

Comprehensive genomic characterization of head and neck squamous cell carcinomas

The Cancer Genome Atlas Network*

The Cancer Genome Atlas profiled 279 head and neck squamous cell carcinomas (HNSCCs) to provide a comprehensive landscape of somatic genomic alterations. Here we show that human-papillomavirus-associated tumours are dominated by helical domain mutations of the oncogene *PIK3CA*, novel alterations involving loss of *TRAF3*, and amplification of the cell cycle gene *E2F1*. Smoking-related HNSCCs demonstrate near universal loss-of-function *TP53* mutations and *CDKN2A* inactivation with frequent copy number alterations including amplification of 3q26/28 and 11q13/22. A subgroup of oral cavity tumours with favourable clinical outcomes displayed infrequent copy number alterations in conjunction with activating mutations of *HRAS* or *PIK3CA*, coupled with inactivating mutations of *CASP8*, *NOTCH1* and *TP53*. Other distinct subgroups contained loss-of-function alterations of the chromatin modifier *NSD1*, *WNT* pathway genes *AJUBA* and *FAT1*, and activation of oxidative stress factor *NFE2L2*, mainly in laryngeal tumours. Therapeutic candidate alterations were identified in most HNSCCs.

HNSCCs affect ~600,000 patients per year worldwide¹. They are characterized by phenotypic, aetiological, biological and clinical heterogeneity. Smoking is implicated in the rise of HNSCC in developing countries, and the role of human papillomavirus (HPV) is emerging as an important factor in the rise of oropharyngeal tumours affecting non-smokers in developed countries². Despite surgery, radiation and chemotherapy, approximately half of all patients will die of the disease. Risk stratification for HNSCC is by anatomic site, stage and histological characteristics of the tumour. Except for HPV status, numerous molecular and clinical risk factors that have been investigated have limited clinical utility.

Published genome-wide profiling studies of HNSCC^{3,4} are limited to single platforms. To generate an integrated genomic annotation of molecular alterations in HNSCC, The Cancer Genome Atlas (TCGA) has undertaken a comprehensive multi-platform characterization of 500 tumours with the a priori hypothesis of detecting somatic variants present in at least 5% of samples. Here, we report the results for analyses from the first 279 patients with complete data.

Samples and clinical data

The cohort consists primarily of tumours from the oral cavity ($n = 172$ out of 279, 62%), oropharynx ($n = 33$ out of 279, 12%), and laryngeal sites ($n = 72$ out of 279, 26%) (Supplementary Information section 1, Supplementary Table 1.1 and Supplementary Data 1.1). Most patients were male ($n = 203$ out of 279, 73%) and heavy smokers (mean pack years = 51). Samples were classified as HPV-positive using an empiric definition of >1,000 mapped RNA sequencing (RNA-Seq) reads, primarily aligning to viral genes *E6* and *E7* (Supplementary Information section 1.2 and Supplementary Fig. 1.1). The HPV status by mapping of RNA-Seq reads was concordant with the genomic, sequencing and molecular data, and indicated that 36 tumours were HPV(+) and 243 were HPV(-) (Supplementary Information section 1.2, Supplementary Fig. 1.1 and Supplementary Data 1.2). Of 33 oropharyngeal tumours, 64% were positive for HPV, compared to 6% of 246 non-oropharyngeal tumours. Molecular HPV signatures were identified using microRNA

(miRNA), DNA methylation, gene expression and somatic nucleotide substitutions (Supplementary Information section 1.2 and Supplementary Figs 1.1–1.3). HPV(+) tumours exhibited infrequent mutations in *TP53* or genetic alterations in *CDKN2A*. We evaluated outcome by site, stage, HPV status, molecular subtypes and putative biomarkers (Supplementary Information section 1.3 and Supplementary Figs 1.4 and 1.5). HPV(+) and interestingly patients with HPV(-), *TP53* wild-type tumours demonstrated favourable outcomes compared to *TP53* mutants and 11q13/*CCND1*-amplified tumours.

DNA and RNA structural alterations

Most tumours demonstrated copy number alterations (CNAs) including losses of 3p and 8p, and gains of 3q, 5p and 8q chromosomal regions (Fig. 1a, Supplementary Fig. 2.1 and Supplementary Information section 2) resembling lung squamous cell carcinomas (LUSCs)⁵ (Fig. 1a and Supplementary Figs 2.1 and 2.2). HNSCC genomes showed high instability with a mean of 141 CNAs (amplifications or deletions) from microarray data and 62 structural aberrations (chromosomal fusions) per tumour by 'high coverage' whole-genome sequencing ($n = 29$) (Supplementary Information section 2.2). We observed 39 regions of recurrent copy number loss and 23 regions of recurrent copy number gain ($q < 0.1$, Supplementary Data 2.1 and 2.2). Both HPV(+) and (-) tumours contained recurrent focal amplifications for 3q26/28, a region involving squamous lineage transcription factors *TP63* and *SOX2* and the oncogene *PIK3CA* (Fig. 1b and Supplementary Fig. 2.3).

HPV(+) tumours were distinguished by novel recurrent deletions ($n = 5$ out of 36, 14%) and truncating mutations ($n = 3$ out of 36, 8%) of TNF receptor-associated factor 3 (*TRAF3*) (Supplementary Figs 2.3 and 2.4, and Supplementary Data 2.1). *TRAF3* is implicated in innate and acquired anti-viral responses⁶ including Epstein-Barr, HPV and human immunodeficiency virus (HIV)⁷⁻⁹, while loss promotes aberrant NF- κ B signalling¹⁰. Although *TRAF3* inactivation has been reported in haematological malignancies and nasopharyngeal carcinoma^{11,12}, to our knowledge this is the first evidence linking *TRAF3* to HPV-associated

*Lists of participants and their affiliations appear at the end of the paper.

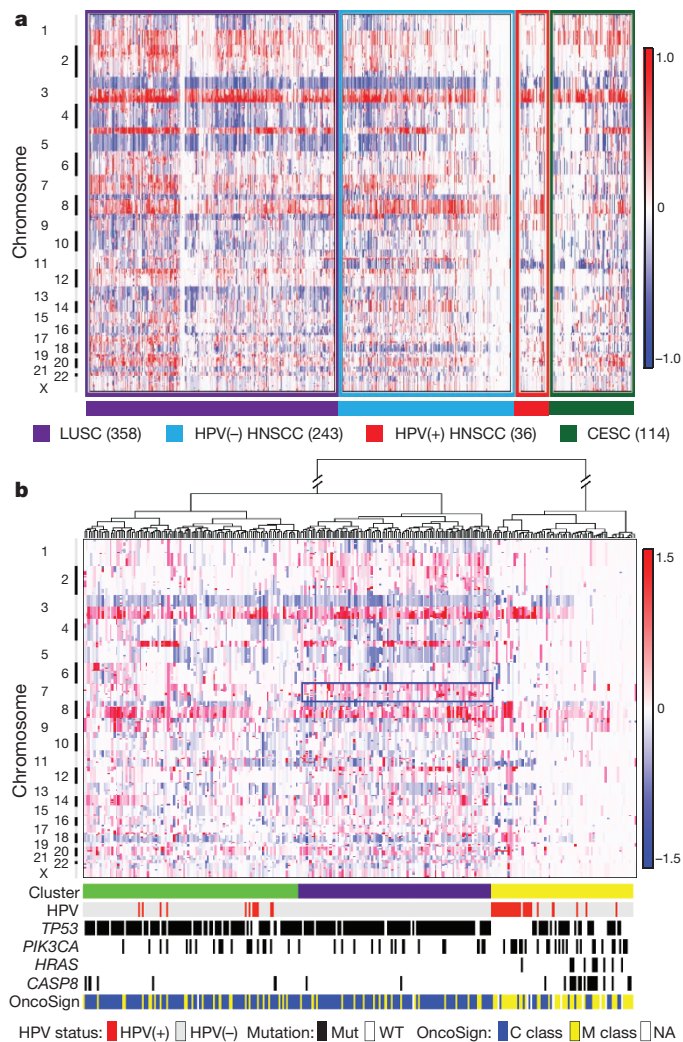


Figure 1 | DNA copy number alterations. **a**, Copy number alterations by anatomic site and HPV status for squamous cancers. Lung squamous cell carcinoma (LUSC, $n = 358$) and cervical squamous cell carcinoma (CESC, $n = 114$). **b**, Unsupervised analysis of copy number alteration of HNSCC ($n = 279$) with associated characteristics. The rectangle indicates chromosome 7 amplifications in the purple cluster. NA, not available.

carcinomas. HPV(+) tumours were also notable for focal amplification of *E2F1* and an intact 9p21.3 region containing the *CDKN2A* gene commonly deleted in HPV(-) tumours.

