Shao Lina (Orcid ID: 0000-0002-6650-1698)

ERG amplification is a secondary recurrent driver event in myeloid malignancy with complex karyotype and *TP53* mutations

Winston Y. Lee[1], Efrain A. Gutierrez-Lanz[1], Hong Xiao[1], David McClintock[1], May P. Chan[1], Dale L. Bixby[2], Lina Shao[1].

[1]Department of Pathology, Michigan Medicine, University of Michigan, Ann Arbor, Michigan 48109-2800, USA.

[2]Division of Hematology and Medical Oncology, Department of Medicine, Michigan Medicine, University of Michigan, Ann Arbor, Michigan 48109-2800, USA.

Short title: *ERG* amplification in myeloid neoplasms

Keywords: acute myeloid leukemia, complex karyotype, ERG, amplification, iAMP21, TP53 mutation

Corresponding author:

Lina Shao, MD, PhD.
Department of Pathology and Clinical Laboratories
Michigan Medicine
University of Michigan
2800 Plymouth Rd, Building 35
Ann Arbor, MI 48109-2800
linashao@med.umich.edu

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/gcc.23027

Abstract

ERG is a transcription factor encoded on chromosome 21q22.2 with important roles in hematopoiesis and oncogenesis of prostate cancer. *ERG* amplification has been identified as one of the most common recurrent events in acute myeloid leukemia with complex karyotype (AML-CK). In this study, we uncover 3 different modes of *ERG* amplification in AML-CK. Importantly, we present evidence to show that *ERG* amplification is distinct from intrachromosomal amplification of chromosome 21 (iAMP21), a hallmark segmental amplification frequently encompassing *RUNX1* and *ERG* in a subset of high-risk B-lymphoblastic leukemia. We also characterize the association with *TP53* aberrations and other chromosomal aberrations, including chromothripsis. Lastly, we show that ERG amplification can initially emerge as subclonal events in low grade myeloid neoplasms. These findings demonstrate that *ERG* amplification is a recurrent secondary driver event in AML and raise the tantalizing possibility of *ERG* as a therapeutic target.

Keywords: acute myeloid leukemia, complex karyotype, ERG, amplification, iAMP21, TP53 mutation

Introduction

Ets-related-gene (ERG), a member of ETS family transcription factors located on chromosome 21q22.2, plays multifaceted roles in development and oncogenesis. During normal development, the pattern of ERG expression is important in orchestrating chondrogenesis, angiogenesis, and hematopoiesis.[1–3] In hematopoiesis, appropriate ERG expression is required for homeostatic maintenance of hematopoietic stem cells and in the differentiation of B-cells and megakaryocytes.[4–11]

ERG can also act as an oncogene. In approximately half of the prostate cancers, translocations that fuse the androgen-responsive *TMPRSS2* promoter with *ERG* gene lead to androgen-induced ERG overexpression.[12] In acute myeloid leukemia (AML), aberrant *ERG* expression is involved in leukemogenesis.[13] Overexpression of *ERG* in AML is associated with a worse prognosis.[14,15] In a subset of AML, copy number amplification of genomic regions containing ERG has been demonstrated as a mechanism in driving *ERG* overexpression.[16,17] Rarely, translocations involving *ERG* and a partner gene (such as, *FUS*) can lead to the production of a chimeric protein with neomorphic gain-of-function.[18]

Intrachromosomal amplification of a 5.1MB genomic region on chromosome21q22 (iAMP21) that includes *RUNX1* and *ERG* defines a clinically aggressive subset of B-lymphoblastic lymphoma (B-ALL) with high relapse risk.[19,20] While the amplified genes in iAMP21 generally associates with increased expression, how this amplification drives oncogenesis remains unclear.[21,22] *RUNX1* amplification, which is almost always present in iAMP21, is readily detectable by routine clinical fluorescent *in situ* hybridization (FISH) study and has been adopted as the working definition for identifying iAMP21 in B-ALL.[19] Using this definition, a recent report described a cohort of AML with possible iAMP21 that appears to draw some parallels with its counterpart in B-ALL.[23] However, it remains unclear how this cohort of AML with iAMP21 relates to B-ALL with iAMP21 and AML with *ERG* amplification.

ERG amplification has been identified as one of the most common recurrent events in acute myeloid leukemia with complex karyotype (AML-CK).[16,17,24] By definition, AML-CK harbors 3 or more unrelated chromosomal aberrations and can arise in a de novo fashion, but may also arise in patients with an antecedent myelodysplastic syndrome (MDS), myelodysplastic/ myeloproliferative neoplasms, and therapy-related myeloid neoplasms.[25] The presence of complex karyotype is strongly associated with TP53 mutations, as demonstrated in cohorts of AML and MDS.[26–29] Therefore, AML-CK categorically represents the evolutionary convergence among these types of myeloid malignancies. These chromosomal aberrations, although highly variable, are non-random with recurring large structural changes, including deletions of 5q and 7q.[17,30–32]

Using comprehensive cytogenetic techniques, including karyotyping, FISH studies, and high-density copy number microarrays, we define *ERG* amplification and its associated genetic abnormalities in myeloid malignancies. Importantly, we demonstrate that *ERG* amplification can exhibit three distinct patterns

sharing features with genetic aberrations associated with *TP53* mutations. Furthermore, we demonstrate the initial emergence of *ERG* amplification as low level subclonal events in myelodysplastic syndromes. These findings support that *ERG* amplification is a recurrent secondary event in the evolution of myeloid malignancy.

Methods

Case selection

This is a retrospective single institution study conducted with the approval of internal review board (IRB# HUM0043196 and HUM00160360). Sixteen cases with *ERG* amplification were identified by cross-referencing our pathology archive with our institutional cohort (325 cases) of myeloid malignancy with high-density copy number array studies from 06/2013 to 03/2020. Fourteen cases of AML-CK without *ERG* amplification was also identified as a comparison cohort for ERG immunohistochemistry and survival analysis. Then the diagnoses, cytogenetic study and medical records were reviewed and reclassify based on the 2016 WHO criteria when necessary. The AML cohort from TCGA was used as a validation cohort, and the publicly available data, including diagnosis, age, karyotype, SNP6.0 copy number microarray results, and pertinent mutations, were generated by TCGA Research Network: https//www.cancer.gov/tcga. The TCGA data was initially interrogated and subsequently obtained through cBioportal for Cancer Genomics.[33,34]

Karyotypic and fluorescent in situ hybridization (FISH) analyses

At least 20 G-banded metaphase cells were obtained from overnight and/or 24-hour cultures using standard techniques. Interphase and metaphase FISH analysis was performed, where available, using *ERG* break-apart probes (Empire Genomics, Williamsville, NY). Two hundred nuclei were scored for the copy number of ERG signals. Metaphase cells with *ERG* amplification were analyzed to locate the amplified *ERG* signals. Karyotypic and FISH results were interpreted according to the International System for Human Cytogenetic Nomenclature (ISCN 2016). FISH images were captured utilizing a Leica DMRA microscope with the Cytovision Imaging system (Leica Microsystems, Buffalo Grove, IL).

High density genomic copy number microarray

Genomic DNA was extracted from bone marrow aspirate samples using the QIAamp DNA mini extraction kit (Qiagen, Germantown, MD) according to the manufacturer's instructions. Cytoscan HD array (ThermoFischer, Waltham, MA) was performed according to the manufacturer's instruction. The analysis was performed using the Chromosome Analysis Suite provided by the manufacturer and R-based Easy Copy Number pipeline (EaCON; https://github.com/gustaveroussy/EaCoN), based on Affymetrix Power tools (ThermoFischer, Waltham, MA) and ASCAT.[35] Copy number change was included only if involving greater than 35 consecutive probes and displaying a log R ratio (LRR) with a magnitude equal to or greater than 0.2. Chromothripsis is defined as equal to or greater than 10 alternating copy number changes.[36] The copy number change was plotted using the karyoploteR package.[37]

Overall survival analysis

The overall survival analysis was based on the Kaplan-Meier method using the survival and survminer packages in R. Between-group comparison was performed using a log-rank test. The overall survival is defined between the time of the initial diagnosis and the time of death as the primary endpoint. The data is right censored if the patient is alive at the time of the analysis.

Immunohistochemical staining and analysis

Immunohistochemical study for ERG protein expression was performed on 4-micron sections of formalin fixed paraffin embedded decalcified bone marrow cores using a rabbit monoclonal anti-ERG antibody (clone EPR3864) on Ventana Benchmark XT automated staining system (Ventana Medical Systems, Tucson, AZ). The stained slides were digitized using Leica-Aperio GT450 (Leica Biosystems, Buffalo Grove, IL). The H-score analysis was performed using Quantitative Pathology and Bioimage Analysis

toolsets (QuPath).[38–40] Briefly, the degree of nuclear staining intensity was assigned into 4 categories with negative staining as 0+ and highest intensity as 3+. The percentage of nuclei of each intensity level is calculated as: H-score = 1x (% cells 1+) + 2x(% cells 2+) +3x(% cell 3+). The statistical comparison of the H-scores between AML with and without *ERG* amplification was performed using Kruskal-Wallis rank sum test.

Results

Clinical characteristics of patients with myeloid neoplasms with ERG amplification.

We identified 16 cases (~4.9% of cases) with segmental aberration of chromosome 21 that harbors *ERG* amplification. Cases with entire or partial trisomy 21 were excluded. A summary of the clinical characteristics is recorded in Table 1. The median age of the cohort is 67 years old (range: 44-79 years old) with a male to female ratio of 1.28. Six patients presented with AML with myelodysplasia related changes (AML-MRC). Among the 2 patients presented with de novo myelodysplastic syndrome (MDS), one patient ultimately succumbed to AML-MRC. Five patients, previously treated with cytotoxic chemotherapy for a prior non-myeloid malignancy, presented with therapy related myeloid neoplasms, including therapy related AML (n=3) and therapy related myelodysplastic syndrome (n=2). Additionally, 3 patients with a history of JAK2 p.V617F positive myeloproliferative neoplasm (MPN) were diagnosed with AML/ blast crisis.

