ARTICLE

Targeted RNA-sequencing for the quantification of measurable residual disease in acute myeloid leukemia

Laura W. Dillon,[1] Sheida Hayati,[1,2] Gregory W. Roloff,[1] Ilker Tunc,[3] Mehdi Pirooznia,[3] Antonina Mitrofanova[2] and Christopher S. Hourigan[1]

[1]Laboratory of Myeloid Malignancies, Hematology Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD; [2]Department of Health Informatics, Rutgers School of Health Professions, Rutgers, The State University of New Jersey, Newark, NJ and [3]Bioinformatics and Computational Biology Core Facility, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, USA

Ferrata Storti Foundation

Haematologica 2019
Volume 104(2):297-304

ABSTRACT

Great effort is spent on developing therapies to improve the dire outcomes of those diagnosed with acute myeloid leukemia. The methods for quantifying response to therapeutic intervention have however lacked sensitivity. Patients achieving a complete remission as defined by conventional cytomorphological methods therefore remain at risk of subsequent relapse due to disease persistence. Improved risk stratification is possible based on tests designed to detect this residual leukemic burden (measurable residual disease). However, acute myeloid leukemia is a genetically diverse set of diseases, which has made it difficult to develop a single, highly reproducible, and sensitive assay for measurable residual disease. Here we present the development of a digital targeted RNA-sequencing-based approach designed to overcome these limitations by detecting all newly approved European LeukemiaNet molecular targets for measurable residual disease in acute myeloid leukemia in a single standardized assay. Iterative modifications and novel bioinformatics approaches resulted in a greater than 100-fold increase in performance compared with commercially available targeted RNA-sequencing approaches and a limit of detection as low as one leukemic cell in 100,000 cells measured, which is comparable to quantitative polymerase chain reaction analysis, the current gold standard for the detection of measurable residual disease. This assay, which can be customized and expanded, is the first demonstrated use of high-sensitivity RNA-sequencing for measurable residual disease detection in acute myeloid leukemia and could serve as a broadly applicable standardized tool.

Introduction

Despite the achievement of a "complete remission" following therapy, patients with acute myeloid leukemia (AML) remain at risk of relapse because of the persistence of disease that is not detected by conventional cytomorphological methods.[1] Risk stratification of such patients is possible based on tests designed to detect this residual leukemic burden (termed measurable residual disease; MRD).[1-6] The importance of MRD testing in AML for prognostic risk stratification has become increasingly evident, such that the AML response criteria were substantively updated in 2017 by the introduction of the category of complete remission without MRD.[4]

A variety of technologies, such as real-time quantitative polymerase chain reaction (qPCR) analysis and flow cytometry, focusing on the identification of recurrent molecular abnormalities or leukemia-associated immune phenotypes, have been developed for MRD detection in AML.[7] The mutational and clonal heterogeneity of AML provides a wide variety of targets for molecular MRD tracking,

Correspondence:
hourigan@nih.gov

Received: July 26 2018.
Accepted: August 28, 2018.
Pre-published: August 31, 2018.

doi:10.3324/haematol.2018.203133

Check the online version for the most updated information on this article, online supplements, and information on authorship & disclosures: www.haematologica.org/content/104/2/297

however not all targets can reliably serve as a surrogate of disease burden. For example, mutations found in preleukemic founder clones can persist at significant levels even during complete remission[8-11] and have also been observed in healthy individuals without hematologic malignancies with increasing age.[12-15] To help harmonize AML MRD detection efforts, the European LeukemiaNet (ELN) recently released consensus guidelines for MRD detection in AML, including recommended markers.[6]

Despite advances in the field, it has remained difficult to develop a single MRD detection assay which is highly reproducible and sensitive, has limited operator dependence, and is capable of quantifying multiple targets simultaneously. To overcome these limitations, we developed a multi-gene, targeted RNA-sequencing-based method for the sensitive detection and quantification of MRD in AML, which we present here. This novel assay has a minimal sequencing budget requirement and outperforms a commercially available, conventional myeloid-targeted RNA-sequencing assay. It has a demonstrated limit of detection as low as one leukemic cell in 100,000 cells measured, so its performance equals that of current gold-standard, single-target qPCR MRD assays.

Methods

Cell lines and clinical samples

Leukemia cell lines positive for fusion genes targeted by the AML MRD panel were cultured according to the supplier's guidelines. Peripheral blood samples were collected from a healthy adult donor and from a 46-year old female with monocytic AML who underwent myeloablative matched related donor allogeneic stem cell transplantation (NHLBI protocol # 07-H-0113). Additional samples from patients were collected in local institutional review board-approved biobanking protocols by collaborators. The samples were processed and RNA isolated as described in the *Online Supplementary Appendix*.

Preparation and analysis of targeted RNA-sequencing libraries for acute myeloid leukemia measurable residual disease detection

Targeted RNA-sequencing libraries were prepared and sequenced as described in the *Online Supplementary Appendix*. In short, unique molecular identifier (UMI) assignment and complementary DNA (cDNA) generation were performed on 250 ng of RNA using the SuperScript IV First Strand Synthesis System (Invitrogen, Carlsbad, CA, USA) and a pool of 100 nM of each barcoded (BC) primer. The barcoded cDNA was subjected to eight cycles of amplification with a pool of 100 nM of each limited amplification (LA) primer and 600 nM RS2 primer, followed by final library amplification and sample indexing. Single-end 150 bp sequencing was performed on a Miseq or Hiseq 2500 platform (Illumina) using the QIAseq Read 1 Primer 1 custom primer (Qiagen, Germany). Raw sequencing FASTQ files were processed, aligned, and panel targets called as outlined in the *Online Supplementary Appendix*.

