DE GRUYTER OPEN

Research article

DOI: 10.2478/rrlm-2014-0003

A new assay to identify recurrent mutations in acute myeloid leukemia using next-generation sequencing

O nouă metodă de identificare a mutaţiilor recurente în leucemia acută mieloblastică folosind secvenţierea de nouă generaţie

Daniel Coriu[1,2], Dumitru Jardan[1,3*], Cerasela Jardan[1, 2], Rodica Tălmaci[1,2], Mihaela Dragomir[1,2], Anca Coliţă[1,2]

1. University of Medicine and Pharmacy "Carol Davila", Bucharest, Romania;
2. Center of Hematology and Bone Marrow Transplantation, Fundeni Clinical Institute, Bucharest, Romania
3. SC Medlife SA, Bucharest, Romania

Abstract

***Introduction**: Acute myeloid leukemia (AML) is a heterogeneous disease characterized by a late onset (it is rare in children), aggressive phenotype and dismal prognosis especially in patients in the older group (>65 years). For risk stratification of patients standard cytogenetic is used along with molecular techniques for point mutation identification. Here we describe a new method using next generation sequencing for identification of mutation in 5 AML recurrently mutated genes - RUNX1, FLT3, DNMT3A, IDH1 and IDH2. **Materials and methods**: Samples from 40 patients with normal karyotype AML referred to Fundeni Clinical Institute were sequenced. Primer design was performed using LaserGene Genomics suit. Next generation sequencing was performed on MiSeq (Illumina) and results were analyzed using LaserGene Genomics suit. Results of next generation sequencing were compared to Sanger sequencing. **Results**: No additional mutations were identified in samples from nine patients presenting FLT3-ITD and/or NPM1 mutations. In 25 out of 31 patients with normal karyotype and no FLT3-ITD and NPM1 mutations, we identified mutations in one of the 5 aforementioned genes. All these mutations identified by next generation sequencing were confirmed using the classical Sanger sequencing. **Conclusions**: We validated a very useful method for mutation identification in AML patients using next generation sequencing. There are many advantages exhibited by this method: it is more cost efficient and it has a higher sensitivity of mutation detection than Sanger sequencing, it has been described as being quantitative and in our case it allowed risk stratification for most of the normal karyotype AML samples which were FLT3-ITD and NPM1 negative.*

***Keywords**: next generation sequencing, AML, recurrent mutations*

Rezumat

Introducere: Leucemia acută mieloblastică (LAM) este o boală heterogenă caracterizată prin debut la vârstă avansată, fenotip agresiv şi prognostic nefavorabil în special în grupul de vârstă de peste 65 de ani. Pentru stratificarea pacienţilor în grupe de risc se utilizează citogenetica clasică împreună cu metodele moleculare pentru

***Corresponding author**: Dumitru Jardan, Address: Department of Hematology, Fundeni Clinical Institute, Bucharest, Romania, phone:+40735300999, e-mail: d.jardan@gmail.com

identificarea mutaţiilor punctiforme. În acest articol descriem o nouă metodă de identificare a mutaţiilor în 5 gene implicate în LAM: RUNX1, FLT3, DNMT3A, IDH1 şi IDH2 utilizând secvenţierea de nouă generaţie. Materiale şi metode: Au fost secvenţiate probe de la 40 de pacienţi cu LAM cu cariotip normal internaţi în Institutul Clinic Fundeni. Design-ul de primeri a fost efectuat utilizând LaserGene Genomics suit. Secvenţierea de nouă generaţie a fost efectuată pe platforma MiSeq de la Illumina. Rezultatele au fost analizate utilizând LaserGene Genomics suit. Rezultatele obţinute prin secvenţierea de nouă generaţie au fost comparate cu secvenţierea Sanger. Rezultate: Nu au fost identificate mutaţii adiţionale în probele de la nouă pacienţi pozitivi pentru mutaţiile FLT3-ITD şi / sau NPM1. În probele de la 25 din 31 de pacienţi, cu cariotip normal şi fără mutaţii FLT3-ITD şi NPM1, au fost identificate mutaţii în una din cele 5 gene studiate. Toate aceste mutaţii, identificate prin secvenţierea de nouă generaţie, au fost confirmate prin metoda de secvenţiere clasică Sanger. Concluzii: În acest studiu am validat o metodă de identificare a mutaţiilor apărute la pacienţii cu LAM utilizând secvenţierea de nouă generaţie. Această metodă prezintă o serie de avantaje: este mai ieftină ca in cazul secvenţierii Sanger, prezintă o sensibilitate crescută pentru detectarea mutaţiilor, a fost descrisă ca fiind cantitativă şi în cazul nostru a permis stratificarea în grupe de risc a majorităţii pacienţilor cu cariotip normal şi fără mutaţiile FLT3-ITD şi NPM1.