Multimodal mechanisms of ERG amplification

Among the 16 cases, 14 cases demonstrate variably sized segmental amplification of chromosome 21q22 (**Figure 1A**). The average length of the segment is 2,603KB with a range of 247KB to 6,558KB. All segments of amplification exhibit allelic imbalance and harbor *ERG* and *ETS2* genes. *RUNX1* and *U2AF1* are neighboring genes that are also known to associate with myeloid neoplasms.[41] In this cohort, *RUNX1* and *U2AF1* do not consistently co-amplify with *ERG*. *RUNX1* amplification is present in 3 cases; whereas *U2AF1* amplification in 2 cases. Interestingly, the regions flanking the amplified *ERG* containing segments can exhibit variable copy number alterations. For example, in patient 6, the flanking regions exhibit copy number loss. In contrast, patient 16, the flanking regions demonstrate lesser degrees of amplifications than the *ERG* segment.

To validate ERG copy number gain, we performed interphase and metaphase FISH studies on 14 cases with available materials using a break-apart probe set against 5' and 3' ends of *ERG*. The remaining materials from patients 8 and 10 were insufficient for further analysis. The FISH studies confirmed *ERG* amplifications and excluded rearrangements in all cases (**Table 2**). Importantly, by correlating karyotypic, interphase, and metaphase FISH analyses, three patterns of *ERG* amplification emerge from the FISH studies (**Figure1B, 1C, and 1D**). The first pattern is characterized by the presence of tandem homogenous staining region on metaphase FISH, consistent with intrachromosomal amplification. In the second pattern, *ERG* amplification is detected as increased marker chromosomes (containing centromeres) in three cases. In the third pattern, which is detected only in patient 9, the *ERG* amplification manifests as double minute chromosomes (without centromeres).

Emergence of ERG amplification as a subclonal event in MDS

The two cases of de novo MDS (patients 5 and 12) exhibit subthreshold segmental copy number amplification involving *ERG* with corresponding allelic imbalance by copy number array analysis (**Figure 2**). In patient 5, FISH studies detected *ERG* amplification as marker chromosomes in 16% of nucleated cells; and in patient 12, as tandem homogenous staining region in 19% of nucleated cells (**Table 2**). In contrast, the patients with above threshold LRR all harbor the amplification in >32% of nucleated cells by FISH studies. Therefore, in patient 5 and 12, *ERG* amplifications are subclonal events.

Association with chromosomal structural variations

All patients, except for patient 12, have complex karyotype (with >= 3 chromosomal abnormalities as defined in WHO 2016) with 14 cases harbors >= 5 abnormalities. These complex changes in karyotypes are also reflected in the patterns of copy number changes in high-density copy number microarrays (**Figure 3**). Importantly, we identified chromothripsis in 9 cases. Patient 12 harbors del20q as the sole chromosomal abnormality on karyotype. As discussed above, the FISH and copy number array studies provide evidence of subclonal *ERG* amplification in this patient. Structural alteration of chr5q is the most frequent event, including 9 cases with 5q deletion and 1 case with loss of heterozygosity (LOH) of 5q without copy number alteration. Loss of 7q is the second most frequent event, as detected in 6 cases.

Association with TP53 alterations

Among the 6 cases with available next generation sequencing results, 4 cases demonstrate pathogenic TP53 mutations (Table 2). Through the copy number array and karyotypic analysis, we have also identified 5 cases with 17p loss and 2 cases with 17p LOH involving *TP53*. Altogether, a total of 9 cases harboring copy number alteration, LOH, and/or pathogenic mutations involving *TP53*, were identified. These findings highlight the association of *TP53* alteration with *ERG* amplification.

ERG amplification in TCGA-AML cohort

To validate the above findings, we interrogated the TCGA-AML cohort for *ERG/ETS2* amplification using the same selection criteria. Eight of 198 cases (approximately 4%) demonstrated *ERG* amplifications with similar segmental patterns seen in our cohort (**Figure 4**). *RUNX1* is co-amplified in 3 cases and *U2AF1* in 7 cases. The flanking regions also show variable patterns of copy number alterations. Five of the 8 cases are associated with complex karyotype (**Table 3**). Chromosome 5q deletions are still the most frequent structural variations (n=6); whereas, deletion 7q is present in 3 cases. The association with *TP53* alteration is also remarkable. All 6 cases with 5q deletion harbor pathogenic *TP53* mutations as well as copy number alterations.

ERG protein expression in AML

ERG amplification in AML is associated with increased *ERG* transcript levels.[16,17] Therefore, we hypothesize that the ERG protein will likely be increased as well. Immunohistochemical stains using monoclonal anti-ERG antibodies were performed on sections of marrow cores from cohorts of AML-CK with and without *ERG* amplification and assessed semiquantitatively by H-score using QuPath image analysis. Although the median H-score of AML with *ERG* amplification is higher than the cohort without ERG amplification (**Figure 5**), the difference between the two groups is not statistically significant ($p = 0.40$; Kruskal-Wallis test). Therefore, whether ERG amplification leads to increased protein expression remains inconclusive. However, it is important to note that the AML with highest ERG IHC H-score (250 of 300) also harbors the highest copy number amplification of ERG in the form of double minute chromosomes (Pt 9)..

Outcome and survival analysis

Twelve patients presented with a diagnosis of AML, among whom 7 received induction chemotherapy and 4 venetoclax and decitabine. One patient (pt 3), who received allogenic hematopoietic stem cell transplant following treatment with decitabine/ venetoclax, is alive beyond 13 months and is undergoing treatment for relapsed disease. Four patients presented with MDS, including 2 with therapy related myeloid neoplasms (pt 1 and 11) and 2 with *de novo* MDS (pt 5 and 12). Patients 1 and 11 were placed on supportive care due to poor performance status and co-morbidities. Interestingly, patients 5 and 12 are the only two cases in our cohort that harbor subclonal ERG amplification. Patient 5 initially presented MDS with multilineage dysplasia and ring sideroblasts and a very high risk IPSS-RA score (6.9). Within 7 months, the patient expired shortly after transformation to AML. Patient 12 is the only one with simple karyotypic abnormality. However, we do not have follow-up clinical information after diagnosis. The patient's death was discovered by searching public database. Comparison of the overall survival between our cohort of AML with *ERG* amplification (median survival: 2.8 months) and patients with AML-

CK without *ERG* amplification (median survival: 2.0 months) shows no significant difference (supplementary figure 1; p = 0.40; log rank test).

Discussion

Our study demonstrates *ERG* amplification is a recurrent event that occurs in a subset of AML with complex karyotype, consistent with prior studies.[16–18] We demonstrated that *ERG* amplification can emerge as low level subclones in MDS and become dominant clones in AML. For first time, we describe the presence of *ERG* amplification in AML transformed from MPN. Our analysis reveals that the amplified regions exhibit great variability in sizes with frequent *ETS2* co-amplifications. *ETS2*, another member of ETS family transcription factors, locates approximately 139kb telomerically from *ERG* and is the closest neighboring coding gene. Due to its proximity, *ETS2* frequently co-amplifies with *ERG.*[16,17] The biological function of *ETS2* remains poorly understood. Interestingly, *ETS2* functionally interacts with ERG in the transcriptional control of megakaryopoiesis.[9,11] There is also limited data to suggest that AML with increased *ETS2* expression portends adverse prognosis.[42] In previously reported cohorts, minimal common region analysis identifies *ERG* as the only gene that is consistently amplified despite frequent *ETS2* co-amplification.[16,17] In addition to complex karyotype and frequent *TP53* mutations, we also noted high incidence of chromothripsis.

The interphase and metaphase FISH studies show that *ERG* amplification can occur in three patterns: intrachromosomal amplification, supernumerary marker chromosomes, and double minutes. These patterns, encompassing in cis, in trans, and extrachromosomal DNA amplification, suggest that *ERG* amplification arises through multiple different mechanisms.[43–45] For example, tandem homogenous staining region can result from bridge-fusion-break (BFB) cycles, and double minutes from chromothripsis. Similar patterns of amplification involving other genes have been described in AML with complex karyotype and *TP53* mutations.[30,46,47] However, a large scale analysis across a large selection of tumor types using whole genome sequencing demonstrates that chromothripsis can co-occur with other types of complex rearrangements, such as BFB cycles.[48] Therefore, whole genome sequencing studies will be needed to further dissect the mechanism of *ERG* amplification and its association with other structural abnormalities.

Despite sharing overlapping regions on chromosome 21q22, *ERG* amplification in AML is distinctly different from iAMP21 as defined in B-ALL. In addition to different patterns of copy number alterations (see above), *RUNX1* amplification, a clinical marker for iAMP21 B-ALL,[19] is infrequently present in our cohort. Secondly, iAMP21 in B-ALL is not known to associate with complex karyotypes and/or *TP53* mutations.[22,49,50] In contrast, our data strongly suggest that *ERG* amplification arises as a secondary driver. Interestingly, a recently described cohort of AML with iAMP21 that is solely defined by *RUNX1* FISH studies shares similar clinicopathologic features with our AML cohort. Both cohorts are comprised of therapy-related myeloid neoplasms and AML-MRC harboring complex karyotype, in particular deletions of chromosomes 5 and 7, and frequent *TP53* mutations.[23] The strong association of *TP53* mutations and complex karyotype in MDS and AML has been well described.[26–29] Although the cohort as described by Xie *et al* likely includes a significant number of cases with *ERG* amplification, confirmatory studies will be needed to ascertain the relationship to AML with *ERG* amplification.

