METHOD OF PREDICTING THE RESPONSIVENESS TO A CANCER THERAPY

Field of Invention

The invention relates generally to the field of oncology. In particular, the invention relates to a method of predicting the responsiveness of a cancer subject to a cancer therapy.

Background

Platinum-based chemotherapeutic compounds, such as cisplatin, carboplatin and oxaliplatin, are commonly used in the treatment of cancers that include testicular, ovarian, cervical, lung, liver, colorectal, gastric and head and neck cancer. These compounds bind to DNA and induce DNA adducts, which can lead to apoptosis in cancer cells, thus eliminating cancer cells in the patient. There is, however, a problem with poor responsiveness towards these compounds in some patients. This may be due to intrinsic or acquired resistance to such compounds in some tumours, the mechanism of which is not clearly understood. It is therefore preferable to be able to screen for patients who are likely to be responsive to treatment with these compounds. Alternative treatment strategies may have to be adopted for patients who are unlikely to respond well to such therapies.

Accordingly, it is generally desirable to overcome or ameliorate one or more of the above-mentioned difficulties.

Summary

In one embodiment is a method of predicting the responsiveness of a cancer subject to a cancer therapy, the method comprising determining the level of RNA-editing events in a sample obtained from the subject, wherein an increased level of RNA-editing events as compared to a reference indicates that the subject is likely to be responsive towards the cancer therapy.

Disclosed herein is a method of predicting the responsiveness of a cancer subject to a platinum-based chemotherapy, the method comprising determining the level of adenosine to inosine (A-to-I) RNA-editing events in 29 genes as described in Table 1 in a cancer sample obtained from the subject, wherein an increased level of adenosine to inosine (A-to-I) RNA-editing events in the 29 genes as compared to a reference indicates that the subject is likely to be responsive towards the platinum-based chemotherapy, wherein the subject is suffering from gastric cancer, liver cancer, esophageal cancer or lung cancer.

In one embodiment is a method of treating a cancer subject found likely to be responsive to a cancer therapy, the method comprising: a) determining the level of RNA-editing events in a sample obtained from the subject, wherein an increased level of RNA editing events as compared to a reference indicates that the subject is likely to be responsive towards the cancer therapy; and b) treating the subject found likely to be responsive to the cancer therapy.

In one embodiment is a method of stratifying a subject into a likely responder or non-responder of a cancer therapy, the method comprising determining the level of RNA-editing events in a sample obtained from the subject, wherein an increased level of RNA editing events as compared to a reference indicates that the subject is a likely responder of the cancer therapy and wherein an unchanged or decrease level of RNA-editing events indicates that the subject is a likely non-responder of the cancer therapy.

Disclosed herein is a method of stratifying a subject into a likely responder or non-responder of a platinum-based chemotherapy, the method comprising determining the level of adenosine to inosine (A-to-I) RNA-editing events in 29 genes as described in Table 1 in a cancer sample obtained from the subject, wherein an increased level of adenosine to inosine (A-to-I) RNA editing events in the 29 genes as compared to a reference indicates that the subject is a likely responder of the platinum-based chemotherapy and wherein an unchanged or decrease level of adenosine to inosine (A-to-I) RNA-editing events indicates that the subject is a likely non-responder of the cancer therapy, wherein the subject is suffering from gastric cancer, liver cancer, esophageal cancer or lung cancer.

Disclosed herein is a method for managing treatment of a cancer subject with a cancer therapy, the method comprising a) selecting a cancer subject for treating with the therapy on the basis that the subject is a likely responder to the therapy, or selecting a subject for not treating with the therapy on the basis that the subject is a likely non-responder to the therapy; and b) treating or not treating the subject with the therapy based on the selection, wherein the selection is based on the stratification method as defined herein.

In one embodiment is a method for predicting treatment outcome of a cancer subject with a cancer therapy, the method comprising determining the level of RNA-editing events in a sample obtained from the subject, wherein an increased level of RNA editing events as compared to a reference indicates that the subject is likely to have a positive treatment outcome with the cancer therapy.

Disclosed herein is a method for predicting treatment outcome of a cancer subject with a platinum-based chemotherapy, the method comprising determining the level of adenosine to inosine (A-to-I) RNA-editing events in 29 genes as described in Table 1 in a cancer sample obtained from the subject, wherein an increased level of adenosine to inosine (A-to-I) RNA editing events as compared to a reference indicates that the subject is likely to have a positive treatment outcome with the cancer therapy, wherein the subject is suffering from gastric cancer, liver cancer, esophageal cancer or lung cancer.

In one embodiment is a method of identifying a subject who is suitable for treatment with a cancer therapy, the method comprising determining the level of RNA-editing events in a sample obtained from the subject, wherein an increased level of RNA editing events as compared to a reference indicates that the subject is likely to have a positive treatment outcome with the cancer therapy and therefore suitable for treatment with the cancer therapy.

Disclosed herein is a method of identifying a subject who is suitable for treatment with a platinum-based chemotherapy, the method comprising determining the level of adenosine to inosine (A-to-I) RNA-editing events in 29 genes as described in Table 1 in a cancer sample obtained from the subject, wherein an increased level of adenosine to inosine (A-to-I) RNA editing events in the 29 genes as compared to a reference indicates that the subject is likely to have a positive treatment outcome with the cancer therapy

and therefore suitable for treatment with the cancer therapy, wherein the subject is suffering from gastric cancer, liver cancer, esophageal cancer or lung cancer.

In one embodiment is a kit for predicting the responsiveness of a cancer subject to a cancer therapy, the kit comprising one or more pairs of primers as described in Table 2 for determining the level of RNA-editing events in a sample obtained from the subject.

In one embodiment is an array for predicting the responsiveness of a cancer subject to a cancer therapy, the array comprising one or more nucleic acid probes for determining the level of RNA-editing events in a sample obtained from the subject.

Brief Description of Drawings

Embodiments of the present invention are hereafter described, by way of non-limiting example only, with reference to the accompanying drawings in which:

Figure 1: RNA editing as a prognostic factor in advanced gastric cancer. A) Number of high confidence RNA editing sites detected in the indicated across the patient samples. **B)** Unbiased clustering of patient samples based on RNA editing levels of 780 editing sites detected in all patient samples (k-means, n = 2). **C)** Overall survival plot for high and low editing group of patients (log rank test).

Figure 2: Response to chemotherapy predicts overall survival in advanced GC patients. Overall survival plot for 3 response groups (PD, SD and PR) (log rank test).

Figure 3: Selection of a panel of 50 RNA editing sites as a 29-gene RNA editing (RE) signature for predicting overall response in advanced GC. A) Depiction of the Pearson correlations between editing frequency of individual sites and overall response to chemotherapy. Editing sites which show significant correlation (p <0.05) are displayed with their gene symbols, including 50 positively (boxed by blue dotted line) and 3 negatively correlated sites. As 3 response groups (PD, SD and PR) are defined as 0, 1, and 2 respectively, positive correlation denotes the higher the editing level the higher the response value (i.e. better response). **B**) k-means clustering of patients (n = 55 with available overall response information) based on RNA editing levels of 50 sites. Two clusters which have relatively “higher” and “lower” editing levels across 50 sites

show significant association with “responders (PR)” and “non-responders (SD & PD)”, respectively. **C**) Disease enrichment analysis results of RE signature genes.

Figure 4: Derivation of an RNA Editing (RE) score to accurately predict chemotherapy response in advanced GC. A) Receiver operating characteristic (ROC) curve showing the performance of RE score in prediction of responders. TPR = true positive rate, FPR = false positive rate, AUC = area under curve. **B)** The distribution of derived RE scores for each response group is given as violin plot. The dots represent the individual score for each patient. A cut-off value 0.4 is shown in dashed line which classifies the responders and non-responders with 84% accuracy. An analysis of the specificity, sensitivity, accuracy and precision of the assay is also provided at this cut-off value.

