

# Plasma cell-free DNA hydroxymethylation profiling reveals anti-PD-1 treatment response and resistance biology in non-small cell lung cancer

Gulfem D Guler , Yuhong Ning, Ceyda Coruh, Giuliana P Mognol, Tierney Phillips, Maryam Nabiyouni, Kyle Hazen, Aaron Scott, Wayne Volkmuth, Samuel Levy 

**To cite:** Guler GD, Ning Y, Coruh C, *et al.* Plasma cell-free DNA hydroxymethylation profiling reveals anti-PD-1 treatment response and resistance biology in non-small cell lung cancer. *Journal for ImmunoTherapy of Cancer* 2024;**12**:e008028. doi:10.1136/jitc-2023-008028

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/jitc-2023-008028>).

Accepted 30 November 2023



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

ClearNote Health Inc, San Diego, California, USA

## Correspondence to

Dr Samuel Levy;  
slevy@clearnotehealth.com

## ABSTRACT

**Background** Treatment with immune checkpoint inhibitors (ICIs) targeting programmed death-1 (PD-1) can yield durable antitumor responses, yet not all patients respond to ICIs. Current approaches to select patients who may benefit from anti-PD-1 treatment are insufficient. 5-hydroxymethylation (5hmC) analysis of plasma-derived cell-free DNA (cfDNA) presents a novel non-invasive approach for identification of therapy response biomarkers which can tackle challenges associated with tumor biopsies such as tumor heterogeneity and serial sample collection.

**Methods** 151 blood samples were collected from 31 patients with non-small cell lung cancer (NSCLC) before therapy started and at multiple time points while on therapy. Blood samples were processed to obtain plasma-derived cfDNA, followed by enrichment of 5hmC-containing cfDNA fragments through biotinylation via a two-step chemistry and binding to streptavidin coated beads. 5hmC-enriched cfDNA and whole genome libraries were prepared in parallel and sequenced to obtain whole hydroxymethylome and whole genome plasma profiles, respectively.

**Results** Comparison of on-treatment time point to matched pretreatment samples from same patients revealed that anti-PD-1 treatment induced distinct changes in plasma cfDNA 5hmC profiles of responding patients, as judged by Response evaluation criteria in solid tumors, relative to non-responders. In responders, 5hmC accumulated over genes involved in immune activation such as interferon (IFN)- $\gamma$  and IFN- $\alpha$  response, inflammatory response and tumor necrosis factor (TNF)- $\alpha$  signaling, whereas in non-responders 5hmC increased over epithelial to mesenchymal transition genes. Molecular response to anti-PD-1 treatment, as measured by 5hmC changes in plasma cfDNA profiles were observed early on, starting with the first cycle of treatment. Comparison of pretreatment plasma samples revealed that anti-PD-1 treatment response and resistance associated genes can be captured by 5hmC profiling of plasma-derived cfDNA. Furthermore, 5hmC profiling of pretreatment plasma samples was able to distinguish responders from non-responders using T cell-inflamed gene expression profile, which was previously identified by tissue RNA analysis.

## WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ While immune checkpoint inhibitor (ICI) treatment can be very effective, it does not work for all patients. Non-invasive methods to predict and monitor ICI response with increased precision can improve clinical implementation of ICI treatments for better patient care.

## WHAT THIS STUDY ADDS

⇒ 5-hydroxymethylation (5hmC) analysis of plasma-derived cell-free DNA (cfDNA) revealed differential profiles for anti-programmed death-1 (PD-1) responders relative to non-responders prior to therapy start as well as after initiation of treatment. Furthermore, our results demonstrate that anti-PD-1 response/resistance genes, previously identified by others through tissue gene expression analysis, can be captured by 5hmC analysis in plasma samples without the need for any tumor biopsy.

## HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ 5hmC analysis of plasma-derived cfDNA provides a novel method for non-invasive prediction and monitoring of ICI response in non-small cell lung cancer.

**Conclusions** These results demonstrate that 5hmC profiling can identify response and resistance associated biological pathways in plasma-derived cfDNA, offering a novel approach for non-invasive prediction and monitoring of immunotherapy response in NSCLC.

## BACKGROUND

Lung cancer is the most common cancer worldwide. It is also the leading cause of cancer-related deaths globally, estimated to claim 1.8 million lives annually.<sup>1</sup> Most patients with lung cancer are diagnosed with advanced disease.<sup>2</sup> 85–90% of all lung cancer cases are characterized as non-small cell lung cancer (NSCLC).<sup>2</sup> Among the treatment options available for NSCLC, immunotherapy has

cemented its place in the management of advanced stage disease, achieving better patient outcomes in both first line setting and beyond, as mono or combination therapy.<sup>3,4</sup>

Immune checkpoint inhibitors (ICIs) targeting programmed death-1 (PD-1) and programmed death-ligand 1 (PD-L1) have achieved remarkable success not only in lung cancer but across 17 different cancer types.<sup>5</sup> ICIs work by reversing tumor microenvironment-induced T-cell inhibition and thereby restoring antitumor immunity.

While ICIs can elicit remarkably effective and durable responses, only a subset of patients show benefit.<sup>6</sup> Measurement of PD-L1 expression in tumors was initially adopted in the clinic as a patient selection biomarker for anti-PD-1.<sup>7</sup> Yet, PD-L1 expression frequently fails to identify all patients who can benefit from ICI treatment as demonstrated by non-response in PD-L1 expressing patients, as well as significant overall survival benefit observed irrespective of PD-L1 expression.<sup>8,9</sup> Additional biomarkers that have emerged to improve patient selection for ICIs include assessment of tumor-infiltrating lymphocytes,<sup>10</sup> T-cell-inflamed gene-expression profile,<sup>11,12</sup> and tumor mutational burden (TMB).<sup>13</sup> Combinatorial use of such biomarkers serves as molecular readouts for the various components of the cancer-immunity cycle,<sup>14</sup> thereby providing mechanistic insights that can indeed enrich for ICI-responding patients. However, the current array of biomarkers not only fall short in accurately identifying ICI responders, but also present practical challenges associated with availability of sufficient tumor biopsy material, tumor heterogeneity as well as technical complexities for standardization of assay analysis and interpretation. Combination of multiple biomarker assays to improve accuracy of predicting responders also increases process complexity and cost. Furthermore, the requirement for invasive tumor biopsy in patients can prove challenging for immunotherapy response monitoring. Therefore, non-invasive and robust biomarkers which can improve ICI patient selection and response monitoring to avoid unnecessary ICI treatment in patients lacking response remain to be established.<sup>3,15</sup>

Liquid biopsy approaches using plasma-derived cell-free DNA (cfDNA) present key advantages for biomarker discovery and development in ICI response prediction and monitoring setting. Circulating cfDNA is composed of DNA fragments found in the blood that originate from genomes of dying cells of different tissues and blood cells, reflecting cell turnover under normal/healthy conditions as well as altered homeostasis caused by disease.<sup>16</sup> Unlike traditional tissue biopsies that can only sample locally and at accessible tumor sites, plasma-derived cfDNA can capture tumor heterogeneity with circulating tumor DNA originating from single as well as multiple tumor sites. Liquid biopsies also provide access to cfDNA originating from the tumor microenvironment and immune cells.<sup>17,18</sup> Moreover, non-invasive approaches like liquid biopsies allow collection of serial samples over time, enabling

dynamic monitoring of treatment response and resistance as well as disease recurrence.<sup>19–21</sup>

Liquid biopsy applications for immunotherapy response prediction have so far largely focused on detection of circulating tumor DNA (ctDNA), the majority of which relies on technologies that measure variant allele frequency of mutations detected a priori in tumor samples.<sup>22</sup> ctDNA levels of pretreatment samples were found to be prognostic in the context of multiple treatment modalities including immunotherapy.<sup>21,23</sup> Changes in ctDNA levels after treatment was also shown to track with response.<sup>21,23</sup> Furthermore, TMB measurements from plasma have been associated with immunotherapy response.<sup>24–26</sup> Overall, limitations on ctDNA detection, whether technological or biological, as is the case for low ctDNA-shedding tumors, present challenges for wider clinical implementation. Unlocking the full potential of liquid biopsies for immunotherapy response prediction requires incorporation of response and resistance-associated biological signals that originate from not only tumor cells but also other cell types such as the immune cells in the tumor microenvironment and even in the larger context of immune system.

Epigenetic modifications in cfDNA, such as 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC), present an opportunity to incorporate signals from tumor and other cell types that influence response to therapy. While widely investigated for liquid biopsy-based cancer detection,<sup>27–30</sup> epigenetic modifications, in particular 5hmC, remain underused in liquid biopsy approaches for immunotherapy response prediction and monitoring. 5hmC is a stable epigenetic mark that originates from oxidation of 5mC by the ten-eleven translocation family dioxygenases.<sup>31</sup> Unlike 5mC, 5hmC has been shown to generally mark transcriptionally permissive chromatin state and is positively correlated with transcriptional activation, particularly for tissue-specific and cell type-specific genes.<sup>32,33</sup> In the context of the immune system, dynamic changes in 5hmC profiles were observed to be important for immune cell differentiation and function.<sup>34,35</sup> As such, 5hmC signatures in plasma-derived cfDNA provide a rich medium for biomarker discovery for various applications spanning from early detection to treatment selection and response monitoring.

