

COMPOSITIONS AND METHODS FOR ASSESSING NUCLEIC ACIDS IN A BIOLOGICAL SAMPLE

FIELD OF THE INVENTION

5

The present invention relates generally to the field of molecular biology. In particular, the present invention relates to biomarkers, reagents and methods for use in the pre-analytical workflow for the assessment of nucleic acids in biological samples.

10 BACKGROUND OF THE INVENTION

The analysis of circulating nucleic acids have recently shown considerable promise as a non-invasive approach to identify and measure disease biomarkers for the diagnosis and treatment of various diseases.

15

However, the analysis of such circulating nucleic acids can be dramatically impacted by the pre-analytical variables. In studies where changes in the nucleic acid expression levels in the bodily fluids are subtle, particularly at disease onset, the biomarker signal to noise ratio are often low. Therefore, confounders introduced by preanalytical conditions may significantly impact the reliability of the data obtained from a sample or the diagnostic accuracy of a test to detect disease in an individual.

20

Improper blood processing may result in cellular contamination or haemolysis of plasma or serum which renders variation in nucleic acid profile unrelated to pathological conditions. Residual platelet contamination is another significant confounding source of circulating nucleic acids. Additionally, assessing the condition of the nucleic acids in a sample is useful for ensuring that the samples with degraded nucleic acids can be identified to reduce potential errors. The application of nucleic acid detection and quantification in the clinical or research labs requires highly reliable and reproducible system to minimise the effect of such confounders. Therefore, proper quality control as well as reliable blood sampling and processing steps become critical preanalytical steps.

25

30

SUMMARY OF INVENTION

In one aspect, there is provided a method of assessing the quality of a sample for use in a nucleic acid analysis, the method comprises determining the level of one or more biomarker in the sample, wherein the one or more biomarker is selected from the group consisting of hsa-miR-28-5p, hsa-miR-1973, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, and hsa-miR-618.

35

In some embodiments, the quality of the sample is determined by a) the level of platelet activation and/or aggregation in the sample, b) the level of haemolysis in the sample, and/or c) the level of stability of nucleic acid in the sample.

40

In some embodiments, a) is determined by measuring/detecting/determining the level of one or more biomarker selected from the group consisting of hsa-miR-28-5p and hsa-miR-1973. In some embodiments, the method comprises determining the level of one of the following: hsa-miR-28-5p; or
5 hsa-miR-1973; or hsa-miR-1973 and hsa-miR-28-5p.

In some embodiments, b) is determined by measuring/detecting/determining the level of one or more biomarker selected from the group consisting of hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a. In some embodiments, the method comprises determining the level of one of the following: hsa-miR-
10 20b-5p; hsa-miR-363-3p; hsa-miR-451a; hsa-miR-20b-5p and hsa-miR-363-3p; hsa-miR-20b-5p and hsa-miR-451a; hsa-miR-363-3p and hsa-miR-451a; or hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a.

In some embodiments, c) is determined by measuring/detecting/determining the level of one or more biomarker is selected from hsa-miR-143-3p and hsa-miR-618. In some embodiments, the method comprises determining the level of: hsa-miR-143-3p; hsa-miR-618; or hsa-miR-143-3p and hsa-miR-
15 618.

In another aspect, there is provided a method of determining the level of platelet activation and/or aggregation in a sample, the method comprises determining the level of one or more biomarker in the sample, wherein the one or more biomarker is selected from hsa-miR-28-5p and hsa-miR-1973.
20

In some embodiments, the method comprises determining the level of one of the following: hsa-miR-28-5p; hsa-miR-1973; or hsa-miR-1973 and hsa-miR-28-5p.
25

In yet another aspect, there is provided a method of determining the level of haemolysis in a sample, the method comprises determining the level of one or more biomarker in the sample, wherein the one or more biomarker is selected from hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a.

In some embodiments, the method comprises determining the level of one of the following: hsa-miR-20b-5p; hsa-miR-363-3p; hsa-miR-451a; hsa-miR-20b-5p and hsa-miR-363-3p; hsa-miR-20b-5p and hsa-miR-451a; hsa-miR-363-3p and hsa-miR-451a; or hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a.
30

In yet another aspect, there is provided a method of determining the level of stability of nucleic acid in a sample, the method comprises determining the level of one or more biomarker in the sample, wherein the one or more biomarker is selected from hsa-miR-143-3p and hsa-miR-618.
35

In some embodiments, the method comprises determining the level of: hsa-miR-143-3p; hsa-miR-618; or hsa-miR-143-3p and hsa-miR-618.
40

In some embodiments, the method further comprises determining the level of one or more biomarker selected from hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720.

- 5 In some embodiments, the method further comprises comparing the level of hsa-miR-1973 and/or hsa-miR-28-5p with that of hsa-miR-1290 and/or hsa-miR-10b-5p to determine the level of platelet activation and/or aggregation in the sample.

- 10 In some embodiments, the method further comprises comparing the level of one or more of hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a with that of hsa-miR-1290 and/or hsa-miR-10b-5p to determine the level of haemolysis in the sample.

- 15 In some embodiments, the method further comprises comparing the level of hsa-miR-143-3p and hsa-miR-618 with that of hsa-miR-1290 and/or hsa-miR-720 to determine the level of stability of nucleic acid in the sample.

- 20 In some embodiments, the method comprises determining the level of at least one biomarker in each of the following group: (a) hsa-miR-28-5p and hsa-miR-1973; (b) hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a; and (c) hsa-miR-143-3p and hsa-miR-618.

- In some embodiments, the method comprises determining the level of hsa-miR-28-5p, hsa-miR-1973, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p and hsa-miR-618.

- 25 In some embodiments, the method comprises determining the level of biomarkers hsa-miR-28-5p, hsa-miR-1973, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, hsa-miR-618, hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720.

- 30 In some embodiments, the method comprises comparing the level of two or more biomarkers to determine a score, wherein the score is compared to that of a control for determining the quality of the sample.

In some embodiments, the sample is a blood sample or a non-cellular bodily fluid sample obtained from a subject, optionally the sample is a plasma and/or serum.

- 35 In yet another aspect, there is provided a kit for use in the method as disclosed herein, the kit comprises an isolated set of probes and/or reagents capable of determining the expression level of the one or more biomarker selected from the group consisting of hsa-miR-1973, hsa-miR-28-5p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, and hsa-miR-618.

In some embodiments, the kit further comprises an isolated set of probes and/or reagents capable of determining the expression level of one or more biomarker selected from the group consisting of hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720.

- 5 Also disclosed are isolated sets of probes capable of detecting at least one biomarker associated with platelet activation and/or aggregation, haemolysis or nucleic acid degradation in a sample.

Also disclosed are isolated sets of probes capable of detecting and/or determining the level of at least one biomarker selected from hsa-miR-1973 and hsa-miR-28-5p for use in determining the level of
10 platelet activation and/or aggregation in a sample. In some examples, the isolated set of probes may further comprise at least one probe for detecting and/or determining the level of at least one biomarker selected from hsa-miR-1290 and hsa-miR-10b-5p. In some examples, the isolated set of probes comprises probes for detecting and/or determining the level of hsa-miR-1973, hsa-miR-28-5p, hsa-miR-1290 and hsa-miR-10b-5p.

15 Also disclosed are isolated sets of probes capable of detecting and/or determining the level of at least one biomarker selected from hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a for use in determining the level of haemolysis in a sample. In some examples, the isolated set of probes further comprises at least one probe for detecting and/or determining the level of at least one biomarker
20 selected from hsa-miR-1290 and hsa-miR-10b-5p. In some examples, the isolated set of probes comprises probes for detecting and/or determining the level of hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-1290 and hsa-miR-10b-5p.

Also disclosed are isolated set of probes capable of detecting and/or determining the level of at least
25 one biomarker selected from hsa-miR-143-3p and hsa-miR-618 for use in determining the level of stability of nucleic acid in a sample, optionally wherein the sample is a stored sample. In some examples, the isolated set of probes may further include at least one probe for detecting and/or determining the level of at least one biomarker selected from hsa-miR-1290 and hsa-miR-720, optionally wherein the sample is a stored sample. In some examples, the isolated set of probes
30 comprises probes for detecting and/or determining the level of hsa-miR-143-3p, hsa-miR-618, hsa-miR-1290 and hsa-miR-720, optionally wherein the sample is a stored sample.

Also disclosed are isolated set of probes capable of detecting and/or determining the level of at least one biomarker selected from hsa-miR-1973, hsa-miR-28-5p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, hsa-miR-618, hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720. In some
35 examples, the isolated set of probes comprises probes for detecting and/or determining the level of hsa-miR-1973, hsa-miR-28-5p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, hsa-miR-618, hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720.

Also disclosed are kits for determining the level of platelet activation and/or aggregation in a sample, the kit comprising reagents for detecting and/or determining the expression level of at least one biomarker selected from hsa-miR-28-5p and hsa-miR-1973.

- 5 In some examples, reagents may include, but is not limited to, probes, primers, aptamers, antibodies, buffers, and enzymes.

In some examples, the kit further comprises reagents for detecting and/or measuring the level of hsa-miR-1290 and hsa-miR-10b-5p. In some examples, the kit comprises reagents for detecting and/or
10 determining the level of hsa-miR-1973, hsa-miR-28-5p, hsa-miR-1290 and hsa-miR-10b-5p. In some examples, such reagents may include, but is not limited to, probes, primers, aptamers, antibodies, buffers, enzymes, and the like.

Also disclosed are kits for determining the level of haemolysis in a sample, the kit comprising an
15 isolated set of probes and/or reagents for detecting and/or determining the expression level of at least one biomarker selected from hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a. In some examples, the kit comprises reagents for detecting and/or determining the level of hsa-miR-1290 and hsa-miR-10b-5p. In some examples, the kit comprises reagents for detecting and/or determining the level of hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-1290 and hsa-miR-10b-5p. In
20 some examples, such reagents may include probes, primers, aptamers, antibodies, buffers and enzymes.

Also disclosed are kits for determining the level of stability of nucleic acid in a sample, the kit comprising reagents for detecting and/or determining the expression level of at least one biomarker
25 selected from hsa-miR-143-3p and hsa-miR-618, optionally wherein the sample is a stored sample. In some examples, the kit further comprises reagents for detecting and/or determining the level of hsa-miR-1290 and hsa-miR-720. In some examples, the kit comprises reagents for detecting and/or determining the level of hsa-miR-143-3p, hsa-miR-618, hsa-miR-1290 and hsa-miR-720. In some examples, such reagents may include probes, primers, aptamers, antibodies, buffers and enzymes.

30 Also provided herein is a method of determining the level of platelet activation and/or aggregation in a sample, the method comprises determining the level of one or more biomarker in the sample, wherein the one or more biomarker is selected from hsa-miR-1973 and hsa-miR-28-5p. In some embodiments, the method comprises determining the level of one or two biomarker in the sample,
35 wherein the biomarker is hsa-miR-1973; hsa-miR-28-5p; or hsa-miR-1973 and hsa-miR-28-5p. In some embodiments, the method further comprises determining the level of one or more reference or normaliser biomarker selected from hsa-miR-1290 and hsa-miR-10b-5p, such as hsa-miR-1290; hsa-miR-10b-5p; or hsa-miR-1290 and hsa-miR-10b-5p. In some embodiments, the method further comprises comparing the level of hsa-miR-1973 and/or hsa-miR-28-5p with that of hsa-miR-1290
40 and/or hsa-miR-10b-5p to determine the level of platelet activation and/or aggregation in the sample.

Also provided herein is a method of determining the level of haemolysis in a sample, the method comprises determining the level of one or more biomarker in the sample, wherein the one or more biomarker is selected from hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a. In some
5 embodiments, the method comprises determining the level of one, two or three biomarkers. In some embodiments, the method comprises determining the level of hsa-miR-20b-5p; hsa-miR-363-3p; hsa-miR-451a; hsa-miR-20b-5p and hsa-miR-363-3p; hsa-miR-20b-5p and hsa-miR-451a; hsa-miR-363-3p and hsa-miR-451a; or hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a in the sample. In some
10 embodiments, the method further comprises determining the level of one or more reference or normaliser biomarker selected from hsa-miR-1290 and hsa-miR-10b-5p, such as hsa-miR-1290; hsa-miR-10b-5p; or hsa-miR-1290 and hsa-miR-10b-5p. In some embodiments, the method further comprises comparing the level of one or more of hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a with that of hsa-miR-1290 and/or hsa-miR-10b-5p to determine the level of haemolysis in the sample.

15 Also provided herein is a method of determining the level of stability of nucleic acid in a sample, the method comprises determining the level of one or more biomarker in the sample, wherein the one or more biomarker is selected from hsa-miR-143-3p and hsa-miR-618. In some embodiments, the method comprises determining the level of hsa-miR-143-3p; hsa-miR-618; or hsa-miR-143-3p and hsa-miR-618 in the sample. In some embodiments, the method further comprises determining the
20 level of one or more reference or normaliser biomarker selected from hsa-miR-1290 and hsa-miR-720, such as hsa-miR-1290; hsa-miR-720; or hsa-miR-1290 and hsa-miR-720. In some embodiments, the method further comprises comparing the level of hsa-miR-143-3p and/or hsa-miR-618 with that of hsa-miR-1290 and/or hsa-miR-720 to determine the level of stability of nucleic acid in the sample.

25 In some embodiments of the methods, the method further comprises comparing the level of two or more biomarker to determine a score for determining the level of platelet activation and/or aggregation, stability of nucleic acid or level of haemolysis in a sample. In some embodiments, the method comprises comparing the level of two or more biomarkers, wherein one of the biomarkers is a
30 reference or normaliser biomarker. In some embodiments, the score is compared to that of a control for determining the level of platelet activation and/or aggregation, stability of nucleic acid or level of haemolysis in a sample.

Also provided are the methods for performing quality control on a sample and/or assessing quality of
35 a sample for use in nucleic acid analysis, comprising determining in the sample the occurrences of one or more conditions that may alter the nucleic acid content in the sample.

In some embodiments, such methods for performing quality control on a sample and/or assessing quality of a sample for use in nucleic acid analysis comprises detecting the presence of at least one of
40 the following conditions in the sample: (a) level of platelet activation and/or aggregation; (b) level of

haemolysis; and/or (c) level of stability of nucleic acid, wherein should one or more of the above conditions is identified in the sample, a person with the relevant skill in the art would be able to make the judgement to: (a) not to proceed with nucleic acid analysis for the sample; (b) to make adjustments in the data analysis for the affected sample to adjust for the biasness of the nucleic acid profile in said sample; and/or (c) to undertake other corrective measures necessary to derive useful information on the nucleic acid content of the sample.

Also disclosed are methods for performing quality control on a sample and/or assessing quality of a sample for use in nucleic acid analysis, wherein the method comprises determining in the sample the level of one or more biomarkers selected from hsa-miR-1973, hsa-miR-28-5p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, hsa-miR-618, hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720. In some embodiments, the method comprises determining the level of at least one biomarker in each of the following group: (a) hsa-miR-1973 and hsa-miR-28-5p; (b) hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a; and (c) hsa-miR-143-3p and hsa-miR-618. In some embodiments, the method comprises determining the level of at least one biomarker in each group, such as least 3, 4, 5, 6 or 7 biomarkers.

In some embodiments of the method, the method further comprises determining the level of one or more reference or normaliser biomarker selected from hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720. In some embodiments, the method further comprises determining the level of one, two or three biomarkers, such as hsa-miR-1290; hsa-miR-10b-5p; hsa-miR-720; hsa-miR-1290 and hsa-miR-10b-5p; hsa-miR-1290 and hsa-miR-720; hsa-miR-10b-5p and hsa-miR-720; or hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720.

In some embodiments, the method comprises determining the level of hsa-miR-1973, hsa-miR-28-5p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, hsa-miR-618, hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720.

In some embodiments, the method further comprises comparing the level of two or more biomarkers to determine a score for determining the quality of a sample. In some embodiments, the score is compared to that of a control for determining the quality of the sample.

In some embodiments, such methods for performing quality control on a sample and/or assessing quality of a sample for use in nucleic acid analysis comprises detecting in the sample the level of platelet activation and/or aggregation in the sample by detecting the level of at least one biomarker selected from hsa-miR-1973 and hsa-miR-28-5p compared to a control. In a further embodiment, such a method may further include detecting the level of at least one biomarker selected from hsa-miR-1290 and hsa-miR-10b-5p.

