT-C--TAGGGGCAACC.....	39
HPV13	---GA-TT---C---G.....GA--T-G--CACAGGCAACC.....	39
PCPV1	---GA-TTACAGG---CCT.....G-A-ATT---C-GCAAAAGCA.....A	43
HPV34	---C...--TT--A.....--T--TTG.....	27
RhPV1E5a	--CCGCC-CC--T-GC-G.....--C-GG--GTGTTT.....	33

GT repeat for HPV51 and HPV26

	-> <-	
most-likelyTTTTACTTTGCTTTTGTGTGTGTTTGTGTGTgTGCTTTGTTAGgCCG...CTTTTGTCT	86
HPV42	GCCAGAAACTGCTGC-----T-G-TTA-AGT---TG--GC-----TGTTG-G---G...TG-ATTAG---A-A	119
HPV61	..GATACTCCTGCTGC--ATAC--TTC-G....GC--CCTATCT-T-C-CAC--A--GCATTT--GGTG-	116
HPV26GTTTATGTG---G-G-----G-G-----TGC---T-G-T-GTT---A.....	101
HPV51ATTTTGCTA--G-G-----G-----C-TGTGT---T-G-T-GTT---...--AC---	101
HPV30	TTGTTAGTTTTGTG-----G-----G--T-----CC-----C---C-ATA-GT-----	134
HPV66ATTTGTGTG--GCG-----T---C-----CA-T--GTC-----	98
HPV18GTG-A--A-G.....-A-G---C.....CA--C.....	71
HPV45--CTG-G--C--.....-A-G---C...--AA--C...--G--A	71
HPV39GGTG-G--TA-A.....-A-A-A-----AA--C.....	71
HPV70GCTG-A---A.....-A-A-A-----AG--C.....	71
HPV59G---GTG-----G.....-G-----AA--C.....	71
HPV16CTGGCGTGC---G-----CT-----C---C-AT-A-TA-GTCCG--GC-TT-	101
HPV35hTTGCTGTGC---GT-G-----CT-----CT---TC-GC--GTA-GTTCG--A----	98
HPV31GTGCTATGC---G-----CTAC-A-T-----TC---C-TA-GTCCA--G----	101
HPV52A-TT-GCTTA-G.....GTG--T-----GCAG-GC-----C-----	71
HPV33--A--T-G-TT--A.....--C--A-CC...-TA--AT-AC-T--T...T-AA-A--	71
HPV58--A---G-TT--A.....-C-----C...ATT---T-GC---A...T-GG---	74
HPV6bE5aACCAGCACAA--CA---GCCTG--ATAA-TGCA--TGT---A--T-----CATCATA--A-TG-	104
HPV11E5aACATTGATA--GCCTG--GTTA-----GCA--TGCA--A--TA--C-----TATTGTA--A-AA-	104
HPV44	..AGCACATCATTAC--GCC-G--GTAA--.....GCCC---T---A-----CATTGTAT-AA-TA-	104
HPV55	..AGCACATCATTAC--GCC-G--GTAC--.....GCCC---T---A--T-----CATTGTAT-AA-TA-	104
HPV13	..AGCAAGTCATTAC--GCC---GTAA--.....GCAC-TACA----TG-A----CATTATAACAA-AT-	104
PCPV1	CCAGTCAATCATTTG--GCC---A-TAA-----GC-C-TAC---A--T---C-C--TATTATAA-AC-TA-	110
HPV34--G-CT--G-----A-A---C...CTG-C---G...T--...--GCAG-	71
RhPV1E5aTTGCTGCTG-----G---G--C-----GTG--T-----GCAC-----GA-T-----GC-T--	89

E5 SuperGroup A Nuc-Aln

* coordinate 4098 in HPV16R

most-likely	TACATTTACATGCAATT.....TTTTTAACAAAACAaCAA	240
HPV42	--ACA----TA....A.....CA----G---T-----	285
HPV61	GTTCG--TGCA...--A.....-G-A-CC-CTTG.....	267
HPV26	-----TT--...-C-.....G-AA-GGT--TTGCTACT	255
HPV51	-----C--T--...-CA.....-CTTTTG---	252
HPV30	--T--GCC----TG-G.....-GG----T---CAC----	291
HPV66	--T--GC----TT-G.....-GG----T---C--TTTG	255
HPV18	-G---A-----T--A.....-G-CTTT-C-G.....	219
HPV45	-----A-G--...GC-.....-ACAC--C-T-----	219
HPV39	-G---AG--TG----G.....GA-A-G...-T.....	216
HPV70	----CA-C-T-...TCA.....G-A-ATG-T.....	213
HPV59	--T--GCC-----A.....C-G-CC...-T-----	219
HPV16	-----AC-----CGC.....-TTAC-...	249
HPV35h	-TA-CGCT-----CAA.....-A---GG--GT---G...	249
HPV31	-T--AC-----TC-.....-GT-----G	252
HPV52	-T--C-GT-----CAG.....-A---GG--C---TG---	225
HPV33	-TA----T-----CAG.....CA-A-G---C-----G-G	225
HPV58	-TA----T-----CAA.....-AC-----CC-----G-C	228
HPV6bE5a	----C-ACT--.....G--ACC---C-G-----	273
HPV11E5a	----CA--T-C.....G-GCA---GC-----	273
HPV44	-----C--.....G-ACAT---C-----G--	276
HPV55	-----CT--.....G-ACAT...-C---G--	273
HPV13	----CCA-T--.....-ACA....-C-----G--	273
PCPV1	----CA-T--.....-GCA---TC-G-----	282
HPV34	-----AC-----T-CC.....-GCA-GT--C-.....	222
RhPV1E5a	----CC-G-----CC-CAGTGTGG---AC-GC-G-ATGATG	249

E5 SuperGroups B-E Nuc-Aln

```

poly-A signal
for HPV41 <-
SuperE.con aat?t?at?accaaacta??gctG?Tga?atTata?t???T???G?C???T???T??TTGGAGGACTTGG 179
HPV41 --CGCT--G-G-----CAC---C---T-G----T-AA 237
CRPV ---T-G-AA--A----A-TT--A-A-A-A---C-GCAGTA-GGCA-C-TTGG-GTTT-CT----- 271

GroupE1.con T?CAATATGG?AGT?TGGGA?T?TT??TGGGAGGTTTGGG?ATTGG?AC?G?CA??GG?? ??GG?GG??G?? 239
HPV1aR -G-----C---C-----G-T--TT-----C-----A--A-C--GA--CTCT--A--AA-AAT 292
HPV63 -A-----G---T-----A-A--CC-----T-----T--T-G--AG--TGGG--T--CC-GTA 190

E5b.con AA?TAG 152
HPV6b-E5b --T--- 219
HPV11-E5b --C--- 225

SuperE.con AATCAGTAGTGCCGGAGGTTCTGGGGGTCGACTAG 214
CRPV ----- 306

GroupE1.con TGGTTATAC?CC?CT?GGTGAGGGTGGTGGGGTTAGAGTTGCTACTCGTCCAACCTCCAGTAA 298
HPV1aR -----T--C--C----- 354
HPV63 -----A--T--A- 207

```

This page intentionally left blank.