Published in final edited form as:

Cancer. 2017 April 01; 123(7): 1184–1193. doi:10.1002/cncr.30459.

Differential mutation frequencies in metastatic cutaneous squamous cell carcinomas versus primary tumors

Ayşe Selen Yilmaz, MS^{1,2}, Hatice Gulcin Ozer, PhD^{1,2}, Jessica Gillespie, BA^{2,3}, Dawn C. Allain, MS^{2,4}, Madison N. Bernhardt, BS^{2,4}, Karina Colossi Furlan, MD^{2,3}, Leticia Tambasco Castro, MD^{2,3}, Sara B. Peters, MD^{2,5}, Priyadharsini Nagarajan, MD, PhD⁶, Stephen Kang, MD^{2,7}, O. Hans Iwenofu, MD^{2,8,10}, Thomas Olencki, DO^{2,9,10}, Theodoros N. Teknos, MD^{2,7,10}, and Amanda Ewart Toland, PhD^{2,3,4,9,*}

¹Department of Biomedical Informatics

²The Ohio State Wexner Medical Center, The Ohio State University, Columbus, OH

³Department of Cancer Biology and Genetics

⁴Department of Internal Medicine, Division of Human Genetics

⁵Department of Pathology, Division of Dermatopathology

⁶Department of Pathology, University of Texas MD Anderson Cancer Center, Houston, TX

⁷Department of Otolaryngology-Head and Neck Surgery, Arthur G. James Cancer Hospital and Richard J. Solove Research Institute

⁸Department of Pathology and Laboratory Medicine

⁹Division of Medical Oncology

¹⁰Comprehensive Cancer Center, and Arthur G. James Cancer Hospital and Richard J. Solove Research Institute

Abstract

BACKGROUND—Exome and targeted sequencing studies have identified potential driver mutations for a variety of tumor types. Cutaneous squamous cell carcinoma (cSCC) is one of the most highly mutated cancers but is typically associated with low rates of metastasis and high survival rates. Nevertheless, metastatic cSCC is a significant health threat; up to 8800 individuals die yearly from this disease.

METHODS—As it is difficult to predict which cSCCs are more likely to metastasize, and because there are no targeted therapies specifically designated for metastatic cSCC, we performed exome and/or targeted sequencing of 18 metastatic and 10 primary cSCCs to identify mutations that were

*Correspondence: 460 W. 12th Avenue, Columbus, OH 43210; Ph: 614-247-8185; Amanda.toland@osumc.edu.

Conflict of Interest: The authors state no conflict of interest

AUTHOR CONTRIBUTIONS:

Conceptualization: TO, TNT, AET; **Methodology:** ASY, HGO, JG; **Investigation:** JG; **Formal Analysis:** ASY, HGO, AET; **Data Curation:** ASY, DCA, MNB, KCF, LCT, AET; **Resources:** DCA, MNB, SBP, PN, SK, OHI, TO, TNT; **Visualization:** LCT, AET; **Writing-original draft:** ASY, AET; **Writing-review and editing:** ASY, HGO, JG, DCA, MNB, KCF, LCT, SBP, PN, SK, OHI, TO, TNT, AET; **Supervision:** AET; **Project Administration:** AET; **Funding acquisition:** TNT, TO, AET

more frequent in metastatic tumors and might be targeted for therapeutic benefit. We compared our results to published sequencing results of an additional 223 primary tumors and 68 metastatic cSCCs.

RESULTS—We identified genes showing higher mutation frequencies in metastatic cSCC relative to primary tumors including the chromatin remodeling gene *KMT2D* and the classic skin tumor suppressor *TP53* which was mutated in 54% of primary tumors relative to 85% of metastatic tumors ($p < 0.0001$).

CONCLUSIONS—These studies uncover potential pathways important in metastatic cSCC that broaden our understanding of the biology contributing to aggressive tumor behavior and may lead to new therapeutic strategies.

Keywords

Cutaneous squamous cell carcinoma; metastasis; exome sequencing

INTRODUCTION

Although only a small proportion of cutaneous squamous cell carcinomas (cSCC) metastasize, approximately 3900–8800 individuals in the United States die of metastatic cSCC each year.¹ Known predictors of metastasis include recurrence, tumors arising in scars, thickness greater than 6mm, poor differentiation, perineural invasion, invasion beyond subcutaneous fat, and anatomic site such as the lip, temple and ear.^{2–5} Immunosuppressed patients such as transplant recipients have higher rates of metastasis.⁴ Treatment efforts are hampered by an inability to predict which lesions are more likely to recur or metastasize as known correlates do not have high specificity for prognosis.

TP53 was first identified as being frequently mutated in cSCC in the 1990's; all subsequent exome and targeted deep-sequencing studies to date for cSCC have confirmed this finding.⁶ Exome and targeted deep-sequencing of primary cSCCs led to the discovery that inactivating mutations in *NOTCH1* and other NOTCH pathway genes occur in over 75% of tumors.^{7–9} More recently, a study of primary cSCCs identified novel mutations in the previously unrecognized kinetochore gene *KNSTRN* in 19% of cSCCs.¹⁰

In addition to sequencing primary cSCCs, two studies have assessed mutation frequencies in metastatic cSCCs.^{11,12} Many of the same genes were also found to be mutated. Exome sequencing studies performed for other types of mucosal SCC tumors including those arising from the esophagus and head and neck reveal that these genes are also mutated in non-cutaneous SCCs suggesting similar pathways leading to tumor development. Due to the similarities of specific genes being mutated across different types of SCC, a “squamousness” mutation profile has been proposed.¹³

To identify mutations unique to metastatic cSCCs, we performed exome sequencing studies on metastatic samples from twelve individuals, six of whom had primary tumor available for comparative analysis. Furthermore, we performed targeted sequencing studies of an additional ten primary cSCCs and six unrelated metastatic samples. We integrated and combined these data with publicly available exome and targeted deep-sequencing studies in

cSCC. Collectively, these studies reveal frequently mutated genes and pathways that are more prevalent in metastatic cSCCs, With further validation studies these may prove to be biomarkers of aggressive disease and potential therapeutic targets.

