

METHOD OF PREDICTING FOR BENEFIT FROM IMMUNE CHECKPOINT INHIBITION THERAPY

CROSS-REFERENCE TO RELATED APPLICATIONS

- 5 [0001] This application claims the benefit of priority of Singapore application No. 10201811546W, filed 21 December 2018, the contents of it being hereby incorporated by reference in its entirety for all purposes.

FIELD OF THE INVENTION

- 10 [0002] The present invention relates generally to the field of cancer. In particular, the present invention relates to the use of a method for selecting patients for immune checkpoint inhibition therapy.

BACKGROUND OF THE INVENTION

- 15 [0003] Among various treatment modalities for cancer, immune checkpoint inhibition (ICI) has made significant breakthroughs in several tumour types. In ICI therapy, immune checkpoint inhibitors such as pembrolizumab and nivolumab block interactions between the immune checkpoint receptor PD-1 and its ligands, reducing negative costimulatory signals and increasing T-effector cell function to elicit anti-tumour responses.
- 20 [0004] While beneficial in certain tumour types, several recent ICI phase III trials have also proved unsuccessful, particularly in solid epithelial tumours. Responses to ICI therapy are observed only in a minority of patients. Therefore, there is a need for robust biomarkers that can predict patient responses to immunotherapy. Currently, the most developed ICI biomarkers are PD-L1 expression, microsatellite instability, and tumour mutation burden.
- 25 These biomarkers are positive predictive biomarkers aimed at identifying patients who will respond to ICI therapy. However, controversies surrounding these biomarkers have been raised and ICI responses in biomarker-negative populations have been observed. These observations highlight a complementary requirement for negative predictive biomarkers for ICI therapy, which can identify tumours that are likely to be resistant to immunotherapy.
- 30 However, negative predictive biomarkers of ICI remain poorly described.
- [0005] Promoters are genomic *cis*-regulatory elements upstream of transcription start sites (TSSs) which function to initiate transcription. Promoter activity is epigenetically regulated,

and more than half of all human genes have multiple promoters, which can be selectively activated as a consequence of normal biological function or disease state. The use of alternative promoters can produce distinct 5' untranslated regions (UTRs) and first exons, enhancing mRNA and protein isoform diversity. In cancer, alternative promoters can generate cancer-specific isoforms with oncogenic properties. Recently, studies have demonstrated that tumours may utilize alternative promoters as a mechanism of immune-editing and evasion. Epigenetically-driven alternative promoter utilization is thus a potential mechanism of resistance to ICI therapy.

[0006] There is therefore a need to develop alternative methods based on alternative promoter utilization to predict patient response to immunotherapy or ICI therapy.

SUMMARY

[0007] In one aspect, there is provided a method of identifying a patient suffering from cancer as benefiting or not benefiting from immune checkpoint inhibition (ICI) therapy comprising the steps of:

- a) measuring an expression level of one or more preselected markers in a cancerous biological sample obtained from the patient;
- b) identifying a differentially expressed alternative promoter based on the expression level of the one or more preselected markers measured in step (a);
- c) calculating an alternative promoter usage score;
- d) identifying the patient as benefiting or not benefiting from ICI therapy using the alternative promoter usage score.

[0008] In another aspect, there is provided a method of determining the prognosis of a patient suffering from cancer comprising the steps of:

- a) measuring an expression level of one or more preselected markers in a cancerous biological sample obtained from the patient;
- b) identifying a differentially expressed alternative promoter based on the expression level of the one or more preselected markers measured in step (a);
- c) calculating an alternative promoter usage score;
- d) determining the prognosis of the patient using the alternative promoter usage score.

[0009] In another aspect, there is provided a biomarker for excluding a patient suffering from cancer from ICI therapy, wherein the biomarker is an alternative promoter usage score

that is above the alternative promoter usage scores of one or more reference samples, wherein the alternative promoter usage score is the sum of the differentially expressed alternative promoters at each alternative promoter site.

5

DEFINITIONS

[0010] As used herein, the term “promoter” refers to a region of DNA that initiates transcription of a gene. A promoter may be a major promoter, a minor promoter or an alternative promoter. A major promoter is a promoter that is the most frequently used for the transcription of a gene.

10 [0011] As used herein, the terms “alternate promoter” and “alternative promoter” refer to a region of DNA that initiates transcription of a gene at an alternative transcription start site than the major promoter or the minor promoter.

[0012] The term “gained promoter” as used herein refers to a promoter that is gained or has increased activity in a cancerous biological sample compared to a non-cancerous biological sample. A gained promoter may be a promoter in the cancerous biological sample that is not present in the non-cancerous biological sample. A gained promoter may be a promoter in the cancerous biological sample that has an increased promoter activity compared to the promoter in the non-cancerous biological sample.

15 [0013] The term “lost promoter” as used herein refers to a promoter that is lost or has decreased activity in a cancerous biological sample compared to a non-cancerous biological sample. A lost promoter may be a promoter that is present in the non-cancerous biological sample but not in the cancerous biological sample. A lost promoter may be a promoter in the cancerous biological sample that has a decreased promoter activity compared to the promoter in the non-cancerous biological sample.

20 [0014] The term “immunotherapy” in the context of cancer refers to a form of therapy that involves the modulation of the immune system to treat cancer. Modulation of the immune system may involve the activation or inactivation of the immune system. This may involve the use of components of the immune system such as antibodies, cytokines and vaccines.

[0015] As used herein, the term “immune checkpoint inhibition therapy” refers to a form of cancer immunotherapy which targets key regulators of the immune system. Certain proteins made by some types of cancer cells and immune cells help keep immune responses in check and can keep T cells from killing cancer cells. Certain cancer cells use these proteins

25
30

to evade immunosurveillance. The term “immune checkpoint inhibition therapy” refers to a therapy that involves the use of factors that block these proteins, restoring immune system function and allowing T cells to mount an effective antitumour response.

[0016] The term “prognosis”, for the purposes of this application, refers to a prediction of the probable course and outcome of a clinical condition or disease. The term “prognosis” does not refer to the ability to predict the course or outcome of a condition with 100% accuracy. Instead, the term “prognosis” refers to the probability that a certain course or outcome will occur in a subject exhibiting a given condition, when compared to those individuals not exhibiting the condition. A “prognosis” can be made regarding one or more clinical outcomes, for example, the rate of progression of the disease in the subject, severity of the disease, survival rate, survival time, likelihood of metastasis, likelihood of disease recurrence or response to a therapeutic intervention.

[0017] The term “epigenetic”, for the purposes of this application, refers to changes that regulate gene expression without altering an individual’s DNA sequence. Examples of epigenetic alterations are DNA methylation, histone modifications and RNA-associated silencing. Epigenetic alterations play a vital role in disease development through control of gene expression.

[0018] The term “biomarker” in the context of this specification refers to a measurable indicator of a biological condition or disease. Biomarkers include but are not limited to substances, structures or processes that can be measured in body tissues and fluids.

[0019] As used herein, the terms “alternate promoter usage score” and “alternative promoter usage score” refer to a measure of alternative promoter utilization. The score can be expressed as an “AP” score or an “APB” score. The term “AP” in the context of AP_{high} and AP_{low} refers to alternate promoter utilization. An AP_{high} score indicates a high alternative promoter utilization and an AP_{low} score indicates a low alternative promoter utilization. The term “APB” in the context of APB_{high} and APB_{low} refers to alternate promoter utilization burden. An APB_{high} score indicates a high alternative promoter utilization and an APB_{low} score indicates a low alternative promoter utilization.

BRIEF DESCRIPTION OF THE DRAWINGS

[0020] The invention will be better understood with reference to the detailed description when considered in conjunction with the non-limiting examples and the accompanying drawings, in which:

5 [0021] **Fig. 1** shows the alternate promoter utilization burden (APBscore) across 33 tumour types. The APBscore was calculated in 10,393 samples across 33 tumour types. Each point represents the APBscore of a sample (y-axis). Tumour types are on the x-axis, ordered from lowest to highest median APBscore of the respective tumour type. The red horizontal bars are the median APBscore of the tumours, while the green horizontal bars represent the
 10 median APBscore of the corresponding normal tissue (provided only if at least 10 normal samples were available for analysis). The 80th centile cut-off is shown as this was found to be the most ideal cut-point to dichotomize samples into the APB_{high} and APB_{low} groups.

