Biology of Human Tumors

Clinical Cancer Research

Mutational Landscape of Aggressive Cutaneous Squamous Cell Carcinoma

Curtis R. Pickering[1], Jane H. Zhou[2], J. Jack Lee[3], Jennifer A. Drummond[4], S. Andrew Peng[3], Rami E. Saade[1], Kenneth Y. Tsai[5,6], Jonathan L. Curry[5], Michael T. Tetzlaff[2], Stephen Y. Lai[1], Jun Yu[3], Donna M. Muzny[4], Harshavardhan Doddapaneni[4], Eve Shinbrot[4], Kyle R. Covington[4], Jianhua Zhang[7], Sahil Seth[7], Carlos Caulin[1], Gary L. Clayman[1], Adel K. El-Naggar[2], Richard A. Gibbs[4,8], Randal S. Weber[1], Jeffrey N. Myers[1], David A. Wheeler[4], and Mitchell J. Frederick[1]

Abstract

Purpose: Aggressive cutaneous squamous cell carcinoma (cSCC) is often a disfiguring and lethal disease. Very little is currently known about the mutations that drive aggressive cSCC.

Experimental Design: Whole-exome sequencing was performed on 39 cases of aggressive cSCC to identify driver genes and novel therapeutic targets. Significantly, mutated genes were identified with MutSig or complementary methods developed to specifically identify candidate tumor suppressors based upon their inactivating mutation bias.

Results: Despite the very high-mutational background caused by UV exposure, 23 candidate drivers were identified, including the well-known cancer-associated genes *TP53*, *CDKN2A*, *NOTCH1*, *AJUBA*, *HRAS*, *CASP8*, *FAT1*, and *KMT2C* (*MLL3*). Three novel candidate tumor suppressors with putative links to cancer or differentiation, *NOTCH2*, *PARD3*, and *RASA1*, were also identified as possible drivers in cSCC. *KMT2C* mutations were associated with poor outcome and increased bone invasion.

Conclusions: The mutational spectrum of cSCC is similar to that of head and neck squamous cell carcinoma and dominated by tumor-suppressor genes. These results improve the foundation for understanding this disease and should aid in identifying and treating aggressive cSCC. *Clin Cancer Res; 20(24); 6582–92. ©2014 AACR.*

Introduction

Cutaneous squamous cell carcinoma (cSCC) is the second most frequent cancer among Caucasians with an incidence of approximately one million cases per year (1). cSCC (25%) and basal cell carcinoma (BCC; 75%) are the major subtypes of nonmelanoma skin cancer (2). Most cSCC arise in the head and neck region because it is frequently exposed to sunlight and its ensuing UV radiation-induced DNA damage, which is the major etiologic factor (2). Immunosuppression, usually associated with organ transplantation, elevates the risk of developing cSCC by over 100-fold (3). Although cSCCs frequently respond well to conventional treatments, including electrodessication and curettage, cryosurgery, wide local excision, and radiotherapy, 3% to 5% of these tumors recur (4). According to a recent large study, patients with cSCC have a 3.7% risk of metastasis and 2.1% risk of disease-specific death (5). Clinically, aggressive cSCCs are characterized by frequent and multiple recurrences necessitating large surgical excisions, increased tendency for regional metastasis, and significant disease-related mortality (6). When aggressive or highly invasive cSCC occurs in the head and neck, surgical treatment can have profound functional, cosmetic, and psychosocial effects, sometimes leading to loss of an eye, ear, or a nose. This may require significant reconstruction and diminish quality of life.

In a large prospective study (6), a primary tumor size greater than or equal to 4 cm, the presence of perineural invasion (PNI), or invasion beyond the subcutaneous tissue were all associated with aggressive cSCC and

[1]Department of Head and Neck Surgery, The University of Texas MD Anderson Cancer Center, Houston, Texas. [2]Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, Texas. [3]Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, Texas. [4]Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas. [5]Department of Dermatology, The University of Texas MD Anderson Cancer Center, Houston, Texas. [6]Department of Immunology, The University of Texas MD Anderson Cancer Center, Houston, Texas. [7]Department of Bioinformatics and Institute for Applied Cancer Science, The University of Texas MD Anderson Cancer Center, Houston, Texas. [8]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

C.R. Pickering and J.H. Zhou contributed equally to this article.

D.A. Wheeler and M.J. Frederick contributed equally to this article.

Corresponding Authors: David A. Wheeler, Director, Cancer Genomics, Human Genome Sequencing Center, 1619.01N, Baylor College of Medicine, Houston, TX 77005. Phone: 713-798-7206; Fax: 713-798-4373; E-mail: wheeler@bcm.edu. Mitchell J. Frederick, Department of Head and Neck Surgery, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Unit 123, Houston, TX 77007. Phone: 713-792-4708; Fax: 713-745-2234; E-mail: mfrederi@mdanderson.org

doi: 10.1158/1078-0432.CCR-14-1768

Translational Relevance

The mutational spectrum of aggressive cutaneous squamous cell carcinoma contains a signature of exposure to UVB, which should aid in the definitive diagnosis for squamous tumors and metastases with ambiguous site of origin. In addition, the mutational landscape is dominated by tumor-suppressor genes, resembling that of head and neck squamous cell carcinoma, but includes novel candidate drivers. Many of the mutated genes are related to differentiation pathways. *KMT2C* mutations are associated with poor outcome and could represent a new biomarker for aggressive disease. They also suggest an epigenetic component to this disease that could possibly be targeted. Mutations in *HRAS* and *STK19* are candidate oncogenic events, but are not yet targetable. These findings expand our knowledge of this disease and should aid in the development of genomically driven treatments.

significantly decreased 3-year disease-specific survival. The American Joint Committee on Cancer (AJCC) has also recognized aggressive features of cSCC that lead to upstaging the disease and are associated with increased risk of recurrence or metastasis, including invasion of bone, tumor sizes greater than 2 cm, or presence of at least two high-risk factors such as poor differentiation, PNI, depth of invasion greater than 2 mm, occurrence at a high-risk site (i.e., ear or lip), or Clark level greater than or equal to 4 (7).

Although some attention has been given to targeting the *EGFR*, there are no standard effective treatments beyond surgery and radiation for cSCC (8, 9). There is an urgent need to identify new therapeutic targets for this group of patients. Knowledge regarding the genetic underpinning of this cancer remains largely rudimentary. Specific mutations in cSCC have been identified in *TP53* (10), *NOTCH* receptors (11, 12), and *RAS* (13).

Efforts to characterize the genetic landscape of cSCC have been hampered by very high background mutation rates associated with UV damage (12, 14, 15), which can be five to 15 times greater than what is found for noncutaneous tumors. The extraordinarily high background mutation rate makes it difficult to identify driver mutations from passengers. Although there have been some recent reports of exome data for cSCC (11, 12) and BCC (15), these cohorts were heterogeneous and small, making it difficult to identify potentially novel tumor drivers.

In the present study, we examined exomic mutations in a cohort of patients with aggressive cSCC. We hypothesized that the analysis of genomic data from a larger cohort of patients with clinically aggressive cSCC disease would permit more definitive characterization of the mutations that contribute to overall disease progression in this subset of disease with poorer prognosis.

Materials and Methods

Tissue processing

Fresh-frozen surgically resected tumor and patient-matched normal lymphocytes were obtained from consented patients treated for cSCC of the head and neck region at The University of Texas MD Anderson Cancer Center (Houston, TX), under an Institutional Review Board approved protocol. Frozen tissue was embedded in optimal cutting temperature (OCT) compound and completely sectioned. Tissue was washed once in PBS before isolating genomic DNA using an ArchivePure DNA purification kit (5Prime).

Library construction

Genomic DNA samples were constructed into Illumina paired-end precapture libraries according to the manufacturer's protocol (Illumina Multiplexing_SamplePrep_Guide_1005361_D) with modifications as described in the BCM-HGSC Illumina Barcoded Paired-End Capture Library Preparation protocol that is accessible from the HGSC website (https://hgsc.bcm.edu/sites/default/files/documents/Illumina_Barcoded_Paired-End_Capture_Library_Preparation.pdf).

