Evolutionary Action Score of TP53 Identifies High-Risk Mutations Associated with Decreased Survival and Increased Distant Metastases in Head and Neck Cancer

David M. Neskey1, Abdullah A. Osman2, Thomas J. Ow3, Panagiota Katsonis4, Thomas McDonald5, Stephanie C. Hicks6, Teng-Kuei Hsu4, Curtis R. Pickering2, Alexandra Ward2, Ameeta Patel7, John S. Yordy7, Heath D. Skinner8, Uma Giri9, Daisuke Sano10, Michael D. Story7,14, Beth M. Beadle11, Adel K. El-Naggar12, Merrill S. Kies13, William N. William15, Carlos Caulin2, Mitchell Frederick2, Marek Kimmel5, Jeffrey N. Myers2, and Olivier Lichtarge4

Abstract

TP53 is the most frequently altered gene in head and neck squamous cell carcinoma, with mutations occurring in over two-thirds of cases, but the prognostic significance of these mutations remains elusive. In the current study, we evaluated a novel computational approach termed evolutionary action (EAp53) to stratify patients with tumors harboring TP53 mutations as high or low risk, and validated this system in both in vivo and in vitro models. Patients with high-risk TP53 mutations had the poorest survival outcomes and the shortest time to the development of distant metastases. Tumor cells expressing high-risk TP53 mutations were more invasive and tumorigenic and they exhibited a higher incidence of lung metastases. We also documented an association between the presence of high-risk mutations and decreased expression of TP53 target genes, highlighting key cellular pathways that are likely to be dysregulated by this subset of p53 mutations that confer particularly aggressive tumor behavior. Overall, our work validated EAp53 as a novel computational tool that may be useful in clinical prognosis of tumors harboring p53 mutations. Cancer Res; 75(7); 1527–36. ©2015 AACR.

Introduction

Head and neck squamous cell carcinoma (HNSCC) is the sixth most common cancer worldwide and accounts for over 45,000 new cases annually in the United States (1,2). Because TP53 is the most frequently mutated gene in HNSCC, genomic alterations in this gene are key events in the development and progression of this disease (3–6).

Multiple studies have demonstrated that TP53 mutations are prognostic for poor outcomes in HNSCC, yet molecular testing for TP53 alterations has not become routine in clinical practice (7–11). Although several classification systems have been described, the main limitation of TP53 as a prognostic biomarker is the lack of a reliable system to accurately assess the functional and clinical impact of specific mutations (10). Whereas most alterations involving tumor suppressor genes render them non-functional through truncating mutations or deletions, p53 is unique in that there is a strong selection bias for missense mutations, particularly within the DNA-binding domain. P53 mutation can result in loss of wild-type functions through either the loss of DNA-binding activity of p53 responsive elements or a dominant-negative effect where the mutated allele binds and inhibits the remaining functional wild-type allele (12). Moreover, some mutant p53 display oncogenic properties, termed 'gain of function' (GOF), which are independent of wild-type p53 (wp53) function (13). Accordingly, GOF p53 mutants can enhance cell transformation, increase tumor formation in mice, and confer cellular resistance to chemotherapy (14,15).
Although this GOF activity has been well characterized in cancer for five ‘‘hotspot’’ or frequently altered p53 amino acids, 175, 245, 248, 273, and 282, our work indicates that non ‘‘hotspot’’ mutations can also confer GOF activity (16). Therefore, we hypothesized that there is a subset of mutations that are particularly deleterious to p53 function resulting in a GOF phenotype and are associated with adverse outcomes in patients with HNSCC.

In an effort to predict which TP53 mutations are highly deleterious, we extended the evolutionary trace (ET) approach, an extensively validated method to identify key functional or structural residues in proteins (17). This is achieved by assigning every sequence position a grade of functional sensitivity to sequence variations, defined by whether its evolutionary substitutions correlate with larger or smaller phylogenetic divergences. Residues with large ET grades typically cluster structurally into evolutionary ‘‘ hotspots’’ that overlap and predict functional sites (18). In large-scale validation studies, motifs made of top-ranked ET residues predict function in protein structures (19), accurately enough to anticipate enzyme substrates (20).

We have hypothesized that the ET method would assess the impact of TP53 missense mutations. The impact should be greater when the mutated residues are more evolutionarily sensitive to sequence variations, i.e., have a larger ET grade, and also when the amino acid change is least conservative, so the mutational impact is the largest. These two components were computed and combined into a single score, called evolutionary action (EA; ref. 21). This action has been shown to correlate linearly with loss of protein function in test systems and with morbidity in Mendelian diseases, as well as apply across protein coding variations population-wide. To apply this EA to TP53 mutations in HNSCC, we further developed a scoring system (EAp53) to stratify TP53 missense mutations into high and low risk.