Here, we investigated serially obtained plasma 5hmC profiles of patients with NSCLC who received anti-PD-1 therapy to identify putative predictive biomarkers and biomarkers for treatment monitoring. Anti-PD-1 treatment induced 5hmC changes over thousands of genomic loci, which were distinct in responding and non-responding patient groups as revealed from on therapy serial 5hmC profiles. In responding patients, 5hmC accumulated over genes that are associated with interferon gamma response and inflammation on anti-PD-1 treatment. In non-responders, 5hmC accumulation was observed over genes associated with epithelial-mesenchymal transition. 5hmC-based changes over these regions were also observed at earlier time points than the radiologic response time point

for both responders and non-responders. Comparison of 5hmC profiles of plasma obtained from responding and non-responding patients prior to any anti-PD-1 treatment revealed a large number of genomic loci, mostly mapping to genic regions, that correlated with response status. These differential 5hmC loci identified at pretreatment time point were enriched for inflammatory response genes in patients who responded to anti-PD-1 treatment. Significantly, mean 5hmC levels were higher in responders relative to non-responders over the 18 genes that define T-cell inflamed signature, which was previously determined to aid in anti-PD-1 response using tissue gene expression analysis.<sup>11 12 36</sup> Overall, our study shows that plasma 5hmC profiles of patients with NSCLC can serve as a valuable molecular signal to identify biomarkers for anti-PD-1 treatment response prediction and monitoring.

## METHODS

### Clinical cohort and study design

This study was performed using plasma obtained from subjects with NSCLC who underwent pembrolizumab or nivolumab monotherapy. 31 patients were enrolled at participating sites in Germany with written informed consent for use of blood specimens for archival storage and retrospective analyses. Total of 151 blood samples were collected as approved by the Institutional Review Boards (IRBs) responsible at each site. The study protocol submission, IRB approval and specimen handling across all sites were managed by Indivumed GmbH (Hamburg, Germany), who provided clinical research specimen handling services.

### Plasma collection

Whole blood specimens were collected by routine venous phlebotomy in Cell-Free DNA BCT tubes according to the manufacturer's protocol (Streck, La Vista, Nebraska, USA) (<https://www.streck.com/collection/cell-free-dna-bct/>). Tubes were maintained at 15°C to 25°C until plasma isolation. Plasma was isolated within 24 hours of phlebotomy by centrifugation of whole blood at 1600×g for 10 min at room temperature, followed by transfer of the plasma layer to a new tube for centrifugation at 1600×g for 10 min. Plasma was then aliquoted and stored at −80°C.

### cfDNA isolation

cfDNA was isolated from plasma samples using the MagMAX Cell-Free DNA Isolation Kit (Thermo Fisher Scientific, Waltham, Massachusetts, USA) following the manufacturer's protocol with automated run on HAMILTON STAR liquid handlers (HAMILTON Company, Reno, Nevada, USA). All cfDNA eluates were quantitated by Molecular Devices's SpectraMax Plate Readers using the PicoGreen dsDNA quantitation assay (Thermo Fisher Scientific, Waltham, Massachusetts, USA). TapeStation 4200 capillary electrophoresis (Agilent Technologies, Santa Clara, California, USA) was

employed to ensure the absence of contaminating high molecular weight DNA emanating from white blood cell lysis.

### 5hmC enrichment assay and 5hmC/whole genome sequencing library preparation

cfDNA was normalized to 10 ng total input for each assay and ligated to sequencing adapters. 5hmC bases were biotinylated via a two-step chemistry and subsequently enriched by binding to Dynabeads M270 Streptavidin (Thermo Fisher Scientific, Waltham, Massachusetts, USA). 5hmC enriched fragments were then processed to obtain 5hmC libraries. Whole genome libraries were prepared in parallel from the same input cfDNA material without the 5hmC enrichment steps. All libraries were quantitated by Molecular Devices's SpectraMax Plate Readers using the PicoGreen dsDNA quantitation assay (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and normalized to 1 ng/μL prior to pooling. Library pools were quantitated by Qubit dsDNA High Sensitivity Assay (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and normalized in preparation for sequencing.

### DNA sequencing and alignment

DNA sequencing was performed according to manufacturer's recommendations with 75 base-pair, paired-end sequencing using a NovaSeq instrument with V.2 reagent chemistry (Illumina, San Diego, California, USA). Data was collected using NovaSeq System Suite V.2.2.04. Raw data processing and demultiplexing was performed using V.2.20.0.422 of Illumina's bcl2fastq software to generate sample specific FASTQ output. Sequencing reads were aligned to the hg38 reference genome using BWA-MEM2 (V.2.2.1) with default parameters.<sup>37</sup> Sequencing data quality was assessed using Picard.<sup>38</sup>

### Peak detection

BWA-MEM2 (V.2.2.1) read alignments were employed to identify regions or peaks of dense read accumulation that mark the location of a hydroxymethylated cytosine residue. Prior to identifying peaks, binary alignment map (BAM) files containing the locations of aligned reads were filtered for poorly mapped (mapping quality (MAPQ) < 30) and not properly paired reads using SAMtools.<sup>39</sup> 5hmC peak calling was carried out using MACS2 with default p value < 1.00e-5 threshold for narrow peak settings as previously described.<sup>27</sup> Identified 5hmC peaks residing in "blacklist regions" as defined elsewhere (<https://sites.google.com/site/anshulkundaje/projects/blacklists>) and residing on chromosomes X, Y and mitochondrial genome were also removed using bedtools.<sup>40</sup> Computation of genomic feature enrichment overlapping 5hmC peaks was performed using the software HOMER (<http://homer.ucsd.edu/homer/>) with default parameters.

### Differential 5hmC peak analysis

For the purpose of identifying peaks with differential 5hmC representation, first, a bed file with a union of all



identified peaks was generated and overlapping peaks were merged. Raw counts over the peaks were normalized by transforming raw counts to CPM (counts per million). Weakly represented peaks (CPM>2 in less than 10 samples) were removed before differential analysis. To identify differential 5-hydroxymethylation induced on immunotherapy treatment, a Wilcoxon signed-rank test, which is a non-parametric paired test used for comparison of two groups measurement, was used to compare plasma collected at baseline time point with the plasma collected while the patients were on treatment at time of radiologic response. Wilcoxon rank-sum test, which is a non-parametric test for unpaired two groups measurement, was applied to select 5hmC regions with the most significant differences between responder and non-responder cohorts by comparing fold change in 5hmC occupancy normalized to baseline, termed  $\zeta$ , as calculated by  $((T_R - T_0)/T_0)$ . Regions with p value<0.05 and a delta  $\zeta$  of at least 1.5 between responders and non-responders were retained for further analysis. 5hmC-based molecular response ( $MR_{5hmC}$ ) was calculated by leave-one-out cross validation method using the equation:

$$MR^{5hmC} = \mu_x - \mu_y$$

where  $\mu_x$  and  $\mu_y$  are the mean of all  $x_i$ ,  $i=1, \dots, n_i$ , and  $y_j$ ,  $j=1, \dots, n_j$ , respectively. The following formula was used to compute both  $x_i$  and  $y_j$ :

$$\frac{T_Q - T_0}{T_0}$$

where genomic loci with increased 5hmC is positively and negatively correlated with treatment response are used for  $x_i$  and  $y_j$ , respectively, while  $T_0$  and  $T_Q$  denote CPM for pretreatment and CPM for on-treatment time points, respectively. Genomic loci  $x_i$  and  $y_j$  were determined by the training set that included all the samples except the samples from the left-outpatient, which were then independently tested using the equation described above. This process was iterated 31 times to independently calculate 5hmC molecular response scores for each patient.  $MR_{5hmC}$  scores were calculated blinded of the patient response.

### Differential 5hmC gene analysis

To reliably identify gene bodies with differential 5hmC representation between the tumor and normal groups, we followed this procedure: (1) remove genes that map to sex chromosomes; (2) remove genes that are weakly represented by requiring >3CPM in at least 10 samples); (3) normalize the gene representation distributions using TMM (trimmed mean of M values). We used the R package edgeR<sup>41</sup> for differential analysis and used Benjamini's and Hochberg's procedure to account for multiple hypothesis testing and to quantify false discovery rate (FDR). In this report, we use adjusted p value and FDR interchangeably.

### Gene set enrichment analysis

mSigDB gene sets that were enriched in 5hmC representation was determined using gene set enrichment analysis

(GSEA) software<sup>42</sup> by analysis of preranked gene list based on fold change in 5hmC CPM over gene bodies.

