CLAIMS

1. A method of base calling nucleobases of two or more polynucleotide sequence portions, the method comprising:

(a) obtaining first intensity data comprising a combined intensity of a first signal obtained based upon a respective first nucleobase of at least one first polynucleotide sequence portion and a second signal obtained based upon a respective second nucleobase of at least one second polynucleotide sequence portion;

(b) obtaining second intensity data comprising a combined intensity of a third signal obtained based upon the respective first nucleobase of the at least one first polynucleotide sequence portion and a fourth signal obtained based upon the respective second nucleobase of the at least one second polynucleotide sequence portion;

(c) selecting one of a plurality of classifications based on the first and the second intensity data, wherein each classification represents a possible combination of respective first and second nucleobases; and

(d) based on the selected classification, base calling the respective first and second nucleobases,

wherein said polynucleotide sequence portions have been selectively processed such that an intensity of the signals obtained based upon the respective first nucleobase is greater than an intensity of the signals obtained based upon the respective second nucleobase.

2. The method of claim 1, wherein the first and second signals and/or the third and fourth signals are obtained substantially simultaneously.

3. The method of claim 1 or claim 2, wherein selecting the classification based on the first and second intensity data comprises selecting the classification based on the combined intensity of the first and second signals and the combined intensity of the third and fourth signals.

4. The method of any one preceding claim, wherein the plurality of classifications comprises sixteen classifications, each classification representing one of sixteen unique combinations of first and second nucleobases.

5. The method of any one preceding claim, wherein the polynucleotide sequence portions have been selectively processed such that, during sequencing, a greater number of the first polynucleotide sequence portions are capable of generating a signal

than a number of the second polynucleotide sequence portions that are capable of generating a signal.

6. The method of claim 5, wherein a ratio between the number of the first polynucleotide sequence portions capable of generating a signal and the number of the second polynucleotide sequence portions capable of generating a signal is between 1.25:1 to 5:1, preferably between 1.5:1 to 3:1, more preferably about 2:1.

7. The method of any one preceding claim, wherein the first signal, second signal, third signal and fourth signal are generated based on light emissions associated with the respective nucleobase.

8. The method of claim 7, wherein the obtained signals are generated by:

contacting a plurality of polynucleotide molecules comprising the first and second polynucleotide sequence portions with first primers for sequencing the first polynucleotide sequence portions and second primers for sequencing the second polynucleotide sequence portions;

extending the first primers and the second primers by contacting the polynucleotide molecules with labeled nucleobases to form first labeled primers and second labeled primers;

stimulating the light emissions from the first and second labeled primers; and

detecting the light emissions at a sensor.

9. The method of claim 8, wherein:

the first and second signals are based on light emissions detected in a first range of optical frequencies;

the third and fourth signals are based on light emissions detected in a second range of optical frequencies; and

wherein the first range of optical frequencies and the second range of optical frequencies are not identical.

10. The method of claim 8 or claim 9, wherein the plurality of polynucleotide molecules are attached to a substrate, and wherein the light emissions from the first labeled primers and the light emissions from the second labeled primers are emitted from the same region or substantially overlapping regions of the substrate.

11. The method of any one of claims 8 to 10, wherein the light emissions detected at the sensor are spatially unresolved.

12. The method of claim 11, wherein the sensor is configured to provide a single output based upon the first and second signals.

13. The method of any one of claims 8 to 12, wherein the sensor comprises a single sensing element.

14. The method of any one of claims 8 to 13, wherein each polynucleotide molecule comprises one or more copies of the first polynucleotide sequence portion and one or more copies of the second polynucleotide sequence portion.

15. The method of any one of claims 1 to 13, wherein the first and second polynucleotide sequence portions are respective portions of different polynucleotide molecules.

16. The method of any one preceding claim, wherein the at least one first polynucleotide sequence portion and the at least one second polynucleotide sequence portion are present in a cluster.

17. The method of any one preceding claim, wherein the one of the plurality of classifications is selected based on the first and the second intensity data using a Gaussian mixture model.

18. The method of any one preceding claim, further comprising, based on the selected classification, determining that the second polynucleotide sequence portion is modified relative to the first polynucleotide sequence portion at a location associated with the first and second nucleobases.

19. A data processing system comprising means for carrying out a method comprising:

(a) obtaining first intensity data comprising a combined intensity of a first signal obtained based upon a respective first nucleobase of at least one first polynucleotide sequence portion and a second signal obtained based upon a respective second nucleobase of at least one second polynucleotide sequence portion;

(b) obtaining second intensity data comprising a combined intensity of a third signal obtained based upon the respective first nucleobase of the at least one first polynucleotide sequence portion and a fourth signal obtained based upon the respective second nucleobase of the at least one second polynucleotide sequence portion;

(c) selecting one of a plurality of classifications based on the first and the second intensity data, wherein each classification represents a possible combination of respective first and second nucleobases; and

(d) based on the selected classification, base calling the respective first and second nucleobases,

wherein said polynucleotide sequence portions have been selectively processed such that an intensity of the signals obtained based upon the respective first nucleobase is greater than an intensity of the signals obtained based upon the respective second nucleobase.

20. A computer-readable storage medium comprising instructions which, when executed by a processor, cause the processor to carry out a method comprising:

(a) obtaining first intensity data comprising a combined intensity of a first signal obtained based upon a respective first nucleobase of at least one first polynucleotide sequence portion and a second signal obtained based upon a respective second nucleobase of at least one second polynucleotide sequence portion;

(b) obtaining second intensity data comprising a combined intensity of a third signal obtained based upon the respective first nucleobase of the at least one first polynucleotide sequence portion and a fourth signal obtained based upon the respective second nucleobase of the at least one second polynucleotide sequence portion;

(c) selecting one of a plurality of classifications based on the first and the second intensity data, wherein each classification represents a possible combination of respective first and second nucleobases; and

(d) based on the selected classification, base calling the respective first and second nucleobases,

wherein said polynucleotide sequence portions have been selectively processed such that an intensity of the signals obtained based upon the respective first nucleobase is greater than an intensity of the signals obtained based upon the respective second nucleobase.