CLAIMS

1. A method for discovering a disease biomarker comprising:

(i) providing a first cDNA sample from a subject with a disease and a second cDNA sample from a subject without the disease;

(ii) normalizing the first and the second cDNA samples;

(iii) sequencing the normalized first and second cDNA samples; and

(iv) comparing the sequencing output for the first and second cDNA samples to discover a disease biomarker; in which:

the first and second cDNA samples comprise double stranded cDNA templates, each template having a known 5' pre-attached adapter and a known 3' pre-attached adapter;

and in which normalizing the first and second cDNA samples comprises:

(i) denaturing the cDNA sample to produce single stranded cDNA templates;

(ii) re-associating the cDNA sample to produce a mixture of post-association single stranded cDNA templates and post-association double stranded cDNA templates;

(iii) annealing a 5' adapter complex to the 5' pre-attached adapter of at least one post-association single stranded cDNA template, and annealing a 3' adapter complex to the 3' pre-attached adapter of the same post-association single stranded cDNA template, in which each adapter complex comprises at least one oligonucleotide;

(iv) ligating an oligonucleotide from the 5' adapter complex to the 5' pre-attached adapter of the post-association single stranded cDNA template and ligating an oligonucleotide from the 3' adapter complex to the 3' pre-attached adapter of the same post-association single stranded cDNA template; and

(v) selectively amplifying the cDNA sample using primers specific to the ligated oligonucleotides.

2. A method for diagnosing a disease in a subject comprising:

(i) providing a cDNA sample from the subject;

(ii) normalizing the cDNA sample; and

(iii) sequencing the normalized cDNA sample, in which the sequencing output is used to identify whether the subject has the disease; in which:

the cDNA sample comprises double stranded cDNA templates, each template having a known 5’ pre-attached adapter and a known 3’ pre-attached adapter;

and in which normalizing the cDNA sample comprises:

(i) denaturing the cDNA sample to produce single stranded cDNA templates;

(ii) re-associating the cDNA sample to produce a mixture of post-association single stranded cDNA templates and post-association double stranded cDNA templates;

(iii) annealing a 5’ adapter complex to the 5’ pre-attached adapter of at least one post-association single stranded cDNA template, and annealing a 3’ adapter complex to the 3’ pre-attached adapter of the same post-association single stranded cDNA template, in which each adapter complex comprises at least one oligonucleotide;

(iv) ligating an oligonucleotide from the 5’ adapter complex to the 5’ pre-attached adapter of the post-association single stranded cDNA template and ligating an oligonucleotide from the 3’ adapter complex to the 3’ pre-attached adapter of the same post-association single stranded cDNA template; and

(v) selectively amplifying the cDNA sample using primers specific to the ligated oligonucleotides.

3. A method according to Claim 2, in which:

(a) the method comprises comparing the sequencing output for the normalized cDNA sample to one or more reference sequences or to the sequencing output of one or more control samples, optionally in which the one or more control samples are from one or more subjects with and/or without the disease;

(b) the cDNA sample is obtained from a biological fluid or a fluid or lysate generated from a biological material, optionally in which the cDNA sample is obtained from blood; and/or

(c) the method further comprises reporting to the subject the outcome of the method.

4. A method according to Claim 1, 2 or 3 in which the disease is cancer.

5. A method according to any preceding claim, in which normalizing comprises increasing the amount of low abundance cDNA within each cDNA sample.

6. A method according to any preceding claim, in which:

(A) the 5’ adapter complex is a front oligonucleotide dimer comprising:

(i) a front lig-oligonucleotide for ligating to the 5’ pre-attached adapter of the (post-association) single stranded cDNA template; and

(ii) a front link-oligonucleotide for annealing to the 5’ pre-attached adapter and the front lig-oligonucleotide, the front link-oligonucleotide comprising a region complementary to the 5’ pre-attached adapter and a region complementary to the front lig-oligonucleotide,

such that, on annealing, an end of the front lig-oligonucleotide is adjacent an end of the 5’ pre-attached adapter to enable ligation of the front lig-oligonucleotide to the 5’ pre-attached adapter at a ligation site; and

(B) the 3’ adapter complex is a back oligonucleotide dimer comprising:

(i) a back lig-oligonucleotide for ligating to the 3’ pre-attached adapter of the (post-association) single stranded cDNA template; and

(ii) a back link-oligonucleotide for annealing to the 3’ pre-attached adapter and the back lig-oligonucleotide, the back link-oligonucleotide comprising a region complementary to the 3’ pre-attached adapter and a region complementary to the back lig-oligonucleotide,

such that, on annealing, an end of the back lig-oligonucleotide is adjacent an end of the 3’ pre-attached adapter to enable ligation of the back lig-oligonucleotide to the 3’ pre-attached adapter at a ligation site, optionally in which:

(A) the front link-oligonucleotide comprises:

(i) a template overhang region at an end of the front link-oligonucleotide proximal the region complementary to the 5’ pre-attached adapter, the template overhang region being non-complementary to a corresponding region of the (post-association) single stranded cDNA template; and/or

(ii) a lig-oligonucleotide overhang region at an end of the front link-oligonucleotide proximal the region complementary to the front lig-oligonucleotide, the lig-oligonucleotide overhang region being non-complementary to a corresponding region of the front lig-oligonucleotide; and/or

(B) the back link-oligonucleotide comprises:

(i) a template overhang region at an end of the back link-oligonucleotide proximal the region complementary to the 3' pre-attached adapter, the template overhang region being non-complementary to a corresponding region of the (post-association) single stranded cDNA template; and/or

(ii) a lig-oligonucleotide overhang region at an end of the back link-oligonucleotide proximal the region complementary to the back lig-oligonucleotide, the lig-oligonucleotide overhang region being non-complementary to a corresponding region of the back lig-oligonucleotide.

7. A method according to any preceding claim, in which sequencing comprises the use of long read sequencing.

8. A method according to any of Claims 1 or 4 to 7, in which:

(a) the first cDNA sample and the second cDNA sample are obtained from biological fluid or a fluid or lysate generated from a biological material, optionally in which the first cDNA sample and the second cDNA sample are obtained from blood;

(b) the method further comprises, prior to step (i), extracting RNA from biological fluid or a fluid or lysate generated from a biological material from the subject with a disease and from the subject without the disease and synthesizing cDNA using the RNA as a template; and/or

(c) the disease biomarker is a cDNA sequence that is present in the first cDNA sample but not in the second cDNA sample or is present in the second cDNA sample but not in the first cDNA sample.

9. A method according to any of Claims 2 to 7, further comprising, prior to step (i), extracting RNA from biological fluid or a fluid or lysate generated from a biological material from the subject and synthesizing cDNA using the RNA as a template.

10. A method according to any of Claims 2 to 7 or 9, in which the cDNA sample is obtained from blood by a method for processing a blood sample comprising:

(i) storing the blood sample at -15°C or below, optionally within 12 hours of collection;

(ii) thawing the blood sample at 5 to 30°C for at least 1 hour;

(iii) extracting RNA from the thawed blood sample; and

(iv) synthesizing cDNA using the extracted RNA as a template.

11. A method according to any of Claims 1 or 4 to 8, in which the first cDNA sample and the second cDNA sample are obtained from blood by a method for processing a blood sample comprising:

(i) storing the blood sample at -15°C or below, optionally within 12 hours of collection;

(ii) thawing the blood sample at 5 to 30°C for at least 1 hour;

(iii) extracting RNA from the thawed blood sample; and

(iv) synthesizing cDNA using the extracted RNA as a template.

12. A method according to Claim 10 or 11, in which:

(i) the blood sample is stored at -20°C or below;

(ii) the blood sample is thawed at 18 to 25°C; and/or

(iii) the blood sample is thawed for 1 to 4 hours.

13. Use in a method for discovering a cancer biomarker of an oligonucleotide dimer composition for selective amplification of single stranded cDNA by ligation of an oligonucleotide to a 5’ and a 3’ end of a post-association single stranded cDNA template having known 5’ and 3’ pre-attached adapters, in which the composition comprises:

(A) a front oligonucleotide dimer comprising:

(i) a front lig-oligonucleotide for ligating to the 5’ pre-attached adapter of the post-association single stranded cDNA template; and

(ii) a front link-oligonucleotide for annealing to the 5’ pre-attached adapter and the front lig-oligonucleotide, the front link-oligonucleotide comprising a region complementary to the 5’ pre-attached adapter and a region complementary to the front lig-oligonucleotide,

such that, on annealing, an end of the front lig-oligonucleotide is adjacent an end of the 5’ pre-attached adapter to enable ligation of the front lig-oligonucleotide to the 5’ pre-attached adapter at a ligation site; and

(B) a back oligonucleotide dimer comprising:

(i) a back lig-oligonucleotide for ligating to the 3’ pre-attached adapter of the post-association single stranded cDNA template; and

(ii) a back link-oligonucleotide for annealing to the 3’ pre-attached adapter and the back lig-oligonucleotide, the back link oligonucleotide comprising a region complementary to the 3’ pre-attached adapter and a region complementary to the back lig-oligonucleotide,

such that, on annealing, an end of the back lig-oligonucleotide is adjacent an end of the 3’ pre-attached adapter to enable ligation of the back lig-oligonucleotide to the 3’ pre-attached adapter at a ligation site, optionally in which:

(A) the front link-oligonucleotide comprises:

(i) a template overhang region at an end of the front link-oligonucleotide proximal the region complementary to the 5’ pre-attached adapter, the template overhang region being non-complementary to a corresponding region of the post-association single stranded cDNA template; and/or

(ii) a lig-oligonucleotide overhang region at an end of the front link-oligonucleotide proximal the region complementary to the front lig-oligonucleotide, the lig-oligonucleotide overhang region being non-complementary to a corresponding region of the front lig-oligonucleotide; and/or

(B) the back link-oligonucleotide comprises:

(i) a template overhang region at an end of the back link-oligonucleotide proximal the region complementary to the 3’ pre-attached adapter, the template

overhang region being non-complementary to a corresponding region of the post-association single stranded cDNA template; and/or

(ii) a lig-oligonucleotide overhang region at an end of the back link-oligonucleotide proximal the region complementary to the back lig-oligonucleotide, the lig-oligonucleotide overhang region being non-complementary to a corresponding region of the back lig-oligonucleotide.