In some embodiments, such methods for performing quality control on a sample and/or assessing quality of a sample for use in nucleic acid analysis comprises detecting in the sample the level of haemolysis in the sample by detecting the level of at least one biomarker selected from hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a compared to a control. In a further embodiment, such a method may further include detecting the level of at least one biomarker selected from hsa-miR-1290 and hsa-miR-10b-5p.

In some embodiments, such methods for performing quality control on a sample and/or assessing quality of a sample for use in nucleic acid analysis comprises detecting in the sample the level of stability of nucleic acid in the sample by detecting the level of at least one biomarker selected from hsa-miR-143-3p and hsa-miR-618 compared to a control. In a further embodiment, such a method may further include detecting the level of at least one biomarker selected from hsa-miR-1290 and hsa-miR-720.

In a further aspect, the present disclosure refers to a kit for use in performing quality control on a sample and/or assessing quality of a sample for use in nucleic acid analysis, comprising reagents for detecting the expression of at least one biomarker selected from the group consisting of hsa-miR-1973, hsa-miR-28-5p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p and hsa-miR-618. In some embodiments, the kit further comprises reagents for detecting the expression of at least one biomarker selected from the group consisting of hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720. In some embodiments, such reagents may include probes, primers, aptamers, antibodies, buffers and enzymes.

In another aspect, disclosed herein are kits for use in any of the methods, the kit comprises an isolated set of probes and/or reagents capable of determining the level of the one or more biomarkers of the preceding claims.

In some examples, the kits as disclosed herein may further comprise instructions for the reagents to be used for performing quality control on a sample and/or assessing quality of a sample for use in nucleic acid analysis. Therefore, the kits may further comprise instructions for using the reagents to assess the presence of at least one of the following conditions in the sample: (a) level of platelet activation and/or aggregation; (b) level of haemolysis; and/or (c) level of stability of nucleic acid.

Also provided herein are methods for collecting a biological sample (e.g., blood) for nucleic acid analysis, such methods comprises: (a) collecting an initial first volume of the biological sample; and subsequently, (b) collecting a second volume of the biological sample in a separate container, for use in nucleic acid analysis. For the avoidance of doubt, the initial first volume and second volume of biological samples are not to be combined for nucleic acid analysis and preferentially, be collected in discrete containers. As examples, the containers may be, but not limited to, tubes, flasks, bags and

compartments within containers. In some embodiments, the initial first volume of the biological sample may be discarded and the volume to be discarded may be at least 1 millilitre (mL). In another aspect, the initial first volume of the biological sample to be discarded may be approximately 1-2mL.

5 Also provided herein are methods of processing a biological sample for nucleic acid analysis, such methods comprising at least two (2) centrifugation steps to provide a first and second supernatants, respectively. In some embodiments, the first centrifugation may be carried out at a speed of from 1,000g to 5,000g, such as from 1,000g to 4,000g, such as from 1,000g to 2,000g, such as 1,500g. In another embodiment, the second centrifugation would be carried out at a speed of from 1,500g to 5,000g, such as from 1,500g to 4,000g, such as from 1,500g to 3,000g, such as 2,500g. In another
10 embodiment, the centrifugation time for the first and second centrifugation may be independently selected from 1 min to 30 min, such as from 5 min to 20 min, such as 15 min.

In some embodiments, the methods of processing a biological sample for nucleic acid analysis may
15 include cryopreservation of the first supernatant after the first centrifugation, with the second centrifugation being carried out following the thawing of the cryopreserved first supernatant (i.e., after freeze/thaw).

Also disclosed is a method of collecting and processing a biological sample (e.g., blood) for nucleic
20 acid analysis comprising: (a) collecting an initial first volume of biological sample; (b) collecting a second volume of biological sample in a separate container; (c) subjecting the second volume of the biological sample to a first centrifugation to provide a first supernatant; and (d) subjecting the first supernatant to a second centrifugation to provide a second supernatant. In some examples, the first centrifugation may be carried out at a speed of from 1,000g to 5,000g, such as from 1,500g to 4,000g,
25 such as from 1,500g to 3,000g, such as 1,500g. In some examples, the second centrifugation may be carried out at a speed of 1,500g to 5,000g, such as from 1,500g to 4,000g, such as from 1,500g to 3,000g, such as 2,500g. In some examples, the centrifugation time for the first and second centrifugation may be independently selected from 1 min to 30 min, such as from 5 min to 20 min, such as 15 min. In some examples, the first supernatant may be cryopreserved after the first
30 centrifugation, with the second centrifugation performed following the thawing of the cryopreserved first supernatant. In some examples, the method may further comprise determining in the second supernatant, the level of at least one biomarker selected from hsa-miR-1973, hsa-miR-28-5p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, hsa-miR-618, hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720.

35 In some embodiments, the sample is selected from a tissue sample (e.g. biopsy tissue sample), a blood sample, or a non-cellular bodily fluid sample obtained from a subject, optionally the sample is a plasma and/or serum.

In the embodiments, the biomarker is microRNA (miRNA). In some embodiments, the one or more biomarkers are detected/determined by methods known in the art such as, but is not limited to, sequencing, nucleic acid hybridisation, microarray and nucleic acid amplification. In some examples, nucleic acid amplification may include, but is not limited to, a reverse transcription-quantitative
5 polymerase chain reaction (RT-qPCR), reverse transcription-polymerase chain reaction (RT-PCR), quantitative polymerase chain reaction (qPCR), a locked nucleic acid PCR, a clustered regularly interspaced short palindromic repeat (CRISPR)-based assay, isothermal amplification assay, and the like.

10 DEFINITIONS

As used herein, "biomarker" or "marker" can refer to a gene, protein, or nucleic acid whose level of expression or concentration in a sample is altered compared to that of a control. A nucleic acid can refer to a deoxyribonucleic acid (DNA) or a ribonucleic acid (RNA). The term RNA may further refer to
15 a messenger RNA (mRNA), ribosomal RNA (rRNA), transfer RNA (tRNA) or small RNAs such as microRNAs (miRNAs), PIWI-interacting RNAs (piRNAs), small nucleolar RNAs (snoRNAs), small interfering RNAs (siRNAs), Y RNA and small nuclear RNAs (snRNAs). As used herein, circulating biomarkers shall refer to cell-free biomarkers found in blood and may include nucleic acids, extracellular vesicles, proteins and metabolites. The term circulating nucleic acids may refer to
20 circulating tumour DNA, mRNA, miRNA or other small RNAs found in blood.

As used herein, the term "nucleic acid analysis" can refer to the detection or measurement of nucleic acids (e.g., miRNA, circulating DNA, mRNA or other small RNA) using sequencing (e.g. Sanger sequencing, bisulfite sequencing or next generation sequencing), nucleic acid hybridization (e.g.,
25 northern blot), microarray and nucleic acid amplification (e.g., quantitative reverse transcription polymerase chain reaction (qRT-PCR), reverse transcription polymerase chain reaction (RT-PCR), quantitative polymerase chain reaction (qPCR), a locked nucleic acid (LNA) real-time PCR, a CRISPR-based assay, or isothermal amplification assay. An isothermal amplification assay can, for example, include, but is not limited to, a nicking endonuclease amplification reaction (NEAR) assay, a
30 transcription mediated amplification (TMA) assay, a loop-mediated isothermal amplification (LAMP) assay, a helicase-dependent amplification (HDA) assay, a clustered regularly interspaced short palindromic repeat (CRISPR) assay, or a strand displacement amplification (SDA) assay. In some aspect, nucleic acid analysis may refer to the detection or measurement of circulating nucleic acids (e.g., miRNA).

35 As used herein, a control refers to a level of expression or concentration of a biomarker that is indicative or correlated with a different outcome compared to the outcome of interest. For example, a biomarker can be a nucleic acid whose level of expression or concentration is altered (e.g., increased or decreased) due to a phenomenon (e.g., degradation, storage, haemolysis, platelet activation,
40 centrifugation, etc) compared to that of a control which had not undergone the alteration. It is

understood by anyone with the relevant skill in the art that comparing the level of expression in the control does not necessarily entail obtaining a control sample and testing said sample at the same time as the test subject. In some embodiments, said control could be a control sample incorporated in the kit or a threshold set to represent the range of expression of the biomarker where expression levels falling into this range would identify a test sample as having undergone a phenomenon (e.g., degradation, storage, haemolysis, platelet activation, centrifugation, etc) that led to the alteration of the measured biomarker.

As used herein, "level" or "expression level" of a biomarker refers to the concentrations of the biomarkers in the sample and may be expressed in any suitable forms of concentration known to the person skilled in the art, such as ng/ μ L, copy number, log(concentration), log2 copies/mL, threshold cycle (Ct) or quantification cycle (Cq) from qPCR and Ct or Cq power of 2.

As used herein, "sample", "biological sample" or "biological material" is intended to include any sampling of cells, tissues, or bodily fluids in which expression of a biomarker can be detected. Examples of such samples and biological samples include, but are not limited to, biopsies, smears, blood, lymph, urine, saliva, or any other bodily secretion or derivative thereof. Blood can, for example, include whole blood, plasma, serum, or any derivative of blood. In some embodiments, the sample or biological sample is whole blood, and/or a non-cellular bodily fluid. In some embodiments, the non-cellular bodily fluid may comprise serum or plasma. Samples can be obtained from a subject by a variety of techniques, which are known to those skilled in the art.

As used herein, the term "differential expression" refers to the measurement of a component (such as a cellular component) in comparison to a control or another sample, and thereby determining the difference in, for example concentration, presence or intensity of said component (such as said cellular component). The result of such a comparison can be given in the absolute, that is a component is present in the samples and not in the control, or in the relative, that is the expression or concentration of component is increased or decreased compared to the control. The terms "increased" and "decreased" in this case can be interchanged with the terms "upregulated" and "downregulated" which are also used in the present disclosure.

As used herein, determining the stability of the nucleic acid in a sample (e.g. a stored sample) refers to the assessment of a sample to determine whether the nucleic acid content in said sample has been altered (chemically and/or structurally) due to prolonged storage or incorrect storage condition leading to degradation or alteration of at least some of the nucleic acid within the sample.

As used herein, determining the level of haemolysis in a sample refers to assessing the extent of breakdown of erythrocytes, commonly referred to as red blood cells, resulting in the release of haemoglobin into the surrounding fluid in the sample. The determination of level of haemolysis may be performed using any common methods known in the art. Common methods to determine

haemolysis may include, but is not limited to, visual inspection, quantification of cell-free haemoglobin level (e.g. using a human haemoglobin enzyme-linked immunosorbent assay (ELISA)), and the like.

5 As used herein, determining the level of platelet activation and/or aggregation in a sample refers to measuring the level of platelet-to-platelet (thrombocytes) adhesion to aggregation in the sample, or the extent of clot formation. The level of platelet activation can be assessed by methods known in the art, such as, but is not limited to, flow cytometry (to detect the activated platelets directly), ELISA (to detect markers related to platelet activation (e.g. thromboxane A2 and B2 (TXA2 and TXB2)), and the like.

10

As used herein, performing quality control on a sample and/or assessing the quality of a sample intended for nucleic acid analysis may refer to determining if a sample has undergone degradation due to prolonged storage, incorrect storage condition or improper sample handling procedures leading to degradation of at least some of the nucleic acid within the sample, hence altering the detectable level of at least some of the nucleic acids contained within the sample. In other examples, performing quality control on a sample and/or assessing the quality of a sample for use in nucleic acid analysis may refer to determining whether a sample has undergone alterations to its nucleic acid level due to improper sample handling, prolonged storage or incorrect storage condition to create variations that are unrelated to the original condition of the sample. Examples of such alterations could be the release of cellular nucleic acids due to cellular contamination, activation of cellular response (e.g., platelet activation or aggregation) or haemolysis leading to contamination of the plasma or serum. It is understood by someone with knowledge in the art that, in some cases, such alterations may result in unacceptably low signal-to-noise ratio, which may render such samples unsuitable for analysis as the nucleic acid expression data could be uninformative or biased, leading to possible misinterpretation or misidentification of relevant biomarkers of interest. Where one or more of the above conditions is identified in the sample during the quality control process, the person with the relevant skill in the art would be able to make the judgement to: (a) not to proceed with nucleic acid analysis for the sample; (b) to make adjustments in the data analysis for the affected sample to adjust for the biasness of the nucleic acid profile in said sample; and/or (c) to undertake other corrective measures necessary to derive useful information on the nucleic acid content of the sample.

15
20
25
30

As used herein, "probe" refers to any molecule or agent that is capable of selectively detecting an intended target biomolecule, for example, by binding directly or indirectly to the target biomolecule. The target molecule can be a biomarker, for example, a nucleotide transcript or a protein encoded by or corresponding to a biomarker. Probes can be synthesized by one of skill in the art, or derived from appropriate biological preparations, in view of the present disclosure. Probes can be specifically designed to be labelled. Examples of molecules that can be utilized as probes include, but are not limited to, RNA, DNA (e.g., primers), proteins, peptides, antibodies, aptamers, affibodies, and organic molecules. In the case of a probe designed for the detection of a nucleic acid biomarker, such a

35

probe may be directed to the target region, the complementary nucleic acid sequence on the reverse strand or copies of the same generated via an amplification process.

As used herein, the term "(statistical) classification" refers to the problem of identifying to which of a set of categories (sub-populations) a new observation belongs, on the basis of a training set of data containing observations (or instances) whose category membership is known. An example is assigning a diagnosis to a given patient as described by observed characteristics of the patient (gender, blood pressure, presence or absence of certain symptoms, etc.). In the terminology of machine learning, classification is considered an instance of supervised learning, i.e., learning where a training set of correctly identified observations is available. The corresponding unsupervised procedure is known as clustering and involves grouping data into categories based on some measure of inherent similarity or distance. Often, the individual observations are analysed into a set of quantifiable properties, known variously as explanatory variables or features. These properties may variously be categorical (e.g., "A", "B", "AB" or "O", for blood type), ordinal (e.g., "large", "medium" or "small"), integer-valued (e.g., the number of occurrences of a part word in an email) or real-valued (e.g., a measurement of blood pressure). Other classifiers work by comparing observations to previous observations by means of a similarity or distance function. An algorithm that implements classification, especially in a concrete implementation, is known as a classifier. The term "classifier" sometimes also refers to the mathematical function, implemented by a classification algorithm, which maps input data to a category.

As used herein, the term "pre-trained" or "supervised (machine) learning" refers to a machine learning task of inferring a function from labelled training data. The training data can consist of a set of training examples. In supervised learning, each example is a pair consisting of an input object (typically a vector) and a desired output value (also called the supervisory signal). A supervised learning algorithm, that is the algorithm to be trained, analyses the training data and produces an inferred function, which can be used for mapping new examples. An optimal scenario will allow for the algorithm to correctly determine the class labels for unseen instances. This requires the learning algorithm to generalize from the training data to unseen situations in a "reasonable" way.

As used herein, the term "score" refers to an integer or number, that can be determined mathematically, for example by using computational models known in the art, which can include but are not limited to, SMV, as an example, and that is calculated using any one of a multitude of mathematical equations and/or algorithms known in the art for the purpose of statistical classification. Such a score is used to enumerate one outcome on a spectrum of possible outcomes. The relevance and statistical significance of such a score depends on the size and the quality of the underlying data set used to establish the results spectrum. For example, a blind sample may be input into an algorithm, which in turn calculates a score based on the information provided by the analysis of the blind sample. This results in the generation of a score for said blind sample. Based on this score, a decision can be made, for example, how likely the patient, from which the blind sample was obtained,

has cancer or not. The ends of the spectrum may be defined logically based on the data provided, or arbitrarily according to the requirement of the experimenter. In both cases the spectrum needs to be defined before a blind sample is tested. As a result, the score generated by such a blind sample, for example the number "45" may indicate that the corresponding patient has cancer, based on a spectrum defined as a scale from 1 to 50, with "1" being defined as being cancer-free and "50" being defined as having cancer.

BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing summary, as well as the following detailed description of preferred embodiments of the present application, will be better understood when read in conjunction with the appended drawings. It should be understood, however, that the application is not limited to the precise embodiments shown in the drawings.

