Method for Rapidly Constructing Amplicon Library through One-step Process

The present invention claims priority of a Chinese patent application filed with the China Patent Office by the Genetron Health(Beijing) Co., Ltd. on April 5, 2017, with the application number of 201710218529.4, and entitled "method for rapidly constructing amplicon library through one-step process". The entire content of this application are incorporated herein by reference.

TECHNICAL FIELD

The invention relates to a field of biotechnology, and in particular to a method for rapidly constructing amplicon library through one-step process.

BACKGROUND OF RELATED ART

Next-generation sequencing (NGS) has been widely used in disease research, diagnosis and treatment in recent years due to its high throughput, high sensitivity, and high automation. Compared with traditional detection method, NGS technology can achieve multi-gene parallel detection and save samples. Besides, it has higher sensitivity which can restore the panoramic view of tumor variation in a more realistic way. However, the traditional method for constructing an amplicon library in the Life NGS platform is cumbersome, requires PCR amplification, digestion, addition, and purification, and takes about 5 hours. Further, because of the need to open the lid in a multi-step operation, the library is easily contaminated and the library loss rate is high. In addition, in the traditional method of constructing the amplicon library, the cost of establishing a library for a single sample is relatively high, which is about 200-1000 RMB per case.

The information disclosed in background part is only intended to enhance an understanding of the general background of the invention, and should not be taken as an acknowledgement or any form of suggestion that this information forms the prior art already known to a person skilled in the art.

SUMMARY

The object of the present disclosure is to provide a method for rapidly constructing an amplicon library through one-step process. The method can construct the amplicon library by one-step PCR in a simple and rapid manner, and since the barcode is introduced before the start of PCR, the possibility of cross-contamination between samples and libraries is greatly reduced, and the requirements of the experimental site partition can be

simplified. The method also controls the cost of establishing a single sample library at 30 RMB percase.

To achieve the above object, the present disclosure provides a method for constructing an amplicon library of a DNA sample, comprising the following steps.

Step 1: synthesizing a primer combination for constructing an amplicon library of a DNA sample, the primer combination of the amplicon library that is used to construct the DNA sample includes:

a forward fusion primer that is designed according to a target amplicon. The forward fusion primer comprises a first linker sequence (Bridge sequence) and a specific forward primer sequence designed according to the target amplicon, both of which are arranged in the order of 5' to 3';

a first linker sequence (Bridge sequence) and a specific forward primer sequence designed according to the target amplicon arranged in the order of 5' to 3';

a reverse fusion primer that is designed according to the target amplicon. The reverse fusion primer comprises a second linker sequence (trP1 sequence) and a specific reverse primer sequence designed according to the target amplicon, both of which are arranged in the order of 5' to 3';

a forward universal primer which comprises a third linker sequence (A sequence), a barcode sequence and a first linker sequence arranged in the order of 5' to 3'; and

a reverse universal primer which comprises a universal sequence (Uni sequence) and a second linker sequence arranged in the order of 5' to 3';

Step 2: constructing a PCR reaction system for the DNA samples, and mixing the forward fusion primers designed according to the target amplicon, the reverse fusion primers designed according to the target amplicon, the forward universal primers and the reverse universal primers together, to serve as a primer combination in the PCR reaction system;

Step 3: performing PCR.

In one aspect, the present disclosure provides a method for constructing an amplicon library of a DNA sample, comprising the following steps:

1) synthesizing a primer combination which is used for constructing an amplicon library of a DNA sample, wherein the primer combination of the amplicon library used to construct the DNA sample includes:

an forward fusion primer designed according to a target amplicon, the forward fusion primer comprising a first linker sequence (Bridge sequence) and a specific forward primer sequence designed according to the target amplicon, both of which are arranged in the order of 5' to 3';

a reverse fusion primer designed according to the target amplicon, the reverse fusion primer comprising a second linker sequence (trP1 sequence) and a specific reverse primer sequence designed according to the target amplicon, both of which are arranged in the order of 5' to 3';

an forward universal primer comprising a third linker sequence (A sequence), a barcode sequence and a first linker sequence arranged in the order of 5' to 3'; and

a reverse universal primer comprising a universal sequence (Uni sequence) and a second linker sequence arranged in the order of 5' to 3';

2) constructing a PCR reaction system for the DNA sample, and mixing the forward fusion primers designed according to the target amplicon, the reverse fusion primers designed according to the target amplicon, the forward universal primers and the reverse universal primers together, to serve as a primer combination in the PCR reaction system; and

3) performing PCR.

In an embodiment of the present disclosure, the first linker sequence comprises a sequence of SEQ ID: 1, and the nucleotide sequence of the sequence of SEQ ID: 1 is GGCATACGTCCTCGTCTA.

In an embodiment of the present disclosure, the second linker sequence comprises a sequence of SEQ ID: 2, and the nucleotide sequence of the sequence of SEQ ID: 2 is TCTATGGGCAGTCGGTGAT.

In an embodiment of the present disclosure, the third linker sequence comprises a sequence of SEQ ID:3, and a nucleotide sequence of the sequence of SEQ ID:3 is CCATCTCATCCCTGCGTGTCTCCGACTCAG.

In an embodiment of the present disclosure, the universal sequence comprises a sequence of SEQ ID: 4, and a nucleotide sequence of the sequence of SEQ ID: 4 is CCACTACGCCTCCGCTTTCCTC.

In an embodiment of the present disclosure, in the primer combination for constructing an amplicon library of the same DNA sample, the barcode sequence in the forward universal primer is the same. In the primer combinations for constructing amplicon libraries of the different DNA samples, the barcode sequences in the forward universal primers are different. The barcode sequence corresponds to the sample. The barcode sequence is different among different samples. As long as different samples can be distinguished, the barcode sequence is not specific and its sequence can be changed.

In an embodiment of the present disclosure, the concentration of the forward fusion primer designed according to any one of the target amplicon, the concentration of reverse fusion primer designed according to any one of the target amplicon, the concentration of forward universal primer, and the concentration of reverse universal primer are all 100 μM.

In an embodiment of the present disclosure, when the number of target amplicons in the same PCR reaction is greater than 1, the forward fusion primer designed according to a target amplicon is a combination of forward fusion primers designed according to each target amplicon, the reverse fusion primer designed according to the target amplicon is a combination of reverse fusion primers designed according to each target amplicon.

In an embodiment of the present disclosure, the molar ratio of the forward fusion primer designed according to any one of the target amplicon to the reverse fusion primer designed according to the target amplicon is 1:1; the molar ratio of the forward universal primer to the reverse universal primer is 1:1. The specific amount of forward universal primers and reverse universal primers should be adjusted according to the number of target amplicons during PCR amplification. For example, when PCR amplification, 5 target amplicons need to be amplified and 22 target amplicons need to be amplified, the specific amount of the forward universal primer and the reverse universal primer may be different, and a specific amount of the forward universal primer and the reverse universal primer may be determined by those skilled in the art according to conventional techniques in the art.

In an embodiment of the present disclosure, the DNA sample is genomic DNA.

In an embodiment of the present disclosure, the genomic DNA is extracted from a tissue sample or a formalin-fixed and paraffin-embedded sample.

In an embodiment of the present disclosure, the target amplicon comprises at least one selected from the group consisting of 22 target amplicons:

Chr2:29432588-29432707 (Hg19) amplicon of the ALK gene, the sequence of which is shown in SEQ ID:5:TAAGGGACAAGCAGCCACACCCCATTCTTGAGGGGCTGAGGTGGAAGAGACAGGCCCGGAGGGGTGAGGCAGTCTTTACTCACCTGTAGATGTCTCGGGCCATCCCGAAGTCTCCAATCTTGGCCACTCTTCCAGGGCCTGGACAGGTCAAGAGGCAGT;

Chr2: 29443616-29443730 (Hg19) amplicon of the ALK gene, the sequence of which is shown in SEQ ID:6:CGGAGGAAGGACTTGAGGTCTCCCCCCGCCATGAGCTCCAGCAGGATGAACCGGGGCAGGG ATTGCAGGCTCACCCCAATGCAGCGAACAATGTTCTGGTGGTTGAATTTGCTGCAGAGCAGAGAG GGATGTAACCAAAATTAACTGAGCTGAGTCTGG;

Chr7: 140453091-140453197(Hg19) amplicon of the BRAF gene, the sequence of which is shown in SEQ ID:7:CCTCAATTCTTACCATCCACAAAATGGATCCAGACAACTGTTCAAACTGATGGGACCCACTC CATCGAGATTTCACTGTAGCTAGACCAAAATCACCTATTTTTACTGTGAGGTCTTCATGAAGAAA TATATCTGAGGTGTAGTAAGTAAAGGAAAACAGTAG;

