Supplemental Materials and Methods

Immunohistochemical analysis. Immunohistochemistry was performed on human paraffin embedded samples fixed in B5 (Bio-Optica) for 2 hours; bone tissues were decalcified in ethylenediaminetetraacetic acid (Osteodec; Bio-Optica) for 5 to 6 hours. Antigen retrieval was carried out by microwaving in 0.1 mM ethylenediaminetetraacetic acid, pH 8.0. Cytoplasmic nucleophosmin was revealed using a mouse anti-NPM1 monoclonal antibody (mAb, clone 376, generated by B.F.).[1] The antibody/antigen interaction was revealed by the alkaline phosphatase anti-alkaline phosphatase (APAAP) technique.

Western blot and detection of NPM1 mutated protein by western blot analysis. Patients' samples were tested for expression of NPM1 mutant protein by western blotting (WB), as previously described.[2] Briefly, lysates from 1 to 2 x 10^6 cells from either peripheral blood or bone marrow of patients were used for analysis. Lysates from the human leukemic cell lines OCI-AML3, carrying *NPM1* mutation A[3], and *NPM1*-wild type U937 were used as positive and negative control, respectively. The following antibodies were used as indicated: anti-NPM1 mutant rabbit polyclonal antibody (produced by B.F.)[4]; anti-pan NPM1 mouse monoclonal antibody (clone 376) (directed against the N-terminus of NPM1, produced by B.F.); anti-NPM1 wild type mouse monoclonal antibody (directed against the C- terminus of NPM1 wild-type, #32-5200, Invitrogen) (supplemental Figure 1). After incubation with the appropriate secondary antibody, the signal was revealed by enhanced chemoluminescence (Immobilon Crescendo Western HRP substrate, Millipore).

Genomic DNA *NPM1* fragment analysis. Genomic DNA was extracted from the bone marrow aspirates with Maxwell Promega. The DNA was quantified using a ND-1000 apparatus (Nanodrop Technologies, Wilmington, DE, USA). For NPM1 fragments analysis,

genomic DNA was amplified with the NPM-Forward: FAM-5'-AGGACAGCCAGATATCAACTGTTAC-3' and primer NPM-Reverse: 5'-AGTTAACTCTCTGGTGGTAGAATGAAA-3'. Reactions of 25 µl contained 10 ng of genomic DNA, primers (0.8 µM each), deoxynucleoside-5′-triphosphates (0.25 mM, each), MgCl2 (1.5 mM), 6,25% of Dimethyl sulfoxide, 0.4 U of Taq-Gold DNA polymerase, and 10X buffer (AmpliTaq DNA polymerase, Applied Biosystem Inc., Foster City, CA). Cycling conditions for NPM1 were as follows: 1 cycle, 7 minutes at 94°C; 30 cycles, 60 seconds at 95°C, 40 seconds at 55°C, and 90 seconds at 72°C; and 1 cycle, 5 minutes at 72°C. 1 µl Polymerase-chain-reaction (PCR) product were mixed with 9 µl of HiDi formamide (Applied Biosystems, Inc.) and 0.3 µl Liz Size Standard (Applied Biosystems, Inc.) and heated to 95°C for 2 minutes. The samples were run on an ABI 3500 Genetic Analyzer T (Applied Biosystems, Inc.). PCR products and internal standards were detected using filter set D. Raw data were analyzed with GeneMapper v4.0 software (Applied Biosystems, Inc.). The allele WT corresponds to 236 bp while an additional peak of 4 bp identifies the mutated allele.

