The mutational landscape of adenoid cystic carcinoma

Allen S Ho1,2,15, Kasthuri Kannan1,15, David M Roy1,15, Luc G T Morris1,2, Ian Ganly1,2, Nora Katabi3, Deepa Ramaswami1, Logan A Walsh1, Stephanie Eng1, Jason T Huse1,3, Jianan Zhang1, Igor Dolgalev1, Kety Huberman1, Adriana Heguy1, Agnes Viale4, Marija Drobnjak5, Margaret A Leversha6, Christine E Rice6, Bhuvanesh Singh2, N Gopalakrishna Iyer7, C Rene Leemans8, Elisabeth Bloemena9, Robert L Ferris10, Raja R Seethala11, Benjamin E Gross12, Yupu Liang12, Rileen Sinha12, Luke Peng1, Benjamin J Raphael13, Sevin Turcan1, Yongxing Gong1, Nikolaus Schultz14, Seungwon Kim10, Simion Chiosea11, Jatin P Shah2, Chris Sander12, William Lee12 & Timothy A Chan1,14

Adenoid cystic carcinomas (ACCs) are among the most enigmatic of human malignancies. These aggressive salivary gland cancers frequently recur and metastasize despite definitive treatment, with no known effective chemotherapy regimen. Here we determined the ACC mutational landscape and report the exome or whole-genome sequences of 60 ACC tumor-normal pairs. These analyses identified a low exonic somatic mutation rate (0.31 non-silent events per megabase) and wide mutational diversity. Notably, we found mutations in genes encoding chromatin-state regulators, such as SMARCA2, CREBBP and KDM6A, suggesting that there is aberrant epigenetic regulation in ACC oncogenesis. Mutations in genes central to the DNA damage response and protein kinase A signaling also implicate these processes. We observed MYB-NFIB translocations and somatic mutations in MYB-associated genes, solidifying the role of these aberrations as critical events in ACC. Lastly, we identified recurrent mutations in the FGF-IGF-P13K pathway (30% of tumors) that might represent new avenues for therapy. Collectively, our observations establish a molecular foundation for understanding and exploring new treatments for ACC.

ACCs are malignancies that cause substantial morbidity and mortality. Typically arising from salivary glands, they are characterized by unpredictable growth, extensive perineural invasion and high rates of metastasis, ultimately resulting in low survival rates. Treatment remains limited to surgery and radiation, and no systemic agent of metastasis, ultimately resulting in low survival rates. Treatment needs to understand its molecular basis and guide the development of effective therapies.

We sequenced the entire exome (n = 55) or genome (n = 5) of 60 ACC samples with matched normal DNA (Table 1). This approach generated 2,221 Gb of mapped sequence, with 92.4% of the target sequence covered to at least 10× depth. Sequencing generated mean exome and genome coverages of 106× and 37×, respectively. To ensure the accuracy of our massively parallel sequencing data, we conducted

Table 1: ACC whole-exome and whole-genome sequencing statistics

| Metric | Whole exome | Whole genome |
|--------|-------------|--------------|
| Tumor-normal pairs sequenced | 55 | 5 |
| Total sequenced (Gb) | 971.5 | 1249.4 |
| Mean fold tumor target coverage (range) | 100× (23–167×) | 38× (34–45×) |
| Mean fold normal target coverage (range) | 112× (25–185×) | 35× (33–37×) |
| Mean somatic mutation rate per megabase (range) | 0.4 (0.06–0.8) | 0.3 (0.2–0.5) |
| Mean number of validated nonsynonymous mutations per sample (range) | 11.8 (1–36) | 11.8 (5–17) |
| Mean number of transcribed noncoding mutations per sample (UTRs) | 5.7 (1–13) | 2.8 (1–6) |
| Total number of structural rearrangements | NA | 17 |
| Total number of genic rearrangements preserving the frame of translation | NA | 1 |

1Human Oncology and Pathogenesis Program, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. 2Head and Neck Service, Department of Surgery, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. 3Department of Pathology, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. 4Genomics Core, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. 5Pathology Core, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. 6Molecular Cytogenetics Core, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. 7Department of Surgical Oncology, National Cancer Center Singapore, Singapore. 8Department of Otolaryngology–Head and Neck Surgery, YU University Medical Center, Amsterdam, The Netherlands. 9Department of Oral and Maxillofacial Surgery and Oral Pathology, YU University Medical Center, Academic Centre for Dentistry, Amsterdam, The Netherlands. 10Department of Otolaryngology–Head and Neck Surgery, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA. 11Department of Pathology, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA. 12Computational Biology Center, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. 13Department of Computer Science, Brown University, Providence, Rhode Island, USA. 14Department of Radiation Oncology, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. 15These authors contributed equally to this work. Correspondence should be addressed to T.A.C. (chant@mskcc.org).

Received 26 February; accepted 25 April; published online 19 May 2013; doi:10.1038/ng.2643
Figure 1 Mutational landscape of ACC. (a) Number of validated nonsynonymous somatic mutations per sample across 60 ACC cases. (b) Representative list of recurrent nonsynonymous somatic mutations. Multiple mutations within the same gene in a given sample were only counted once. (c) Pathways affected by driver mutations identified by CHASM. Log values of Bonferroni false discovery rate (FDR)-corrected P values for pathway enrichment are shown. (d) Location of ACC alterations in key chromatin-remodeling genes. HSA, helicase-SANT–associated domain; Zn-TAZ, TAZ zinc-finger domain; DUF, domain of unknown function; Kat11, histone acetylation protein domain; ZZ, ZZ-type zinc-finger domain; CREB, cAMP response element–binding domain; SNF2_N, SNF2 family N-terminal domain; helicase C, helicase-conserved C-terminal domain; TPR, tetratricopeptide-repeat domain; JmjC, JmjC hydroxylase domain; FtsJ, FtsJ-like methyltransferase domain. (e) Location of ACC alterations in key established cancer genes. PI3K, PI3K p85β, PI3K p85-binding domain; PI3K_RBD, PI3K Ras-binding domain; PI3K_A, PI3K accessory domain; EGF-Ca, calcium-binding EGF domain; hEGF, human growth factor–like EGF domain; Ank repeat, ankyrin repeat; DspC, dual-specificity phosphatase, catalytic domain; p53 TAD, p53 transactivation motif; p53 tetramer, p53 tetramerization motif; LMR, LM-12/Notch repeat domain; PTEN_C2, phosphatase and tensin homolog of C2; NOD, nucleotide-binding oligomerization domain.