Cuvinte cheie: *Secvenţiere de nouă generaţie, LAM, mutaţii recurente*

Received: 18th November 2013; Accepted: 8th February 2014; Published: 16th February 2014.

Introduction

Acute myeloid leukemia (AML) is a heterogeneous disease characterized by a late onset (it is rare in children), aggressive phenotype and dismal prognosis especially in patients in the older group (>65 years) (1). Due to its clinical evolution AML attracted a lot of attention as it has (as most of the hematological neoplasms) easily accessible neoplasm tissue. Thus the first cancer sample to have its genome sequenced was an AML sample from a woman (2). In that study authors discovered many new mutations: a part of them being passenger mutations and a part of them being driver mutations as they were implicated in well known signaling networks associated with cancer. Later eight pairs of presentation and relapse samples were sequenced allowing discovery of new driver mutations associated with disease progression (3).

AML research led to discovery of mutations in many genes and allowed development of risk stratification algorithm, which incorporate these mutations as prognostic factors (4). First mutations were identified as fusion genes by standard cytogenetics. The most common fusion genes in AML are: *PML-RARa, CBFb-MYH11, MLL-AF9, AML1-ETO* (1). Use of molecular techniques, besides cytogenetics, led to discovery of a whole new array of mutations: *MLL-PTD* (partial tandem duplication in *MLL*), *FLT3-ITD* (internal tandem duplication in *FLT3*), *FLT3-TKD* (tyrosine kinase domain mutation in *FLT3*), tetranucleotide insertion in *NPM1* exon 12 (coding for C-terminal region of *NPM1*), mutations in *CEBPA* and *RUNX1*. In the last couple of years still new recurrent mutations were identified in a number of genes: *DNMT3A*, *IDH1* and *IDH2* (5,6,7).

Risk stratification in AML is primarily influenced by cytogenetic analyses (1). Use of cytogenetics for risk stratification allows evaluation of ~50% of the patient as the rest of them present a normal karyotype (1). For a long time, patients with normal cytogenetics were classified as intermediate risk group although it was clear that some these patients have a favorable clinical evolution and some a dismal one (8). Introduction of molecular techniques allowed risk stratification for normal karyotype patients using gene mutations as prognostic factors (1, 8). *MLL-PTD, FLT3-ITD, FLT3-TKD* and *NPM1* mutations are among the first gene mutations with recognized

prognostic impact. Their investigation allows an improved and more accurate patient stratification. The use of information regarding additional gene mutations provides an exciting opportunity for risk stratification of almost all AML patients, and ultimately for an improved clinical management of these patients (8).

In our laboratory, we routinely analyze AML patients using cytogenetic and molecular techniques. Thus besides a standard karyotype, we perform identification of 4 most common fusion genes by RT-PCR: *PML-RARa, CBCb-MYH11, AML1-ET*O and *MLL-AF9*. Also, we perform mutational analysis to identify following mutation – *FLT3-I*TD and tetranucleotide insertion in *NPM1* exon 12. For risk stratification purposes, karyotype is not always conclusive as it can identify some chromosome abnormalities, which are rare or were not described in literature and as such they do not provide any prognostic value. Also in ~10% of cases no metaphases could be obtained. RT-PCR for fusion genes can stratify ~35% of the patients as they present one of the above listed fusions (9). Still, in case of *AML1-ETO* and *CBFb-MYH11* additional analysis of *c-KIT* mutation should be performed as ~30% of those patients will present a mutation in *c-KIT* (10, 11). Identification of mutation in *c-KIT* in those patients changes prognosis for the patients from good to intermediate (11). Mutations in *FLT3* and *NPM1* could further stratify 20-25% of patients. Thus in our case use of all the above techniques allows risk stratification of about 60% to 65% of the cases (many *FLT3-ITD* mutations are described in *PML-RARa* patients thus the above percentages are not directly additive).