HPV(-) tumours featured novel co-amplifications of 11q13 (*CCND1*, *FADD* and *CTTN*) and 11q22 (*BIRC2* and *YAP1*), which also contain genes implicated in cell death/NF- κ B and Hippo pathways. HPV(-) tumours featured novel focal deletions in the nuclear set domain gene (*NSD1*) and tumour suppressor genes (for example, *FAT1*, *NOTCH1*, *SMAD4* and *CDKN2A*; Supplementary Fig. 2.3). Recurrent focal amplifications in receptor tyrosine kinases (for example, *EGFR*, *ERBB2* and *FGFR1*) also predominated in HPV(-) tumours. Notably, unsupervised clustering analysis of CNAs identified a mutually exclusive subset of predominantly oral cavity tumours with reduced CNAs, a pattern recently described in cancer as ‘M’ class (tumours driven by mutation rather than CNA)¹³ (Fig. 1b). This subset in particular contained a new three-gene pattern of activating mutations in *HRAS*, frequently with inactivating *CASP8* mutations, and wild-type *TP53*. We confirmed a previously reported favourable clinical outcome in tumours with few CNAs¹⁴. The three-gene constellation of wild-type *TP53* with mutant *HRAS* and *CASP8* suggested an alternative tumorigenesis pathway involving RAS and/or alterations in cell death/NF- κ B¹⁵. Unsupervised

analysis also suggested that clustering was a function of chromosome 7 amplification (including the *EGFR* locus) in a manner that largely excluded HPV(+) tumours.

To detect additional structural alterations, we interrogated whole-genome and RNA-Seq data (Supplementary Information section 3, Supplementary Data 3.1 and Supplementary Fig. 3.1). Known fusion oncogenes reported in solid tumours including those involving the *ALK*, *ROS* or *RET* genes were not observed in HNSCC. Previously reported *FGFR3-TACC3* fusions were present in two HPV(+) tumours (Supplementary Fig. 3.2). Only 1 out of 279 patients showed evidence of the type III isoform of *EGFR* (vIII), previously described in HNSCC¹⁶ (Supplementary Fig. 3.3). Although our investigation did not identify additional novel oncogenic fusions, several tumours demonstrated exon 1 of *EGFR* or *FGFR3* fused to non-recurrent partners, suggesting potential promoter swaps for the partner genes (Supplementary Data 3.1). A low prevalence of an alternative *MET* transcript with skipped exon 14 was identified in two HPV(-) tumours (Supplementary Fig. 3.4); this finding was reported to be an activating event in non-small cell lung cancer¹⁷. Structural alterations (homozygous deletions, intra- and inter-chromosomal fusions) were more commonly associated with loss of function in tumour suppressor genes, most prominently *CDKN2A* (Supplementary Figs 3.5 and 3.6), followed by *TP53*, *RBI*, *NOTCH1* and *FAT1* (Supplementary Figs 3.7–3.9), than with protein-coding fusion events. RNA-Seq data (Supplementary Data 3.3) demonstrated evidence of alternative splicing in genes not previously described in HNSCC including kallikrein 12 (*KLK12*) (Supplementary Fig. 3.11), as well as genes such as *TP63* with known importance in HNSCC (Supplementary Fig. 3.12).

By DNA analysis, most HPV(+) tumours demonstrated clear evidence of host genome integration, usually in a single genomic location per sample and almost always in association with amplifications of the host genome (Supplementary Fig. 3.10 and Supplementary Data 3.2). Interrogation of RNA transcripts confirmed transcription across the viral-human integration locus. However, none of the genes involved were recurrent, suggesting no single driver mechanism related to HPV integration. Similarly, none of the integration sites involved the *MYC* gene as reported in HPV(+) cell lines¹⁸.

Somatic mutations

Whole-exome sequencing identified somatically mutated genes, many located in regions of CNAs and annotated in the COSMIC database¹⁹ (Fig. 2). The mean sequencing coverage across targeted bases was 95 \times , with 82% of target bases above 30 \times coverage. In 279 samples, 12,159 synonymous somatic variants, 37,061 non-synonymous somatic variants, and 2,579 germline single base substitutions from the single nucleotide polymorphism database (dbSNP)²⁰ were detected (Supplementary Information section 4). Targeted re-sequencing of 394 unique regions (Supplementary Fig. 4.1) validated 99% of mutations. Interrogation of RNA for expression of the mutated alleles confirmed the variant in 86% of cases (Supplementary Information section 3.2 and Supplementary Fig. 3.1). In contrast to previous reports, the mutation rates did not differ by HPV status, although transversions at CpG sites were more frequent in HPV(-) tumours and a predominance of TpC mutations were noted in HPV(+) cases³ (Supplementary Fig. 1.1). Mutations were statistically enriched in 11 genes (Fig. 2). Among inactivating mutations (premature termination of the protein by nonsense, frameshift or splice-site mutations), four genes segregated exclusively or predominantly in HPV(-) tumours. Two were associated with cell cycle and survival (*CDKN2A* ($P < 0.01$) and *TP53* ($P < 0.01$)) and two were linked to Wnt/ β -catenin signalling (*FAT1* ($P < 0.01$) and *AJUBA* ($P = 0.14$))^{21,22}. We observed *TP53* mutation among HPV(-) samples at higher rates (86%) than have been previously reported¹⁹, while only 1 out of 36 HPV(+) cases had a non-synonymous *TP53* mutation. Previously unreported somatic mutations and deletions of *AJUBA* were primarily 5' inactivating events and clustered missense mutations in the functional LIM domain (Supplementary Fig. 4.2). *AJUBA* is a

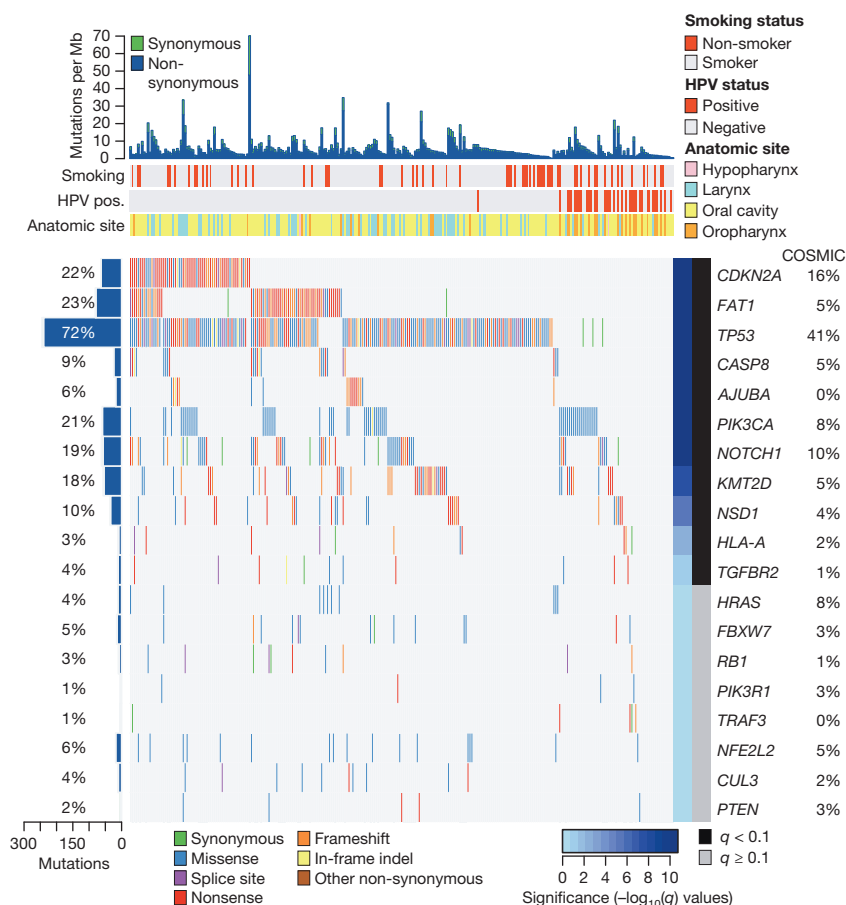


Figure 2 | Significantly mutated genes in HNSCC. Genes (rows) with significantly mutated genes (identified using the MutSigCV algorithm; $q < 0.1$) ordered by q value; additional genes with trends towards significance are also shown. Samples (columns, $n = 279$) are arranged to emphasize mutual exclusivity among mutations. Left, mutation percentage in TCGA. Right, mutation percentage in COSMIC ('upper aerodigestive tract' tissue). Top, overall number of mutations per megabase. Colour coding indicates mutation type.

