Cancer Research

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

David M. Neskey[1], Abdullah A. Osman[2], Thomas J. Ow[3], Panagiotis Katsonis[4], Thomas McDonald[5], Stephanie C. Hicks[6], Teng-Kuei Hsu[4], Curtis R. Pickering[2], Alexandra Ward[2], Ameeta Patel[2], John S. Yordy[7], Heath D. Skinner[8], Uma Giri[9], Daisuke Sano[10], Michael D. Story[7,14], Beth M. Beadle[11], Adel K. El-Naggar[12], Merrill S. Kies[13], William N. William[13], Carlos Caulin[2], Mitchell Frederick[2], Marek Kimmel[5], Jeffrey N. Myers[2], and Olivier Lichtarge[4]

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.*

[1]Department of Otolaryngology Head and Neck Surgery, Hollings Cancer Center, Medical University of South Carolina, Charleston, South Carolina. [2]Department of Head and Neck Surgery, The University of Texas M. D. Anderson Cancer Center, Houston, Texas. [3]Department of Otolaryngology Head and Neck Surgery, Albert Einstein School of Medicine, Yeshiva University, New York, New York. [4]Department of Human and Molecular Genetics, Baylor College of Medicine, Houston, Texas. [5]Department of Statistics, Rice University, Houston, Texas. [6]Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, Massachusetts. [7]Radiation Oncology, UT Southwestern Medical Center, Dallas, Texas. [8]Department of Thoracic Radiation Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas. [9]Department of Experimental Radiation Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas. [10]Department of Otolaryngology-Head and Neck Surgery, Yokahama University, Yokahama, Japan. [11]Department of Head and Neck Radiation Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas. [12]Department of Pathology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas. [13]Department of Thoracic/Head and Neck Medical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas. [14]Department of Radiation Oncology, UT Southwestern Medical Center, Dallas, Texas.

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

D.M. Neskey, A.A. Osman, J.N. Myers, and O. Lichtarge contributed equally to this article.

Corresponding Author: Jeffrey N. Myers, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX 77030. Phone: 713-745-2667; Fax: 713-794-4662; E-mail: jmyers@mdanderson.org

doi: 10.1158/0008-5472.CAN-14-2735

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 nonfunctional 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 (wtp53) 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 (γ) and the fitness phenotype (φ) will be related by $\varphi = f(\gamma)$, 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 ($\partial f/\partial \gamma_i$) and the magnitude of the substitution ($\Delta\gamma_i$). The term $\partial f/\partial \gamma_i$ 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 ($\Delta\gamma_i$) 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

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 pBabe retroviral expression vector (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:chr17:7579670_7579709del). 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, 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-diphenyltetrazolium 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

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

		Training set			Validation set					
		Overall survival			Overall survival		Progression-free survival		Time to distant metastases	
Prognostic variable	Total	HR 95% CI	*P* value	Total	HR 95% CI	*P* value	HR 95% CI	*P* value	HR 95% CI	*P* value
Age	168	1.41 (1.03–1.94)	0.0319	96	1.21 (0.88–1.68)	0.2449	1.13 (0.83–1.55)	0.4272	0.65 (0.38–1.10)	0.1055
EAp53 status										
Low-risk (ref)[a]	86	1		64						
High-risk	82	2.25[b]	0.008	32	1.71 (1.02–2.89)	0.0433	1.75 (1.03–2.85)	0.0381	2.10 (0.87–5.07)	0.0998
T stage										
≤ 2 (ref)	47	1		23						
>2	121	1.14 (0.69–1.90)	0.6108	73	1.87 (0.99–3.54)	0.0539	1.74 (0.96–3.16)	0.0700	3.80 (0.88–16.44)	0.0746
N stage										
<2 (ref)	123	1		45						
≥2b	45	1.45 (0.86–2.43)	0.1601	51	2.28 (1.35–3.87)	0.0021	2.20 (1.33–3.63)	0.0020	4.87 (1.63–14.56)	0.0046
Site: larynx										
No (ref)	116	1		84						
Yes	52	0.97 (0.59–1.59)	0.9005	12	0.63 (0.30–1.33)	0.2271	0.74 (0.36–1.51)	0.4083	1.40 (0.46–4.20)	0.5530
Site: pharynx										
No (ref)	160			82						
Yes	8	2.35 (0.94–5.87)	0.0673	14	1.40 (0.75, 2.64)	0.2941	1.20 (0.64–2.25)	0.5646	0.55 (0.13–2.36)	0.4191
Positive margins										
No (ref)	NA			85						
Yes	NA	NA	NA	11	1.55 (0.76–3.16)	0.2272	1.49 (0.73–3.02)	0.2685	1.40 (0.41–4.78)	0.5898
Extracapsular extension										
No (ref)	NA			51						
Yes	NA	NA	NA	45	1.43 (0.87–2.36)	0.1616	1.35 (0.83–2.19)	0.2220	2.72 (1.10–6.76)	0.0310