Real-time quantitative polymerase chain reaction

Expression of the "type A" mutation of the nucleophosmin gene (*NPM1* mutA) and the *RUNX1-RUNX1T1* fusion gene was determined for cell dilutions and patient samples by qPCR using the *ipsogen* NPM1 mutA Muta*Quant* kit (Qiagen, cat# 677513) or *ipsogen* RUNX1-RUNX1T1 kit (Qiagen, cat# 675013), respectively, according to the manufacturer's instructions, and the Rotor-Gene Q 5plex HRM (Qiagen). A standard curve for each gene was obtained from the plasmid serial dilutions and used to calculate the copy number for each sample. A patient's sample was considered positive for *NPM1* mutA or *RUNX1-RUNX1T1* if the copy number detected was contained within the standard curve and the water control was negative.

Digital droplet polymerase chain reaction

CBFB-MYH11 type A expression was determined with digital droplet PCR (ddPCR) by converting the established European Against Cancer (EAC) assay for *CBFB-MYH11* type A to the Raindance platform (RainDance Technologies, Billerica, MA, USA), as described in the *Online Supplementary Appendix*.

ArcherDx Myeloid FusionPlex library preparation and analysis

Anchored multiplex PCR-based enrichment RNA-sequencing libraries were generated from 250 ng of RNA using the ArcherDx Myeloid FusionPlex assay for Illumina (ArcherDx, Boulder, CO, USA) and analyzed using Archer Analysis software version 5.1.3, as described in the *Online Supplementary Appendix.*

Statistical analysis

Data were analyzed using Prism statistical software (v.7.0b, GraphPad software, La Jolla, CA, USA).

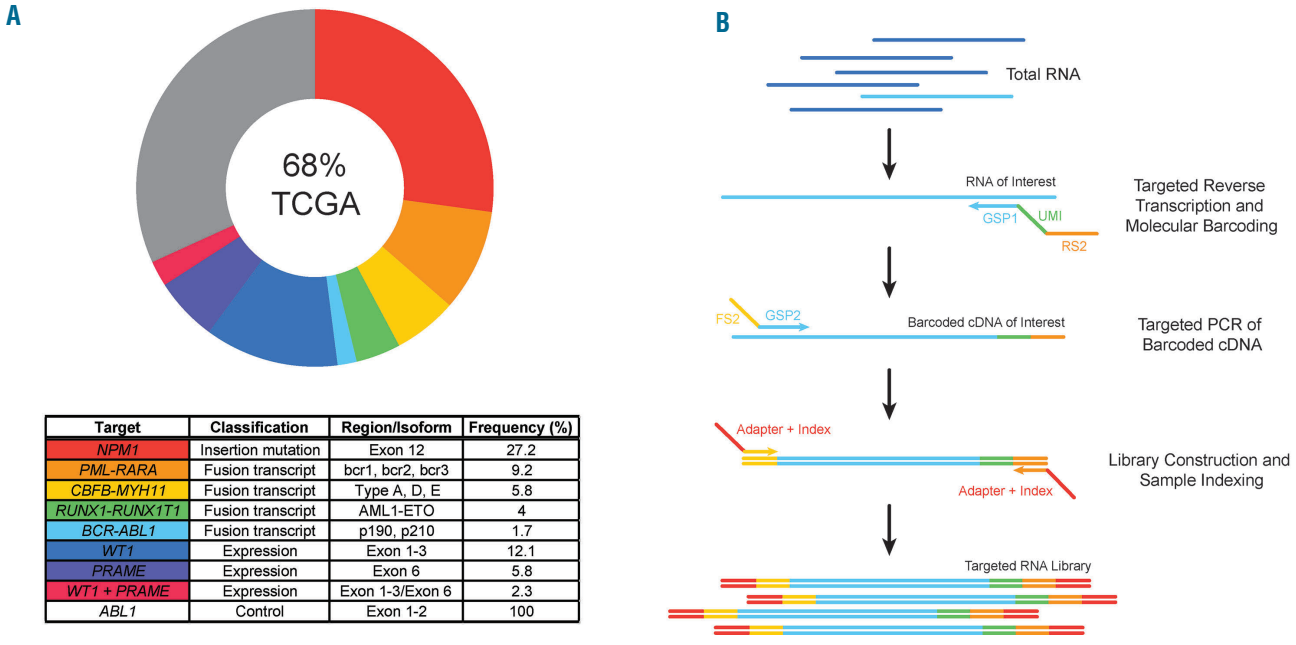
Results

Targeted RNA-sequencing panel target selection for acute myeloid leukemia measurable residual disease detection

Targets for this multi-gene RNA-sequencing panel were chosen from those recommended for AML MRD detection by the ELN[6] and for which standardized qPCR assays for MRD detection have already been well established[5,16-19] (Figure 1A). These targets include: (i) recurrent chimeric fusion transcripts *PML-RARA*, *CBFB-MYH11*, *RUNX1-RUNX1T1*, and *BCR-ABL1*, present in 20.8% of AML patients in The Cancer Genome Atlas (TCGA) dataset;[20] (ii) the recurrent insertion site in exon 12 of *NPM1*, which is found in an additional 27.2% of patients; and lastly (iii) aberrant expression of *WT1* and *PRAME* transcripts, which may be used for MRD detection for up to another 20.2% of patients not covered by the preferred fusion or mutated *NPM1* targets. Based on the TCGA AML cohort, *WT1* and *PRAME* could also serve as a secondary tracking target, for orthogonal validation, in over 80% of those patients who also express fusion or *NPM1* mutant transcripts. Similar to the established qPCR assays, wild-type *ABL1* transcript expression was included as a normalizing control.[16]

Development of the targeted RNA-sequencing method for acute myeloid leukemia measurable residual disease detection

In developing a targeted RNA-sequencing strategy that would be optimal for MRD detection, we took into account several important factors including: (i) the need to optimize target capture efficiency while minimizing the amount of RNA required, as this resource is often limited when dealing with patient specimens; (ii) the incorporation of strategies for the digital quantification of transcripts; (iii) the need to minimize sequencing burden; and (iv) the generation of a workflow that is simple, efficient, and easy to adopt.