## RESULTS

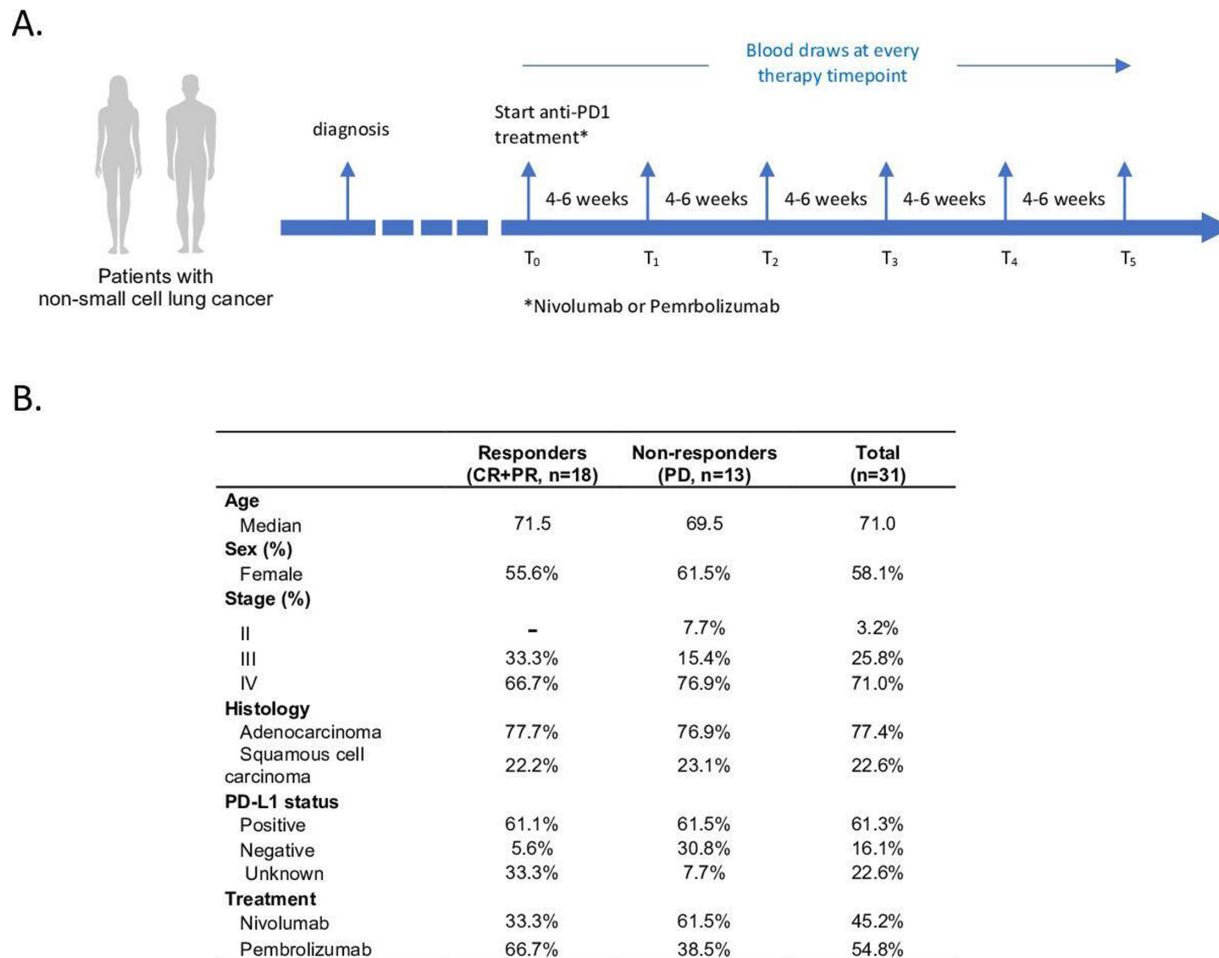
### Clinical cohort and study design

We assembled a cohort of 31 patients with NSCLC who received anti-PD-1 immunotherapy agents pembrolizumab or nivolumab as monotherapy to investigate whether 5hmC profiles of plasma-derived cfDNA could inform on immunotherapy response, (figure 1A). The median age for the study cohort was 71, with female subjects representing 58.1% of the cohort (figure 1B). Majority of the patients were patients with late stage cancer (96.8%) with adenocarcinomas and squamous cell carcinomas constituting 77.4% and 22.6% of the study cohort, respectively. Blood samples were collected at baseline prior to starting treatment and after commencement of treatment in 4–6 week intervals (figure 1A). The complete cohort consisted of 31 patients and 151 blood samples in total, with 3–6 blood samples collected for each patient prior to anti-PD-1 therapy administration at each treatment cycle (figure 1A). Response to therapy was evaluated by radiologic imaging. Of the 31 patients, 18 were partial or complete responders (CR+PR) per RECIST (Response Evaluation Criteria in Solid Tumors) V.1.1, while 13 patients showed progressive disease (PD) (figure 1B).

cfDNA samples isolated from plasma were subjected to 5hmC enrichment assay. Libraries prepared from 5hmC-enriched cfDNA were sequenced along with whole genome libraries which were sequenced at low pass as input controls. Genomic regions enriched for 5hmC were identified by using a peak detection algorithm employing read density. The number of peaks discovered per million unique reads in a given sample ranged from 2,102 to 10,090 with a median of 5,635 (online supplemental figure 1A). The majority of peaks localized to introns and intergenic regions, 60% and 30%, respectively (figure 2A). Consistent with previous publications,<sup>20–24</sup> 5hmC peaks were most statistically enriched over genic regions as determined by the ratio of 5hmC bases in each genomic feature relative to the total number of bases in the same feature genome-wide (online supplemental figure 1B).

### Responders and non-responders show distinct hydroxymethylation changes in cancer and immune response genes after anti-PD-1 treatment

To examine whether there were any changes in plasma-derived 5hmC profiles after anti-PD-1 treatment, we first performed paired differential analysis using 5hmC fragment counts per gene by comparing the time point of radiologic response ( $T_R$ ) to baseline time point that was collected before treatment start ( $T_0$ ). Patients who had PD, as determined by radiologic

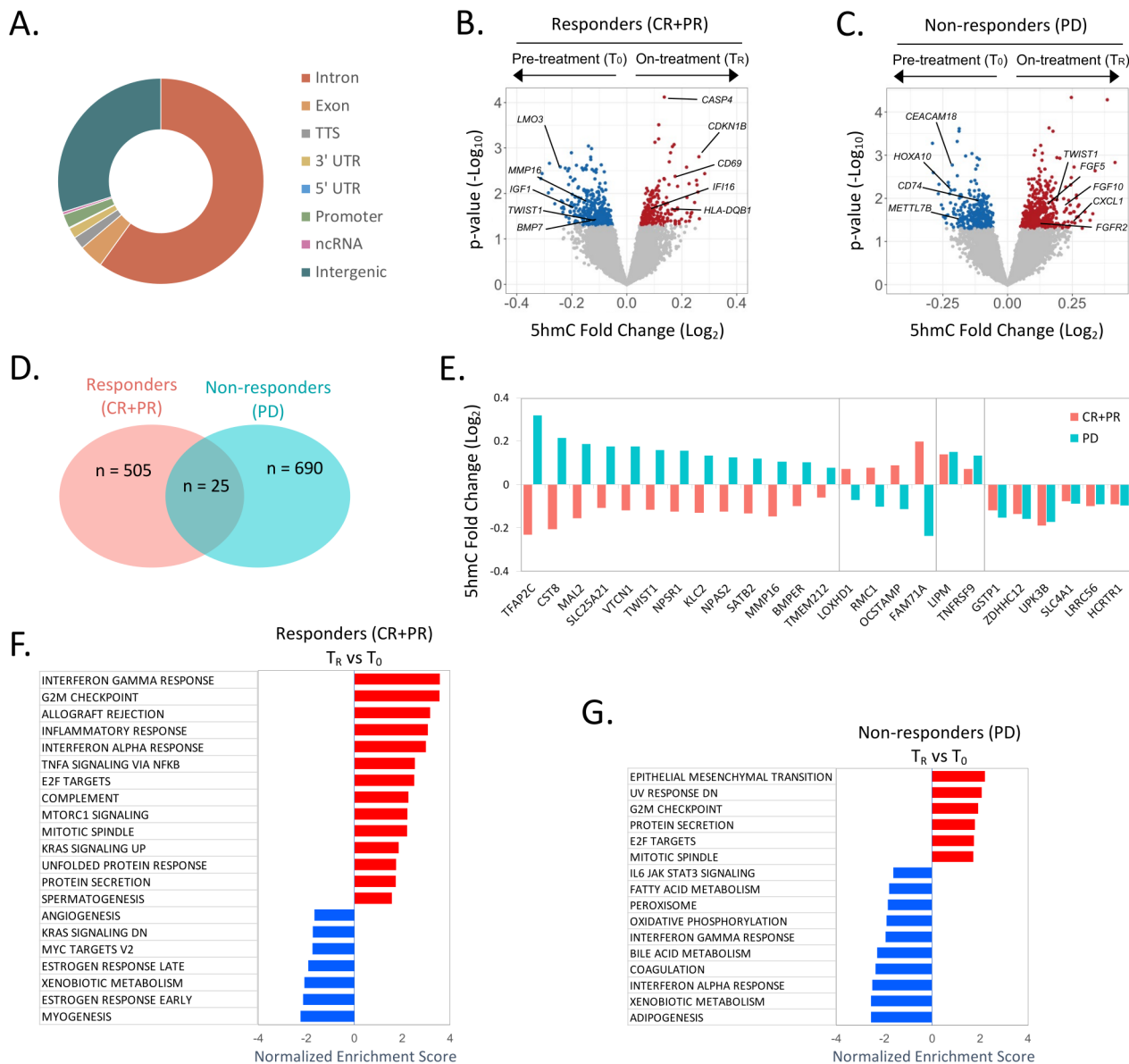


**Figure 1** Study design and clinical characteristics of study cohort. (A) Schematic showing study design. (B) Clinical characteristics of the patient cohort. CR, complete responders; PD, progressive disease; PD-1, programmed death-1; PR, partial responders.

