Title: The importance of ancestry to understanding tumor mutation burden in cancer

Authors: Alexandra Snyder*[1], Nicholas McGranahan[2]

Affiliations:

[1] Generate Biomedicines, Boston, MA, USA. asnyder@generatebiomedicines.com

[2] Cancer Research UK Lung Cancer Centre of Excellence, University College London Cancer Institute, London, UK. Cancer Genome Evolution Research Group, University College London Cancer Institute, London, UK.

Abstract: In the current issue of *Cancer Cell,* Nasser and colleagues find that in solid tumors, tumor-only sequencing leads to an overestimate of the biomarker tumor mutation burden (TMB), particularly in patients of African or Asian ancestry. Correction of the TMB estimate improves the correlation between TMB and response to immunotherapy.

Text:

Molecularly-informed treatments of advanced cancers have led to impressive clinical benefit in a subset of patients; however, long-term survival and cure remain elusive, particularly in patients with advanced solid tumors. Molecular alterations that are tumor type agnostic and can easily be measured provide an opportunity to test therapies across histologies and offer the promise of access to novel therapies. To this end, the field has seen pan-tumor FDA approvals for patients whose tumors harbor *NTRK* fusions, microsatellite instability or high tumor mutation burden (TMB-H). TMB-H describes a subset of tumors that harbor a high number of non-synonymous variants per megabase of tumor sequenced DNA. The hypothesis that TMB-H tumors are more likely to be recognized and rejected by the immune system and thus more likely to respond to immunotherapy with checkpoint blockade was rooted in preclinical evidence (Matsushita et al., 2012) and tested across multiple tumor types (Cristescu et al., 2018, Litchfield et al, 2021, Samstein et al., 2019, Yarchoan et al., 2017). While the pan-tumor and tumor-specific cut-off to define TMB-H remain areas of debate (Chan et al., 2019), as a first attempt at translating this concept to patient benefit, in 2020 the FDA approved the treatment with the anti-PD-1 agent pembrolizumab for patients with unresectable or metastatic TMB-H (≥10 mutations/megabase) solid tumors that have progressed following prior treatment and who have no satisfactory alternative treatment options (Marcus et al., 2021).

In this issue of *Cancer Cell,* Nasser and colleagues **(*insert citation when ready*)** hypothesize that the absence of germline sequence in tumor-only panels could adversely impact the estimation of TMB, leading to an overestimation that particularly impacts patients whose ancestries are underrepresented in current databases. The authors start with an evaluation of 8,193 patients from the DFCI/PROFILE cohort, highlighting that a proportion of patients' self-reported ancestry is either missing or does not necessarily match the DNA sequencing based inferred ancestry. Next, they perform a simulated comparison of whole-exsome sequencing (WES) to a tumor-only panel by comparing the TMB called for 3,618 TCGA samples versus application of the methods used for panel testing (Oncopanel). This analysis suggests that TMB is inflated in the simulated panel test in all cases, but this inflation is greater in non-Europeans.

To further explore this concept, across >120,000 patients whose tumors had undergone either tumor-only or tumor plus germline panel sequencing, the authors find again that a greater number of patients of non-European ancestry featured tumor alterations that are suspected to represent germline variants. Some limitations of this portion of the analysis includes the lack of complete overlap between genes measured in different panels, which the authors control for by

per-variant variant allele frequency (VAF) distributions, as well as self-reporting of race, an imperfect proxy for ancestry. In a subset of 498 patients whose tumors had undergone tumor plus germline panel testing, which is used as the ground truth, they find that tumor-only estimation of TMB again shows an inflation of TMB values more pronounced in non-Europeans. The authors apply these learnings back to the original DFCI/PROFILE cohort to develop a metric they call corrected TMB, or TMB-c, the application of which removes the putative inflation of TMB in 5 out of 7 examined cancers.

Finally, the authors apply these concepts to clinical outcomes. Noting that all patients are predicted to have some “TMB inflation” using tumor-only methods, patients of Asian or African ancestry would be predicted to more than patients of European ancestry. To correct for this, the authors develop a system to categorize tumors as “true” or “false” for TMB-high or low; for example, a “false TMB-high” tumor is one that is erroneously called as having a high TMB, despite likely having a low TMB in reality. Tumors from 1,840 patients treated with ICI at DFCI with seven cancer types and 234 patients with NSCLC treated with ICI at MSKCC are classified thus and their outcomes examined. The true TMB-high patients experience an overall survival approximately one year longer than those patients with false TMB-high or true TMB-low. Relevant to the current approval, it is important to note that no patients are observed who are misclassified as TMB-low, i.e. exceptionally few tumors are incorrectly called as TMB-low. This implies that few patients miss the opportunity to benefit from a checkpoint inhibitor due to erroneous TMB classification, although some patients may be given treatment that is unlikely to succeed, a concept considered when deciding what TMB cut-off should be advanced for clinical use (Marcus et al., 2021).