Target	Classification	Region/Isoform	Frequency (%)
NPM1	Insertion mutation	Exon 12	27.2
PML-RARA	Fusion transcript	bcr1, bcr2, bcr3	9.2
CBFB-MYH11	Fusion transcript	Type A, D, E	5.8
RUNX1-RUNX1T1	Fusion transcript	AML1-ETO	4
BCR-ABL1	Fusion transcript	p190, p210	1.7
WT1	Expression	Exon 1-3	12.1
PRAME	Expression	Exon 6	5.8
WT1 + PRAME	Expression	Exon 1-3/Exon 6	2.3
ABL1	Control	Exon 1-2	100

Figure 1. The targeted RNA-sequencing assay for measurable residual disease detection can be applied to over two-thirds of patients with acute myeloid leukemia. (A) The acute myeloid leukemia (AML) measurable residual disease (MRD) RNA-sequencing assay comprises eight targets, including insertion mutations, fusion transcripts, and wild-type transcript expression. The frequency of each target was determined using AML cases in The Cancer Genome Atlas (TCGA) for whom clinical information, mutation analysis, and RNA-sequencing data were available (n=173). The expression of insertions and fusion transcripts was evaluated first. For the remaining patients, *WT1* and *PRAME* wild-type transcript overexpression was evaluated and defined as overexpressed if greater than the mean of the entire cohort. (B) The AML MRD RNA-sequencing assay begins with targeted reverse transcription utilizing a pool of primers consisting of a gene-specific region (GSP1), random 12-nucleotide unique molecular identifier (UMI), and conserved sequence region (RS2) per target. The resulting barcoded complementary DNA (cDNA) is subjected to limited amplification using a reverse primer complementary to the RS2 sequence and a pool of forward primers consisting of a gene-specific region (GSP2) and conserved sequence region (FS2) per target. The targeted amplicons are then subjected to amplification, library construction, and sample indexing for Illumina sequencing.

The final assay design addresses these factors by utilizing a pool of target-specific primers containing 12 nucleotide UMIs (*Online Supplementary Table S1*) which capture and individually tag RNA molecules of interest during reverse transcription, followed by targeted PCR of the barcoded cDNA, and library construction (Figure 1B). Unlike most targeted RNA-sequencing approaches, this simplified design reduces protocol steps while maximizing utilization of the RNA input by performing target enrichment during the reverse transcription step, as opposed to after cDNA generation. The concurrent addition of molecular barcodes during this first step also allows for a digital output, increasing the accuracy of transcript quantification. Additionally, the amplicon-based enrichment design for fusion detection limits the sequencing requirements, since only fusion transcripts and not wild-type transcripts, will be amplified. Finally, with a total of only three steps, the hands-on time is minimized, allowing for the entire protocol to be completed in less than a day.

Validation of assay performance and limit of detection

To assess the sensitivity and dynamic range of the AML MRD RNA-sequencing panel, cell lines expressing fusion transcripts or patient cells positive for the *NPM1* mutA insertion mutation (94% blasts) were serially diluted (1:10 to 1:100,000) into healthy adult donor peripheral blood mononuclear cells and RNA was isolated. A total of 250 ng of RNA from each dilution was subjected to targeted RNA-sequencing library preparation and sequencing. Sequencing files were processed by extracting the UMI from each read, alignment to the human genome, and clustering of sequences which correspond to the intended panel targets (*Online Supplementary Figure S1*). Targets were quantified by the number of unique UMIs, with a library-specific UMI cutoff value established to eliminate background due to sequencing errors (*Online Supplementary Figure S2*).

The expression of all assay targets exhibited significant correlation (linear regression $r^2 \geq 0.97$) and leukemic cells could be detected at a level as low as one leukemic cell in 100,000 healthy donor cells (Figure 2). For the fusion and mutated *NPM1* transcripts, detection sensitivity was between 1:10,000 to 1:100,000. Many of the cell lines expressing fusion transcripts also displayed aberrant *WT1* and/or *PRAME* transcript expression, which was also highly correlated and for which the assay showed varying degrees of sensitivity, ranging from 1:1,000 to 1:100,000. Detection of all targets was highly reproducible across replicates.

Determination of assay sequencing requirements

Variations in sequencing depth depending on the leukemic burden present in a sample and baseline error rates between sequencing platforms are important elements which can affect assay performance and reliability. Additionally, sequencing read requirements play an important role in the feasibility of assay adaptation into practice.

To address these factors, we first examined the impact of sequencing depth and platform on assay detection met-