The goals of this study were to evaluate the ability of EAp53 to identify a subset of TP53 mutations in HNSCC that are associated with the worst patient outcomes and to validate the impact of these mutations in laboratory-based models. We found that EAp53 could identify mutations with GOF phenotypes, termed high risk, that are highly prognostic of poor overall survival, progression-free survival, and the development of distant metastasis in two patient cohorts. Furthermore, high-risk mutations were found to be associated with increased cellular invasion, tumorigenicity, and propensity for distant metastases in both in vitro and in vivo models, thereby associating functionally significant TP53 mutations with outcomes of patients harboring high-risk TP53 alterations. These oncogenic p53 mutations also had a distinct mRNA expression profile, which suggests high-risk p53 mutants regulate unique cellular pathways at the transcriptional level. These findings highlight the need for further evaluation of the EAp53 scoring system as prognostic biomarker in both retrospective and prospective datasets. Finally, this study emphasizes the need for continued investigation into the cellular pathways driving the oncogenic phenotype of these high-risk p53 mutations, which could lead to the identification of novel therapeutics targets and ultimately personalization of cancer treatment based on p53 mutational status.

Patients and Methods
Patient selection and tissue procurement for the EAp53 method

Training set. A cohort of patients was identified from The Cancer Genome Atlas (TCGA) HNSC project that had human papilloma virus (HPV)–negative tumors (see supplementary Methods for additional details) and underwent surgical resection alone (n = 103) or surgery followed by postoperative radiotherapy (n = 65). Patient and tumor characteristics along with outcome data were extracted from TCGA HNSC Supplementary Data.

Validation set. A cohort of patients with HNSCC treated with surgery followed by postoperative radiotherapy at The University of Texas MD Anderson Cancer Center (UTMDACC) from 1992 to 2002 was identified (n = 96). Clinical records were reviewed retrospectively, and TP53 gene status was determined according to a protocol approved by the Institutional Review Board at UTMDACC. Patients who received chemotherapy were excluded. Clinical and pathologic factors were recorded, including patient age, sex, T and N stage, surgical margin status, and extracapsular lymph node extension (ECE). Survival outcomes including overall and disease-free survival along with time to distant metastases were also determined.

DNA isolation

Samples were isolated using three different methods depending on the platform used to perform the sequencing of TP53. The techniques used either snap-frozen tumor sample or formalin-fixed, paraffin-embedded tissue, and the detailed description of the extraction method is included in Supplementary Methods.

TP53 sequencing

As mentioned above, three different techniques were utilized to determine TP53 sequence. In each assay, the coding regions and surrounding splice sites from exons 2 to 11 of the TP53 gene were evaluated via direct sequencing from genomic DNA, and a detailed description is included in Supplementary Methods.

Calculation of the EAp53 scores

The EA scores for each TP53 mutation were calculated based on a simple model of the phenotype–genotype relationship (21), which hypothesizes that protein evolution is a continuous and differentiable process. Accordingly, the genotype (g) and the fitness phenotype (f) will be related by $f = g^{-i} g^{f}$, and the phenotypic impact of any mutation at residue i (evolutionary action) will be the product of two terms: the sensitivity of p53 function to residue variations ($\Delta g_i$) and the magnitude of the substitution ($\Delta f_i$). The term $\Delta f_i/\Delta g_i$ was measured by importance ranks of the Evolutionary Trace method (22,23), according to which, residues that vary among closer homologous sequences are ranked less important than those that only vary among distant homologous sequences. The magnitude of the substitution ($\Delta f_i$) was measured by ranks of amino acid substitution odds (24), however, these odds were computed for different deciles of the evolutionary trace grade at the substituted position. We normalized the product to become percentile scores for p53 protein, for example, an EA score of 68 implied that the impact was higher than 68% of all possible amino acid substitutions in p53.

Statistical classification by EAp53

Missense mutations were scored by EAp53 from 0 to 100 with higher scores representing more deleterious alterations. wtp53 sequences were scored as zero because this is the normally functioning protein; see Supplementary Data for additional description of scoring system. In addition to obtain scores for

Patients selection and tissue procurement for the EAp53 method
individual p53 mutations, an EAp53 server is available at http://mammoth.bcm.tmc.edu/EAp53.

Univariate Cox proportional hazard models were used to estimate HRs and their corresponding P values for all risk factors in the training set. The optimal threshold for EAp53 to stratify patients between favorable and poor outcomes was identified using the training dataset, and the P value for the estimated HR was adjusted (25). The threshold discretizes EA into low-risk and high-risk patients.

Multivariate Cox proportional hazards model was created with all risk factors except the discretized EA factor. Rank-based procedure was used based on the Cox proportional hazards model to determine the best threshold for EA in a multivariate model (26). A Cox proportional hazard model using all risk factors as above along with the discretized EA factor was created to estimate HRs for each risk factor and their P values. The P value for the discretized EA factor was adjusted to account for using the data to first determine the cutoff point. We removed covariates from the model one by one and repeated the above procedure until a model was developed, containing risk factors that have HRs that were significant, and the P value for the coefficient of the factors in the final model was < 0.05.

