



Supplementary Figure 1. Correlation between BLAST statistics. **(A)** Scatterplot of BLAST maximum bit score and \log_{10} E-value for the study samples ($n = 154$) showed a perfect linear negative-correlation ($R^2 = 1.0$). The regression line (\log_{10} E-value = $9.41 + -0.301 * \text{max bit score}$) shown as (—) may be employed for prediction of E-values based on bit scores. E-values $<2.13\text{E-}176$ (\log_{10} -176) equivalent to max bit scores >614 (↓) were rounded and reported as 0 and interpreted as *identical* between the query and hit sequences in the BLAST report. Extrapolation of the fitted line allows for prediction of actual E-values based on bit scores (range 200-1,200) as shown. **(B)** Scatterplot of BLAST maximum bit score and maximum identity (%) for the study samples ($n = 154$) showed high curvilinear correlation ($R^2 = 0.999$). The least squares regression curve (max identity % = $37.02 * \text{max bit score}^{0.1416}$) (—) based on the *power* equation model may be employed for prediction of max identity based on bit score. Bit scores <531 equivalent to an E-value $>3.793\text{E-}151$ (\log_{10} -150) and max identity of $<90\%$ (↑) may serve as a threshold for quality checks by manual verification of HPV genotype. Of note, one sample was excluded from the plot due to zero BLAST hits.