In this study we describe a new Next Generation Sequencing method which allows identification of mutations in hot-spots of 5 genes: *FLT3* (non-ITD mutation), *RUNX1*, *DNMT3A*, *IDH1* and *IDH2*.

Materials and methods

Patients and samples

In the first phase normal blood DNA was used for assay development. After initial validation peripheral blood samples at presentation from 40 patients with normal karyotype AML were used (these patients were also tested for *FLT3-ITD* and *NPM1* mutations). Samples were obtained from patients referred to Fundeni Clinical Institute, Hematology Department in the period between January 2010 and December 2012. This study was conducted according with local and national ethical rules.

DNA extraction was done using TRIZOL reagent, using manufacturer recommendations (Life technologies).

Primer design

Hot-spots in *FLT3, RUNX1, DNMT3A, IDH1* and *IDH2* were determined using mutational data from COSMIC database (http://cancer.sanger.ac.uk/cancergenome/ projects/cosmic/). Primer design was done using LaserGene Genomics Suite (DNASTAR). Primers were designed so that amplicons are about 150 bp long. We have designed primers for one region of *DNMT3A*, one region of *IDH1*, one region of *FLT3*, two regions of *IDH2* and 5 regions of *RUNX1*. Primer sequences are listed in *Table I*.

PCR amplification

In the first phase we did singleplex PCR for library preparation and testing of amplification efficiency for all of the primers. In the second phase we tested and used a multiplexed assay, which allows amplification of all the regions in 2 tubes. For library preparation in single and multiplex reactions 100 ng of DNA per reaction was used. FastStart Taq DNA polymerase was used (Roche Applied Science).

PCR program: one cycle - 10 minutes at 95°C and 35 cycles - 1 minute at 95°C, 1 minute at 60°C, and 1 minute at 72°C with no extension step.

Table I. Primer sequences for one region of DNMT3A, one region of IDH1, one region of FLT3, two regions of IDH2 and 5 regions of RUNX1 genes.

Gene region	Aminoacids sequenced	Primer sequence
DNMT3A-exon23	880-912	primer-f AGTCCACTATACTGACGTCTCCAA primer-r CTACCTCAGTTTGCCCCCAT
IDH1-exon4	109-138	primer-f TATTCTGGGTGGCACGGTCT primer-r CACATTATTGCCAACATGACTTAC
IDH2-exon4_fr1	136-163	primer-f1 GCTGAAGAAGATGTGGAAAAGTC primer-r1 AATGGTGATGGGCTTGGTCC
IDH2-exon4_fr2	145 -178	primer-f2 CTATCCGGAACATCCTGGGG primer-r2 GGATCCCCTCTCCACCCTG
FLT3_exon20	833-847	primer-f TGGTGAAGATATGTGACTTTGGAT primer-r CATGATAACGACACAACACAAAAT
RUNX1_exon3-1	54-72	primer-f GAGCCCAGGCAAGATGAGC primer-r CGGCCAGCACCTCCACCAT
RUNX1_exon3-2	82-115	primer-f CCGCAGCATGGTGGAGGTG primer-r TCTCCGGGCCAGTACCTTGAA
RUNX1_exon4	140-169	primer-f ATGGCTGGCAATGATGAAAAC primer-r CAAGCATAGTTTTGACAGATAACG
RUNX1_exon5	188-204	primer-f TTCACAAACCCACCGCAAGT primer-r GGTGTACCAGCCCCAAGTG
RUNX1_exon7	297-322	primer-f ATCCATTGCCTCTCCTTCTG primer-r CTCAGCTGCAAAGAATGTGTT

