Genomic analysis of metastatic cutaneous squamous cell carcinoma

Yvonne Y. Li¹,², Glenn J. Hanna³, Alvaro C. Laga⁴, Robert I. Haddad¹,⁵,⁶, Jochen H. Lorch¹,⁵,⁶,*, Peter S. Hammerman¹,⁵,⁶*

Affiliations
¹Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA; ²Broad Institute, Cambridge, MA; ³Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA; ⁴Department of Dermatology, Brigham and Women's Hospital; ⁵Department of Medicine, Brigham and Women's Hospital, Boston, MA; ⁶Department of Medicine, Harvard Medical School, Boston, MA.

*These authors contributed equally to this article.

Running Title: Genomics of metastatic cutaneous squamous carcinomas

Key words: squamous cell carcinomas, cancer genomics, skin cancer

Grant Support: PSH is supported by NCI K08 CA163677

Corresponding authors
Peter Hammerman, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, MA 02215. Phone: 617-632-6049. Fax: 617-632-5786. E-mail: Peter_Hammerman@dfci.harvard.edu.
Jochen Lorch, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, MA 02215. Phone: 617-632-3090. Fax: 617-632-4448. E-mail: Jochen_Lorch@dfci.harvard.edu.

Disclosure of Potential Conflicts of Interest

J.H. Lorch has received research support from Novartis. No potential conflicts of interest were disclosed by the other authors.

Word count: 5,243

4 figures and 2 tables
Statement of Translational Relevance

The vast majority of cutaneous squamous cell carcinomas (cSCCs) are treated effectively with simple surgical excision. However, in approximately 5% of cases metastatic disease develops and is associated with very poor clinical outcomes. There are no therapies approved by the FDA with a specific indication for metastatic cSCC and development of novel agents has been slow, likely due to a limited knowledge of the molecular basis of this disease. Here, we performed a next-generation sequencing study of 29 individuals with metastatic cSCC to describe the key genomic alterations in cSCC and enumerate potential therapeutic targets. We identify multiple genes which display recurrent mutation, amplification, and deletion in this disease, including several alterations which have been or are being pursued as therapeutic targets in other cancer types. Together, our data present an initial genomic portrait of metastatic cSCCs and suggest that patients with this disease may benefit from biomarker-associated therapeutic agents under evaluation in other cancers, namely squamous cell carcinomas of the lung and head and neck.
Abstract

Purpose: A rare 5% of cutaneous squamous cell carcinomas metastasize, lack FDA-approved therapies, and carry a poor prognosis. Our aim was to identify recurrent genomic alterations in this little-studied population of metastatic cSCCs.

Experimental Design: We performed targeted sequencing of 504 cancer-associated genes on lymph node metastases in 29 patients with cSCC and identified mutations and somatic copy number alterations associated with metastatic cSCC. We determined significantly mutated, deleted and amplified genes and associated genomic alterations with clinical variables.

Results: The cSCC genome is heterogeneous with widely varying numbers of genomic alterations and does not appear to be associated with HPV. We found previously identified recurrently altered genes (TP53, CDKN2A, NOTCH1/2) but also a wide spectrum of oncogenic mutations affecting RAS/RTK/PI3K, squamous differentiation, cell cycle, and chromatin remodeling pathway genes. Specific mutations in known oncogenic drivers and pathways were correlated with inferior patient outcomes. Our results suggest potential therapeutic targets in metastatic cSCC including PIK3CA, FGFR3, BRAF, and EGFR, similar to those reported in SCCs of the lung and head and neck, suggesting that clinical trials could be developed to accrue patients with SCCs from multiple sites of origin.

Conclusions: We have genomically characterized a rare cohort of 29 metastatic cSCCs and identified a diverse array of oncogenic alterations that can guide future studies of this disease.

Introduction

Nonmelanoma skin cancers are the most common type of cancer in the United States, with over 3.5 million new cases diagnosed annually (1). Cutaneous squamous cell carcinoma (cSCC) comprises 20% of these cases, and its incidence is continuing to rise (2). 95% of cSCCs are curable with surgical resection; however, 5% metastasize – usually to nearby lymph nodes - leading to a 3-year disease-free survival rate of 56% (3) and a 5-year survival rate of 25-35% (4-7). Therapies for patients with metastatic cSCCs are lacking and have been limited by a lack of knowledge of the genomic alterations that drive metastatic cSCCs. In addition, there are no validated molecular biomarkers predictive of disease behavior or treatment response.

Numerous risk factors for the development of cSCC have been identified, including exposure to ultraviolet radiation, ionizing agents, and chemical carcinogens (8). Approximately 65% of cSCCs arise from premalignant precursor conditions such as actinic keratosis (9). Organ transplant recipients on immunosuppression regimens...
are 65-times more likely to develop cSCC (10). Human papillomavirus (HPV) infection has also been associated with increased risk in developing cSCC (11) and patients on chemotherapy targeting BRAF frequently develop cSCCs with RAS mutations (12). The risk factors for developing metastases are less characterized, with an analysis of 615 patients showing only tumor thickness associated with a significant risk of metastasis (13). A second study found that a combination of risk factors (tumor diameter, differentiation histology, perineural invasion, and tumor invasion) improved upon previous staging systems to predict clinically aggressive cSCCs with poor outcome (14).