Chr7: 55241604-55241726(Hg19) amplicon of the EGFR gene, the sequence of which is shown in SEQ ID:8: TGACCCTTGTCTCTGTGTTCTTGTCCCCCCCAGCTTGTGGAGCCTCTTACACCCAGTGGAGAAGCT CCCAACCAAGCTCTCTTGAGGATCTTGAAGGAAACTGAATTCAAAAAGATCAAAGTGCTGGGCT CCGGTGCGTTCGGCACGGTGTATAAGGTAAGGTCCCTGG;

Chr7: 55242398-55242513 (Hg19) amplicon of the EGFR gene, the sequence of which is shown in SEQ ID:9: ACAATTGCCAGTTAACGTCTTCCTTCTCTCTCTGTCATAGGGACTCTGGATCCCAGAAGGTGAGA AAGTTAAAATTCCCGTCGCTATCAAGGAATTAAGAGAAGCAACATCTCCGAAAGCCAACAAGGA AATCCTCGATGTGAGTTTCTGCTTTGCTGTGT;

Chr7: 55248970-55249096 (Hg19) amplicon of the EGFR gene, the sequence of which is shown in SEQ ID:10:GAAGCCACACTGACGTGCCTCTCCCTCCCTCCAGGAAGCCTACGTGATGGCCAGCGTGGAC AACCCCCACGTGTGCCGCCTGCTGGGCATCTGCCTCACCTCCACCGTGCAGCTCATCACGCAGCT CATGCCCTTCGGCTGCCTCCTGGACTATGTCCGGGAACAC;

Chr7: 55259505-55259621 (Hg19) amplicon of the EGFR gene, the sequence of which is shown in SEQ ID:11:CCGCAGCATGTCAAGATCACAGATTTTGGGCTGGCCAAACTGCTGGGTGCGGAAGAGAAA GAATACCATGCAGAAGGAGGCAAAGTAAGGAGGTGGCTTTAGGTCAGCCAGCATTTTCCTGACA CCAGGGACCAGGCTGCCTTCCCACTAGCTGTATTGTTTA;

Chr17: 37880969-37881082 (Hg19) amplicon of the ERBB2 gene, the sequence of which is shown in SEQ ID:12:CATACCCTCTCAGCGTACCCTTGTCCCCAGGAAGCATACGTGATGGCTGGTGTGGGCTCCC CATATGTCTCCCGCCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTTATG CCCTATGGCTGCCTCTTAGACCATGTCCG;

Chr12: 25380261-25380363 (Hg19) amplicon of the KRAS gene, the sequence of which is shown in SEQ ID:13:AGTCCTCATGTACTGGTCCCTCATTGCACTGTACTCCTCTTGACCTGCTGTGTCGAGAATAT CCAAGAGACAGGTTTCTCCATCAATTACTACTTGCTTCCTGTAGGAATCCTGAGAAGGGAGAAAC ACAGTCTGGATTATTACAGTGCA;

Chr12: 25398183-25398310 (Hg19) amplicon of the KRAS gene, the sequence of which is shown in SEQ ID:14:AAAGAATGGTCCTGCACCAGTAATATGCATATTAAAACAAGATTTACCTCTATTGTTGGAT CATATTCGTCCACAAAATGATTCTGAATTAGCTGTATCGTCAAGGCACTCTTGCCTACGCCACCA GCTCCAACTACCACAAGTTTATATTCAGTCATTTTCAGCAGGCCTT;

Chr7: 116340233-116340335 (Hg19) amplicon of the MET gene, the sequence of which is shown in SEQ ID:15:TCGATCTGCCATGTGTGCATTCCCTATCAAATATGTCAACGACTTCTTCAACAAGATCGTCA ACAAAAACAATGTGAGATGTCTCCAGCATTTTTACGGACCCAATCATGAGCACTGCTTTAATAGG GTAAGTCACATCAGTTCCC;

Chr7: 116411880-116412005 (Hg19) amplicon of the MET gene, the sequence of which is shown in SEQ ID:16:CCATGATAGCCGTCTTTAACAAGCTCTTTCTTTCTCTCTGTTTTAAGATCTGGGCAGTGAAT TAGTTCGCTACGATGCAAGAGTACACACTCCTCATTTGGATAGGCTTGTAAGTGCCCGAAGTGTA AGCCCAACTACAGAAATGGTTTCAAATGAATCTGTAGACTACCGAGCT;

Chr7: 116417426-116417546 (Hg19) amplicon of the MET gene, the sequence of which is shown in SEQ ID:17:ATGTTACGCAGTGCTAACCAAGTTCTTTCTTTTGCACAGGGCATTTTGGTTGTGTATATCAT GGGACTTTGTTGGACAATGATGGCAAGAAAATTCACTGTGCTGTGAAATCCTTGAACAGTAAGTG GCATTTTATTTAACCATGGAGTATACTTTTGTGGTTTGCAAC;

Chr7: 116423399-116423499 (Hg19) amplicon of the MET gene, the sequence of which is shown in SEQ ID:18:CAGTCAAGGTTGCTGATTTTGGTCTTGCCAGAGACATGTATGATAAAGAATACTATAGTGT ACACAACAAAACAGGTGCAAAGCTGCCAGTGAAGTGGATGGCTTTGGAAAGTCTGCAAACTCAA AAGTTTACCACCAAGTCAGATGTG;

Chr1: 115256507-115256586 (Hg19) amplicon of the NRAS gene, the sequence of which is shown in SEQ ID:19:TTCGCCTGTCCTCATGTATTGGTCTCTCATGGCACTGTACTCTTCTTGTCCAGCTGTATCCAG TATGTCCAACAAACAGGTTTCACCATCTATAACCACTTGTTTTCTGTAAGAATCCTGGGGGTG;

Chr1: 115258651-115258755 (Hg19) amplicon of the NRAS gene, the sequence of which is shown in SEQ ID:20:TGAGAGACAGGATCAGGTCAGCGGGCTACCACTGGGCCTCACCTCTATGGTGGGATCATAT TCATCTACAAAGTGGTTCTGGATTAGCTGGATTGTCAGTGCGCTTTTCCCAACACCACCTGCTCCA ACCACCACCAGTTTGTACTCAG;

Chr3: 178936056-178936179 (Hg19) amplicon of the PIK3CA gene, the sequence of which is shown in SEQ ID:21:GGAAAATGACAAAGAACAGCTCAAAGCAATTTCTACACGAGATCCTCTCTCTGAAATCACT GAGCAGGAGAAAGATTTTCTATGGAGTCACAGGTAAGTGCTAAAATGGAGATTCTCTGTTTCTTT TTCTTTATTACAGAAAAAATAACTGAATTTGGCTGATCTCAGCATGTT;

Chr3: 178952000-178952092 (Hg19) amplicon of the PIK3CA gene, the sequence of which is shown in SEQ

ID:22:ATGCCAGAACTACAATCTTTTGATGACATTGCATACATTCGAAAGACCCTAGCCTTAGATA AAACTGAGCAAGAGGCTTTGGAGTATTTCATGAAACAAATGAATGATGCACATCATGGTGGCTG GACAACAAAAATGGATTG;

Chr17: 7577027-7577154 (Hg19) amplicon of the TP53 gene, the sequence of which is shown in SEQ ID:23:CTTCTTGTCCTGCTTGCTTACCTCGCTTAGTGCTCCCTGGGGGCAGCTCGTGGTGAGGCTCC CCTTTCTTGCGGAGATTCTCTTCCTCTGTGCGCCGGTCTCTCCCAGGACAGGCACAAACACGCACC TCAAAGCTGTTCCGTCCCAGTAGATTACCACTACTCAGGATAGGAAAAGAG;

Chr17: 7577507-7577613 (Hg19) amplicon of the TP53 gene, the sequence of which is shown in SEQ ID:24:GCAAGTGGCTCCTGACCTGGAGTCTTCCAGTGTGATGATGGTGAGGATGGGCCTCCGGTTC ATGCCGCCCATGCAGGAACTGTTACACATGTAGTTGTAGTGGATGGTGGTACAGTCAGAGCCAAC CTAGGAGATAACACAGGCCCAAGA;

Chr17: 7578182-7578298 (Hg19) amplicon of the TP53 gene, the sequence of which is shown in SEQ ID:25:CCCCAGTTGCAAACCAGACCTCAGGCGGCTCATAGGGCACCACCACACTATGTCGAAAAGT GTTTCTGTCATCCAAATACTCCACACGCAAATTTCCTTCCACTCGGATAAGATGCTGAGGAGGGG CCAGACCTAAGAGCAATCAGTGAGGAATCAGAGG;

Chr17: 7578389-7578537 (Hg19) amplicon of the TP53 gene, the sequence of which is shown in SEQ ID:26:ACCATCGCTATCTGAGCAGCGCTCATGGTGGGGGCAGCGCCTCACAACCTCCGTCATGTGC TGTGACTGCTTGTAGATGGCCATGGCGCGGACGCGGGTGCCGGGCGGGGGTGTGGAATCAACCC ACAGCTGCACAGGGCAGGTCTTGGCCAGTTGGCAAAACATCTTGTTGAGGGCAGGGGAGTACTG.

