## SUPPLEMENTAL FIGURE LEGENDS

**Supplemental Figure 1:** Cumulative proportion of patients with platelet and neutrophil recovery in patients with at least one of the $D^{R882}TAS$ mutations compared with those without any $D^{R882}TAS$ mutations. P values are calculated using the log-rank test.

**Supplemental Figure 2:** Cumulative proportion of patients with platelet and neutrophil recovery in patients with or without mutations in *NPM1* (A), *NRAS* (B), *KRAS* (C), and *FLT3* (D). P values are calculated using the log-rank test.

**Supplemental Figure 3:** The variant allele frequency (VAF) of *DNMT3A* R882 and non-R882 mutations during remission. Each data point represents an individual *DNMT3A* mutation identified at the time of diagnosis. The horizontal dash line indicates the cutoff used to distinguish between "high" versus "low" VAF. P value is calculated using the Mann Whitney test.

**Supplemental Figure 4:** Cumulative proportion of patients with platelet ($\geq 50x10^9/L$) and neutrophil recovery ($\geq 0.5x10^9/L$) over time in patients with or without measurable/minimal residual disease (MRD) as detected by flow cytometry. P values are calculated using the log-rank test.

## SUPPLEMENTAL METHODS

### Treatment regimens

Patients were treated with one of the following regimens: 1) "3+7" (daunorubicin 60mg/m$^2$/day x 3 days and cytarabine 100mg [age ≥ 60] or 200mg [age < 60]/m$^2$/day by continuous infusion x 7 days), 2) "FLAG-IDA" (G-CSF 300µg/day SC x 6 days, idarubicin 10mg/m$^2$/day x 3 days, fludarabine 30mg/m$^2$/day x 5 days and cytarabine 2g/m$^2$/day over 4 hours x 5 days), or 3) "MEC" (mitoxantrone 10mg/m$^2$/day x 5 days, etoposide 100mg/m$^2$/day x 5 days and cytarabine 1g [age ≥ 60] or 1.5g [age < 60]/m$^2$/day by continuous infusion x 4 days). At our institution, FLAG-IDA is the preferred frontline chemotherapy regimen for newly diagnosed patients with secondary AML and/or unfavorable cytogenetics (if known prior to the initiation of treatment) to maximize CR rates after 1 cycle of induction. Due to the small number of patients treated with MEC (n = 5), they were combined with those treated with FLAG-IDA into a single group for analysis.

### AML classification

AML patients were classified based on cytogenetic abnormalities. Cytogenetic abnormalities that involved t(16;16)(p13.1;q22), inv(16)(p13.1;q22), or t(8;21) (q22;q22.1) with or without other cytogenetic abnormalities were considered favorable risk. Monosomies or deletions of chromosomes 5, 7, 17 [-5, -7, -17, del(5q), del(7q), abn(17p)]; abnormalities of the long arm of chromosome 3(q21;q26), t(6;9)(p23;q34), or t(9;22)(q34;q11); abnormalities involving the long arm of chromosome 11 (abn 11q23) [with the exception of t(9;11)(p21.3;q23.3) which was considered intermediate risk]; monosomy karyotype; or complex cytogenetic abnormalities

(defined as three or more unrelated chromosomal abnormalities) were considered unfavorable risk. Other cytogenetic abnormalities and normal cytogenetics were designated intermediate risk. Cytogenetic analysis was unsuccessful in 20 patients and not performed for one patient. AML with myelodysplasia-related changes (AML-MRC) was defined as per the 2016 revision to the World Health Organization classification of acute leukemia[1].