In addition to AML, *ERG* amplification was also present as subclonal events in 2 cases of MDS in Pt 5 and 12. These *ERG* amplifications were first noted as subthreshold copy number gain on copy number arrays and later confirmed by FISH studies to be present in 15-19% of nucleated cells. In patient 5, the subclonal *ERG* amplification is present in the backdrop of a dominant clone with complex karyotype, including deletion in chromosome 5q. The disease, initially diagnosed as MDS with ring sideroblasts and multilineage dysplasia, progressed rapidly. The patient expired shortly after transformation to AML within 7 months of diagnosis. As for patient 12, the dominant clone carries deletion 20q as the sole karyotypic abnormality. However, we have limited information on the course of his disease. Overall, these findings suggest that *ERG* amplification can emerge as a low level secondary event in the early phase of disease evolution.

We assessed ERG protein levels in AML with *ERG* amplification using immunohistochemical stain and evaluate semi-quantitatively using H-score. All cases of AML with *ERG* amplification exhibit intense nuclear staining for ERG protein. However, the increase in expression was not statistically significant when compared to AML-CK without *ERG* amplification. We should point out that this result remains inconclusive for the following reasons. Firstly, our immunohistochemical stain may lack the dynamic range to sufficiently detect the difference. Significant physiologic ERG expression has been observed in hematopoietic progenitor cells and the expression dissipates during maturation and differentiation. Prior studies have shown that *ERG* amplification can lead to further increase *ERG* RNA transcript level in AML.[16,17] Second, we were unable to exclusively score ERG protein level in the myeloid blasts. Our cohorts of AML exhibit significant maturation. Therefore, the marrow cores are a mixture of maturing myeloid, erythroid, megakaryocytic elements, and variable number of blasts, ranging from 30% to 80%. Our computational method can exclude megakaryocytes by nuclear size. However, it is impossible to morphologically distinguish early myeloid and erythroid precursors from blasts in our stained sections. Cases with higher blast fraction will likely to have scored higher for ERG staining. Lastly, our study does not address the possibility that ERG expression may be increased through epigenetic mechanisms in non-amplified cases. Therefore, we believe additional studies using multiplex labeling techniques to separate blasts from maturing hematopoietic elements will provide further insights.

AML-CK represents evolutionary convergence of different types of myeloid malignancies. The adverse prognosis of this disease can be attributed to its onset in patients with advanced age and poor performance status, and the ineffectiveness of standard chemotherapy due to its inherent genetic heterogeneity and complexity. Therefore, identification of recurrent secondary driver events may offer opportunities for targeted therapy. *ERG* amplification in our cohort and others has emerged as one of the most common recurrent drivers in AML-CK and serve as a potential target for therapy.[16–18] Furthermore, the discovery of oncogenic ERG overexpression in prostate cancer has already set off pre-clinical development of specific ERG inhibitors.[51–53] Therefore, it would be intriguing to explore the utility of ERG inhibitors in treating AML with *ERG* amplification.

Acknowledgments: None

Funding information

This study was supported by the Department of Pathology, University of Michigan, Ann Arbor, MI.

Conflict of interest Disclosures
The authors disclose no conflict of interest.

Author's contribution
WL and LS designed the research study, analyzed, interpreted the data, and prepared the manuscript. WL, LS, EG, HX, MC, and DM performed data acquisition and analysis. DB interpreted the analysis and revised the manuscript. All authors read and approved the final manuscript.

References

1. Lefebvre V, Bhattaram P. Vertebrate skeletogenesis. Curr Top Dev Biol 2010;90(C):291–317.
2. Shah A V., Birdsey GM, Randi AM. Regulation of endothelial homeostasis, vascular development and angiogenesis by the transcription factor ERG. Vascul Pharmacol 2016;863–13.
3. Dzierzak E, Bigas A. Blood Development: Hematopoietic Stem Cell Dependence and Independence. Cell Stem Cell 2018;22(5):639–651.
4. Knudsen KJ, Rehn M, Hasemann MS, et al. ERG promotes the maintenance of hematopoietic stem cells by restricting their differentiation. Genes Dev 2015;29(18):1915–1929.
5. Ng AP, Loughran SJ, Metcalf D, et al. Erg is required for self-renewal of hematopoietic stem cells during stress hematopoiesis in mice. Blood 2011;118(9):2454–2461.
6. Loughran SJ, Kruse EA, Hacking DF, et al. The transcription factor Erg is essential for definitive hematopoiesis and the function of adult hematopoietic stem cells. Nat Immunol 2008;9(7):810–819.
7. Taoudi S, Bee T, Hilton A, et al. ERG dependence distinguishes developmental control of hematopoietic stem cell maintenance from hematopoietic specification. Genes Dev 2011;25(3):251–262.
8. Ng AP, Coughlan HD, Hediyeh-zadeh S, et al. An Erg-driven transcriptional program controls B cell lymphopoiesis. Nat Commun;11(1):.
9. Stankiewicz MJ, Crispino JD. ETS2 and ERG promote megakaryopoiesis and synergize with alterations in GATA-1 to immortalize hematopoietic progenitor cells. Blood 2009;113(14):3337–3347.
10. Heuston EF, Keller CA, Lichtenberg J, et al. Establishment of regulatory elements during erythro-megakaryopoiesis identifies hematopoietic lineage-commitment points. Epigenetics and Chromatin 2018;11(1):1–18.
11. Kruse EA, Loughran SJ, Baldwin TM, et al. Dual requirement for the ETS transcription factors Fli-1 and Erg in hematopoietic stem cells and the megakaryocyte lineage. Proc Natl Acad Sci U S A 2009;106(33):13814–13819.
12. Tomlins SA, Laxman B, Varambally S, et al. Role of the TMPRSS2-ERG gene fusion in prostate cancer. Neoplasia 2008;10(2):177–188.
13. Martens JHA. Acute myeloid leukemia: A central role for the ETS factor ERG. Int J Biochem Cell Biol 2011;43(10):1413–1416.
14. Marcucci G, Baldus CD, Ruppert AS, et al. Overexpression of the ETS-related gene, ERG, predicts a worse outcome in acute myeloid leukemia with normal karyotype: A Cancer and Leukemia Group B study. J Clin Oncol 2005;23(36):9234–9242.
15. Metzeler KH, Dufour A, Benthaus T, et al. ERG expression is an independent prognostic factor and allows refined risk stratification in cytogenetically normal acute myeloid leukemia: A comprehensive analysis of ERG, MN1, and BAALC transcript levels using oligonucleotide microarrays. J Clin Oncol 2009;27(30):5031–5038.
16. Baldus CD, Liyanarachchi S, Mrozek K, et al. Acute myeloid leukemia with complex karyotypes and abnormal chromosome 21: Amplification discloses overexpression of APP, ETS2, and ERG genes. Proc Natl Acad Sci 2004;101(11):3915–3920.
17. Weber S, Haferlach C, Jeromin S, et al. Gain of chromosome 21 or amplification of chromosome arm 21q is one mechanism for increased ERG expression in acute myeloid leukemia. Genes, Chromosom Cancer 2016;55(2):148–157.
18. Sotoca AM, Prange KHM, Reijnders B, et al. The oncofusion protein FUS-ERG targets key hematopoietic regulators and modulates the all-trans retinoic acid signaling pathway in t(16;21) acute myeloid leukemia. Oncogene 2016;35(15):1965–1976.
19. Harrison CJ. Blood Spotlight on iAMP21 acute lymphoblastic leukemia (ALL), a high-risk pediatric disease. Blood 2015;125(9):1383–1386.
20. Heerema NA, Carroll AJ, Devidas M, et al. Intrachromosomal amplification of chromosome 21 is associated with inferior outcomes in children with acute lymphoblastic leukemia treated in contemporary standard-risk children's oncology group studies: A report from the children's oncology group. J Clin Oncol 2013;31(27):3397–3402.
21. Strefford JC, Van Delft FW, Robinson HM, et al. Complex genomic alterations and gene expression in acute lymphoblastic leukemia with intrachromosomal amplification of chromosome 21. Proc Natl Acad Sci U S A 2006;103(21):8167–8172.

