



Figure S1: Cumulative pair-wise distance frequencies for HPV16 genes and control region: Pair-wise distances (substitutions per site) are calculated for the full-genome, LCR and all ORFs of reference sequences. Horizontal plain grey line represent the pair-wise distance 95th percentile. E2-E4 (E2 minus E4) stands for the E2 gene non-overlapping with the E4 gene.