Systematic Screening of Promoter Regions Pinpoints Functional Cis-Regulatory Mutations in a Cutaneous Melanoma Genome

Rebecca C. Poulos1,2, Julie A.I. Thoms1,2, Anushi Shah1,2, Dominik Beck1,2, John E. Pimanda1,2,3, and Jason W.H. Wong1,2

Abstract

With the recent discovery of recurrent mutations in the TERT promoter in melanoma, identification of other somatic causal promoter mutations is of considerable interest. Yet, the impact of sequence variation on the regulatory potential of gene promoters has not been systematically evaluated. This study assesses the impact of promoter mutations on promoter activity in the whole-genome sequenced malignant melanoma cell line COLO-829. Combining somatic mutation calls from COLO-829 with genome-wide chromatin accessibility and histone modification data revealed mutations within promoter elements. Interestingly, a high number of potential promoter mutations (n = 23) were found, a result mirrored in subsequent analysis of TCGA whole-melanoma genomes. The impact of wild-type and mutant promoter sequences were evaluated by subcloning into luciferase reporter vectors and testing their transcriptional activity in COLO-829 cells. Of the 23 promoter regions tested, four mutations significantly altered reporter activity relative to wild-type sequences. These data were then subjected to multiple computational algorithms that score the cis-regulatory altering potential of mutations. These analyses identified one mutation, located within the promoter region of NDUF9, which encodes the mitochondrial NADH dehydrogenase (ubiquinone) 1 beta subcomplex 9, to be recurrent in 4.4% (19 of 432) of TCGA whole-melanoma exomes. The mutation is predicted to disrupt a highly conserved SP1/KLF transcription factor binding motif and its frequent co-occurrence with mutations in the coding sequence of NF1 supports a pathologic role for this mutation in melanoma. Taken together, these data show the relatively high prevalence of promoter mutations in the COLO-829 melanoma genome, and indicate that a proportion of these significantly alter the regulatory potential of gene promoters.

Implications: Genomic-based screening within gene promoter regions suggests that functional cis-regulatory mutations may be common in melanoma genomes, highlighting the need to examine their role in tumorigenesis. Mol Cancer Res; 13(8); 1218–26. ©2015 AACR.

Introduction

Cancers develop when certain somatic mutations are acquired by individual cells (1). However, most tumors harbor a number of different genetic and epigenetic aberrations (2, 3). This means that identifying all driver mutations, and distinguishing them from passenger mutations, remains a major challenge (4, 5). This problem is exacerbated in the noncoding genome, as a lack of selective pressure in these regions contributes to the acquisition of higher numbers of mutations. In the past, research into cancer driver mutations has typically focused on protein-coding mutations, as the functional consequence of noncoding mutations is difficult to determine. Furthermore, availability of datasets required for analysis of noncoding mutations has been limited.

However, recurrent mutations in the TERT promoter were recently identified in melanoma and other cancers (6, 7). The TERT promoter mutations were the first recurrent cis-regulatory somatic point mutations identified in cancer that alter gene expression (6). These mutations drive cancer by generating a transcription factor binding motif for E-twenty-six (ETS) transcription factors, with corresponding increases in promoter activity and expression of TERT (6). Remarkably, mutations at the promoter of TERT are found in as many as 85% of metastatic melanomas and are also found at high frequency in many other cancers, including glioblastoma (62%) and bladder cancer (59%; ref. 7).

With advances in sequencing technology (8–11), whole cancer genomes are being sequenced at a rapid pace and hundreds of sequenced samples are available for analysis from The Cancer Genome Atlas (TCGA), the Wellcome Trust Sanger Institute Cancer Genome Project, and the International Cancer Genome Consortium (12). This has led to a number of studies that have focused on mapping the cis-regulatory mutational landscape across cancers (13, 14). These analyses have primarily used recurrence to determine those mutations with potential driver roles in cancer (13, 14). However, while numerous recurrent cis-regulatory mutations have been identified, thus far, only a single promoter mutation in addition to the TERT promoter mutations, in the promoter of the SDHD gene, has been linked to changes in gene expression (13). Nevertheless, from these studies, it is...
evident that mutations at promoters can be found in many cancers. It appears that promoter mutations may even be more prevalent in melanomas than other cancers. For example, one recent study analyzing mutations within 500 bp of a transcription start site (TSS) that were recurrent in five or more cancer samples, found that 78% of those mutations identified were from melanomas (14).