Illumina sequencing and copy number

Four precapture libraries were pooled together and hybridized in solution to the HGSC VCRome 2.1 design1 (42 Mb; NimbleGen) according to the manufacturer's protocol NimbleGen SeqCap EZ Exome Library SR User's Guide (Version 2.2) with minor revisions. Exomes were sequenced on an Illumina HiSeq 2000 platform to an average coverage of $\times$115. For both tumor and normal samples, >80% of bases achieved a q20 quality and $\times$20 coverage. Details regarding library preparation and coverage for all samples are provided in the Supplementary Methods and Supplementary Table S4. Sequence analysis was performed using the HGSC Mercury analysis pipeline (https://www.hgsc.bcm.edu/software/mercury) to call mutations and generate BAM files (Supplementary Methods). Finalized BAM files generated from whole-exome sequencing were then used to generate copy-number data using an in house R package, BEDTools (16), and ABSOLUTE (17) to estimate the absolute copy number based on ploidy and purity (Supplementary Methods).

Significantly mutated genes

MutSigCV v1.4, which corrects for gene-specific background rates, was run on maf files that included flanking regions. IntOGen v2.3.0 software (Universitat Pompeu Fabra), which examines whether genes are enriched for potential impactful mutations beyond what is expected by chance (18), was run as an online package. Two additional algorithms were developed to recognize a bias toward inactivating mutations (see Supplementary Methods).

Statistical analyses

Wilcoxon rank sum tests and Fisher exact tests were used to test associations between individual mutation and

continuous/ordinal variables and nominal variables, respectively. For analysis of primary site, tumors from the ear or lip were grouped into one high-risk site category, and the remaining cases grouped as either preauricular, scalp, periorbital (including temple), cheek, or other. Pearson correlation coefficients and *P* values were provided for assessing associations between total number of mutations and continuous/ordinal variables. Kruskal–Wallis tests were used to test associations between total number of mutations and nominal variables. Kappa coefficient correlations computed for each pair of genes and the results were summarized in a matrix plot. We used the Cox model score test to assess whether individual genes or total number of mutations are related with any of the time-to-event outcomes. All analyses were performed using SAS software (version 9.3; SAS Institute).

Results

Patient characteristics

To comprehensively characterize the somatic mutations in aggressive cSCC, we performed whole-exome sequencing on DNA from snap-frozen tumors and matched normal blood from 39 patients. These cases were considered aggressive because they presented with regional or distant metastasis or had at least one of the features associated with increased mortality previously described by Clayman and colleagues (Fig. 1; Supplementary Tables S1 and S2; ref. 6). All cSCCs studied arose on the head or neck, with the majority from the ear, preauricular, scalp, or periorbital regions: 38.5% were recurrent, 25.6% were persistent, and 35.9% were previously untreated. Seven of the samples were from metastatic sites. The majority of cSCCs (71.8%) had invaded beyond the subcutaneous space, 48.7% had PNI present, 43.6% were poorly differentiated, and 100% of evaluable cases had Clark levels $\geq$4.

Exome sequencing

Before sequencing, samples were evaluated by a trained pathologist and found to have an average tumor cellularity of 49.4% $\pm$ 22.4% as assessed visually (Supplementary Table S3). All specimens had greater than 10% tumor nuclei by pathology and sequencing analysis. Samples were sequenced to $\times$115 average coverage (Supplementary Table S4). A total of 108,034 somatic alterations were detected in 16,588 genes (Supplementary Table S5). The total number of mutations/patient ranged from 86 to 12,112. A median of 61.2 mutations/Mb was detected in this cohort (Fig. 2A). This mutation frequency is one of the highest mutation rates ever detected, is more than four times as high as the rate in melanoma (14), and is higher than the rate in other squamous tumor types (Fig. 2A). These tumors did not contain functionally relevant *POLE* mutations (19). In addition, the tumors were largely clonal. Less than 15% of mutated genes were found to have a low allele fraction (Supplementary Table S6), and only four samples had a subclonal genome fraction >15% (Supplementary Table S3). Although many of the tumors were previously treated there was no difference in the number (Supplementary Fig. S1A) or types (Supplementary Fig. S1B) of mutations between treated and untreated tumors. Instead, the vast majority of mutations appear to be caused by exposure to UV light, as is expected for skin tumors. UVB exposure is known to cause C>T transitions often following a pyrimidine base. When averaged across the cohort 75% of events were C>T transitions (Fig. 2B and C) and 87% of those were at a C following a pyrimidine base. In addition, 5.6% of events were dinucleotide polymorphisms (DNP), which are another signature of UV exposure (Fig. 2B and C). A few select tumor types are shown for comparison (Fig. 2B).

Interestingly, four of the tumors did not appear to have a signature of UVB exposure (Fig. 2C). They had a much lower rate of C>T mutations (average 39% C>T) and a lower overall number of mutations (average 294). This mutation signature is more similar to that of human papilloma virus–negative head and neck squamous cell carcinoma (HNSCC), with a C>T frequency around 40% (Fig. 2B; ref. 20). These tumors represent all of the nose tumors in the cohort. Subsequent review of these clinical

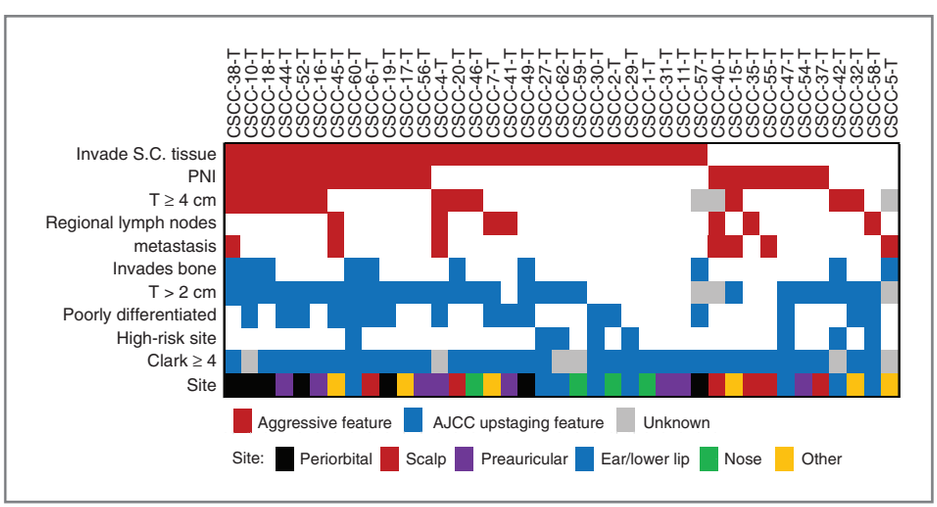

Figure 1. Aggressive features of cSCC. Aggressive features used to define this cohort are shown in red. Features that result in upstaging by AJCC criteria are shown in blue.