[0022] **Fig. 2** shows the alternate promoter utilization in gastric cancer (discovery cohort, $n = 24$). In particular, (A) shows a heatmap of alternate promoter utilization in discovery
 15 cohort of patients treated with nivolumab or pembrolizumab. Transcript with higher than 4-fold expression level compared to the median level in all tumour and mapping to the previously identified gain alternative promoter site were considered as gained alternative promoter (marked red in the heatmap). Transcript with lower than 4-fold expression level compared to the median level in all tumour and mapping to the previously identified lost
 20 alternative promoter site were considered as lost alternative promoter (marked blue in the heatmap). (B) shows the association between AP_{high(disc)} group vs AP_{low(disc)} group with T-cell immune correlates. AP_{high(disc)} group are in red, whereas those in AP_{low(disc)} group are in blue. Depicted are the expression of T-cell markers CD8A ($P = 0.059$) and the T-cell cytolytic markers GZMA ($P = 0.025$) and PRF1 ($P = 0.011$). AP_{high(disc)} group shows lower expression
 25 of immune markers. (C) shows a Kaplan Meier curve of progression free survival comparing AP_{high} group vs AP_{low} group in the discovery cohort.

[0023] **Fig. 3** shows the alternate promoter utilization in gastric cancer (pembrolizumab trial cohort, $n = 37$). In particular, (A) shows a heatmap of alternate promoter utilization. Transcript with higher than 4-fold expression level compared to the median level in all
 30 tumour and mapping to the previously identified gain alternative promoter site were considered as gained alternative promoter (marked red in the heatmap). Transcript with lower than 0.25-fold expression level compared to the median level in all tumour and mapping to

the previously identified lost alternative promoter site were considered as lost alternative promoter (marked blue in the heatmap). (B) shows a graph of alternative promoter utilization scores. Alternative promoter utilization score is calculated as the sum of gained and lost alternative promoter in each sample. High alternate promoter utilization was defined as those
 5 > 66th centile. (C) shows the association between AP_{high} group vs AP_{low} group with T-cell immune correlates. AP_{high} group are in red, whereas those in AP_{low} group are in blue. Depicted are the expression of T-cell markers CD8A ($P = 0.0037$) and the T-cell cytolytic markers GZMA ($P = 0.0055$) and PRF1 ($P = 0.016$). AP_{high} group shows lower expression of immune markers. (D) is a waterfall plot of response to pembrolizumab according to AP_{high}
 10 (red) and AP_{low} (blue) subgroups. Y axis represents percentage of maximum tumour reduction assessed according to RECIST 1.1 criteria.

[0024] **Fig. 4** shows the survival curves based on alternate promoter utilization. In particular, (A) shows a Kaplan Meier curve of progression free survival comparing AP_{high} group vs AP_{low} group. (B) shows a Swimmer plot where the x-axis represents the duration of
 15 pembrolizumab therapy for each patient. AP_{high} (red) and AP_{low} (blue) subgroups are depicted. (C) shows a Kaplan Meier curve of progression free survival comparing TCGA subtypes split by AP_{high} group vs AP_{low} group.

[0025] **Fig. 5** shows the bioinformatics workflow for APBscore algorithm. Somatic promoter regions (gain and loss) identified in gastric cancer (2732 regions) are mapped
 20 against all promoters identified in Pan-Cancer analysis from TCGA (113,076 promoters). In total, 4672 promoters were selected, and relative promoter activity used for calculation of APBscore. Tumours are dichotomized into APB_{high} and APB_{low} at the 80th centile of APBscore.

[0026] **Fig. 6** shows the correlation of APBscore groups with markers of cytolytic T-cell activity. In particular, (A) shows the association between APB_{high} group vs APB_{low} group (all
 25 tumour samples, pan-cancer) with T-cell immune correlates. Tumours were dichotomized into APB_{high} and APB_{low} at the 80th centile of APBscore. APB_{high} group are in red, whereas those in APB_{low} group are in blue. Depicted are the expression of T-cell markers CD8A and the T-cell cytolytic markers GZMA and PRF1. APB_{high} group shows lower expression of
 30 immune markers (Wilcoxon test $p < 0.0001$ (***) for CD8A, GZMA and PRF1). (B) shows the association between APB_{high} group vs APB_{low} group with CD8A, GZMA and PRF1 for select tumour types: bladder cancer (BLCA), breast cancer (BRCA), cervical squamous cell

carcinoma (CESC), esophageal carcinoma (ESCA), head and neck squamous cell carcinoma (HNSC), liver hepatocellular carcinoma (LIHC), lung squamous cell carcinoma (LUSC), ovarian cancer (OV) and stomach adenocarcinoma (STAD). APB_{high} group are in red, whereas those in APB_{low} group are in blue. (***) = $p < 0.0001$, (**) = $p < 0.001$, (*) = $p < 0.01$, NS = $p > 0.05$, Wilcoxon test).

[0027] **Fig. 7** shows the differential gene expression in APBscore groups highlighting immune genes. In this volcano plot of immune genes correlated with APB_{high} and APB_{low}, the x-axis is the $\log_2$ fold change ($\log_2FC$) of gene expression (RSEM) between APB_{high} and APB_{low}. The y-axis is the $-\log_{10}$ adjusted p-value results (Bonferroni correction). Significant genes defined as $\log_2FC > 1$ and adjusted p value < 0.05 are expressed as red dots. The top genes ($\log_2FC > 1.3$ and $-\log_{10}$ adjusted p value > 200) are labelled.

[0028] **Fig. 8** shows the association between APBscore groups and other immune-correlates. In particular, the figure shows a volcano plot of $\sim 20,000$ genes in the PanCanAtlas correlated with APB_{high} and APB_{low}. The x-axis is the $\log_2$ fold change ($\log_2FC$) of gene expression (RSEM) between APB_{high} and APB_{low}. The y-axis is the $-\log_{10}$ adjusted p-value results (Bonferroni correction). Immune genes are expressed as red dots. Equal split of overall gene expression is seen between both groups, with a skew of under-expression of immune genes in the APB_{high} group.

[0029] **Fig. 9** shows the correlation of lymphocyte infiltration signature score with APBscore. The lymphocyte infiltration signature is dichotomized by APBscore groups. The APB_{low} group has a statistically higher score compared to APB_{high} (Wilcoxon test $p < 0.0001$).

[0030] **Fig. 10** shows the correlation of IFN- γ response signature with APBscore. The IFN- γ response signature as defined by dichotomized by APBscore groups. The APB_{low} group has a statistically higher score compared to APB_{high} (Wilcoxon test $p < 0.0001$).

[0031] **Fig. 11** shows the correlation between APBscore and objective response rates (ORR) with pembrolizumab in 19 tumour types selected due to correlation with *CD8A*, *GZMA* and *PRF1*. The x-axis is the median APBscore per tumour type and the y-axis is the ORR of the tumour type with single-agent pembrolizumab as described in published trials. The size of the circle corresponds to the number of patients treated in the clinical trial selected for ORR assessment. The color of the circle corresponds to the number of samples within the TCGA cohort used for assessment of APBscore. COAD includes both COAD and READ (colon and rectal cancer).

DETAILED DESCRIPTION OF THE PRESENT INVENTION

[0032] In a first aspect, the present invention refers to method of identifying a patient suffering from cancer as benefiting or not benefiting from immune checkpoint inhibition (ICI) therapy comprising the steps of:

- 5 a) measuring an expression level of one or more preselected markers in a cancerous biological sample obtained from the patient;
- b) identifying a differentially expressed alternative promoter based on the expression level of the one or more preselected markers measured in step (a);
- c) calculating an alternative promoter usage score;
- 10 d) identifying the patient as benefiting or not benefiting from ICI therapy using the alternative promoter usage score.

[0033] It will generally be understood to one of skill in the art that the cancerous biological sample may be a fresh, frozen, fixed or preserved sample.

[0034] In one embodiment, the patient suffering from cancer has undergone or is undergoing ICI therapy. In a preferred embodiment, the patient suffering from cancer has not undergone ICI therapy.

[0035] The cancerous biological sample may be taken from a patient diagnosed with one or more cancers which may include but are not limited to carcinoma, sarcoma and melanoma. The biological sample may be a cell, tissue or fluid sample.

20 [0036] In some embodiments, the cancerous biological sample may be taken from a patient suffering from one or more cancers which include but are not limited to adrenocortical carcinoma, bladder urothelial carcinoma, breast invasive carcinoma, cervical squamous cell carcinoma and endocervical adenocarcinoma, cholangiocarcinoma, colon adenocarcinoma, diffuse large B-cell lymphoma, esophageal carcinoma, glioblastoma multiforme, head and
25 neck squamous cell carcinoma, kidney chromophobe, kidney renal clear cell carcinoma, kidney renal papillary cell carcinoma, acute myeloid leukaemia, brain lower grade glioma, liver hepatocellular carcinoma, lung adenocarcinoma, lung squamous cell carcinoma, mesothelioma, ovarian serous cystadenocarcinoma, pancreatic adenocarcinoma, pheochromocytoma and paraganglioma, prostate adenocarcinoma, rectum adenocarcinoma,
30 sarcoma, skin cutaneous melanoma, stomach adenocarcinoma, testicular germ cell tumours, thyroid carcinoma, thymoma, uterine corpus endometrial carcinoma, uterine carcinosarcoma, uveal melanoma, anal carcinoma and combinations thereof.