Automated Sanger sequencing of *NPM1* gene. For NPM1 sequencing analysis, genomic DNA previously extracted was amplified using primers NPM1–F (5′-TTAACTCTCTGG TGGTAGAATGAA-3′) and NPM1–R (5′-CAAGACTATTTGCCATTCCTAAC-3′). The total reactions volume of 50 µl contained 100 ng of genomic DNA, primers (10 pmol each), deoxynucleoside-5′-triphosphates (10 mM, each), MgCl2 (1.5 mM), 2.5 U of Taq-Gold DNA polymerase, and 10X buffer (AmpliTaq DNA polymerase, Applied Biosystem Inc., Foster City, CA). Cycling conditions for NPM1 were as follows: 1 cycle, 5 minutes at 94°C; 40 cycles, 30 seconds at 94°C, 60 seconds at 55°C, and 60 seconds at 72°C; and 1 cycle, 5 minutes at 72°C. The PCR products were purified by enzymatic cleaning (ExoSAP-IT Applied Biosystems) and sequenced directly with primer NPM1-R2 (5′-GGCATTTTTGGACAACACA-3 ′) using the Big Dye Terminator V1.1 Cycle Sequencing Kit

(Applied Biosystems) according to the protocol standard. The Big Dye products were purified by Centers-Sep 96-well plates (Princeton Separations, Inc). For the sequencing reaction is used the ABI 3500 Genetic Analyzer (Applied Biosystems, Inc.) and standard conditions present in the kit protocol used for the marking. The raw data obtained from the sequencer are assembled, edited, analyzed by means of dedicated software (Sequencing Analysis v7.0 Applied Biosystems, Inc) in order to obtain the nucleotide sequence.

Targeted deep sequencing of *NPM1* exon 1 to 12. Genomic DNA was subjected to deep targeted sequencing of *NPM1* gene (TruSeq Custom Panel Custom Panel-ILLUMINA), using 12-25 ng for all samples. Libraries, generated according to the manufacturer's instructions, were sequenced on an Illumina MiSeq instrument for a range of 2x151- 2x251 cycles, using MiSeq Reagent Kit v2 or v3 (median sample coverage: 8074X). Bioinformatics variant calling was performed with MiSeq Reporter software (version 2.5.1) with default settings, and variant annotation was performed with Illumina Variant Studio 2.0. *NPM1* variants were then subjected to the further following filters, and retained only if: i) they passed the default filters; ii) they were present at an allele frequency of at least >5%; iii) they were predicted to change the gene coding sequence or involved the conserved splice-site (i.e., the 4 nucleotides surrounding the exon-intron junction); iv) they were not present in the EVS (Exome Variant Server) of healthy people database at a frequency >1%; v) the total coverage depth at the concerned position was at least 200 reads.

Targeted deep sequencing of myeloid NGS panel. Genomic DNA from patients samples was subjected to deep targeted sequencing of 54 genes frequently mutated in myeloid neoplasms (TruSight Myeloid Sequencing Panel Custom Panel-ILLUMINA), when indicated. Libraries, generated from 50 ng of DNA according to the manufacturer's instructions, were sequenced on an Illumina MiSeq instrument for 2x151cycles, using MiSeq Reagent Kit v3

(median sample coverage: 4911X). Bioinformatics variant calling was performed as described in the previous section (“Targeted deep sequencing of NPM1 exon 1 to 12”). However, *BCORL1* F111L variant was discarded because it is present in the ExAC (Exome Aggregation Consortium) database of healthy people at an allele frequency of 100%.

RNA isolation and reverse-transcription (RT)-PCR.

For sequencing of *NPM1/CCDC28A* fusion transcript in PG pt.4, total RNA was extracted from BM cells with the AllPrep Mini Kit (Qiagen) according to the manufacturer's protocol. RNA was reverse transcribed to generate cDNA for PCR reaction using the Thermoscript thermostable reverse transcriptase (Invitrogen, Thermo Fisher Scientific) following manufacturer’s instructions. Briefly, 300 ng of total RNA were first assembled in a 12 µl reaction with 50 ng random examer primers and heated at 65 °C for 5 min to denature RNA and primers. The reaction was then placed on ice and combined to the remaining reagents (5X buffer, dNTPs, DTT, RNase inhibitor, Thermoscript RT) in a 20 µl final volume. The reaction proceeded for 10 min at 25°C followed by 35 min at 55°C and ended with a 5 min incubation at 85°C. PCR amplification was carried out using primers that amplified 1000 bp fragments. The PCR mix contained 0.75 µL of 0.3 µM primers, 12.5 µL of HiFi ready Master mix (Kapa Biosystems), 2 µL of template cDNA and RNase free water to 25 µL. PCR reaction was performed in a 2720 thermal cycler (Applied Biosystem) with the following program: 95°C for 3 minutes, followed by 30 cycles of 98°C for 20 seconds, 58°C for 15 seconds, 72°C for 15 seconds, final extension at 72°C for 1 min. PCR products were then stored at 4°C. The following primers were used: i) *NPM1/CCDC28A* forward, 5′-CTGAGGCCCCAGAACTATCT-3’; ii) *NPM1/CCDC28A* reverse, 5′-GTCGCTGTGAAGTTGACTGT-3’. The appropriate volume of loading dye (6X Thermo Fisher Scientific) was added to PCR products, then loaded to 1% agarose gel (TopVision Agarose, Thermo Fisher Scientific) in TAE 1X buffer. The PCR products were identified by