[a]Low risk is defined at EAp53 score of less than 74.39 in the training set and less than 75 in the validation set and includes wtp53.
[b]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.

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 Biosystem) 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 (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.

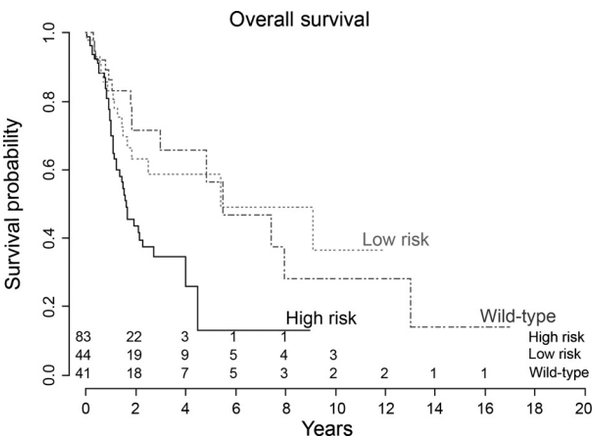

Figure 1.
EAp53 identifies two functionally distinct groups of p53 mutations. Patients with tumors harboring high-risk EAp53 ($n = 83$) mutations have a decreased overall survival relative to low-risk EAp53 ($n = 44$) mutations and wtp53 ($n = 41$). In contrast, patients with low-risk EAp53 mutations appear to have similar survival outcomes to wtp53.

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

Dataset	Characteristic	HR 95% CI	*P* value	Adjusted for	Model likelihood ratio *P* value
Training					
Overall survival	EA score ≥77.78	2.30[a]	0.0087	Age, pharyngeal site	0.0002
Validation					
Overall survival	EA score ≥75	1.79 (1.06–3.04)	0.0300	N stage ≥2b, laryngeal site	0.0002
Disease-free survival	EA score ≥75	1.94 (1.16–3.25)	0.0116	N stage ≥2b	0.0003
Time to distant metastases	EA score ≥75	2.35 (0.97–5.69)	0.0591	N stage ≥2b	0.001

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.

[a]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.

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 ($P = 0.008$; 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 ($P = 0.009$; 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 ($n = 63$), $P = 0.030$ and $P = 0.011$ (Tables 1 and 2). Furthermore, this survival difference appeared to be associated with the development of distant metastases ($P = 0.059$; 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 ≥ 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, and 0.09, 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, and 0.03, 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, and 0.002, respectively; Fig. 3B). In contrast with the increased cellular

