IJC
International Journal of Cancer

Size matters: Dissecting key parameters for panel-based tumor mutational burden analysis

Ivo Buchhalter[1,2], Eugen Rempel[1], Volker Endris[1], Michael Allgäuer[1], Olaf Neumann[1], Anna-Lena Volckmar[1], Martina Kirchner[1], Jonas Leichsenring[1], Amelie Lier[1], Moritz von Winterfeld[1], Roland Penzel[1], Petros Christopoulos[3,4],‡ Michael Thomas[3,4],‡ Stefan Fröhling[5,6], Peter Schirmacher[1,6], Jan Budczies[1,6†] and Albrecht Stenzinger[1,6†]

[1]Institute of Pathology, University Hospital Heidelberg, Heidelberg, Germany
[2]Division of Applied Bioinformatics, German Cancer Research Center (DKFZ), Heidelberg, Germany
[3]Department of Thoracic Oncology, Thoraxklinik at Heidelberg University Hospital, Heidelberg, Germany
[4]Translational Lung Research Center Heidelberg (TLRC-H), Heidelberg, Germany
[5]Department of Translational Oncology, National Center for Tumor Diseases (NCT) Heidelberg and German Cancer Research Center, Heidelberg, Germany
[6]German Cancer Consortium (DKTK), Heidelberg Partner Site, Germany

Tumor mutational burden (TMB) represents a new determinant of clinical benefit from immune checkpoint blockade that identifies responders independent of PD-L1 expression levels and is currently being explored in clinical trials. Although TMB can be measured directly by comprehensive genomic approaches such as whole-genome and exome sequencing, broad availability, short turnaround times, costs and amenability to formalin-fixed and paraffin-embedded tissue support the use of gene panel sequencing for approximating TMB in routine diagnostics. However, data on the parameters influencing panel-based TMB estimation are limited. Here, we report an extensive *in silico* analysis of the TCGA data set that simulates various panel sizes and compositions. We demonstrate that panel size is a critical parameter that influences confidence intervals (CIs) and cutoff values as well as important test parameters including sensitivity, specificity, and positive predictive value. Moreover, we evaluate the Illumina TSO500 panel, which will be made available for TMB estimation, and propose dynamic, entity-specific cutoff values based on current clinical trial data. Optimizing the cost–benefit ratio, our data suggest that panels between 1.5 and 3 Mbp are ideally suited to estimate TMB with small CIs, whereas smaller panels tend to deliver imprecise TMB estimates for low to moderate TMB (0–30 muts/Mbp), connected with insufficient separation of hypermutated tumors from non-hypermutated tumors.

Introduction

During the process of carcinogenesis, tumor cells accumulate mutations some of which lead to the formation of tumor-specific neoantigens.[1] These neoantigens are recognized by a patient's immune system as non-self and cancerous outgrowth is kept in check. If the aberrant cell population persists, an equilibrium may form. The immune system automatically selects for better adapted tumor clones that often hijack the

Key words: tumor mutational burden, TMB, mutational load, panel, NGS, sequencing
Additional Supporting Information may be found in the online version of this article.
†Shared last authorship.
‡Members of the German Center for Lung Research (DZL), Heidelberg, Germany.
Conflict of interest: AS is a consultant/advisory board member of AstraZeneca, Bristol-Myers Squibb, Novartis and Thermo Fisher Scientific and received speaker's honoraria from BMS, MSD, Roche, Illumina, AstraZeneca, Novartis and Thermo Fisher as well as research funding from Chugai and BMS. PS received advisory board honoraria from Pfizer, Roche, Novartis and AstraZeneca as well as speaker's honoraria and research funding from Roche, AstraZeneca and Novartis. MT received advisory board honoraria from Novartis, Lilly, BMS, MSD; Roche, Celgene, Takeda, AbbVie and Boehringer, speaker's honoraria from Lilly, MSD, Takeda, research funding from AstraZeneca, BMS, Celgene, Roche and travel grants from BMS, MSD, Novartis and Boehringer.
Grant sponsor: German Cancer Consortium (DKTK); **Grant numbers:** NA; **Grant sponsor:** German Cancer Consortium
DOI: 10.1002/ijc.31878
History: Received 16 Jul 2018; Accepted 10 Sep 2018; Online 21 Sep 2018
Correspondence to: Albrecht Stenzinger, Institute of Pathology, University Hospital Heidelberg, Im Neuenheimer Feld 224 69120 Heidelberg, Germany, Tel.: 49-622156-34380, Fax: 49-622156-5251, E-mail: albrecht.stenzinger@med.uni-heidelberg.de; or Jan Budczies, Institute of Pathology, University Hospital Heidelberg, Im Neuenheimer Feld 224 69120 Heidelberg, Germany, Tel.: 49-622156-32757, Fax: 49-622156-5251, E-mail: jan.budczies@med.uni-heidelberg.de

What's new?

While the immune system can fight cancer, tumor clones may hijack the body's own mechanisms of dampening immune responses by expressing immune-checkpoint proteins. Tumor mutational burden (TMB) is an emerging selection biomarker for patients who would benefit from immune checkpoint inhibitor therapy. Unification of TMB estimation and driver mutation detection in a single panel sequencing assay is highly desirable but challenging, however. Combinatorial calculations and simulations in the TCGA dataset show that panel sizes greater than 1.5 Mbp are recommendable, and–although not immunogenic–synonymous and nonsense mutations can be included in TMB estimation together with cancer-type specific correction factors.

body's own mechanisms of dampening immune responses, for example, by expressing immune-checkpoint proteins such as PD-L1.[2,3] The clinical relevance of these insights is underlined by the successful implementation of immune checkpoint inhibitors against CTLA-4 (cytotoxic T-lymphocyte-associated protein 4), PD-1 (programmed cell death protein 1) and PD-L1 (programmed death-ligand 1) in various cancers including cutaneous melanoma,[4,5] non–small-cell lung cancer (NSCLC),[6,7] kidney cancer,[8] bladder cancer,[9] head and neck cancers,[10] Hodgkin's lymphoma[11] and Merkel cell carcinoma.[12]

However, not all patients benefit from immune checkpoint inhibitor therapy; therefore, availability of biomarkers to guide patient management is crucial.[13–15] Currently, the most widely used biomarker is PD-L1 expression assessed by immunohistochemistry.[16,17] Immunohistochemical assays are rather straightforward to implement, robust, and offer a short turnaround time. However, there is accumulating evidence that the predictive power of PD-L1 expression is limited, reflected by the FDA designation of a complementary rather companion diagnostic tool for several indications and cancer types.[18,19] This clinical observation is well in line with core biological considerations,[20] which suggest that a single biomarker is unlikely to reflect the multifaceted and complex interplay between host and tumor.[21]

A tumor biomarker receiving increasing attention is overall somatic mutational load or tumor mutational burden (TMB). The rationale here is that accumulation of more mutations is associated with a higher chance of new immunogenic peptides (neoantigens) being presented to the immune system.[22,23] Indeed, higher TMB was associated with clinical benefit in a retrospective exploratory analyses of patients with melanoma receiving CTLA-4 blockade and patients with NSCLC receiving PD-1 blockade.[24–26] In the CheckMate (CM) 026 trial testing the first-line nivolumab in Stage IV or recurrent NSCLC, there was no significantly prolonged progression-free survival with anti PD-1 therapy compared to standard chemotherapy in patients with PD-L1 expression level of 5% or more.[27] However, in an exploratory *post hoc* analysis, it was found that patients with higher TMB showed better response and longer median progression-free survival with PD-1 blockade compared to chemotherapy and that patients with both high TMB and high PD-L1 expression responded particularly well to immunotherapy. In line with these findings, in small-cell lung cancer,[28] TMB was also found to be independently predictive for response to combinatorial checkpoint inhibition and PD-L1 blockade.

