haematologica
Journal of the European Hematology Association

Novel deep targeted sequencing method for minimal residual disease monitoring in acute myeloid leukemia

by Esther Onecha, Maria Linares, Inmaculada Rapado, Yanira Ruiz-Heredia, Pilar Martinez-Sanchez, Teresa Cedena, Marta Pratcorona, Jaime Perez Oteyza, Pilar Herrera, Eva Barragan, Pau Montesinos, Jose Antonio Garcia Vela, Elena Magro, Eduardo Anguita, Angela Figuera, Rosalia Riaza, Pilar Martinez-Barranco, Beatriz Sanchez-Vega, Josep Nomdedeu, Miguel Gallardo, Joaquin Martinez-Lopez, and Rosa Ayala

Haematologica 2018 [Epub ahead of print]

Citation: Esther Onecha, Maria Linares, Inmaculada Rapado, Yanira Ruiz-Heredia, Pilar Martinez-Sanchez, Teresa Cedena, Marta Pratcorona, Jaime Perez Oteyza, Pilar Herrera, Eva Barragan, Pau Montesinos, Jose Antonio Garcia Vela, Elena Magro, Eduardo Anguita, Angela Figuera, Rosalia Riaza, Pilar Martinez-Barranco, Beatriz Sanchez-Vega, Josep Nomdedeu, Miguel Gallardo, Joaquin Martinez-Lopez, and Rosa Ayala. Novel deep targeted sequencing method for minimal residual disease monitoring in acute myeloid leukemia.
Haematologica. 2018; 103:xxx
doi:10.3324/haematol.2018.194712

Novel deep targeted sequencing method for minimal residual disease monitoring in acute myeloid leukemia

AUTHORS

Esther Onecha,[1,2] Maria Linares,[1,2] Inmaculada Rapado,[1,2,3] Yanira Ruiz-Heredia,[1,2] Pilar Martinez-Sanchez,[1] Teresa Cedena,[1,2,3,4] Marta Pratcorona,[5] Jaime Perez Oteyza,[6] Pilar Herrera,[7] Eva Barragan,[4,8] Pau Montesinos,[4,8] Jose Antonio Garcia Vela,[9] Elena Magro,[10] Eduardo Anguita,[11] Angela Figuera,[12] Rosalia Riaza,[13] Pilar Martinez-Barranco,[14] Beatriz Sanchez-Vega,[1,2] Josep Nomdedeu,[5] Miguel Gallardo,[2]* Joaquin Martinez-Lopez,[1,2,3,4]* and Rosa Ayala[1,2,3,4]*

*M.G., J.M-L. and R.A. contributed equally to this study

RUNNING HEAD

Deep sequencing method for MRD monitoring in AML

AUTHORS AND AFFILIATIONS

[1]Hematology Department, Hospital Universitario 12 de Octubre, Madrid, Spain; [2]Hematological Malignancies Clinical Research Unit, CNIO, Madrid, Spain; [3]CIBERONC, Spain; [4]Complutense University, Madrid, Spain; [5]Hematology Department, Hospital Santa Creu i Sant Pau, Barcelona, Spain; [6]Hematology Department, Hospital Universitario Sanchinarro, Madrid, Spain; [7]Hematology Department, Hospital Universitario Ramon y Cajal, Madrid, Spain; [8]Hematology Department, Hospital Universitario La Fe, Valencia, Spain; [9]Department of Hematology, Hospital Universitario de Getafe, Madrid, Spain; [10]Hematology Department, Hospital Universitario Principe de Asturias, Madrid, Spain; [11]Hematology Department, Hospital Clínico San Carlos, IdISSC, UCM, Madrid, Spain; [12]Hematology Department, Hospital Universitario de la Princesa, Madrid, Spain; [13]Hematology Department, Hospital Universitario Severo Ochoa, Madrid, Spain; [14]Hematology Department, Hospital Universitario Fundación Alcorcón, Madrid, Spain.

CORRESPONDING AUTHOR
Rosa Ayala, MD, PhD
Hematology Department
Hospital Universitario 12 de Octubre
Edificifio de Atención Ambulatoria, Planta 6 Bloque D
Avenida de Córdoba s/n, 28041, Madrid, Spain
Email: rosam.ayala@salud.madrid.org
Phone: +34 91 779 27 88
Fax: +34 91 390 80 00

SCIENTIFIC CATEGORY
Acute myeloid leukemia

KEYWORDS
Acute myeloid leukemia, sequencing, minimal residual disease

Word count: 2931

Abstract word count: 237

Figures and tables: 4 figures, 2 tables

References: 37

Funding
CRIS against Cancer foundation and ISCIII

Abstract

A high proportion of patients with acute myeloid leukemia who achieve minimal residual disease negative status ultimately relapse because a fraction of pathological clones remains undetected by standard methods. We designed and validated a high-throughput sequencing method for minimal residual disease assessment of cell clonotypes with mutations of *NPM1*, *IDH1/2* and/or *FLT3*-SNV. For clinical validation, 106 follow-up samples from 63 patients in complete remission were studied by sequencing, evaluating the level of mutations detected at diagnosis. The predictive value of minimal residual disease status by sequencing, multiparameter flow cytometry, or quantitative PCR was determined by survival analysis. The method achieved a sensitivity of 10^{-4} for single nucleotide variant and 10^{-5} for insertions/deletions and could be used in acute myeloid leukemia patients who carry any mutation (86% in our diagnosis data set). Sequencing–determined minimal residual disease positive status was associated with lower disease-free survival (hazard ratio 3.4, p=0.005) and lower overall survival (hazard ratio 4.2, p<0.001). Multivariate analysis showed that minimal residual disease positive status by sequencing was an independent factor associated with risk of death (hazard ratio 4.54, p =0.005) and the only independent factor conferring risk of relapse (hazard ratio 3.76, p =0.012). This sequencing-based method simplifies and standardizes minimal residual disease evaluation, with high applicability in acute myeloid leukemia. It also improves upon flow cytometry and quantitative PCR to predict acute myeloid leukemia outcome and could be incorporated in clinical settings and clinical trials.

Introduction

Cytogenetic and molecular alterations at diagnosis and response to treatment are the most useful criteria to predict relative risk of relapse in acute myeloid leukemia (AML), and to guide the choice between chemotherapy and hematopoietic stem cell transplantation (HSCT) in first complete remission (CR).[(1)] The definition of CR for AML includes criteria for the identification of patients with poor prognosis using cytomorphological methods.[(2)] But, these studies do not have a good predictive value because most of the CR cases relapse within 3 years of diagnosis.[(3)]

Assessment of minimal residual disease (MRD) is critical in monitoring patients in morphological remission, to inform decisions about further therapy.[(1)] Indeed, several studies have reported MRD status as a stronger predictor of relapse, because patients who are MRD negative have a better prognosis than those who are MRD positive.[(4, 5)] In support of this, recent non-randomized studies from prospective multicenter trials suggested better outcomes when leukemia therapy was selected based on the results of MRD assessment. [(6-8)]

AML is, nevertheless, a biologically complex and heterogeneous disease, which makes MRD testing challenging when compared with other hematological neoplasms such as acute lymphoblastic leukemia or multiple myeloma. The detection of very low levels of MRD by conventional methods such as quantitative (q)-PCR or multiparameter flow cytometry (MFC) provides powerful independent prognostic information. Unfortunately, as described for cytomorphological CR, many patients who achieve MRD negative status relapse as a result of the progression of undetected leukemic cells. The most common method for MRD detection is MFC, with intermediate applicability (70–80%) and limited sensitivity.[9, 10] However, there is no consensus on multi-antibody panels with regards to inter-laboratory performance, and the technique requires a high level of expertise. The other principal MRD monitoring method, qPCR, has good sensitivity (10^{-4}–10^{-6}), but its applicability is limited in up to 40% of patients who present molecular alterations (*RUNX1-RUNX1T1, CBFβ-MYH11* or *NPM1)* at diagnosis.[11]

For the above reasons, new methods with higher sensitivity, specificity, applicability and performance are needed for MRD assessment in AML. Against this background, next-generation sequencing **(**NGS) and digital PCR (dPCR) have recently emerged as potentially promising platforms to assess MRD.[12] Here, we optimized and clinically validated a new deep targeted NGS-based method, supported with dPCR technical validation, for MRD detection and quantification (both small insertion/deletions [InDels] and single nucleotide variants [SNVs]) in AML patients, in an attempt to improve and/or complement the current MRD evaluation techniques, and to establish its potential as a predictor of patient outcome.

Methods

More detailed information can be found in the *Online Supplementary data (1–6).*

Patients and samples

One hundred and ninety patients with *de novo* or secondary non–M3 AML were included in mutational profile screening at diagnosis. We performed a new selection for retrospective MRD assessment using the following criteria: presence of the *NPM1* type A mutation, or SNVs in *FLT3*, *IDH1* and/or *IDH2* at diagnosis, and availability of at least one follow-up genomic (g)-DNA sample.

The MRD approach included 51 (48%) follow-up samples taken at post-induction, and 55 (52%) at post-consolidation time, corresponding to 63 patients diagnosed between 2006 and 2016 (for selection criteria see *Online Supplementary 6* and *Supplementary Table S1*). Patients were treated according to PETHEMA (Programa Español de Tratamientos en Hematología) or CETLAM (Grupo cooperativo de Estudio y Tratamiento de Leucemias Agudas y Mielodisplasias) protocols. The study was conducted according to the

Spanish law 14/2007 of biomedical research, and was approved by the Research Ethics Board of each participating institution. All patients provided informed consent. The main clinical characteristics of patients are summarized in Table 1. All patients achieved CR by cytomorphological criteria after induction therapy (<5% of bone marrow blasts).

To construct calibration curves, commercial (Horizon Discovery, UK) reference standard gDNA was used for somatic SNVs in *IDH1* (R132C) and *IDH2* (R172K). As a further source of gDNA, we used the OCI-AML3 cell line (ACC 582, DSMZ, Germany) with the *NPM1* type A mutation (c.863_864insCCTG) to examine InDels. As OCI-AML3 cells also present a SNV in *DNMT3A* (R882C), this was included only for technical optimization.

Deep targeted sequencing workflow

The sequencing workflow included one first study at diagnosis and a second study at follow-up. Mutational profile screening at diagnosis was done with a custom NGS myeloid panel of 32 genes frequently mutated in myeloid diseases,[13] (*Online Supplementary Table S2*) and *NPM1* analysis was carried out with qPCR.[14]

The specific mutations detected at diagnosis were studied at follow-up. A variety of experimental steps were first tested to define optimal conditions (*Online Supplementary 1*). We established an optimal protocol (Figure 1) that included DNA amplification, library preparation and sequencing as experimental steps (*Online Supplementary 2*).

