RESEARCH COMMUNICATION

Computational Prediction of Nuclear Localization Signals and Structural Characteristics of 91 Types of HPV E6 Proteins

Maryam Esmaeili, Hassan Mohabatkar*

Abstract

Human papillomaviruses (HPVs) are small DNA tumor viruses that replicate and assemble exclusively in the nucleus. Thus their proteins, including E6, must carry nuclear localization signals (NLSs) to enter the nucleus. To analyze and to predict the nuclear localization signals and several post translational modifications by bioinformatics analysis, we obtained 91 E6 protein sequences from available databases. To investigate the localization of these sequences, we used Hum-Ploc software. Homology and alignment of sequences were performed by Blast software and Multalin server respectively. Prediction of N-glycosylation and serine, threonine and tyrosine phosphorylation sites of HPV E6 protein sequences was accomplished with NetNGlyc and NetPhos software. Out of 91 types, the NLSs of 29 types were predicted by signal-3L and signal-CF software. We tried to predict the NLSs of remaining HPV E6 proteins according to the homology of the already predicted NLSs. However, because of considerable variation between E6 protein sequences, we could not classify the NLSs in monopartite or bipartite. According to the results, all NLSs of HPV E6 proteins could be assigned to 11 categories. NLSs of several HPV E6 protein sequences were also determined by experimental studies. Overall, different types of HPV E6 protein in same category show approximately similar pattern in post translational modifications such as N-glycosylation and phosphorylation. Some HPV "early" genes, such as E6, are known to act as oncogenes that promote tumor growth and malignant transformation. Thus more detailed recognition of nuclear localizing sequences and nucleocytoplasmic transport pathway can play a key role in prevention and treatment of HPV infection and related cancers. The results also show that bioinformatics technology can direct and simplify experimental studies.

Key Words: HPV - nuclear localization signal (NLS) - E6 protein - post translational modification - bioinformatics

Asian Pacific J Cancer Prev, **9**, 631-636

Introduction

Human papilomaviruses (HPVs) are icosahedral DNA tumor viruses that infect squamous epithelial cells of the skin or of the anogenital and oropharyngeal mucosa. HPV infection is associated with more than 90% of all cervical cancers, which is the second leading cause of cancer death among women worldwide (Roux and Moroianu, 2000; Zur Hausen, 2000). HPV types are classified as either high or low risk, depending on whether they are associated with malignant or benign lesions, respectively (Zur Hausen, 1991; Gussione et al., 2002). The HPV genome encodes for early proteins including E1, E2, E4, E5, E6, and E7 and for late proteins including L1 and L2 (Harvard et al., 2004).

Among the early proteins, high risk E6 and E7, the two viral oncoproteins, appear to be most important for malignant conversion, as demonstrated by their capacity to immortalize and transform keratinocytes and epithelial cells. Low risk HPV E6 and E7, however, lack such biological activity (Tao et al., 2003; Band et al., 1993; Hawley-Nelson et al., 1989; Liu et al., 1999). Biochemically, high risk E6, but not low risk E6, interacts with E6AP and tumor suppressor protein P53 to induce the ubiquitination-mediated degradation of P53 (Munger et al., 1989; Scheffner et al., 1990; Tao et al., 2003). High risk E7 interacts with tumor suppressor protein pRb to promote cell cycle progression (Munger et al., 1989; Scheffner et al., 1990; Boyer et al., 1996). Thus, interaction with cellular tumor suppressor proteins and perturbation of normal cell cycle controlled by high risk E6 and E7 are the most important influences for malignant conversion (Song et al., 2000; Mantovani and Banks, 2001; Munger et al., 2001).