[0037] In one embodiment, the cancer is a solid cancer. In another embodiment, the solid cancer is a gastro-intestinal cancer. In yet another embodiment, the gastro-intestinal cancer is a gastric cancer. In yet another embodiment, the gastric cancer is of the chromosomally unstable (CIN) and/or genomically stable (GS) subtype. In yet another embodiment, the
5 gastric cancer is metastatic gastric cancer.

[0038] In some embodiments, the preselected marker is one or more nucleic acid molecules. The nucleic acid molecule may be a DNA or an RNA. In some embodiments, the nucleic acid molecule is selected from the group consisting of an mRNA, a cDNA, a microRNA and a genomic DNA. The preselected marker may be one or more genes or one or
10 more transcripts. In some embodiments, the preselected marker is one or more peptides. In another embodiment, the peptide is a post-translationally modified peptide.

[0039] In one embodiment, the preselected marker is a marker that is associated with a somatic promoter region in a cancerous biological sample. In a preferred embodiment, the cancerous biological sample is a gastric cancer sample.

[0040] The expression level of the one or more preselected markers in the cancerous biological sample described herein may be measured using a method selected from the group consisting of whole transcriptome sequencing (WTS), Nanostring analysis, RNA sequencing, and combinations thereof.

[0041] It would be appreciated by a person skilled in the art that the usage of alternative promoters allows initiation of transcription at different transcription start sites. The usage of
20 alternative promoters can therefore influence gene expression in various ways, including varying gene expression levels and/or causing the production of altered mRNA transcripts and protein isoforms.

[0042] In one embodiment, the differentially expressed alternative promoter described
25 herein is identified by comparing the expression level of the one or more preselected markers in the cancerous biological sample with the expression level of the said one or more preselected markers in one or more reference samples to determine an increase or decrease in expression level of said preselected markers between the cancerous biological sample and the one or more reference samples. It will be understood to a person skilled in the art that the
30 expression level of the preselected markers, which may be one or more genes, transcripts or peptides, may be regulated by the promoter and is therefore indicative of the activity of the promoter.

[0043] A differentially expressed alternative promoter may be identified by comparing the expression level of one or more preselected markers in the cancerous biological sample to the expression level of the one or more preselected markers in the one or more reference samples.

[0044] In one embodiment, an increase in the expression level of the one or more preselected markers in the cancerous biological sample compared to the expression level of the one or more preselected markers in the one or more reference samples for a gained promoter indicates a differentially expressed alternative promoter.

[0045] In another embodiment, a decrease in the expression level of the one or more preselected markers in the cancerous biological sample compared to the expression level of the one or more preselected markers in the one or more reference samples for a lost promoter indicates a differentially expressed alternative promoter.

[0046] The differentially expressed alternative promoter may be identified in various ways. In one embodiment, the differentially expressed alternative promoter described herein is identified by a gained promoter with an at least 4-fold increase in the expression level of transcripts when compared to the median expression level of transcripts for the promoter in a panel of tumour samples. In another embodiment, the differentially expressed alternative promoter described herein is identified by a lost promoter with an expression level of the transcripts which is less than 0.25-fold or at least a 4-fold reduction when compared to the median expression level of transcripts for the promoter in a panel of tumour samples. In one embodiment, the median expression level of transcripts for each promoter is the median expression level of the promoter across all samples in a panel of tumour samples.

[0047] In another embodiment, the differentially expressed alternative promoter described herein is identified by a gained promoter with a relative promoter activity greater than the median promoter activity of a panel of tumour samples. In another embodiment, the differentially expressed alternative promoter described herein is identified by a lost promoter with a relative promoter activity less than the median promoter activity of a panel of tumour samples. In one embodiment, the median expression level of transcripts for each promoter is the median expression level of the promoter across all samples in a panel of tumour samples.

[0048] In one embodiment, the gained promoter is a promoter that is gained or has increased activity in a cancerous biological sample compared to a non-cancerous biological sample. The gained promoter may be a promoter in the cancerous biological sample that is not present in the non-cancerous biological sample, or a promoter in the cancerous biological

sample that has an increased promoter activity compared to the promoter in the non-cancerous biological sample.

[0049] In another embodiment, the lost promoter is a promoter that is lost or has decreased activity in a cancerous biological sample compared to a non-cancerous biological sample. A
5 lost promoter may be a promoter that is present in the non-cancerous biological sample but not in the cancerous biological sample, or a promoter in the cancerous biological sample that has a decreased promoter activity compared to the promoter in the non-cancerous biological sample.

[0050] After the identification of differentially expressed alternative promoters as
10 described herein, an alternative promoter usage score is then calculated. In one embodiment, the alternative promoter usage score may be an AP score. In another embodiment, the alternative promoter usage score may be an APBscore. The alternative promoter usage score may be calculated by determining the sum of the differentially expressed alternative promoters as described herein.

[0051] The alternative promoter usage score may be used in a method of identifying a patient suffering from cancer as benefiting or not benefiting from ICI therapy. The method of identifying a patient suffering from cancer as benefiting or not benefiting from ICI therapy may further comprise a step of comparing the alternative promoter usage score as described herein with a reference score to identify the patient as benefiting or not benefiting from ICI
20 therapy.

[0052] In one embodiment, the reference score is a score at the median, tertile or quartile of alternative promoter usage scores from one or more reference samples. In another embodiment, the reference score is a score at the 10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th or 90th percentile of alternative promoter usage scores from one or more reference samples. In a
25 preferred embodiment, the reference score may be a score at the 66th or 80th percentile of alternative promoter usage scores from one or more reference samples.

[0053] In a preferred embodiment, the reference score is an absolute score. The absolute score is a fixed and non-variable score to which the patient's alternative promoter score is compared.

[0054] In one embodiment, an increased alternative promoter usage score compared to the reference score identifies the patient as not benefiting from ICI therapy. In one embodiment, the increased alternative promoter usage score is a score above the 66th percentile of

alternative promoter usage scores from one or more reference samples. In another embodiment, the increased alternative promoter usage score is a score above the 80th percentile of alternative promoter usage scores from one or more reference samples.

[0055] In another embodiment, a decreased alternative promoter usage score compared to the reference score identifies the patient as benefiting from ICI therapy. In one embodiment, the decreased alternative promoter usage score is a score below the 66th percentile of alternative promoter usages scores from one or more reference samples. In one embodiment, the decreased alternative promoter usage score is a score below the 80th percentile of alternative promoter usages scores from one or more reference samples.

[0056] In one aspect, the present invention provides a method of determining the prognosis of a patient who is undergoing or has undergone ICI therapy, the method comprising the steps of:

- a) measuring an expression level of one or more preselected markers in a cancerous biological sample obtained from the patient;
- b) identifying a differentially expressed alternative promoter based on the expression level of the one or more preselected markers measured in step (a);
- c) calculating an alternative promoter usage score;
- d) determining the prognosis of the patient using the alternative promoter usage score.

[0057] In one embodiment, the method of determining the prognosis of a patient who is undergoing or has undergone ICI therapy as described herein further comprises the step of comparing the alternative promoter usage score described herein with a reference score to determine the prognosis of the patient, wherein an increased alternative promoter usage score compared to a reference score indicates a poorer prognosis.

[0058] In one embodiment, the increased alternative promoter usage score is a score above the 66th percentile of alternative promoter usage scores from one or more reference samples. In another embodiment, the increased alternative promoter usage score is a score above the 80th percentile of alternative promoter usage scores from one or more reference samples.

[0059] In another embodiment, the method of determining the prognosis of a patient who is undergoing or has undergone ICI therapy as described herein further comprises the step of comparing the alternative promoter usage score described herein with a reference score to determine the prognosis of the patient, wherein a decreased alternative promoter usage score compared to a reference score indicates a better prognosis.

[0060] In one embodiment, the decreased alternative promoter usage score is a score below the 66th percentile of alternative promoter usages scores from one or more reference samples. In another embodiment, the decreased alternative promoter usage score is a score below the 80th percentile of alternative promoter usages scores from one or more reference samples.