Libraries were sequenced on the Ion Proton System platform (Life Technologies, Thermo Fisher Scientific Inc.) with an estimated depth ≥ 1,000,000 of reads, generating .fastq files. These files were analyzed using a custom bioinformatic pipeline; which leads from the .fastq file and a .csv file that contains information about name identifier, run and barcode identifier, chromosomal position and the variant detected in the diagnosis to be evaluated in the follow-up sample. Through Ensembl Perl API,[15] the aligned mutated sequence and the aligned wild type (wt) sequence are presented in FASTA format (sequences of 40 bp). Finally, we obtained a .csv file containing the name identifier, run and barcode identifier, chromosomal position, the variant, the specific target sequence in FASTA format (mutated forward, mutated reverse, wt forward and wt reverse), the counts of each and the ratio (mutated/wt) in absolute values.

Results

A high percentage of AML patients could benefit from deep sequencing MRD approach

In total, 211 (80%) SNVs and 46 (20%) InDels were detected in the 190 patients analyzed at diagnosis using the NGS custom panel. We detected one variant (SNV or InDel) in 48 (25%) cases, 2 or more variants in 116 (61%) cases and no variants in 26 (14%) cases. In addition, we detected the *NPM1* type A mutation in 53 (28%) patients by qPCR. Genes (*TET2*, *ASXL1*, or *DNMT3A*) with evidence of clonal hematopoiesis of indeterminate potential (CHIP) association were excluded from the analysis.[11] Consequently, 82% of patients in our cohort could benefit from this approach.

Based on those genes reported as potential markers to monitor MRD,[16] and also the availability of follow-up samples, we focused on *IDH1/2* and *FLT3*-SNV. We identified at diagnosis *IDH1* mutations in 13 patients (7%), *IDH2* mutations in 27 patients (14%) and *FLT3*-SNV mutations (18%) in 34 patients.

Deep sequencing MRD has a sensitivity of 10^{-4} for SNVs and 10^{-5} for InDels

To establish the limit of quantification (LOQ) of the method, we used 10-fold serial dilutions of mixed mutated and control DNA. To study prototype InDels, we used gDNA from OCI-AML3 cells (*NPM1* type A) and to study prototype SNVs, we used both gDNA from OCI-AML3 cells (*DNMT3A*) and commercial reference gDNA (*IDH1*/*IDH2)*. As a control, we used a pool of gDNA from ten individuals without somatic mutations in these chromosomal regions. In all cases, initial allele frequency was 50% and a total of six dilutions were carried out to construct a calibration curve, covering a theoretical dynamic range from 10^{-1} to 10^{-7}.

As shown in Figure 2A, B, MRD NGS testing of *NPM1* (InDel) could quantify one mutated cell in the order of 10^{-5}, and in the case of SNVs (*IDH1*, *IDH2* and *DNMT3A*) the LOQ was 10^{-4}, which was reproducible for all SNVs tested.

NGS is more sensitive than dPCR for MRD testing

We compared the sensitivity of sequencing with that of dPCR using the same LOQ dilution protocol. Clone frequency expressed as target concentration (mutated copies/μL in wt copies/μL) gradually decreased with each dilution, reaching an LOQ of 10^{-3} for *NPM1*, *IDH1* and *IDH2* (Figure 2C–D). While both methods showed similar detection limits and good linearity, the LOQ for the sequencing method was one order of magnitude higher than that for dPCR (*IDH1* and *IDH2*), and two orders of magnitude higher for InDels (*NMP1*).

MRD status tested by sequencing has prognosis impact in AML

Median of depth coverage was 401,300 aligned reads (interquartile range 195,100–825,700) for the 88 *NPM1* and 18 SNV (9 *IDH1*, 7 *IDH2*, and 2 *FLT3*) follow-up samples evaluated. We detected no mutated sequence in 13 (12%) samples, 1–5 mutated sequences in 19 (18%) samples, and more than 10 in 74 (70%) samples. The ratio of mutated sequences to wt sequences defined MRD levels. Considering MRD levels from the 106 samples evaluated we established the optimal cutoff to classify MRD status (positive *vs* negative) by ROC curves (*Online Supplementary Figure S1*) at each check-point of MRD evaluation (post-induction [n=51], post-consolidation [n=55], or both together [n=106]).

Survival analysis revealed that positive MRD status (MRD levels > 0.1%) at post induction (n=35) was associated with a significantly lower rate of overall survival (OS) (33% *vs.* 78%; hazard ratio [HR]: 3.5; 95% confidence interval [CI]: 1.1–10.7; p=0.019), but a non-significant lower rate of disease-free survival (DFS) (58% *vs.* 78%; HR: 2.18; 95%CI: 0.63–7.5; p=0.208) (Figure 3A, B). At post-consolidation (n=28), MRD positive status (MRD levels > 0.025%) was associated both with significantly shorter OS (33% *vs.* 81%; HR: 6.0; 95% CI: 1.3–28.7; p<0.001), and significantly shorter DFS (17% *vs.* 94%; HR: 19.6; 95% CI:2.5–155.6; p<0.001) (Figure 3C, D). Also, survival analysis was performed combining post-induction and post-consolidation (n=63), in order to compare survival analysis with MFC and qPCR data sets. We observed that positive MRD status (MRD levels > 0.035%) was associated with a higher risk of relapse (48% *vs.* 81%; HR: 3.4; 95% CI: 1.4–8.5; p=0.005) and death (37% *vs.* 81%; HR: 4.2; 95% CI: 1.6–10.7; p<0.001) (Figure 3E, F). In order to test the power of *NPM1* and SNVs as independent predictive markers, we performed the analysis separately. Evaluating *NPM1* as an MRD marker (n=54), we found that MRD positive status was associated both with significantly shorter OS (43% *vs.* 78%; HR: 3.3; 95%CI: 1.2–8.8; p=0.011), and shorter DFS (57% *vs.* 85%; HR: 2.9; 95%CI: 0.9–7.6; p=0.052), and the similar results are found when we evaluated *IDH1*, *IDH2* or *FLT3*-SNV as MRD markers (n=11). Accordingly, MRD positive status was associated both with significantly shorter OS (17% *vs.* 100%; HR: NA; p=0.041), and shorter DFS (17% *vs.* 75%; HR:6.3; 95%CI:0.7–54; p=0.058).

In univariate Cox analysis (Table 2A), the risk of death was significantly higher in patients with increased age (HR: 1.04; p=0.013), those with *FLT3*-ITD (HR: 3.45; p=0.007), and those with MRD positive status tested by NGS (HR: 4.22; p=0.002). Risk of relapse was significantly higher only in those patients with MRD positive status tested by NGS (HR: 3.4; p=0.008). In multivariate analysis (Table 2B), the risk of death was significantly higher in patients with increased age (HR: 1.05; p=0.004), those with mutated *FLT3*-ITD (HR: 8.87; p=0.001), and those with MRD positive status tested by NGS (HR: 4.54; p=0.005). The risk of relapse was higher only in patients MRD positive patients tested by NGS (HR: 3.76; p=0.012).

MRD tested by sequencing improves prediction of OS and DFS over MFC and qPCR

A positive correlation was found when comparing MRD assessment by NGS *vs* MFC (r=0.47, p=0.005, n=75), and NGS *vs* qPCR (r=0.62, p<0.001, n=80) (*Online Supplementary Figure S2*). There were differences between positive MRD and negative MRD groups of patients tested by MFC, but they were not significant for OS (p=0.193) or DFS (p=0.117) (n=46, Figure 4A). Similarly, differences were observed between positive MRD and negative MRD groups by qPCR of *NPM1*, although significance was not reached for OS (p=0.212) or DFS (p=0.086) (n=46, Figure 4B).

Discussion

We have optimized and validated a high sensitivity NGS method to detect and quantify *NPM1*, *IDH1*, *IDH2* and *FLT3*-SNV mutated sequences at very low allele frequency in follow-up gDNA samples. NGS has demonstrated prognostic value for pre-treatment status in patients with AML,[17] and may also be a useful tool to detect MRD.[18, 19] We first studied the mutational profile of patients with AML using a custom NGS panel to ensure a high applicability (82% of patients). This approach is also a useful screening method to detect all potential MRD markers and to choose those most relevant. The combination of several markers is possible and recommended to overcome limitations of MRD assessment that are due to sub-clonal heterogeneity of AML and to CHIP.[11] Accordingly, our method has the capacity to evaluate multiple markers simultaneously and, considering that 61% of patients in our cohort had two or more genetic alterations this approach is sufficiently robust to monitor MRD in patients even if they present clonal evolution.

Reported variants associated with CHIP are frequently located in *DNMT3A*, *TET2* or *ASXL1* genes, and are detected at the preleukemic phase and at complete AML remission.[20-23] Indeed, any gene could carry both CHIP and non-CHIP variants, and these should be evaluated for each patient. Moreover, studies have shown that genes related to CHIP (*IDH1/2*) are useful for predicting prognosis because in these cases the genetic alterations have been acquired in the leukemic clone and not before.[24]

The sensitivity achieved with this method equates to one mutated cell per 100,000 cells (LOQ 10^{-5}) for *NPM1* and one mutated cell per 10,000 cells (LOQ 10^{-4}) for *IDH1*, *IDH2* and *FLT3*-SNV. This difference in sensitivity is related to the fact that the *NPM1* type A (insCCTG) mutation is rarely generated erroneously by NGS, and the quantification is precise.