centrosomal protein that regulates cell division, vertebrate ciliogenesis and left–right axis determination²³. Additionally, *AJUBA* is subject to EGFR-RAS-MAPK-dependent phosphorylation and implicated in Hippo growth and regeneration pathways conserved from *Drosophila* to mammals^{24,25}, in ataxia-telangiectasia mutated (*ATM*) and *ATM* and Rad-3-related (*ATR*)-mediated DNA damage response²⁶, and tumour invasion and migration²⁷.

A frequently mutated novel gene, the nuclear receptor binding SET domain protein 1 (*NSD1*), was identified in 33 HNSCCs. Alterations included inactivating mutations ($n = 29$) and focal homozygous deletions ($n = 4$). *NSD1* is a histone 3 Lys 36 (H3K36) methyltransferase, similar to *SETD2*, which is frequently mutated in the clear cell variant of renal cell carcinoma, and associated with DNA hypomethylation²⁸. Germline carriers of inactivating mutations in *NSD1* are associated with craniofacial abnormalities (Sotos syndrome), and malignancies including squamous carcinoma, implicating *NSD1* as a tumour suppressor gene²⁹. Interestingly, *NSD1* functions as an oncogene when fused to nucleoporin-98 (*NUP98*) t(5;11)(q35;p15.5) in haematological cancers with increased H3K36 trimethylation levels at *HOXA* genes and accompanying transcriptional activation³⁰. Translocations involving other dedicated H3K36 methyltransferase genes including *WHSC1* (also known as *MMSET* and *NSD2*) are reported in 20% of multiple myelomas. By contrast, *NSD1* loss has been associated with sporadic non-melanoma skin cancers³¹. Significant inactivating mutations were found in genes linked to squamous differentiation including in *NOTCH1* (19%), and other non-significant family members (*NOTCH2* 9%, and *NOTCH3* 5%, $q > 0.1$, non-significant), and the *TP63* target gene *ZNF750* (4%, $q > 0.1$, non-significant), which falls in a significantly deleted peak at 17q25.3. The analysis identified additional mutations including *TRAF3*, *RB1* and *NFE2L2*, among others with q values < 1 (non-significant). The frequently mutated apoptosis gene *CASP8* displayed clustered missense and other inactivating mutations in the first death effector,

intron and caspase peptidase domains. Statistically significant mutations in *KMT2D* (also known as *MLL2*) and *HLA-A* could contribute to defective immunosurveillance. Of known oncogenes, only *PIK3CA* achieved statistical significance ($q < 0.01$). Approximately one-quarter of the mutated *PIK3CA* cases displayed concurrent amplification, with an additional 20% of tumours containing focal amplification without evidence of mutation. Seventy-three per cent of *PIK3CA* mutations localized to Glu542Lys, Glu545Lys and His1047Arg/Leu hotspots that promote activation, with the remaining mutations of uncertain function. Recurrent activating mutations of *HRAS* in the GTPase domain in residues 11–13 approached statistical significance ($q = 0.2$).

We extended our unsupervised genome-wide analysis of significantly mutated genes as well as genes reported in COSMIC to a subgroup analysis by anatomic sites, tumour versus normal status, HPV status and four previously validated gene expression subtypes^{32,33} (Supplementary Information section 5, Supplementary Figs 5.1–5.4 and Supplementary Data 4.1 and 5.1–5.4). Additional mutations included *TRAF3*, *RB1* and *NFE2L2*, among others with q values < 1 , and we observed statistical evidence for mutations of *HRAS* ($q = 0$ in COSMIC subset) and other genes. Sporadic inactivating mutations and deletions of *TGFBR2* were identified primarily in oral cavity tumours, consistent with its role in promoting squamous tumorigenesis in mouse models³⁴. Investigating COSMIC database mutations focused attention on the significant deletion peak at 4q31.3 containing the gene *FBXW7*, a ubiquitin ligase targeting cyclin E and *NOTCH* genes, in which we identified mutations that included recurrent Arg505Gly/Leu substitutions ($n = 14$). Genes with at least one identical mutation previously reported in COSMIC include *SCN9A*, *CHEK2*, *PTCH1* and *PIK3R1*. We further focused on somatic alterations and protein expression that represent plausible therapeutic targets (Fig. 3, Supplementary Information sections 6 and 7, Supplementary Figs 6.1 and 6.2 and Supplementary Data 6.1 and 6.2).

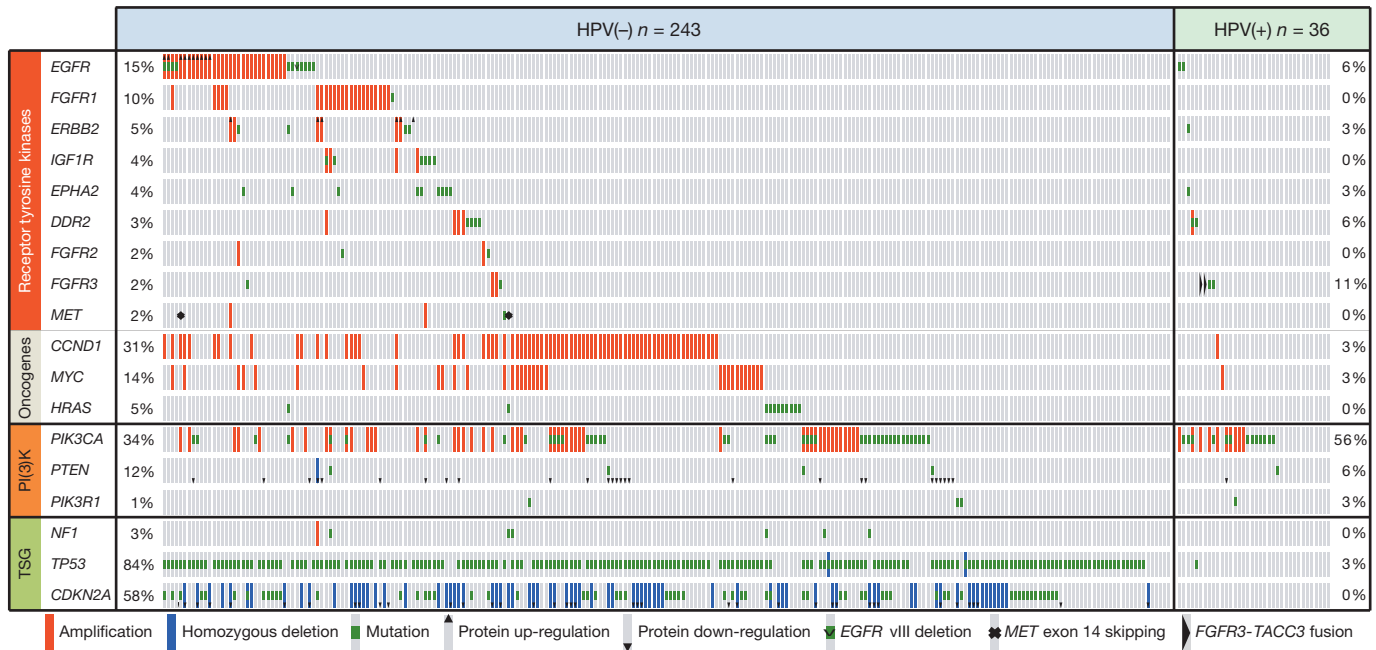


Figure 3 | Candidate therapeutic targets and driver oncogenic events. Alteration events for key genes are displayed by sample ($n = 279$). TSG, tumour suppressor gene.