## Sequencing analysis of diagnostic samples

Analysis of diagnostic samples was performed using the TruSight Myeloid Sequencing Panel (Illumnia, San Diego, CA) on the MiSeq platform (Illumina), as previously described[2,3]. Fifty-four genes were profiled (39 genes tested for hotspot regions; 15 genes tested for complete exonic regions, **Supplementary Table 4**) using amplicon-based library preparation and (2x250bp) paired-end sequencing using 50 ng of input DNA. Sequence data were analyzed by the NextGENe (v.2.3.1, SoftGenetics) and MiSeq Reporter (v2.4.60, MSR, Illumina) software packages. Data files from each sample were uploaded into Bench Lab NGS v4.2 (Agilent Technologies; Santa Clara, CA) for subsequent filtering to prioritize for reporting those variants that passed all MSR quality criteria including depth of coverage of at least 100x and a variant allele frequency (VAF) threshold of >15%. Well documented hotspots that were detected at VAF <15% were verified by an orthogonal method (Sanger sequencing when VAF was between 10-15%, and ddPCR when VAF was between 2-10%). Variants with a global population minor allele frequency (MAF) >1% according to population databases (1000 Genomes Phase 1 release v3.20101123 and Phase 3 release v5.20130502, ESP6500 [Variants in the Exome Sequencing Project ESP6500SI-V2 data set of the exome sequencing project, annotated with SeattleSeqAnnotation 137], Exome

Aggregation Consortium, release 0.3, dbSNP build 147) and/or present in the Advanced Molecular Diagnostic Laboratory internal database of recurring variants were excluded. All reported variants based on the above-described pipeline were considered putative oncogenic mutations and included for downstream analysis.

**Sequencing analysis of remission samples**

Peripheral blood samples were collected during remission in PAXgene Blood DNA Tubes (PreAnalytiX, Hombrechtikon, Switzerland). DNA was extracted according to manufacturer's instructions. Analysis of DNA samples to detect persistent mutations was performed using error-corrected hybrid capture sequencing with a custom panel (Nimblegen, Roche, Basel, Switzerland) consisting of 267 genomic regions from 37 genes (11 with full exon coverage) with a genomic footprint of 83,397 bp (**Supplementary Table 5**). The relatively small genomic footprint of this panel allowed for economic sequencing at high coverage. Illumina-compatible libraries were constructed from 100 ng of sheared peripheral blood leukocyte genomic DNA using the Covaris M220 sonicator (Covaris, Woburn, MA, USA) and the KAPA HyperPrep Kit (#KK8504, Kapa Biosystems, Wilmington, MA, USA). Following end repair and A-tailing, adapter ligation was performed for 16 hours at 4ºC using 100-fold molar excess of adapters. Agencourt AMPure XP beads (Beckman-Coulter) were used for library clean up, and ligated fragments were amplified by PCR for 6 cycles using 0.5 µM universal and indexed primers. Four libraries were pooled together in a single capture hybridization following the Nimblegen protocol. Following hybrid capture at 47ºC for 72 hours, the captured DNA fragments were enriched with 12 cycles of PCR. Paired-end 2x125 bp sequencing was performed on an Illumina HiSeq 2500 instrument with 8

libraries multiplexed into each lane to achieve mean on-target coverage of ~60 000X. FASTQ files were mapped to the human reference genome hg19 using BWA (v 0.7.12), processed using the Genome Analysis ToolKit (GATK) IndelRealigner and Base Quality Score Recalibration (BQSR) (v 3.4-46), and sorted by genome position and indexed using SAMtools (v 1.3)[4-6]. To assist with accurate mutation calls at low variant allele frequencies, we employed an error suppression technique involving duplex unique molecular identifiers (UMIs) as previously described[7]. For single nucleotide variants, mutation calls were derived from duplex consensus sequences with ≥1 supporting read considered a true variant. For indels, mutation calls were derived from Vardict with standard settings[8]. The lower detection limit of this technique was approximately 0.01% VAF.