Although these studies used whole-exome sequencing (WES), recent reports demonstrated the clinical utility of TMB approximation by targeted panel NGS.[29–31] Because WES is not feasible in larger patient population in current clinical practice due to its high costs, substantial turn-around times and limited availability of fresh unfixed tissue in routine diagnostics ideally required for comprehensive sequencing, a targeted sequencing approach utilizing widely available formalin-fixed paraffin-embedded (FFPE) clinical specimens is preferable and most promising.

Besides reports on the Foundation Medicine (FMI) and the MSKCC-IMPACT panels,[31–33] data on the framework and content of a reliable panel-based TMB estimation in a laboratory-developed test setting are preliminary and very limited.[34–36] Moreover, data on the TMB panel that are currently being codeveloped by technical and pharmaceutical industry in a strategic partnership are lacking.[37] As many laboratories worldwide are currently considering implementation of TMB measurement in their workflows rather than outsourcing to commercial providers, additional guidance based on scientific data is needed.

Using the TCGA data set, we have carried out an extensive *in silico* analysis to address (*i*) general aspects and requirements of panel sequencing–based TMB estimation as well as (*ii*) the performance of specific gene panels.

Materials and Methods

Cancer types investigated in our study

Adrenocortical Carcinoma (ACC), Bladder Urothelial Carcinoma (BLCA), Breast invasive carcinoma (BRCA), Cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), Colon adenocarcinoma (COAD), Esophageal Carcinoma (ESCA), Glioblastoma multiforme (GBM), Head and Neck squamous cell carcinoma (HNSC), Kidney Chromophobe (KICH), Kidney renal clear cell carcinoma (KIRC), Kidney Renal Papillary Cell Carcinoma (KIRP), Brain Lower Grade Glioma (LGG), Liver hepatocellular carcinoma (LIHC), Lung adenocarcinoma (LUAD), Lung squamous cell carcinoma (LUSC), Ovarian serous cystadenocarcinoma (OV), pancreatic adenocarcinoma (PAAD), Pheochromocytoma and Paraganglioma (PCPG), Prostate Adenocarcinoma (PRAD),

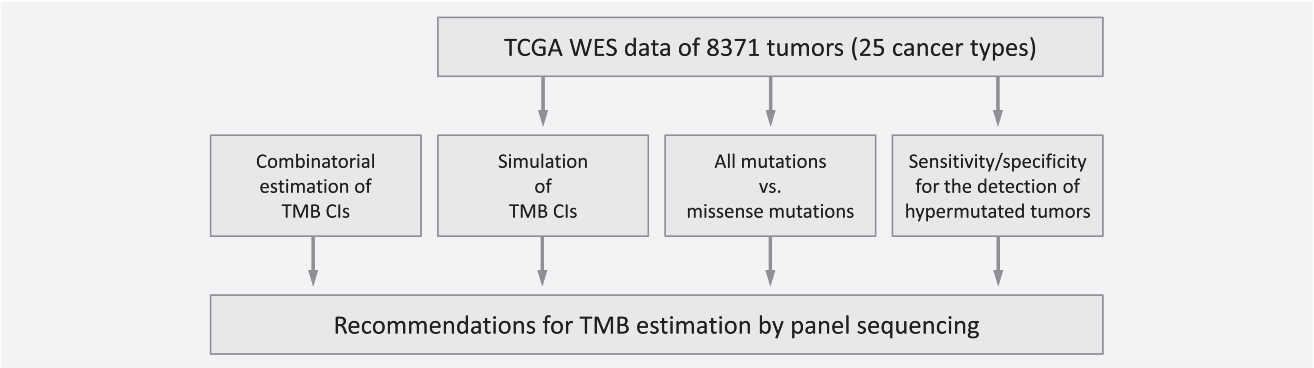

Figure 1. Flowchart visualizing the analysis workflow comprising combinatorial calculations and simulations of sequencing panels in the TCGA data. CI, confidence interval.

Rectum adenocarcinoma (READ), Skin Cutaneous Melanoma (SKCM), Stomach adenocarcinoma (STAD), Papillary Thyroid Carcinoma (THCA), Uterine Corpus Endometrial Carcinoma (UCEC) and Uterine Carcinosarcoma (UCS).

Confidence intervals for TMB estimates

Assuming that each base in the tumor genome is mutated with the same fixed probability, estimation of TMB can be combinatorial modeled based on the binomial distribution with the panel length as sample size and the number of detected mutations as the number of successes. Under this assumption, confidence intervals (CIs) of TMB were calculated using the Clopper–Pearson method as it is implemented in the R package binom. In the panel sequencing simulation study in the WES data of the pan-cancer TCGA cohort, empirical CIs were estimated using a sliding window including 300 tumors of similar TMB.

TCGA mutation data

Somatic mutation calls were downloaded from the TCGA-legacy repository for 23 major cancer types. We chose to use the legacy data because all target files used in our study are created for the hg19 (GRCh37) reference genome. To assure that the data can be compared within and between cohorts, we only used mutation calls provided by the Broad institute (*.oxoG.snp.capture.tcga.vcf.gz). All PASS somatic variants were extracted and annotated using Annovar[38] and Gencode 19[39] To make sure that the target regions considered were the same for all samples, we only considered mutations falling into regions included in the Agilent SureSelect V4 enrichment kit. The microsatellite instability (MSI) status for TCGA-COAD was obtained from supplemental data from Cortes-Ciriano *et al.*[40] Access to TCGA data was granted by the NIH (project number: 15058).

Regions of interest (panels)

The regions of commercial panels were either extracted from the vendors homepage (TST170) or generously provided by the vendor prior to publication (TSO500). Overlapping target regions, e.g., region used for mutation calling and structural variant detection, were merged. Because some of the commercial panels include intronic regions not covered by the Agilent SureSelect V4 enrichment kit, these regions were excluded from the panel length calculation. Additionally, we simulated panels of various sizes. For this we first calculated the mean target region of the Agilent SureSelect V4 enrichment kit and then randomly drew n target regions with n = target_panel_size/mean_target_region.

Calculation of TMB thresholds to separate hypermutated and non-hypermutated tumors

The TMB thresholds were calculated specifically for each cancer type based on the threshold of 243 missense mutations defined in the CM 026 trial.[27] Here, we took into account all point mutations instead of only the missense mutations and calculated cancer type–specific thresholds as follows: THR = 243 × sum(all mutations)/sum(missense mutations). The sums are counting all mutations detected in the tumors of the specific cancer type. The significance of enrichment of polymerase ε (POLE)–mutated and of MSI-high tumors in the set of hypermutated tumors was assessed by the two-sided Fisher's exact test.

Results

The analysis is segmented to address the core aspects that influence TMB approximation by panel sequencing in clinical practice. Figure 1 provides an overview on these segments, which are detailed in the subsequent paragraphs.

