

Stability of viral and viral-cellular fusion transcripts in cervical cancers

Franziska Ehrig, Norman Häfner, Corina Driesch, Katrin Beer, Matthias Dürst

Department of Gynecology, Jena University Hospital – Friedrich Schiller University Jena, 07747 Jena, Germany

Supplementary File S1

Nucleotide sequences of analyzed transcripts

1. Episome-derived transcript HPV16epi-plusUTR

```
ATGCACCAAA AGAGAACTGC AATGTTTCAG GACCCACAGG AGCGACCCAG
AAAGTTACCA CAGTTATGCA CAGAGCTGCA AACAACTATA CATGATATAA
TATTAGAATG TGTGTACTGC AAGCAACAGT TACTGCGACG TGAGGTGTAT
TAACTGTCAA AAGCCACTGT GTCCTGAAGA AAAGCAAAGA CATCTGGACA
AAAAGCAAAG ATTCCATAAT ATAAGGGGTC GGTGGACCGG TCGATGTATG
TCTTGTTGCA GATCATCAAG AACACGTAGA GAAACCCAGC TGTAATCATG
CATGGAGATA CACCTACATT GCATGAATAT ATGTTAGATT TGCAACCAGA
GACAACTGAT CTCTACTGTT ATGAGCAATT AAATGACAGC TCAGAGGAGG
AGGATGAAAT AGATGGTCCA GCTGGACAAG CAGAACCGGA CAGAGCCCAT
TACAATATTG TAACCTTTTG TTGCAAGTGT GACTCTACGC TTCGGTTGTG
CGTACAAGC ACACACGTAG ACATTCGTAC TTTGGAAGAC CTGTTAATGG
GCACACTAGG AATTGTGTGC CCCATCTGTT CTCAGAAACC ATAACTACC
ATGGCTGATC CTGCAGAGCA ACGAAGTATC CTCTCCTGAA ATTATTAGGC
AGCACTTGGC CAACCACCCC GCCGCGACCC ATACCAAAGC CGTCGCCTTG
GGCACCGAAG AAACACAGAC GACTATCCAG CGACCAAGAT CAGAGCCAGA
CACCGGAAAC CCCTGCCACA CCACTAAGTT GTTGACACAGA GACTCAGTGG
ACAGTGCTCC AATCCTCACT GCATTTAACA GCTCACACAA AGGACGGATT
AACTGTAATA GTAACACTAC ACCCATAGTA CATTTAAAAG GTGATGCTAA
TACTTTAAAA TGTTTAAAGAT ATAGATTTAA AAAGCATTGT ACATTGTATA
CTGCAGTGTC GTCTACATGG CATTGGACAG GACATAATGT AAAACATAAA
AGTGCAATTG TTACACTTAC ATATGATAGT GAATGGCAAC GTGACCAATT
TTTGTCTCAA GTTAAAATAC CAAAAACTAT TACAGTGTCT ACTGGATTTA
TGTCTATATG ACAAATCTTG AACTGTCATC CACAACATTA CTGGCGTGCT
TTTTGCTTTG CTTTGTGTGC TTTTGTGTGT CTGCCTATTA ATACGTCCGC
TGCTTTTGTC TGTGTCTACA TACACATCAT TAATAATATT GGTATTACTA
TTGTGGATAA CAGCAGCCTC TCGTTTTAGG TGTTTTATTG TATATATTAT
ATTTGTTTTAT ATACCATTAT TTTTAATACA TACACATGCA CGCTTTTAA
TTACATAATG TATATGTACA TAATGTAATT GTTACATATA ATTGTTGTAT
ACCATAACTT ACTATTTTTT CTTTTTTTATT TTCATATATA ATTTTTTTTTT
TTGTTTGTTT GTTTGTTTTT TAATAAAA
```

Legend:

HPV-16 E6*I	nt 83-226, 409-559
HPV-16 E7	nt 562-858
HPV-16 E1	nt 865-880
HPV-16 E4	nt 3358-3620
HPV-16 E5	nt 3850-4099
Poly-A signals	AATAAA or ATTAAA
ATTTA	"AUUUA" motif, ARE (AU-rich element)
nt in black	non-coding regions e.g. UTR

2. Episome-derived transcript HPV16epi-minusUTR

ATGCACCAAA AGAGAACTGC AATGTTTCAG GACCCACAGG AGCGACCCAG
AAAGTTACCA CAGTTATGCA CAGAGCTGCA AACAACTATA CATGATATAA
TATTAGAATG TGTGTACTGC AAGCAACAGT TACTGCGACG TGAGGTGTAT
TAACTGTCAA AAGCCACTGT GTCCTGAAGA AAAGCAAAGA CATCTGGACA
AAAAGCAAAG ATTCCATAAT ATAAGGGGTC GGTGGACCGG TCGATGTATG
TCTTGTTGCA GATCATCAAG AACACGTAGA GAAACCCAGC **TGTAATCATG**
CATGGAGATA CACCTACATT GCATGAATAT ATGTTAGATT TGCAACCAGA
GACAACTGAT CTCTACTGTT ATGAGCAATT AAATGACAGC TCAGAGGAGG
AGGATGAAAT AGATGGTCCA GCTGGACAAG CAGAACCGBA CAGAGCCCAT
TACAATATTG TAACCTTTTG TTGCAAGTGT GACTCTACGC TTCGGTTGTG
CGTACAAAGC ACACACGTAG ACATTTCGTAC TTTGGAAGAC CTGTTAATGG
GCACACTAGC AATTGTGTGC CCCATCTGTT CTCAGAAACC **ATAATCTACC**
ATGGCTGATC CTGCAGAGCA ACGAAGTATC CTCTCCTGAA ATTATTAGGC
AGCACTTGGC CAACCACCCC GCCGCGACCC ATACCAAAGC CGTCGCCTTG
GGCACCGAAG AAACACAGAC GACTATCCAG CGACCAAGAT CAGAGCCAGA
CACCGGAAAC CCCTGCCACA CCACTAAGTT GTTGACACAGA GACTCAGTGG
ACAGTGCTCC AATCCTCACT GCATTTAACA GCTCACACAA AGGACGGATT
AACTGTAATA GTAACACTAC ACCCATAGTA **CATTTAAAAG** GTGATGCTAA
TACTTTAAAA TGTTTAAAGAT ATAGATTTAA AAAGCATTGT ACATTGTATA
CTGCAGTGTC GTCTACATGG CATTGGACAG GACATAATGT AAAACATAAA
AGTGCAATTG TTACACTTAC ATATGATAGT GAATGGCAAC GTGACCAATT
TTTGTCTCAA GTTAAAATAC CAAAAACTAT TACAGTGTCT **ACTGGATTTA**
TGTCTATATG ACAAATCTTG ATRACTGCATC CACAACATTA CTGGCGTGCT
TTTTGCTTTG CTTTGTGTGC TTTTGTGTGT CTGCCTATTA ATACGTCCGC
TGCTTTTGTC TGTGTCTACA TACACATCAT TAATAATATT GGTATTACTA
TTGTGGATAA CAGCAGCCTC TCGTTTAGG TGTTTTATTG TATATATTAT
ATTTGTTTAT ATACCATTAT TTTTAATACA TACACATGCA CGCTTTTAA
TTACATAA

Legend:

HPV-16 E6*I nt 83-226, 409-559
HPV-16 E7 nt 562-858
HPV-16 E1 nt 865-880
HPV-16 E4 nt 3358-3620
HPV-16 E5 nt 3850-4099
Poly-A signals **AATAAA or ATTAAA**
ATTTA "AUUUA" motif, ARE (AU-rich element)
nt in black non-coding regions e.g. UTR

