Digital gene expression analysis of NSCLC-patients reveals strong immune pressure, resulting in an immune escape under immunotherapy

Michael Wessolly1,2*, Susann Stephan-Falkenau3, Anna Streubel3, Marcel Wiesweg4, Sabrina Borchert1,2, Elena Mairinger1, Jens Kollmeier5, Henning Reis1,6, Torsten Bauer5, Kurt Werner Schmid1, Thomas Mairinger3, Martin Schuler2,4 and Fabian D. Mairinger1,2

Abstract

Background: Immune checkpoint inhibitors (ICIs) are currently one of the most promising therapy options in the field of oncology. Although the first pivotal ICI trial results were published in 2011, few biomarkers exist to predict their therapy outcome. PD-L1 expression and tumor mutational burden (TMB) were proven to be sometimes-unreliable biomarkers. We have previously suggested the analysis of processing escapes, a qualitative measurement of epitope structure alterations under immune system pressure, to provide predictive information on ICI response. Here, we sought to further validate this approach and characterize interactions with different forms of immune pressure.

Methods: We identified a cohort consisting of 48 patients with advanced non-small cell lung cancer (NSCLC) treated with nivolumab as ICI monotherapy. Tumor samples were subjected to targeted amplicon-based sequencing using a panel of 22 cancer-associated genes covering 98 mutational hotspots. Altered antigen processing was predicted by NetChop, and MHC binding verified by NetMHC. The NanoString nCounter® platform was utilized to provide gene expression data of 770 immune-related genes. Patient data from 408 patients with NSCLC were retrieved from The Cancer Genome Atlas (TCGA) as a validation cohort.

Results: The two immune escape mechanisms of PD-L1 expression (TPS score) (n = 18) and presence of altered antigen processing (n = 10) are mutually non-exclusive and can occur in the same patient (n = 6). Both mechanisms have exclusive influence on different genes and pathways, according to differential gene expression analysis and gene set enrichment analysis, respectively. Interestingly, gene expression patterns associated with altered processing were enriched in T cell and NK cell immune activity. Though both mechanisms influence different genes, they are similarly linked to increased immune activity.

Conclusion: Pressure from the immune system will lay the foundations for escape mechanisms, leading to acquisition of resistance under therapy. Both PD-L1 expression and altered antigen processing are induced similarly by pronounced immunoactivity but in different context. The present data help to deepen our understanding of the underlying mechanisms behind those immune escapes.

*Correspondence: michael.wessolly@uk-essen.de
1 Institute of Pathology, University Hospital Essen, University of Duisburg-Essen, Hufelandstrasse 55, 45147 Essen, Germany
Full list of author information is available at the end of the article

© The Author(s). 2022. Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.
Introduction

In 2018, 9.6 million people died from cancer or its associated ailments. Based on data from 2015, cancer mortality rates have already surpassed mortality rates from strokes and coronary disease in people below the age of 70 in most western countries. With lung cancer ranking as the deadliest cancer (18.4% of the deaths in 2018) it poses a serious health issue nowadays and in the near future [1–3]. Non-small cell lung cancer (NSCLC) is the most common lung cancer accounting for about 80% of cases. The most prominent subtypes in this category are lung adenocarcinoma (AdC) and lung squamous-cell carcinoma (SCC) [4]. Tobacco exposure is considered the most prominent risk factor for developing lung cancer. It is estimated that 80% of all NSCLC cases are associated with tobacco smoking in various countries including the United States [5].

Recently, immune checkpoint inhibitors (ICIs) have seen frequent use in clinical setups. As a subset of immunotherapy, ICIs are monoclonal antibodies directed against negative regulatory molecules either on immune cells (PD-1 and CTLA-4) or on tumor cells (PD-L1) [6–8]. Though promising results were shown in various clinical studies, the main problem of primary resistance remains. Though significant improvements to clinical benefits have been observed with ICIs, there still remains a large unmet medical need to improve therapy responses [9–11]. Therefore, biomarkers are urgently needed to evaluate patient’s suitability for immunotherapy.

The most frequently discussed biomarker and the only one routinely applied in the clinic is PD-L1 immunohistochemistry. Among available scoring systems, the tumor proportional score (TPS) focusses on PD-L1 expression on tumor cells, while the immune cell (IC) score counts PD-L1 expression on immune cells. Both methods can be summarized with the combined positive score (CPS). Many studies indeed show higher effectiveness of ICIs in patients with high expression levels of PD-L1 (evaluated by TPS scoring) [7, 12–14]. Despite these successes, there are still unexplored caveats concerning PD-L1 expression, some studies displayed equivocal results [15, 16], marking the need for generalized cut-offs regarding the usage of PD-L1 as a biomarker.

NSCLCs are generally considered to have a high rate of somatic mutations compared to other malignancies, thereby increasing the odds of new tumor neoepitopes being generated and presented to immune cells via MHC class I [17–19]. According to neoepitope hypothesis [9, 20, 21] this tumor mutational burden (TMB) is associated with increased tumor immunogenicity and, to a further extend, can serve as a predictive marker for effectiveness of immunotherapy [22–24]. Convincing results suggest that tumors with high TMB respond better to immunotherapy, resulting in the FDA approval of pembrolizumab (a-PD-1 antibody) application in TMB-high solid tumors based on the KEYNOTE-158 study [25, 26]. However, there are still unexplored caveats regarding TMBs usage as a predictive biomarker for immunotherapy. Its clinical utility, particularly in the context of chemo-immunotherapy combinations, has been put into question by several studies [27–31]. In addition, there is no methodological consensus how to define low or high TMB [15, 32, 33]. Though, it should be noted that attempts were made at creating a general applicable assay to determine high and low TMB. Perhaps, the most prominent candidate is the Foundation medicine CDx assay, which was also approved by the FDA and used in the KEYNOTE-158 study [26].