In an embodiment of the present disclosure, the specific forwardforward primer sequence designed according to the Chr2:29432588-29432707 (Hg19) amplicon of the ALK gene is shown as SEQ ID: 27: ACTGCCTCTTGACCTGTCC; the specific reverse primer sequence designed according to the Chr2:29432588-29432707 (Hg19) amplicon of the ALK gene is shown as SEQ ID: 28: TAAGGGACAAGCAGCCACAC.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr2: 29443616-29443730 (Hg19) amplicon of ALK gene is shown as SEQ ID: 29: CCAGACTCAGCTCAGTTAATTTTGG; the specific reverse primer sequence designed according to the Chr2: 29443616-29443730 (Hg19) amplicon of the ALK gene is shown as SEQ ID: 30: CGGAGGAAGGACTTGAGGT.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr7: 140453091-140453197(Hg19) amplicon of BRAF gene is shown as SEQ ID: 31: CTACTGTTTTCCTTTACTTACTACACCTC; the specific reverse primer sequence designed according to the

Chr7: 140453091-140453197(Hg19) amplicon of the BRAF gene is shown as SEQ ID: 32: CCTCAATTCTTACCATCCACAAAATGG.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr7: 55241604-55241726(Hg19) amplicon of EGFR gene is shown as SEQ ID: 33: TGACCCTTGTCTCTGTGTTCTTG; the specific reverse primer sequence designed according to the Chr7: 55241604-55241726(Hg19) amplicon of the BRAF gene is shown as SEQ ID: 34: CCAGGGACCTTACCTTATACACC.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr7: 55242398-55242513 (Hg19) amplicon of EGFR gene is shown as SEQ ID:35: ACAATTGCCAGTTAACGTCTTCC; the specific reverse primer sequence designed according to the Chr7: 55242398-55242513 (Hg19) amplicon of the EGFR gene is shown as SEQ ID: 36: ACACAGCAAAGCAGAAACTCAC.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr7: 55248970-55249096 (Hg19) amplicon of EGFR gene is shown as SEQ ID: 37: GAAGCCACACTGACGTGC; the specific reverse primer sequence designed according to the Chr7: 55248970-55249096 (Hg19) amplicon of the EGFR gene is shown as SEQ ID: 38: GTGTTCCCGGACATAGTCCAG.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr7: 55259505-55259621 (Hg19) amplicon of EGFR gene is shown as SEQ ID: 39: CCGCAGCATGTCAAGATCACA; the specific reverse primer sequence designed according to the Chr7: 55259505-55259621 (Hg19) amplicon of the EGFR gene is shown as SEQ ID: 40: TAAACAATACAGCTAGTGGGAAGGC.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr17: 37880969-37881082 (Hg19) amplicon of ERBB2 gene is shown as SEQ ID: 41: CATACCCTCTCAGCGTACCC; the specific reverse primer sequence designed according to the Chr17: 37880969-37881082 (Hg19) amplicon of the ERBB2 gene is shown as SEQ ID: 42: CGGACATGGTCTAAGAGGCAG.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr12: 25380261-25380363 (Hg19) amplicon of KRAS gene is shown as SEQ ID: 43: TGCACTGTAATAATCCAGACTGTGT; the specific reverse primer sequence designed according to the Chr12: 25380261-25380363 (Hg19) amplicon of the KRAS gene is shown as SEQ ID: 44: AGTCCTCATGTACTGGTCCCTC.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr12: 25398183-25398310 (Hg19) amplicon of KRAS gene is shown as SEQ ID: 45: AAGGCCTGCTGAAAATGACTGA; the specific reverse primer sequence designed according to the Chr12: 25398183-25398310 (Hg19) amplicon of the KRAS gene is shown as SEQ ID: 46: AAAGAATGGTCCTGCACCAGTA.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr7: 116340233-116340335 (Hg19) amplicon of MET gene is shown as SEQ ID: 47: TCGATCTGCCATGTGTGCATT; the specific reverse primer sequence designed according to the Chr7: 116340233-116340335 (Hg19) amplicon of the MET gene is shown as SEQ ID: 48: GGGAACTGATGTGACTTACCCT.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr7: 116411880-116412005 (Hg19) amplicon of MET gene is shown as SEQ ID: 49: CCATGATAGCCGTCTTTAACAAGC; the specific reverse primer sequence designed according to the Chr7: 116411880-116412005 (Hg19) amplicon of the MET gene is shown as SEQ ID: 50: AGCTCGGTAGTCTACAGATTCATTT.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr7: 116417426-116417546 (Hg19) amplicon of MET gene is shown as SEQ ID: 51: ATGTTACGCAGTGCTAACCAAG; the specific reverse primer sequence designed according to the Chr7: 116417426-116417546 (Hg19) amplicon of the MET gene is shown as SEQ ID: 52: GTTGCAAACCACAAAAGTATACTCCA.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr7: 116423399-116423499 (Hg19) amplicon of MET gene is shown as SEQ ID: 53: CAGTCAAGGTTGCTGATTTTGGTC; the specific reverse primer sequence designed according to the Chr7: 116423399-116423499 (Hg19) amplicon of the MET gene is shown as SEQ ID: 54: CACATCTGACTTGGTGGTAAACTT.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr1: 115256507-115256586 (Hg19) amplicon of NRAS gene is shown as SEQ ID: 55: CACCCCCAGGATTCTTACAGAAAA; the specific reverse primer sequence designed according to the Chr1: 115256507-115256586 (Hg19) amplicon of the NRAS gene is shown as SEQ ID: 56: TTCGCCTGTCCTCATGTATTGG.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr1: 115258651-115258755 (Hg19) amplicon of NRAS gene is shown as SEQ ID: 57: CTGAGTACAAACTGGTGGTGGT; the specific reverse primer sequence designed according to the Chr1:

115258651-115258755 (Hg19) amplicon of the NRAS gene is shown as SEQ ID: 58: TGAGAGACAGGATCAGGTCAGC.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr3: 178936056-178936179 (Hg19) amplicon of PIK3CA gene is shown as SEQ ID: 59: GGAAAATGACAAAGAACAGCTCAAAG; the specific reverse primer sequence designed according to the Chr3: 178936056-178936179 (Hg19) amplicon of the PIK3CA gene is shown as SEQ ID: 60: AACATGCTGAGATCAGCCAAATTC.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr3: 178952000-178952092 (Hg19) amplicon of PIK3CA gene is shown as SEQ ID: 61: ATGCCAGAACTACAATCTTTTGATGAC; the specific reverse primer sequence designed according to the Chr3: 178952000-178952092 (Hg19) amplicon of the PIK3CA gene is shown as SEQ ID: 62: CAATCCATTTTTGTTGTCCAGCC.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr17: 7577027-7577154 (Hg19) amplicon of TP53 gene is shown as SEQ ID: 63: CTCTTTTCCTATCCTGAGTAGTGGTAATC; the specific reverse primer sequence designed according to the Chr17: 7577027-7577154 (Hg19) amplicon of the TP53 gene is shown as SEQ ID: 64: CTTCTTGTCCTGCTTGCTTACC.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr17: 7577507-7577613 (Hg19) amplicon of TP53 gene is shown as SEQ ID: 65: TCTTGGGCCTGTGTTATCTCCTAG; the specific reverse primer sequence designed according to the Chr17: 7577507-7577613 (Hg19) amplicon of the TP53 gene is shown as SEQ ID: 66: GCAAGTGGCTCCTGACCTG.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr17: 7578182-7578298 (Hg19) amplicon of TP53 gene is shown as SEQ ID: 67: CCTCTGATTCCTCACTGATTGCTC; the specific reverse primer sequence designed according to the Chr17: 7578182-7578298 (Hg19) amplicon of the TP53 gene is shown as SEQ ID: 68: CCCCAGTTGCAAACCAGAC.

In an embodiment of the present disclosure, the specific forward primer sequence designed according to the Chr17: 7578389-7578537 (Hg19) amplicon of TP53 gene is shown as SEQ ID: 69: CAGTACTCCCCTGCCCTCAA; the specific reverse primer sequence designed according to the Chr17: 7578389-7578537 (Hg19) amplicon of the TP53 gene is shown as SEQ ID: 70: ACCATCGCTATCTGAGCAGC.