Genomic characterization of cSCC has mostly been performed on small cohorts of samples. Exome analysis of 8 and 11 primary patient tumors, respectively, identified a large mutational burden of 33.3 mutations per megabase (Mb) of coding sequence, recurrent TP53 mutations and copy number loss (15), and recurrent NOTCH family loss-of-function mutations (16). SNP array analysis of 60 tumors identified loss of heterozygosity at 3p and 9p in 65-75% of the samples (17). Targeted analysis of the CDKN2A locus in 40 samples identified alterations (mutation, copy loss, promoter methylation) in 76% of cases (18). Microarray comparison of 10 actinic keratosis and 30 cSCC samples identified several MAPK pathway genes significantly overexpressed in the malignant samples (19). Similar findings were reported by studies involving larger cohorts of primary cSCCs: targeted sequencing of the known NOTCH1/2, TP53, CDKN2A, and RAS genes on 132 cSCCs that developed sporadically and 39 cSCCs that developed after BRAF-inhibitor treatment (20), and exome sequencing of 39 clinically aggressive cSCC primaries (21). Recently, missense mutations in the kinetochore-associated protein KNSTRN has emerged as a novel potential driver of cSCC, recurring in approximately 19% of cSCC cases (22). Genomic understanding of metastatic cSCCs is limited, though VEGFA overexpression has been linked to lymphatic metastasis in mouse models (23).

The evaluation of biomarker-driven targeted therapies in cSCCs has been limited. Most trials are exploring EGFR-targeted therapy, as advanced tumors often show upregulated EGFR expression without RAS mutations (24, 25) - observations similar to those made in SCCs of the head and neck and lung. However, some studies have found no correlation of EGFR overexpression with the malignant phenotype (26). Clinical activity of EGFR antagonists in cSCCs has been observed, with a surprising 18% complete response rate in a phase II trial of gefitinib (27), suggesting that further refinement of the subset of cSCC patients likely to respond to EGFR therapy is needed. A more comprehensive understanding of metastatic SCC is necessary to identify genomic characteristics and target pathways for this aggressive disease. Here, we sequenced 29 cSCC lymph node metastases to search for recurrent genomic alterations and better define potential avenues for clinical trial development and therapy.

**Methods**

**Sample selection and sequencing**
Cases of cSCC with lymph node metastases were identified from the Dana-Farber Cancer Institute-Harvard Cancer biorespository in accordance with standards established by the Institutional Review Board. All cases underwent a secondary review by a Board Certified Dermatopathologist who verified the diagnosis and identified the optimal portions of the section for isolation of tumor DNA and DNA from adjacent normal areas. Tissue from these areas was isolated from the FFPE block using a small bore punch biopsy needle and the resultant cores were used for DNA isolation using the Qiagen FFPE DNA extraction kit. DNA was quantified and quality controlled by Nanodrop and pico-Green assays prior to library construction.

Samples were sequenced using the OncoPanelv2 platform (28, 29), a targeted Illumina sequencing strategy aimed to simultaneously detect mutations, translocations and copy-number variations in archived clinical tumor specimens. Targeted sequencing was achieved by designing RNA baits to capture the exons of 504 genes with relevance to cancer. The bait set was augmented with specific intronic sequences to detect translocations often involved in cancer. Sequencing was performed using 100bp reads on an Illumina HiSeq 2500. The reads were aligned to human reference genome b37 using Picard and the Firehose pipeline at the Broad Institute. The BAM files are in the process of being submitted to dbGAP.

Relevant de-identified clinical data were abstracted from the patient charts in accordance with an IRB approved protocol.

**Variant calling**

Variant calling (SNVs, indels) was performed using the Firehose pipeline running Mutect (30) and filtering out OxoG artifacts. We also removed likely germline mutations that were previously seen in both dbSNP build 134 and 1000 Genome data using Oncotator (http://www.broadinstitute.org/oncotator/) (31-35). Significance analysis was conducted using MutsigCV, Mutsig2.0, and Mutsig1.5, which incorporate different methods of calculating background mutation rates. Mutsig 1.5 estimates background rate using synonymous mutations. Mutsig2.0 estimates enrichment of mutations at evolutionarily conserved positions and the clustering of mutations at gene hotspots. Finally, MutSigCV considers gene expression, replication time, and chromatin state when calculating background rate. Given that we started with a set of cancer genes, we took a less stringent approach to the analysis: we ran all three versions of MutSig and considered the most significant value from the three methods.

We considered mutations overlapping positions in the COSMIC database more likely to be cancer-associated. To lower the noise of this analysis, we only considered mutations seen in at least three cancer samples in COSMIC. For nonsynonymous mutations in oncogenes, we performed a detailed literature search to determine whether these mutations had previously been functionally validated in vitro.
Copy number analysis was performed using Nexus7.5 (BioDiscovery Inc) after calculating the sequencing coverage using GATK tools. Coverages were normalized over GC-content using lowess regression, and log2Ratios of coverage were calculated using a best fit reference that resulted in the lowest variance. CNAs were called using the following NGS settings: significance threshold of 1E-4; no maximum contiguous probe spacing; at least six probes per segment; CNAs with a log ratio greater than 0.3 were called gains and greater than 0.6 were called high gains; the single copy loss threshold was set at -0.5 and high loss was -1. X/Y chromosomes were not analyzed.

Recurrent copy number changes were detected using GISTIC 2.0 (36). We provided the segments covered by the Oncopanel platform, and set segments for the rest of the genome to be copy neutral. We then used GISTIC to search for peaks of copy number recurrence in covered areas of the genome. To reduce noise generated by the many discontinuous segments, which would more easily appear significant against the neutral background of the untargeted genome, we chose to apply GISTIC’s arm-level peel correction which has previously been used in a similar setting where multiple discontinuous segments were causing noisy GISTIC peaks (37). We also increased the minimum segment size from 4 to 6 to encourage joining of Oncopanel segments.

**Pathway analysis**

To calculate recurrent percentages, we considered an alteration to be activating if it landed in a known oncogene and was either a known activating mutation based on literature search or highly amplified. Similarly, we considered alterations to be inactivating if they were nonsense mutations or homozygous deletions. Missense or other nonsynonymous mutations in COSMIC were taken into consideration as a mutations of unknown functional effect but potentially associated with cancer. Missense mutations not present in COSMIC and were not previously validated in the literature were not included.

