



Published in final edited form as:

*Science*. 2011 August 26; 333(6046): 1157–1160. doi:10.1126/science.1208130.

## The Mutational Landscape of Head and Neck Squamous Cell Carcinoma

Nicolas Stransky<sup>1,\*</sup>, Ann Marie Egloff<sup>2,\*</sup>, Aaron D. Tward<sup>1,3,4,\*</sup>, Aleksandar D. Kostic<sup>1</sup>, Kristian Cibulskis<sup>1</sup>, Andrey Sivachenko<sup>1</sup>, Gregory V. Kryukov<sup>1,5</sup>, Michael Lawrence<sup>1</sup>, Carrie Sougnez<sup>1</sup>, Aaron McKenna<sup>1</sup>, Erica Shefler<sup>1</sup>, Alex H. Ramos<sup>1</sup>, Petar Stojanov<sup>1</sup>, Scott L. Carter<sup>1</sup>, Douglas Voet<sup>1</sup>, Maria L Cortés<sup>1</sup>, Daniel Auclair<sup>1</sup>, Michael F. Berger<sup>1</sup>, Gordon Saksena<sup>1</sup>, Candace Guiducci<sup>1</sup>, Robert Onofrio<sup>1</sup>, Melissa Parkin<sup>1</sup>, Marjorie Romkes<sup>6</sup>, Joel L. Weissfeld<sup>7</sup>, Raja R. Seethala<sup>8</sup>, Lin Wang<sup>8</sup>, Claudia Rangel-Escareño<sup>9</sup>, Juan Carlos Fernandez-Lopez<sup>9</sup>, Alfredo Hidalgo-Miranda<sup>9</sup>, Jorge Melendez-Zajgla<sup>9</sup>, Wendy Winckler<sup>1</sup>, Kristin Ardlie<sup>1</sup>, Stacey B. Gabriel<sup>1</sup>, Matthew Meyerson<sup>1,5,10,11</sup>, Eric S. Lander<sup>1,5,12</sup>, Gad Getz<sup>1</sup>, Todd R. Golub<sup>1,5,11,13,14,†</sup>, Levi A. Garraway<sup>1,5,10,11,‡,†</sup>, and Jennifer R. Grandis<sup>2,15,‡,†</sup>

<sup>1</sup>The Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

<sup>2</sup>Department of Otolaryngology, University of Pittsburgh and University of Pittsburgh Cancer Institute, Pittsburgh, PA 15213, USA

<sup>3</sup>Department of Otology and Laryngology, Harvard Medical School, Boston, MA 02114, USA

<sup>4</sup>Department of Otolaryngology, Massachusetts Eye and Ear Infirmary, Boston, MA 02114, USA

<sup>5</sup>Harvard Medical School, Boston, MA 02115, USA

<sup>6</sup>Department of Medicine, University of Pittsburgh and University of Pittsburgh Cancer Institute, Pittsburgh, PA 15261, USA

<sup>7</sup>Department of Epidemiology, University of Pittsburgh and University of Pittsburgh Cancer Institute, Pittsburgh, PA 15232, USA

<sup>8</sup>Department of Pathology, University of Pittsburgh and University of Pittsburgh Cancer Institute, Pittsburgh, PA 15213, USA

<sup>9</sup>Instituto Nacional de Medicina Genómica, Mexico City, 01900, Mexico

<sup>10</sup>Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02115, USA

<sup>11</sup>Center for Cancer Genome Discovery, Dana-Farber Cancer Institute, Boston, MA 02115, USA

<sup>12</sup>Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA

<sup>13</sup>Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, MA 02115, USA

<sup>14</sup>Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA

‡To whom correspondence should be addressed. Levi\_Garraway@dfci.harvard.edu (L.A.G.); grandisjr@upmc.edu (J.R.G.).

\*These authors contributed equally to this work.

†These authors contributed equally to this work.

Supporting Online Material

[www.sciencemag.org/cgi/content/full/science.1208130/DC1](http://www.sciencemag.org/cgi/content/full/science.1208130/DC1)

Materials and Methods

SOM Text

Figs. S1 to S10

Tables S1 to S11

References (35–66)

<sup>15</sup>Department of Pharmacology, University of Pittsburgh and University of Pittsburgh Cancer Institute, Pittsburgh, PA 15213, USA

## Abstract

Head and neck squamous cell carcinoma (HNSCC) is a common, morbid, and frequently lethal malignancy. To uncover its mutational spectrum, we analyzed whole-exome sequencing data from 74 tumor-normal pairs. The majority exhibited a mutational profile consistent with tobacco exposure; human papilloma virus was detectable by sequencing of DNA from infected tumors. In addition to identifying previously known HNSCC genes (*TP53*, *CDKN2A*, *PTEN*, *PIK3CA*, and *HRAS*), the analysis revealed many genes not previously implicated in this malignancy. At least 30% of cases harbored mutations in genes that regulate squamous differentiation (e.g., *NOTCH1*, *IRF6*, and *TP63*), implicating its dysregulation as a major driver of HNSCC carcinogenesis. More generally, the results indicate the ability of large-scale sequencing to reveal fundamental tumorigenic mechanisms.