MATERIALS and METHODS

Study Participants and Samples

All studies were approved by the OSU Institutional Review Board. For individuals with a diagnosis of metastatic cSCC or primary non-metastatic cSCC, we collected formalin fixed paraffin embedded tissue samples. Inclusion criteria included being immunocompetent (i.e not a transplant recipient), having sufficient tumor tissue and a source of normal, non-tumor DNA, and a pathologist confirmed primary cutaneous SCC of non-lip origin. For targeted sequencing studies, individuals with a diagnosis of primary cSCC were selected from an existing sample collection. Hematoxylin and eosin slides for each sample were reviewed by a pathologist. Areas of normal skin, primary tumor and metastatic tumor were marked for coring. Tissue cores for tumors were taken from regions with greater than 70% tumor cells and non-necrotic regions. Tissue DNA was extracted from cores or was from macrodissected sections as described previously.¹⁴ Peripheral blood DNA, when available, was isolated in the OSU Human Genetics sample bank.

Whole exome library generation and sequencing

We successfully generated exome libraries for metastatic samples from twelve individuals along with six primary tumors giving rise to the metastatic samples and two unrelated primary tumors using Illumina Nextera Rapid Capture Exome Library Prep kit which captures approximately 37Mb of coding sequence (Supplementary Table 2). Exome libraries were prepared following the Illumina Nextera Rapid Capture Enrichment Guide with the following exceptions: The first PCR amplification was increased to 11 cycles. As our DNA consisted of small fragments, the bead ratio in the first PCR clean-up was changed to 1:1 to eliminate fragments smaller than ~175bp. This led to a loss of DNA so each sample was run in duplicate and pooled before the first hybridization step. DNAs were barcoded to allow for pooling of 6–8 samples per lane. Libraries were sequenced on an Illumina HiSeq2500 with paired-end 100bp reads. e.

Data analysis exome studies

Analyses were first completed on the sets of matched normal DNA, primary tumor and metastatic samples, focusing on genes shown previously to be highly mutated in primary cSCCs, genes in epigenetic pathways, and genes on panels used for identifying mutations in therapeutic targets. Data processing and analysis included QC filtering, alignment, variant calling and variant annotation. For QC filtering we used FASTQC for the read quality evaluation. We then trimmed adapter sequences on the 3' and 5' ends and filtered based on base quality scores (<20) and nucleotide distributions using Cutadapt Software.¹⁵ After preprocessing the sequence reads using a filtering index, we performed sequence alignment to the human reference genome, Grch37 build, using Burrows Wheeler Algorithm (BWA) aligner version 0.7.5 for paired-end data.¹⁶ For variant calling, duplicate reads were removed using Picard 1.94. GATK 2.8.1 was used to realign the sequences around the known indels

to improve alignment.¹⁷ Mutect 1.1.4 software was used to make the variant calling using paired samples.¹⁸ Indels were identified using GATK somatic indel detector version 1.4. Variant annotation was carried out on GenomOncology's GenomAnalytics platform, using Ensembl69 and dbSNP138 databases. The Broad Institute Integrative Genomic Viewer was used to confirm positives. Sequencing performance was assessed by measuring the number of reads, mapped reads, target base coverage, and read quality using Picard HsMetrics tool. Somatic exonic variants (inframe, frameshift, nonsense, missense) were included in mutation counts. Read cutoffs using the default Mutect read cutoffs were 14 for tumor samples and 8 for normal samples.

Targeted-sequencing and data analysis

Targeted sequencing using the Ion AmpliSeq Comprehensive Cancer Panel which consists of complete exon coverage for 409 genes from the Wellcome Trust Sanger Institute Cancer Gene Census was successfully completed on six additional metastatic samples plus one metastatic sample included in the original exome sequencing studies (Supplementary Table 5). Additionally, we included paired normal and SCC DNA from ten individuals with primary non-metastatic cSCCs. qPCR was utilized to ensure sufficient amplifiable material. Libraries were run on an Ion ProtonTM.

Analysis for the Ion AmpliSeq Comprehensive Cancer Panel was performed using ThermoFisher Ion Reporter software. For paired analyses, normal samples when available were used to filter germline variants. For samples lacking a paired normal, variants present in dbSNP with a frequency of greater than 0.1% and variants not in the COSMIC database present at 40% or more of the reads were eliminated from analyses. Somatic mutations had to be present in at least 9% of sequence reads with a minimum total read count of 20. Nonsense, missense, and frameshift mutations were included.

Statistics

A Fisher's exact test was used to calculate p-values for differences in mutation frequency between primary tumors and metastatic tumors. Hochberg multiple comparisons testing was used to correct for multiple tests. To determine if any one study was driving the observed differences, we performed a "take one out analysis" for all genes with p-values of 0.05 or less.

RESULTS

Participant demographics for exome study

To identify mutations that are present in metastatic cSCCs, we performed exome sequencing on samples from twelve individuals (Supplementary Table 1). Metastatic samples were from a variety of sites but over half were from lymph nodes (Table 1). The primary cSCCs giving rise to metastases were most often located on the face or ear (Table 1). Study participants were primarily males with an average age of 72.3 years at diagnosis (Table 1; Supplementary Tables 1, 2). Samples for exome sequencing included ten formalin-fixed paraffin-embedded metastatic cSCC DNA samples with a paired germline DNA sample from blood or histologically normal tissue and two additional metastatic samples for which

normal tissue was unavailable. In addition to metastatic tumors, we also obtained tissue from the primary cSCC thought to be giving rise to the metastasis for six of the twelve individuals and independent primary tumors from two individuals (Supplementary Tables 1 and 2).