In extensive validation of nearly every candidate somatic mutation identified (2,751 variant calls; Supplementary Fig. 1) using targeted resequencing (Supplementary Fig. 2 and Supplementary Table 1), we identified a mean of 22 somatic mutations per sample, corresponding to approximately 0.31 non-silent mutations per megabase (Fig. 1). This rate is quite low compared to most adult solid tumors, such as head and neck squamous cell carcinoma\(^6,7\) and colon cancer\(^8\), but similar to hematological malignancies and neuroblastoma\(^9,10\). The transition/transversion ratio (Ti/Tv) was 1.1, similar to that observed across tumors (Figs. 1b and 2). We employed CHASM, a widely used approach for distinguishing driver and passenger mutations\(^11\), to identify multiple potential driver mutations, including ones in PI3KCA, TP53, PTEN, SMARCA2, KDM6A and CREBBP (Supplementary Tables 4 and 5). Analysis of these genes with driver mutations showed marked enrichment in pathways involved in chromatin remodeling, DNA damage response, MYB-MYC, protein

mutation frequency correlated with solid histology (Wilcoxon rank-sum test \(P = 4.0 \times 10^{-5}\)), and MYB translocations occurred in 57% of samples (34/60).

We validated 710 distinct nonsynonymous mutations across 643 genes (1–36 mutations per tumor; Fig. 1a, Supplementary Fig. 4 and Supplementary Table 3), with substantial mutational heterogeneity observed across tumors (Figs. 1b and 2). We employed CHASM, a widely used approach for distinguishing driver and passenger mutations\(^11\), to identify multiple potential driver mutations, including ones in PI3KCA, TP53, PTEN, SMARCA2, KDM6A and CREBBP (Supplementary Tables 4 and 5). Analysis of these genes with driver mutations showed marked enrichment in pathways involved in chromatin remodeling, DNA damage response, MYB-MYC, protein...
kinase A (PKA) and phosphatidylinositol 3-kinase (PI3K) signaling (Fig. 1c–e). For example, despite low overall mutation frequency, 35% of ACC tumors were mutated in chromatin regulators. These data suggest that the ACC mutational landscape is characterized by combinations of MYB pathway alterations and mutations in specific biological processes (Fig. 2). Notably, a small subset (n = 8) was observed with no CHASM-designated driver mutations. It is possible that some mutations in these tumors are drivers not called by CHASM or that other non-exonic alterations are important.

We used exome and genome sequencing data to characterize the ACC copy number landscape. We analyzed somatic copy number variations (CNVs) using ExomeCNV15 and found high concordance with a subset (n = 12) that additionally underwent array-based analysis. GISTIC2.0 (ref. 16) identified recurrent high-level losses at 6q24, 12q13 loss, PIK3CA, and PTEN (Fig. 3a and Supplementary Table 6)5. Samples with 14q loss were more likely to have solid histology (Fisher’s exact test \( P = 2.0 \times 10^{-2} \)), whereas samples with 6q24 loss were enriched for advanced tumor stage \( (P = 4.0 \times 10^{-2}) \). Expression array analysis on 23 ACC tumors found no distinct subgroups (Supplementary Fig. 5). Genes harboring driver mutations were confirmed to be expressed in ACC tumors (Supplementary Table 7).

Whole-genome paired-end sequencing identified numerous structural variations (Fig. 3b and Supplementary Table 8), and the existence of 17 structural variations across 5 samples was confirmed using PCR (Supplementary Fig. 6 and Supplementary Table 9). When also taking into account FISH data, MYB translocations were the only recurrent structural variations detected. However, we cannot exclude the possibility that there are less common recurrent translocations17. Notably, one sample harbored a tandem duplication within FGFR2. Similar alterations comprising dimerized, constitutively active variants have been described in hematological malignancies18. We did not detect a fusion transcript from this tumor using RT-PCR, but this does not rule out an alternative transcript configuration. Larger numbers of tumors will need to be analyzed to further characterize the incidence of this alteration and its impact. The remaining variants were not predicted to be in frame. Collectively, few CNVs or structural variations were found across the cohort, demonstrating the relatively ‘quiet’ nature of the ACC genome. Thus, our data indicate that the major structural variations in ACC are MYB translocations and recurrent deletions on 6q24, 12q13 and 14q.

We observed that the diversity of identified ACC alterations converged on several well-defined pathways (Fig. 4). A prominent feature of the ACC mutational landscape is the presence of multiple mutations targeting chromatin-remodeling genes (35% incidence; Figs. 1d and 4a). Such alterations are increasingly recognized as having key roles in oncogenesis19,20 and have been reported in various other tumors10,21 but not in ACC. Among ACC alterations, somatic mutations were substantially enriched in chromatin-state modifiers \( (Q = 4.5 \times 10^{-3}) \). We identified multiple aberrations in the SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin (SMARC) family,
The majority of CREBBP mutations clustered in the region encoding the critical KAT11 histone acetyltransferase domain (p.Arg1446Cys, p.Ile1453Asn and p.Trp1472Ser). We also observed numerous missense mutations in KDM6A (7%), also known as UTX, which encodes a histone demethylase. The functional impact of these KDM6A mutations was evaluated via a well-characterized assay for trimethylation at lysine 27 of histone H3 (H3K27me3). We observed abrogation of demethylase activity in cells overexpressing mutant KDM6A but not in those overexpressing wild-type KDM6A (Fig. 5a,b). Moreover, whereas wild-type KDM6A suppressed growth, mutants either lost the ability to suppress growth or, in some cases, augmented it (dominant phenotype).