using a DNA size marker (1 Kb Plus Ladder, Thermo Fisher Scientific) (supplemental Figure 7). Gel electrophoresis was analysed by UVidoc Imager (Uvitec). PCR products were then purified using the QIAquick PCR Purification kit (Quiagen), quantified with a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific) and outsourced for sequencing (Eurofins Genomics). Sequences were visualized by SnapGene software (not shown).

Cell lines, cell culture and transfection procedures. The OCI-AML3 human AML cell line (DMSZ, Germany) harbouring NPM1 mutation A was previously described.[3] NIH-3T3 murine fibroblasts (ATCC, USA) were cultured in DMEM supplemented with 10% bovine calf serum, 1% glutamine, and 1% antibiotics. For transfection purposes, cells were seeded overnight on glass coverslips and transiently transfected with the indicated plasmid using Lipofectamine 2000 (Invitrogen) following the manufacturer's instructions. After 24h incubation, cells were used for either immunofluorescence or WB analysis.

pEGFP-C1-NPM1 plasmid constructs and cell transfection. The pEGFP-C1-NPM1wt and pEGFP-C1-NPM1mA constructs had been previously generated in our laboratory.[4] These were used as templates to generate the following constructs based on the sequence of the new mutations and translocations identified in the single patients: pEGFP-C1-NPM1_exon 5 (PG pt.1); pEGFP-C1-NPM1_exon 5 (PG pt.2); pEGFP-C1-NPM1_exon 5 (PG pt.3); pEGFP-C1-NPM1_exon 5 (MLL pt.4); pEGFP-C1-NPM1/RPP30 (MLL pt.1); pEGFP-C1-NPM1/SETBP1 (MLL pt.2); pEGFP-C1-NPM1/CCDC28A (MLL pt.3, and PG pt.4). Generation of the specific plasmids and constructs cloning were commissioned to GeneScript (Piscataway, USA) once provided with the plasmid pEGFP-C1-NPM1wt as template and the sequences related to the newly identified mutations/translocations. For ectopic expression of the GFP-NPM1 fusion proteins, the murine fibroblasts cell line NIH-3T3 were transiently transfected with plasmids: p-EGFP-C1 empty vector, p-EGFP-C1-

NPM1wt and p-EGFP-C1-NPM1mutA, as controls, and the p-EGFP-C1-NPM1_new mutants or new fusion transcripts. The subcellular localization of the GFP-NPM1 fusion proteins was analyzed by either fluorescence or confocal microscopy; GFP-NPM1 fusion proteins were also evaluated by Western blotting for expression, molecular weight and reactivity with different anti-NPM1 specific antibodies, or antibodies against the new fusion partner, as indicated (supplemental Figures 3, 4 and 10). Treatment with the specific Crm1/XPO1 inhibitor leptomycin B (20 ng/ml for 5 hours) (Merck Biosciences Ltd.) of NIH-3T3 overexpressing the new GFP-NPM1 fusion proteins was used to evaluate the NES dependence of their subcellular localization, as previously reported.[4]