Identification of Significantly Mutated Genes in Metastatic cSCC

Ultraviolet light (UV)-related cancers such as cSCC and melanoma have been reported to have the highest mutational burden of any cancers.^{7,9,19,20} After excluding any variants in dbSNP and synonymous changes, samples in our exome study showed 1 to 135 mutations per Mb sequenced. The average mutation rate was 11.8 mutations per Mb of exomic DNA sequenced after removal of the tumor with 135 mutations per Mb. The average number of reads per sample was 66 million with an 82% mapping rate and 93% paired reads (Supplementary Table 2).

Initial analyses focused on genes previously reported to have been mutated in cSCC and “potentially targetable” genes as found on clinical sequencing panels. As expected, many of the common mutations observed in our samples had been identified in previous studies of cSCCs (Figure 1, Table 2, Supplementary Table 3)^{7,8,10}. Of previously reported genes, *TP53* was mutated in 58% and NOTCH pathway genes were mutated in 66% of the metastatic samples respectively. *CDKN2A* was mutated in only two of the metastatic samples (17%), which is lower than previous studies. Our analyses also revealed high mutation rates in epigenetic regulators: *KMT2D* (also called *MLL2*) was mutated in 8/12 (67%) of metastatic samples (three truncating mutations) and *KAT6A* (also known as *MOZ*) was mutated in 4/12 (33%) of samples (two truncating mutations). Since both *KMT2D* and *KAT6A* are important for epigenetic regulation and histone modifications, we assessed the frequency of mutations in other epigenetic and histone modifiers. Epigenetic genes found to be mutated in our metastatic samples include *KMT2C* (58%), *SETD2* (50%), *ARID2* (17%), *EP300* (33%), *KMT2A* (33%), *TET2* (8%), *KDM6A* (8%) and *CREBBP* (17%).

A second category of genes exhibiting frequent mutations in our exome study were in DNA repair pathways. From the exome sequencing 4/12 (33%) metastatic tumors had missense mutations in *ATR*. Three metastatic tumors (25%) had mutations in *FANCD2* including one with a truncating mutation. Other DNA repair genes showing mutation frequencies over 10% include *BLM* (17%) and *ATM* (25%). Additional cancer-related genes found to be mutated in the metastatic samples included *PIK3CA*, *PIK3CG*, *FAT1*, *KIT* and *ERRB4* (Figure 1; Table 2; Supplementary Table 3). Finally, we observed high mutation frequency in genes previously reported to be mutated in metastatic cSCC including *NFI* (42%), *SMARCA4* (33%), and *RIPK2* (25%), (Supplementary Table 3).¹²

Differences in mutation frequency between paired primary and metastatic samples

We were curious if any of the mutations for which we had paired primary and metastatic samples were specific to the metastatic tumor. For the six primary/metastatic sample pairs, many of the mutations observed in the metastatic sample were present in the primary tumor and vice versa, but none were 100% concordant (Figure 1) (Supplementary Tables 3, 4). Two primary tumors (289199T and 387481T) for which we also sequenced a metastatic sample from the same individual did not have any mutations in common with the metastatic

tumor, suggesting that they were not the cSCC giving rise to the metastasis (Supplementary Tables 3, 4). For commonly mutated genes there was no pattern of specific genes that were predominantly mutated in one type of sample versus the other, although the numbers of primary-metastatic pairs are too small to make meaningful comparisons.

Targeted sequencing

No other studies have yet performed exome or targeted sequencing of primary non-metastatic and metastatic cSCCs simultaneously. To determine if there were differences in mutation frequency between primary non-metastatic cSCCs and metastatic tumors, we performed targeted exon sequencing using the Ion AmpliSeq Comprehensive Cancer Panel which includes 409 genes. Targeted sequencing was carried out rather than additional exome sequencing due to cost constraints and because this panel included the majority of the frequently mutated genes from our exome studies except for *FAT1* so we would be able to compare results.

Study inclusion for the targeted sequencing study was similar to our exome study. We also included one metastatic sample, 351152, for which we also performed exome sequencing. For the targeted sequencing study, more of the metastatic tumors originated from cSCCs on the face (83%) compared to the non-metastatic primary tumors which arose on the face (40%) or limb (50%)(Table 1). There was a trend towards younger age of diagnosis in individuals with primary tumors (mean 62.5 years) versus individuals with metastatic tumors (mean 74.5 years).

For the targeted sequencing, the mean depth was 274 reads with an average of 97.91% on target reads and a uniformity of 79% (Supplementary Table 5). Gene mutation frequencies in the metastatic samples in the targeted sequencing study were similar to those in the exome study, suggesting that these two approaches yielded similar gene coverage (Table 2; Supplementary Table 6). Comparison of mutation frequency between metastatic samples from both studies (n=18) and primary tumors (n=10) showed similar mutation frequencies for genes except *TP53* which only had mutations in two of the ten primary tumors (p-value=0.02); this is much lower than previous reports for primary cSCC. This is unlikely to be due to issues with the capture technology as all six of the six metastatic samples in the targeted study had *TP53* mutations, but may be due to lower read counts for some exons. For the one individual, 351152, for which we performed exome sequencing as well as targeted sequencing for the metastatic sample, there was a high degree of concordance between exome and targeted sequencing data. One *EP300* mutation was present at a low frequency (6% of reads) in the metastatic sample and paired primary tumor (1%) on exome analyses but was not seen in the targeted sequencing study. Of 71 missense, nonsense and/or frameshift mutations identified by the targeted sequencing in the sample from individual 351152, 67 were also detected by whole exome analysis (94.3%). Four additional mutations were observed in the metastatic sample by both targeted and exome sequencing, but not in any reads of the primary tumor, suggesting that these were either present at very low levels in the primary tumor, were present in another part of the primary tumor not sampled, or were unique to the metastatic sample.