Other related mutations included ones in genes encoding proteins with histone acetyltransferase or deacetylase activity (EP300, ARID4B, ARID5B and BRD1) and histone methyltransferase or demethylase activity (FTSJ1 and MLL3), as well as in histones themselves (HIST1H2AL and HIST1H1E). EP300 binds CREBBP as a transcriptional coactivator that regulates cell proliferation and differentiation and harbors a splice-site mutation affecting its KAT11 histone acetylation domain (p.Gly1429?). HIST1H2AL encodes a core component of the nucleosome, and HIST1H1E encodes a protein that is known to link histones for compaction into higher-order structures. Notably, histone gene mutations have been implicated in cancers such as pediatric gliomas, targeting epigenetic regulation in other disorders, similar treatments for patients with ACC with chromatin derangements might hold promise.

We observed a second group of mutations in genes involved in the DNA damage response (q = 5.6 × 10−3; Fig. 4b). In TP53 (5% incidence), one missense (p.Pro151Ser) and two nonsense (p.Arg213* and p.Arg342*) variants were identified that affect the central binding
and tetramer motifs (Fig. 1e). Similarly, UHRF1, encoding a ubiquitin ligase involved in the TP53-dependent DNA damage checkpoint, was altered in 8% of cases (two mutations and three homozygous deletions). TXNIP, encoding a protein frequently repressed in cancers, was noted to have a frameshift insertion affecting its arrestin domain (p.Leu129fs). We also identified tumors that harbored two missense mutations within a particular gene (ATM, BRCA1 or DCLRE1A, a DNA cross-link repair gene). We moreover found markedly decreased expression of key TP53 transcriptional targets specifically in tumors with TP53 pathway alterations (binomial test $P = 1.0 \times 10^{-4}$; Fig. 5c).

Previous work has implicated MYB as an ACC fusion oncogene with NFIB via a t(6;9) translocation (Fig. 4c). Indeed, we verified that 57% of our cohort had the MYB-NFIB translocation by FISH analysis. We extended this initial finding and observed further MYB pathway dysregulation, with alterations found in an additional 8% of cases. Specifically, splice-site and coding mutations involving exon 10 of MYB were seen, which potentially disrupt its leucine-rich negative regulatory domain and trigger constitutive activation. Five additional MYB homozgyous deletions were identified in CNV analysis (4 of 5 tumors were positive for the MYB-NFIB translocation), although these deletions may be secondary alterations occurring around the translocation. Other mutations identified in the MYB pathway included ones present in MYCBP2, MGA and MCM4. In NFIB, we observed two truncating mutations affecting the CTF/NFI transcription modulation domains (p.Tyr249* and p.Pro390fs), in addition to four homozygous deletions (all tumors were positive for the MYB-NFIB translocation). Our data highlight MYB as an active oncogenic partner in fusion transcripts in ACC but also suggest a separate role for NFIB, given the presence of mutations specific to this gene.

Multiple genes involved in the PKA pathway were found to be mutated (27% incidence; $q = 4.2 \times 10^{-5}$). We identified recurrent mutations in RYR3 (7%) and RYR2 (2%), both of which encode tetrameric intracellular calcium channels implicated in breast cancer progression. Mutations were also observed in PKA pathway–associated genes that are known or potential tumor suppressors, including in tyrosine phosphatases (PTPRG, PTPRH, PTPRJ and PTPRK). Nonsense (PTPRG and PTPRH) and frameshift (PTPRK) mutations abrogated the phosphatase domains of the encoded proteins (p.Glu736*, p.Trp602* and p.Leu457fs, respectively), and an additional four samples had homozygous deletions of PTPRK. PKA signaling components are known to facilitate CREBBP recruitment via calcium influx and CREB phosphorylation, which regulates multiple critical developmental processes involved in tumorigenesis.

Mutations in genes important in the fibroblast growth factor (FGF)–insulin-like growth factor (IGF)–PI3K pathway were also identified in 30% of our cohort ($q = 2.4 \times 10^{-2}$; Fig. 4d). Three tumors harbored separate missense mutations in PIK3CA (Fig. 1e). Another tumor had two mutations in PTEN affecting the catalytic domain (p.Arg130fs and p.Lys144Gln). All five lesions occurred in COSMIC v61 (Catalogue of Somatic Mutations in Cancer) hotspots. FOXO3 (ref. 46) was also altered in 7% of tumors. We further identified mutations in receptor kinase and PI3K–associated genes, including in FGFR6, FGFR4, IGFBP2 and ILK17RD. FGF and IGF proteins are potent, oncogenic PI3K activators, whereas ILK17RD seems to be a PI3K inhibitor through its interaction with FGFR3. The samples with PI3K–associated mutations all had solid histology, considered the most aggressive ACC variant (Fisher’s exact test $P = 1.6 \times 10^{-3}$). We further evaluated samples with altered PI3K using immunohistochemistry to detect phosphorylated AKT and PRAS40. Functional activation of the PI3K pathway was observed in all ACC tumors harboring PIK3CA or PTEN mutations but not in tumors with wild-type forms of these genes (Fig. 5d and Supplementary Fig. 7). This finding is supported by the presence of significantly enriched PI3K signatures in ACC tumors with mutant PI3K identified via GSEA (Gene Set
Enrichment Analysis) \( P < 1.0 \times 10^{-3}; \) Supplementary Fig. 8a). Taken together, these findings delineate a previously undescribed subset of patients with ACC who we hypothesize might benefit from agents targeting this pathway.50