**MRD analysis of post-induction bone marrow samples by flow cytometry**

A standardized 3-tube 10-colour leukemia panel (**Supplementary Table 6**) was applied to assess leukemia-associated and differences from normal immunophenotypes[9]. The antibody tubes include 3 core markers (CD45, CD34 and CD33) in addition to lineage and maturation markers allowing the detection of cross-lineage antigen expression, asynchronous, and altered antigen expression across the different lineages. Bone marrow samples were processed within 24 hours after collection using a stain, lyse and wash technique as previously described[9]. A minimum of 250,000 events were acquired on a Navios cytometer ™ (Beckman Coulter, Miami, FL, USA). List mode files were analyzed using Kaluza software 1.3 (Beckman Coulter) or Infinicyt Software 1.7 (Cytognos, Salamanca, Spain). Aberrant antigen expression was documented in the diagnostic samples and electronically saved in reference images with the analyzed data files. All subclones

exceeding 10% of the total leukemic cells were monitored in the follow-up samples. Residual disease was calculated as percentage of viable cells, determined by light scatter features, based on a minimum of 50 clustered events. All samples with > 0.1% leukemic events were considered as MRD positive.

## References for Supplemental Methods:


1. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood.* 2016;127(20):2391-2405.
2. Thomas M, Sukhai MA, Zhang T, et al. Integration of Technical, Bioinformatic, and Variant Assessment Approaches in the Validation of a Targeted Next-Generation Sequencing Panel for Myeloid Malignancies. *Arch Pathol Lab Med.* 2017;141(6):759-775.
3. Spiegel JY, McNamara C, Kennedy JA, et al. Impact of genomic alterations on outcomes in myelofibrosis patients undergoing JAK1/2 inhibitor therapy. *Blood Adv.* 2017;1(20):1729-1738.
4. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25(16):2078-2079.
5. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25(14):1754-1760.
6. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011;43(5):491-498.
7. Newman AM, Lovejoy AF, Klass DM, et al. Integrated digital error suppression for improved detection of circulating tumor DNA. *Nat Biotechnol.* 2016;34(5):547-555.
8. Lai Z, Markovets A, Ahdesmaki M, et al. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. *Nucleic Acids Res.* 2016;44(11):e108.
9. Jafari K, Tierens A, Rajab A, Musani R, Schuh A, Porwit A. Visualization of Cell Composition and Maturation in the Bone Marrow Using 10-Color Flow Cytometry and Radar Plots. *Cytometry B Clin Cytom.* 2018;94(2):219-229.

**Supplemental Table 1:** Patient, disease, and treatment characteristics of the study cohort.

| **Characteristics** | **N = 323** |
|---|---|
| **Age (years), median (range)** | 58 (18 – 82) |
| **Gender, N (%)** | |
| Male | 174 (54) |
| Female | 149 (46) |
| **AML classification, N (%)** | |
| *De novo* | 272 (84) |
| AHD | 27 (8) |
| Therapy-related | 24 (7) |
| **AML with MRC, N (%)** | |
| No | 255 (79) |
| Yes | 68 (21) |
| **Cytogenetic risk group, N (%)** | |
| Favorable | 34 (11) |
| Intermediate | 198 (61) |
| Unfavorable | 61 (19) |
| Unsuccessful | 30 (9) |
| **Chemotherapy regimen, N (%)** | |
| 3+7 | 255 (79) |
| FLAG-IDA or MEC | 68 (21) |
| **CBC values at time of diagnosis, median (range)** | |
| Blasts, x $10^9$/L | 1.5 (0 - 414.8) |
| WBC count, x $10^9$/L | 8.1 (0.3 - 423.3) |
| ANC, x $10^9$/L | 1.1 (0 - 70.9) |
| Hemoglobin, g/L | 91 (43 – 175) |
| MCV, fL | 95.9 (57.6 - 136.8) |
| RDW, % | 16.3 (12.9 - 36.3) |
| Platelets, x $10^9$/L | 60 (4 - 1052) |

Abbreviations: AHD, antecedent hematologic disorder; MRC, myelodysplasia-related changes; CBC, complete blood count; WBC, white blood cell; ANC, absolute neutrophil count; MCV, mean corpuscular volume; RDW, red blood cell distribution width.

