

Compare TMB Scores and MSI Statuses from QIAseq Tumor Mutational Burden Panels

December 20, 2024

Sample to Insight -

QIAGEN Aarhus A/S  $\cdot$  Kalkværksvej 5, 11.  $\cdot$  DK - 8000 Aarhus C  $\cdot$  Denmark digitalinsights.qiagen.com  $\cdot$  ts-bioinformatics@qiagen.com

## Compare TMB Scores and MSI Statuses from QIAseq Tumor Mutational Burden Panels

This tutorial uses the capabilities of CLC Genomics Workbench with the Biomedical Genomics Analysis plugin to calculate reliable and consistent TMB score and MSI status values.

QIAGEN offers the QIAseq Tumor Burden Mutation panel assay (DHS-8800Z) covering a significantly larger region of the genome than other Targeted DNA panels and including gene families, which increases the difficulty of variant calling especially with regards to specificity. Through a series of tools and filters, the Identify TMB Status template workflow has the ability to accurately call variants and to compute a TMB score (total number of somatic mutations in a defined region of a tumor genome) that can be used for classification as TMB low, intermediate or high. High TMB scores have been shown to be correlated with improved patient response rates to immune check-point inhibitors [A. Stenzinger et al., 2019].

In addition, the DHS-8800Z QIAseq Targeted DNA panel covers loci useful for Microsatellite Instability Status inference. This status measures the statistical variation of the length distribution of several microsatellite loci, and compares the statistical variation of the test sample with a normal samples' baseline (included in QIAGEN Reference Data Set available for download in the Workbench). If the proportion of unstable microsatellite loci is higher than the predefined threshold, then the sample is considered unstable. Many patients with MSI-High tumors have had a positive response to immunotherapy treatments. In addition, The US Food and Drug Administration (FDA) has granted accelerated approval to pembrolizumab for pediatric and adult patients with microsatellite instability (MSI)-High, leading to the first approved treatment for any solid tumor with this biomarker irrespective of the tumor's origin [Chang et al., 2018].

This tutorial covers:

- Importing filtered variant tracks and read mappings into the Workbench.
- Running the Calculate TMB Score tool on adjusted target regions to output TMB Score reports for each sample.
- Running the Detect QIAseq MSI Status workflow on these samples for comparison.

We use a subset of the Identify TMB Status template workflow results for 5 samples as example data. Analyses in this tutorial should run within 20 min on a standard laptop. If you wish to run the Identify TMB Status template workflow, which takes longer to run, the original reads can be found on our website.

**Prerequisites** For this tutorial, you must be working with the *CLC Genomics Workbench* with the Biomedical Genomics Analysis plugin installed. How to install plugins is described in the Workbench manual.

**Download and import data** Go through the following steps to download and import the data into the Workbench.

1. Download the sample data from our website and save it on your computer. For each sample, we provide:



- A Read Mapping of the UMI Reads (=)
- A Variant track: The **Filtered variant** track (**P**) includes all somatic variants that remained after the filtering performed in the Identify TMB Status Template workflow and that is input Calculate TMB Score tool.
- 2. Start the workbench.
- 3. Import the reads via the toolbar: Import ((1)) | Standard Import.
- 4. Select the TMB\_MSI\_tutorial.zip file and leave the option to "Automatic import". Click Next.
- 5. Select the location in the Navigation Area of your Workbench where the data should be saved, and click **Finish**.

In addition, you will need to import the Reference Data Set needed to run the tool and workflow demoed in this tutorial.

- 1. Click on the button References in the top right corner of the Workbench.
- 2. In the QIAGEN Sets tab, under the Reference Data Sets section, choose the QIAseq TMB Panels hg38 and click on **Download** (figure 1).