However, with relatively low numbers (<30) of whole-genome sequenced samples currently publicly available for individual tumor types, assessment of the functional impact of recurrent mutations, or indeed nonrecurrent mutations, remains challenging. A number of computational approaches are now available to functionally annotate and prioritize mutations with potential cis-regulatory function (15–18). However, to our knowledge, experimental analyses that systematically evaluate the functional consequence of potential cis-regulatory mutations have not yet been performed in any cancer genome. Therefore, the significance of all promoter mutations within any individual cancer genome remains unexplored.

In this study, we combine computational and experimental approaches to comprehensively survey the landscape of somatic promoter mutations in the genome of the malignant melanoma cell line COLO-829 (19). By using a whole-genome sequenced cell line, we have been able to adopt an unbiased approach and experimentally assess the functional consequence of promoter mutations in a cell-specific manner. We have identified that there are multiple mutations that perturb promoter activity in the COLO-829 cell line. In doing so, we have identified a mutation in the promoter of NDUFB9 that is recurrent in cutaneous melanomas, highlighting the need for further investigation of promoter mutations in melanoma development.

**Materials and Methods**

Further details of materials and methods are supplied in Supplementary Methods.

**Mutation calls**

The mutation calls used to identify putative somatic promoter mutations in the COLO-829 malignant melanoma cell line were obtained from the Catalogue of Somatic Mutations in Cancer (COSMIC) database (19–21). In addition, mutations were called in 34 whole-genome sequenced metastatic skin cutaneous melanoma tumor and matched normal samples from TCGA using the Strelka pipeline (22). For 432 whole-exome sequenced cutaneous melanoma samples, data were obtained from TCGA, and base calls were determined at specified sites using SAMtools mpileup (23).

**Identification of candidate COLO-829 promoter mutations**

DNase-seq data were obtained for COLO-829 melanoma (GEO: GSM1008571) and normal penis foreskin melanocyte primary cells (GEO: GSE18927) from the Encyclopedia of DNA Elements (ENCODE; ref. 24) and the Human Epigenome Atlas (25), respectively. In addition, normal penis foreskin melanocyte primary cell ChIP-seq data were obtained from the Human Epigenome Atlas (25) for the H3K4me3 histone modification (GEO: GSE16368). COLO-829 mutation coordinates were overlapped with COLO-829 DNase-seq and normal melanocyte H3K4me3 histone mark ChIP-seq regions as depicted in Fig. 1A. Putative promoter mutations were identified as those within these regions (26) and within ±1 kb of a TSS as determined by RefSeq gene annotations (27).

**Mutation analysis**

MutSigCV data (28) were plotted in Fig. 2, by the frequency of each given measure of expression, replication time, and noncoding mutation count. The relevant values corresponding to each of the COLO-829 promoter mutations were also plotted.

For each COLO-829 promoter mutation identified, conservation was determined using the PhastCons score (29) for the mutation itself, a 15 bp flanking region (mutation ± 7 bp) and the DNase I hypersensitive (DHS) peak in which the mutation lies. Transcription factor motifs created or removed by each mutation were identified by use of the OncoCis tool (15), which utilizes transcription factor motifs from the JASPAR database (30).

**Reporter assays**

Wild-type and mutant DHS regions flanking each COLO-829 putative promoter mutation were amplified by polymerase chain reaction (PCR) from COLO-829 melanoma and HCC1143 breast cancer genomic DNA (Supplementary Table S3; see Supplementary Table S6 for PCR primers). To produce reporter constructs, regions were cloned upstream of the firefly luciferase gene in the promoter-less vector pGL2 Basic (Promega Corporation). Any single-nucleotide polymorphisms (SNP) in reporter constructs are recorded in Supplementary Table S5.