Comparison of mutation frequencies across exome and targeted sequencing studies

As this study is relatively small considering the number of samples assessed, we wanted to determine how our results compared to previous exome and targeted sequencing studies. Mutation summaries for published exome- and targeted-sequencing studies of metastatic cSCC, primary non-metastatic cSCC and sun-exposed histologically normal skin were evaluated for mutation frequency in genes found to be frequently mutated in our samples (Table 2).^{7–12,21} Samples were considered to have a mutation if, relative to a matched normal DNA, the tumor had a nonsense, frameshift, missense, or splice-site mutation. We discarded non-synonymous changes that were present in dbSNP at a frequency over 1% unless they were reported in COSMIC or other somatic mutation databases. We used the authors' cutoff values for inclusion as a mutation in their analysis which may differ from our criteria. There was a range of reported mutation frequencies across the studies for some genes. For example, the frequency of samples with *TP53* varied between studies from 20% in primary cSCCs from our study to 95% of metastatic cSCCs in one series.¹¹ Differences between studies may be a reflection of sample numbers, read depths, sequence analysis pipelines, or coverage of region. Overall, the mutation frequency of genes assessed in our study was similar to previous reports.

We calculated the frequency of mutations found in each gene from the five published studies plus data from our study. The exact number of samples per gene varied widely depending on whether the gene was included on each study's targeted sequencing panel and whether whole exome analyses was completed. Mutation data was available from 42 to 233 primary cSCCs and 79 to 86 metastatic cSCCs depending on the gene. Multiple genes showed significant mutation frequency differences between primary cSCCs and metastatic cSCCs after multiple comparisons testing (Table 3). Interestingly, *TP53* showed a significantly higher mutation frequency in metastatic tumors (85%) compared to primary tumors (54%; p-value <0.0001; corrected p-value <0.002). This result remained significant (p<0.0001) even when data from our primary tumors were removed (since they showed a much lower mutation frequency than reported in the literature) or when any other study was removed from analysis. *CDKN2A* also showed higher mutation frequency (42%) in metastatic samples relative to primary tumors (29%; p-value=0.04), but this was not significant after accounting for multiple comparisons testing, nor did this result remain significant when other studies were removed. *KMT2D*, a gene involved in chromatin remodeling showed a higher mutation rate in metastatic tumors with 62% of metastatic tumors having at least one mutation compared to only 31% in primary tumors (p-values 0.001; corrected p-value 0.03). Other genes showing statistically significant differences in mutation frequency included *PIK3CG* (7% of primary tumors versus 29% of metastatic tumors; p-value <0.0001; corrected p-value <0.002). However, this difference was totally dependent on the study by Lee et al (2014)¹⁰ in which exome or targeted-sequencing was performed on 104 tumors, but none had a *PIK3CG* mutation. *BRAF* also showed borderline differences (5% of primary tumors versus 13% of metastatic tumors; p-value=0.04). There were more *FAT1* mutations in primary tumors (37%) relative to metastatic tumors (22%), but this was not significant after adjustment for multiple comparisons or when other studies were removed from analysis. DNA repair gene *ATR* showed a trend towards a difference, being mutated in six of 42 (14%) primary tumors versus 25 of 86 (29%) metastatic cSCCs (p-value =0.08).

DISCUSSION

Here, we present data from whole exome- and/or targeted-sequencing for 18 metastatic cSCCs and ten primary non-metastatic cSCCs. Findings from this study are consistent with previous exome analyses of SCCs; *TP53*, NOTCH pathway genes, *CDKN2A*, PI3K/AKT pathway genes, and epigenetic regulators were mutated at high rates in both non-metastatic primary and metastatic samples.

Primary tumors giving rise to metastases showed a high concordance of gene mutations with the metastatic sample suggesting that many mutations are occurring prior to metastasis. The concept of mutations occurring early in skin tumor development is consistent when comparing primary tumor results to histologically normal sun-exposed skin; genes showing frequent mutations in the primary tumors, including *TP53*, *NOTCH1* epigenetic regulators and DNA repair genes have been found previously in a substantial proportion of histologically normal sun-exposed skin samples.^{9,21}

We identified multiple genes with differential mutation frequencies between metastatic and non-metastatic primary tumors. Of note, *TP53*, a highly characterized gene in cSCC, showed a statistically higher rate of mutation in metastatic tumors (85%) relative to primary non-metastatic cSCCs (54%). Further studies are needed to understand the implications of this finding.

Genes important in chromatin remodeling and histone modification showed high rates of mutations across all samples but only *KMT2D* showed significantly higher rates of mutation in the metastatic tumors (62%) relative to non-metastatic tumors (31%). *KMT2D* is a histone methyltransferase that regulates H3 lysine 4 and is mutated other cancers such as cutaneous melanoma (19–29%), esophageal SCC (14–19%), and head and neck SCC (11–16%).^{19,20,22–26} *KMT2D* mutated cells *in vitro* show increased transcriptional stress and genomic instability.²⁴ Other epigenetic regulators *EP300*, *SETD2*, *KMT2C*, *KMT2D*, *KDM6A* and *KDM6B*, are mutated in a variety of lung, esophageal and head and neck SCCs suggesting that common epigenetic pathways drive squamous cell carcinogenesis, and these may also correlate with aggressiveness.^{13,22,28–30} A previous study of 39 metastatic cSCCs identified *KMT2C* mutations in 69% of tumors which were highly correlated with bone metastasis.¹¹ Sixty-seven percent of our metastatic samples had mutations in this gene, but there were no significant differences in *KMT2C* mutation frequency between non-metastatic primary and metastatic tumors across the published studies. *CREBBP* is a transcriptional co-activator, with a high degree of similarity to p300, which couples chromatin remodeling to transcription factor recognition and has histone acetyltransferase activity.³¹ In previous studies, 35–36% of primary cSCCs had *CREBBP* missense mutations, but it wasn't clear if these were functional.^{9,32} We observed *CREBBP* mutations in 28% of our metastatic samples and 40% of primary tumors (Table 2) but observed no significant differences between metastatic and non-metastatic cSCCs from published studies (Table 3), suggesting similar mutation frequencies in metastatic and non-metastatic tumors. Collectively, these studies suggest that mutations in some epigenetic and chromatin regulators may be associated with metastatic cSCC and that, generally, epigenetic pathways are a common mutational target in cSCC.