Despite not meeting significance by CHASM analysis, alterations were observed in the Notch signaling pathway in 13% of samples (Fig. 4e). In addition to recurrent NOTCH1 (5% incidence, including three missense and one nonsense) and FOXP2 (3%) mutations, predicted functional mutations were found in DTX4, FBXW7 and CNTN6. FBXW7 (p.Arg465His alteration identified) is a tumor suppressor that targets c-myc and NOTCH1 for degradation.31 NOTCH1 has a dual role in cancer, depending on cellular context, as either an oncogene or a tumor suppressor.47,52-55. Analysis of ACC samples with altered NOTCH1 using GSEA56 showed a trend toward enrichment of Notch signaling (Supplementary Fig. 8b). These alterations require further characterization but may have therapeutic implications.57

Notably, we observed mutations in various genes involved in cell metabolism processes. Of particular interest are recurrent alterations in HSPG2 (7%), which encodes a basement membrane–based proteoglycan (perlecan) known to modulate tumorigenic growth factors, including FGF. FGF signaling promotes mitogenesis and angiogenesis,58 whereas deletion of perlecan heparan sulfate domains has been shown to impair tumorigenesis.59-61. ACC produces high perlecan levels within its pseudocystic cribriform structures,62,63 and maximum perlecan synthesis occurs during the tumor’s proliferative phase.64

Additionally, we identified a hotspot IDH1 mutation affecting the encoded catalytic domain (p.Arg132His),33,65-67.

ACC is unique in its propensity for perineural invasion and distant metastasis. We noted numerous mutations in genes encoding cell adhesion molecules (29% incidence). We found mutations affecting the active domains of factors responsible for neuronal axon guidance cues, including in NTNG1, SEMA3G and SEMA5A. Netrins and semaphorins are implicated in neuronal invasion68 and metastasis69. Additionally, mutations in protocadherin genes were observed, including truncating mutations in FAT3. FAT4 has been hypothesized to be a tumor suppressor in gastric adenocarcinoma.70 We examined FAT4 through RNA interference (RNAi)-mediated knockdown in human salivary cells (HSG and human salivary adenocarcinoma cells (HSY). Experiments were performed in triplicate. Two RNAi constructs were used for FAT4 knockdown (KD). \( * P < 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001, \) NS, not significant (ANOVA). Error bars, \( \pm 1 \) s.e.m.

In summary, we report a large-scale ACC analysis and highlight new ACC driver mutations within corresponding pathways. Convergent alterations include pathway derangements involving chromatin remodeling, DNA damage response, MYB-MYC and tyrosine kinase signaling (Fig. 4). Our discovery of genomic alterations in targetable pathways suggests potential avenues for novel treatments to address a typically chemoresistant malignancy. Verified ACC cell lines71,72 are needed to further substantiate the clinical usefulness of the mutations identified here. In total, our data provide insights into the genetic framework underlying ACC oncogenesis and establish a foundation for identifying new therapeutic strategies.
METHODS

Methods and any associated references are available in the online version of the paper.

Accession code. Sequencing and genotype data have been deposited in the NCBI database of Genotypes and Phenotypes (dbGaP) under accession phs000612.v1.p1.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We thank A. Kaye and S. Thomas for technical assistance. We are indebted to R.J. Wong (Memorial Sloan-Kettering Cancer Center), P.K. Brindle (St. Jude Children’s Research Hospital) and O.J. Baker (State University of New York at Buffalo) for generously providing cell lines and constructs. We are additionally grateful to J.O. Boyle, S.G. Patel, A.R. Shaia, R.J. Wong, A.L. Ho, M.C. Mariano and L. Lopez for facilitating tissue procurement and contributing valuable input. This work was supported by the US National Institutes of Health (NIH) (ROI CA154767 (T.A.C.), R21 DE023229 (T.A.C.), 5T32 CA009685 (A.S.H.)), the Geoffrey Beene Foundation (T.A.C.), the Starr Cancer Consortium (T.A.C.) and the Louis Gerstner Fellowship (T.A.C.), the Starr Cancer Consortium Medical Research Fellows Program. The cBio Portal was supported in part by the NIH National Resource for Network Biology (GM103504). T.A.C. would like to thank Jeff and Marnie Kaufman and the Adenoid Cystic Carcinoma Research Foundation for the support of our work and inspiration to see it through for the benefit of patients.

AUTHOR CONTRIBUTIONS

T.A.C., A.S.H., K.K. and W.L. designed the experiments. D.M.R., A.S.H., S.E., A.V., A.H., K.H., D.R., L.A.W., J.T.H., J.Z., N.G.I., Y.G., M.A.L. and C.E.R. performed the experiments. A.S.H., W.L., T.A.C., K.K., D.M.R., L.G.T.M. and S.T. analyzed the data. N.K. and R.R.S. provided histopathological confirmation of sample purity. A.S.H., T.A.C. and W.L. wrote the manuscript. A.S.H., I.G., L.G.T.M., M.D., J.P.S., B.S., N.G.I., C.R.L., E.B., R.L.F., R.R.S., S.C. and S.K. contributed new reagents and provided tissues for analysis. W.L., K.K., A.S.H., T.A.C., L.P., B.J.R., I.D., Y.L., R.S., N.S., B.E.G. and C.S. conducted the bioinformatics analysis.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Adelstein, D.J., Koyfman, S.A., El-Naggar, A.K. & Hanna, E.Y. Biology and management of salivary gland cancers. Semin. Radiat. Oncol. 22, 245–253 (2012).
2. Agulnik, M., et al. Phase II study of lapatinib in recurrent or metastatic epidermal growth factor receptor and/or erbB2 expressing adenoid cystic carcinoma and non adenoid cystic carcinoma malignant tumors of the salivary glands. J. Clin. Oncol. 25, 3978–3984 (2007).
3. Holst, V.A., Marshall, C.E., Moskaluk, C.A. & Frierson, H.F. Jr. KIT protein expression and analysis of c-kit gene mutation in adenoid cystic carcinoma. Mod. Pathol. 12, 956–960 (1999).
4. Laurie, S.A., Ho, A.L., Fury, M.G., Sherman, E. & Pfister, D.G. Systemic therapy in the management of metastatic or locally recurrent adenoid cystic carcinoma of the salivary glands: a systematic review. Lancet Oncol. 12, 815–824 (2011).
5. Persson, M. et al. Clinically significant copy number alterations and complex rearrangements of MYB and NFIB in head and neck adenoid cystic carcinoma. Genes Chromosom. Cancer 51, 800–817 (2012).
6. Agrawal, N. et al. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. Science 333, 1154–1157 (2011).
7. Stratyski, N. et al. The mutational landscape of head and neck squamous cell carcinoma. Science 333, 1157–1160 (2011).
8. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. Nature 487, 330–337 (2012).