Plasmids were transfected into COLO-829 cells. The cell line was obtained from the Peter MacCallum Cancer Centre (March 2014; ref. 31) prior to plasmid preparation. COLO-829 authentication was performed by validating the presence of COLO-829 mutations as reported in the COSMIC database (20, 21), against those identified from promoter amplification and sequencing. Reporter and normalization plasmids (pRL-TK; Promega Corporation) were cotransfected into COLO-829 cells using Lipofectamine 2000 (Life Technologies), with reporter gene activity measured after 48 hours. All constructs were tested in quadruplicate in at least two independent experiments. A mutant promoter region was determined to have different activity from wild-type in cases in which the mutant produced altered promoter activity in the same direction over three or more experiments, and was statistically significant at least once (P < 0.05, by unpaired t test).

**NDUFB9 promoter mutation and gene analysis**

RNA-sequencing data for metastatic melanomas were obtained from TCGA and plotted in Fig. 4C. Mutant and wild-type samples were identified by determining mutation calls from whole-exome sequencing data using SAMtools mpileup (23).

Co-occurrence (see Fig. 4D) was analyzed among samples with and without the NDUFB9 promoter mutation, using samples and mutation calls from TCGA. Samples were deemed mutated if they had at least one nonsilent protein-coding mutation in each gene from a list of commonly mutated melanoma genes used in previous research (32). Significant associations (P < 0.05) were calculated by the two-tailed Fisher exact test, using counts of mutations in each designated gene for NDFUB9 wild-type and mutant promoter groupings.
Results and Discussion

Identification of promoter mutations in the COLO-829 melanoma genome

Somatic mutation calls for the whole-genome sequenced COLO-829 cell line (19) were first obtained from the COSMIC database (20, 21). In total, there were 32,901 unique somatic mutation sites (Fig. 1A). We developed a method to identify mutations that lie in promoter regions. Peak calling in COLO-829 DNase-seq and normal melanocyte ChIP-seq data identified 146,280 DHS peaks and 26,275 H3K4me3 peaks (Fig. 1A). To identify promoters, we overlapped mutation calls with DHS and H3K4me3 peaks as shown in Fig. 1A, and determined which of those were within ±1 kB of a RefSeq annotated TSS (27). We chose to prioritize mutations within DHS peaks as these regions represent the most active regulatory regions. The use of H3K4me3 data further enabled us to define cell type–specific gene promoters. We found 26,035 COLO-829 DHS peaks to overlap a melanocyte H3K4me3 region (Fig. 1A), with 59% of those being within ±1 kB of a TSS (Supplementary Fig. S1A). Supporting our method of genome-wide promoter designation, genes with a promoter DHS peak in COLO-829 were found to be significantly more highly expressed than those without (P < 0.0001, by unpaired t test; Supplementary Fig. S1B). To further support our methodology, when using Fantom5 (33) TSS annotations rather than RefSeq (27) annotations for TSS, we found there to be a high degree of overlap in the mutations identified (Supplementary Fig. S1C).

Using COLO-829 somatic mutation calls, we found 31 mutations that were within a COLO-829 DHS and melanocyte H3K4me3 peak. We performed a bootstrap analysis to determine the likelihood of finding this number of putative promoter mutations in the COLO-829 cell line. This analysis (Supplementary Fig. S1D) revealed that the number of promoter mutations in COLO-829 (n = 31) was significantly lower than the median number of mutations (n = 41) that occurred at random in 1,000 iterations. This finding is consistent with previous research, which indicates that somatic mutation density is reduced in regulatory DNA due to the accessibility of DHS regions to global genome repair machinery (34).