Importantly, one of the limitations that motivates this study – the relative lack of representation of diverse groups in clinical trials – impacts the application of the concept of corrected TMB (TMB-c) to this cohort: while TMB-c correlates with outcomes in patients of European ancestry, it does not in those patients with Asian or African ancestry, who represent a small minority of the total cohort.

The relevance of this work to the field can be understood in the context of methods for measuring TMB clinically, and the datasets that underpin the existing approval for checkpoint blockade. TMB can be calculated using several methods (Fig. 1). The gold standard TMB calculation involves using whole exome sequencing (WES) of both tumor and matched normal tissue. In this case, somatic mutations are determined by using the tumor sample mapped to a standard reference genome, and any germline mutations are removed by virtue of also being present in the matched normal tissue. However, given that WES is not routinely used in the

clinic due to both cost and time-constraints, manufacturers and academic groups have developed gene-panels that seek to measure TMB both with and without matched normal tissue. In the context of tumor-only panels, bioinformatic techniques are required to distinguish somatic from germline mutations and to obtain an estimate of tumor TMB; in particular, databases of normal tissue sequences are used to remove common germline mutations. Notably, these databases are predominantly derived from analysis of genomes from patients of European ancestry. Likewise, the standard reference genome used to map sequence reads and call mutations is based on a European ancestry genome. Therefore, as demonstrated by Nasser and colleagues, further work is needed to globalize the current databases of genomic sequence data, to permit a detailed exploration of the interplay between germline ancestry, somatic alterations and response to treatment.

The approval for the treatment of TMB-H tumors with checkpoint blockade is rooted in datasets involving both types of assays. Using the FoundationMedicine tumor-only panel (F1CDx), Marabelle and colleagues reported an overall response rate (ORR) of 29% among 790 patients with diverse solid tumors treated with pembrolizumab (Marabelle et al., 2020). Cristescu and colleagues performed WES on 1,772 tumors from pembrolizumab-treated patients with 24 tumor types and found an ORR of 31.4% among TMB-H patients (Marcus et al., 2021, Cristescu et al., 2022). However, patient ancestry was not reported for either study, so it is unknown how ancestry may have impacted TMB estimates.

As novel therapies based on genomic testing become part of standard treatments, to bring these advancements to patients in need, the field must both improve access to care and by deepen our understanding of how these assays apply to diverse populations. This is particularly relevant given that alterations are commonly assessed using tumor-only and circulating tumor DNA (ctDNA)-based assays. As all therapies come with associated risks, the correct categorization of TMB-H and other molecular alterations is essential. Nasser and colleagues highlight the importance of this issue; coupled with efforts by diagnostic test-providers, this work should contribute to improvement in the field's ability to bring the best therapeutic options to patients.

Acknowledgements

N.M is a Sir Henry Dale Fellow, jointly funded by the Wellcome Trust and the Royal Society (Grant Number 211179/Z/18/Z), and also receives funding from Cancer Research UK Lung

Cancer Centre of Excellence, Rosetrees, and the NIHR BRC at University College London Hospitals

Declaration of Interests

N.M. has received consultancy fees and has stock options in Achilles Therapeutics. N.M.. holds European patents relating to targeting neoantigens (PCT/EP2016/ 059401), identifying patient response to immune checkpoint blockade (PCT/ EP2016/071471), determining HLA LOH (PCT/GB2018/052004), predicting survival rates of patients with cancer (PCT/GB2020/050221).

A.S. was previously an employee of, and still holds stock in, Merck, Inc. A.S. is currently an employee of Generate Biomedicines.

REFERENCES

Chan T.A., Yarchoan M., Jaffee E., Swanton C., Quezada SA, Stenzinger A., Peters S., Ann Oncol. (2019). Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. Jan 1;30(1):44-56. doi: 10.1093/annonc/mdy495.

Cristescu R., Aurora-Garg D., Albright A., Xu L., Liu X.Q., Loboda A., Lang L., Jin F., Rubin E.H., Snyder A., Lunceford J. (2022). Tumor mutational burden predicts the efficacy of pembrolizumab monotherapy: a pan-tumor retrospective analysis of participants with advanced solid tumors. J Immunother Cancer. Jan;10(1):e003091. doi: 10.1136/jitc-2021-003091.