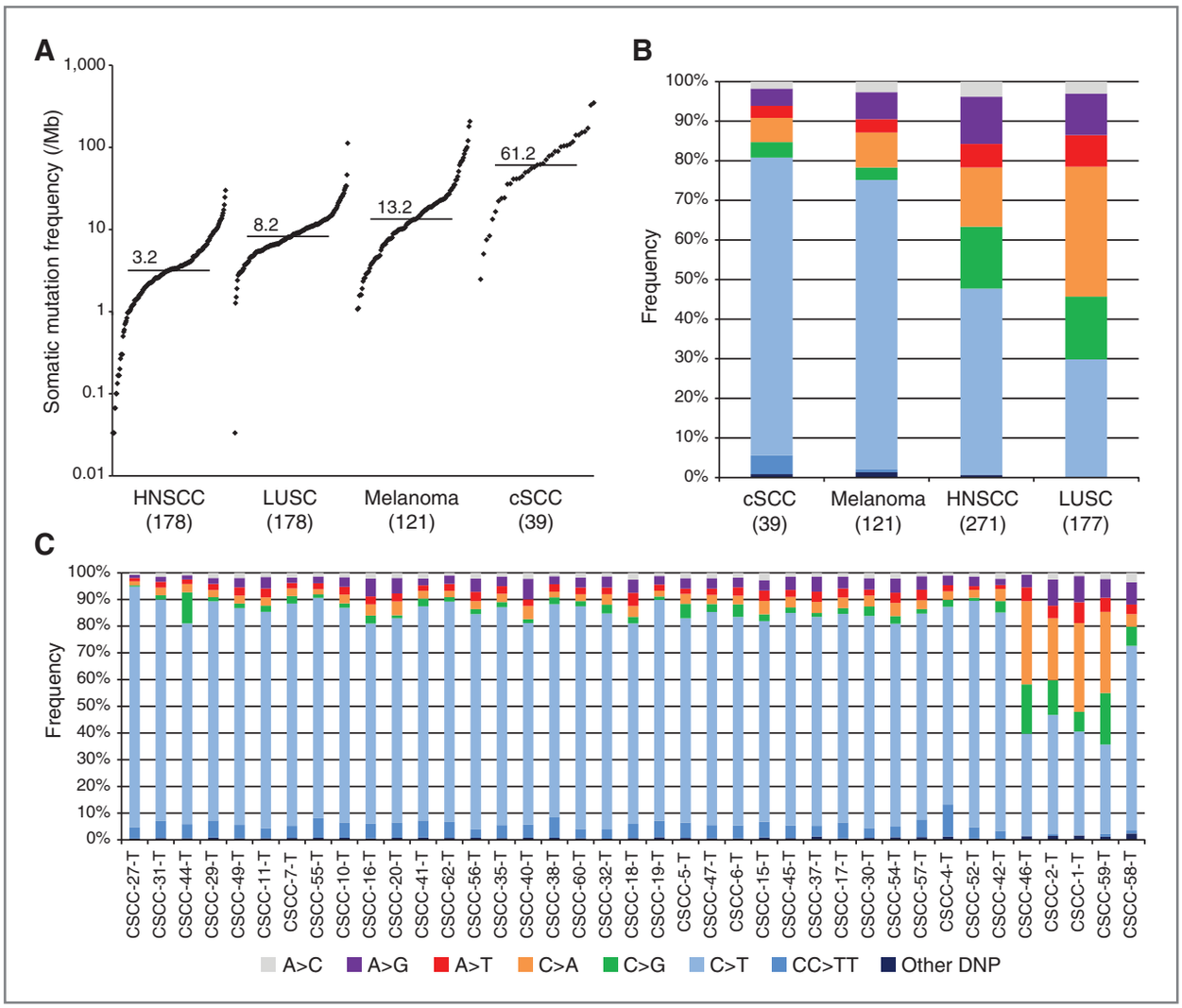

Figure 2. Mutation frequency and types in cSCC. A, mutation frequency in cSCC compared with other tumor types. Median value is shown and indicated by the horizontal line. Non-cSCC samples are TCGA data from Lawrence et al. (24). B, mutation types in cSCC compared with other tumor types. Mutation type frequencies were calculated for each sample and then averaged across the cohort to eliminate bias from highly mutated samples. HNSCC and LUSC are TCGA data from Kandoth et al. (20) Melanoma data are from Hodis et al. (14). C, mutation types by patient.

histories was ambiguous to the source epithelium for the lesion, cutaneous, or mucosal. It is possible that these tumors arose from the mucosal surface of the nose but grew out to the skin surface.

Given the high background rate of mutations in cSCC it is difficult to identify candidate driver genes. Previous sequencing studies of cSCC did not include enough cases for statistical determination of drivers and relied on frequency of mutations alone. We performed MutSig analysis (21) and identified 11 genes with a q value of <0.1 (Table 1 and Supplementary Table S7); however, two of these (*RBM46* and *DCLK1*) had low allelic fractions, suggesting that they may not be true drivers (Supplementary Table S6). Because of the high background rate, we undertook additional methods to identify candidate driver genes. Analysis using the Integrative Onco Genomics (IntOGen) package (18), which scores genes according to the cumulative predicted functional impact of missense and other nonsynonymous mutations, identified 292 candidate genes below the FDR cutoff of 0.1. Only seven of the genes detected as candidate drivers by IntOGen were also significant by MutSig. Because of the large number of significantly mutated genes detected by IntOGen (Supplementary Table S8), it is likely that many were false positives or passengers. We therefore sought out additional methods to identify drivers that would have increased specificity and not be affected greatly by a high-mutation background. Tumor-suppressor genes frequently have a high proportion of inactivating mutations, and this signature has been used by others to identify putative tumor suppressors that may be drivers from cancer-sequencing data (22). We developed two algorithms to detect a bias toward inactivating mutations. The

Table 1. Genes significantly mutated in cSCC by multiple methods compared with other cancer types

	cSCC (39)					TCGA HNSC HPV (−) (243)	TCGA LSCC (178)	TCGA SKCM (278)
	MutSig *q* value	IntOGen FM *q* value	χ^2 FDR	Multinomial FDR	Patients number (%)	Patients number (%)	Patients number (%)	Patients number (%)
TP53	0[a]	0[a]	5.28E−21[a]	<1.0E−05[a]	37 (94.9%)	203 (83.5%)[a]	161 (90.4%)[a]	46 (16.5%)[a]
CDKN2A	9.34E−07[a]	3.48E−13[a]	2.81E−12[a]	2.70E−04[a]	17 (43.6%)	63 (25.9%)[a]	32 (18.0%)[a]	42 (15.1%)[a]
PEG10	3.66E−06[a]	ND	9.69E−01	1	9 (23.1%)	2 (0.8%)	3 (1.7%)	5 (1.8%)
NOTCH2	1.85E−03[a]	2.79E−07[a]	4.20E−08[a]	1.00E−04[a]	20 (51.3%)	12 (4.9%)	13 (7.3%)	22 (7.9%)
NOTCH1	2.10E−02[a]	0[a]	1.04E−05[a]	2.00E−05[a]	23 (59.0%)	49 (20.2%)[a]	15 (8.4%)	9 (3.2%)
HRAS	2.10E−02[a]	2.55E−02[a]	9.69E−01	1	8 (20.5%)	11 (4.5%)[a]	5 (2.8%)[a]	3 (1.1%)
BBS9	2.10E−02[a]	6.61E−01	3.22E−01	6.22E−01	9 (23.1%)	5 (2.1%)	7 (3.9%)	10 (3.6%)
CASP8	2.93E−02[a]	1.47E−02[a]	2.96E−01	2.30E−01	9 (23.1%)	24 (9.9%)[a]	2 (1.1%)	9 (3.2%)
DCLK1	3.39E−02[a]	ND	9.69E−01	1	17 (43.6%)	3 (1.2%)	4 (2.2%)	10 (3.6%)
RBM46	3.67E−02[a]	ND	1.03E−01	1.22E−01	13 (33.3%)	0 (0.0%)	1 (0.6%)	15 (5.4%)
AJUBA	9.99E−02[a]	7.40E−12[a]	6.82E−03[a]	8.72E−02[a]	7 (17.9%)	17 (7.0%)[a]	0 (0.0%)	0 (0.0%)
SNX25	1	2.12E−02[a]	3.95E−05[a]	1.66E−02[a]	7 (17.9%)	4 (1.6%)	3 (1.7%)	2 (0.7%)
EIF2D	1	5.43E−02[a]	3.35E−05[a]	8.81E−02[a]	2 (5.1%)	0 (0.0%)	1 (0.6%)	0 (0.0%)
PARD3	1	6.04E−02[a]	1.03E−02[a]	4.00E−04[a]	12 (30.8%)	7 (2.9%)	8 (4.5%)	15 (5.4%)
OPN3	1	1.28E−02[a]	2.26E−03[a]	3.80E−01	4 (10.3%)	0 (0.0%)	2 (1.1%)	4 (1.4%)
FBX021	1	2.13E−02[a]	2.44E−02[a]	1.22E−01	5 (12.8%)	2 (0.8%)	2 (1.1%)	6 (2.2%)
DCLRE1A	1	4.49E−02[a]	2.44E−02[a]	1.61 E−01	5 (12.8%)	2 (0.8%)	1 (0.6%)	1 (0.4%)
COBLL1	1	5.43E−02[a]	3.22E−01	6.90E−02[a]	9 (23.1%)	3 (1.2%)	5 (2.8%)	18 (6.5%)
RASA1	1	1.21E−01	8.93E−05[a]	4.19E−03[a]	5 (12.8%)	13 (5.3%)[a]	10 (5.6%)	3 (1.1%)
SEC31A	1	2.14E−01	6.28E−02[a]	8.76E−02[a]	7 (17.9%)	4 (1.6%)	5 (2.8%)	6 (2.2%)
ZNF644	1	2.42E−01	3.95E−05[a]	8.72E−02[a]	6 (15.4%)	3 (1.2%)	3 (1.7%)	5 (1.8%)
KMT2C	1	3.20E−01	1.14E−03[a]	9.79E−03[a]	15 (38.5%)	21 (8.6%)	29 (16.3%)	45 (16.2%)
FAT1	1	ND	4.29E−10[a]	1.00E−05[a]	17 (43.6%)	64 (26.3%)[a]	26 (14.6%)	27 (9.7%)
KMT2D	1	7.80E−01	1.20E−01	1.50E−02[a]	27 (69.2%)	43 (17.7%)[a]	36 (20.2%)[a]	62 (22.3%)
NFE2L2	1	ND	ND	ND	0 (0.0%)	17 (7.0%)[a]	27 (15.2%)[a]	3 (1.1%)
PIK3CA	1	4.09E−01	3.22E−01	4.03E−01	4 (10.3%)	45 (18.5%)[a]	28 (15.7%)[a]	10 (3.6%)
RAC1	1	9.43E−01	6.95E−01	ND	2 (5.1%)	8 (3.3%)[a]	2 (1.1%)	20 (7.2%)[a]
BRAF	1	1	9.69E−01	1	7 (17.9%)	3 (1.2%)	8 (4.5%)	140 (50.4%)[a]
NRAS	1	1	6.95E−01	ND	2 (5.1%)	1 (0.4%)	0 (0.0%)	86 (30.9%)[a]
STK19	1	4.02E−01	9.69E−01	1	8 (20.5%)	3 (1.2%)	4 (2.2%)	4 (3.4%)[a]