Our method, as with any NGS method, has an intrinsic error rate that limits its sensitivity for most SNVs to 1–2% of all reads. This limitation can nevertheless be overcome by virtue of the scalable nature of NGS.[16]

Thus, we boosted NGS sensitivity by increasing the amount of DNA by PCR prior to sequencing, which increased the depth of coverage to one million reads. By also optimizing the bioinformatic analysis, we focused the search for the precise variant in order to eliminate random sequencing errors, enhancing the specificity of the technique and reducing the computational time. To the best of knowledge, our NGS method presents possibly the highest sensitivity reported for NGS in AML.[18, 19, 24-27]

dPCR is a relatively novel technique for precise and absolute quantification of nucleic acids, which is based on limiting partitions of the PCR volume and Poisson statistics.[28] It is also an extremely sensitive technique, with a high specificity due to the detection of mutant alleles.[29] However, when we compared the same standard dilutions in NGS and dPCR, NGS afforded a 2-log increment in LOQ for InDels (*NPM1*) and a 1-log increment for SNVs (*IDH1/2*), with the sensitivity of dPCR for InDels similar to that reported in a previously published study (10^{-2}).[30] Compared with NGS, dPCR is a faster measurement technique but, as it is focused, it requires allele-specific primers that can complicate the experimental procedure, and a high number of parallel experiments are needed to raise the sensitivity, increasing the cost of the assay. Additionally, although it is possible to multiplex dPCR, unfortunately only a few targets can be monitored simultaneously within each sample.[29] Another advantage of NGS technology is that it does not require calibration curves in each assay, and the results are reported in absolute values, facilitating its standardization.

The NGS method described in this report showed comparable sensitivities (10^{-4} for SNVs and 10^{-5} for InDels) to MFC methods in those cases with immunophenotyphically aberrant populations.[10, 31] Although our method showed a similar sensitivity to that of qPCR, it does not require oligonucleotides that hybridize specifically to a particular sequence, so all nucleotides in the amplified region can be studied. Consequently, the NGS test is capable of detecting all *NPM1* subtype mutations in the same assay.

We found positive correlations when MRD levels were evaluated by NGS *vs* MFC and *vs* qPCR, but not with the expected results. In the case of MFC, this could be explained, in part, because *NPM1* mutations are usually associated with monocytic subtype-AML, which frequently presents more difficulties for identifying MRD by MFC. Indeed, Salipante et al[27] described that the level of success of MFC depends greatly on the immunophenotype of the abnormal blasts and how to discriminate them from background regenerative blasts. Moreover, due to the lack of standardization, MFC shows substantial variability across laboratories, including that of sample processing, instrument configuration, number of events, and training of pathologists.[32] The lack of a strong correlation between NGS and qPCR could be explained by the nature of the sample (sequencing uses gDNA whereas qPCR uses cDNA). Although RNA overexpression allows a higher sensitivity of detection, RNA levels do not correlate with the number of tumoral cells, in contrast to

mutated DNA. Accordingly, mutated DNA is more representative of the tumoral burden than is overexpression of mutated RNA.[33] It should be noted that the prediction of survival and progression of AML using MRD NGS was improved over the other methodologies employed, at least in the cohorts evaluated.

Finally, survival analysis showed that MRD positive status tested by NGS was associated with a higher risk of relapse and death and that MRD negative status at post-consolidation was associated with a longer OS and DFS; as according to recently published studies.[23] Supporting these findings, previous studies reported that an MRD check-point at post-consolidation could be the best moment for analysis because a better prediction is observed.[8, 34-37] Cox regression multivariate analyses confirmed that MRD positive status by sequencing was the only factor with significant risk prediction of relapse ($p=0.012$).

In conclusion, we have optimized a new targeted sequencing method with high sensitivity for MRD evaluation with applicability for a high percentage of AML patients, improving the capacity to predict the AML outcome over MFC or qPCR in our cohort.

References

1. 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.

2. Cheson BD, Bennett JM, Kopecky KJ, et al. Revised recommendations of the International Working Group for Diagnosis, Standardization of Response Criteria, Treatment Outcomes, and Reporting Standards for Therapeutic Trials in Acute Myeloid Leukemia. J Clin Oncol. 2003;21(24):4642-4649.

3. Dohner H, Weisdorf DJ, Bloomfield CD. Acute Myeloid Leukemia. N Engl J Med. 2015;373(12):1136-1152.

4. 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.

5. Kronke J, Schlenk RF, Jensen KO, et al. Monitoring of minimal residual disease in NPM1-mutated acute myeloid leukemia: a study from the German-Austrian acute myeloid leukemia study group. J Clin Oncol. 2011;29(19):2709-2716.

6. Rubnitz JE, Inaba H, Dahl G, et al. Minimal residual disease-directed therapy for childhood acute myeloid leukaemia: results of the AML02 multicentre trial. Lancet Oncol. 2010;11(6):543-552.

7. Balsat M, Renneville A, Thomas X, et al. Postinduction Minimal Residual Disease Predicts Outcome and Benefit From Allogeneic Stem Cell Transplantation in Acute Myeloid Leukemia With NPM1 Mutation: A Study by the Acute Leukemia French Association Group. J Clin Oncol. 2017;35(2):185-193.

8. Zhu HH, Zhang XH, Qin YZ, et al. MRD-directed risk stratification treatment may improve outcomes of t(8;21) AML in the first complete remission: results from the AML05 multicenter trial. Blood. 2013;121(20):4056-4062.

9. Freeman SD, Virgo P, Couzens S, et al. Prognostic relevance of treatment response measured by flow cytometric residual disease detection in older patients with acute myeloid leukemia. J Clin Oncol. 2013;31(32):4123-4131.

10. Terwijn M, van Putten WL, Kelder A, et al. High prognostic impact of flow cytometric minimal residual disease detection in acute myeloid leukemia: data from the HOVON/SAKK AML 42A study. J Clin Oncol. 2013;31(31):3889-3897.

11. 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.

12. Grimwade D, Freeman SD. Defining minimal residual disease in acute myeloid leukemia: which platforms are ready for "prime time"? Hematology Am Soc Hematol Educ Program. 2014;2014(1):222-233.

13. Cedena MT, Rapado I, Santos-Lozano A, et al. Mutations in the DNA methylation pathway and number of driver mutations predict response to azacitidine in myelodysplastic syndromes. Oncotarget. 2017;8(63):106948-106961.

14. Gorello P, Cazzaniga G, Alberti F, et al. Quantitative assessment of minimal residual disease in acute myeloid leukemia carrying nucleophosmin (NPM1) gene mutations. Leukemia. 2006;20(6):1103-1108.

15. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. Genome Biol. 2016;17(1):122.

16. Tomlinson B, Lazarus HM. Enhancing acute myeloid leukemia therapy - monitoring response using residual disease testing as a guide to therapeutic decision-making. Expert Rev Hematol. 2017;10(6):563-574.

17. Patel JP, Gonen M, Figueroa ME, et al. Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. N Engl J Med. 2012;366(12):1079-1089.

18. Thol F, Kolking B, Damm F, et al. Next-generation sequencing for minimal residual disease monitoring in acute myeloid leukemia patients with FLT3-ITD or NPM1 mutations. Genes Chromosomes Cancer. 2012;51(7):689-695.

19. Kohlmann A, Nadarajah N, Alpermann T, et al. Monitoring of residual disease by next-generation deep-sequencing of RUNX1 mutations can identify acute myeloid leukemia patients with resistant disease. Leukemia. 2014;28(1):129-137.

20. Metzeler KH, Herold T, Rothenberg-Thurley M, et al. Spectrum and prognostic relevance of driver gene mutations in acute myeloid leukemia. Blood. 2016;128(5):686-698.

21. Bullinger L, Dohner K, Dohner H. Genomics of Acute Myeloid Leukemia Diagnosis and Pathways. J Clin Oncol. 2017;35(9):934-946.

22. Koeffler HP, Leong G. Preleukemia: one name, many meanings. Leukemia. 2017;31(3):534-542.

23. Jongen-Lavrencic M, Grob T, Hanekamp D, et al. Molecular Minimal Residual Disease in Acute Myeloid Leukemia. N Engl J Med. 2018;378(13):1189-1199.

24. 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.

25. Getta BM, Devlin SM, Levine RL, et al. Multicolor Flow Cytometry and Multigene Next-Generation Sequencing Are Complementary and Highly Predictive for Relapse in Acute Myeloid Leukemia after Allogeneic Transplantation. Biol Blood Marrow Transplant. 2017;23(7):1064-1071.

26. Debarri H, Lebon D, Roumier C, et al. IDH1/2 but not DNMT3A mutations are suitable targets for minimal residual disease monitoring in acute myeloid leukemia patients: a study by the Acute Leukemia French Association. Oncotarget. 2015;6(39):42345-42353.

27. Salipante SJ, Fromm JR, Shendure J, et al. Detection of minimal residual disease in NPM1-mutated acute myeloid leukemia by next-generation sequencing. Mod Pathol. 2014;27(11):1438-1446.

28. Brunetti C, Anelli L, Zagaria A, et al. Droplet Digital PCR Is a Reliable Tool for Monitoring Minimal Residual Disease in Acute Promyelocytic Leukemia. J Mol Diagn. 2017;19(3):437-444.

29. Roloff GW, Lai C, Hourigan CS, et al. Technical Advances in the Measurement of Residual Disease in Acute Myeloid Leukemia. J Clin Med. 2017;6(9).

30. Minervini A, Francesco Minervini C, Anelli L, et al. Droplet digital PCR analysis of NOTCH1 gene mutations in chronic lymphocytic leukemia. Oncotarget. 2016;7(52):86469-86479.

31. Ossenkoppele G, Schuurhuis GJ. MRD in AML: does it already guide therapy decision-making? Hematology Am Soc Hematol Educ Program. 2016;2016(1):356-365.

32. Cruz NM, Mencia-Trinchant N, Hassane DC, et al. Minimal residual disease in acute myelogenous leukemia. Int J Lab Hematol. 2017;39 Suppl 1:53-60.

33. Duployez N, Nibourel O, Marceau-Renaut A, et al. Minimal residual disease monitoring in t(8;21) acute myeloid leukemia based on RUNX1-RUNX1T1 fusion quantification on genomic DNA. Am J Hematol. 2014;89(6):610-615.

34. Jacobsohn DA, Tse WT, Chaleff S, et al. High WT1 gene expression before haematopoietic stem cell transplant in children with acute myeloid leukaemia predicts poor event-free survival. Br J Haematol. 2009;146(6):669-674.

35. Hourigan CS, Gale RP, Gormley NJ, et al. Measurable residual disease testing in acute myeloid leukaemia. Leukemia. 2017;31(7):1482-1490.

36. Rossi G, Carella AM, Minervini MM, et al. Optimal time-points for minimal residual disease monitoring change on the basis of the method used in patients with acute myeloid leukemia who underwent allogeneic stem cell transplantation: a comparison between multiparameter flow cytometry and Wilms' tumor 1 expression. Leuk Res. 2015;39(2):138-143.

37. Jourdan E, Boissel N, Chevret S, et al. Prospective evaluation of gene mutations and minimal residual disease in patients with core binding factor acute myeloid leukemia. Blood. 2013;121(12):2213-2223.