While TMB mainly hints towards the development of potential tumor neoepitopes, it fails to inform which epitopes are generated or if they are bound by the MHC class I complex and could potentially activate immune cells. In previous works we identified altered epitope processing [34, 35] as an important mechanism for tumor immune escape. In particular, those patients characterized by simultaneous PD-L1 expression and high abundance of altered processing showed significantly decreased overall survival.

Study aim

According to the emerging hallmarks of cancer, tumors need to develop tactics to evade the immune system once it exerts a strong selection pressure towards the tumor [36]. This early pressure can be identified by high expression of genes associated with immune response [37]. In this study, we sought to explore if varying shapes of immune pressure cause the development of different escape mechanisms. This may have deep implications for therapies focused on enhancing the immune response.

Material and methods

Demographic data and study design

Forty-eight patients were selected for the current study (Fig. 1), they were either diagnosed with advanced/recurrent lung adenocarcinoma (AdC, n = 23) or lung squamous-cell carcinoma (SCC, n = 25). Diagnostic criteria were based on the World Health Organization (WHO) classification of lung tumors [38]. Between 2012 and
In 2016, all necessary patient data were collected at the Helios Klinikum Emil von Behring, Berlin, Germany. Patients were included, if sufficient follow-up to estimate overall and progression-free survival, adequate amount of tumor material, was available and biomarkers stratifying for targeted therapy were absent. All patients lacked oncogenic drivers in EGFR, ALK and ROS1. PD-L1 levels were determined by immunohistochemical analysis. The appropriate antibody QR-1 was provided by Quaratet, Potsdam, Germany. It has been validated in routine diagnostics to have similar staining activity to E1L3N and 28–8 antibodies. Positive cell detection was defined as membranous stained tumor cells relative to all tumor cells. Tumors were considered PD-L1 positive, if their TPS was at least 1%.

First-line therapy for all patients consisted of chemotherapy. Most patients (45/48) received doublet therapy with one platinum-based component (cis–/carboplatin) and pemetrexed, gemcitabine or vinorelbine. In 28 patients, radiation therapy was applied additionally (Table 1). Each patient received nivolumab (anti-PD-1 antibody) as mono-immunotherapy at least as second-line treatment.

In addition, data from 408 NSCLC cases were downloaded from The Cancer Genome Atlas website (TCGA) [39, 40]. Though whole-exome sequencing was performed with these samples, only the genes covered in the targeted panel for the 48 patients mentioned above were considered for analysis. Many patients in this validation cohort were diagnosed with stage I-II NSCLCs, therefore the preferred therapeutic approach consisted of surgical resection and supportive radio-or chemotherapy.

**Nucleic acid preparation**

Tumor samples were isolated, fixed with formalin and embedded into paraffin (Formalin-fixed paraffin embedded, FFPE). Based on an eosin and hematoxylin stained slide the tumor area was marked and the amount of tumor cells was determined in the target area. FFPE sections were prepared by using the “Microm HM340E” microtome (Thermo Fisher Scientific). Cut tissue slides were stored at $-20\,^\circ\text{C}$ until RNA isolation [3].

Two sections of each FFPE block ($10\,\mu\text{m}$ thickness) were used for semi-automatic isolation with the Maxwell purification system (Maxwell RSC RNA FFPE Kit, AS1440, Promega). The purification was performed according to the manufacturer’s instructions. RNA was eluted in $50\,\mu\ell$ RNase-free water and stored at $-80\,^\circ\text{C}$.

RNA concentration was measured using a Qubit 2.0 fluorometer (Life Technologies) appertaining the RNA broad-range assay. RNA integrity was assessed using a Fragment Analyzer (Advanced Analytical Inc., Ames, IA, USA) appertaining DNF-489 standard sensitivity RNA analysis kit.

**Digital gene expression analysis**

Digital gene expression analysis was performed using the NanoString nCounter® platform (NanoString

---

**Fig. 1** Study design. The figure displays the methodology used within the study procedure.
Technologies, Inc., Seattle, USA) with the NanoString nCounter® PanCancer Immune Profiling Panel. The panel covers 770 genes, which are involved in various immune pathways, including the activation of the innate and adaptive immune response, cell migration and the activity of immune checkpoints, as well as 40 reference genes for biological normalization purposes. Probes were hybridized to 100 ng of total RNA input for 20 h at 65°C and put into the nCounter® PrepStation. The post-hybridization processing was performed by the nCounter® Max/Flex System using the high-sensitivity protocol and the cartridge was scanned and read on the DigitalAnalyzer at 555 FOV [41].

NanoString data processing and normalization
NanoString data processing was done with the R statistical programming environment (v4.0.3) [42]. Considering the counts obtained for positive control probe sets, raw NanoString counts for each gene were subjected to a technical factorial normalization, carried out by subtracting the mean counts plus two-times standard deviation from the CodeSet inherent negative controls. Subsequently, a biological normalization using the included mRNA reference genes was performed. Additionally, all counts with $p > 0.05$ after one-sided $t$-test versus negative controls plus 2x standard deviations were interpreted as not expressed to overcome basal noise [41].