We found that 23 of the 31 COLO-829 mutations were within ±1 kB of a TSS, and we therefore selected these for further analysis (Fig. 1A; Supplementary Table S1). Of these 23 mutations, five mapped to putative bidirectional promoters, providing a total of 28 candidate associated genes (Fig. 1A; Supplementary Table S1). A summary of somatic promoter mutations and their location in the COLO-829 genome is depicted in the plot at Fig. 1B.
Analysis of genomic context of COLO-829 promoter mutations

To determine whether the 23 promoter mutations occur in regions of the genome that are generally mutated more frequently, we compared the expression, replication time, and region mutation frequency of the COLO-829 promoter mutations against genomic rates generated by MutSigCV (28). Results indicate that the majority of COLO-829 promoter mutations lie within highly expressed regions with short replication times (Fig. 2A and B). The mutations were also within regions that are among the most common noncoding mutation rates in the genome (Fig. 2C). This suggests that the promoter mutations are not enriched within highly variable genomic regions, which are typically late replicating and contain lowly expressed genes (28).

Finally, in order to establish whether COLO-829 is similar to, and therefore broadly indicative of, metastatic melanoma samples, we analyzed mutations within 34 whole-genome sequenced melanoma datasets from TCGA. We found that the number of promoter mutations ranged from 1 to 548 mutations per sample (median 80; Supplementary Table S2). When normalized to the total number of mutations across the genome, we found the promoter mutation rate of COLO-829 was slightly lower than the median rate across all melanomas (0.0007 versus 0.0013 promoter mutations/total mutation). This suggests that while promoter mutations are generally even more frequent in melanoma samples compared with COLO-829, the findings from our study should still be broadly representative of melanoma samples. However, further research will need to firmly establish whether the frequency of functional promoter mutations found in COLO-829 is truly representative of melanoma genomes more generally.

Annotation and validation of promoter mutations in COLO-829

To eliminate false-positive mutation calls, we validated each putative promoter mutation by interrogating COLO-829 whole-genome as well as Sanger sequencing data from PCR amplified

**Figure 2.**
COLO-829 promoter mutations plotted against data points from MutSigCV analysis. Putative COLO-829 promoter mutations (gray bars) are plotted, along with the frequency of occurrence in the genome (per 100 kb region) of each given measurement for (A) expression level (B) DNA replication time and (C) noncoding mutation count from MutSigCV data (28). DNA replication time is expressed on a scale of 100 (early) to 1,500 (late).

**Figure 3.**
Validation of the functional consequence of COLO-829 promoter mutations by reporter assays. Representative results from one luciferase reporter assay experiment for each reporter construct tested. Results from wild-type (wt) and mutant (mt) constructs are adjacent, labeled according to the gene to which the promoter region is associated. Fold change is calculated against the average of replicate wt values. Promoter regions with no activity (luciferase activity <2 times that of the promoter-less vector pGL2 Basic) are indicated by a cross. Only statistically significant differences (by unpaired t test) are indicated; *, P < 0.05; **, P < 0.01; ***, P < 0.001.
Of 23 putative promoter mutations, we could validate 20, with these being within ±1 kB of 25 different genes (Table 1). We next performed a series of analyses to annotate each COLO-829 putative promoter mutation and infer their functional impact. Half of the 20 validated COLO-829 putative promoter mutations were deemed to be highly conserved, as they lie within a 15 bp region (mutation ±7 bp) with a greater average PhastCons conservation score (29) than the surrounding COLO-829 DHS region (Table 1). Furthermore, 19 of the 20 validated mutations alter transcription factor motifs (Table 1) based on mutation motif analysis using OncoCis (15). See Fig. 1B for a summary of the annotation of each promoter mutation.

Having identified 20 validated mutations with the potential to alter promoter activity, the functional consequence of each mutation was tested using reporter assays in COLO-829 cells. Wild-type and mutant promoter regions were isolated for 23 out of 25 genes (Supplementary Table S3). For potential bidirectional promoters, regions were cloned in both directions and promoter activity was determined.