E6 cooperates with activated ras in the transformation and immortalization of baby mouse kidney cells (Storey and Banks, 1993; Sherman and Schlegel, 1996) and baby rat kidney cells (Liu et al., 1994; Sherman and Schlegel, 1996). High risk E6 is known to be functionally involved in regulation of gene transcription and can interact with other transcription factors and coactivators including p300/CBP (Patal et al., 1999; Roux and Moroianu, 2000; Tao et al., 2003), IRF3 (Ronco et al., 1993; Roux and Moroianu, 2000; Tao et al., 2003), and c-Myc (Gross-

*Department of Biology, College of Science, Shiraz University, Shiraz, Iran *For correspondence: mohabat@shirazu.ac.ir*

Mesilaty et al., 1998; Roux and Moroianu, 2000; Tao et al., 2003). The multifunctional activity of E6 is not restricted to the nucleus. This protein can act as a regulator of signal transduction through interactions with cytoplasmic proteins such as E6BP, E6TP (Tao et al., 2003) and thus results in modification of several cellular activities, including differentiation, signal transduction, cytoskeletal structure, cell polarity, and DNA replication (Thomas et al., 1999; Roux and Moroianu, 2000; Tao et al., 2003;). HPV E6 proteins contain four zinc-binding motifs, Cys-X-X-Cys; forming two Cys-Cys fingers that bind zinc directly (Kanda et al., 1991; Sherman and Schlegel, 1996).

These features of E6 categorize it as a nucleophilic protein, like nuclear proteins in general. The E6 oncoprotein localizes to the nucleus and/or in the cytoplasm (Kanda et al., 1991; Lechner et al., 1994; Sherman and Schlegel, 1996; Harvard et al., 2004) and bind nonspecifically to DNA (Malion et al., 1987; Sherman and Schlegel, 1996).The E6 proteins must carry nuclear localization signals (NLSs) to enter the nucleus. The basic para diagram for nuclear import is that the NLS cargo is bound (either directly or indirectly via an adaptor) by an import receptor in the cytoplasm, translocated through the nuclear pore complex (NPC) and released inside the nucleus. The receptor and the adaptor are then recycled back to the cytoplasm in a form competent for another round of import (Roux and Moroianu, 2000). The structural and functional characteristics of HPV E6 proteins would suggest that the nuclear import of these play a key role in HPV infection and production. So it is important to investigate the NLSs of HPVs.

To date, approximately 130 HPV types have been isolated but only a few NLSs of HPV E6 proteins have been experimentally determined. In this study, we tried to predict nuclear localization signals and several post translational modifications of 91 types of HPV E6 proteins by bioinformatics analysis.

Materials and Methods

Sequences

The HPV E6 protein sequences were obtained from the following databases:

http://www.ncbi.nlm.nih.gov/
http://beta.uniprot.org/uniprot/
http://ca.expasy.org

Prediction of NLS of HPV E6 proteins

To investigate the localization of these sequences, Hum-Ploc server was used which was available at: http://chou.med.harvard.edu/bioinf/Hum-Ploc (Chou and Shen, 2008; Chou and Shen, 2007; Chou, 2005; Shen and Chou, 2005).

To identify potential NLSs in protein sequences that localized in nucleus, PredictNLS, Signal-3L and Signal-CF web servers were applied which were available at: http://chou.med.harvard.edu/bioinf/PredictNLS (Cokol et al., 2000); http:// chou.med.harvard.edu/bioinf/Signal-3L (Shen and Chou, 2007; Chou and Shen, 2007; Chou, 2001; Shen and Chou, 2006); http://chou.med.harvard.edu/bioinf/Signal-CF (Chou and Shen, 2007; Chou, 2001; Shen and Chou, 2006).

A useful tool for analyzing the homology of all types of HPV E6 protein sequences is BLAST (Gasteiger et al., 2003) available at: http://au.expasy.org/tools/blas

To align the sequences, the following web server was used: http://bioinfo.genopole-toulouse.prd.fr/multalin/multalin (Corpet, 1988).

Prediction of post translational modifications of HPV E6 protein sequences

Prediction of N-glycosylated and serine, threonine and tyrosine phosphorylation sites of HPV E6 protein sequences were performed by NetNGlyc and NetPhos software (Blom et al., 1999).

