Detection of FLT3 (FMS-like Tyrosine Kinase) Internal Tandem Duplication (ITD) Mutation using Next Generation Sequencing Technology and Nested PCR

- Improving detection sensitivity of FLT3-ITD in AML -

2016년 2월

서울대학교 대학원

의과대학 협동과정 중앙생물학 전공

김 대윤
차세대 시퀀싱 및 중첩 중합 효소 연쇄 반응 기법을 이용한 FLT3 (FMS-like Tyrosine Kinase) 내부 일렬 중복 유전변이의 민감도 향상을 위한 검출 방법 고안

- 급성골수성백혈병에서 FLT3-ITD의 검출 민감도 향상-

Feb. 2016

서울대학교 대학원
Department of Cancer Biology
College of Medicine
KIM DAE YOON
Abstract

Introduction:

Sensitive detection of internal tandem duplication (ITD) mutation of FLT3 is very important in acute myeloid leukemia. Conventional PCR methods are not satisfactory in detecting relevant mutations in patients harboring the mutations. To increase detection sensitivity of FLT3-ITD, I developed new detection algorithm using next generation sequencing (NGS) data. I validated results using nested polymerase chain reaction (PCR) methods. I compared results of NGS data, nested PCR and conventional PCR methods.

Methods:

First, using whole exome sequencing data of 81 AML patients, I applied calling algorithm for FLT3-ITD. Briefly, to detect ITDs with NGS data, the reads are aligned to a reference sequence (UCSC hg19), with BWA which is a read aligner allowing soft-clipping. Some reads can be an
indication of the occurrence of ITD and BWA aligns those reads as soft-clipped.

Second, I designed two types of primer for Nested PCR. The first primer was targeted wildly for between exon14 and exon15 of FLT3 gene. Nested PCR primer was designed to target previously reported regions in which ITD mutations were frequently found. PCR reactions of two steps were performed using the PCR primers sequentially.

**Results:**

In these 81 patients, FLT3-ITD was positive only in 7 patients when tested by conventional PCR methods. When NGS detection method was applied, FLT3-ITD was positive in 11 patients (11/81, 13%). When validation was performed using nested PCR, FLT3-ITD was confirmed in all 11 patients. Nested PCR detected additional 4 patients positive for FLT3-ITD in this population. For 65 patients, FLT3-ITD was negative by both NGS and nested PCR method. One patient with FLT3-ITD by conventional PCR was not defined clearly by nested PCR and NGS method because of smear band and low depth coverage (coverage depth of exon 14 was 42). Overall, NGS method improved
sensitivity of FLT3-ITD detection by 57% in this population compared with currently used conventional PCR method. The concordance rate of NGS method and nested PCR was 95% (77/81).

**Conclusions:**

FLT3-ITD is a very important genetic factor, guiding a therapeutic direction for AML patients. Here, I have developed more sensitive alternative detection methods for FLT3-ITD based on NGS. All of detection results were validated by nested PCR. NGS method is not only more sensitive than conventional PCR but also capable of determining FLT3-ITD size and amount in AML patients.

---

**Keywords:** FLT3, ITD mutation, Acute Myeloid Leukemia, Next Generation Sequencing, Nested PCR, Genetic maker

**Student Number:** 2014-21156
# Contents

Abstract.................................................................2

Contents...........................................................................5

List of tables and figures...............................................6

List of Abbreviations......................................................7

Introduction......................................................................8

Material and methods....................................................12

Results..........................................................................19

Discussion.....................................................................22

References.....................................................................32

Abstract in Korean.........................................................35
List of figures and tables

Figure 1. IGV view of read alignments mapped to the position of FLT3-ITD (chr13:28608227) (Sample ID: Patient 6)……………………………………..27

Figure 2. The procedure of FLT3-ITD detection with NGS data…………………………………………………………28

Figure 3. Correlation analysis between blast percent and FLT-ITD burden………………………………………30

Figure 4. Comparison of NGS Tools………………..31

Table 1. Overall result of positive patient by NGS method and nested PCR………………………………………..29
List of abbreviations

AML: Acute myeloid leukemia
FLT3: FMS-like tyrosine kinase
ITD: internal tandem duplication
CR: Complete remission
NGS: Next generation sequencing
WES: Whole exome sequencing
ICGC: International Genome Consortium
TCGA: The Cancer Genome Atlas
SV: structural variations
HCT: hematopoietic cell transplantation
Introduction

The presence of FLT3 (FMS-like tyrosine kinase) internal tandem duplication (ITD) mutation in the juxtamembrane (JM) domain was previously reported by Nakao and colleagues [1]. Afterwards, correlation between prognosis and FLT3-ITD has been widely studied in acute myeloid leukemia (AML).