22. Li Y, Schwab C, Ryan SL, et al. Constitutional and somatic rearrangement of chromosome 21 in acute lymphoblastic leukaemia. Nature 2014;508(1):98–102.
23. Xie W, Xu J, Hu S, et al. iAMP21 in acute myeloid leukemia is associated with complex karyotype, TP53 mutation and dismal outcome. Mod Pathol [Epub ahead of print].
24. Mareschal S, Palau A, Lindberg J, et al. Challenging conventional karyotyping by next-generation karyotyping in 281 intensively treated patients with AML. Blood Adv 2021;5(4):1003–1016.
25. Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H TJ, editor. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues (IARC 2017). Revised 4th. Lyon: International Agency for Research on Cancer; 2017.
26. Papaemmanuil E, Gerstung M, Malcovati L, et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood 2013;122(22):3616–3627.
27. Rücker FG, Schlenk RF, Bullinger L, et al. TP53 alterations in acute myeloid leukemia with complex karyotype correlate with specific copy number alterations, monosomal karyotype, and dismal outcome. Blood 2012;119(9):2114–2121.
28. Haase D, Stevenson KE, Neuberg D, et al. TP53 mutation status divides myelodysplastic syndromes with complex karyotypes into distinct prognostic subgroups. Leukemia 2019;33(7):1747–1758.
29. Bernard E, Nannya Y, Hasserjian RP, et al. Implications of TP53 allelic state for genome stability, clinical presentation and outcomes in myelodysplastic syndromes. Nat Med 2020;26(10):1549–1556.
30. Rücker FG, Schlenk RF, Bullinger L, et al. TP53 alterations in acute myeloid leukemia with complex karyotype correlate with specific copy number alterations, monosomal karyotype, and dismal outcome. Blood 2012;119(9):2114–2121.
31. Rücker FG, Bullinger L, Schwaenen C, et al. Disclosure of candidate genes in acute myeloid leukemia with complex karyotypes using microarray-based molecular characterization. J Clin Oncol 2006;24(24):3887–3894.
32. Mrózek K, Eisfeld AK, Kohlschmidt J, et al. Complex karyotype in de novo acute myeloid leukemia: typical and atypical subtypes differ molecularly and clinically. Leukemia 2019;33(7):1620–1634.
33. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. Sci Signal 2013;6(269):pl1–pl1.
34. Cerami E, Gao J, Dogrusoz U, et al. The cBio Cancer Genomics Portal: An open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012;2(5):401–404.
35. Van Loo P, Nordgard SH, Lingjærde OC, et al. Allele-specific copy number analysis of tumors. Proc Natl Acad Sci U S A 2010;107(39):16910–16915.
36. Stephens PJ, Greenman CD, Fu B, et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell 2011;144(1):27–40.
37. Gel B, Serra E. KaryoploteR: An R/Bioconductor package to plot customizable genomes displaying arbitrary data. Bioinformatics 2017;33(19):3088–3090.
38. Bankhead P, Loughrey MB, Fernández JA, et al. QuPath: Open source software for digital pathology image analysis. Sci Rep 2017;7(1):1–7.
39. Hirsch FR, Varella-Garcia M, Bunn PA, et al. Epidermal Growth Factor Receptor in Non–Small-Cell Lung Carcinomas: Correlation Between Gene Copy Number and Protein Expression and Impact on Prognosis. J Clin Oncol 2003;21(20):3798–3807.
40. John T, Liu G, Tsao M-S. Overview of molecular testing in non-small-cell lung cancer: mutational analysis, gene copy number, protein expression and other biomarkers of EGFR for the prediction of response to tyrosine kinase inhibitors. Oncogene 2009;28(S1):S14–S23.
41. Graubert TA, Shen D, Ding L, et al. Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. Nat Genet 2012;44(1):53–57.
42. Fu L, Fu H, Wu Q, et al. High expression of ETS2 predicts poor prognosis in acute myeloid leukemia and may guide treatment decisions. J Transl Med 2017;15(1):1–9.
43. Albertson DG. Gene amplification in cancer. Trends Genet 2006;22(8):447–455.
44. Verhaak RGW, Bafna V, Mischel PS. Extrachromosomal oncogene amplification in tumour pathogenesis and evolution. Nat Rev Cancer 2019;19(5):283–288.
45. Tanaka H, Watanabe T. Mechanisms Underlying Recurrent Genomic Amplification in Human Cancers. Trends in Cancer 2020;6(6):462–477.

46. Rücker FG, Dolnik A, Blätte TJ, et al. Chromothripsis is linked to TP53 alteration, cell cycle impairment, and dismal outcome in acute myeloid leukemia with complex karyotype. Haematologica 2018;103(1):e17–e20.
47. Bochtler T, Granzow M, Stölzel F, et al. Marker chromosomes can arise from chromothripsis and predict adverse prognosis in acute myeloid leukemia. Blood 2017;129(10):1333–1342.
48. Cortés-Ciriano I, Lee JJK, Xi R, et al. Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing. Nat Genet 2020;52(3):331–341.
49. Moorman A V., Richards SM, Robinson HM, et al. Brief report: Prognosis of children with acute lymphoblastic leukemia (ALL) and intrachromosomal amplification of chromosome 21 (iAMP21). Blood 2007;109(6):2327–2330.
50. Rand V, Parker H, Russell LJ, et al. Genomic characterization implicates iAMP21 as a likely primary genetic event in childhood B-cell precursor acute lymphoblastic leukemia. Blood 2011;117(25):6848–6855.
51. Wang X, Qiao Y, Asangani IA, et al. Development of Peptidomimetic Inhibitors of the ERG Gene Fusion Product in Prostate Cancer. Cancer Cell 2017;31(4):532-548.e7.
52. Mohamed AA, Xavier CP, Sukumar G, et al. Identification of a small molecule that selectively inhibits ERG-positive cancer cell growth. Cancer Res 2018;78(13):3659–3671.
53. Wang S, Kollipara RK, Srivastava N, et al. Ablation of the oncogenic transcription factor ERG by deubiquitinase inhibition in prostate cancer. Proc Natl Acad Sci U S A 2014;111(11):4251–4256.

Figure legends:

Figure 1. Characterization of *ERG* amplification and other regions of chromosome 21 in myeloid neoplasms. (A) Copy number alteration (CNA) of chromosome 21 including *ERG* amplification, as assessed by high-density copy number microarray. ERG amplifications are flanked by complex patterns of CNAs on chromosome 21 and do not always include RUNX1. * denotes cases (Pt 5 and 12) with subthreshold *ERG* amplifications with log R ratio (LRR) <0.2. Examples demonstrating three patterns of ERG amplifications as characterized by metaphase karyotyping (B), and interphase (C) and metaphase (D) FISH analyses. A dual color break-apart probe set flanking the *ERG* gene locus with green centromeric and orange telomeric probes was used in FISH studies. The first pattern, as seen in pt 3, is consistent with *in cis ERG* amplification. Karyotype shows a loss of 5q, 12, 17, and 21, as well as an addition of a large marker chromosome, which harbors *in cis* ERG amplification (~9x) as confirmed with interphase and metaphase. The second patterns, as seen in pt 4, is consistent with *ERG* amplification (8-9x) by supernumerary marker chromosomes. Karyotype demonstrates a loss of 5q, 13, 18, and 22 and multiple small marker chromosomes, each harbor a copy of *ERG*, as confirmed with interphase and metaphase FISH studies. The third pattern, as seen in pt 9, is consistent with *ERG* amplification (~19x) by double minute chromosomes. The karyotype shows a loss of 9, 18, and 20q, and a gain of 21 as well as numerous double minute chromosomes, which harbor *ERG,* as confirmed by interphase and metaphase FISH.

Figure 2. Detection of subclonal *ERG* amplification in myelodysplastic syndrome. Pt 5 and 12 shows subthreshold *ERG* amplifications with LRR <0.2. However, subtle changes in LRR tracings (A), corroborated by loss of heterozygosity as seen on the B-allele frequency (BAF) tracings (B), along chromosome 21 suggest *ERG* amplifications are small subclonal events, which are subsequently confirmed with *ERG* FISH studies.

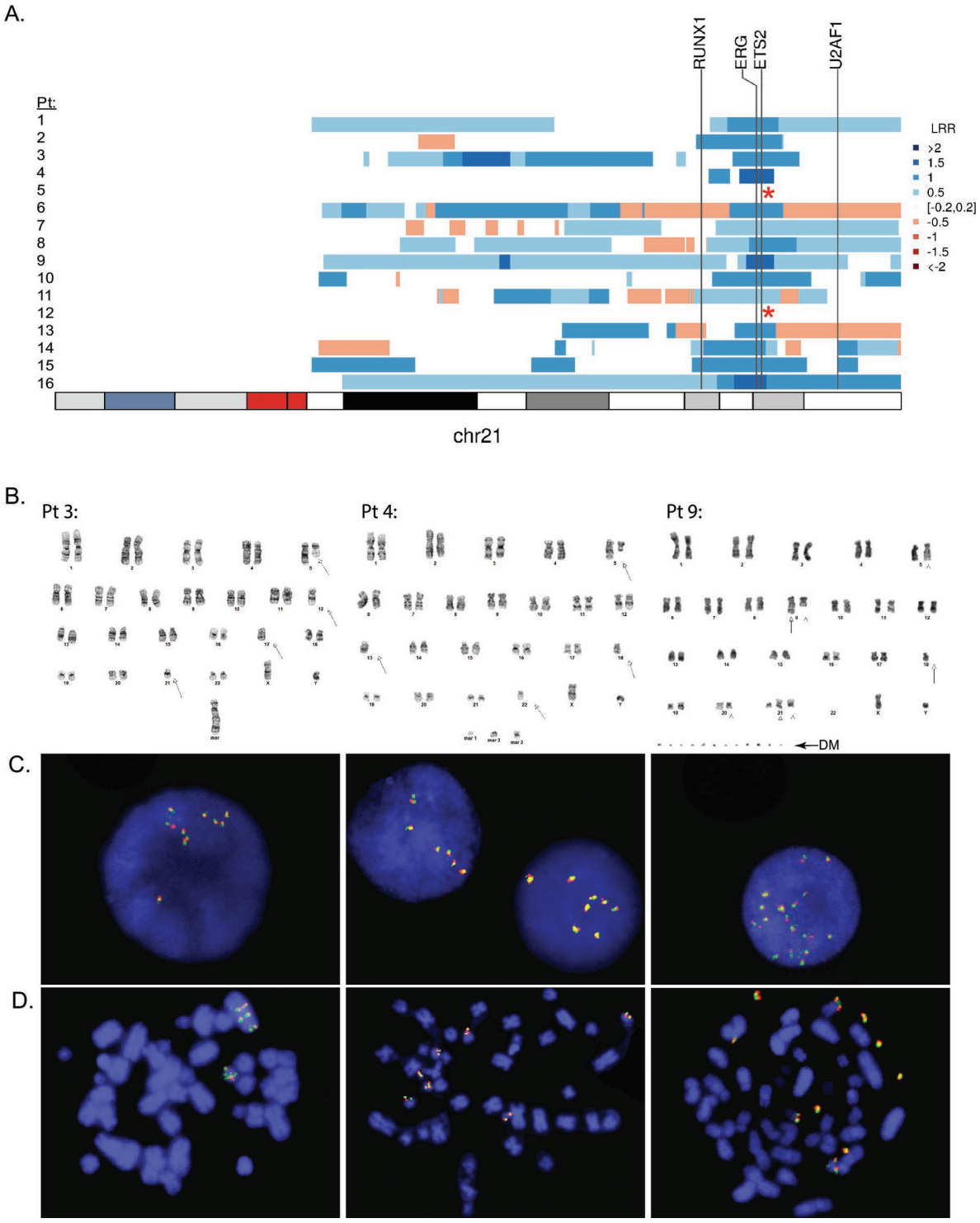
Figure 3. Genome wide copy number alterations in myeloid neoplasms with *ERG* amplification as assessed by copy number arrays. *ER*G amplification is associated with complex large copy number alterations, including frequent loss of chromosome 5q.