KNSTRN was previously found to be mutated in 19% of cSCCs.¹⁰ Surprisingly, we did not observe mutations in *KNSTRN*. To ensure that we were not missing any mutations in our exome analyses, we specifically looked at our genomic coverage in this region and BAM file sequences. We had an average read depth of 30–50X for exons of this gene, including the “hotspot” mutation, but did not identify any *KNSTRN* mutations. This could be due to different mutation profiles between non-metastatic and metastatic cSCCs or due to the smaller number of samples in our study. Our data is consistent with other studies of metastatic cSCCs that did not identify any mutations in *KNSTRN*.¹²

One difference of the current study with previous reports is the inclusion of six matched primary and normal DNA samples for the analyses and the high level of mutation concordance between these related tumors. The high concordance rate between the paired primary and metastatic samples suggests that the exome sequencing data had high sensitivity for detecting mutations. The power of this study and our findings was increased by incorporating data from five additional studies. For many of the genes, the mutation rate in our samples was similar to previous studies. Interestingly, data from this study is similar to cSCC mutation data in mice. Two studies have recently performed whole exome sequencing on primary and metastatic tumors in mice.^{33,34} Although some of the murine tumors were induced by chemicals instead of UV-light many of the same genes mutated in our samples were shown to be mutated in the metastatic mouse cSCCs including *Trp53* (30%), *Notch1* (30%), *Kmt2c* (26%), *Kmt2d* (22%), and *Fat1* (19%).³³

There are limitations to this study, which include the small number of samples sequenced in our series. Thus, we may be missing mutations that are meaningful in a small subset of samples or in genes that are underrepresented. Additionally, this study did not assess for copy number changes. As *CDKN2A* shows a high degree of copy number loss in cSCCs, this may explain the relatively low mutation rate in this gene in our study.¹² Another limitation is that histologically normal adjacent tissue was used for ten of the 18 metastatic cases as a source of normal DNA. Studies have shown a high-degree of mutation in histologically normal sun-exposed skin.^{9,21} Thus, somatic mutations present in greater than 9% of reads in both the adjacent paired normal and tumor would not have been identified due to our variant filtering process. It is well established that tumors are highly heterogeneous; as we only sampled a portion of the tumor we may be missing mutations unique to other parts of the tumor. Only one sample in our series was sequenced by both exome and targeted sequencing; thus, it is difficult to directly compare sensitivity and specificity of the two approaches. However, for that sample there was a high degree of concordance (94%) for the targeted sequencing relative to the exome data. Finally, there are limitations to aggregating data from multiple independent studies as there are differences in variant calling pipelines including number of reads required to “call” a mutation and read depths. This was evident by the statistically significant differences in *PIK3CG* mutation rates being driven by one study. Further studies are important to confirm the findings seen here as well, particularly the genes showing borderline differences in mutation frequency. Studies are also warranted to determine the functional role of mutations in these genes in metastasis.

In summary, exome and targeted sequencing of 18 metastatic cSCCs, 10 primary tumors and a pooled analysis of other studies consisting of up to 223 cSCCs and 68 metastatic cSCCs revealed frequent mutations in genes important in chromatin remodeling, DNA repair and the PI3K/AKT pathway. Analyses showed that *KMT2D* and *TP53* had higher mutation frequencies in the metastatic cSCCs relative to non-metastatic cSCC. Future studies dissecting the biological relevance of these mutations on aggressive phenotypes identify biomarkers that predict metastasis and potentially lead to new therapeutic targets.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The Cooperative Human Tissue Network, OSU Tissue Archives Service, Department of Otolaryngology-Head and Neck Surgery, David Lambert, and Ilene Latimer aided in sample ascertainment. The Human Genetics Sample Bank processed blood DNAs, the OSU CCC Solid Tumor group provided histology support. Pearly Yan of the OSU CCC Genomics Shared Resource performed the exome sequencing. Cynthia Timmers and the OSU CCC Solid Tumor Shared Resource were instrumental in generating libraries for the targeted sequencing. The OSU Pharmacogenomics Core Lab conducted the targeted sequencing.

Funding Sources: This work was supported in part by a Pelotonia Idea award, the National Institutes of Health (1R01 CA134461-01, 5R03 CA173788, P30 CA0160858) and the American Cancer Society (RSG-07-083-1-MGO). The targeted sequencing was supported by a generous gift from Dan and Margy Smith.