9. Zhang, J. et al. The genetic basis of early T-cell precursor acute lymphoblastic leukemia. Nature 481, 157–163 (2012).
10. Pugh, T.J. et al. Medulloblastoma extender sequence uncovering subtype-specific somatic mutations. Nature 488, 106–110 (2012).
11. Molenar, J.J. et al. Sequencing of neuroblastoma identifies chromothripsis and defects in neurotigenes genes. Nature 483, 589–593 (2012).
12. Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. Nature 489, 519–525 (2012).
13. DePristo, M.A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. 43, 491–498 (2011).
14. Carter, H. et al. Cancer-specific high-throughput annotation of somatic mutations: computational prediction of driver missense mutations. Cancer Res. 69, 6660–6667 (2009).
15. Sathirapongsasuti, J.F. et al. Exome sequencing-based copy-number variation and loss of heterozygosity detection. ExomeCNV. Bioinformatics 27, 2648–2654 (2011).
16. Merzel, C.H. et al. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol. 12, R4 (2011).
17. Vek穉, H. et al. DNA copy number gains at loci of growth factors and their receptors in salivary gland adenoid cystic carcinoma. Clin. Cancer Res. 13, 3133–3139 (2007).
18. Wolffe, A.P. Chromatin remodeling: why it is important in cancer. Oncogene 20, 2993–2998 (2001).
19. Chi, P., Aliis, C.D. & Wang, G.G. Covalent histone modifications—miswritten, misinterpreted and mis-erased in human cancers. Nat. Rev. Cancer 10, 457–469 (2010).
20. Dimpf, G.L. et al. Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. Nature 463, 360–363 (2010).
21. Van Houw, J.D. et al. Heterozygous missense mutations in SMARC2 cause Nicolaides-Baraitser syndrome. Nat. Genet. 44, 445–449 (2012).
22. Tsurusaki, Y. et al. Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. Nat. Genet. 44, 376–378 (2012).
23. Pasqualucci, L. et al. Inactivating mutations of acetyltransferase genes in B-cell lymphoma. Nature 471, 189–195 (2011).
24. Flaus, A., Martin, D.M., Barton, G.J. & Owen-Hughes, T. Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. Nucleic Acids Res. 34, 2887–2905 (2006).
25. Hu, Y. et al. RECL05l/RecQ5 helicase regulates homologous recombination and suppresses tumor formation via disruption of Rad51 presynaptic filaments. Genes Dev. 21, 3073–3084 (2007).
26. Masse, J.E. et al. The S. cerevisiae architectural HMGB protein NHP6A complexed with DNA DNA and protein configurationnal changes under binding. J. Mol. Biol. 323, 263–284 (2002).
27. Mullighan, C.G. et al. CREBBP mutations in relapsed acute lymphoblastic leukemia. Nature 471, 235–239 (2011).
28. Reva, B. Predicting the functional impact of mutations: application to cancer genomics. Nucleic Acids Res. 39, e118 (2011).
29. Spangle, J.M. & Munger, K. The human papillomavirus type 16 E6 oncoprotein activates mTORC1 signaling and increases protein synthesis. J. Virol. 84, 9398–9407 (2010).
30. Iyer, N.G. et al. p300 regulates p53-dependent apoptosis after DNA damage in colorectal cancer cells by modulation of PUMA/p21 levels. Proc. Natl. Acad. Sci. USA 101, 7386–7391 (2004).
31. Parsons, D.W. et al. An integrated genomic analysis of human glioblastoma multiforme. Science 321, 1807–1812 (2008).
32. Jones, S. et al. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. Science 321, 1801–1806 (2008).
33. Oisen, E.A. et al. Phase IIb multicenter trial of vorinostat in patients with persistent, progressive, or treatment refractory cutaneous T-cell lymphoma. J. Clin. Oncol. 25, 3109–3115 (2007).
34. Fenaux, P. et al. Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes, a randomised, open-label, phase III study. Lancet Oncol. 10, 223–232 (2009).
35. Tian, A.L. et al. UHRF1 depletion causes a G2/M arrest, activation of DNA damage response and apoptosis. Biochem. J. 435, 175–185 (2011).
36. Cerami, E.G. et al. Pathway Commons, a web resource for biological pathway data. Nucleic Acids Res. 39, D685–D690 (2011).
37. Egloff, A.M. & Grandis, J.R. Molecular pathways: context-dependent approaches to notch targeting as cancer therapy. Clin. Cancer Res. 18, 5188–5195 (2012).
38. Persson, M. et al. Recurrent fusion of MYB and NFIB transcription factor genes in carcinomas of the breast and head and neck. Proc. Natl. Acad. Sci. USA 106, 18740–18744 (2009).
39. Zhang, L. et al. Functional SNP in the microRNA-367 binding site in the 3'UTR of the calcium channel ryodine receptor gene 3 (RYR3) affects breast cancer risk and calcification. Proc. Natl. Acad. Sci. USA 108, 13653–13658 (2011).
798

54. Wang, J. et al. Notch promotes radioresistance of glioma stem cells. Stem Cells 28, 17–28 (2010).

55. Allen, T.D., Rodriguez, E.M., Jones, K.D. & Bishop, J.M. Activated Notch1 induces lung adenomas in mice and cooperates with Myc in the generation of lung adenocarcinoma. Cancer Res. 65, 8530–8537 (2005).