Figure 4. *ERG* amplification associated with complex genome-wide copy number alterations is validated by the TCGA cohort of AML. (A) Genome wide patterns of copy number alterations as assessed by SNP6.0 array include frequent chromosome 5q loss. (B) segmental *ERG* amplification on chr21 in the TCGA cohort is accompanied by a similar pattern of flanking CNA changes, as seen in the UM cohort.

Figure 5. ERG protein expression in AML with complex karyotype as assessed by immunohistochemistry (IHC). (A) Examples of intense ERG staining (left panel) of patient 9 with *ERG* amplification by double minute chromosomes, and moderate to weak ERG staining (right panel) of a patient with AML-CK without *ERG* amplification. (B) Computational H-scoring for staining intensity showing red as intense, orange as moderate, yellow as weak, and blue as negative. (C) H-scores of ERG IHC is more intense in AML-CK with *ERG* amplification (ERGAmp; n=11) than AML without ERG amp (NoAmp; n=13) but did not reach statistical significance (p =0.40; Kruskal-Wallis rank sum test)

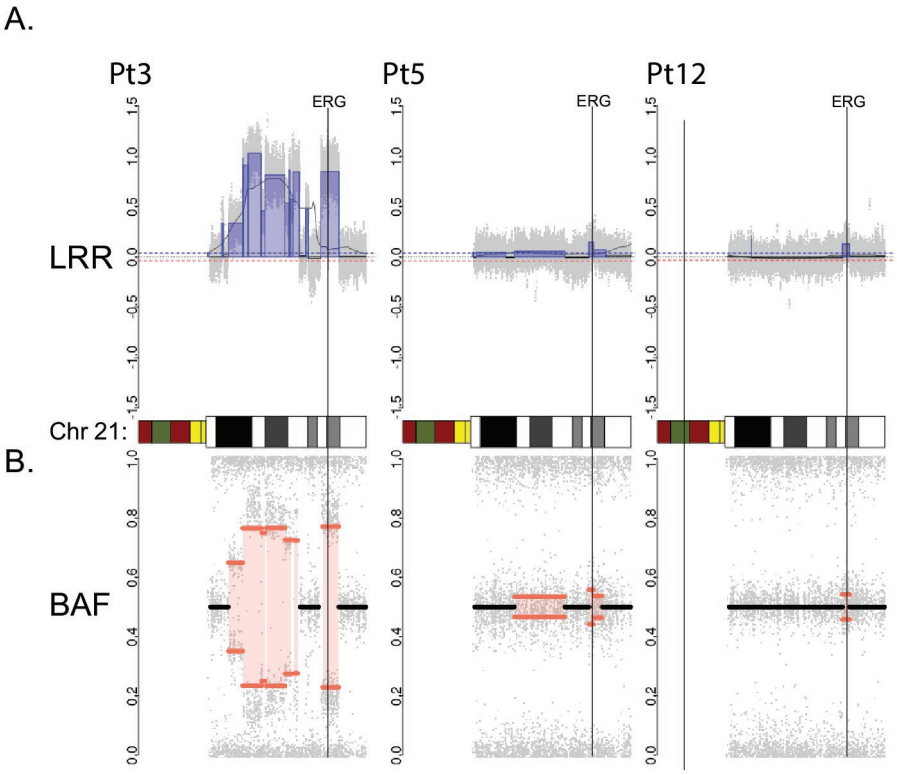
Supplementary figure 1. Overall survival analysis of acute myeloid leukemia with complex karyotype with *ERG* amplification (ERGamp) and without *ERG* amplification (No ERGamp) shows poor survival without significant differences (p =0.4; log rank test).