NOTE: TCGA data were collected from cBioPortal. TCGA significance was determined by MutSig in the pan-cancer analysis of Lawrence et al. (24).

Abbreviation: ND, not determined.

[a]Statistically significant genes.

first method treats all inactivating mutations (i.e., splice, frameshift, and nonsense) equally and uses the global frequency of inactivating and noninactivating mutations actually observed in the cohort along with a χ^2 statistic to calculate the likelihood that the inactivating to nonactivating ratio for a specific gene is significantly higher than expected by chance. This χ^2 method identified a list of 24 candidate tumor-suppressor genes (Supplementary Table S9), including five well-known tumor suppressors that were among the top 10 most significant genes detected (Table 1). Although the χ^2 approach can identify mutated genes with an inactivation bias, it is possible that certain genes are more predisposed to nonsense mutations because of their specific codon usage and the abnormally high rate of C>T (or G>A) mutations due to the UV signature. We therefore developed an additional method that takes into account the gene-specific codon usage and impact of UV signature to detect inactivation bias, using a different statistic based upon a multinomial probability model. Interestingly, the multinomial method largely identified a subset of the same genes found by the χ^2 analysis (Table 1 and Supplementary Table S10). The results of MutSig, IntOGen, χ^2, and multinomial analysis appear in Table 1, which includes genes identified as significant by MutSig (i.e., the gold standard) or by at least two of the other methods. *TP53*, *CDKN2A*, *NOTCH2*, *NOTCH1*, and *AJUBA* were identified as significant by all four methods, whereas *SNX25*, *EIF2D*, and *PARD3* were significant by the three alternative methods but not by

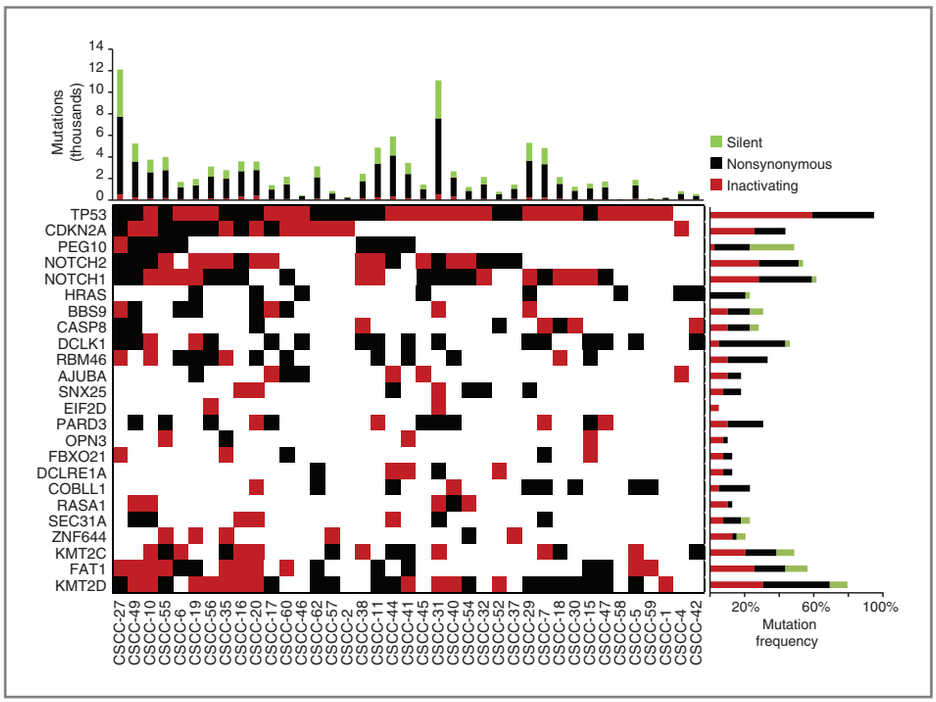

Figure 3. Key mutations in cSCC. Total number of mutations per patient is shown on the top. Inactivating mutations include nonsense, frameshift, and splice site events. Mutation frequency for each gene is shown on the right.

MutSig. Furthermore, the known tumor-suppressor genes *FAT1* and *KMT2C (MLL3)* and a putative tumor-suppressor *RASA1* (22, 23) were significant by the two methods detecting inactivation bias. The distribution for these mutations among the cohort can be seen in Fig. 3 along with the frequency of inactivating, missense, and silent mutations.

Interestingly, eight of the top 23 genes were also found to be significantly mutated in HNSCC (Table 1; ref. 24). Only four genes were significant in lung squamous cell carcinoma (LUSC) and two genes were significant in melanoma (SKCM; ref. 24). The genes common to cSCC and HNSCC are *TP53*, *CDKN2A*, *NOTCH1*, *HRAS*, *CASP8*, *AJUBA*, *RASA1*, *FAT1*, and *KMT2D*. The presence of moderately frequent mutations in eight common genes suggests that the biology of cSCC may be similar to that of HNSCC.

In both cSCC and HNSCC, NOTCH1 alterations appear to be inactivating (Fig. 4) because the missense mutations cluster in the EGF-like repeats responsible for ligand binding, and the truncating mutations are distributed throughout the gene but not clustered in the C-terminal PEST domain, in contrast with what is found for T-cell acute lymphoblastic leukemia (ALL). *NOTCH1* and *NOTCH2* were similarly mutated in >50% of cSCC cases and more than 30% of the mutations are inactivating. We recently showed that *NOTCH1* is a tumor-suppressor gene in HNSCC (25) but the role of *NOTCH2* in cancer is poorly understood. *NOTCH2* mutations are not statistically significant in HNSCC by MutSig, but do have a high inactivating mutation ratio in this cancer as well (Fig. 4).