FLT3 (FMS-like tyrosine kinase) internal tandem duplication (ITD) is well known as one of the poor genetic factors which leads to early relapse and eventual treatment failure. Acute myeloid leukemia (AML) is an aggressive myeloid hematologic malignancy. Although, many patients with AML are able to achieve complete remission (CR) with conventional induction chemotherapy, a significant proportion of patients will suffer relapse. Especially, the rates of relapse are significantly higher in patients with FLT3 (FMS-like tyrosine kinase 3) internal tandem duplication (ITD) mutations [2-9].

For these reasons, detection of FLT3 (FMS-like tyrosine kinase 3) internal tandem duplication (ITD) mutations in
patients with acute myeloid leukemia (AML) is of tantamount importance in predicting the overall prognosis and guiding optimal treatment plan including allogeneic transplantation approaches.

Despite the importance of detecting FLT3-ITD variants, low sensitivity of FLT3-ITD detection compared to other AML-associated changes, such as AML-ETO and inv (16), had always been a barrier to treatment success [8].

A major reason in difficulties of detecting FLT3-ITD is heterogeneity of cancer cells. The sensitivity of PCR assays is determined by the amount of template DNA and the numbers of PCR cycle. In the FLT3-ITD assay, however, increasing the number of PCR cycle will not enhance the sensitivity of FLT3-ITD detection. Even though mutant allele was amplified by increasing PCR cycle, wild-type allele will be amplified greater than mutant allele by increasing PCR cycle because of the relative abundance of wild-type allele.

Another reason of difficulties is diversity of insertion size. If both mutant and wild-type allele were amplified equally, there will be no problem. However, generally shorter
template has a competitive advantage in PCR cycle [8]. The repeated sequence with mutant allele inserted will require longer time to complete a PCR cycle. This bias in PCR will be present even the population ratio of mutant allele to wild type is 1:1 in patient specimens, but typically this ratio will be less than 1. In conclusion, despite the urgent need for detecting FLT3-ITD mutations, current methods have their own limitations of low yield stemming from the inherent problems shown above.

Currently, novel technologies such as next generation sequencing (NGS) are revolutionizing cancer research. Based on these modern technologies, worldwide collaborative efforts, such as International Genome Consortium (ICGC) [10] and The Cancer Genome Atlas (TCGA) [11] projects are successfully launched. Due to these efforts, huge NGS data have been being produced on a daily basis.

Commonly used platforms supplied by Roche, Illumina and Life Technologies for genome analysis produce shorter reads than the traditional capillary sequencing [12]. Calling for structural variations (SVs) using the shorter reads have several limitations. Difficulties in mapping
reads to the genome may not enable accurate identification of structural variations (SVs) including ITD variation. Even though paired-end mapping can detect a few base pairs in the breakpoint position, it can hardly detect changes exceeding a few base pairs [13]. It is useful only when the entire sequence is contained within the DNA fragment [14]. Hence, it has been difficult to detect large-sized ITD mutations (30bp to 100bp) by NGS technology compared with point mutation or small indel mutation.

Despite difficulties in developing algorithm, NGS analysis is enabling us to calculate exact tumor burden and insertion length by counting alternative reads coverage. Therefore, detection algorithm of structural variant including FLT3-ITD using NGS data should be a deserving and powerful tool.

Especially, in cases of AML in which mutant allele length and burden have a huge impact on clinical outcome, exact measurement of allele length and burden is more important than any other diseases [15].
Material and Method

Patients and Sample preparation

Bone marrow samples were acquired from 81 patients with de novo acute myeloid leukemia (AML) diagnosed between June 2000 and December 2014. The clinical and laboratory characteristics of the patients were described in Table 1. The median age of patients was 50.39 years (IQR, 35-62) with a male to female ratio of 43:38.

