Permanent Abstract Number: 3167
Tue 4/2/2019, 8 AM, Poster Section 21, #20

# Improving and Standardizing TMB Assay Performance

Matthew G. Butler[1], Yves Konigshofer[1], Lequan Nguyen[1], Shikha Kalotra[1], Omo Clement[2], Russell K. Garlick[2], and Bharathi Anekella[1]

[1]SeraCare Life Sciences, Inc., Gaithersburg, MD, and [2]SeraCare Life Sciences, Inc., Milford, MA

AACR 2019 March 29 – April 3, 2019 Atlanta, Georgia

## INTRODUCTION

Determining and using the most effective and safest treatment is of great importance in cancer disease management. Checkpoint inhibitors that target immune regulatory molecules such as PD-1, PD-L1, and CTLA-4, have successfully improved PFS and OS - but only in some patients and for some cancers. In the case of PD-1 and PD-L1 inhibitors, it is believed that PD-L1 expression by a solid tumor allows it to escape attack from the immune system and that by inhibiting the PD-L1/PD-1 interaction, immune evasion is no longer possible. This may then cause some of the mutations that give rise to expressed neoantigens to act as targets for T cells and allow for the gradual elimination of the tumor. Since the risk for developing autoimmune adverse events is not insignificant with the use of checkpoint inhibitors, determining which patients are likely to respond favorably to their use is imperative. Similarly, expanding approved uses to new indications also benefits from the identification of populations that are most likely to show improvement in PFS and OS - generally, through clinically-validated biomarkers.

Current indications for the use of PD-1 and PD-L1 inhibitors rely on cancer type and several biomarkers. One of these is the expression (or level of expression) of PD-L1 on the tumor. Another - usually in colorectal cancer - is the presence of microsatellite instability (MSI), which indicates that DNA is not being copied with high fidelity and resulting mutations may lead to the emergence of potential neoantigens. While MSI can be assessed by the PCR amplification of several loci, the presence or absence of neoantigens cannot, and they can also be created in the absence of MSI. This has given rise to a potential biomarker - tumor mutational burden (TMB) - that is an assessment of the number of relevant mutations in a tumor. TMB measurement is challenging since different targeted next-generation sequencing (NGS) panels look at different regions and percentages of the genome and use different criteria for what constitutes a relevant mutation. Presently, there is poor correlation between different TMB assays at mutation levels that may be relevant for a companion diagnostic where patients may be denied treatment.

Here, we describe the characterization of a panel of reference materials for the assessment, harmonization, and improvement of TMB measurements by NGS assays.

## MATERIALS AND METHODS

### TMB Reporting

TMB values here are expressed as likely somatic mutations per 1 mega base pairs (Mbp). Whenever possible, the denominator (Mbp) is the number of genomic locations where a mutation – if present in the data – could be included in the TMB calculation.

### Cell Line DNA

Tumor cell lines along with matched "normal" lymphoblastoid cell lines from the same donors were sourced from repositories such as ATCC (Manassas, VA). Genomic DNA was extracted from cells using Gentra® Puregene® Cell Kits (QIAGEN, Hilden, Germany). Genomic DNA was extracted from cells embedded in FFPE blocks using QIAamp® DNA FFPE Tissue Kits (QIAGEN).

### NGS Analysis

Whole exome sequencing (WES) – the current gold standard for TMB measurement – was performed by fragmenting the genomic DNA and then preparing libraries using SureSelect® Human All Exon v6+COSMIC kits (Agilent Technologies, Santa Clara, CA). WES sequencing was performed on a NovaSeq® (Illumina, San Diego, CA) at Macrogen (Rockville, MD). Resulting FASTQ, VCF, etc. files were analyzed further.

A custom 2 Mbp SureSelect panel (Agilent Technologies) was used to generate additional libraries that were sequenced on a MiSeq® (Illumina).