Panel size and TMB cutoff points

Sequencing of the entire genome or exome facilitates direct measurement of TMB in a given tumor. However, in case of targeted NGS, TMB needs to be extrapolated from the number of mutations detected in the sequenced region. To obtain an overview on the interdependence of TMB cutoff points and panel sizes, we calculated ratios of genome and panel lengths and their association to mutation measurement. Using this scaling approach, we can determine how many genomic

Table 1. Relation of mutation numbers in the genome and mutation numbers in sequencing panels of different size

Panel size (Mbp)	Number of mutations in the genome[1] corresponding to 1 mutation in the panel[2]	Number of mutations in the panel corresponding to 25,000 mutations in the genome[2]
0.05	54,000	1
0.2	13,500	2
0.533[3]	5,066	5
1	2,700	10
1.5	1,800	14
1.942[4]	1,391	18
33[5]	82	306

[1]2,700 Mbp were used as an approximation to the region of the human genome where mutations can be called.
[2]Rounded off to the next integer.
[3]Illumina TST170 panel.
[4]Illumina TSO500 panel.
[5]Human exome (protein coding regions).

mutations correspond to a single mutation in the panel region (Table 1). *Vice versa*, a TMB cutoff point to detect hypermutated tumors by whole-genome sequencing (e.g., 25,000 mutations/genome) can be transformed into cutoff points for panels of different sizes. As shown in Table 1, for NGS panels smaller than 1 Mbp, the thresholds for the separation of hypermutated from non-hypermutated tumors fall below 10 mutations. Thus, small panels will be inaccurate in the classification of tumors with TMB close to the threshold.

Estimation of TMB confidence intervals

To quantify the precision of TMB estimation for NGS panels of different sizes, we calculated CIs (interval estimates containing the true TMB value with 95% confidence) for different TMBs by combinatorics based on the binomial distribution and by simulations in WES data of the pan-cancer TCGA cohort (Fig. 2). Considering the log scale display, the graphics shows a strongly increasing relative error of TMB (in percentage) for increasing TMB. For TMB = 10 mutations (muts)/Mpb, the combinatorial 95% CIs ranged between 3.4–22.8, 6.1–15.5, and 7.4–13.2 muts/Mbp for panel sizes of 0.533, 1.942, and 5 Mbp, respectively. As shown in Figure 2*a*, using a 1.942 Mbp panel instead of a 0.533 Mbp panel considerably shortened the range of the CIs with a length of 9.5 muts/Mbp instead of 19.4 muts/Mbp for TMB = 10 muts/Mbp, while the additional benefit of using a 5 Mbp panel was moderate (length of CI: 5.8 Mbp for TMB = 10 muts/Mbp).

Simulating a real-world scenario, we projected WES data from 8,371 tumors from TCGA to the Illumina panels TST170 (0.533 Mpb) and TSO500 (1.942 Mbp). Empirical confidence estimated using a sliding window (Figs. 2*b* and 2*c*) were very similar to those obtained by the combinatorial method confirming that combinatorial estimation of CIs is a good approximation to the real-world situation. Collectively, these data demonstrate that the precision of TMB estimation using targeted NGS considerably depends

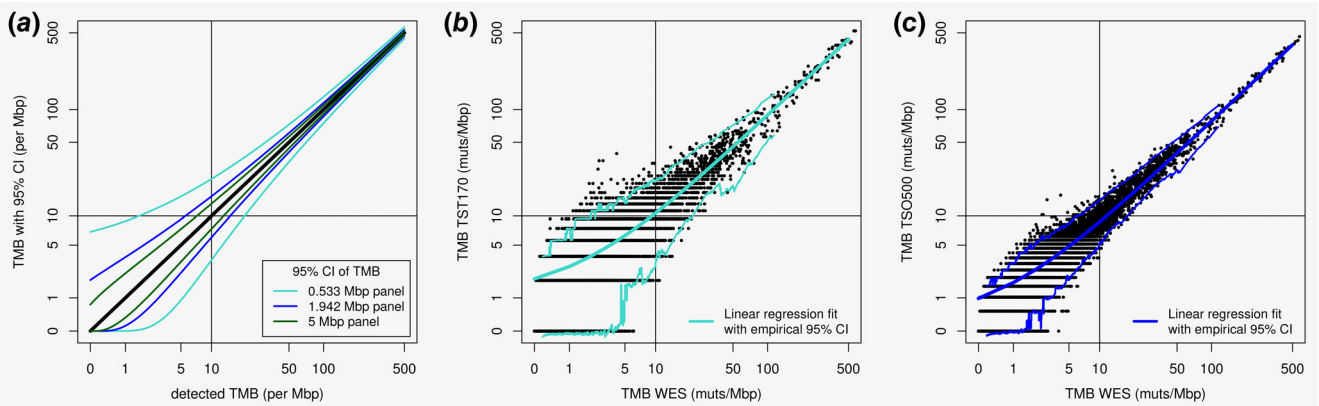

Figure 2. Precision of TMB estimation for sequencing panels of different size. (*a*) Theoretical confidence intervals (CIs) for detected TMB = 0, ..., 500 muts/Mbp. Binomial distribution CIs were calculated using the Clopper–Pearson method. The CIs were derived under the hypothesis that mutations are homogenously distributed in the genome, that is, each base in the tumor genome is mutated with the same probability and represents the upper limit for the precision that can be obtained. (*b*) Empirical CIs estimated from a simulation of the TST170 panel in the pan-cancer TCGA data (8,371 tumors). (*c*) Empirical CIs estimated from a simulation of the TSO500 panel in the pan-cancer TCGA data.

on the panel size for panels typically used in routine diagnostics.

Inclusion of synonymous mutations in the calculation of TMB

The definition of TMB used in the exploratory analysis of the CM 026 trial included only missense mutations.[27] However, for moderately sized sequencing panels, it is tempting to include all mutations instead of only the subset of missense mutations. Leaving both technical and biological considerations regarding, for example, calling algorithms and antigenicity probabilities aside, the underlying rationale of this approach is to increase the combinatorial precision of TMB estimation by including a higher number of mutations. By contrast, from a biological standpoint, there is no rationale to include synonymous mutations that are silent and do not contribute to immunogenicity. Therefore, we analyzed the relation of the number of non-missense mutations (including synonymous mutations, nonsense mutations and splice site mutations) and the number of missense mutations in the tumors from TCGA (Fig. 3 and Supporting Information Figure S1). For many of the cancer types including COAD, LUAD, skin cutaneous melanoma (SKCM) and uterine corpus endometrial carcinoma (UCEC), non-missense and missense mutations correlated very strongly ($R \geq 0.97$). For these cancer types, we propose to include all-point mutations in the calculation of TMB to enhance the precision of the estimation. Following this approach, the cutoff points for detection of hypermutated tumors need to be scaled by a factor of a = sum (all mutations)/sum (missense mutations) with sums counting all mutations in the tumors of the cancer type under consideration. The scaling factors (shown in Fig. 3 and Supporting Information Figure S1) were similar