DNA extraction

Bone marrow samples which were collected at the time of diagnosis and relapse were used to extract DNA samples. DNA extraction was performed using QuickGene DNA whole blood kit S(KURABO INDUSTRY, Japan) and Norgen blood genomic DNA isolation kit (Norgen Biotek Corp, cat#46300) according to the manufacturer’s recommendations.

WES and alignment

For whole exome sequencing (WES), I prepared and sequenced Exome using the Solexa sequencing technology.
platform (HiSeq2000, Illumina, SanDiego, CA, USA) following the manufacturer’s instructions. I randomly sheared 3ug of genomic DNA using Covaris System to generate about 150-bp inserts. The fragmented DNA was end-repaired using T4 DNA polymerase and Klenow polymerase, and Illumina paired-end adaptor-oligonucleotides were ligated to the sticky ends. I analyzed the ligation mixture by electrophoresis on an agarose gel, sliced and purified fragments with 200-250bp sizes. Purified DNA library was hybridized with SureSelect Human All Exons V3 and V5 probes set (Agilent, Santa Clara, CA, USA) to capture 50Mb targeted exons following manufacture’s instruction. I prepared the HiSeq2000 paired-end flow cell to the manufacturer’s protocol using captured exome library. Clusters of PCR colonies were then sequenced on the HiSeq2000 platform using recommended protocols from the manufacturer.

FASTQ files were aligned to human reference (human_g1k_v 37.fasta) by using the Burrows-Wheeler aligner (BWA-0.7.5) [16] to make SAM file. SortSam in Picard-tools-1.68 was used to convert to BAM file and this was sorted by chromosomes and processed through a PCR
duplicate marking process, which enables the Genome Analysis ToolKit [17](GATK-1.6.5) to ignore duplicates in subsequent processing. I performed a local realignment prior to recalibration, which gives the most accurate quality scores for each sample. Recalibration was performed to increase recalibration accuracy. All processes have default value.

Conventional PCR

I reviewed FLT3-ITD positivity in Electronic Medical Record (EMR) of Seoul National University Hospital. The method I performed for the diagnosis is shown below.

I designed primer to target FLT3 gene juxtmembrane domain exon 14. Forward primer sequence was 3’-CCCTTCCCTTTTATCCAAGA-5’ and reverse sequence was 3’-AACTGTGCCTCCCATTTTTG-5’. Composition of PCR mixture was as shown below: DW 37ul, 10X buffer 5ul, 10mM dNTP 1ul, Forward primer(10pmol/ul) 2ul, reverse primer(10pmol/ul) 2ul, Taq(5U/ul) 0.25 and DNA (100ng/ul) 3ul. Total 50ul mixture was used for PCR reaction. The cycle of PCR reaction was 30 cycles with the following condition: DNA denaturation condition was
30sec with 94°C and primer annealing condition was 30sec with 60°C and primer extension was 1 min with 72°C. The length of PCR product was 574bp.

**Nested PCR**

Nested PCR primer was designed to two type of primer. Primary primer was targeted at FLT3 between exon 14 and 15. Then, I designed second primer to more closely target the region at exon 14 which is commonly harboring ITD. Primary PCR product size was approximately 238bp and second PCR product size was 112bp. I used Macime PCR PreMix Kit (i-Tag iNtRon BIOTECHNOLOGY, lot# 32324447) contained DNA polymerase 2.5U, dNTPs 2.5mM each, Reaction buffer 1X and gel loading buffer 1X for PCR reaction. The cycle of PCR reaction was 25 cycles with following conditions: DNA denaturation condition was 30sec with 94°C and primer annealing condition was 30sec with 60°C and primer extension was 30sec with 72°C.

**Method for FLT3-ITD detection with NGS**

To detect ITDs with NGS data, the reads are aligned to a reference sequence (UCSC hg19), with BWA [1] which is a read aligner allowing soft-clipping (i.e., aligning only a
portion of read sequence with the rest unaligned). Due to duplication of a sequence, some reads were aligned to the position of ITD including breakpoint. Such reads can be an indication of the occurrence of ITD and BWA aligns those reads as soft-clipped (Figure 1). The procedure of FLT3-ITD detection is briefly described below (Figure 2):

1. Scan the positions in the region of FLT3 gene (e.g., chr13:28577000-28676800) where reads are aligned as soft-clipped. To avoid false alignments, I abandoned the reads whose soft clipped portion size is less than 10 bps.

