

CLAIMS

What is claimed is:

1. A method of identifying microsatellite instability in a biological sample comprising:

obtaining a plurality of signals by detecting fluorescence of fragments comprising nucleic acid sequences obtained using the biological sample wherein each signal corresponds to one of a plurality of different microsatellite loci;

determining one or more signal features for each of the plurality of signals; and

applying one or more classifiers to one or more of the signal features of the plurality of microsatellite loci to identify whether the biological sample is microsatellite instability high, microsatellite instability low, or microsatellite stable.

2. The method of claim 1, further comprising applying one or more classifiers to one or more signal features corresponding to the signal for each individual microsatellite locus in the plurality of different microsatellite loci to identify whether each individual microsatellite locus is microsatellite unstable or microsatellite stable and combining these determinations across loci to determine a microsatellite status of the biological sample.

3. The method of claim 1 or claim 2, wherein the applying one or more classifiers comprise comparing a signal feature derived from the biological sample and a signal feature derived from one or more samples of non-cancerous tissue.
4. The method of claim 1 or claim 2, wherein at least one classifier comprises a fragment size threshold.
5. The method of claim 1 or claim 2, wherein at least one classifier comprises a fragment size interval.
6. The method of claim 1 or claim 2, wherein applying the one or more classifiers comprises evaluating a peak count within a specified size interval.
7. The method of claim 1 or claim 2, wherein applying the one or more classifiers comprises a evaluating a relative peak count between tumor and normal tissues within a specified size interval.
8. The method of claim 1 or claim 2, wherein applying the one or more classifiers comprises evaluating a peak envelope count within a specified size interval.

9. The method of claim 1 or claim 2, wherein applying the one or more classifiers comprises evaluating a peak envelope separation within a specified size interval.

10. The method of claim 1 or claim 2, wherein applying the one or more classifiers comprises evaluating a peak envelope separation within a specified size interval.

11. The method of claim 1 or 2, wherein applying the one or more classifiers comprises evaluating a shift in one or more peak locations within a specified size interval.

12. The method of claim 1 or 2, wherein applying the one or more classifiers comprises analyzing a peak pattern input.

13. The method of claim 12, wherein the peak pattern input comprises two or more values of: peak amplitudes along the fragment size axis, peak locations along the fragment size axis; peak amplitudes relative to a largest peak, peak locations relative to a largest peak, peak envelope peak amplitudes, peak envelope peak

locations, peak envelope peak widths, or peak metrics relative to peak envelope metrics.

14. The method of claim 2, further comprising:

assigning the biological sample a high microsatellite instability status when a percentage of the microsatellite loci is determined to be microsatellite unstable is above a first predetermined threshold, a low microsatellite instability status if the percentage of microsatellite loci is determined to be microsatellite unstable is above a second predetermined threshold but below the first determined threshold, or a microsatellite stable status if the percentage of microsatellite loci determined to be microsatellite unstable is below the second predetermined threshold.

15. The method of claim 2, further comprising:

analyzing the signal features to assign either a stable value or an unstable value to each of the microsatellite loci;

calculating a weighted sum across the assigned stable and unstable values of the microsatellite loci; and

assigning the biological sample a high microsatellite instability status if the weighted sum across the microsatellite loci exceeds a first predetermined threshold, assigning the biological sample a low microsatellite instability status if

the weighted sum across the microsatellite loci exceeds a second predetermined threshold but not the first predetermined threshold, or a microsatellite stable status if the weighted sum across the microsatellite loci is less than the second predetermined threshold.

16. The method of claim 15, wherein the weighted sum is calculated using one or more classification functions which map the plurality of signal features to three distinct output values.

17. A method for identifying microsatellite instability in a biological sample, comprising:

obtaining a plurality of signals by detecting fluorescence of fragments comprising nucleic acid sequences amplified from the biological sample, the nucleic acid sequences corresponding to a plurality of different microsatellite loci wherein each signal corresponds to one of a plurality of different microsatellite loci; and

analyzing the plurality of signals using one or more classifiers to identify whether the biological sample has high microsatellite instability, low microsatellite instability, or is microsatellite stable.

18. The method of claim 17, wherein the classifier comprises a non-linear classification function.

19. The method of claim 18, wherein the non-linear classification function comprises a multi-layer artificial neural network.
20. The method of claim 17, wherein the classifier comprises a deep learning neural network.
21. A computer program product comprising executable code stored in a non-transitory computer readable medium executable on one or more computer processors to identify microsatellite instability in a biological sample, the executable code comprising one or more computer readable instructions for:
- obtaining a plurality of signals from a capillary electrophoresis genetic analysis instrument, wherein the signals are detected from fluorescence of fragments comprising nucleic acid sequences amplified from the biological sample via polymerase chain reaction, the nucleic acid sequences corresponding to a plurality of different microsatellite loci wherein each signal corresponds to one of a plurality of different microsatellite loci; and
 - analyzing each of the plurality of signals using one or more classifiers to identify whether the biological sample has a microsatellite instability high, microsatellite instability low, or microsatellite stable status.