Fusion plasmids Rev(1.4)–eGFP. To assess their efficiency, the sequence of the individual newly discovered NES has been cloned in the fusion protein Rev(1.4)–eGFP, as previously described.[5] Generation of the specific plasmids and constructs cloning were commissioned to GeneScript (Piscataway, USA) once provided with the plasmid pRev(1.4)**–**eGFP as template and the sequences related to the newly identified NES motifs. The pRev(1.4)**–**eGFP is a construct encoding a mutated form of the HIV Rev protein, fused with eGFP.[6] The wild-type Rev protein is able of nucleus-cytoplasm shuttling due to a NES and a Nuclear Localization Signal (NLS). The mutant protein Rev(1.4) contains a NLS, but lacks the NES, thus it accumulates in the nucleus. The re-introduction of a NES sequence between the pRev(1.4) and the eGFP restores its shuttling activity. Inserting the different newly identified NES sequences in this system allows to compare the efficiency of the different NES. A plasmid expressing the sequence of the original Rev protein NES (pRev(1.4)-NES3-eGFP) was used as positive control for efficient NES.

Here, the introduced sequences were:

i) F, 5'-GTG AAA CTC TTA GCG GAG GAT GTG AAA CTC TTA AGT ATA TCT GGA AAG; R, 3'- CTT TCC AGA TAT ACT TAA GAG TTT CAC ATC CTC CGC TAA GAG TTT

CAC, corresponding to the amino acid sequence *V*KL*L*AED*V*K*L*LS*I*SGK, including the intramolecular NES either ***V***-X_2-***L***-X_3-***V***-X_1-***L*** or **L**-X_3-**V**-X_2-**L**-X_1-**I**, identified for the new NPM1 exon 5 mutant from PG pt.1;

ii) F, 5'-TTA GTA GCT GTG GAG GAA GAT GCA GAG TCA GAA GAT GAA GAG GAG GAG GAT GTG AAA CTC TTA AAA; R, 3'- TTT TAA GAG TTT CAC ATC CTC CTC CTC TTC ATC TTC TGA CTC TGC ATC TTC CTC CAC AGC TAC TAA, corresponding to the amino acid sequence LVAVEEDAESEDEEEEDVKLLK, including the C-terminal part of the new NPM1 exon 5 mutant from PG pt.2;

iii) F, 5'-GTG AAA CTC TTA AGT ATA TCT GGA AAG CTT TCT GCC TTA AGT ATA TCT GGA AAG; R, 3'- CTT TCC AGA TAT ACT TAA GGC AGA AAG CTT TCC AGA TAT ACT TAA GAG TTT CAC, corresponding to the amino acid sequence VKLLS*I*SGK*L*SA*L*S*I*SGK, including the intramolecular NES **I**-X_3-**L**-X_2-**L**-X_1-**I**, identified for the new NPM1 exon 5 mutant from PG pt.3;

iv) F, 5'-GTG CAA CTC TTA AGT GGG CTG CAA CTC TTA AGT ATA TCT GGA AAG; R, 3'-CTT TCC AGA TAT ACT TAA GAG TTG CAG CCC ACT TAA GAG TTG CAC, corresponding to the amino acid sequence VQ*L*SG*L*QL*L*S*I*SGK, including the intramolecular NES either ***L***-X_2-***L***-X_2-***L***-X_1-***I*** or **L**-X_3-**L**-X_2-**L**-X_1-**I**, identified for the new NPM1 exon 5 mutant from MLL pt.4;

v) F, 5'-ATA GAA AAA AAA CTA GAA AAA CTG CTT TTG GAA TTA TCT CTA CAG; R, 3'- CTG TAG AGA TAA TTC CAA AAG CAG TTT TTC TAG TTT TTT TTC TAT, corresponding to the amino acid sequence *I*EKK*L*EK*L*L*L*E*L*S*L*Q, including the C-terminal NES either ***I***-X_3-***L***-X_2-***L***-X_1-***L*** or **L**-X_3-**L**-X_2-**L**-X_1-**L**, identified for the new NPM1/RPP30 fusion protein from MLL pt.1;

vi) F, 5'-ATG CCG GTG CTG GAA AAA TGC ATC GAC CTG CCC AGC AAA AGA; R, 3'- ATG CCG GTG CTG GAA AAA TGC ATC GAC CTG CCC AGC AAA AGA, corresponding to the amino acid sequence M*P*V*L*EKC*I*D*L*PSKR, including the

intramolecular NES **M**-X_2-**L**-X_3-**I**-X_1-**L**, identified for the new NPM1/SETBP1 fusion protein from MLL pt.2;