**Figure 4.** Analysis of the NDUFB9 putative promoter mutation in COLO-829. A, location of the NDUFB9 putative promoter mutation (chr8:125,551,344 C>T) within the 5′-untranslated region (UTR) along with its position relative to COLO-829 DHS, melanocyte H3K4me3, and HCT116 SP1 ChIP-seq peaks. PhastCons conservation within the region and the SP1/Krüppel-like transcription factor (KLF) consensus binding motif are shown. B, PhastCons conservation score (29) for the NDUFB9 promoter mutation, 15 bp region (mutation ±7 bp) and surrounding DHS region. C, boxplot indicating no significant difference (n.s.) by unpaired t-test in NDUFB9 gene expression within wild-type (n = 302) and mutant (n = 16) samples available from TCGA. D, co-occurrence plot for the NDUFB9 promoter mutation and any nonsilent protein-coding mutations in commonly mutated melanoma-associated genes.
assessed independently. As shown in Fig. 3, we identified four gene candidates by reporter assays in which the mutant plasmid construct had significantly altered promoter activity from wild-type \( (P < 0.05, \text{by unpaired } t \text{ test}) \). These genes are *HIST1H3H*, *NDUFB9*, *PSMC6*, and *SPSB3*, with each harboring a single promoter mutation in COLO-829 (Table 1). When applying a Benjamini and Hochberg False Discovery Rate multiple testing correction, all mutations except for the *PSMC6* mutation remain significant.

We hypothesized that single-base mutations in highly conserved regions that altered binding motifs, would be the most likely to lead to a change in expression when compared to wild-type. However, these analyses did not accurately predict which mutations would have functional consequences as subsequently demonstrated by altered promoter activity in reporter assays. Of the nine mutations that altered transcription factor binding motifs and were deemed to be highly conserved (Table 1), only two altered promoter activity in reporter assays (those in the *NDUFB9* and *SPSB3* promoters). Two mutations, which altered promoter activity in reporter assays (those in the *HIST1H3H* and *PSMC6* promoters), were not highly conserved (Table 1) and thus had not been predicted to be functional. However, it was evident that across all three tools, many mutations found not to perturb promoter activity also had relatively high scores. This means that while the tools were generally quite sensitive in detecting functional promoter mutations, they are not particularly specific. It is important to note, however, that the mutations selected for analysis here are only those which were already found to fall within a promoter region of a gene. The majority of mutations, if not all, would fall within ENCODE DHS regions and transcription factor binding sites. Thus, most of the mutations would have a high background score relative to all mutations across a genome. Therefore, the specificity of the tools considered here is likely to be much higher when applied in genome-wide analysis of mutations.

It is vital to account for tissue specificity when identifying active promoter regions (15). In our analysis, we utilized matched-tumor COLO-829 DNase-seq data to identify putative promoter regions. Such a methodology prioritizes the identification of mutations in existing and novel DHS regions. However, our analysis is limited in its ability to identify mutations that remove a DHS site. To explore this, we conducted an analysis using melanocyte DNase-seq data in place of COLO-829 DNase-seq.

![Table 1. Annotation of putative promoter mutations in COLO-829 according to conservation, transcription factor motif alteration, and scores attributed by RegulomeDB, FATHMM-MKL, and FunSeq](mcr.aacrjournals.org/content/13/8/1223/F1.large.jpg)