5 [0061] In another aspect, the present invention provides a biomarker for excluding a patient suffering from cancer from ICI therapy, wherein the biomarker is an alternative promoter usage score that is above the alternative promoter usage score of one or more reference samples, wherein the alternative promoter usage score is the sum of the differentially expressed alternative promoters at each alternative promoter site.

10 [0062] In one embodiment, the present invention provides a biomarker for excluding a patient suffering from cancer for ICI therapy, wherein the biomarker is an alternative promoter usage score that is above the 66th percentile of alternative promoter usage scores of one or more reference samples, wherein the alternative promoter usage score is the sum of the differentially expressed alternative promoters at each alternative promoter site

15 [0063] In another embodiment, the present invention provides a biomarker for selecting a patient suffering from cancer for ICI therapy, wherein the biomarker is an alternative promoter usage score that is below the 66th percentile of alternative promoter usage scores of one or more reference samples, wherein the alternative promoter usage score is the sum of the differentially expressed alternative promoters at each alternative promoter site.

20 [0064] In another aspect, the present invention provides a biomarker for excluding a patient suffering from cancer from ICI therapy, wherein the biomarker is an alternative promoter usage score that is above the 80th percentile of alternative promoter usage scores of one or more reference samples, wherein the alternative promoter usage score is the sum of the differentially expressed alternative promoters at each alternative promoter site.

25 [0065] In another embodiment, the present invention provides a biomarker for selecting a patient suffering from cancer for ICI therapy, wherein the biomarker is an alternative promoter usage score that is below the 80th percentile of alternative promoter usage scores of one or more reference samples, wherein the alternative promoter usage score is the sum of the differentially expressed alternative promoters at each alternative promoter site.

30 [0066] The one or more reference samples may be selected from the group consisting of one or more tumour samples obtained from one or more different patients, one or more non-

cancerous samples obtained from one or more different patients, one or more non-cancerous samples obtained from the same patient and combinations thereof. In one embodiment, the one or more non-cancerous samples obtained from the same patient is tissue adjacent to the cancerous tissue. It will be understood by the person skilled in the art that multiple combinations of these reference samples may be used in the methods of the invention.

[0067] In another aspect, the present invention refers to a method of determining if a patient suffering from cancer should not be treated with immune checkpoint inhibitor (ICI) therapy, the method comprising the steps of:

- a) measuring an expression level of one or more preselected markers in a cancerous biological sample obtained from the patient;
- b) identifying a differentially expressed alternative promoter based on the expression level of the one or more preselected markers measured in step (a);
- c) calculating an alternative promoter usage score;
- d) identifying the patient as not benefiting from ICI therapy using the alternative promoter usage score;
- e) treating the patient with an alternative therapy when said patient is identified as not benefiting from ICI therapy.

[0068] In another aspect, the present invention refers to a method of treating a patient suffering from cancer, the method comprising the steps of:

- a) measuring an expression level of one or more preselected markers in a cancerous biological sample obtained from the patient;
- b) identifying a differentially expressed alternative promoter based on the expression level of the one or more preselected markers measured in step (a);
- c) calculating an alternative promoter usage score;
- d) identifying the patient as benefiting or not benefiting from ICI therapy using the alternative promoter usage score;
- e) treating the patient identified as benefiting from ICI therapy according to step (d) with an immune checkpoint inhibitor, or treating the patient identified as not benefiting from ICI therapy according to step (d) with an alternative therapy.

[0069] In one embodiment, ICI therapy may include but is not limited to ipilimumab, nivolumab, pembrolizumab, atezolizumab, avelumab, durvalumab, cemiplimab, spartalizumab, sintilimab, camrelizumab and tislelizumab.

[0070] In another embodiment, the alternative therapy is any therapy other than ICI therapy that is suitable for treating cancer. For example, alternative therapies may include chemotherapy, radiation therapy, stem cell transplant, surgery, hormone therapy and targeted therapy. Targeted therapies involve the use of drugs that interfere with specific molecules
5 necessary for tumour growth and progression. An alternative therapy may also include ceasing ICI therapy in a patient undergoing ICI therapy.

[0071] The invention illustratively described herein may suitably be practiced in the absence of any element or elements, limitation or limitations, not specifically disclosed herein. Thus, for example, the terms "comprising", "including", "containing", etc. shall be
10 read expansively and without limitation. Additionally, the terms and expressions employed herein have been used as terms of description and not of limitation, and there is no intention in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the
15 present invention has been specifically disclosed by preferred embodiments and optional features, modification and variation of the inventions embodied therein herein disclosed may be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of this invention.

[0072] The invention has been described broadly and generically herein. Each of the
20 narrower species and subgeneric groupings falling within the generic disclosure also form part of the invention. This includes the generic description of the invention with a proviso or negative limitation removing any subject matter from the genus, regardless of whether or not the excised material is specifically recited herein.

[0073] Other embodiments are within the following claims and non-limiting examples. In
25 addition, where features or aspects of the invention are described in terms of Markush groups, those skilled in the art will recognize that the invention is also thereby described in terms of any individual member or subgroup of members of the Markush group.

EXPERIMENTAL SECTION

30 [0074] Non-limiting examples of the invention and comparative examples will be further described in greater detail by reference to specific Examples, which should not be construed as in any way limiting the scope of the invention.

[0075] **Materials and Methods**

[0076] REMARK criteria for validation of tumour biomarkers was followed in this study.

[0077] Clinical cohorts

[0078] *Discovery cohort*

5 [0079] Consecutive patients with metastatic gastric cancer treated with nivolumab or pembrolizumab treatment at Samsung Medical Centre, Seoul, Korea were included in this cohort. ICIs were administered as salvage treatment in patients who failed to at least one cytotoxic regimen. Nivolumab 3 mg/kg was administered as a 1-hour infusion every 2 weeks and Pembrolizumab 200 mg was administered as a 30-minute intravenous infusion every
10 3 weeks until disease progression or unacceptable toxicity. Ethics approval was obtained, and all patients provided written informed consent before archival tumour tissue specimens from primary tumours were collected and prospectively followed up for survival data.

[0080] *Pembrolizumab trial cohort*

[0081] Patients with histologically proven metastatic and/or recurrent gastric
15 adenocarcinoma that had failure of at least 1 line of chemotherapy that included platinum/fluoropyrimidine were enrolled in this study. The trial was conducted in accordance with the Declaration of Helsinki and the Guidelines for Good Clinical Practice (ClinicalTrials.gov identifier: NCT#02589496). The trial protocol was approved by the Institutional Review Board of Samsung Medical Center (Seoul, Korea) and all patients
20 provided written informed consent before enrolment. Pembrolizumab 200 mg was administered as a 30-minute intravenous infusion every 3 weeks until documented disease progression, unacceptable toxicity, or up to 24 months. Tumour responses were evaluated every two cycles according to RECIST 1.1 criteria.

[0082] Nanostring analysis

25 [0083] NanoString nCounter Reporter CodeSets were designed for 80 recurrent somatic alternate promoter related genes, as well as immune-related genes corresponding to intra-tumoural cytolytic activity (CYT), cytokines and immune checkpoints. At least two probes were designed for each gene to measure the expression of canonical and alternate promoter-driven transcripts. A canonical probe at the 5' transcript marked by unaltered H3K4me3, and
30 an alternate probe at the 5' transcript of the somatic promoter. Data analysis was performed using the vendor-provided nCounter software (nSolver). Raw counts were normalized using the geometric mean of the internal positive control probes included in each CodeSet.

[0084] RNA sequencing

[0085] Tumour tissues were obtained between day -42 and day 1 prior to initiation of study treatment. If tumour content was estimated as more than 40% after thorough pathological review, tumour DNA and RNA were extracted from freshly obtained tissues
5 using a QIAamp Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The concentrations and 260/280 and 260/230 nm ratios were measured with an ND1000 spectrophotometer (Nanodrop Technologies, Thermo-Fisher Scientific, MA, USA) and then further quantified DNA/RNA using a Qubit fluorometer (Life Technologies, CA, USA).

10 [0086] RNA transcriptomic analysis

[0087] RNAseq data was aligned to GENCODE v19 transcript annotation using TopHat and FPKM abundance measure were generated using Cufflinks. Transcripts were then merged across all samples and normalized using Cuffnorm. To analyze alternative promoter-associated expression, RNAseq reads were mapped against the genomic location previously
15 identified by epigenomic profiling. RNAseq mapping to these epigenome-defined promoter regions were then quantified, normalized by promoter length and by library size. Finally, fold changes in expression at each promoter site were computed between each tumour and the median expression level across all tumour samples.