We next compared a few genes significantly mutated in related tumor types but not in cSCC. Mutations in the oxidative stress gene *NFE2L2* were first described in LUSC (15% of cases; ref. 26) and later found in HNSCC (7%; ref. 24). No mutations in *NFE2L2* were found in this cSCC cohort. *PIK3CA* is also significantly mutated in LUSC (16%) and HNSCC (19%), but was mutated only five times in 4 cSCC patients (10%) and was not statistically significant (Table 1). In addition, the mutations included two inactivating mutations and no mutations in the classical hotspots (E545 and H1047; Supplementary Table S5).

Melanoma is characterized by frequent hotspot mutations in *BRAF* and *NRAS* (14). No hotspot mutations in *BRAF* or *NRAS* were observed in cSCC. However, hotspot mutations in *RAC1* and *STK19* have been reported in melanoma and were found in this cohort. We observed one P29S mutation in *RAC1*, and five mutations around D89 in *STK19*. These included three D89N, one E88K, and one P90S mutations (Supplementary Table S5).

We also identified two new candidate tumor-suppressor genes for cSCC. *PARD3* and *RASA1* were mutated in 31% and 13% of patients with cSCC (Table 1), with 33% and 66% of their mutations predicted to truncate or eliminate the proteins (Fig. 3), respectively. *RASA1* was identified in pan-cancer analyses and HNSCC as a candidate tumor-suppressor gene because of its high inactivation mutation ratio (22, 24).

Kappa analysis was performed to identify correlations between the mutations in each gene (Supplementary Table S11). *HRAS* was highly correlated with *AJUBA* (0.423, $P = 0.008$) and inversely associated with *TP53*

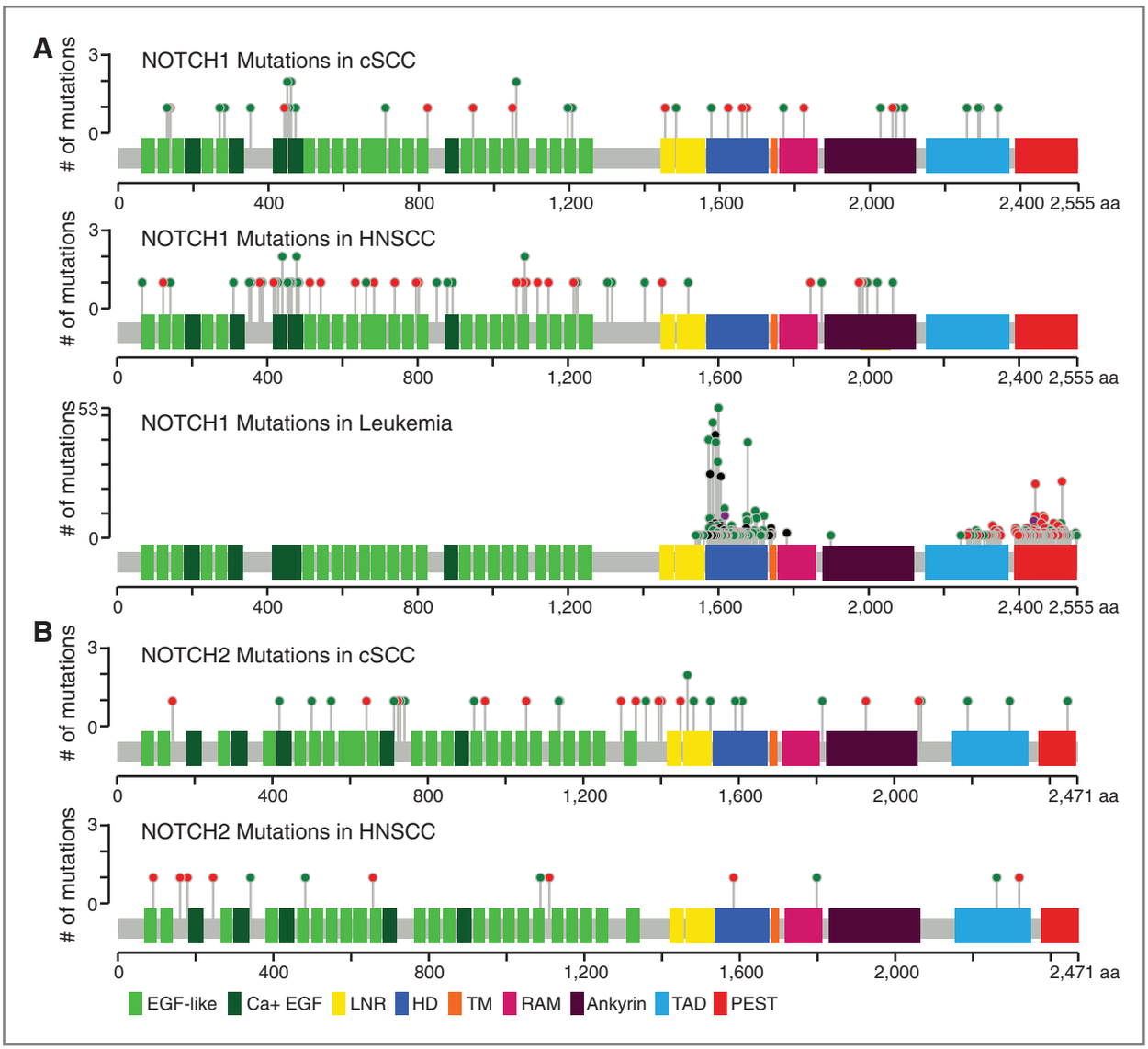

Figure 4. Mutations in *NOTCH1* and *NOTCH2* appear to be inactivating in cSCC. A, the spectrum and location of *NOTCH1* mutations observed in cSCC, HNSCC (TCGA data), are compared with T-cell ALL (COSMIC data) in which *NOTCH1* alterations are instead activating. In both cSCC and HNSCC, most missense mutations cluster in the N-terminal EGF-like repeats in which ligand binding occurs, and truncating mutations are distributed throughout the gene, whereas, in leukemia truncating mutations are confined to the C-terminal PEST domain responsible for degradation of activated intracellular NOTCH1 and missense mutations cluster in the heterodimerization domain where they cause ligand-independent activation. B, *NOTCH2* alterations in cSCC and HNSCC have a similar pattern because missense mutations are clustered in the EGF-like domains and truncating mutations are scattered throughout the gene, suggesting the alterations inactivate function. Red dots, truncating mutations (splice, frameshift, or nonsense); green dots, missense mutations; black dots, inframe insertions or deletions; and purple dots, residues affected by different mutation types. NOTCH domains are indicated with different colors. Light green, EGF-like repeat; dark green, calcium-binding EGF domain; yellow, LNR (negative regulatory) repeat; blue, heterodimerization domain (HD); orange, transmembrane (TM) region; pink, RAM domain; purple, ankyrin repeats; blue, transactivation domain (TAD); and red, proline glutamic acid serine–threonine rich (PEST) domain.

(−0.107, $P = 0.004$). *TP53* and *HRAS* are also inversely correlated in HNSCC.

Because four tumors lacked a UV signature, we repeated the MutSig analysis while excluding those tumors. The top 12 most significant genes remained at the top of the list; however, one new gene was added as the 10th most significant gene, *RIPK4* ($q = 0.053$). This gene is quite interesting because it encodes a serine/threonine kinase essential for keratinocyte differentiation (27). *RIPK4* was mutated in 28% of the tumors with a UV signature, with all mutations clustering in either exon 2 or exon 8, which encode the kinase and ankyrin repeat domains, respectively. There was also a high ratio of nonsense, frameshift, and splice mutations (35%) that was nearly significant by the two methods detecting inactivation bias, suggesting selection for inactivation of the gene in cSCC.