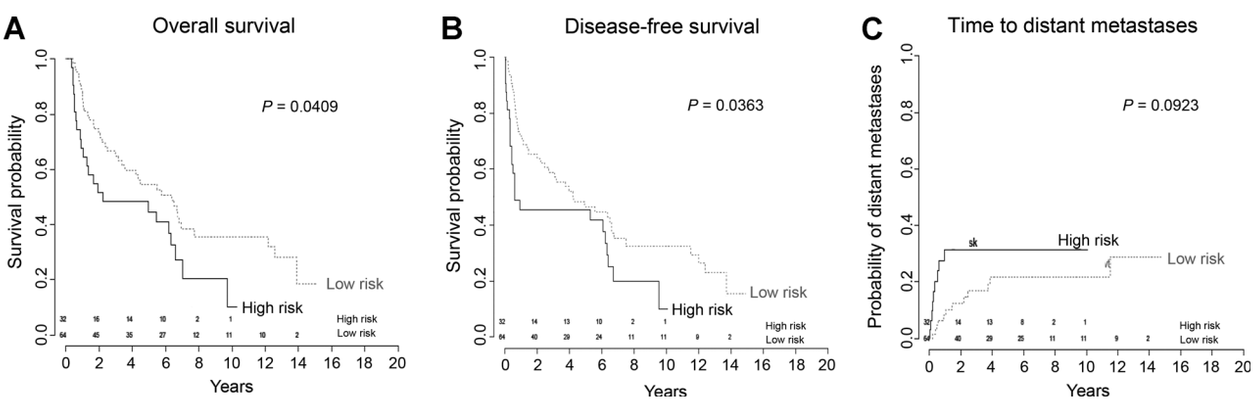

Figure 2.
EAp53 can identify patients with decreased survival outcomes in HNSCC. Log-rank tests of Kaplan–Meir survival plots for a cohort of patients treated with surgery followed by radiotherapy validated that EAp53 can identify patients with high-risk mutations that are associated with a decreased 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).

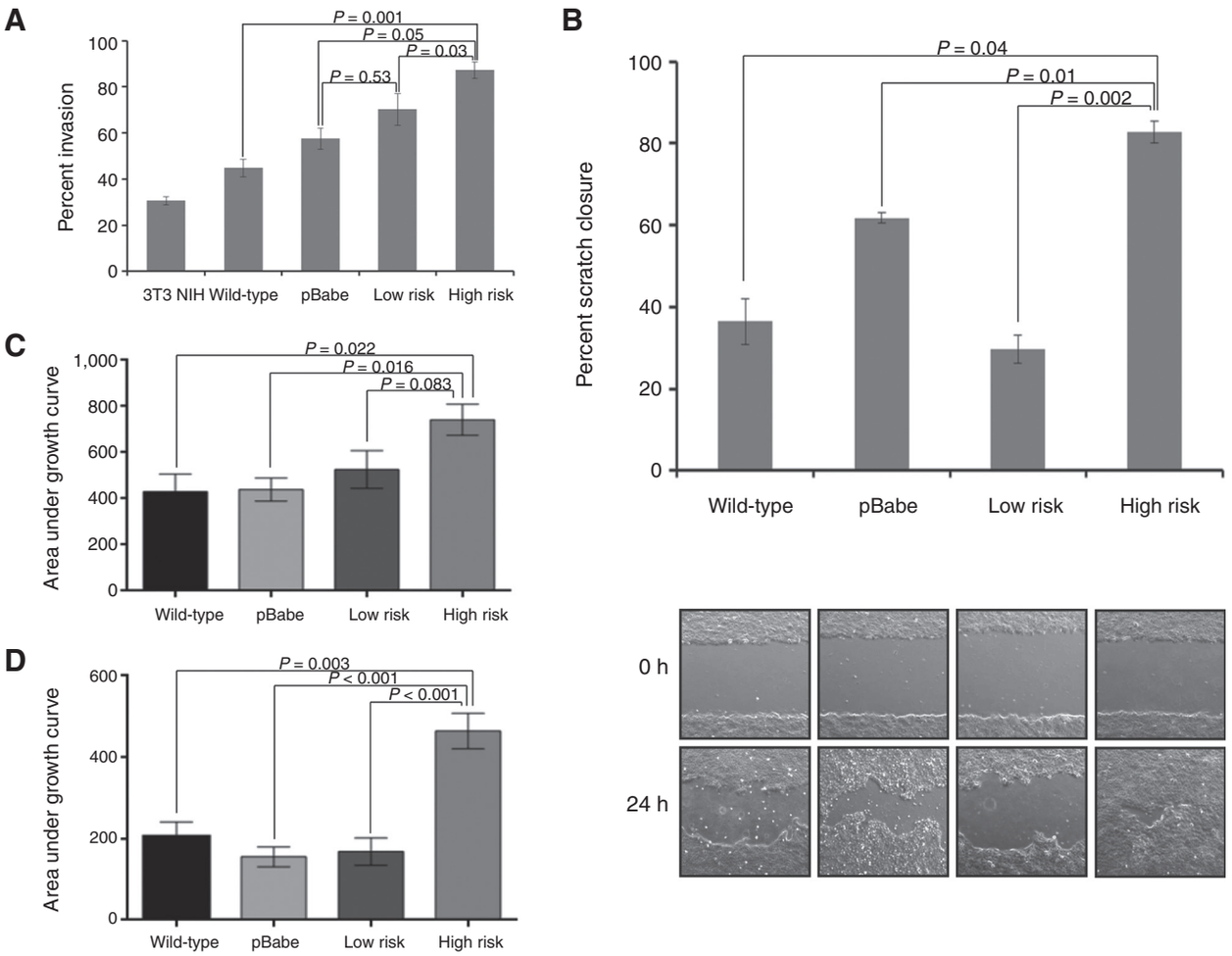

Figure 3.