**Statistical analysis**

Statistical analyses to test for correlation between genomic and clinical features were performed using standard R packages. We used the Fisher’s exact test for discrete variables, the log-rank test for continuous variables, and the Bonferonni method of multiple testing correction.

**Results**

**Clinical characteristics of the metastatic cSCC cohort**

We sequenced DNA from 29 cSCCs, 26 with matched normal skin, to determine somatic copy number alterations (CNAs) and mutations (SNVs) in these tumors. All samples were lymph node metastases, with available clinical and survival data...
The primary tumors were predominantly of the head and neck, with the parotid gland being the most frequent site of metastasis. There were 19 males and 10 females, with median age of 74 at diagnosis of metastatic cSCC. 11 of the patients subsequently developed recurrent disease, with an average progression free survival of 37 months. 12 patients (41%) were recurrence-free at 3 years, slightly lower than the previously observed disease-free survival rate of 56%. The samples were also independently validated to be HPV-negative by a combination of p16 immunohistochemistry and hybrid-capture based DNA sequencing of HPV E6 and E7 genes (Pathogenica).

**Landscape of genomic alterations**

We performed targeted sequencing of 504 cancer-associated genes on the cohort to an average fold coverage of 82x (range: 25-166x) in the tumor samples and 69x (range 15-219x) in the normal samples, and identified somatic SNVs and CNAs. C→T transition mutations were the dominant substitution, constituting 67% of the mutation spectrum, consistent with the role of UV light exposure in this disease. UV light damages DNA by forming covalent links between adjacent pyrimidines (38), consistent with our observation that 87% of the C→T transitions occurred after a pyrimidine. We did not observe a high rate of the TpC→G mutation type, which has been previously described in HPV and other virally-driven cancers (39).

A genomic overview of SNVs and CNAs is shown in Figure 1A. The two stacked histograms show the number of each type of SNV and CNA per sample and the coverage plot shows the average sequencing depth achieved per sample. The 26 paired samples had fewer SNVs on average (59 nonsynonymous mutations per sample) compared to the three unpaired samples (117 nonsynonymous mutations per sample), and likely more accurately reflect the mutation rate in this tumor type given an enhanced ability to filter contaminating germline events. The average mutation rate across the 504 sequenced genes was 33 per Mb in the paired samples, varying highly (4-117 per Mb) depending on the sample.

In contrast to SNV rates, CNA rates did not appear dependent on the presence of a matched normal sample when following standard filtering procedures optimized to remove germline copy number variants. The number of genes with copy number alterations also varied highly: 2 samples had no genes altered, whereas 6 samples had over 200 genes altered. There was no correlation between the total number of SNVs and CNAs in each sample and the depth of coverage, suggesting that the variation may be biologically based and not confounded by tumor cellularity.

Overall, metastatic cSCC appears to be a genomically complex and heterogenous disease, with large differentials in mutation rate and allelic imbalance across the samples. The full list of SNVs and CNAs are supplied in Supplementary Tables 1 and 2.

**Overview of SNV alterations**
Figure 1B depicts recurrently mutated genes in metastatic cSCC which exhibited statistical evidence of selection for mutation as determined by the Mutsig algorithm (40) or which did not reach statistical significance but have well-annotated roles in other cancer types. Mutsig is a computational tool that examines gene-specific background mutation rates and assigns significance based on whether a gene is mutated at a probability higher than chance given the mutational patterns observed in the dataset. The table is ordered by number of recurrences and include previously identified cSCC tumor suppressors: TP53, CDKN2A, and NOTCH1/2, including both truncating mutations and mutations at sites previously annotated in the Catalog of Somatic Mutations in Cancer (COSMIC) database (41). TP53 was mutated in 23 of 29 (79%) samples and CDKN2A was altered by both mutation and homozygous loss in 14 samples (48%). These findings, in addition to the lack of HPV sequences detected in the tumor DNA, agreed with the independent validation of a lack of HPV in our sample cohort. Lastly, NOTCH genes showed inactivating mutations in seven samples (24%) but if we included nonsynonymous SNVs of unknown functional significance, the rate increased to 69%, similar to the 75% rate noted previously (16).

RIPK4, a regulator of squamous epithelial differentiation, has been previously reported as recurrently mutated in head and neck SCCs (42). RIPK4 was also recurrently altered in our cSCC cohort, with mutations in seven samples (24%). Two of these mutations were truncating, suggesting a recurrent inactivation of this gene. Another reported tumor suppressor in head and neck SCC is SMAD4, a gatekeeper gene that maintains genomic stability (43). Haploinsufficiency of SMAD4 is thought to lead to genome instability as well as metastasis and inflammation. In our cohort, SMAD4 had COSMIC mutations in two samples (7%).

There were known gain-of-function oncogene mutations in 11 of 29 samples (38%), though recurrent events were rare in our cohort (Fig. 3A heatmap, Table 2). Two cases had BRAF mutations (G464R/G469R) - G469R, for example, has been reported in 1% of BRAF-mutated melanomas (44). One case had a KRAS G12C mutation and another had an EGFR S720F mutation; both are rare mutations previously identified in other types of SCC including lung, anal, and tonsil (35, 45). An additional case had an FGFR3 transmembrane domain G380R mutation that renders the protein constitutively active and is known to cause the autosomal dominant disease achondroplasia (46). One case had a KIT exon 11 E562D mutation; exon 11 mutations are prevalent in 66% of gastrointestinal stromal tumors (47). Four additional mutations have also been validated as activating: HRAS G13D is common in bladder, thyroid, and kidney cancers (48); ERBB4 E563K is one of the ERBB4 mutations reported in 19% of melanoma patients (49); and EZH2 Y641S is one of the most common recurrent mutations in certain types of lymphoma (50). The functional significance of PIK3CA P471L and HGF E199K is unclear; however, these two mutations were also observed in two of the 11 cSCC samples previously analyzed by next-generation sequencing (15). Mutations in the coiled-coil domain of CARD11 have been described in diffuse large B-cell lymphoma (51) and recently, mutations in the CARD domain have been shown to disrupt CARD11 autoinhibition
and activate the protein (52). Two of the CARD11 mutations we observed (E24K, D199N) are located in this domain. It was interesting to note that nearly all of these mutations are mutually exclusive (Figure 3A), in which each activating mutation belongs to a distinct tumor.