Copy-number alterations

Copy-number values were calculated from the exome-sequencing coverage data and adjusted for purity and ploidy by using the ABSOLUTE algorithm. Large regions of copy-number gain were frequently detected (in >25% of samples) on chromosomes 7, 8q, 9q, 14, and 20, and regions of loss were detected on 3p, 4, 5q, 8p, 9p, 11, 17p, 18, 19, and 21 (Supplementary Fig. S2 and Supplementary Table S12). The *CCND1* region of chromosome 11q was also focally amplified. In the copy-number data, as with the mutation data, there are many similarities between cSCC and HNSCC. For example, both tumor types have losses in 3p, 5q, 8p, 18, and 21 and gains in 3q, 5p, 8q, 14, and 20 (25).

Clinical significance

To begin addressing the importance of genetic alterations in cSCC, the top candidate genes along with the total number of mutations per patient were analyzed for clinicopathologic associations. This analysis correlated 29 different clinical ordinal characteristics, three continuous clinical variables, and eight different measurable parameters related to time intervals or patient status (Supplementary Table S2). Although all of the patients in this cohort were selected based upon a clinical diagnosis of cSCC, to maintain clinical uniformity the four nasal cases with ambiguous site of origin were removed from the clinical analyses. As these studies were exploratory in nature, multiple testing corrections were not applied so that sensitivity could be maximized. A summary matrix of all *P* values obtained for the clinical parameters tested by gene appears in Supplementary Table S13. The total number of mutations observed per tumor did not correlate with any of the genes or clinicopathologic parameters examined except for histologic subtype ($P = 0.02$). Tumors classified as acantholytic had a median number of mutations (3589) that was roughly 1.5 times greater than tumors with no specific histologic subtype (2295) and more than triple the median number of mutations (1,033) in tumors with sarcomatoid or adenosquamous histology.

Tumors with *AJUBA* mutations were positively correlated with depth of invasion ($P = 0.02$), and on average invaded with a depth (16.0 ± 6.4 mm) almost twice that of tumors lacking the mutation (8.4 ± 5.6 mm). The presence of PNI, a known aggressive feature, was positively associated with

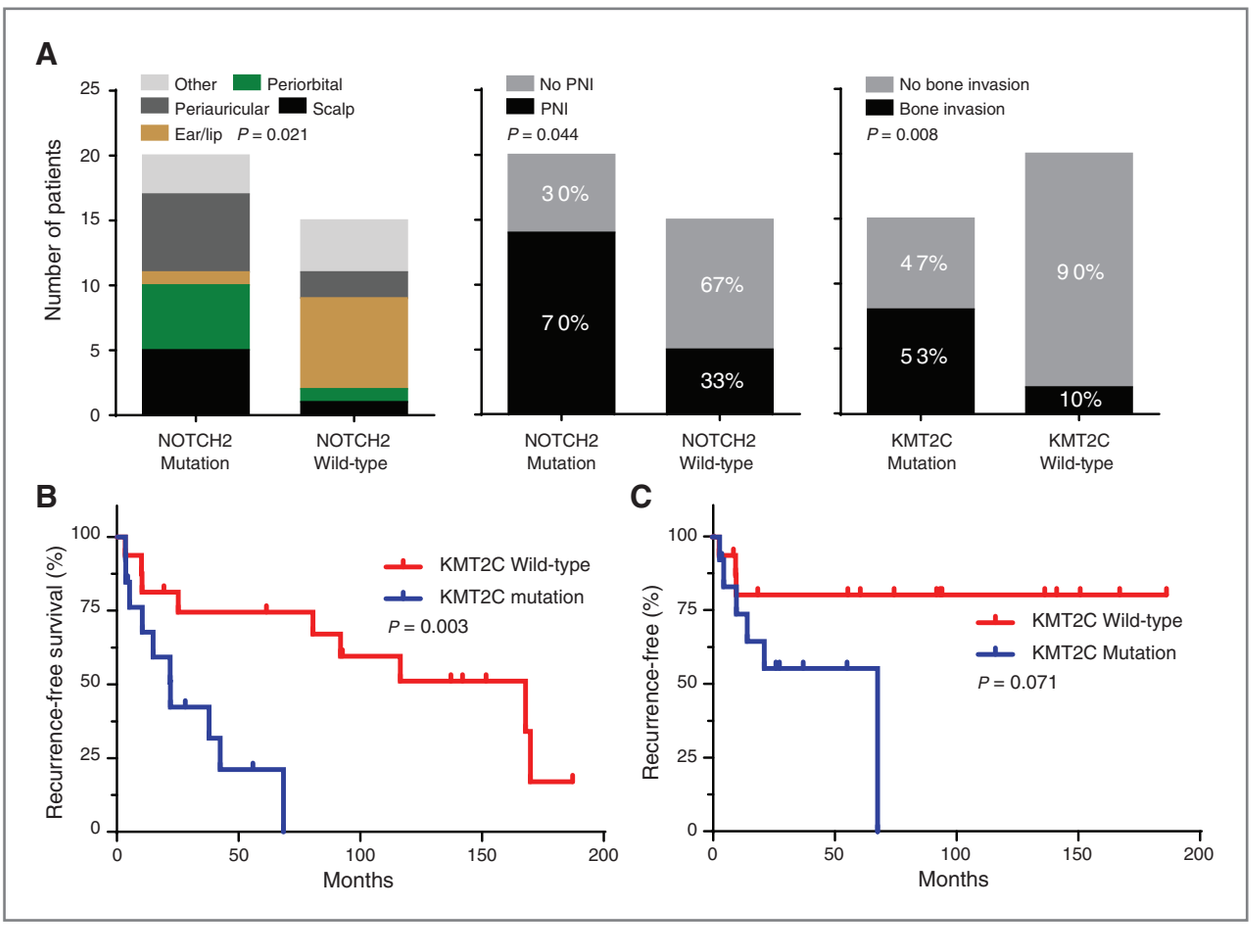

Figure 5. Mutation status and clinical parameters. A, association between mutation status and various clinical parameters. B, *KMT2C* mutation is associated with shorter time to recurrence. C, *KMT2C* mutation is associated with worse recurrence-free survival. Censored events are indicated by a vertical bar.

mutations in *NOTCH2* ($P = 0.04$; Fig. 5A). Approximately 70% of patients with *NOTCH2* mutations had PNI present compared with just 33% of patients with no *NOTCH2* mutation. Interestingly, *NOTCH2* mutations were also associated with primary tumor site ($P = 0.04$), as the presence of *NOTCH2* mutation was more common in cSCCs arising in the scalp or periorbital regions compared with the ear (Fig. 5A). The increased tendency for tumors with *NOTCH2* mutations to have PNI may be independent from tumor site, as there was no significant association between tumor site and PNI ($P = 0.19$).

There was a highly significant positive association between bone invasion and *KMT2C* mutations ($P = 0.008$). Only 10% of patients with wild-type *KMT2C* had bone invasion, compared with 53% of patients with *KMT2C* mutation (Fig. 5A). Consistent with the positive association between *KMT2C* and bone invasion, patients with *KMT2C* mutation had significantly shorter recurrent-free survival times ($P = 0.003$) with a median recurrent survival of 21.6 months compared with 167.5 months for patients with wild-type *KMT2C* (Fig. 5B). The HR for recurrence or death for patients with *KMT2C* mutation was 5.16 (95% confidence interval, 1.55–17.18) compared with those whose tumors were wild-type. Similarly, patients with tumors harboring *KMT2C* mutation had trends toward shorter time to disease recurrence ($P = 0.07$), and shorter overall survival (OS; $P = 0.09$). Poor prognosis of patients with *KMT2C* mutation appeared to be independent from bone invasion, as patients with bone invasion did not have shortened recurrent-free survival times ($P = 0.98$).