3. Fusion transcript HPV16-MAP4

ATGCACCAAA AGAGAACTGC AATGTTTCAG GACCCACAGG AGCGACCCAG
AAAGTTACCA CAGTTATGCA CAGAGCTGCA AACAACTATA CATGATATAA
TATTAGAATG TGTGTACTGC AAGCAACAGT TACTGCGACG TGAGGTGTAT
TAACTGTCAA AAGCCACTGT GTCCTGAAGA AAAGCAAAGA CATCTGGACA
AAAAGCAAAG ATTCCATAAT ATAAGGGGTC GGTGGACCGG TCGATGTATG
TCTTGTTGCA GATCATCAAG AACACGTAGA GAAACCCAGC TGTAATCATG
CATGGAGATA CACCTACATT GCATGAATAT ATGTTAGATT TGCAACCAGA
GACAACTGAT CTCTACTGTT ATGAGCAATT AAATGACAGC TCAGAGGAGG
AGGATGAAAT AGATGGTCCA GCTGGACAAG CAGAACCGBA CAGAGCCCAT
TACAATATTG TAACCTTTTG TTGCAAGTGT GACTCTACGC TTCGGTTGTG
CGTACAAGC ACACACGTAG ACATTTCGTAC TTTGGAAGAC CTGTTAATGG
GCACACTAGG AATTGTGTGC CCCATCTGTT CTCAGAAAACC ATAACTTACC
ATGGCTGATC CTGCAGCCAT TAAGACTGAG GGAAAACCTG CAGAAGTCAA
GAAGATGACT GCAAAGTCTG TACCAGCTGA CTTGAGTCGC CCAAAGAGCA
CCTCCACCAG TTCCATGAAG AAAACCACCA CTCTCAGTGG GACAGCCCCC
GCTGCAGGGG TGGTTCCAG CCGAGTCAAG GCCACACCCA TGCCCTCCCG
GCCCTCCACA ACTCCTTTCA TAGACAAGAA GCCCACCTCG GCCAAAACCA
GCTCCACCAC CCCCCTGGCTC AGCCGCCTGG CCACCAATAC TTCTGCTCCT
GATCTGAAGA ATGTCCGCTC CAAGGTTGGC TCCACGGAAA ACATCAAGCA
TCAGCCTGGA GGAGGCCGGG CCAAAGTAGA GAAAAAACA GAGGCAGCTG
CTACAACCCG AAAGCCTGAA TCTAATGCAG TCACTAAAAC AGCCGGCCCA
ATTGCAAGTG CACAGAAACA ACCTGCGGGG AAAGTCCAGA TAGTCTCCAA
AAAAGTGAGC TACAGCCATA TTCAGTCCAA GTGTGGTTCC AAGGACAATA
TTAAGCATGT CCCTGGAGGT GGTAATGTTT AGATTTCAGAA CAAGAAAGTG
GACATCTCTA AGGTCTCCTC CAAGTGTGGG TCTAAGGCTA ACATCAAGCA
CAAGCCTGGT GGAGGAGATG TCAAGATTGA AAGTCAGAAG TTGAACTTCA
AGGAGAAGGC CCAGGCCAAG GTGGGATCCC TCGATAATGT GGGCCACCTA