Half (48%) of the samples had truncating or COSMIC mutations in one or more chromatin remodeling genes. CREBBP and EP300 are histone acetyltransferases and have truncating mutations in 6 and 3 samples, respectively. EP300 is a known transcriptional coactivator of NOTCH pathway genes. Notably, the truncating mutations in NOTCH1, NOTCH2, NOTCH4, and EP300 are mutually exclusive across the samples. MLL2, a histone methyltransferase that is frequently mutated in non-Hodgkin lymphomas (53), demonstrated nonsense mutations in five samples. Three members of ARID family gene transcription factors (ARID1A, ARID2, and ARID5B) had likely inactivating mutations in five samples, and ARID2 was the mostly recurrently inactivated chromatin modifying enzyme with truncating mutations in 10% of cases. The SWI/SNF complex member SMARCA4 had a splice site mutation in one sample. The EZH2 activating mutation Y641S as mentioned above was also seen in one sample. Thus epigenetic dysregulation may be a recurrent oncogenic mechanism in metastatic cSCC.

Overview of copy number alterations

Copy number alterations in the 504 cancer-associated genes were analyzed using GISTIC, which finds recurrent gains and losses against a multi-factored background (including length, amplitude, known fragile sites, surrounding sequence context, among other factors) (36). We observed 25 significantly amplified and 11 significantly deleted genes using a standard GISTIC q-value threshold of 0.25 (Fig. 2). Peaks that cluster together (i.e. around MYC) suggest a potentially broader event whereas isolated peaks containing only one gene (i.e. TP63) may indicate a more focal event.

The most significantly recurrent loss was at 9p21, including the cell cycle regulators CDKN2A and CDKN2B, which showed loss in 6 samples (21%). Numerous genes were recurrently gained across the samples, including the MYC and EGFR oncogenes. TP63 was amplified in seven samples (24%), and has been previously observed at a similar frequency in lung SCCs (35). TP63 has also been identified as an oncogene involved in squamous cell differentiation in mouse SCC models (54). We did not observe high-level (more than one copy) amplification of PIK3CA or SOX2, additional genes on chromosome 3q that have been implicated in the pathogenesis of squamous cell carcinomas (35).

When we focused on only the high-level amplifications, the most recurrently altered oncogene was LAMA5, in four samples. LAMA5 may be associated with the metastatic nature of this patient cohort, as it is strongly expressed and promotes migration in melanoma cells (55).
Though there were tumor suppressors recurrently amplified, such as FANCC or SDHB, these may be passenger events for nearby genes that were not targeted in our hybrid capture panel.

**Pathway overview and potential therapeutic targets**

Many of most significant and functionally characterized somatic alterations we identified belong to cancer signaling pathways. We were able to categorize the SNVs and CNAs described above into four major categories: the RAS/RTK/PI3K pathway, cell cycle pathway, squamous cell differentiation pathway, and chromatin remodeling genes. We then examined the well-characterized pathways in a detailed supervised analysis to identify additional altered genes which could impact these core signaling pathways. We only included alterations that appeared to be pathway-activating (known activating mutation or high-level amplification), pathway-inactivating (nonsense mutation or homozygous loss), or likely functional (present in multiple COSMIC tumor samples).

The majority of the activating mutations affected genes in the RAS-RAF-MEK-ERK and PI3K/AKT pathways (Fig. 3A): the receptor tyrosine kinases FGFR3, KIT, EGFR, ERBB4; receptor ligand HGF; RAS family members KRAS, HRAS; RAF family member BRAF, MTOR, and PI3K family member PIK3CA. Aside from an activating mutation, EGFR was also significantly recurrently amplified (Fig. 2), though only one sample had a high-level gain (Fig. 3A). One sample had a nonsense STK11 mutation in addition to a TP53 nonsense mutation. STK11 negatively regulates the PI3K/AKT pathway via AMPK-TSC1/2-mTOR and loss of both TP53 and STK11 has been shown to induce lung SCC in mouse models (56). Two samples had a COSMIC mutation in the tumor suppressor PTEN, also a negative regulator of the PI3K/AKT pathway. Though only one was a truncating mutation, the other was also a likely inactivating mutation as both are seen in numerous (~60) samples in COSMIC. NF1 is a negative regulator of RAS and had a COSMIC mutation in 3 samples (10%). The heatmap in Figure 3A illustrates the trend towards mutual exclusivity for these pathway alterations.

Aside from TP53 and CDKN2A, we found recurrent alterations in other cell cycle pathway genes including RB1, MYC, CDK4, CDK6, and CCND1 (Fig. 3B). MYC was amplified in ten samples, though only highly amplified in one, while CCND1 was amplified in four. Two samples each had a high-level gain in CDK4 and CDK6, respectively. One sample had a nonsense mutation in ATR, for which truncating mutations are recurrent in endometrial cancers and associated with a poorer overall survival (57).

Numerous genes involved in the squamous cell differentiation pathway were recurrently altered in our cohort (Fig. 3C). TP63 amplification, NOTCH1 and NOTCH2 mutations have been previously reported in lung as well as head and neck SCCs (35, 42). Inactivation of the NF-kB pathway, which is required for keratinocyte differentiation in vivo (58), is also implicated via mutations in RIPK4, amplifications...
of TP63 (transcriptional activator of RIPK4) and NFKBIA. There are also amplifications and activating mutations in CARD11 that appear to activate the NF-kB pathway based on previous reports (59). SMAD4 is a gatekeeper gene in head and neck SCC, as shown in studies where knockout mice developed spontaneous head and neck SCC (43) or induced differentiation of mammary epithelial cells into squamous epithelial cells, leading to SCC (60). In our cohort, one sample in which we detected no TP53, NOTCH 1/2/4, or oncogenic activating mutations had a CDKN2A COSMIC missense mutation and a SMAD4 nonsense mutation.

