WHAT IS CLAIMED IS:

1. A method for determining a replication factor C subunit 1 (RFC1) repeat expansion status comprising:

under control of a hardware processor:

(a) receiving a plurality of sequence reads generated from a sample obtained from a subject;

(b) aligning the plurality of sequence reads to a sequence graph to generate a plurality of aligned sequence reads, wherein the sequence graph represents a locus of RFC1 and comprises a repeat sequence representation flanked by non-repeat sequences of the locus of RFC1, and wherein the plurality of aligned sequence reads comprises the plurality of sequence reads and alignments of the plurality of sequence reads to the sequence graph;

(c) determining a number of occurrences of a plurality of repeat sequences in aligned sequence reads of the plurality of aligned sequence reads using a first occurrence threshold and a first quality threshold;

(d) determining a frequency indication of a number of occurrences of a pathogenic repeat sequence relative to a total number of occurrences of the plurality of repeat sequences;

(e) determining a repeat expansion status at the locus of RFC1 of the subject using the frequency indication of the number of occurrences of the pathogenic repeat sequence relative to the total number of occurrences of the plurality of repeat sequences.

2. The method of claim 1, wherein the repeat expansion at the locus of RFC1 is at about chr4:39348424 of hg38, or a corresponding position of another reference genome sequence.

3. The method of any one of claims 1-2, comprising: determining the subject has zero, one, or two alleles with a repeat expansion at the locus of RFC1 using the plurality of aligned sequence reads.

4. The method of any one of claims 1-3, wherein the subject has two alleles with repeat expansion at the locus of RFC1.

5. The method of claim 4, wherein determining the repeat expansion status at the locus of RFC1 comprises:

determining the frequency indication of the number of occurrences of the pathogenic repeat sequence relative to the total number of occurrences of the plurality of repeat sequences is greater than a first status threshold; and

determining the repeat expansion status at the locus of RFC1 as pathogenic status.

6. The method of claim 4, wherein determining the repeat expansion status at the locus of RFC1 comprises:

determining the frequency indication of the number of occurrences of the pathogenic repeat sequence relative to the total number of occurrences of the plurality of repeat sequences is less than or equal to a first status threshold and greater than or equal to a second status threshold; and

determining the repeat expansion status at the locus of RFC1 as carrier status.

7. The method of claim 4, wherein determining the repeat expansion status at the locus of RFC1 comprises:

determining the frequency indication of the number of occurrences of the pathogenic repeat sequence relative to the total number of occurrences of the plurality of repeat sequences is less than or equal to a first status threshold and greater than or equal to a second status threshold;

determining a frequency indication of a number of occurrences of the pathogenic repeat sequence and a sequence with a high sequence similarity to the pathogenic repeat sequence is greater than a third status threshold, optionally wherein the pathogenic repeat sequence and the sequence similar to the pathogenic repeat sequence differs by one base;

determining the frequency indication of the number of occurrences of the pathogenic sequence is greater than a fourth status threshold; and

determining the repeat expansion status at the locus of RFC1 as carrier status.

8. The method of claim 4, wherein determining the repeat expansion status at the locus of RFC1 comprises:

determining the frequency indication of the number of occurrences of the pathogenic repeat sequence relative to the total number of occurrences of the plurality of repeat sequences is less than a second status threshold; and

determining the repeat expansion status at the locus of RFC1 as benign status.

9. The method of any one of claims 1-3, wherein the subject has one allele of RFC1 with repeat expansion at the locus of RFC1.

10. The method of claim 9,

wherein determining the number of occurrences of the plurality of repeat sequences in the aligned sequence reads of the plurality of sequence reads comprises: determining the number of occurrences of the plurality of repeat sequences in the aligned sequence reads of the plurality of aligned sequence reads that are from the one allele of RFC1 with repeat expansion at the locus of RFC1; and

wherein determining the repeat expansion status at the locus of RFC1 comprises:

determining the frequency indication of the number of occurrences of the pathogenic repeat sequence relative to the total number of occurrences of the plurality of repeat sequences is greater than a first status threshold; and

determining the repeat expansion status at the locus of RFC1 as pathogenic status.

11. The method of claim 9,

wherein determining the number of occurrences of the plurality of repeat sequences in the aligned sequence reads of the plurality of sequence reads comprises: determining the number of occurrences of the plurality of repeat sequences in the aligned sequence reads of the plurality of aligned sequence reads that are from the one allele of RFC1 with repeat expansion at the locus of RFC1; and

wherein determining the repeat expansion status at the locus of RFC1 comprises:

determining the frequency indication of the number of occurrences of the pathogenic repeat sequence relative to the total number of occurrences of the plurality of repeat sequences is less than or equal to a first status threshold; and

determining the repeat expansion status at the locus of RFC1 as carrier status.

12. The method of any one of claims 1-3, wherein the subject has zero allele with repeat expansion at the locus of RFC1, and wherein the repeat expansion status at the locus of RFC1 is benign status.

13. The method of any one of claims 4-11, wherein the repeat expansion at the locus of RFC1 comprises greater than a threshold total copies of one or more repeat sequences.

14. A system for determining a repeat expansion status of a gene of interest comprising:

non-transitory memory configured to store executable instructions; and

a hardware processor in communication with the non-transitory memory, the hardware processor programmed by the executable instructions to perform:

(a) receiving a plurality of sequence reads generated from a sample obtained from a subject;

(b) aligning the plurality of sequence reads to a sequence graph to generate a plurality of aligned sequence reads, wherein the sequence graph represents a locus of a gene of interest and comprises a repeat sequence representation flanked by non-repeat sequences of the locus of the gene, and wherein the plurality of aligned sequence reads comprises the plurality of sequence reads and alignments of the plurality of sequence reads to the sequence graph;

(c) determining a number of occurrences of a plurality of repeat sequences in aligned sequence reads of the plurality of aligned sequence reads using a first occurrence threshold and a first quality threshold;

(d) determining a frequency indication of a number of occurrences of a pathogenic repeat sequence relative to a total number of occurrences of the plurality of repeat sequence(e) determining a repeat expansion status at the locus of the gene of interest of the subject using the frequency indication of the number of occurrences of the pathogenic repeat sequence relative to the total number of occurrences of the plurality of repeat sequences.

15. The system of claim 14, wherein the gene of interest is replication factor C subunit 1 (RFC1).

16. The system of claim 15, wherein the locus of the gene of interest with repeat expansion is at about chr4:39348424 of hg38, or a corresponding position of another reference genome sequence.

17. The system of any one of claims 15-16, wherein the repeat expansion is associated with or causes a disease, optionally wherein the disease is cerebellar ataxia, neuropathy, and vestibular areflexia syndrome (CANVAS).

18. The system of any one of claims 15-17, wherein the repeat sequence representation is AARRG.

19. The system of any one of claims 15-18, wherein the pathogenic repeat sequence is AAGGG or ACAGG.

20. The system of any one of claims 14-19, wherein the hardware processor is programmed by the executable instructions to perform: determining the subject has zero, one, or two alleles with a repeat expansion at the locus of the gene of interest using the plurality of aligned sequence reads.