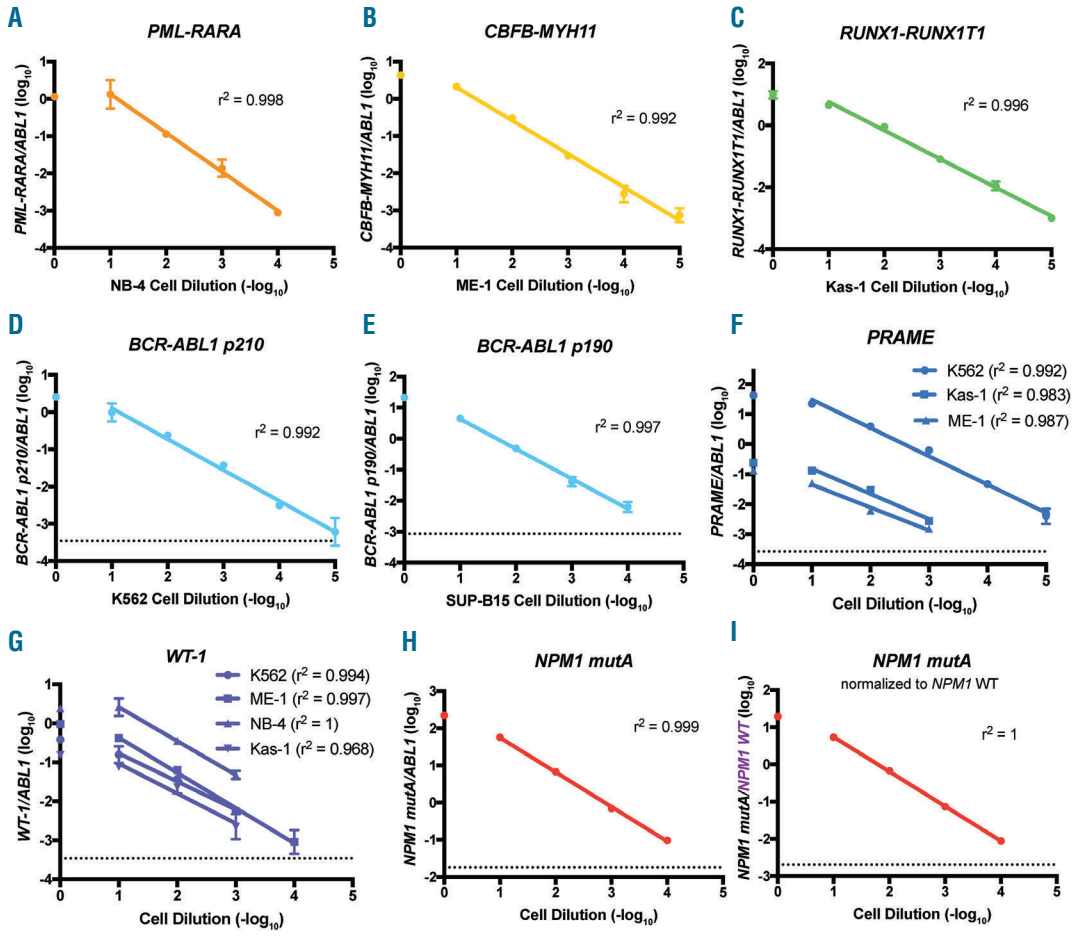


Figure 2. Evaluation of assay performance by target. Cell lines or patient cells positive for assay targets (A) *PML-RARA*, (B) *CBFB-MYH11*, (C) *RUNX1-RUNX1T1*, (D) *BCR-ABL1* p210 isoform, (E) *BCR-ABL1* p190 isoform, (F) *PRAME*, (G) *WT-1*, and (H) *NPM1* mutA insertion mutation were serially diluted in peripheral blood mononuclear cells from a normal healthy individual and RNA was isolated. The resulting RNA was subjected to acute myeloid leukemia (AML) measurable residual disease (MRD) RNA-sequencing library preparation and sequencing. The data are presented as a ratio of the number of target copies/*ABL1* copies (log_{10}) on the y-axis and the cell dilution ($-log_{10}$) on the x-axis. Each data point represents two replicates, except *NPM1* mutA (n=1), and error bars represent the standard deviation. A linear regression was calculated for each dilution series (excluding the 100% leukemia samples and any samples for which the target was not detected) and the coefficient of determination (r^2) indicated in the graph area. *NPM1* mutA data were normalized to either (H) *ABL1* copies or (I) wild-type *NPM1* copies, both of which exhibited similar results and a strong correlation. If background values were detected in the normal healthy individual, an average of the two libraries sequenced is represented as a dashed line.

rics. A set of libraries generated on a serial dilution of K562 cells (positive for *BCR-ABL1* p210 fusion transcript, *WT1*, and *PRAME* expression) in normal healthy peripheral blood mononuclear cells were subjected to sequencing on both the Illumina MiSeq and HiSeq 2500 platforms at an average depth of 3.6 million and 38.6 million reads, respectively. Regardless of platform or sequencing depth, the ratio of each target relative to *ABL1* expression was highly reproducible and significantly correlated (Pearson correlation r = 1.00) (*Online Supplementary Figure S3*), confirming the robustness of the bioinformatics pipeline across different depths and sequencers.

Next, we determined the minimum number of reads necessary to retain assay sensitivity across the dynamic range. To do so, libraries obtained from the serial dilution of K562 cells or *NPM1* mutA patient cells in healthy donor peripheral blood mononuclear cells were down-sampled stepwise and separately subjected to the bioinformatics pipeline (*Online Supplementary Figure S4*). We determined that as few as 1 million reads are needed for the detection of fusion transcripts and 3 million reads for the detection of *NPM1* insertion mutations (due to high wild-type *NPM1* transcript expression). Thus, with a minimal sequencing read requirement equivalent to a micro or nano flow cell on the MiSeq platform, this assay is able to retain up to a 5-log dynamic range.

Comparison of the performance of the study assay with that of current gold-standard, single-target methodologies

Next, we compared the performance of the AML MRD RNA-sequencing panel to that of the current gold-standard qPCR assays for single-mutation MRD detection. The same RNA from the serial dilution of positive control cell lines used to generate targeted RNA-sequencing libraries was assessed by qPCR or ddPCR for *NPM1* mutA (Figure 3A), *RUNX1-RUNX1T1* (Figure 3B), or *CBFB-MYH11* type A (Figure 3C) mutations. Across all three mutations, a comparable limit of detection was observed for both qPCR and ddPCR (right graph) compared to tar-