In short, though no highly recurrent mutation was identified, we found that mutations activating the RAS/RTK/PI3K, cell cycle, and squamous cell differentiation pathways are recurrent across the samples and present opportunities for biomarker driven clinical trials in this patient cohort.

Genomic or clinical factors correlated to prognosis

In an exploratory analysis, we searched for correlations between the significantly altered genes in Figure 1B and the clinical factors listed in Table 1 such as recurrence and PFS. There was a trend for immunocompromised patients to have recurrence and worse prognosis (log rank p-value of 0.017); however, the small number of immunocompromised patients (n=4) limits this observation. Similarly, the lack of a validation cohort in the literature significantly dampens enthusiasm for clinical correlations found in our work.

No single gene significantly correlated with a clinical factor. A broader analysis of all genes altered more than three times in our cohort revealed hypothesis-forming associations between ARID5B and CARD11 alterations and PFS (Supplementary Fig. 1). The association of CARD11 activating mutations or amplifications with a better prognosis is interesting given that these alterations likely activate the NF-kB pathway and promote differentiation.

Given that alterations causing RAS pathway activation, cell cycle pathway inactivation, squamous cell differentiation, and chromatin remodeling gene inactivation are recurrent across the metastatic cSCC samples, we assessed whether these pathways correlated with a clinical factor. There was no correlation between the cell cycle alterations (in TP53 or CDKN2A) nor the squamous differentiation alterations (in TP63, NOTCH1, or NOTCH2) and prognosis or other clinical variables. In contrast, both the RAS/RTK/PI3K pathway and chromatin remodeling mutations were significantly correlated with a worse prognosis, and a combination of both types of mutations increased the significance of the correlation, suggesting that these may be independent predictors (Fig. 4). We chose functionally or clinically relevant RAS/RTK/PI3K alterations that are known or very likely to be activating the pathway (those circled in Figure 3A and also present in Table 2) with the exception of the EGFR and ERBB4 mutations, which were instead associated with a long PFS. The average PFS for RAS/RTK/PI3K pathway-mutated samples without EGFR/ERBB4 mutations was 12 months, for non-RAS/RTK/PI3K pathway-mutated
samples was 50 months, and for the EGFR/ERBB4 samples was 79 months. Similarly, we chose chromatin remodeling mutations that were likely to be functionally relevant; this included truncating mutations in ARID2 and NF2, and missense mutations in EZH2 and SMARCB1 that were previously seen in COSMIC. The chromatin remodeling gene mutations correlated with a worse progression-free-survival and suggest that epigenetic dysregulation plays a role in metastatic cSCC.

Analysis of overall survival data supported the correlation among samples with mutations in chromatin modifiers or mutations with chromatin modifiers and/or RAS/RTK/PI3K and poor outcome, though the correlation among RAS/RTK/PI3K and poor outcome was not supported (Supplementary Fig. 2). A much larger sample size would be needed to characterize these observations further.

**Discussion**

The cohort in this study represents the rare 5% of cSCC tumors that have metastasized and have a poor clinical prognosis. Of the 29 patients studied, 11 exhibited recurrence within an average of 24 months. However, the actual times to recurrence varied from 1 month to 78 months, and some patients are still recurrence-free at 130 months. This suggests that there may be genomic or clinical features that can distinguish between these two types of prognoses within the metastatic cohort. The frequency of recurrence also underscores the need for improved therapeutic options for this patient population.

We identified recurrent somatic mutations and copy number alterations in metastatic cSCCs. The top three recurrently altered genes are TP53, CDKN2A, and NOTCH1/2/4, at frequencies similar to previous reports of both cSCCs (15, 16, 21, 22) and squamous cancers from other sites such as lung or HPV-negative head and neck (31, 32, 42, 61). TP53 was mutated in 79% of samples, CDKN2A altered in 48%, and NOTCH1/2/4 in 69%. The prevalence of somatic TP53 mutations is in concordance with the HPV-negative assessment of the samples; as the E6 protein of HPV binds to TP53 and marks it for degradation – an independent mechanism from mutational inactivation. Unlike the TP53 and CDKN2A mutations, the majority of NOTCH mutations were missense mutations not present in COSMIC. Thus, given the high mutation rate in this tumor type (33 mutations per Mb cancer-associated coding sequence), it may be more conservative to estimate that NOTCH family members are inactivated in ~25% of our metastatic cSCC cohort. It should also be mentioned that TP53, CDKN2A, and NOTCH genes can be inactivated by mechanisms other than somatic mutation and deletion and that the rates of loss of these genes may be higher than that observed in the context of our analysis.

Oncogenic alterations activating the RAS/RTK/PI3K pathway were present in 45% of samples and – aside from EGFR/ERBB4 mutations - significantly correlated with a worse PFS. Currently, the principal target being evaluated in clinical trials in cSCC is EGFR with some clinical activity reported to date but no prospectively validated biomarker for patient selection. In our dataset we observed two samples with
potential EGFR activation: one with a rare EGFR activating mutation and a second with high-level amplification. However, numerous samples had activations in other receptor tyrosine kinases (KIT, FGFR3, ERBB4), downstream kinases (KRAS, HRAS, BRAF), and genes in the PI3K/AKT pathway (MTOR, PIK3CA, PTEN, STK11). These potential targets are currently being investigated in clinical trials of other tumor types and we feel that including patients with cSCCs should be considered. Given that 1) many of these alterations converge on key downstream mediators of cellular survival and proliferation such as MEK and mTOR, 2) recent data from other groups showing that combined BRAF and MEK inhibition blocked proliferation in a mouse model of cSCC (62), and 3) mTOR-based inhibitors reduced the risk of developing cSCCs in immune-compromised patients (63), we feel that evaluation of such strategies for patients with cSCCs is warranted.