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 Y236C, 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).

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, or low-risk p53 mutations (Supplementary Table S4; Fig. 3C and D).

High-risk *TP53* 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

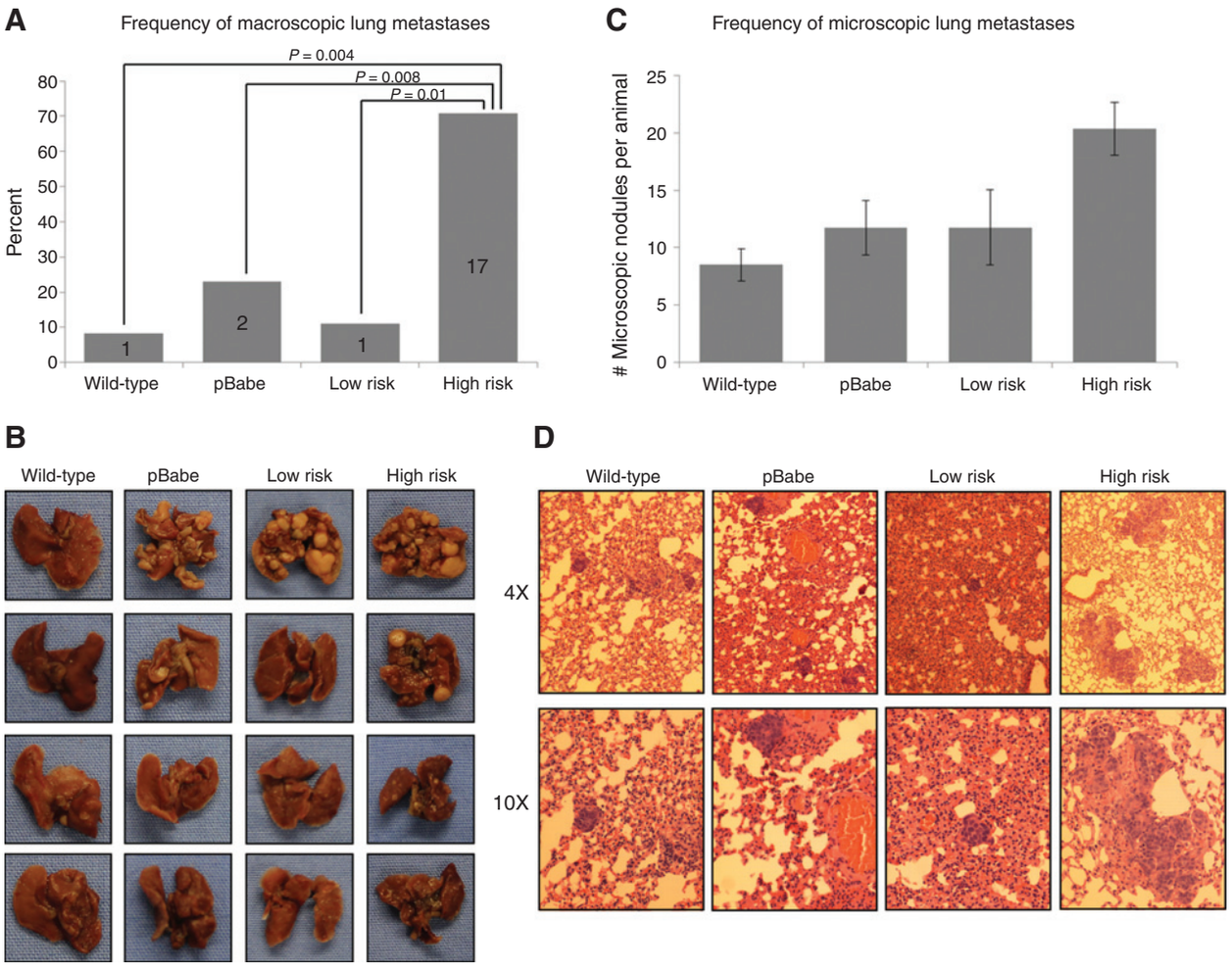

Figure 4.
High-risk *TP53* mutations are associated with increased development of lung metastases. UMSCC1 and PCI13 isogenic cell lines harboring either wtp53, pBabe, low-, or high-risk mutations were injected into tail vein of nude mice. The low- and high-risk series are a composite of two mutations, A161S or Y236C, and three mutations, R175H, C238F, or G245D, respectively, and the results represent two independent experiments (A). Data labels in columns correspond to number of animals in each group. B, macroscopic nodules can be appreciated in the representative images from the lungs of mice top from the pBabe, low-, and high-risk groups. C and D, the frequency of microscopic lung metastases was assessed by hematoxylin and eosin in animals from the wtp53, pBabe, low-risk, and high-risk groups.