References

1. Karia PS, Han J, Schmults CD. Cutaneous squamous cell carcinoma: estimated incidence of disease, nodal metastasis, and deaths from disease in the United States, 2012. *J Am Acad Dermatol.* 2013; 68:957–966. [PubMed: 23375456]
2. Rowe DE, Carroll RJ, Day CL. 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; 26:976–990. [PubMed: 1607418]
3. Moore BA, Weber RS, Prieto V, et al. Lymph node metastases from cutaneous squamous cell carcinoma of the head and neck. *Laryngoscope.* 2005; 115:1561–1567. [PubMed: 16148695]
4. Brantsch KD, Meisner C, Schönlisch B, et al. Analysis of risk factors determining prognosis of cutaneous squamous cell carcinoma: a prospective study. *Lancet Oncol.* 2008; 9:713–720. [PubMed: 18617440]
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; 149:541–547. [PubMed: 23677079]
6. Brash DE. Sunlight and the onset of skin cancer. *Trends Genet.* 1997; 13:410–414. [PubMed: 9351343]
7. Durinck S, Ho C, Wang NJ, et al. Temporal dissection of tumorigenesis in primary cancers. *Cancer Discov.* 2011; 1:137–143. [PubMed: 21984974]
8. Wang NJ, Sanborn Z, Arnett KL, et al. Loss-of-function mutation in Notch receptors in cutaneous and lung squamous cell carcinoma. *Proc Natl Acad Sci U S A.* 2011; 108:17761–17766. [PubMed: 22006338]
9. South AP, Purdie KJ, Watt SA, et al. NOTCH1 mutations occur early during cutaneous squamous cell carcinogenesis. *J Invest Dermatol.* 2014; 134:2630–2638. [PubMed: 24662767]
10. Lee CS, Bhaduri A, Mah A, et al. Recurrent point mutations in the kinetochore gene *KNSTRN* in cutaneous squamous cell carcinoma. *Nat Genet.* 2014; 46:1060–1062. [PubMed: 25194279]
11. Pickering CR, Zhou JH, Lee JJ, et al. Mutational landscape of aggressive cutaneous squamous cell carcinoma. *Clin Cancer Res.* 2014; 20:1–11.

12. Li YY, Hanna GJ, Laga AC, Haddad RI, Lorch JH, Hammerman PS. Genomic analysis of metastatic cutaneous squamous cell carcinoma. *Clin Cancer Res.* 2015; 21:1–10.
13. Schwaederle M, Elkin SK, Tomson BN, Carter JL, Kurzrock R. Squamousness: next-generation sequencing reveals shared molecular features across squamous tumor types. *Cell Cycle.* 2015; 14:2355–2361. [PubMed: 26030731]
14. Dworkin AM, Tseng SY, Allain DC, Iwenofu OH, Peters SB, Toland AE. Merkel cell polyomavirus in cutaneous squamous cell carcinoma of immunocompetent individuals. *J Invest Dermatol.* 2009; 129:2868–2674. [PubMed: 19554019]
15. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 2011; 17:10–12.
16. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009; 25:1754–1760. [PubMed: 19451168]
17. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010; 20:1297–1303. [PubMed: 20644199]
18. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnology.* 2013; 31:213–219.
19. Hodis E, Watson IR, Kryukov GV, et al. A landscape of driver mutations in melanoma. *Cell.* 2012; 150:251–263. [PubMed: 22817889]
20. Cancer Genome Atlas Network. Genomic classification of cutaneous melanoma. *Cell.* 2015; 161:1681–1696. [PubMed: 26091043]
21. Martincorena I, Roshan A, Gerstung M, et al. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science.* 2015; 348:880–886. [PubMed: 25999502]
22. Gao YB, Chen ZL, Li JG, et al. Genetic landscape of esophageal squamous cell carcinoma. *Nat Genet.* 2014; 46:1097–1102. [PubMed: 25151357]
23. Lin DC, Hao JJ, Nagata Y, et al. Genomic and molecular characterization of esophageal squamous cell carcinoma. *Nat Genet.* 2014; 46:467–473. [PubMed: 24686850]
24. Lee JJ, Sholl LM, Lindeman NI, et al. Targeted next-generation sequencing reveals high frequency of mutations in epigenetic regulators across treatment-naïve patient melanomas. *Clin Epigenet.* 2015; 7:59.
25. Stransky N, Egloff AM, Tward AD, et al. The mutational landscape of head and neck squamous cell carcinoma. *Science.* 2011; 333:1157–1160. [PubMed: 21798893]
26. Chung CH, Guthrie VB, Masica DLJ, et al. Genomic alterations in head and neck squamous cell carcinoma determined by cancer gene-targeted sequencing. *Ann Oncol.* 2015; 26:1216–1223. [PubMed: 25712460]
27. Kantidakis T, Saponaro M, Mitter R, et al. Mutation of cancer driver MLL2 results in transcriptional stress and genome instability. *Genes Dev.* 2016; 30:408–420. [PubMed: 26883360]
28. Kim Y, Hammerman PS, Kim J, et al. Integrative and comparative genomic analysis of lung squamous cell carcinomas in East Asian patients. *J Clin Oncol.* 2014; 32:121–128. [PubMed: 24323028]
29. Martin D, Abba MC, Molinolo AA, et al. The head and neck cancer cell oncogenome: A platform for the development of precision molecular therapies. *Oncotarget.* 2014; 4:8906–8923.
30. Song Y, Li L, Ou Y, et al. Identification of genomic alterations in oesophageal squamous cell carcinoma. *Nature.* 2014; 509:91–95. [PubMed: 24670651]
31. Vo N, Goodman RH. CREB-binding protein and p300 in transcriptional regulation. *J Biol Chem.* 2001; 276:13505–13508. [PubMed: 11279224]
32. Watt SA, Purdie KJ, den Breems NY, et al. CREBBP mutation in human cutaneous squamous cell carcinoma. *Exp Dermatol.* 2016; 25:650–651. [PubMed: 27094574]
33. Nassar D, Latil M, Boeckx B, Lambrechts D, Blanpain C. Genomic landscape of carcinogen-induced and genetically-induced mouse skin squamous cell carcinoma. *Nat Med.* 2015; 21:946–954. [PubMed: 26168291]

