

Supplementary Data File List

S1.1: This file contains clinical and demographic information for all patients.

S1.2: This file contains the results of molecular analyses used to determine the HPV status of all patients. These include data from in situ hybridization, p16 staining, RNA and DNA sequencing, and the MassArray assay.

S1.3: This file contains counts of base changes seen in APOBEC and smoking mutation signatures for HPV(+) and HPV(-) subjects. Fisher's exact test p-values are also shown.

S2.1: This file contains results from GISTIC analyses for head and neck (HNSC), lung, and cervical squamous cell carcinoma. Additional analyses for HNSC were conducted by HPV status and tumor site.

S2.2: This file contains the results of analyses of frequency comparisons of significantly reoccurring DNA copy number alterations by HPV status and tumor site. Fisher's exact test p-values are also shown.

S3.1: This file contains the results of fusion detection analyses performed with MapSplice.

S3.2: This file contains information about HPV viral integration sites based on the analysis of RNA and DNA sequencing data.

S3.3: This file contains the results of SigFuge analyses to detect differential expression of multiple gene isoforms. Uncorrected p-values are shown.

S4.1: This file contains the results of MutSig analyses to detect significantly mutated genes. Additional analyses were conducted by gene expression subtype, tumor site, and HPV status. Mutation counts by site and HPV status are also shown, as are Fisher's exact test p-values.

S4.2: This file contains information about structural rearrangements identified by BreakDancer and Meerkat.

S5.1: This file contains the results of SAM analyses to identify differentially expressed genes. False discovery rate q-values are shown for tumor vs. normal, as well as comparisons based on tumor site, HPV status, and smoking status.

S5.2: This file contains the results of SAMseq analyses to identify differentially expressed miRNAs. False discovery rate q-values and other summary statistics are shown for tumor vs. normal, as well as comparisons based on tumor site, HPV status, smoking status, and miRNA subtype.

S5.3: This file contains the results of analyses that identified epigenetically silenced genes based on gene expression levels in methylated and unmethylated samples. Test statistics, and corrected and uncorrected p-values are also shown, as are correlations of methylation and expression levels.

S5.4: This file contains the results of analyses to identify differentially methylated genes. Test statistics, and corrected and uncorrected p-values are shown for comparisons based on tumor site, HPV status, and smoking status.

S6.1: This file contains information about the 160 antibodies that were used in the reverse phase protein array analyses.

S6.2: This file lists the barcodes for the n = 200 samples for which reverse phase protein array analyses were performed.

S7.1: This file contains p-values for the mutual exclusivity modules analyses presented in Figure S7.1. In addition, uncorrected and corrected Fisher's exact test p-values are shown for the associations presented in Figures 4A and 4B.

S7.2: This file summarizes information about the subtypes identified by the RNA, miRNA, methylation, reverse phase protein array, and PARADIGM analyses. Two-way tables show counts for all pairs of subtypes. Fisher's exact test p-values are also presented.

S7.3: This file identifies patients that exhibit alterations in the pathways described in Figures 5, S7.15 part 1, and S7.15 part 2. For each patient, specific alterations are shown based on output from the cBioPortal.