In an embodiment of the present disclosure, the target amplicons are the following 22 species:

Chr2:29432588-29432707 (Hg19) amplicon of the ALK gene, the sequence of which is shown in SEQ ID:5;

Chr2: 29443616-29443730 (Hg19) amplicon of the ALK gene, the sequence of which is shown in SEQ ID:6;

Chr7: 140453091-140453197(Hg19) amplicon of the BRAF gene, the sequence of which is shown in SEQ ID:7;

Chr7: 55241604-55241726(Hg19) amplicon of the EGFR gene, the sequence of which is shown in SEQ ID:8;

Chr7: 55242398-55242513 (Hg19) amplicon of the EGFR gene, the sequence of which is shown in SEQ ID:9;

Chr7: 55248970-55249096 (Hg19) amplicon of the EGFR gene, the sequence of which is shown in SEQ ID:10;

Chr7: 55259505-55259621 (Hg19) amplicon of the EGFR gene, the sequence of which is shown in SEQ ID:11;

Chr17: 37880969-37881082 (Hg19) amplicon of the ERBB2 gene, the sequence of which is shown in SEQ ID:12;

Chr12: 25380261-25380363 (Hg19) amplicon of the KRAS gene, the sequence of which is shown in SEQ ID:13;

Chr12: 25398183-25398310 (Hg19) amplicon of the KRAS gene, the sequence of which is shown in SEQ ID:14;

Chr7: 116340233-116340335 (Hg19) amplicon of the MET gene, the sequence of which is shown in SEQ ID:15;

Chr7: 116411880-116412005 (Hg19) amplicon of the MET gene, the sequence of which is shown in SEQ ID:16;

Chr7: 116417426-116417546 (Hg19) amplicon of the MET gene, the sequence of which is shown in SEQ ID:17;

Chr7: 116423399-116423499 (Hg19) amplicon of the MET gene, the sequence of which is shown in SEQ ID:18;

Chr1: 115256507-115256586 (Hg19) amplicon of the NRAS gene, the sequence of which is shown in SEQ ID:19;

Chr1: 115258651-115258755 (Hg19) amplicon of the NRAS gene, the sequence of which is shown in SEQ ID:20;

Chr3: 178936056-178936179 (Hg19) amplicon of the PIK3CA gene, the sequence of which is shown in SEQ ID:21;

Chr3: 178952000-178952092 (Hg19) amplicon of the PIK3CA gene, the sequence of which is shown in SEQ ID:22;

Chr17: 7577027-7577154 (Hg19) amplicon of the TP53 gene, the sequence of which is shown in SEQ ID:23;

Chr17: 7577507-7577613 (Hg19) amplicon of the TP53 gene, the sequence of which is shown in SEQ ID:24;

Chr17: 7578182-7578298 (Hg19) amplicon of the TP53 gene, the sequence of which is shown in SEQ ID:25; and

Chr17: 7578389-7578537 (Hg19) amplicon of the TP53 gene, the sequence of which is shown in SEQ ID:26.

In an embodiment of the present disclosure, the molar ratio of the combination of the forward fusion primers designed according to the above 22 target amplicons, the combination of the reverse fusion primers designed according to the above 22 target amplicons, the forward universal primer and the reverse universal primer is: 0.1-0.3: 0.1-0.3: 0.5-1: 0.5-1, for example, 0.1:0.1:0.5:0.5.

In an embodiment of the present disclosure, the molar ratio of the forward fusion primer designed according to the Chr2:29432588-29432707 (Hg19) amplicon of ALK gene, the forward fusion primer designed according to the Chr2: 29443616-29443730 (Hg19) amplicon of ALK gene, the forward fusion primer designed according to the Chr7: 140453091-140453197(Hg19) amplicon of BRAF gene, the forward fusion primer designed according to the Chr7: 55241604-55241726(Hg19) amplicon of EGFR gene, the forward fusion primer designed according to the Chr7: 55242398-55242513 (Hg19) amplicon of EGFR gene, the forward fusion primer designed according to the Chr7: 55248970-55249096 (Hg19) amplicon of EGFR gene, the forward fusion primer designed according to the Chr7: 55259505-55259621 (Hg19) amplicon of EGFR gene, the forward fusion primer designed according to the Chr17: 37880969-37881082 (Hg19) amplicon of ERBB2 gene; the forward fusion primer designed according to the Chr12: 25380261-25380363 (Hg19) amplicon of KRAS gene, the forward fusion primer designed according to the Chr12: 25398183-25398310 (Hg19) amplicon of KRAS gene; the forward fusion primer designed according to the Chr7: 116340233-116340335 (Hg19) amplicon of MET gene, the forward fusion primer designed according to the Chr7: 116411880-116412005 (Hg19) amplicon of MET gene, the forward fusion primer designed according to the Chr7: 116417426-116417546 (Hg19) amplicon of MET gene, the forward fusion primer designed according to the Chr7: 116423399-116423499 (Hg19) amplicon of MET gene, the forward fusion primer designed according to the Chr1: 115256507-115256586 (Hg19) amplicon of NRAS gene, the forward fusion primer designed according to the Chr1: 115258651-115258755 (Hg19) amplicon of NRAS gene, the forward fusion primer designed according to the Chr3: 178936056-178936179 (Hg19) amplicon of PIK3CA gene, the forward

fusion primer designed according to the Chr3: 178952000-178952092 (Hg19) amplicon of PIK3CA gene, the forward fusion primer designed according to the Chr17: 7577027-7577154 (Hg19) amplicon of TP53 gene, the forward fusion primer designed according to the Chr17: 7577507-7577613 (Hg19) amplicon of TP53 gene, the forward fusion primer designed according to the Chr17: 7578182-7578298 (Hg19) amplicon of TP53 gene, and the forward fusion primer designed according to the Chr17: 7578389-7578537 (Hg19) amplicon of TP53 gene is: 1:2:1:4:2:1:2:4:2:2:2:2:1:4:2:2:2:2:4:2:4:2.

In an embodiment of the present disclosure, the PCR reaction system includes the following components:

PCR master mix	10μl;
DNA sample	1-8μl total 20 ng;
Primer combination for constructing an amplicon library of the same DNA sample	2μl;
DNAase-free H_2O	making up to 20μl。

In an embodiment of the present disclosure, the PCR master mix is KAPA HiFi PCR Kits 2×.

In an embodiment of the present disclosure, the reaction procedure for performing PCR is:

Temperature	Time	Number of cycles
98°C	30s	
98°C	10s	22 cycles
60°C	90s	
72°C	90s	
72°C	10min	
4°C	----	

In an embodiment of the present disclosure, after the PCR reaction, a step of purifying the PCR amplification product is also included.

Compared with the prior art, the present disclosure has the following advantages:

The method disclosed in the present disclosure is based on the design of the PGM platform, and can effectively amplify multiple target regions (amplicons) at the same time. In the process of constructing the library , the present disclosure only involves one round of PCR reaction and one round of product purification steps, which greatly simplifies the experimental operation of the existing commercial kit (such as PCR process, purification step, digestion and joints, etc.) Step), and saves the construction time. The entire library construction process only takes 2.5 hours (including DNA and RNA of the same sample).

Effectively elimination of sample and library contamination is achieved. The significant simplified operation process makes the library construction process more secure and reliable, and the reduction of operation process and steps effectively eliminates the library pollution that may be caused during the library construction process.

Streamlined bioinformatics analysis process is obtained. The amplicon library obtained by the method has a single structure and reliable data, and the DNA strand composition of the obtained library is simple and clear, and the subsequent bioinformatics analysis is more simplified.

After the library is constructed, the library is only needed to be quantified by the instrument "Qubit 2.0", which eliminates quantification step by the instrument "qPCR". Therefore, the library construction time is shortened and corresponding operation steps are reduced, and the experimental errors that may be caused by the cumbersome experimental process are avoided.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a distribution diagram of an amplification product detected after completion of construction of an amplicon library in Example 1 of the present invention.

Figure 2 is a related parameter of 22 amplicons in the library obtained in Example 1 of the present invention.

DETAILED DESCRIPITON OF THE EMBODIMENTS

The specific embodiments of the present invention are described in detail below with reference to the accompanying drawings, but it is understood that the scope of the present invention is not limited by the specific embodiments.

Example 1

The samples to be tested are six FFPE samples (ie, formalin-fixed and paraffin-embedded samples, FFPE stands for Formalin-Fixed and Parrffin-Embedded), four of which are FFPE samples from patients with non-small cell lung cancer, and two are FFPE samples from patients without cancer. One-step method is used to construct an amplicon DNA library from 6 FFPE samples using a specific designed fusion primer. The

specific process is as follows:

1. Extraction of genomic DNA:

The genomic DNA in the FFPE sample is extracted using the “Qiagen FFPE DNA Kit”. The detailed steps for extraction can be referred to the kit instructions. The genomic DNA is dissolved in “Tris-HCl” buffer, quality of the extracted DNA is detected using “Nano Drop”. After the concentration of the sample DNA is detected using the instrument “Qubit3.0”, each genomic DNA sample is diluted to a concentration of 20 ng/μl.