34. McCreery MQ, Halliwill KD, Chin D, et al. Evolution of metastasis revealed by mutational landscapes of chemically induced skin cancers. *Nat Med.* 2015; 21:1514–1520. [PubMed: 26523969]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 1. Mutation frequency in metastatic cSCC samples
A. The percentage of metastatic samples that were found to have a mutation by exome sequencing of 26 genes previously shown to be mutated in primary or metastatic cSCCs are indicated. Sample 351152 is not included in the frequency as its mutations were included in the targeted sequencing analyses. **B.** Mutations in 26 genes for 12 metastatic samples and paired tumors are shown. As indicated in the key, mutations shared in both the metastatic sample and the primary tumor giving rise to the metastasis are indicated in purple, mutations only seen in the metastatic sample are indicated in red, mutations only observed in the paired primary tumor are indicated in blue. Mutations seen in samples for which a paired tumor was not available are indicated in gray. Truncating mutations are denoted by an “*”.

Fig. 2. Targeted-sequencing of metastatic and primary cSCC samples

A. The percentage of seven metastatic samples found to have a mutation by targeted sequencing analysis of 21 commonly mutated genes is indicated. (This includes 351152 which was exome sequenced). Genes are ranked by the percentage of samples with a potentially functional mutation (nonsense, frameshift, splice-site or missense). As indicated in the key, samples with one or more missense mutations are indicated in light red, and samples with at least one truncating (nonsense or frameshift) mutation are indicated in dark red. **B.** The percentage of ten primary cSCCs found to have a mutation in one of 21 commonly mutated genes by targeted sequencing is indicated. Genes are ranked by the percentage of samples with a potentially functional mutation (nonsense, frameshift, splice-site or missense). As denoted in the key, samples with one or more missense mutations are indicated in light blue, and samples with at least one truncating (nonsense or frameshift) mutation are indicated in dark blue.

Table 1

Characteristics of Study Participants and Tumors

Characteristic	Descriptor	Exome Metastatic n of 121 (%)	Targeted Metastatic n of 6 (%)	Targeted Primary SCC n of 10 (%)
Sex	Male	7 (58%)	5 (83%)	7 (70%)
	Female	5 (42%)	1 (17%)	3 (33%)
Mean age at diagnosis (range)		72.3 years (50–89)	74.5 years (65–87)	62.5 years (40–84)
Primary cSCC Site	Face/Ear/Neck	9 (75%)	5 (83%)	4 (40%)
	Limb	2 (17%)	1 (17%)	5 (50%)
	Trunk	0 (0%)	0 (0%)	1 (10%)
	Other/Not Specified	1 (8%)	0 (0%)	0 (0%)
Differentiation status	Well	1 (8%)	2 (29%)	3 (30%)
	Moderately	5 (42%)	1 (17%)	1 (10%)
	Poorly	4 (33%)	2 (33%)	1 (10%)
	Not noted	2 (17%)	0 (0%)	5 (50%)
Invasive Status of primary	Not noted	3 (25%)	0 (0%)	2 (20%)
	In situ/Bowen's	0 (0%)	0 (0%)	3 (30%)
	Invasive	7 (58%)	5 (83%)	5 (50%)
	Metastatic	2 (17%)	1 (17%)	0 (0%)
Mean age at metastasis (range)		72.6 years (51–89)	74.5 years (65–87)	63.5 years (58 & 69)
Metastasis Site ^I	Lymph Node	6 (50%)	1 (17%)	2 (20%)
	Parotid Gland	0 (0%)	3 (50%)	N/A
	Tonsil	1 (8%)	0 (0%)	N/A
	Face	1 (8%)	0 (0%)	N/A
	Muscle/Soft Tissue	2 (17%)	0 (0%)	N/A
	Bone	1 (8%)	0 (0%)	N/A
	Lip	1 (8%)	0 (0%)	N/A
	Not specified	2 (17%)	2 (33%)	N/A

n, number; %, percentage; Dx, diagnosis; NA, not applicable.

^I cSCC that metastasized to multiple sites is counted more than once so the number is greater than number of tumors