Figure 1



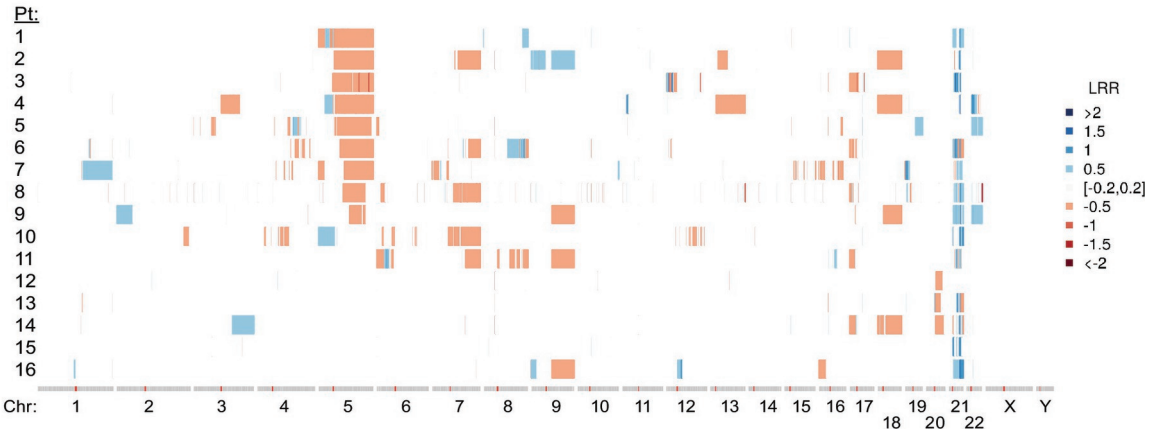
gcc_23027_figure 1v3.eps

Figure 2



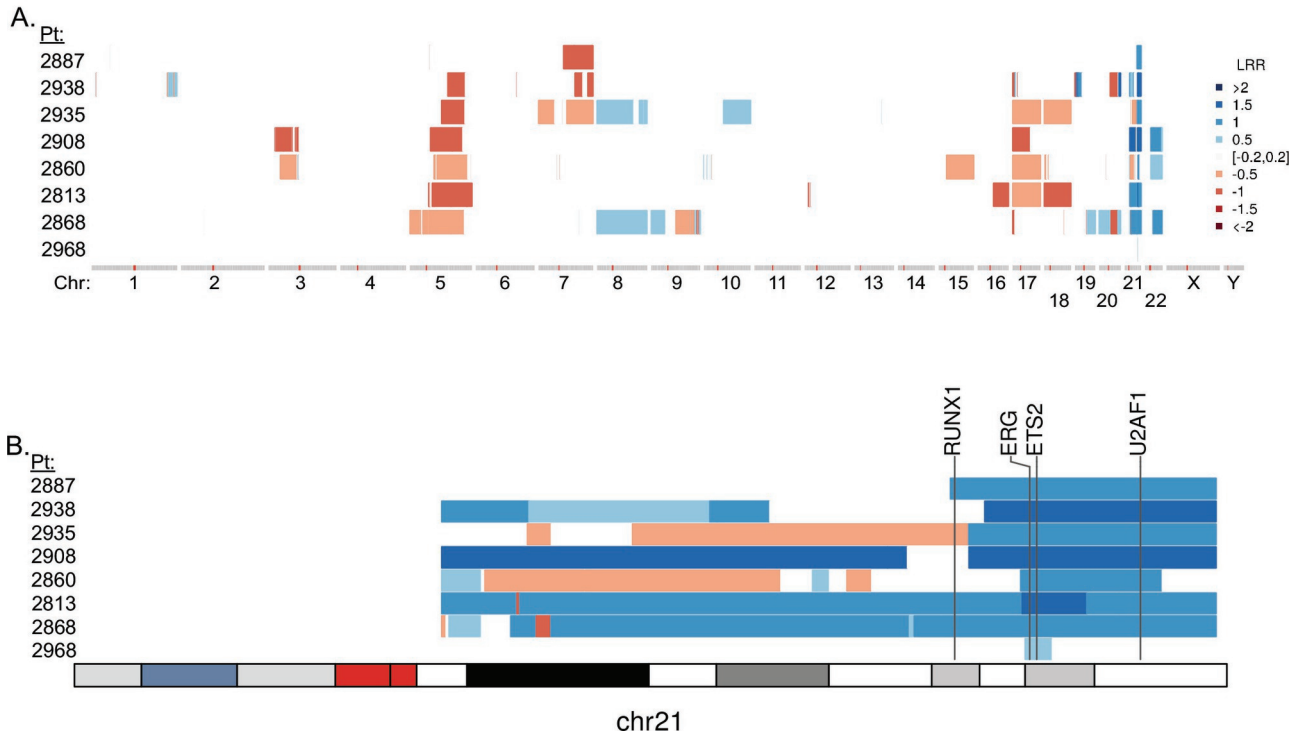
gcc_23027_figure 2v3.eps

Figure 3



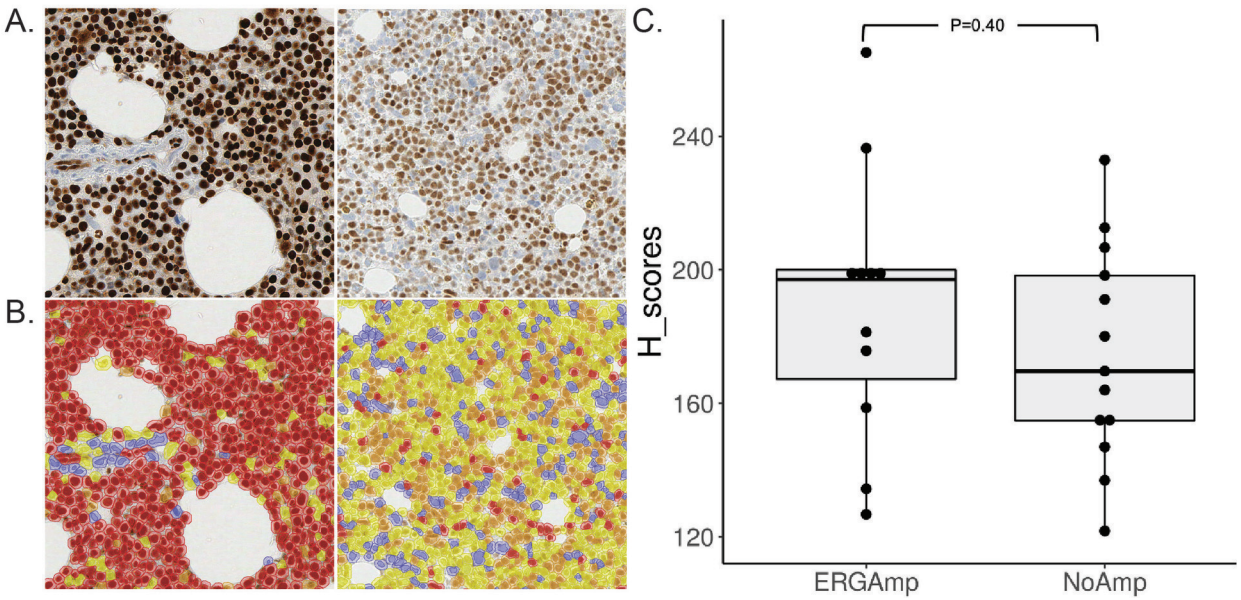
gcc_23027_figure 3v3.eps

Figure 4



gcc_23027_figure 4v3.eps

Figure 5.