[0088] PDL1 immunohistochemistry analysis, MSI status, EBV status, TCGA subtyping
20 and tumour mutational burden were based on classifications used for a phase II study of single agent pembrolizumab in metastatic gastric cancer.

[0089] Statistical analysis

[0090] Associations of clinicopathologic features to histologic sub-classification was performed using Fisher's Exact Test. Progression-free survival (PFS) was calculated from the
25 time of first dose of pembrolizumab to the time of disease progression or death, and overall survival (OS) was calculated from time of first dose of pembrolizumab or nivolumab to time of death. Kaplan-Meier (KM) curves and log rank test were used for survival analysis. The hazard ratio (HR) and its 95% confidence interval (CI) were evaluated for each analysis using Cox proportional hazards regression model. All analyses were done using R (3.4.1). In the
30 validation cohort, samples with alternate promoter usage score greater than the 66th percentile were defined as high alternate promoter utilization (AP_{high}) and remaining as low alternate promoter utilization (AP_{low}).

[0091] PDL1 Immunohistochemistry

[0092] PDL1 immunohistochemistry was performed using the Dako PD-L1 IHC 22C3 pharmDx kit (Agilent Technologies). PD-L1 protein expression was determined using CPS, which was the number of PD-L1 staining cells (tumour cells, lymphocytes, macrophages) divided by the total number of viable tumour cells, multiplied by 100.

[0093] MSI status

[0094] Tumour tissue MSI status was determined by both IHC for MLH1 and MSH2 and PCR analysis of 5 markers with mononucleotide repeats.

[0095] EBV status subtypes

10 [0096] EBV status was determined by EBV-encoded small RNA (EBER) in situ hybridization.

[0097] Tumour mutational burden

[0098] Mutational load was determined from whole exome sequencing analysis. Mutational load for a subject was defined as the number of somatic non-synonymous SNVs that passed all filters. Somatic mutations were annotated with variant effect predictors. Mutational load was calculated as the number of non-synonymous SNVs in the tumour exome data. The ML-H threshold for tissue was set as the upper tertile.

[0099] TCGA subtype definition

[00100] Gastric cancer subtypes defined by TCGA was based on DNA genomic alterations. 20 These groups included EBV(+), MSI-H, CIN and genome stable tumours, which lack CIN and are heavily enriched in the diffuse histologic subtype. As a proxy for CIN, EBV(-), MSS tumours were stratified into CIN and genome stable based on their TP53 status. Mutational signature analysis was performed using the deconstructSigs package (v1.6.0) in R.

[00101] Dataset

25 [00102] Promoter activity was inferred from RNA-Seq data available from the PanCanAtlas of the TCGA consisting of 10393 samples (9668 tumour and 725 normal samples) across 33 tumour types.

[00103] APBscore algorithm

[00104] The initial epigenetic promoter alteration study in gastric cancer identified 2732 30 somatic promoter regions (2053 gained and 679 lost) and transcriptomic expression of these isoforms were used to create the APBscore algorithm.

[00105] For the TCGA data, GENCODE v19 annotation was used to determine the set of promoters. The overlapping first exons of each TSS were combined to obtain a set of promoters. Activity of each promoter was then quantified using junction reads aligning into the first introns of the constituting transcripts. The total junction read counts were then
 5 normalized across the entire dataset. The log₂ transformed normalized read counts were used for promoter activity in further downstream analysis. Gene expression estimates were obtained by summing up the activities of each promoter per gene. Each promoter's activity was then normalized by gene expression to obtain relative promoter activities. In total 113,076 promoters were identified.

10 [00106] Somatic promoter regions identified in gastric cancer were then selected from this set of 113,076 promoters. 4672 promoters (3263 in gained region, 1409 in lost regions) were located within the 2732 somatic promoter regions identified in gastric cancer. Of these 4672 promoters, the median relative promoter activity of every promoter across all tumour samples (n=9668) was calculated. For every sample, the APBscore was calculated as the number of
 15 gained promoters with relative promoter activity greater than median and lost promoters with relative promoter activity less than median (**Fig. 1**). APBscore at various centile cut-offs (10th, 20th, 30th ... , 90th) were used to dichotomize tumours into APB_{high} and APB_{low}. The 80th centile was determined as the ideal cut-off for the APB_{high} and APB_{low} groups and used for further downstream analysis. Normal samples were only added into the cohort for the
 20 analyses involving normal tissue and were not used in the calculation of median scores and cut-offs.

[00107] Immune correlates

[00108] Transcriptomic expression levels of *CD8A*, *GZMA* and *PRFI* and a selection of more than 700 genes including immune checkpoints, markers of various immune cell types,
 25 genes related to the adaptive and innate immune response and antigens was selected and extracted from the TCGA dataset from the Broad GDAC Firehose. Immune subtypes, other immune signatures and TMB were extracted from a Pan-Cancer immune landscape analysis. Progression-free survival (PFS) and overall survival (OS) including censorship data was extracted from the Pan-Cancer immune landscape analysis.

30 [00109] Immune Checkpoint Inhibitor clinical data

[00110] Data from clinical trials involving immune checkpoint inhibitors is rapidly expanding, with drugs being tested as single agent and in combination with other therapies.

Pembrolizumab remains the immune checkpoint inhibitor that has been tested across the broadest spectrum of tumour types with clinical trial data being reported. Therefore, an extensive literature review of single agent pembrolizumab trials was conducted and the largest studies that were reported for each tumour type were selected. Objective response rates (ORR) from these trials were then correlated with APBscores of each tumour type. TMB was extracted from the Pan-Cancer immune landscape analysis and *PD-L1* transcript expression was extracted from the Broad GDAC Firehose.

[00111] Statistical Analysis

[00112] Wilcoxon rank-sum test was used to compare expression levels of *CD8A*, *GZMA* and *PRFI* between APB_{high} and APB_{low} groups. The “Rtsne” R package was used to generate T-SNE plots. Pearson’s test was used for correlation between APBscore and ORR of pembrolizumab. Kaplan-Meier (KM) curves and log rank test were used for survival analysis. The hazard ratio (HR) and its 95% confidence interval (CI) were evaluated for each analysis using Cox proportional hazards regression model. Network mapping of somatic promoter gene function was done using the Reactome Pathway Database. All analyses were done using R (3.5.2).

[00113] Validation of estimation of promoter activity using RNA-Seq

[00114] The accuracy of promoter activity estimation using RNA-Seq was validated by bench-marking against other “gold-standard” measures such as H3K4me3 ChIP-Seq and CAGE tag data available on public data-sets. H3K4me3 levels correlated strongly with RNA-Seq promoter activity, suggesting that epigenetic-based and transcript-based estimates of promoter activity were consistent (Kruskal-Wallis $p < 0.001$). The findings were confirmed with CAGE-Tag data, with unique promoters identified on RNA-Seq having higher CAGE-Tag support.

[00115] The present algorithm was also compared with other bioinformatic methods utilized in RNA-Seq quantification and first exon read counts. The present algorithm yielded results similar to these algorithms with high levels of correlation (Pearson’s correlation coefficient > 0.85). Overall, this analysis demonstrates that our approach enables the quantitative, robust, and reproducible estimation of promoter activity from RNA-Seq data.

[00116] **Example 1: Alternate promoter utilization in metastatic gastric cancer**

[00117] The first cohort consisted of 24 metastatic gastric cancer patients treated with nivolumab and pembrolizumab (29 subjects were initially included, with 24 tumour samples

passing quality control for sufficient tissue for Nanostring analysis). A customized Nanostring panel was used to measure transcripts associated with either the canonical or alternate promoter. Differentially expressed alternative promoters were defined as a promoter site showing $< 0.25x$ fold change (for lost somatic promoters) or $> 4x$ fold change in expression level (for gained somatic promoters) over the median across all samples. Using this algorithm, it was found that a third of the tumours (8/24) displayed high alternate promoter utilization in more than 10% of the sites ($> 8/80$). This group was defined as AP_{high} while the rest were defined as AP_{low} (**Fig. 2A**).

[00118] Measurement of cytolytic T-cell activity had previously been described by studying expression of *CD8A* (CD8+ tumour infiltrating lymphocytes), granzyme A (*GZMA*) and perforin 1 (*PRF1*). The AP_{low} group demonstrated significantly increased expression of *GZMA* ($P = 0.025$), *PRF1* ($P = 0.011$) and *CD8A* ($P = 0.059$) when compared to the AP_{high} group suggesting increased cytotoxic T-cell activity in the AP_{low} group (**Fig. 2B**). These findings are concordant with earlier results described in early gastric cancer, demonstrating that alternate promoter utilization in metastatic gastric cancers is also inversely related to anti-tumour immunity. Notably, despite the small sample size, heterogenous treatment regimens, and non-trial based nature of this cohort, there was a trend for patients with AP_{high} tumours to have worse progression-free survival (PFS) compared to patients with AP_{low} tumours (129 days vs 389 days, HR 1.96 95% CI: 0.55 to 6.93, $P = 0.29$, **Fig. 2C**). Based on these findings, this hypothesis was further tested in a separate cohort of uniformly-treated patients.