2. For each position, check if the soft-clipped reads aligned to the position support a breakpoint of an ITD. For example, if the left-end portion of a read sequence is soft-clipped (and its right-end is properly aligned), the soft-clipped sequence is searched in the reference sequence ranged from the end position of the read alignment. As ITD size is smaller, the soft-clipped sequence is observed closer to the end position. Similarly, if a right-end portion is soft-clipped, its sequence is searched in the reference sequence ranged before the start position.
If the soft-clipped sequence is found in the reference sequence, it is considered as an ITD position candidate. In case of left-end soft-clipped reads, the sequence from the alignment start position to the end position of soft-clipped sequence in reference is considered duplicated sequence of an ITD.

ITD often includes insertion between duplications. In this case, searching the entire soft-clipped sequence does not work. To address the issue, I repeated searching soft-clipped sequence of smaller size from the end, until the size is less than 10bps (Figure 2-b).

3. For each candidate ITD position, I counted reads which support the occurrence of ITD. I first generated ITD sequence given predicted duplication from the previous step among the reads aligned to the predicted ITD position, I counted the reads which are aligned to the breakpoint in the generated sequence. If the number of reads is more than 3, then, ITD is called.

**Comparison of NGS Tools**

I performed NGS Tools to compare detecting power. Additionally, 3 tools developed for SV detection were used
in this study. (Pindel[18], Genomon-ITDetectector[19] and Breakmer [20]). 81 of aligned bam files were used equally for the tools, setting a default option. CollectInsertSizeMetrics module of Picard was used for calculating insertion size. In case of Pindel, I defined short insertion as duplicate changes if the short insertion is located in FLT3 gene locus.
Results

I performed three difference methods to detect FLT3-ITD variant in AML patient DNA samples. The one of them was a conventional PCR commonly used to determine whether the patient has FLT3-ITD. Another one was nested PCR which I designed to improve detection rate of FLT3-ITD. The last one is WES which is not specifically designed to detect the FLT3-ITD. I compared these three methods and evaluated sensitivity of FLT3-ITD detection rate.

To compare detection sensitivity, 81 AML WES data were used. Mean mapped depth coverage of FLT3 exon 14 (recurrently occurred FLT3-ITD region) is 71. Distribution of mapped depth coverage is 16.65 to 237.08. Detection rate of FLT3-ITD by WES method is 13% (11/81). In this cohort, positive rate of nested PCR was 18% (15/81), which was more sensitive. In contrast, conventional PCR detected only 7% (6/81) showing least sensitivity.

Between nested PCR and WES, concordance rate was 95% (79/81). All of 11 samples with FLT3-ITD detected by WES were validated by nested PCR. 4 patients were
additionally detected by nested PCR. In these 4 patients, blast percentages were 9.7%, 81.9%, 14.8% and 51.9% respectively. When compared WES and conventional PCR, 5 patients showed discrepancy and among them, 1 patient who had abnormal cytogenetic changes, was not tested by conventional PCR. The blast percentages of remaining 4 patients were over 80% (89.8%, 93%, 85.7% and 80.3% respectively). In contrast, tumor burdens harboring FLT3-ITD, determined by coverage depth of exon14 per variant read (Tumor burden=variant allele/exon14 coverage), are not as high as observed blast percentage, implying that FLT3-ITD faction did not correlate with actual blast proportion. To analyze the association, I performed correlation analysis. There was no correlation between tumor burden with FLT3-ITD and blast percentage (r=0.248) (Figure 3).

All of FLT3-ITD mutations occurred within exon 14 which has been reported in a recurrently mutated region. Distribution of insertion start site of FLT3-ITD was chr13:28608200 to 28608267 (reference sequence: GRCh37). Inserted lengths of FLT3-ITD were various. The largest insertion size is 113bp and the smallest is 21bp.
Interestingly, FLT3-ITD mutations of 2 patients (patient1, patient7) showed that mutant alleles of different sizes were inserted. Duplicated sequence of 60bp and 21bp size were inserted in two different positions (chr13:28608247 and chr13:28608262) respectively in patient7. Patient 1 had alleles of different sizes (39bp and 75bp) inserted in the same position (chr13:28608227) (Table 1).

