Claims

1. A method for detecting a genetic variant in a region of interest in a DNA sample comprising:

   (i) establishing, for a given sequencing platform, sequencing process and sequencing depth, a threshold frequency at or above which the genetic variant must be observed in sequencing results of amplification reactions to assign a positive determination for the presence of the genetic variant in a given amplification reaction, wherein the threshold frequency is based on the distribution of the number of reads supporting the genetic variant expected to be observed in the sequencing results of amplification reactions due to amplification and sequencing error (read count distribution);
   (ii) partitioning the DNA sample into a plurality of replicate amplification reactions, so that, for the genetic variant, the mean number of amplifiable template molecules of the region of interest in a replicate amplification reaction is fewer than the reciprocal of the threshold frequency determined in step (i);
   (iii) performing the amplification reactions of step (ii) and sequencing the products of amplification reactions;
   (iv) based on step (i) and the results of step (iii), determining the presence/absence of the genetic variant in each replicate amplification reaction; and
   (v) integrating the results of (iv) to determine the presence/absence of the genetic variant in the region of interest in the DNA sample.

2. The method according to claim 1, wherein for the genetic variant to be determined as being present in the region of interest in the DNA sample in step (v), a positive determination for the presence of the genetic variant must be made in more than one replicate amplification reaction in step (iv), optionally wherein a positive determination for the presence of the genetic variant must be made in at least 3 replicate amplification reactions in step (iv).

3. The method according to any one of the preceding claims, wherein the threshold frequency of step (i) is the frequency at which the cumulative distribution function value for the probability of observing the genetic variant at that frequency is 0.99 or greater, 0.995 or greater, 0.999 or greater, 0.9999 or greater, or 0.99999 or greater.

4. The method according to any one of the preceding claims, wherein threshold frequency of step (i) is determined using a binomial, over-dispersed binomial, Beta, Normal, Exponential or Gamma probability distribution model.

5. The method according to any one of the preceding claims, wherein step (i) comprises:

   (a) based on the read count distribution for a plurality of genetic variants, establishing a plurality of threshold frequencies at or above which the genetic variants should be observed in sequencing results of amplification reactions to assign a positive determination for the presence of the genetic variant in a given amplification reaction; and
   (b) based on step (a), establishing an overall threshold frequency at or above which a genetic variant must be observed in sequencing results of a given amplification reaction to assign a positive determination for the presence of the genetic variant in that amplification reaction, which is the threshold frequency at which at least 90%, 95%, 97.5%, 99% or more of the plurality of threshold frequencies determined in step (a) are less than this value.

6. The method according to any one of the preceding claims, wherein the mean number of amplifiable template molecules to be present in a replicate amplification reaction is the number at which when the genetic variant is present in a single amplifiable template molecule of a replicate amplification reaction, the probability that a positive determination will be made for that replicate is 0.9 or greater.

7. The method according to any one of the preceding claims, wherein the method comprises a step of determining, for a given sequencing platform, sequencing process and sequencing depth, the distribution of the number of reads supporting the genetic variant or a plurality of genetic variants expected to be observed in the sequencing results of amplification reactions due to amplification and sequencing error (read count distribution),
optionally wherein the method of determining the read count distribution further comprises sequencing a DNA sample multiple times to determine the read count distribution for the genetic variant or a plurality of genetic variants.

8. The method according to any one of claims 1 to 6, wherein the read count distribution is based on sequencer and/or polymerase error rates, optionally taking sequence context into account.

9. The method according to any one of claims 1 to 5, wherein the method comprises 'looking-up' a reference value or plurality of reference values for the read count distribution for the genetic variant or a plurality of genetic variants in a database, chart, table, list, catalogue, index, directory, or register, optionally wherein the reference value or plurality of reference values were determined by sequencing a DNA sample multiple times, to determine the read count distribution for the genetic variant or a plurality of genetic variants.

10. The method according to any one of the preceding claims, wherein following partitioning of the DNA sample in step (ii), each replicate amplification reaction has more than a single amplifiable template molecule for the region of interest per replicate amplification reaction, optionally wherein the mean number of amplifiable template molecules per replicate amplification reaction is more than 1 and less than 1000, more than 2 and less than 1000, or more than 5 and less than 1000.

11. The method according to any of the preceding claims, which is capable of detecting a genetic variant which is present within the population of amplifiable template molecules of a DNA sample at a frequency of less than 2%, less than 1%, less than 0.5%, less than 0.2%, less than 0.1% or less than 0.05%.

12. The method according to any one of the preceding claims, wherein partitioning the DNA sample into a plurality of replicate amplification reactions comprises diluting the DNA sample and aliquoting into replicate amplification reactions.

13. The method according to any one of the preceding claims, wherein replicate amplification reactions are performed in parallel and sequencing of the products of amplification reactions is performed in parallel,
optionally wherein amplification reactions are performed using one or more primers flanking the region of interest and which integrate sample and/or amplification reaction replicate specific identifier sequences into the products of amplification,
and further optionally wherein the primers integrate sequence adapters into the products of amplification.

14. The method according to any one of the preceding claims, wherein a plurality of regions of interest are analysed in parallel.

15. The method according to any one of the preceding claims, wherein the method further comprises:

    (vi) determining the frequency of the genetic variant in the DNA sample.

16. Use of the method according to any one of claims 1 to 15:

    to detect and/or quantify circulating tumour DNA in a sample obtained from a subject;
    in an in vitro method of diagnosis or prognosis or monitoring a disease;
    in a method of selecting a patient for a therapy;
    in a method of identifying genetic variation predictive of susceptibility, resistance or response to a therapy; or
    to monitor or evaluate response to a therapy.