[00119] **Example 2: Alternate promoter utilization as a predictor of response and survival with pembrolizumab treatment**

[00120] For the second cohort, transcriptomic data from the phase II study described earlier was used. Transcriptomic data from pretreatment biopsy samples and matched clinical data was available for 37 subjects and used for analysis. The median age was 57 years, 73% were male ($N = 27$), 4 (11%) were EBV positive and 4 were MSI (11%) with the rest defined as CIN or GS TCGA subtype. Complete or partial responses to therapy was seen in 11 subjects (30%). Using 2732 somatic alternate promoter sites previously identified in gastric cancer [10], differentially expressed alternative promoters were defined similar to the first cohort ($< 0.25x$ fold change for known somatically lost promoters or $> 4x$ fold change for known somatically gained promoters). Notably, good concordance between RNAseq and Nanostring

platforms for assessment of alternate promoter utilization has previously been shown. The sum of differentially expressed sites in each sample was calculated to define an alternate promoter usage score (**Fig. 3A**). Scores ranged from 37 to 426 (median 136) (**Fig. 3B**). Using data from the first cohort to guide cut-off points, the AP_{high} group was defined as samples

5 >66% centile (n = 13), while AP_{low} constituted the remaining samples.

[00121] The AP_{high} group had no statistically significant differences in clinicopathological characteristics compared to the AP_{low} group for age, gender or histological subtype. No differences were also detected between the two groups between TCGA subtypes, mutational load, and PDL1 CPS scores (**Table 1**). The AP_{low} group demonstrated significantly increased

10 expression of *CD8A* ($P = 0.0037$), *GZMA* ($P = 0.0055$) and *PRF1* ($P = 0.016$) when compared to the AP_{high} group suggesting increased cytotoxic T-cell activity in the AP_{low} group (**Fig. 3C**). Objective response rate (ORR), defined as either partial or complete response to therapy, was higher in the AP_{low} group compared to AP_{high} group (10/24 vs 1/13, $P = 0.03$) (**Fig. 3D**). Of note, in the AP_{high} group, the only response was in an MSI subtype

15 tumour. Median progression free survival (PFS) was 55 days in the AP_{high} group compared to 180 days in the AP_{low} group (logrank $P = 0.0076$) (**Fig. 4A, 5B**). The AP_{low} group had 17% EBV (n = 4) and 12% MSI (n = 3) TCGA subtype samples, while the AP_{high} group had only 8% MSI (n = 1) and no EBV samples (**Fig. 4C**). As previously shown, PFS between the various TCGA subtypes were different ($P = 0.0026$), with the MSI and EBV subtypes having

20 a significantly longer survival (491 days (MSI/EBV) vs 80 days (CIN/GS)). Notably, amongst the CIN/GS subtype, PFS was also statistically significantly different between the AP_{low} and AP_{high} groups (48 days (CIN/GS AP_{high}) vs 161 days (CIN/GS AP_{low}), $P = 0.0019$) (**Fig. 4D**). Based on preliminary overall survival data, a trend towards improved survival was seen in the AP_{low} group (340 days vs 292 days, $P = 0.16$). Multivariate analysis of

25 clinicopathologic and alternate promoter utilization revealed high alternate promoter utilization as an independent predictive factor for PFS with pembrolizumab (HR 0.29, (95% CI 0.099–0.85), $P=0.024$) (**Table 2**).

[00122] **Table 1.** Differences in characteristics between AP_{high} and AP_{low} groups.

Characteristics	AP Groups		P-value
	AP _{High} (n = 13)	AP _{Low} (n = 24)	
	n (%)	n (%)	
Age, years			
<57	6 (46)	11 (46)	1

≥57	7 (54)	13 (54)	
Gender			
Male	11 (85)	16 (67)	0.44
Female	2 (15)	8 (33)	
Primary tumour location			
Body	6 (46)	17 (71)	0.17
Antrum/Others	7 (54)	7 (29)	
Histology			
Poorly differentiated	5 (38)	15 (62)	0.34
Well/Moderately differentiated	5 (38)	5 (21)	
Signet Ring/Others	3 (23)	4 (17)	
Immune Signature			
High	5 (38)	14 (58)	0.31
Low	8 (62)	10 (42)	
Mesenchymal subtype (ACRG)			
Non-mesenchymal	9 (69)	22 (92)	0.16
Mesenchymal	4 (31)	2 (8)	
Mutational load			
Low	3 (23)	13 (54)	0.13
Moderate	8 (62)	7 (29)	
High	2 (15)	4 (17)	
TCGA Subtype			
CIN	7 (54)	6 (25)	0.26
GS	5 (38)	11 (46)	
EBV	0 (0)	4 (17)	
MSI	1 (8)	3 (12)	
P53 mutation status			
Mutated	6 (46)	17 (71)	0.17
Wild type	7 (54)	7 (29)	
KRAS mutation status			
Mutated	3 (23)	4 (17)	0.68
Wild type	10 (77)	20 (83)	
PDL1 CPS*			
0	8 (62)	6 (29)	0.22
1	3 (23)	5 (24)	
2-9	0 (0)	3 (14)	
≥10	2 (15)	7 (33)	

Fisher test

* 3 samples did not have tissue available for PDL1 CPS scoring by immunohistochemistry, all from AP_{low} group.

[00123] **Table 2.** Univariate and Multivariate Survival Analysis.

Variable	Univariate		Multivariate	
	HR (95% CI)	P Value	HR (95% CI)	P Value
Alternate Promoter Group: High vs Low	0.36(0.17-0.78)	0.0094	0.29 (0.099 – 0.85)	0.024
TCGA subgroups: CIN/GS vs EBV/MSI	0.10(0.024-0.45)	0.0024	0.31 (0.052 – 1.88)	0.20
Age	0.98 (0.95 – 1)	0.061		
Gender : Female vs Male	1.2 (0.52 – 2.5)	0.72		
Primary tumour site: Body vs Antrum/Others	0.61 (0.28 – 1.3)	0.22		
Histology Poorly diff/Signet/others vs Well/Mod	0.98(0.44 – 2.2)	0.97		
Mutational Load High vs Low/Mod	5.67(1.3 24)	0.018	3.46 (0.52 – 23)	0.20
<i>P53</i> mutation status Mutated vs wildtype	0.74 (0.36 – 1.56)	0.43		
<i>KRAS</i> mutation status Mutated vs wildtype	2.72 (0.92 – 7.96)	0.068		
Mesenchymal subtype Mesenchymal vs Non-mesenchymal	0.73(0.29 – 1.8)	0.48		
Immune Signature High vs Low	2.31(1.1 – 4.8)	0.025	1.82 (0.80 – 4.12)	0.15
PDL1 CPS ≥ 1 vs 0	7.62 (3 – 19)	<0.001	3.09 (1.15 – 8.29)	0.025

[00124] **Example 3: Alternate promoter utilization evolution post-treatment with pembrolizumab**

- 5 [00125] Paired biopsy samples were available for eight subjects from the second cohort, providing an opportunity to monitor tumour evolution as a consequence of ICI therapeutic pressure. Post-treatment biopsies were taken from the primary stomach tumour at the point of progression on pembrolizumab. Of these eight subjects, two had partial response (PR), with duration of response of 211 and 491 days (both AP_{low}), one had stable disease (SD), with
- 10 duration of response for 167 days (AP_{low}), and five had progressive disease (PD) (AP_{high} $N = 3$; AP_{low} $N = 2$) as best response. Interestingly, very consistent shifts were observed in the directionality of alternative promoter utilization based on clinical responses. Specifically, tumours with PR and SD exhibited x1.5 or higher increase in alternate promoter usage score in the post-treatment biopsy samples compared to pre-treatment biopsy samples, while all

five tumours with PD exhibited reductions in alternate promoter usage scores in the post-treatment biopsy sample (Fisher's exact test, $P = 0.018$). These results further support a relationship between alternative promoter landscapes and ICI therapeutic pressure.