	HIIII a la l	a state and a state of the stat		Manage <sup>2</sup> / M Workspaces - P	anage ugins Reference Da	Manage ta Workflows	Contact Support
🐻 Manage Reference	Data						×
Download Genomes Public Repositories	QIAGEN Sets Reference Data Library	Custom Sets Reference Data Sets	Imported Data	Man Free space in Free space in	age Reference Da CLC_References temporary folder	ta: Locally location: 33 location: 33	↓ 2.52 GB 2.52 GB
QIAseq TMB Panels hg	38 3.0						₹
QIAseq TMB Pane RefSeq MANE v1.3	els hg38	^	QIAseq TMB Par Version: 3.0, Reference	nels hg38 Data Set	Size o 7	n disk 💭 33.0 MB	^
QIAseq Methyl Pa Ensembl MANE v1.3	anels hg38 3		Coj	by from server Download Delete	Create Cust	om Set	
QIAseq Methyl Pa Ensembl v87	anels hg19			Downloading 6 elements	Cancel Pa	use Resume	
QIAseq RNA Pane	els hg38		Reference Data included:	Version	Download Size	On Disk Size	.
QIAseq UPX 3' Pa	nels hg38		Sequence	hg38_no_alt_analysis_set	660.1 MB	688.5 MB	
OlAseg RNA Eusi	on YP Panels ho38		genes genes	refseq_mane_v1.3_hg38_no_alt_analysis_set	1.4 MB	1.4 MB	
RefSeq GRCh38.p14		~	mrna	refseq_mane_v1.3_hg38_no_alt_analysis_set	3.0 MB	3.0 MB	~
Help						CI	ose

Figure 1: Importing the relevant Reference Data Set.

3. Once imported, the Reference data elements included in the set will be available in the CLC\_References folder in the Navigation Area.

#### Run the Calculate TMB Score tool

Calculating TMB score is better done on target regions with coverage 100X. Target regions with insufficient coverage are automatically removed when using the Identify TMB Status template



workflow. However, when using the Calculate TMB Score on its own, it is recommended to adjust target regions files before hand. To remove the low coverage regions of the original target region file:

1. Start the Create Mapping Graph tool from the Tools menu:

Tools | Utility Tools (🔊) | Tracks (🗟) | Graph Tracks (🗟) | Create Mapping Graph (🔩)

and choose the 5 read mappings to be run in batch mode as shown in figure 2.

Create Mapping Graph	Tracks Select a reads track Navigation Area	Selected elements (5)
<ol> <li>Select a reads track</li> <li>Select graph tracks to create</li> <li>Result handling</li> </ol>	Qr <enter search="" term="">         Image: TMB_MSI       - Image: Sil Mapped UMI Reads         Image: Sil Mapped UMI Reads       - Image: Sil Mapped UMI Reads         Image: Sil Mapped UMI Reads       - Image: Sil Mapped UMI Reads         Image: Sil Mapped UMI Reads       - Image: Sil Mapped UMI Reads         Image: Sil Mapped UMI Reads       - Image: Sil Mapped UMI Reads         Image: Sil Mapped UMI Reads       - Image: Sil Mapped UMI Reads         Image: Sil Mapped UMI Reads       - Image: Sil Mapped UMI Reads         Image: Sil Mapped UMI Reads       - Image: Sil Mapped UMI Reads         Image: Sil Mapped UMI Reads       - Image: Sil Mapped UMI Reads         Image: Sil Mapped UMI Reads       - Image: Sil Mapped UMI Reads         Image: Sil Mapped UMI Reads       - Image: Sil Mapped UMI Reads         Image: Sil Mapped UMI Reads       - Image: Sil Mapped UMI Reads         Image: Sil Mapped UMI Reads       - Image: Sil Mapped UMI Reads         Image: Sil Mapped UMI Reads       - Image: Sil Mapped UMI Reads         Image: Sil Mapped UMI Reads       - Image: Sil Mapped UMI Reads         Image: Sil Mapped UMI Reads       - Image: Sil Mapped UMI Reads         Image: Sil Mapped UMI Reads       - Image: Sil Mapped UMI Reads         Image: Sil Mapped UMI Reads       - Image: Sil Mapped UMI Reads         Image: Sil Mapped UMI Reads<!--</th--><th>▼       Image: Single of the system         Image: Single of the system       Image: Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system</th></enter>	▼       Image: Single of the system         Image: Single of the system       Image: Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system         Image: Single of the system       Single of the system
Help Rese	t	Previous Next Finish Cancel

Figure 2: Select the 5 read mappings in batch.

- 2. Use the default setting for the tool to generate read coverage graph for the mappings.
- 3. Save the Mapping Graph Tracks in the Navigation Area.
- 4. Now open the Identify Graph Threshold Areas tool:

Tools | Utility Tools (🔊) | Tracks (🔚) | Graph Tracks (🖳) | Identify Graph Threshold Areas (🔄)

5. Select in batch the Mapping Graph Tracks you just generated (figure 3).

<ul> <li>Identify Graph Thresho</li> <li>1. Choose where to run</li> </ul>	d Areas
<ol> <li>Select graph track</li> <li>Parameters</li> <li>Result handling</li> </ol>	Qv <enter search="" term="">         Image: Search term&gt;       Image: Search term&gt;         Image: Search term&gt;       Image: Search term&gt;</enter>
Help Res	Previous Next Finish Cancel

Figure 3: Select the 5 mapping graph tracks in batch.