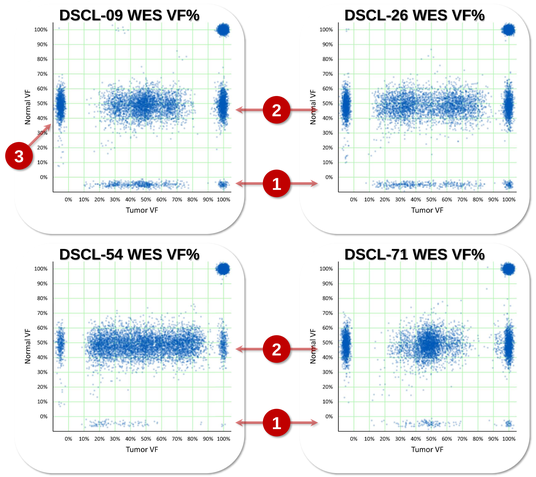
### TMB Analysis

WES data was analyzed using criteria established by the Friends of Cancer Research. Briefly, somatic variants had to be present at a variant allele frequency (VF) of at least 5% and observed in at least three reads at locations sequenced with at least 25 reads. Further, they had to be non-synonymous ("TMB WES Standard") although synonymous variants in coding regions were included for ("TMB WES All"). Additionally, a panel of normals was used to remove variants that may be due to assay and sequencing noise. The overall pipeline used BWA-MEM for sequence alignment (Li, H. and Durbin, R., Bioinformatics, 25, 1754-60), Picard and GATK®4 for data preprocessing (SortSam, AddOrReplaceReadGroups, MarkDuplicates, BaseRecalibrator, ApplyBQSR; Van der Auwera, G. A. et al. Curr. Protoc. Bioinformatics 43,11.10.1-11.10.33), MuTect2 for variant discovery (Cibulskis, K. et al. Nat. Biotechnol. 31, 213–219.), SnpEff for mutation annotation (Cingolani, P. et al. Fly. 6, 80-92.), and SnpSift for variant filtration (Cingolani, P. et al Front. Genet. 3:35)..

Data from the custom 2 Mbp panel was aligned using Bowtie 2 (Langmead, B and Salzberg, SL, Nature Methods, 9(4), 357). Sorted BAM files and pileups were generated with SAMtools (Genome Research Limited, London, UK). Tumor/normal pileups were compared with VarScan® (Koboldt, DC, et al., Genome Research, 22(3), 568-576) using a minimum VF of 1%. The resulting variants were filtered using minimum P values of 1E-6. Variants whose VF were at least 5% in the tumor, at least 25-fold higher in the tumor than normal, and were at locations sequenced with at least 50 reads in both samples were used to calculate TMB. The denominator was the number of bases that had sufficient coverage for comparison – about 2 Mbp.

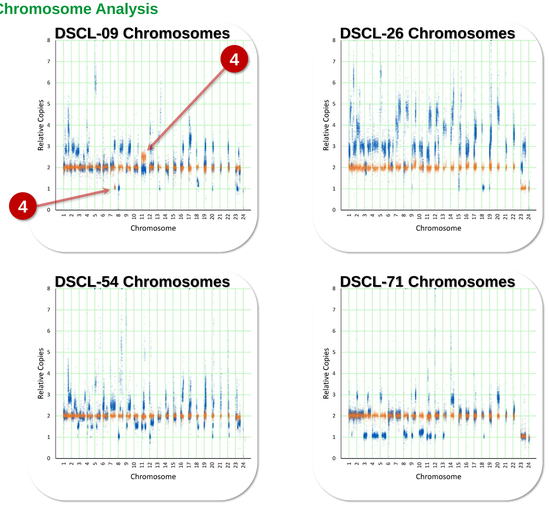
## RESULTS

### TMB by WES



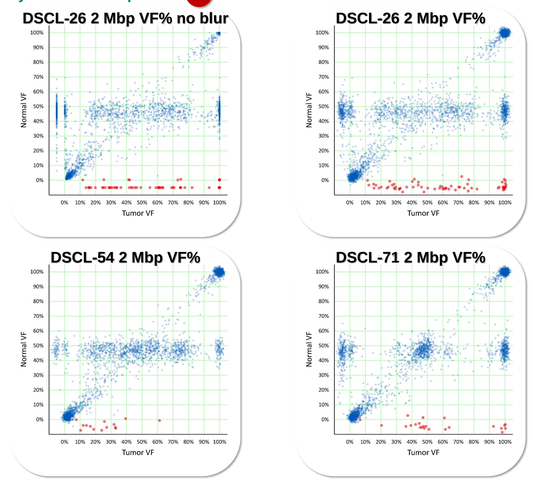

WES data from genomic DNA from tumor and matched normal cell lines was analyzed for variant frequencies in coding regions that were also considered high confidence regions in Genome in a Bottle (GIAB) HG002 v3.3.2. Unique variants were set to center around -5% in the matched cell line. To reveal overlaps, data were randomly blurred +/- 5% VF using a binomial distribution. Somatic variants used for the TMB calculation are found along the horizontal axis.

### Chromosome Analysis




WES data were further analyzed for CNVs by comparing relative coverage within exons and normalizing to the most likely ploidy. Data for the tumor lines are shown in blue. Data for the matched normal lines are shown in orange.

### TMB by Custom 2 Mbp Panel 5




A 2 Mbp panel was used to analyze genomic DNA from tumor/normal matched cell lines. The targets were chosen from genes using GIAB high-confidence regions that also contained no repetitive DNA. Unique variants were set to center around -5% in the matched cell line. To reveal overlaps, data were randomly blurred +/- 5% VF using a binomial distribution – except in the top-left chart. Somatic variants used for the TMB calculation are found along the horizontal axis and are highlighted in red.