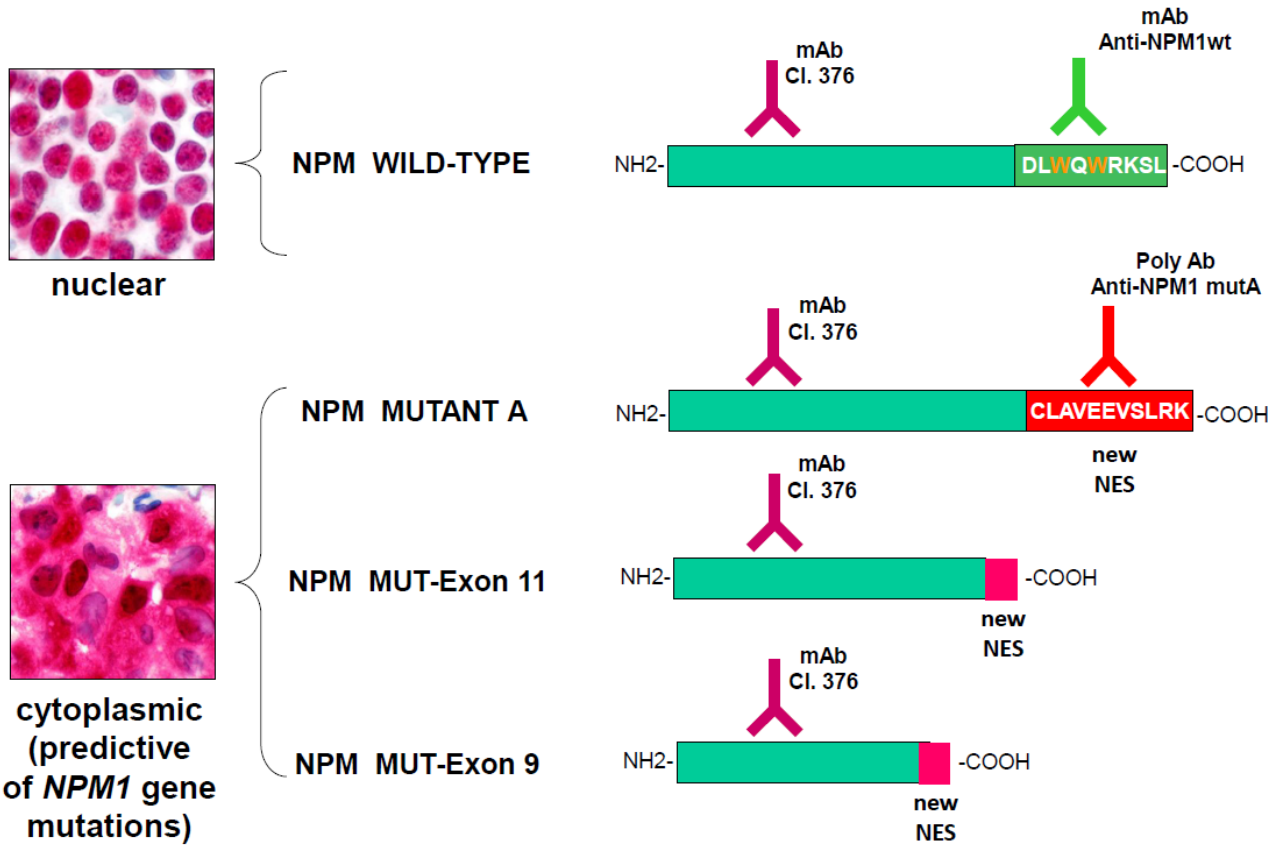
vii) F, 5'-ATG CAG GAG AAA TTA GCT CGC TTG AAT TTG GAG CTC; R, 3'- AGA GCT CCA AAT TCA AGC GAG CTA ATT TCT CCT GCA, corresponding to the amino acid sequence MQEKLARLNLEL, including the first C-terminal NES **M**-X_3-**L**-X_2-**L**-X_1-**L** (1^{st} NES), identified for the new NPM1/CDCC28A fusion protein from MLL pt.3 and PG pt.4; and F, 5'-TTG AAT TCT TCC ATA CAA AAA CTC CAT TTG; R, 3'- CAA ATG GAG TTT TTG TAT GGA AGA ATT CAA, corresponding to the amino acid sequence LNSSIQKLHL, including the second identified C-terminal NES **L**-X_3-**I**-X_2-**L**-X_1-**L** (2^{nd} NES).