I compared result by multiple NGS tools known as good performance for SV (Pindel, Denomon-ITDector and Breakmer). My method called 11 patients positive for FLT3-ITD. However, conventional tools were less sensitive. Pindel and Genomon-ITDector had 72% power of detection (8/11), and 45% detection power was shown by Breakmer, compared to my new method. (Figure 4)
Discussion

Diagnosis of FLT3-ITD is a very important genetic factor, guiding a therapeutic direction for those patients harboring FLT3-ITD mutations. The therapeutic approach for patients with FLT3-ITD has traditionally been intensive induction chemotherapy, followed by consolidative chemotherapy or hematopoietic cell transplantation (HCT). In recent years, multiple small-molecule inhibitors of the FLT3 tyrosine kinase have been studied pre-clinically and clinically [21]. Although FLT3-ITD is an obviously poor prognosis factor of AML, there have been many difficulties because of low sensitivity in detecting these abnormalities.

To solve these problems, I have developed new method based on NGS and validated with nested PCR which is more sensitive than conventional PCR. As a result, 4 more patients were additionally detected by NGS method and validated by nested PCR. Impressively, when I compared mine with several other NGS Tools, my tool is more powerful for detection of FLT3-ITD.
Exceptionally, one patient with FLT3-ITD detected by conventional PCR was not found by NGS and nested PCR both. However, mapping depth of FLT3 was not enough to detect structure variation (mapping depth of exon 14 was 42 x), considering the fact that good quality NGS data has been usually possible in cases of reading depth of above 100 x by Whole Exome Sequencing (WES) data.

Quantitation of tumor burden, one of the typical features by NGS methods, is done giving rise to exact value of tumor cell burden harboring FLT3-ITD. I analyzed correlation analysis between blast percent and tumor burden with FLT-ITD in order to clarify reason of differences in sensitivity and to establish the role of NGS technique in evaluating FLT3-ITD mutations compared to other methods.

Tumor burden estimated by FLT3-ITD and morphologic blast percentage at the time of diagnosis showed little correlation (r=-0.248) and this may explain why some of the patients with FLT3-ITD having higher blast percentage were negative by conventional PCR.
In conclusion, even though, neoplastic white blood cell are observed abundantly in patients’ bone marrow, the proportion of tumor cells with FLT3-ITD may be much less. This phenomenon may reflect the heterogeneity of cancer, which lowers the sensitivity of conventional PCR.

Here I developed new NGS method which was more sensitive than other existing methods. This method was validated by nested PCR. It also enables measurement of FLT3-ITD size and quantitation of FLT3-ITD amount in AML patients.
**Figure legend**

**Figure 1.** IGV view of read alignments mapped to the position of FLT3-ITD (chr13:28608227) whose size is 75 bps. This FLT3-ITD is called with 23 supporting reads. The reads with a series of mismatch bases are soft-clipped reads. (Sample ID: Patient 6)

**Figure 2.** The procedure of FLT3-ITD detection with NGS data. a) Schematic diagram of ITD sequence and NGS short reads support the breakpoint of ITD b) Reference sequence without ITD and alignments of the reads with ITD breakpoint to the reference. c) ITD sequence generated with predicted duplication. To finally call an ITD, I counted reads aligned to the breakpoint of the sequence (supporting reads). d) Flowchart of our FLT3-ITD detection method.