GCC_23027_Figure 5v3 copy.jpg

Figure 1

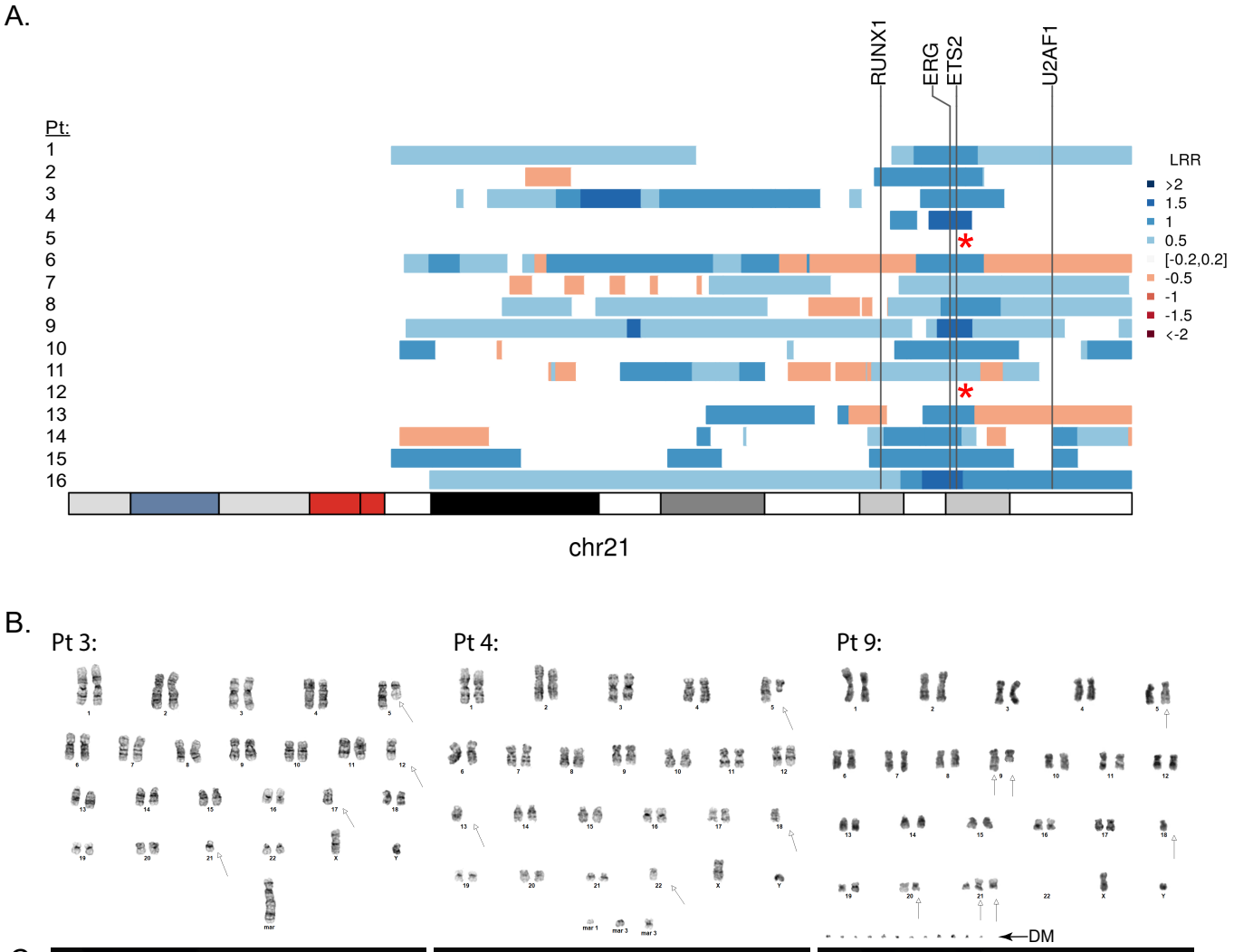


Figure 2

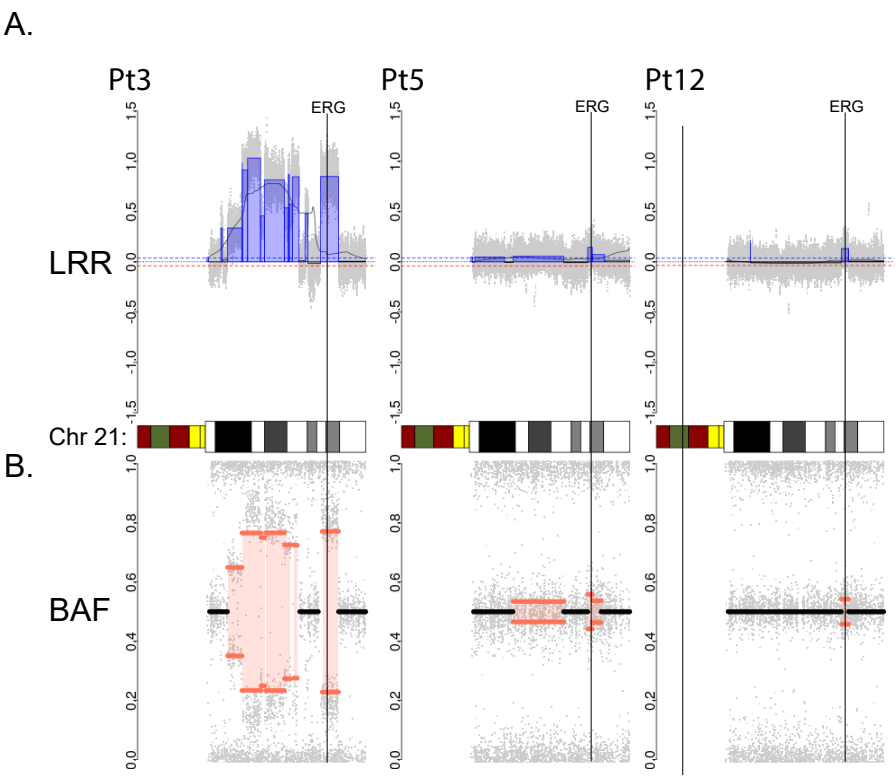


Figure 3

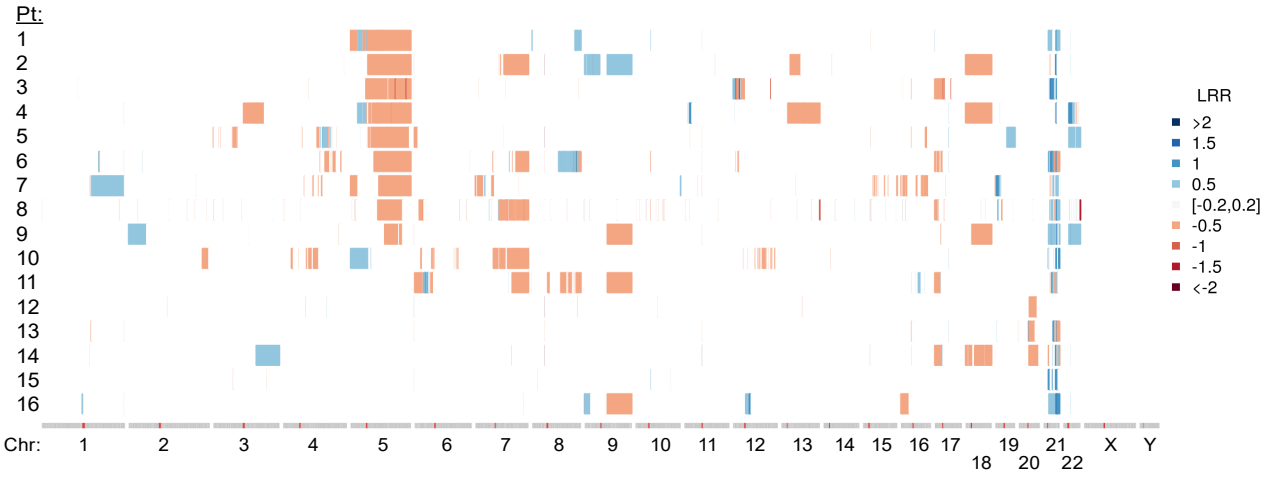


Figure 4

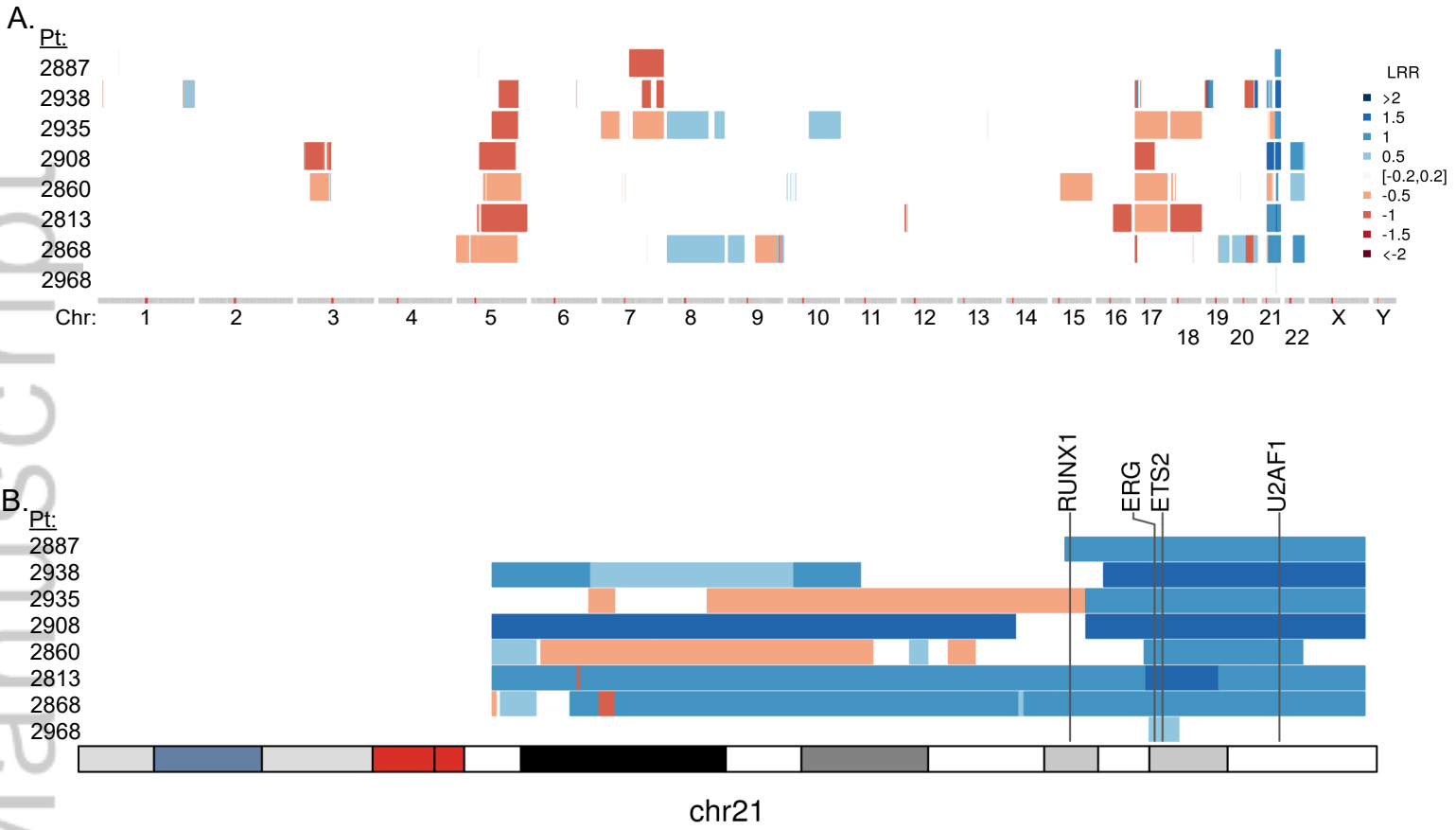


Figure 5.

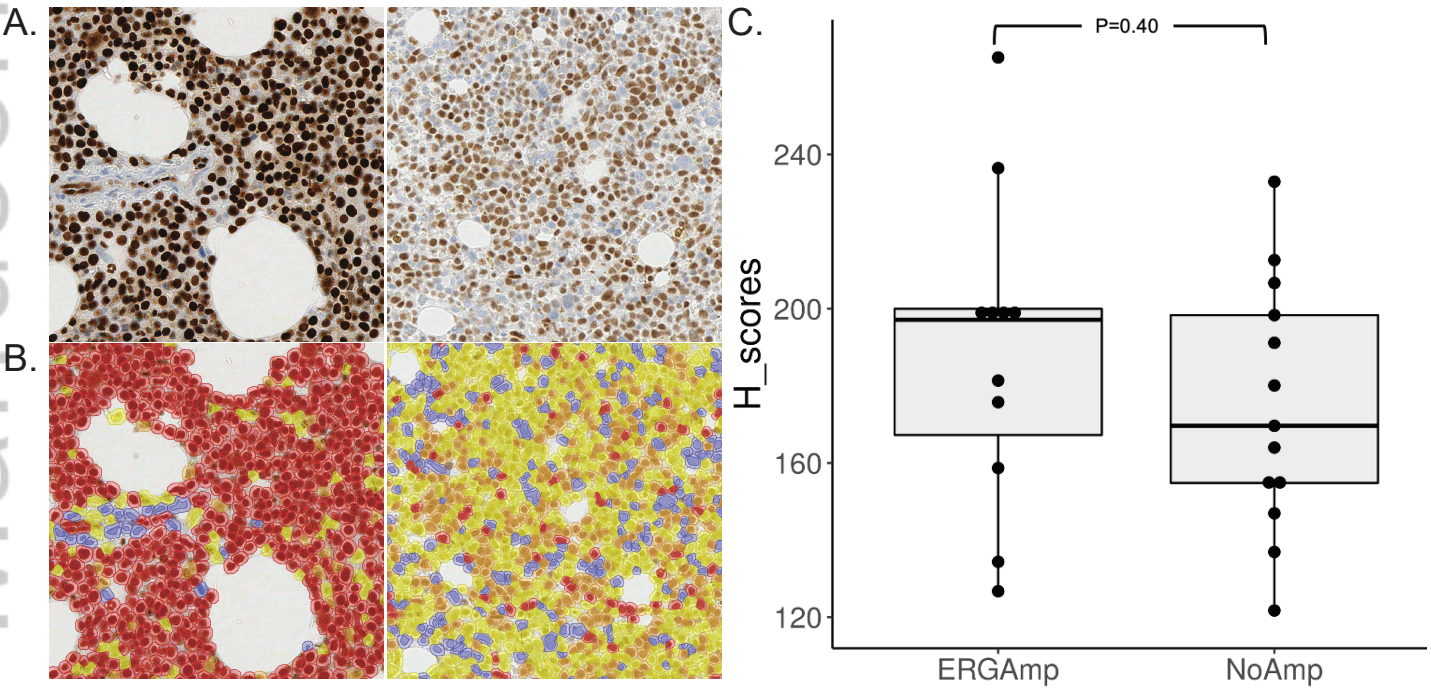


Table 1. Clinical characteristics of patients with myeloid neoplasms with *ERG* amplification

				Prior Malignancy			
Pt	**Age (yrs)**	**Sex**	**Diagnosis**	**Diagnosis**	**Treatment**	**Treatment**	**OS (months)**
1	79	M	t-MDS-EB1	Urothelial carcinoma	cisplatin, etoposide, radiation	Supportive	0.4
2	62	F	AML-MRC	None	None	Supportive	0.1
3	63	M	AML-MRC	None	None	Decitabine/ Venetoclax Bone marrow transplant	>13.1
4	69	M	t-AML	Classic Hodgkin lymphoma	ABVD	FLAG	12.7
5	70	M	MDS-RS-MLD ->AML-MRC	Prostate cancer	Resection only	Azacitidine FLAG	6.3
6	44	F	t-AML	Follicular lymphoma	R-CHOP	FLAG	10.2
7	60	M	t-AML	Multiple myeloma	RVD	FLAG	1.9
8	59	F	AML-MRC	None	None	FLAG	2.4
9	76	M	AML-MRC	None	None	Decitabine/ Venetoclax	4.2
10	52	F	AML-MRC	None	None	7+3	0.2
11	59	M	t-MDS	Pancreatic cancer	5FU, gemcitabine, cisplatin	Supportive	2.1
12	69	M	MDS-MLD	None	None	Unknown	26.1
13	65	F	AML-MRC	None	None	7+3	3.1
14	74	F	AML/ Blast phase	ET, JAK2+	Hydroxyurea and anagrelide	Decitabine	2.6
15	74	F	AML/ Blast phase	ET, JAK2+	Hydroxyurea	Decitabine/ Venetoclax	10.8
16	70	M	AML/ Blast phase	ET, JAK2+;	Hydroxyurea	FLAG, Decitabine/ Venetoclax	2.5