Results

Prediction of NLS of HPV E6 proteins

The full sequences of 91 types of HPV E6 proteins were obtained from available databases. Out of 91 types, the NLSs of 29 types were predicted by Signal-3L and Signal-CF software but PredictNLS software could not predict any NLSs in any of the sequences. Because of considerable variation between E6 protein sequences we could not classify the NLSs in monopartite or bipartite. The NLSs of the remaining HPV types were predicted according to the homology of the previously predicted NLSs.

According to the results, all of NLSs of HPV E6 proteins were categorized in 11 categories (Table 1). These deviations were based on the homology of HPV sequences that were obtained from BLAST and Multalin software. NLSs of several HPV E6 protein sequences were also already determined by experimental studies such as HPV types 16 (Tao et al., 2003). To predict the NLS of several E6 protein sequences, we used the NLS of HPV16 E6 protein that was obtained from experimental studies (Table 2).

Prediction of post translational modifications of HPV E6 protein sequences

Out of 91 types of HPV E6 proteins, N-glycosylation sites (in specific sequence, Asn-Xaa-Ser/Thr) of 41 sequences were predicted by suitable software. Forty eight percent of all sequences contained Isoleucine, Leucine or Threonine in the second position (Xaa) of specific sequence. Eight types of E6 proteins contained more than one N-glycosylated site in their sequences and 16 types of E6 proteins were glycosylated in Asn-20, Asn-21or Asn-22. Out of 29 HPV sequences that servers can predict NLSs for them, only two sequences (HPV7 and HPV36) have glycosylated site in their NLSs. Majority of 91 types of HPVs were phosphorylated on certain Serine, Threonine and Tyrosine amino acids in their sequences. Types of 2, 2a, 27, 27b, 57 and 57c were phosphorylated on Serine amino acids and types of 49, 54 and 76 were phosphorylated on Tyrosine amino acids only. Out of 29 HPV sequences that servers can predict NLSs for them, only four sequences (HPV3, HPV9, HPV10 and HPV17) do not have phosphorylated site in their NLSs.

Table 1. Homology Analysis of NLSs of 91HPV E6 Proteins - Division into 9 Categories

