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;

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

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;

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;