Figure 5. Validation of RE signature in GC cell lines. A) Chemosensitivity of 10 GC cell lines to oxaliplatin. The bars show the average IC_{50} value of 3 biological replicates with standard error of means. **B)** Cell viability assay of 6 GC cell lines cultured with indicated concentration of oxaliplatin for 48 hours. Cells are stained with crystal violet. **C)** Editing levels of 26 sites from the panel of RE signature quantified by Sanger sequencing. Numbers in cells denote the percentage of editing levels and coloring shows the relative gradient across the row (scale). Black cells denote undetectable editing level. The Pearson correlation coefficient between the average IC_{50} values in **A)** and average RNA editing levels in **C)** is r = -0.58.

Figure 6. Validation of RE signature in TCGA STAD cohort. Distribution (violin plots) and performance (ROC curves) of RE scores in TCGA STAD cohort stratified by response type in Stage IV patients. For score calculation, 50 sites in RE signature and their editing levels in TCGA patients are used wherever available and at least 8 sites were required to be edited to calculate the score. Primary response A) refers to the “primary_therapy_outcome” and follow-up response B) denotes the “follow-up_treatment_success” as reported by TCGA. Response groups are merged based on the consensus of the individual treatment outcomes: Non-responder = stable/progressive disease, Responder = partial/complete remission, STAD = stomach adenocarcinoma, TPR = true positive rate, FPR = false positive rate, AUC = area under the curve. p-values are shown for Wilcoxon rank-sum test.

Figure 7. Cell viability assay of GC cell lines. Relative cell viability of 10 GC cell lines treated with the indicated concentrations of oxaliplatin for 72h. Data are shown as mean ±SD, n = 3. Error bars are not drawn for those points where they are shorter than the symbol.

Detailed Description

Disclosed herein is a method of predicting the responsiveness of a cancer subject to a cancer therapy, the method comprising determining the level of RNA-editing events in a sample obtained from the subject, wherein an increased level of RNA-editing events as compared to a reference indicates that the subject is likely to be responsive towards the cancer therapy.

Without being bound by theory, RNA editing, which is an epigenetic mechanism, introduces changes in the RNA sequences encoded by the genome, contributing to "RNA mutations". In humans, the most frequent type of RNA editing is the conversion of adenosine to inosine (A-to-I). Adenosine deaminases acting on RNA (ADAR) are enzymes responsible for binding to double stranded RNA (dsRNA) and converting A to I by deamination. An unbiased transcriptome-wide A-to-I RNA editing analysis was conducted on 104 patients who have advanced GC and enrolled onto first- line palliative chemotherapy by RNA Sequencing (RNA-Seq). It was found that in these patients, higher editing frequency of a panel of 50 editing sites (across 29 gene transcripts) predicted better response to platinum-based chemotherapy. Based on the 29-gene RNA editing signature, an RNA editing score (RE score) was derived. By employing this RE score, it was found that 80% of patients with high RE score achieved partial response (PR) ("responders"), while 85% of patients with low RE score obtained stable disease (SD) or progressive disease (PD) ("non-responders"). This 29-gene RNA editing signature can therefore be used to stratify patients with advanced gastric and predict benefit from chemotherapy. In particular, it can be used to predict clinical benefit of platinum-based chemotherapy in patients with advanced gastric cancer with up to 84% accuracy.

The cancer therapy may refer to chemotherapy, radiotherapy and immunotherapy. In one embodiment, cancer therapy is chemotherapy. Useful classes of anti-cancer agents include chemotherapeutic agents, representative examples of which include antitubulin

agents, auristatins, DNA minor groove binders, DNA replication inhibitors, alkylating agents (e.g., platinum complexes such as cis-platin, mono(platinum), bis(platinum) and tri-nuclear platinum complexes and carboplatin), anthracyclines, antibiotics, antifolates, antimetabolites, calmodulin inhibitors, chemotherapy sensitizers, duocarmycins, etoposides, fluorinated pyrimidines, ionophores, lexitropsins, maytansinoids, nitrosoureas, platinols, pore-forming compounds, purine antimetabolites, puromycins, radiation sensitizers, rapamycins, steroids, taxanes, topoisomerase inhibitors, vinca alkaloids, or the like.

In one embodiment, the cancer therapy is a platinum-based chemotherapy. The cancer therapy may, for example, involve treatment with cisplatin, oxaliplatin, and carboplatin. In one embodiment, the platinum-based chemotherapy is a (platinum-fluoropyrimidine doublet chemotherapy.

The term “responsiveness” to therapy may refer to any one or more of: extending survival (including overall survival and progression free survival); resulting in an objective response (including a complete response or a partial response); or improving signs or symptoms of cancer. In some embodiments, responsiveness may refer to improvement of one or more factors according to the published set of RECIST guidelines for determining the status of a tumor in a cancer patient, i.e., responding, stabilizing, or progressing. A responsive subject may refer to a subject whose cancer(s) show improvement, e.g., according to one or more factors based on RECIST criteria. A non-responsive subject may refer to a subject whose cancer(s) do not show improvement, e.g., according to one or more factors based on RECIST criteria.

In one embodiment, a responsive subject refers to a subject who shows a complete response (CR) or a partial response (PR), while a non-responsive subject is one who does not show improvement and has a stable disease (SD) or progressive disease (PD).

The term “predicting the responsiveness” may also refer to determining the likelihood of a subject who is responsive to a cancer therapy. In one embodiment, there is provided a method of determining the likelihood of a cancer subject who is responsive to a cancer therapy, the method comprising determining the level of RNA-editing events in a sample obtained from the subject, wherein an increased level of RNA-editing events as

compared to a reference indicates that the subject is likely to be responsive towards the cancer therapy.

The terms “cancer” and “cancerous” refer to or describe the physiological condition in mammals that is typically characterized in part by unregulated cell growth. As used herein, the term “cancer” refers to non-metastatic and metastatic cancers, including early stage and late stage cancers. The term “precancerous” refers to a condition or a growth that typically precedes or develops into a cancer. By “non-metastatic” is meant a cancer that is benign or that remains at the primary site and has not penetrated into the lymphatic or blood vessel system or to tissues other than the primary site. Generally, a non-metastatic cancer is any cancer that is a Stage 0, I, or II cancer, and occasionally a Stage III cancer. By “early stage cancer” is meant a cancer that is not invasive or metastatic or is classified as a Stage 0, I, or II cancer. The term “late stage cancer” generally refers to a Stage III or Stage IV cancer but can also refer to a Stage II cancer or a sub-stage of a Stage II cancer. One skilled in the art will appreciate that the classification of a Stage II cancer as either an early stage cancer or a late stage cancer depends on the particular type of cancer.

The term “cancer” includes but is not limited to, breast cancer, large intestinal cancer, lung cancer, small cell lung cancer, gastric (stomach) cancer, liver cancer, blood cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer, cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer, rectal or colorectal cancer, anal cancer, colon cancer, fallopian tube carcinoma, endometrial carcinoma, cervical cancer, vulval cancer, squamous cell carcinoma, vaginal carcinoma, Hodgkin's disease, non-Hodgkin's lymphoma, esophageal cancer, small intestine cancer, endocrine cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue tumor, urethral cancer, penile cancer, prostate cancer, chronic or acute leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumor, glioma, astrocytoma, glioblastoma multiforme, primary CNS lymphoma, bone marrow tumor, brain stem nerve gliomas, pituitary adenoma, uveal melanoma (also known as intraocular melanoma), testicular cancer, oral cancer, pharyngeal cancer or a combination thereof.

In one embodiment, the cancer is selected from the group consisting of gastric cancer, liver cancer, esophageal cancer and lung cancer.

In one embodiment, the cancer is an advanced cancer (such as a stage IV cancer). The cancer may, for example, be an advanced gastric cancer, an advanced liver cancer, an advanced esophageal cancer or an advanced lung cancer.

In one embodiment, the RNA-editing events are adenosine to inosine (A-to-I) RNA-editing events.

In one embodiment, the level of RNA-editing events is determined by RNA sequencing. Alternatively, the level of RNA-editing events may be determined by PCR amplification followed by Sanger sequencing analysis of the purified PCR products or by other techniques that are well known in the art.