Table 2

Somatic mutation frequencies across exome and targeted deep-sequencing studies

Approach/Study Reference	Targeted ²¹	Exome ⁷	Exome/Targeted ¹⁰	Exome/Targeted ⁹	Targeted/Current Study	Targeted ¹²	Exome ¹¹	Exome/Targeted/Current Study
Sample type (n)	Sun exposed skin (n=234) ¹	Primary (n=8)	Primary (n=4/100)	Primary (n=20/91)	Primary (n=10)	Metastatic (n=29)	Metastatic (n=39)	Metastatic (n=12/6) ³
Gene								
<i>TP53</i>	58 (25%)	6 (75%)	47 (42%)	13 (65%)/58 (64%)	2 (20%)	23 (79%)	37 (95%)	7 (58%)/6 (100%)
<i>CDKN2A</i>	3 (1%)	2 (25%)	31 (28%)	9 (45%)/25 (28%)	1 (10%)	14(48%) ²	17 (44%)	2 (17%)/3 (50%)
<i>NOTCH1</i>	186 (79%)	6 (75%)	60 (54%)	8 (40%)/68 (75%)	5 (50%)	14 (48%)	23 (59%)	3 (25%)/3 (50%)
<i>NOTCH2</i>	134 (57%)	4 (50%)	38 (34%)	10 (50%)/57 (63%)	3 (30%)	9 (31%)	20 (51%)	5 (42%)/1 (17%)
<i>NOTCH3</i>	50 (21%)	1 (13%)	24 (21%)	5 (25%)/ND	0 (0%)	5 (17%)	13 (33%)	4 (33%)/0 (0%)
<i>NOTCH4</i>	21 (9%)	3 (37.5)	0/ND	3 (15%)/ND	1 (10%)	5 (17%)	9 (23%)	2 (17%)/1 (17%)
<i>HRAS</i>	10 (4%)	1 (13%)	7 (6%)	4 (20%)/15 (16%)	0 (0%)	1 (3%)	8 (21%)	1 (8%)/1 (17%)
<i>KMT2A</i>	28 (12%)	2 (25%)	0/ND	3 (15%)/ND	2 (20%)	4 (14%)	9 (23%)	4 (33%)/1 (17%)
<i>KMT2C</i>	36 (15%)	4 (50%)	46 (44%)	8 (40%)/ND	3 (30%)	7 (24%)	15 (39%)	7 (58%)/2 (33%)
<i>KMT2D</i>	64 (27%)	0 (0%)	0/ND	8 (40%)/ND	5 (50%)	14 (48%)	27 (69%)	8 (66%)/4 (67%)
<i>EP300</i>	ND	2 (25%)	1 (25%)/ND	5 (25%)/ND	2 (20%)	10 (34%)	10 (26%)	3 (25%)/2 (33%)
<i>EZH2</i>	9 (4%)	1 (13%)	1 (25%)/ND	3 (15%)/ND	0 (0%)	2 (7%)	2 (5%)	1 (8%)/0 (0%)
<i>TET2</i>	ND	0 (0%)	1 (25%)/ND	4 (20%)/ND	4 (40%)	4 (14%)	9 (23%)	2 (17%)/2 (33%)
<i>ARID2</i>	27 (12%)	2 (25%)	0/ND	4 (20%)/ND	1 (10%)	8 (28%)	7 (18%)	2 (17%)/1 (17%)
<i>CREBBP</i>	21 (9%)	1 (13%)	0/ND	7 (35%)/33 (36%)	4 (40%)	8 (28%)	13 (33%)	2 (17%)/3 (50%)
<i>PIK3CA</i>	12 (5%)	1 (13%)	0/ND	1 (5%)/ND	1 (10%)	4 (14%)	5 (13%)	2 (17%)/1 (17%)
<i>PIK3CG</i>	ND	3 (38%)	0/0%	5 (25%)/ND	2 (20%)	8 (28%)	13 (33%)	3 (25%)/1 (17%)
<i>BRAF</i>	6 (3%)	0 (0%)	5 (5%)	0 (0%)/ND	2 (20%)	3 (10%)	7 (18%)	1 (8%)/0 (0%)
<i>ATR</i>	ND	1 (13%)	1 (25%)/ND	2 (10%)/ND	2 (20%)	9 (31%)	11 (28%)	4 (33%)/1 (17%)
<i>ATM</i>	ND	3 (38%)	0/ND	2 (10%)/ND	3 (30%)	10 (34%)	5 (13%)	3 (25%)/1 (17%)
<i>BLM</i>	ND	1 (13%)	0/ND	1 (5%)/ND	3 (30%)	1 (3%)	6 (15%)	2 (17%)/1 (17%)
<i>FAT1</i>	89 (38%)	0 (0%)	41 (37%)	8 (40%)/ND	ND	8 (28%)	4 (10%)	5 (45%)/ND
<i>FANCD2</i>	ND	1 (13%)	0/ND	2 (10%)/ND	1 (10%)	3 (10%)	8 (21%)	3 (25%)/2 (33%)

ND (no reported data/not included on targeted panel)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

¹ 234 biopsies from 4 individuals; some overlap in mutations across adjacent biopsies;
² includes copy number aberrations;
³ the individual sequenced on both exome and targeted sequencing is counted once for the exome sequencing;
⁴ this result was not significant (p-value=0.8) after removing the Lee et al study (10) from analysis.

Table 3

Mutation frequencies in primary versus metastatic tumors

Gene	Primary Tumors (up to n=233) ¹	Metastatic Tumors (up to n=86)	p-value
<i>TP53</i>	126/233 (54%)	73/86 (85%)	<0.0001 ²
<i>CDKN2A</i>	68/233 (29%)	36/86 (42%)	0.04
<i>NOTCH1</i>	147/233 (63%)	43/86 (50%)	0.04
<i>NOTCH2</i>	112/233 (48%)	35/86 (41%)	0.26
<i>NOTCH3</i>	30/142 (21%)	22/86 (26%)	0.51
<i>NOTCH4</i>	7/42 (17%)	17/86 (20%)	0.81
<i>HRAS</i>	27/233 (12%)	11/86 (13%)	0.85
<i>KMT2A</i>	7/42 (17%)	18/86 (21%)	0.64
<i>KMT2C</i>	61/142 (43%)	31/86 (36%)	0.33
<i>KMT2D</i>	13/42 (31%)	53/86 (62%)	0.001 ²
<i>EP300</i>	10/42 (24%)	25/86 (30%)	0.67
<i>EZH2</i>	5/42 (4%)	5/86 (6%)	0.29
<i>TET2</i>	9/42 (21%)	17/86 (20%)	0.81
<i>ARID2</i>	7/42 (17%)	18/86 (21%)	0.64
<i>CREBBP</i>	45/133 (34%)	26/86 (30%)	0.66
<i>PIK3CA</i>	3/42 (7%)	12/86 (14%)	0.38
<i>PIK3CG</i>	10/142 (7%)	25/86 (29%)	<0.0001 ^{2, 3}
<i>BRAF</i>	7/142 (5%)	11/86 (13%)	0.04
<i>ATR</i>	6/42 (14%)	25/86 (29%)	0.08
<i>ATM</i>	8/42 (19%)	19/86 (22%)	0.65
<i>FAT1</i>	49/132 (37%)	17/79 (22%)	0.02
<i>FANCD2</i>	4/42 (10%)	16/86 (19%)	0.21

¹ Targeted studies do not include every gene so denominator varies;² significant after multiple comparisons testing;³ not significant after take one out analyses.