imaging according to RECIST V.1.1 criteria, were categorized as non-responders, whereas CR and PR were grouped as responders. This differential analysis yielded 530 genes with a  $p$  value  $< 0.05$  in responders (figure 2B). Among the genes with increased 5hmC counts were genes associated with tumor immunity or immune cell activation such as *HLA-DQB1* and *CD69*. In contrast, genes with decreased 5hmC counts were enriched for cancer or drug resistance genes such as *IGF1*, *TWIST1* and *MMP16*. In non-responding patients, a similar comparison of on-treatment samples from time of T<sub>R</sub> to paired pretreatment (T<sub>0</sub>) samples identified 715 genes with differential 5hmC (figure 2C). In contrast to responding patients, genes with increased 5hmC included genes that are associated with resistance to checkpoint inhibition based on previous studies of tissue samples, such as fibroblast growth factor (FGF) signaling,<sup>43</sup> while genes with decreased 5hmC included *CD74* with known functions in immune response.<sup>44</sup> Among the genes differentially hydroxymethylated after treatment, only 25 genes were shared between responders and non-responders (figure 2D). The majority of these

genes (17 out of 25) displayed opposite trends between responders and non-responders (figure 2E). Genes that increased in 5hmC in non-responders while decreasing in responders included multiple cancer-associated genes such as *MMP16* and *TWIST1*. Altogether, the minimal overlap between the differentially hydroxymethylated genes in non-responder and responder cohorts suggests that these changes detected in plasma cfDNA after treatment are specific to anti-PD-1 response/resistance.

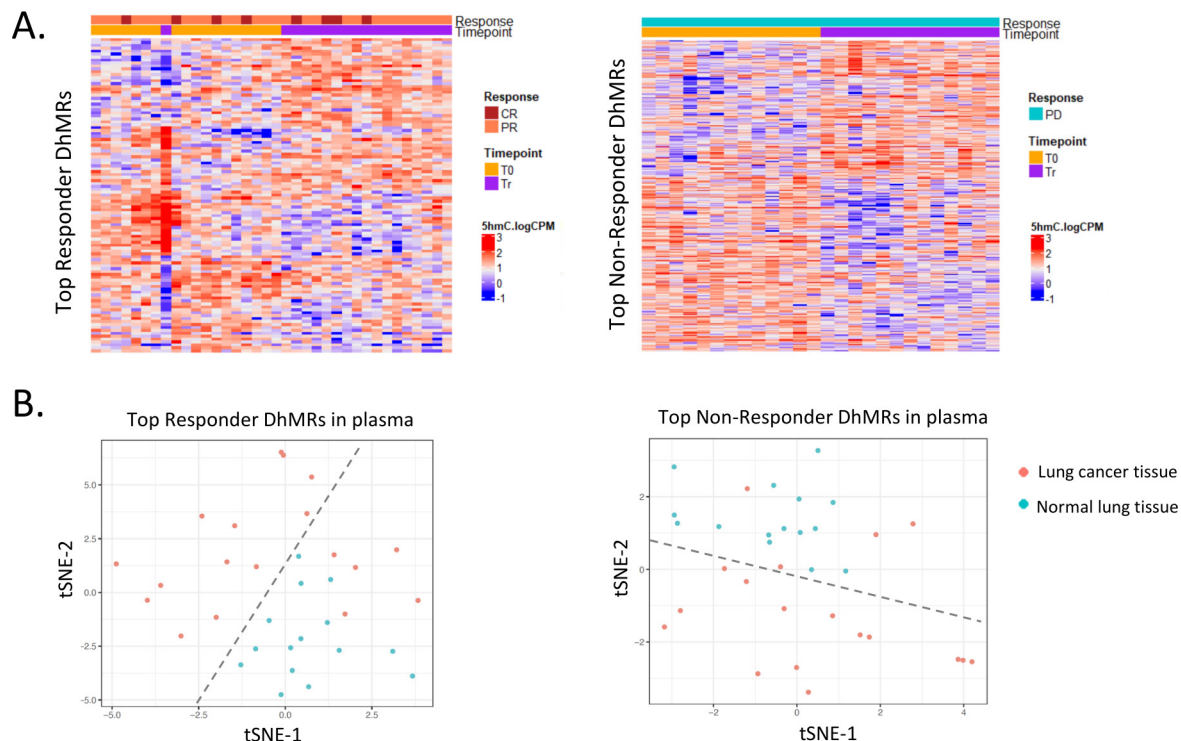
To determine whether 5hmC profiles in plasma-derived cfDNA can inform about the biological response to anti-PD-1 treatment, we performed GSEA using 5hmC counts over gene bodies. In responders, the top enriched pathways were heavily immune-related; such as interferon gamma (IFN- $\gamma$ ) response, inflammatory response, interferon alpha (IFN- $\alpha$ ) response and tumor necrosis factor (TNF)- $\alpha$  signaling via nuclear factor (NF)- $\kappa$ B (figure 2F). IFN- $\gamma$  response, the most significantly enriched pathway in 5hmC analysis of plasma cfDNA, was previously associated with anti-PD-1 response by gene expression analysis of tumor tissue samples



**Figure 2** Distinct changes in plasma-derived cell-free DNA 5hmC profiles on anti-programmed death-1 treatment in responders and non-responders. (A) 5hmC peak overlap with gene and other genomic annotations. (B,C) Volcano plot comparing 5hmC counts over genes at time of response ( $T_R$ ) to pretreatment ( $T_0$ ) in responders (CR+PR) (B) or non-responders (PD) (C). Hypo-hydroxymethylated or hyper-hydroxymethylated genes in on-treatment samples relative to pretreatment with nominal p value  $> 0.05$  are shown in blue or red, respectively. (D) Genes with differential 5hmC counts on treatment relative to pretreatment. (E) 5hmC fold change over the genes identified in both responders and non-responders to have differential hydroxymethylation. (F) GSEA using 5hmC counts over genes to compare baseline ( $T_0$ ) to  $T_R$  in responding patients reveals pathways associated with activation of immune response. (G) GSEA using 5hmC counts over genes to compare baseline ( $T_0$ ) to  $T_R$  in non-responding patients reveals pathways associated with resistance to therapy. CR, complete responders; GSEA, gene set enrichment analysis; ncRNA, non coding ribonucleic acid; PD, progressive disease; PR, partial responders; TTS, transcription termination site; '3'UTR, 3-prime untranslated region; 5'UTR, 5-prime untranslated region; 5hmC, 5-hydroxymethylcytosine.

taken at baseline from patients who went on to be treated with pembrolizumab.<sup>11 12</sup> Among the leading-edge genes that accounted for the enrichment of these immune gene sets were lymphocyte activation marker *CD69* and IFN inducible transmembrane proteins coding genes *IFITM2* and *IFITM1*. For non-responding patients, there were fewer gene sets that significantly changed on treatment. Among the top

enriched gene sets identified in non-responders was epithelial mesenchymal transition (EMT) gene set (figure 2G) with leading-edge subset that included mesenchymal markers *CDH11* and *CDH2*. EMT is known to be associated with resistance to several drugs, including ICIs.<sup>45 46</sup> These findings suggest that inflamed T-cell and EMT signatures found in the tumor microenvironment of responders and



**Figure 3** Anti-programmed death-1 treatment-induced 5hmC changes in plasma-derived cell-free DNA in patients with responding and non-responding non-small cell lung cancer. (A) Heatmaps showing 5hmC fragment counts (logCPM) over top DhMRs identified in responding (left panel) and in non-responding patients (right panel), as determined by thresholding with  $p$  value  $< 0.05$  and  $\log FC > 1.5$ . (B) t-SNE plots showing segregation of lung cancer tissue samples from normal lung tissue controls using top DhMRs identified in plasma samples from responders (left panel) and non-responders (right panel). CPM, counts per million; CR, complete responders; DhMRs, hydroxymethylated regions; FC, fold change; PD, progressive disease; PR, partial responders; t-SNE, t-distributed stochastic neighbor embedding.

non-responders, respectively, can be identified in plasma cfDNA by 5hmC analysis.