<table>
<thead>
<tr>
<th>Mutation coordinate</th>
<th>Associated gene(s)</th>
<th>PhastCons conservation score</th>
<th>Motif created</th>
<th>Motif removed</th>
<th>Transcription factor motif alteration</th>
<th>Score(a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr1:109,825,879 G&gt;A</td>
<td>PSRCCT</td>
<td>0.18 0.42</td>
<td>PPARG:RXRA</td>
<td>ELK1:FEV:ELFS:...</td>
<td>2a 0.9518 2.497</td>
<td></td>
</tr>
<tr>
<td>chr1:118,472,354 G&gt;A</td>
<td>WDR5/G0AP2</td>
<td>0.64 0.34</td>
<td>MZF1_5-13</td>
<td>REL:ELFS:FEV:...</td>
<td>2a 0.6684 1.414</td>
<td></td>
</tr>
<tr>
<td>chr2:10,867,246 C&gt;T</td>
<td>RPE</td>
<td>0.52 0.28</td>
<td>REL:ELFS:FEV:...</td>
<td>2b 0.8269 1.851</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr3:107,941,296 G&gt;A</td>
<td>IFT57</td>
<td>0.98 0.38</td>
<td>ELF1</td>
<td>2b 0.8655 2.626</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr5:82,373,334 C</td>
<td>XRC4/TMEM167A</td>
<td>0.00 0.19</td>
<td>Amnt:SPIB</td>
<td>CREB1</td>
<td>2b 0.3141 2.741</td>
<td></td>
</tr>
<tr>
<td>chr6:27,777,830 A&gt;C</td>
<td>HIST1H3H</td>
<td>0.00 0.12</td>
<td>PPARG:RXRA:Hox9</td>
<td>ELK1:Stat6:Stat6:...</td>
<td>4 0.9504 2.703</td>
<td></td>
</tr>
<tr>
<td>chr6:31,940,123 G&gt;A</td>
<td>STR1B/DM1Z5</td>
<td>0.54 0.10</td>
<td>NAFATC2:NKFI1</td>
<td>ELF1</td>
<td>2a 0.7645 1.256</td>
<td></td>
</tr>
<tr>
<td>chr7:66,461,664 G&gt;A</td>
<td>TYW1</td>
<td>0.00 0.05</td>
<td>ELK1:Sp1</td>
<td>2a 0.5561 0.734</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr8:62,262,591 C&gt;T</td>
<td>GDAPI</td>
<td>0.00 0.16</td>
<td>STAT3</td>
<td>SPI:ELFS:SP2</td>
<td>4 1.825 1.926</td>
<td></td>
</tr>
<tr>
<td>chr8:125,531,344 C&gt;T</td>
<td>NDUF9B3:TAATDNI</td>
<td>0.58 0.35</td>
<td>Klf4:Klf5:Klf6:...</td>
<td>2a 0.7400 3.171</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr11:125,439,130 G&gt;C</td>
<td>E2F4</td>
<td>0.00 0.01</td>
<td>Sp1</td>
<td>4 0.4696 0.946</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr12:41,086,186 C&gt;T</td>
<td>CNTNI</td>
<td>0.00 0.28</td>
<td>NAFATC2:SPIB:ERG...</td>
<td>4 0.3491 1.557</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr14:53,173,816 A&gt;C</td>
<td>PSMC6</td>
<td>0.49 0.53</td>
<td>SPI:ELF1:FLI1:...</td>
<td>2a 0.9514 3.211</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr13:76,618,176 G&gt;A</td>
<td>GPATCHZL</td>
<td>0.54 0.54</td>
<td>THAPI</td>
<td>4 0.8413 1.423</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr13:78,730,433 C&gt;T</td>
<td>IREB2</td>
<td>0.43 0.40</td>
<td>Klf:THAPF1:...</td>
<td>2a 0.9865 3.447</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr16:18,532,640 G&gt;A</td>
<td>NDBP2:SPSB3</td>
<td>0.24 0.12</td>
<td>ELK1</td>
<td>2b 0.0873 1.819</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr19:2,643,275 G&gt;A</td>
<td>ZNF460</td>
<td>0.00 0.00</td>
<td>ELK1</td>
<td>3a 0.0575 2.752</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr20:36,156,522 G&gt;A</td>
<td>BLCAP</td>
<td>0.00 0.30</td>
<td>TFAP2C</td>
<td>4 0.2862 1.260</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr21:45,553,369 G&gt;A</td>
<td>C2orf53</td>
<td>0.00 0.17</td>
<td>NFAT2:FOXO3:Foxo1:...</td>
<td>2b 0.3151 0.800</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr22:30,988,178 G&gt;A</td>
<td>PST1</td>
<td>0.46 0.31</td>
<td>N2e3:N2e3</td>
<td>2b 0.2941 1.701</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Counts: 20 25 10/20 11/20 15/20