Immunofluorescence analysis. For immunofluorescence and confocal microscopy studies, transfected NIH-3T3 cells, grown on glass coverslips, were rinsed in phosphate-buffered saline (PBS) and fixed in 4% paraformaldehyde, pH 7.4 (10 min), air dried and flipped onto standard glass slides with Mowiol mounting medium (Sigma-Aldrich). When indicated, nuclei were stained with 4,6-diamidino-2-phenylindole (DAPI) fluorescent stain in Prolong Gold mounting reagent (Molecular Probes by Life Technologies) following standard procedures. Immunofluorescence images were collected by fluorescence microscope (Olympus IX51 model, Shinjuku, Tokyo, Japan) and processed with CellSens Digital Imaging Software (Olympus). Confocal analysis was done with a Zeiss LSM 800 confocal microscope (Carl Zeiss) using 488-nm (for eGFP) laser line for excitation. In the NES efficiency assay using the pRev(1.4)–eGFP system, a minimum of 100 cells were examined for each experimental condition and cells with or without cytoplasmic fluorescence were counted and their percentage calculated and graphed using Excel (GraphPad Software).

Data Sharing Statement.

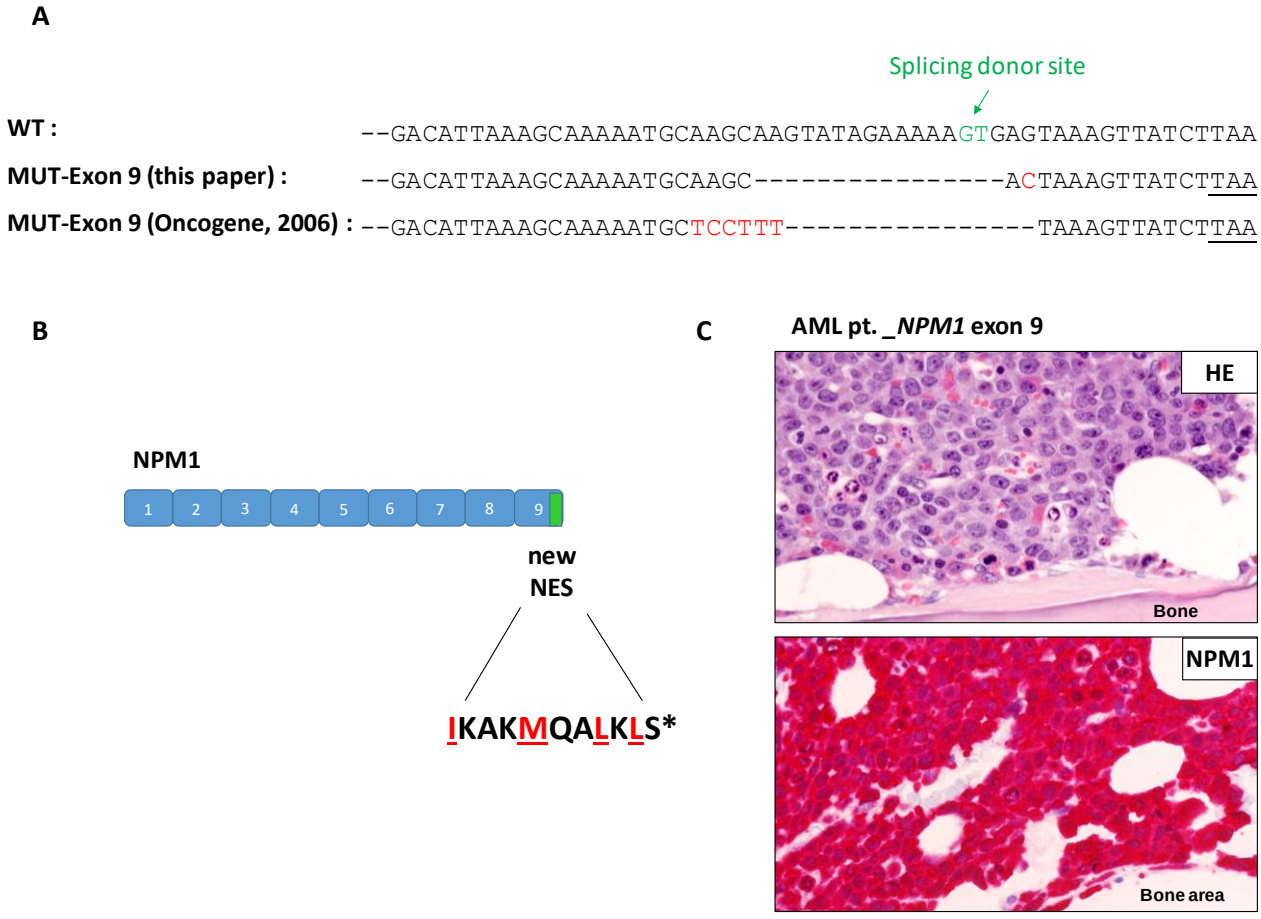
For original data, please contact maria.martelli@unipg.it