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

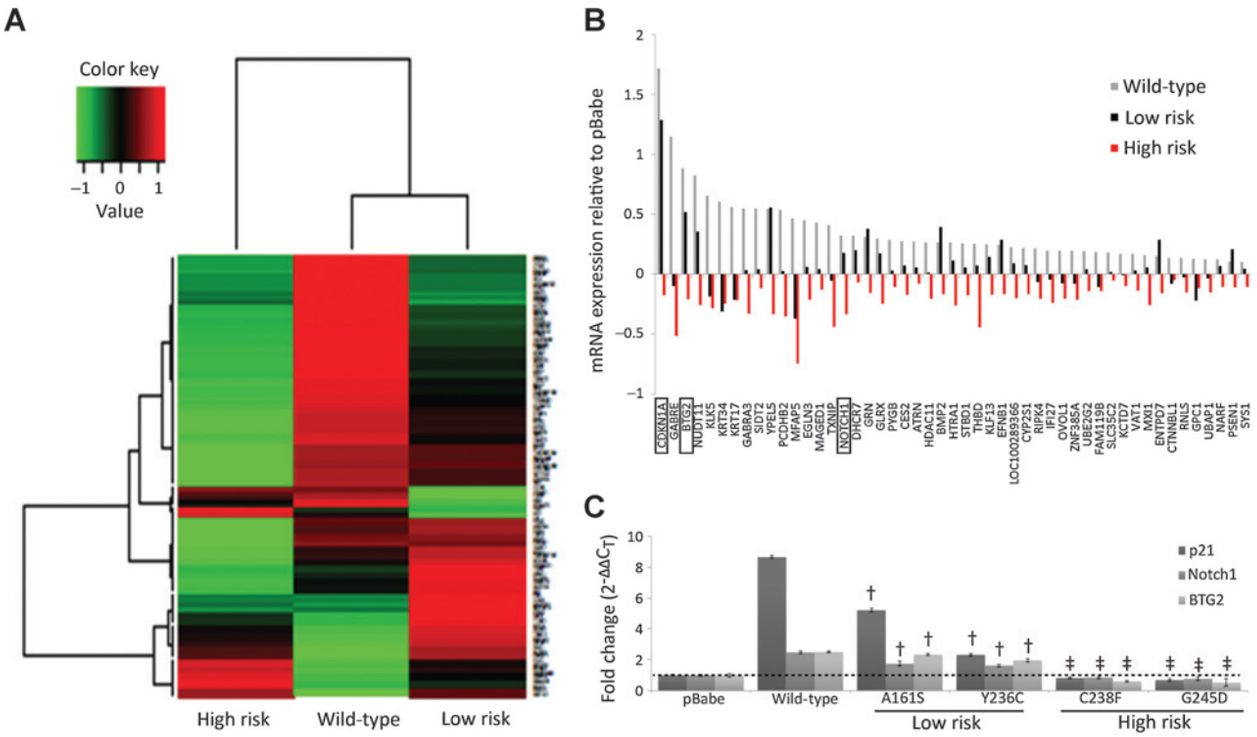

Figure 5.
Suppression of downstream *TP53* target genes by high-risk *TP53* mutations. A, heatmap depicting expression for 102 *TP53* target genes relative to pBabe. B, analysis of the entire expression array relative to the pBabe identified 49 genes that had significantly different levels of expression in both the wild-type and the high-risk mutation. This analysis identified three known *TP53* targets genes, *p21, Notch1*, and *BTG2*, as significantly overexpressed and underexpressed in the wild-type and high-risk mutation, respectively. C, qRT-PCR of *p21, BTG2*, and *Notch1* in wtp53, pBabe, two low-risk mutations, A161s and Y236C, and two high-risk mutations, C238F and G245D. † and ‡, significant overexpression and suppressed expression relative to pBabe, respectively.