---

HNSCC is the sixth most common non-skin cancer in the world, with an incidence of ~600,000 cases per year and mortality rate of ~50% (1). The major risk factors for HNSCC are tobacco use, alcohol consumption, and infection with human papilloma virus (HPV) (2). Despite advances in our knowledge of its epidemiology and pathogenesis, the survival rates for many types of HNSCC have improved little over the past forty years (3). As such, a deeper understanding of HNSCC pathogenesis is needed to promote the development of improved therapeutic approaches.

We performed solution-phase hybrid capture and whole exome sequencing on paired DNA samples (tumors and matched whole blood) from 92 HNSCC patients. Most anatomic sites were represented (oral cavity, oropharynx, hypopharynx, larynx, and sinonasal cavity; Fig. 1C and table S1). Of these patients, 89% and 79% reported a history of tobacco and alcohol use, respectively (table S1). Initially, 14% of all tumors and 53% of oropharyngeal tumors were found positive for HPV based on HPV-16 PCR/in situ hybridization (Fig. 1 and table S1). Tumor copy-number analysis using SNP arrays (fig. S1) replicated previous findings of frequent *CCND1* amplifications, *CDKN2A* deletions, and rarer *MYC*, *EGFR*, *ERBB2*, or *CCNE1* amplifications (4), indicating that the collection is genetically representative of HNSCC.

We achieved 150-fold mean sequence coverage of targeted exonic regions, with 87% of loci covered at >20-fold (figs. S2 and S3 and table S2). We excluded from further analysis 18 tumors in which initial analysis revealed extensive stromal admixture (figs. S3 and S4 and supplemental methods), leaving 74 samples for analysis. We also performed whole genome sequencing (31-fold mean coverage, table S3) on an oropharyngeal tumor and a hypopharyngeal tumor.

On average, 130 coding mutations per tumor were identified, 25% of which were synonymous (Fig. 1A). We queried 321 of these mutations by mass spectrometric genotyping and validated 288 (89.7%). However, the validation rate increased to 95.7% for mutations whose allelic fraction was >20% of total DNA, suggesting that the sensitivity of mass spectrometric genotyping may be reduced in the setting of increased stromal admixture.

The overall HNSCC mutation rate was comparable to other smoking-related malignancies such as small cell lung cancer and lung adenocarcinoma (5, 6). The mutation rate of HPV-positive tumors was approximately half of that found in HPV-negative HNSCC (mean of 2.28 mutations/Mb compared with 4.83 mutations/Mb;  $p = 0.004$ ), consistent with

epidemiologic studies suggestive of biological differences between HPV-positive and HPV-negative disease. The two tumors that underwent whole-genome sequencing harbored 19 (HN\_62469) and 111 (HN\_62699) “high-confidence” somatic rearrangements, respectively (fig. S5 and tables S4 and S5).

Although base mutation rates varied widely (0.59–24 mutations/Mb; Fig. 1A), the average rate of guanosine-to-thymidine (G→T) transversions at non-CpG sites ( $12\% \pm 6\%$ ) was characteristic of tobacco exposure (Fig. 1B). Among patients who reported a smoking history, tumors with the highest fraction of G→T transversions showed a tendency toward increased overall mutation rates ( $p = 0.02$ ) (Fig. 1, B and C). Thus, the G→T transversion frequency may represent a robust readout of “functional” tobacco exposure. We observed differences in mutation rates and G→T transversion frequencies by tumor site even when restricting the analysis to HPV-negative tumors. In particular, HPV-negative laryngeal cancers exhibited higher mutation rates and G→T transversion frequencies compared to HPV-negative cancers found in the oral cavity, oropharynx, hypopharynx, or sinonasal cavity ( $p = 0.008$  and  $p < 0.0001$ , respectively, Fig. 1, A to C, and fig. S8).