**Supplemental Table 2:** Multivariable Cox regression analysis of the association between $D^{R822}TAS$ mutations at diagnosis and time to count recovery

| Variable | Platelet recovery | | | Neutrophil recovery | | |
|---|---|---|---|---|---|---|
| | P-value | HR | 95% CI | P-value | HR | 95% CI |
| **Age** | 0.5963 | 1.0024 | 0.9936 to 1.0112 | 0.2553 | 1.0050 | 0.9964 to 1.0137 |
| **AML type** | | | | | | |
| *De novo* | - | - | - | - | - | - |
| AHD | 0.0666 | 0.6203 | 0.3724 to 1.0333 | 0.2244 | 0.7309 | 0.4407 to 1.2121 |
| Therapy-related | 0.1093 | 0.6322 | 0.3606 to 1.1082 | 0.5242 | 0.8316 | 0.4715 to 1.4667 |
| **AML with MRC** | | | | | | |
| No | - | - | - | - | - | - |
| Yes | **0.0018** | 0.5973 | 0.4323 to 0.8254 | 0.0719 | 0.7450 | 0.5407 to 1.0266 |
| **Cytogenetic risk group** | | | | | | |
| Favorable | - | - | - | - | - | - |
| Intermediate | 0.9549 | 0.9890 | 0.6744 to 1.4504 | 0.4816 | 0.8726 | 0.5971 to 1.2753 |
| Unfavorable | 0.6036 | 0.8810 | 0.5460 to 1.4213 | 0.4416 | 1.2120 | 0.7427 to 1.9778 |
| Unsuccessful | 0.0562 | 0.5958 | 0.3501 to 1.0138 | 0.2105 | 0.7220 | 0.4336 to 1.2022 |
| **Chemotherapy regimen** | | | | | | |
| 3+7 | - | - | - | - | - | - |
| FLAG-IDA or MEC | 0.2639 | 0.7793 | 0.5032 to 1.2070 | 0.1338 | 0.6997 | 0.4387 to 1.1159 |
| **$D^{R882}TAS$ mutation at diagnosis** | | | | | | |
| None | **-** | - | - | - | - | - |
| At least one | **<0.0001** | 0.5618 | 0.4361 to 0.7237 | **0.0048** | 0.7015 | 0.5482 to 0.8977 |

Bold font indicates significance.

Abbreviations: HR, hazard ratio; CI, confidence interval; AHD, antecedent hematologic disorder; MRC, myelodysplasia-related changes.

**Supplemental Table 3:** Multivariable Cox regression analysis of the association between the persistence of mutations in remission and time to count recovery

| Variable | Platelet recovery | | | Neutrophil recovery | | |
|---|---|---|---|---|---|---|
| | P-value | HR | 95% CI | P-value | HR | 95% CI |
| **Age** | 0.0704 | 1.0175 | 0.9986 to 1.0368 | 0.5554 | 1.0054 | 0.9877 to 1.0233 |
| **AML type** | | | | | | |
| *De novo* | - | - | - | - | - | - |
| AHD or therapy-related | 0.8366 | 1.0876 | 0.4897 to 2.4154 | 0.4970 | 1.3231 | 0.5898 to 2.9680 |
| **AML with MRC** | | | | | | |
| No | - | - | - | - | - | - |
| Yes | 0.9335 | 1.0334 | 0.4771 to 2.2384 | 0.7624 | 0.8789 | 0.3807 to 2.0290 |
| **Cytogenetic risk group** | | | | | | |
| Favorable | - | - | - | - | - | - |
| Intermediate | 0.5420 | 1.2743 | 0.5847 to 2.7769 | 0.4246 | 1.3774 | 0.6277 to 3.0229 |
| Unfavorable | 0.4501 | 0.6527 | 0.2158 to 1.9747 | 0.2638 | 1.9667 | 0.6006 to 6.4408 |
| Unsuccessful | 0.3567 | 0.5906 | 0.1928 to 1.8096 | 0.4203 | 1.5455 | 0.5361 to 4.4553 |
| **Chemotherapy regimen** | | | | | | |
| 3+7 | - | - | - | - | - | - |
| FLAG-IDA or MEC | 0.2754 | 0.6140 | 0.2556 to 1.4752 | 0.1756 | 0.5065 | 0.1893 to 1.3553 |
| **MRD by flow cytometry** | | | | | | |
| Negative | **-** | - | - | **-** | - | - |
| Positive | 0.8050 | 1.0888 | 0.5542 to 2.1392 | 0.6057 | 1.1873 | 0.6188 to 2.2780 |
| **Mutation persistence** | | | | | | |
| No persistence | **-** | - | - | **-** | - | - |
| Persistence VAF < 20% | 0.6277 | 0.8713 | 0.4994 to 1.5203 | 0.5942 | 0.8596 | 0.4928 to 1.4995 |
| Persistence VAF > 20% | **0.0212** | 0.3597 | 0.1508 to 0.8582 | **0.0248** | 0.4015 | 0.1810 to 0.8907 |