aMutations with changes in promoter activity from wild-type, as determined by reporter assays (see Fig. 3), are shaded in light gray.
bTranscription factor motif alterations were determined by using the OncoCis tool (15), which utilizes transcription factor motifs generated by the JASPAR database (30).
cScores were obtained for each mutation from RegulomeDB (17), FATHMM-MKL (18), and FunSeq2 (16).
data. Our results are summarized in Supplementary Fig. S3A, where we found that the majority of mutations, including the NDUFB9 putative promoter mutation, could be identified with both COLO-829 and melanocyte data (n = 17). However, n = 6 mutations were identified from COLO-829 DNase-seq only, and n = 4 mutations were identified from melanocyte DNase-seq only. To determine whether the four mutations present only from melanocyte DNase-seq could potentially have removed a COLO-829 DHS region and might warrant further analysis, we analyzed RNA-sequencing data from melanoma cell lines in the Cancer Cell Line Encyclopedia (CCLE; ref. 35). We found that all associated genes were expressed in COLO-829 at similar levels as in other melanoma cell lines (Supplementary Fig. S3B). While these mutations still warrant further analysis in order to determine whether they may yet have an impact on promoter activity, this finding suggests that the mutations are unlikely to be responsible for the complete removal of a DHS region in the COLO-829 cell line and a corresponding loss of expression of the associated gene.

Instead, there is some evidence that the DHS peak may have shifted in COLO-829, possibly due to the presence of the promoter mutation (Supplementary Fig. S3C), but this would need further investigation. Additionally, we also performed an analysis using DNase-seq and H3K4me3 ChIP-seq data from other common cell lines (Supplementary Fig. S3D and S3E). We found that as many as 48% (n = 11, from GM12878 lymphoblastoid cells; Supplementary Fig. S3D) of COLO-829 promoter mutations could be missed if using data from an unmatched cell line from a different cancer type. Some additional putative promoter mutations may be identified by analyzing DNase-seq and H3K4me3 ChIP-seq data from other cell lines. We performed an analysis of putative promoter mutations in DHS peaks (150 bp) that were not within 500 bp (±175 bp) of a COLO-829 DHS peak using data from different cell lines (Supplementary Fig. S3E).

We found a handful (n = 8) of additional mutations, many of which were unique to a single cell line. With the exception of a single mutation (chr7:75,834,729 at ZNF717; Supplementary Fig. S3E), all of these COLO-829 mutations fell outside of DHS regions from both COLO-829 and melanocytes. This means that it is unlikely that these mutations would be functional in either cell type and so do not warrant further testing. These results highlight the importance of using matched molecular data, while also potentially incorporating unmatched data to answer more targeted research questions.

**Evaluating the significance of the recurrent NDUFB9 promoter mutation**

It is interesting to note that somatic promoter mutations with functional consequences were reasonably prevalent—occurring in four of 23 (17.4%) promoter regions tested in COLO-829 (Fig. 3). As this study is to our knowledge, the first to systematically screen by experimentation a panel of promoter mutations in a single genome, an expected background rate of functional “mutations” by which to compare our findings is not available. However, given the few gene expression-altering promoter mutations that have yet been identified to date in cancer, a rate of 17.4% of promoter mutations being functional can be considered reasonably high, and therefore, it is likely that most promoter mutations, even those that are functional, are likely to be passenger mutations in the context of cancer. This is consistent with observations of nonsynonymous mutations within coding regions of the genome, where the majority are also generally believed to be passenger mutations. Nevertheless, as with coding mutations, there remains a possibility that any of the four functional promoter mutations identified in the COLO-829 genome are important in the pathogenesis of melanomas.

In order to determine whether there was a driver role among any of the four mutations altering promoter activity in COLO-829, we examined other cutaneous melanoma samples to identify recurrence—a key characteristic of driver mutations. Using the mutation calls from the 34 melanoma whole genomes available in the TCGA, we found that only the NDUFB9 promoter mutation was recurrent (2 of 34 samples). Because mutation calls at promoter regions can often be made from whole-exome sequence data, as sequencing reads frequently span up to 100 bp past intended capture regions (36), we further examined whole-exome sequencing data from cutaneous melanoma samples from the TCGA. We found that the four candidate promoter mutation sites were covered to varying extents (Supplementary Table S4). The sequencing data revealed that the NDUFB9 promoter mutation was still the only mutation that was recurrent across other cutaneous melanomas, present in 4.4% of all samples (19 of 432; Supplementary Table S4).