Next, the threshold established in the training dataset was applied to the validation dataset to classify TP53 mutations as either low or high risk. Using survival time, univariate Cox proportional hazard models were built for each risk factor using the validation dataset to get estimates for HRs and their P values. No adjustments to the P values were necessary. Cox proportional hazard model was also built containing all the risk factors including the EAp53 threshold determined in the test set and reducing the model by removing the covariate with the largest P value in a stepwise manner until the final model only contained significant risk factors.

We used the discretized EA data to perform log-rank tests to determine differences in time to death between low-risk and high-risk TP53 patients. The above analyses for the validation dataset were reproduced with disease-free survival as an outcome and time to metastases as an outcome.

Site-directed mutagenesis and TP53 constructs

Mutations with a variety of EA scores (Supplementary Table S3) were produced using a QuickChange-II Site-Directed Mutagenesis Kit (Agilent Technologies), starting with wild-type human p53 cDNA with polymorphic region 72 containing arginine previously cloned into a pBABEpuro; Addgene). Primer sequences used for mutagenesis are shown in Supplementary Table S3. All mutagenesis steps were performed according to the manufacturer’s protocol. Mutations were confirmed by Sanger sequencing at the MD Anderson Cancer Center DNA core facility.

Generation of the HNSCC stable cell lines

Two HNSCC cell lines UMSCC1 and PCI13 were selected for their lack of p53 expression due to a splice-site in UMSCC1 (hg19:chr17:7578370C>T) and a deletion in PCI13 (hg19:chr17:7579670G>757970del). UMSCC1 was provided by Dr. Thomas Carey (University of Michigan). PCI13 was acquired from Dr. Jennifer Grandis (University of Pittsburgh). The genomic identities of the cell lines were authenticated using short tandem repeat analysis (27). Cells stably expressing TP53 constructs were generated as described previously, and a detailed description of the technique is included in Supplementary Methods (28).

Immunoblotting

Western blotting was performed using standard techniques previously described (27), and primary antibodies to anti-p53 (Santa Cruz Biotechnology; sc-126), anti-p21 (Calbiochem; OP64), and beta-actin (sc81178; Santa Cruz Biotechnology) were used.

Invasion assay

BD BioCoat Matrigel Invasion Assays were used following the manufacturer’s protocol, and a detailed description is included in Supplementary Methods. Each cell line was run in triplicate. 3T3 NIH cells were used as an invasion control as these cells characteristically migrate but do not invade. The low risk and high risk represent a composite of three mutations, F134C, A161S, and Y236C, and four mutations, R175H, H179Y, C238F, and G245D, respectively, and the results represent two independent experiments.

Cell proliferation assay

Cell proliferation was determined using a 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyterazolium bromide (MTT) assay as previously described (29). Briefly, the cells were seeded at different densities and grown in a medium containing 10% FBS in 96-well tissue culture plates. After a 24-hour attachment period, the plates were assigned to different time points to obtain linear optical density (OD). Cells were then incubated for 3 hours in medium containing 2% FBS and 0.25 mg/mL MTT, after which the cells were lysed in 200 μl dimethylsulfoxide (DMSO) to release the formazan. The conversion of MTT to formazan was quantified with an EL-808 96-well plate reader (BioTek Instruments) set at an absorbance of 570 nm. The OD values were then obtained and analyzed to determine the percentage of cell viability.

Orthotopic nude mouse model of oral cavity cancer

All animal experimentation was approved by the Animal Care and Use Committee of the UTMDACC. Our orthotopic nude mouse model of oral cavity cancer has been previously validated, and a detailed description of the technique is included in Supplementary Methods (30). The results represent three independent experiments with the low- and high-risk mutations, including two mutations (F134C and A161S) and four mutations (R175H, H179Y, C238F, and G245D), respectively.

Tail vein model

The experimental metastatic tail vein injection model was performed using standard techniques previously described, and a detailed description is included in Supplementary Methods (31). The results represent four independent experiments with PCI13 and UMSCC1 cell lines harboring low-risk mutations: F134C or A161S, high-risk mutations: R175H, C238F or G245D, pBabe, or wtP53.

mRNA expression arrays

Total RNA was isolated from cell lines by using Tri-reagent and hybridized to Affymetrix GeneChip Human Exon 1.0ST Arrays (Affymetrix) according to the manufacturer’s instructions, and a detailed description is included in Supplementary Methods. The
One hundred and two p53 target genes were identified and validated through rigorous literature review (32). The difference in expression between pBabe and each of other groups was calculated, and heatmaps were generated depicting these expression patterns target gene and 49 selected genes.