#### Differentially hydroxymethylated loci identified in cfDNA can separate normal lung from lung cancer tissue

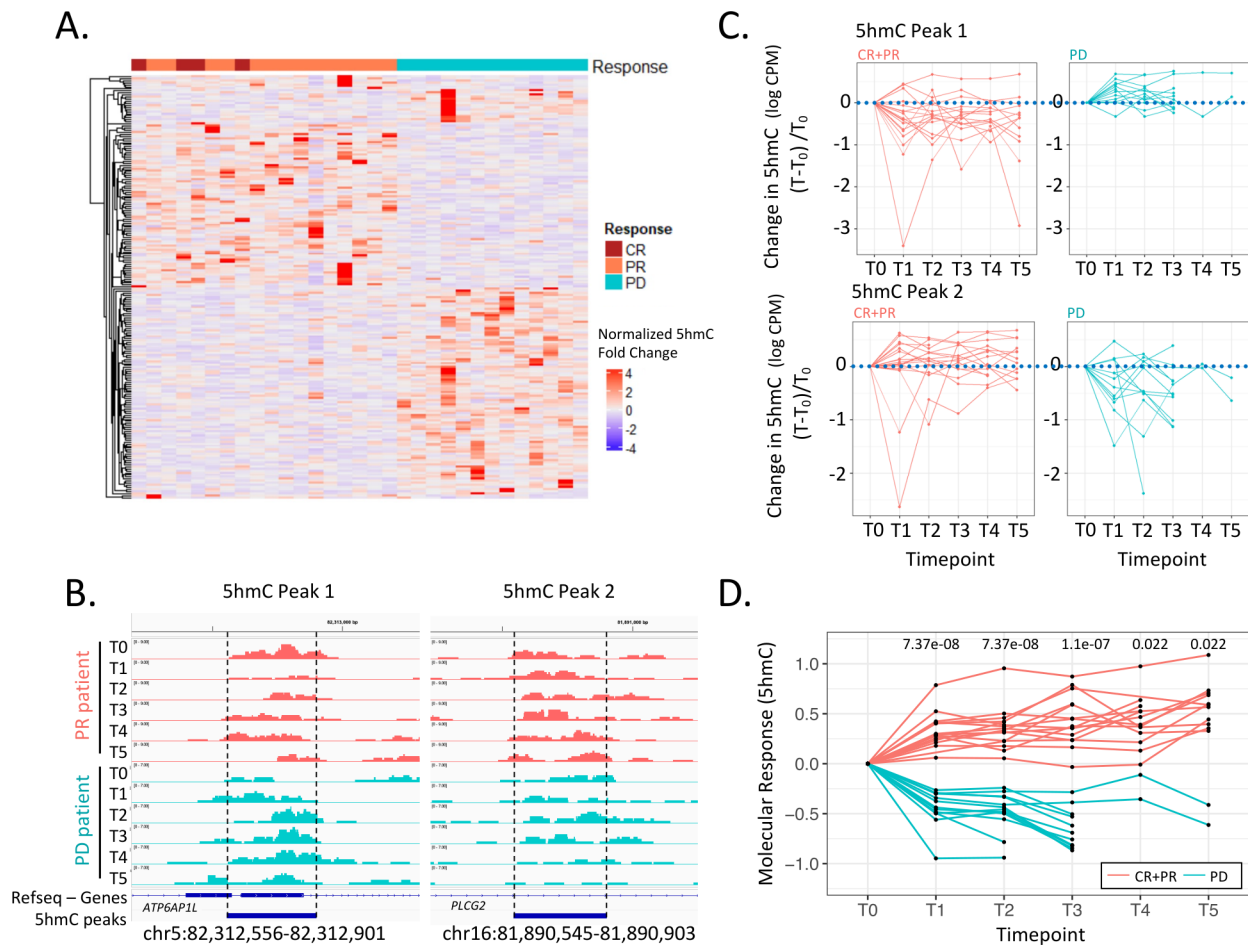
Differential hydroxymethylated genes in cfDNA identified in this study have previously been associated with ICI response and resistance in tissue transcriptome analysis. Therefore, we sought to ask whether the changes observed in cfDNA hydroxymethylome are at least partially derived from the hydroxymethylome of tumor-derived DNA. For this purpose, we first identified the differentially hydroxymethylated regions (DhMRs) on anti-PD-1 treatment by comparing 5hmC profiles at  $T_R$  to  $T_0$  in responders and in non-responders 152 and 301 DhMRs with a  $p$  value below 0.05 and a fold change exceeding 1.5 were identified by Wilcoxon rank-sum test in responders (figure 3A, left panel) and non-responders (figure 3A, right panel), respectively. These DhMRs which were separately identified in responders and non-responders were specific and non-overlapping (online supplemental figure 2A,B). To investigate whether these DhMRs identified in plasma were informative with regards to 5hmC profile changes in tumor tissue, we generated 5hmC profiles of 15 normal lung tissues and 18 lung cancer tissues. Both top responder and

non-responder DhMRs identified in plasma were able to separate tumor lung tissue from normal lung tissue as visualized by t-distributed stochastic neighbor embedding (figure 3B). These results suggest that the DhMRs identified in plasma are congruent with changes in 5hmC profiles of lung tumor tissue.

#### Anti-PD-1 therapy response monitoring with plasma-derived 5hmC profiles

The observation that the anti-PD-1 treatment-induced 5hmC changes were significantly non-overlapping between responders and non-responders suggests that such regions could potentially be used to monitor treatment response. To test this idea, first, genomic loci with the most significant changes between the responding and non-responding cohorts were determined by calculating treatment induced fold change in 5hmC occupancy normalized to baseline ( $(T_R - T_0)/T_0$ ) (termed  $\zeta$ ) and then identifying the regions with the most statistically significant changes by applying a threshold  $p$  value of 0.05 and difference of  $\zeta$  at least 1.5. This resulted in identification of 154 regions with decreased 5hmC and 129 regions with increased 5hmC in responders relative to non-responders (figure 4A). Majority of these loci mapped to gene features as exemplified by the two 5hmC peak loci visualized using the genome browser Integrative Genomics Viewer (IGV) (figure 4B). Examination of all study time points showed





**Figure 4** Anti-PD-1 therapy response monitoring using plasma-derived cell-free DNA. (A) Heatmap showing anti-PD-1 treatment-induced log fold change in 5hmC counts over 283 genomic loci in responder and non-responder cohorts. (B) Genome browser screenshots over two distinct DhMRs with opposite behaviors between responding (coral) and non-responding (teal) patients as identified in panel A. Peak 1 showing a decrease of 5hmC levels in a representative CR PR patient and an increase of 5hmC levels in a representative PD patient over time, where T0 is baseline and T5 is time of response. Peak 2 showing opposite 5hmC changes between CR PR and PD. (C) 5hmC fold change over two example genomic loci in responding (coral) and non-responding (teal) patients over the course of treatment. (D) Spider plot showing overall 5hmC-based molecular response over all treatment-induced DhMRs ( $p < 0.05$ ) in responding (coral) and non-responding (teal) patients at each time point throughout treatment. Molecular response was calculated independently for each patient by leave-one-out cross validation as described in methods. Wilcoxon rank-sum test p values are shown. CR, complete responders; DhMRs, hydroxymethylated regions; PD, progressive disease; PD-1, programmed death-1; PR, partial responders; 5hmC, 5-hydroxymethylcytosine.

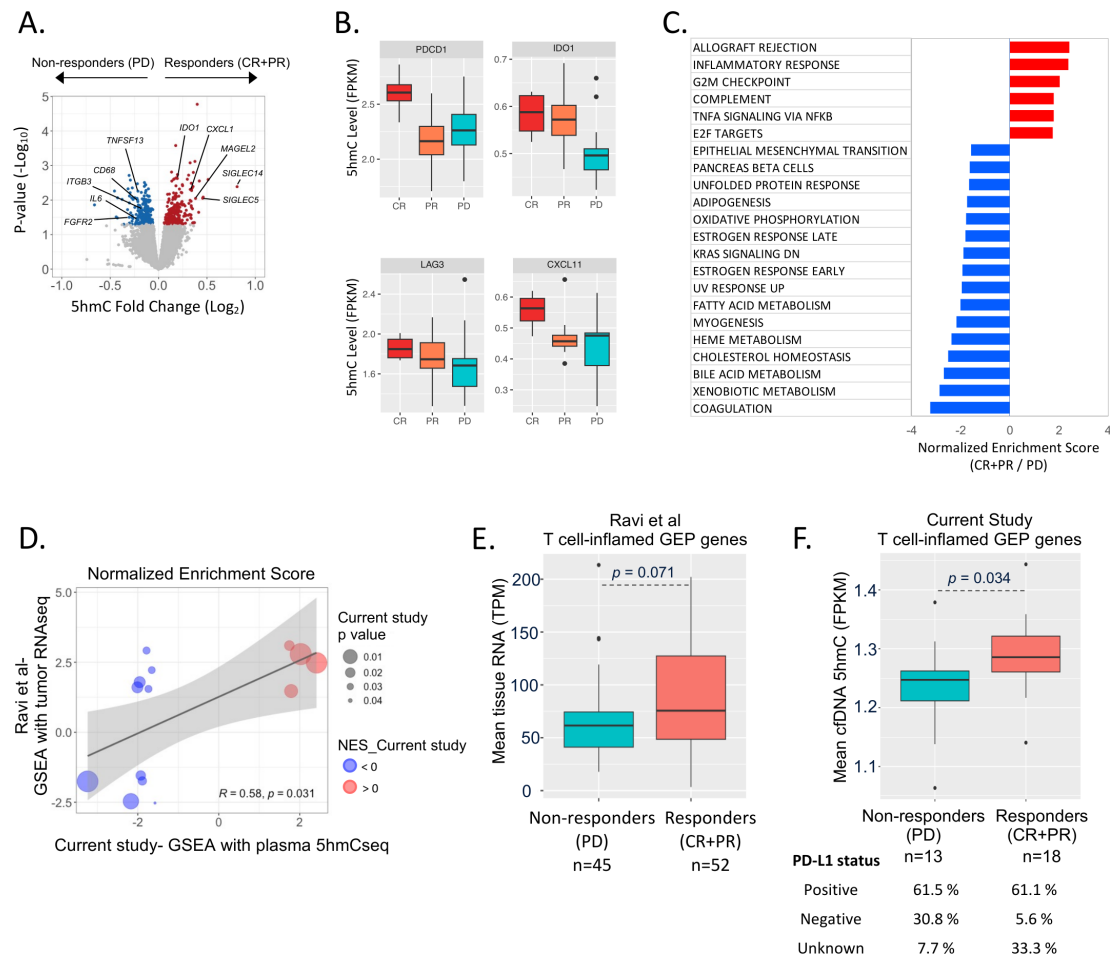
that the changes in 5hmC levels observed at time of radiologic response can be detected as early as 4–6 weeks from therapy start (figure 4C). Based on this observation, we next investigated the molecular response elicited by anti-PD-1 treatment on the plasma hydroxymethylome as measured by cumulative change in 5hmC counts over the top differential regions whose hydroxymethylation levels correlated either negatively or positively with response as identified in figure 4A. To this end, we tested prediction of response by leave-one-out cross validation, where all but one patient was used to determine top differential 5hmC regions, then an  $MR5_{hmC}$  was calculated for the patient samples that were left out blinded of the patient response. This was repeated 31 times for each patient. We found that  $MR5_{hmC}$  can distinguish responders from non-responders as early as the first time point, which was 4–6 weeks from therapy start (figure 4D). These results suggest