Acknowledgments

This study was supported by the Subdirección General de Investigación Sanitaria (Instituto de Salud Carlos III, Spain) grants PI13/02387 and PI16/01530, and the CRIS against Cancer foundation grant 2014/0120. M.L. holds a postdoctoral fellowship of the Spanish Ministry of Economy and Competitiveness (FPDI-2013-16409). P.R.P. holds a postdoctoral fellowship of the Spanish of Instituto de Salud Carlos III: Contrato Predoctoral de Formación en Investigación en Salud i-PFIS (IFI 14/00008).

Authorship contributors

Contribution: E.O. collected samples, performed experiments, analyzed and interpreted data, and wrote the manuscript. M.L. analyzed and interpreted data, and wrote the manuscript. I.R. analyzed and interpreted data. P.R.P. collected samples and analyzed and interpreted data. P.M-S. collected samples and clinical data. T.C. analyzed and interpreted data experiments. M.P., J.P.O., P.P., E.B., P.M., J.A.G.V., E.M., E.A., A.F., R.R. and P.M-B., collected samples and clinical data. B.S.V. performed experiments. J.N. and M.G. supervised research, analyzed and interpreted data, and wrote the manuscript. R.A. and J.M.L. collected

samples and clinical data, designed and supervised research and experiments, analyzed and interpreted data, and wrote the manuscript. All authors prepared the report and approved the final version.

Conflict-of-interest disclosure: The author declares no competing financial interests.

Role of funding source

Funding source not have additional role in this study.

Ethics committee approval

The study was conducted according to the Declaration of Helsinki, and the protocol was reviewed and approved by the institutional review board/independent ethics committee.

Correspondence: Rosa Ayala, Hematology Department, Hospital Universitario 12 de Octubre, Edificifio de Atención Ambulatoria, Planta 6 Bloque D, Avenida de Córdoba s/n, 28041, Madrid, Spain; email: rosam.ayala@salud.madrid.org.

TABLES

Table 1. Main characteristics of AML patients included in the MRD study

PATIENTS (n = 63)	
Follow-up sample type	
BM	58 (92%)
PB	5 (8%)
Sex	
Male	21 (33%)
Female	42 (67%)
Age at diagnosis	
Median	54 (IQR, 41.5–66.0)
Blasts at diagnosis	
Median count	69 (IQR, 51.0–81.0)
Leukocytes at diagnosis	
Median count ($_{\times}10^{9}$/L)	15.7 (IQR,12.2–20.24)
AML secondary	
No	59 (94%)
Yes	4 (6%)
Cytogenetic risk	
Favorable	25 (40%)
Intermediate	36 (57%)
Adverse	2 (3%)
***FLT3*-ITD**	
FLT3 negative	49 (78%)
FLT3 positive	14 (22%)
***FLT3*-TKD**	
FLT3 negative	60 (95%)
FLT3 positive	3 (5%)
NPM1	
NPM1 negative	6 (10%)
NPM1 positive	57 (90%)
HSCT	
No	42 (67%)
allo-HSCT	7 (11%)
auto-HSCT	14 (22%)
Relapse	
No	42 (67%)
Yes	21 (33%)
Death	
No	40 (63%)
Yes	23 (37%)
Treatment[*]	
3+7 regimen	50 (80%)
Flugaza	8 (13%)
Mylotarg	2 (3%)
Panobidara	3 (4%)

BM indicates bone marrow; PB, peripheral blood; AML, acute myeloid leukemia; ITD, internal tandem duplications; TKD, tyrosine kinase domain; allo-HSCT, allogeneic HSCT and auto-HSCT, autologous HSCT. [*]3+7 regimen of chemotherapy: one or two induction cycles of cytarabine and idarubicin during

seven and three days, respectively; and two or three consolidation cycles at high doses of cytarabine, twice a day for three alternates days followed by allo- or auto-HSCT. The remainder of patients were included in others clinical trials (Mylotarg, NTC0104104; Flugaza (NCT02319135); Panobidara, NCT00840346). Clinical data were collected in the following Spanish AML epidemiological registries: NCT01700413, NCT02006004, NCT00464217, NCT02607059, NCT01041040 and NCT01296178.

Table 2. Cox regression analysis

A.

	Risk of Death		**Risk of Relapse**	
	HR (95%CI)	**p value**	**HR (95%CI)**	**p value**
Sex (female vs male)	1.20 (0.50–2.83)	0.682	0.94 (0.37–2.44)	0.906
Age per year	1.04 (1.00–1.07)	0.013 *	1.03 (0.99–1.06)	0.069
Blasts at dx (%)	1.00 (0.99–1.02)	0.667	1.01 (0.99–1.03)	0.532
Leukocytes at dx ($\times 10^9$/l)	1.01 (0.99–1.01)	0.418	1.00 (0.99–1.01)	0.508
Favorable vs adverse (ELN risk)	0.67 (0.08–5.43)	0.714	0.75 (0.09–6.00)	0.786
Interm. vs adverse (ELN risk)	1.03 (0.13–7.86)	0.976	1.02 (0.13–7.82)	0.988
Mutated *FLT3*-ITD	3.45 (1.40-8.52)	0.007 *	2.37 (0.86–6.51)	0.095
Allo-HSCT vs intensive qt	1.35 (0.40–4.57)	0.634	1.78 (0.41–7.78)	0.44
Allo-HSCT vs auto-HSCT	0.29 (0.05–1.74)	0.176	0.64 (0.11–3.77)	0.629
MRD$^+$ by MFC	2.10 (0.67–6.62)	0.203	2.40 (0.77–7.46)	0.130
MRD$^+$ by qPCR	2.51 (0.56–11.2)	0.228	5.01 (0.64–38.8)	0.123
MRD$^+$ by NGS	4.22 (1.66–10.7)	0.002 **	3.41 (1.37–8.48)	0.008 **

B.

	Risk of Death		**Risk of Relapse**	
	HR (95%CI)	**p value**	**HR (95%CI)**	**p value**
Age per year	1.05 (1.02–1.09)	0.004 *	1.03 (0.99–1.07)	0.061
Sex (female vs male)	0.84 (0.33–2.17)	0.720	1.25 (0.44–3.52)	0.671
Leukocytes at dx ($\times 10^9$/l)	1.01 (0.99–1.03)	0.219	1.07 (0.99–1.02)	0.481
Favorable vs adverse (ELN risk)	13.75 (0.84–226.1)	0.067	7.09 (0.37–134.15)	0.192
Interm. vs adverse (ELN risk)	11.22(0.82–154.2)	0.071	5.86 (0.39–86.84)	0.203
Mutated *FLT3*-ITD	8.87 (2.54–30.95)	0.001 **	4.18 (1.11–15.69)	0.034
MRD$^+$ by NGS	4.54 (1.58–13.03)	0.005 **	3.76 (1.34–10.54)	0.012 *

(A) Univariate Cox regression analysis of each prognostic factor influencing the risk of relapse and risk of death of AML patients. (B) Multivariate Cox regression analysis evaluating the most relevant factors detected in univariate analysis. Abbreviations are explained in Table 1. CI indicates confidence interval; Dx, diagnosis; ELN, European Leukaemia Net; HR, hazard ratio; MFC, multiparametric flow cytometry; MRD, minimal residual disease; NGS, next-generation sequencing; qt, chemotherapy; *p values are considered significant (< 0.05), ** (< 0.01).

Figure legends

Figure 1. Workflow of NGS-MRD method

DNA amplification, library preparation and sequencing experimental workflow. gDNA is amplified by qPCR using specific primers. Library preparation is carried out in four steps: end repair, adaptor ligation, size selection, and PCR amplification. The library is then sequenced. A custom bioinformatic pipeline analyzes the obtained sequences. The results are expressed as a ratio of sequences mutated among wild-type sequences.

Figure 2. Calibration curve of MRD in serial dilutions

Top, 10-fold dilution curve for the assessment of sensitivity of sequencing in (A) InDels, using OCI-AML3 gDNA with 50% *NPM1* type A mutation ($R^2 = 0.98$); and in (B) SNV, using OCI-AML3 gDNA with 50% mutated *DNMT3A* ($R^2 = 0.98$), and gDNA with 50% mutated *IDH1* or *IDH2* from a commercial standard ($R^2 = 0.91$, $R^2 = 0.98$, respectively). *Bottom*, same 10-fold dilution curves for the assessment of sensitivity of dPCR in InDels (C, $R^2 = 0.98$); and in SNV (D, $R^2 = 0.91$ for *IDH1* and $R^2 = 0.98$ for *IDH2*).

Vertical red bars indicate LOQ according to the sample. Clone frequency is expressed as target concentration as mutated copies/μL in wild-type copies/μL. Negative control are included in the calibration curve and presented levels below the corresponding values of LOQ.

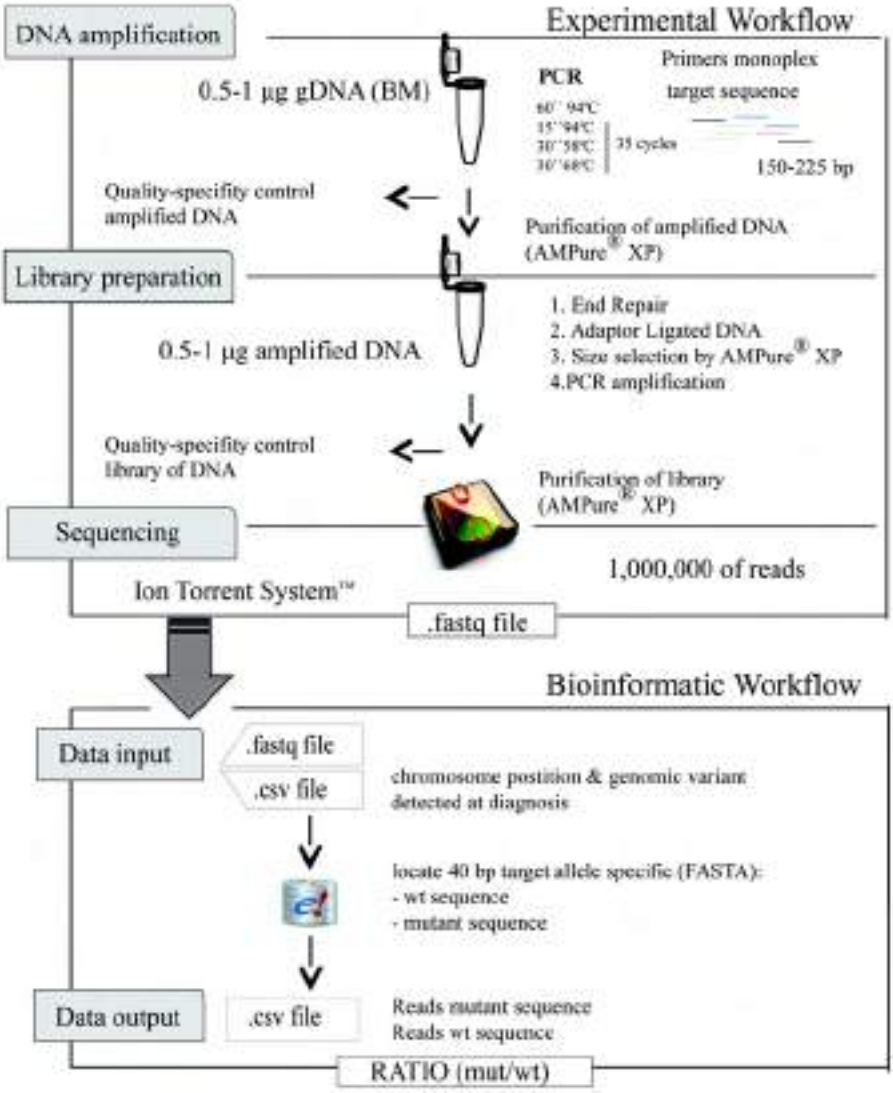
Figure 3. Analysis of OS and DFS in AML patients stratified according to MRD levels by sequencing