Quantitative RT-PCR analyses

Validation of the mRNA expression array for two TP53 target genes (p21 and Notch1) was performed by quantitative RT-PCR (qRT-PCR). Total RNA was isolated from HNSCC cell lines using the TRIzol method. Reverse transcription was performed using the high capacity cDNA Reverse Transcription Kit (Applied Biosystems) according to the manufacturer's protocol, and a detailed description is included in Supplementary Methods.

Results

Patient characteristics and follow-up in the training and validation sets

Analysis of the TCGA HNSCC cohort identified 168 patients with non–HPV-associated tumors harboring either missense mutations or wtp53 for our training cohort. The majority of these patients presented with advanced T stage (70%), whereas only 30% of patients presented with advanced neck disease (Supplementary Table S1). The median survival time was 2.96 years, and median follow-up time was 2.49 years.

Ninety-six patients treated at the MD Anderson Cancer Center for HNSCC through surgical resection and postoperative radiotherapy comprised an independent validation set for the EAp53 scoring system. The majority of patients had advanced T stage (76%) or N stage (54%), and nearly half had aggressive pathologic features (Supplemental Table S1). The median survival time and median follow-up time were 5.49 and 7.66 years, respectively.

TP53 mutation in the patient cohorts

One hundred and forty-three TP53 missense mutations were identified in 127 patients in the training set, whereas 50 missense mutations were identified in 47 patients in the validation cohort (Supplementary Tables S2 and S3). The majority of these missense mutations were within the DNA-binding domain. Furthermore, 34% and 38% of the mutations were located in the well-characterized TP53 hotspot sites, 175, 245, 248, 273, 282, in the training and validation sets, respectively.

Table 1 Univariate analysis of overall outcomes in the training and validation sets

<table>
<thead>
<tr>
<th>Prognostic variable</th>
<th>Total</th>
<th>HR 95% CI</th>
<th>P value</th>
<th>Total</th>
<th>HR 95% CI</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall survival</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training set</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>168</td>
<td>1.41 (1.03–1.94)</td>
<td>0.0319</td>
<td>96</td>
<td>1.21 (0.88–1.68)</td>
<td>0.2449</td>
</tr>
<tr>
<td>EAp53 status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-risk (ref)*</td>
<td>86</td>
<td>1</td>
<td></td>
<td>64</td>
<td>1.71 (1.02–2.89)</td>
<td>0.0381</td>
</tr>
<tr>
<td>High-risk</td>
<td>82</td>
<td>2.25**</td>
<td>0.008</td>
<td>32</td>
<td>1.75 (1.03–2.85)</td>
<td>0.0381</td>
</tr>
<tr>
<td>T stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 2 (ref)</td>
<td>47</td>
<td>1</td>
<td></td>
<td>23</td>
<td>1.87 (0.99–3.54)</td>
<td>0.0359</td>
</tr>
<tr>
<td>&gt;2</td>
<td>121</td>
<td>1.14 (0.69–1.90)</td>
<td>0.6108</td>
<td>73</td>
<td>1.74 (0.96–3.16)</td>
<td>0.0700</td>
</tr>
<tr>
<td>N stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;2 (ref)</td>
<td>123</td>
<td>1</td>
<td></td>
<td>45</td>
<td>2.28 (1.35–3.87)</td>
<td>0.0021</td>
</tr>
<tr>
<td>≥2b</td>
<td>45</td>
<td>1.45 (0.86–2.43)</td>
<td>0.1601</td>
<td>51</td>
<td>2.20 (1.33–3.63)</td>
<td>0.0020</td>
</tr>
<tr>
<td>Site: larynx</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (ref)</td>
<td>116</td>
<td>1</td>
<td></td>
<td>84</td>
<td>0.63 (0.30–1.33)</td>
<td>0.2271</td>
</tr>
<tr>
<td>Yes</td>
<td>52</td>
<td>0.97 (0.59–1.59)</td>
<td>0.9005</td>
<td>12</td>
<td>0.74 (0.36–1.51)</td>
<td>0.4083</td>
</tr>
<tr>
<td>Site: pharynx</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (ref)</td>
<td>160</td>
<td>1</td>
<td></td>
<td>82</td>
<td>1.40 (0.75–2.64)</td>
<td>0.2941</td>
</tr>
<tr>
<td>Yes</td>
<td>8</td>
<td>2.35 (0.94–5.87)</td>
<td>0.0673</td>
<td>14</td>
<td>1.20 (0.64–2.25)</td>
<td>0.5646</td>
</tr>
<tr>
<td>Positive margins</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (ref)</td>
<td>85</td>
<td>1</td>
<td></td>
<td>85</td>
<td>1.55 (0.76–3.16)</td>
<td>0.2272</td>
</tr>
<tr>
<td>Yes</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>11</td>
<td>1.49 (0.75–3.02)</td>
<td>0.2685</td>
</tr>
<tr>
<td>Extracapsular extension</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (ref)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>51</td>
<td>1.43 (0.87–2.36)</td>
<td>0.1667</td>
</tr>
<tr>
<td>Yes</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>45</td>
<td>1.35 (0.83–2.19)</td>
<td>0.2220</td>
</tr>
</tbody>
</table>

*Low-risk is defined as EAp53 score of less than 74.39 in the training set and less than 75 in the validation set and includes wtp53.