2. Design and synthesize primers:

a forward fusion primer is designed according to the target amplicon. The forward fusion primer includes a first linker sequence and a specific forward primer sequence that is designed according to the target amplicon, both of which are arranged in the order of 5' to 3';

a reverse fusion primer is designed according to the target amplicon. The reverse fusion primer includes a second linker sequence and a specific reverse primer sequence designed according to the target amplicon, both of which are arranged in the order of 5' to 3';

a forward universal primer comprises a third linker sequence, barcode sequence and a first linker sequence arranged in the order of 5' to 3'; and

a reverse universal primer comprises a universal sequence and a second linker sequence arranged in the order of 5' to 3'.

In the primer combination for constructing the amplicon library of the DNA sample, the information of the specific forward primer sequence and the specific reverse primer sequence designed according to the target amplicon are as follows:

Information of different target amplicons is given in the table below, and the specific forward primer sequence “Special Primer Start” and the specific reverse primer sequence “Special Primer End” designed for these amplicons are also given. Sequences of forward fusion primers designed according to the target amplicon, reverse fusion primers designed according to the target amplicon, forward universal primers, and reverse universal primers are also given. Puf representes an alternative forward universal primer and Pur representes a reverse universal primer.

Gene/Amplicons	Chr	Amp Start	Ins Start	Ins End	Amp End	length	Primer name	Special Primer Start	Primer name	Special Primer End
ALK-001	2	29432569	29432588	29432707	29432727	159	Pspf-1	ACTGCCTCTTGACCTGTCC	pspr-1	TAAGGGACAAGCAGCCACAC
ALK-002	2	29443591	29443616	29443730	29443749	159	Pspf-2	CCAGACTCAGCTCAGTTAATTTTGG	pspr-2	CGGAGGAAGGACTTGAGGT
BRAF-001	7	140453062	140453091	140453197	140453224	163	Pspf-3	CTACTGTTTTCCTTTACTTACTACACCTC	pspr-3	CCTCAATTCTTACCATCCACAAAATGG
EGFR-001	7	55241581	55241604	55241726	55241749	169	Pspf-4	TGACCCTTGTCTCTGTGTTCTTG	pspr-4	CCAGGGACCTTACCTTATACACC
EGFR-002	7	55242375	55242398	55242513	55242535	161	Pspf-5	ACAATTGCCAGTTAACGTCTTCC	pspr-5	ACACAGCAAAGCAGAAACTCAC
EGFR-003	7	55248952	55248970	55249096	55249117	166	Pspf-6	GAAGCCACACTGACGTGC	pspr-6	GTGTTCCCGGACATAGTCCAG
EGFR-004	7	55259484	55259505	55259621	55259646	163	Pspf-7	CCGCAGCATGTCAAGATCACA	pspr-7	TAAACAATACAGCTAGTGGGAAGGC
ERBB2-001	17	37880949	37880969	37881082	37881103	155	Pspf-8	CATACCCTCTCAGCGTACCC	pspr-8	CGGACATGGTCTAAGAGGCAG
KRAS-001	12	25380236	25380261	25380363	25380385	150	Pspf-9	TGCACTGTAATAATCCAGACTGTGT	pspr-9	AGTCCTCATGTACTGGTCCCTC
KRAS-002	12	25398161	25398183	25398310	25398332	172	Pspf-10	AAGGCCTGCTGAAAATGACTGA	pspr-10	AAAGAATGGTCCTGCACCAGTA
MET-001	7	116340212	116340233	116340335	116340357	146	Pspf-11	TCGATCTGCCATGTGTGCATT	pspr-11	GGGAACTGATGTGACTTACCCT
MET-002	7	116411856	116411880	116412005	116412030	175	Pspf-12	CCATGATAGCCGTCTTTAACAAGC	pspr-12	AGCTCGGTAGTCTACAGATTCATTT
MET-003	7	116417404	116417426	116417546	116417572	169	Pspf-13	ATGTTACGCAGTGCTAACCAAG	pspr-13	GTTGCAAACCACAAAAGTATACTCCA

Gene/Amplicons	Chr	Amp Start	Ins Start	Ins End	Amp End	length	Primer name	Special Primer Start	Primer name	Special Primer End
MET-004	7	116423375	116423399	116423499	116423523	149	Pspf-14	CAGTCAAGGTTGCTGATTTTGGTC	pspr-14	CACATCTGACTTGGTGGTAAACTT
NRAS-001	1	115256483	115256507	115256586	115256608	126	Pspf-15	CACCCCCAGGATTCTTACAGAAAA	pspr-15	TTCGCCTGTCCTCATGTATTGG
NRAS-002	1	115258629	115258651	115258755	115258777	149	Pspf-16	CTGAGTACAAACTGGTGGTGGT	pspr-16	TGAGAGACAGGATCAGGTCAGC
PIK3CA-001	3	178936030	178936056	178936179	178936203	174	Pspf-17	GGAAAATGACAAAGAACAGCTCAAAG	pspr-17	AACATGCTGAGATCAGCCAAATTC
PIK3CA-002	3	178951973	178952000	178952092	178952115	143	Pspf-18	ATGCCAGAACTACAATCTTTTGATGAC	pspr-18	CAATCCATTTTTGTTGTCCAGCC
TP53-001	17	7576998	7577027	7577154	7577176	179	Pspf-19	CTCTTTTCCTATCCTGAGTAGTGGTAATC	pspr-19	CTTCTTGTCCTGCTTGCTTACC
TP53-002	17	7577483	7577507	7577613	7577632	150	Pspf-20	TCTTGGGCCTGTGTTATCTCCTAG	pspr-20	GCAAGTGGCTCCTGACCTG
TP53-003	17	7578158	7578182	7578298	7578317	159	Pspf-21	CCTCTGATTCCTCACTGATTGCTC	pspr-21	CCCCAGTTGCAAACCAGAC
TP53-004	17	7578369	7578389	7578537	7578557	189	Pspf-22	CAGTACTCCCCTGCCCTCAA	pspr-22	ACCATCGCTATCTGAGCAGC

Primer name	Primer sequence
Pspf-1	GGCATACGTCCTCGTCTAACTGCCTCTTGACCTGTCC
Pspf-2	GGCATACGTCCTCGTCTACCAGACTCAGCTCAGTTAATTTTGG
Pspf-3	GGCATACGTCCTCGTCTACTACTGTTTTCCTTTACTTACTACACCTC
Pspf-4	GGCATACGTCCTCGTCTATGACCCTTGTCTCTGTGTTCTTG

Primer name	Primer sequence
Pspf-5	GGCATACGTCCTCGTCTAACAATTGCCAGTTAACGTCTTCC
Pspf-6	GGCATACGTCCTCGTCTAGAAGCCACACTGACGTGC
Pspf-7	GGCATACGTCCTCGTCTACCGCAGCATGTCAAGATCACA
Pspf-8	GGCATACGTCCTCGTCTACATACCCTCTCAGCGTACCC
Pspf-9	GGCATACGTCCTCGTCTATGCACTGTAATAATCCAGACTGTGT
Pspf-10	GGCATACGTCCTCGTCTAAAGGCCTGCTGAAAATGACTGA
Pspf-11	GGCATACGTCCTCGTCTATCGATCTGCCATGTGTGCATT
Pspf-12	GGCATACGTCCTCGTCTACCATGATAGCCGTCTTTAACAAGC
Pspf-13	GGCATACGTCCTCGTCTAATGTTACGCAGTGCTAACCAAG
Pspf-14	GGCATACGTCCTCGTCTACAGTCAAGGTTGCTGATTTTGGTC
Pspf-15	GGCATACGTCCTCGTCTACACCCCCAGGATTCTTACAGAAAA
Pspf-16	GGCATACGTCCTCGTCTACTGAGTACAAACTGGTGGTGGT
Pspf-17	GGCATACGTCCTCGTCTAGGAAAATGACAAAGAACAGCTCAAAG
Pspf-18	GGCATACGTCCTCGTCTAATGCCAGAACTACAATCTTTTGATGAC
Pspf-19	GGCATACGTCCTCGTCTACTCTTTTCCTATCCTGAGTAGTGGTAATC
Pspf-20	GGCATACGTCCTCGTCTATCTTGGGCCTGTGTTATCTCCTAG
Pspf-21	GGCATACGTCCTCGTCTACCTCTGATTCCTCACTGATTGCTC
Pspf-22	GGCATACGTCCTCGTCTACAGTACTCCCCTGCCCTCAA
pspr-1	TCTATGGGCAGTCGGTGATTAAGGGACAAGCAGCCACAC
pspr-2	TCTATGGGCAGTCGGTGATCGGAGGAAGGACTTGAGGT
pspr-3	TCTATGGGCAGTCGGTGATCCTCAATTCTTACCATCCACAAAATGG
pspr-4	TCTATGGGCAGTCGGTGATCCAGGGACCTTACCTTATACACC
pspr-5	TCTATGGGCAGTCGGTGATACACAGCAAAGCAGAAACTCAC
pspr-6	TCTATGGGCAGTCGGTGATGTGTTCCCGGACATAGTCCAG
pspr-7	TCTATGGGCAGTCGGTGATTAAACAATACAGCTAGTGGGAAGGC
pspr-8	TCTATGGGCAGTCGGTGATCGGACATGGTCTAAGAGGCAG