Type	Accession No.	NLS
Category 1		
HPV1A	P06929	1MATPIRTVRQLSE13*
HPV63	Q07855	MDLTSVHSVRDLSS
Category 2		
HPV02	A3F905	1MHTRAGMSEENPCPRNIFLLCKEYGLEL-EDLRLLCVWCKRPLSEA45*
HPV2A	P25484	MSEENPCPRNIFLLCKEYGLELEDLRLLC-VWCKRPLSEA
HPV27	P36808	MRTRAGMSEENPCPRNIFLLCKQYGL-ELEDLRLLCVYCRRALSDA
HPV27b	Q4H2F3	MRTRAGMSEENPCPRNIFLLCKQYGLELEDL-RLLCVYCRRALSDA
HPV57	P22158	MSEENPCPRNIFLLCREYGLELEDLRIL-CVYCKRPLSDA
HPV57c	A8CG50	MSAENPCPRNIFLLCREYGLELEDLRILCV-YCKRPLSDA
Category 3		
HPV03	P36799	1MAVAMSMDANCPKNIFLLCRNTGIGFDDLRLHC-IFCTKQLTTTELQAFALREL53*
HPV28	P50802	MDDQRPKNIFLLCRDSGISFDDLRL-HCIFCAKVLTTAELSAFALREL
HPV94	Q705D2	MSMGAQEPRNIFLLCRNCGISFEDLRL-CCVFCTKQLTVAELTAFALREL
HPV77	O56946	MSTSDGRPKNIFLLCRDSGIPFDDLRLL-CVFCTKELTTAELAAFTIREL
HPV29	P50803	MSRGDGYPKNIFLLCRDSGVPFEDL-RLQCVFCTKELTSPELAAFCIREL
HPV10	P36802	MSMGAQEPRNILLLCRNCGIPLEDLR-LCCIFCTKQLTAAELAAFALREL
Category 4		
HPV04	Q07854	1MADGRPATLDDFCRRFDISFFDLRLTCIFCSHT-VDLADLALFYLKKLSLVFRGNCYYACC60*
HPV95	Q705D9	MANGRPTTLDEYCRRFDISFFDLRLPCIFCFHPV-DLAELASFYIKKLSLVFRGSCYYACC
HPV65	Q07856	MADGRPAALDDFCRRFDISFFDLHLTCIFC-SHTVDLQDLASFYLKKLSLVFRGGCYYACC
HPV88	A8R8M7	MQMEDSYLPKRLDDFCSIFNISLFDVHLPCIFCGYIL-DLQQLGSFYQKQLSLVWRSGACFACC
HPV60	Q80941	MQMEEDRFPTTVADYCSEFDIPLKDLKLKCVFC-RFYLTEQQLAAFYIKNLKLVWKNRYCFACC
HPV48	Q80920	MEPQFPTDLDSYCKYFNISFFDLVLKCIFCK-FSVSIVDLASFHNKRLSVIWRDNTPFACC
HPV50	Q80927	MEPQRAKNLVDYCKQQQISFFELELQCLFCKF-VITLPDLASFHCKKLALVYRDGIAFAAC
Category 5		
HPV05	P06930	1MAEGAEHQQKLTEKDKAELPLSIRDLAEALG31*
HPV5b	P26556	MAEGAEHQQKLTEKDKAELPSTIRDLAETL
HPV36	P50810	MAEQASEQQNITEKEKEQLPLTIKGLSESL
HPV47	P22422	MAQKALEQTTVKE - EKLELPTTIRGLAQLL
HPV08	P06428	MDGQDKASYLDTNKDELPSTIKELAAAL
HPV12	P36803	MAQQA-DQQTVTD-STPELPTTIKELADL
HPV07	P36800	1MSARCGSTARTLFELCDQCNITLPTLQ-INCIFCNSILQTAEVLAFA46*
HPV6B	P06462	MESANASTSATTIDQLCKTFNLSM-HTLQINCVFCKNALTTAEIYSYA
HPV11	P04019	MESKDASTSATSIDQLCKTFNLSLHTL-QIQCVFCRNALTTAEIYAYA
HPV13	Q02269	MESANASTPAKTIDQLCKECNLSMHSLQ-ILCVFCRKTLSTAEVYAFQ
HPV6A	Q84291	MESANASTSATTIDQLCKTFNLSMH-TLQINCVFCKNALTTAEIYSYA
HPV40	P36812	MSARCGSQARTLYELCDQCNITLPTL-QIDCVFCKTVLKTAEVLAFA
HPV43	P19709	MSARSCSQNARTIFELCDECNITLP-TLQIGCIFCKKWLLTTEVLSFA
HPV44	P19710	MESANASTSAQSIDQLCKECNIPMH-NLQILCVFCRKTLSTAEVYSFA
HPV32	P36810	MASTSASSQPSTLYQLCKDFGLTL-RNLQICCIWCKNHLTSAEAYAYH