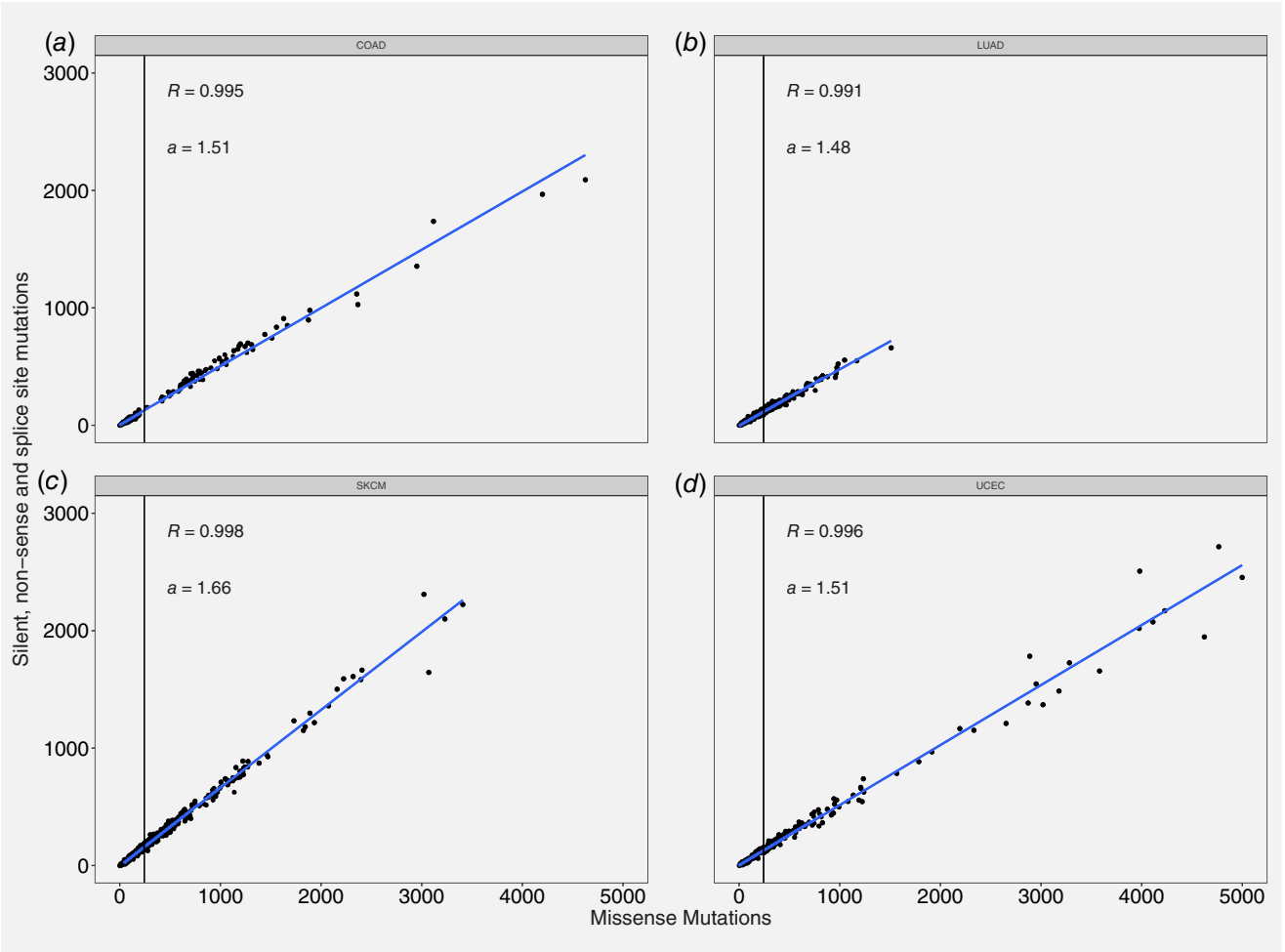

Figure 3. Relation between the number of all exonic point mutations (including silent mutations) and the number of missense mutations. The number of non-missense mutations (silent mutations, nonsense mutations and mutations at splice sites) correlated strongly with the number of missense mutations in lung adenocarcinoma (LUAD, *a*), colon adenocarcinoma (COAD, *b*), skin cutaneous melanoma (SKCM, *c*) and uterine endometrial cell cancer (UCEC, *d*) and many other cancer types (Supporting Information Fig. S1). The scaling factors a = sum (all point mutations)/sum(missense mutations) were similar (between 1.44 and 1.54) for all the cancer types with the single exception of a slightly higher scaling factor for the SKCM (a = 1.66).

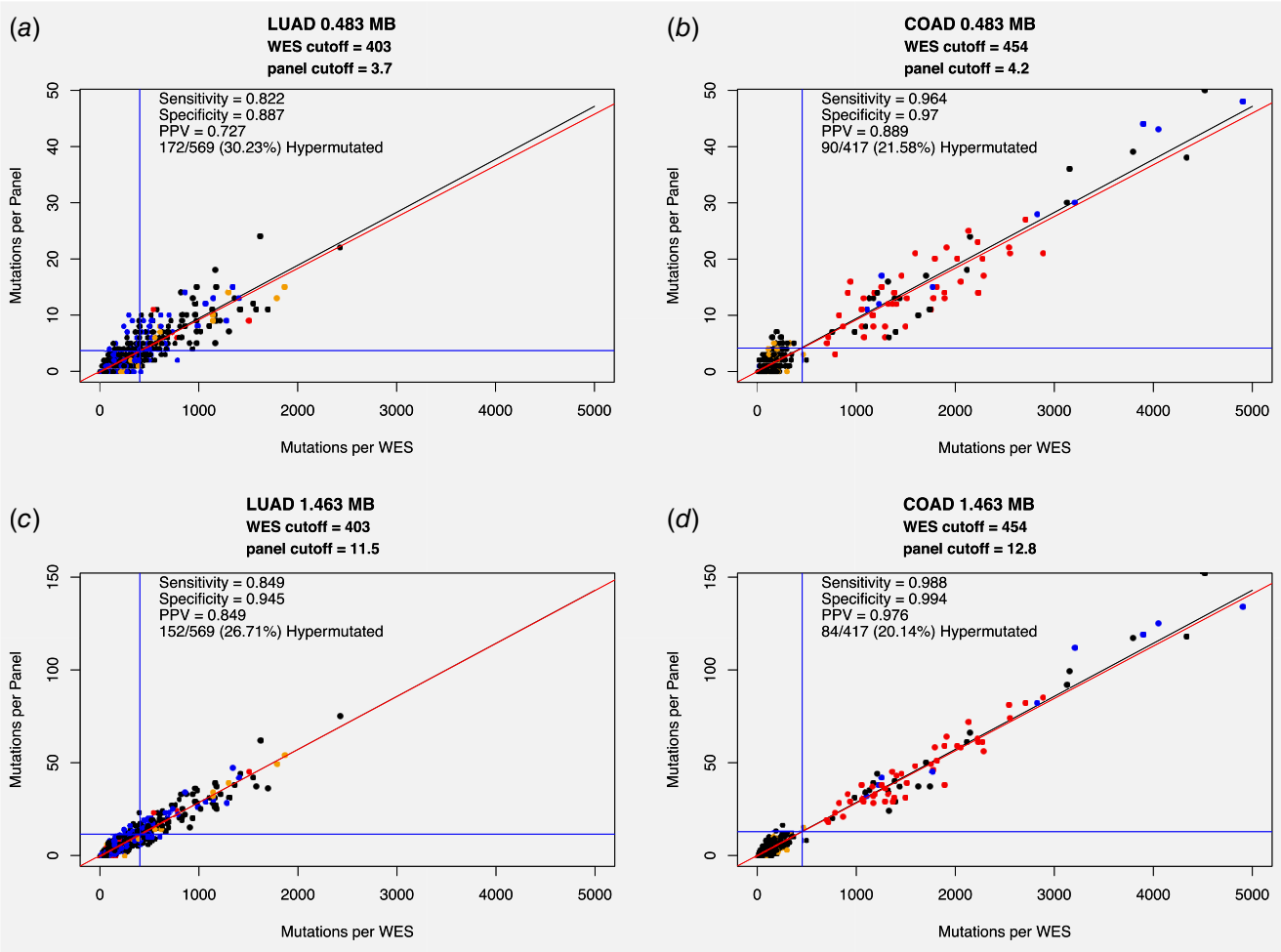

Figure 4. Detection of hypermutated tumors by simulated panels of 0.5 and 1.5 Mbp. All-point mutations in the coding region were included in the calculation of TMB. Cancer type–specific cutoff points were used for the separation of hypermutated tumors. The cutoff points were calculated from the threshold determined in CM 026 (n = 243) by scaling with the a-factors determined before. (*a*, *c*) Lung adenocarcinoma (LUAD): blue = KRAS mutated, orange = EGFR mutated, red = EGFR amplified. (*b*, *d*) Colon adenocarcinoma (COAD): blue = POLE mutated, orange = MSI-low/MSS, red = MSI-high. Black diagonal: Expected relation between panel TMB and WES TMB (slope = panel size/exome size). Red diagonal: Linear regression to the data points. Blue vertical line: WES cutoff point. Blue horizontal line: panel sequencing cutoff point obtained by transforming the WES cutoff point with the linear regression model.