Notwithstanding the overall apparent correlation between G→T transversions and mutation rates, several “outlier” tumors showed elevated mutation rates despite a low fraction of G→T transversions. Some of these tumors contained mutations in one or more DNA repair genes. Strikingly, both HNSCC tumors with the highest mutation rates occurred in non-smokers (Fig 1). These results raise the possibility that some HNSCC tumors may contain genetic alterations that promote elevated mutation rates apart from the effects of tobacco (SOM).

To explore the biological basis of HNSCC in an unbiased manner, we used the MutSig algorithm (7) to identify genes harboring more mutations than expected by chance, given the total number of mutations detected. This analysis revealed 39 genes with high statistical significance (False discovery rate  $q < 0.1$ ; figs. S6 and S7 and tables S6 and S7). Compared to recent cancer genome projects such as ovarian cancer and multiple myeloma (7, 8), our analysis of HNSCC revealed a larger number of significantly mutated genes. However, the majority of mutated genes did not reach statistical significance (table S6), suggesting that many may contain passenger events. Thus, we hypothesized that the MutSig algorithm identified an enriched set of genes that likely underwent positive selection during tumorigenesis. Toward this end, numerous significant genes had previously been implicated in HNSCC, including *TP53*, *CDKN2A*, *HRAS*, *PTEN*, and *PIK3CA* (4) ( $q < 0.1$ ) providing support for the validity of the approach. *TP53*, the most commonly mutated gene in HNSCC, was also disrupted by a 100 kb deletion detected by whole genome sequencing, and validated with a focal copy number change detected by SNP array (fig. S9). However, most significantly mutated genes had not previously been implicated in HNSCC.

To explore their biological significance, we first considered mutated HNSCC genes that also undergo frequent genetic alterations in other cancers. *NOTCH1* was particularly noteworthy: point mutations affecting this gene occurred in 11% of the HNSCC tumors (Figs. 2 and 3 and tables S6 and S7), and focal deletions were seen in two additional tumors (Fig. 2). Previous evidence from animal models had implicated Notch dysregulation in cutaneous squamous cell carcinoma (9), but somatic *NOTCH1* mutations had not previously been identified in squamous malignancies. In addition, we found non-synonymous point mutations in *NOTCH2* or *NOTCH3* in 11% of the samples (Figs. 2 and 3 and table S6) and a focal deletion of *NOTCH3* in one additional case (Fig. 2). Whereas *NOTCH1* contains activating mutations in T-cell acute lymphoblastic leukemia and chronic lymphocytic leukemia (10, 11) and *NOTCH2* contains activating mutations in diffuse large B-cell

lymphoma (12), the mutations in HNSCC appeared to be loss-of-function mutations, consistent with those recently described for myeloid leukemia (13).

Several *NOTCH1* nonsense mutations in HNSCC are predicted to generate truncated proteins that lack the C-terminal ankyrin repeat domain, a region critical for transactivation of target genes (Fig. 2) (14). Five additional mutations (four missense and one in-frame deletion) cluster in highly conserved residues situated within or nearby the extracellular ligand binding domain (Fig. 2). Two others are splice-site mutations that may generate truncated proteins or delete critical functional residues (e.g., ligand binding or activation by proteolytic cleavage; Fig. 2). Together, these findings suggested that NOTCH dysregulation—and more generally mechanisms governed by NOTCH signaling—contribute to the genesis or progression of HNSCC.

To further interpret the mutations identified in HNSCC, we looked for functionally related “gene sets” harboring an excess of mutations. For this purpose, we considered an expanded list of 76 genes ( $q < 0.25$ ; table S7) and looked for enrichment in functional gene sets. The highest-scoring gene set contained genes related to epidermal development (table S8). The significantly mutated genes ( $q < 0.25$ ) in this gene set included *NOTCH1*, *IRF6*, and *TP63*. These genes are all clearly related to squamous differentiation. The most abundant *TP63* protein product in squamous epithelia, known as  $\Delta Np63$ , promotes renewal of basal keratinocytes by a mechanism that requires downregulation of *NOTCH1* and *CDKN2A* (15–17). *IRF6*, in turn, has been implicated in the proteasomal degradation of  $\Delta Np63$  (18). Furthermore, terminal differentiation in squamous epithelia is induced in response to genotoxic stress by a mechanism involving p53-dependent transactivation of *NOTCH1*—an activity antagonized by  $\Delta Np63$  (19). Because HNSCC involves transformation of the squamous epithelial lineage, which is histologically similar to the epidermis, these findings led us to hypothesize that mutations in such genes disrupt a stratified squamous development/differentiation program in precursor cells of this malignancy.