Primer name	Primer sequence
pspr-9	TCTATGGGCAGTCGGTGATAGTCCTCATGTACTGGTCCCTC
pspr-10	TCTATGGGCAGTCGGTGATAAAGAATGGTCCTGCACCAGTA
pspr-11	TCTATGGGCAGTCGGTGATGGGAACTGATGTGACTTACCCT
pspr-12	TCTATGGGCAGTCGGTGATAGCTCGGTAGTCTACAGATTCATTT
pspr-13	TCTATGGGCAGTCGGTGATGTTGCAAACCACAAAAGTATACTCCA
pspr-14	TCTATGGGCAGTCGGTGATCACATCTGACTTGGTGGTAAACTT
pspr-15	TCTATGGGCAGTCGGTGATTTCGCCTGTCCTCATGTATTGG
pspr-16	TCTATGGGCAGTCGGTGATTGAGAGACAGGATCAGGTCAGC
pspr-17	TCTATGGGCAGTCGGTGATAACATGCTGAGATCAGCCAAATTC
pspr-18	TCTATGGGCAGTCGGTGATCAATCCATTTTTGTTGTCCAGCC
pspr-19	TCTATGGGCAGTCGGTGATCTTCTTGTCCTGCTTGCTTACC
pspr-20	TCTATGGGCAGTCGGTGATGCAAGTGGCTCCTGACCTG
pspr-21	TCTATGGGCAGTCGGTGATCCCCAGTTGCAAACCAGAC
pspr-22	TCTATGGGCAGTCGGTGATACCATCGCTATCTGAGCAGC
puf-1	CCATCTCATCCCTGCGTGTCTCCGACTCAGCTTGACACCGCGGCATACGTCCTCGTCTA
puf-2	CCATCTCATCCCTGCGTGTCTCCGACTCAGTTGGAGGCCAGCGGCATACGTCCTCGTCTA
puf-3	CCATCTCATCCCTGCGTGTCTCCGACTCAGTGGAGCTTCCTCGGCATACGTCCTCGTCTA
puf-4	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCAGTCCGAACGGCATACGTCCTCGTCTA
puf-5	CCATCTCATCCCTGCGTGTCTCCGACTCAGTAAGGCAACCACGGCATACGTCCTCGTCTA
puf-6	CCATCTCATCCCTGCGTGTCTCCGACTCAGTTCTAAGAGACGGCATACGTCCTCGTCTA
puf-7	CCATCTCATCCCTGCGTGTCTCCGACTCAGTCCTAACATAACGGCATACGTCCTCGTCTA
puf-8	CCATCTCATCCCTGCGTGTCTCCGACTCAGCGGACAATGGCGGCATACGTCCTCGTCTA
puf-9	CCATCTCATCCCTGCGTGTCTCCGACTCAGTTGAGCCTATTCGGCATACGTCCTCGTCTA
puf-10	CCATCTCATCCCTGCGTGTCTCCGACTCAGCCGCATGGAACGGCATACGTCCTCGTCTA
pur	CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGGTGAT

The first linker sequence is GGCATACGTCCTCGTCTA, the second linker sequence is TCTATGGGCAGTCGGTGAT, the third linker sequence is CCATCTCATCCCTGCGTGTCTCCGACTCAG, and the universal sequence is CCACTACGCCTCCGCTTTCCTC.

3. Form a PCR reaction system. The specific PCR reaction system is as follows:

PCR reaction system component	content
KAPA HiFi PCR Kits 2×	10μl
Genomic DNA (10 ng/μl itself)	2μl
Primer combination for constructing an amplicon library of the same DNA sample	2μl
DNAase free H_2O	make up to the total of 20μl
Total	20μl

Primer combinations for constructing an amplicon library of the same DNA sample are prepared by the following methods: (1) the forward universal primer, the reverse universal primer, and each forward fusion primer designed according to the 22 target amplicons and each reverse fusion primer are dissolved in water to a concentration of 100 μM; (2) 22 forward fusion primers with a serial number ranging from small to large are respectively mixed with a concentration of 100μM, and the molar ratio is 1:2:1:4:2:1:2:4:2:2:2:2:2:4:2:2:2:2:4:2:4:2, so as to obtain the forward fusion primer combination, and 22 reverse fusion primers with a concentration of 100 μM are respectively mixed with the corresponding forward fusion primers in equal volume to obtain a reverse fusion primer combination, and then the forward fusion primer combination and the reverse fusion primer combination are mixed in equal volume; (3) mixing in equal volume of forward universal primers and reverse universal primers with concentrations of 100 μM; (4) the forward fusion primer combination, the reverse fusion primer combination, the forward universal primer and the reverse universal primer are mixed according to a molar ratio of 0.1:0.1:0.5:0.5, so that the amplicons for constructing the DNA sample are obtained. Six different sets of samples to be tested need to correspond to primer combinations containing six different barcode sequence tags.

4. Carry out the PCR program. The PCR instrument is the 2720 Thermal Cycler of Applied Bio-system. The PCR reaction procedure is as follows:

Temperature	Time	Number of cycles
98℃	30s	
98℃	10s	22 Cycles
60℃	90s	
72℃	90s	

72℃	10min	
4℃	----	

5. After the PCR reaction, purification is carried out using “Agencourt AMPure XP Kit” (Cat. No. A63880/A63881/A63882) from Beckman Coulter company. The steps are as follows:

1) take out the Agencourt AMPure XP Kit 30 minutes in advance, rotate the magnetic beads in the Kit thoroughly, and keep the EP tube at room temperature.

2) After the completion of the PCR reaction, the magnetic beads are rotated again sufficiently, and 20 μl of magnetic beads are added to the system, repeatedly blow 5 times or more, or rotate thoroughly, and allow the Kit to be placed at room temperature for 5 minutes.

3) Transfer the EP tube to the magnetic stand and keep for 5 minutes until the solution is clarified. Carefully remove the supernatant with a pipette, taking care not to touch the beads.

4) Add 100 μl of freshly prepared 80% ethanol solution to each tube, and place the EP tube on the magnetic stand and rotate two turns, keep it for 5 minutes, and discard the supernatant.

5) Repeat step 4) once.

6) Open the EP tube and keep it at room temperature to make the liquid volatilize completely. Make sure the surfaces of the magnetic beads are dull, and be careful not to over-dry the magnetic beads.

7) Remove the EP tube from the magnetic stand, add 30 μl of PCR-grade purified water, rotate and mix, and keep it for 10 minutes at room temperature.

8) Place the EP tube on the magnetic stand for 2 minutes or until the solution is clarified. Carefully suck the supernatant from the side away from the magnet with a pipette, taking care not to touch the beads.

At this point, the amplicon library is constructed. Figure 1 shows the distribution of amplified products detected by Agilent 2200 TapeStation Systems after the completion of the library. The abscissa is the length of the fragment, the ordinate is the signal intensity (FU), and the lower peak is the 25 bp position marker, the upper peak is a 1500 bp position marker. As shown in Figure 1, the PCR products obtained by PCR amplification are concentrated in the range of 241-271 bp. Figure 1 shows that the experimental results are consistent with the experimental design. From Figure 1, the size of the constructed library and the library concentration can be judged.

6. Sequencing and results analysis

The amplicon library is obtained by the fusion primer one-step method. The amplicon sequencing is performed using the chip 318 of the Ion PGM platform, and the data amount of each library is 50 M bps.

The average sequencing depth of each sample is not less than 1600X, and the single amplicon sequencing depth reached 600X. The obtained sequencing results are shown in Fig. 2. From Fig. 2, it is possible to further analyze whether or not each amplicon of the 22 amplicons is amplified and the amplification uniformity of each amplicon.