Next generation sequencing and selection of mutations
After DNA isolation on a Maxwell® 16 Research (Promega Corporation, Madison, USA) as recommended in the manufacturer’s protocol, all tumor samples were sequenced using a small panel of 22 genes and 92 amplicons covering hotspots characteristic for NSCLC (Colon Lung v2 AmpliSeq Panel by Thermo Scientific, Waltham, MA, USA).

Non-synonymous mutations with a coverage above 500 reads and an allelic frequency above 3.0% were included into the analysis. Variants with an allelic frequency below 3.0% were filtered out and regarded as artifacts due to formalin fixation. Considering the percentage of tumor cells, the mutations validated needed to be detectable in at least 10% of the tumor sample.

The influence of mutations on proteasomal cleavage was predicted by the machine learning tool NetChop 3.1 [43, 44]. The binding of the resulting epitopes to MHC class I was subsequently simulated by NetMHC 4.0 [45, 46], also based on convolutional neural networks. The whole procedure is described in detail in our previous works [34].

Explorative data analysis
Explorative data analysis was performed in the R programming environment (v 4.0.3) [42]. The Shapiro-Wilk-test was applied to test for normal distribution of the data [47]. For dichotomous variables either the Wilcoxon Mann-Whitney rank sum test (non-parametric) or two-sided students t-test (parametric) was applied [48].
For ordinal variables with more than two groups, either the Kruskal-Wallis test (non-parametric) or ANOVA (parametric) was used to detect group differences. Double dichotomous contingency tables were analyzed using Fisher’s Exact test. To test dependency of ranked parameters with more than two groups, the Pearson’s Chi-squared test was used. Correlations between metric variables were tested by using the Spearman’s rank correlation test as well as the Pearson’s product moment correlation coefficient for linear modeling. The pathway package can visualize the relation of differentially expressed genes to various signaling pathways. The pathway interactions were provided by the Kyoto Encyclopedia of Genes and Genomes (KEGG) [49]. Significant pathway associations were identified by gene set enrichment analysis using the WEB-based GEne SeT AnaLysis Toolkit (WebGestalt) [50, 51]. Each run was executed with 1000 permutations. The selected database containing pathway information was KEGG. Finally, all associations were ranked according to the false discovery rate ($p < 0.05$).

**Results**

**Signs of altered epitope processing were frequently identified in the discovery cohort**

Based on the NetChop analysis, around 40% of all identified non-synonymous mutations were associated with changes in proteasomal processing, which also lead to structural changes of presented epitopes. NetMHC analysis revealed that only 11% (of 366 altered epitopes) were still capable to either bind MHC class I molecules in a sufficient manner or trigger an impactful anti-cancer immune response.

**Differential gene expression analysis**

To ascertain the possible influence of altered processing on the immune response based on mRNA expression data, patients were categorized by PD-L1 expression on tumor cells (TPS > 1%) and evidence of non-synonymous mutations leading to altered epitope processing. Accordingly, we formed three groups, 1) PD-L1 positive tumors ($n = 18$), 2) tumors with altered epitope processing ($n = 10$), and 3) tumors exhibiting both immune escape mechanisms ($n = 6$) and such showing neither mechanism ($n = 14$).

The differentially expressed genes varied strongly between tumors showing an immune escape based on altered epitope processing and those showing PD-L1 expression. 98% and 90% of all differentially expressed genes were specific for each group, respectively (supplemental Figs. 1 and 2 and Fig. 2A). Interestingly, tumors exhibiting both immune escape mechanisms presented a similar gene expression pattern as PD-L1 positive tumors (supplemental Fig. 3, Fig. 2B). Comparing samples with only one mechanism to those presenting both mechanisms, 24% and 35% of genes showed similar expression patterns. 37 (56%) additional genes, which are not shared with the single positive groups, showed specific over-expression in the double-positive group (supplemental Figs. 1, 2 and 3, Fig. 2D).

In the validation cohort, 33% (AdCs) and 62% (SCCs) of all non-synonymous mutations were associated with altered processing. Altered processing significantly affects expression of more genes compared to the discovery cohort ($n = 161$ vs $20$). Still, 45% of their respective genes were overlapping (supplemental Figs. 4 and 16).

**Gene set enrichment analysis**

Gene set enrichment analysis was performed among all three described patient groups in the discovery (Fig. 3A, B, C, supplemental Figs. 5, 6 and 8) and validation cohort (Fig. 3D, supplemental Fig. 7). Gene expression in association with the pathways “MicroRNAs in cancer”, “GnRH signaling pathway” and “Cell cycle” was significantly decreased in patients displaying PD-L1 expression ($p < 0.05$) (Fig. 3A, supplemental Fig. 5), while genes in association with “Primary immunodeficiency”, “Staphylococcus aureus infection” and “Systemic lupus erythematosus” were enriched. “Bladder Cancer”, “GnRH signaling”, “Tight junction interaction”, and “Sphingolipid signaling” were decreased in association with the occurrence of altered processing (Fig. 3B, supplemental Fig. 6).