FIG. 1 shows the hemoglobin level, platelet activation, and miRNA expression levels in blood samples under different processing methods: serum with a clotting duration of 30 min (serum 30); serum with a clotting duration of 60 min (serum 60); serum with clotting duration of 90 min (serum 90); normal plasma samples (plasma); platelet-poor-plasma (PPP). (A) Haemolysis level measured by human hemoglobin enzyme-linked immunosorbent assay (ELISA). The hemoglobin concentration was similar among all serum samples but higher in both plasma and PPP samples. (B) Platelet activation level determined by human thromboxane B2 ELISA. TXB2 level increased as clotting duration increased. TXB2 concentration in both plasma and PPP was minimal as platelet activation did not take place. (C) The miRNA expression examined by miRNA profiling. The average expression level of all measured miRNAs was used to indicate the expression level for a sample. miRNA level increased when clotting duration increased in serum. miRNA expression was highest in plasma while PPP had comparable expression level to serum. (D) Correlation between thromboxane B2 level and miRNA expression level of serum 30, serum 60, and serum 90 (Pearson $r = 0.6425$, $P = 0.00013$).

FIG. 2 shows the principal component analysis (PCA) scatter plot from 356 miRNA profiling after global normalization. (A) initial PCA by sample type. There was similarity between Plasma, Serum 60 and Serum 90. PPP was closer to Serum 30. (B) Subsequent PCA by sample types. (C) initial PCA by individuals (subject 1 to 10; subject 9 appeared as outlier and was removed from the PCA analysis) did not show clustering effect. (D) Subsequent PCA by individuals. PC1, principal component 1; PC2, principal component 2; Var, variance.

FIG. 3 shows the platelet-related biomarkers (n=10). (a) hsa-miR-1973 and (b) hsa-miR-28-5p showed higher expression in plasma than PPP. The expression in serum also correlated well with TXB2 level (platelet activation). (c) hsa-miR-1290 and (d) hsa-miR-10b-5p did not show differential expression in plasma and PPP. There was no correlation with TXB2 in serum samples.

FIG. 4 depicts the expression of miRNA markers and thromboxane B2 (TBX2) in different sample types. A–D: Platelet-related markers are hsa-miR-1973 (A) and hsa-miR-28-5p (B), and reference markers are hsa-miR-1290 (C) and hsa-miR-10b-5p (D). E: TBX2 level is shown. miRNA expression level in log2 copy numbers per milliliter of samples (Log2 copies/mL) was determined by miRNA profiling. TBX2 was measured by enzyme-linked immunosorbent assay. Refer to **Table 3** for the *P* value from one-way analysis of variance. *n* = 10.

FIG. 5 depicts the expression of platelet-related biomarkers in different sample types. Blood was collected from 12 healthy volunteers and processed into serum (60 minutes of clotting), plasma, and platelet-poor plasma (PPP). A–D: The expression of platelet-related biomarkers hsa-miR-1973 (A) and hsa-miR-28-5p (B) and reference markers hsa-miR-1290 (C) and hsa-miR-10b-5p (D) in serum (60 minutes of clotting), plasma and PPP. The results indicate that PPP had the lowest expression of platelet-related biomarkers, A and A, among the different samples. Statistical analyses were performed with one-way analysis of variance. ***P* < 0.01, ****P* < 0.001, and *****P* < 0.0001.

FIG. 6 depicts the expression profile of miRNA in red blood cells (RBCs), platelet-poor plasma (PPP) spiked with different concentration of RBC lysate (hemo 1, hemo 2, and hemo 3) and non-spiked PPP (no hemo). The expression of the RBC-related markers hsa-miR-20b-5p, hsa-miR-363-3p, and hsa-miR-451a increased when the RBC concentration increased. The expression of the reference markers hsa-miR-1290 and hsa-miR-10b-5p did not show difference in all non-spiked and spiked PPP samples. *n* = 4.

FIG. 7 depicts the expression profile of miRNA in red blood cells (RBCs), platelet-poor plasma (PPP) spiked with different concentration of RBC lysate (hemo 1, hemo 2, and hemo 3) and non-spiked PPP (no hemo). A: The difference in miRNA expression (Δ) in RBC and spiked PPP against non-spiked PPP. Each bar represents an individual miRNA. B: Expression of hsa-miR-1290, hsa-miR-10b-5p, hsa-miR-20b-5p, hsa-miR-363-3p, and hsa-miR-451a in RBCs, spiked PPP, and non-spiked PPP. miRNA expression level in log2 copy numbers per milliliter of samples (Log2 copies/mL) was determined by miRNA profiling. *n* = 4.

FIG. 8 shows the comparison of sample quality from different sizes of tubes (*n*=4). A: Hemoglobin levels (measured by ELISA) were significantly higher in discard tubes when compared to 6-mL tube and 10-mL tube (* *p*<0.05). B: Haemolysis scores: mean raw threshold cycle (Cq) of red blood cell-related miRNAs minus mean Cq of reference miRNAs. No significant difference was seen in haemolysis score for tubes with different sizes. C: Platelet scores: mean Cq of platelet-related miRNAs minus mean Cq of reference miRNAs. No significant difference was seen in Platelet score for tubes with different sizes. Statistical analyses were performed with one-way analysis of variance (A) and paired t-test (B and C).

FIG. 9 shows the comparison of sample quality from different PPP processing methods (n=5). Refer to Table 5 for processing conditions (Con1 to Con6). A: Haemolysis score: mean raw threshold cycle (Cq) of red blood cell-related miRNAs minus mean Cq of reference miRNAs. The haemolysis score was slightly decreased in lower first spin speed (Con5 and Con6). B: Platelet scores: mean Cq of platelet-related miRNAs minus mean Cq of reference miRNAs. Increase second spin speed resulted in better removal of platelet (Con 1 to Con3). However, Con4 showed the lowest platelet score among all the groups. This could be due to lysis of platelet under high-speed second spin. Con2 and Con3 had better score performance overall. Statistical analyses were performed with one-way ANOVA compared to Con1 (* $p \leq 0.05$).

FIG. 10 shows the comparison of sample quality from post freeze-thaw spin. A: Haemolysis scores: mean raw threshold cycle (Cq) of red blood cell-related miRNAs minus mean Cq of reference miRNAs. The haemolysis scores remained unchanged at all conditions. B: Platelet scores: mean Cq of platelet-related miRNAs minus mean Cq of reference miRNAs. The platelet score showed that the best condition was Con2. For FT (Freeze-thaw) samples, the second spin upon thawing (Con4) can partially remove platelets, hence increasing the platelet score. Statistical analyses were performed with one-way ANOVA compared to Con1 (**** $p \leq 0.0001$, ** $p \leq 0.001$, n=4).

FIG. 11 shows the comparison of sample quality from delayed processing (n=6). A: Hemoglobin level by Hb ELISA. The assay showed an elevated hemoglobin level in on ice samples. B: Haemolysis scores: mean raw threshold cycle (Cq) of red blood cell-related miRNAs minus mean Cq of reference miRNAs. The haemolysis score was consistent with Hb ELISA and showed significant difference in on ice samples. C: Platelet scores: mean Cq of platelet-related miRNAs minus mean Cq of reference miRNAs. The platelet score slightly decreased in on ice samples indicating lysis of platelets. Statistical analyses were performed with one-way ANOVA compared with baseline. (* $p \leq 0.05$, n=6).

FIG. 12 shows the mean expression levels (Log2 copies/mL) of all detectable miRNAs (252 miRNAs) in serum after storage at various temperature: (A) 25 °C, (B) 4 °C, (C) -20 °C, and (D) -80 °C. Data are expressed as medians with 95% confidence intervals (CIs). Statistical analyses were performed with one-way ANOVA compared with Day 0. (**** $p \leq 0.0001$, *** $p \leq 0.001$, n=4) Expression of serum miRNA at 25 °C and 4 °C significantly decreased on day 3 and day 30, respectively.

FIG. 13 shows the mean expression levels (Log2 copies/mL) of all detectable miRNAs (235 miRNAs) in PPP after storage at various temperature: (A) 25 °C, (B) 4 °C, (C) -20 °C, (D) -80 °C. Data are expressed as medians with 95% CIs. Statistical analyses were performed with one-way ANOVA. (*** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$, n=4) Expression of PPP miRNA at 25 °C and 4 °C significantly decreased starting on day 3 and day 14, respectively.

FIG. 14 shows the regression coefficient of miRNA against expression in Log2 copies/mL for (A) serum and (B) PPP at 4 °C. Each marker represents one individual miRNA. hsa-miR-143-3p and hsa-

miR-618 were highlighted because they were chosen to be stability markers, whereas hsa-miR-1290 and hsa-miR-720 were highlighted as reference markers. Stability markers were identified as miRNAs negatively correlated to storage time, decent expression in samples and poses minimal correlation with both haemolysis and platelet. Regression coefficient was obtained from linear regression performed between miRNA expression level in serum or PPP samples and storage duration (days) at 4°C. There is a negative correlation trend between degradation and miRNA expression level in (A) serum and (B) PPP.

FIG. 15 shows (A and B): The distribution of serum regression coefficient against guanine-cytosine (GC) content of miRNA (Pearson $R = 0.1631$, $P = 0.0047$) (A) and length of miRNA sequence (Pearson $R = -0.1215$, $P = 0.0357$) (B). (C and D): The distribution of platelet-poor plasma (PPP) regression coefficient against GC of miRNA (Pearson $R = 0.1812$, $P = 0.0017$) (C) and length of miRNA sequence (Pearson $R = -0.1906$, $P = 0.0009$) (D). Regression coefficient (y axis) was obtained from linear regression performed between miRNA expression level in serum or PPP samples and storage duration (days) at 4°C.

FIG. 16 depicts a schematic workflow of recommended blood collection, processing, and storage conditions. PPP, platelet-poor plasma; RBCs, red blood cells.

20 DETAILED DESCRIPTION OF INVENTION

Circulating nucleic acids such as miRNAs were identified in human plasma and serum as well as in other body fluids such as urine or tears. Such nucleic acids are involved in post-transcriptional gene regulation of cellular processes such as cell differentiation, proliferation, and apoptosis. It was also discovered that circulating nucleic acid expression levels were altered in multiple pathological conditions, forming 'signatures' in various diseases including diabetes, heart diseases, drug-induced liver injury and cancers. Routine biopsies from tissues are too invasive and not practical for the diagnosis and prognosis of diseases. The research direction was then focused on circulating nucleic acid biomarkers.

However, changes in the nucleic acid expression level in the blood might be subtle during the onset of disease and biomarker signal/noise ratio are often low¹. Therefore, preanalytical conditions remain to be the major confounding factors in studies involving circulating nucleic acids such as miRNAs²⁻⁴. All types of blood cells contain large quantities of nucleic acids. Improper blood processing may result in cellular contamination or haemolysis of plasma/serum which renders variation in circulating nucleic acid profile unrelated to pathological conditions⁵. The residual platelet contamination is also a significant confounding source of circulating nucleic acids such as miRNAs⁶⁻⁷. The application of circulating nucleic acid detection and quantification in the clinical labs requires highly reliable and reproducible system. Therefore, quality control of the blood sampling and processing steps become critical.

miRNAs are non-coding RNAs of 21~22 nucleotides long which are synthesized in the cell nucleus. miRNAs have been shown to be relatively stable in storage and therefore the degradation of miRNAs in the sample could be a potentially useful indication to reflect the degradation rate of the nucleic acids present in the sample. By identify stability-related miRNA biomarkers, a panel of biomarkers can be established to provide a measurement of the stability of the nucleic acids in samples to be used for nucleic acid analysis.

The present invention relates to isolated sets of probes suitable for use as a quality control panel in measuring key parameters that are relevant for determining the condition of sample. The isolated sets of probes may comprise the following:

- (a) isolated sets of probes capable of detecting at least one biomarker associated with platelet activation and/or aggregation in a sample;
- (b) isolated sets of probes capable of detecting at least one biomarker associated with haemolysis in a sample; and/or
- (c) isolated sets of probes capable of detecting at least one biomarker associated with nucleic acid degradation in a sample.

The establishment of the above sample quality control panel using haemolysis-related, platelet-related and stability-related miRNAs allow the screening of ideal samples to be used in downstream nucleic acid analysis for diagnostics, research, and measurement. The use of the panel can also be further expanded to study the biasness of different cohorts, which may improve sampling procedures and research quality.

In one example, the present disclosure relates to isolated sets of probes capable of detecting at least one biomarker selected from hsa-miR-1973 and hsa-miR-28-5p for use in determining the level of platelet activation and/or aggregation in a sample. In some examples, the isolated set of probes for use in determining the level of platelet activation and/or aggregation in a sample may further include at least one probe for detecting at least one biomarker selected from the group of hsa-miR-1290 and hsa-miR-10b-5p for use as a reference gene or a normaliser. In some examples, the isolated set of probes for use in determining the level of platelet activation and/or aggregation in a sample may comprise a set of probes for detecting hsa-miR-1973, hsa-miR-28-5p, hsa-miR-1290 and hsa-miR-10b-5p.

In one example, the present disclosure may also relate to isolated sets of probes capable of detecting at least one biomarker selected from hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a for use in determining the level of haemolysis in a sample. In some examples, the isolated set of probes for use in determining the level of haemolysis in a sample may further comprise at least one probe for detecting at least one biomarker selected from the group of hsa-miR-1290 and hsa-miR-10b-5p for use as a reference gene or a normaliser. In some examples, the isolated set of probes for use in

determining the level of haemolysis in a sample may comprise a set of probes for detecting hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-1290 and hsa-miR-10b-5p.

5 In one example, the present disclosure may also relate to isolated sets of probes capable of detecting at least one biomarker selected from hsa-miR-143-3p and hsa-miR-618 for use in determining the level of stability of nucleic acid in a sample, optionally wherein the sample is a stored sample. In some examples, the isolated set of probes for use in determining the stability of nucleic acid in a sample may further comprise at least one probe for detecting at least one biomarker selected from the group of hsa-miR-1290 and hsa-miR-720 for use as a reference gene or a normaliser, optionally wherein
10 the sample is a stored sample. In some examples, the isolated set of probes for use in determining the stability of nucleic acid in a sample may comprise a set of probes for detecting hsa-miR-143-3p, hsa-miR-618, hsa-miR-1290 and hsa-miR-720.

In another example, kits for determining the occurrence of platelet activation and/or aggregation in a sample are described, the kit comprising reagents (for example probes, primers, aptamers, antibodies, buffers and enzymes) for detecting the expression of at least one biomarker selected from hsa-miR-1973 and hsa-miR-28-5p. Such a kit for determining the level of platelet activation and/or aggregation in a sample may further comprise reagents for detecting hsa-miR-1290 and hsa-miR-10b-5p. In another example, the kit for use in determining the level of platelet activation and/or
15 aggregation in a sample comprises reagents for detecting hsa-miR-1973, hsa-miR-28-5p, hsa-miR-1290 and hsa-miR-10b-5p.

In another example, kits for determining the level of haemolysis in a sample are described, the kits comprising reagents (for example probes, primers, aptamers, antibodies, buffers and enzymes) for
25 detecting the expression of at least one biomarker selected from hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a. Such a kit for determining the level of haemolysis in a sample may further comprise reagents for detecting hsa-miR-1290 and hsa-miR-10b-5p. In another example, the kit for use in determining the level of haemolysis in a sample comprises reagents for detecting hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-1290 and hsa-miR-10b-5p.

30 In another example, kits for determining the level of stability of nucleic acid in a sample are described, the kits comprising reagents (for example probes, primers, aptamers, antibodies, buffers and enzymes) for detecting the expression of at least one biomarker selected from the group consisting of hsa-miR-143-3p and hsa-miR-618. Such a kit for determining the stability of nucleic acid in a sample may further comprise reagents for detecting hsa-miR-1290 and hsa-miR-720. In some examples, the
35 kit for use in determining the stability of nucleic acid in a sample comprises reagents for detecting hsa-miR-143-3p, hsa-miR-618, hsa-miR-1290 and hsa-miR-720. The sample described herein may be a stored sample.