(between 1.44 and 1.54) for all the cancer types with the single exception of a slightly higher scaling factor for SKCM (a = 1.66).

TMB cutoffs for different sizes of simulated panels

To simulate sequencing panels of sizes 0.5, 1, 1.5, 2, 3, 5 and 10 Mbp, respectively, we randomly collected amplicons across the exome. We then calculated the number of high-confidence somatic single nucleotide variants (SNVs) falling into these simulated target regions for all samples. When plotting the numbers of mutations found in the panels against the mutations found in the entire exome, we observed a linear relationship across all tumor types analyzed (Fig. 4 and Supporting Information Figure S2). Although the data points scattered around the theoretical diagonal (black) line, we additionally plotted the linear regression curve (red line).

Currently, there is no consensus on optimal cutoff points to separate tumors of high- and low-mutational burden. Instead of using a fixed cutoff point (such as 25,000 mutations per genome or 10 muts/Mbp), we used a dynamic cancer type specific threshold based on the cutoff of 243 missense mutations from the exploratory analysis of the CM 026 study.[27] To increase the precision of TMB estimation, we included all mutations in the analysis (instead of considering missense mutations only) and therefore scaled the cutoff of 243 by the cancer type–specific factors obtained before (denoted as "a" in Fig. 3). The resulting thresholds were similar for all examined tumor types and ranged from 403 (LUAD) to 457 (PRAD) with an average of 427.9 and a median of 425.5 somatic point mutations (Supporting Information Table S1). In a second step, cutoffs of WES data were converted to cutoffs for the panels data by the linear regression fit. As exemplified in

Figure 4, increasing the size of the panels changed sensitivity and specificity as well as the positive predictive value (PPV). Moreover, larger gene panels were associated with higher cutoff values thus increasing robustness and reliability of a given assay.

For two of the tested tumor types, we investigated associations between TMB and other clinically relevant molecular parameters. We narrowed in on a simulated panel of ~1.5 Mbp and samples classified as hypermutated or non-hypermutated based on the panel-specific threshold determined in Fig. 3. For LUAD, we looked at analyzed the mutation status of KRAS and EGFR. For COAD, we analyzed the mutation status of POLE as well as the MSI status. Although we did not find any significant enrichment of KRAS or EGFR mutations in high- or low-mutational load group of the LUAD samples, tumors with mutations in POLE and MSI-high samples were highly enriched in the hypermutated fraction of COAD with $p = 8.64 \times 10^{-14}$ and $p = 3.58 \times 10^{-35}$, respectively (Fig. 4*d*). In line with this finding, we found an enrichment of MSI-low cases in the non-hypermutated fraction with $p = 0.0067$. As visualized in Figures 4*b* and 4*d*, smaller gene panels fall short of their larger counterparts even in COAD cases harboring TMB levels that follow a bimodal distribution caused by MSI and mutant POLE.

TMB detection using TST170 and TSO500

Different types of commercial, community-designed and customer-designed sequencing panels are currently used for the detection of driver mutations in specific cancer types or across cancer types. Additionally new panels designed to discriminate between low- and high-mutational burden are emerging. Here, we compared the performance of the Illumina TST170 (0.533 Mbp) and TSO500 (1.942 Mbp, refer to Supporting Information Table S2 for basic characteristics) panels for the detection of TMB. Interestingly, plotting the number of somatic mutations discovered in these panels against the number of somatic mutations found in the entire exome showed a deviation from the expected pattern with more mutations being detected by panels than expected (Fig. 5). Therefore, the threshold between hypermutated and non-hypermutated tumors was slightly higher than expected by the panel size compared to the exome size, namely 5 and 16 for the TST170 and TSO500 panels, respectively.

Influence of hot spot mutations on TMB measurement

One concern when determining the TMB using only a small subset of the genome is a bias of the result due to the selection panel region. For example, usage of panels designed to detect recurrent driver mutations ("hotspot mutations") in cancer samples could lead to an overestimation of TMB. To investigate this possibility, we excluded all hot spot mutations from the analysis the COAD cohort. We defined hot spots as positions being mutated in at least 20 tumors in a recent study of more than 1,100 colorectal tumors.[41] Although there was a slight decrease in the hypermutated cases detected by the panel (103 with *vs.* 97 without hot spot mutations), the overall picture was left unaltered (Figs. 5*e* and 5*f*).

Influence of panel size and composition on the detection of TMB-high tumors

We simulated 100 random sequencing panels for each of the sizes 0.5, 1 and 1.5 Mbp and compared their performance for the detection of hypermutated tumors in terms of sensitivity, specificity and PPV (see Fig. 6). For LUAD, the median of sensitivity increased from 0.78 *via* 0.82 to 0.85, the median of specificity increased from 0.89 *via* 0.93 to 0.94 and the median of the PPV increased from 0.73 *via* 0.81 to 0.84. For COAD, the median of sensitivity increased from 0.98 *via* 0.99 to 0.99, the median of specificity increased from 0.97 *via* 0.99 to 0.99 and the median of the PPV from 0.89 *via* 0.96 to 0.98. There was also a clear trend to narrower ranges of sensitivity and specificity with increasing panel size (i.e., for larger panels, the performance depended less on the particular panel chosen).

Discussion

Recent results of the CM 026 and 227 trials suggest that TMB may become a new diagnostic biomarker to predict response to checkpoint blockers either alone or in combinatorial regimens in NSCLC patients.[27,28] Data obtained from trials and other studies including different tumor entities support this notion.[24–31,42–47] Very interestingly, CM 026 and 227 also suggest that TMB is a parameter that is independent from PD-L1 expression and identifies additional patients who may benefit from immune therapy. Besides WES, targeted sequencing of FFPE material with variable panel sizes was used to estimate TMB,[31,33–36,48] and lately CM 227[30] showed the clinically utility of FFPE-based TMB analysis as determined by the FMI panel. To provide a wide availability of testing, laboratory developed tests to determine the TMB needs to be implemented. However, current data that can provide guidance in this context are scarce and preliminary.

Addressing this need, we carried out combinatorial calculations and extensive simulations on the basis of TCGA mutation data. For small panel sizes (<1 Mbp), we observed large CIs for lower and moderate TMBs (between 0 and 30 mutations per Mbp). For example, when using a 0.5 Mpb panel, the combinatorial 95% CI did not exclude a TMB of 10 mutations per Mbp, which is a currently proposed threshold for high TMB, for a tumor harboring only 5 mutations per Mbp in reality. These data suggest that TMB estimation using small gene panels can be highly imprecise and thus clinically suboptimal for patient stratification and response prediction. In terms of precision, the benefit of increasing a 0.5 Mbp panel to a 1.5 Mbp panel was large, whereas the benefit of increasing a 1.5 Mbp panel to a 5 Mbp panel was moderate.