Integrated genome analysis and pathways

Correlative genetic alteration analysis identified numerous pairwise significant findings (Supplementary Information section 7 and Supplementary Fig. 7.1). In particular, co-amplification of 11q13 containing *CCND1*, *FADD* and *CTTN* and a narrow segment of 11q22 containing the genes with equal evidence for *YAP1* and *BIRC2* was further characterized (Supplementary Fig. 7.2). Chromosome 11q22 was focally but rarely amplified in the absence of co-amplification of 11q13. This novel finding suggests that the selection pressure for this co-amplification stems from the interaction of *BIRC2* with *FADD* and the caspase cascade that inhibits cell death. Notably, the vast majority of tumours with the 11q13 amplification had large deletions in the telomeric region of 11q22, including other genes known to be important in cell death in cancer such as *ATM* and *CASP1*, 4, 5 and 12. Amplification of 11q13 was anti-correlated with *CASP8* mutations, suggesting an alternative function of *CASP8* and *FADD* in cell death/NF- κ B activation³⁵.

We investigated whether clinical factors, single gene alterations and statistically significant pairwise gene correlations (Supplementary Fig. 7.1 and Supplementary Data 7.1) might segregate previously defined molecular subtypes and/or anatomic sub-sites (Supplementary Data 5.1–5.4). We confirmed reported gene expression subtypes (atypical (24%), mesenchymal (27%), basal (31%) and classical (18%)), and assessed the subtypes for enrichment of somatic alterations^{32,33} (Supplementary Data 7.1). Notably, *TP53* mutation, *CDKN2A* loss of function, chromosome 3q amplification, alteration of oxidative stress genes (*KEAP1*, *NFE2L2* or *CUL3*), heavy smoking history (Supplementary Table 1.1) and larynx sub-site co-occurred in most classical subtype tumours (Fig. 4a and Supplementary Information section 7.2), similar to LUSC⁵ (Supplementary Figs 5.1 and 5.2). Collectively, these findings suggest that the *NFE2L2* oxidative stress pathway is a tobacco-related signature across anatomic tumour sites. By contrast, the basal subtype demonstrated inactivation of *NOTCH1* with intact oxidative stress

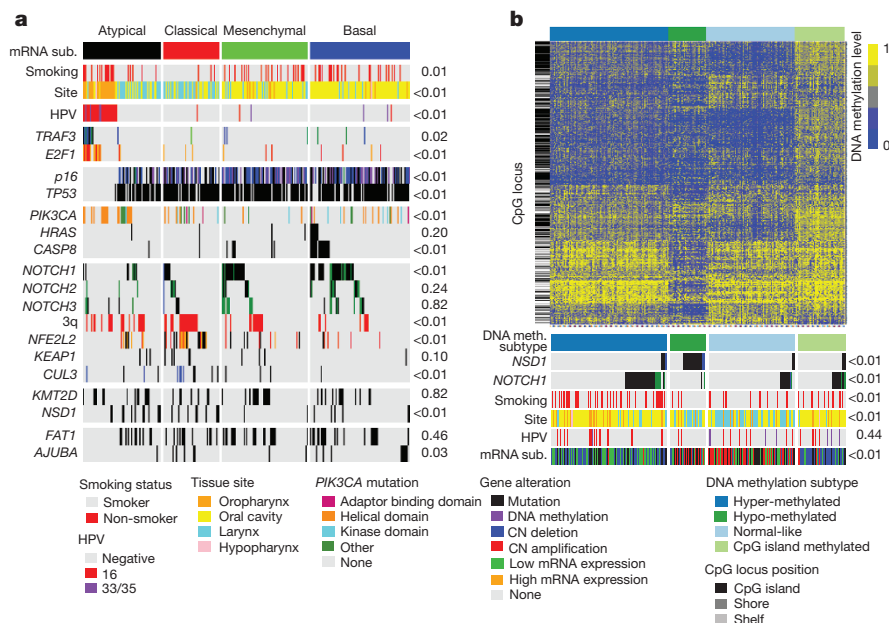
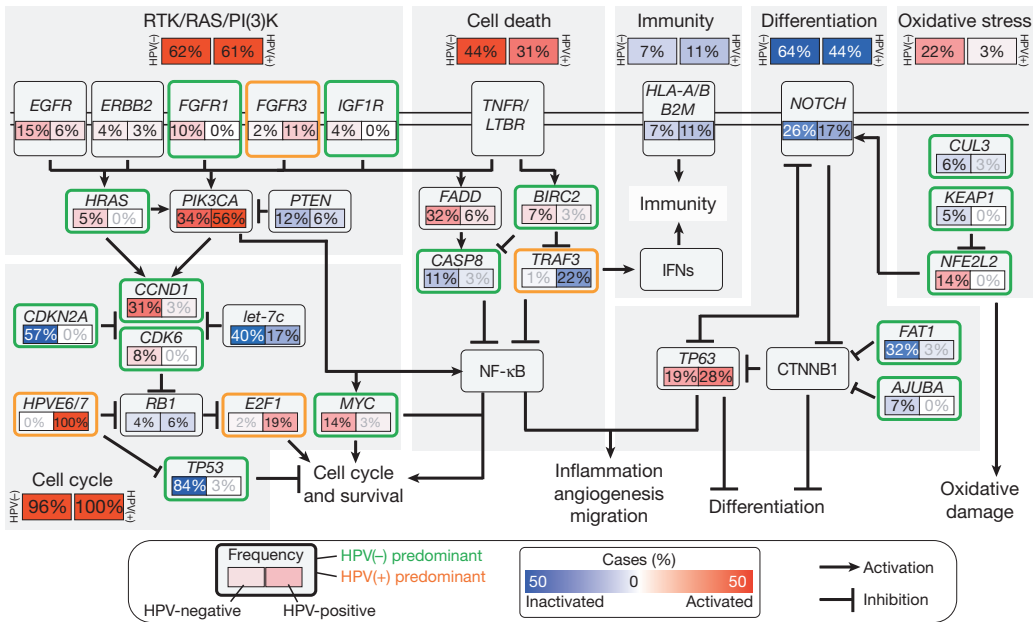


Figure 4 | Integrated analysis of genomic alterations. **a, b**, Samples ($n = 279$) are displayed in columns and grouped by gene expression (**a**) or methylation (**b**) subtype (sub.). Unadjusted two-sided Fisher's exact test *P* values assess the association of each genomic alteration. Methylation probe location of CpG islands, shores and shelves are shown on the left of **b**. Annotation shows HPV status and subtype (16, 33 and 35). CN, copy number.



signalling and fewer alterations of chromosome 3q. Analysis of the 3q locus highlighted a marked relative decrease of *SOX2* expression in basal tumours relative to all other HNSCC and tumour adjacent normal samples (Supplementary Fig. 5.4), supporting the interaction of transcription factors *SOX2*, *TP63*, *NFE2L2* and *NOTCH1* as driving differences between expression subtypes. Additionally, the basal subtype included most tumours with the *HRAS*–*CASP8* co-mutation and most co-amplified 11q13/q22 tumours. These findings along with *HRAS* mutations implicate disrupted cell death as a major alteration in this subtype³⁶ (Supplementary Fig. 7.2). The atypical subtype was characterized by a lack of chromosome 7 amplifications (Supplementary Fig. 5.3), enrichment of HPV(+) tumours with activating mutations in exon 9 that contains the *PIK3CA* helical domain. By contrast, the mesenchymal subtype showed high levels of alteration in innate immunity genes, in particular high expression of natural killer cell marker *CD56* and a low frequency of HLA class I mutations (Supplementary Fig. 7.3). Among the significantly mutated genes, *TP53* ($P < 0.001$), *CASP8* ($P = 0.01$), *NSD1* ($P = 0.01$) and *CDKN2A* (0.06) were the most differentially mutated across anatomic sites (Supplementary Data 4.1). Most *CASP8* mutations (22 out of 24, 92%) were in oral cavity tumours, whereas *TP53*, *NSD1* and *CDKN2A* demonstrated decreased mutation rates in oropharyngeal tumours relative to other sites.

