Supplementary Methods

Human tissue samples. The tumor tissue samples from patients with HNSCC and normal samples were obtained from patients surgically treated in the Department of Otolaryngology-Head and Neck Surgery at Johns Hopkins Medical Institutions, Baltimore, using appropriate informed consent obtained after institutional review board approval. Microdissection of frozen tumor tissue was done to assure that >80% of tissue contained HNSCC. The normal tissues consisted of tissues obtained from non-cancer affected control patients that underwent uvulopalatopharyngoplasty (UPPP). After review by a pathologist, a section of dissected mucosal layer from discarded UPPP specimens was immediately frozen in liquid nitrogen. All specimens were stored at -80 °C until processing. For microarray analysis, a cohort including 44 HNSCC tumors and 25 normal mucosa tissues was used (**Supplementary Table I**). For quantitative real-time PCR analysis, in addition to the above cohort used for microarray analysis, two separate cohorts with one comprising of 31 HNSCC tumor tissues and 17 normal mucosa tissues and the other 63 HNSCC tumor tissues and 30 normal mucosa were also used.

DNA extraction and RNA isolation. DNA was extracted from all samples by digestion with 50 μg/mL proteinase K (Boehringer, Mannheim, Germany) in the presence of 1% SDS at 48 °C overnight, followed by phenol/chloroform extraction and ethanol precipitation. DNA sample quality was assessed by gel electrophoresis (17). Total RNA was isolated with trizol reagent according to the manufacturer's instructions, and then purified with an RNeasy Kit (Qiagen, Germantown, MD). The quality and integrity of the extracted RNA was assessed by Agilent 2100 Bioanalyzer (Agilent, Palo Alto, CA).

HPV analysis. The HPV status was determined as described previously (18). In brief, specific primers and probes have been designed to amplify the E6, E7 regions of HPV16. All the samples were run in duplicate. Primers and probes to a housekeeping gene (β-actin) were run in duplicate and parallel to normalize input DNA. Samples in which 2 results were not concordant were repeated twice in duplicate and were usually due to failed PCR in one of the initial reactions. Each reaction was run 50 cycles. By using serial dilutions, standard curves were developed for the HPV 16 viral copy number using CaSki (American Type Culture Collection) cell line genomic DNA, known to have 600 copies per genome (6.6 pg of DNA per genome). Standard curves were developed for HPV16 E6 and E7, using serial dilutions of DNA extracted from CaSki cells with 50,000, 5,000, 500, 50, and 5 pg of DNA. Standard curves were developed as well for the β-

actin housekeeping gene (2 copies per genome), using the same serial dilutions of the CaSki genomic DNA. This additional step allowed for relative quantification of the input DNA level and final quantity as the number of viral copies per genome per cell. HPV copy number more than 1 copy per cell for tumor samples were regarded as positive. For saliva samples, any amplified sample with HPV E6 or E7 amplification with at control β-actin amplification of 10 ng was regarded as positive.

Copy Number Analysis. The cohort of 44 HNSCC tumor tissues and 25 normal mucosa samples were run on Affymetrix SNP6.0 arrays for copy number analysis. All arrays were run according to the manufacturers' instructions. DNA processing, preparation, hybridization and chip scanning were performed by the Johns Hopkins Microarray Core Facility. The data was normalized using the crlmm package and genome build HG18 annotations (19). Briefly, data was loaded, quantile normalized, and copy number estimates made for each nonpolymorphous probe. The gene level copy number estimate was made by averaging all such probes between the transcription start and transcription end sites defined in the UCSC genome browser refFlat file from build HG18. Where there were multiple annotations for a gene due to build characteristics, the measurements showing the maximum variance were retained for that gene, providing CNV estimates for 17377 genes. The 25 normal mucosa samples and 44 HNSCC samples were then analyzed with the Tibshirani and Hastie outlier statistic (20) and compared to 1.74 million values generated by permutation of sample labels to generate an empirical p-value using the approximation of Smyth and Phipson (21). Outlier samples were defined in the standard way.