**The lack of a 95% confidence interval is because the cutpoint methods used and referenced can determine significance of the hazard ratio in the form of a P value, but a mathematical derivation of the confidence interval has not yet been determined.

Figure 1. EAp53 identifies two functionally distinct groups of p53 mutations. Patients with tumors harboring high-risk EAp53 (n = 83) mutations have a decreased overall survival relative to low-risk EAp53 (n = 44) mutations and wtp53 (n = 41). In contrast, patients with low-risk EAp53 mutations appear to have similar survival outcomes to wtp53.
Validation and training EAp53 can identify patients with decreased survival outcomes in HNSCC. Log-rank tests of Kaplan-Meier analysis in the training set revealed that the low EAp53 score, high EAp53 score, and wtp53 mutants were assessed as risk factors for poor overall survival in the training set. The initial comparison analyzed three groups independently, low EAp53, high EAp53, and wtp53. Univariate analysis in the training set revealed that the low EAp53 score mutations, i.e. low risk, and wild type were not statistically different, whereas the high EAp53 score mutations, termed high-risk mutations, appeared to be distinct from the other two groups (Table 1 and Fig. 1). Given the similar outcomes, patients with tumors having low-risk mutations were combined with wtp53. After univariate analyses, an EA threshold of greater than 74.39 identified a group of 82 patients with high-risk mutations associated with a significantly decreased overall survival compared with 86 patients in the low-risk group \( (P = 0.008; \text{Table 1}) \). A rank-based procedure based on the Cox proportional hazard ratios in a multivariate model was then applied to the training set, and identified an EA threshold of 77.78 \( (P = 0.009; \text{Table 2}) \). Because this model relies on a threshold specific to the TCGA data, the range from 74.39 to 77.78 should contain the valid threshold. Therefore, we applied a threshold of 75 to the validation set to determine low- and high-risk individuals. With this threshold established, this prognostic model was validated in an independent dataset, and patients with HNSCC harboring high-risk mutations \((n = 33)\) had significantly decreased overall and disease-free survival relative to the low-risk group \((n = 63)\). \[ P = 0.030 \text{ and } P = 0.011 \text{ (Tables 1 and 2)} \]. Furthermore, this survival difference appeared to be associated with the development of distant metastases \((P = 0.059; \text{Table 2}) \). In addition, when patients that achieved locoregional control were analyzed, a log-rank test revealed tumors with high-risk mutations still associated with increased rate of distant metastases \((P = 0.00006) \). In addition, in a multivariate analysis using backward variable selection, the final model identified N stage \( \geq N2b, P = 0.0064 \) and high-risk mutations, \( P = 0.05 \), as predictors of distant metastases.

Survival analysis using the Kaplan–Meier method revealed that high-risk TP53 mutations were associated with poor outcomes for overall survival, disease-free survival, and distant metastasis \((P = 0.04, 0.04, \text{and } 0.09, \text{respectively; Fig. 2A–C}) \).

**High-risk TP53 mutations are more invasive in vitro and tumorigenic in vivo**

To determine the impact of the high-risk TP53 mutations in preclinical models, several low- and high-risk TP53 mutant transcripts were stably expressed in two HNSCC cell lines that do not endogenously express p53 due to bi-allelic alterations including somatic mutations (Supplementary Fig. S1). Cell lines expressing high-risk TP53 mutations, including R175H, H179Y, C238F, and G245D, were significantly more invasive than cells expressing wtp53, p53 null (pBabe), or low-risk mutations, A161S, F134C, and Y236C \((P = 0.001, 0.05, \text{and } 0.03, \text{respectively; Fig. 3A}) \). Assessment of cell motility by scratch assay revealed that cells bearing high-risk TP53 mutants, R175H, G245D, and C238F, were significantly more motile than cells expressing wtp53, pBabe, or low-risk TP53 mutation, A161S \((P = 0.04, 0.01, \text{and } 0.002, \text{respectively; Fig. 3B}) \). In contrast with the increased cellular

### Table 2. Cox proportional hazard regression after backward variable selection and adjusting for EA score for overall survival in the training set

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>HR 95% CI</th>
<th>P value Adjusted for</th>
<th>Model likelihood ratio P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall survival</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EA score ( \geq 77.78 )</td>
<td>( 2.30^* )</td>
<td>0.0087 Age, pharyngeal site</td>
<td>0.0002</td>
</tr>
<tr>
<td>Overall survival</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EA score ( \geq 75 )</td>
<td>( 1.79 (1.06–3.04) )</td>
<td>0.0500 N stage ( \geq 2b ), laryngeal site</td>
<td>0.0002</td>
</tr>
<tr>
<td>Disease-free survival</td>
<td>( 1.94 (1.16–3.25) )</td>
<td>0.0176 N stage ( \geq 2b )</td>
<td>0.0003</td>
</tr>
<tr>
<td>Time to distant metastases</td>
<td>( 2.35 (0.97–5.69) )</td>
<td>0.0591 N stage ( \geq 2b )</td>
<td>0.001</td>
</tr>
</tbody>
</table>

**NOTE:** The threshold of 75 was set and the same analysis was performed on the validation set for overall survival, disease-free survival, and time to distant metastasis in the validation set.