The term "sample" may refer to any sample derived from or containing cells, organisms (bacteria, viruses), lysed cells or organisms, cellular extracts, nuclear extracts, components of cells or organisms, extracellular fluid, media in which cells or organisms are cultured in vitro, blood, plasma, serum, gastrointestinal secretions, urine, ascites, homogenates of tissues or tumors, synovial fluid, feces, saliva, sputum, cyst fluid, amniotic fluid, cerebrospinal fluid, peritoneal fluid, lung lavage fluid, semen, lymphatic fluid, tears, pleural fluid, nipple aspirates, breast milk, external sections of the skin, respiratory, intestinal, and genitourinary tracts, and prostatic fluid.

In one embodiment, the sample is a cancer sample. The method as defined herein may comprise analyzing RNA editing frequencies of editing sites in endoscopic biopsies of the primary tumor prior to chemotherapy initiation, in order to evaluate the RNA editing profile for each patient.

In one embodiment, the method comprises determining the level of RNA-editing events in one or more genes as described in Table 1. The method may comprise determining the level of RNA-editing events in 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28 or 29 genes listed in Table 1. In one embodiment, the method comprises determining the level of RNA-editing events in all genes as described in Table 1.

In one embodiment, the method comprises determining the level of RNA-editing events in one or more genes selected from the group consisting of CCNYL1, PHACTR4, LIMD1, RRP36, MRPL30, CPT1A, RBM39, ORC2, AHR, METTL7A, PAICS, NOP14, SLC35F5, SRFBP1, PGAM5, MAVS, LONP2, EIF2AK2, GATC, ZNF587, CTSS, ARSD, FGD5-AS1, SNRPD3, MED18, NEAT1, NUP43, NDUFS1 and TMPO.

In one embodiment, the method comprises determining the level of RNA-editing events in CCNYL1, PHACTR4, LIMD1, RRP36, MRPL30, CPT1A, RBM39, ORC2, AHR, METTL7A, PAICS, NOP14, SLC35F5, SRFBP1, PGAM5, MAVS, LONP2, EIF2AK2, GATC, ZNF587, CTSS, ARSD, FGD5-AS1, SNRPD3, MED18, NEAT1, NUP43, NDUFS1 and TMPO.

In one embodiment, the method comprises determining the level of RNA-editing events in one or more RNA-editing sites as described in Table 1. The method may comprise determining the level of RNA-editing events in 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 or 50 RNA-editing sites listed in Table 1. In one embodiment, the method comprises determining the level of RNA-editing events in all 50 RNA-editing sites described in Table 1.

In one embodiment, the method comprises determining a score (e.g. a Z score) based on the level of RNA-editing events in the one or more RNA-editing sites.

For each RNA-editing site, the score (e.g. Z score or z-transformation) may be calculated based on the formula:

$$z = \frac{x - \mu}{\sigma},$$

where x is the RNA-editing level of each RNA-editing site, μ is the mean RNA-editing level of all the RNA-editing sites and σ is the standard deviation of the editing level of all RNA-editing sites. The individual z-scores for the RNA-editing sites may be averaged to obtain an overall Z score for each patient.

The patient with a score that is above or below a cut-off value may be regarded as high or low editing patient, respectively. The cut-off value may be adjusted in each cohort.

The score may be compared to a reference with a cut-off value. The cut-off value may be cancer or cohort specific. The cut-off value may be between -1 to 1. The cut-off value may, for example, be -1, -0.95, -0.9, -0.85, -0.8, -0.75, -0.7, -0.65, -0.6, -0.55, -0.5, -0.45, -0.4, 0.35, -0.3, -0.25, -0.2, -0.15, -0.1, -0.05, 0, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95 or 1.0.

In one embodiment, a score that is above the cut-off value indicates that the subject is likely to have a complete response (CR) or partial response (PR), while a score of less or equal than the cut-off value indicates that the subject is likely to have a stable disease (SD) or progressive disease (PD).

In one embodiment, the score has a cut-off of 0.4, wherein a score of >0.4. indicates that the subject is likely to have a complete response (CR) or partial response (PR), while a score of ≤0.4 indicates that the subject us likely to have a stable disease (SD) or progressive disease (PD).

In one embodiment, the score has a cut-off of 0.4, wherein a score of >0.4. indicates that the subject is likely to have a partial response (PR), while a score of ≤0.4 indicates that the subject us likely to have a stable disease (SD) or progressive disease (PD).

In one embodiment, the method comprises determining the level of RNA-editing events with one or more pairs of primers as described in Table 2.

In one embodiment, the method comprises treating the cancer subject with the cancer therapy.

In one embodiment, the method as defined herein also predict the likelihood of resistance to a cancer therapy. The method may comprise determining the level of RNA-editing events in a sample obtained from the subject, wherein an increased level of RNA-editing events as compared to a reference indicates that the subject is less likely to be resistant to the cancer therapy.

Disclosed herein is a method of treating a cancer subject found likely to be responsive to a cancer therapy, the method comprising: a) determining the level of RNA-editing events in a sample obtained from the subject, wherein an increased level of RNA editing

events as compared to a reference indicates that the subject is likely to be responsive towards the cancer therapy; an b) treating the subject found likely to be responsive to the cancer therapy.

The term “treating” as used herein may refer to (1) preventing or delaying the appearance of one or more symptoms of the disorder; (2) inhibiting the development of the disorder or one or more symptoms of the disorder; (3) relieving the disorder, i.e., causing regression of the disorder or at least one or more symptoms of the disorder; and/or (4) causing a decrease in the severity of one or more symptoms of the disorder.

Treatment of a subject may comprise administering to the subject a cancer therapy such as a platinum-based cancer therapy. The term “administering” may refer to contacting, applying, injecting, transfusing or providing an effective amount of a cancer therapy to a subject.

The term “effective amount” as defined herein is meant the administration of an amount of agent to an individual in need thereof, either in a single dose or as part of a series, that is effective for that elicitation, treatment or prevention. The effective amount will vary depending upon the health and physical condition of the individual to be treated, the taxonomic group of individual to be treated, the formulation of the composition, the assessment of the medical situation, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials.

The term “subject” as used throughout the specification is to be understood to mean a human or may be a domestic or companion animal. While it is particularly contemplated that the methods of the invention are for treatment of humans, they are also applicable to veterinary treatments, including treatment of companion animals such as dogs and cats, and domestic animals such as horses, cattle and sheep, or zoo animals such as primates, felids, canids, bovids, and ungulates. The “subject” may include a person, a patient or individual, and may be of any age or gender.

Disclosed herein is a method of stratifying a subject into a likely responder or non-responder of a cancer therapy, the method comprising determining the level of RNA-editing events in a sample obtained from the subject, wherein an increased level of RNA editing events as compared to a reference indicates that the subject is a likely responder

of the cancer therapy and wherein an unchanged or decrease level of RNA-editing events indicates that the subject is a likely non-responder of the cancer therapy.

Disclosed herein is a method for managing treatment of a cancer subject with a cancer therapy, the method comprising a) selecting a cancer subject for treating with the therapy on the basis that the subject is a likely responder to the therapy, or selecting a subject for not treating with the therapy on the basis that the subject is a likely non-responder to the therapy; and b) treating or not treating the subject with the therapy based on the selection, wherein the selection is based on the stratification method as defined herein.

Disclosed herein is a method for predicting treatment outcome of a cancer subject with a cancer therapy, the method comprising determining the level of RNA-editing events in a sample obtained from the subject, wherein an increased level of RNA editing events as compared to a reference indicates that the subject is likely to have a positive treatment outcome with the cancer therapy.

Disclosed herein is a method of identifying a subject who is suitable for treatment with a cancer therapy, the method comprising determining the level of RNA-editing events in a sample obtained from the subject, wherein an increased level of RNA editing events as compared to a reference indicates that the subject is likely to have a positive treatment outcome with the cancer therapy and therefore suitable for treatment with the cancer therapy.

Disclosed herein is a kit for predicting the responsiveness of a cancer subject to a cancer therapy, the kit comprising one or more pairs of primers as described in Table 2 for determining the level of RNA-editing events in a sample obtained from the subject.