Methylation Array Analysis. Bisulfite conversion of genomic DNA from the above cohort was done with the EZ DNA methylation Kit (Zymo Research, D5002) by following manufacturer's protocol with modifications for Illumina Infinium Methylation Assay in the Johns Hopkins microarray core. Bisulfite-converted genomic DNA was analyzed using Illumina's Infinium Human Methylation27 Beadchip Kit (WG-311-1202). Beadchip contains 27,578 CpG loci converting more than 14, 000 human RefSeq genes at single-nucleotide resolution. Chip process and data analysis were performed by using reagents provided in the kit and following manufacturer's manual. Data were extracted and summarized using BeadStudio v3.0 software. Beta estimates were generated from the U and M readings for each probe. We have applied the same statistical analysis used for copy number looking for hypomethylation outliers in Notch pathway genes defined in the

Kyoto Encyclopedia of Genes and Genomes (22). NOTCH genes with a P value <0.05 were considered significant.

Gene expression profiling. RNA isolated from the cohort of 44 HNSCC tumor tissues and 25 normal mucosa samples were run on Affymetrix HuEx 1.0 GeneChips for expression analysis. All arrays were run according to the manufacturers' instructions. cDNA processing, preparation, hybridization and chip scanning were performed by the Johns Hopkins Microarray Core Facility. Analysis was performed in R with standard packages for expression (23-25), including RMA normalization and core summary for gene level estimates. Final gene level expression values were produced by choosing the highest mean expression levels among all probesets linked to the same gene for expression. This yielded expression estimates for 16330 genes. To look at the expression of Notch pathway genes on a tumor sample basis, we compared the expression level of each gene in each individual tumor to the distribution of expression levels in normal mucosa samples assuming a normal distribution. We declared a gene as significantly differentially expressed if the tumor expression level was 2σ away from the mean normal expression level, effectively setting a threshold of α = 0.05.

Heat map. Heat maps of differential gene expression were created using the R-package gplots. The visualized data comprised gene by sample expression levels, and expression was normalized to *Z*-score on a gene-by-gene basis to convert all genes to the same scale.

Gene set analysis. Gene Set Analysis was performed using a mean rank gene set enrichment test (26), as provided in the limma R package (27). Initial gene ranks were generated using an Empirical Bayes statistic. Significance of the Notch signaling pathway gene set (KEGG database) or Nguyen_Target_Set (Molecular Signature Database from the Broad Institute) (28) were measured. Recently, Nguyen identified 85 differentially expressed genes as a gene set (Named as Nguyen_*NOTCH1*_Targets in Molecular Signatures Database of the Broad Institute) of Notch downstream targets that are concomitantly modulated by activated Notch1 in mouse and human primary keratinocytes (28). With this gene set, we also assessed the expression status of Notch signaling pathway in the above cohort of HNSCC tumors.

Targeted and whole-exome sequencing. A total of 52 genes (**Supplementary Table II**) were selected from the COSMIC mutation database (v58), 29 of which were selected for targeted hotspot sequencing, and 23 for sequencing across all gene exons. The exome sequencing was performed by