The results of sequencing are analyzed by data processing and bioinformatics analysis to obtain mutations in the detected genes. The data processing process includes conversion, quality control and sequence alignment of the sequencing data (reference genome is NCBI GRCh37/Hg19), mutation site analysis and other processes, and the mutation information of the detected samples is obtained through data processing analysis.

The actual sample collection is as follows: Among the FFPE samples of 6 subjects, no tumor-related mutations are detected in 2 normal human samples, among the 4 FFPE samples of tumor patients, p.R248W mutation is detected in Sample1, p.T790M mutation is detected in Sample2, p.G12A mutation is detected in Sample3, and p.E545K mutation is detected in Sample4. This result is consistent with the results of the sanger test. The practical applicability and good specificity of the present invention are fully illustrated.

The foregoing descriptions of specific embodiments of the present invention are presented for purposes of illustration and description. It is not intended to be exhaustive or to limit the invention to the precise form disclosed, and obviously many modifications and variations are possible in light of the above teaching. The exemplary embodiments are chosen and described in order to explain certain principles of the invention and their practical application, to thereby enable others skilled in the art to make and utilize various exemplary embodiments of the present invention, as well as various alternatives and modifications thereof. The invention is to be limited only by the claims provided below and equivalents thereof.

Claims

1. A method for constructing an amplicon library of a DNA sample, comprising the following steps:

1) synthesizing a primer combination which is used for constructing an amplicon library of a DNA sample, wherein the primer combination of the amplicon library used to construct the DNA sample includes:

an forward fusion primer designed according to a target amplicon, the forward fusion primer comprising a first linker sequence (Bridge sequence) and a specific forward primer sequence designed according to the target amplicon, both of which are arranged in the order of 5' to 3';

a reverse fusion primer designed according to the target amplicon, the reverse fusion primer comprising a second linker sequence (trP1 sequence) and a specific reverse primer sequence designed according to the target amplicon, both of which are arranged in the order of 5' to 3';

an forward universal primer comprising a third linker sequence (A sequence), a barcode sequence and a first linker sequence arranged in the order of 5' to 3'; and

a reverse universal primer comprising a universal sequence (Uni sequence) and a second linker sequence arranged in the order of 5' to 3';

2) constructing a PCR reaction system for the DNA sample, and mixing the forward fusion primers designed according to the target amplicon, the reverse fusion primers designed according to the target amplicon, the forward universal primers and the reverse universal primers together, to serve as a primer combination in the PCR reaction system; and

3) performing PCR.

2. The method for constructing an amplicon library of a DNA sample according to claim 1, wherein the first linker sequence comprises a sequence of SEQ ID: 1; the second linker sequence comprises a sequence of SEQ ID 2; the third linker sequence comprises a sequence of SEQ ID: 3; the universal sequence comprises a sequence of SEQ ID: 4.

3. The method for constructing an amplicon library of a DNA sample according to claim 1, wherein in a primer combination for constructing a plurality of amplicon libraries of the same DNA sample, the barcode sequences in the forward universal primers are the same; in a primer combination for constructing amplicon libraries of different DNA samples, the barcode sequences in the forward universal primers are different.

4. The method for constructing an amplicon library of a DNA sample according to claim 1,

wherein when the number of target amplicon in a same PCR reaction is greater than one, the forward fusion primer designed according to the target amplicon is a combination of forward fusion primers designed according to each target amplicon, the reverse fusion primer designed according to the target amplicon is a combination of reverse fusion primers designed according to each target amplicon.

5. The method for constructing an amplicon library of a DNA sample according to claim 1, wherein the DNA sample is genomic DNA, the genomic DNA is extracted from a tissue sample or a formalin-fixed paraffin-embedded sample.

6. The method for constructing an amplicon library of a DNA sample according to claim 1, wherein the target amplicon comprises at least one selected from the group consisting of 22 target amplicons:

Chr2:29432588-29432707 (Hg19) amplicon of the ALK gene, the sequence of which is shown in SEQ ID:5;

Chr2: 29443616-29443730 (Hg19) amplicon of the ALK gene, the sequence of which is shown in SEQ ID:6;

Chr7: 140453091-140453197(Hg19) amplicon of the BRAF gene, the sequence of which is shown in SEQ ID:7;

Chr7: 55241604-55241726(Hg19) amplicon of the EGFR gene, the sequence of which is shown in SEQ ID:8;

Chr7: 55242398-55242513 (Hg19) amplicon of the EGFR gene, the sequence of which is shown in SEQ ID:9;

Chr7: 55248970-55249096 (Hg19) amplicon of the EGFR gene, the sequence of which is shown in SEQ ID:10;

Chr7: 55259505-55259621 (Hg19) amplicon of the EGFR gene, the sequence of which is shown in SEQ ID:11;

Chr17: 37880969-37881082 (Hg19) amplicon of the ERBB2 gene, the sequence of which is shown in SEQ ID:12;

Chr12: 25380261-25380363 (Hg19) amplicon of the KRAS gene, the sequence of which is shown in SEQ ID:13;

Chr12: 25398183-25398310 (Hg19) amplicon of the KRAS gene, the sequence of which is shown in SEQ ID:14;

Chr7: 116340233-116340335 (Hg19) amplicon of the MET gene, the sequence of which is shown in SEQ ID:15;

Chr7: 116411880-116412005 (Hg19) amplicon of the MET gene, the sequence of which is shown in SEQ ID:16;

Chr7: 116417426-116417546 (Hg19) amplicon of the MET gene, the sequence of which is shown in SEQ ID:17;

Chr7: 116423399-116423499 (Hg19) amplicon of the MET gene, the sequence of which is shown in SEQ ID:18;

Chr1: 115256507-115256586 (Hg19) amplicon of the NRAS gene, the sequence of which is shown in SEQ ID:19;

Chr1: 115258651-115258755 (Hg19) amplicon of the NRAS gene, the sequence of which is shown in SEQ ID:20;

Chr3: 178936056-178936179 (Hg19) amplicon of the PIK3CA gene, the sequence of which is shown in SEQ ID:21;

Chr3: 178952000-178952092 (Hg19) amplicon of the PIK3CA gene, the sequence of which is shown in SEQ ID:22;

Chr17: 7577027-7577154 (Hg19) amplicon of the TP53 gene, the sequence of which is shown in SEQ ID:23;

Chr17: 7577507-7577613 (Hg19) amplicon of the TP53 gene, the sequence of which is shown in SEQ ID:24;

Chr17: 7578182-7578298 (Hg19) amplicon of the TP53 gene, the sequence of which is shown in SEQ ID:25; and

Chr17: 7578389-7578537 (Hg19) amplicon of the TP53 gene, the sequence of which is shown in SEQ ID:26.

7. The method for constructing an amplicon library of a DNA sample according to claim 6, wherein,

the specific forward primer sequence designed according to the Chr2:29432588-29432707 (Hg19) amplicon of the ALK gene is shown as SEQ ID: 27; the specific reverse primer sequence designed according to the Chr2:29432588-29432707 (Hg19) amplicon of the ALK gene is shown as SEQ ID: 28;

the specific forward primer sequence designed according to the Chr2: 29443616-29443730 (Hg19) amplicon of ALK gene is shown as SEQ ID: 29; the specific reverse primer sequence designed according to the Chr2: 29443616-29443730 (Hg19) amplicon of the ALK gene is shown as SEQ ID: 30;

the specific forward primer sequence designed according to the Chr7: 140453091-140453197(Hg19) amplicon of BRAF gene is shown as SEQ ID: 31; the specific reverse primer sequence designed according to the Chr7: 140453091-140453197(Hg19) amplicon of the BRAF gene is shown as SEQ ID: 32;

the specific forward primer sequence designed according to the Chr7: 55241604-55241726(Hg19) amplicon of EGFR gene is shown as SEQ ID: 33; the specific reverse primer sequence designed according to the Chr7: 55241604-55241726(Hg19) amplicon of the BRAF gene is shown as SEQ ID: 34.

8. The method for constructing an amplicon library of a DNA sample according to claim 6, wherein,

the specific forward primer sequence designed according to the Chr7: 55242398-55242513 (Hg19) amplicon of EGFR gene is shown as SEQ ID:35; the specific reverse primer sequence designed according to the Chr7: 55242398-55242513 (Hg19) amplicon of the EGFR gene is shown as SEQ ID: 36;

the specific forward primer sequence designed according to the Chr7: 55248970-55249096 (Hg19) amplicon of EGFR gene is shown as SEQ ID: 37; the specific reverse primer sequence designed according to the Chr7: 55248970-55249096 (Hg19) amplicon of the EGFR gene is shown as SEQ ID: 38;

the specific forward primer sequence designed according to the Chr7: 55259505-55259621 (Hg19) amplicon of EGFR gene is shown as SEQ ID: 39; the specific reverse primer sequence designed according to the Chr7: 55259505-55259621 (Hg19) amplicon of the EGFR gene is shown as SEQ ID: 40.