Discussion

We have generated the first list of significantly mutated candidate driver genes in aggressive cSCC. The background mutation rate in cSCC is so high that more than half of all genes were mutated and 218 genes were mutated in at least half of the patients in this study. We report what are likely to be many of the most important drivers in aggressive cSCC. Currently, very few studies have comprehensively examined the mutational landscape of cSCC. These prior studies were unable to achieve statistical significance and little clinical information exists about their cohorts. Thus, it is presently unknown whether the candidate drivers we identified are enriched in aggressive cSCC compared with cSCC with a more benign clinical course.

Six of the top genes in aggressive cSCC, *TP53*, *CDKN2A*, *NOTCH1*, *NOTCH2*, *HRAS*, and *FAT1* were previously reported by South and colleagues (12) in a cohort comprised of 20 cSCCs derived mostly from patients with immunosuppression secondary to organ transplantation. It is interesting to note that cSCCs that arise in organ transplant recipients are frequently aggressive, suggesting a common biology despite differences in predisposing risk factors. Interestingly, another group has recently characterized the genomic landscape of BCC (15), and also observed frequent mutations in *NOTCH1*, *NOTCH2*, and *TP53*. However, *PTCH1* was the only gene in BCC found to be significantly mutated. *PTCH1* mutations occurred in 75% of BCC tumors and 70% of the alterations were inactivating. In this cSCC cohort, *PTCH1* mutations occurred in just 17% of patients and only two mutations were inactivating, suggesting that *PTCH1* is not a driver in aggressive cSCC.

The mutational spectrum of cSCC is quite similar to that of HNSCC. Eight of the top mutated genes are shared between these tumor types, both derived from stratified squamous epithelia. The major mutational difference between cSCC and HNSCC is the UV signature. We identified four tumors from the nose that lacked the UV signature. Subsequent review of those patient records identified ambiguity in the source of the tumor with respect to the skin or mucosa of the nose. Some mucosal tumors in the nose can grow through the skin and present as an apparently cutaneous lesion, especially for large tumors. Similarly, nodal or parotid metastases in the head and neck region may have uncertainty regarding whether the tumor originated in skin or mucosa. These data suggest that the UV signature may be able to aid clinicians in making a definitive diagnosis in these cases.

Although high frequencies of both *NOTCH1* and *NOTCH2* mutations in cSCC were previously reported by two groups (11, 12), we were able to show for the first time that both genes are significantly mutated in cSCC by using MutSig. Recently, we demonstrated that *NOTCH1* behaves functionally as an *in vitro* and in *vivo* tumor suppressor in HNSCC (25). A similar role is likely in cSCC, because conditional knockout of *NOTCH1* in mouse skin predisposes animals to skin tumors (28). A role for *NOTCH2* in cancers is less clear because mice with conditional knockout of *NOTCH2* are not predisposed to tumors (29). However, activated *NOTCH2* can arrest the growth of keratinocytes (29) *in vitro* and combined inactivation of both *NOTCH1* and *NOTCH2* can more severely alter differentiation of skin than loss of *NOTCH1* alone (29). Collectively, the data suggest *NOTCH2* activation may inhibit tumor growth and future functional studies are needed.

It is possible that the cell of origin for the tumor may influence the relative roles of *NOTCH1* and *NOTCH2*. *NOTCH1* and its presumed ligand *JAG1* are expressed in the lower to middle epidermal layers, whereas expression of *NOTCH2* and its presumed ligands, *JAG2* and *DLL1*, are confined to the basal cells of skin (29). *NOTCH1* and *NOTCH2* may both be a barrier to carcinogenesis in some systems (human), whereas *NOTCH1* may be the primary barrier for other systems (mouse).

An interesting and novel candidate driver gene we identified was *RIPK4*, which is known to control keratinocyte differentiation (27). Inactivating mutations in this gene are associated with a severe autosomal recessive lethal disease in humans (30) known as Popliteal Pterygium Syndrome (also Bartsocas-Papas syndrome) that affects the face, limbs, and genitalia. Knockout of *RIPK4* in mice produces a similar neonatal lethal syndrome accompanied by defective epidermal differentiation, including keratinocyte hyperplasia with expanded spinous and granular layers (27). The

clustering of *RIPK4* mutations within the kinase and ankyrin repeat domains, strongly suggests that the mutations were nonrandom and support a hypothesis that *RIPK4* is a putative tumor suppressor for aggressive cSCC.

Many of the identified genes are related to differentiation signaling. These include *NOTCH1*, *NOTCH2*, *FAT1*, *AJUBA*, *CASP8*, and *RIPK4*. Several of these genes were also identified in HNSCC and linked to differentiation there as well. This suggests a common and important barrier to tumorigenesis in these squamous epithelial tumors.

Another interesting candidate tumor-suppressor gene in cSCC is *RASA1*. *RASA1* belongs to a family of RAS GTPase activating proteins, many of which have been implicated as tumor suppressors in cancer because they function to negatively regulate prooncogenic *RAS* (23). RAS GTPase family members with confirmed tumor-suppressor function include *NF1*, *DAB2IP*, and *RASAL2*, which are frequently inactivated in tumors through genomic loss, mutation, or epigenetic silencing (23). Inactivation of these genes has been proposed to explain activation of the RAS pathway in tumors that do not harboring specific *RAS* mutations. The role of *RASA1* in cancer has not been clearly defined, despite the fact that it is frequently inactivated by mutation in many other tumor types (22).

A primary goal of this study was to identify new targetable genes for the treatment of cSCC, because there are very few nonsurgical options for patients with aggressive disease. Unfortunately, we did not identify any easily targetable events. The most frequently altered genes are nearly all tumor-suppressor genes, similar to what was found in HNSCC (25). The most obvious oncogene identified is *HRAS*, which has proven difficult to target. An interesting novel target is *STK19*. The pattern of mutational clustering suggests that they may activate the kinase. Although little is currently known about the function of *STK19* and which drugs can target it, it is hoped that this genetic characterization of cSCC will lead to identification of driver pathways that will be targetable. Even if mutated tumor-suppressor genes are not directly targetable, it is possible that the pathways they define can be targetable through other codependent genes or that a synthetic lethality can be identified.

In this cohort of aggressive cSCC, we found frequent inactivating mutations in *KMT2C*, which encodes a component of a histone methylation complex involved in transcriptional regulation. Inactivating *KMT2C* mutations have been reported for a number of tumors, including cancers of the stomach (31), bladder (32), and breast (33). In the TCGA stomach cancer dataset there is a trend toward reduced OS in patients with *KMT2C* mutation compared with patients who are wild-type (median 13 vs. 59 months). In this aggressive cSCC cohort, patients with *KMT2C* mutations had significantly shorter recurrent-free survival, shorter time to recurrence, and were more likely to have bone invasion. The data support a role for *KMT2C* in the aggressive behavior of cSCC.

One of the challenges both patients and clinicians face with aggressive head and neck cSCC is the invasive tumor behavior and the innumerable recurrences that eventually require extensive surgeries; often affecting function and form. Therefore, the stakes are high in identifying the subset of cSCC that will manifest an unfavorable biology. Understanding the genomic signatures of aggressive cSCC provides an opportunity to intensify upfront therapies to prevent the morbid consequences of treating advanced disease. These results set the stage for understanding and recognizing the key drivers in this disease. The goal is to use this information to explore potential biomarkers that predict aggressive cSCC and identify genomically targeted therapies that are effective for those patients.

Disclosure of Potential Conflicts of Interest

R.A. Gibbs is a consultant/advisory board member for GE Clarient. No potential conflicts of interest were disclosed by the other authors..