Asuragen (Austin, Texas) (29-31). Gene-specific primers were designed to amplify products up to 200 bp. Genomic DNA from HNSCC tumor tissues and lymphocytes was fragmented to an average size of ~4 kb using the Covaris S220 (Covaris, Woburn, MA). All fragmented DNA and FFPE specimens were evaluated for the extent of fragmentation following analysis using E-gels (Life Technologies). Fragmented genomic DNA (500 ng) was then merged with an emulsified ~2000 member primer library (SuraSeqTM 7500) using the RainDance RDT 1000 platform. The RDT 1000 instrument is a microfluidic chip-based platform that incorporates microdroplet-based technology to amplify hundreds to thousands of genomic loci with high specificity and uniformity (Tewhey R, Nat Biotech, 2009, 27, 1025). Templates within merged droplets were next amplified using the following PCR conditions: 1 cycle of 94 °C for 2 min, 55 cycles of (94 °C for 15 sec; 54 °C for 15 sec; 68 °C for 30 sec), 1 cycles of 68 °C for 10 min and 4 °C hold. Following emulsion breaking, the resulting PCR products were purified using Qiagen MinElute kit according to the manufacturer's instructions. A fraction of the purified PCR products was examined for size and quantity using the Agilent Bioanalyzer Lab-on-a-chip DNA 12000 and the Nanodrop, respectively. Subsequently, a tagging PCR reaction was performed to append unique barcode sequences to the gene-specific products from each sample and to add adapters specific for sequencing on the Illumina GAIIx platform. Samples were fingerprinted for multiplex sequencing using one of 48 barcodes (Illumina). Purified products (10 ng) were tagged using the following conditions: 1 cycle of 94 °C for 2 min, 10 cycles of (94 °C for 30 sec; 56 °C for 30 sec; 68 °C for 1 min), 1 cycle of 68 °C for 10 min and 4 °C hold. These PCR products were then pooled and purified using Qiagen MinElute PCR purification kit, and quantified using the KAPA Library Quant kit (KAPA Biosystems, Cape Town, South Africa) following the manufacturer's instructions. All samples were normalized to 8.6 nM and pools of 15 samples per lane were prepared. Flow cell preparation and data acquisition were completed using Illumina's recommended protocols. Paired-end sequencing runs (2x151) were performed using the Illumina GAIIx platform.

The sequence read data generated from the GAIIx were demultiplexed, adapter and primer sequences were removed, and trimmed to retain high quality data (q20 or higher). The sequence read data were filtered, aligned and variant scores calculated as previously described (Hadd et al 2012). We retained only the high coverage regions for analysis, high coverage being defined by having greater than 10% of the sample median coverage and above 100 reads. We also flagged the loci that are known to have frequent false positive calls in our library. For each set of matched samples, we filtered out variants that were present in the lymphocytes as

putative germline variants, or as sample specific systematic error. In our analysis, if the matched normal loci is found to have greater than 1% variant reads, or the variant score difference is within the 99.5% percentile of all pairwise differences for non-annotated loci across the matched pair, the background is considered high and the variant is removed from the list. Variants were annotated using gene structure from the NCBI RefSeq transcript set. Coding base substitutions were classified as missense, nonsense, splice site, or silent.

Quantitative real-time RT-PCR. The same RNA samples used for microarray analysis, as well as the RNA samples from an independent HNSCC cohort as described above were assessed for *HES1* and *HEY1* expression levels using quantitative real-time RT-PCR (Taqman). Reverse transcription was performed with random hexamer primers and Superscript II reverse transcriptase (Invitrogen Corp.) as described previously (17). Quantitative RT-PCR was then carried out on the Applied Biosystems 7900 Sequence Detection Instrument (Applied Biosystems, Carlsbad, CA). The primers and probes (Integrated DNA technologies, San Diego, CA, USA) used in the analysis are available upon request. The ratio between the values of the gene of interest and the reference gene Actin or GAPDH were obtained by Taqman analysis and used as a measure for representing the relative quantity of mRNA expression levels in a particular sample. Fluorogenic PCRs were carried out in a reaction volume of 10 μL of 200 nmol/L of each primer, 100 nmol/L of probe, 0.375 unites of platinum Taq Polymerase (Invitrogen), 100 μmol/L of ROX Reference Dye (Invitrogen), 8.4 mmol/L ammonium sulfate, 33.5 mmol/L Trizma (Sigma), 3.35 mmol/L magnesium chloride, 5 mmol/L mercaptoethanol, and 0.05% dimethyl sulfoxide. Each real-time RT-PCR reaction consisted of 1.5 μL of diluted cDNA solution. Thermal cycling was initiated with a first denaturation step at 95°C for 2 minutes, followed by 50 cycles of 95°C for 15 seconds and 60°C for 1 minute. Each reaction was done in triplicate; the average of the triplicate was considered for analysis. The triplicate reactions also provided evidence of reproducibility of the individual reactions. Standard curves were generated which allow direct quantification of gene of interest in each sample. In our analysis, serial dilutions of cDNA product for each gene of interest were used for constructing the calibration curves on each plate.