5hmC profiles obtained from plasma contain treatment-induced changes that could serve as molecular response markers with potential for earlier detection than radiologic response. Altogether, our findings demonstrate that 5hmC-based changes identified in plasma samples of patients with NSCLC can be used for anti-PD-1 therapy response monitoring.

#### Anti-PD-1 responders and non-responders show distinct pretreatment plasma cfDNA 5hmC profiles that associate with known response and resistance biology

To investigate whether responders had distinct baseline cfDNA 5hmC profiles relative to non-responders, we compared 5hmC profiles at time points collected prior to treatment start (T0). This differential 5hmC analysis identified 482 genes with  $p$  value  $< 0.05$  (figure 5A). In responders relative to non-responders, 5hmC was





**Figure 5** Differential hydroxymethylation in plasma-derived cfDNA in patients with responding compared with non-responding non-small cell lung cancer. (A) Volcano plot showing differential 5hmC analysis comparing plasma-derived 5hmC profiles prior to anti-PD-1 treatment in responders relative to non-responders. (B) Differential 5hmC levels (FPKM) over immune modulating genes associated with anti-PD-1 response. (C) Hallmark genesets gene set enrichment analysis (GSEA) using 5hmC counts over genes to compare baseline (T0) 5hmC profiles in responding versus non-responding patients reveals enrichment of inflammatory and immune response enrichment in responders. Red indicates enrichment in responders. Blue indicates enrichment in non-responders. (D) Correlation of normalized enrichment scores for gene sets with p value < 0.05 in both current study and Ravi *et al* as determined by analysis of 5hmCseq or RNA seq, respectively. Pearson correlation and associated p value is reported. E–F. Mean baseline RNA levels expressed as TPM from Ravi *et al* (E) or mean baseline 5hmC levels expressed as FPKM from current study (F) over the 18 genes in the T-cell inflamed gene expression profile (GEP), in responding (CR+PR, coral) and non-responding (PD, teal) patients. P value is calculated by the Wilcoxon rank-sum test. Boxplots are shown with Tukey whiskers. Boxplot center line and bounds represent median and 25th to 75th percentile, respectively, while dots represent outliers. cfDNA, cell-free DNA; CR, complete responders; FPKM, fragments per kilobase per million mapped fragments; NES, normalized enrichment score; PD, progressive disease; PD-1, programmed death-1; PD-L1, programmed death ligand-1; PR, partial responders; RNA-seq, RNA sequencing; TPM, transcripts per million mapped reads; 5hmC, 5-hydroxymethylcytosine.

higher over genes with known functions in immune response such as sialic acid-binding immunoglobulin-like lectin 5<sup>47</sup> and *CXCL1* (figure 5A). On the other hand, non-responders had elevated 5hmC over genes associated with metastasis/EMT or resistance to anti-PD-(L)1 therapy such as *FGFR2*<sup>48</sup> and *IL6*<sup>49</sup> (figure 5A). Inspection of 5hmC levels over genes that are known to be involved in modulating T-cell responses including *PDCD1*, *IDO1*, *LAG3* and *CXCL11* revealed an overall increase in 5hmC in responders, particularly in CR patients, relative to non-responders (figure 5B).

Next, GSEA was employed to investigate biological processes that can be distinguished by comparing

pretreatment cfDNA 5hmC profiles in responding patients to pretreatment cfDNA profiles in non-responders. This analysis revealed several gene sets that are associated with an inflamed immune state such as allograft rejection, inflammatory response and TNF- $\alpha$  signaling via NF- $\kappa$ B to be upregulated in responders relative to non-responders (figure 5C). Among the leading-edge genes that drove enrichment of these immune-related gene sets were major histocompatibility complex II gene *HLA-DQA1* and non-classical major histocompatibility complex gene *CD1D*, both of which are involved in antigen presentation. Notably, among the gene sets that were upregulated in non-responders was EMT genes, which were previously

associated with resistance to anti-PD-1 treatment,<sup>45</sup> with leading-edge genes that included EMT marker *CDH2* and *WNT5A*, the protein product of which promotes EMT. Furthermore, genes that are downregulated on KRAS activation were downregulated in responding patients, suggesting KRAS activation in responders. Interestingly, KRAS mutations were previously correlated with response and better outcomes on immunotherapy treatment.<sup>50 51</sup>

Given the presence of both tumor-derived and non-tumor-derived DNA in circulating cfDNA, we sought to compare the biological processes implicated by 5hmC profiling of plasma-derived cfDNA to those uncovered by RNA analysis of tissue samples. For this purpose, we employed a data set from a recently published comprehensive study that profiled tissue samples from patients with NSCLC to investigate anti-PD-1 treatment response.<sup>52</sup> In this data set, 152 patient samples had associated gene expression data, from which pretreatment tumor samples were profiled from patients who were treated with anti-PD-1 agents. Comparison of GSEA results from the tissue RNA analysis with that of cfDNA 5hmC analysis revealed positive correlation specifically for gene sets with higher significance (figure 5D), such as allograft rejection, G2M checkpoint and coagulation gene sets.

Previous analyses of tissue RNA had uncovered and independently validated a signature termed T cell-inflamed gene expression profile (GEP) which incorporates RNA abundance for 18 genes to distinguish responders to anti-PD-1 treatment from non-responders.<sup>11 12 36</sup> To evaluate the relative levels of plasma cfDNA 5hmC and tissue RNA over the T cell-inflamed GEP in responders and non-responders, we first examined the mean RNA levels over the 18-genes comprising T cell-inflamed GEP. Comparison of 52 responders (CR+PR) relative to 45 non-responders (PD) from Ravi *et al* (2023) revealed increased RNA abundance for T cell-inflamed GEP in responders (figure 5E), consistent with previous reports. Strikingly, mean 5hmC levels of T cell-inflamed GEP genes, as measured in plasma cfDNA, were also higher in responders compared with non-responders (figure 5F), validating that increased expression over the T cell-inflamed GEP can indeed distinguish responders from non-responders. Next, we investigated whether 5hmC measurements over the 18-T-cell inflamed GEP genes in plasma cfDNA can distinguish patients with better overall survival. Only a subset of patients had follow-up for a minimum of 36 months post-treatment start (total n=19). For survival analysis, these patients were separated into two groups according to their mean 5hmC value over the 18 T-cell inflamed GEP genes, where patients above or below the median of the full cohort were assigned to high 5hmC (n=9) or low 5hmC (n=10) groups, respectively. As expected, based on the results that showed higher 5hmC levels over the T-cell inflamed GEP genes for responders relative to non-responders (figure 5F), a trend separating the two groups in terms of overall survival was observed, however it did not reach statistical significance with a p value of 0.11 in this smaller subset of patients (online supplemental

figure 3). Overall, our results show that the expression of these tissue-relevant genes in predicting therapy response can also be captured by increased hydroxymethylation using plasma cfDNA without the need for tissue samples. Altogether, these findings indicate that immune relevant signatures consistent with immunotherapy response can be identified in plasma using 5hmC profiles.

## DISCUSSION

ICIs can exert significant and durable antitumor effects, although in a subpopulation of patients. Current strategies to classify patients who would benefit most from these treatments are not sufficiently reliable.<sup>3</sup> Here, we explored the potential for using 5hmC-based changes that are detectable in plasma-derived cfDNA to predict and monitor anti-PD-1 treatment response in patients with NSCLC. Altogether, our results indicate that cfDNA 5hmC profiles contain loci that are correlated with anti-PD-1 response that could be exploited for biomarker discovery for patient selection and therapy response monitoring for ICIs in treatment of lung cancer.