Further inspection of recurrent mutations identified eleven additional genes carrying disruptive mutations that function in the squamous differentiation program. The evidence includes mouse knockouts with defects in squamous epithelial differentiation (*Notch1*, *Notch2*, *Irf6*, *Tp63*, *Ripk4*, *Cdh1*, *Ezh2*, and *Dicer1*) (Fig. 3A) (20–25); human germline mutations causing orofacial clefting syndromes (*IRF6*, *TP63*, *CDH1*, and *MLL2*) (26) and knockdown or deregulated expression leading to a differentiation block and increased proliferation in cultured human keratinocytes (*TP63*, *NOTCH1*, *IRF6*, *MED1*) (15, 27). Thus, many mutated genes in HNSCC may govern squamous differentiation. These mutations may promote an immature and more proliferative basal-like phenotype, consistent with known stages of progression and markers of differentiation in HNSCC (Fig. 3B).

We also found recurrent mutations in less well-characterized genes. For example, mutations in *SYNE1* and *SYNE2* were observed in 20% and 8% of HNSCC samples, respectively (fig. S7 and tables S6 and S7). These genes have been implicated in the regulation of nuclear polarity (28), a process that operates upstream of NOTCH1 in squamous epithelia (Fig. 3B) (29). *RIMS2* and *PCLO* mutations were seen in 11% and 12% of cases, respectively; the corresponding proteins mediate calcium sensing (30), another crucial process for terminal squamous differentiation (20).

Beyond the genes directly involved in squamous differentiation, we found mutations involving two apoptosis-related genes: *CASP8* (8%) and *DDX3X* (4%) (fig. S7 and table S7). Thus, suppression of apoptosis may also contribute to HNSCC pathogenesis, perhaps in concert with disrupted squamous maturation (Fig. 3B). The histone methyltransferases *PRDM9* (11%) and *EZH2* (6%) also show highly significant mutation rates.

Viral infection by HPV figures prominently into the etiology of a subset of HNSCC, and is most frequently detected by in situ hybridization (ISH) or p16 immunohistochemistry. We reasoned that HNSCC genome sequencing might also offer a robust HPV detection method. We therefore utilized the PathSeq algorithm (31) and a viral sequence database to identify HNSCC sequencing reads that aligned to HPV genomes. We observed HPV-16 sequence reads in 14 tumors (19%) (range: 1–40,000 reads), 11 of which were also positive by HPV-16 PCR ( $p < 0.0001$ ; table S9). The three tumors that were HPV-negative by PCR had very low HPV-16 sequence read counts (fig. S10); this may reflect reduced HPV dosage or technical contamination. We observed an inverse correlation between HPV status (determined by sequencing) and *TP53* mutation, as shown previously ( $p = 0.006$ ) (32). These data underscore the potential utility of massively parallel sequencing to detect both human and non-human etiologic agents in tumor specimens.

Given that NOTCH pathway inhibitors have entered clinical trials, the discovery of loss-of-function *NOTCH1* mutations in HNSCC may have important therapeutic implications. A recent clinical trial of a gamma secretase inhibitor (which inhibits NOTCH) was halted in part due to an increased frequency of skin cancers in the treatment arm (33). This clinical observation is consistent with those mouse models, in which cutaneous knockout of *Notch1* promotes skin tumor formation (24). Our results suggest that patients taking gamma secretase inhibitors may require monitoring for the development of both cutaneous and head/neck squamous malignancies.

Despite the anatomical distinctions that dominate current clinical management of HNSCC, our results point to several unifying features at the molecular level. For example, *TP53* inactivation—either through somatic mutation or HPV infection—appears nearly universal in this malignancy. The present study suggests that disruption of the squamous differentiation program may represent an additional overarching feature that occurs by numerous genetic mechanisms across tumors from multiple anatomic sites. Thus, HNSCC pathogenesis may involve a maturation arrest or a lineage dependency similar to that seen in other cancer types (34). However, HNSCC appears to be unusual in that the mutational etiology is diverse, in contrast to leukemia and prostate cancer where developmental pathologies appear to be caused by lesions in only a few target genes. Rational therapeutic avenues targeting this block in squamous differentiation may require synthetic lethal approaches to identify specific cellular dependencies arising from NOTCH inactivation, TP63 alteration, or other events that deregulate the program. Finally, our results demonstrate that whole exome sequencing of large numbers of tumor/normal pairs should enable fundamental new insights into tumor biology that are relevant to many human cancers.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