In contrast, “Endocytosis” and more interestingly the immune pathways “T cell receptor signaling”, “IL-17 signaling”, and “Natural killer cell mediated cytotoxicity” were enriched in association with altered epitope processing. Regarding increased expression, the combined patient group showed an enrichment in “Primary immunodeficiency”, “Staphylococcus aureus infection”, “Systemic lupus erythematosus” (overlap with PD-L1 expression, Fig. 3A, C, supplemental Figs. 5 and 8), and “Endocytosis” (Overlap with altered processing, Fig. 3B, C, supplemental Figs. 6 and 8). Regarding decreased expression, “GnRH signaling pathway” and “MicroRNAs in cancer” overlapped with the PD-L1 group, while “Tight junction”, “Sphingolipid signaling pathway”, and “Bladder cancer” overlapped with altered epitope processing. In addition to that, the combined group was uniquely associated with increased expression in the following pathways: “Complement system and coagulation cascades”, “Cytokine-cytokine receptor interaction”, and “Antigen processing and presentation”. Strikingly, this patient group also displayed less signs of “Apoptosis” (Fig. 3C, supplemental Fig. 8).
Various immunological pathways show signs of upregulation in all patient groups

We primarily observed changes in three immunological pathways which are considered important for anti-tumor immune activity: T cell receptor signaling (Fig. 5), natural killer cell mediated cytotoxicity (Fig. 4) and signaling of T helper cells (types 1 and 2, supplemental Figs. 9, 10, 11 and 12). Tumors with PD-L1 expression as well as those featuring processing escapes exhibited signs of increased gene expression in all three pathways prior to treatment. PD-L1 expression is seemingly associated with upregulated gene expression within those immunological signaling pathways (Figs. 4A and 5A and supplemental Fig. 9). In contrast, the pathway-associated gene expression was also upregulated in patients displaying altered processing (Figs. 4B and 5B and supplemental Fig. 10), but not as prominent. In the validation cohort altered processing was also linked to increased gene expression in
those pathways (Figs. 4C and 5C, supplemental Fig. 11). If patients showed signs of both mechanisms, pathway-associated genes were even more highly upregulated (Figs. 4D and 5D and supplemental Fig. 12).

**Discussion**

Before discussing the main conclusions that could be drawn from our work, the small group sizes need to be outlined. Altered processing occurred only in 10 patients, while PD-L1 overexpression was present in 18 patients. Both mechanisms are present in 6 patients. The compared subgroups are rather imbalanced regarding their size. Therefore, further conclusions need to be drawn carefully.

Just by observing the gene expression profiles of each group they are mostly different from each other (Fig. 2 and supplemental Figs. 1, 2, 3 and 4). Despite the differences outlined by differential gene expression analysis, all groups are seemingly related to overexpression of genes in various immunological pathways. The gene expression associated with natural killer cell mediated immune response, T cell receptor signaling and T helper cell type 1 and type 2 signaling is increased in patients with PD-L1 overexpression or altered processing. The combined group (Figs. 4D and 5D, supplemental Fig. 12) indicates that both mechanisms are similarly induced by strong gene expression in these immunological pathways. However, comparatively less genes are expressed in patients with altered processing (NK cell signaling: 43 vs 23, T cell signaling: 27 vs 15) than in patients with PD-L1 overexpression (Figs. 4A, B and 5A, B, supplemental Figs. 9 and 10).

The patient groups were correlated to the immune signature proclaimed by Ayers, Ribas and McClanahan in 2017 [37]. This signature is linked to T cell-associated inflammation, which leads to interferon gamma secretion (IFNγ) and furthermore to increased PD-L1 expression. The patient group, displaying both escape mechanisms apparently has the strongest correlation to the immune signature (supplemental Figs. 14 and 15). One caveat of this finding is again the above-mentioned low group size. Furthermore, despite the correlation of the combined group being close to significant ($p = 0.05$), the data points of this group and the group with PD-L1 overexpression
strongly overlap. This could be a hint that PD-L1 overexpression exerts the bigger influence within the combined group. However, PD-L1 expression alone does not display the strongest correlation with the immune signature. Only by including the group displaying altered processing they score the highest. This might give us a further hint that when faced with a strong immune response, it might be effective for the tumor to select both mechanisms for an effective immune escape.

PD-L1 is often expressed on somatic cells, in order to regulate an overshooting immune response. Furthermore, increased PD-L1 expression also serves as an immune escape mechanism for tumors, since it allows them to shut down an immune response [10, 11]. However, this also allows PD-L1 to be used as a predictive biomarker for anti-PD-1/PD-L1 immunotherapy [12], which itself counters the downregulation of the immune response induced by PD-L1. Altered epitope processing has previously been investigated by our group [34, 35, 52] and it was suggested as an additional immune escape mechanism, termed processing escape (supplemental Fig. 13). Processing escapes can mechanistically
work in two ways. First, the epitope sequences are prolonged/shortened because mutations change the cleavage patterns of the proteasome. As a consequence, they are unsuitable for MHC binding and are no longer presented on the cell surface. Other research has shown that immune dominant epitopes can be removed from the epitope repertoire by a process called immune editing [9, 53, 54]. Second, the epitopes are still capable of successful MHC class I binding but are less effective in activating T cells due to their altered form.

We assume differences in the temporal aspect of the evolution regarding the two escape mechanisms. PD-L1 expression is an immediate response against pressure by the immune system. It serves as a reactive mechanism by the tumor to induce an immune escape (Figs. 4A and 5A and supplemental Fig. 9). However,
it must be noted that PD-L1 expression is not exclusively induced by a direct immune response. Therefore, the activity of oncological factors like EGFR, Ras, MAPK, EML4-ALK, MET and PI3K-Akt as well as IFNγ and HIF-1 may also be associated with increased PD-L1 expression. Despite displaying high PD-L1 expression, patients do not necessarily show durable therapy responses upon ICB application [55–57]. This should also be taken into consideration if PD-L1 as a biomarker is concerned. However, this may not fully apply to our study as patients should lack other signs of immune activity as well. Contrarily to this, we found many hints of immune activity based on gene expression analysis. The expression of cytokines, complement factors and factors associated with the complement system, antigen processing and antigen presentation are increased in the risk group (Fig. 3C). Effector molecules of NK cells like granymes and perforins are also expressed (Fig. 4). T cell receptor activity is high (Fig. 5), while the Ras-MAPK cascade lacks expression if patients show either PD-L1 expression, signs of altered epitope processing or both (Fig. 4A, B, D). These examples point towards an immune activity within the tumor, which makes immune activity as the major driver of PD-L1 expression more likely. This is especially true when considering that patients were negative for oncogenic drivers in EGFR, ALK and ROS1.

