High tumor amplification burden is associated with TP53 mutations in the pan-cancer setting

Rushikesh S. Joshi,† Amelie Boichard,† Jacob J. Adashek, and Razelle Kurzrock

†School of Medicine, University of California San Diego, La Jolla, CA, USA; ‡Department of Molecular Cancer Genetics, University of Strasbourg Hospitals, Strasbourg, France; §Department of Oncology, The Sidney Kimmel Comprehensive Cancer Center, The Johns Hopkins Hospital, Baltimore, MD, USA; ¶MCW Cancer Center and Genomic Sciences and Precision Medicine Center, Medical College of Wisconsin, WIS, USA; #Worldwide Innovative Network (WIN) for Personalized Cancer Therapy, Paris, France

ABSTRACT
Next-generation sequencing data is fundamentally changing the clinical management of patients with cancer. The most frequent genomic alterations in malignancy are mutations and amplifications, with a subset of tumors having multiple amplifications – “amplificators”. We sought to understand the molecular correlates of high tumor amplification burden in a pan-cancer context. Using both national registries and a single-institution dataset, our results demonstrate that cancers with TP53 mutations (as compared to those with wild-type TP53) exhibited significantly higher tumor amplification burden across all datasets. Amplifications, generally associated with overexpression, may be potentially actionable secondary consequences of TP53 mutations.

Background
Advances in next-generation sequencing (NGS) techniques have begun to revolutionize our fundamental understanding of disease, especially cancer. The identification of specific genomic anomalies has enabled the development of molecular and immune marker-specific drugs as treatment options for various cancers.1–5 For instance, some patients with high tumor mutational burden appear more responsive to immune checkpoint blockade.1,6

However, the underlying biology driving certain genomic alteration patterns remains to be elucidated. As an example, gene amplification refers to an increase in the number of copies of a specific gene and is a prominent manifestation of genomic instability in mammalian cells.7–9 Gene amplifications are often present in cancer cells and can be the cause of RNA or protein overexpression.8,9 The occurrence of gene amplifications in early stages of cancer and the amplification of multiple genes in some tumors may suggest an underlying genomic etiology.10,11 While the mechanisms behind gene amplification have not been empirically determined, they are generally understood to be the result of DNA double-stranded breaks, impaired DNA replication, or dysfunction in the DNA repair machinery.10 Interestingly, we have observed a group of patients, dubbed ‘amplificators’, who have large numbers of gene amplifications, with or without concomitant large numbers of deleterious mutations.

In this study, we reviewed the medical records of 1,891 patients seen at the University of California, San Diego (UCSD) Moores Center for Personalized Cancer Therapy, and additionally explored 7,246 tumor samples from The Cancer Genome Atlas (TCGA). We show an association between TP53 mutations and a high number of oncogenic gene amplifications. TP53 is a tumor suppressor, designated the “guardian of the genome” because of its crucial role in maintaining genomic integrity (Figure 1).11–16 Although TP53 alterations are considered difficult to drug, their secondary effects, such as amplifications, might be important in that the resultant overexpression levels may be actionable.

Results and discussion
Exploratory analysis was performed to identify patients deemed to have higher proportions of gene amplifications. To define the “amplificator” phenotype, we examined samples that expressed the top 10% of tumor amplification burden, across all samples, and based on two different sequencing panels. The two panels used for sequencing were whole genome sequencing (WGS) and a panel of 315 common oncogenes from the FoundationOne CDx gene panel by Foundation Medicine (FM) (https://www.foundationmedicine.com). A total of 7,246 patient samples were included from TCGA and 1,891 patients treated at Moores Cancer Center at UCSD and sequenced by FM.