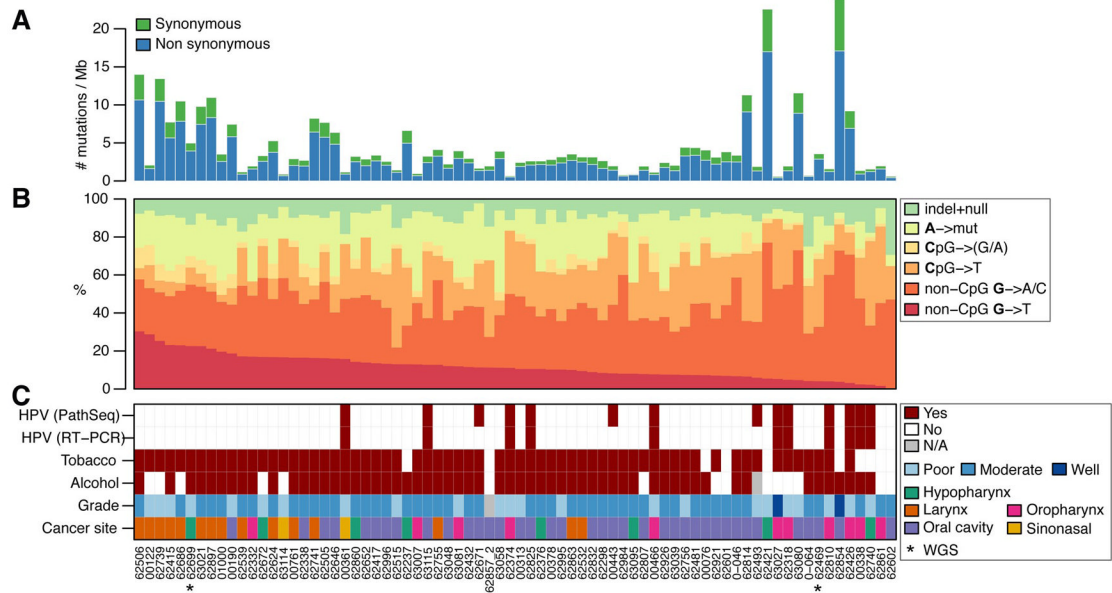
## Acknowledgments

We thank the Broad Institute Sequencing Platform for the generation of sequencing data, the staff of the Biological Samples Platform and the Genetic Analysis Platform at the Broad Institute, and J. Aster for helpful discussions. This work was conducted as part of the Slim Initiative for Genomic Medicine, a joint U.S.-Mexico project funded by the Carlos Slim Health Institute, and is part of a global effort in collaboration with the International Cancer Genome Consortium (ICGC). The work was also supported by grants from the National Human Genome Research Institute (S.G., E.S.L.), the National Cancer Institute (M.M., S.G., L.A.G., J.R.G., A.E.), the Starr Cancer Consortium (L.A.G.), the Novartis Institutes for BioMedical Research (N.S.), the Howard Hughes Medical Institute (T.R.G.), the American Cancer Society (J.R.G.) and the NIH Director's New Innovator Award (L.A.G.). L.A.G. and M.M. are paid consultants for and equity holders in Foundation Medicine, a genomics-based oncology diagnostics company, and are paid consultant for Novartis. M.F.B. is a paid consultant for Foundation Medicine. E.S.L. is an equity holder in Foundation Medicine. The BAM files and SNP array data are available in dbGaP under accession no. phs000370.v1.p1.