Analysis of OS for induction data set (A), for consolidation data set (C), and both together (E); and for DFS for induction data set (B), for consolidation data set (D), and both together (F). At post–induction check-point (n=35) the cutoff used was 0.001 for OS and DFS. At post-consolidation check–point (n=28) the cut off used was 0.00026 for OS and DFS. At both check-point (all data set) the cut off used was 0.00035 (n=63) for OS and DFS. Number of censored patients with respect to the stratified groups and the number at risk is indicated. **P* values are considered significant (< 0.05), ** (< 0.01).

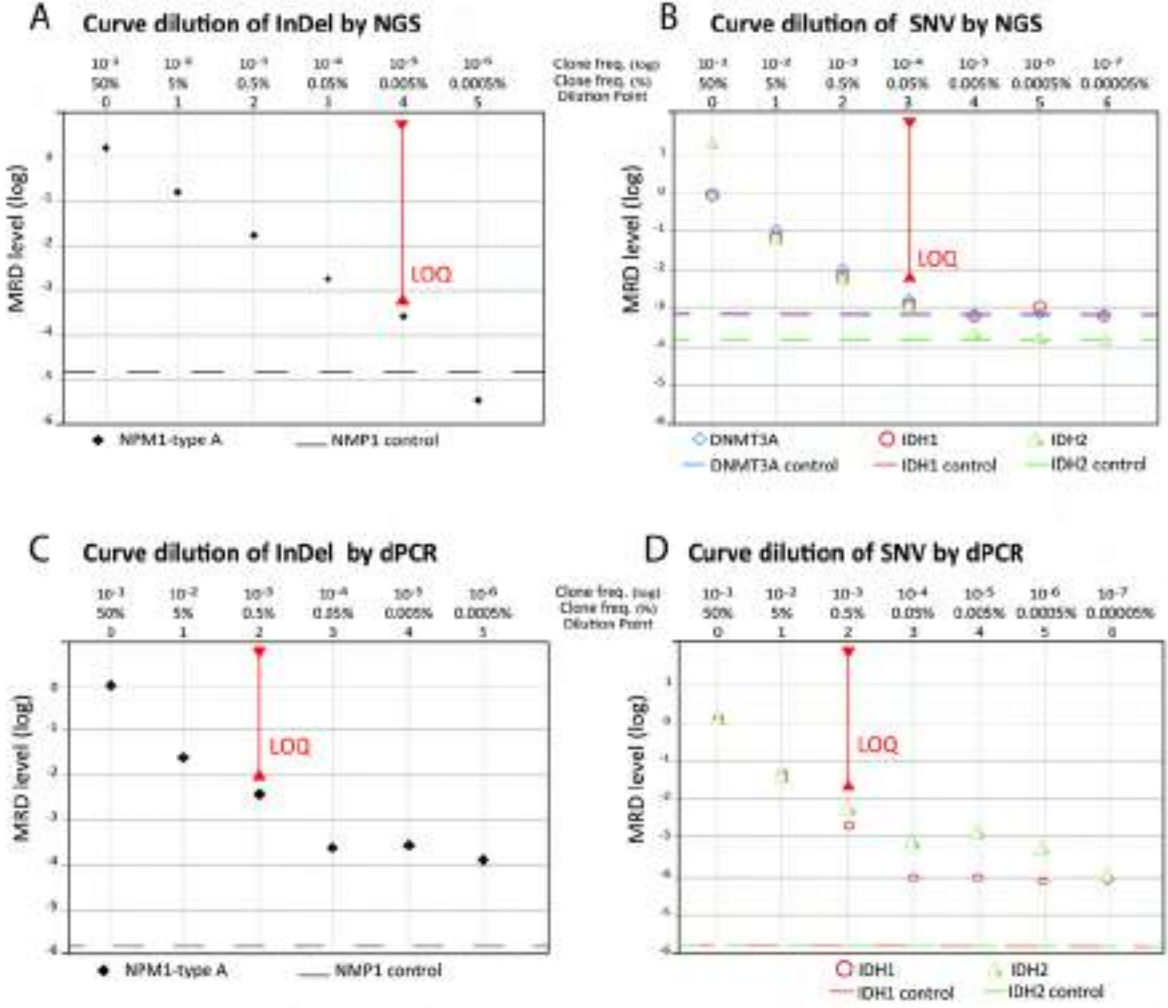
Figure 4. Prognosis analysis of OS and DFS in AML patients stratified according to MRD levels by conventional methods

Kaplan-Meier plots of (A) OS and (B) DFS with respect to MFC analysis and (C) OS and (D) DFS with respect to qPCR analysis. Number of censored patients with respect to each stratified group and number at risk is indicated. *p values are considered significant (< 0.05), ** (< 0.01).

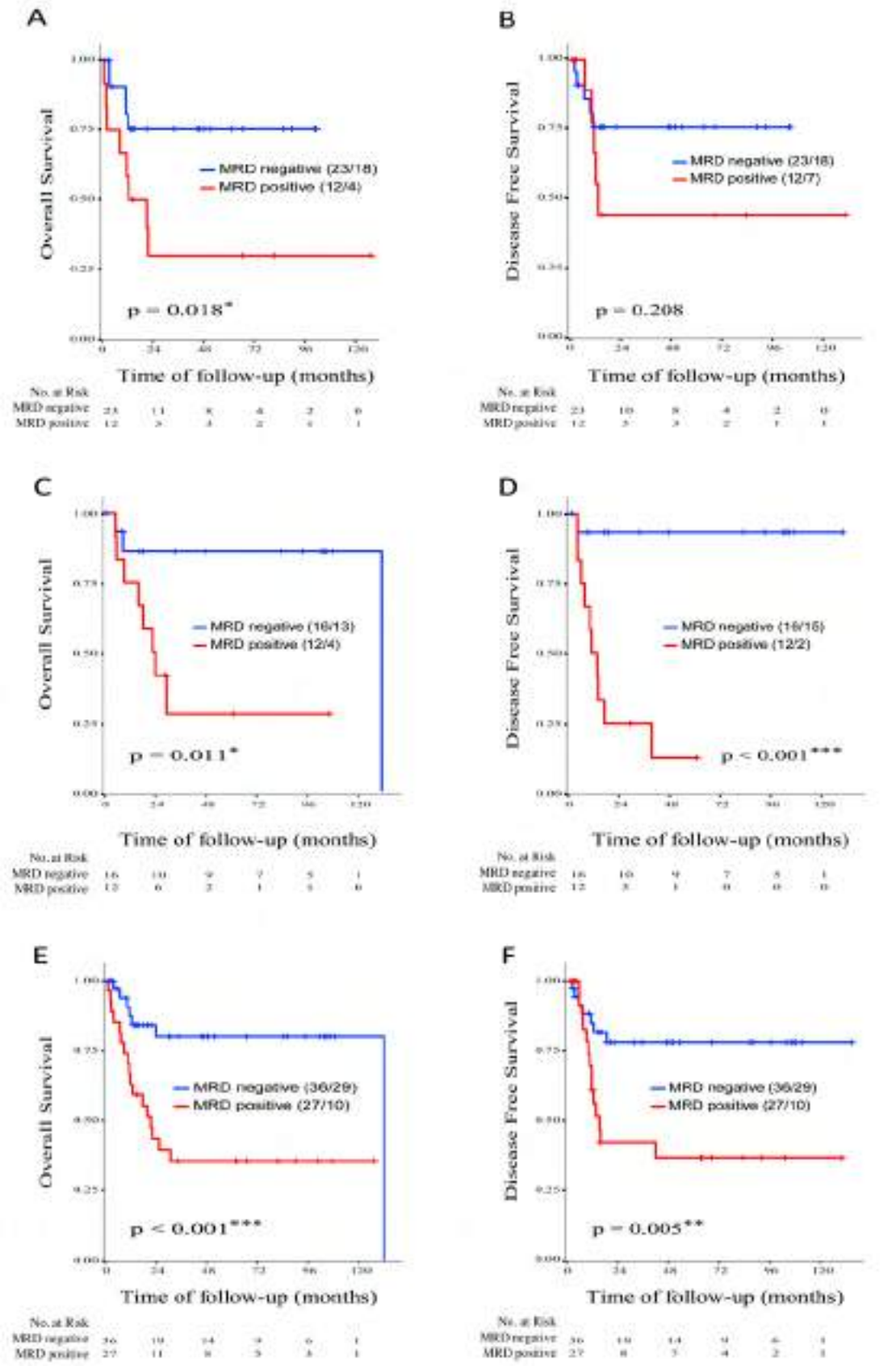
Onecha.E et al.
Figure 1.