In Table 1, the differences between amplificators and non-amplificators, as determined by both WGS and the FM gene panel were tabulated for the TCGA cohort, while the UCSD cohort amplificator phenotype was determined solely by the genes in the FM panel. Within the TCGA cohort, amplificators, as defined by both WGS and the FM panel, exhibited a higher
The tumor suppressor *TP53* has long been implicated in the development of diverse cancers. It is the most commonly mutated gene in cancer and has diverse functions important to oncogenesis.\(^\text{12-17}\) Unfortunately, *TP53* alterations are considered difficult to target from a therapeutic standpoint.\(^\text{12-17}\) Groups have, however, reported an increase in the expression of vascular endothelial growth factor (VEGF) in a pan-cancer analysis as well as improvement in outcome of patients who receive VEGF/VEGFR inhibitors when their tumors harbor deleterious *TP53*
Table 1. Amplifiers and non-amplifiers in both the TCGA and UCSD datasets.

| Information for TCGA and for UCSD Cancer Cohort |
|-----------------------------------------------|
| Total samples N (% total) | TP53-mutated samples N (% total) | Number of WGS amplifications per sample, mean [CI95%] | Number of FM-panel amplifications per sample, mean [CI95%] |
| TCGA Data* |
| All samples | 7246 (100) | 2897 (40) | 344 [331–357] | 5.1 [4.9–5.3] |
| Non-amplifiers (WGS) | 6519 (100) | 2448 (38) | 190 [184–197] | 3.2 [3.1–3.3] |
| WGS amplifiers | 727 (100) | 449 (62) | 1720 [1662–1778] | 22.4 [21.5–23.2] |
| Non-amplifiers (FM) | 6514 (100) | 2416 (37) | 207 [199–215] | 3.0 [2.9–3.1] |
| FM-panel amplifiers | 732 (100) | 481 (66) | 1562 [1503–1621] | 24.0 [23.2–24.7] |
| UCSD** |
| All samples | 1891 (100) | 824 (44) | - | 2.5 [2.2–2.7] |
| Non-amplifiers (FM) | 1706 (100) | 707 (41) | - | 1.2 [1.1–1.3] |
| FM-panel amplifiers | 185 (100) | 117 (63) | - | 14.2 [13.4–15.0] |

Amplifications in TP53 mutated vs. wild-type samples in the TCGA and UCSD datasets

| Total samples N (% total) | Number of WGS amplifications per sample Mean [CI95%] | Number of FM panel amplifications per sample Mean [CI95%] | p-value |
|--------------------------|---------------------------------|---------------------------------|---------|
| TCGA Data* |
| TP53 mutated samples | 2897 (40%) | 516 [491–542] | 6.58E-83 | 7.7 [7.3–8.0] | 9.28E-96 |
| TP53 wild-type samples | 4349 (60%) | 281 [958–958] | 229 [215–243] | 7 [0–6981] | 3.4 [3.2–3.6] | 0 [0–73] | 9.28E-96 |
| UCSD** |
| TP53 mutated samples | 824 (44%) | - | - | 3.3 [3.0–3.7] | 3.39E-13 |
| TP53 wild-type samples | 1067 (56%) | - | - | 1.8 [1.5–2.0] | 0 [0–38] | 3.39E-13 |

In the TCGA data, a total of 7,246 samples that had copy number variation (CNV) and mutation data were curated from 11,245 possible TCGA samples across all cancer cohorts.

1The phenotype “WGS amplificator” corresponds to tumors presenting a high number of amplifications considering the whole genome (top 10% amplification burden, within the whole genome): All p < 0.0001.

2The phenotype “FM-panel amplificator” corresponds to tumors presenting a high number of amplifications considering only genes included in the Foundation One panel (top 10% amplification burden, within the 315 genes of the Foundation One panel manufactured by Foundation Medicine): All p < 0.0001.

**In the UCSD data, a total of 1,891 samples sequenced from patients treated at Moores Cancer Center in La Jolla, CA, were analyzed across all cancer types.