Library preparation

After PCR amplification of the regions of interest we performed amplicon clean-up using Agencourt AMPure XP kit (Beckman-Coulter) following manufacturer's instructions. A second round of PCR was done using Nextera Indexing primers (Illumina) to introduce indexes for sample demultiplexing, following manufacturer's instructions. After indexing we performed a second round of amplicon clean-up using Agencourt AMPure XP kit (Beckman-Coulter) following manufacturer's instructions. To quantify purified library we used Qubit 2.0 Fluorometer and dsDNA HS Assay (LifeTechnologies).

Sanger Sequencing

For assay validation we compared results obtained using next generation sequencing with Sanger sequencing. We sequenced amplicons using the same primers as presented in *table I*. For

sequencing we used BigDye terminator version 3.1 chemistry and ABI 3130 machine (LifeTechnologies). No difference was observed between Sanger and next-generation sequencing regarding identified mutations.

Next Generation Sequencing

For Next Generation Sequencing we used MiSeq second generation machine (Illumina). Miseq machine uses a SBS chemistry (sequencing by synthesis). We used MiSeq Reagent Nano Kit v2 which provide ~ 1 million of clusters. Library quantification and loading into reagent cartridge was done following manufacturer's instruction.

Data analysis

Generated data was analyzed using LaserGene Genomics Suite version 11.2 (DNASTAR) and visualized either using SeqManPro package (part of the LaserGene Genomics Suite) or Integrated Genomic Viewer (Broad Institute). Human reference build hg19 was used.

Coverage varied between 20000x to 50000x as estimated by Integrated Genomic Viewer. Minimum variant frequency for clinical reporting was set to 5%.

Results

In the first phase of assay development we used singleplex PCR for library preparation using normal blood DNA. This library was pooled and sequenced on both ABI 3130 and MiSeq machines. As expected there were no differences in sequence (we obtained a wild-type sequence as expected). After that we developed a multiplex assay which allows library amplification in 2 tubes – we multiplexed primers for *DNMT3A*-exon23, *IDH1*-exon4, *IDH2*-exon4_fr1, *RUNX1*-exon3-fr1 and *RUNX1*-exon5 in tube one and *IDH2*-exon4_fr2, *FLT3*-exon20, *RUNX1*-exon3-fr2, *RUNX1*-exon4 and *RUNX1*-exon7 (*Table I*) in the second tube (note all Sanger sequencing was performed in singleplex mode). In the multiplex library preparation mode all amplicons exhibited approximately the same amplification efficiency thus from now on we used only multiplex assay for library preparation. After this initial validation we used 40 clinical samples from patients with AML for mutation identification. Nine of the 40 patients presented mutations *FLT3-ITD* and/or in *NPM1* – 2 samples were *FLT3-ITD* positive and *NPM1* negative, 5 samples were *FLT3-ITD* negative and *NPM1* positive and 2 samples were positive for both mutations. After sequencing on Miseq, we identified 25 patients with mutations in one of the 5 genes tested. Identified mutations are listed in *Table II*. No additional mutations were identified in nine samples presenting *FLT3-ITD* and/or *NPM1* mutations. All these mutations identified

Table II. Identified mutations of DNMT3A, IDH1, FLT3, IDH2 and RUNX1 genes, in patients with AML.

Gene	Mutation	Numer of samples
DNMT3A	R882C	2
	R882H	5
	K906E	1
FLT3	Y482H	1
	D839N	1
IDH1	R132C	2
	R132H	3
IDH2	R140Q	2
	R172K	2
RUNX1	R162K	1
	R166L	1
	R201Q	1
	R201X	1
	R204X	1
	R204Q	1

by next generation sequencing were confirmed using the classical Sanger sequencing.