CCTGCAGGAG GTGCTGTGAA GACTGAGGGC GGTGGCAGCG AGGCTCCTCT
GTGTCCGGGT CCCCCTGCTG GGGAGGAGCC GGCCATCTCT GAGGCAGCGC
CTGAAGCTGG CGCCCCACTG TCAGCCAGTG GCCTCAATGG CCACCCACC
CTGTGAGGGG GTGGTGACCA AAGGGAGGCC CAGACCTTGG ACAGCCAGAT
CAGGAGACAA GCATCTAATG ATGACATTCT GGTCTCGTCT TCCGTCTCCC
CCGTGTTCCC CTCTTGTCTC CCCTGTTCCC CTCTCCCTTC CCTCCTCCCA
TGTCACTGCA GATTGAGACC TACAGGCTGA CGTTCCGGGC AAATGCCAGG
GCCCCACCG ACCACGGGGC CGACATTGTC TCCCGCCCCC CACACTTCCC
TGGCGGCCCC AACTCGGGCT CCCGGGTCTT TGGCCCCCTT TCCCGGGCTG
TCCACTAGAC CAGTGAGCGC TTGGGCGCCG TGCTGGGCAG CCCGCTAGGC
TCGCCTTCCC TCCTGCTTTG CGTGCCCGGG GCAGCAGCAG CCCTGCCCCA
CACCTCCTCT CACTCCCCAG CCTGGGCCCA TCTCCCTGCT TTGGTCTTGC
CCCATCACTG CGCCACTGCT CCGTGGAGGA GGTGGGAGG GGGTTGGGGT
GGTTGAGGCT AAGTTGGGAT CTAGGAGAGG AGAACCAGAT TCTATCCTCA
TCTTTTTTTG GTTCTTTGGT CCAAACCCAA AAGAAACTGA CATGCCCTCC
CTTCTCCCTG GATCTACCTG GAGGGAAGAG TGGAGGTGGA TTCCGAGTGG
TGACAGGACG CTGACCGTGG AGCTTAAGCC ACTGCCTCTC CCTCTGGTCC
CACAAATGGG CGCCCCCCCC TCCCCATGCA GGTGGTGTCTG GGCCCTTCTT
GCTGCCCTGC CCCAAGTTGG GGGTCAGTGC TGCCTGTCCC CATGCTTAAC
ATACCCGCTC AGCTGCTGTC ACATTTTCT TGTTTTGTCC TTTTATTTTT
TTCTAATAAC CTA AAAACTG CAAAATAGT TCTGCAGGTT GAAGCCATGT
CTACATGAAA GTCCTCAGTA AGTGTTAGAG GGAACAGGGC GGAGATATCC
TTATGCCACC CCCGCTGGAG GATGTGGGCA GCTTAGGGCC CTGGAGGCGG
TGCGGCAGGG AAGAGGGGTG CAGAGGCTGT GGCTGGTGGAG CCGGTGAGGC
ACACAAGGGG CCCTTGGAGC GTGGACTGGT TGGTTTTGCC ATTTTGTGTG
GTGTATGCTG CTTTTCTTTT CTAACCAAGA GGCTGGTTTTT GGCATCTCTG
TCCCATTCCC TGGGATCTGG TGGTCAGCCC TAGGATAAAA AGCCAGGGCT
GGAGAACAAG AAAGGGCCAG GAGATGGAAT TC