Abbreviations: CI95% = 95% confidence interval; FM = Foundation Medicine; N = number; TCGA = The Cancer Genome Atlas; UCSD = University of California San Diego; WGS = whole genome sequencing.

alterations as a possible therapeutic proxy for targeting harmful TP53 alterations.13,15,17

Importantly, mutations in the TP53 tumor suppressor increase genomic instability, corroborating the reputation of TP53 as the “guardian of the genome.”12 Interestingly, TP53 also likely plays a role in transcriptional regulation.18 This inherent process of upregulating and downregulating various aspect of mRNA production may directly impact the tumor suppression function of this protein.18,19 It may be that TP53 impacts mRNA expression through many TP53-dependent pathways that directly impact transcriptional regulation and via indirect transcriptional regulation (such as by virtue of secondary amplifications) as well as through posttranscriptional regulation.18

Our data suggest that a subset of cancers have high tumor amplification burden, and these tumors are significantly more likely to bear TP53 mutations than those with lower tumor amplification burden. In contrast, BRAF and KRAS alterations correlated with decreased tumor amplification burden in the TCGA dataset. A limitation of our findings is that it is unclear why BRAF and KRAS alterations would correlate with a decreased tumor amplification burden. It is also unclear why specific tumor types such as breast cancer and ovarian serous carcinomas are especially likely to have an amplificator phenotype, though the latter could be due to the fact that high-grade ovarian serous carcinomas demonstrate TP53 anomalies in about 90% of cases.20 TP53 mutations may correlate with high tumor amplification burden because these mutations impair genomic stability as evidenced by loss-of-function TP53 mutations shown to be associated with increased mutation rate.21 Since amplifications (which generally [but not always] cause overexpression)22 may be pharmacologically tractable in some cases, targeting them may be an indirect way to impact the consequences of TP53 mutation-related genomic instability.

Methods

Two distinct datasets were used for the statistical analysis. The first dataset was retrieved from the publicly available
repository, The Cancer Genome Atlas (TCGA) (https://portal.gdc.cancer.gov/), which is a cohort of sequenced cancer samples from patients. Our second dataset was composed of a cohort of patients who had been treated and sequenced using the FoundationOne CDx gene panel (Foundation Medicine, Inc., Cambridge, MA) (https://corpsite.foundationmedicine.com/genomic-testing) (Supplemental Table S1) at UCSD. All studies were conducted under the auspices of an Internal Review Board (IRB) Committee-approved protocol (NCT02478931) and any investigational trials for which the patient gave consent.

Data collected and reviewed retrospectively for this study from the UCSD cohort included genomic information from sequencing results detailing amplifications and mutations present (with gene localization) across all cancer types and using the genes in the FM panel. Data collected from TCGA included demographic information such as age (years), sex, primary cancer diagnosis, number of amplifications, and mutation status for all of the FM panel genes. Within the TCGA dataset, patient samples from all cancer types were queried.

Descriptive statistics were tabulated to describe patient sample information, comparing our amplificators to non-amplificators in both the TCGA cohort, as well as our institutional dataset. For the TCGA cohort, statistical summaries were stratified into sub-groups based on amplificator vs. non-amplificator phenotype, as determined by both WGS and the FM gene panels (Supplemental Tables S2 and S3). Descriptive information included number of patients matching the criteria for amplificator or non-amplificator phenotype, total number of TPS3 mutated samples present in each sub-group, average number of WGS amplifications per sample, and average number FM gene amplifications per sample. Similar analysis was conducted in the UCSD cohort; however, amplificators were determined solely based on number of amplifications present in the FM gene panel as WGS was not utilized for these patients. Additionally, for both cohorts, a second analysis was conducted to describe summary statistics based on TPS3 mutation status. Patient samples were stratified into TPS3 mutant and TPS3 wild-type sub-groups. Within these two sub-groups, the average number of WGS amplifications and FM-panel amplifications per sample was determined and compared using student’s t-tests. Similar analysis was conducted for the UCSD dataset comparing only number of FM panel amplifications per sample in the TPS3 mutant and wild-type sub-groups.