Disclosed herein is an array for predicting the responsiveness of a cancer subject to a cancer therapy, the array comprising one or more nucleic acid probes for determining the level of RNA-editing events in a sample obtained from the subject. The one or more nucleic acid probes may be one that is selected from Table 2. The one or more nucleic acid probes may be optionally conjugated to a detectable label.

Disclosed herein is a composition comprising one or more nucleic acid probes for determining the level of RNA-editing events in a sample obtained from a subject. The one or more nucleic acid probes may be one that is selected from Table 2. The one or more nucleic acid probes may be optionally conjugated to a detectable label. The composition may comprise a cancer sample that is obtained from a subject.

As used herein, "and/or" refers to and encompasses any and all possible combinations of one or more of the associated listed items, as well as the lack of combinations when interpreted in the alternative (or).

As used in this application, the singular form "a," "an," and "the" include plural references unless the context clearly dictates otherwise. For example, the term "an agent" includes a plurality of agents, including mixtures thereof.

Throughout this specification and the statements which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

The reference in this specification to any prior publication (or information derived from it), or to any matter which is known, is not, and should not be taken as an acknowledgment or admission or any form of suggestion that that prior publication (or information derived from it) or known matter forms part of the common general knowledge in the field of endeavour to which this specification relates.

Those skilled in the art will appreciate that the invention described herein in susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications which fall within the spirit and scope. The invention also includes all of the steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations of any two or more of said steps or features.

Certain embodiments of the invention will now be described with reference to the following examples which are intended for the purpose of illustration only and are not intended to limit the scope of the generality hereinbefore described.

EXAMPLES

Methods

Gastric cancer cohort

A total of 104 patients from the 3G (**G**enomic **G**uided therapy for **G**astric cancer) Trial were involved in this study. The patients were diagnosed with metastatic or recurrent GC and enrolled onto the first line palliative chemotherapy (platinum-fluoropyrimidine doublet chemotherapy regime). For each eligible patient, fresh endoscopic biopsies of the primary tumor *in situ* within 3 weeks prior to treatment initiation were used for RNA-Seq analysis. Treatment response to chemotherapy (PD: progressive disease; SD: stable disease; PR: partial response) was evaluated by radiologists who were blinded to the study, and assessed as the best overall response as per RECIST (version 1.1) criteria by CT scans or MRI done at 6-week intervals. At the cut-off date, in this cohort of 104 patients, overall survival and tumor response data had been obtained for 54 and 55 patients, respectively, where a total of 50 patients had both overall survival and tumor response data.

RNA sequencing

A total of 1 μg of total RNA was used to create libraries with Illumina TruSeq Stranded Total RNA Library Prep Kit (Illumina) according to manufacturer's instructions. Library fragment size was determined using the DNA 1000 Kit on the Agilent Bioanalyzer (Agilent Technologies). Libraries were quantified by qPCR using the KAPA Library Quantification Kit (KAPA Biosystems). Libraries were pooled in equimolar and cluster generation was performed on the Illumina cBOT system (Illumina). Sequencing (150bp pair-end) was performed on the Illumina HiSeq 3000 system at the Duke-NUS Genome Biology Facility, according to manufacturer's protocol (Illumina).

Identification of RNA editing events

A bioinformatics pipeline adapted from a previously published method was used to identify RNA editing events from RNA-Seq data by using the CSI NGS Portal (https://csibioinfo.nus.edu.sg/csingsportal). For each sample, raw reads with adapters were trimmed by using Trimmomatic (v0.38) retaining the reads with >= 35 bases and average read quality score > 20 after trimming. Clean reads were mapped to the reference human genome (*hg19*) with a splicing junction database generated from transcript annotations derived from UCSC, RefSeq, Ensembl and GENCODE (v19) by using Burrows–Wheeler Aligner with default parameters (*bwa mem*, v0.7.17-r1188). To retain high quality data, PCR duplicates were removed (*samtools markdup -r*, v1.9) and the reads with mapping quality score < 20 were discarded. Junction-mapped reads were then converted back to the genomic-based coordinates. An in-house perl script was utilized to call the variants from samtools pileup data and the sites with at least two supporting reads were initially retained. The candidate events were filtered by removing the single nucleotide polymorphisms reported in different cohorts (1000 Genomes Project), NHLBI GO Exome Sequencing Project (https://evs.gs.washington.edu/EVS/), dbSNP (v150) and excluding the sites within the first six bases of the reads caused by imperfect priming of random hexamer during cDNA synthesis. For the sites not located in Alu elements, i.e. short repetitive DNA sequences abundant in the human genome, the candidates within the four bases of a splice junction on the intronic side, and those residing in the homopolymeric regions and in the simple repeats were all removed. Candidate variants located in the reads that map to the non-unique regions of the genome by using BLAST-like alignment tool were also excluded. At last, only A-to-G editing sites based on the strand information from the strand-specific RNA-Seq data were considered for all the downstream analyses. The genomic regions of the editing variants and the associated genes were annotated by using ANNOVAR (v2018) with the UCSC *refGene* table annotation. The same pipeline was applied on the TCGA STAD (The Cancer Genome Atlas - stomach adenocarcinoma) cohort.

To identify high confidence and common editing events, stringent filtering criteria were applied. Specifically, each editing site was required to have a coverage of at least 20 reads and editing frequency higher than 0.1 (10%) in all the samples. This resulted in 780 high confidence editing sites shared by 104 samples of our GC cohort. For TCGA STAD cohort, these thresholds were not applied for the validation of RE signature in order to include more sites, as the number of high-confidence editing sites were relatively fewer in TCGA due to the lower sequencing depth. In TCGA, the inventors

included only those samples with at least 8 out of 50 sites in the RE signature were found to be edited. Then for each sample all these edited sites were used to calculate the RE score.

Clustering analyses and heatmaps

The clustering and heatmap analyses were performed by using R package "superheat". The RNA editing levels of the corresponding sites were used as the distance matrix to perform the k-means clustering for the heatmaps in an unsupervised manner (scale = TRUE, n.clusters.rows = 2, clustering.method = "kmeans"). The side plots were drawn using "scattersmooth" parameter with the "lm" method based on the average editing value of the corresponding rows or columns calculated by "rowMeans2" and "colMeans2" functions from the "matrixStats" package, respectively.

Correlation of RNA editing with response to chemotherapy

Pearson correlation analysis was performed by using "cor.test" function in R with "pearson" method. For each of 780 editing sites, Pearson correlation coefficient (r) and associated p-value were calculated between RNA editing levels and the overall response to chemotherapy across 55 patients with available data. The editing frequencies calculated from the RNA-Seq data were used as the first vector. As the second vector, the overall response data was used after transforming the original categorical variables into numerical variables (PD = 0, SD = 1, PR = 2) so that the correlation analysis can be performed. P-values were used to assess the editing sites with significant correlations at $p < 0.05$ threshold, which resulted in 50 positively and 3 negatively correlated sites.

RE score calculation and RE signature

To predict the chemotherapy outcome based on RNA editing, a RE score was developed based on the "z-score" using 50 sites that showed significant positive correlation with the overall response. First, z-transformation was performed for each site based on the RNA editing levels. Then the samples were ranked by using the average z-score across the sites. The samples above and below a cut-off value (0.4) were regarded as the high and low editing groups, respectively. The statistical measures were calculated based on the prediction of the responders in the cohort. Disease enrichment analysis of RE signature genes were performed by using Enrichr web server with "PheWeb 2019" database.

Gastric cancer cell lines and validation of the RE signature

A total of 10 cell lines were used for the validation of the RE signature. AGS, SNU5, SNU16 and NCI-N87 were purchased from the American Type Culture Collection (Manassas, VA) between 2019-2020. MKN1, MKN7, MKN28, MKN45 and MKN74 were obtained from Japanese Collection of Research Bioresources Cell Bank between 2019-2020. YCC11 cell line was provided by Singapore Gastric Cancer Consortium (SGCC) in 2014. All the cell lines were tested negative for mycoplasma. SNU5 cells were cultured with IMDM (Gibco BRL, Grand Island, NY, USA) supplemented with 20% FBS (Gibco BRL), and all the other cell lines were cultured in Roswell Park Memorial Institute (RPMI) medium (Gibco BRL) supplemented with 10% fetal bovine serum (FBS) (Gibco BRL). Cells were grown in a humidified incubator with 5% CO_2 at 37°C.

