THAT WHICH IS CLAIMED:

1. A method of analyzing data for one or more samples, the data for each sample being obtained from a component separation and tandem mass spectrometer system comprising a separation portion, including a liquid chromatograph, a gas chromatograph, a supercritical fluid chromatograph, or a capillary electrophoresis analyzer, and a first mass spectrometry step or provision (MS1) and a second mass spectrometry step or provision (MS2), wherein the data from the MS1 includes MS1 sample components and the data from the MS2 includes MS2 sample components, said method comprising:

analyzing, for each sample, a data set for a plurality of MS2 sample components to determine a precursor ion mass-to-charge ratio (m/z) and a retention index (RI) for each MS2 sample component;

comparing the precursor ion m/z and the RI for each MS2 sample component to precursor ion mass-to-charge ratios and retention indices of known compounds and removing any MS2 sample component from the data set corresponding to one of the known compounds, with remaining MS2 sample components in the data set being candidate MS2 sample components;

forming component clusters across the candidate MS2 sample components, each component cluster including candidate MS2 sample components each having the precursor ion m/z and the RI within respective ranges for the component cluster;

for each sample within each component cluster, retrieving one or more MS1 sample components within the respective precursor ion m/z and RI ranges for the component cluster;

for each component cluster, determining at most one consensus MS1 sample component, by aggregating the one or more MS1 sample components within the respective precursor ion m/z and RI ranges, to represent each sample, and associating the consensus MS1 sample component with a corresponding consensus MS2 sample component, consensus precursor ion m/z, and consensus RI determined by aggregating the MS2 sample components and associated precursor ion m/z and the RI within the component cluster;

for each component cluster, determining whether the consensus MS1 sample component indicates a molecular ion or a derivative relationship to the molecular ion;

grouping component clusters according to consensus RI and selecting one or more component clusters from each group of component clusters, the one or more component clusters being candidate component clusters; and

correlating the consensus precursor ion m/z, the consensus RI, the consensus MS1 sample component, and the consensus MS2 sample component of the candidate component clusters with an unknown compound in the samples.

2. The method of claim 1, wherein removing any MS2 sample component from the data set corresponding to one of the known compounds comprises removing any of a known molecular ion, a

background artifact, a chromatographic artifact, or a derivative relationship of the known molecular ion, a mass relationship of the known molecular ion, the background artifact, or the chromatographic artifact, corresponding to one of the known compounds, from the data set.

3. The method of claim 2, wherein removing the derivative relationship of the known molecular ion or the mass relationship of the known molecular ion from the data set comprises removing any of an in-source fragment, an adduct, an isotope, a dimer, an oligomer, a multiple charged species, an adduct of an isotope, or an oligomer of an isotope, from the data set.

4. The method of claim 1, comprising determining precursor ion mass-to-charge ratios and retention indices of the known compounds from the MS1 sample components thereof.

5. The method of claim 4, comprising associating any of the precursor ion m/z, the RI, the MS1 sample components, and the MS2 sample components of each known compound with the respective known compound in a database comprising an ion data repository.

6. The method of claim 1, comprising associating the precursor ion m/z, the RI, MS1 sample components, and the MS2 sample components of each candidate component cluster with the unknown compound in a database comprising an ion data repository.

7. The method of claim 1, comprising determining the retention index by normalizing a retention time for each MS2 sample component.

8. The method of claim 1, wherein forming component clusters across the candidate MS2 sample components comprises forming component clusters across the candidate MS2 sample components using divisive clustering or agglomerative clustering.

9. The method of claim 1, comprising sorting the candidate MS2 sample components by precursor ion mass prior to forming the component clusters across the candidate MS2 sample components.

10. The method of claim 9, comprising sorting the component clusters by retention index to generate secondary component clusters.

11. The method of claim 1, wherein the step of forming component clusters across the candidate MS2 sample components is repeated two or more times.

12. The method of claim 1, wherein comparing the consensus MS1 sample component for each component cluster, across the component clusters, is performed across a plurality of the samples.

13. The method of claim 12, comprising removing any component clusters that are present in less than 5% of the plurality of samples.

14. The method of Claim 1, comprising analyzing a plurality of possible derivative relationships or possible mass relationships of the MS1 sample components prior to associating the consensus MS1 sample component with a corresponding consensus precursor ion m/z.

15. The method of claim 14, wherein analyzing the plurality of possible derivative relationships or the plurality of possible mass relationships comprises analyzing the plurality of possible derivative relationships or possible mass relationships using symbolic algebra in a single run to generate computer code for analyzing a set of specified derivative relationships or mass relationships.

16. The method of claim 1, comprising removing any component clusters lacking a plausible molecular formula, prior to grouping the component clusters.

17. The method of claim 1, comprising removing any component clusters lacking statistical significance from metadata of the data of the one or more samples being analyzed, prior to grouping the component clusters.

18. The method of claim 1, comprising prioritizing candidate component clusters having statistical significance with respect to metadata of the data of the one or more samples being analyzed as candidate unknown compounds.

19. An apparatus for analyzing data for one or more samples, the data for each sample being obtained from a component separation and tandem mass spectrometer system comprising a separation portion, including a liquid chromatograph, a gas chromatograph, a supercritical fluid chromatograph, or a capillary electrophoresis analyzer, and a first mass spectrometry step or provision (MS1) and a second mass spectrometry step or provision (MS2), the apparatus comprising a processor and a memory storing executable instructions that, in response to execution by the processor, cause the apparatus to at least perform the method steps of any one of Claims 1 to 19.

20. A computer program product for analyzing data for one or more samples, the data for each sample being obtained from a component separation and tandem mass spectrometer system comprising a separation portion, including a liquid chromatograph, a gas chromatograph, a supercritical fluid

chromatograph, or a capillary electrophoresis analyzer, and a first mass spectrometry step (MS1) and a second mass spectrometry step (MS2), the computer program product comprising at least one non-transitory computer readable storage medium having computer-readable program code stored thereon, the computer-readable program code comprising program code for performing the method steps of any one of Claims 1 to 19.

21. A computer program product for analyzing data for one or more samples, the data for each sample being obtained from a component separation and tandem mass spectrometer system comprising a separation portion, including a liquid chromatograph, a gas chromatograph, a supercritical fluid chromatograph, or a capillary electrophoresis analyzer, and a first mass spectrometry step (MS1) and a second mass spectrometry step (MS2), the computer program product comprising computer-readable program code for performing the method steps of any one of Claims 1 to 19.