6. Set the lower threshold at 100 to exclude regions with coverage below 100x, disable the upper threshold and specify the target region reference track for the TMB panel available in the CLC\_Reference folder (figure 4). You will find it in the QIAseq TMB Panels hg38 folder, in the subfolder named **target\_regions**. Make sure you select the "Reference Data" tab (highlighted in red) to be able to locate easily the reference tracks that belong to a given Reference Data Set. The same tracks are also present in the "Navigation area" tab but organized differently.



. Choose where to run	Parameters	
. Select graph track	Threshold parameters	Navigation Area Reference Data Selected elements (1)
	Window size 1	G Asseq TMB Panels hg38 (RefSeq GRCh38.p12 (no A China
. Batch overview	Use lower threshold	B genes
Parameters	Lower threshold 100.0	🕀 🚔 mrna
		😥 🔂 cds
. Result handling	Use upper threshold	🐵 🔮 dbsnp_tmb
	Upper threshold 1000.0	target_primers     E
		Inset_regions
	Region parameters	DHS-8800Z target regions
	No restriction	B. B mispriming_events
	Evolute apportated	masking_regions
		🕀 🖻 msi_loci 🔍
	<ul> <li>Include annotated only</li> </ul>	
	Region track * DHS-8800Z_target_regions	
		OK Cancel

Figure 4: Set up the tool to generate target region tracks of minimum coverage 100x.

You can now start the Calculate TMB Score tool from the Tools menu:

## Tools | Biomedical Genomics Analysis ( ) | Oncology Score Estimation ( ) | Calculate TMB Score ( )

Because each sample has now its own target region file, we cannot work in batch mode anymore, but you will need to repeat the following steps for each sample:

1. Select a variant track as seen in figure 5.

Choose where to run	Select variant track(s) Navigation Area	Selected elements (1)
Select variant track(s)	Q- <enter search="" term=""></enter>	₩         S1 Variants passing filters
Specify settings	TMB_MSI	
Configure filters	S11 Variants passing filters	
Result handling	S2 Variants passing filters	
	Batch	

Figure 5: Select a variant track to generate the TMB Score report.

2. Now select from the Navigation Area the adjusted Target regions track you just generated for the sample (see figure 6). Then specify the Masking regions track and Exon regions (mRNA track) specific to the DHS-8800Z panel. These files are included in the CLC\_References folder of the Navigation Area once you have download the Reference Data Set. You will find them in the QIAseq TMB Panels hg38 folder, in the subfolders named **masking\_regions** and **mrna**. Again, select the "Reference Data" tab (highlighted in red) to be able to locate easily the reference tracks that belong to a given Reference Data Set.

Also, enable the option "Enable TMB status detection using thresholds". We will leave the default thresholds, as these were chosen based on internal benchmark analyses of lung cancer cell lines and different tissue cancer samples. Given the lack of standardization of methods and the heterogeneity of tumor mutation burden across many tumor types, it is difficult to establish cutoff values. Thresholds should be set according to the samples analysed.

3. In the Configure filters dialog, leave all filters as they are set by default as these have



□ Target regions	
□ Target regions	
Target regions 🚓 S1 Mapped UMI Reads (threshold)	Ø
Exon regions 🚓 Homo_sapiens_refseq_mane_v1.3_RNA Masking regions 🚓 tmb_masking_regions	ର୍ଷ ଷ୍
Detection thresholds           Image: Detection thresholds           Image: TMB status detection using thresholds	
Maximum score for low TMB status     10.0       Minimum score for high TMB status     15.0	
	Exon regions →: Homo_sapiens_refseq_mane_v1.3_RNA Masking regions →: tmb_masking_regions Detection thresholds ☑ Enable TMB status detection using thresholds Maximum score for low TMB status 10.0 Minimum score for high TMB status 15.0

Figure 6: Select the target, exon and masking regions tracks. We also choose to enable the TMB status in this tutorial.

been configured to optimize sensitivity and specificity in selecting variants included in the TMB score calculation. Select from the Reference Data tab the Variant Database **DHS-6600Z-8800Z\_dbsnp\_tmp\_v151\_refseq** from the QIAseq TMB Panels hg38 folder (figure 7)).