**Figure 3.** Correlation analysis between blast percent and FLT-ITD burden. Eleven patients with FLT3-ITD by WES were used for this analysis.
**Figure 4.** Comparison of NGS Tools. Blue color indicated positive results by our method. Green, yellow and red indicated each Tools’ results which are compared with my NGS tool.
Figure 1. IGV view of read alignments aligned to the position of FLT3-ITD
Figure 2. The procedure of FLT3-ITD detection with NGS data

**a** Alternative sequence with ITD

- Read 1
- Read 2

**b** Reference sequence

- Searching read fragment

**c** ITD sequence

- Perfect match
- Supporting Read
- Soft-clipped portion
- Insertion
- Breakpoint

**d** Sorted BAM

- Scan soft-clipping positions
- Identify ITD position candidates
- Count reads to support duplication (breakpoint)
  - ≥ min. # of reads
- ITD call

VCF

 Seoul National University
<table>
<thead>
<tr>
<th>Patient</th>
<th>ITD size</th>
<th>Start pos (exon#)</th>
<th>#ALTReads</th>
<th>Nested</th>
<th>Conventional</th>
<th>Depth of exon14</th>
</tr>
</thead>
<tbody>
<tr>
<td>patient 1</td>
<td>39/75</td>
<td>28608227(14)</td>
<td>9/5</td>
<td>P</td>
<td>P</td>
<td>65.49</td>
</tr>
<tr>
<td>patient 2</td>
<td>30</td>
<td>28608267(14)</td>
<td>32</td>
<td>P</td>
<td>P</td>
<td>84.93</td>
</tr>
<tr>
<td>patient 3</td>
<td>49</td>
<td>28608217(14)</td>
<td>18</td>
<td>P</td>
<td>P</td>
<td>37.91</td>
</tr>
<tr>
<td>patient 4</td>
<td>52</td>
<td>28608219(14)</td>
<td>6</td>
<td>P</td>
<td>N</td>
<td>94.12</td>
</tr>
<tr>
<td>patient 5</td>
<td>113</td>
<td>28608200(14)</td>
<td>18</td>
<td>P</td>
<td>ND</td>
<td>106.38</td>
</tr>
<tr>
<td>patient 6</td>
<td>51</td>
<td>28608262(14)</td>
<td>34</td>
<td>P</td>
<td>P</td>
<td>71</td>
</tr>
<tr>
<td>patient 7</td>
<td>60/21</td>
<td>28608247/28608262(14)</td>
<td>3/5</td>
<td>P</td>
<td>N</td>
<td>84.65</td>
</tr>
<tr>
<td>patient 8</td>
<td>51</td>
<td>28608216(14)</td>
<td>4</td>
<td>P</td>
<td>N</td>
<td>121.11</td>
</tr>
<tr>
<td>patient 9</td>
<td>45</td>
<td>28608262(14)</td>
<td>9</td>
<td>P</td>
<td>P</td>
<td>29.59</td>
</tr>
<tr>
<td>patient 10</td>
<td>78</td>
<td>28608239(14)</td>
<td>4</td>
<td>P</td>
<td>P</td>
<td>31.7</td>
</tr>
<tr>
<td>patient 11</td>
<td>60</td>
<td>28608218(14)</td>
<td>50</td>
<td>P</td>
<td>N</td>
<td>61.73</td>
</tr>
<tr>
<td>patient 12</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>P</td>
<td>ND</td>
<td>61.73</td>
</tr>
<tr>
<td>patient 13</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>P</td>
<td>ND</td>
<td>18.62</td>
</tr>
<tr>
<td>patient 14</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>P</td>
<td>N</td>
<td>97.45</td>
</tr>
<tr>
<td>patient 15</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>P</td>
<td>N</td>
<td>115.43</td>
</tr>
<tr>
<td>patient 16</td>
<td>-</td>
<td>-</td>
<td>ND</td>
<td>P</td>
<td></td>
<td>42.51</td>
</tr>
</tbody>
</table>
Figure 3. Correlation analysis between blast percent and FLT-ITD burden.

![Correlation Analysis Graph](image)
### Figure 4. Comparison of NGS methods

<table>
<thead>
<tr>
<th>Patient</th>
<th>Nested</th>
<th>New</th>
<th>Pindel</th>
<th>Genomon ITDector</th>
<th>Breakmer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient 1</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>X</td>
</tr>
<tr>
<td>Patient 2</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Patient 3</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Patient 4</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>X</td>
<td>O</td>
</tr>
<tr>
<td>Patient 5</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>X</td>
</tr>
<tr>
<td>Patient 6</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Patient 7</td>
<td>O</td>
<td>O</td>
<td>X</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Patient 8</td>
<td>O</td>
<td>O</td>
<td>X</td>
<td>O</td>
<td>X</td>
</tr>
<tr>
<td>Patient 9</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>X</td>
</tr>
<tr>
<td>Patient 10</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Patient 11</td>
<td>O</td>
<td>O</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Patient 12</td>
<td>O</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Patient 13</td>
<td>O</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Patient 14</td>
<td>O</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Patient 15</td>
<td>O</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>