Immunohistochemistry. 56 HNSCC and 11 non-cancer formalin-fixed and paraffin-embedded samples were obtained from the Head and Neck Tissue Bank at Johns Hopkins and were used to construct a tissue microarray under JHU-IRB approved protocols. The 5 μm cuts were used for slide preparation. Immunostaining was carried out on Bond-Leica autostaining system (Leica Microsystems) using standard

immunohistochemistry (IHC) protocol. IHC protocol incorporated heat-induced antigen retrieval with citrate buffer (pH 6.0) followed by peroxide-blocking step and primary antibody incubation for 15 minutes with rabbit polyclonal antibody against HES1 (Millipore, AB5702, dilution 1: 200), for 30 minutes with rabbit polyclonal antibody against HEY1 (Abcam, ab22614, dilution 1: 200) or 30 minutes with goat polyclonal antibody against cleaved/activated NOTCH1 (Santa Cruz Biotechnology, sc-6014, dilution1: 400). Reaction was developed with biotin-free Bond polymer detection system (Leica Microsystems). 3,3′-Diaminobenzidine (DAB) chromogen substrate for used for visualization of reaction. Slides were counterstained with hematoxylin, dehydrated, and cover slipped. Slides were scanned with 20x resolution. The tissues were analyzed using Aperio software. The staining was categorizes as strong, moderate or none/weak for each individual tissue and averaged for tissue quadruplicates.

Cell Culture. Human HNSCC cell lines UPCI-SCC090 (090), SCC61 and SCC15 were used in experiments. 090 was received from Dr. Susanne Gollin, University of Pittsburgh. SCC61 was received from Dr. Ralph Weichselbaum (University of Chicago), and SCC15 was purchased from American Type Culture Collection (ATCC). All cell lines were finger-printed to confirm the authenticity. Cell line 090 was cultured in DMEM medium supplemented with 10% FBS and 1% penicillin/streptomycin (Corning). Cell lines SCC61 and SCC15 were cultured in DMEM/F12 medium supplemented with 10% FBS, 0.4 µg/ml Hydrocortisone and 1% penicillin/streptomycin (Corning). Cell growth conditions were maintained at 37°C in an atmosphere of 5% CO2. Cell line RNA was extracted as described above for primary tissues.

Transient Transfection, NOTCH inhibition and Cell Proliferation Assay. ON-TARGETplus Pool of siRNAs against *NOTCH1*, *HEY1* non-targeting Pool of siRNA (Thermo Scientific) was used to downregulate the expression of *NOTCH1*, *HEY1* or used as control. Cells were seeded in 96-well plates and allowed to grow until the cells were approximately 70% confluent. Cells were transfected with siRNA using Lipofectamine RNAiMAX Reagent (Invitrogen). *NOTCH1* was inhibited by Gamma Secretase Inhibitor XXI,Compound E (GSI-XXI, EMD Millipore). Cell metabolic activity was determined every 24 hours using the CCK-8 colorimetric assay (Dojindo). Values are mean ± SEM for pentaplicates of cultured cells. The transfection efficiency was confirmed by qRT-PCR as described above for primary tissues with primers and probes for *NOTCH1*, or *HEY1* and normalized to *GAPDH* at 48 hours time point.

Analysis of the TCGA HNSCC Data. The TCGA datasets, including DNA copy number datasets from 288 HNSCC tumor tissues, RNA-seq expression datasets from 279 HNSCC tumor tissues and 37 adjacent normal tissues, and *NOTCH1* mutation data sets from 281 HNSCC tumor tissues, were obtained from the public access data portal (32). The DNA copy number, RNA expression, and *NOTCH1* mutation in HNSCCs were assayed using Affymetrix 6.0 SNP arrays, RNAseq, and Whole genome sequencing respectively. Among these data sets, 279 HNSCC tumor tissues had full DNA copy number, RNA-seq expression, and *NOTCH1* mutation data available, and hereby were used in our study.

Statistical Analysis. All RT-PCR analyses utilized an unpaired Student's t test to determine statistical significance between experimental variables. All statistical tests were two sided without multiple testing correction. A p value threshold of $\alpha = 0.05$ was used to indicate statistical significance. All computations were done in R 2.15.2.