Supplemental Figure 1



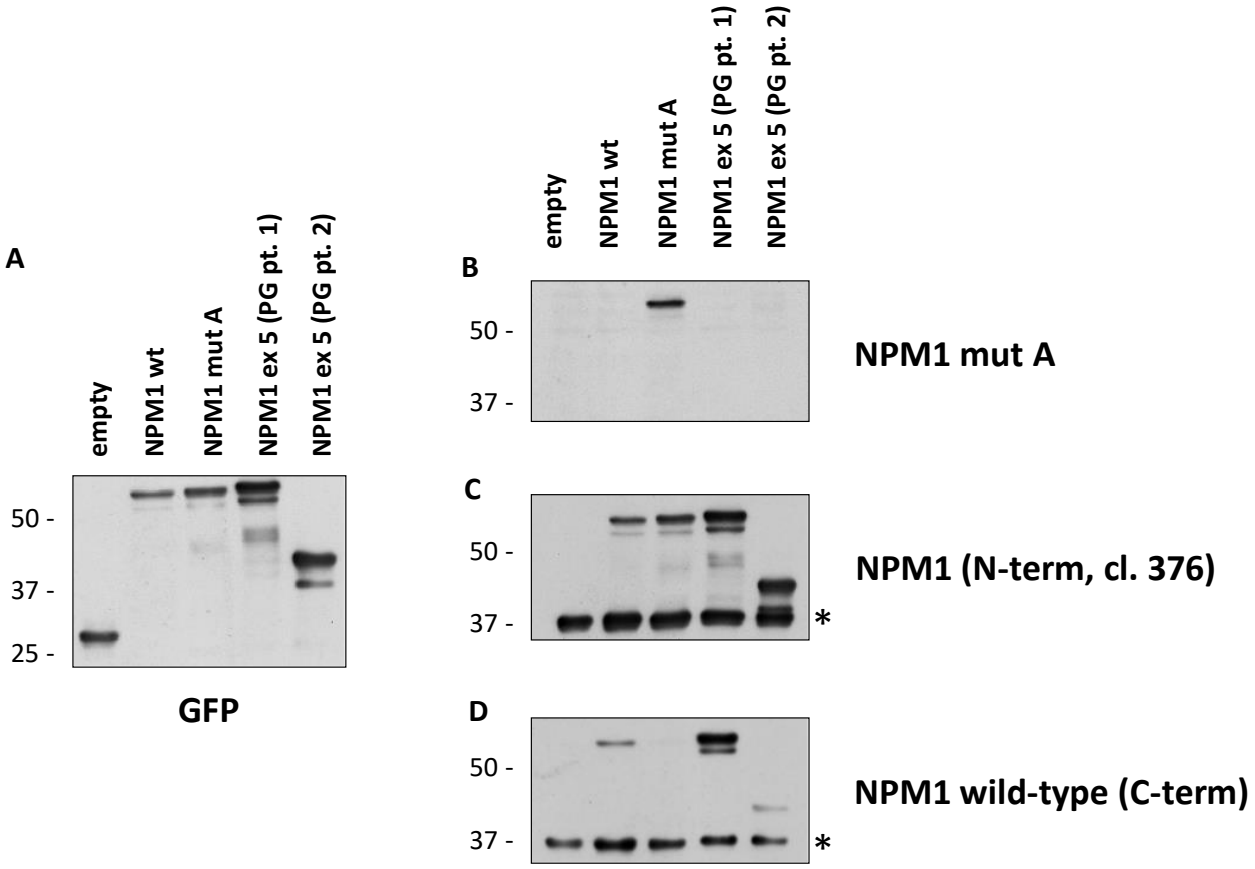
Supplemental Figure 1. Panel of anti-NPM1 antibodies used for Western blotting studies. The anti-pan NPM1 mouse monoclonal antibody (directed against the N-terminus of NPM1) (mAb Cl. 376) is able to recognize either NPM1 wild-type or all the NPM1 mutants, including NPM1 MUT-Exon 11 or MUT-Exon 9. NPM1 wild-type protein (*upper panel*) is detected in the nucleus by immunohistochemistry with mAb Cl. 376, whilst NPM1 mutant proteins (*lower panel*) are detected in the cytoplasm of leukemic cells. The anti-NPM1 wild type mouse monoclonal antibody (mAb Anti-NPM1wt) is specific for the C-terminus of the wild-type protein. The anti-NPM1 mutant rabbit polyclonal antibody (poly Ab Anti-NPM1 mutA) is directed against the C-terminus of the NPM1 mutant A and is able to recognize specifically only NPM1 mutant A and other exon-12 mutants[2], but not NPM1 MUT-Exon 11 or MUT-Exon 9.