17. McKenna, A., et al., *The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA


초 록

서론: 급성골수성 백혈병에 있어서 FLT3 유전자의 내부 중복 삼입 서열을 예민하게 감지 하는 것은 매우 중요하다. 하지만 구조적인 문제점이나 방법론 등에 의해 내부 중복 삼입 서열을 정확하게 감지하는 것에는 어려움이 있다. 우리는 FLT3 유전자를 정확하고 민감하게 감지하기 위해 차세대 염기서열 시퀀스 기법을 이용한 새로운 알고리즘을 고안하였고 이 결과를 중첩 중합효소 연쇄 반응을 이용하여 검증 및 비교분석 하였다.

방법: 81명의 급성골수성 백혈병 환자의 전장 엑솜 유전체 데이터를 이용하여 고안된 FLT3 유전자의 내부중복 삼입 서열 감지 알고리즘에 대입하였다. 간단히 알고리즘의 방법론에 대해 말하자면, 시퀀싱된 유전자 조각들은 BWA 를 이용하여 인간 기준서열에 배열된다. 이렇게 배열된 유전자 조각들 중 기준서열과 맞지 않은 조각들은 부리진 조각의 표시로 남게 되고 이를 검출 하여 내부 중복서열이라 정의한다. 이렇게 정의된 결과를 중첩 중합효소 연쇄반응을 이용하여 검증하였다. 첫번째 프라이머는 FLT3 유전자의 엑손 14번과
15번을 대상으로 하여 제작되었다. 이를 이용하여 중합효소 연쇄 반응을 진행하였으며 이 중복 DNA를 이용하여 중첩 중합효소 연쇄반응을 진행하였다. 중첩 프라이머는 FLT3 유전자의 내부 중복 삽입 서열이 빈번하게 관찰되는 양손 14번을 대상으로 하여 제작되었다.

결과: 81명의 급성골수성 백혈병 환자의 실험결과 FLT3 유전자의 내부 중복 삽입 서열이라 진단받은 환자는 7명이었다. 하지만 새로 고안된 차세대 염기서열 시퀀싱을 기반으로 한 방법으로 적용하였을 때, 11명의FLT3 내부 중복 삽입서열이 발견되었다 (11/81, 13%). 중첩 중합효소 연쇄반응을 이용하여 검증하였을 때, 11명 모두 FLT3 유전자의 내부 중복 삽입서열이 보여졌다. FLT3 유전자의 내부 중복 삽입 서열이 발견된 환자의 경우 중첩중합효소 연쇄반응을 이용한 실험은 희미한 실험 결과로 인해 확인이 불가하였으며, 차세대 염기서열을 이용한 방법으로는 음성이었다. 하지만 이 환자의 차세대 염기서열 시퀀싱 실험결과, FLT3 유전자의 양손 14번에 위치한 유전체 조각의 깊이가 깊지 않아 정확한 판단을 할 수 없었다.
결론: FLT3 유전자의 내부 중복 삽입 서열은 급성골수성 백혈병 환자의 치료 방법을 결정하는 중요한 유전인자 중 하나이다. 본 연구는 이를 민감하게 검색하기 위해 차세대 염기서열 시퀀싱 방법을 이용한 검출 방법을 고안했다. 이 방법은 FLT3 유전자의 내부 중복 삽입 서열을 민감하게 검출 할 수 있을 뿐만 아니라 일반적인 PCR 방법으로 정확히 측정 할 수 없는 삽입 서열의 길이와 정량이 가능한 장점을 가지고 있는 것이 특징이다.

주요어: FLT3 유전자의 내부 중복 삽입 서열 돌연변이, 급성 골수성 백혈병, 차세대 염기서열 시퀀싱, 중첩 중합효소 연쇄반응, 유전 인자
학 번: 2014-21156