9. The method for constructing an amplicon library of a DNA sample according to claim 6, wherein,

the specific forward primer sequence designed according to the Chr17: 37880969-37881082 (Hg19) amplicon of ERBB2 gene is shown as SEQ ID: 41; the specific reverse primer sequence designed according to the Chr17: 37880969-37881082 (Hg19) amplicon of the ERBB2 gene is shown as SEQ ID: 42;

the specific forward primer sequence designed according to the Chr12: 25380261-25380363 (Hg19) amplicon of KRAS gene is shown as SEQ ID: 43; the specific reverse primer sequence designed according to the Chr12: 25380261-25380363 (Hg19) amplicon of the KRAS gene is shown as SEQ ID: 44;

the specific forward primer sequence designed according to the Chr12: 25398183-25398310 (Hg19) amplicon of KRAS gene is shown as SEQ ID: 45; the specific reverse primer sequence designed according to the Chr12: 25398183-25398310 (Hg19) amplicon of the KRAS gene is shown as SEQ ID: 46.

10. The method for constructing an amplicon library of a DNA sample according to claim 6, wherein,

the specific forward primer sequence designed according to the Chr7: 116340233-116340335 (Hg19) amplicon of MET gene is shown as SEQ ID: 47; the specific reverse primer sequence designed

according to the Chr7: 116340233-116340335 (Hg19) amplicon of the MET gene is shown as SEQ ID: 48;

the specific forward primer sequence designed according to the Chr7: 116411880-116412005 (Hg19) amplicon of MET gene is shown as SEQ ID: 49; the specific reverse primer sequence designed according to the Chr7: 116411880-116412005 (Hg19) amplicon of the MET gene is shown as SEQ ID: 50;

the specific forward primer sequence designed according to the Chr7: 116417426-116417546 (Hg19) amplicon of MET gene is shown as SEQ ID: 51; the specific reverse primer sequence designed according to the Chr7: 116417426-116417546 (Hg19) amplicon of the MET gene is shown as SEQ ID: 52.

11. The method for constructing an amplicon library of a DNA sample according to claim 6, wherein,

the specific forward primer sequence designed according to the Chr7: 116423399-116423499 (Hg19) amplicon of MET gene is shown as SEQ ID: 53; the specific reverse primer sequence designed according to the Chr7: 116423399-116423499 (Hg19) amplicon of the MET gene is shown as SEQ ID: 54;

the specific forward primer sequence designed according to the Chr1: 115256507-115256586 (Hg19) amplicon of NRAS gene is shown as SEQ ID: 55; the specific reverse primer sequence designed according to the Chr1: 115256507-115256586 (Hg19) amplicon of the NRAS gene is shown as SEQ ID: 56;

the specific forward primer sequence designed according to the Chr1: 115258651-115258755 (Hg19) amplicon of NRAS gene is shown as SEQ ID: 57; the specific reverse primer sequence designed according to the Chr1: 115258651-115258755 (Hg19) amplicon of the NRAS gene is shown as SEQ ID: 58.

12. The method for constructing an amplicon library of a DNA sample according to claim 6, wherein,

the specific forward primer sequence designed according to the Chr3: 178936056-178936179 (Hg19) amplicon of PIK3CA gene is shown as SEQ ID: 59; the specific reverse primer sequence designed according to the Chr3: 178936056-178936179 (Hg19) amplicon of the PIK3CA gene is shown as SEQ ID: 60;

the specific forward primer sequence designed according to the Chr3: 178952000-178952092 (Hg19) amplicon of PIK3CA gene is shown as SEQ ID: 61; the specific reverse primer sequence designed according to the Chr3: 178952000-178952092 (Hg19) amplicon of the PIK3CA gene is shown as SEQ ID: 62;

the specific forward primer sequence designed according to the Chr17: 7577027-7577154 (Hg19) amplicon of TP53 gene is shown as SEQ ID: 63; the specific reverse primer sequence designed according to the Chr17: 7577027-7577154 (Hg19) amplicon of the TP53 gene is shown as SEQ ID: 64.

13. The method for constructing an amplicon library of a DNA sample according to claim 6, wherein,

the specific forward primer sequence designed according to the Chr17: 7577507-7577613 (Hg19) amplicon of TP53 gene is shown as SEQ ID: 65; the specific reverse primer sequence designed according to the Chr17: 7577507-7577613 (Hg19) amplicon of the TP53 gene is shown as SEQ ID: 66;

the specific forward primer sequence designed according to the Chr17: 7578182-7578298 (Hg19) amplicon of TP53 gene is shown as SEQ ID: 67; the specific reverse primer sequence designed according to the Chr17: 7578182-7578298 (Hg19) amplicon of the TP53 gene is shown as SEQ ID: 68;

the specific forward primer sequence designed according to the Chr17: 7578389-7578537 (Hg19) amplicon of TP53 gene is shown as SEQ ID: 69; the specific reverse primer sequence designed according to the Chr17: 7578389-7578537 (Hg19) amplicon of the TP53 gene is shown as SEQ ID: 70.

14. The method for constructing an amplicon library of a DNA sample according to claim 1, wherein the target amplicons are the following 22 species:

Chr2:29432588-29432707 (Hg19) amplicon of the ALK gene, the sequence of which is shown in SEQ ID:5;

Chr2: 29443616-29443730 (Hg19) amplicon of the ALK gene, the sequence of which is shown in SEQ ID:6;

Chr7: 140453091-140453197(Hg19) amplicon of the BRAF gene, the sequence of which is shown in SEQ ID:7;

Chr7: 55241604-55241726(Hg19) amplicon of the EGFR gene, the sequence of which is shown in SEQ ID:8;

Chr7: 55242398-55242513 (Hg19) amplicon of the EGFR gene, the sequence of which is shown in SEQ ID:9;

Chr7: 55248970-55249096 (Hg19) amplicon of the EGFR gene, the sequence of which is shown in SEQ ID:10;

Chr7: 55259505-55259621 (Hg19) amplicon of the EGFR gene, the sequence of which is shown in SEQ ID:11;

Chr17: 37880969-37881082 (Hg19) amplicon of the ERBB2 gene, the sequence of which is shown in SEQ ID:12;

Chr12: 25380261-25380363 (Hg19) amplicon of the KRAS gene, the sequence of which is shown in SEQ ID:13;

Chr12: 25398183-25398310 (Hg19) amplicon of the KRAS gene, the sequence of which is shown in SEQ ID:14;

Chr7: 116340233-116340335 (Hg19) amplicon of the MET gene, the sequence of which is shown in SEQ ID:15;

Chr7: 116411880-116412005 (Hg19) amplicon of the MET gene, the sequence of which is shown in SEQ ID:16;

Chr7: 116417426-116417546 (Hg19) amplicon of the MET gene, the sequence of which is shown in SEQ ID:17;

Chr7: 116423399-116423499 (Hg19) amplicon of the MET gene, the sequence of which is shown in SEQ ID:18;

Chr1: 115256507-115256586 (Hg19) amplicon of the NRAS gene, the sequence of which is shown in SEQ ID:19;

Chr1: 115258651-115258755 (Hg19) amplicon of the NRAS gene, the sequence of which is shown in SEQ ID:20;

Chr3: 178936056-178936179 (Hg19) amplicon of the PIK3CA gene, the sequence of which is shown in SEQ ID:21;

Chr3: 178952000-178952092 (Hg19) amplicon of the PIK3CA gene, the sequence of which is shown in SEQ ID:22;

Chr17: 7577027-7577154 (Hg19) amplicon of the TP53 gene, the sequence of which is shown in SEQ ID:23;

Chr17: 7577507-7577613 (Hg19) amplicon of the TP53 gene, the sequence of which is shown in SEQ ID:24;

Chr17: 7578182-7578298 (Hg19) amplicon of the TP53 gene, the sequence of which is shown in SEQ ID:25; and

Chr17: 7578389-7578537 (Hg19) amplicon of the TP53 gene, the sequence of which is shown in SEQ ID:26.

15. The method for constructing an amplicon library of a DNA sample according to claim 14, wherein the molar ratio of the combination of the forward fusion primers designed according to the 22 target amplicons of claim 14, the combination of the reverse fusion primers designed according to the above 22 target amplicons, the forward universal primer and the reverse universal primer is:

0.1-0.3: 0.1-0.3: 0.5-1: 0.5-1.

16. The method for constructing an amplicon library of a DNA sample according to claim 15, wherein the PCR reaction system includes the following components:

PCR master mix, 10μl;

DNA sample, 1-8μl total 20 ng;

Primer combination for constructing an amplicon library of the same DNA sample, 2μl;

DNAase-free H2O, making up to 20μl.