14. A system or test kit for discovering a disease biomarker, comprising:

(a) one or more testing devices for normalizing a first cDNA sample from a subject with a disease and a second cDNA sample from a subject without the disease and sequencing the normalized first and second cDNA samples;

(b) a processor; and

(c) storage medium comprising a computer application that, when executed by the processor, is configured to:

(i) access the determined sequence(s) for the first and second cDNA samples on the one or more testing devices

(ii) calculate whether there is a cDNA sequence that is present in the first cDNA sample but not in the second cDNA sample or is present in the second cDNA sample but not in the first cDNA sample; and

(iii) output from the processor the result of step (ii);

in which the one or more testing devices comprises an oligonucleotide dimer composition for selective amplification of single stranded cDNA by ligation of an oligonucleotide to a 5’ and a 3’ end of a single stranded cDNA template having known 5’ and 3’ pre-attached adapters, in which the oligonucleotide dimer composition comprises:

(A) a front oligonucleotide dimer comprising:

(i) a front lig-oligonucleotide for ligating to the 5’ pre-attached adapter of the post-association single stranded cDNA template; and

(ii) a front link-oligonucleotide for annealing to the 5’ pre-attached adapter and the front lig-oligonucleotide, the front link-oligonucleotide comprising a

region complementary to the 5’ pre-attached adapter and a region complementary to the front lig-oligonucleotide,

such that, on annealing, an end of the front lig-oligonucleotide is adjacent an end of the 5’ pre-attached adapter to enable ligation of the front lig-oligonucleotide to the 5’ pre-attached adapter at a ligation site; and

(B) a back oligonucleotide dimer comprising:

(i) a back lig-oligonucleotide for ligating to the 3’ pre-attached adapter of the post-association single stranded cDNA template; and

(ii) a back link-oligonucleotide for annealing to the 3’ pre-attached adapter and the back lig-oligonucleotide, the back link oligonucleotide comprising a region complementary to the 3’ pre-attached adapter and a region complementary to the back lig-oligonucleotide,

such that, on annealing, an end of the back lig-oligonucleotide is adjacent an end of the 3’ pre-attached adapter to enable ligation of the back lig-oligonucleotide to the 3’ pre-attached adapter at a ligation site.

15. A system or test kit for diagnosing a disease in a subject, comprising:

(a) one or more testing devices for normalizing a cDNA sample from a subject and sequencing the normalized cDNA sample

(b) a processor; and

(c) storage medium comprising a computer application that, when executed by the processor, is configured to:

(i) access the determined sequence(s) for the cDNA sample on the one or more testing devices

(ii) calculate whether a cDNA sequence is present or absent, in which the presence or absence of the cDNA sequence is associated with the disease; and

(iii) output from the processor whether a subject has the disease;

in which the one or more testing devices comprises an oligonucleotide dimer composition for selective amplification of single stranded cDNA by ligation of an oligonucleotide to a 5’ and a 3’ end of a single stranded cDNA template having

known 5’ and 3’ pre-attached adapters, in which the oligonucleotide dimer composition comprises:

(A) a front oligonucleotide dimer comprising:

(i) a front lig-oligonucleotide for ligating to the 5’ pre-attached adapter of the post-association single stranded cDNA template; and

(ii) a front link-oligonucleotide for annealing to the 5’ pre-attached adapter and the front lig-oligonucleotide, the front link-oligonucleotide comprising a region complementary to the 5’ pre-attached adapter and a region complementary to the front lig-oligonucleotide,

such that, on annealing, an end of the front lig-oligonucleotide is adjacent an end of the 5’ pre-attached adapter to enable ligation of the front lig-oligonucleotide to the 5’ pre-attached adapter at a ligation site; and

(B) a back oligonucleotide dimer comprising:

(i) a back lig-oligonucleotide for ligating to the 3’ pre-attached adapter of the post-association single stranded cDNA template; and

(ii) a back link-oligonucleotide for annealing to the 3’ pre-attached adapter and the back lig-oligonucleotide, the back link oligonucleotide comprising a region complementary to the 3’ pre-attached adapter and a region complementary to the back lig-oligonucleotide,

such that, on annealing, an end of the back lig-oligonucleotide is adjacent an end of the 3’ pre-attached adapter to enable ligation of the back lig-oligonucleotide to the 3’ pre-attached adapter at a ligation site.