Moreover, cutoff values depend on the size of the gene panel. For example, a 15 muts/Mbp cutoff determined by one specific panel X cannot be automatically transferred to a

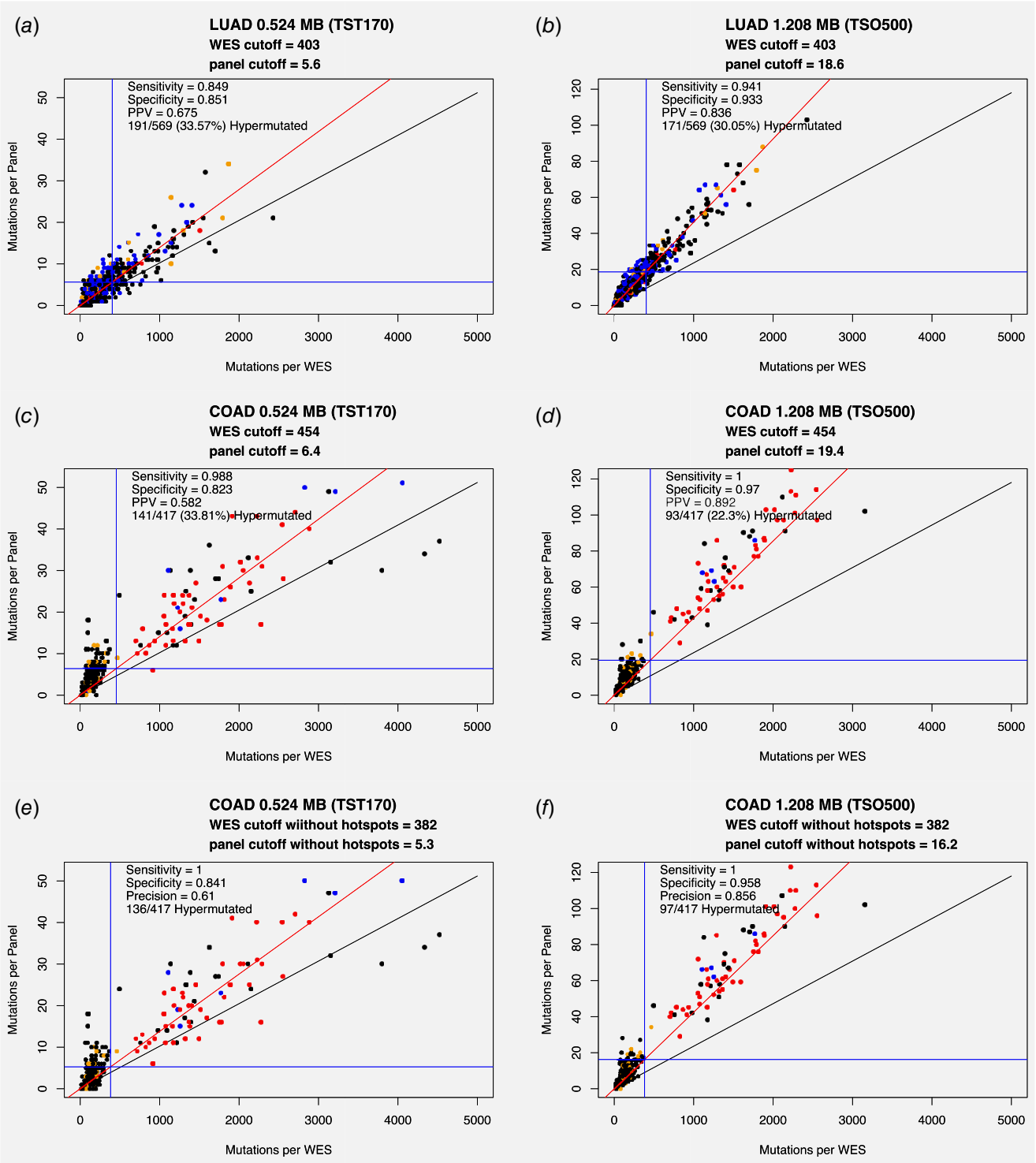

Figure 5. Detection of hypermutated tumors by the Illumina TST170 and TSO500 panels. All-point mutations in the coding region were included in the calculation of TMB, and the CM 026 threshold (n = 243) scaled with the cancer type–specific scaling factors estimated before was used at cutoff point for the separation of hypermutated tumors. (*a*, *b*) LUAD, Lung adenocarcinoma. (*c*, *d*) COAD, Colorectal adenocarcinoma. (*e*, *f*) Same as (*c*, *d*), but with the mutational hotspots removed. Black diagonal represents expected relation between panel TMB and WES TMB (slope = panel size/exome size). Red diagonal represents linear regression to the data points. Blue vertical line indicates WES cutoff point. Blue horizontal line indicates panel sequencing cutoff point obtained by transforming the WES cutoff point with the linear regression model.