Cristescu R., Mogg R., Ayers M., Albright A., Murphy E., Yearley J., Sher X., Liu X.Q., Lu H., Nebozhyn M., Zhang C., Lunceford J.K., Joe A., Cheng J., Webber A.L., Ibrahim N., Plimack E.R., Ott P.A., Seiwert T.Y., Ribas A., McClanahan T.K., Tomassini J.E., Loboda A., Kaufman D. (2018). Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. Science. Oct 12;362(6411):eaar3593. doi: 10.1126/science.aar3593.

Litchfield K., Reading J.L., Puttick C., Thakkar K., Abbosh C., Bentham R., Watkins T.B.K., Rosenthal R., Biswas D., Rowan A., Lim E., Al Bakir M., Turati V., Guerra-Assunção J.A., Conde L., Furness A.J.S., Saini S.K., Hadrup S.R., Herrero J., Lee S.H., Van Loo P., Enver T., Larkin J., Hellmann M.D., Turajlic S., Quezada S.A., McGranahan N., Swanton C. (2021) Meta-analysis of tumor- and T cell-intrinsic mechanisms of sensitization to checkpoint inhibition. Cell. 2021 Feb 4;184(3):596-614.e14. doi: 10.1016/j.cell.2021.01.002. Epub 2021 Jan 27

Marabelle A., Fakih M., Lopez J., Shah M., Shapira-Frommer R., Nakagawa K., Chung H.C., Kindler H.L., Lopez-Martin J.A., Miller W.H. Jr, Italiano A., Kao S., Piha-Paul S.A., Delord J.P., McWilliams R.R., Fabrizio D.A., Aurora-Garg D., Xu L., Jin F., Norwood K., Bang Y.J. . (2021) Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study. Lancet Oncol 21(10):1353-1365. doi: 10.1016/S1470-2045(20)30445-9. Epub 2020 Sep 10.

Marcus L., Fashoyin-Aje L.A., Donoghue M., Yuan M., Rodriguez L., Gallagher P.S., Philip R., Ghosh S., Theoret M.R., Beaver J.A., Pazdur R., Lemery S.J. (2021) FDA Approval Summary: Pembrolizumab for the Treatment of Tumor Mutational Burden-High Solid Tumors. Clin Cancer Res. Sep 1;27(17):4685-4689. Doi: 10.1158/1078-0432.CCR-21-0327. Epub 2021 Jun 3.

Matsushita H., Vesely M.D., Koboldt D.C., Rickert C.G., Uppaluri R., Magrini V.J., Arthur C.D., White J.M., Chen Y.S., Shea L.K., Hundal J., Wendl M.C., Demeter R., Wylie T., Allison J.P., Smyth M.J., Old L.J., Mardis E.R., Schreiber R.D. (2012) Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. Nature. Feb 8;482(7385):400-4. Doi: 10.1038/nature10755.

Nasser citation when ready

Samstein R.M., Lee C.H., Shoushtari A.N., Hellmann M.D., Shen R., Janjigian Y.Y., Barron D.A., Zehir A., Jordan E.J., Omuro A., Kaley T.J., Kendall S.M., Motzer R.J., Hakimi A.A., Voss M.H., Russo P., Rosenberg J., Iyer G., Bochner B.H., Bajorin D.F., Al-Ahmadie H.A., Chaft J.E., Rudin C.M., Riely G.J., Baxi S., Ho A.L., Wong R.J., Pfister D.G., Wolchok J.D., Barker C.A., Gutin P.H., Brennan C.W., Tabar V., Mellinghoff I.K., DeAngelis L.M., Ariyan C.E., Lee N., Tap W.D., Gounder M.M., D'Angelo S.P., Saltz L., Stadler Z.K., Scher H.I., Baselga J., Razavi P.,

Klebanoff C.A., Yaeger R., Segal N.H., Ku G.Y., DeMatteo R.P., Ladanyi M., Rizvi N.A., Berger M.F., Riaz N., Solit D.B., Chan T.A., Morris L.G.T.. (2019) Tumor mutational load predicts survival after immunotherapy across multiple cancer types. Nat Genet Feb;51(2):202-206. doi: 10.1038/s41588-018-0312-8. Epub 2019 Jan 14.

Figure 1. Matched tumor normal sequencing vs. tumor only sequencing. In top panel, a matched normal is used to filter germline mutations (red and green). In bottom panel, a germline data base is used. In this case, one germline mutation (green) is erroneously interpreted as somatic, leading to overestimate of TMB.