Unsupervised analysis of gene expression by HPV status and of reverse-phase protein arrays (Supplementary Information section 6), DNA methylation (Supplementary Information section 8), and miRNA platforms (Supplementary Information section 9, Supplementary Table 9.1, Supplementary Figs 7.4–7.9 and Supplementary Data 7.1) showed high correlation across platforms ($P < 0.01$; Fig. 4, Supplementary Information section 7.9 and Supplementary Data 7.2) and coordinated alterations of genes including the epithelial–mesenchymal transition signature³⁷ (Supplementary Figs 7.4, 7.7 and 7.8). However, within the broader cross-platform agreement, individual unsupervised clustering of miRNA, reverse-phase protein arrays and DNA methylation data provided insight into the association of molecular subtypes with single gene alterations. The most notable example was the detection of hypomethylation and loss-of-function mutations of *NSD1*, and wild-type *NOTCH1* in atypical and classical gene expression subtypes (Fig. 4b and Supplementary Table 7.1).

Supervised analyses detected genomic features (miRNA, gene expression and DNA methylation) associated with anatomic site (Supplementary Figs 5.5–5.8 and Supplementary Data 5.1–5.4). A supervised integrated analysis identified target genes that are inversely regulated by

miRNAs in HNSCC (Supplementary Information section 7.4). Among these miRNA–messenger RNA networks, *let-7c-5p* and *miR-100-5p* exhibited a correlation between low copy number and expression. *let-7c-5p* and *miR-100-5p* were decreased in tumours compared to normal (Supplementary Fig. 7.10). For these miRNAs, deletion was highly associated with increased expression of target genes, including the cell cycle regulator *CDK6*, transcription factor *E2F1* (ref. 38), mitosis regulator *PLK1* (ref. 39), and transcription factor *HMG2A* (ref. 40; Supplementary Figs 7.10, Supplementary Tables 7.2 and 7.3).

Integrative bioinformatics analysis identified a limited number of pathways targeted by frequent genome alterations (Fig. 5, Supplementary Information section 7, Supplementary Figs 7.11–7.15 and Supplementary Data 7.3). Among receptor tyrosine kinases, *EGFR/ERBB2* or *FGFR1/3* alterations are the most frequent. Among downstream targets of the receptor tyrosine kinase (RTK)/RAS/phosphatidylinositol-3-OH kinase (PI(3)K) pathway, *PIK3CA* dominates with occasional *HRAS* and *PTEN* alterations. Further downstream, nearly every tumour has alteration of genes governing the cell cycle. The tumour suppressors *TP53* and *CDKN2A*, oncogenes *CCND1* and *MYC*, and the newly identified miRNA *let-7c*, are most often altered in HPV(-) tumours, whereas viral genes *E6*, *E7* and *E2F1* predominate in HPV(+) cases. In addition, we report frequent alterations in genes involved in cell death, NF- κ B-mediated survival, or immunity pathways^{15,35}. Co-amplification of *FADD* \pm *BIRC2*, or *CASP8* \pm *HRAS* mutations define exclusive HPV(-) subsets, whereas *TRAF3* loss characterizes an HPV(+) subset. These alterations along with *PIK3CA* and *TP63* converge on NF- κ B transcription factors that promote cell survival, migration, inflammation and angiogenesis^{41,42}. Furthermore, *TRAF3* and/or *HLA* loss are implicated in deregulation of innate antiviral and adaptive anti-tumour immunity^{43,44}. Further alterations of *NOTCH*, *TP63* and other genes in HPV(-) tumours (*FAT1* and *AJUBA*) recently linked functionally to β -catenin (*CTNNB1*) are also detected^{21,22,45}. Finally, we highlight a previously underappreciated role for a key transcription factor regulator of oxidative stress, *NFE2L2*, and its protein complex partners *CUL3* and *KEAP1* in HPV(-) HNSCCs.

Conclusion

The TCGA study represents the most comprehensive integrative genomic analysis of HNSCC. Loss of *TRAF3*, activating mutations of *PIK3CA*, and amplification of *E2F1* in HPV(+) oropharyngeal cancers point to aberrant activation of NF- κ B, other oncogenic pathways, and cell cycle, as critical in the pathogenesis and development of new

targeted therapies for these tumours. In HPV(−) HNSCCs, mutually exclusive subsets containing amplicons on 11q with *CCND1*, *FADD*, *BIRC2* and *YAP1*, or concurrent mutations of *CASP8* with *HRAS*, also target cell cycle, death, NF- κ B and other oncogenic pathways. Recent studies predict that the inactivation of *AJUBA*, as well as *FAT1* and *NOTCH1*, may converge to uncheck Wnt/ β -catenin signalling, implicated in deregulation of cell polarity and differentiation. The 3q amplicon found in both HPV(+) and (−) HNSCCs includes transcription factors *TP63*, *SOX2* and signal molecule *PIK3CA*, which are also implicated in homeostasis of epithelial stem cells and differentiation. Among these, the biological function and agents targeting BIRCs, PI(3)K, Wnt/ β -catenin and NOTCH are under investigation. Collectively, these findings provide new insights into HNSCC and suggest that shared and unique alterations might be leveraged to accelerate progress in prevention and therapy across tumour types.

Received 19 October 2013; accepted 1 December 2014.