HPV42	P27229	MSGTSASSQPRTLYQLCKEFGLTLRNL-QISCIWCKKHLTGAEVLAYH
HPV54	Q81018	MSATEPHTDQPRTLADLCKVCNIP-MHSLQLPCAFCKKTVCTAEIYA
HPV74	Q8B5X0	MESANASTSAKSIDQLCKDCNIP-MHNLQISCVFCRKTLSCPEVYS
Category 7		
HPV09	P36801	1MYLTEQIMDRPKPRTVKELADTLVI25*
HPV80	O56952	MDRPKPSSVQELADTLCI
HPV15	P36804	MDRPKPFSVQQLADTLCI
HPV17	P36805	MDRPKPQTVRELADTLCI
HPV23	P50776	MQTVHYLSRMCYTKLLMDSTRPLTVQQLSDKLTV
HPV22	P50775	MQPLVVIYALLAYYLSRMGCYSVFMAL-QRPLTVQQLSDKLTV
HPV38	Q80907	MELPKPQTVQQLSDKLTV
HPV38b	Q4JHA8	MELPKPQTVQQLSDKLTV
Category 8		
HPV19	P36806₁	ANAQATEEEIEIVEEGTTAPQVTEPPLPATIAGL-AALLEIPLDDCLVPCNFCGKFLSHLE A61*
HPV14	P28830	MATTDSSTDSADEGPSPKSSYCDTTETKSSFLEPPL-PATICGLANLLEIPLDDCLIPCNFCGNFLTHLEV
HPV21	P28832	MADSSTDSADEGPSPKRRHLEEENTSSFLEPPLPA-TIRDLANLLEIPLDDCLVPCNFCGNFLTHLEV
HPV20	P28831	MATPPSSEDS----------NIGEAKFPILEPPL-PATICGLAKLLEIPLDDCLIPCNFCGNFLTHLEV
HPV25	P28833	MATANAEQSIGP----------PEQAQVIQPPL PATITDLAALLEIPLDDCLVPCNFCGNFLTYLEI
HPV24	P50777	MAQPGKPQSVLELSRLLNIPLDDCVVPCNFC-KRFLSYTEL
HPV93	Q6TY36	MVYKADILMAAVSKPQTITELARCLG-IPLEDALVPCNFCKKFLTYLEV
Category 9		
HPV53	P36815	1MDRQLFENTEERPRTLHQLCEVVN-KPLLELQLGCVFC37*
HPV30	P36809	MAFKFENTGERPRTVHHLCEVQETSLLELQLQCVYC
HPV66	Q4PZJ0	MDSIFSNTQERPRSLHHLSEVLQIPLLDLRLSCVYC
HPV56	P24836	MEPQFNNPQERPRSLHHLSEVLEIPLIDLRLSCVYC
HPV51	P26554	MFEDKRERPRTLHELCEALNVSMHNIQVVCVYC
HPV26	P36807	MFEDPRERPRTLHELCESLNTTLQNLQVQCVYC
HPV69	Q9JH51	MFQDPRERPRTIHELCEALNTPLQSLQVQCVYC
HPV82	Q9IR59	MFEDIRERPRTLHELCEACNTSMHNIQVLCVYC
Category 10		
HPV72	Q81997	1MPMGLHNPTNIWLLCKEIEVDLEDLRITCIFCK-NELTTEELLAIAIKELQIVWR54*
HPV61	Q80948	MGPCNPTNIFLLCKDYEVDFEDLRLTCVFC-KNELTTEELLAFALKELSIVWR
HPV81	Q705E9	MVSTAVMSLGPANPTNLFLLCKECEVDLDDLQL-TCIFCKKELTVGELLSFAIRELNLVWR
HPV62	Q676V2	MTAGPARPTNLFLLCKEYDVDLDDLHL-TCIFCKTDLSAGELLSFAIRELHVVRR
HPV83	Q9WNN0	MSGVRYPTNIFLLCKDCEVDLEDLRL-ICIYCTNELTTAEVLSFAWKELCIKWR
HPV102	Q2VJC0	MSSARYPTNIFVLCKECEVDLEDLRLLC-IYCTTELTAAEVLSFAIKELCIKWD
HPV106	Q2VJC7	MGTYKRDMLLQHKIAMTNGDCCPKNIFLLCKAY-QLDLQDLNITCVFCKTALTEAEVVSFACKELKVVWR
HPV84	Q99FX3	MPNGRYHPTNIFVLCQEYEVEFDDLRL-ICIFCKEELTEGEVLAFAVKELLIVWR
Category 11		
HPV76	O56940	1MARPAKVCELSQLLNIPISSMLIPCNFCTG30*
HPV75	O56934	MARPAKVCELSQLLNIPISQILLPCNFCTG
HPV49	P36813	MARPVKVCELAHHLNIPIWEVLLPCNFCTG
HPV107	A3RKL6	MLMDRPRTIVQLTQHLNIPVEDLLVPCKFCKR