Discussions

We describe here a new method for mutation identification in AML patients using next generation sequencing. This method provides a series of advantages over traditional methods of mutation identification – library preparation can be multiplexed saving money and time, in one sequencing run can be sequenced tens of targets at an affordable price (cost of next generation sequencing is lower than of Sanger sequencing if using sufficient samples per run), sensitivity of mutation detection can go down to 1% which is better than 20% in case of Sanger sequencing, mutation detection is quantitative which can give an idea of mutation clone prevalence (12). Also using some modifications of the technique this method potentially allows for identification of very rare mutations being useful for minimal residual disease (MRD)(13) quantification, which could be very useful in AML because in less than 40% a suitable MRD marker can be identified.

In our settings this method allowed identification of mutations in 25 out of 31 patient samples, which were classified in intermediate risk group as they presented normal karyotype, and no mutations in *FLT3-ITD* and *NPM1*. Adding this method of mutation detection to our existing panel of investigations would significantly improve the risk stratification of AML patients. We performed mutation identification on a relatively small number of samples and according to literature some of the mutations in these genes overlap and the actual number of patient for which we could obtain risk stratification would actually be lower (5). But this problem could be solved by adding identification of mutations in some other genes like *CEBPA* (5).

This method allowed identification of three unusual mutations – K906E in *DNMT3A*, which is a mutation, not found in the literature (most of the mutations in this gene affect position R882). This mutation will probably affect structure and function of the protein as PolyPhen algorithm is listing this mutation as probably damaging with a score of 0.999 (sensitivity: 0.14; specificity: 0.99) (http://genetics.bwh.harvard.edu/pph2/). Also other interesting mutations were identified - R201X and R204X in *RUNX1*. These are unusual mutations as they truncate resulting protein and thus probably no active protein is produced, but these were previously described (14).

From clinical point of view, identification of mutations in these five genes is very important – mutations in *DNMT3A*, *FLT3* and *RUNX1* have an associated poor prognosis whereas mutations in *IDH1* and *IDH2* have an associated good prognosis (8). Also *IDH1* and *IDH2* were described as useful targets for small molecule inhibitors and drug development companies conduct an active research in this area thus making these two genes potential markers for personalized medicine in AML (15).

Conclusions

We tested a very useful method for mutation identification in AML patients using next generation sequencing. There are many advantages exhibited by this method: it is more cost efficient than Sanger sequencing, it has a higher sensitivity of mutation detection, it has been described as being quantitative and in our case it allowed risk stratification for most of the normal karyotype AML samples which were *FLT3-ITD* and *NPM1* negative. Still next generation sequencing should be mainly regarded for research use as the FDA gold standard is Sanger sequencing.

In comparison to other methods of mutation identification – genotyping methods (as opposed to sequencing) – this method allows identification of all point mutations in the analyzed regions with sufficient sensitivity to be clinically relevant.

The potential of next generation sequencing to identify rare mutation opens a new way for MRD detection using gene mutations as MRD markers and this will make screening for as many gene mutations as we can a priority in AML diagnosis.

This work was supported by the grants PN 41-087 /PN2-099 from the Romanian Ministry of Research and Technology.

List of abbreviations:

AML – acute myeloid leukemia
FLT3-ITD - internal tandem duplication in FLT3
FLT3-TKD - tyrosine kinase domain mutation in FLT3
MLL-PTD - partial tandem duplication in MLL
MRD – Minimal residual disease

Conflict of interest

Authors declare no conflict of interest.