[00126] **Example 4: Alternate promoter utilization across multiple tumour types**

- 5 [00127] To investigate if the applicability of these findings to other tumour types besides gastric cancer, a recently described algorithm was used to infer promoter activity in the PanCanAtlas RNA-seq database of 10,393 samples (9668 tumour and 725 normal samples) across 33 tumour types (**Table 3**). Briefly, using Gencode (release 19) annotations, a set of 113,076 possible promoters was compiled. Promoter activity was inferred by quantifying the
- 10 expression initiated at each promoter using unique junctional reads, assuming that isoforms with identical or very close TSSs are regulated by the same promoter.

[00128] **Table 3.** TCGA cohort description and sample size.

Tumour type	TCGA code	Tumour samples (n)	Normal samples (n)	Tumour + normal (n)
Adrenocortical carcinoma	ACC	79	0	79
Bladder Urothelial Carcinoma	BLCA	405	19	424
Breast invasive carcinoma	BRCA	1095	112	1207
Cervical squamous cell carcinoma and endocervical adenocarcinoma	CESC	306	3	309
Cholangiocarcinoma	CHOL	36	9	45
Colon adenocarcinoma	COAD	281	41	322
Diffuse Large B-cell Lymphoma	DLBC	48	0	48
Esophageal carcinoma	ESCA	185	11	196
Glioblastoma multiforme	GBM	165	5	170
Head and Neck squamous cell carcinoma	HNSC	504	44	548
Kidney Chromophobe	KICH	66	25	91
Kidney renal clear cell carcinoma	KIRC	530	72	602
Kidney renal papillary cell carcinoma	KIRP	291	32	323
Acute Myeloid Leukemia	LAML	171	0	171
Brain Lower Grade Glioma	LGG	528	0	528
Liver hepatocellular carcinoma	LIHC	373	50	423
Lung adenocarcinoma	LUAD	506	59	565
Lung squamous cell carcinoma	LUSC	500	51	551
Mesothelioma	MESO	87	0	87
Ovarian serous cystadenocarcinoma	OV	307	0	307
Pancreatic adenocarcinoma	PAAD	179	4	183
Pheochromocytoma and Paraganglioma	PCPG	184	3	187
Prostate adenocarcinoma	PRAD	498	52	550
Rectum adenocarcinoma	READ	95	10	105
Sarcoma	SARC	263	2	265

Skin Cutaneous Melanoma	SKCM	472	1	473
Stomach adenocarcinoma	STAD	415	35	450
Testicular Germ Cell Tumours	TGCT	156	0	156
Thyroid carcinoma	THCA	509	59	568
Thymoma	THYM	120	2	122
Uterine Corpus Endometrial Carcinoma	UCEC	177	24	201
Uterine Carcinosarcoma	UCS	57	0	57
Uveal Melanoma	UVM	80	0	80
Combined		9668	725	10393

[00129] Of the 113,076 promoters, 4672 promoters mapped to 2732 tumour-associated promoter regions previously defined in gastric cancer. It was hypothesized that alterations in this set of promoters might represent a generalized pan-tumour response to host immune recognition, not limited to gastric cancer. To test this hypothesis, the 4672 promoters were used to compute APB levels (APBscore) for each tumour (**Fig. 5**). As shown in **Fig. 1**, a wide range of APBscores were observed both within and between tumour types. The median APBscore was 178 for the entire cohort (range: 46 to 241). Tumour types with the lowest median APBscores included thyroid cancer (THCA) (median 165, range: 54 to 207), diffuse large B cell lymphoma (DLBC) (median: 166, range: 108 to 196) and lung adenocarcinoma (LUAD) (median: 167, range: 141 to 215). In contrast, glioblastoma (GBM) (median: 207, range: 172 to 241), low grade glioma (LGG) (median: 204, range: 130 to 230) and testicular germ cell tumours (TGCT) (median: 189, range: 83 to 218) had the highest median APBscores.

[00130] Tumour APBscores were correlated with markers of T-cell cytolytic activity: *CD8A*, *GZMA* and *PRF1*. Similar to other pan-cancer TCGA studies, the primary analyses were conducted on the entire cohort of TCGA samples agnostic of tumour-type, but tumour-type specific results were also provided. Previously, it was observed that dichotomizing tumours at the median or top tertile yielded significant correlations with *CD8A*, *GZMA* and *PRF1*. However, as these studies were done with smaller sample sizes, it was opted to re-establish the ideal cut-off for dichotomizing tumours, with respect to correlation with these three genes. APBscore thresholds were thus tested at various centiles (10th, 20th, 30th ..., 90th) to dichotomize the APB_{high} and APB_{low} groups. For the entire cohort (all tumour types), at all nine cut-offs for dichotomy, APB_{high} tumours had significantly lower levels of *CD8A*, *GZMA* and *PRF1* compared to APB_{low} tumours ($p < 0.0001$) (**Fig. 6A**). The 80th centile was found to have the strongest correlation with the three genes, with 19 out 33 tumour types correlating

significantly with *CD8A*, *GZMA* and *PRF1*. For subsequent analysis, tumours were dichotomized into the APB_{high} and APB_{low} groups at the 80th centile (APBscore cutoff: 190) of the entire cohort. Importantly, previous findings in stomach cancer (STAD) were robustly replicated in this analysis, even using a different technique for interrogating alternate promoters (**Fig. 6B**). To address potential confounding by tumour-type, a tumour-type specific analysis was also conducted, where each tumour-type was dichotomized into APB_{high} and APB_{low} groups at the median APBscore of that specific tumour type. For 20 of the 33 tumour types, there was significant correlation with *CD8A*, *GZMA* and *PRF1*. Eight tumour types did not have a significant correlation at any cut-off: kidney chromophobe (KICH), DLBC, thymoma and thymic carcinoma (THYM), kidney renal cell cancer (KIRC), acute myeloid leukemia (LAML), cholangiocarcinoma (CHOL), uterine carcinoma sarcoma (UCS), pheochromocytoma and paraganglioma (PCPG). Of these eight, CHOL, KICH and UCS have relatively small sample sizes ($n < 100$), while DLBC, THYM, LAML are of hematologic/immune origin.

[00131] Notably, compared to the set of 4672 promoters originally defined in gastric cancer used to define APBscore, similar strengths of correlation to *CD8A*/*GZMA*/*PRF1* expression were not observed if APBscore was inferred using all 113,076 identified promoters, or other similarly sized promoter subsets obtained from permutation testing (empirical $p < 0.001$). This observation supports the hypothesis that although the APBscore was originally derived from gastric cancer, the APBscore algorithm may apply across multiple tumour types possibly reflecting a conserved pan-cancer response to host immunity. To investigate specific functions of the tumour-associated promoter isoforms, 570 promoters (12% of 4672) upregulated in at least 100 tumours (~ 1% of cohort of 9668 tumours) were analyzed. Different tumour types utilized specific clusters of promoter isoforms, and network analysis of gene functions revealed that alternate promoters affect genes with diverse roles. This may suggest that on a genome-wide scale, alternate promoter selection in tumours may be driven less by intrinsic gene function and more by extrinsic selective pressures (e.g. host anti-tumour immunity). Tumour APBscores were also compared to non-malignant tissues. Of 725 matched normal tissue samples analyzed in the PanCanAtlas, only 5% ($n=36$ normal samples) were classified as APB_{high} (at the 80th centile cut-off). When normal and tumour samples were clustered by alternate promoter usage, normal samples clustered close to, but distinct from their corresponding tumour type. In 15 out of the 16 tumour types with at least

10 normal samples, tumour samples were more likely to be classified as APB_{high} compared to normal samples (Fisher's exact, $p < 0.001$).

[00132] **Example 5: Alternate promoter utilization between tumor and normal samples**

5 [00133] To further explore interactions between alternate promoter utilization and tumour immunity, the analyses were expanded to study relationships between APBscore and a broad spectrum of ~700 immune-related genes covering multiple immune cell types, immune checkpoints and antigens. The majority of these genes (78%) exhibited significantly higher expression in APB_{low} tumours compared to APB_{high} tumours (**Fig. 7**), suggesting that that

10 APB_{high} tumours have an immune-depleted phenotype. Differential expression analysis of all ~20,000 genes available in the PanCanAtlas confirmed that the down-regulation of genes in the APB_{high} subgroup was restricted to only the immune genes (**Fig. 8A**). In a Pan-Cancer immune landscape study using TCGA samples, tumours were classified into six immune subtypes characterized by distinct immune signatures - wound healing, IFN- γ dominant, inflammatory, lymphocyte depleted, immunologically quiet, and TGF- β dominant.