56. Xie, Y., Liao, J.D., Zhou, Z., et al. The Fbxw7/hCdc4 tumor suppressor in human neurofibromatosis type 1. Oncogene 21, 406–414 (2002).

57. Sarmento, L.M. & Barata, J.T. Therapeutic potential of Notch inhibition in T-cell acute lymphoblastic leukemia: rationale, caveats and promises. Expert Rev. Anticancer Ther. 11, 1403–1415 (2011).

58. Iozzo, R.V. & Murdoch, A.D. Proteoglycans of the extracellular environment: clues from the gene and protein side offer novel perspectives in molecular diversity and function. J. Biol. Chem. 278, 33232–33238 (2003).

59. Tran, P.K. et al. Increased intimal hyperplasia and smooth muscle cell proliferation in transgenic mice with heparan sulfate-deficient perlecan. Circ. Res. 94, 550–558 (2004).

60. Zhou, Z. et al. Impaired angiogenesis, delayed wound healing and retarded tumor growth in perlecan heparan sulfate-deficient mice. Cancer Res. 64, 4699–4702 (2004).

61. Savoré, C. et al. Perlecan knockdown in metastatic prostate cancer cells reduces heparin-binding growth factor responses in vitro and tumor growth in vivo. Clin. Exp. Metastasis 22, 377–390 (2005).

62. Munakata, R., Irie, T., Cheng, J., Nakajima, T. & Saku, T. Pseudocyst formation by adenoid cystic carcinoma cells in collagen gel culture and in SCID mice. J. Oral Pathol. Med. 25, 441–448 (1996).

63. Cheng, J., Saku, T., Okabe, H. & Furutmay, H. Basement membranes in adenoid cystic carcinoma. An immunohistochemical study. Cancer 69, 2631–2640 (1992).

64. Kimura, S. et al. Perlecan (heparan sulfate proteoglycan) gene expression reflected in the characteristic histological architecture of salivary adenoid cystic carcinoma. Virchows Arch. 437, 122–128 (2000).

65. Yan, H. et al. IDH1 and IDH2 mutations in gliomas. N. Engl. J. Med. 360, 765–773 (2009).

66. Figueroa, M.E. et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. Cancer Cell 18, 553–567 (2010).

67. Dang, L. et al. Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. Nature 462, 739–744 (2009).

68. Binnardi, N.O. et al. Plexin-B1 and semaphorin 4D cooperate to promote perineural invasion in a RhoA/ROK-dependent manner. Am. J. Pathol. 180, 1232–1242 (2012).

69. Rodrigues, S., De Wever, O., Bruyneel, E., Rooney, R.J. & Gespach, C. Opposing roles of netrin-1 and the dependence receptor DCC in cancer cell invasion, tumor growth and metastasis. Oncogene 26, 5615–5625 (2007).

70. Zang, Z.J. et al. Exome sequencing of gastric adenocarcinoma identifies recurrent somatic mutations in cell adhesion and chromatin remodeling genes. Nat. Genet. 44, 570–574 (2012).

71. Pruchanone, J., Olthoff, K., Yuan, J., Stoeckel, D. & Tetsu, O. Genetic profiling reveals cross-contamination and misidentification of 6 adenoid cystic carcinoma cell lines. ACC2, ACC3, ACCM, ACCNS, ACCS and CAC2. PLoS ONE 4, e6040 (2009).

72. Anonymous. RETRACTED: Nef from SIVmac239 decreases proliferation and migration of adenoid-cystic carcinoma cells and inhibits angiogenesis. Oral Oncol. 47 (2011) 847–854] 48, 95 (2012).

73. Krzywinski, M. et al. Circos: an information aesthetic for comparative genomics. Genome Res. 19, 1639–1645 (2009).
ONLINE METHODS

Tumor samples. Primary tumor samples and matched normal specimens (peripheral blood or tissue) were obtained with written informed consent per approved institutional review board protocols (Supplementary Note). Specimens were snap frozen in liquid nitrogen immediately after surgical resection and stored at −80 °C. Sections stained with hematoxylin and eosin were prepared, and diagnosis was confirmed by a dedicated head and neck pathologist. Microdissection was performed to ensure >70% tumor purity. DNA was extracted using the DNeasy Blood & Tissue kit (Qiagen) and quantified with the Quant-iT PicoGreen dsDNA assay (Invitrogen). RNA was extracted using the RNeasy Mini kit (Qiagen) and quantified using a Nanodrop spectrophotometer (Thermo-Fisher). Confirmation of tumor with matched normal DNA samples was performed by mass spectrometric fingerprint genotyping (Sequenom) as previously described44.

Whole-exome capture and sequencing. Whole-exome capture libraries were constructed via Agilent SureSelect 51Mb target enrichment. Approximately 2–3 µg of genomic DNA from each tumor and normal specimen was sheared and ligated to barcoded sequencing adapters. Enriched exome libraries were sequenced on a HiSeq 2000 platform (Illumina) to generate paired-end reads (2 × 76 bp) (Supplementary Table 1).