Legend:

HPV-16 E6*I nt 83-226, 409-559
HPV-16 E7 nt 562-858
HPV-16 E1 nt 865-880
cellular exons
Poly-A signals AATAAA or ATTAAA
ATTTA "AUUUA" motif, ARE (AU-rich element)
nt in black non-coding regions e.g. UTR

4. Fusion transcript HPV16-EIF1

ATGCACAAA AGAGAAGTGC AATGTTTCAG GACCCACAGG AGCGACCCAG
AAAGTTACCA CAGTTATGCA CAGAGCTGCA AACAACTATA CATGATATAA
TATTAGAATG TGTGTACTGC AAGCAACAGT TACTGCGACG TGAGGTGTAT
TAACTGTCAA AAGCCACTGT GTCCTGAAGA AAAGCAAAGA CATCTGGACA
AAAAGCAAAG ATTCCATAAT ATAAGGGGTC GGTGGACCGG TCGATGTATG
TCTTGTTGCA GATCATCAAG AACACGTAGA GAAACCCAGC TGTAATCATG
CATGGAGATA CACCTACATT GCATGAATAT ATGTTAGATT TGCAACCAGA
GACAACTGAT CTCTACTGTT ATGAGCAATT AAATGACAGC TCAGAGGAGG
AGGATGAAAT AGATGGTCCA GCTGGACAAG CAGAACCGBA CAGAGCCCAT
TACAATATTG TAACCTTTTG TTGCAAGTGT GACTCTACGC TTCGGTTGTG
CGTACAAAGC ACACACGTAG ACATTTCGTAC TTTGGAAGAC CTGTTAATGG
GCACACTAGG AATTGTGTGC CCCATCTGTT CTCAGAAACC ATAACTACC
ATGGCTGATC CTGCAGACCC CTTTGCTGAT GCAAGTAAGG GTGATGACCT
GCTTCCTGCT GGCACAGAGG ATTATATCCA TATAAGAATT CAACAGAGAA
ACGGCAGGAA GACCCTTACT ACTGTCCAAG GGATCGCTGA TGATTACGAT
AAAAAGAAAC TAGTGAAGGC GTTTAAGAAA AAGTTTGCCT GCAATGGTAC
TGTAATTGAG CATCCGGAAT ATGGAGAAGT AATTCAGCTA CAGGGTGACC
AACGCAAGAA CATATGCCAG TTCCTCGTAG AGATTGGACT GGCTAAGGAC
GATCAGCTGA AGGTTTCATGG GTTTTAAGTG CTTGTGGCTC ACTGAAGCTT
AAGTGAGGAT TTCCTTGCAA TGAGTAGAAT TTCCCTTCTC TCCCTTGTC
CAGGTTTAAA AACCTCACAG CTTGTATAAT GTAACCATTT GGGGTCCGCT
TTTAACTTGG ACTAGTGTAA CTCCTTCATG CAATAAACTG AAAAGAGCCA
TGCTGTCTAG TCTTGAAGTC CCTCATTTAA ACAGAGGTCA AGCAATAGGC
GCCTGGCAGT GTCAAGCCTG AAACCAAGCA ATACCGTCAT GTTTCAGCCA
AGCCCAGAGC CCTAAGATTA CAAACAATA TGGCCGGAAC CTCCTCAGCT
CTCCCTCTGC AGAGTTCCCT ACCCTAAGAG AATGTTACCA CCTGAACAGT
CCTCGGTGAA TCTGAGAGGA GAGGATGGGG TAAGGCAGAA GCACCAGCTG
TACTACTAGA AGGAGCTTTT GGTGGTAGAT CCCCTGGTGT CTCCAACCTG
ACTAGGTGGA CAGAGCTCAA AGAGGCCCTC TTACCGCTAG CGAGGTGATA
GGACATCTGG CTTGCCACAA AGGTCTGTTC GACCAGACAT ATCCTAGCTA
AGGGATGTCC AAACATCAGA ATGTGAGGCC AACCTTCTAT CAGAGTTAAA
CTTTTGACAA GGAACAAAT CTCAAACTGA TCCATCAGTC ATGTAGCTAG
CTGTAGAGCT TGCAACTTAA TAGCAGCAGC TGCCCAATGC CATGTGAAGT
AACAACTGG TTTTTGGTTT TTTTTTCCCC TTCAGTTTTA ATGTTATGTG
TAATGTATTT AAACCCTTAT TTAATAAAA CTTGTTTTCA GAAATACCTG
ACTTGCA

Legend:

HPV-16 E6*I nt 83-226, 409-559
HPV-16 E7 nt 562-858
HPV-16 E1 nt 865-880
cellular exons
Poly-A signals AATAAA or ATTAAA
ATTTA "AUUUA" motif, ARE (AU-rich element)
nt in black non-coding regions e.g. UTR