Further analysis was performed to assess the number of mutations in common cancer genes (Supplemental Table S1) across samples with the top 10% of amplifications against samples with the bottom 90% of amplifications. The top four

| Table 2. Comparing frequency of mutated genes in samples with amplification burden in the top 10% of samples vs. those in the bottom 90% using WGS for the FM panel (see also Supplemental Table S1 for the FM gene list) for both TCGA and UCSD datasets. |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| Gene            | Top 10% of amplifications | Bottom 90% of amplifications | Odds Ratio (CI) | Bonferroni adjusted p-value |
| TPS3****        | 449/727 (61.8%)         | 2448/6519 (37.6%) | 2.7 (2.3–3.1)   | <0.00001 |
| BRAF            | 17/727 (2.3%)           | 591/6519 (9.1%)  | 0.2 (0.1–0.4)   | <0.0001  |
| KRAS            | 34/727 (4.7%)           | 611/6519 (9.4%)  | 0.5 (0.3–0.7)   | 0.008     |
| GATA3           | 29/727 (4.0%)           | 159/6519 (2.4%)  | 1.7 (1.1–2.5)   | Not significant |
| TCGA dataset comparing frequency of mutated genes in samples with amplifications in the top 10% vs. those in the bottom 90% using WGS* |
| TPS3****        | 481/732 (65.7%)         | 2416/6514 (37.1%) | 3.3 (2.8–3.8)   | <0.003 |
| BRAF            | 16/732 (2.2%)           | 592/6514 (9.1%)  | 0.2 (0.1–0.4)   | <0.003  |
| KRAS            | 25/732 (3.4%)           | 620/6514 (9.5%)  | 0.3 (0.2–0.5)   | <0.003  |
| GATA3           | 32/732 (4.4%)           | 156/6514 (2.4%)  | 1.9 (1.3–2.7)   | Not significant |
| UCSD dataset comparing frequency of mutated genes in samples with amplifications in the top 10% vs. those in the bottom 90% using FM panel*** |
| TPS3****        | 117/185 (63.2%)         | 707/1706 (41.4%) | 2.4 (1.8–3.3)   | <0.003 |
| BRAF            | 4/185 (2.2%)            | 107/1706 (6.3%)  | 0.3 (0.1–0.9)   | Not significant |
| KRAS            | 32/185 (17.3%)          | 259/1706 (15.2%) | 1.2 (0.8–1.7)   | Not significant |
| GATA3           | 11/185 (5.9%)           | 35/1706 (2.1%)   | 3.0 (1.5–6.0)   | Not significant |

*Using the TCGA cohort, which consists of 7,246 cancer samples, the 90th percentile for number of amplifications using WGS was calculated to be 1,016. The 90th percentile for number of amplifications using the FM panel was calculated to be 15.

** The FM panel includes 321 genes (see supplemental Table S1).

*** Using the UCSD dataset, which consists of 1,891 cancer samples, the 90th percentile for number of amplifications (FM panel) was calculated to be 9.

**** Requires to number of samples with designated gene mutation/total samples in that subgroup. For instance, within TCGA database, 727 samples were in the top 10% for amplification burden (“amplificators”); of these 727 samples, 449 (65.7%) had a TPS3 mutation.

Abbreviations: FM = Foundation Medicine; TCGA = The Cancer Genome Atlas; UCSD = University of California San Diego; WGS = whole genome sequencing.
genes reported in our analysis were TP53, BRAF, KRAS, and GATA3. Odds ratios and Bonferroni adjusted p-values were then calculated to compare mutation burden between the two amplification sub-groups. This analysis was performed using both the WGS and FM information for the TCGA dataset and only FM-panel genes for the UCSD dataset.

For the TCGA dataset, copy numbers were measured using whole-genome microarray. Gene-level focal copy number variations (CNVs) were normalized using data from all TCGA cohorts (« pan-cancer » data set) and estimated using the GISTIC2 threshold method,23 where the values −2, −1, 0, 1, and 2 represented homozygous deletion, single-copy deletion, diploid normal copy, low-level amplification, and high-level amplification, respectively. Only high-level amplification (+2) was considered for this analysis.

For the UCSD dataset, copy numbers were measured using gene-panel capture sequencing (FoundationOne CDx test, Foundation Medicine, Inc.) and gene-level amplifications were reported when the number of copies exceeded 6.

All statistical analysis was conducted using a combination of Microsoft Excel version 16.42 (Microsoft Corporation, Redmond, Washington, USA) and R version 3.6.1 (R Foundation for Statistical Computing, Vienna, Austria).