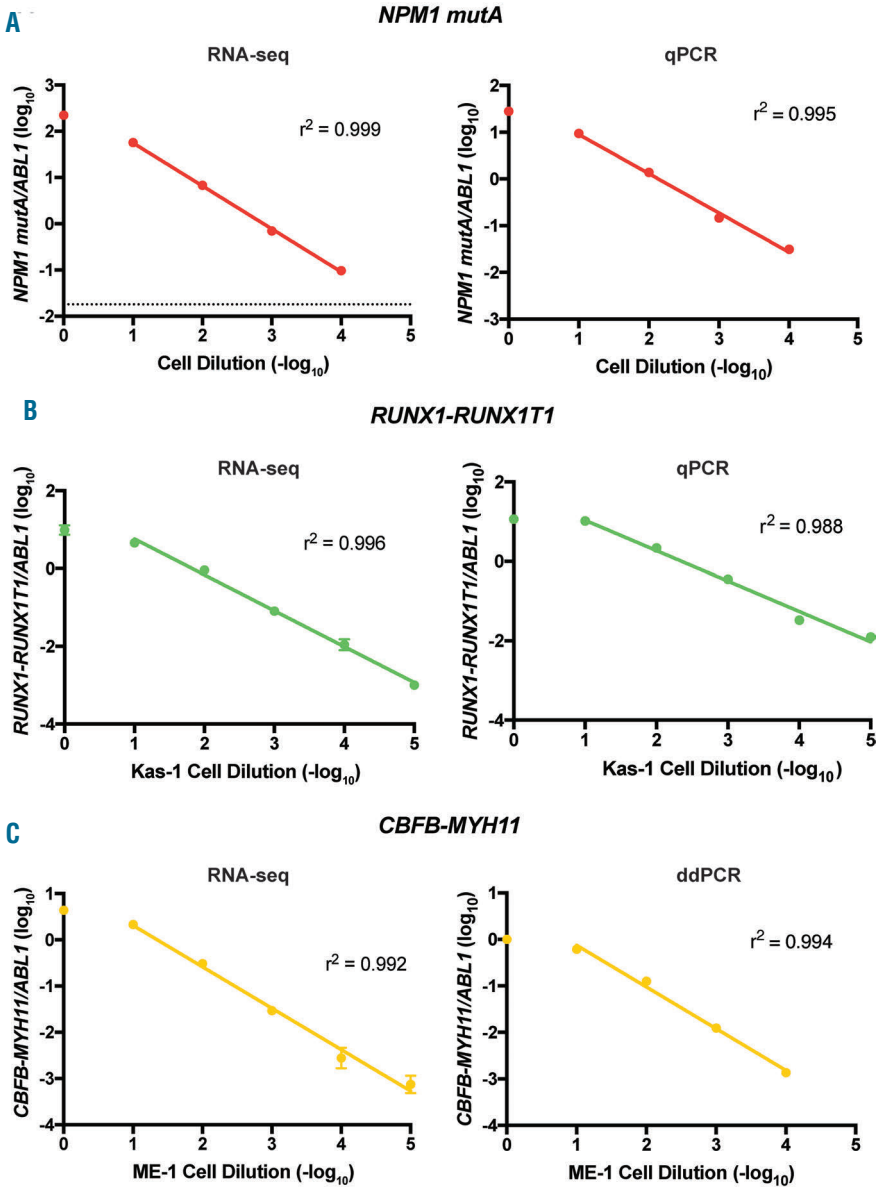


Figure 3. Comparison of the performance of the acute myeloid leukemia measurable residual disease targeted RNA-sequencing assay with that of current gold-standard, single-target methodologies. Gold-standard quantitative real-time polymerase chain reaction (qPCR) or digital droplet polymerase chain reaction (ddPCR) (right) techniques performed on the serial dilution samples for (A) *NPM1* mutA, (B) *RUNX1-RUNX1T1*, and (C) *CBFB-MYH11* exhibited comparable performance to that of the acute myeloid leukemia (AML) measurable residual disease (MRD) targeted RNA-sequencing (RNA-seq) technique (left). The data are presented as a ratio of the number of target copies/*ABL1* copies (log_{10}) on the y-axis and the cell dilution ($-log_{10}$) on the x-axis. A linear regression was calculated for each dilution series (excluding the 100% leukemia samples and any samples for which the target was not detected) and the coefficient of determination (r^2) indicated in the graph area.

geted RNA-sequencing (left graph), thus confirming that the AML MRD RNA-sequencing assay on its own serves as an adequate replacement for current, single-target MRD detection assays.

Comparison of the performance of the study assay with that of a commercially available myeloid targeted RNA-sequencing assay

The Myeloid FusionPlex assay from ArcherDx is a commercially available targeted RNA-sequenced assay which also utilizes molecular barcodes and is designed to detect and identify fusions, point mutations, and expression level changes in a panel of 84 genes associated with malignancies of myeloid origin. In contrast to the assay presented here, the FusionPlex technology utilizes an anchored multiplex-PCR enrichment (AMP-E) approach, which is intended for fusion discovery in a bulk tumor setting.

The performance of the Myeloid FusionPlex assay for the detection of MRD was assessed for *NPM1* mutA and *CBFB-MYH11* mutations using 250 ng of the same RNA as used in the AML MRD RNA-sequencing assay. We confirmed that the Myeloid FusionPlex assay can detect these mutations, but at varying frequencies (*Online Supplementary Table S2*). Compared to the Myeloid FusionPlex assay, our AML MRD RNA-sequencing assay exhibits a 1,000-fold increased limit of detection for *NPM1* mutA and 100-fold increased limit of detection for *CBFB-MYH11* (Figure 4), indicating that our assay outperforms a commercially available targeted RNA-sequencing assay for MRD-level detection in AML.