Discussion

A nuclear localizing sequence (NLS) is an amino acid sequence, which acts like a tag on the exposed surface of a protein. This sequence is used to target the protein to the cell nucleus through the nuclear pore complex (NPC) and to direct a newly synthesized protein into the nucleus via its recognition by cytosolic nuclear transport receptors. Traditionally, in order to identify an NLS experimentally, both of the facts should be considered routinely. The candidate should be deleted to disrupt the nuclear import

Table 2. Tentative NLSs Based on Homology Sequences

Type	Accession No.	NLS		
HPV16	P03126	ERPRKL	KCLKFYSK	KQRHLDKKQR
HPV34	P36811	ERPYKL	PCLLFYSK	KQRHVDENKR
HPV35	P27228	ERPYKL	KCLKFYSK	KQRHLEEKKR
HPV73	Q82005	ERPYKL	PCLKFYSK	KQKHVDEKKR
HPV31	P17386	ERPRKL	KCLRFYSK	KQRHLDKKKR
HPV67	A4U7F2	EKPRNL	QCLRLLSK	KQRHVDRKKR
HPV33	P06427	EKPRTL	LCLRFLSK	KKRHVDLNKR
HPV58	P26555	EKPRTL	KVCLRLLSK	KKRHVDLNKR
HPV39	P24835	ERPYKL	SCIKFYAKIR	KLRHLNSKRR
HPV68	P54667	ERPYKL	SCIKFYAKIR	KLRHLTTKRR
HPV70	P50804	ERPYKL	KCIKFHAKVR	KLRHVNTKRR
HPV97	Q1AHS5	KRPYKL	KCLTFFSKIRE	KYKHLKDKRR
HPV45	P21735	QRPYKL	CIDFYSRIRELR	KRRHLKDKRR
HPV18	P06463	RRPYKL	KCIDFYSRIRE	KLRHLNEKRR
HPV52	P36814	TRPRTL	MCLRFLSK	KERHVNANKR
me180	P27962	ERPYKL	SCIKFYAK	KLRHLNSKRR
HUMAN	Q9Y4Y4	QRPYKL	KCIDFYSR	KRRHLKDKRR

of the NLS; secondly, a non-nuclear protein will be imported into the nucleus if fused to the NLS (Yang et al., 2006; Tinland et al., 1992). These evidences show that identification of the NLS motifs of approximately 130 types of HPVs by experiments is very difficult as if only NLS of HPV16 E6 protein was obtained from experimental studies (Tao et al., 2003). The bioinformatics analysis could be useful for predicting new NLSs. However the similar genomic positions of the E6 open reading frames (ORFs) in all papillomaviruses, as well as, the conserved pattern of Cysteine doublet motifs, suggest that E6 proteins may share common functions (Corossman et al., 1988) but The E6 proteins of all papillomaviruses are poorly conserved. There fore, we could not classify E6 proteins into monopartite or bipartite by bioinformatics investigation. However, experimental studies could demonstrate HPV16 E6 protein between residues 120 and 151 (Sherman and Schlegel, 1996), 7 and 12, 65 and 72, and 115 and 124 (Tao et al., 2003), harbored nuclear localization signals (NLSs), as deletion of the regions could abolish nuclear staining (Tao et al., 2003, Sherman and Schlegel, 1996). But the available servers could not determine NLSs of HPV16 E6 protein. Thus we noted that these results could not always consistent with experimental results.

Some papillomaviruses proteins have been reported to be post–translationally modified by glycosylation and phosphorylation. N-linked glycosylation is one of the most common post translational modifications of protein in the exocytic pathway of eukaryotic cells. Viruses utilize host cell glycosylation machinery to synthesize and process oligosaccharides attached to viral glycoproteins. Glycosylated proteins play an important role in viral particle assembly and in adherence and penetration of the virus into the target cell (Zhou et al., 1992).There are few reports about the post translational modifications of HPV16 E6 capsid protein L1 and L2 but the role of glycosylation is still unknown in early proteins of HPVs. The phosphorylation of the early gene products HPV18 E7 and HPV16 E7, HPV1 E4, HPV16 E2 and HPV11 E2 are well known (Fang et al., 1998) but their role are still unknown (Zhou et al., 1992). Overall, different types of HPV E6 protein in same category show approximately similar pattern in post translational modifications such as N-glycosylation and phosphorylation. In conclusion, identification of NLSs of HPV E6 protein would play a key role in prevention and treatment of HPV infection and allied cancers. The results also show that the different types classified in the same category could share post translational modifications and similar nucleoplasmic transport pathway. Finally bioinformatics technology can direct and simplify experimental studies.