Gx Calculate TMB Score		×
1. Choose where to run	Configure filters	
2. Select variant track(s)	Quality filters	
3. Specify settings	Minimum average quality	25.0
4. Configure filters	Minimum QUAL Minimum coverage	100
5. Result handling	Minimum count	2
	Minimum frequency (%)	5.0
	Minimum read position test probability Minimum read direction test probability	0.01
	Caractina Steam	
Sec. 1	Maximum frequency (%) 95.0	
(DEP)	Variant databases	nb_v151_refseq
Sec.	Non-synonymous filters	
011	Non-synonymous filter	
64100 V.C.		
ANEX		
Help Rese	et	Previous Next Finish Cancel

Figure 7: Select the dbsnp variant database to filter out germline variants.

4. In the last dialog, choose the location where you want your reports to be saved and click **Run**.

Redo these steps for each sample, changing the variant track and the corresponding target regions file at each iteration.



The tool will output two files for each sample: the **Variants passing filters (TMB Somatic)** track, containing the variants that were used to calculate the TMB score. The tool applies various quality, germline and non-synonymous filters to identify relevant variants.

The **TMB** report () includes filtering statistics and the calculated TMB score and status. The TMB status is considered low if the TMB score is lower than 10; intermediate if the TMB score is between 10 and 15; and high if the TMB score is larger than 15. Note that these generic category cut-offs are only indicative and should not be used as cancer-specific classifications without proper validation by the user.

In addition to the TMB status, the report lists the length of the combined target regions, counts of various types of variants, and a value assigned to the tumor mutational burden calculated as the number of mutations per Mb. The quality filters statistics recapitulates how many variants were removed by the various filters applied by the tool. Finally, the report shows the frequency distributions of input and somatic variants.

### Run the workflow using QIAseq Panel Analysis Assistent

We will now run the TMB and MSI Panel (DHS-8800Z) workflow on all 5 samples with the QIAseq Panel Analysis Assistent.

1. Open the assistent here:

#### Workflows | Template Workflows | QIAseq Panel Analysis Assistant ()

and choose the Targeted TMB/MSI tab.

2. Click on the arrow to the right of the TMB and MSI Panel (DHS-8800Z) workflow, select the option **Somatic, Illumina (MSI)** as seen on figure 8 and click **Run**.

。 QIAseq Panel Analysis A	Assistant				×
<enter search="" term=""></enter>					₹
Targeted DNA	^	Tumor Mutational Burden Panel (DHS-6600Z)	•	Panel des	cription
Targeted DNA Pro		TMB and MSI Panel (DHS-8800Z) 🔻		Panel des	cription
Targeted DNA Ultra		Somatic, Illumina (TMB) Somatic, Illumina (MSI)			
Targeted Methyl					
Targeted RNAscan					
Targeted RNA					
RNA Fusion XP					
UPX 3' RNA	~				
Help			Close	More	Run

Figure 8: Select the 8800Z panel and choose the MSI workflow.

- 3. Check the batch mode (as highlighted in figure 9) before selecting the read mappings called **Mapped UMI Reads** that should be analyzed before clicking **Next**.
- 4. Finally, in the last wizard step, choose to **Save** the results of the workflow and specify a location in the Navigation Area before clicking **Finish**.



G. Somatic, Illumina (MSI)	for TMB and MSI Panel (DHS-8800Z)				×
1. Choose where to run	Select read mapping <ul> <li>Select from Navigation Area</li> </ul>				
2. Select Workflow Input	O Select files for on-the-fly import:	CLC Format			
3. Result handling	Navigation Area		s	elected elements (5)	
4. Save location for new elements	Q• <enter search="" term=""> TMB MSI →\$\$\$ 511 Mapped U →\$\$\$\$ 13 Mapped U →\$\$\$\$ 13 Mapped U →\$\$\$\$\$ 2 Mapped U ↓ \$\$\$\$\$ (4 Manned II) ♥ Batch</enter>	MI Reads MI Reads AI Reads AI Reads AI Reads		<ul> <li>S11 Mapped UMI Reads</li> <li>S13 Mapped UMI Reads</li> <li>S1 Mapped UMI Reads</li> <li>S2 Mapped UMI Reads</li> <li>S2 Mapped UMI Reads</li> <li>S4 Mapped UMI Reads</li> </ul>	
Help Reset	t		Previous	Next Finish	Cancel

Figure 9: Select the sequencing reads. Remember to run them in batch mode!

The workflow will output 5 MSI reports. You can now open side by side for each sample the TMB and the MSI reports (as we did for S2 in figure 10). For S2 its MSI status is stable, and its TMB score is low.