References

1. Vardiman JW, Thiele J, Arber DA, Brunning RD, Borowitz RD, Porwit A, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood. 2009 July 30;114(5):937-951. DOI: 10.1182/blood-2009-03-209262
2. Ley TJ, Mardis ER, Ding L, Fulton B, McLellan MD, Chen K, et al. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. Nature. 2008 Nov 6;456(7218):66-72. DOI: 10.1038/nature07485
3. Ding L, Ley TJ, Larson DE, Miller CA, Koboldt DC, Welch JS, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature. 2012 Jan 11;481(7382):506-10. DOI: 10.1038/nature10738
4. Damm F, Heuser M, Morgan M, Wagner K, Görlich K, Großhennig A, et al. Integrative prognostic risk score in acute myeloid leukemia with normal karyotype. Blood. 2011 April 28;117(17):4561-4568. DOI: 10.1182/blood-2010-08-303479
5. Patel JP, Gönen P, Figueroa ME, Fernandez H, Sun Z, Racevskis J, et al. Prognostic Relevance of Integrated Genetic Profiling in Acute Myeloid Leukemia. N Engl J Med. 2012;366:1079-1089. DOI: 10.1056/NEJMoa1112304
6. Patel KP, Ravandi F, Ma D, Paladugu A, Barkoh BA, Medeiros LJ, et al. Acute Myeloid Leukemia With IDH1 or IDH2 Mutation Frequency and Clinicopathologic Features. Am J Clin Pathol. 2011 Jan;135(1):35-45. DOI: 10.1309/AJCPD7NR2RMNQDVF
7. Ley TJ, Ding L, Walter MJ, McLellan MD, Lamprecht T, Larson DE, et al. DNMT3A Mutations in Acute Myeloid Leukemia. N Engl J Med. 2010;363:2424-2433. DOI: 10.1056/NEJMoa1005143
8. Estey EH. Acute myeloid leukemia: 2012 update on diagnosis, risk stratification, and management. Am J Hematol. 2012 Jan;87(1):89-99. DOI: 10.1002/ajh.22246
9. van Dongen JJ, Macintyre EA, Gabert JA, Delabesse E, Rossi V, Saglio G, et al. Standardized RT-PCR analysis of fusion gene transcripts from chromosome aberrations in acute leukemia for detection of minimal residual disease. Report of the BIOMED-1 Concerted Action: investigation of minimal residual disease in acute leukemia. Leukemia. 1999 Dec;13(12):1901-28. DOI: 10.1038/sj.leu.2401592
10. Cairoli R, Beghini A, Grillo G, Nadali G, Elice F, Ripamonti CR, et al. Prognostic impact of c-KIT mutations in core binding factor leukemias: an Italian retrospective study. Blood 2006. 107(9):3463-8. DOI: 10.1182/blood-2005-09-3640
11. Care RS, Valk PM, Goodeve AC, Abu-Duhier FM, Geertsma-Kleinekoort W, Wilson GA, et al. Incidence and prognosis of c-KIT and FLT3 mutations in core binding factor (CBF) acute myeloid leukaemias. Br J Haematol. 2003 Jun;121(5):775-7. DOI: 10.1046/j.1365-2141.2003.04362.x
12. Ladetto M, Bruggemann M, Ferrero S, Pepin F, Drandi D, Monitillo L, et al. Next-Generation Sequencing and Real-Time Quantitative PCR for Minimal Residual Disease (MRD) Detection Using the Immunoglobulin Heavy Chain Variable Region: A Methodical Comparison in Acute Lymphoblastic Leukemia (ALL), Mantle Cell Lymphoma (MCL) and Multiple Myeloma (MM). ASH 2012, abstract 788.
13. Schmitt MW, Kennedy S R, Salk JJ, Fox E J, Hiatt JB, Loeb LA. Detection of ultra-rare mutations by next-generation sequencing. Proc Natl Acad Sci USA. 2012 Sep 4;109(36):14508-13. DOI: 10.1073/pnas.1208715109
14. Gaidzik VI, Bullinger L, Schlenk RF, Zimmermann AS, Röck J, Paschka P, et al. RUNX1 mutations in acute myeloid leukemia: results from a comprehensive genetic and clinical analysis from the AML study group. J Clin Oncol. 2011 Apr 1;29(10):1364-72. DOI: 10.1200/JCO.2010.30.7926
15. Wang F, Travins J, DeLaBarre B, Penard-Lacronique V, Schalm S, Hansen E, et al. Targeted Inhibition of Mutant IDH2 in Leukemia Cells Induces Cellular Differentiation. Science. 2013 May 3;340(6132):622–6. DOI: 10.1126/science.1234769