*The lack of a 95% confidence interval is because the cutoff points used and referenced can determine significance of the hazard ratio in the form of a P value, but a mathematical derivation of the confidence interval has not yet been determined.

## Determination of EAp53 threshold, risk models, and survival analysis

Age, disease site, T-stage, N-stage, and TP53 missense mutations were assessed as risk factors for poor overall survival in the training set. The initial comparison analyzed three groups independently, low EAp53 score, high EAp53 score, and wtp53. Univariate analysis in the training set revealed that the low EAp53 score mutations, i.e. low risk, and wild type were not statistically different, whereas the high EAp53 score mutations, termed high-risk mutations, appeared to be distinct from the other two groups (Table 1 and Fig. 1). Given the similar outcomes, patients with tumors having low-risk mutations were combined with wtp53. After univariate analyses, an EA threshold of greater than 74.39 identified a group of 82 patients with high-risk mutations associated with a significantly decreased overall survival compared with 86 patients in the low-risk group \( (P = 0.008; \text{Table 1}) \). A rank-based procedure based on the Cox proportional hazard ratios in a multivariate model was then applied to the training set, and identified an EA threshold of 77.78 \( (P = 0.009; \text{Table 2}) \). Because this model relies on a threshold specific to the TCGA data, the range from 74.39 to 77.78 should contain the valid threshold. Therefore, we applied a threshold of 75 to the validation set to determine low- and high-risk individuals. With this threshold established, this prognostic model was validated in an independent dataset, and patients with HNSCC harboring high-risk mutations \((n = 33)\) had significantly decreased overall and disease-free survival relative to the low-risk group \((n = 63)\). \[ P = 0.030 \text{ and } P = 0.011 \text{ (Tables 1 and 2)} \]. Furthermore, this survival difference appeared to be associated with the development of distant metastases \((P = 0.059; \text{Table 2}) \). In addition, when patients that achieved locoregional control were analyzed, a log-rank test revealed tumors with high-risk mutations were still associated with increased rate of distant metastases \((P = 0.00006) \). In addition, in a multivariate analysis using backward variable selection, the final model identified N stage \( \geq N2b, P = 0.0064 \) and high-risk mutations, \( P = 0.05 \), as predictors of distant metastases.

Survival analysis using the Kaplan–Meier method revealed that high-risk TP53 mutations were associated with poor outcomes for overall survival, disease-free survival, and distant metastasis \((P = 0.04, 0.04, \text{and } 0.09, \text{respectively; Fig. 2A–C}) \).

High-risk TP53 mutations are more invasive in vitro and tumorigenic in vivo

To determine the impact of the high-risk TP53 mutations in preclinical models, several low- and high-risk TP53 mutant transcripts were stably expressed in two HNSCC cell lines that do not endogenously express p53 due to bi-allelic alterations including somatic mutations (Supplementary Fig. S1). Cell lines expressing high-risk TP53 mutations, including R175H, H179Y, C238F, and G245D, were significantly more invasive than cells expressing wtp53, p53 null (pBabe), or low-risk mutations, A161S, F134C, and Y236C \((P = 0.001, 0.05, \text{and } 0.03, \text{respectively; Fig. 3A}) \). Assessment of cell motility by scratch assay revealed that cells bearing high-risk TP53 mutants, R175H, G245D, and C238F, were significantly more motile than cells expressing wtp53, pBabe, or low-risk TP53 mutation, A161S \((P = 0.04, 0.01, \text{and } 0.002, \text{respectively; Fig. 3B}) \). In contrast with the increased cellular...

---

**A** Overall survival

**B** Disease-free survival

**C** Time to distant metastases

*Figure 2.*

EAp53 can identify patients with decreased survival outcomes in HNSCC. Log-rank tests of Kaplan–Meier survival plots for a cohort of patients treated with surgery followed by radiotherapy validated that EAp53 can identify patients with high-risk mutations that are associated with a decreased overall and disease-free survival, \( P = 0.041 \) and 0.036, respectively (A and B). In addition, high-risk EAp53 mutations appear to be associated with an increased rate of distant metastases, \( P = 0.092 \) (C).
invasion and motility observed in the high-risk TP53 mutations, the rate of proliferation was independent of TP53 status (Supplementary Fig. S2). From these results, we concluded that the expression of high-risk mutations is associated with greater migration and invasion, suggesting a GOF phenotype.