The alterations identified in our cohort exhibit similarities to other SCCs studied to date. Somatic alterations activating PIK3CA, HRAS, TP63, CCND1, EGFR, MYC, and inactivating TP53, CDKN2A, NOTCH1, NOTCH2, RIPK4, and SMAD4 have been previously described in head and neck SCC (42, 61). In particular, mutations of TP53 almost exclusively occurred in HPV-negative head and neck SCCs, which is consistent with the HPV-negative nature of our cohort. Lung SCC has also been noted to be similar to HPV-negative head and neck SCC, with mutations in PIK3CA, PTEN, TP53, CDKN2A, NOTCH1, and HRAS (35), which are shared in metastatic cSCC. In esophageal SCC, mutations in TP53, CDKN2A, and NOTCH1 are also recurrent, and a recent study has identified frequent dysregulation in RTK-MAPK-PI3K signaling, G1-S cell cycle regulation, and epigenetic modification (64). Numerous other recurrent alterations have been identified in SCCs – for example, ASCL4 loss-of-function mutations and FOXP1 focal deletions in lung SCC; however, these genes were not part of our targeted panel.

Recent exome sequencing of a cohort of 39 clinically aggressive cSCC primary tumors that presented with metastases found no clinically targetable oncogenes, though nonsynonymous mutations in the oncogenes HRAS and STK19 were identified (21). Similarly, targeted sequencing of 100 cSCC primary tumors confirmed previous rates of recurrent tumor suppressor mutations but did not identify recurrent oncogenic mutations aside from KNSTRN (22). In our targeted sequencing cohort of 29 metastatic tumors, we found gain-of-function mutations in 12 oncogenes across 13 samples, including clinically targetable BRAF, FGFR3, PIK3CA and EGFR mutations. In addition, two kinase mutations in our cohort, PIK3CA P471L and HGF E199K, and NOTCH4 W309* were previously identified in a cohort of 11 primary cSCCs, suggesting that these mutations may have functional roles in metastatic cSCCs. Along these lines, other oncogenic mutations may also be recurrent at a low prevalence. Thus, larger cohort studies are necessary to identify both recurrent and more unique oncogenic alterations. Assessment of the whole genome, transcriptome, and methylome on a future cohort may also identify relevant structural variations, alterations in other cancer-associated genes, mutations in non-coding regions, or methylated genes. However, comprehensive
transcriptome sequencing may be challenging given the rarity of metastatic cSCC cases.

In short, we have sequenced a rare cohort of metastatic cSCCs and identified a wide spectrum of oncogenic mutations in known oncogenes and tumor suppressors novel to this tumor type. These mutations mostly fell in RAS/RTK/PI3K pathway and chromatin remodeling genes, and appeared to be significantly correlated with PFS within metastatic cSCC patients. The results of our study suggest that agents currently undergoing investigation in clinical studies for other cancer types (such as MEK/mTOR/FGFR/BRAF/PI3K inhibitors) should be considered for individuals with cSCC, and given the similarity among the genomic alterations found in cSCC, HPV-negative HNSCC and lung SCC, that it may be prudent to include patients with SCCs of various sites of origin in clinical studies. Many of the mutations we identified in metastatic cSCCs were not previously seen in genomic studies of primary cSCCs; however, more studies of larger cohorts will be needed to differentiate the genomic events important to each type of tumor. Given the rare nature of this cohort our analysis was limited by the quality of material available for analysis and sub-optimal sequencing coverage in some of the samples may have limited our ability to detect important genomic alterations. Further, the use of more global analysis techniques such as whole-exome, whole-genome or transcriptome sequencing on larger cohorts of patients with cutaneous SCCs will be needed to provide a more complete understanding of the most critical genomic alterations in this disease.

Acknowledgements

We would like to thank the Dana-Farber Cancer Institute Center for Genome Discovery (P. van Hummelen, A. Thorner, M. Ducar, B. Wollison) for their assistance with sample processing and analysis and the Pathology Core (C. Namgyal) at the Brigham and Women’s Hospital for assistance with the tissue blocks.

References


Table 1: Cohort description table

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>19</td>
</tr>
<tr>
<td>female</td>
<td>10</td>
</tr>
<tr>
<td><strong>Age at diagnosis of nodal metastasis in years (median; range)</strong></td>
<td>74; 48-92</td>
</tr>
<tr>
<td><strong>Immune status</strong></td>
<td></td>
</tr>
<tr>
<td>immunocompromised</td>
<td>4</td>
</tr>
<tr>
<td>not immunocompromised</td>
<td>25</td>
</tr>
<tr>
<td><strong>Smoking</strong></td>
<td></td>
</tr>
<tr>
<td>yes</td>
<td>12</td>
</tr>
<tr>
<td>no</td>
<td>17</td>
</tr>
<tr>
<td><strong>Recurrence</strong></td>
<td></td>
</tr>
<tr>
<td>yes</td>
<td>11</td>
</tr>
<tr>
<td>no</td>
<td>18</td>
</tr>
<tr>
<td><strong>Received radiation therapy</strong></td>
<td></td>
</tr>
<tr>
<td>yes</td>
<td>6</td>
</tr>
<tr>
<td>no</td>
<td>23</td>
</tr>
<tr>
<td><strong>Prior diagnosis of cSCC</strong></td>
<td></td>
</tr>
<tr>
<td>yes</td>
<td>12</td>
</tr>
<tr>
<td>no</td>
<td>17</td>
</tr>
<tr>
<td><strong>Progression-free survival in months (median; range)</strong></td>
<td>37; 1-130</td>
</tr>
<tr>
<td><strong>Overall survival in months (median; range)</strong></td>
<td>60; 7-155</td>
</tr>
</tbody>
</table>