Validation of assay performance in patient specimens

Finally, we examined the performance of the AML MRD RNA-sequencing assay applied to patient samples. RNA was isolated from diagnostic peripheral blood or bone marrow samples of patients clinically annotated to harbor mutations included in our assay and assessed the samples by AML MRD targeted RNA-sequencing. In all cases, the annotated mutation was detectable using our assay (Table 1). For *NPM1* mutA and *RUNX1-RUNX1T1-*

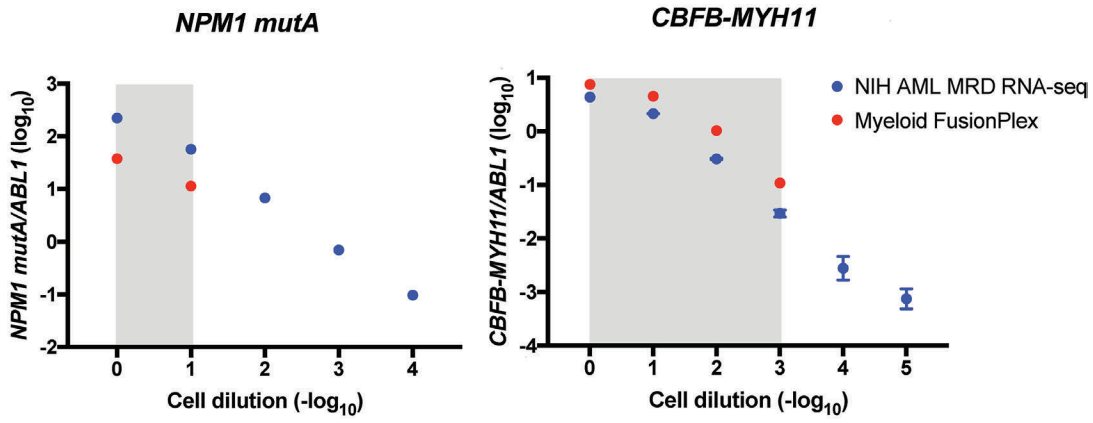


Figure 4. Comparison of the performance of the acute myeloid leukemia measurable residual disease targeted RNA-sequencing assay to that of the commercially available Myeloid FusionPlex assay. *NPM1* mutA and *CBFB-MYH11* serial dilution samples were analyzed using the ArcherDx Myeloid FusionPlex assay. The ratio of target copies/*ABL1* copies (log_{10}) is plotted on the y-axis and the cell dilution ($-log_{10}$) on the x-axis for both the National Institutes of Health acute myeloid leukemia measurable residual disease targeted RNA-sequencing assay (NIH AML MRD RNA-Seq) (blue) and the Myeloid FusionPlex assay (red). The shaded areas indicate the cell dilution frequencies in which the mutation was detectable by both methodologies.

Table 1. Validation of mutation detection in diagnostic patient samples by acute myeloid leukemia measurable residual disease targeted RNA-sequencing.

Patient	Sample type	Annotated mutation	Mutation detected by RNA-seq	Gene expression detected by RNA-seq
A	PB	*NPM1* mut	*NPM1* mutA	*WT1, PRAME*
B	BM	*NPM1* mut	*NPM1* mutA	*WT1*
C	PB	*NPM1* mut	*NPM1* mutA	*WT1, PRAME*
D	PB	*RUNX1-RUNX1T1*	*RUNX1-RUNX1T1*	*WT1, PRAME*
E	BM	*PML-RARA*	*PML-RARA bcr1*	*WT1, PRAME*
F	BM	*PML-RARA*	*PML-RARA bcr1*	*WT1, PRAME*
G	BM	*PML-RARA*	*PML-RARA bcr3*	*WT1*
H	BM	*BCR-ABL1 p210*	*BCR-ABL1 p210*	*WT1, PRAME*
I	BM	*BCR-ABL1 p210*	*BCR-ABL1 p210*	*WT1*
J	BM	*CBFB-MYH11*	*CBFB-MYH11 type D*	*WT1, PRAME*

RNA-seq: RNA-sequencing; PB: peripheral blood; BM: bone marrow.

positive patient samples, we also confirmed the detectability of these mutations by qPCR (*Online Supplementary Table S3*). Furthermore, using our AML MRD RNA-sequencing assay we were able to specifically define the isoform present in each sample, including the rarer *CBFB-MYH11* type D isoform (*Online Supplementary Figure S5*), and to identify the overexpression of *WT1* or *PRAME* as potential secondary tracking targets (Table 1).

To test the ability of this assay to track disease burden in a patient and detect relapse relative to standard clinical assessments, we performed AML MRD targeted RNA-sequencing on RNA isolated from peripheral blood mononuclear cells of an *NPM1* mutA-positive AML patient at various time-points during treatment. *NPM1* mutA transcripts were detected at the time of pathological cytomorphological complete remission prior to allogenic stem cell transplantation. While *NPM1* mutA transcript expression was not detected immediately after transplantation, an increase in *NPM1* mutA transcript was detected at two subsequent time-points, up to 3 months prior to clinical relapse (Figure 5).

Collectively, these findings confirm that the AML MRD targeted RNA-sequencing assay can identify all of the included targets in both peripheral blood and bone marrow from patients harboring these mutations and is capable of detecting MRD in patient samples during clinical complete remission.

Discussion

Here we present the development of a targeted RNA-sequencing method for the detection and measurement of MRD in AML. This assay, which can cover over two-thirds of patients in a single standardized assay (Figure 1), is highly specific, sensitive, and resilient to variations in sequencing depth and platform used during data collection. MRD status is now integrated into response criteria for AML,[4] with ELN consensus guidelines for measurement now available.[6] The assay presented here detects all of the molecular targets included in these guidelines.

To date, molecular MRD detection in AML has primarily focused on single target qPCR assays.[18] The use of next-generation sequencing for AML MRD detection has begun to emerge but has primarily focused on DNA as the starting material.[21-27] Several features of the AML MRD RNA-sequencing assay presented here in theory make it ideal for MRD detection, including: (i) the use of next-genera-