- In some examples, the kits as disclosed herein may further comprise instructions on how to use the reagents comprised within the kits. In some examples, the kits may further comprise instructions to use the reagents in methods as described herein. In some examples, the kits may comprise instructions to use the reagents in methods of performing quality control on a sample and/or assessing quality of a sample for nucleic acid analysis. For example, the kits may further comprise instructions to use the reagents in methods of determining the level of platelet activation and/or aggregation in a sample, or in methods of determining the level of haemolysis in a sample, or in methods of determining the level of stability of nucleic acid in a sample.
- 10 In one aspect, there is provided a method of assessing the quality of a sample for use in a nucleic acid analysis, the method comprises determining the level of one or more biomarker in the sample, wherein the one or more biomarker is selected from the group consisting of hsa-miR-28-5p, hsa-miR-1973, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, and hsa-miR-618.
- 15 In some examples, the method may comprise determining the level of two or more, three or more, four or more, five or more, six or more, or all seven biomarkers selected from the group consisting of hsa-miR-28-5p, hsa-miR-1973, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, and hsa-miR-618.
- 20 In some examples, the method of assessing the quality of a sample for use in a nucleic acid analysis involve the analysis of various aspects of a sample as described herein. In some embodiments, the quality of the sample is determined by: a) the level of platelet activation and/or aggregation in the sample, b) the level of haemolysis in the sample, and/or c) the level of stability of nucleic acid in the sample.
- 25 In some embodiments, the quality of the sample is determined by a) the level of platelet activation and/or aggregation in the sample, which in turn is determined by measuring/detecting/determining the level of one or more biomarker selected from the group consisting of hsa-miR-28-5p and hsa-miR-1973. In some embodiments, the method comprises determining the level of one of the following: hsa-miR-28-5p; or hsa-miR-1973; or hsa-miR-1973 and hsa-miR-28-5p.
- 30 In another aspect, there is provided a method of determining the level of platelet activation and/or aggregation in a sample, the method comprises determining the level of one or more biomarker in the sample, wherein the one or more biomarker is selected from hsa-miR-28-5p and hsa-miR-1973.
- 35 In some embodiments, the method comprises determining the level of one of the following: hsa-miR-28-5p; hsa-miR-1973; or hsa-miR-1973 and hsa-miR-28-5p.
- Also provided herein is a method of determining the level of platelet activation and/or aggregation in a sample, the method comprises determining the level of one or more biomarker in the sample, wherein
- 40

the one or more biomarker is selected from hsa-miR-1973 and hsa-miR-28-5p. In some embodiments, the method comprises determining the level of one or two biomarker in the sample, wherein the biomarker is hsa-miR-1973; hsa-miR-28-5p; or hsa-miR-1973 and hsa-miR-28-5p. In some embodiments, the method further comprises determining the level of one or more reference or
 5 normaliser biomarker selected from hsa-miR-1290 and hsa-miR-10b-5p, such as hsa-miR-1290; hsa-miR-10b-5p; or hsa-miR-1290 and hsa-miR-10b-5p. In some embodiments, the method further comprises comparing the level of hsa-miR-1973 and/or hsa-miR-28-5p with that of hsa-miR-1290 and/or hsa-miR-10b-5p to determine the level of platelet activation and/or aggregation in the sample, such as comparing the level of:

- 10 hsa-miR-28-5p with that of hsa-miR-1290;
- hsa-miR-28-5p with that of hsa-miR-10b-5p;
- hsa-miR-28-5p with that of hsa-miR-1290 and hsa-miR-10b-5p;
- hsa-miR-1973 with that of hsa-miR-1290;
- hsa-miR-1973 with that of hsa-miR-10b-5p;
- 15 hsa-miR-1973 with that of hsa-miR-1290 and hsa-miR-10b-5p;
- hsa-miR-1973 and hsa-miR-28-5p with that of hsa-miR-1290;
- hsa-miR-1973 and hsa-miR-28-5p with that of hsa-miR-10b-5p; or
- hsa-miR-1973 and hsa-miR-28-5p with that of hsa-miR-1290 and hsa-miR-10b-5p.

20 In some embodiments, the quality of the sample is determined by b) the level of haemolysis in the sample, which in turn is determined by measure/detecting/determining the level of one or more biomarker selected from the group consisting of hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a. In some embodiments, the method comprises determining the level of one of the following: hsa-miR-20b-5p; hsa-miR-363-3p; hsa-miR-451a; hsa-miR-20b-5p and hsa-miR-363-3p; hsa-miR-20b-5p and
 25 hsa-miR-451a; hsa-miR-363-3p and hsa-miR-451a; or hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a.

In yet another aspect, there is provided a method of determining the level of haemolysis in a sample, the method comprises determining the level of one or more biomarker in the sample, wherein the one
 30 or more biomarker is selected from hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a.

In some embodiments, the method comprises determining the level of one of the following: hsa-miR-20b-5p; hsa-miR-363-3p; hsa-miR-451a; hsa-miR-20b-5p and hsa-miR-363-3p; hsa-miR-20b-5p and
 35 hsa-miR-451a; hsa-miR-363-3p and hsa-miR-451a; or hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a.

Also provided herein is a method of determining the level of haemolysis in a sample, the method comprises determining the level of one or more s in the sample, wherein the one or more biomarker is selected from hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a. In some embodiments, the
 40 method comprises determining the level of one, two or three biomarkers. In some embodiments, the

- method comprises determining the level of hsa-miR-20b-5p; hsa-miR-363-3p; hsa-miR-451a; hsa-miR-20b-5p and hsa-miR-363-3p; hsa-miR-20b-5p and hsa-miR-451a; hsa-miR-363-3p and hsa-miR-451a; or hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a in the sample. In some embodiments, the method further comprises determining the level of one or more reference or normaliser biomarker
- 5 selected from hsa-miR-1290 and hsa-miR-10b-5p, such as hsa-miR-1290; hsa-miR-10b-5p; or hsa-miR-1290 and hsa-miR-10b-5p. In some embodiments, the method further comprises comparing the level of one or more of hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a with that of hsa-miR-1290 and/or hsa-miR-10b-5p to determine the level of haemolysis in the sample, such as comparing the level of:
- 10 hsa-miR-20b-5p with that of hsa-miR-1290;
 hsa-miR-20b-5p with that of hsa-miR-10b-5p;
 hsa-miR-20b-5p with that of hsa-miR-1290 and hsa-miR-10b-5p;
 hsa-miR-363-3p with that of hsa-miR-1290;
 hsa-miR-363-3p with that of hsa-miR-10b-5p;
- 15 hsa-miR-363-3p with that of hsa-miR-1290 and hsa-miR-10b-5p;
 hsa-miR-451a with that of hsa-miR-1290;
 hsa-miR-451a with that of hsa-miR-10b-5p;
 hsa-miR-451a with that of hsa-miR-1290 and hsa-miR-10b-5p.
 hsa-miR-20b-5p and hsa-miR-363-3p with that of hsa-miR-1290;
- 20 hsa-miR-20b-5p and hsa-miR-363-3p with that of hsa-miR-10b-5p;
 hsa-miR-20b-5p and hsa-miR-363-3p with that of hsa-miR-1290 and hsa-miR-10b-5p;
 hsa-miR-20b-5p and hsa-miR-451a with that of hsa-miR-1290;
 hsa-miR-20b-5p and hsa-miR-451a with that of hsa-miR-10b-5p;
 hsa-miR-20b-5p and hsa-miR-451a with that of hsa-miR-1290 and hsa-miR-10b-5p;
- 25 hsa-miR-363-3p and hsa-miR-451a with that of hsa-miR-1290;
 hsa-miR-363-3p and hsa-miR-451a with that of hsa-miR-10b-5p;
 hsa-miR-363-3p and hsa-miR-451a with that of hsa-miR-1290 and hsa-miR-10b-5p;
 hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a with that of hsa-miR-1290;
 hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a with that of hsa-miR-10b-5p; or
- 30 hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a with that of hsa-miR-1290 and hsa-miR-10b-5p.

- In some embodiments, the quality of the sample is determined by c) the level of stability of nucleic acid in the sample, which in turn is determined by measuring/detecting/determining the level of one or
- 35 more biomarker is selected from hsa-miR-143-3p and hsa-miR-618. In some embodiments, the method comprises determining the level of: hsa-miR-143-3p; hsa-miR-618; or hsa-miR-143-3p and hsa-miR-618.

In yet another aspect, there is provided a method of determining the level of stability of nucleic acid in a sample, the method comprises determining the level of one or more biomarker in the sample, wherein the one or more biomarker is selected from hsa-miR-143-3p and hsa-miR-618.

- 5 In some embodiments, the method comprises determining the level of: hsa-miR-143-3p; hsa-miR-618; or hsa-miR-143-3p and hsa-miR-618.

Also provided herein is a method of determining the level of stability of nucleic acid in a sample, the method comprises determining the level of one or more biomarker in the sample, wherein the one or
 10 more biomarker is selected from hsa-miR-143-3p and hsa-miR-618. In some embodiments, the method comprises determining the level of hsa-miR-143-3p; hsa-miR-618; or hsa-miR-143-3p and hsa-miR-618 in the sample. In some embodiments, the method further comprises determining the level of one or more reference or normaliser biomarker selected from hsa-miR-1290 and hsa-miR-720, such as hsa-miR-1290; hsa-miR-720; or hsa-miR-1290 and hsa-miR-720. In some
 15 embodiments, the method further comprises comparing the level of hsa-miR-143-3p and/or hsa-miR-618 with that of hsa-miR-1290 and/or hsa-miR-720 to determine the level of stability of nucleic acid in the sample, such as comparing the level of:

- hsa-miR-143-3p with that of hsa-miR-1290;
- hsa-miR-143-3p with that of hsa-miR-720;
- 20 hsa-miR-143-3p with that of hsa-miR-1290 and hsa-miR-720;
- hsa-miR-618 with that of hsa-miR-1290;
- hsa-miR-618 with that of hsa-miR-720;
- hsa-miR-618 with that of hsa-miR-1290 and hsa-miR-720;
- hsa-miR-143-3p and hsa-miR-618 with that of hsa-miR-1290;
- 25 hsa-miR-143-3p and hsa-miR-618 with that of hsa-miR-720; or
- hsa-miR-143-3p and hsa-miR-618 with that of hsa-miR-1290 and hsa-miR-720.

In some embodiments of the methods, the method further comprises comparing the level of two or more biomarker to determine a score for determining the level of platelet activation and/or
 30 aggregation, stability of nucleic acid or level of haemolysis in a sample. In some embodiments, the method comprises comparing the level of two or more biomarkers, wherein one of the biomarkers is a reference or normaliser biomarker. In some embodiments, the score is compared to that of a control for determining the level of platelet activation and/or aggregation, stability of nucleic acid or level of haemolysis in a sample.

35 Also provided are the methods for performing quality control on a sample and/or assessing quality of a sample for nucleic acid analysis, comprising determining in the sample one or more of the following conditions that may alter the nucleic acid content in the sample:

- (a) platelet activation and/or aggregation;
- 40 (b) haemolysis; and/or

(c) degradation of nucleic acids in the sample.

5 In some embodiments of such methods for performing quality control and/or assessing quality of a sample on a sample for nucleic acid analysis comprises detecting in the sample the level of platelet activation and/or aggregation in the sample by detecting the level of at least one biomarker selected from hsa-miR-1973 and hsa-miR-28-5p compared to a control. In some embodiments, such a method may further include detecting the level of at least one biomarker selected from hsa-miR-1290 and hsa-miR-10b-5p as a reference or normaliser.

10 In some embodiments of such methods for performing quality control on a sample and/or assessing quality of a sample for nucleic acid analysis comprises detecting in the sample the level of haemolysis in the sample by detecting the level of at least one biomarker selected from hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a compared to a control. In some embodiments, such a method may further include detecting the level of at least one biomarker selected from hsa-miR-1290 and hsa-miR-10b-5p as a reference or normaliser.

20 In some embodiments of such methods for performing quality control on a sample and/or assessing quality of a sample for nucleic acid analysis comprises detecting in the sample the level of stability of nucleic acid in the sample by detecting the level of at least one biomarker selected from hsa-miR-143-3p and hsa-miR-618 compared to a control. In some embodiments, such a method may further include detecting the level of at least one biomarker selected from hsa-miR-1290 and hsa-miR-720 as a reference or normaliser.

25 In some embodiments, the method further comprises determining the level of one or more biomarker selected from hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720.

30 In some embodiments, the method further comprises comparing the level of hsa-miR-1973 and/or hsa-miR-28-5p with that of hsa-miR-1290 and/or hsa-miR-10b-5p to determine the level of platelet activation and/or aggregation in the sample.

In some embodiments, the method further comprises comparing the level of one or more of hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a with that of hsa-miR-1290 and/or hsa-miR-10b-5p to determine the level of haemolysis in the sample.

35 In some embodiments, the method further comprises comparing the level of hsa-miR-143-3p and hsa-miR-618 with that of hsa-miR-1290 and/or hsa-miR-720 to determine the level of stability of nucleic acid in the sample.

In some embodiments, the method comprises determining the level of at least one biomarker in each of the following group: (a) hsa-miR-28-5p and hsa-miR-1973; (b) hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a; and (c) hsa-miR-143-3p and hsa-miR-618.

- 5 In some embodiments, the method comprises determining the level of hsa-miR-28-5p, hsa-miR-1973, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p and hsa-miR-618.

In some embodiments, the method comprises determining the level of biomarkers hsa-miR-28-5p, hsa-miR-1973, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, hsa-miR-618, hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720.

10

In some embodiments, such methods for performing quality control on a sample and/or assessing quality of a sample for nucleic acid analysis comprise detecting the presence of at least one of the following conditions in the sample: (a) level of platelet activation and/or aggregation; (b) level of haemolysis; and/or (c) level of stability of nucleic acid, wherein should one or more of the above conditions is identified in the sample, the person with the relevant skill in the art would be able to make the judgement to either: (a) not to proceed with nucleic acid analysis for the sample; (b) to make adjustments in the data analysis for the affected sample to adjust for the biasness of the nucleic acid profile in said sample; or (c) to undertake other corrective measures necessary to derive useful information on the nucleic acid content of the sample.

15

20

In one aspect, the present disclosure relates to methods for performing quality control on a sample and/or assessing quality of a sample for use in nucleic acid analysis, wherein the method comprises determining in the sample the level of one or more biomarkers selected from hsa-miR-1973, hsa-miR-28-5p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, hsa-miR-618, hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720. In some embodiments, the method comprises determining the level of at least one biomarker in each of the following group: (a) hsa-miR-1973 and hsa-miR-28-5p; (b) hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a; and (c) hsa-miR-143-3p and hsa-miR-618. In some embodiments, the method comprises determining the level of at least one biomarker in each group, such as least 3, 4, 5, 6 or 7 biomarkers.

25

30

For the avoidance of doubt, while there are two to three biomarkers listed for determining each of the conditions such as the level of platelet activation and/or aggregation, level of haemolysis and level of stability of nucleic acid, it is understood by the person skilled in the art that the use of one or any suitable number of biomarker may be sufficient for determining each condition. Similarly, the use of one or any suitable number of reference or normaliser biomarker for each condition may also be sufficient.