Among the patient selection biomarkers that are currently used in the clinic is PD-L1 immunohistochemistry using tumor biopsy samples. Yet, it has been reported that patients with PD-L1 immunostaining below established thresholds may still benefit from anti-PD-1 treatments. Furthermore, lack of response is also observed in PD-L1 positive patients.<sup>3 7</sup> Both intratumor and inter-tumor heterogeneity with respect to PD-L1 expression may yield such contradictory results as tumor sampling at a single tumor site might not accurately reflect the state of overall PD-L1 expression in a patient's tumor(s). A second important variable is the poor uniformity in the PD-L1 immunohistochemistry antibodies and different thresholds for PD-L1 positivity used.<sup>7</sup> These limitations can be circumvented by liquid biopsy, which captures DNA shed not only by cancer cells from different tumor sites but also by tumor microenvironment and immune cells. In particular, analysis of plasma cfDNA 5hmC, which marks active biological pathways, can provide better understanding of drug response in tumor and immune cells using a single analyte. Indeed, we found that plasma cfDNA 5hmC profiles are significantly different in responders relative to non-responders both prior to treatment (figure 5) as well as after treatment (figure 2). Our analysis revealed 5hmC enrichment over immune relevant genes, indicating increased immune gene activation in responders. Given that ~61% of both cohorts were PD-L1 positive (figure 1B), these results show that 5hmC profiling can provide an independent measurement of immune status that is informative of potential response to immunotherapy.

5hmC profiles in plasma-derived cfDNA have proven to be valuable biomarkers for cancer detection in recent years.<sup>27–29 53</sup> A stable epigenetic mark that is produced on active demethylation of DNA methylcytosines, 5hmC is positively correlated with gene expression and

regulation.<sup>32 35 54</sup> Unlike 5mC which covers large fractions of the genome that are transcriptionally repressed, 5hmC marks a much smaller fraction that is enriched in tissue-specific and cell type-specific genes. Recently it was shown that 5hmC analysis provides the most informative signal for RNA abundance in disease relevant pathways in comparison to WGS (whole genome sequencing) and WGBS (whole genome bisulfite sequencing) data.<sup>54</sup> Furthermore, 5hmC remodeling is a hallmark of immune cell differentiation and activation.<sup>34 35 55</sup> 5hmC profiles can therefore provide distinctive signatures for determination of tissue identity and immune states using cfDNA.<sup>34 35 55</sup>

Plasma-derived cfDNA presents unique advantages as a source for candidate biomarkers for ICI response and treatment monitoring. Non-invasive biomarkers would circumvent the challenges related to tumor biopsies, such as low rates of patient compliance, potential for complications and insufficient material collection. In addition, non-invasive biomarkers can facilitate serial sampling which allows monitoring of dynamic changes during treatment that could be related to immune activation or acquired resistance. In this study, anti-PD-1 therapy response induced distinct changes in plasma cfDNA 5hmC profiles of responders compared with non-responders that can be observed as early as 4–6 weeks (figure 4C,D). Our results show that 5hmC analysis offers a novel non-invasive method for serial monitoring of anti-PD-1 therapy response in plasma.

Among the advantages of 5hmC-based cfDNA analysis is the ability to inform on not only the tumor cells but also other cells that contribute DNA to circulation. While immune cells make up the majority of cfDNA particularly in healthy subjects, tumor cells and other components of the tumor microenvironment also shed DNA into blood in patients with cancer. Methods confined to analysis of tumor-derived cfDNA, such as tumor-informed mutation analysis, necessitate addressing sensitivity challenges resulting from tumor-derived DNA dilution in the rest of cfDNA. Furthermore, such methods also fail to inform on the tumor microenvironment. Yet, the composite nature of cfDNA presents an opportunity to gain insight into immune and tumor status using a single analyte such as 5-hydroxymethylated cfDNA.

Heterogeneity of cfDNA can also present challenges in identification of signals specific to clinical questions at hand. One such challenge in the context of mutational analysis is to distinguish tumor-derived mutations from the ones that originate from clonal hematopoiesis.<sup>56</sup> Additional sequencing of the circulating white blood cells can help distinguish clonal hematopoiesis of indeterminate potential (CHIP)-derived mutations from tumor-derived ones, however at the expense of increased cost. While cfDNA heterogeneity can also present challenges to other types of analytes, tissue-specific signatures in cfDNA, such as 5hmC and 5mC, can help infer origins of cfDNA. Furthermore, because 5hmC is specifically enriched over transcriptionally active genes or regulatory

regions, it can be used to gain insight into active cellular programs and states in the tissues or cells that contribute DNA to circulation. Consistently, 5hmC analysis of cfDNA in this study identified IFN response, inflammatory response in responders as well as EMT in non-responders (figure 2B–G), pointing to both immune cell-derived and tumor cell-derived biology.

Challenges associated with identification of a reliable ICI response prediction biomarker strategy have been at least in part attributed to complex dynamics between the tumor and the immune system. While PD-1-mediated suppression of effector T cells is a major mechanism by which tumors evade the immune system, there are other mechanisms tumors can use for the same purpose, rendering ICI treatment ineffective. Similarly, the capacity of tumors to elicit immune response through mechanisms such as neoantigen expression also proves important for determining ICI treatment response. Therefore, it is paramount to identify biomarkers that can inform on both the tumor and immune states relevant to activation of antitumor immunity. TMB,<sup>13</sup> microsatellite instability status<sup>57 58</sup> and T-cell inflammation signatures<sup>11 12 36</sup> which were developed subsequent to anti-PD-L1 tumor staining as additional biomarkers have improved identification of patients who may benefit from ICI treatment. Yet, combining biomarkers that have multiple analyte requirements presents its own challenges as these approaches will not only increase cost but also may not be practical for reasons such as sample input requirements. Altogether, non-invasive single analyte tests that can inform on both the immune status as well as tumors themselves would be ideal candidates for predictive biomarkers for ICIs. 5hmC offers a single analyte solution to this end, with hydroxymethylome comparisons yielding gene sets that implicate both tumor and immune originating cfDNA.

Comparison of plasma cfDNA profiles of responding and non-responding patients indicate that 5hmC analysis can capture immune activation and tumor progression. Notably, this work showed that cfDNA 5hmC profiling corroborates response signatures previously discovered by tissue analysis such as T cell-inflamed gene expression profile.<sup>11 12</sup> Overall, analysis of 5hmC profiles in patients who received anti-PD-1 treatment showed that 5hmC can correlate with response. Further investigations of these findings in larger study cohorts will enable exploration of the full potential for using 5hmC in predicting and monitoring response to immunotherapy.

**Acknowledgements** The authors would like to thank Dr Stephen Quake and Dr Alan Ashworth for their critical review of the manuscript and providing important suggestions.

**Contributors** Conceptualization and study design: GDG, SL. Methodology: GDG, YN, TP, MN, WV, SL. Data generation and clinical annotation: GDG, YN, TP, MN, KH, AS, WV. Data analysis and interpretation: GDG, YN, CC, GPM, SL. Writing original draft: GDG, YN, CC, GPM, SL. Writing, editing and reviewing: All authors. Acting as guarantor: GDG and SL.

**Funding** This work was supported by ClearNote Health.

**Competing interests** GDG, YN, CC, GPM, TP, MN, KH, AS, WV, and SL are employees of ClearNote Health.



**Patient consent for publication** Consent obtained directly from patient(s).

**Ethics approval** The study was approved by the ethical committee of the Ärztekammer Hamburg under the # PV5134. The ethic committees of Nordrhein, Schleswig-Holstein, Niedersachsen, Thüringen, Baden-Württemberg, Brandenburg, Westfalen-Lippe, Rheinland-Pfalz and Sachsen-Anhalt followed this decision. Participants gave informed consent to participate in the study before taking part.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** All data relevant to the study are included in the article or uploaded as supplementary information. Availability of data and material: Sequenced fragment counts from this study are available as supplemental material.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iDs