To validate our in vitro data, the isogenic p53 mutant cell lines were introduced into an orthotopic nude mouse tongue cancer model. Although neither the rate nor time of tumor formation was different across various p53 statuses, cell lines that expressed high-risk TP53 mutations formed significantly larger tumors compared with cells that expressed wtp53, pBabe, low-risk mutations, A161S or F134C, or high mutations, R175H, H179Y, C238F, G245D, were introduced into the orthotopic model of tongue cancer (C and D).

High-risk p53 mutations have higher propensity for lung metastases

To directly evaluate the contribution of high-risk p53 mutations in the development of distant metastases, we used the experimental metastatic tail vein injection model. Animals injected with tumor cells carrying high-risk mutations, R175H, C238F, or G245D, had significantly more grossly visible pulmonary metastases when compared with those injected with cells with low-risk p53 mutations, F134C or A161S, P = 0.01, pBabe, P = 0.008, and wtp53, P = 0.004 (Fig. 4A and B). In addition, animals injected with high-risk mutant bearing cells had higher numbers of microscopic metastases (Fig. 4C and D).

High-risk p53 mutations have a distinct expression profile of known p53 target genes

In an effort to identify genes or pathways responsible for the increased tumorigenicity and metastases associated with high-risk TP53 mutations, we performed mRNA expression arrays of cells harboring either wtp53, pBabe, low-risk mutation, A161S, or high-risk mutation, C238F. After standardizing gene expression to the control p53 null cell line, pBabe, an analysis of 102

---

**Figure 3.** High-risk TP53 mutations are more invasive in vitro and tumorigenic in vivo. Matrigel invasion assays were performed on PCI13 isogenic HNSCC cell line expressing either high- or low-risk TP53 mutations, wtp53, pBabe empty vector control. A, the low- and high-risk series are a composite of three mutations, F134C, A161S, and Y235C, and four mutations, R175H, H179Y, C238F, G245D, respectively, and the results represent two independent experiments. B, scratch assays were performed on PCI13 isogenic HNSCC cell line expressing wtp53, pBabe empty vector control, low-risk mutation A161S (shown), or high-risk mutations, G245D, R175H, C238F (shown). Percent scratch closure represents two independent experiments, and the mean closure is represented for each p53 status. PCI13 and UMSCC1 isogenic cell lines expressing either wtp53, pBabe, low-risk mutations, A161S or F134C, or high mutations, R175H, H179Y, C238F, or G245D, were introduced into the orthotopic model of tongue cancer (C and D).
p53 target genes revealed wtp53 cells had an elevated level of expression in 61% (62/102) of the targets genes, whereas the high-risk mutant expressing cells suppressed expression in the majority of the same genes. In contrast, the expression pattern of cells expressing a low-risk mutation was more similar to that of wtp53 cells as revealed by the hierarchal clustering (Fig. 5A). Upon further analysis of the entire expression array, 49 genes were differentially expressed with wtp53 and the high-risk mutation having elevated and suppressed levels of expression, respectively, relative to p53 null, pBabe cells (Fig. 5B). Within these 49 genes, three p53 target genes p21, Notch1, and BTG2 were shown to have this differential expression profile. qRT-PCR validated the elevated level of expression of these three genes in cells expressing wtp53 or low-risk p53 mutations, A161S and Y236C, and the significantly decreased expression in cells bearing high-risk mutations, C238F and G245D, relative to cells without p53 expression (Fig. 5C).

Discussion

An appropriate method for classifying TP53 mutations and assessing the functional impact of these alterations has been sought for decades, yet a suitable classification scheme remains elusive. Therefore, we have utilized a novel computational method that considers evolutionary variations to predict the functional impact of TP53 mutations. We hypothesized that this model could identify missense mutations with the greatest functional impact that would translate to decreased patient survival outcomes. Our analysis of two patient datasets along with in vitro and in vivo models supports TP53 mutational status as a prognostic biomarker in HNSCC. Currently, TP53 mutational status is not incorporated into the clinical evaluation of patients with HNSCC, even though previous reports have identified TP53 mutations associated with poor patient outcomes (8, 10, 11). These previous systems either...
heuristically stratified mutations based on their potential impact on the p53 molecule and could not be validated in these two cohorts (Supplementary Tables S5 and S6; Supplementary Figs. S3 and S4) or stratified TP53 mutations by the presence of a copy-number loss using a multitiered genomic approach (11). In contrast, EAp53 is based on a formal model of evolution that has been shown to perform well against other approaches (17,21,33), and has been refined here to specifically evaluate the clinical impact of TP53 mutations. Our results show that patients with HNSCC harboring high-risk TP53 mutations had significantly worse overall survival, which was validated in a second patient cohort.