Acknowledgments

Funded in part by the Joan and Irwin Jacobs Fund and by National Cancer Institute grants P30 CA023100 (RK). Figure 1 created with BioRender.com.

Author contributions

All authors reviewed and approved the final version of this manuscript.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Disclosure statement

RJ and AB have no disclosures or conflicts of interest. JJA serves on the advisory board of CureMatch Inc. RK has received research funding from Biological Dynamics, Boehringer Ingelheim, Debiopharm, Foundation Medicine, Genentech, Grifols, Guardant, Incyte, Konica Minolta, Medimmune, Merck Serono, Omnisq, Pfizer, Sequenom, Takeda, and TopAlliance; as well as consultant and/or speaker fees and/or advisory board for Actuate Therapeutics, AstraZeneca, Bicara Therapeutics, Biological Dynamics, EISAI, EOM Pharmaceuticals, Ilyon, Merck, NeoGenomics, Neomed, Pfizer, Prosprdtx, Roche, TD2/Volatra, Turning Point Therapeutics, X-Biotech; has an equity interest in CureMatch Inc., CureMetrix, and iDByDNA; serves on the Board of CureMatch and CureMetrix; and is a co-founder of CureMatch.

Ethics approval and consent to participate

All studies were conducted with Institutional Review Board Committee approval under IRB NCT02478931 (PREDICT).

Funding

This work was supported by the National Cancer Institute [P30 CA023100].

ORCID

Jacob J. Adashek https://orcid.org/0000-0003-4272-312X
Razelle Kurzrock https://orcid.org/0000-0003-4110-1214

References

1. Goodman AM, Kato S, Bazhenova L, Patel SP, Frampton GM, Miller V, Stephens PJ, Daniels GA, Kurzrock R. Tumor mutational burden as an independent predictor of response to immunotherapy in diverse cancers. Mol Cancer Ther. 2017;16(11):2598–2608. doi:10.1158/1535-7163.MCT-17-0386.
2. Maciejko L, Smalley M, Goldman A. Cancer immunotherapy and personalized medicine: emerging technologies and biomarker-based approaches. J Mol Biomark Diagn. 2017;8(5). doi:10.4172/2155-9929.1000350.
3. Adashek JJ, Subbiah V, Kurzrock R. From tissue-agnostic to n-of-one therapies: (r)evolution of the precision paradigm. Trends Cancer. 2021;7(1):15–28. doi:10.1016/j.trecan.2020.08.009.
4. Kato S, Kim KH, Lim HJ, Boichard A, Nikanjam M, Weihe E, Kuo DJ, Eskander RN, Goodman A, Galanina N, et al. Real-world data from a molecular tumor board demonstrates improved outcomes with a precision n-of-one strategy. Nat Commun. 2020;11(1):4965. doi:10.1038/s41467-020-18613-3.
5. Sicklick JK, Kato S, Okamura R, Patel H, Nikanjam M, Fanta PT, Hahn ME, De P, Williams C, Guido J, et al. Molecular profiling of advanced malignancies guides first-line n-of-1 treatments in the i-predict treatment-naïve study. Genome Med. 2021;13(1):155. doi:10.1186/s13073-021-00969-w.
6. Jardim DL, Goodman A, de Melo Gagliato D, Kurzrock R. The challenges of tumor mutational burden as an immunotherapy biomarker. Cancer Cell. 2021;39(2):154–173. doi:10.1016/j.ccell.2020.10.001.
7. Debatisse M, Malfoy B. Gene amplification mechanisms. Adv Exp Med Biol. 2005;570:343–361.
8. Smith KA, Agarwal ML, Chernov MV, Chernova OB, Deguchi Y, Ishizaka Y, Patterson TE, Poupon MF, Stark GR. Regulation and mechanisms of gene amplification. Philos Trans R Soc Lond B Biol Sci. 1995;347(1319):49–56.
9. Matsui A, Ihara T, Suda H, Mikami H, Semba K. Gene amplification: mechanisms and involvement in cancer. Biomol Concepts. 2013;4(6):567–582. doi:10.1515/bmc-2013-0026.
10. Albertson DG. Gene amplification in cancer. Trends Genet. 2006;22(8):447–455. doi:10.1016/j.tree.2006.06.007.
11. Andor N, Graham TA, Jansen M, Xia LC, Aktipis CA, Petritsch C, Ji HP, Maley CC. Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. Nat Med. 2016;22(1):105–113. doi:10.1038/nm.3984.
12. Donehower LA, Soussi T, Korkut A, Liu Y, Schultz A, Cardenas M, Li X, Babur O, Hsu TK, Lichtarge O, et al. Integrated analysis of tp53 gene and pathway alterations in the cancer genome atlas. Cell Rep. 2019;28(11):3010. doi:10.1016/j.celrep.2019.08.061.
13. Li AM, Boichard A, Kurzrock R. Mutated tp53 is a marker of increased VEGF expression: analysis of 7,525 pan-cancer tissues. Cancer Biol Ther. 2020;21(1):95–100. doi:10.1089/cbi.2019.1665956.
14. Nigro JM, Baker SJ, Preisinger AC, Jessup JM, Hostetter R, Cleafy K, Bigner SH, Davidson N, Baylin S, Devilee P, et al. Mutations in the p53 gene occur in diverse human tumour types. Nature. 1989;342(6250):705–708. doi:10.1038/342705a0.
15. Wheler JJ, Janku F, Naing A, Li Y, Stephen B, Zinner R, Subbiah V, Fu S, Karp D, Falchook GS, et al. TP53 alterations correlate with response to VEGF/vegfr inhibitors: implications for targeted
therapeutics. Mol Cancer Ther. 2016;15(10):2475–2485. doi:10.1158/1535-7163.MCT-16-0196.