- Ferlay, J. *et al.* Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer* **127**, 2893–2917 (2010).
- Ang, K. K. *et al.* Human papillomavirus and survival of patients with oropharyngeal cancer. *N. Engl. J. Med.* **363**, 24–35 (2010).
- Stransky, N. *et al.* The mutational landscape of head and neck squamous cell carcinoma. *Science* **333**, 1157–1160 (2011).
- Agrawal, N. *et al.* Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in *NOTCH1*. *Science* **333**, 1154–1157 (2011).
- The Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489**, 519–525 (2012).
- Oganesyan, G. *et al.* Critical role of TRAF3 in the Toll-like receptor-dependent and -independent antiviral response. *Nature* **439**, 208–211 (2006).
- Karim, R. *et al.* Human papillomavirus (HPV) upregulates the cellular ubiquitinase UCHL1 to suppress the keratinocyte's innate immune response. *PLoS Pathog.* **9**, e1003384 (2013).
- Eliopoulos, A. G. *et al.* CD40-induced growth inhibition in epithelial cells is mimicked by Epstein–Barr Virus-encoded LMP1: involvement of TRAF3 as a common mediator. *Oncogene* **13**, 2243–2254 (1996).
- Imbeault, M. *et al.* Acquisition of host-derived CD40L by HIV-1 *in vivo* and its functional consequences in the B-cell compartment. *J. Virol.* **85**, 2189–2200 (2011).
- Ni, C. Z. *et al.* Molecular basis for CD40 signaling mediated by TRAF3. *Proc. Natl Acad. Sci. USA* **97**, 10395–10399 (2000).
- Chung, G. T. *et al.* Constitutive activation of distinct NF- κ B signals in EBV-associated nasopharyngeal carcinoma. *J. Pathol.* **231**, 311–322 (2013).
- Annunziata, C. M. *et al.* Frequent engagement of the classical and alternative NF- κ B pathways by diverse genetic abnormalities in multiple myeloma. *Cancer Cell* **12**, 115–130 (2007).
- Ciriello, G. *et al.* Emerging landscape of oncogenic signatures across human cancers. *Nature Genet.* **45**, 1127–1133 (2013).
- Smeets, S. J. *et al.* Genetic classification of oral and oropharyngeal carcinomas identifies subgroups with a different prognosis. *Cellular Oncol.* **31**, 291–300 (2009).
- Mayo, M. W. *et al.* Requirement of NF- κ B activation to suppress p53-independent apoptosis induced by oncogenic Ras. *Science* **278**, 1812–1815 (1997).
- Sok, J. C. *et al.* Mutant epidermal growth factor receptor (EGFRvIII) contributes to head and neck cancer growth and resistance to EGFR targeting. *Clin. Cancer Res.* **12**, 5064–5073 (2006).
- Kong-Beltran, M. *et al.* Somatic mutations lead to an oncogenic deletion of met in lung cancer. *Cancer Res.* **66**, 283–289 (2006).
- Popescu, N. C., DiPaolo, J. A. & Amsbaugh, S. C. Integration sites of human papillomavirus 18 DNA sequences on HeLa cell chromosomes. *Cytogenet. Cell Genet.* **44**, 58–62 (1987).
- Forbes, S. A. *et al.* The Catalogue of Somatic Mutations in Cancer (COSMIC). *Curr. Protoc. Hum. Genet.* Chapter 10, Unit 10.11 (2008).
- Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311 (2001).
- Morris, L. G. *et al.* Recurrent somatic mutation of *FAT1* in multiple human cancers leads to aberrant Wnt activation. *Nature Genet.* **45**, 253–261 (2013).
- Haraguchi, K. *et al.* Ajuba negatively regulates the Wnt signaling pathway by promoting GSK-3 β -mediated phosphorylation of β -catenin. *Oncogene* **27**, 274–284 (2008).
- Nagai, Y. *et al.* The LIM protein Ajuba is required for ciliogenesis and left-right axis determination in medaka. *Biochem. Biophys. Res. Commun.* **396**, 887–893 (2010).
- Sun, G. & Irvine, K. D. Ajuba family proteins link JNK to Hippo signaling. *Sci. Signal.* **6**, ra81 (2013).
- Reddy, B. V. & Irvine, K. D. Regulation of Hippo signaling by EGFR-MAPK signaling through Ajuba family proteins. *Dev. Cell* **24**, 459–471 (2013).
- Kalan, S., Matveyenko, A. & Loayza, D. LIM protein Ajuba participates in the repression of the ATR-mediated DNA damage response. *Front. Genet.* **4**, 95 (2013).
- Nola, S. *et al.* Ajuba is required for Rac activation and maintenance of E-cadherin adhesion. *J. Cell Biol.* **195**, 855–871 (2011).
- The Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* **499**, 43–49 (2013).
- Fickie, M. R. *et al.* Adults with Sotos syndrome: review of 21 adults with molecularly confirmed *NSD1* alterations, including a detailed case report of the oldest person. *Am. J. Med. Genet. A.* **155**, 2105–2111 (2011).
- Wang, G. G., Cai, L., Pasillas, M. P. & Kamps, M. P. NUP98–*NSD1* links H3K36 methylation to *Hox-A* gene activation and leukaemogenesis. *Nature Cell Biol.* **9**, 804–812 (2007).
- Quintana, R. M. *et al.* A transposon-based analysis of gene mutations related to skin cancer development. *J. Invest. Dermatol.* **133**, 239–248 (2013).
- Chung, C. H. *et al.* Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression. *Cancer Cell* **5**, 489–500 (2004).
- Walter, V. *et al.* Molecular subtypes in head and neck cancer exhibit distinct patterns of chromosomal gain and loss of canonical cancer genes. *PLoS ONE* **8**, e56823 (2013).
- Lu, S. L. *et al.* Loss of transforming growth factor- β type II receptor promotes metastatic head-and-neck squamous cell carcinoma. *Genes Dev.* **20**, 1331–1342 (2006).
- Oberst, A. & Green, D. R. It cuts both ways: reconciling the dual roles of caspase 8 in cell death and survival. *Nature Rev. Mol. Cell Biol.* **12**, 757–763 (2011).
- Park, S. J. *et al.* Opposite role of Ras in tumor necrosis factor- α -induced cell cycle regulation: competition for Raf kinase. *Biochem. Biophys. Res. Commun.* **287**, 1140–1147 (2001).
- Byers, L. A. *et al.* An epithelial-mesenchymal transition gene signature predicts resistance to EGFR and PI3K inhibitors and identifies Axl as a therapeutic target for overcoming EGFR inhibitor resistance. *Clin. Cancer Res.* **19**, 279–290 (2013).
- Wong, J. V., Dong, P., Nevins, J. R., Mathey-Prevot, B. & You, L. Network calisthenics: control of E2F dynamics in cell cycle entry. *Cell Cycle* **10**, 3086–3094 (2011).
- Sanhaji, M. *et al.* Polo-like kinase 1 inhibitors, mitotic stress and the tumor suppressor p53. *Cell Cycle* **12**, 1340–1351 (2013).
- Morishita, A. *et al.* HMGA2 is a driver of tumor metastasis. *Cancer Res.* **73**, 4289–4299 (2013).
- Bancroft, C. C. *et al.* Effects of pharmacologic antagonists of epidermal growth factor receptor, PI3K and MEK signal kinases on NF- κ B and AP-1 activation and IL-8 and VEGF expression in human head and neck squamous cell carcinoma lines. *Int. J. Cancer* **99**, 538–548 (2002).
- Yang, X. *et al.* Δ Np63 versatily regulates a broad NF- κ B gene program and promotes squamous epithelial proliferation, migration, and inflammation. *Cancer Res.* **71**, 3688–3700 (2011).
- Keating, P. J. *et al.* Frequency of down-regulation of individual HLA-A and -B alleles in cervical carcinomas in relation to TAP-1 expression. *Br. J. Cancer* **72**, 405–411 (1995).
- Esteban, F. *et al.* Lack of MHC class I antigens and tumour aggressiveness of the squamous cell carcinoma of the larynx. *Br. J. Cancer* **62**, 1047–1051 (1990).
- Dotto, G. P. Crosstalk of Notch with p53 and p63 in cancer growth control. *Nature Rev. Cancer* **9**, 587–595 (2009).

Supplementary Information is available in the online version of the paper.

Acknowledgements This study was supported by National Institutes of Health (NIH) grants: P50CA097190, P50CA16672, U54 HG003273, U54 HG003067, U54 HG003079, U24 CA143799, U24 CA143835, U24 CA143840, U24 CA143843, U24 CA143845, U24 CA143848, U24 CA143858, U24 CA143866, U24 CA143867, U24 CA143882, U24 CA143883, U24 CA144025 and R01 CA 095419. Supported by The Bobby F. Garrett Fund for Head and Neck Cancer Research and NIDCD Intramural Projects ZIA-DC-000016, 73 and 74.

Author Contributions The TCGA research network contributed collectively to this study. Biospecimens were provided by the tissue source sites and processed by the biospecimen core resource. Data generation and analyses were performed by the genome sequencing centres, cancer genome characterization centres and genome data analysis centres. All data were released through the data coordinating centre. Project activities were coordinated by the National Cancer Institute and National Human Genome Research Institute project teams. We also acknowledge the following TCGA investigators who made substantial contributions to the project: C.V.W. (manuscript coordinator); V.W. (data coordinator); P.S.H. and D.N.H. (analysis coordinators); Y.W. (data coordinating centre representative); R.A., J.B., T.E.C., J.J.L., M.Parfenov, A.D.T., V.W. and M.D.W. (clinical and HPV analysis); A.D.C. (DNA copy number analysis); J.Cho, A.H., M.Parfenov and M.D.W. (mutation calling); L.C. and L.Danilova (DNA methylation analysis); R.Bowlby, J.P.B., Z.C., H.Cheng, A.H., F.F.L., A.G.R., A.D.S. and C.V.W. (miRNA sequence analysis); L.A.B., L.Diao, P.K., W.L., S.N. and J.W. (reverse-phase protein arrays analysis); L.A.B., L.Diao, S.R.J., P.K., V.W., M.D.W. and N.Z. (mRNA sequence analysis); S.Benz, Z.C., R.L.F., S.N., C.R.P., N.Schultz, T.Y.S., C.V.W. and V.W. (pathway analysis); A.E.N., J.R.G. and D.N.H. (project chairs).

Author Information The primary and processed data used to generate the analyses can be downloaded by registered users from The Cancer Genome Atlas (<https://tcga-data.nci.nih.gov/tcga/tcgaDownload.jsp>, <https://cghub.ucsc.edu/> and https://tcga-data.nci.nih.gov/docs/publications/hnsc_2014/). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.N.H. (hayes@med.unc.edu), J.R.G. (grandisjr@upmc.edu) or A.E.N. (anaggar@mdanderson.org).