35

In some embodiments, the method for performing quality control on a sample and/or assessing quality of a sample for use in nucleic acid analysis comprises determining the level of one or more of the following combination of biomarkers:

- | | |
|----|--|
| | hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-20b-5p; |
| 5 | hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-363-3p; |
| | hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-451a; |
| | hsa-miR-28-5p, hsa-miR-618, hsa-miR-20b-5p; |
| | hsa-miR-28-5p, hsa-miR-618, hsa-miR-363-3p; |
| | hsa-miR-28-5p, hsa-miR-618, hsa-miR-451a; |
| 10 | hsa-miR-1973, hsa-miR-143-3p, hsa-miR-20b-5p; |
| | hsa-miR-1973, hsa-miR-143-3p, hsa-miR-363-3p; |
| | hsa-miR-1973, hsa-miR-143-3p, hsa-miR-451a; |
| | hsa-miR-1973, hsa-miR-618, hsa-miR-20b-5p; |
| | hsa-miR-1973, hsa-miR-618, hsa-miR-363-3p; |
| 15 | hsa-miR-1973, hsa-miR-618, hsa-miR-451a; |
| | hsa-miR-1973, hsa-miR-143-3p, hsa-miR-618, hsa-miR-20b-5p; |
| | hsa-miR-1973, hsa-miR-143-3p, hsa-miR-618, hsa-miR-363-3p; |
| | hsa-miR-1973, hsa-miR-143-3p, hsa-miR-618, hsa-miR-451a; |
| | hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-618, hsa-miR-20b-5p; |
| 20 | hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-618, hsa-miR-363-3p; |
| | hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-618, hsa-miR-451a; |
| | hsa-miR-1973, hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-20b-5p; |
| | hsa-miR-1973, hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-363-3p; |
| | hsa-miR-1973, hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-451a; |
| 25 | hsa-miR-1973, hsa-miR-28-5p, hsa-miR-618, hsa-miR-20b-5p; |
| | hsa-miR-1973, hsa-miR-28-5p, hsa-miR-618, hsa-miR-363-3p; |
| | hsa-miR-1973, hsa-miR-28-5p, hsa-miR-618, hsa-miR-451a; |
| | hsa-miR-1973, hsa-miR-143-3p, hsa-miR-20b-5p, hsa-miR-363-3p; |
| | hsa-miR-1973, hsa-miR-143-3p, hsa-miR-20b-5p, hsa-miR-451a; |
| 30 | hsa-miR-1973, hsa-miR-143-3p, hsa-miR-363-3p, hsa-miR-451a; |
| | hsa-miR-1973, hsa-miR-618, hsa-miR-20b-5p, hsa-miR-363-3p; |
| | hsa-miR-1973, hsa-miR-618, hsa-miR-20b-5p, hsa-miR-451a; |
| | hsa-miR-1973, hsa-miR-618, hsa-miR-363-3p, hsa-miR-451a; |
| | hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-20b-5p, hsa-miR-363-3p; |
| 35 | hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-20b-5p, hsa-miR-451a; |
| | hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-363-3p, hsa-miR-451a; |
| | hsa-miR-28-5p, hsa-miR-618, hsa-miR-20b-5p, hsa-miR-363-3p; |
| | hsa-miR-28-5p, hsa-miR-618, hsa-miR-20b-5p, hsa-miR-451a; |
| | hsa-miR-28-5p, hsa-miR-618, hsa-miR-363-3p, hsa-miR-451a; |
| 40 | hsa-miR-1973, hsa-miR-143-3p, hsa-miR-618, hsa-miR-20b-5p, hsa-miR-363-3p; |

hsa-miR-1973, hsa-miR-143-3p, hsa-miR-618, hsa-miR-20b-5p, hsa-miR-451a;
 hsa-miR-1973, hsa-miR-143-3p, hsa-miR-618, hsa-miR-363-3p, hsa-miR-451a;
 hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-618, hsa-miR-20b-5p, hsa-miR-363-3p;
 hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-618, hsa-miR-20b-5p, hsa-miR-451a;
 5 hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-618, hsa-miR-363-3p, hsa-miR-451a;
 hsa-miR-1973, hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-20b-5p, hsa-miR-363-3p;
 hsa-miR-1973, hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-20b-5p, hsa-miR-451a;
 hsa-miR-1973, hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-363-3p, hsa-miR-451a;
 hsa-miR-1973, hsa-miR-28-5p, hsa-miR-618, hsa-miR-20b-5p, hsa-miR-363-3p;
 10 hsa-miR-1973, hsa-miR-28-5p, hsa-miR-618, hsa-miR-20b-5p, hsa-miR-451a;
 hsa-miR-1973, hsa-miR-28-5p, hsa-miR-618, hsa-miR-363-3p, hsa-miR-451a;
 hsa-miR-1973, hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-618, hsa-miR-20b-5p;
 hsa-miR-1973, hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-618, hsa-miR-363-3p;
 hsa-miR-1973, hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-618, hsa-miR-451a;
 15 hsa-miR-1973, hsa-miR-143-3p hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a;
 hsa-miR-1973, hsa-miR-618, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a;
 hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a;
 hsa-miR-28-5p, hsa-miR-618, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a;
 hsa-miR-1973, hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-618, hsa-miR-20b-5p, hsa-miR-
 20 363-3p;
 hsa-miR-1973, hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-618, hsa-miR-20b-5p, hsa-miR-
 451a;
 hsa-miR-1973, hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-618, hsa-miR-363-3p, hsa-miR-
 451a;
 25 hsa-miR-1973, hsa-miR-143-3p, hsa-miR-618, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-
 451a;
 hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-618, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-
 451a;
 hsa-miR-1973, hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-
 30 451a;
 hsa-miR-1973, hsa-miR-28-5p, hsa-miR-618, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-
 451a; and
 hsa-miR-1973, hsa-miR-28-5p, hsa-miR-143-3p, hsa-miR-618, hsa-miR-20b-5p, hsa-miR-
 363-3p, hsa-miR-451a.

35

In some embodiments of the method, the method further comprises determining the level of one or more reference or normaliser biomarker selected from hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720. In some embodiments, the method further comprises determining the level of one, two or three biomarkers, such as hsa-miR-1290; hsa-miR-10b-5p; hsa-miR-720; hsa-miR-1290 and hsa-miR-10b-

5p; hsa-miR-1290 and hsa-miR-720; hsa-miR-10b-5p and hsa-miR-720; or hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720.

5 In some embodiments, the method comprises determining the expression level of hsa-miR-1973, hsa-miR-28-5p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, hsa-miR-618, hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720.

10 In some embodiments, the method further comprises comparing the level of two or more biomarkers to determine a score for determining the quality of a sample. In some embodiments, the method comprises comparing the level of two or more biomarkers, wherein one of the biomarkers is a reference or normaliser biomarker. In some embodiments, the score is compared to that of a control for determining the quality of the sample.

15 In some embodiments, the method comprises comparing the level of two or more biomarkers to determine a score, wherein the score is compared to that of a control for determining the quality of the sample.

20 In some embodiments, the sample is a blood sample or a non-cellular bodily fluid sample obtained from a subject. In some examples, the sample is a plasma and/or serum.

25 In yet another aspect, there is provided a kit for use in the method as disclosed herein, the kit comprises an isolated set of probes and/or reagents capable of determining the expression level of the one or more biomarker selected from the group consisting of hsa-miR-1973, hsa-miR-28-5p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, and hsa-miR-618.

30 Also disclosed are kits for use in performing quality control on a sample for nucleic acid analysis comprising reagents for detecting the expression of at least one biomarker as described herein. In some examples, the kit comprises reagents for detecting the expression of one or more biomarker selected from the group consisting of hsa-miR-1973, hsa-miR-28-5p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p and hsa-miR-618. In some embodiments, the kit further comprises reagents for detecting the expression of at least one biomarker selected from the group consisting of hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720. In some embodiments, such reagents may include probes, primers, aptamers, antibodies, buffers and enzymes.

35 In another aspect, disclosed herein are kits for use in any of the methods, the kit comprises an isolated set of probes and/or reagents capable of determining the level of the one or more biomarkers of the preceding claims.

Compositions for use in the methods disclosed herein include, but are not limited to, probes, antibodies, affibodies, nucleic acids, and/or aptamers. Preferred compositions can detect the level of expression (e.g., miRNA) of a panel of biomarkers from a biological sample.

5 Any of the compositions can be provided in the form of a kit or a reagent mixture. By way of an example, labelled probes can be provided in a kit for the detection of a panel of biomarkers. Kits can include all components necessary or sufficient for assays, which can include, but is not limited to, detection reagents (e.g., probes), buffers, control reagents (e.g., positive and negative controls), amplification reagents, solid supports, labels, instruction manuals, etc. In some embodiments, the kit
10 comprises a set of probes for the panel of biomarkers and a solid support to immobilize the set of probes. In some embodiments, the kit comprises a set of probes for the panel of biomarkers, a solid support, and reagents for processing the sample to be tested (e.g., reagents to isolate the protein or nucleic acids from the sample).

15 In some examples, the kits as disclosed herein may further comprise instructions for the reagents to be used for performing quality control on a sample and/or assessing quality of a sample for use in nucleic acid analysis. Therefore, the kits may further comprise instructions for using the reagents to assess the presence of at least one of the following conditions in the sample: (a) level of platelet activation and/or aggregation; (b) level of haemolysis; and/or (c) level of stability of nucleic acid.

20 In some embodiments, the method of assessing the quality of a sample for use in a nucleic acid analysis, the method comprises determining the level of hsa-miR-1973 and at least one, at least two, at least three, at least four, at least five, at least six biomarkers selected from hsa-miR-28-5p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, and hsa-miR-618.

25 In some embodiments, the method of assessing the quality of a sample for use in a nucleic acid analysis, the method comprises determining the level of hsa-miR-28-5p and at least one, at least two, at least three, at least four, at least five, at least six biomarkers selected from hsa-miR-1973, hsa-miR-28-5p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, and hsa-miR-618.

30 In some embodiments, the method of assessing the quality of a sample for use in a nucleic acid analysis, the method comprises determining the level of hsa-miR-20b-5p and at least one, at least two, at least three, at least four, at least five, at least six biomarkers selected from hsa-miR-1973, hsa-miR-28-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, and hsa-miR-618.

35 In some embodiments, the method of assessing the quality of a sample for use in a nucleic acid analysis, the method comprises determining the level of hsa-miR-363-3p and at least one, at least two, at least three, at least four, at least five, at least six biomarkers selected from hsa-miR-1973, hsa-miR-28-5p, hsa-miR-20b-5p, hsa-miR-451a, hsa-miR-143-3p, and hsa-miR-618.

40

In some embodiments, the method of assessing the quality of a sample for use in a nucleic acid analysis, the method comprises determining the level of hsa-miR-451a and at least one, at least two, at least three, at least four, at least five, at least six biomarkers selected from hsa-miR-1973, hsa-miR-28-5p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-143-3p, and hsa-miR-618.

5

In some embodiments, the method of assessing the quality of a sample for use in a nucleic acid analysis, the method comprises determining the level of hsa-miR-143-3p and at least one, at least two, at least three, at least four, at least five, at least six biomarkers selected from hsa-miR-1973, hsa-miR-28-5p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a and hsa-miR-618.

10

In some embodiments, the method of assessing the quality of a sample for use in a nucleic acid analysis, the method comprises determining the level of hsa-miR-618 and at least one, at least two, at least three, at least four, at least five, at least six biomarkers selected from hsa-miR-1973, hsa-miR-28-5p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a and hsa-miR-143-3p.

15

Also provided herein are methods for collecting a biological sample (e.g., blood) for nucleic acid analysis, such methods comprises: (a) collecting an initial first volume of biological sample; and subsequently, (b) collecting a second volume of the biological sample in a separate container for use in nucleic acid analysis. For the avoidance of doubt, the initial first volume and second volume of biological samples are not to be combined for nucleic acid analysis and preferentially, be collected in discrete containers. As examples, the containers may be, but not limited to, tubes, flasks, bags and compartments within containers. In one aspect, the initial first volume of the biological sample may be discarded and the volume to be discarded may be at least 1 millilitre (mL). In another aspect, the initial first volume of sample to be discarded may be approximately 1-2mL. In some embodiments, such methods of sample collection may reduce haemolysis and/or contamination (e.g., from skin cell, microorganisms, etc) that may alter the nucleic acid content of the sample. It would be understood by someone skilled in the art that the collection of the initial first volume to be discarded may comprise more than one sample draw, provided that the objective of each sample draw for the initial first volume is performed with the intent of obtaining a higher quality subsequent sample for nucleic acid analysis (i.e. in the second volume). It would also be understood by someone skilled in the art that "discarding" the first sample shall refer to the non-use of said initial sample for nucleic acid analysis. However, such initial samples may be used for purposes other than nucleic acid analysis, particularly for the analysis of circulating nucleic acids (e.g., miRNAs).

20

25

30

Also disclosed herein are methods of collecting a sample for nucleic acid analysis, the method may comprise (a) collecting an initial first volume of the sample; and (b) further collecting a second volume of the sample in a separate container for use in nucleic acid analysis.

35

Also disclosed herein are methods of processing a sample for nucleic acid analysis, the method comprises: (a) subjecting the sample to a first centrifugation at a speed of from 1,000g to 5,000g for a first period of time to provide a first supernatant; and (b) subjecting the supernatant obtained from (a)

40

to a second centrifugation at a speed of from 1,500g to 5,000g for a second period of time to provide a second supernatant.

Also provided herein are methods of processing a biological sample for nucleic acid analysis, such methods comprising at least two (2) centrifugation steps to provide a first and second supernatants, respectively. In some embodiments, the first centrifugation may be carried out at a speed of from 1,000g to 5,000g, such as from 1,000g to 4,000g, such as from 1,000g to 2,000g, such as 1,500g. In another embodiment, the second centrifugation would be carried out at a speed of from 1,500g to 5,000g, such as from 1,500g to 4,000g, such as from 1,500g to 3,000g, such as 2,500g. In another embodiment, the centrifugation time for the first and second centrifugation may be independently selected from 1 min to 30 min, such as from 5 min to 20 min, such as 15 min.

In a further embodiment, the methods of processing a biological sample for nucleic acid analysis may include cryopreservation of the first supernatant after the first centrifugation, with the second centrifugation being carried out following the thawing of the cryopreserved first supernatant (i.e., after freeze/thaw).

Also disclosed herein are methods of collecting and processing a biological sample for nucleic acid analysis comprising: (a) collecting an initial first volume of the biological sample; (b) further collecting a second volume of the biological sample in a separate container; (c) subjecting the second volume of the biological sample to a first centrifugation at a speed of from 1,000g to 5,000g for a first period of time to provide a first supernatant; and (d) subjecting the first supernatant obtained from (c) to a second centrifugation at a speed of from 1,500g to 5,000g for a second period of time to provide a second supernatant. In some examples, the second centrifugation may be performed after at least one freeze/thaw of the sample after the first centrifugation. In some examples, the first and second period of time may be independently selected from 1 min to 30 min, such as from 5 min to 20 min, such as 15 min. In some examples, the method may further comprise: determining in the second supernatant the expression level of one or more miRNAs selected from hsa-miR-1973, hsa-miR-28-5p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, hsa-miR-618, hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720; and optionally determining in the second supernatant the expression level of one or more miRNAs selected from hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720.

In a further aspect, the method of collecting and processing a biological sample (e.g. blood) for nucleic acid comprising: (a) collecting an initial first volume of biological sample; (b) collecting a second volume of biological sample in a separate container; (c) subjecting the second volume of the biological sample to a first centrifugation to provide a first supernatant; and (d) subjecting the first supernatant to a second centrifugation to provide a second supernatant. In some embodiments, the first centrifugation may be carried out at a speed of from 1,000g to 5,000g, such as from 1,500g to 4,000g, such as from 1,500g to 3,000g, such as 1,500g. In another embodiment, the second

centrifugation may be carried out at a speed of 1,500g to 5,000g, such as from 1,500g to 4,000g, such as from 1,500g to 3,000g, such as 2,500g. In another embodiment, the centrifugation time for the first and second centrifugation may be independently selected from 1 min to 30 min, such as from 5 min to 20 min, such as 15 min. In some embodiments, the first supernatant may be cryopreserved after the first centrifugation, with the second centrifugation performed following the thawing of the cryopreserved first supernatant. In some embodiments, the method may further comprise determining in the second supernatant, the level of at least one biomarker selected from hsa-miR-1973, hsa-miR-28-5p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, hsa-miR-618, hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720.

In the embodiments of any of the aspects, the biomarker may be microRNA (miRNA). As used herein, the term “miRNA” or “microRNA” or “miR” refers to an RNA (or RNA analogue) comprising the product of an endogenous, non-coding gene whose precursor RNA transcripts can form small stem-loops from which mature “miRNAs” are cleaved by the endonuclease Dicer. MiRNAs are encoded in genes distinct from the mRNAs whose expression they control. In one example, the term “miRNA” or “microRNA” refers to single-stranded RNA molecules of at least 10 nucleotides and of not more than 35 nucleotides covalently linked together. In one example, the polynucleotides are molecules of 10 to 33 nucleotides in length, or of 15 to 30 nucleotides in length, or of 17 to 27 nucleotides in length, or of 18 to 26 nucleotides in length, i.e., 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 28, 29, 30, 31, 32, 33, 34, or 35 nucleotides in length, not including, optionally, labels and/or elongated sequences (e.g., biotin stretches). The sequences of the miRNAs as described herein are provided in Table 1. As would be appreciated by the skilled person in the art, miRNA is a type of polynucleotide that has sequences comprising letters such as “AUGC.” It will be understood that the nucleotides are in 5'→3' order from left to right and that “A” denotes adenosine, “U” denotes uracil, “G” denotes guanosine, and “C” denotes cytosine, unless otherwise noted. The letters A, U, G, and C can be used to refer to the base themselves, to the nucleosides, or to the nucleotides comprising the bases, as is standard in the art.