The editing levels of 26 out of 50 editing sites in these 10 cell lines were quantified by Sanger sequencing. The sites were selected as the top 10 sites which showed the highest change in editing level between high and low editing group as defined in the RNA-Seq data, and additional 16 sites that are randomly picked from the remaining sites in the panel of RE signature. Drug response of the cell lines to oxaliplatin (Sigma-Aldrich) were assessed by IC_{50} values using MTT (3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) assay. Briefly, GC cell lines were seeded in 96-well plates with 2.5×10^3 to 10×10^3 cells/well according to their growth rate. After 72-h oxaliplatin drug treatment, 10 μL MTT substrate (Sigma-Aldrich) was added into each well followed by 3-hour incubation. MTT substrate was then removed and cells were lysed by addition of 100 μL MTT stop solution. Absorbance at 570 nm was measured using Tecan microplate reader.

For the cell viability assay, 6 suitable cell lines were used, where the cells were seeded in 6-well plate with 3×10^4 to 15×10^4 cells/well and cultured with indicated concentration of oxaliplatin for 48 hours. Cells were stained with crystal violet (Sigma-Aldrich) for colony visualization.

Statistical analyses

As a rule of thumb, all the available features and the samples were included in the respective analyses as long as the data being investigated were non-missing. Initial clustering analysis was performed in an unbiased manner based on the editing levels of

all the sites identified in all the samples. The downstream analyses were performed by using a panel of 50 sites that are selected solely based on the correlation value of their editing levels obtained from RNA-Seq data and the overall response to chemotherapy assessed by clinicians and radiologists. RE score was derived based on the editing levels of the selected genes in all the samples with response data available. No manual selection or exclusion was applied on the genes, editing sites or samples for this study, unless limited by the availability of data under investigation, e.g. 55 out of 104 samples had response data. These criteria were also applied to the cell lines in the same way as per patient samples. For the experimental validations, unless otherwise indicated, the data are presented as the mean ± standard error of mean (SEM) of three independent experiments. Wilcoxon rank-sum test was applied when comparing distributions, Pearson correlation coefficient was reported for paired correlation analyses, and log-rank test p-value was shown for the survival analyses. A p-value of less than 0.05 was considered to be statistically significant, and the type of the statistical test applied was indicated appropriately.

Ethics approval and consent to participate

Samples were collected from the cancer centres in Singapore (n = 37) and South Korea (n = 67) between 2010 - 2018. A written informed consent was obtained from all the patients prior to the enrolment to the trial, and the study was done in accordance with the Declaration of Helsinki and International Conference on Harmonisation and Good Clinical Practice guidelines. The protocol was approved by the Institutional Review Board at each study site and complied with local laws and regulations.

Availability of data and materials

RNA-Editing pipeline is available online at the CSI NGS Portal (https://csibioinfo.nus.edu.sg/csingsportal).

Example 1

Higher editing frequency of a panel of 50 editing sites in tumors predicts better response to palliative platinum-based chemotherapy in patients with advanced gastric cancer

An unbiased transcriptome-wide A-to-I RNA editing analysis was conducted using RNA-Seq data of 104 gastric cancer (GC) patients who have advanced, metastatic or recurrent GC and enrolled onto first-line palliative chemotherapy (platinum-fluoropyrimidine doublet chemotherapy regime). For each eligible patient, fresh endoscopic biopsies of the primary tumor in situ within 3 weeks prior to treatment initiation were used for RNA-Seq analysis. Treatment response to platinum-based chemotherapy was assessed by blinded radiologists and outcomes were locally reviewed. At the cut-off date, in this cohort of 104 patients, overall survival (OS) and tumor response data had been obtained for 54 and 55 patients, respectively. A total of 50 patients have both OS and tumor response data.

Based on the RNA editing analysis, patients could be stratified into two groups based on the global RNA editing level (calculated based on 780 editing sites detected in all patient samples, **Fig. 1A**), and patients demonstrating the higher editing level had better survival (**Fig. 1B, C**).

Due to a significant correlation between tumor response to chemotherapy and OS time among 50 patients (**Fig. 2**), it was next investigated if RNA editing frequency in tumors can predict patient's overall response to chemotherapy. It was found that higher editing frequencies of a panel of 50 editing sites (across 29 gene transcripts) were significantly correlated with better response (**Fig. 3A**). This 29-gene RNA editing signature was next used to predict the likelihood of response to chemotherapy. Strikingly, 60% of patients in "high-editing" group had achieved partial response (PR) ("responders"), while only 22.5% of patients in "low-editing" group were responders (**Fig. 3B**).

Based on the 29-gene editing signature, an RNA editing score (**RE score**) was derived. The performance of RE score was tested by a ROC curve (AUC = 0.77) (**Fig. 4A**). When a cut-off score of 0.4 was set, patients were predicted to be responders (PR) with 84% accuracy, (**Fig. 4B**), suggesting that employing a RE score based on the 29-gene RNA editing signature could accurately stratify responders and non-responders to chemotherapy.

The RNA editing signature was next validated in 10 commercially available GC cell lines. First, the drug response of each cell line to oxaliplatin was assessed by IC_{50} (the half maximal inhibitory concentration) values (Figure 5A, Figure 7) and further

confirmed by cell viability assay (Figure 5B). Then, the RNA editing levels of 26 randomly selected sites from 50 sites of the signature was quantified in the same 6 cell lines by Sanger sequencing (Figure 5C). A negative correlation between the average IC_{50} values and average RNA editing levels ($r = -0.58$) was found, implying that GC cells with higher editing levels of the RNA editing signature sites demonstrate higher chemosensitivity, which was consistent with our observation in the patient samples.

To infer whether our RE signature is representative of different chemotherapy regimens, the RE score was applied in a subset of patients with stage IV gastric cancer in the TCGA STAD (stomach adenocarcinoma) cohort. These patients were treated with multiple combinations of drugs, along with multiple data points reporting for primary and follow-up treatment outcome. When the RE score was applied on these patients, it was observed that responders had significantly higher levels of editing in the panel of 50 sites compared to non-responders (**Figure 6A-B**), confirming the observation in the earlier advanced GC cohort, despite the diverse drug regimens between the two cohorts. These results suggest that the RE signature is coherent as a predictive marker which is independent of different chemotherapy regimens.