Bold font indicates significance.

Abbreviations: HR, hazard ratio; CI, confidence interval; AHD, antecedent hematologic disorder; MRC, myelodysplasia-related changes; MRD, minimal/measurable residual disease; VAF, variant allele frequency.

**Supplemental Table 4:** List of genes and regions sequenced by the TruSight Myeloid Sequencing Panel.

| Gene Target | Region (exon) | Gene Target | Region (exon) | Gene Target | Region (exon) | Gene Target | Region (exon) |
|---|---|---|---|---|---|---|---|
| *ABL1* | 4-6 | *DNMT3A* | full | *KDM6A* | full | *RAD21* | full |
| *ASXL1* | 12 | *ETV6* | full | *KIT* | 2, 8-11, 13, 17 | *RUNX1* | full |
| *ATRX* | 8-10, 17-31 | *EZH2* | full | *KRAS* | 2, 3 | *SETBP1* | 4 (partial) |
| *BCOR* | full | *FBZW7* | 9 - 11 | *MLL* | 5 - 8 | *SF3B1* | 13-16 |
| *BCORL1* | full | *FLT3* | 14, 15, 20 | *MPL* | 10 | *SMC1A* | 2, 11, 16, 17 |
| *BRAF* | 15 | *GATA1* | 2 | *MYD88* | 3, 4, 5 | *SMC3* | 10, 13, 19, 23, 25, 28 |
| *CALR* | 9 | *GATA2* | 4 - 6 | *NOTCH1* | 26-28, 34 | *SRSF2* | 1 |
| *CBL* | 8, 9 | *GNAS* | 8, 9 | *NPM1* | 12 | *STAG2* | full |
| *CBLB* | 9, 10 | *HRAS* | 2, 3 | *NRAS* | 2, 3 | *TET2* | 3-11 |
| *CBLC* | 9, 10 | *IDH1* | 4 | *PDGFRA* | 12, 14, 18 | *TP53* | 2 - 11 |
| *CDKN2A* | full | *IDH2* | 4 | *PHF6* | full | *U2AF1* | 2, 6 |
| *CEBPA* | full | *IKZF1* | full | *PTEN* | 5, 7 | *WT1* | 7, 9 |
| *CSF3R* | 14-17 | *JAK2* | 12, 14 | *PTPN11* | 3, 13 | *ZRSR2* | full |
| *CUX1* | full | *JAK3* | 13 | | | | |

**Supplemental Table 5:** List of genes and regions sequenced in remission using a custom 37-gene error-corrected NGS platform.