Table 2. Karyotype, array and mutation results in patients with ERG amplification

Pt	Karyotype	Copy number alteration		ERG AMP (% Nuc./ pattern)	TP53		Other pathogenic mutations	Chr Thrps	Pertinent Negatives
		5q	7q		CNA	Mutations			
1	46-47,XY,-5,-21,+2-3mar[cp12]/46-48,sl,+8[5]/ 48-49,sl,+8,+?21,+?21[cp3]	Loss	ND	3-6x(74%)/ Marker chr	LOH	NA	NA	5	CEBPA, FLT3 ITD, FLT3 D835, NPM1, KIT D816V, and IDH 1/2
2	42-45,XX,-5,-7,+9,+11,add(11)(q13),del(13)(q 12q22),add(14)(p11.2),-18,-21,-22,+0-2mar[cp 17]/46,XX[3]	Loss	Loss	3-10x(89%)/ THSR	ND	TP53:p.F134L (35%); TP53:c.673-1G> 1 Splice variant (35%)	NA	ND	KIT816V, NPM1, IDH1/2, FLT3, CEBPA
3	44,XY,inv(2)(p11.2q13)c,der(5;17)(p10;q10),-1 2,-21,+mar[14]/44,sl,del(7)(q31q36)[4]/88,slx2[2]	Loss	Loss	4-12x(88.5%)/ THSR	Loss	TP53:p.R175H (87.2%)	NA	12p, 21	NPM1, CEBPA, FLT3
4	46-48,XY,add(5)(q11.2),-13,-18,-22,+3-5mar[c p3]/46-47,sl,del(3)(q11q25)[cp13]/46,XY[3]	Loss	ND	3-9x(85%)/ Marker chr	ND	NA	NA	22	CEBPA, FLT3 ITD, FLT3 D835, NPM1, KIT D816V, and IDH 1/2
5	44-46,XY,der(2)ins(2;6)(q23;p24p12)add(2)(q 23),der(3)t(3;12)(p24;p13),add(4)(q12),der(5;2 2)(p10;q10),-6,der(12)t(3;12)t(?6;12)(q12;q24) ,der(19)dup(19)(q13.1q13.4)add(19)(q13.4),+ mar[cp17]/43-44,sl,add(X)(p11.2),-der(2)ins(2; 6)add(2),+add(2)(q32),-der(3)t(3;12),+3,+6,-7,- der(12)t(3;12)t(6;12),+add(12)(p13),add(19)(p 13),add(?21)(p11.2),+del(?22)(q11.2q13),-mar [cp3]	Loss	ND	3-6x(16%)/ Marker chr	LOH	NA	NA	4	IDH1/2
6	44-45,XX,add(1)(q21),add(4)(q28),der(5)t(1;5) (q25;q13),-7,add(8)(q24),der(13)t(13;17)(p11. 2;q12),-17,i(21)(q10),+0-1mar[cp6]/44,sl,del(1 2)(p11.2p13)[cp12]/46,XX[2]	Loss	Loss	3-8x(81%)[#]/ NA	Loss	NA	NA	21	NA
7	42-45,X,-Y,del(4)(q13q25),-5,add(7)(p11.2),t(1 3;13)(q12;q14),-15,-16,-16,add(19)(p13),-21,d er(21)t(1;21)(q12;p11.2)ins(21;?)(p11.2;?),+1- 4mar[cp11]/41-42,sl,add(3)(q27),inv(9)(p21q3 3),add(11)(p11.2),add(14)(p11.2),-17,del(20)(q11.2q13.1)[cp9]	Loss	ND	3-6x(68%)/ Marker chr	ND	NA	NA	15,16, 21	CEBBA, FLT3 ITD, FLT3 D835, NPM1, KIT D816V, and IDH 1/2
8	44-46,XX,del(5)(q13q33),del(6)(p23p25),der(7 ;17)(p10;q10), -19,-21,+1-3mar[cp20]	Loss	Loss	5x*/ NA	Loss	TP53:p.Y220C (69.4%)	NF1:p.I679fs(10.8%)	ND	NA

9	44-45,XY,add(5)(q22),add(9)(q12),add(9)(q34),-18,del(20)(q11q13),+21,dic(21;22)(p11.2;p11.2)x2,6~27dmin[cp13]/ 46,XY[7]	Loss	ND	3-33x(91.5%)/ Double minutes	ND	NA	NA	ND	IDH1, IDH2, NPM1, FLT3 or CEBPA
10	45-47,XX,der(2)t(2;4)(q37;q21),-4,+i(5)(p10),add(6)(p11.2),-7,-12,der(15)t(7;15)(p13;p11.2),add(21)(p11.2),+3-4mar[cp16]/92-93,slx2[2]/46,XX[2]	LOH	Loss	7x*/ NA	ND	NA	NA	12	NA
11	42,XY,del(6)(p12p25),-8,add(9)(q13)x2,-16,-21,-22,+3mar[1]/42,sl,add(3)(q29),del(7)(q32q36),-17[cp8]/ 42,sdl,del(4)(q13q35)[cp5]/ 46,XY[6]	ND	Loss	3-6x(57.5%)#/ NA	Loss	NA	NA	6p, 21	NA
12	46,XY,del(20)(q11.2q13.3)[17]/46,XY[3]	ND	ND	3-4x(19%)/ THSR	ND	NA	NA	ND	NA
13	46,XX,del(20)(q11.2q13.3),add(21)(q22)[3]/47,sl,+mar[15]/46,XX[2]	ND	ND	3-7x(32%)/ THSR	ND	ND	BCOR:p.R1217* (36.8%) SRSF2:p.G93_R94del (11%) SMC1A:p.R468Q(48%)	ND	FLT3 and CEBPA
14	44,XX,add(17)(p11.2),-18,der(20)t(6;20)(q22;q13.1),-21[16]/44,sl,del(4)(q21q35)[4]	ND	ND	3-13x(93.5%)/ THSR	Loss	NA	NA	21	NA
15	47-49,XX,ins(3;13)(p25;q13q32),add(5)(q22),del(13)(q12q22),+1-3mar[cp10]/46,XX[10]	ND	ND	3-6x (69.0%)/ Marker chr	ND	ND	ASXL1:p.G646fs(35.2%) JAK2:p.V617F(33.8%)	ND	NA
16	44-45,XY,-9,der(12)t(12;21)(q21;q22),der(21)t(12;21)(q21;q22)dup(12)(q24q21)[cp14]/44-46,sl,add(16)(p11.2)[cp7]	ND	ND	5-10x(92%), >7x(76.5%) THSR	ND	TP53:p.248Q (39%), TP53:p.H179R (39%), TP53:pG245C (5.0%)	JAK2:p.V617F(65%) TET2:p.T624fs(47%) TET2:p.S1023fs(3.6%)	ND	IDH1/2 and FLT3

*ERG amplification is estimated based on cytogenetic array results as there were insufficient remaining materials for FISH studies.
#Pattern of ERG amplification cannot be determined due to the unavailability of metaphase spread for FISH studies.
ChrThrps = Chromothripsis. THSR = Tandem homogenous staining region. Marker chr = marker chromosome. NA= Not available. ND= Not detected.

Table 3. Genetic results and survival in TCGA cohort

Pt	Diagnosis	OS (Months)	Karyotype	5q	7q	TP53		Other pathogenic mutations
						CNA	Mutations	
2813	AML	1.0	44-45,X, -Y, -5,add(16)(q22), -17, -18,iso(21),+mars[cp5]/82-84,XX, -Y, -3, -4, -11, -12, -19, -21,+21[cp5]	Loss	Normal	Loss	P53:p.C176Y	ND
2860	AML	14.0	44,XX,t(4;11)(q21;q23), -5, -7,add(12)(q24),add(18)(q23),del(20)(q12)[3]/43,XX,del(3)(p12),der(3)t(3;3)(p21;q2?7),psudic(5;7)(q13;p22), -10, -15, -17,add(18)(p11.2), -21,+22,+mar[cp14]/46,XX[3]	Loss	Normal	Loss	P53.p.M40Lfs*7	ND
2868	AML	5.0	46,XY, -5,+8,del(9)(q22),add(10)(q26),der(15;19)(q10;q10),add(17)(p11.2), -20, -21,add(21)(p11),add(22)(q13),+3mar[20]	Loss	Normal	Loss	P53:X126_splice	PRDM1:p.I99F
2887	AML	NA	46,XX,del(7)(q11.2)[20]	Normal	Loss	Normal	ND	IDH1:p.R132C, NRAS:p.G12D
2908	AML	1.0	46~49,XY,del(3)(p14),del(5)(p11.2q33),del(17)(q21q21),add(21)(p11.2),+22,mar[cp20]	Loss	Normal	Loss	P53:p.C141W, P53:p.Q317*,	DNMT3A:p.R882C, TET2:p.R1216*
2935	AML	2.0	44~47,XY,del(5)(q22q35)[20], -7[14], -8[6],der(12)t(10;12)(p11.2q21)[2],add(14)(p12)[11], -17[13],der(17)t(10;17)(q11.2;p13)[14], -18[7],add(18)(p11.2)[7], -21[10],i(21)(q10)[4], -22[4],+mar[10],+mar1x2[6][cp20]	Loss	Loss	Loss	P53:pR248Q	NF1:p.R1241*
2938	AML	10.0	45,X,-Y[3]/46,XY [17]	Loss	Loss	Amp	P53:p.H179R, P53:p.R342Efs*3	DNMT3A:p.M315*, MSH3:p.X752_splice, RASA1:p.X830_splice
2968	AML	15.1	46,XY[20]	Normal	Normal	Normal	ND	U2AF1:p.S34F, NRAS:p.Q61H, DNMT3A:p.E447*, BCOR:p.K839Qfs*5