Table 1

chromosome	position	ref	mut	gene	gene strand	gene type	Func_refGene	GeneDetail_refGene
chr2	208619479	A	G	CCNYL1	+	protein_coding	UTR3	NM_001142300:c.*1057A>G;NM_152523:c.*1057A>G;NM_001330218:c.*1057A>G
chr1	28825419	A	G	PHACTR4	+	protein_coding	UTR3	NM_001350159:c.*2358A>G;NM_001350158:c.*2358A>G;NM_001350160:c.*2358A>G;NM_001350161:c.*2358A>G;NM_001048183:c.*2358A>G;NM_023923:c.*2358A>G
chr3	45722300	A	G	LIMD1	+	protein_coding	ncRNA_intronic	.
chr6	42997578	A	G	RRP36	+	protein_coding	downstream	dist=241
chr2	99812857	A	G	MRPL30	+	protein_coding	UTR3	NM_145212:c.*689A>G
chr11	68523472	T	C	CPT1A	-	protein_coding	UTR3	NM_001876:c.*1640A>G
chr20	34301191	T	C	RBM39	-	protein_coding	intronic	.
chr3	45721917	A	G	LIMD1	+	protein_coding	ncRNA_intronic	.
chr2	201774264	T	C	ORC2	-	protein_coding	downstream	dist=630
chr3	45721934	A	G	LIMD1	+	protein_coding	ncRNA_intronic	.
chr7	17384778	A	G	AHR	+	protein_coding	UTR3	NM_001621:c.*2090A>G
chr12	51324496	A	G	METTL7A	+	protein_coding	UTR3	NM_014033:c.*563A>G
chr3	45722381	A	G	LIMD1	+	protein_coding	ncRNA_intronic	.
chr4	57327082	A	G	PAICS	+	protein_coding	UTR3	NM_006452:c.*1378A>G;NM_001079524:c.*1378A>G;NM_001079525:c.*1378A>G
chr6	42997538	A	G	RRP36	+	protein_coding	downstream	dist=201
chr12	51324467	A	G	METTL7A	+	protein_coding	UTR3	NM_014033:c.*534A>G
chr4	2940164	T	C	NOP14	-	protein_coding	ncRNA_intronic	.
chr2	114466200	T	C	SLC35F5	-	protein_coding	UTR3	NM_001330314:c.*6595A>G;NM_001330315:c.*6595A>G;NM_001330316:c.*6521A>G;NM_025181:c.*6595A>G
chr5	121362874	A	G	SRFBP1	+	protein_coding	UTR3	NM_152546:c.*53A>G
chr12	133297904	A	G	PGAM5	+	protein_coding	UTR3	NM_001170544:c.*490A>G;NM_001170543:c.*490A>G
chr7	17384832	A	G	AHR	+	protein_coding	UTR3	NM_001621:c.*2144A>G
chr20	3852829	A	G	MAVS	+	protein_coding	UTR3	NM_001206491:c.*6035A>G;NM_020746:c.*6035A>G
chr16	48388245	A	G	LONP2	+	protein_coding	ncRNA_intronic	.
chr2	37329283	T	C	EIF2AK2	-	protein_coding	intergenic	dist=2896;dist=3001
chr12	120899075	A	G	GATC	+	protein_coding	UTR3	NM_176818:c.*1313A>G
chr19	58372048	A	G	ZNF587	+	protein_coding	UTR3	NM_032828:c.*540A>G;NM_001204817:c.*540A>G
chr12	133298339	A	G	PGAM5	+	protein_coding	UTR3	NM_001170544:c.*925A>G;NM_001170543:c.*925A>G
chr2	114465355	T	C	SLC35F5	-	protein_coding	UTR3	NM_001330314:c.*7440A>G;NM_001330315:c.*7440A>G;NM_001330316:c.*7366A>G;NM_025181:c.*7440A>G
chr1	150703692	T	C	CTSS	-	protein_coding	UTR3	NM_001199739:c.*1830A>G;NM_004079:c.*1830A>G
chrX	2824434	T	C	ARSD	-	protein_coding	UTR3	NM_001669:c.*878A>G
chr2	99812978	A	G	MRPL30	+	protein_coding	UTR3	NM_145212:c.*810A>G
chr7	17384887	A	G	AHR	+	protein_coding	UTR3	NM_001621:c.*2199A>G
chr3	14986098	T	C	FGD5-AS1	-	antisense_RNA	ncRNA_exonic	.
chr22	24969092	A	G	SNRPD3	+	protein_coding	UTR3	NM_001278656:c.*1147A>G;NM_004175:c.*1147A>G
chr19	58375195	A	G	ZNF587	+	protein_coding	UTR3	NM_032828:c.*3687A>G;NM_001204817:c.*3687A>G
chr19	58372768	A	G	ZNF587	+	protein_coding	UTR3	NM_032828:c.*1260A>G;NM_001204817:c.*1260A>G
chr1	28661576	A	G	MED18	+	protein_coding	UTR3	NM_017638:c.*95A>G;NM_001127350:c.*95A>G
chr7	17384527	A	G	AHR	+	protein_coding	UTR3	NM_001621:c.*1839A>G
chr11	65208818	A	G	NEAT1	+	lincRNA	ncRNA_exonic	.
chr2	208619749	A	G	CCNYL1	+	protein_coding	UTR3	NM_001142300:c.*1327A>G;NM_152523:c.*1327A>G;NM_001330218:c.*1327A>G
chr7	17384792	A	G	AHR	+	protein_coding	UTR3	NM_001621:c.*2104A>G
chr6	150045858	T	C	NUP43	-	protein_coding	UTR3	NM_198887:c.*2247A>G
chr4	2940462	T	C	NOP14	-	protein_coding	ncRNA_exonic	.
chr2	206987652	T	C	NDUFS1	-	protein_coding	downstream	dist=151
chr6	150046432	T	C	NUP43	-	protein_coding	UTR3	NM_198887:c.*1673A>G
chrX	2824979	T	C	ARSD	-	protein_coding	UTR3	NM_001669:c.*333A>G
chr3	45722287	A	G	LIMD1	+	protein_coding	ncRNA_intronic	.
chr1	28825370	A	G	PHACTR4	+	protein_coding	UTR3	NM_001350159:c.*2309A>G;NM_001350158:c.*2309A>G;NM_001350160:c.*2309A>G;NM_001350161:c.*2309A>G;NM_001048183:c.*2309A>G;NM_023923:c.*2309A>G
chr19	58372069	A	G	ZNF587	+	protein_coding	UTR3	NM_032828:c.*561A>G;NM_001204817:c.*561A>G
chr12	98943103	A	G	TMPO	+	protein_coding	UTR3	NM_001307975:c.*1467A>G;NM_001032284:c.*1467A>G;NM_001032283:c.*1467A>G