| Gene Target | Region (exon) | Gene Target | Region (exon) | Gene Target | Region (exon) | Gene Target | Region (exon) |
|---|---|---|---|---|---|---|---|
| *ASXL1* | 12 | *FLT3* | 14, 15, 20 | *MLL* | 5 - 8 | *SF3B1* | 13-16 |
| *BCOR* | full | *GATA2* | 2 - 6 | *MYC* | 2 | *SRSF2* | 1 |
| *BRAF* | 15 | *GNAS* | 11 | *NPM1* | 12 | *STAG2* | full |
| *CBL* | 8, 9 | *IDH1* | 4 | *NRAS* | 2, 3 | *TET2* | 3-11 |
| *CCND2* | 5 | *IDH2* | 4 | *PHF6* | full | *TP53* | 2 - 11 |
| *CDKN2A* | 1, 2 | *JAK2* | 12, 14 | *PTEN* | 5, 7 | *U2AF1* | 2, 6 |
| *CEBPA* | Full | *KDM6A* | full | *PTPN11* | 3, 13 | *U2AF2* | 6 |
| *DNMT3A* | Full | *KIT* | 2, 8-11, 13, 17 | *RAD21* | full | *WT1* | 7, 9 |
| *ETV6* | Full | *KRAS* | 2, 3 | *RUNX1* | full | *ZRSR2* | full |
| *EZH2* | Full | | | | | | |

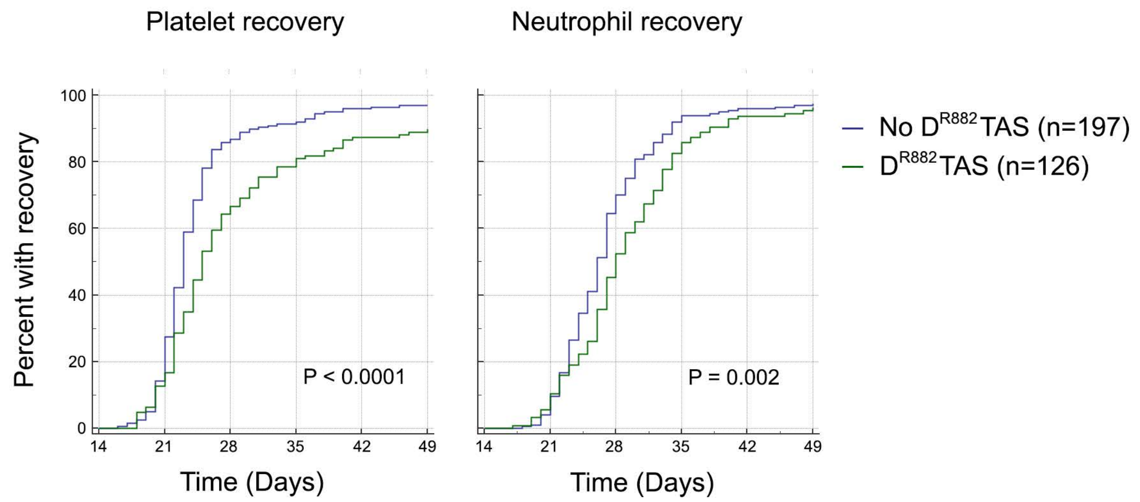
**Supplemental Table 6:** Antibody panels for measurable/minimal residual disease (MRD) detection by flow cytometry

| | FITC | PE | ECD | PC5.5 | PeCy7 | APC | APCAF700 | APCAF750 | PB | KO |
|---|---|---|---|---|---|---|---|---|---|---|
| **Tube 1** | CD65 | CD13 | CD14 | CD33 | CD34 | CD117 | CD7 | CD11b | CD16 | CD45 |
| **Tube 2** | CD36 | CD64 | CD56 | CD33 | CD34 | CD123 | CD19 | CD38 | HLADR | CD45 |
| **Tube 3** | CD71 | CD11c | CD4 | CD33 | CD34 | CD2 | CD10 | CD235a | CD15 | CD45 |