Gulferm D Guler <http://orcid.org/0009-0005-4743-7749>

Samuel Levy <http://orcid.org/0000-0002-4444-5103>

## REFERENCES

- Sung H, Ferlay J, Siegel RL, *et al*. Global cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021;71:209–49.
- Surveillance E, Results E. SEER\*Stat database: incidence - SEER research data, 9 registries, Nov 2019 sub (1975–2017) - linked to County attributes - time dependent (1990–2017) income/Rurality, 1969–2017 counties, National Cancer Institute, DCCPS, surveillance research program, released April 2020. *Submitted* 2019.
- Wang M, Herbst RS, Boshoff C. Toward personalized treatment approaches for non-small-cell lung cancer. *Nat Med* 2021;27:1345–56.
- Shields MD, Marin-Acevedo JA, Pellini B. Immunotherapy for advanced non-small cell lung cancer: A decade of progress. *Am Soc Clin Oncol Educ Book* 2021;41:1–23.
- Upadhaya S, Neffelin ST, Hodge JP, *et al*. Combinations take centre stage in Pd1/Pd1 inhibitor clinical trials. *Nat Rev Drug Discov* 2021;20:168–9.
- Reck M, Rodríguez-Abreu D, Robinson AG, *et al*. Five-year outcomes with Pembrolizumab versus chemotherapy for metastatic non-small-cell lung cancer with PD-L1 tumor proportion score  $\geq 50$ . *JCO* 2021;39:2339–49.
- Gibney GT, Weiner LM, Atkins MB. Predictive biomarkers for Checkpoint inhibitor-based Immunotherapy. *Lancet Oncol* 2016;17:e542–51.
- 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.
- Rodríguez-Abreu D, Powell SF, Hochmair MJ, *et al*. Pemetrexed plus platinum with or without Pembrolizumab in patients with previously untreated metastatic Nonsquamous NSCLC: protocol-specified final analysis from KEYNOTE-189. *Ann Oncol* 2021;32:881–95.
- Tumeh PC, Harview CL, Yearley JH, *et al*. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* 2014;515:568–71.
- Ayers M, Lunceford J, Nebozhyn M, *et al*. IFN- $\gamma$ -related mRNA profile predicts clinical response to PD-1 blockade. *J Clin Invest* 2017;127:91190:2930–40.
- Cristescu R, Mogg R, Ayers M, *et al*. Pan-tumor Genomic biomarkers for PD-1 Checkpoint blockade-based Immunotherapy. *Science* 2018;362:eaar3593.
- Rizvi NA, Hellmann MD, Snyder A, *et al*. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 2015;348:124–8.
- Chen DS, Mellman I. Oncology meets Immunology: the cancer-immunity cycle. *Immunity* 2013;39:1–10.
- Bindal P, Gray JE, Boyle TA, *et al*. Biomarkers of therapeutic response with immune Checkpoint inhibitors. *Ann Transl Med* 2021;9:1040.
- Jahr S, Hentze H, Englisch S, *et al*. DNA fragments in the blood plasma of cancer patients: Quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Res* 2001;61:1659–65.
- Mouliere F, Chandrananda D, Piskorz AM, *et al*. Enhanced detection of circulating tumor DNA by fragment size analysis. *Sci Transl Med* 2018;10:eaat4921.
- Pereira B, Chen CT, Goyal L, *et al*. Cell-free DNA captures tumor heterogeneity and driver alterations in rapid Autopsies with pre-treated metastatic cancer. *Nat Commun* 2021;12:3199.
- Joosse SA, Pantel K. Circulating DNA and liquid biopsies in the management of patients with cancer. *Cancer Res* 2022;82:2213–5.
- Garcia-Murillas I, Schiavon G, Weigelt B, *et al*. Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. *Sci Transl Med* 2015;7:302ra133.
- Bratman SV, Yang SYC, lafolla MAJ, *et al*. Personalized circulating tumor DNA analysis as a predictive biomarker in solid tumor patients treated with Pembrolizumab. *Nat Cancer* 2020;1:873–81.
- Stadler J-C, Belloum Y, Deitert B, *et al*. Current and future clinical applications of ctDNA in Immuno-oncology. *Cancer Res* 2022;82:349–58.
- Nabet BY, Esfahani MS, Moding EJ, *et al*. Noninvasive early identification of therapeutic benefit from immune Checkpoint inhibition. *Cell* 2020;183:363–76.
- 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;24:1441–8.
- Khagi Y, Goodman AM, Daniels GA, *et al*. Hypermutated circulating tumor DNA: correlation with response to Checkpoint inhibitor-based Immunotherapy. *Clin Cancer Res* 2017;23:5729–36.
- Aggarwal C, Thompson JC, Chien AL, *et al*. Baseline plasma tumor Mutation burden predicts response to Pembrolizumab-based therapy in patients with metastatic non-small cell lung cancer. *Clin Cancer Res* 2020;26:2354–61.
- Song C-X, Yin S, Ma L, *et al*. 5-Hydroxymethylcytosine signatures in cell-free DNA provide information about tumor types and stages. *Cell Res* 2017;27:1231–42.
- Li W, Zhang X, Lu X, *et al*. 5-Hydroxymethylcytosine signatures in circulating cell-free DNA as diagnostic biomarkers for human cancers. *Cell Res* 2017;27:1243–57.
- Guler GD, Ning Y, Ku C-J, *et al*. Detection of early stage Pancreatic cancer using 5-Hydroxymethylcytosine signatures in circulating cell free DNA. *Nat Commun* 2020;11:5270.
- Shen SY, Singhania R, Fehringer G, *et al*. Sensitive tumour detection and classification using plasma cell-free DNA Methylomes. *Nature* 2018;563:579–83.
- Tahiliani M, Koh KP, Shen Y, *et al*. Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in mammalian DNA by MLL partner Tet1. *Science* 2009;324:930–5.
- Cui X-L, Nie J, Ku J, *et al*. A human tissue map of 5-Hydroxymethylcytosines exhibits tissue specificity through gene and enhancer modulation. *Nat Commun* 2020;11:6161.
- Szulwach KE, Li X, Li Y, *et al*. Integrating 5-Hydroxymethylcytosine into the Epigenomic landscape of human embryonic stem cells. *PLoS Genet* 2011;7:e1002154.
- Nestor CE, Lentini A, Hägg Nilsson C, *et al*. 5-Hydroxymethylcytosine remodeling precedes lineage specification during differentiation of human Cd4(+) T cells. *Cell Rep* 2016;16:559–70.
- Tsiouplis NJ, Bailey DW, Chiou LF, *et al*. TET-mediated epigenetic regulation in immune cell development and disease. *Front Cell Dev Biol* 2021;8:623948.
- Ott PA, Bang Y-J, Piha-Paul SA, *et al*. T-cell-inflamed gene-expression profile, programmed death ligand 1 expression, and tumor mutational burden predict efficacy in patients treated with Pembrolizumab across 20 cancers: KEYNOTE-028. *J Clin Oncol* 2019;37:318–27.
- Li H, Durbin R. Fast and accurate long-read alignment with burrows-Wheeler transform. *Bioinformatics* 2010;26:589–95.
- Institute B. “Picard Toolkit.” broad Institute, Github repository”. 2019. Available: <http://broadinstitute.github.io/picard/>
- Li H, Handsaker B, Wysoker A, *et al*. The sequence alignment/map format and Samtools. *Bioinformatics* 2009;25:2078–9.
- Quinlan AR, Hall IM. Bedtools: a flexible suite of utilities for comparing Genomic features. *Bioinformatics* 2010;26:841–2.



- 41 Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of Digital gene expression data. *Bioinformatics* 2010;26:139–40.
- 42 Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102:15545–50.
- 43 Adachi Y, Kamiyama H, Ichikawa K, et al. Inhibition of FGFR Reactivates IFN $\gamma$  signaling in tumor cells to enhance the combined antitumor activity of Lenvatinib with anti-PD-1 antibodies. *Cancer Res* 2022;82:292–306.
- 44 Su H, Na N, Zhang X, et al. The biological function and significance of Cd74 in immune diseases. *Inflamm Res* 2017;66:209–16.
- 45 Mariathasan S, Turley SJ, Nickles D, et al. TGF $\beta$  attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* 2018;554:544–8.
- 46 Hugo W, Zaretsky JM, Sun L, et al. Genomic and Transcriptomic features of response to anti-PD-1 therapy in metastatic Melanoma. *Cell* 2016;165:35–44.
- 47 Vuchkovska A, Glanville DG, Scurti GM, et al. Siglec-5 is an inhibitory immune Checkpoint molecule for human T cells. *Immunology* 2022;166:238–48.
- 48 Ranieri D, Rosato B, Nanni M, et al. Expression of the Fgfr2 Mesenchymal splicing variant in epithelial cells drives epithelial-Mesenchymal transition. *Oncotarget* 2016;7:5440–60.
- 49 Huseni MA, Wang L, Klementowicz JE, et al. Cd8+ T cell-intrinsic IL-6 signaling promotes resistance to anti-PD-L1 Immunotherapy. *Cell Rep Med* 2023;4:100878.
- 50 Liu C, Zheng S, Jin R, et al. The superior efficacy of anti-PD-1/PD-L1 Immunotherapy in KRAS-mutant non-small cell lung cancer that correlates with an inflammatory phenotype and increased Immunogenicity. *Cancer Lett* 2020;470:95–105.
- 51 Lauko A, Kotecha R, Barnett A, et al. Impact of KRAS Mutation status on the efficacy of Immunotherapy in lung cancer brain metastases. *Sci Rep* 2021;11:18174.
- 52 Ravi A, Hellmann MD, Arniella MB, et al. Genomic and Transcriptomic analysis of Checkpoint blockade response in advanced non-small cell lung cancer. *Nat Genet* 2023;55:807–19.
- 53 Haan D, Bergamaschi A, Friedl V, et al. Epigenomic blood-based early detection of Pancreatic cancer employing cell free DNA. *Clinical Gastroenterology and Hepatology* 2023;21:1802–1809.
- 54 Sjöström M, Zhao SG, Levy S, et al. The 5-Hydroxymethylcytosine landscape of prostate cancer. *Cancer Res* 2022;82:3888–902.
- 55 Tsagaratou A, Åijö T, Lio C-WJ, et al. Dissecting the dynamic changes of 5-Hydroxymethylcytosine in T-cell development and differentiation. *Proc Natl Acad Sci USA* 2014;111:E3306–15.
- 56 Abbosh C, Swanton C, Birnbak NJ. Clonal Haematopoiesis: a source of biological noise in cell-free DNA analyses. *Ann Oncol* 2019;30:358–9.
- 57 Overman MJ, McDermott R, Leach JL, et al. Nivolumab in patients with metastatic DNA mismatch repair-deficient or Microsatellite instability-high colorectal cancer (Checkmate 142): an open-label, Multicentre, phase 2 study. *Lancet Oncol* 2017;18:1182–91.
- 58 Zhao P, Li L, Jiang X, et al. Mismatch repair deficiency/Microsatellite instability-high as a Predictor for anti-PD-1/PD-L1 Immunotherapy efficacy. *J Hematol Oncol* 2019;12:54.