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 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
Development of methodology: D.M. Neskey, A.A. Osman, T.J. Ow, P. Katsonis, A. Patel, H.D. Skinner, 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, J.S. Yordy, U. Giri, M.D. Story, M.S. Kies
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): D.M. Neskey, A.A. Osman, P. Katsonis, T. McDonald, S.C. Hicks, T.-K. Hsu, C.R. Pickering, B.M. Beadle, W.N. William, C. Caulin, M. Frederick, M. Kimmel, J.N. Myers, O. Lichtarge
Writing, review, and/or revision of the manuscript: D.M. Neskey, A.A. Osman, T.J. Ow, P. Katsonis, S.C. Hicks, J.S. Yordy, H.D. Skinner, B.M. Beadle, W.N. William, C. Caulin, M. Frederick, 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, J.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 GM066099, and NSF DBI 1356569 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 R90 DKO71505).

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.

References

1. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin 2005;55:74–108.
2. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA Cancer J Clin 2011;61:69–90.
3. Stransky N, Egloff AM, Tward AD, Kostic AD, Cibulskis K, Sivachenko A, et al. The mutational landscape of head and neck squamous cell carcinoma. Science 2011;333:1157–60.
4. Agrawal N, Frederick MJ, Pickering CR, Bettegowda C, Chang K, Li RJ, et al. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. Science 2011;333:1154–7.
5. Petitjean A, Mathe E, Kato S, Ishioka C, Tavtigian SV, Hainaut P, et al. Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database, R16, November 2012. Hum Mutat 2007;28:622–9.
6. Hayes D, Grandis JR, ElNaggar A. Comprehensive genomic characterization of squamous cell carcinoma of the head and neck in the Cancer Genome Atlas [abstract]. In: Proceedings of the Annual Meeting of the American Association for Cancer Research.
7. Erber R CC, Homann N, et al. TP53 DNA contact mutations are selectively associated with allelic loss and have a strong clinical impact in head and neck cancer. Oncogene 1998:1671–79.
8. Poeta ML, Manola J, Goldwasser MA, Forastiere A, Benoit N, Califano JA, et al. TP53 mutations and survival in squamous-cell carcinoma of the head and neck. N Engl J Med 2007;357:2552–61.
9. Koch WM, Brennan JA, Zahurak M, Goodman SN, Westra WH, Schwab D, et al. p53 mutation and locoregional treatment failure in head and neck squamous cell carcinoma. J Natl Cancer Inst 1996;88:1580–6.
10. Lindenbergh-van der Plas M, Brakenhoff RH, Kuik DJ, Buijze M, Bloemena E, Snijders PJ, et al. Prognostic significance of truncating TP53 mutations in head and neck squamous cell carcinoma. Clin Cancer Res 2011;17:3733–41.
11. Gross AM, Orosco RK, Shen JP, Egloff AM, Carter H, Hofree M, et al. Multi-tiered genomic analysis of head and neck cancer ties TP53 mutation to 3p loss. Nature Genetics 2014.
12. Willis A, Jung EJ, Wakefield T, Chen X. Mutant p53 exerts a dominant negative effect by preventing wild-type p53 from binding to the promoter of its target genes. Oncogene 2004;23:2330–8.
13. Brosh R, Rotter V. When mutants gain new powers: news from the mutant p53 field. Nat Rev Cancer 2009;9:701–13.
14. Dittmer D, Pati S, Zambetti G, Chu S, Teresky AK, Moore M, et al. Gain of function mutations in p53. Nat Genetics 1993;4:42–6.
15. Muller PA, Vousden KH. p53 mutations in cancer. Nat Cell Biol 2013;15:2–8.
16. Xie TX, Zhou G, Zhao M, Sano D, Jasser SA, Brennan RG, et al. Serine substitution of proline at codon 151 of TP53 confers gain of function activity leading to anoikis resistance and tumor progression of head and neck cancer cells. Laryngoscope 2013;123:1416–23.
17. Lichtarge O, Wilkins A. Evolution: a guide to perturb protein function and networks. Curr Opin Struct Biol 2010;20:351–9.
18. Madabushi S, Yao H, Marsh M, Kristensen DM, Philippi A, Sowa ME, et al. Structural clusters of evolutionary trace residues are statistically significant and common in proteins. J Mol Biol 2002;316:139–54.
19. Erdin S, Ward RM, Venner E, Lichtarge O. Evolutionary trace annotation of protein function in the structural proteome. J Mol Biol 2010;396:1451–73.
20. Amin SR, Erdin S, Ward RM, Lua RC, Lichtarge O. Prediction and experimental validation of enzyme substrate specificity in protein structures. Proc Natl Acad Sci U S A 2013;110:E4195–202.
21. Katsonis P, Lichtarge O. A formal perturbation equation between genotype and phenotype determines the evolutionary action of protein-coding variations on fitness. Genome Res 2014;24:2050–8.
22. Lichtarge O, Bourne H, Cohen F. An evolutionary trace method defines binding surfaces common to protein families. J Mol Biol 1996;257:342–58.
23. Mihalek I, Res I, Lichtarge O. A family of evolution-entropy hybrid methods for ranking protein residues by importance. J Mol Biol 2004;336:1265–82.
24. Henikoff S, Henikoff J. Amino acid substitution matrices from protein blocks. Proc Natl Acad SciU S A 1992;89:10915.
25. Contal C, O'Quigley J. An application of changepoint methods in studying the effect of age on survival in breast cancer. Comput Stat Data An 1999;30:253–70.
26. Klein JP, Wu JT. Discretizing a continuous covariate in survival studies. In: Balakrishnan N, Rao CR, editors. Handbook of Statistics 23: advances in survival analysis. New York: Elsevier; 2004. p. 27–42.
27. Zhao M, Sano D, Pickering CR, Jasser SA, Henderson YC, Clayman GL, et al. Assembly and initial characterization of a panel of 85 genomically validated cell lines from diverse head and neck tumor sites. Clin Cancer Res 2011;17:7248–64.