## References and Notes

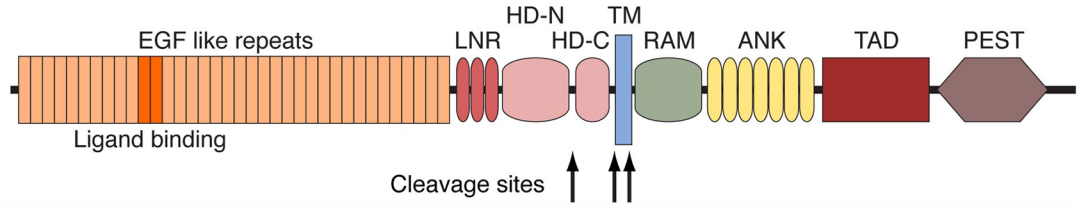
1. Ferlay J, et al. *Int J Cancer*. 2010; 127:2893. [PubMed: 21351269]
2. Argiris A, Karamouzis MV, Raben D, Ferris RL. *Lancet*. 2008; 371:1695. [PubMed: 18486742]
3. Gupta S, Kong W, Peng Y, Miao Q, Mackillop WJ. *Int J Cancer*. 2009; 125:2159. [PubMed: 19569190]
4. Leemans CR, Braakhuis BJ, Brakenhoff RH. *Nat Rev Cancer*. 2011; 11:9. [PubMed: 21160525]
5. Lee W, et al. *Nature*. 2010; 465:473. [PubMed: 20505728]
6. Pleasance ED, et al. *Nature*. 2010; 463:184. [PubMed: 20016488]
7. Chapman MA, et al. *Nature*. 2011; 471:467. [PubMed: 21430775]
8. The Cancer Genome Atlas Research Network. *Nature*. 2011; 474:609. [PubMed: 21720365]
9. Koch U, Radtke F. *Curr Top Dev Biol*. 2010; 92:411. [PubMed: 20816403]
10. Puente XS, et al. *Nature*. 2011; 475:101. [PubMed: 21642962]
11. Sok JC, et al. *Clin Cancer Res*. 2006; 12:5064. [PubMed: 16951222]
12. Lee SY, et al. *Cancer Sci*. 2009; 100:920. [PubMed: 19445024]
13. Klinakis A, et al. *Nature*. 2011; 473:230. [PubMed: 21562564]
14. Kovall RA, Blacklow SC. *Curr Top Dev Biol*. 2010; 92:31. [PubMed: 20816392]
15. Nguyen BC, et al. *Genes Dev*. 2006; 20:1028. [PubMed: 16618808]
16. Okuyama R, et al. *Oncogene*. 2007; 26:4478. [PubMed: 17237812]
17. Su X, et al. *EMBO J*. 2009; 28:1904. [PubMed: 19494829]
18. Moretti F, et al. *J Clin Invest*. 2010; 120:1570. [PubMed: 20424325]
19. Yugawa T, et al. *Cancer Res*. 2010; 70:4034. [PubMed: 20442293]
20. Blanpain C, Fuchs E. *Nat Rev Mol Cell Biol*. 2009; 10:207. [PubMed: 19209183]
21. Chidgey M, et al. *J Cell Biol*. 2001; 155:821. [PubMed: 11714727]
22. Dumortier A, et al. *PLoS One*. 2010; 5:e9258. [PubMed: 20174635]
23. Ezhkova E, et al. *Genes Dev*. 2011; 25:485. [PubMed: 21317239]
24. Nicolas M, et al. *Nat Genet*. 2003; 33:416. [PubMed: 12590261]
25. Tinkle CL, Lechler T, Pasolli HA, Fuchs E. *Proc Natl Acad Sci USA*. 2004; 101:552. [PubMed: 14704278]
26. OMIM. [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)
27. Tu CL, Chang W, Xie Z, Bikle DD. *J Biol Chem*. 2008; 283:3519. [PubMed: 18065418]
28. Luke Y, et al. *J Cell Sci*. 2008; 121:1887. [PubMed: 18477613]
29. Williams SE, Beronja S, Pasolli HA, Fuchs E. *Nature*. 2011; 470:353. [PubMed: 21331036]
30. Fujimoto K, et al. *J Biol Chem*. 2002; 277:50497. [PubMed: 12401793]
31. Kostic AD, et al. *Nat Biotechnol*. 2011; 29:393. [PubMed: 21552235]
32. Westra WH, et al. *Clin Cancer Res*. 2008; 14:366. [PubMed: 18223210]
33. Extnance A. *Nat Rev Drug Discov*. 2010; 9:749. [PubMed: 20885394]
34. Garraway LA, Sellers WR. *Nat Rev Cancer*. 2006; 6:593. [PubMed: 16862190]
35. Amin W, et al. *BMC Cancer*. 2009; 9:396. [PubMed: 19912644]
36. Kostic AD, et al. *Nat Biotechnol*. 2011; 29:393. [PubMed: 21552235]
37. Moberg M, Gustavsson I, Gyllensten U. *J Clin Microbiol*. 2003; 41:3221. [PubMed: 12843067]
38. The Cancer Genome Atlas Research Network. *Nature*. 2008; 455:1061. [PubMed: 18772890]
39. Gnirke A, et al. *Nat Biotechnol*. 2009; 27:182. [PubMed: 19182786]
40. Fisher S, et al. *Genome Biol*. 2011; 12:R1. [PubMed: 21205303]
41. Depristo MA, et al. *Nat Genet*. 2011; 43:491. [PubMed: 21478889]
42. Chapman MA, et al. *Nature*. 2011; 47:467. [PubMed: 21430775]
43. Li H, Durbin R. *Bioinformatics*. 2010; 26:589. [PubMed: 20080505]
44. Reich M, et al. *Nat Genet*. 2006; 38:500. [PubMed: 16642009]
45. The Cancer Genome Atlas Research Network. *Nature*. 2011; 474:541.