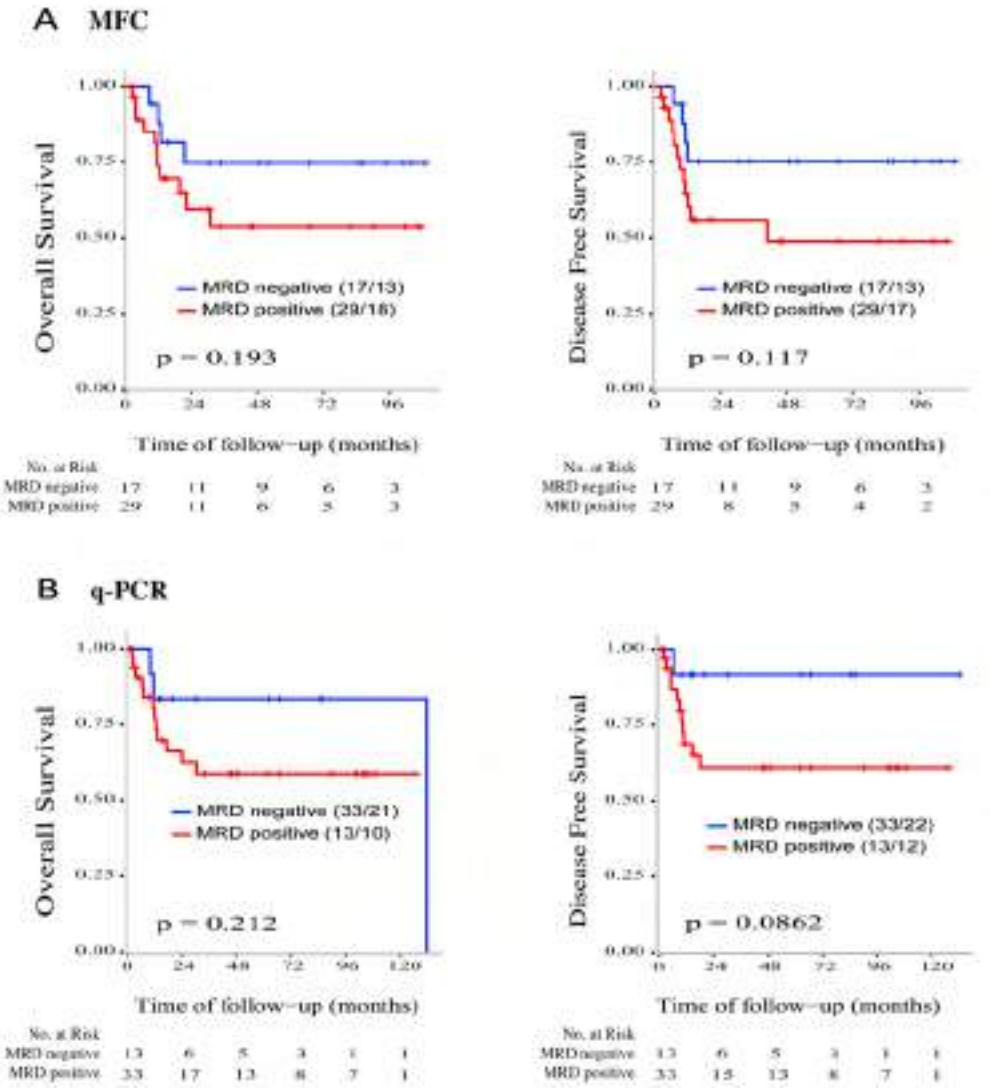
Onecha.E et al.
Figure 2.



Onecha.E et al.
Figure 3.



Onecha.E et al.
Figure 4.



Supplementary data

Supplementary 1. Conditions tested during the set-up of the NGS-based method

We tested a variety of methods to find optimal conditions to detect and quantify mutations at very low allele frequency in follow-up gDNA samples.

As a first approach, we used the same conditions as those in the diagnosis protocol, with 10 ng of gDNA, selected Ampliseq primers and the Ion AmpliSeq DNA & RNA Library Preparation workflow with an expected deep coverage of 500,000 reads. In a second approach, we used a higher DNA concentration (30–50 ng), higher specificity and quality primers (TIB MOLBIOL, Roche Diagnostics, SL) with a more robust polymerase (Platinum® PCR SuperMix High Fidelity), and the "Prepare Amplicon Libraries without Fragmentation Using the Ion Plus Fragment Library Kit" (Thermo Fisher Scientific, Inc.) and its workflow, testing a wide range of internal conditions. The coverage of sequencing was increased to 1,000,000 reads, however, the sensitivity was not increased.

Supplementary 2. Conditions of the optimal NGS-based method

DNA extraction was performed in a Maxwell®16 MDx instrument (Promega Biotech Iberica, SL) and quantified on a Qubit®2.0 Fluorometer (Invitrogen™, Thermo Fisher Scientific Inc., WA, USA).

The same primer pairs (*Supplementary Table S3*) used at diagnosis were used to amplify 0.5–1 μg of gDNA of patient samples (3 μg for calibration curve assays) by PCR using Platinum™*Taq* DNA Polymerase High Fidelity (Invitrogen™) with the following conditions: 60 seconds at 94ºC for initial denaturation, followed by 35 cycles of 15 seconds at 94°C for denaturation, 30 seconds at 58ºC for annealing and 30 seconds at 68ºC for extension. The final volume was 100 μL (79.6 μL DNA–H_2O, 10 μL 10× High Fidelity PCR Buffer, 4 μL 50 nM MgSO4, 2 μL 10 mM dNTP Mix (NZYTech, Lda, Lisbon, Portugal), 0.4 μL DNA polymerase (5U/μL), and 2 μL each of 10 μM forward and reverse primers. Libraries were constructed using NEBNext® Fast DNA Library Prep Set for Ion Torrent™ (New England Biolabs Inc., Ipswich, MA, USA). Specificity and quantification of the final product, both for amplified DNA and amplified libraries, was analyzed with the Agilent Bioanalyser 2100 (Agilent Technologies, Palo Alto, CA, USA).

The *IDH1* and *IDH2* dilution curves allowed us established the LOD of NGS at 10^{-4}, based on mean + 2.5 SD ratio from alternative 1 and alternative 2 results (*Supplementary Table S4*). In the same way, based on mean + 2.5 SD mutated aligned reads from alternative 1 and alternative 2, a technical cutoff was established

at 70 mutated aligned reads with a minimum coverage of 100,000 readings aligned, and a prognosis value of this cutoff was validated by survival analyses (*Supplementary Figure S3*).

Supplementary 3. Digital PCR of *NMP1* and *IDH1/2* mutations

dPCR for 10-fold dilutions curves of *NPM1*, *IDH1* and *IDH2* mutated gDNA was performed with specific primers and probes. Allele frequency was calculated as the ratio of mutated copies to wild-type copies/µL. dPCR assays were performed using QuantStudio™ 3D Digital PCR System using the FAM™/VIC® TaqMan® Assay (Applied Biosystems™, Thermo Fisher, La Jolla CA, USA) to study *NPM1* type A (c.863_864insTCTG), *IDH1* (c.394C/T) and *IDH2* (c.515G/A). A final volume of 14.5 µL (7.5 µL of PCR Master Mix 2×, 0.75 µL TaqMan® Assay 20× and 6.75 µL of gDNA at 50 ng/µL) was loaded into a QuantStudio™ 3D Digital PCR Chip v2 (Thermo Fisher), and amplified by PCR using the GeneAmp® 9700 system (Thermo Fisher). PCR was performed with the following conditions: 10 minutes at 96ºC for initial denaturation, 39 cycles of 2 minutes at 56–60ºC followed by 30 seconds at 98ºC, and a final 2 minutes step at 60ºC. After the PCR, each chip was read individually using the QuantStudio™ 3D Digital PCR Instrument (Thermo Fisher Scientific, Inc), which generates a file (.eds) containing the processed image data that is then interpreted using QuantStudio™ 3D AnalysisSuite Software (Thermo Fisher Scientific, Inc).

Supplementary 4. MRD monitoring of *NMP1* by qPCR

Detection and quantification of mutated *NPM1* transcripts were performed by allele-specific qPCR according to the procedure described by Gorello,[1] using RNA as starting sample. The protocol to detect *NPM1* by RT-PCR was performed in a final volume of 10 µl: 1.5 µL of H_2O + 0.5 µL of cProbe-LNA 4 µM (5´- 6FAM-ACCAAGAGGCT+A+T+TC+A+A– –BBQ -3´, Isogen Life Science) + 0.5 µL cNPM-F (10 µM, Isogen Life Science), 5´-GAAGAATTGCTTCCGGATGACT-3´+ 0.5 µL cNPM–mutA-R (10 µM, Isogen Life Science), 5´-CTTCCTCCACTGCCAGACAGA-3´+ 5 µL of Taq Man Fast Advanced Master Mix (Applied Biosystems) + 2 µL of cDNA. Amplification conditions were: 2 min at 50ºC for enzyme activation, 20 seconds at 95ºC for initial enzyme inactivation and AmpliTaq polymerase activation, followed by 40 cycles of 60 seconds at 95ºC for denaturation plus 20 seconds at 60ºC for annealing. We used the ABI PRISM 7900 Sequence Detection System (Applied Biosystems) for sample amplification and analysis.

For normalization of the expression of mutated *NPM1*, *GUS-β* expression was used as a control. MRD positive status was considered as the presence of *NPM1* copies > 0.00001 after therapy.[2]

Supplementary 5. MRD monitoring by MFC

After erythrocyte lysis, follow-up bone marrow samples were analyzed using a panel of monoclonal antibodies for the detection of the same immunophenotypic alterations described at diagnosis.[3] In our study, 10/75 (13%) samples evaluated by MCF were determined with MCF of 8 colours and the remaining 65/75 (87%) were determined with MCF of 4 colours. MRD positive status by flow cytometry was considered as the presence of AML cells greater than 0.001 at post-therapy.[2]

Supplementary 6. Statistical analyses

Contingency tables were used to analyse associations between categorical variables using Fisher´s test or Chi-square test for statistical significance. Student´s t-test was used to compare averages of continuous variables between groups. The concordance between sequencing, MFC and qPCR was analysed in log space using the Spearman correlation test. ROC (receiver operating characteristic) curves were employed to establish the cutoff value to predict survival by the NGS method, by MFC or by qPCR; however, for MFC and qPCR, the sensitivity and specificity achieved were comparable or less than those using the standard thresholds for MRD detections in AML and finally we used these (data not shown). For survival analysis, the endpoints examined were disease-free survival (DFS) and overall survival (OS), from the starting point of the treatment. In the cases that several samples from the same patient were evaluated, the one in which the lowest MRD levels were detected was selected for survival analysis. Survival curves were calculated according to the Kaplan-Meier method, and the log-rank test was used for estimation of survival and differences between groups. Univariate and multivariate analysis were performed using the Cox regression model; the most relevant variables for univariate analysis were: sex, age, blasts at diagnosis, leukocytes at diagnosis, cytogenetic risk (ELN recommendation; groups: favorable, intermediate and adverse), mutated *FLT3*-ITD, hematopoietic stem cell transplantation (HSCT) (groups: allo-HSCT, auto-HSCT and therapy), and MRD status by each technique (MFC, qPCR, NGS). Variables included in the multivariate analysis were chosen based on the results obtained in the univariate analysis and those with greater prognostic relevance in AML: sex, age, leukocytes at diagnosis, cytogenetic risk, mutated *FLT3*-ITD and MRD status by NGS.