Table 1. Sequences of the miRNA biomarkers described herein (Accession numbers from miRbase).

miRNA	Accession No.	Sequence	SEQ ID NO
hsa-miR-1973	MIMAT0009448	ACCGUGCAAAGGUAGCAUA	1
hsa-miR-28-5p	MIMAT0000085	AAGGAGCUCACAGUCUAUUGAG	2
hsa-miR-20b-5p	MIMAT0001413	CAAAGUGCUCAUAGUGCAGGUAG	3
hsa-miR-363-3p	MIMAT0000707	AAUUGCACGGUAUCCAUCUGUA	4
hsa-miR-451a	MIMAT0001631	AAACCGUUACCAUACUGAGUU	5
hsa-miR-143-3p	MIMAT0000435	UGAGAUGAAGCACUGUAGCUC	6
hsa-miR-618	MIMAT0003287	AAACUCUACUUGUCCUUCUGAGU	7
hsa-miR-1290	MIMAT0005880	UGGAUUUUUGGAUCAGGGA	8
hsa-miR-10b-5p	MIMAT0000254	UACCCUGUAGAACCGAAUUUGUG	9
hsa-miR-720	Not available	UCUCGCUGGGGCCUCCA	10

In some aspect, provided herein are DNA-, RNA-, and protein-based detection methods that either directly or indirectly detect the biomarkers described herein. The present invention also provides compositions, reagents, and kits for such diagnostic purposes. The diagnostic methods described

5 herein may be qualitative or quantitative. Quantitative diagnostic methods may be used, for example, to compare a detected biomarker level to a cut-off or threshold level. Where applicable, qualitative or quantitative diagnostic methods can also include amplification of target, signal, or intermediary. In some embodiments of the above aspects, the one or more biomarkers are detected/determined by methods selected from sequencing, nucleic acid hybridisation, microarray and nucleic acid

10 amplification such as a reverse transcription-quantitative polymerase chain reaction (RT-qPCR), reverse transcription-polymerase chain reaction (RT-PCR), quantitative polymerase chain reaction (qPCR), a locked nucleic acid PCR, a clustered regularly interspaced short palindromic repeat (CRISPR)-based assay, or isothermal amplification assay.

15 It is known in the art that, biomarkers can be combined to form a biomarker panel and a predictive score may be generated, for example using a linear model. From such a score, the cut-off threshold to determine if the quality of the sample is acceptable for further analysis may be derived. An example would be to calculate such a risk score using logistic regression, a form of linear model. The prediction score may also be calculated using a classification algorithm selected from the group

20 comprising support vector machine algorithm, logistic regression algorithm, multinomial logistic regression algorithm, Fisher's linear discriminant algorithm, quadratic classifier algorithm, perceptron algorithm, k-nearest neighbours algorithm, artificial neural network algorithm, random forests algorithm, decision tree algorithm, naive Bayes algorithm, adaptive Bayes network algorithm, and ensemble learning method combining multiple learning algorithms. The challenge in the field pertains

25 to identifying relevant biomarkers, such as the relevant stability related biomarkers, that could be reliably applied to identify a sample in which the nucleic acid has degraded and is not useful for downstream analysis. Where relevant biomarkers could be identified via exhaustive and well-designed studies, it would be within the skill of someone aware of the state of the art to apply the measured level of the relevant biomarkers in such statistical models to generate such a score.

30 Examples of such mathematical methods used to perform the calculations disclosed herein, for example, the calculation of a prediction score, can be, but are not limited to, support vector machine algorithm, logistic regression algorithm, multinomial logistic regression algorithm, Fisher's linear discriminant algorithm, quadratic classifier algorithm, perceptron algorithm, k-nearest neighbours algorithm, artificial neural network algorithm, random forests algorithm, decision tree algorithm, naive

35 Bayes algorithm, adaptive Bayes network algorithm, and ensemble learning method combining multiple learning algorithms. In one example, the calculation of the prediction score may be calculated using linear models and support vector machine algorithms.

The invention illustratively described herein may suitably be practiced in the absence of any element

40 or elements, limitation or limitations, not specifically disclosed herein. Thus, for example, the terms

"comprising", "including", "containing", etc. shall be 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
5 modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the 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.

10

The invention has been described broadly and generically herein. Each of the 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
15 herein.

Throughout this disclosure, certain embodiments may be disclosed in a range format. It should be understood that the description in range format is merely for convenience and brevity and should not be construed as an inflexible limitation on the scope of the disclosed ranges. Accordingly, the
20 description of a range should be considered to have specifically disclosed all the possible sub-ranges as well as individual numerical values within that range. For example, description of a range such as from 1 to 6 should be considered to have specifically disclosed sub-ranges such as from 1 to 3, from 1 to 4, from 1 to 5, from 2 to 4, from 2 to 6, from 3 to 6 etc., as well as individual numbers within that range, for example, 1, 2, 3, 4, 5, and 6. This applies regardless of the breadth of the range.

25

Other embodiments are within the following statements, claims and non-limiting examples. In 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.

30

An exemplary embodiment of the present invention is provided in the following example. While the exemplary embodiment is described in sufficient detail to enable those skilled in the art to practice the invention, it should be understood that the present invention is not limited to this example.

35 **EXAMPLES**

Example 1: Identification of platelet-related and red blood cell-related miRNA

Platelet-related miRNA as major confounding factor in miRNA profile in non-haemolysed samples

40

To examine the changes of miRNA expression level in blood samples under different processing methods, serum samples with clotting duration of 30 min, 60 min and 90 min (serum 30, serum 60 and serum 90, respectively), plasma and Plate-let-poor-plasma (PPP) were examined. PPP is plasma samples which went through double-spin to eliminate platelets. Blood from 10 healthy subjects were collected for this study. Human Haemoglobin ELISA was used to measure the level of cell-free haemoglobin in samples as an indicator of haemolysis. The haemoglobin concentration remained low and did not vary in the serum after different clotting duration. Both plasma and PPP showed higher haemoglobin concentration than the serum samples, but the results were similar to each other (FIG. 1A). As a haemoglobin level of below 500 µg/mL was considered acceptable, no significant haemolysis was observed for those samples. Thromboxane B2 (TXB2) ELISA was used to measure the level of TXB2 in serum as an indicator of platelet activation and aggregation. The TXB2 level increased when serum clotting duration increased, (e.g. the TXB2 level correlated to serum clotting duration up to 90 mins), indicating that more platelets were activated in longer time. Both plasma and PPP showed low TXB2 concentration as platelet aggregation and activation did not take place (FIG. 1B). The miRNA expression level varied in different sample types (FIG. 1C). Interestingly the average expression of detectable miRNAs (i.e. 356 miRNAs) correlated with TXB2 concentration in serum samples (Pearson $r = 0.6425$, $P = 0.00013$) (FIG. 1D). This finding is consistent with the results reported by other studies, highlighting that activated platelets secrete miRNA during the coagulation process. The miRNA expression was highest in plasma because the large amount of residue platelets inside plasma. PPP had a low mean miRNA expression compared with plasma, serum 60 and serum 90, but closer to serum 30. One-way analysis of variance was performed on haemoglobin, TXB2, and miRNA expression level from different sample type. Results are shown in Table 2.

Table 2. One-way ANOVA between sample types

Comparison	Haemoglobin	Thromboxane B2	miRNA level
	<i>p</i> value	<i>p</i> value	<i>p</i> value
Serum30 vs. Serum60	0.9793	0.719	0.0002
Serum30 vs. Serum90	0.5613	0.0007	<0.0001
Serum30 vs. Plasma	0.0074	0.8642	<0.0001
Serum30 vs. PPP	0.0675	0.8878	0.5951
Serum60 vs. Serum90	0.8781	0.0257	0.2667
Serum60 vs. Plasma	0.0014	0.1865	0.0024
Serum60 vs. PPP	0.0163	0.2076	<0.0001
Serum90 vs. Plasma	<0.0001	<0.0001	0.3239
Serum90 vs. PPP	0.001	<0.0001	<0.0001
Plasma vs. PPP	0.9118	>0.9999	<0.0001

The similarity of the miRNA profiles in different sample types was further investigated using principal component analysis (PCA). The PCA showed that after undergoing global normalization, the miRNA profiles were clustered by sample type. There were similarities in the miRNA profiles of plasma,

Serum 60 and Serum 90. However, the miRNA profile in PPP was closer to Serum 30 (FIG. 2A, C). Unlike the sample types, the individual PCAs found no clustering (FIG. 2B, D). This finding implied a disparate miRNA expression level in different sample types even within the same individual. Therefore, the pre-analytical and sample types of consideration highly impact the reliability of miRNA expression as a diagnostic or research tool because individual signatures may be easily masked by cofounding variables.

Platelet-related miRNA

The miRNA expressions in serum 30, serum 60 and serum 90 were profiled and analysed. Linear regression was performed to find correlation between miRNAs and platelet activation from a short to long clotting duration. The miRNAs that had P value < 0.05 were identified as biomarkers which are strongly correlated with platelet activation. Of 356 miRNAs, two miRNAs, hsa-miR-1973 and hsa-miR-28-5p, were chosen as platelet-related markers due to the strong correlation with platelet activation and reasonable expression in serum. (FIG. 3A, B). The expression of hsa-miR-1973 and hsa-miR-28-5p in serum samples correlated well with the level of TXB2 (Pearson $r = 0.5693$ and 0.5917 , $P = 0.00102$ and 0.0006 , respectively). Plasma samples showed the highest expression of these two miRNAs, while the expression in PPP remained the lowest. Two miRNAs which showed the least correlation with TXB2, hsa-miR-1290 and hsa-miR-10b-5p (Pearson $r = 0.0049$ and -0.2061 , $P = 0.9796$ and 0.2745 , respectively), were chosen to be reference gene for platelet-related markers (FIG. 3C, D). No correlation between these markers and TXB2 was observed. Different representations of the expression levels of platelet-related markers, reference markers, and TXB2 level in different sample types are as shown in FIG. 4A-E.

One-way analysis of variance was performed to compare the expression levels of these markers in different samples. The P values are given in Table 3. Both hsa-miR-1973 and hsa-miR-28-5p had significant differences in plasma and PPP samples, whereas hsa-miR-1290 and hsa-miR-10b-5p did not have differences. In a separate study, the expression of these miRNAs was also analyzed in serum (with 60 minutes of clotting), plasma, and PPP. The same trend was seen in the platelet-related markers and reference markers (FIG. 5A-D).

Table 3. Comparison of miRNAs and TXB2 among different sample types

Comparison	P values				
	miR-1973	miR-28-5p	miR-1290	miR-10b-5p	TXB2
Serum 30 versus serum 60	<0.0001	<0.0001	0.117	0.7238	0.1115
Serum 30 versus serum 90	<0.0001	<0.0001	0.3082	0.9999	0.0397
Serum 30 versus plasma	<0.0001	<0.0001	0.141	0.9824	0.1444
Serum 30 versus PPP	0.0444	0.0081	0.1015	0.798	0.255
Serum 60 versus serum 90	0.0035	0.1125	0.9789	0.976	0.1077
Serum 60 versus plasma	<0.0001	0.0026	0.0046	0.3641	0.0107

Comparison	P values				
	miR-1973	miR-28-5p	miR-1290	miR-10b-5p	TXB2
Serum 60 versus PPP	<0.0001	<0.0001	0.0045	>0.9999	0.0077
Serum 90 versus plasma	0.0001	0.0402	0.0244	0.9962	0.0168
Serum 90 versus PPP	<0.0001	<0.0001	0.0167	0.9904	0.0179
Plasma versus PPP	<0.0001	<0.0001	0.6852	0.2308	0.988

PPP, platelet-poor plasma; TXB2, thromboxane B2.

RBC-related biomarkers

- 5 Using the above samples from ten individual subjects, hemolyzed samples were generated by spiking red blood cell lysate into PPP at the concentration of 0 % v/v (no hemo), 0.05 % v/v (hemo 1), 0.2 % v/v (hemo 2) and 1 % v/v (hemo 3). The difference of expression of 277 detectable miRNAs in red blood cell lysate (RBC) and spiked PPP is plotted in **FIG. 7A**. Majority of the miRNAs had a positive correlation with the concentration of RBC lysate (i.e. increased in expression when the concentration of RBC lysate increased). To find the most prominent markers related to haemolysis, the fold change between 'no hemo' and 'RBC' were analysed. A higher fold change indicated larger increase of miRNA expression contributed by the RBC compared to background PPP. The miRNAs which have fold change larger than 8 in RBC lysate compared to PPP and had good expression in all samples were shortlisted. Hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a were identified as RBC-related biomarkers from these criteria (**FIG. 7B**). Hsa-miR-1290 and hsa-miR-10b-5p were used as reference genes as the expression in PPP was higher than in RBC. A histogram with the SD of RBC-related miRNAs is given in **FIG. 6**. One-way analysis of variance was performed to determine the significance of these markers (**Table 4**).
- 20 Platelet-related and RBC-related miRNAs can be used to evaluate the quality of processed plasma samples. Expression of the markers was normalized against reference genes during data analysis to generate a haemolysis score and a platelet score. The geometric mean was used to calculate the aggregated value for two platelet-related miRNAs, three RBC-related miRNAs and the two reference genes. To be more specific, the haemolysis score was calculated by normalizing the expression of RBC-related miRNAs against the reference gene ($Cq_{RBC} - Cq_{norm}$), whereas the platelet score was calculated by normalizing the expression of platelet-related miRNA against the reference gene ($Cq_{platelet} - Cq_{norm}$). Higher scores indicated that samples had lesser miRNA contamination from RBCs or platelets.

30 **Table 4.** One-way ANOVA between RBC, non-spiked PPP and spiked PPP

Comparison	hsa-miR-20b-5p	hsa-miR-363-3p	hsa-miR-451a	hsa-miR-1290	hsa-miR-10b-5p
	p value	p value	p value	p value	p value
RBC vs. No hemo	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

Comparison	hsa-miR-20b-5p	hsa-miR-363-3p	hsa-miR-451a	hsa-miR-1290	hsa-miR-10b-5p
	p value	p value	p value	p value	p value
RBC vs. Hemo1	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
RBC vs. Hemo2	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
RBC vs. Hemo3	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
No hemo vs. Hemo1	<0.0001	<0.0001	0.0002	0.9997	0.9987
No hemo vs. Hemo2	<0.0001	<0.0001	<0.0001	0.9911	0.6986
No hemo vs. Hemo3	<0.0001	<0.0001	<0.0001	0.9268	0.0881
Hemo1 vs. Hemo2	<0.0001	0.0005	0.0097	0.9689	0.5391
Hemo1 vs. Hemo3	<0.0001	<0.0001	<0.0001	0.8605	0.0537
Hemo2 vs. Hemo3	<0.0001	<0.0001	<0.0001	0.9962	0.5875

Example 2: Plasma Processing consideration

To systematically establish a reliable blood processing workflow, the effects of the stepwise processing of PPP were calculated using the haemoglobin level, haemolysis score and platelet score. The various considerations included the size of the blood tubes, the centrifugation protocol, post freeze-thaw spinning and whole blood storage.

Tube size

An often-neglected aspect of the tube choice is the size of the tube. The varying sizes have different heights, which lead to different travelling distances for the blood cells to reach the bottom of the tubes during centrifugation. We wanted to investigate if the red blood cells and platelets removal were similar in different tube sizes.

Blood was collected from four healthy donors, each in both 6-mL and 10-mL tubes. A standardized centrifugation protocol of 1500 g for 15 min was applied on all samples. The initial 1 to 2 ml of blood for each blood draw was discarded to avoid haemolysis. Haemoglobin ELISA was performed to assess the haemolysis level in the samples. Plasma from both 6-mL and 10-mL tubes was tested to investigate the impact of tube sizes. The discard tubes (containing the first drawn blood) from two of the subjects were also tested. The result showed that haemoglobin concentration was significantly higher in the discard tubes compared with subsequent tubes ($p < 0.05$) (**FIG. 8A**). This result supports the use of an initial discard tube for miRNA analysis to avoid blood haemolysis. There were no significant differences in the haemoglobin levels between the 6-mL tubes and 10-mL tubes.