1 Summary		
Sample name	S2 Variants passing filters	=
TMB status	Low	
Length of target regions (bp)	1,212,980	·
Variants inside target regions and after quality filters	51	
Germline variants	50	
Somatic variants	1	
Non-coding somatic variants	0	
Synonymous somatic variants	0	
Non-synonymous somatic variants	1	
Tumor mutational burden (mutations/Mb)	0.82	
Tumor mutational burden (mutations/Mb	0.82	
Tumor mutational burden (mutations/Mb)	0.82	₽.C
Tumor mutational burden (mutations/Mb Montput: Second Sec	0.82     S2 Mapped UMI Reads ×     omial distribution method. determined with at least 50.0% testable loci.	
Tumor mutational burden (mutations/Mb MSI_report-S2 Mapped UMI Reads × The MSI status was detected using Multin The MSI status of the sample can only be Sample name	0.82         Image: S2 Mapped UMI Reads ×         Image: S2 Mapped UMI Reads ×         Image: S2 Mapped UMI Reads ×         S2 Mapped UMI Reads	
Tumor mutational burden (mutations/Mb MSI_report-S2 Mapped UMI Reads × The MSI status was detected using Multin The MSI status of the sample can only be Sample name Number of loci	0.82 S2 Mapped UMI Reads × omial distribution method. determined with at least 50.0% testable loci. S2 Mapped UMI Reads 120	
Tumor mutational burden (mutations/Mb MSI_report-S2 Mapped UMI Reads × The MSI status was detected using Multin The MSI status of the sample can only be Sample name Number of loci Number of unstable loci	0.82 S2 Mapped UMI Reads × omial distribution method. determined with at least 50.0% testable loci. S2 Mapped UMI Reads 120 17	
Tumor mutational burden (mutations/Mb MSI_report-S2 Mapped UMI Reads × The MSI status was detected using Multin The MSI status of the sample can only be Sample name Number of loci Number of unstable loci Number of stable loci	0.82 S2 Mapped UMI Reads × omial distribution method. determined with at least 50.0% testable loci. S2 Mapped UMI Reads 120 17 97	
Tumor mutational burden (mutations/Mb MSI_report-S2 Mapped UMI Reads × The MSI status was detected using Multin The MSI status of the sample can only be Sample name Number of loci Number of unstable loci Number of stable loci Number of not testable loci	0.82 S2 Mapped UMI Reads × omial distribution method. determined with at least 50.0% testable loci. S2 Mapped UMI Reads 120 17 97 6	
Tumor mutational burden (mutations/Mb MSI_report-S2 Mapped UMI Reads × The MSI status was detected using Multin The MSI status of the sample can only be Sample name Number of loci Number of unstable loci Number of stable loci Percentage of unstable loci	0.82 S2 Mapped UMI Reads × omial distribution method. determined with at least 50.0% testable loci. S2 Mapped UMI Reads 120 17 97 6 14.91	
Tumor mutational burden (mutations/Mb MSI_report-S2 Mapped UMI Reads × The MSI status was detected using Multin The MSI status of the sample can only be Sample name Number of loci Number of unstable loci Number of stable loci Number of not testable loci Percentage of unstable loci MSI status	0.82 S2 Mapped UMI Reads × omial distribution method. determined with at least 50.0% testable loci. S2 Mapped UMI Reads 120 17 97 6 14.91 MSS	

Figure 10: TMB and MSI reports opened simultaneously.

By opening the other samples in a similar way, you will notice that TMB and MSI are not always correlated: Samples 11 and 13 are both MSI-high, but one has a low TMB score (S13), while the other has a high TMB score (S11). The same situation is happening for the MSS samples S1 (low TMB score) and S4 (high TMB score). While the consequences of TMB score and MSI status



on potential therapies are still under investigation, QIAGEN panels and bioinformatics solution for standardized and consistent assessment of both.



## **Bibliography**

- [A. Stenzinger et al., 2019] A. Stenzinger, J. A., Maas, J., Stewart, M., Merino, D., Wempe, M., and Dietel, M. (2019). Tumor mutational burden standardization initiatives:recommendations for consistent tumor mutational burdenassessment in clinical samples to guide immunotherapytreatment decisions. *Genes C*, pages 1–11.
- [Chang et al., 2018] Chang, L., Chang, M., Chang, H., and Chang, F. (2018). Microsatellite instability: A predictive biomarker for cancer immunotherapy. *Applied I*, 26(2):e15–e21.