CLAIMS

1. A method of detecting a pathogen in a biological sample, comprising:

receiving sequence data from a sequencing device and generated from a biological sample;

identifying k-mers in the sequence data that have an exact match in a hash table that is initialized with a first set of k-mers comprising pathogen k-mers in a genome of the pathogen and with a second set of k-mers comprising control k-mers derived from a human genome;

providing a positive detection output for the biological sample based at least in part on a first count of exact matches of the k-mers in the sequence data with the first set being above a first threshold and a second count of exact matches of the k-mers in the sequence data with the second set being above a second threshold; and

terminating identifying the k-mers of the biological sample in response to providing the positive detection output, wherein the positive detection output is provided while sequence data from the biological sample is being generated.

2. The method of claim 1, wherein the k-mers in the sequence data, the first set, and the second set are of a fixed size that is greater than 24 nucleotides.

3. The method of claim 1, wherein the first set of k-mers is a subset of all k-mers in the genome of the pathogen.

4. The method of claim 3, wherein the subset is based on sufficient nonsimilarity to the human genome.

5. The method of claim 1, wherein the control k-mers are based on mRNA sequences expressed in a tissue of the biological sample.

6. The method of claim 1, wherein the first set of k-mers comprises variants in the genome of the pathogen.

7. The method of claim 1, wherein the first set is larger than the second set.

8. The method of claim 1, wherein the first set comprises k-mers from a plurality of different pathogens, and wherein the positive detection output for the biological sample is indicative of the pathogen of the plurality of different pathogens.

9. The method of claim 1, comprising aligning the sequence data to the genome of the pathogen based on the positive detection output.

10. The method of claim 9, comprising identifying sequence variants of the pathogen in aligned sequence data.

11. The method of claim 1, comprising providing a recommendation for treatment based on the positive detection output.