Haemolysis score was calculated by normalizing the expression of RBC-related miRNAs against reference gene ($Cq_{RBC} - Cq_{norm}$). Platelet score was calculated by normalising the expression of platelet-related miRNA against reference genes ($Cq_{plate} - Cq_{norm}$). Higher scores indicate samples with

less miRNA contamination from red blood cells or platelets. There was no significant difference in haemolysis score for plasma collected in different tube sizes (FIG. 8B). Similarly, there was no significant difference in platelet score for plasma collected in different tube sizes (FIG. 8C).

5 Centrifugation protocol

As the size of the blood tubes did not affect sample quality, the PPP quality obtained using different centrifugation protocols was examined. To generate PPP, the 2-step centrifugation protocols were tested with blood samples from five healthy individuals. Lower speed (1000 g or 1500 g) was used for the first spin in the original vacutainer. After the first spin, the plasma was aspirated out into 15-mL tubes. For the second spin, different spin speeds were experimented, varying from 1500 g to 5000 g. The centrifugation speeds and durations were demonstrated in Table 5.

The haemolysis score indicated a slight decrease in lower first spin (Con5 and Con6), although this was not significant (FIG. 9A). This indicates a higher first spin is slightly more effective in removing red blood cells. However, the second spin speed did not affect the haemolysis score (Con1 to Con4). The platelet score, however, was more affected by the second spin speed (FIG. 9B). A slight upward trend was observed when the second spin speed increased, regardless of first spin speed. However, Con4, which had the highest second spin speed at 5000 g showed the lowest platelet score, with a statistically significant difference compared with Con1. This trend was deduced to be due to the lysis of platelets under a high spin speed. It was concluded that Con2 or Con3 (first spin at 1500 g, 15 min and second spin at 2000 g to 3000 g, 15 min) could be the ideal blood processing protocol for miRNA assays from the aspect of minimally confounding haemolysis and platelet factors. Therefore, a standardized dual-spin protocol of 1500 g for 15 min as first spin and 2500 g for 15 min as second spin, both at room temperature, was established for the processing of PPP.

Table 5. The condition of PPP processing.

Condition	First spin	Second spin
Con1	1500 g, 15 min	1500 g, 20 min
Con2	1500 g, 15 min	2000 g, 15 min
Con3	1500 g, 15 min	3000 g, 15 min
Con4	1500 g, 15 min	5000 g, 5 min
Con5	1000 g, 15 min	1500 g, 20 min
Con6	1000 g, 15 min	3000 g, 15 min

Post freeze-thaw (FT) spin

As many studies may use stored samples which did not go through a second spin, an experiment was conducted to investigate whether sample quality can be improved by a post freeze-thaw spin to remove the residual platelets. The fresh and freeze-thaw conditions were demonstrated Table 6.

Con1 represents single spin without going a freeze-thaw cycle, which served as the baseline in this study; Con2 represents the dual-spin protocol that was developed in the section “Centrifugation Protocol” above; Con3 represents the single-spin protocol that had undergone one freeze-thaw cycle, whilst Con4 represents the dual-spin protocol with the second spin performed only after one freeze-thaw cycle.

The haemolysis score did not change significantly under the different conditions (FIG. 10A). This is consistent with previous finding (FIG. 9A) whereby the red blood cell removal is more affected by the first spin. In FIG. 10B, Con2, which represents the standard dual-spin protocol, generated the highest platelet score. The platelet score of Con3 was similar to Con1, which suggested no difference in platelet amount with one freeze-thaw cycle. The post-FT second spin in Con4, interestingly, was able to increase the platelet score. It was concluded that for studies using stored samples, the confounded platelet-related miRNA could still be partially removed by post freeze-thaw spin.

Table 6. Conditions of the PPP processing

Conditions	First spin	Second spin	Freeze-thaw
Con1	Fresh ²	NIL	NIL
Con2	Fresh ²	Fresh ²	NIL
Con3	Fresh ²	NIL	1 FT
Con4	Fresh ²	Post-FT ³	1 FT

²The status ‘Fresh’ indicated that the centrifugation was done on freshly drawn samples without going through a freeze-thaw cycle.

³Post-FT indicated that the centrifugation was performed after a freeze-thaw (FT) cycle.

Example 3: Stability biomarkers for sample storage

Whole blood storage

The plasma processing after blood draw can be delayed in the clinical labs for various reasons. Hence, there is a need to find the best storage condition for whole blood in blood tubes in case of delay. Consequently, PPP processed immediately (baseline), PPP processed after storage at room temperature for up to 7 hours (RT), and PPP processed after storage on ice for up to 7 hours (on ice) were compared. The haemoglobin levels in RT samples were comparable with the baseline samples (FIG. 11A). The samples on ice had elevated haemoglobin levels which indicated increased haemolysis. The profile is consistent with the haemolysis score (FIG. 11B). The samples on ice also had a slight decrease in the platelet score, which could be due to lysis of platelets; however, this finding was not significant (FIG. 11C). It was concluded that the storage of whole blood at room temperature is more ideal than storage on ice for up to 7 hours.

Real-time stability of miRNA in plasma and serum

- A stability study was conducted to assess the integrity of miRNA in PPP and serum samples after storage at various temperature. This study is important for evaluating the reliability of biomarker discovery in biobank samples. Blood was drawn from four healthy volunteers and processed into PPP and serum. All samples were kept at 25 °C, 4 °C, -20 °C and -80 °C for various duration. RNA was extracted at each timepoint with spike-ins of synthetic RNA functioning as the workflow normalization. The miRNA profiling was performed at the end of the study.
- The mean expression level of all detectable miRNAs (252 miRNAs) in serum samples stored at 25 °C decreased significantly and drastically after three days (FIG. 12A). A similar level of degradation in serum samples stored at 4 °C was seen on day 30, although the decreasing trend in miRNA expression started as early as day 3 (FIG. 12B). Serum samples were much more stable at -20 °C and -80 °C, and only a slight decrease was seen starting from day 180 for -20 °C storage and day 360 for -80 °C storage (FIG. 12C and D). A similar trend was seen in PPP (235 detectable miRNAs), whereby significant degradation was seen in samples stored at 25 °C after 3 days (FIG. 13A) and samples stored at 4 °C after 14 days (FIG. 13B). Similarly, PPP samples were more stable under storage at -20 °C and -80 °C (FIG. 13C and D).
- Regression analysis was performed on expression against duration (days) at 4 °C for serum and PPP samples (FIG. 14A and B). Regression coefficient was obtained and plotted against the average expression of the miRNAs. Most miRNA decreased in expression over time and reflected as a negative value in regression coefficient. Among the detected miRNAs, hsa-miR-143-3p (coefficient = -0.1029) and hsa-miR-618 (coefficient = -0.0494) were identified as stability markers in serum (FIG. 14A). Both miRNAs had expression that was negatively correlated with storage time, yet reasonable for detection, even at day 30 (highlighted in FIG. 14A, B). A similar trend was seen in PPP profile for hsa-miR-143-3p (FIG. 14B). hsa-miR-1290 (coefficient = 0.0283) and hsa-miR-720 (coefficient = 0.0064) were identified as reference markers from serum profile because their expression was minimally impacted by storage time (highlighted in FIG. 14A, B). Similar trend for these two miRNAs was seen in the PPP samples. These miRNAs were chosen as stability or reference markers not only because of their abundant expression, but also the low correlation with both haemolysis and platelets. This was to ensure that the use of stability markers and references had a minimal confounding effect as a result of the presence of platelet and haemolysis.
- The result of the regression analysis implied that the decay of miRNAs can be induced by both storage time and temperature. A correlation analysis was performed using the regression coefficient and expression on miRNAs. The degradation of miRNA is significantly correlated with the expression of miRNAs in serum (Pearson $R = -0.3464$, $P < 0.0001$) and PPP (Pearson $R = -0.2970$, $P < 0.0001$). The correlation analysis was extended to the percentage of GC and length of miRNA. The result indicates that the degradation of miRNA is also correlated with the percentage of GC of the sequence

and the length of miRNA in both serum and plasma (see *R* and *P* value in **FIG. 15**). Initial analysis showed that the regression coefficient has no correlation with the GC content of the sequence and the length of miRNA, but subsequent analysis of the same data suggested otherwise.

5 It has been well defined in recent studies that cellular remnants are an important consideration factor in miRNA studies. Haemolysis, which involves the rupture of erythrocytes, causes a significant increase in the miRNA levels. Results of this study indicate that the vast majority of miRNAs are released by haemolysis. Platelet activation in the serum could also lead to significant increase in miRNAs. Different miRNAs are affected to different degrees by haemolysis and platelet activation or
10 presence. In addition, haemolysis and platelet activation or presence influence a large overlap of common miRNAs found in biofluids, and the two processes are not orthogonal to each other. Thus, if blood collection and processing are not well controlled, it would be extremely difficult to normalize the effect of haemolysis and platelet activation or presence in a post analysis step. Results of this study suggest that haemolysis has a larger effect compared with platelet activation or presence, and platelet
15 activation or presence only becomes the major confounding factor in samples under circumstances where haemolysis is well controlled. A longer platelet activation results in similar miRNA profile in the serum compared with that in the plasma, without any platelet removal from a second spin, suggesting that miRNAs are secreted outside the platelets in the clotting process, and the presence of platelets in the plasma makes the sample equivalent to long-clotted serum. The miRNA in PPP remained low
20 because of its platelet deprivation. Although serum clotting duration is difficult to control in practice, PPP appeared to be the most suitable sample type.

This study found that heterogenous processing conditions, such as centrifugation speed, will impact miRNA measurement. As the presence of red blood cells and platelets may significantly affect the
25 levels of circulating miRNAs and alter the outcome in biomarker discovery studies, additional centrifugation step will help to eliminate red blood cells and platelets. Skin cells may contaminate the first blood draw with tissue-specific miRNAs. Hence, the discard the first tube (~1 mL) was recommended in the blood collection protocol. In this study we have identified the use of PPP as ideal samples to be used in miRNA research. We have also established blood processing protocols which
30 could minimize preanalytical variables factors and allow the use of relative miRNA expression in various research. The application of dual-spin can be practically implemented in a clinical setting to obtain a high-quality PPP sample for miRNA analysis. The schematic workflow of recommended blood processing is shown in **FIG. 16**. A post freeze-thaw spin of stored plasma could partially remove confounding platelets and other cell debris. This study demonstrated that, in the case of delayed
35 processing, whole blood samples can be stored at room temperature for up to 7 hours for the least haemolysis- and platelet-confounding impacts. This finding is consistent with previous studies that reported that on-ice storage of whole blood causes changes in plasma quality and damage to platelet membrane integrity.

The miRNA in serum or plasma is stable for up to 1 year if stored at -20 °C or -80 °C based on the real-time stability studies. Contrary to previous findings whereby miRNA is remarkably stable in plasma, the results indicate that miRNA decays in two distinct phases. The first phase occurs a few days after the blood draw, and the second phase occurs after a long period of storage. It is possible that previous publications are observing the second phase in stored samples. On the basis of the comprehensive analysis of 356 miRNAs, the result indicates that degradation is correlated with the abundance, percentage of GC, and length of miRNA in both serum and PPP samples.

The establishment of sample quality control panel using RBC-related, platelet-related and stability-related miRNAs will allow the screening of ideal samples to be used in downstream miRNA research and measurement. The impact of haemolysis- and platelet-confounding factors is associated with the biological signal. If the disease of interest generates strong changes of related miRNA into the circulating, it is likely that the signal can tolerate higher noise that comes from the confounding factor. However, if the disease signal is subtle, the sample QC must be stringent to prevent the noise from overwhelming the disease signal. It is possible for the user laboratories to define an acceptable range of this panel based on their application. Users can follow the Clinical and Laboratory Standard Institute guideline EP07-A2 to generate samples of different QC scores using platelet or RBC spike-in. The cutoff can then be identified at which point miRNA of interest shows significant changes compared with the control. In some studies, not only the absolute score but also the score difference between groups matters. For example, in a case-control study that aims to compare miRNA expression between groups, the imbalance of the QC score will have to be taken into consideration in statistical analysis or machine learning. Therefore, the use of the panel can be further expanded to study the degree of bias of different cohorts, which may improve research quality.

It is possible to transfer the QC panel onto a different technology platform. A previous study found a good correlation among the MiRXES platform, miRNA sequencing, as well as other qPCR systems. However, by switching to a different technology, some of the miRNAs in the panel may change because of the limit of detection, limit of blank, linearity, precision, and reproducibility of the platform. Thus, it is advantageous to use a sensitive and robust platform for discovery and validation. Besides, if the biological signal of the marker is strong, the signal would overcome the noise induced by platform change and the markers would be more consistent despite a technology change. Therefore, the abundance of the biomarkers in the QC panel was included as one of the consideration factors.

There are other biological variables that may affect the miRNA profile, such as diurnal variation, fasting versus non-fasting state at blood collection, and other endogenous interferences in the blood, which are not within the scope of this study. Because the QC panel was identified using samples from healthy individuals, some of these markers could be affected by disease conditions. Therefore, the QC panel should be used with caution. Nonetheless, this study focused on sample handling, which assessed common clinical processing variables in a systematic manner to compare the profile of hundreds of miRNAs. The panel of 356 miRNAs enabled the finding of more precise biomarkers that

reflect the preanalytical confounding factors. The miRNA fold change in disease signature could be relatively small. Hence, given the drastic changes caused by the preanalytical confounding factors, a true disease signature could be easily overwhelmed and obscured by the blood-processing variables. This study emphasizes the importance of a standard processing workflow and good sample quality in clinical research for a reliable and reproducible result.

Standardization of preanalytical variables represents a critical step in the clinical implementation as well as research studies. Although there are blood collection tubes with preservative for the stabilization of cell-free nucleic acid, they are often costly. In addition, different blood collection systems may also alter the analytical output. Therefore, the development of more preanalytical solutions will benefit the development of miRNA panels as diagnostic and research tools.

Materials and Method

Blood sample collection and processing

To evaluate the miRNA profile of different sample type, ten healthy non-fasting male or female subjects between 25 and 60 years of age who were of Chinese or Malay descent were recruited. Blood was drawn into three serum tubes and two K2EDTA plasma tubes for each subject in randomized drawing sequence to avoid experimental bias. A BD Vacutainer Safety-Lok blood collection set with 21-gauge needles was used (Becton Dickinson, Plymouth, UK, catalogue no. 367281). For plasma processing, blood was collected in K2-EDTA BD Vacutainer tube (Becton Dickinson, Plymouth, UK, catalogue no. 367525) and gently inverted a few times according to manufacturer's protocol. Plasma tubes were processed by centrifugation at 1500 g for 15 min. The supernatants were transferred to new microtubes which were subsequently stored at -80 °C until further analysis. The red blood cell pellets were collected into microtube for further experiments. To generate platelet-poor-plasma (PPP), the spun plasma was transferred into a 15 mL tube which were centrifuged second time at 2500 g for 15 min to remove platelets. The top 2/3 of the supernatants were aliquoted into microtubes and stored at -80 °C. For serum samples, blood was collected into serum separator tube (Becton Dickinson, Plymouth, UK, catalogue no. 367814) and left in upright position for 30 min, 60 min or 90 min at room temperature to allow clotting. After clotting, the blood tubes were centrifuged at 1500 g for 15 min. The supernatants were transferred into microtubes and stored at -80 °C until further analysis.

To assess the impact of blood tube size for plasma, four healthy, non-fasting male or female subjects between 25 and 60 years of age were recruited. Blood was drawn into discard tube for up to 1 mL and then subsequently drawn into 6 mL EDTA tube (Becton Dickinson, Plymouth, UK, catalogue no. 367863) and 10 mL EDTA tube (Becton Dickinson, Plymouth, UK, no. 366643). Both tubes were centrifuged at 1500 g for 15 min.