5. Fusion transcript HPV16-FAM110B

ATGCACCAA AGAGAAGTGC AATGTTTCAG GACCCACAGG AGCGACCCAG
AAAGTTACCA CAGTTATGCA CAGAGCTGCA AACAACTATA CATGATATAA
TATTAGAATG TGTGTACTGC AAGCAACAGT TACTGCGACG TGAGGTGTAT
TAACTGTCAA AAGCCACTGT GTCCTGAAGA AAAGCAAAGA CATCTGGACA
AAAAGCAAAG ATTCCATAAT ATAAGGGGTC GGTGGACCGG TCGATGTATG
TCTTGTTGCA GATCATCAAG AACACGTAGA GAAACCCAGC TGTAATCATG
CATGGAGATA CACCTACATT GCATGAATAT ATGTTAGATT TGCAACCAGA
GACAACTGAT CTCTACTGTT ATGAGCAATT AAATGACAGC TCAGAGGAGG
AGGATGAAAT AGATGGTCCA GCTGGACAAG CAGAACCAGG CAGAGCCCAT
TACAATATTG TAACCTTTTG TTGCAAGTGT GACTCTACGC TTCGGTTGTG
CGTACAAAGC ACACACGTAG ACATTTCGTAC TTTGGAAGAC CTGTTAATGG
GCACACTAGG AATTGTGTGC CCCATCTGTT CTCAGAAACC ATAACTTACC
ATGGCTGATC CTGCAGCTCT GCCTGCAGTT GAACACAAAG ACCTGGAGGA
GACTCCCACCT CCTTCAGGGA GAAGAAAGGA GGCAGCGAAG CAGATTAAG
GAACTTGTGG CTTGTGGCCT TTTTGTGCAA GGGCCGCTG GAAGGGGAGA
AAAGTGTATG AAAGCCACCG TGGCGGTTAA TGAGCTGTGA GATGAGGCGC
TGCTGCGGCC GCTGCTGCTG ACACTCGCTC CCAGGCTGCA CCCGCCGCC
TTGGCGCTTC ATGTACATGT GTCTATTCAG GCCTTGCGGA GCGGCCAGA
AGAGCATCCA ACACGGCTGC CGGGGAAAGA CCGCCACCA TGCCACGGA
GACCCTACAG ACAGGTAGCA TGGTGAAGCC GGTGAGCCCC GCGGGCACCT
TCACCTCTGC TGTGCCCTG CGCATCCTGA ACAAGGGGCC AGACTACTTC
CGCAGGCAGG CCGAGCCCAA CCCCAGAGG CTCAGCGCCG TGGAGAGGCT
GGAGGCCGAC AAGGCCAAGT ACGTCAAGAG CCAGGAGGTG ATCAACGCCAA
GCAGGAGCCC GTGAAGCCCG CCGTGCTGGC CAAGCCCCCG GTGTGCCCGG
CTGCCAAGCG CGCACTGGGC AGCCCCACGC TCAAAGTGTT CGGCAACCAC
GCCAAGACCG AGAGCGGCGT GCAGAGGGAG AACCTGAAGC TGGAGATCCT
GAAGAACATC ATCAATAGCT CCGAGGGCTC TAGCTCGGGC TCGGGGCACA
AGCACAGCTC CCGCAACTGG CCGCCCCACC GGTGCGAAGC CACTGACCTG
CACCGTCACT CTTTCGCGGA GTCCCTGAAG GTCTACCCCA CGCAGGGCCG
CAGGAGCCCC CAGGAGGGCG GCTCCCACGT GGGCAGGAGA CTGCTGGAGC
AGTCAGCCGA GTCCTTCCCT CACGTGTCCC ACAGCTCTTC GGACATCCGC
AAGGTGACCA GCGTGAAGCC CCTCAAGGCC ATCCCCTGCA GTAGCTCTGC
CCCTCCCCTG CCTCCCAAGC CCAAAATCGC AGCCATCGCC TCCATGAAGT
CCCCCGAGGC CGACCCTGTG GAACCAGCTT GTGGAGTCAG CCGAAGACCC
TCCCTCCAGC GGTCTAAGTC AGACTTGAGT GACAGATATT TCCGAGTGA
CGCGGACGTG GAGAGGTTT TCAACTACT GTGGACTGGA CCCGGAAGAG
CTGGAAAACC TGGGAATGGA AAACCTTGCA AGGGCTAATT CTGACATAAT