Acknowledgement

Support of this study by Shiraz University is acknowledged.

References

American Joint Committee on Cancer Breast (2002). In: AJCC Band V, Dalal S, Delmolino L, et al (1993). Enhanced degradation of p53 protein in HPV-6 and BPV-1 E6-immortalized human mammary epithelial cells. *EMBO J*, **12**, 1847-52.

Blom N, Gammeltoft S, Brunak S (1999). Sequence- and structure-based prediction of eukaryotic protein phosphorylation sites. *J Mol Biol*, **294**, 1351-62.

Boyer SN, Wazer DE, Band V (1996). E7 protein of human papillomavirus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway. *Cancer Res*, **56**, 4620–24.

Chou KC (2001). Using subsite coupling to predict signal peptides. *Protein Eng*, **14**, 75-9.

Chou KC (2005). Using amphiphilic pseudo amino acid composition to predict enzyme subfamily classes. *Bioinformatics*, **21**, 10-9.

Chou KC, Shen HB (2007). Hum-mPLoc: an ensemble classifier for large-scale human protein subcellular location prediction by incorporating samples with multiple sites. Biochem *Biophys Res Comm*, **355**, 1006-11.

Chou, KC, Shen HB (2007). Signal-CF: A subsite-coupled and window-fusing approach for predicting signal peptides. *Biochem Biophys Res Comm*, **357**, 633-40.

Chou KC, Shen, HB (2008). Cell-PLoc: A package of web-servers for predicting subcellular localization of proteins in various organisms. *Nature Protocols*, **3**, 153-62.

Cokol M, Nair R, Rost B (2000). Finding nuclear localization signals. *EMBO Reports*, **1**, 411-15.

Corossman R, Mora S, Laimins L (1988). Intracellular localization and DNA-binding properties of human papillomavirus type 18 E6 protein expressed with a Baculovirus vector. *J Virol*, **63**, 366-74.

Corpet F, (1988) Multiple sequence alignment with hierarchical clustering. *Acids Res*, **16**, 10881-90.

Fang NX, Frazer IH, Zhou J, et al (1998). Post translational modifications of recombinant human papillomavirus type 6b major capsid protein. *Virus Res*, **60**, 113-23.

Gasteiger E, Gattiker A, Hoogland C, et al (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res*, **31**, 3784-8.

Gross-Mesilaty S, Reinstein E, Bercovich B, et al (1998). Basal and human papillomavirus E6 oncoprotein-induced degredation of Myc proteins by the ubiquitin pathway. *Proc Nalt Acad Sci USA*, **95**, 8058-63.

Gussione E, Massimi P, Bernat A, et al (2002). Comparative analysis of the intracellular location of the high and low-

risk human papillomavirus oncoproteins. *Virology*, **293**, 20-5.

Harvard L, Rahmouni S, Boniver J, et al (2004). High levels of p105 (NFKB1) and p100 (NFKB2) proteins in HPV16-transformed keratinocytes: role of E6 and E7 oncoproteins. *Virology*, **931**, 357-66.

Hawley-Nelson P, Vousden KH, Hubbert NL, et al (1989). HPV16 E6 and E7 proteins cooperate to immortalize human foreskin keratinocytes. *EMBO J*, **8**, 3905-10.

Kanda T, Watanabe S, Zanma S, et al (1991). Human papillomavirus type 16 E6 proteins with glycine substitution for cysteine in the metal-binding motif. *Virology*, **185**, 536-43.

Lechner MS, Mack DH, Finicle AB, et al (1992). Human papillomavirus E6 proteins bind P53 in vivo and abrogate P53 mediated repression of transactivation. *EMBO J*, **11**, 3045-52.