Table 2: All functionally validated or likely activating mutations identified in metastatic cSCC samples

<table>
<thead>
<tr>
<th>Gene</th>
<th>Mutation</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>BRAF</em></td>
<td>G464R</td>
<td>clinically-relevant activating (rare in melanoma)</td>
</tr>
<tr>
<td><em>BRAF</em></td>
<td>G469R</td>
<td>clinically-relevant activating (1% of melanomas)</td>
</tr>
<tr>
<td><em>KRAS</em></td>
<td>G12C</td>
<td>clinically-relevant activating</td>
</tr>
<tr>
<td><em>FGFR3</em></td>
<td>G380R</td>
<td>clinically-relevant activating (94% of achondroplasia)</td>
</tr>
<tr>
<td><em>KIT</em></td>
<td>E562D</td>
<td>functionally validated activating (exon 11 mutation in 60% of GISTs)</td>
</tr>
<tr>
<td><em>HRAS</em></td>
<td>G13D</td>
<td>functionally validated activating</td>
</tr>
<tr>
<td><em>EGFR</em></td>
<td>S720F</td>
<td>functionally validated activating (5% of NSCLCs)</td>
</tr>
<tr>
<td><em>ERBB4</em></td>
<td>E563K</td>
<td>functionally validated activating</td>
</tr>
<tr>
<td><em>EZH2</em></td>
<td>Y641S</td>
<td>functionally validated activating</td>
</tr>
<tr>
<td>Gene</td>
<td>Mutation</td>
<td>Function Description</td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
<td>----------------------------------------------------------</td>
</tr>
<tr>
<td>MTOR</td>
<td>S2215F</td>
<td>(22% of FLs) functionally validated activating</td>
</tr>
<tr>
<td>PIK3CA</td>
<td>P471L</td>
<td>likely activating (same mutation in (15))</td>
</tr>
<tr>
<td>HGF</td>
<td>E199K</td>
<td>likely activating (same mutation in (15))</td>
</tr>
<tr>
<td>CARD11</td>
<td>E24K</td>
<td>likely activating (gain of function mutations in CARD domains in vitro)</td>
</tr>
<tr>
<td>CARD11</td>
<td>D199N</td>
<td>likely activating (gain of function mutations in CARD domains in vitro)</td>
</tr>
</tbody>
</table>

**Figure Legends**

Figure 1 – Integrated view of selected recurrently altered genes

1A - Genomic overview of sequencing and variant calling. The top three plots show the distribution of CNA types across the samples, distribution of SNV types across the samples, and the average coverage of tumor samples and their matched normal samples where available.

1B - Heatmap representation of selected recurrently altered genes. CNAs are colored in red for high-level amplification events and green for homozygous deletion events. For simplicity, low-level CNA events are not shown. SNVs are colored by type in purple, beige, or blue, and also labeled: I for insertion or deletion (indel), S for missense, C for COSMIC, * for truncating, and O for other types of nonsynonymous mutation (splice site, non-stop). Significantly mutated genes as determined by Mutsig CV are those whose q-scores pass threshold of 0.1 (or –logQ-value greater than 1) on the left-hand plot. The genes are listed in order of decreasing number of alterations across the samples, as shown on the right-hand plot.

Figure 2 – Overview of copy number alterations

GISTIC plot showing the most recurrently gained (red) and lost (blue) loci in metastatic cSCCs. Peaks are considered significant if they pass the q-value threshold of 0.25. The majority of peaks contain only one gene, as we determined CNAs using 504 cancer-associated genes.

Figure 3 – Recurrently altered pathways in metastatic cSCC

Pathway diagrams depicting the percentage of samples with alterations in 3A - RAS/RAF/MEK/ERK and PI3K/AKT signaling
3B - cell cycle, and
3C - squamous cell differentiation

Alterations are classified as activating (high-level amplification or known activating mutation colored red), inactivating (homozygous loss or truncating mutation colored blue), or potentially cancer associated (COSMIC mutation colored white). For each pathway, we show integrated heatmaps (similar to Fig 1B) to show the detailed alteration pattern of each gene; however, we now also include light red and
light green to represent low-level CNAs. Note that each heatmap is sorted independently across the samples, to best illustrate the pattern of mutations, such as mutual exclusivity or concurrence.

Figure 4 – Alterations in RTK/RAS/PI3K pathway and chromatin remodeling genes associated with progression-free survival (PFS) in metastatic cSCC. Kaplan-Meier survival curves of metastatic cSCC patients comparing patients with or without mutation in A) RTK/RAS/PI3K pathway, B) chromatin remodeling genes, or C) both.
Figure 1

A)

B)
Figure 2

GISTIC q-value

chromosome

SDHB
TP63
DAXX
EGFR
MYC
FANCC
NFkB1A
NOTCH3, TPM4
LAM5

GISTIC q-value

CDKN2A
CDKN2B
Figure 3

A) RAS/RTK/PI3K pathway

- **Pathway legend**
  - Red: high-level amplification or activating mutation
  - Blue: homozygous loss or inactivating mutation
  - Empty: COSMIC mutation of unknown effect

```
Pathway:

<table>
<thead>
<tr>
<th>Gene</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>HGF</td>
<td>2%</td>
</tr>
<tr>
<td>EGFR</td>
<td>2%</td>
</tr>
<tr>
<td>ERBB4</td>
<td>2%</td>
</tr>
<tr>
<td>FGFR3</td>
<td>2%</td>
</tr>
<tr>
<td>KIT</td>
<td>2%</td>
</tr>
<tr>
<td>MET</td>
<td>2%</td>
</tr>
<tr>
<td>RAS</td>
<td>5%</td>
</tr>
<tr>
<td>NF1</td>
<td>10%</td>
</tr>
<tr>
<td>NF2</td>
<td>17%</td>
</tr>
<tr>
<td>KRAS</td>
<td>2%</td>
</tr>
<tr>
<td>HRAS</td>
<td>2%</td>
</tr>
<tr>
<td>BRAF</td>
<td>10%</td>
</tr>
<tr>
<td>AKT</td>
<td>2%</td>
</tr>
<tr>
<td>STK11</td>
<td>2%</td>
</tr>
<tr>
<td>PTEN</td>
<td>2%</td>
</tr>
<tr>
<td>PIK3CA</td>
<td>5%</td>
</tr>
</tbody>
</table>
```