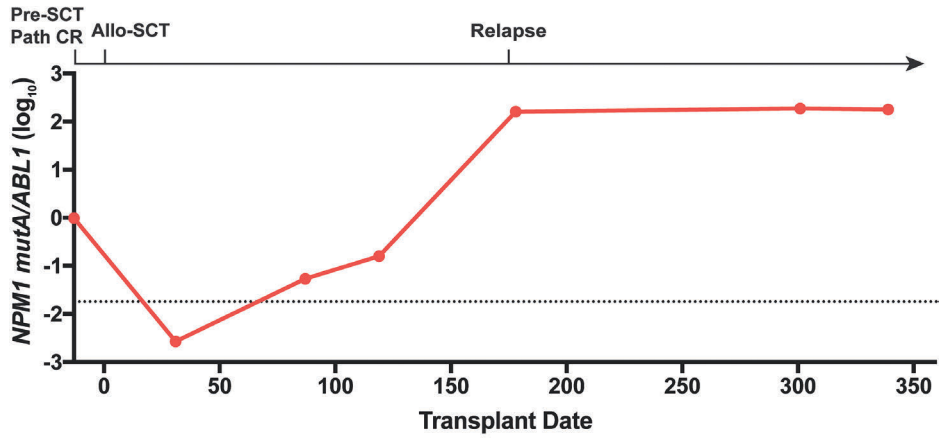


Figure 5. Evaluation of assay performance in serial samples from a patient. RNA isolated from peripheral blood mononuclear cells of a leukemia patient positive for the *NPM1* mutA insertion mutation at serial time-points was subjected to acute myeloid leukemia (AML) measurable residual disease (MRD) RNA-sequencing library preparation and sequencing. The data are presented as the ratio of *NPM1* mutA copies/*ABL1* copies (log_{10}) on the y-axis and the date of sample collection relative to allogeneic stem cell transplantation (in days) on the x-axis. Important clinical time-points are depicted above the graph. Clinical relapse occurred on day 175 while MRD was detectable by targeted RNA-sequencing on days 87 and 119. The background ratio of *NPM1* mutA copies/*ABL1* copies (log_{10}) detected in the normal donor is represented as a dashed line. Pre-SCT: before stem cell transplantation; Path CR: pathological cytomorphological complete remission; Allo-SCT: allogeneic stem cell transplantation.

tion sequencing allows for both target multiplexing and flexibility in identifying mutations that could vary between patients (variations in fusion breakpoints or insertion sequences, etc.); (ii) use of primers targeting recurrent AML abnormalities greatly increases the sensitivity of the assay over bulk RNA- or DNA-sequencing; (iii) RNA as the starting material allows for the simultaneous examination of mutations/fusions and changes in transcript expression; (iv) RNA increases the limit of detection over that provided by DNA if the transcript is expressed at a level greater than the genomic equivalent per cell; (v) UMIs allow for the absolute quantification of target levels; and (vi) targeted primers during reverse transcription increase the limit of detection and allow for efficient use of the starting material. However, it is conceivable that this RNA-sequencing assay could be complemented in the future by the use of a DNA-based AML MRD next-generation sequencing approach (tracking, for example, somatic mutations in *TP53*, *IDH1*, *IDH2*, *FLT3*, etc.) which together would allow coverage of almost all cases of AML.

Utilizing cell lines and patient samples expressing the targets included in our assay, we demonstrated that the AML MRD RNA-sequencing assay has a sensitivity for residual disease detection down to as low as 1 in 100,000 cells (Figure 2). Importantly, this detection limit is well below the threshold of 1 in 1,000 cells currently suggested by the ELN MRD consortium[6] and is comparable to that of the gold standard single-target molecular techniques (Figure 3). Furthermore, due to the MRD-focused design of our assay, it can achieve up to a 1,000-fold greater sensitivity than that of the most similar targeted RNA-sequencing assay for myeloid malignancies available on the commercial market (Figure 4). However, since RNA expression levels can vary from one patient to another, it is important to note that sensitivity levels can vary and a detection limit of 1 in 100,000 cells may not always be attainable. This is a general difficulty affecting all RNA-based methods of MRD detection, including qPCR, and is reflected in the recommendation that molecular relapse be determined by the progression of trends across multiple time-points in an individual patient rather than by a single landmark assessment.[6]

Several important considerations for the clinical utility of an MRD assay include sample input requirements, cost, time, and ease of assay adaptation. Each of these factors were considered in the assay design. The use of targeted primers and the addition of UMIs during the reverse transcription step allow for maximal usage of the starting material while simplifying the workflow (Figure 1B), which can be completed in a single day and is easy to adopt and/or automate. Additionally, with a minimal sequencing requirement of only 1-3 million reads (*Online Supplementary Figure S4*), a single patient sample can be assessed for immediate results or multiplexed on larger scale runs to minimize costs (*Online Supplementary Table S4*). Importantly, this multiplex assay and analytic workflow is both flexible and expandable. The design allows for all targets to be screened at diagnosis, with the ability to tailor analysis to a subset of patient-specific targets at later time-points to further minimize sequencing costs. Additional targets can easily be added to the assay design, potentially allowing for the detection of other AML MRD markers, chimerism,[28,29] and HLA loss[30] for those poor-risk AML patients who are not optimally covered by this assay but who will often undergo allogeneic stem cell transplantation.

While we confirmed the feasibility of use of this test in patient blood and bone marrow (Table 1, Figure 5), future work is needed to test the utility of this technique in large cohorts of patients and to determine the specific impact of MRD detection on AML patient outcomes in this setting. Overall, we believe that this UMI-based RNA-sequencing assay provides a high-throughput, reproducible, and broadly applicable tool for standardized detection of residual disease in patients with AML.

Acknowledgments

This work was supported by the Intramural Research Program of the National Heart, Lung, and Blood Institute (NHLBI) of the National Institutes of Health (NIH). GWR is

supported by the Medical Research Scholars Program, a public-private partnership supported jointly by the NIH, and by generous contributions to the Foundation for the NIH by the Doris Duke Charitable Foundation (Grant #2014194). The authors thank Samuel J. Rulli and Melanie Hussong of Qiagen, Inc. for their insight and advice during the assay development. The authors also thank Yuesheng Li, Yan Luo, and Patrick Burr of the NHLBI DNA Sequencing and Genomics Core for providing next-generation sequencing services and troubleshooting issues that arose during data collection. Finally, the authors also thank A. John Barrett of the National Heart, Lung and Blood Institute, Paul Liu of the NHLBI, and Gabriel Ghiaur of the Sidney Kimmel Cancer Center at Johns Hopkins for kindly providing patient samples or cell lines.