Processing escapes, on the other hand, are forced by natural selection, requiring multiple generation cycles. Therefore, they indeed seem like a slow, more adaptive approach to combat the immune response (Figs. 4B and 5B and supplemental Fig. 10).

There are relevant consequences that could be drawn from the study in regards to immune escape mechanisms. PD-L1 is already used as a biomarker, and high PD-L1 expression supports the use of ICI monotherapy. While anti-PD-1/PD-L1 therapy may seem effective at first, the underlying mechanism of processing escapes may lead then to acquired resistance, since both mechanisms are synergistic. It should also be noted that in our study, tumor biopsies were taken before initiation of therapy. Their gene expression signature already shows high immune pressure from the beginning, which leads to the induction of PD-L1 expression and processing escapes later down the line. Therefore, both mechanisms may serve as a combined biomarker already at initial diagnosis.

**Conclusion**

Pressure of the immune system during carcinogenesis lays the foundations for escape mechanisms and future resistance to therapy. Despite different genes influencing each respective mechanism (PD-L1 expression and processing escapes), it seems both can be shaped by a strong immune response, which results in simultaneous activation of both and thereby detrimental clinical outcomes as demonstrated in previous works. Identification of the underlying mechanisms of immune silencing may improve patient selection for immunotherapy.

**Abbreviations**

ADC: Adenocarcinoma; ANOVA: Analysis of variance; AS: Amino acid; COXPH: Cox proportional-hazards model; CTL: Cytotoxic T-lymphocyte; CTLA-4: Cytotoxic T-lymphocyte-associated Protein 4; DDCT: Double dichotomous contingency tables; DNA: Deoxyribonucleic acid; ER: Endoplasmic reticulum; FDR: False discovery rate; FFPE: Formalin-fixed paraffin-embedded tissue; HLA: Human leucocyte antigen; H&E stain: Hematoxylin and eosin stain; ICI: Immune checkpoint inhibitor; IC50: Half maximal inhibitory concentration; IEDB: Immune epitope data base; IFNγ: Interferon gamma; I/O: Immune‑oxygen; Mb: Megabase; MHC: Major histocompatibility complex; mRNA: Messenger ribonucleic acid; NGS: Next-Generation Sequencing; NSCLC: Non-small cell lung cancer; OS: Overall survival; PD-L1: Programmed cell death protein 1; PD-L1: Programmed cell death 1 ligand 1; SCC: Squamous-cell carcinoma; SNP: Single nucleotide polymorphism; TAP: Transporter associated with antigen processing; TCGA: The Cancer Genome Atlas; TCR: T cell receptor; TMB: Tumor mutational burden; TH1: T helper cell type 1; TH2: T helper cell type 2; UICC: Union internationale contre le cancer.

**Supplementary Information**

The online version contains supplementary material available at https://doi.org/10.1186/s12885-021-09111-w.

**Additional file 1: Suppl. Figure 1.** Differential gene expression analysis of all patients showing increased PD-L1 expression. The log-fold changes between each state (PD-L1 positive or negative) are plotted against the p-value, displaying significant differences in expression between each state. Grey: Differential gene expression does not differ significantly between each state. Red: Top 10 most differentially expressed genes (p < 0.01).

**Suppl. Figure 2.** Differential gene expression analysis of all patients showing signs of altered epitope processing. The log-fold changes between each state (altered processing is present or not) are plotted against the p-value, displaying significant differences in expression between each state. Grey: Differential gene expression does not differ significantly between each state. Red: Top 10 most differentially expressed genes (p < 0.05).

**Suppl. Figure 3.** Differential gene expression analysis of all patients showing signs of altered epitope processing in the validation cohort. The log-fold changes between each state (altered processing is present or not) are plotted against the p-value, displaying significant differences in expression between each state. Grey: Differential gene expression does not differ significantly between each state. Red: Top 10 most differentially expressed genes (p < 0.01).

**Suppl. Figure 4.** Differential gene expression analysis of all patients showing signs both altered epitope processing and PD-L1 expression. The log-fold changes between each state (both mechanisms are present or not) are plotted against the p-value, displaying significant differences in expression between each state. Grey: Differential gene expression does not differ significantly between each state. Red: Top 10 most differentially expressed genes (p < 0.01).

**Suppl. Figure 5.** To gain further insight into which biological processes are affected by differential gene expression in patients showing PD-L1 expression, a gene set enrichment analysis was performed. The gene ontology (GO) analysis is part of the greater enrichment analysis. The analysis was performed to estimate the correlation between a patient group/escape mechanism and certain biological functions (red chart), cellular components (blue chart) and molecular functions (green chart).

**Suppl. Figure 6.** To gain further insight into which biological processes are affected by differential gene expression in patients showing signs of altered epitope processing (discovery cohort), a gene set enrichment analysis was performed.
The gene ontology (GO) analysis is part of the greater enrichment analysis. The analysis was performed to estimate the correlation between a patient group/escape mechanism and certain biological functions (red chart), cellular components (blue chart) and molecular functions (green chart).