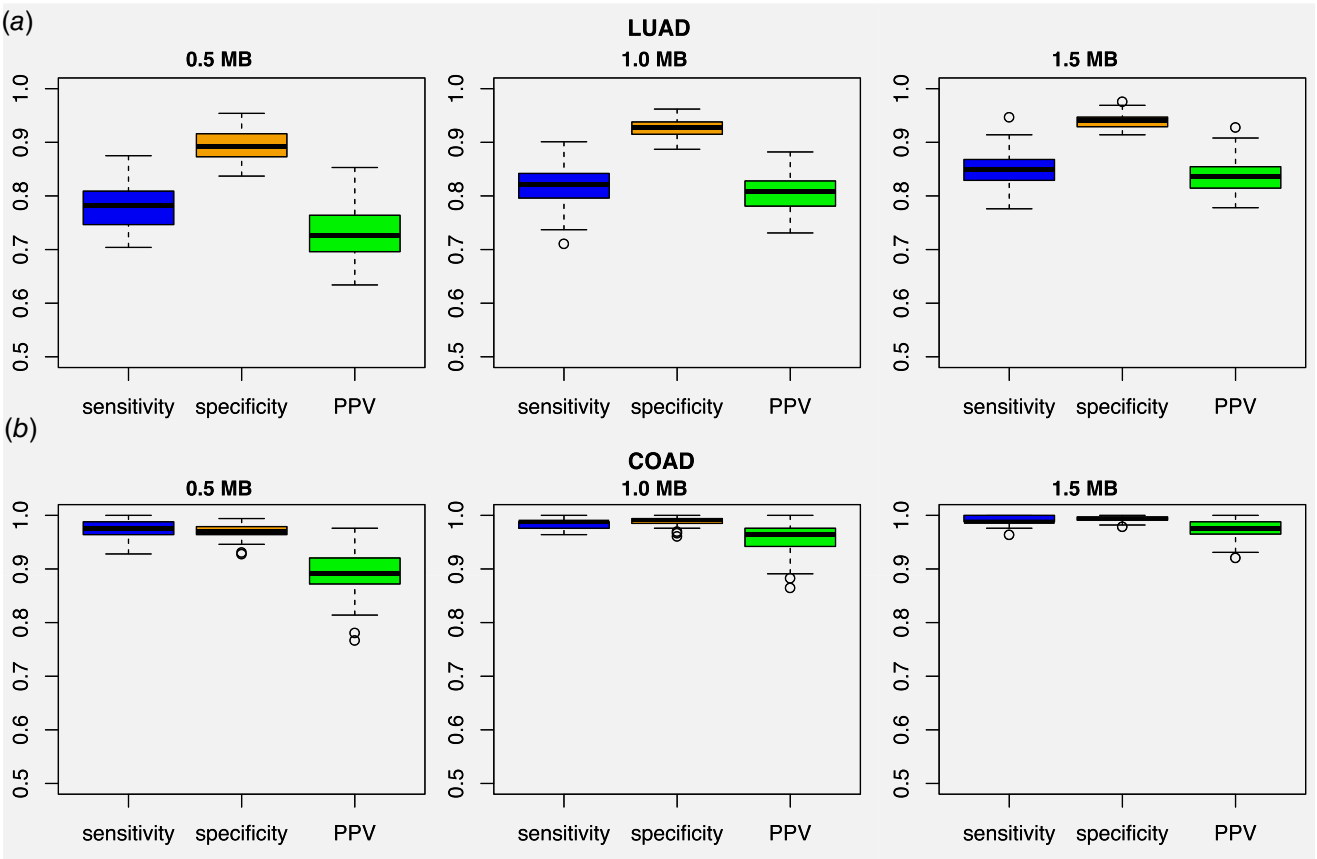

Figure 6. Sensitivity, specificity and positive predictive value (PPV) for the detection of hypermutated tumors with 100 random sequencing panels. For each of the sizes 0.5, 1.0 and 1.5 Mbp, 100 random sequencing panels were simulated. Sensitivity, specificity and PPV increased for increasing panel sizes, while the variance of these quantities in the 100 simulated panels decreased. (*a*) Lung adenocarcinoma (LUAD). (*b*) Colon adenocarcinoma (COAD).

different panel Y of different size and composition. In general, with increasing panel size, cutoff values increase correspondingly so that very small gene panels of several kbp have cutoff values lower than 10 mutations to determine TMB high cases while larger gene panels will be associated with higher thresholds. Although very small panels are unable to separate signal from noise (almost every NSCLC case will show at least one mutated gene in a panel directed against prevalent mutations in NSCLC-related genes), larger gene panels are associated with reasonable cutoff values that help identify true signals from background noise in routine diagnostics.

Importantly, we observed a significant increase in sensitivity, specificity and PPV for the detection of hypermutated tumors using 1.5 Mbp panels compared to 0.5 and 1 Mbp panels in simulations in the TCGA data. Even in tumor types like colorectal adenocarcinoma where the otherwise continuous variable TMB shows a largely bimodal distribution due to the underlying biology (associated with POLE and MSI status), the size effect of gene panels leading to improved core test parameters was still influential. For all other tumor entities showing a unimodal distribution of TMB across samples, this effect becomes even more relevant arguing for the use of sufficiently large gene panels to reliably determine mutational load in different cancer types.

In line with these data, our *in silico* analysis of commercially available panels also demonstrates that Illumina's TST170 is too small for precise TMB estimation, while the TSO500 panel shows good performance data. Based on the CM 026 results, we here calculated dynamic cutoff values for both panels in all major cancer entities. Unlike the MSI-H/dMMR (deficient DNA mismatch repair) status of a tumor, which was approved by the FDA as a histology-agnostic biomarker for prediction of response to pembrolizumab[49,50] based on several CM trials (016, 164, 012, 028 and 158), our data suggest that TMB cutoffs should not only be panel-specific but also be entity-specific. The latter notion is in line with data by Chalmers *et al.*[32]

Including all mutations instead of only missense mutations in the calculation of TMB is a strategy to enhance precision, as measuring a higher number of mutations reduces the error in TMB estimation. Therefore, we analyzed if the number of other mutations (including silent mutations) is proportional to the number of missense mutations in major cancer types. High correlations were obtained for many cancer types

supporting that inclusion of all mutations is possible without changing the biological interpretation of TMB. Including all mutations, the number of mutations contributing to TMB could be enhanced by a factor between 1.44 and 1.54 for most cancer types and by a factor of 1.66 for melanoma. Cutoff points for the separation of hypermutated tumors when originally based on the load of missense mutations need to be shifted accordingly.

The TCGA data set provided a unique well-controlled environment for the *in silico* analysis carried out here and highlighted the strength of the TCGA data also in the field of routine diagnostics-related research. We acknowledge, however, that our analysis did not include important wet-lab factors that will influence TMB measurement. These include but are not restricted to DNA quality and quantity, coverage and read depth of sequencing platforms, the influence of germline subtraction without sequencing blood as well as the impact of C > T transitions introduced by formalin fixation. In this context, the threshold of the variant allele frequency defined for the inclusion of mutations in the calculation of TMB is crucial and requires further discussions. Adding to the complexity, McGranahan recently reported that clonality influences clinical outcome: although clonal mutations that contribute to TMB were associated with clinical benefit, subclonal mutations were less likely to elicit a durable clinical response to PD-1 blockade.[51] As therapy regimens can influence the clonal architecture of tumors, TMB estimates may dynamically change during the course of disease and may be associated with variable predictive power depending on algorithms filtering clonality. Wet-lab studies need to address these and other aspects in the near future.

TMB as predictive biomarker for immune therapies and in particular blockage of the PD-1/PD-L1 axis is expected to enter routine clinical diagnostics very soon. Recent data suggest that TMB can also be measured in circulating tumor DNA and has clinical utility.[52–54] Therefore, TMB panels, workflows and evaluation methods urgently need to be optimized to handle clinical FFPE tissue samples. Here, we contributed to this process by combinatorial calculations and extensive simulations in the TCGA data. Our analyses showed that "size does matters" with an optimal panel size being reached between 1.5 and 3 Mbp considering the benefit–cost ratio and that inclusion of all point mutations (instead of only missense mutations) in the TMB calculation is possible and recommendable to enhance precision.

Acknowledgements

This work was supported by the German Cancer Consortium (DKTK; to PS, JB and AS). The results shown here are in whole or part based upon data generated by the TCGA Research Network: http://cancergenome.nih.gov/.

References

1. Gubin MM, Artyomov MN, Mardis ER, et al. Tumor neoantigens: building a framework for personalized cancer immunotherapy. *J Clin Investig* 2015;125:3413–21.
2. Schreiber RD, Old LJ, Smyth MJ. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. *Science* 2011;331:1565–70.
3. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer* 2012;12:252–64.
4. Hodi FS, O'Day SJ, McDermott DF, et al. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med* 2010;363:711–23.
5. Wolchok JD, Kluger H, Callahan MK, et al. Nivolumab plus ipilimumab in advanced melanoma. *N Engl J Med* 2013;369:122–33.
6. Borghaei H, Paz-Ares L, Horn L, et al. Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N Engl J Med* 2015;373:1627–39.
7. Brahmer J, Reckamp KL, Baas P, et al. Nivolumab versus docetaxel in advanced squamous-cell non-small-cell lung cancer. *N Engl J Med* 2015;373:123–35.
8. Motzer RJ, Escudier B, McDermott DF, et al. Nivolumab versus Everolimus in advanced renal-cell carcinoma. *N Engl J Med* 2015;373:1803–13.
9. Powles T, Eder JP, Fine GD, et al. MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic bladder cancer. *Nature* 2014;515:558–62.
10. Bauml J, Seiwert TY, Pfister DG, et al. Pembrolizumab for platinum- and Cetuximab-refractory head and neck cancer: results from a single-arm, phase II study. *J Clin Oncol* 2017;35:1542–9.
11. Ansell SM, Lesokhin AM, Borrello I, et al. PD-1 blockade with nivolumab in relapsed or refractory Hodgkin's lymphoma. *N Engl J Med* 2015;372:311–9.
12. Kaufman HL, Russell J, Hamid O, et al. Avelumab in patients with chemotherapy-refractory metastatic Merkel cell carcinoma: a multicentre, single-group, open-label, phase 2 trial. *Lancet Oncol* 2016;17:1374–85.
13. Topalian SL, Taube JM, Anders RA, et al. Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. *Nat Rev Cancer* 2016;16:275–87.
14. Brahmer JR, Tykodi SS, Chow LQ, et al. Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *N Engl J Med* 2012;366:2455–65.
15. Goodman AM, Kato S, Bazhenova L, et al. Tumor mutational burden as an independent predictor of response to immunotherapy in diverse cancers. *Mol Cancer Ther* 2017;16:2598–2608.
16. Herbst RS, Soria JC, Kowanetz M, et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature* 2014;515:563–7.
17. Taube JM, Klein A, Brahmer JR, et al. Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clin Cancer Res* 2014;20:5064–74.
18. Jorgensen JT. Companion and complementary diagnostics: clinical and regulatory perspectives. *Trends Cancer* 2016;2:706–12.
19. Scheerens H, Malong A, Bassett K, et al. Current status of companion and complementary diagnostics: strategic considerations for development and launch. *Clin Transl Sci* 2017;10:84–92.
20. Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. *Nature* 2017;541:321–30.
21. Thorsson V, Gibbs DL, Brown SD, et al. The immune landscape of cancer. *Immunity* 2018;48:812.e14–30.e14.
22. Grizzi G, Caccese M, Gkountakos A, et al. Putative predictors of efficacy for immune checkpoint inhibitors in non-small-cell lung cancer: facing the complexity of the immune system. *Expert Rev Mol Diagn* 2017;17:1055–1069.
23. Schumacher TN, Kesmir C, van Buuren MM. Biomarkers in cancer immunotherapy. *Cancer Cell* 2015;27:12–4.
24. Rizvi NA, Hellmann MD, Snyder A, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 2015;348:124–8.