In addition to the development and validation of a novel method for stratifying TP53 mutations, we evaluated several low- and high-risk p53 mutations in two isogenic HNSCC cell lines in both in vitro and in vivo models. These results corroborated that EAp53 can identify a subset of high-risk mutations that are more invasive and tumorigenic than the loss of p53, implying an oncogenic or GOF phenotype. In addition, as seen in the patient cohort, the high-risk mutations had a greater propensity for distant metastases in an experimental metastatic model. Furthermore, the low-risk mutations were most similar to wtp53 in both in vivo and in vitro models, suggesting that these alterations may maintain some residual tumor suppressive functions. Although previous classification systems in HNSCC have identified mutations associated with poor outcomes, these studies have not correlated the potential functional impact of these mutations at the cellular level (7,8,10). In addition, previous functional studies in yeast that determined the transactivational activity of TP53 mutations classified most alterations as nonfunctional, including many hotspot mutations considered to have GOF properties (34). Given the limitations of these previous studies, we were able to identify patients with tumors harboring high-risk mutations associated with decreased survival and confirm these same alterations have oncogenic properties in preclinical assays.

Analysis of mRNA expression arrays revealed that cells with wtp53 had elevated levels of expression in the majority of TP53 target genes, whereas cells bearing high-risk mutations had suppressed levels of expression in these same genes, which may partially explain the GOF properties observed with these mutations. The low-risk mutations modulated the expression of TP53 target genes more similarly to wtp53 again, implying some residual wild-type function. Suppressed target gene expression in mutations that have GOF phenotypes corroborates previous studies that have postulated alteration of the p53 transcriptome is a potential mechanism of GOF activities (35,36).

In conclusion, the EAp53 system appears to identify high-risk mutations associated with decreased survival and increased development of distant metastases in patients with HNSCC, which is corroborated in both in vitro and in vivo studies of invasion, tumorigenicity, and development distant metastases. To adopt EAp53 into clinical practice, it will be necessary to confirm the prognostic utility in prospective clinical trials with patients with HNSCC managed with the therapeutic standard of care.
Furthermore, further investigation into the cellular pathways driving the oncogenic phenotype of these high-risk p53 mutations is necessary, which could lead to the identification of novel therapeutics targets and ultimately personalization of cancer treatment based on p53 mutational status.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors’ Contributions

Conception and design: D.M. Neskey, A.A. Osman, T.J. Ow, P. Katsonis, A. Patel, D. Sano, B.M. Beadle, M. Kimmel, J.N. Myers, O. Lichtarge


Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): D.M. Neskey, A.A. Osman, T.J. Ow, C.R. Pickering, A. Ward, A. Patel, I.S. Yordy, L. Giri, M.D. Story, M.S. Kies


Writing, review, and/or revision of the manuscript: D.M. Neskey, N.A. Osman, T.J. Ow, P. Katsonis, S.C. Hicks, I.S. Yordy, H.D. Skinner, B.M. Beadle, W.N. William, C. Caulin, M. Frederic, M. Kimmel, J.N. Myers, O. Lichtarge

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): D.M. Neskey, A. Ward, A. Patel, I.S. Yordy, B.M. Beadle, M.S. Kies, J.N. Myers

Study supervision: A.A. Osman, M. Kimmel, J.N. Myers, O. Lichtarge

Other (performed experiments): A.A. Osman

Other (provided materials and reviewed pathology): A.K. El-Naggar

Grant Support

This work was supported by the UTMDACC PANTHEON program, the UTMDACC Sisters Institution Network Fund, the National Institute of Health Specialized Program of Research Excellence Grant (P50CA097007), the NIH R01 DE14613, Cancer Prevention and Research Institute of Texas (RP120258), National Research Science Award Institutional Research Training Grant (T32CA60374), the National Institute of Health Program Project Grant (CA06294), and the Cancer Center Support Grant (CA016672). O. Lichtarge gratefully acknowledges support from NIH R01 GM079656 and R01 GM086699, and NSF DBI 1356566 and DBI 0851393. P. Katsonis was supported by the Pharmacoinformatics Training Program of the Keck Center of the Gulf Coast Consortia (NIH grant no. 5 R00 DK071505).

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

Received September 18, 2014; revised November 19, 2014; accepted December 2, 2014; published OnlineFirst January 29, 2015.
Evolutionary Action Score of TP53 Identifies High-Risk Mutations Associated with Decreased Survival and Increased Distant Metastases in Head and Neck Cancer


Updated version
Access the most recent version of this article at:
doi:10.1158/0008-5472.CAN-14-2735

Supplementary Material
Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2015/02/03/0008-5472.CAN-14-2735.DC1

Cited articles
This article cites 32 articles, 10 of which you can access for free at:
http://cancerres.aacrjournals.org/content/75/7/1527.full#ref-list-1

Citing articles
This article has been cited by 1 HighWire-hosted articles. Access the articles at:
http://cancerres.aacrjournals.org/content/75/7/1527.full#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.