B) cell cycle pathway

```
Pathway:

<table>
<thead>
<tr>
<th>Gene</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>MYC</td>
<td>5%</td>
</tr>
<tr>
<td>CDKN2A</td>
<td>24%</td>
</tr>
<tr>
<td>CDKN2B</td>
<td>18%</td>
</tr>
<tr>
<td>TP53</td>
<td>10%</td>
</tr>
<tr>
<td>ATR</td>
<td>1%</td>
</tr>
<tr>
<td>CCND1</td>
<td>2%</td>
</tr>
<tr>
<td>CDK4</td>
<td>2%</td>
</tr>
<tr>
<td>CDK6</td>
<td>1%</td>
</tr>
<tr>
<td>RB1</td>
<td>1%</td>
</tr>
</tbody>
</table>
```

C) squamous differentiation pathway

```
Pathway:

<table>
<thead>
<tr>
<th>Gene</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOTCH1</td>
<td>17%</td>
</tr>
<tr>
<td>NOTCH2</td>
<td>7%</td>
</tr>
<tr>
<td>NOTCH4</td>
<td>3%</td>
</tr>
<tr>
<td>TP63</td>
<td>2%</td>
</tr>
<tr>
<td>RIPK4</td>
<td>7%</td>
</tr>
<tr>
<td>CARD11</td>
<td>7%</td>
</tr>
<tr>
<td>CREBBP</td>
<td>2%</td>
</tr>
<tr>
<td>EP300</td>
<td>2%</td>
</tr>
<tr>
<td>NF1</td>
<td>10%</td>
</tr>
<tr>
<td>NF2</td>
<td>2%</td>
</tr>
</tbody>
</table>
```

**Heatmap legend**

- Red: likely pathway-activating mutation
- Gray: likely pathway-inactivating mutation
- Blue: high-level loss
- Green: high-level gain
- Pink: one copy gain
- Orange: one copy loss
- Brown: copy-neutral
- Yellow: indel
- Black: missense mutation
- Cyan: missense mutation
- Green: frameshift mutation
- Red: truncating mutation
- Purple: frameshift mutation
- Brown: nonsense mutation
- Green: in-frame deletion
- Red: in-frame insertion
- Grey: in-frame insertion
- Orange: in-frame deletion
- Yellow: in-frame deletion
- Red: frameshift deletion
- Orange: frameshift insertion
- Green: frameshift insertion
- Brown: frameshift insertion

**CNAs legend**

- Red: loss
- Blue: gain
- Green: copy-neutral
- Yellow: indel
- Orange: missense mutation
- Purple: frameshift mutation
- Grey: frameshift deletion
- Red: frameshift insertion
- Orange: frameshift insertion
- Green: frameshift insertion
- Brown: frameshift insertion
- Pink: frameshift deletion
- Cyan: frameshift deletion
- Green: nonsense mutation
- Yellow: nonsense mutation
- Red: frameshift insertion
- Orange: frameshift insertion
- Grey: frameshift insertion
- Red: frameshift deletion
- Orange: frameshift deletion
- Grey: frameshift deletion
- Pink: frameshift deletion
- Cyan: frameshift deletion
- Green: frameshift deletion
- Orange: frameshift deletion
- Grey: frameshift deletion
- Red: frameshift deletion
- Orange: frameshift deletion
- Grey: frameshift deletion
- Pink: frameshift deletion
- Cyan: frameshift deletion
- Green: frameshift deletion
- Orange: frameshift deletion
- Grey: frameshift deletion
- Red: frameshift deletion
- Orange: frameshift deletion
- Grey: frameshift deletion
- Pink: frameshift deletion
- Cyan: frameshift deletion
- Green: frameshift deletion
- Orange: frameshift deletion
- Grey: frameshift deletion
- Red: frameshift deletion
- Orange: frameshift deletion
- Grey: frameshift deletion
- Pink: frameshift deletion
- Cyan: frameshift deletion
- Green: frameshift deletion
- Orange: frameshift deletion
- Grey: frameshift deletion
- Red: frameshift deletion
- Orange: frameshift deletion
- Grey: frameshift deletion
- Pink: frameshift deletion
- Cyan: frameshift deletion
- Green: frameshift deletion
- Orange: frameshift deletion
- Grey: frameshift deletion
- Red: frameshift deletion
Figure 4

(A) Progression-free survival probability with log-rank p-value of 0.00289.
- No such mutation (n=19)
- RAS/RAF pathway activating mutation (n=10)

(B) Progression-free survival probability with log-rank p-value of 0.01498.
- No such mutation (n=22)
- Chromatin gene functional mutation (n=7)

(C) Progression-free survival probability with log-rank p-value of 0.0006.
- No such mutation (n=14)
- RAS/RAF pathway or chromatin gene mutation (n=15)
Clinical Cancer Research

Genomic analysis of metastatic cutaneous squamous cell carcinoma

Yvonne Y Li, Glenn J. Hanna, Alvaro C. Laga, et al.

Clin Cancer Res Published OnlineFirst January 14, 2015.

Updated version
Access the most recent version of this article at:
doi:10.1158/1078-0432.CCR-14-1773

Supplementary Material
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2015/01/15/1078-0432.CCR-14-1773.DC1

Author Manuscript
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.