25. Snyder A, Makarov V, Merghoub T, et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N Engl J Med* 2014;371:2189–99.
26. Van Allen EM, Miao D, Schilling B, et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* 2015; 350:207–11.
27. Carbone DP, Reck M, Paz-Ares L, et al. First-line nivolumab in stage IV or recurrent non-small-cell lung cancer. *N Engl J Med* 2017;376:2415–26.
28. Hellmann MD, Callahan MK, Awad MM, et al. Tumor mutational burden and efficacy of nivolumab monotherapy and in combination with Ipilimumab in small-cell lung cancer. *Cancer Cell* 2018;33:853–61. e4.
29. Rosenberg JE, Hoffman-Censits J, Powles T, et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. *Lancet* 2016;387:1909–20.
30. Hellmann MD, Ciuleanu TE, Pluzanski A, et al. Nivolumab plus Ipilimumab in lung cancer with a high tumor mutational burden. *N Engl J Med* 2018;378:2093–104.
31. Rizvi H, Sanchez-Vega F, La K, et al. Molecular determinants of response to anti-programmed cell death (PD)-1 and anti-programmed death-ligand 1 (PD-L1) blockade in patients with non-small-cell lung cancer profiled with targeted next-generation sequencing. *J Clin Oncol* 2018;36: 633–41.
32. Zehir A, Benayed R, Shah RH, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med* 2017;23:703–13.
33. Campesato LF, Barroso-Sousa R, Jimenez L, et al. Comprehensive cancer-gene panels can be used to estimate mutational load and predict clinical benefit to PD-1 blockade in clinical practice. *Oncotarget* 2015;6:34221–7.
34. Garofalo A, Sholl L, Reardon B, et al. The impact of tumor profiling approaches and genomic data strategies for cancer precision medicine. *Genome Med* 2016;8:79.
35. Roszik J, Haydu LE, Hess KR, et al. Novel algorithmic approach predicts tumor mutation load and correlates with immunotherapy clinical outcomes using a defined gene mutation set. *BMC Med* 2016;14:168.
36. Bristol-Myers Squibb Company and Illumina, Inc.. Bristol-Myers Squibb and Illumina announce strategic collaboration to develop and commercialize companion diagnostics for Bristol-Myers Squibb's oncology immunotherapies; New York & San Diego; 2018 Apr 13. https://news.bms.com/press-release/partnering-news/bristol-myers-squibb-and-illumina-announce-strategic-collaboration-dev. Accessed 19/06/2018.
37. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010;38:e164.
38. Harrow J, Frankish A, Gonzalez JM, et al. GENCODE: the reference human genome annotation for the ENCODE project. *Genome Res* 2012;22:1760–74.
39. Cortes-Ciriano I, Lee S, Park WY, et al. A molecular portrait of microsatellite instability across multiple cancers. *Nat Commun* 2017;8:15180.
40. Yaeger R, Chatila WK, Lipsyc MD, et al. Clinical sequencing defines the genomic landscape of metastatic colorectal cancer. *Cancer Cell* 2018;33: 125–36. e3.
41. Fehrenbacher L, Spira A, Ballinger M, et al. Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial. *Lancet* 2016;387: 1837–46.
42. Ramalingam S, Hellmann M, Awad M, et al. Tumor Mutational Burden (TMB) as a biomarker for clinical benefit From dual immune checkpoint blockade with Nivolumab + Ipilimumab in first-line non-small cell lung cancer: identification of TMB cutoff from CheckMate 568. Paper presented at: American Association for Cancer Research (AACR) Annual Meeting; 2018 Apr 14–18, Chicago, USA.
43. Hellmann MD, Nathanson T, Rizvi H, et al. Genomic features of response to combination immunotherapy in patients with advanced non-small-cell lung cancer. *Cancer Cell* 2018;33: 843–52. e4.
44. Galsky MD, Saci A, Szabo PM, et al. 848PD: impact of tumor mutation burden on nivolumab efficacy in second-line urothelial carcinoma patients: exploratory analysis of the phase ii checkmate 275 study. *Annals Oncol* 2017;28: mdx371.003-mdx371.003.
45. Balar AV, Galsky MD, Rosenberg JE, et al. Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: a single-arm, multicentre, phase 2 trial. *Lancet* 2017;389:67–76.
46. Powles T, Duran I, van der Heijden MS, et al. Atezolizumab versus chemotherapy in patients with platinum-treated locally advanced or metastatic urothelial carcinoma (IMvigor211): a multicentre, open-label, phase 3 randomised controlled trial. *Lancet* 2018;391:748–57.
47. Frampton GM, Fichtenholtz A, Otto GA, et al. Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. *Nat Biotechnol* 2013;31: 1023–31.
48. Oncology Center of Excellence (OCE). FDA grants accelerated approval to pembrolizumab for first tissue/site agnostic indication. Silver Spring, MD: U.S. Food and Drug Administration. 2017 May 23. https://www.fda.gov/Drugs/InformationOnDrugs/ApprovedDrugs/ucm560040.htm. Accessed 19/06/2018.
49. Lemery S, Keegan P, Pazdur R. First FDA approval agnostic of cancer site - when a biomarker defines the indication. *N Engl J Med* 2017; 377:1409–12.
50. Chalmers ZR, Connelly CF, Fabrizio D, et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med* 2017;9:34.
51. McGranahan N, Furness AJ, Rosenthal R, et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* 2016;351:1463–9.
52. Gandara DR, Kowanetz M, Mok TSK, et al. Blood-based biomarkers for cancer immunotherapy: tumor mutational burden in blood (bTMB) is associated with improved Atezolizumab. *Ann Oncol* 2017;28.
53. Gandara DR, Paul SM, Kowanetz M, et al. Blood-based tumor mutational burden as a predictor of clinical benefit in non-small-cell lung cancer patients treated with atezolizumab. *Nat Med* 2018;24:1441–1448.
54. Volckmar AL, Sultmann H, Riediger A, et al. A field guide for cancer diagnostics using cell-free DNA: from principles to practice and clinical applications. *Genes Chromosomes Cancer* 2018;57: 123–39.