22. The computer program product of claim 21, wherein at least one classifier comprises an artificial intelligence generated classifier.

23. A system for identifying microsatellite instability in a biological sample using a capillary electrophoresis genetic analysis instrument, comprising:

one or more computer processors connected to a non-transitory computer readable medium storing one or more computer readable instructions that, when executed by the one or more computer processors:

obtain a plurality of signals from a capillary electrophoresis genetic analysis instrument, wherein the signals are detected from fluorescence of fragments comprising nucleic acid sequences amplified from the biological sample via polymerase chain reaction, the nucleic acid sequences corresponding to a plurality of different microsatellite loci wherein each signal corresponds to one of a plurality of different microsatellite loci; and

analyze each of the plurality of signals using one or more classifiers to identify whether the biological sample has a microsatellite instability high, microsatellite instability low, or microsatellite stable status;

a memory connected to at least one of the one or more processors for storing one or more of the signal features; and

a user device display connected to the memory and configured to display one or more of the signal features.

24. The system of claim 23, wherein at least one classifier comprises an artificial intelligence generated classifier.

25. A kit for identifying microsatellite instability in a biological sample, the kit comprising:

a plurality of polymerase chain reaction (PCR) primers configured to flank a plurality of microsatellite loci of a biological sample such that, when the PCR primers and the biological sample are combined and subjected to an amplification process, fluorescently labeled DNA fragments are generated comprising the plurality of microsatellite loci, wherein at least some of the plurality of microsatellite loci are different from others of the plurality of microsatellite loci; and

a computer program product embedded in a non-transitory computer readable medium comprising executable instruction code that, when executed by one or more processors causes the one or more processors to perform processing comprising:

using fluorescent data obtained from the plurality of fluorescently labelled microsatellite loci to classify microsatellite instability of the biological sample.

26. The kit of claim 25, wherein execution of the executable instruction code causes the one or more processors to perform processing comprising:

applying one or more locus-specific algorithms to the fluorescent data to generate locus-specific results classifying microsatellite instability of corresponding specific loci of the plurality of microsatellite loci; and

using the locus-specific results to classify the microsatellite instability of the biological sample.

27. A method of treating a patient having a tumor with an anti-tumor drug, the anti-tumor drug being more likely to be effective when cells of the tumor exhibit high microsatellite instability, the method comprising:

obtaining a biological sample from the patient, the biological sample comprising cells of the tumor;

determining whether the cells of the tumor exhibit high microsatellite instability using the method of any one of claims 1-20; and

if the microsatellite instability of the cells of the tumor is determined to be high, administering the anti-tumor drug to the patient to treat the tumor.

28. A method of treating a patient having a tumor with an anti-tumor drug, the anti-tumor drug being more likely to be effective when cells of the tumor exhibit high microsatellite instability, the method comprising:

obtaining a biological sample from the patient, the biological sample comprising cells of the tumor;

determining whether the cells of the tumor exhibit high microsatellite instability using the computer program product of any one of claims 21-22; and

if the microsatellite instability of the cells of the tumor is determined to be high, administering the anti-tumor drug to the patient to treat the tumor.

29. A method of treating a patient having a tumor with an anti-tumor drug, the anti-tumor drug being more likely to be effective when cells of the tumor exhibit high microsatellite instability, the method comprising:

obtaining a biological sample from the patient, the biological sample comprising cells of the tumor;

determining whether the cells of the tumor exhibit high microsatellite instability using the system of any one of claims 23-24; and

if the microsatellite instability of the cells of the tumor is determined to be high, administering the anti-tumor drug to the patient to treat the tumor.

30. A method of treating a patient having a tumor with an anti-tumor drug, the anti-tumor drug being more likely to be effective when cells of the tumor exhibit high microsatellite instability, the method comprising:

obtaining a biological sample from the patient, the biological sample comprising cells of the tumor;

determining whether the cells of the tumor exhibit high microsatellite instability using the kit of any one of claims 25-26; and

if the microsatellite instability of the cells of the tumor is determined to be high, administering the anti-tumor drug to the patient to treat the tumor.

31. The method of claim 1, wherein applying the one or more classifiers comprises analyzing peak pattern input comprising two or more values of: peak amplitudes along the fragment size axis, peak locations along the fragment size axis; peak amplitudes relative to a largest peak, peak locations relative to a largest peak, peak envelope peak amplitudes, peak envelope peak locations, peak envelope peak widths, or peak metrics relative to peak envelope metrics.

32. The method of claim 31, further comprising measuring the deviations of peak pattern input for the biological sample relative to peak pattern input from normal tissue samples.

33. The method of claim 31, further comprising measuring the deviations of peak pattern input for the biological sample relative to nominal peak pattern values calculated across a population of individuals without cancer.

34. The method of claim 31, further comprising measuring the deviations of peak pattern input for the biological sample relative to peak pattern input from normal tissue samples, as compared to nominal peak pattern values calculated across a population of individuals without cancer.

35. The method of claim 31, further comprising computing a difference signal between normalized data from tumor cells of the biological sample and normalized data from normal non-tumor cells of the biological sample and deriving peak patterns based on these difference signals.