References

1. Hourigan CS, Gale RP, Gormley NJ, Ossenkoppele GJ, Walter RB. Measurable residual disease testing in acute myeloid leukaemia. Leukemia. 2017;31(7):1482-1490.
2. Araki D, Wood BL, Othus M, et al. Allogeneic hematopoietic cell transplantation for acute myeloid leukemia: time to move toward a minimal residual disease-based definition of complete remission? J Clin Oncol. 2016;34(4):329-336.
3. Buckley SA, Wood BL, Othus M, et al. Minimal residual disease prior to allogeneic hematopoietic cell transplantation in acute myeloid leukemia: a meta-analysis. Haematologica. 2017;102(5):865-873.
4. Dohner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood. 2017;129(4):424-447.
5. Goswami M, McGowan KS, Lu K, et al. A multigene array for measurable residual disease detection in AML patients undergoing SCT. Bone Marrow Transplant. 2015;50(5):642-651.
6. Schuurhuis GJ, Heuser M, Freeman S, et al. Minimal/measurable residual disease in AML: a consensus document from the European LeukemiaNet MRD Working Party. Blood. 2018;131(12):1275-1291.
7. Zhou Y, Wood BL. Methods of detection of measurable residual disease in AML. Curr Hematol Malig Rep. 2017;12(6):557-567.
8. Grimwade D, Ivey A, Huntly BJ. Molecular landscape of acute myeloid leukemia in younger adults and its clinical relevance. Blood. 2016;127(1):29-41.
9. Corces-Zimmerman MR, Hong WJ, Weissman IL, Medeiros BC, Majeti R. Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. Proc Natl Acad Sci USA. 2014;111(7):2548-2553.
10. Jan M, Snyder TM, Corces-Zimmerman MR, et al. Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. Sci Transl Med. 2012;4(149):149ra118.
11. Shlush LI, Zandi S, Mitchell A, et al. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. Nature. 2014;506(7488):328-333.
12. Genovese G, Kahler AK, Handsaker RE, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. N Engl J Med. 2014;371(26):2477-2487.
13. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. N Engl J Med. 2014;371(26):2488-2498.
14. Steensma DP, Bejar R, Jaiswal S, et al. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. Blood. 2015;126(1):9-16.
15. Young AL, Challen GA, Birmann BM, Druley TE. Clonal haematopoiesis harbouring AML-associated mutations is ubiquitous in healthy adults. Nat Commun. 2016;7: 12484.
16. Beillard E, Pallisgaard N, van der Velden VH, et al. Evaluation of candidate control genes for diagnosis and residual disease detection in leukemic patients using 'real-time' quantitative reverse-transcriptase polymerase chain reaction (RQ-PCR) - a Europe Against Cancer program. Leukemia. 2003;17(12):2474-2486.
17. Cilloni D, Renneville A, Hermitte F, et al. Real-time quantitative polymerase chain reaction detection of minimal residual disease by standardized WT1 assay to enhance risk stratification in acute myeloid leukemia: a European LeukemiaNet study. J Clin Oncol. 2009;27(31):5195-5201.
18. Gabert J, Beillard E, van der Velden VH, et al. Standardization and quality control studies of 'real-time' quantitative reverse transcriptase polymerase chain reaction of fusion gene transcripts for residual disease detection in leukemia - a Europe Against Cancer program. Leukemia. 2003;17(12):2318-2357.
19. Ivey A, Hills RK, Simpson MA, et al. Assessment of minimal residual disease in standard-risk AML. N Engl J Med. 2016;374(5):422-433.
20. Cancer Genome Atlas Research N, Ley TJ, Miller C, et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med. 2013;368(22):2059-2074.
21. Gaksch L, Kashofer K, Heitzer E, et al. Residual disease detection using targeted parallel sequencing predicts relapse in cytogenetically normal acute myeloid leukemia. Am J Hematol. 2018;93(1):23-30.
22. Klco JM, Miller CA, Griffith M, et al. Association between mutation clearance after induction therapy and outcomes in acute myeloid leukemia. JAMA. 2015;314(8):811-822.
23. Levis MJ, Perl AE, Altman JK, et al. A next-generation sequencing-based assay for minimal residual disease assessment in AML patients with FLT3-ITD mutations. Blood Adv. 2018;2(8):825-831.
24. Malmberg EB, Stahlman S, Rehammar A, et al. Patient-tailored analysis of minimal residual disease in acute myeloid leukemia using next-generation sequencing. Eur J Haematol. 2017;98(1):26-37.
25. Morita K, Kantarjian HM, Wang F, et al. Clearance of somatic mutations at remission and the risk of relapse in acute myeloid leukemia. J Clin Oncol. 2018;36(18):1788-1797.
26. Roloff GW, Lai C, Hourigan CS, Dillon LW. Technical advances in the measurement of residual disease in acute myeloid leukemia. J Clin Med. 2017;6(9).
27. Rothenberg-Thurley M, Amler S, Goerlich D, et al. Persistence of pre-leukemic clones during first remission and risk of relapse in acute myeloid leukemia. Leukemia. 2018;32(7):1598-1608.
28. Aloisio M, Licastro D, Caenazzo L, et al. A technical application of quantitative next generation sequencing for chimerism evaluation. Mol Med Rep. 2016;14(4):2967-2974.
29. Bornhauser M, Oelschlaegel U, Platzbecker U, et al. Monitoring of donor chimerism in sorted CD34+ peripheral blood cells allows the sensitive detection of imminent relapse after allogeneic stem cell transplantation. Haematologica. 2009;94(11):1613-1617.
30. Ahci M, Toffalori C, Bouwmans E, et al. A new tool for rapid and reliable diagnosis of HLA loss relapses after HSCT. Blood. 2017;130(10):1270-1273.