ATCCCTCAAC TTCCGCAGCG CAAGCATGAT CAGCTCAGAC TGTGAACAGT
CTCAGGACAG TAACAGTGAC CTTAGAAATG ATGACAGTGC CAATGACCGC
GTGCCGTATG GCATTTCTGC CATCGAAAGA AATGCTAGAA TCATCAAGTG
GTTATATAGC ATCAAACAAG CTAGAGAGTC ACAGAAGGTC TCCCATGTGT
AAGACAGTGC GTGGAAAGGA GGGAGCGTGG GTCTCTGTGC TTACGCAGTT
GTGAAGGTTG TGAGGTCTCC TTGAAGTGTT GTGAGCTTCT CCACTCTTTG
TGTTGCTTGT TGTGCAATGT TTTCAAGTTG CATGCTTGTC AACATGGGGA
CTGACCTGGC TTCCACGTCA CCATCGCTAC AGAGCCTGGC TGAACACGCG
TCATCTGCCA AAAGTTTCAG GCCAGAATCT AATTAGGGCT TTGCTCGTAA
GCAAAACTGT TTACATGAAA ATGGTATACA ACTTCTTCGA TATTTATGTG
GTTCTTGTGT TTTTATATCC CGCGCTATTG TGTCTTAATT AAAAGTTTAA
TCAACTGGCT TTTAAGTTGA GAGCAGAATC TTTTGAAT AAAAAGCCAC
TTTGCTACA TGTGAGACTA TTCCAGTGGG ACAGTAAATG GGATCTGCTA
CCAAACAGAT AATGACTGAC CGAACATAGA ACCAGGCTGG GTTTTCTGTG
GAATAAAGTG GTGAGCCTGA CGTTAATCGT TTACATGTTG AAATTCTCTT
GCCACTTAGT GAAGCAGTCA TTGGGGTAAC AATGCTACGT TTTGTGTCCG
ATTCAGTGTG TAAACGTTGG TGGTTCTATG TAATACTGAC CCCAAGGAAA
GACGATCAGA TACAAAGAAA GATGGTTAAA TCTAGGAGGT TGGGTGTTTG
GGGGTTTTGT TGATGTTGTT CTGATTGGGA TAATTCAGAA ATTCCTAATA
TTATCTTTTC ACAGAATGAC CACCCTAAAA CAAATACTGA ATGCCTTTAA
AAAAAAAAAA AGTTGTGGTT TTTTGTTTTT TTTTTTTTTT TTTTTGGTTCG
AGAACTACTA ATTTTCATGCT ATTTCTTTCT CTCTCCCTTT GTTGTATTTT
GATTTGGGGG AAGGGAGTGA GGCTTGTGCA GATTTGAGCT ATCTGAAGAC
ATGAACAGAA GTCAAATACC AGCCATATCA GTAGAGCCTT CCATTCAAAA
GTGAACACAC TTGGTAGCTG AACCATATCT GAGCAGCGTG ATTCTGTTGT

CTTGATATG TTATTCTCTC AGGCTGTGGA TCTTTGTTAC AGCTCTAGGA
AACTGTAGAA ATACAATGCT ATGTAGCTTT TCCACCCCAT GGTCTCTAGT
GCTGCCTATC AACTTGTCCC CTTTTTTCCC AATAACCTGT CTTCCAGGTG
GTACAGTTAG CTGTCACTCA GCTGACACCA TGATGTGGCA GCAGAGAGGG
AAACCTACAA GTGGTTTGCC TCATTGCCTT TGCCACATCT GAAGTTCTCA
GCAGCACTAC CTTAGACTTC ATGAGCTAAT AGGAACTTT TTATGGTGTA
AATGCTGTAA GACTTTGTAC ATACTTCAGT TGTTATGAAA TCTTTAAAGA
AAAAAAGAAA AAGTTACGCT **AATAAATTGC** TGTGGTGCAG G

Legend:

HPV-16 E6*I nt 83-226, 409-559
HPV-16 E7 nt 562-858
HPV-16 E1 nt 865-880
cellular exons
Poly-A signals **AATAAA or ATTAAA**
ATTTA "AUUUA" motif, ARE (AU-rich element)
nt in black non-coding regions e.g. UTR