To find the best centrifugation protocol to produce PPP, five healthy, non-fasting male or female subjects between 25 and 60 years of age were recruited. Blood was drawn into discard tube for up to 1 mL, and subsequently drawn into 6 mL EDTA tube, with 3 tubes for each subject. Of these, one of the blood tubes was centrifuged at 1000 g for 15 min and the other two were centrifuged at 1500 g for 15 min. The supernatant was aspirated from each tube and dispensed into 15 mL tubes for the second spin. Various centrifugation conditions (1500 g for 20 min, 2000 g for 15 min, 3000 g for 15 min, and 5000 g for 5 min) were used in the second spin.

From the above two studies, a dual spin protocol for the processing of platelet-poor plasma was established. EDTA tubes were centrifuged at 1500 g for 15 min at room temperature. Afterward, plasma was transferred into 15 mL tube for the second spin at 2500 g for 15 min at room temperature. The top two-third of the supernatant from the second tube were considered to be PPP and transferred into cryotubes for use.

To study the effect of post freeze-thaw spin, blood was collected in K2-EDTA tubes from four healthy, non-fasting male or female healthy subjects between 25 and 60 years of age. The K2-EDTA tubes were processed using 1500 g centrifugation speed for 15 min. The supernatant from each tube was dispensed into four different tubes, one of them underwent RNA extraction immediately, one of them was centrifuged again at 2500 g for 15 min before RNA extraction, and two of the tubes were stored at -80 °C and subjected to a freeze-thaw cycle. After thawing, only one of the freeze-thawed tubes went through a second centrifugation at 2500 g for 15 min. Both tubes underwent RNA extraction after a freeze-thaw cycle.

For the whole blood storage test, three tubes of blood were drawn from six healthy, non-fasting male or female individuals between 25 and 60 years of age. For each individual, one tube was subjected to immediate processing following dual spin protocol (Baseline). One tube was stored at room temperature in an upright position for up to 7 hours (room temperature) before being subjected to the dual-spin protocol. One tube was stored on ice in upright position for up to 7 hours (on ice) before subjected to the dual-spin protocol.

For real-time stability test, blood from four healthy, non-fasting male or female subjects between 25 and 60 years of age was drawn into discard tubes, for up to 1 mL, and subsequently switched to serum tubes and K2-EDTA tubes. The serum processing was performed by allowing 60 minutes for clotting before centrifugation at 1500 g for 15 min. The storage conditions of the serum samples were 25 °C for 3 days; 4 °C for 3, 7 and 30 days; -20 °C for 3, 7, 30, 90, 180 and 360 days; -80 °C for 3, 7, 30, 90, 180 and 360 days. Day 0 represented freshly processed serum. The PPP processing was done using the dual-spin protocol. The storage conditions of the PPP samples were 25 °C for 3 and 7 days; 4 °C for 3, 7, 14 and 30 days; -20 °C for 3, 7, 14, 30, 60, 90, 180, 270 and 360 days; -80 °C for 3, 7, 14, 30, 60, 90, 180, 270 and 360 days. Day 0 represented freshly processed PPP. For multiple

blood tubes collected from the same individual, the tubes and processing conditions were randomized and did not follow collection order.

Red blood cell lysis

5 Red blood cell pellets were washed with Phosphate Buffered Saline for two times. The ACK Lysis buffer (Gibco Biosciences, Dublin, Ireland, catalogue no. A1049201) was added to the pellets at an equivalent volume. The mixture was gently resuspended and swirled for 30 seconds. The mixture was then centrifuged for 10 min at 2000 x g, and the supernatant containing the red blood cell lysate was
10 collected. The lysed red blood cells were spiked into PPP to generate contrived samples with abundant haemolysis-related miRNA.

Haemolysis and platelet activation parameters

15 Haemolysis was assessed by quantification of cell-free haemoglobin. Levels of cell-free haemoglobin in the plasma or serum were measured by Human Haemoglobin ELISA kit (Abcam, Cambridge, MA) following protocol from the product insert. Platelet activation was assessed by the quantification of Thromboxane B2 (TXB2) by ELISA (R&D Systems, Minneapolis, MN) according to manufacturer's protocol. The TXB2 ELISA is specific to TXB2 in cell culture supernatant, serum, plasma, and urine,
20 with a sensitivity of 0.31 ng/mL.

RNA isolation

25 RNAs were isolated from 200 µL of serum or plasma using Maxwell RSC miRNA Plasma and Serum Kit on Promega Maxwell RSC 48 Instrument (Promega, Wisconsin, USA) according to the manufacturer's protocol. Three spike-ins of synthetic short RNAs with distinct sequences from endogenous human miRNAs were added to the lysis buffer as controls to normalize workflow variations. The resulting miRNA was eluted using 50 µL of nuclease free water.

miRNA profiling

The miRNA profiling was performed following an established protocol in MiRXES Lab (Zou R, *et al.*, Cancers (Basel) **2021**, 13:1-12). A total of 356 miRNAs were profiled using the multiplexed RT-qPCR platform developed in MiRXES Lab Pte. Ltd. Isolated miRNAs underwent reverse transcription using
35 an in-house reverse transcription system and modified stem-loop RT primer pools (MiRXES, Singapore) on a Veriti Thermal Cycler (Applied Biosystem, Waltham, MA, USA) according to manufacturer's instructions. For each RT reaction, a standard consisted of a serial of six 10-fold dilutions of synthetic miRNA and two no-template controls (NTCs) were reverse transcribed on the same plate. The cDNA was then pre-amplified using a 14-cycle PCR reaction with Augmentation
40 Primer Pools (MiRXES, Singapore) on the Veriti Thermal Cycler. Single PCR was performed on the

amplified cDNA samples using a miRNA-specific qPCR assay and ID3EAL miRNA qPCR Master Mix according to the manufacturer's instruction (MiRXES, Singapore). The qPCR reaction for each sample was performed with technical duplicates on the QuantStudio 5 Real-Time PCR System (Applied Biosystem, Waltham, MA, USA). Raw threshold cycle (Cq) values were calculated using the
5 QuantStudio Design and Analysis software 1.4.3 with automatic baseline setting and a threshold of 0.4.

In miRNA profiling, absolute expression of each miRNA (number of copies present) was calculated by interpolation of sample Cq values with synthetic miRNA standard curves after correcting for variations
10 in RT-qPCR efficiency with spike-ins. Any miRNA with Cq number greater than that of the no template controls (NTCs) were deemed undetectable and removed from analysis. The data underwent normalization with spike-ins to exclude variation during RNA isolation and RT-qPCR.

RT-qPCR for selected panel

15 The miRNAs included in the panel were hsa-miR-1973 and hsa-miR-28-5p as platelet-related biomarkers; hsa-miR-20b-5p, hsa-miR-363-3p, and hsa-miR-451a as RBC-related biomarkers; and hsa-miR-1290 and hsa-miR-10b-5p as reference marker for both RBC-related and platelet-related biomarkers.

20 RNA was reverse transcribed using ID3EAL cDNA synthesis reagents (MiRXES) with a modified stem-loop reverse transcription primer pool assembled for the panel. A total of 5 µL of extracted RNA was mixed with ID3EAL miRNA reverse transcription buffer, ID3EAL reverse transcriptase, and reverse transcription primer pool in a total reaction volume of 15 µL. The reaction mixture was
25 incubated at 42 °C for 30 minutes, followed by 95 °C for 5 minutes to inactivate the reverse transcriptase on a C1000 Touch Thermal Cycler (Bio-Rad).

Real-time quantitative PCR was performed using ID3EAL miRNA qPCR reagents (MiRXES) with specific pairs of primers for each of the specific targets. Each cDNA samples was diluted 10 times
30 with nuclease-free water and added in duplicate into a 384-well plate (Applied Biosystems, Foster City, CA). PCR amplification was performed in a total reaction volume of 15 µL containing cDNA, 1 x ID3EAL miRNA qPCR master mix, 1 x ID3EAL miRNA qPCR primers (MiRXES), topped up with nuclease-free water, qPCR amplification and detection were performed on QuantStudio 5 Real-Time PCR System (Thermo Fisher Scientific) with the following cycling conditions: 95 °C for 10 minutes, 40
35 °C for 5 minutes, followed by 40 cycles of 95 °C for 10 seconds and 60 °C for 30 seconds (optical reading). Raw cycle threshold (Cq) values were calculated using QuantStudio Design and Analysis software version 1.5 and a threshold of 0.4. Cq values was used for the interpretation of the results.

Statistical Analysis

40

In miRNA profiling, all the expression levels were presented in log2 copy numbers per millilitre of samples (Log2 copies/mL). The miRNA profile was obtained by normalizing individual miRNAs with their overall expression in all samples (Mestdagh P, *et al.*, *Genome Biol.* 2009, 10:R64). The miRNAs underwent Z-score standardization before principal component analysis (PCA) was performed. Statistical analysis was performed using MATLAB software version R2019b (MathWorks Inc., Natick, MA).

A linear regression model was used to investigate the correlation among miRNA expression, haemolysis, platelets, stability and storage. To examine if platelet-related miRNA is a major confounding factor, linear regression was performed between expression level of detectable miRNA and TXB2 level in serum samples. To identify platelet-related miRNA, linear regression was performed between serum clotting duration and expression level of all detectable miRNAs.

Those miRNAs with p value < 0.05 in the regression analysis were inferred as having high correlation and identified as potential biomarkers. To assess real-time stability of miRNA in plasma and serum, linear regression was performed between duration (days) at 4 °C and level of expression of all detectable miRNAs. All linear regression was performed using MATLAB R2019b. Correlation analysis was performed using GraphPad Prism software version 9.2.0 (GraphPad Software Inc., San Diego, CA).

The RBC score and platelet score were calculated using ΔCq (average Cq of biomarkers minus average Cq of reference). Higher scores indicate samples with less haemolysis or platelet content. One-way ANOVA was performed to determine if there is significant difference between groups.

Before sample collection, sample size calculation was performed using MATLAB R2019b with 80% power. One-way analysis of variance or paired t-test was performed using GraphPad Prism 9.2.0 to determine whether there was a significant difference between groups.

References

1. Uratani, R. et al. Diagnostic potential of cell-free and exosomal MicroRNAs in the identification of patients with high-risk colorectal adenomas. *PLoS One* 11, 1–16 (2016).
2. Faraldi, M. et al. and storage conditions. *30*, 1–13 (2020)
3. Risberg, B. et al. Effects of Collection and Processing Procedures on Plasma Circulating Cell-Free DNA from Cancer Patients. *J. Mol. Diagnostics* 20, 883–892 (2018)
4. Binderup, H. G. et al. Quantification of microRNA levels in plasma – Impact of preanalytical and analytical conditions. *PLoS One* 13, 1–13 (2018)
5. Zhao, H. et al. Effects of preanalytic variables on circulating MicroRNAs in whole blood. *Cancer Epidemiol. Biomarkers Prev.* 23, 2643–2648 (2014)
6. Cheng, H. H. et al. Plasma Processing Conditions Substantially Influence Circulating microRNA Biomarker Levels. *PLoS One* 8, 1–11 (2013)

7. Edelstein, L. C. et al. MicroRNAs in platelet production and activation. *J. Thromb. Haemost.* 11, 340–350 (2013)

CLAIMS

1. A method of assessing the quality of a sample for use in a nucleic acid analysis, the method comprises determining the level of one or more biomarker in the sample, wherein the one or more biomarker is selected from the group consisting of hsa-miR-28-5p, hsa-miR-1973, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, and hsa-miR-618.
2. The method of claim 1, wherein the quality of the sample is determined by:
 - a) the level of platelet activation and/or aggregation in the sample,
 - b) the level of haemolysis in the sample, and/or
 - c) the level of stability of nucleic acid in the sample.
3. The method of claim 1 or 2, wherein a) is determined by measuring/detecting/determining the level of one or more biomarker selected from the group consisting of hsa-miR-28-5p and hsa-miR-1973.
4. The method of claim 3, wherein the method comprises determining the level of one of the following:
 - hsa-miR-28-5p; or
 - hsa-miR-1973; or
 - hsa-miR-1973 and hsa-miR-28-5p.
5. The method of claim 1 or 2, wherein b) is determined by measuring/detecting/determining the level of one or more biomarker selected from the group consisting of hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a.
6. The method according to claim 5, wherein the method comprises determining the level of one of the following:
 - hsa-miR-20b-5p;
 - hsa-miR-363-3p;
 - hsa-miR-451a;
 - hsa-miR-20b-5p and hsa-miR-363-3p;
 - hsa-miR-20b-5p and hsa-miR-451a;
 - hsa-miR-363-3p and hsa-miR-451a; or
 - hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a.
7. The method of claim 1 or 2, wherein c) is determined by measuring/detecting/determining the level of one or more biomarker is selected from hsa-miR-143-3p and hsa-miR-618.

8. The method according to claim 7, wherein the method comprises determining the level of:

hsa-miR-143-3p;
hsa-miR-618; or
hsa-miR-143-3p and hsa-miR-618.

9. A method of determining the level of platelet activation and/or aggregation in a sample, the method comprises determining the level of one or more biomarker in the sample, wherein the one or more biomarker is selected from hsa-miR-28-5p and hsa-miR-1973.

10. The method according to claim 9, wherein the method comprises determining the level of one of the following: hsa-miR-28-5p; hsa-miR-1973; or hsa-miR-1973 and hsa-miR-28-5p.

11. A method of determining the level of haemolysis in a sample, the method comprises determining the level of one or more biomarker in the sample, wherein the one or more biomarker is selected from hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a.

12. The method according to claim 11, wherein the method comprises determining the level of one of the following: hsa-miR-20b-5p; hsa-miR-363-3p; hsa-miR-451a; hsa-miR-20b-5p and hsa-miR-363-3p; hsa-miR-20b-5p and hsa-miR-451a; hsa-miR-363-3p and hsa-miR-451a; or hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a.

13. A method of determining the level of stability of nucleic acid in a sample, the method comprises determining the level of one or more biomarker in the sample, wherein the one or more biomarker is selected from hsa-miR-143-3p and hsa-miR-618.

14. The method according to claim 13, wherein the method comprises determining the level of: hsa-miR-143-3p; hsa-miR-618; or hsa-miR-143-3p and hsa-miR-618.

15. The method according to any one of the preceding claims, wherein the method further comprises determining the level of one or more biomarker selected from hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720.

16. The method according to claim 15, wherein the method further comprises comparing the level of hsa-miR-28-5p and/or hsa-miR-1973 with that of hsa-miR-1290 and/or hsa-miR-10b-5p to determine the level of platelet activation and/or aggregation in the sample.

17. The method according to claim 16, wherein the method further comprises comparing the level of one or more of hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a with that of hsa-miR-1290 and/or hsa-miR-10b-5p to determine the level of haemolysis in the sample.

18. The method according to claim 16, wherein the method further comprises comparing the level of hsa-miR-143-3p and hsa-miR-618 with that of hsa-miR-1290 and/or hsa-miR-720 to determine the level of stability of nucleic acid in the sample.

19. The method according to claim 1, wherein the method comprises determining the level of at least one biomarker in each of the following group:

- (a) hsa-miR-28-5p and hsa-miR-1973;
- (b) hsa-miR-20b-5p, hsa-miR-363-3p and hsa-miR-451a; and
- (c) hsa-miR-143-3p and hsa-miR-618.

20. The method according to claim 1 or 19, wherein the method comprises determining the level of hsa-miR-28-5p, hsa-miR-1973, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p and hsa-miR-618.

21. The method according to claim 15, wherein the method comprises determining the level of biomarkers hsa-miR-28-5p, hsa-miR-1973, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, hsa-miR-618, hsa-miR-1290, hsa-miR-10b-5p and hsa-miR-720.

22. The method according to any one of the preceding claims, wherein the method comprises comparing the level of two or more biomarkers to determine a score, wherein the score is compared to that of a control for determining the quality of the sample.

23. The method according to any one of the preceding claims, wherein the sample is a blood sample or a non-cellular bodily fluid sample obtained from a subject, optionally the sample is a plasma and/or serum.

24. A kit for use in the method according to any one of the preceding claims, the kit comprises an isolated set of probes and/or reagents capable of determining the expression level of the one or more biomarker selected from the group consisting of hsa-miR-1973, hsa-miR-28-5p, hsa-miR-20b-5p, hsa-miR-363-3p, hsa-miR-451a, hsa-miR-143-3p, and hsa-miR-618.