Statistical analysis was performed using the R statistical software platform. All p values were two-sided, with statistical significance defined as a p–value of 0.05 or less.

Supplementary Table S1. Samples and patients evaluated.

Follow-up samples included in the study and their correlation patient, as well as evaluation time. In those

patients where a single sample was studied the patient is noted with the letter M. If several samples were studied per patient, these are listed numerically (M1, M2, etc.), and the sample selected for the analysis of survival is indicated. The levels of MRD in P3, P9, P38 and P62 patients were evaluated by studying both *NPM1* and *IDH1*. The sample selected for survival analysis is indicated. Two patients were removed from the study because of a missed follow-up.

Supplementary Table S2. Genes included in the NGS panel

Genes sequenced by NGS grouped by biological function, the chromosome where it is located, genomic coordinates (start–end) of region sequenced, the number of amplicons that the gene covers, the region of the gene that encompasses all the amplicons expressed as a percentage, and the number of exons.

Supplementary Table S3. Sequences of primers for MRD assay

Specific primer sequences (TIB MOLBIOL, Roche Diagnostics, SL) taken from the custom AML panel used at diagnosis (Ion AmpliSeq™, Thermo Fisher Scientific, Inc) for *DNMT3A* (used only for optimization), *IDH1*, *IDH2*, and *FLT3*; or from the commercial panel (Ion AmpliSeq™ AML Panel) in the case of *NPM1*.

Supplementary Table S4. VAF of dilution curves

Table represents the counts of aligned reads, both of the target sequence, wt sequence and the other two possible alternatives (sequences not mutated), the ratio (mutated aligned sequences/wt aligned sequences), and the fluctuation of the ratio with respect to the mutated sequence [Δlog(ratio)]; according to *IDH1* **(A)** and *IDH2* dilution curves **(B).** The LOD (10^{-4}) was established based on ratio mean + 2.5 SD from alternative 1 and alternative 2 results.

Supplementary Figure S1. ROC curves

Plots show the sensitivity or true positive rate (TPR) in the y-axis against 1-specificity or the false positive rate (FPR) in the x-axis, at various threshold settings. ROC curves determined the optimal cutoff level that maximizes sensitivity and specificity for the cases evaluated at each check-point for both OS and DFS studies. For OS the sensitivity and the specificity achieved was 0.69 and 0.77 at post-induction, 0.73 and 0.91 at post-consolidation, and 0.71 and 0.67 at both together. For DFS the sensitivity and the specificity

achieved was 0.77 and 0.60 at post-induction, 0.76 and 0.89 at post-consolidation, and 0.72 and 0.67 at both together. The area under the curve (AUC) is annotated.

Supplementary Figure S2. Correlation of levels of MRD measure by NGS and conventional methods

Correlation between NGS *vs* MFC (left) and correlation between NGS *vs* qPCR (right) detected by Spearman test; cases with available data for these tests were included. A significant positive correlation were found in both cases: NGS *vs* MFC (r=0.41, p=0.003), and NGS *vs* qPCR (r=0.46, p<0.001).

Supplementary Figure S3. Prognostic value of technical cutoff

A, OS curves of patients stratified according to MRD status based on technical cutoff (70 aligned mutated reads). The group categorized as MRD negative had greater OS than the group categorized as MRD positive (HR: 2.55 (1.00–6.46), p=0.049). **B,** DFS curves of patients stratified under same criteria, the MRD negative group had greater DFS than the group categorized as MRD positive (HR: 3.18 (1.16–8.69), p=0.024. Number of censored patients with respect to the stratified groups and the number at risk is indicated. **P* values are considered significant (< 0.05), ** (< 0.01).

References

1. Gorello P, Cazzaniga G, Alberti F, Dell'Oro MG, Gottardi E, Specchia G, et al. Quantitative assessment of minimal residual disease in acute myeloid leukemia carrying nucleophosmin (NPM1) gene mutations. Leukemia. 2006;20(6):1103-8.
2. Schuurhuis GJ, Heuser M, Freeman S, Bene MC, Buccisano F, Cloos J, et al. Minimal/measurable residual disease in AML: a consensus document from the European LeukemiaNet MRD Working Party. Blood. 2018;131(12):1275-91.
3. Tomlinson B, Lazarus HM. Enhancing acute myeloid leukemia therapy - monitoring response using residual disease testing as a guide to therapeutic decision-making. Expert Rev Hematol. 2017;10(6):563-74.

Onecha.E et al.
Supplementary Table S1.

Patient	Marker	1I		2I	1C		2C		3C			Selected Sample		
P1	NPM1	M1	**M2**				M3					M2	Induction (n=35)	Survival Analysis (n=63)
P2	NPM1	**M1**			M2		M3					M1		
P3	NPM1	**M1**					M2					M1		
	IDH1	M										–		
P4	NPM1	**M1**			M2							M1		
P5	NPM1	**M1**			M2		M3					M1		
P6	NPM1	**M1**							M2			M1		
P7	NPM1	**M**										M		
P8	NPM1	**M**										M		
P9	IDH1	M										–		
	NPM1	**M**										M		
P10	NPM1	**M**										M		
P11	NPM1	**M**										M		
P12	NPM1	**M1**			M2							M1		
P13	NPM1	**M**										M		
P14	NPM1	**M**										M		
P15	IDH2	**M**										M		
P16	NPM1	**M**										M		
P17	NPM1	M1	**M2**									M2		
P18	NPM1	**M**										M		
P19	NPM1	**M**										M		
P20	NPM1	**M**										M		
P21	NPM1	**M**										M		
P22	NPM1	**M**										M		
P23	NPM1	**M**										M		
P24	NPM1	**M**										M		
P25	NPM1	**M**										M		
P26	NPM1	**M**										M		
P27	FLT3			**M**								M		
	NPM1			M								M		
P28	IDH2	**M**										M		
P29	NPM1	**M**										M		
P30	FLT3	**M**										M		
P31	NPM1	**M**										M		
P32	NPM1	**M**										M		
P33	NPM1	**M**										M		
P34	NPM1	**M**										M		
P35	NPM1	**M1**			M2		M3					M1		
P36	IDH1				M1	**M2**						M2	Consolidation (n=28)	
P37	NPM1	M1							M2			M2		
P38	IDH1						**M1**		M2			M1		
	NPM1						M1		M2			–		
P39	IDH2						M1		M2	M3	**M4**	M4		
P40	NPM1								**M1**	M2		M1		
P41	NPM1	M1			**M2**							M2		
P42	NPM1				**M1**				M2			M1		
P43	NPM1						M1	**M2**				M2		
P44	NPM1	M1							**M2**			M2		
P45	NPM1	M1							**M2**			M2		
P46	NPM1	**M1**							M2			M1		
P47	NPM1								**M**			M		
P48	NPM1								**M**			M		
P49	NPM1	**M1**			M2							M1		
P50	NPM1								**M**			M		
P51	NPM1								**M**			M		
P52	NPM1				**M**							M		
P53	IDH2								**M**			M		
P54	NPM1								**M**			M		
P55	NPM1								**M**			M		
P56	NPM1	M1			**M2**							M2		
P57	NPM1				**M**							M		
P58	NPM1				**M**							M		
P59	NPM1	M1			**M2**		M3					M2		
P60	NPM1	M1			**M2**							M2		
P61	NPM1				**M**							M		
P62	IDH1	M1					**M2**					M2		
	NPM1	M1					M2					–		
P63	NPM1	M1			**M2**							M2		
–	NPM1	M1			M2							–		
–	NPM1								M			–		
		n=51			n=55									
		n = 106												

Onecha.E et al.
Supplementary Table S2.

	Gene	Chr	Start	End	Amplicons	Coverage (%)	Exons
Trancription factor	ETV6	Chr 12	11802955	12044078	20	94	8
	RUNX1	Chr 21	36164534	36421235	18	69	10
Signaling molecular	EPOR	Chr 19	11488559	11495009	21	93	8
	FLT3	Chr 13	28578144	28644774	33	97	24
	HRAS	Chr 11	532539	534348	10	83	5
	JAK2	Chr 9	5021946	5126885	57	97	23
	SH2B3	Chr 12	111855922	111886159	15	64	7
Epigenetic Regulation	DNMT3A	Chr 2	25457019	25523119	51	91	25
	IDH1	Chr 2	209101751	20811633	22	98	8
	IDH2	Chr 15	90627407	90634952	23	87	11
	TET2	Chr 4	106155047	106197701	64	99	10
	ASXL1	Chr 20	30954090	31025087	52	91	13
	KDM6A	Chr X	44732713	44970702	64	93	29
	KMT2A	Chr 11	118339409	118392950	145	96	37
	MPL	Chr 1	43803438	43818424	30	92	12
	PHF6	Chr X	133511597	133559416	22	98	11
Transcriptional regulation	CBL	Chr 11	119077153	119170540	41	93	16
	EZH2	Chr 7	148504653	148544423	44	99	21
	KIT	Chr 4	55524151	55604786	51	99	22
	KRAS	Chr 12	25362621	25398385	10	83	5
	NRAS	Chr 1	115251095	115258874	9	100	4
	CALR	Chr 19	13049314	13055076	23	86	9
Splicing	SF1	Chr 11	64552722	64545911	30	80	19
	SF3A1	Chr 22	30730553	30752852	37	94	18
	SF3B1	Chr 2	198256947	198294851	66	97	26
	SRSF2	Chr 17	74732208	74733231	5	70	2
	U2AF35	Chr 21	44513107	44524598	15	87	10
	ZRSR2	Chr X	15808511	15841407	26	97	11
	PRPF40B	Chr 12	50024310	50037977	54	95	26
Tumor supressor	PTEN	Chr 10	89624161	89725315	21	93	9
	TP53	Chr 17	7572847	7579960	23	93	13
	VHL	Chr 3	10183314	10195319	27	55	3

Onecha.E et al.
Supplementary Table S3.