Owing to its recurrence, the NDUFB9 promoter mutation (chr8:125,551,344 C>T) was investigated in more detail to determine whether there was further evidence for a driver role in melanoma. The protein encoded by NDUFB9 is part of both the mitochondrial dysfunction and oxidative phosphorylation (OXPHOS) key canonical pathways. In reporter assays, activity of the mutant NDUFB9 promoter decreased compared to wild-type (Fig. 3). This is consistent with bioinformatic analysis that revealed that the mutation removes an SP1/Krüppel-like factor (KLF) binding motif (Fig. 4A) and lies in a highly conserved region (Fig. 4B). SP1 is a ubiquitously expressed transcription factor that binds to GC-rich regions of many promoters (37). Although no SP1 ChIP-seq data were available for COLO-829 or melanocytes, examination of SP1 ChIP-seq data in the HCT116 colorectal cancer cell line shows a clear peak at the promoter of NDUFB9 (Fig. 4A). As this promoter mutation substitutes a highly conserved cytosine to thymine, it is highly plausible that this mutation would result in a reduction in SP1 binding at the NDUFB9 promoter, accounting for the decrease observed in mutant promoter activity in reporter assays (Fig. 3).

To establish the effect of the mutation within the context of melanoma samples, TCGA RNA-sequencing data were analyzed. However, we found that there was no significant difference in NDUFB9 gene expression between wild-type and mutant cohorts (Fig. 4C). Nevertheless, the mutation does co-occur significantly in melanoma with nonsilent NF1 coding mutations (Fig. 4D). Co-occurrence with promoter mutations has previously been observed, with TERT and DPH3 promoter mutations both significantly co-occurring with NF1 and BRAF mutations, respectively (14). NF1 mutations are common in melanoma, resulting in the deregulation of extracellular signaling kinase pathways (38) and have been shown to be a key mediator of BRAF inhibitor resistance (39). Coding mutations of NF1 appear to play a role in the downregulation of NDUFB9 expression, with cooperative mutation of the NDUFB9 promoter showing the greatest decrease in NDUFB9 expression (Supplementary Figs. S4A and S4B). However, NDUFB9 expression does not appear to significantly correlate with patient survival (P = 0.18; Supplementary Fig. S4C), meaning that the impact of the
downregulation of NDUFB9 requires further investigation. However, there is some evidence that NDUFB9 is upregulated in response to UVA damage (40) and therefore it may be possible that downregulation of NDUFB9, via cooperative mutation of its promoter and NF1, is a first step in increasing the sensitivity of melanocytes to UV damage.

**Concluding Remarks**

In summary, we found that promoter mutations are common in melanoma, with functionally relevant somatic mutations that alter promoter activity potentially arising quite frequently. We identified four somatic mutations within the COLO-829 cell line that altered promoter activity when tested in these cells. One such mutation is a recurrent mutation in the NDUFB9 promoter region that may potentially play a role in melanoma development through a compound mutational pattern involving NF1. This study is the first of its kind to comprehensively survey somatic promoter mutations in a melanoma genome. By validating the functional consequence of each mutation experimentally in an unbiased manner, we have also evaluated the use of existing bioinformatics tools to prioritize functional noncoding mutations. The findings of this study provide a foundation for future work into the clinical significance of promoter mutations in human cancers.

**Disclosure of Potential Conflicts of Interest**

No potential conflicts of interest were disclosed.

**References**


Systematic Screening of Promoter Regions Pinpoints Functional Cis-Regulatory Mutations in a Cutaneous Melanoma Genome

Rebecca C. Poulos, Julie A.I. Thoms, Anushi Shah, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/1541-7786.MCR-15-0146

Supplementary Material
Access the most recent supplemental material at:
http://mcr.aacrjournals.org/content/suppl/2015/06/18/1541-7786.MCR-15-0146.DC1

Cited articles
This article cites 40 articles, 17 of which you can access for free at:
http://mcr.aacrjournals.org/content/13/8/1218.full.html#ref-list-1

Citing articles
This article has been cited by 2 HighWire-hosted articles. Access the articles at:
/content/13/8/1218.full.html#related-urls

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.