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported Licence. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons licence, users will need to obtain permission from the licence holder

to reproduce the material. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-sa/3.0>

The Cancer Genome Atlas Research Network (Participants are arranged by area of contribution and then by institution.)

Genome sequencing centre: Broad Institute Michael S. Lawrence¹, Carrie Sougnez¹, Lee Lichtenstein¹, Kristian Cibulskis¹, Eric Lander¹, Stacey B. Gabriel¹, Gad Getz^{1,2}

Genome characterization and data analysis centres: BC Cancer Agency Adrian Ally³, Miruna Balasundaram³, Inanc Birol³, Reanne Bowlby³, Denise Brooks³, Yaron S. N. Butterfield³, Rebecca Carlsen³, Dean Cheng³, Andy Chu³, Noreen Dhalla³, Ranabir Guin³, Robert A. Holt³, Steven J. M. Jones³, Darlene Lee³, Haiyan I. Li³, Marco A. Marra³, Michael Mayo³, Richard A. Moore³, Andrew J. Mungall³, A. Gordon Robertson³, Jacqueline E. Schein³, Payal Sipahimalani³, Angela Tam³, Nina Thiessen³, Tina Wong³;
Harvard Medical School/Brigham & Women's Hospital/MD Anderson Cancer Center Alexei Protopopov⁴, Nety Santos⁵, Semin Lee⁵, Michael Parfenov⁵, Jianhua Zhang⁴, Harshad S. Mahadeshwar⁴, Jiabin Tang⁴, Xiaojia Ren⁵, Sahil Seth⁴, Psalm Haseley⁵, Dong Zeng⁴, Lixing Yang⁵, Andrew W. Xu⁵, Xingzhi Song⁴, Angeliki Pantazi⁵, Christopher A. Bristow⁴, Angela Hadjiapanayis^{5,6}, Jonathan Seidman⁵, Lynda Chin^{1,4}, Peter J. Park^{5,6,7}, Raju Kucherlapati^{5,6};
The University of Texas MD Anderson Cancer Center Rehan Akbani⁴, Tod Casavant⁴, Wenbin Liu⁴, Yiling Lu⁴, Gordon Mills⁴, Thomas Motter⁴, John Weinstein⁴, Lixia Diaol⁴, Jing Wang⁴, You Hong Fan⁴;
University of Kentucky Jinze Liu⁸, Kai Wang⁸;
University of North Carolina at Chapel Hill J. Todd Auman⁹, Saianand Balu⁹, Thomas Bodenheimer⁹, Elizabeth Buda⁹, D. Neil Hayes⁹, Katherine A. Hoadley⁹, Alan P. Hoyle⁹, Stuart R. Jefferys⁹, Corbin D. Jones⁹, Patrick K. Kimes⁹, Yufeng Liu⁹, J. S. Marron⁹, Shaowu Meng⁹, Piotr A. Mieczkowski⁹, Lisle E. Mose⁹, Joel S. Parker⁹, Charles M. Perou⁹, Jan F. Prins⁹, Jeffrey Roach⁹, Yan Shi⁹, Janae V. Simons⁹, Darshan Singh⁹, Matthew G. Soloway⁹, Donghui Tan⁹, Umadevi Veluvolu⁹, Vonn Walter⁹, Scot Waring⁹, Matthew D. Wilkerson⁹, Junyuan Wu⁹, Ni Zhao⁹;
Broad Institute Andrew D. Cherniack¹, Peter S. Hammerman^{1,10}, Aaron D. Tward^{1,5}, Chandra Sekhar Pedamallu^{1,10}, Gordon Saksena¹, Joonil Jung¹, Akinyemi I. Ojesina^{1,10}, Scott L. Carter¹, Travis I. Zack^{1,10,11}, Steven E. Schumacher^{1,10}, Rameen Beroukhi^{1,5,10}, Samuel S. Freeman¹, Matthew Meyerson^{1,5,10}, Juok Cho¹, Lynda Chin^{1,4}, Gad Getz^{1,2}, Michael S. Noble¹, Daniel DiCara¹, Hailei Zhang¹, David I. Heiman¹, Nils Gehlenborg^{1,5}, Doug Voelt¹, Pei Lin¹, Scott Frazer¹, Petar Stojanov^{1,10}, Yingchun Liu¹, Lihua Zou¹, Jaegil Kim¹, Carrie Sougnez¹, Stacey B. Gabriel¹, Michael S. Lawrence¹;
Baylor College of Medicine Donna Muzny¹², Harshavardhan Doddapaneni¹², Christie Kovar¹², Jeff Reid¹², Donna Morton¹², Yi Han¹², Walker Hale¹², Hsu Chao¹², Kyle Chang¹², Jennifer A. Drummond¹², Richard A. Gibbs¹², Nipun Kakkar¹², David Wheeler¹², Liu Xi¹²;
Memorial Sloan-Kettering Cancer Center Giovanni Ciriello¹³, Marc Ladanyi¹³, William Lee¹³, Ricardo Ramirez¹³, Chris Sander¹³, Ronglai Shen¹³, Rileen Sinha¹³, Nils Weinhold¹³, Barry S. Taylor¹³, B. Arman Aksoy¹³, Gideon Dresdner¹³, Jianjiong Gao¹³, Benjamin Gross¹³, Anders Jacobsen¹³, Boris Reva¹³, Nikolaus Schultz¹³, S. Onur Sumer¹³, Yichao Sun¹³, Timothy A. Chan¹³, Luc G. Morris¹³;
University of California Santa Cruz/Buck Institute Joshua Stuart¹⁴, Stephen Benz¹⁴, Sam Ng¹⁴, Christopher Benz¹⁴, Christina Yau¹⁴;
Johns Hopkins University/Sidney Kimmel Comprehensive Cancer Center Stephen B. Baylin¹⁵, Leslie Cope¹⁵, Ludmila Danilova¹⁵, James G. Herman¹⁵;
University of Southern California Moiz Bootwalla¹⁶, Dennis T. Maglinte¹⁶, Peter W. Laird¹⁶, Timothy Triche Jr¹⁶, Daniel J. Weisenberger¹⁶, David J. Van Den Berg¹⁶

Disease working group: Nishant Agrawal¹⁷, Justin Bishop¹⁷, Paul C. Boutros¹⁸, Jeff P. Bruce¹⁹, Lauren Averett Byers⁴, Joseph Califano¹⁷, Thomas E. Carey²⁰, Zhong Chen²¹, Hui Cheng²¹, Simion I. Chiosea²², Ezra Cohen²³, Brenda Diergaarde²², Ann Marie Egloff²², Adel K. El-Naggar⁴, Robert L. Ferris²², Mitchell J. Frederick⁴, Jennifer R. Grandis²², Yan Guo²⁴, Robert I. Haddad¹⁰, Peter S. Hammerman^{1,10}, Thomas Harris²⁵, D. Neil Hayes⁹, Angela B. Y. Hui¹⁹, J. Jack Lee⁴, Scott M. Lippman²⁶, Fei-Fei Liu²⁷, Jonathan B. McHugh²⁰, Jeff Myers⁴, Patrick Kwok Shing Ng⁴, Bayardo Perez-Ordóñez²⁸, Curtis R. Pickering⁴, Michael Prystowsky²⁵, Marjorie Romkes²², Anthony D. Saleh²¹, Maureen A. Sartor²⁰, Raja Seethala²², Tangy U. Seiwert²³, Han Si²¹, Aaron D. Tward^{1,5}, Carter Van Waes²¹, Daryl M. Waggott¹⁸, Maciej Wiznerowicz²⁹, Wendell G. Yarbrough³⁰, Jiexin Zhang⁴, Zhixiang Zuo²³