Liu Y, Chen JJ, Gao Q, et al (1999). Multiple functions of human papillomavirus type 16 E6 contribute to the immortalization of mammary epithelial cells. *J Virol*, **73**, 7297-307.

Liu A, Ghahi J, Ostrow RS, et al (1994). The E6 of gene of human papillomavirus type 16 is sufficient for transformation of baby rat kidney cells in cotransfection with activated IIa-ras. *Virology*, **201**, 388-96.

Malion RG, Wojciechowicz D, Defendi, V (1987). DNA binding activity papillomavirus proteins. *J Virol*, **61**, 1655-60.

Mantovani F, Banks L (2001). The human papillomavirus E6 protein and its contribution to malignant progression. *Oncogene*, **20**, 7874-87

Munger K, Phelps WC, Bubb V, et al (1989). The E6 and E7 genes of the human papillomavirus type 16 together are necessary and sufficient for transformation of primary human keratinocytes. *J Virol*, **63**, 4417-21.

Munger K, Basile JR, Duensing S, et al (2000). Human papillomavirus types 16 E6 and E7 contribute differently to carcinogenesis. *Virology*, **267**, 141-50.

Munger K, Basile JR, Duensing S, et al (2001). Biological activities and molecular targets of the human papillomavirus E7 oncoprotein. *Oncogene*, **20**, 7888-98.

Patal D, Huang SM, Baglia LA (1999). The E6 protein of human papillomavirus type 16 binds to and inhibits co-activation by CBP and p300. *EMBO J*, **18**, 5061-72.

Ronco LV, Kkarpora AY, Vididal M, et al (1993). Human papillomavirus 16 E6oncoprotein bind to interferon regulatory factor 3 and inhibits its transcriptional activity. *Genes Dev*, **12**, 2061-72.

Roux GL, Moroianu J (2002). Nuclear entry of high-risk human papillomavirus type 16 E6 oncoprotein occurs via several pathways. *J Virol*, **77**, 2330-3.

Scheffner M, Werness BA, Huibregtse JM, et al (1990). The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell*, **63**, 1129–36.

Sherman L, Schlegel R (1996). Serum- and calcium –induced differentiation of human keratinocytes is inhibited by the E6 oncoprotein of human papillomavirus type 16. *J Virol*, , 3269-79.

Shen HB, Chou KC (2006). Ensemble classifier for protein folding pattern recognition. *Bioinformatics*, **22**, 1717-22.

Shen HB, Chou KC (2007). Signal-3L: a 3-layer approach for predicting signal peptides. *Biochem Biophys Res Comm*, **363**, 297-303.

Storey A, Banks L (1993). Human papillomavirus type 16 E6 gene cooperates with EJ-ras to immortalize primary mouse cells. *Oncogene*, **8**, 919-24.

Tao M, Kruhlak M, Xia S, et al (2003). Signals that dictate nuclear localization of human papillomavirus type 16 oncoprotein E6 in living cells. *J Virol*, **77**, 13232-47.

Tinland B, et al (1992). The T-DNA-linked VirD2 protein contains tow distinct functional nuclear localization signals. *Proc Natl Acad Sci USA*, **89**, 7442-6.

Thomas M, Pim D, Banks L (1999). The role of the E6-P53 interaction in the molecular pathogenesis of HPV. *Oncogene*, **18**, 7690-700.

Yang J, Wang YL, Si LS (2006). Predicting the nuclear localization signals of 107 types of HPV L1 proteins by bioinformatic analysis. *Genomics Proteomics Bioinformatics*, **4**, 34-41.

Zhou J, Sun X, Frazer I (1992). Glycosylation of human papillamavirus type 16 L1 protein. *Virology*, **194**, 210-8.

Zur Hausen H (1991). Human papillomaviruses in the pathogenesis of anogenital cancer. *Virology*, **184**, 9-13.

Zur Hausen H (2000). Papillomaviruses causing cancer: evasion from host-cell control in early events carcinogenesis. *J Nalt Cancer Inst*, **92**, 690-8.