46. Berger MF, et al. *Nature*. 2011; 470:214. [PubMed: 21307934]
47. Getz G, et al. *Science*. 2007; 317:1500. [PubMed: 17872428]
48. Benjamini Y, Hochberg Y. *J R Stat Soc Ser B (Methodological)*. 1995; 57(1):11.
49. Robinson JT, et al. *Nat Biotechnol*. 2011; 29:24. [PubMed: 21221095]
50. Fujita PA, et al. *Nucleic Acids Res*. 2011; 39:D876. [PubMed: 20959295]
51. Kozomara A, Griffiths-Jones S. *Nucleic Acids Res*. 2011; 39:D152. [PubMed: 21037258]
52. Sherry ST, et al. *Nucleic Acids Res*. 2001; 29:308. [PubMed: 11125122]
53. Griffith OL, et al. *Nucleic Acids Res*. 2008; 36:D107. [PubMed: 18006570]
54. UniProt Consortium. *Nucleic Acids Res*. 2011; 39:D214. [PubMed: 21051339]
55. Adzhubei IA, et al. *Nat Methods*. 2010; 7:248. [PubMed: 20354512]
56. Forbes SA, et al. *Nucleic Acids Res*. 2011; 39:D945. [PubMed: 20952405]
57. Ding L, et al. *Nature*. 2008; 455:1069. [PubMed: 18948947]
58. Beroukhim R, et al. *Nature*. 2010; 463:899. [PubMed: 20164920]
59. Paez JG, et al. *Nucleic Acids Res*. 2004; 32:e71. [PubMed: 15150323]
60. Mills AA. *Curr Opin Genet Dev*. 2006; 16:38. [PubMed: 16359856]
61. Wong AJ, et al. *Proc Natl Acad Sci USA*. 1992; 89:2965. [PubMed: 1557402]
62. Sok JC, et al. *Clin Cancer Res*. 2006; 12:5064. [PubMed: 16951222]
63. Lopez F, et al. *Head & Neck*. 2011; 33:145. [PubMed: 20848437]
64. Kee Y, D'Andrea AD. *Genes Dev*. 2010; 24:1680. [PubMed: 20713514]
65. Kutler DI, et al. *Arch Otolaryngol Head Neck Surg*. 2003; 129:106. [PubMed: 12525204]
66. Ahel I, et al. *Nature*. 2008; 451:81. [PubMed: 18172500]



**Fig. 1.** Mutation rates, base substitution frequencies, and rearrangements in head and neck cancers. **(A)** Rate of synonymous and non-synonymous mutations, expressed in mutations per megabase of covered target sequence. Non-synonymous mutation rates range from 0.43 to 17.1 mutations/megabase (mean = 3.3). **(B)** Breakdown of individual base substitution rates used for mutation significance, for the same samples as panel (A). The samples were ordered by the rate of G→T transversions, which are indicative of smoking-induced mutations. **(C)** Key clinical parameters for the samples described in panels (A) and (B) (table S10). The first row indicates HPV detection by sequencing, the second row indicates HPV detection by real-time PCR.





Regions of **NOTCH1** activating mutations:

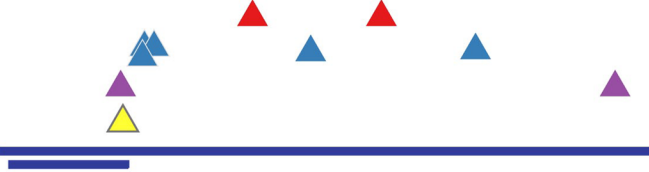
T-ALL  
CLL



Regions of alterations in HNSCC:

**NOTCH1**

Nonsense  
Missense  
Splice site  
Indel  
Deletion



**NOTCH2**

Nonsense  
Missense

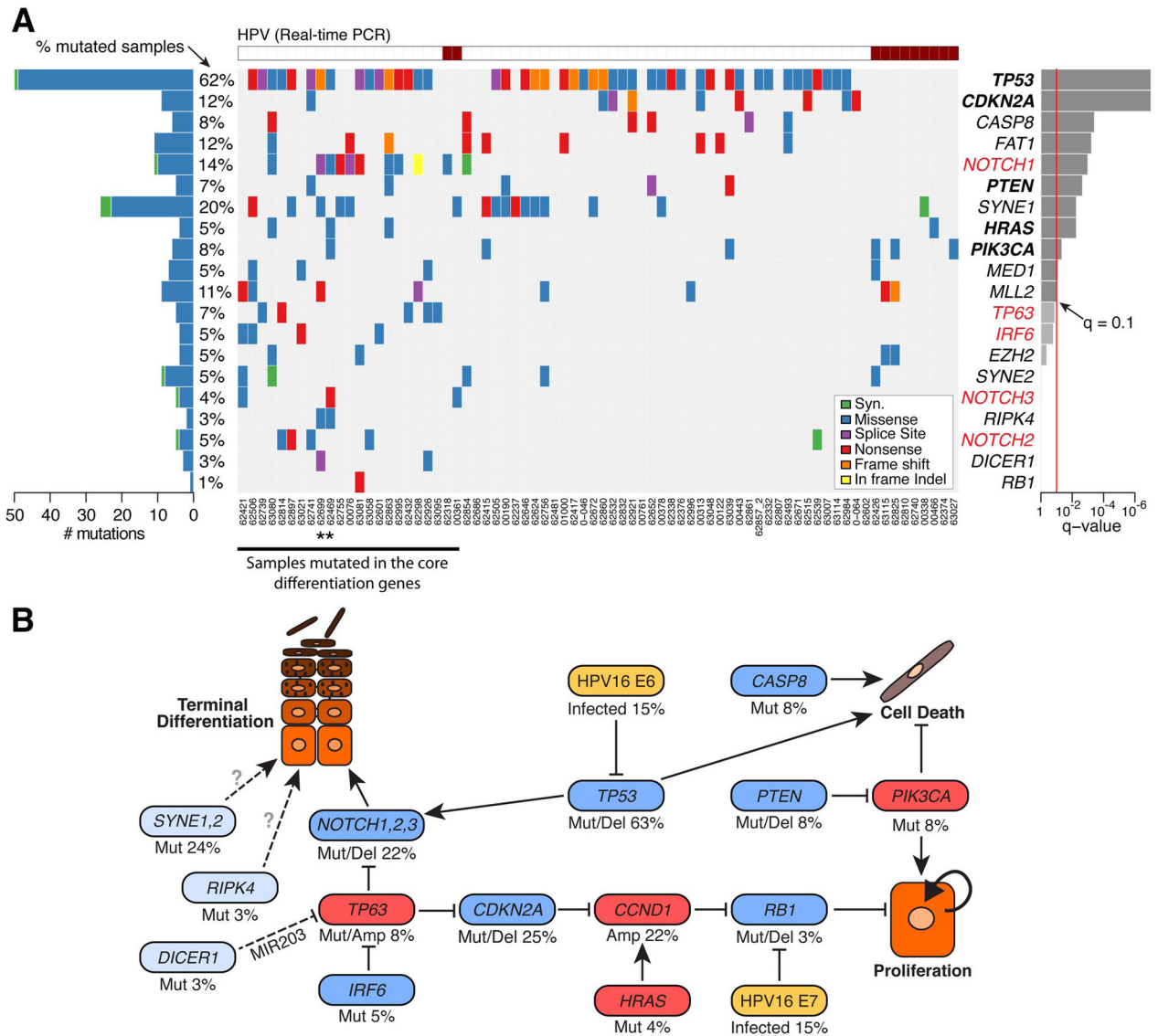


**NOTCH3**

Nonsense  
Missense  
Deletion



**Fig. 2.** NOTCH gene mutations identified in head and neck cancer. A. schematic diagram of the domain structure of *NOTCH1*. (domain structures of *NOTCH2* and *NOTCH3* are similar). All nonsense mutations occur upstream of the TAD domain, which is required for transactivation of target genes. Each arrowhead represents a single point mutation in an individual tumor, of the class indicated to the left.



**Fig. 3.** Genetic disruption of a squamous differentiation program in head and neck cancers. **(A)** Heatmap representation of individual mutations present in a series of 74 tumors, represented in columns. (Top) HPV status by tumor. (Middle) Matrix of mutations in individual genes by type of mutation and tumor. (Left) Number of mutations in each gene. Percentages represent the fraction of tumors harboring at least one mutation in the specified gene. (Right) Deleted recurrently mutated genes ranked by q-value. Genes that define the core differentiation cluster are listed in red. **(B)** Proposed partial wiring diagram of the molecular circuitry of HNSCC. Blue, loss of function; red, gain of function. Numbers listed beneath each protein represent the fraction of tumors harboring mutations, amplifications, or deletions in the corresponding genes.