### TMB Value Comparison 6

| Cell Line | Type | TMB WES Standard | TMB WES All | TMB 2 Mbp |
|---|---|---|---|---|
| DSCL-09 | Lung Cancer | 25.85 | 33.09 | N/A |
| DSCL-26 | Lung Cancer | 19.93 | 25.98 | 30.43 |
| DSCL-54 | Breast Cancer | 4.35 | 6.00 | 6.79 |
| DSCL-71 | Lung Cancer | 6.84 | 8.63 | 11.04 |

WES and 2 Mbp datasets were analyzed for TMB (somatic mutations per Mbp). For WES data, the "TMB WES Standard" calculation used only non-synonymous variants in coding regions while the "TMB WES All" calculation included all and resulted in ~29% higher values. The "TMB 2 Mbp" calculation included all variants and also included variants in introns. Our WES data were found to be nearly identical to independent WES data for the same matched cell lines (data not shown).

### Genes that were Mutated in Cell Lines 7

| Unique Mutated Genes in WES | Genes Mutated in 2 Lines | Genes Mutated in 3 Lines | Genes Mutated in 4 Lines |
|---|---|---|---|
| 1422 | 94 | 5 | 2 |

The WES data were analyzed further to determine which genes contained non-synonymous somatic mutations and whether there were common ones between the four different cell lines. The majority of genes were found to have non-synonymous somatic mutations in only one cell line, while TP53 (which is commonly mutated in cancer) and PCLO were mutated in all four cell lines. At the same time, based on their TMB scores, the odds of encountering a gene that would be mutated in all four cell lines is around one in a thousand, so PCLO may have just been observed by chance.

## DISCUSSION

1. Different tumor lines contain different distributions of somatic variants. For three of the lines (DSCL-09/26/71), the VFs closely resemble those of germline variants, suggesting that many of the somatic variants appeared prior to aneuploidy. However, in line DSCL-54, the low VF of many somatic variants suggests that they appeared after the emergence of substantial aneuploidy.
2. Tumor lines show signs of aneuploidy as evidenced by heterozygous variants at VF other than 50% and by over- and under-representation of reads coming from different parts of the genome. The presence of substantial aneuploidy in some cancers may present a challenge for TMB assessment with targeted panels that only look at a limited number of genes because having an extra copy of a gene (or more) may increase the likelihood of that gene being mutated somewhere in one copy. At the same time, the duplication of many of the genes in targeted cancer panels may provide a growth advantage (the exceptions being tumor suppressors).
3. Tumor lines also show considerable Loss of Heterozygosity as evidenced by germline variants from the normal lines that are missing in the matched tumor lines.
4. One of the normal lines appears to be missing part of chromosome 7 and containing extra chromosome 11. The fact that the chromosome 7 deletion is present in all lymphoblastoid cells suggests that this appeared in a B cell progenitor. Interestingly, the tumor appears to contain a duplication of a nearby region on chromosome 7.
5. A custom 2 Mbp panel resulted in similar VF distributions as WES, albeit with fewer datapoints. There was significant random noise in both paired samples at and below 5% VF, but unique variants could be identified. Apparent variants on a 0% to 100% VF diagonal were predominantly INDELs that may have appeared due to sequencing errors. The vast majority of unique variants were not INDELs, so omitting indels may improve TMB assessment – except when cells are MSI high and may be more prone to the generation of INDELs. These samples had not been fixed with formalin, but any elevation in background variants of the tumor sample due to fixation artifacts could have led to the emergence of apparent variants and an apparent elevation of TMB.
6. TMB calculated from WES data using all variants in coding regions generally resulted in similar values to TMB calculated from a custom 2 Mbp panel that also included introns. It should be noted that the custom panel used regions of genes from all over the genome and not a targeted subset of genes, which may make such an approach more robust in the presence of aneuploidy. However, this approach also makes comparison to TMB calculated with only non-synonymous variants more difficult. At the same time, a TMB score of 8.63 vs. 11.04 may be sufficiently different to affect patient stratification.
7. A large number of genes were found to be mutated in the four cell lines. While some genes were mutated in coding regions in more than one line, ~93% of genes with non-synonymous mutations were unique. Since different targeted cancer panels look at different genes, differences in sequenced genes may also lead to differences in TMB (in addition to effects due to aneuploidy). For example, TTN (titin) was mutated in two of the four lines in ~108 kbp of coding regions. With six and seven mutations in that gene, respectively, that leads to TMB estimates of 55 and 65. However, this gene is also likely present at more than two copies. There were also no non-synonymous mutations in it in the other two lines.

## CONCLUSIONS

- A panel of four matched tumor/normal cell lines was created that spans a low to high range of TMB values and can also be used in tumor-only workflows. Ideally, orthogonal methods of TMB assessment should obtain similar values.
- TMB values obtained by WES and a custom 2 Mbp panel were generally similar, and may be more similar than data from some targeted cancer panels (data not shown yet). However, ± 1 mutation = ± 0.5 TMB (and ± 1 TMB if only 1 Mbp).
- Given the aneuploid nature of many cancers, assessing TMB by targeted cancer panels may increase discordance to WES TMB due to aneuploidy and sampling.
- Differences between different targeted cancer panels may be significant enough to affect stratification based on TMB values. At the same time, it is unknown whether TMB by WES is the most relevant measure of TMB.