16. Hollstein M, Rice K, Greenblatt MS, Soussi T, Fuchs R, Sorlie T, Hovig E, Smith-Sorensen B, Montesano R, Harris CC. Database of p53 gene somatic mutations in human tumors and cell lines. Nucleic Acids Res. 1994;22(17):3551–3555.

17. Leslie KK, Filiaci VL, Mallen AR, Thiel KW, Devor EJ, Moxley K, Richardson D, Mutch D, Secord AA, Tewari KS, et al. 2021. Mutated p53 portends improvement in outcomes when bevacizumab is combined with chemotherapy in advanced/recurrent endometrial cancer: an nrg oncology study. Gynecol Oncol. 161(1):113–121.

18. Sullivan KD, Galbraith MD, Andrysik Z, Espinosa JM. Mechanisms of transcriptional regulation by p53. Cell Death Differ. 2018;25(1):133–143. doi:10.1038/cdd.2017.174.

19. Adashek JJ, Kato S, Parulkar R, Szeto CW, Sanborn JZ, Vaske CJ, Benz SC, Reddy SK, Kurzrock R. Transcriptomic silencing as a potential mechanism of treatment resistance. JCI Insight. 2020;5(11). doi:10.1172/jci.insight.134824.

20. Ahmed AA, Etemadmoghadam D, Temple J, Lynch AG, Riad M, Sharma R, Stewart C, Fereday S, Caldás C, Defazio A, et al. Driver mutations in tp53 are ubiquitous in high grade serous carcinoma of the ovary. J Pathol. 2010;221(1):49–56. doi:10.1002/path.2696.

21. Burtness B, Deneke A, Baca Y, Serebriiskii I, Parker MI, Nicolas E, Lee JW, Wise-Draper T, Sukari A, Somer BG, et al. Correlation of tumor mutational burden (tmb) with cdkn2a and tp53 mutation in HPV-negative head and neck squamous cell carcinoma (hnsc). J Clin Oncol. 2020;38(15_suppl):6552. doi:10.1200/JCO.2020.38.15_suppl.6552.

22. Boichard A, Lippman SM, Kurzrock R. Therapeutic implications of cancer gene amplifications without mRNA overexpression: silence may not be golden. J Hematol Oncol. 2021;14(1):201. doi:10.1186/s13045-021-01211-1.

23. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. Gistic2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol. 2011;12(4):R41. doi:10.1186/gb-2011-12-4-r41.