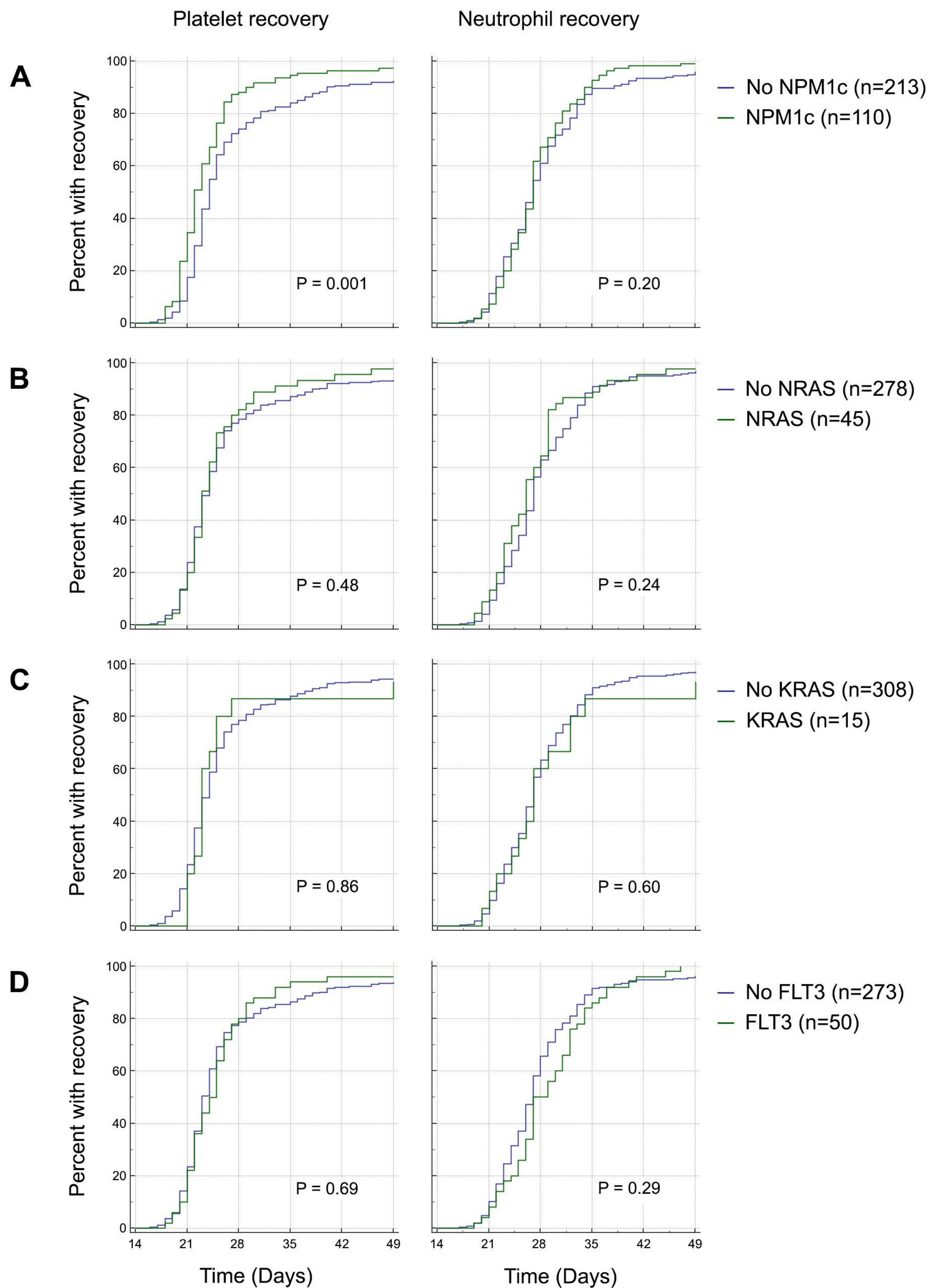
NOTE. All antibodies are from Beckman Coulter (Miami, FL, USA).

Abbreviations: FITC, fluorescein isothiocyanate; PE, Phycoerythrin; ECD, PE-Texas red; PC5.5, PE-Cychrome 5.5; APC, allophycocyanin; APCAF700, APC-Alexafluor 700; APCAF750, APC-Alexafluor 750; PB, Pacific Blue; KO, Krome Orange.

## Supplemental Figure 1



## Supplemental Figure 2



**Supplemental Figure 3**

P = 0.002

Variant Allele Frequency (%)

50
40
30
20
10
0

DNMT3A R882
DNMT3A non-R882

# Supplemental Figure 4