Biospecimen core resource: International Genomics Consortium Ken Burnett³¹, Daniel Crain³¹, Johanna Gardner³¹, Kevin Lau³¹, David Mallery³¹, Scott Morris³¹, Joseph Paulauskis³¹, Robert Penny³¹, Candace Shelton³¹, Troy Shelton³¹, Mark Sherman³¹, Peggy Yena³¹;
Nationwide Children's Hospital Aaron D. Black³², Jay Bowen³², Jessica Frick³², Julie M. Gastier-Foster³², Hollie A. Harper³², Kristen Leraas³², Tara M. Lichtenberg³², Nilsa C. Ramirez³², Lisa Wise³², Erik Zmuda³²

Data coordinating centre: Julien Baboud³³, Mark A. Jensen³³, Ari B. Kahn³³, Todd D. Pihl³³, David A. Pot³³, Deepak Srinivasan³³, Jessica S. Walton³³, Yunhu Wan³³

Project office: Robert A. Burton³⁴, Tanja Davidsson³⁵, John A. Demchok³⁵, Greg Eley³⁶, Martin L. Ferguson³⁵, Kenna R. Mills Shaw³⁵, Bradley A. Ozenberger³⁷, Margi Sheth³⁵, Heidi J. Sofia³⁷, Roy Tarnuzzer³⁵, Zhining Wang³⁵, Liming Yang³⁵, Jean Claude Zenklusen³⁵

Tissue source sites: Analytical Biological Services Charles Saller³⁸, Katherine Tarvin³⁸;
Fred Hutchinson Cancer Research Center Chu Chen³⁹;
Georgia Regents University Roni Bollag⁴⁰, Paul Weinberger⁴⁰;
Greater Poland Cancer Centre Wojciech Golusiński²⁹, Pawel Golusiński²⁹, Matthew Ibbts²⁹, Konstanty Korski²⁹, Andrzej Mackiewicz²⁹, Wiktoria Suchorska²⁹, Bartosz Szybiak²⁹, Maciej Wiznerowicz²⁹;
International Genomics Consortium Ken Burnett³¹, Erin Curley³¹, Johanna Gardner³¹, David Mallery³¹, Robert Penny³¹, Troy Shelton³¹, Peggy Yena³¹;
Indiana University Simon Cancer Center Christina Beard⁴¹, Colleen Mitchell⁴¹, George Sandusky⁴¹;
Johns Hopkins University Nishant Agrawal¹⁷, Julie Ahn¹⁷, Justin Bishop¹⁷, Joseph Califano¹⁷, Zubair Khan¹⁷;
Princess Margaret Cancer Centre Jeff P. Bruce¹⁹, Angela B. Y. Hui¹⁹, Jonathan Irish²⁷, Fei-Fei Liu²⁷, Bayardo Perez-Ordóñez²⁸, John Waldron²⁷;
Ontario Institute for Cancer Research Paul C. Boutros¹⁸, Daryl M. Waggott¹⁸;
The University of Texas MD Anderson Cancer Center Jeff Myers⁴, William N. William Jr⁴;
University of California San Diego Scott M. Lippman²⁶;
University of Miami Sophie Egea⁴², Carmen Gomez-Fernandez⁴², Lynn Herbert⁴²;
University of Michigan Carol R. Bradford²⁰, Thomas E. Carey²⁰, Douglas B. Chepeha²⁰, Andrea S. Haddad²⁰, Tamara R. Jones²⁰, Christine M. Komarck²⁰, Mayya Malakh²⁰, Jonathan B. McHugh²⁰, Jeffrey S. Moyer²⁰, Ariane Nguyen²⁰, Lisa A. Peterson²⁰, Mark E. Prince²⁰, Laura S. Rozek²⁰, Maureen A. Sartor²⁰, Evan G. Taylor²⁰, Heather M. Walline²⁰, Gregory T. Wolf²⁰;
University of North Carolina at Chapel Hill Lori Boice⁹, Bishamjit S. Chera⁹, William K. Funkhouser⁹, Margaret L. Gulley⁹, Trevor G. Hackman⁹, D. Neil Hayes⁹, Michele C. Hayward⁹, Mei Huang⁹, W. Kimryn Rathmell⁹, Ashley H. Salazar⁹, William W. Shockley⁹, Carol G. Shores⁹, Leigh Thorne⁹, Mark C. Weisler⁹, Sylvia Wrenn⁹, Adam M. Zanation⁹;
University of Pittsburgh Simion I. Chiosea²², Brenda Diergaarde²², Ann Marie Egloff²², Robert L. Ferris²², Marjorie Romkes²², Raja Seethala²²;
Vanderbilt University Brandee T. Brown²⁴, Yan Guo²⁴, Michelle Pham²⁴;
Yale University Wendell G. Yarbrough³⁰

¹The Eli and Edythe L. Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts 02142, USA. ²Massachusetts General Hospital, 55 Fruit Street, Boston, Massachusetts 02114, USA. ³Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia V5Z 4S6, Canada. ⁴The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. ⁵Harvard Medical School, Boston, Massachusetts 02115, USA. ⁶Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. ⁷Boston Children's Hospital, Boston, Massachusetts 02115, USA. ⁸University of Kentucky, Lexington, Kentucky 40506, USA. ⁹University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. ¹⁰Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, Massachusetts 02215, USA. ¹¹Harvard University, Boston, Massachusetts 02115, USA. ¹²Baylor College of Medicine, Houston, Texas 77030, USA. ¹³Memorial Sloan-Kettering Cancer Center, 1275 York Ave, New York, New York 10065, USA. ¹⁴University of California Santa Cruz, Santa Cruz, California 95064, USA. ¹⁵The Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins University, Baltimore, Maryland 21287, USA. ¹⁶University of Southern California, 1450 Biggy Street, NRT G511, Los Angeles, California 90033, USA. ¹⁷Johns Hopkins Medical Institutions, Baltimore, Maryland 21205, USA. ¹⁸Ontario Institute for Cancer Research 101 College Street, Suite 800, Toronto, Ontario M5G 0A3, Canada. ¹⁹Ontario Cancer Institute, Toronto, Ontario M5G 2M9, Canada. ²⁰University of Michigan Comprehensive Cancer Center, Ann Arbor, Michigan 48109, USA. ²¹National Institute on Deafness and Other Communication Disorders, National Institutes of Health, Bethesda, Maryland 20892, USA. ²²University of Pittsburgh, Pittsburgh, Pennsylvania 15213, USA. ²³The University of Chicago, Chicago, Illinois 60637, USA. ²⁴Vanderbilt University Medical Center 7209 Medical Center East, South Tower Nashville, Tennessee 37232, USA. ²⁵Albert Einstein College of Medicine, Bronx, New York 10461, USA. ²⁶University of California San Diego, San Diego, California 92093, USA. ²⁷University of Toronto, Toronto, Ontario M5G 2M9, Canada. ²⁸University Health Network, Toronto, Ontario M5G 2C4, Canada. ²⁹Greater Poland Cancer Centre, Garbary 15, 61-866, Poznań, Poland. ³⁰Yale School of Medicine, 800 Howard Avenue, 4th Floor Yale Physicians Building, New Haven, Connecticut 06520-8041, USA. ³¹International Genomics Consortium, 445 North 5th Street, Phoenix, Arizona 85004, USA. ³²The Research Institute at Nationwide Children's Hospital, Columbus, Ohio 43205, USA. ³³SRA International, Fairfax, Virginia 22033, USA. ³⁴Frederick National Laboratory for Cancer Research, 1050 Boyles Street, Frederick, Maryland 21702, USA. ³⁵National Cancer Institute, 31 Center Drive 3A20, Bethesda, Maryland 20892, USA. ³⁶Scimentis, LLC, Atlanta, Georgia 30666, USA. ³⁷National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, USA. ³⁸Analytical Biological Services, Inc., 701 Cornell Drive, Wilmington, Delaware 19801, USA. ³⁹Fred Hutchinson Cancer Research Center, Mailstop M5-C800, 1100 Fairview Avenue North, Seattle, Washington 98109, USA. ⁴⁰Georgia Regents University, Augusta, Georgia 30912, USA. ⁴¹Indiana University Simon Cancer Center, Indianapolis, Indiana 46202, USA. ⁴²University of Miami, Sylvester Comprehensive Cancer Center, 1550 NW 10th Avenue, Miami, Florida 33136, USA.