**Suppl. Figure 7.** To gain further insight into which biological processes are affected by differential gene expression in patients displaying both mechanisms (PD-L1 expression and altered epitope processing), a gene set enrichment analysis was performed. The gene ontology (GO) analysis is part of the greater enrichment analysis. The analysis was performed to estimate the correlation between a patient group/escape mechanism and certain biological functions (red chart), cellular components (blue chart) and molecular functions (green chart).

**Suppl. Figure 8.** To gain further insight into which biological processes are affected by differential gene expression in patients displaying both mechanisms (PD-L1 expression and altered epitope processing), a gene set enrichment analysis was performed. The gene ontology (GO) analysis is part of the greater enrichment analysis. The analysis was performed to estimate the correlation between a patient group/escape mechanism and certain biological functions (red chart), cellular components (blue chart) and molecular functions (green chart). **Suppl. Figure 9.** KEGG pathway analysis of T helper cell (subtype 1 and 2) differentiation in patients expressing PD-L1. The plots were created via the pathway package in R. Genes are either strongly expressed (red) or their expression is reduced (green).

**Suppl. Figure 10.** KEGG pathway analysis of T helper cell (subtype 1 and 2) differentiation in patients showing signs of altered epitope processing (validation cohort). The plots were created via the pathway package in R. Genes are either strongly expressed (red) or their expression is reduced (green).

**Suppl. Figure 11.** KEGG pathway analysis of T helper cell (subtype 1 and 2) differentiation in patients showing signs of altered epitope processing (validation cohort). The plots were created via the pathway package in R. Genes are either strongly expressed (red) or their expression is reduced (green).

**Suppl. Figure 12.** KEGG pathway analysis of T helper cell (subtype 1 and 2) differentiation in patients showing both signs of altered epitope processing and PD-L1 expression. The plots were created via the pathway package in R. Genes are either strongly expressed (red) or their expression is reduced (green).

**Suppl. Figure 13.** Activation of cytotoxic lymphocytes by tumor neoepitopes under immune checkpoint therapy. The figure additionally highlights the role of altered epitope processing. Through high mutational load, mutations change protosomal cleavage patterns leading to structural changes or disruption of the original epitope. Regardless, the immunogenicity of the tumor neoepitopes is lowered. The T cell does not become active and immune checkpoint inhibition is rendered ineffective, since it cannot promote weak or absent signaling.

**Suppl. Figure 14.** Expression of IFN gamma associated genes (based on [37]) in association with patients expressing PD-L1 or showing signs of altered processing. FALSE/TRUE: Patients displaying neither mechanism (green), TRUE/FALSE: Patients display signs of altered epitope processing, but no signs of PD-L1 expression (red). FALSE/TRUE: Patients displaying signs of PD-L1 expression, but without any signs of altered epitope processing (blue). TRUE/TRUE: Patients show signs of both mechanisms (violet). NS: Not significant. p = 0.05926.

**Suppl. Figure 15.** Expression of an extended mRNA based immune signature (based on [37]) in association with patients expressing PD-L1 or showing signs of altered processing. FALSE/TRUE: Patients displaying neither mechanism (green), TRUE/FALSE: Patients display signs of altered epitope processing, but no signs of PD-L1 expression (red). FALSE/TRUE: Patients displaying signs of PD-L1 expression, but without any signs of altered epitope processing (blue). TRUE/TRUE: Patients show signs of both mechanisms (violet). NS: Not significant. p = 0.05623.

**Suppl. Figure 16.** The discovery cohort (“Sample”, red) and the validation cohort (dark red) were compared regarding gene expression in association with altered epitope processing.

**Authors’ contributions**

Conceptualization, MW, FM. Methodology, MWE, SB, and FM. Software, MWE and FM. Validation, MWE, FM, SSF, AS, MWI. Formal analysis, MWE, FM, SSF. Investigation, FM, HR, EM, BS, SSF, AS. Resources, TB, KWS, TM, MS. Data curation, MWE, JK, SSF, MWI. Writing–original Draft preparation, MWE, FM. Writing–review and editing. All listed authors. Visualisation, MWE and FM. Supervision, FM. Project administration, MS, FM. Funding acquisition, HR, MS, FM. All authors have read and agreed to the published version of the manuscript.

**Funding**

Open Access funding enabled and organized by Projekt DEAL. This work was funded by Bristol Myers Squibb through its International Immuno–Oncology Network.

**Availability of data and materials**

Data regarding this work will be made available upon a reasonable request to the corresponding author.

**Declarations**

**Ethics approval and consent to participate**

The study was conducted retrospectively and was approved by the Ethics Committee of the Medical Faculty of the University Duisburg-Essen (identifier: 13-5382-BO). The investigations conform to the principles of the declaration of Helsinki. All patient data were anonymized to make sure that their identity cannot be assumed. The Ethics Committee of the Medical Faculty of the University Duisburg-Essen waived the need for a written, informed consent as it was no longer necessary due most patients being deceased at the time of data collection. Furthermore, it was no longer possible to obtain the consent retrospectively, as all patient data have been anonymized.

**Consent for publication**

Not applicable.