Analysis pipeline and mutation annotation. Reads were aligned to the hg19 genome build (GRCh37) using the Burrows-Wheeler Aligner (BWA)75 (Supplementary Fig. 1). Further indel realignment, base-quality score recalibration and duplicate-read removal were performed using the Genome Analysis Toolkit (GATK) v2.2 (ref. 76). Somatic Sniper v1.0.0 (ref. 77) and GATK Somatic Indel Detector76 were used to generate single-nucleotide variation (SNV) and indel calls, respectively, using standard, default parameters. MuTect78 was used to confirm SNV calls with the conditions described previously79. Baseline filters (depth of 3x coverage in both tumor and normal specimens, >97% normal allelic fraction, >10% tumor allelic fraction) were chosen. We excluded germline variants found in the 1000 Genomes Project80, ESP5400 (National Heart, Lung, and Blood Institute (NHLBI) GO Exome Sequencing Project) and dbsNP132 (ref. 81). Further filtering was based on mapping quality score, average read length and strand bias, as previously described77. Resulting mutations were annotated based on RefSeq (Release 55) using Annovar82. All candidates were manually inspected via Integrative Genomics Viewer (IGV) v2.1 (ref. 83) (Supplementary Fig. 4). The mean numbers of somatic mutations per sample (16 nonsynonymous, 6 synonymous) comprised exonic SNVs and indels. The incidence of non-silent mutations per megabase (0.31) refers to nonsynonymous mutations found within the exon-capture target regions (51 Mb).

Validation of mutations. Nearly all SNVs or indels were resequenced via the SOLID (Life Technologies) or MiSeq (Illumina) platforms to validate putative mutations. Regions encompassing putative candidates were amplified with PCR using Kapa Fast Hot-Start Taq Polymerase (KapaBiosystems) with whole-genome amplified genomic DNA generated using the Repli-G kit (Qiagen). For preparation of samples for the SOLID platform, we sheared 3 µg of DNA and selected 200- to 220-bp fragments by agarose gel electrophoresis. DNA was purified and ligated to P1 and P2 adaptors according to the manufacturer's instructions. After PCR amplification, DNA was quantified and size confirmed by Bioanalyzer DNA 1000 assay (Agilent Technologies). Approximately 300–500 ng of DNA per sample was used for 24-h hybridization with custom SureSelect baits (Agilent Technologies) according to the manufacturer's instructions. Captured DNA was separated from free DNA using the SureSelect protocol. After emulsion PCR and bead preparation, each sample was loaded into one octet of a SOLID plate for paired-end sequencing, which was followed by BWA75 alignment, GATK v2.2 (ref. 76) variant detection and manual inspection on IGV 2.1.

For MiSeq sequencing, we purified amplified templates using AMPure (Beckman Coulter). Purified PCR products were sequenced bidirectionally with M13 forward and reverse primers and BigDye Terminator kit v3.1 (Applied Biosystems). Dye terminators were removed using the CleanSEQ kit (Beckman Coulter), and the resulting amplicons were barcoded and prepared for sequencing using the Nextera DNA Sample Prep kit (Illumina). Libraries were sequenced using the paired-end 150-cycle protocol followed by BWA75 alignment, GATK v2.2 (ref. 76) variant detection and manual inspection on IGV 2.1 (Supplementary Table 3).

Whole-genome capture and sequencing. Library construction and whole-genome sequencing were performed as previously described84. Briefly, 5 µg each of native DNA from tumor and matched normal samples was sheared separately and ligated to standard paired-end adaptors, and products were run on a gel and excised. Bands were purified using MinElute columns (Qiagen) and enriched by PCR amplification (ten cycles). Paired-end sequencing (2 × 101 bp) was performed on the HiSeq 2000 platform (Supplementary Table 8). The HumanOmni 2.5-8v1 BeadChip kit (Illumina) was used for genotyping.

Structural variation and copy number analysis. Structural variations were identified using CREST (CLipping REVeals tStructure)85, which employs soft-clipped reads to directly map the breakpoints of structural variations. Samples were analyzed using the paired analysis module. Primers were designed for the 1,000-bp regions surrounding the predicted structural variation breakpoints using Primer3 (ref. 86) (Supplementary Table 9). We used a standard touchdown PCR protocol, and products were run on a gel, excised, purified using the NucleoSpin Gel and PCR Clean-up kit (Clontech) and sequenced. Candidates were considered to be germline if found in both the tumor and normal samples (Supplementary Fig. 6).

CNVs were calculated with ExomeCNV15, a statistical method to detect somatic CNVs using depth-of-coverage information from mapped short sequence reads. A subset of the 60-sample cohort (n = 7) was also run on Affymetrix Genome-Wide SNP 6.0 arrays for confirmation. Analysis was performed with the aroma.affymetrix package87, and segmentation was performed using the CBS algorithm in the DNAcopy R package88. For quality control, the subset of samples (n = 5) genotyped on HumanOmni 2.5-8v1 BeadChips (as part of whole-genome sequencing) was also analyzed for CNVs using tQCN89 for normalization, with OncoSNP v1.2 then used to characterize CNVs and LOH events90. Copy number clustering was performed on segmented copy number data. A unified breakpoint profile (region × sample matrix) was derived by combining breakpoints across all samples and determining the minimal common regions of change. Unified breakpoint profiles were computed using the Bioconductor package CNTools91. Hierarchical clustering was performed using the R function hclust, with Manhattan distance and Ward's agglomeration method. Gene-level copy number changes across samples were identified using GISTIC2.0 (Genomic Identification of Significant Targets in Cancer)92 (Supplementary Table 6). Homozygous deletions were defined as −2 or lower, whereas high-level amplifications were defined as 2 or higher.

Gene expression profiling. Approximately 300 ng of RNA was extracted from fresh-frozen tissue, and RNA integrity was assessed using the Eukaryote Total RNA Nano Assay on the Agilent 2100 Bioanalyzer. Expression profiling of viable RNA was performed using the Human HT-12 Expression BeadChip array (Illumina). Analysis was performed with Partek Genomics Suite v6.5.