28. Pickering CR, Zhang J, Yoo SY, Bengtsson L, Moorthy S, Neskey DM, et al. Integrative genomic characterization of oral squamous cell carcinoma identifies frequent somatic drivers. Cancer Discov 2013;3:770–81.
29. Sano D, Fooshee DR, Zhao M, Andrews GA, Frederick MJ, Galer C, et al. Targeted molecular therapy of head and neck squamous cell carcinoma with the tyrosine kinase inhibitor vandetanib in a mouse model. Head Neck 2011;33:349–58.
30. Sano D, Myers JN. Xenograft models of head and neck cancers. Head Neck Oncol 2009;1:32.
31. Ketcham AS, Wexler H, Minton JP. Experimental study of metastases. JAMA 1966;198:157–64.
32. Parikh N, Hilsenbeck S, Creighton CJ, Dayaram T, Shuck R, Shinbrot E, et al. Effects of TP53 mutational status on gene expression patterns across 10 human cancer types. J Pathol 2014;232:522–33.
33. Katsonis P, Koire A, Wilson SJ, Hsu TK, Lua RC, Wilkins AD, et al. Single nucleotide variations: biological impact and theoretical interpretation. Protein Sci 2014;23:1650–66.
34. Kato S, Han SY, Liu W, Otsuka K, Shibata H, Kanamaru R, et al. Understanding the function-structure and function-mutation relationships of p53 tumor suppressor protein by high-resolution missense mutation analysis. Proc Natl Acad Sci U S A 2003;100:8424–9.
35. Vikhanskaya F, Lee MK, Mazzoletti M, Broggini M, Sabapathy K. Cancer-derived p53 mutants suppress p53-target gene expression–potential mechanism for gain of function of mutant p53. Nucleic Acids Res 2007;35:2093–104.
36. Kakudo Y, Shibata H, Otsuka K, Kato S, Ishioka C. Lack of correlation between p53-dependent transcriptional activity and the ability to induce apoptosis among 179 mutant p53s. Cancer Res 2005;65:2108–14.