**Competing interests**

Marcel Wiesweg reports honoraria from Boehringer Ingelheim, Novartis, Roche and Takeda and research funding from Bristol Myers Squibb and Takeda. Jens Kollmeier reports a consulting and advisory role without personal honoraria for Roche, Boehringer Ingelheim, Bristol Myers Squibb, MSD and Takeda. Henning Reis reports a consulting and advisory role for Bristol Myers Squibb; honoraria from Roche and Bristol Myers Squibb; travel support from Philips, Roche and Bristol Myers Squibb; research funding from Bristol Myers Squibb and share ownership from Bayer. Martin Schuler reports consultancy for AstraZeneca, Boehringer Ingelheim, Bristol Myers Squibb, Celgene, Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG), Lilly and Novartis; honoraria for CME presentations from Alesion, Boehringer Ingelheim, Celgene, GlaxoSmithKline, Lilly and Novartis; research funding to the institution from Roehringer Ingelheim, Bristol Myers Squibb and Novartis and other support from Universität Duisburg-Essen (patents). All remaining authors declare no conflict of interest.

**Author details**

1 Institute of Pathology, University Hospital Essen, University of Duisburg-Essen, Hufelandstrasse 55, 45147 Essen, Germany. 2 German Cancer Consortium (DKTK), Partner Site University Hospital Essen, Hufelandstrasse 55, 45147 Essen, Germany. 3 German Cancer Center, University Hospital Essen, Hufelandstrasse 55, 45147 Essen, Germany. 4 Department of Tissue Diagnostics, Helios Klinikum Emil von Behring, Berlin, Germany. 5 Department of Medical Oncology, West German Cancer Center, University Hospital Essen, Hufelandstrasse 55, 45147 Essen, Germany. 6 Dr. Senckenberg Institute of Pathology, University Hospital Frankfurt, Goethe University Frankfurt, Frankfurt, Germany.

**Received:** 24 February 2021  **Accepted:** 14 December 2021  **Published online:** 07 January 2022
References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424.

2. Omran AR. The epidemiologic transition: a theory of the epidemiology of population change. Milbank Mem Fund Q. 1971;49(4):509–38.

3. Bray F, Jemal A, Grey N, Ferlay J, Forman D. Global cancer transitions according to the Human Development Index (2008-2030): a population-based study. Lancet Oncol. 2012;13(8):790–801. https://doi.org/10.1016/S1470-2045(12)70211-5. Epub 2012 Jun 1.

4. Molina Jr, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. Mayo Clin Proc. 2008;83(5):584–94.

5. Alberg AJ, Brock MV, Ford JG, Samet JM, Spivack SD. Epidemiology of lung cancer: diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. Chest. 2013;143(5 Suppl):e209S–e295S.

6. Brahamer J, Reckamp KL, Baas P, Eberhardt WE, Poddubskaya E, et al. Nivolumab versus docetaxel in advanced squamous-cell non-small-cell lung cancer. N Engl J Med. 2015;373(2):123–35.

7. Herbst RS, Sanabria J, Ramalingam S, Kim DW, Felip E, Hrushesky WJ, et al. Evolution of neoantigen landscape during immune checkpoint blockade in non-small cell lung cancer. Cancer Discov. 2017;7(3):264–76.

8. Fehrenbacher L, Spira A, Ballinger M, Kowanetz M, Vansteenkiste J, Mazerie B, Muraro L, 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(10027):1540–50.

9. Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A, et al. Impact of DNA damage response and repair (DDR) gene mutations on response to immunotherapy in diverse cancers. Mol Cancer Ther. 2018 Sep 4.

10. Riaz N, Havel JJ, Makarov V, Desrichard A, Urba WJ, Sims JS, et al. Tumor and microenvironment evolution during immunotherapy with Pembrolizumab. Cell. 2017;171(4):939–e416.

11. Fares CM, Van Allen EW, Drake CG, Allison JP, Hu-Lieskovan S. Mechanisms of resistance to immune checkpoint blockade: why does checkpoint inhibitor immunotherapy not work for all patients? Am Soc Clin Oncol Educ Book. 2019;39:147–64.

12. Herbst RS, Morgenstern D, Boshoff C. The biology and management of lung cancer: results from the phase II KEYNOTE-158 study. J Clin Oncol. 2020;38(1):1–10.

13. Aurélien FM, Fakh J, Lopez M, Shah R, Shapira-Frommer K, Nakagawa HC, Chung Hedy L, Kindler JA, Lopez-Martin WH, Miller A, Italiano S, Kao SA, Piña-Paul J-P, Delord RR, McWilliams DA, Fabrizio D, Aurora-Garg L, Xu F, Jin K, Norwood YJ. Association of tumour mutational burden with outcomes in patients treated with pembrolizumab: prospective biomarker analysis of the multicohort open-label phase 2 KEYNOTE-158 study. Lancet Oncol. 2020;21(10):1535–65. https://doi.org/10.1016/S1470-2045(20)30445-9.

14. Rizvi NA, Cho BC, Reimnuth N, Lee KH, Luft A, Ahn M-J, et al. Durvalumab with or without tremelimumab vs standard chemotherapy in first-line treatment of metastatic non-small-cell lung cancer: the MYS-TIC phase 3 randomized clinical trial. JAMA Oncol. 2020;6(5):661–74.

15. Hellmann MD, Paz-Ares L, Bernabe Caro R, Zurawski B, Kim S-W, Carcereny Costa E, et al. Nivolumab plus ipilimumab in advanced non–small-cell lung cancer. N Engl J Med. 2019;381(21):2020–31.

16. Marabelle A, Fakh M, Lopez J, Shah M, Shapira-Frommer R, Nakagawa K, et al. Association of tumour mutational burden with outcomes in patients with selected advanced solid tumours treated with pembrolizumab in KEYNOTE-158. Ann Oncol. 2019;30:v477–8.