Table 2

	Primer Name	primer Sequence	SEQ ID	Locus	Tm	Sequences of partial amplicon containing editing sites	SEQ ID	product size	sequencing primer	Cycles	validation
ARSD	ARSD-3UTR-F	AATGCCAGGACTGTGAGAGAGG	1	chrX:2824979	60	gttttagtagagacA/Ggagtttcacca	83	545		40cycle	yes
	ARSD-3UTR-R1	TTGGCACAATGACGAACCCT	2						ARSD-3UTR-R1		
	ARSD-3UTR-F2	AGAAATGGCTTCAGGCTGCT	3	chrX:2824434	60	tgggaggctA/Gagtgggaggatcac	84	335	ARSD-3UTR-F2	50cycle	SNP
	ARSD-3UTR-R	GTCTTACTCTGTCTCCCAGCCC	4								
EIF2AK2	EIF2AK2-3UTR-F	ATGAAACTCTACCCATCACCCAGC	5	chr2:37329283	60	tcctgcctcA/Gccgcttgaatagc	85	399		50cycle	yes
	EIF2AK2-3UTR-R	CCAGTAGGGCCGGATATTGTAATCC	6						EIF2AK2-3UTR-R		
MAVS	MAVS-3UTR-F	ACTCCTAGCATTCCTCACTCTGTTG	7	chr20:3852829	60	aggtggtggatcatacacttA/Gtaatcccagcactt	86	660	MAVS-3UTR-F	50cycle	yes
	MAVS-3UTR-R	GATGTTGGCACACTGGGCAG	8								
CCNYL1	CCNYL1-3UTR-F6	GGCACTGCAGCATGAGTATAAAT	9	chr2:208619479	60	ccatcctctagcctcA/Ggccacacaagta	87	259	CCNYL1-3UTR-F6	50cycle	yes
	CCNYL1-3UTR-R7	GCTTAATATTCAGGAGACAGGTGCC	10								
	CCNYL1-3UTR-F7	GGCACCTGTCTCCTGAATATTAAGC	11	chr2:208619749	60	ataggaatagctactA/Gcactccaacctg	88	285	CCNYL1-3UTR-F7	50cycle	yes
	CCNYL1-3UTR-R8	CATTTTTCTAAGAGATGGGGTCTTG	12								
NOP14	NOP14-3UTR-F7	GGCAAGGAACTGGACATTACCTC	13	chr4:2940462	60	ctcctgggttcaA/Ggcgattctcctgcctcagc	89	188	NOP14-3UTR-F7	50cycle	yes
	NOP14-3UTR-R8	CCTCCCGAGTAGTTGGGACTAC	14						NOP14-3UTR-R8		
				chr4:2940164				161			
									NOP14-3UTR-R9		
	NOP14-3UTR-F4	CAATCAGTTCCTGGCGAGGAT	15	chr4:2940462	60	ttgctttcagcccA/Ggagttcgaga	90	386	NOP14-3UTR-F4	50cycle	yes
	NOP14-3UTR-R5	CCTTGGACTCAGCTCAAGCA	16								
PAICS	PAICS-3UTR-F	GTTCTCCCGAAGCCTATCCATG	17	chr4:57327082	60	gcggaagttgcagtgA/Ggcctagatcat	91	492	PAICS-3UTR-F	50cycle	yes
	PAICS-3UTR-R	CATGGACAGCGAACATAAAGCTC	18								
NEAT1	NEAT1-3UTR-F	GCTGCTGTTTCCTGTCCCTTTATG	19	chr11:65208818	60	ttgtttgttttagtA/Gatgtgggagt	92	657	NEAT1-3UTR-F	50cycle	yes
	NEAT1-3UTR-R	GAACCAGACACACAATGCGAGC	20								
CPT1A	CPT1A-3UTR-F	TTGGAGTCGATGGGAAGGTGC	21	chr11:68523472	62	gcctcccA/Gagtagctg	93	402	CPT1A-3UTR-R	40cycle	yes
	CPT1A-3UTR-R	TTTGGCCCATAAAGGTGGGA	22								
FGD5-AS1	FGD5-AS1-F5	AAAAGTTGGGGTCTCGCTGT	23	chr3:149866098	60	agtgttaggattA/Gcaggcgtgagccacc	94	131	FGD5-AS1-F5	50cycle	yes
	FGD5-AS1-R5	GAGTACACAACCAGGCAGGG	24								
LIMD1	LIMD1-edt-2-F	tccttatgccaagaatcatgtcct	25	chr3:45722287	60	gtggatcacttA/Gaggtcaggag	95	624	LIMD1-edt-2-F	40cycle	yes
	LIMD1-edt-2-R	agtttgctgactccacacaca	26	chr3:45722300		ggagttA/Gagaccagcctga	96				yes
				chr3:45722381		cgcacgcctgtA/Gatcccagcta	97				yes
	LIMD1-3UTR-F1	TCCATGGGCTTCTGTCACAC	27	chr3:45721917	60	ctcccA/Gcctcagcctccc	98	884	LIMD1-SEQ-F	50cycle	yes
	LIMD1-edt-1-R	tgattctttggcataaaggaggg	28	chr3:45721934		tcccgagtA/Ggccgggaccacagg	99		caggacctcactctgttgcc (SEQ ID NO: 136)		yes
NUP43	NUP43-3UTR-F1	TGGTAAGGTTTTGCTTTCTGTCC	29	chr6:150046432	60	ttttgattagtA/Ggagacgacagg	100	504	NUP43-3UTR-F1	50cycle	yes
	NUP43-3UTR-R1	TGGCATGAAATGTAATGCTGTT	30								
				chr6:150045858	60	gcaggcatctatA/Gatcccagcta	101	219			
ORC2	ORC2-3UTR-F2	GAGTGCAGTGGCGTGATCTA	31	chr2:201774264	60	ttgaatttttA/Gcagagatgg	102	236	ORC2-3UTR-F2	50cycle	yes
	ORC2-3UTR-R2	CATAGTGGCTCACCCCTGAT	32								
PGAM5	PGAM5-3UTR-F	TGTGGGGACTTGAAAGAGGC	33	chr12:133297904	60	ggctggggcacA/Gtggctcacacctgta	103	260	PGAM5-3UTR-F	40cycle	yes
	PGAM5-3UTR-R5	GGGAATACAGGTGTGAGCCA	34								
	PGAM5-3UTR-F5	CGCCCACCACCATACCTG	35	chr12:133298339	60	catacctggctA/Gatttttgtatttttta	104	298	PGAM5-3UTR-R	40cycle	yes
	PGAM5-3UTR-R	TAGTCCCCTCGGCCCCTAAAG	36								
RBM39	RBM39-intronic-F2	GCGTTATCCTTGGTCACACTGC	37	chr20:34301191	60	tctctgggatttaA/Ggtgaaccacctc	105	1000	RBM39-intronic-R2	40cycle	yes
	RBM39-intronic-R2	GGCGGAATCTGCAAACCTGTA	38								
RRP36	RRP36-3UTR-F	TGGTTCATTGCGTCCTCGTG	39	chr6:42997578	60	CTAGTTTGATCTTA/GAACTCTTGGCCTC	106	812	RRP36-3UTR-F	40cycle	yes
	RRP36-3UTR-R	GCCTGAGGATTTTCCAGACCA	40	chr6:42997538		CGTATTTTTA/GGTAGAGATAGAGT	107				yes
SLC35F5	SLC35F5-3UTR-F	CTCTGTTGCCCCCTGGTTAC	41	chr2:114466200	60	ctcctgagttcaA/Ggtgattctcctg	108	657	SLC35F5-3UTR-F	40cycle	yes
	SLC35F5-3UTR-R	GCAGTCTAGATGGCTGGGGA	42								
	SLC35F5-3UTR-F3	GGAGGAATGCAGACAGGACAT	43	chr2:114465355	60	attaggagttcaaA/Gaccagtgtgaccaac	109	545	SLC35F5-3UTR-F3	40cycle	yes
	SLC35F5-3UTR-R3	GGATGACTGCAAAGAAAGGCAA	44								