Authors' Contributions

Conception and design: C.R. Pickering, J.J. Lee, J.N. Myers, D.A. Wheeler, M.J. Frederick

Development of methodology: C.R. Pickering, J. Yu, G.L. Clayman, R.A. Gibbs, M.J. Frederick

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): C.R. Pickering, J.J. Lee, R.E. Saade, K.Y. Tsai, J.L. Curry, M.T. Tetzlaff, D.M. Muzny, H. Doddapaneni, R.A. Gibbs, M.J. Frederick

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): C.R. Pickering, J.A. Drummond, S.A. Peng, R.E. Saade, K.Y. Tsai, S.Y. Lai, J. Yu, E. Shinbrot, K.R. Covington, J. Zhang, S. Seth, C. Caulin, A.K. El-Naggar, R.A. Gibbs, D.A. Wheeler, M.J. Frederick

Writing, review, and/or revision of the manuscript: C.R. Pickering, J.H. Zhou, J.J. Lee, K.Y. Tsai, J.L. Curry, M.T. Tetzlaff, S.Y. Lai, J. Yu, K.R. Covington, A.K. El-Naggar, R.S. Weber, J.N. Myers, M.J. Frederick

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): J.H. Zhou, J.A. Drummond, R.E. Saade, M.J. Frederick

Study supervision: R.A. Gibbs, J.N. Myers, D.A. Wheeler, M.J. Frederick

Other (collected the patient's clinical, pathologic, and demographic data; reviewed all the pathologic slides; acquisition of the samples; made and reviewed all the frozen section slides of the samples for quality control; isolated the DNA from the samples for the sequencing): J.H. Zhou

Other (provide specimens for analysis): A.K. El-Naggar

Acknowledgments

The authors thank the patients for their courage and generosity. The authors thank the members of the Myers laboratory for moral support, technical support, and helpful discussions.

Grant Support

This work was supported by the Cancer Prevention Research Institute of Texas grant RP100233; NIH/National Institute of Dental and Craniofacial Research grant RC2DE020958; NIH Specialized Program of Research Excellence grants P50CA097007; Cancer Center Support Grant P30CA0CA16672, and the Pantheon Program. M.J. Frederick is supported in part as a fellow of the Sheikh Khalifa Bin Zayed Al Nahyan Institute for Personalized Cancer Therapy. J.H. Zhou is supported by NIH grant T32CA163185.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received July 9, 2014; revised September 30, 2014; accepted October 1, 2014; published OnlineFirst October 10, 2014.

References

1. Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. [Erratum appears in CA Cancer J Clin. 2011 Mar-Apr;61(2):133-4]. CA Cancer J Clin 2010;60:277–300.
2. Madan V, Lear JT, Szeimies RM. Non-melanoma skin cancer. Lancet 2010;375:673–85.
3. Rangwala S, Tsai KY. Roles of the immune system in skin cancer. Br J Dermatol 2011;165:953–65.
4. Rowe DE, Carroll RJ, Day CL Jr. Prognostic factors for local recurrence, metastasis, and survival rates in squamous cell carcinoma of the skin, ear, and lip. Implications for treatment modality selection. J Am Acad Dermatol 1992;26:976–90.
5. Schmults CD, Karia PS, Carter JB, Han J, Qureshi AA. Factors predictive of recurrence and death from cutaneous squamous cell carcinoma: a 10-year, single-institution cohort study. JAMA Dermatol 2013;149:541–7.
6. Clayman GL, Lee JJ, Holsinger FC, Zhou X, Duvic M, El-Naggar AK, et al. Mortality risk from squamous cell skin cancer. J Clin Oncol 2005;23:759–65.
7. Warner CL, Cockerell CJ. The new seventh edition American Joint Committee on Cancer staging of cutaneous non-melanoma skin cancer: a critical review. Am J Clin Dermatol 2011;12:147–54.
8. Lewis CM, Glisson BS, Feng L, Wan F, Tang X, Wistuba II, et al. A phase II study of gefitinib for aggressive cutaneous squamous cell carcinoma of the head and neck. Clin Cancer Res 2012;18:1435–46.
9. Wollina U. Cetuximab in non-melanoma skin cancer. Expert Opin Biol Ther 2012;12:949–56.
10. Brash DE, Rudolph JA, Simon JA, Lin A, McKenna GJ, Baden HP, et al. A role for sunlight in skin cancer: UV-induced p53 mutations in squamous cell carcinoma. Proc Natl Acad Sci U S A 1991;88:10124–8.
11. Durinck S, Ho C, Wang NJ, Liao W, Jakkula LR, Collisson EA, et al. Temporal dissection of tumorigenesis in primary cancers. Cancer Discov 2011;1:137–43.
12. South AP, Purdie KJ, Watt SA, Haldenby S, den Breems NY, Dimon M, et al. NOTCH1 mutations occur early during cutaneous squamous cell carcinogenesis. J Invest Dermatol 2014;24:154.
13. Su F, Viros A, Milagre C, Trunzer K, Bollag G, Spleiss O, et al. RAS mutations in cutaneous squamous cell carcinomas in patients treated with BRAF inhibitors. N Engl J Med 2012;366:207–15.
14. Hodis E, Watson IR, Kryukov GV, Arold ST, Imielinski M, Theurillat JP, et al. A landscape of driver mutations in melanoma. Cell 2012;150:251–63.
15. Jayaraman SS, Rayhan DJ, Hazany S, Kolodney MS. Mutational landscape of basal cell carcinomas by whole-exome sequencing. J Invest Dermatol 2014;134:213–20.
16. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2010;26:841–2.
17. Carter SL, Cibulskis K, Helman E, McKenna A, Shen H, Zack T, et al. Absolute quantification of somatic DNA alterations in human cancer. Nat Biotechnol 2012;30:413–21.
18. Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, Tamborero D, Schroeder MP, Jene-Sanz A, et al. IntOGen-mutations identifies cancer drivers across tumor types. Nat Methods 2013;10:1081–2.
19. Shinbrot E, Henninger EE, Weinhold N, Covington KR, Goksenin AY, Schultz N, et al. Exonuclease mutations In DNA Polymerase Epsilon reveal replication strand specific mutation patterns and human origins of replication. Genome Res. 2014 Sep 16. [Epub ahead of print].
20. Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational landscape and significance across 12 major cancer types. Nature 2013;502:333–9.
21. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature 2013;499:214–8.
22. Davoli T, Xu AW, Mengwasser KE, Sack LM, Yoon JC, Park PJ, et al. Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. Cell 2013;155:948–62.
23. Maertens O, Cichowski K. An expanding role for RAS GTPase activating proteins (RAS GAPs) in cancer. Adv Biol Regul 2014;55:1–14.
24. Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature 2014;505:495–501.
25. Pickering CR, Zhang J, Yoo SY, Bengtsson L, Moorthy S, Neskey DM, et al. Integrative genomic characterization of oral squamous cell carcinoma identifies frequent somatic drivers. Cancer Discov 2013;3:770–81.
26. Cancer Genome Atlas Research N. Comprehensive genomic characterization of squamous cell lung cancers. Nature 2012;489:519–25.
27. Holland P, Willis C, Kanaly S, Glaccum M, Warren A, Charrier K, et al. RIP4 is an ankyrin repeat-containing kinase essential for keratinocyte differentiation. Curr Biol 2002;12:1424–8.
28. Nicolas M, Wolfer A, Raj K, Kummer JA, Mill P, van Noort M, et al. Notch1 functions as a tumor suppressor in mouse skin. Nat Genet 2003;33:416–21.
29. Massi D, Panelos J. Notch signaling and the developing skin epidermis. Adv Exp Med Biol 2012;727:131–41.
30. Kalay E, Sezgin O, Chellappa V, Mutlu M, Morsy H, Kayserili H, et al. Mutations in RIPK4 cause the autosomal-recessive form of popliteal pterygium syndrome. Am J Hum Genet 2012;90:76–85.
31. Je EM, Lee SH, Yoo NJ, Lee SH. Mutational and expressional analysis of MLL genes in gastric and colorectal cancers with microsatellite instability. Neoplasma 2013;60:188–95.
32. Gui Y, Guo G, Huang Y, Hu X, Tang A, Gao S, et al. Frequent mutations of chromatin remodeling genes in transitional cell carcinoma of the bladder. Nat Genet 2011;43:875–8.
33. Ellis MJ, Ding L, Shen D, Luo J, Suman VJ, Wallis JW, et al. Whole-genome analysis informs breast cancer response to aromatase inhibition. Nature 2012;486:353–60.