17. Paz-Ares L, Rango S, Novello C, Salmos B, Cheng Y, Gadgeel S, et al. LBA80Pembrolizumab (pembrol) plus platinum-based chemotherapy (chemo) for metastatic NSCLC. tissue TMB (tTMB) and outcomes in KEYNOTE-021, 189, and 407. Ann Oncol. 2019;30:v177–8.

18. Stenzinger A, Endris V, Budczies J, Merkelbach-Bruse S, Kazdal D, Dietmaier W, et al. Harmonization and standardization of panel-based tumor mutational burden measurement: real-world results and recommendations of the quality in pathology study. J Thorac Oncol. 2020;15(7):1177–89.

19. Samstein RM, Lee C-H, Shouhurst AN, Hellmann MD, Shen R, Janjigian YY, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. Nat Genet. 2019;51(2):202–6.

20. Willis C, Fiander M, Tran D, Kytosky B, Thomas J-M, Calderon F, et al. Tumor mutational burden in lung cancer: a systematic literature review. Oncotarget. 2019;10(61):6604–22.

21. Wessolly M, Stepahan-Falkenau S, Streubeil A, Werner R, Borchert S, Griff S, et al. Novel epitope quality-based immune escape mechanism reveals patient’s suitability for immune checkpoint inhibition. Cancer Manag Res. 2020;12:7881–90.

22. Wessolly M, Walter RFH, Vollbrecht C, Werner R, Borchert S, Schmeller J, et al. Processing escape mechanisms through altered proteasomal cleavage of epitopes affect immune response in pulmonary neuroendocrine tumors. Technol Cancer Res Treat. 2018;17:1533038118814818.

23. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646–74.

24. Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR, et al. A novel epitope quality-based immune escape mechanism of mCRC reveals patient’s suitability for immune checkpoint inhibition. Cancer Manag Res. 2020;12:8381–90.

25. Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JHM, Beasley MB, et al. The 2015 World Health Organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification. J Thorac Oncol. 2015;10(9):1243–60.
39. Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. Nature. 2012;489(7417):519–25. https://doi.org/10.1038/nature11404. Epub 2012 Sep 9. Erratum in: Nature. 2012;491(7563):288. Rogers, Kristen [corrected to Rodrigues, Kristen].

40. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. Nature. 2014;511(7511):S43–50. https://doi.org/10.1038/nature13385. Epub 2014 Jul 9. Erratum in: Nature. 2014;514(7521):262. Rogers, K [corrected to Rodrigues, K]. Erratum in: Nature. 2018;559(7715):E12.

41. Mairinger F, Bankfalvi A, Schmid KW, Mairinger E, Mach P, Walter RF, et al. Digital immune-related gene expression signatures in high-grade serous ovarian carcinoma: developing prediction models for platinum response. Cancer Manag Res. 2019;11:9571–83.

42. Team RC. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2020.

43. Kesmir C, Nussbaum AK, Schild H, Detours V, Brunak S. Prediction of proteasome cleavage motifs by neural networks. Protein Eng. 2002;15(4):287–96.

44. Nielsen M, Lundegaard C, Lund O, Keymir C. The role of the proteasome in generating cytotoxic T-cell epitopes: insights obtained from improved predictions of proteasomal cleavage. Immunogenetics. 2003;57(1–2):33–41.

45. Nielsen M, Lundegaard C, Worning P, Lauemøller SL, Lamberth K, Buus S, et al. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. Protein Sci. 2003;12(5):1007–17.

46. Andreatta M, Nielsen M. Gapped sequence alignment using artificial neural networks: from the MHC class I system. Bioinformatics (Oxford, England). 2016;32(4):511–7.

47. Royston JP. Algorithm AS 181: the W test for normality. J R Stat Soc: Ser C: Appl Stat. 1982;31(2):176–80.

48. Bauer DF. Constructing confidence sets using rank statistics. J Am Stat Assoc. 1972;67(339):687–90.

49. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27–30.

50. Zhang B, Kirov S, Snoddy J. WebGestalt: an integrated system for exploring gene sets in various biological contexts. Nucleic Acids Res. 2005;33(Web Server issue):W741–8.

51. Liao Y, Wang J, Jæhning EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. Nucleic Acids Res. 2019;47(W1):W199–w205.

52. Wessolly M, Mairinger F. Processing escapes: a new perspective on immune escape mechanisms; 2017. p. 555662.

53. Nejo T, Matsushita H, Karasaki T, Nomura M, Saito K, Tanaka S, et al. Reduced Neoantigen expression revealed by longitudinal multiomics as a possible immune evasion mechanism in glioma. Cancer Immunol Res. 2019;7(7):1148–61.

54. Vesely MD, Schreiber RD. Cancer immunoediting: antigens, mechanisms, and implications to cancer immunotherapy. Ann N Y Acad Sci. 2013;1284(1):1–5.

55. Yi M, Jiao D, Xu H, Liu Q, Zhao W, Han X, et al. Biomarkers for predicting efficacy of PD-1/PD-L1 inhibitors. Mol Cancer. 2018;17(1):129.

56. Ribas A, Hu-Lieskovan S. What does PD-L1 positive or negative mean? J Exp Med. 2016;213(13):2835–40.

57. Yi M, Niu M, Xu L, Luo S, Wu K. Regulation of PD-L1 expression in the tumor microenvironment. J Hematol Oncol. 2021;14(1):10.

**Publisher’s Note**
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.