ZNF587	ZNF578-3UTR-F4	GGCTCATGCTTGTAAACCTGATG	45	chr19:58372048	60	ccaggagttA/Ggagatcag	110	217	ZNF578-3UTR-F4	40cycle	yes
	ZNF578-3UTR-R4	TGCACACCACCATATTACCATAGTG	46	chr19:58372069		ctgggcaacatA/Ggccagacc	111				yes
	ZNF578-3UTR-F5	CTATCTTACCAAGAGGGAGTATGG	47	chr19:58372768	60	ggcacaatctcA/Ggctcattgc	112	303	ZNF578-3UTR-F5	50cycle	yes
	ZNF578-3UTR-R5	TCAGACTGGGCAAAAAGGT	48								
	ZNF578-3UTR-F6	CAGCTATATGAAGAATCTTGGCCAG	49	chr19:58375195	60	ggcggatcacA/Gaggtcagg	113	133	ZNF578-3UTR-F6	60cycle	yes
	ZNF578-3UTR-R8	CCAGGTTCAACCGATTTTCTGG	50								
METTL7A	METTL7A-3UTR-ed-144F	GGGAAGAAAGAGTTTTGCTCTTG	51	chr12_51324467	60	cccgggttcaA/Ggtgattcac	114	670	METTL7A-3UTR-ed-144F	50cycle	yes
	METTL7A-3UTR-ed-813R	TGGGAGACAATAGAAAATGGCTA	52	chr12_51324496		gcctcctgagtA/Ggctggtatta	115				yes
SNRPD3	SNRPD3-3UTR-F	TCAGCACAGTGGCTTGCTT	53	chr22_24969092	60	GGATCACTTGA/GGCCTGGGA	116	352	SNRPD3-3UTR-F	40cycle	yes
	SNRPD3-3UTR-R	GGCTGCTTTTCCAGTTTCTGT	54								
SRFBP1	SRFBP1-3UTR-F	AGTGCCTCTTTCTGCAAACTT	55	chr5_121362874	60	TTTTTTTTA/GAGACAGGA	117	228		40cycle	yes
	SRFBP1-3UTR-R	AGTAATAGCTGCAGAGCCTGG	56						SRFBP1-3UTR-R		
CTSS	CTSS-3UTR-F	GGCTCCTTCTCCATAAAGCAC	57	chr1:150703692	60	agtaaagacA/Gggttttgcca	118	554		50cycle	yes
	CTSS-3UTR-R	AGAAGGCATTCATAGGCTTTAACA	58						CTSS-3UTR-R		
NDUFS1	NDUFS1-edt-F	CTCCCTTGGAATGAGGGAGGT	59	chr2_206987652	60	AACATGGTGAAA/GCCCCATCTCTA	119	630		50cycle	yes
	NDUFS1-edt-R	GGGCTGAGTCTTTAGTGAAAGC	60						NDUFS1-edt-R		
CYP20A1	CYP20A1-3UTR-F	GGAGGGCTCAATCAACAGGT	61	chr2_204169105	60	TATTTTTAGTA/GGAGACGGGG	120	1167		50cycle	check in other cell
	CYP20A1-3UTR-R	AGGTTGGTAGCATTCTTTTTCCT	62						CYP20A1-3UTR-R		
PHACTR4	PHACTR4-3UTR-F	CTGTGGGTCCCTGTTATCTGT	63	chr1_28825370	60	ccaaggcggtcA/Gggagtttgaga	121	444	PHACTR4-3UTR-F	50cycle	yes
	PHACTR4-3UTR-R	TCAAGGAGATGTGGATTCAGG	64	chr1_28825419		tctctactaA/Gaaatacaaa	122				yes
AC079447.1,MRPL30	MRPL30-3UTR-F	GGGTCTACTGCCACCTTAGC	65	chr2_99812857	60	ACCTCCCA/GGGCTCAAGT	123	416	MRPL30-3UTR-F	50cycle	yes
	MRPL30-3UTR-R	ATGCAAAGAGCCACTCACAGA	66	chr2_99812978		TGTTGTCCAA/GGCTGGTCT	124				yes
GATC	GATC-3UTR-F	CCAAACTGCCAAAGTTGGTT	67	chr12: 120899075	60	attttttttA/Gagtagaggtggg	125	464		35cycle	yes
	GATC-3UTR-R	CGTAAACTCCAAGAGCGGGT	68						GATC-3UTR-R		
LONP2	LONP2-intronic-F	GGATCTAAGCATGTCCACTCTACA	69	chr16_48388245	60	GCCTCAGCCTCCCAA/GGTAGCTG	126	745		50cycle	yes
	LONP2-intronic-R	TCTGCAAGGCAGAGATGTGAAA	70						LONP2-intronic-R		
MED18	MED18-3UTR-F	CCACATTTGGGGCCTATCCT	71	chr1:28661576	60	ttttgagacA/Ggagtctcgc	127	397		40cycle	yes
	MED18-3UTR-R	GCCTGGCAATGGTGAAATGG	72						MED18-3UTR-R		
SPPL2A	SPPL2A-intergenic-F5	GCGGATCATGGAGAATGGGGTA	73	chr15_50995764	60	GCCTGACCA/GACATGGTGAA	128	263	SPPL2A-intergenic-F5	60cycle	yes
	SPPL2A-intergenic-R5	AGCTGGGGTTACAGGCGA	74								
TMPO	TMPO-3UTR-F3	GAGATGGAGTCTCACTCTTGTCAC	75	chr12_98943103	60	ccaggttcaA/Ggcaattttcct	129	483	TMPO-3UTR-R:tcatccacaaagataaccatttcaa (SEQ ID NO: 137)	50cycle	yes
	TMPO-3UTR-R3	ACATCTTCGGCTTCTTCCTTTTTAG	76								
AC019117.3,AHR	AHR-3UTR-seq-F	ACTGGAGTGCAGTGGCAC	77	chr7_17384527	60	CTGGGATTA/GCAGGCGTG	130	260	AHR-3UTR-seq-F	50cycle	yes
	AHR-3UTR-R2	TGCAAAATGTGTCTAAGCTCTTCA	78								
	AHR-3UTR-F2	ATGGAATCCCAAACTGAGATGT	79	chr7_17384792		AAATTA/GGCTGGGC	131	272	AHR-3UTR-F2	50cycle	yes
	AHR-3UTR-seq-R	GGAGACGGAGTCTCGGTCTATTG	80	chr7_17384832		CTACTCA/GGGAGGC	132				yes
				chr7_17384887		AGAGCCAA/GGATCGCCC	133				yes
				chr7_17384778		CTCTACTAA/GAAATACA	134				yes
STK4	STK4-3UTR-F	ACCTGCCAGATGCCAGAAAT	81	chr20_43707709	60	AAAAAAAAAATTA/GGTCGGGC	135	723	STK4-3UTR-Seq-R: AGACGGGGTCTCATTCTGTCA (SEQ ID NO: 138)	35cycle	yes
	STK4-3UTR-R	AGATGGGGTGAGGTGGAGAT	82								

CLAIMS

1. A method of predicting the responsiveness of a cancer subject to a platinum-based chemotherapy, the method comprising determining the level of adenosine to inosine (A-to-I) RNA-editing events in 29 genes as described in Table 1 in a cancer sample obtained from the subject, wherein an increased level of adenosine to inosine (A-to-I) RNA-editing events in the 29 genes as compared to a reference indicates that the subject is likely to be responsive towards the platinum-based chemotherapy, wherein the subject is suffering from gastric cancer, liver cancer, esophageal cancer or lung cancer.

2. The method of claim 1, wherein the platinum-based chemotherapy is a platinum-fluoropyrimidine doublet chemotherapy.

3. The method of claim 1 or 2, wherein the cancer subject is suffering from an advanced cancer.

4. The method of any one of the preceding claims, wherein the level of A-to-I RNA-editing events is determined by RNA sequencing.

5. The method of any one of the preceding claims, wherein the method comprises determining the level of A-to-I RNA-editing events in one or more RNA-editing sites at positions as described in Table 1.

6. The method of claim 5, wherein the method comprises determining a score based on the level of A-to-I RNA-editing events in the one or more RNA-editing sites.

7. The method of claim 6, wherein the score is a z score calculated based on the formula: $z = \frac{x - \mu}{\sigma}$, wherein where x is the RNA-editing level of each RNA-editing site, μ is the mean RNA-editing level of all the RNA-editing sites and σ is the standard deviation of the editing level of all RNA-editing sites, wherein the z score has a cut-off of 0.4, wherein a z score of >0.4. indicates that the subject is likely to have a complete response (CR) or partial response (PR), while a z score of ≤0.4

indicates that the subject is likely to have a stable disease (SD) or progressive disease (PD).

8. A method of stratifying a subject into a likely responder or non-responder of a platinum-based chemotherapy, the method comprising determining the level of adenosine to inosine (A-to-I) RNA-editing events in 29 genes as described in Table 1 in a cancer sample obtained from the subject, wherein an increased level of adenosine to inosine (A-to-I) RNA editing events in the 29 genes as compared to a reference indicates that the subject is a likely responder of the platinum-based chemotherapy and wherein an unchanged or decrease level of adenosine to inosine (A-to-I) RNA-editing events indicates that the subject is a likely non-responder of the cancer therapy, wherein the subject is suffering from gastric cancer, liver cancer, esophageal cancer or lung cancer.

9. A method for managing treatment of a cancer subject with a platinum-based chemotherapy, the method comprising a) selecting a cancer subject for treating with the therapy on the basis that the subject is a likely responder to the therapy, or selecting a subject for not treating with the therapy on the basis that the subject is a likely non-responder to the therapy; and b) wherein the subject is to be treated or not treated with the therapy based on the selection, wherein the selection is based on the stratification method of claim 8, wherein the subject is suffering from gastric cancer, liver cancer, esophageal cancer or lung cancer.

10. A method for predicting treatment outcome of a cancer subject with a platinum-based chemotherapy, the method comprising determining the level of adenosine to inosine (A-to-I) RNA-editing events in 29 genes as described in Table 1 in a cancer sample obtained from the subject, wherein an increased level of adenosine to inosine (A-to-I) RNA editing events as compared to a reference indicates that the subject is likely to have a positive treatment outcome with the cancer therapy, wherein the subject is suffering from gastric cancer, liver cancer, esophageal cancer or lung cancer.

11. A method of identifying a subject who is suitable for treatment with a platinum-based chemotherapy, the method comprising determining the level of adenosine to inosine (A-to-I) RNA-editing events in 29 genes as described in Table 1 in a cancer

sample obtained from the subject, wherein an increased level of adenosine to inosine (A-to-I) RNA editing events in the 29 genes as compared to a reference indicates that the subject is likely to have a positive treatment outcome with the cancer therapy and therefore suitable for treatment with the cancer therapy, wherein the subject is suffering from gastric cancer, liver cancer, esophageal cancer or lung cancer.