FISH. Tissue microarrays were constructed for tumors with available formalin-fixed, paraffin-embedded tissue. Unstained frozen slides were generated for all other tumors. FISH was performed using a three-color probe mix consisting of BACs for 5′ MYB (RP11-614H6, RP11-104D9; green), 3′ MYB (RP11-3323N12, RP11-1060C14; orange) and 3′ NF1B (RP11-413D24, RP11-589C16; red) (BACPAC Resources and Wellcome Trust Sanger Institute93) (Supplementary Fig. 3). DNA was prepared by standard alkaline lysis and labeled by nick translation with fluorochrome-conjugated dUTPs (Enzo Life Sciences). Paraffin-embedded and frozen slide sections were processed as previously described93.

Driver-passenger mutation and pathway analysis. Driver-passenger mutation analysis was performed with CHASM v1.0.5 (Cancer-specific High-throughput Annotation of Somatic Mutations)14, a random forest–based classification algorithm that predicts whether tumor-derived somatic missense mutations are important contributors to cancer cell fitness. The default set of 52 features was used with a Benjamini-Hochberg FDR cutoff of 0.35 (Supplementary Table 4).
Pathway analyses for enrichment were performed using IPA (Winter Release 2012, Ingenuity Systems), DAVID v6.7 (Database for Annotation, Visualization and Integrated Discovery)\(^4\) and hypergeometric distribution, with potential CHASM-identified drivers used as input. Enriched gene sets reflecting Gene Ontology\(^5\) biological processes, Kyoto Encyclopedia of Genes and Genomes (KEGG, Release 64.0) and Biocarta pathways were corrected for multiple-hypothesis testing using the Benferroni method. Aβerrant signaling pathways (Fig. 4) were constructed referencing IPA and Pathway Commons\(^8\).

**Expression data** were analyzed using GSEA v2.0.10 (ref. 56). Functional consequences were also evaluated using the Memorial Sloan-Kettering Cancer Center cBioPortal for Cancer Genomics\(^96,97\).

**Cell culture.** Cell lines were cultured with the following: 293T and COS7 cell lines, DMEM supplemented with 10% FBS; HSG, DMEM supplemented with 10% FBS; HSY, DMEM supplemented with 10% FBS; HTB-41, McCoy supplemented with 10% FBS; and HFF-1, DMEM supplemented with 15% FBS. All cell lines were obtained from ATCC except for HSG (O.J. Baker), HSY (R.J. Wong) and MEFs (P.K. Brindley).

**FA74 knockdown, transfection and viral transduction.** FA74 knockdown was performed using small interfering RNA (siRNA) or short hairpin RNA (shRNA) (Dharmacon). siRNAs were transfected into cells in antibiotic-free medium using Lipofectamine RNAiMAX (Invitrogen), medium was changed after 6 h and cells were harvested after 48 h. pGIPZ shRNAs were transfected into 293T cells, along with plasmids encoding packaging and envelope proteins. Viruses were harvested from cell culture supernatant at 48, 72 and 96 h after transfection. Cells were transduced with 8 µg/ml polybrene. Two independent RNAi sequences were used for knockdown. The fold change of relative knockdown was normalized to mRNA levels of respective controls.

**KDM6A mutagenesis and overexpression.** Plasmid encoding Flag-tagged KDM6A was obtained (AddGene, 17438), and site-directed mutagenesis was performed with QuikChange II XL (Stratagene). We transfected 293T and COS7 cells using FuGene HD (Promega).

**Immunofluorescence.** COS7 cells were plated on poly-κ-l-lysine coverslips (BD Biosciences) and transfected with plasmids encoding wild-type or mutant Flag-tagged KDM6A. After a 36-h incubation, cells were fixed in 4% paraformaldehyde in PBS, permeabilized in 0.5% Triton X-100 in PBS and blocked in 10% FBS. Primary antibodies included those to Flag (Sigma, F3165; 1:50 dilution) and H3K27me3 (Millipore, 07-449; 1:50 dilution). Secondary antibodies included Alexa Fluor 488–conjugated antibody to rabbit IgG and Alexa Fluor 568–conjugated antibody to rabbit IgG (Molecular Probes; 1:200 dilutions). Images were acquired using an SP5 confocal microscope (Leica).

**Growth curves.** Growth curve assays were performed in triplicate, and cell numbers were quantified using the Vi-Cell XR Cell Viability Analyzer (Beckman Coulter) or in real time in quadruplicate with the xCELLigence System (Roche). xCELLigence plates were seeded with 5,000 or 20,000 cells per well. Growth was reported as the cell index, a measure of impedance reflecting the number of viable, adherent cells.

**TP53 signaling pathway PCR array.** Approximately 200 ng of RNA extracted from fresh-frozen tissue was used to generate cDNA with SuperScript III (Invitrogen). Gene expression was quantified using the RT² Profiler PCR Array Human p53 Signaling Pathway (SA Bioscience). Reactions were performed using a Mastercycler ep Realplex 4s and Realplex software (Eppendorf). The \(2^{-\Delta\Delta C_T}\) method was used to calculate \(\Delta C_T\) values. In all tissue-specific comparisons, individual gene expression levels in tumor specimens were compared with those in normal salivary tissue. Comparisons were made between TP53 pathway wild-type and TP53 pathway altered samples. Expression levels of all 84 TP53 target genes were normalized against the average levels of the 2 most stably expressed reference genes (GAPDH and HPRT1). Gene expression was considered below the limit of detection if the \(C_T\) value was 35.0 or greater.

**PI3K immunohistochemistry.** We stained 5-μm formalin-fixed, paraffin-embedded tissue sections for phosphorylated Akt (Cell Signaling Technology, 4060; 1:100 dilution) on Benchmark Ultra (Ventana) using the OmniMap DAB anti-Rabbit Detection kit (Ventana, 760-149) and multimer detection system. We similarly stained 5-μm formalin-fixed, paraffin-embedded tissue sections for phosphorylated PRA540 (Cell Signaling Technology, 2997; 1:40 dilution) using biotinylated secondary antibody (Vector Labs, BA-1000; 1:300 dilution) and a streptavidin detection system (Ventana).