

herein provide very wide coverage of reasonable methods to automatically make MSI status calls in many cases. The methods described can also be used to assign a confidence metric to the calls, for example, by reporting the proximity of calculated results to decision thresholds, which, in turn, can be used to focus human review efforts on those cases where the automated MSI assessment is less confident.

[0011a] An aspect of the present invention provides 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.

[0011b] Another aspect of the present invention provides 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.

[0011c] Another aspect of the present invention provides 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.

[0011d] Another aspect of the present invention provides 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.

[0011e] Another aspect of the present invention provides 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.

[0012] Embodiments of the invention used to detect microsatellite instability in a biological sample are disclosed. Signal data is received from a capillary electrophoresis genetic analysis instrument, wherein the signal data is measured from fluorescence of fragments comprising nucleic acid sequences amplified from the biological sample via polymerase chain reaction. The nucleic acid sequences correspond to a plurality of different microsatellite loci. Different loci can exhibit different signal characteristics. At a particular locus, a hierarchy of analysis methods can be applied that may also be peculiar to the characteristics of the signal data at that locus. For example, a three-level hierarchy could be described as follows: A first processing algorithm is implemented to obtain a first determination, based on the signal data, regarding instability of one or more first microsatellite loci of the plurality of different microsatellite loci. A second processing algorithm is implemented to obtain a second determination, based on the signal data, regarding instability of one or more second microsatellite loci of the plurality of different microsatellite loci. A third processing algorithm is then implemented to measure