GENE	PRIMERS
IDH1	Fw, 5´–AAGAATAAAACACATACAAGTTGGAAATTTCT–3´
	Rv, 5´–GAGAAGCCATTATCTGCAAAAATATCCC–3´
IDH2	Fw, 5´–ACAAAGTCTGTGGCCTTGTACTG–3´
	Rv, 5´–CTGGACCAAGCCCATCACCAT–3´
NPM1	Fw, 5´–GTTAACTCTCTGGTGGTAGAATGAAAAATAGA–3´
	Rv, 5´–GATATCAACTGTTACAGAAATGAAATAAGACG–3´
FLT3	Fw, 5´– TTGGAAACTCCCATTTGAGATCATATTCAT–3´
	Rv, 5´–TCTATCTGCAGAACTGCCTATTCCTAA –3´
DNMT3A	Fw, 5´–GATGACTGGCACGCTCCAT–3´
	Rv, 5´–GCTGTGTGGTTAGACGGCTTC–3´

Onecha.E et al.
Supplementary Table S4.

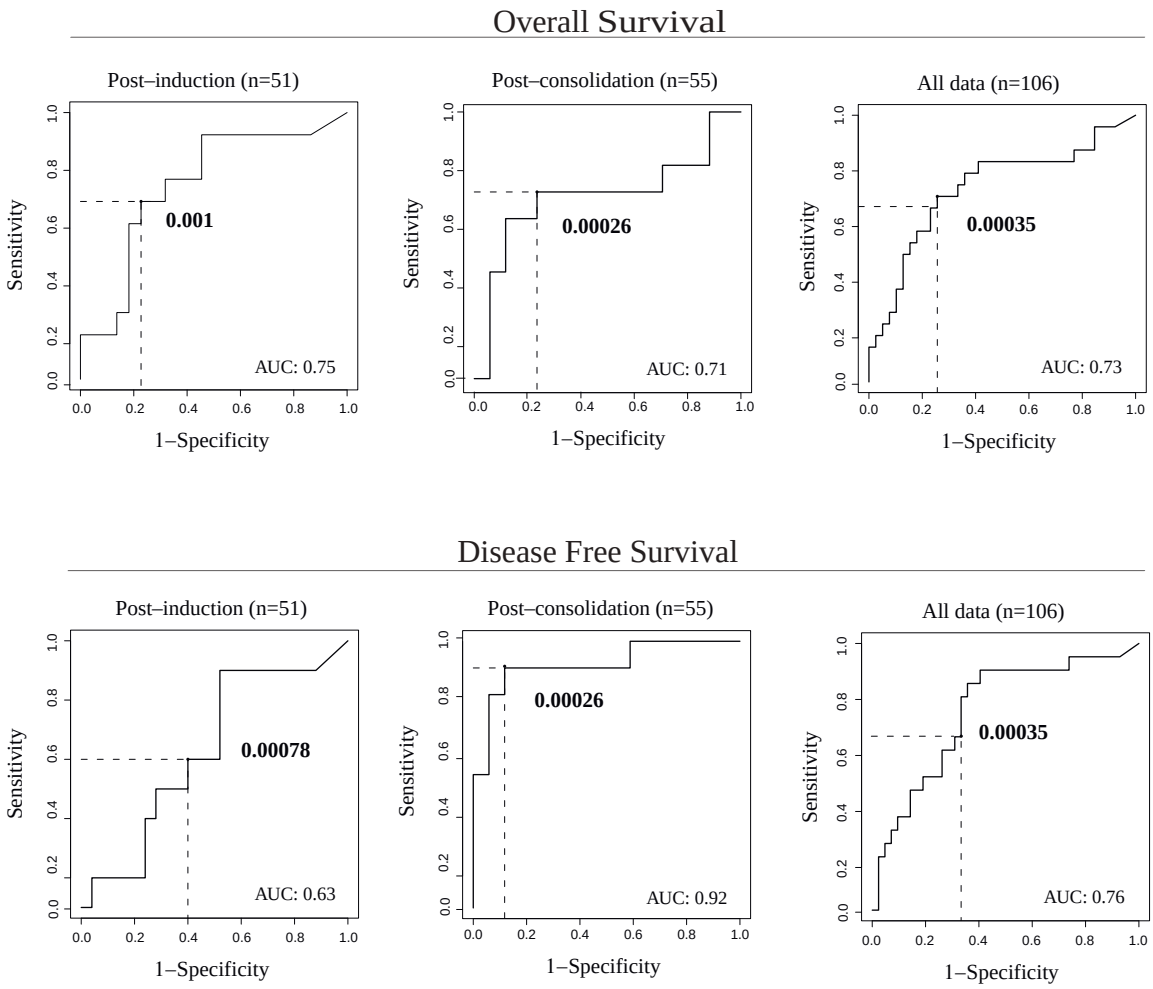
A) IDH1 Dilution Curve

		Mutated sequence (c.515G>A)			Alternative sequence 1 (c.515 G>T)				Alternative sequence 2 (c.515 G>C)			
Dil	Aligned no Reads	Aligned Mutated Reads	Ratio	log(ratio)	Aligned Alt.1 Reads	Ratio	log(ratio)	Δlog(ratio)	Aligned Alt.2 Reads	Ratio	log(ratio)	Δlog(ratio)
1	46,955	1,198,507	2.55[1]	1.41	50	1.06[-3]	-2.97	-4.38	57	1.21[-3]	-2.92	-4.32
10^{-1}	1,346,647	100,279	7.45[-3]	-1.13	19	1.41[-5]	-4.85	-3.72	6	4.46[-6]	-5.35	-4.22
10^{-2}	1,716,364	11,065	6.45[-4]	-3.19	44	2.56[-5]	-4.59	-2.4	1	5.83[-7]	-6.23	-4.04
10^{-3}	1,987,343	1,958	9.85[-5]	-3.01	30	1.51[-5]	-4.82	-1.81	1	5.03[-7]	-6.3	-3.29
10^{-4}	1,607,631	368	2.29[-4]	-3.64	35	2.18[-5]	-4.66	-1.02	0	0	–	–
10^{-5}	2,101,916	341	1.62[-4]	-3.79	35	1.67[-5]	-4.78	-0.99	0	0	–	–
10^{-6}	1,532,677	203	1.32[-4]	-3.88	19	1.24[-5]	-4.91	-1.03	0	0	–	–
Control	1,391,100	268	1.92[-4]	-3.72	19	1.36[-5]	-4.87	-1.15	1	7.19[-7]	-6.15	-2.43

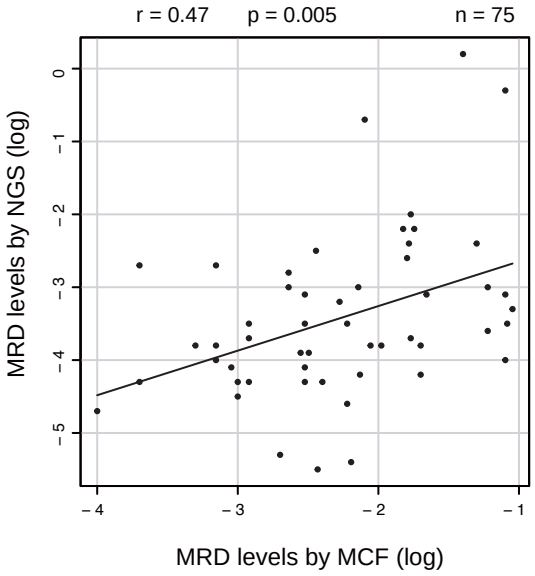
B) IDH2 Dilution Curve

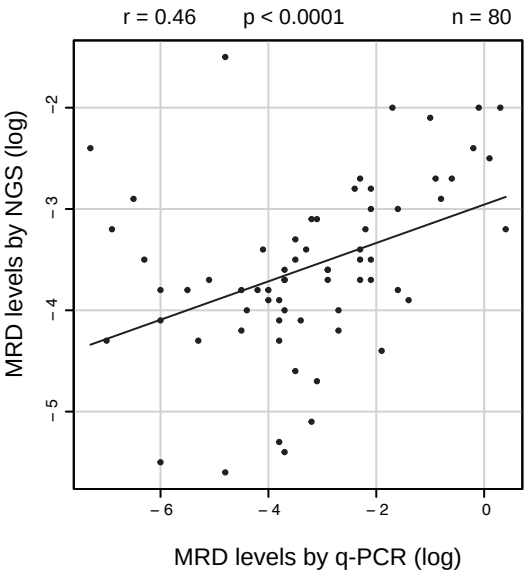
		Mutated sequence (c.394 C>T)			Alternative sequence 1 (c.394 C>A)				Alternative sequence 2 (c.394 C>G)			
Dil	Aligned no Reads	Aligned Mutated Reads	Ratio	log(ratio)	Aligned Alt.1 Reads	Ratio	log(ratio)	Δlog(ratio)	Aligned Alt.2 Reads	Ratio	log(ratio)	Δlog(ratio)
1	179,315	199,266	1.11	0.05	6	3.35[-5]	-4.48	-4.52	4	2.23[-5]	-4.65	-4.70
10^{-1}	319,918	24,128	7.54[-2]	-1.12	5	1.56[-5]	-4.81	-3.68	0	–	–	–
10^{-2}	369,661	2,687	7.26[-3]	-2.14	2	5.41[-6]	-5.27	-3.13	1	2.71[-6]	-5.57	-3.43
10^{-3}	330,661	372	1.12[-3]	-2.95	2	6.05[-6]	-5.22	-2.27	0	0	–	–
10^{-4}	222,345	120	5.39[-4]	-3.27	2	9.00[-6]	-5.05	-1.78	0	0	–	–
10^{-5}	288,670	348	1.21[-3]	-2.92	5	1.73[-5]	-4.76	-1.84	0	0	–	–
10^{-6}	411,043	265	6.45[-4]	-3.19	8	1.95[-5]	-4.71	-1.52	0	0	–	–
Control	391,730	425	1.08[-3]	-2.97	6	1.53[-5]	-4.81	-1.85	1	2.55[-6]	-5.59	-2.63

Onecha.E et al.
Supplementary Figure S1.



Onecha.E et al.
Supplementary Figure S2.





Onecha.E et al.

Supplementary Figure S3.