15 When APBscore was correlated with these immune subtypes, it was found that lymphocyte depleted or immunologically quiet subtypes were more likely to be APB_{high} tumours (Fisher's exact $p < 0.0001$) (**Fig. 8B**). Notably, lymphocyte depleted APB_{high} tumours comprised a diversity of tumour subtypes, while immunologically quiet APB_{high} tumours were largely dominated by

20 LGGs. In contrast, IFN- γ and inflammatory subtypes were more likely to be APB_{low} tumours (Fisher's exact $p < 0.0001$) (**Fig. 8B**), and APB_{low} tumours were also more likely to have higher lymphocyte infiltration signature scores and IFN- γ response signatures ($p < 0.0001$, Wilcoxon test) (**Fig. 9, Fig. 10**). Importantly, APBscores were uncorrelated with TMB levels ($r = 0.02$, $p = 0.02$), suggesting distinct processes drive alternate promoter utilization and the

25 acquisition of DNA somatic mutations.

[00134] **Example 6: Alternate promoter utilization and survival outcomes**

[00135] The association between APBscores with progression-free survival (PFS) and overall survival (OS) data obtained from the TCGA Pan-Cancer analysis was investigated. To avoid confounding at the pan-cancer level where survival outcomes may be influenced by

30 intrinsic tissue- or site-specific properties, individual tumour-type specific analyses were conducted. When analysed at the tumour type-specific level, the majority of tumour types did not exhibit any correlations between APBscore and PFS or OS. Only a few tumour types

exhibited survival differences between the APB groups (KIRC, LGG, LUAD, THYM). Multivariate analysis of APBscore and disease subtype also suggests that APBscore is not an independent prognostic predictor of survival for patients who have not undergone ICI therapy.

5 [00136] **Example 7: Alternate promoter utilization and pembrolizumab resistance**

[00137] To test if APB_{high} tumours are more resistant to ICIs as previously shown in metastatic gastric cancer, the association of APBscore with objective response rates (ORRs) for pembrolizumab reported in independent studies were quantified. Across all the tumour types, a significant negative correlation between ORR and APBscore (Pearson's $R = -0.46$, $p = 0.025$) was observed. Similar correlations were not observed with *CD8A* ($r = 0.18$, $p=0.4$),
10 *GZMA* ($r = 0.17$, $p=0.4$), or *PRFI* ($r = 0.3$, $p=0.1$). Consistent with previous studies, positive associations were observed with TMB ($r=0.53$, $p=0.0078$) and *PD-L1* expression ($r=0.42$, $r=0.042$). These results thus suggest that APBscore and TMB are likely complementary in identifying distinct subsets of tumours, with TMB selecting types with better ICI response,
15 and APBscore identifying those with lack of response since APBscore and TMB are not correlated. Interestingly, when only tumour types exhibiting significant APBscore correlations with *CD8A*, *GZMA* and *PRFI* were selected for the analysis (19 tumour types), the correlations between APBscore and ORRs improved in strength and significance: Pearson's $R = -0.55$, $p = 0.019$ (**Fig. 11**). Taken collectively, these results extend previous
20 findings shown in gastric cancer and establish a pan-cancer relationship between APBscore and pembrolizumab resistance.

[00138] **Equivalents**

The foregoing examples are presented for the purpose of illustrating the invention and should not be construed as imposing any limitation on the scope of the invention. It will readily be
25 apparent that numerous modifications and alterations may be made to the specific embodiments of the invention described above and illustrated in the examples without departing from the principles underlying the invention. All such modifications and alterations are intended to be embraced by this application.

Claims

1. A method of identifying a patient suffering from cancer as benefiting or not benefiting from immune checkpoint inhibition (ICI) therapy comprising the steps of:
 - a) measuring an expression level of one or more preselected markers in a cancerous biological sample obtained from the patient;
 - b) identifying a differentially expressed alternative promoter based on the expression level of the one or more preselected markers measured in step (a);
 - c) calculating an alternative promoter usage score;
 - d) identifying the patient as benefiting or not benefiting from ICI therapy using the alternative promoter usage score.
2. The method of claim 1, wherein the preselected marker is a nucleic acid molecule or a peptide.
3. The method of claim 1, wherein the preselected marker is a marker that is associated with a somatic promoter region in a cancerous biological sample.
4. The method of any one of claims 1 to 3, wherein the differentially expressed alternative promoter in step (b) is identified by comparing the expression level of the one or more preselected markers in the cancerous biological sample with the expression level of the said one or more preselected markers in one or more reference samples to determine an increase or decrease in expression level of said preselected markers between the cancerous biological sample and the one or more reference samples.
5. The method of claim 4, wherein an increase in the expression level of the one or more preselected markers in the cancerous biological sample compared to the expression level of the one or more preselected markers in the one or more reference samples for a gained promoter indicates a differentially expressed alternative promoter, and wherein a decrease in the expression level of the one or more preselected markers in the cancerous biological sample compared to the expression level of the one or more

preselected markers in the one or more reference samples for a lost promoter indicates a differentially expressed alternative promoter.

- 5 6. The method of claim 5, wherein the gained promoter is a promoter that is gained or increased in a cancerous biological sample compared to a non-cancerous biological sample and wherein the lost promoter is a promoter that is lost or decreased in a cancerous biological sample compared to a non-cancerous biological sample.
- 10 7. The method of any one of claims 4 to 6, wherein the expression level of the one or more preselected markers in the cancerous biological sample is compared with a percentile expression level of said one or more preselected markers in the one or more reference samples, or the mean expression level of said one or more preselected markers in the one or more reference samples, or both.
- 15 8. The method of claim 7, wherein i) an increase in expression level of a preselected marker compared with the median expression level for a gained promoter, or ii) a decrease in expression level of a preselected marker compared with the median expression level for a lost promoter indicates a differentially expressed alternative promoter.
- 20 9. The method of any one of claims 1 to 8, wherein the alternative promoter usage score in step (c) is calculated by determining the sum of the differentially expressed alternative promoters identified in step (b).
- 25 10. The method of any one of claims 1 to 9, wherein the alternative promoter usage score in step (c) is calculated by determining the sum of the number of gained promoters with at least a 4-fold increase in the expression level of transcripts and the number of lost promoters with an expression level lower than 0.25-fold when compared to the median expression level of transcripts.
- 30 11. The method of any one of claims 1 to 9, wherein the alternative promoter usage score in step (c) is calculated by determining the sum of the number of gained promoters

with a relative promoter activity greater than median and the number of lost promoters with relative promoter activity less than median.

- 5 12. The method of any one of claims 1 to 11, further comprising the step of comparing the alternative promoter usage score calculated in step (c) with a reference score to identify the patient as benefiting or not benefiting from ICI therapy, wherein an increased alternative promoter usage score compared to the reference score identifies the patient as not benefiting from ICI therapy.
- 10 13. The method of claim 12, wherein the reference score is the alternative promoter usage score from one or more reference samples.
- 15 14. The method of claim 12 or 13, wherein the increased alternative promoter usage score is a score above the 66th percentile of alternative promoter usage scores from one or more reference samples.
- 20 15. The method of claim 14, wherein the increased alternative promoter usage score is a score above the 80th percentile of alternative promoter usage scores from one or more reference samples.
- 25 16. A method of determining the prognosis of a patient who is undergoing or has undergone ICI therapy, the method comprising the steps of:
- a) measuring an expression level of one or more preselected markers in a cancerous biological sample obtained from the patient;
 - b) identifying a differentially expressed alternative promoter based on the expression level of the one or more preselected markers measured in step (a);
 - c) calculating an alternative promoter usage score;
 - d) determining the prognosis of the patient using the alternative promoter usage score.
- 30 17. The method of claim 16, further comprising the step of comparing the alternative promoter usage score calculated in step (c) with a reference score to determine the

prognosis of the patient, wherein an increased alternative promoter usage score compared to a reference score indicates a poorer prognosis.

5 18. The method of claim 17, wherein the reference score is the alternative promoter usage score from one or more reference samples.

10 19. The method of claim 17 or 18, wherein the increased alternative promoter usage score is a score above the 66th percentile of alternative promoter usage scores from one or more reference samples.

20. The method of claim 19, wherein the increased alternative promoter usage score is a score above the 80th percentile of alternative promoter usage scores from one or more reference samples.

15 21. A biomarker for excluding a patient suffering from cancer from ICI therapy, wherein the biomarker is an alternative promoter usage score that is above the alternative promoter usage score of one or more reference samples, wherein the alternative promoter usage score is the sum of the differentially expressed alternative promoters at each alternative promoter site.

20 22. The biomarker of claim 21, wherein the biomarker is an alternative promoter usage score that is above the 66th percentile of alternative promoter usage scores of one or more reference samples.

25 23. The biomarker of claim 21, wherein the biomarker is an alternative promoter usage score that is above the 80th percentile of alternative promoter usage scores of one or more reference samples.

30