Supplemental Figure 2

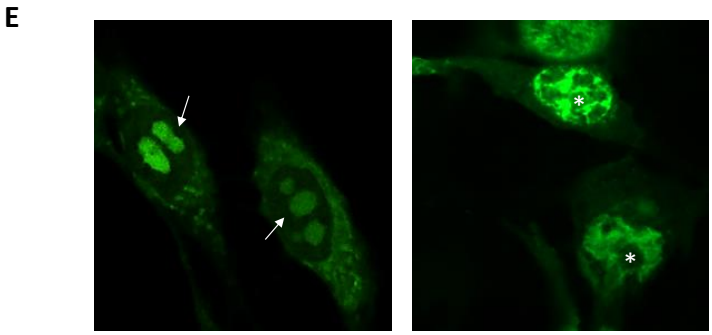


Supplemental Figure 2. A new *NPM1* gene mutation involving exon 9. A. Alignment of genomic sequences corresponding to the *NPM1* exon 9 showing a new mutation in the splicing donor site of exon 9, consisting in the loss of 16 nucleotides at the 3′ end of exon 9 and the presence of 1 new nucleotide (C, in red), as compared with mutation in exon 9 described by Mariano et al (Oncogene 2006; 25(31):4376-80). Nucleotide sequence analysis revealed also in our case a frameshift with formation of a stop codon (TAA, underlined). **B.** Translation of nucleotide sequence revealed a newly acquired NES motif (IKAKMQALKLS) at the C-terminus of the truncated protein. **C.** Immunohistochemical analysis on bone marrow trephine showing diffuse infiltration by leukemic blasts (HE, *upper*) with aberrant cytoplasmic positivity for NPM1 (NPM1, *lower*). HE: hematoxylin-eosin staining; NPM1: mouse monoclonal clone 376 anti-NPM1 antibody staining, APAAP technique; hematoxylin counterstaining. Images were collected using an Olympus B61 microscope and a UPlanApo 40x/0.85; Camedia 4040, Dp_soft Version 3.2; and Adobe Photoshop CC 2019.

Supplemental Figure 3



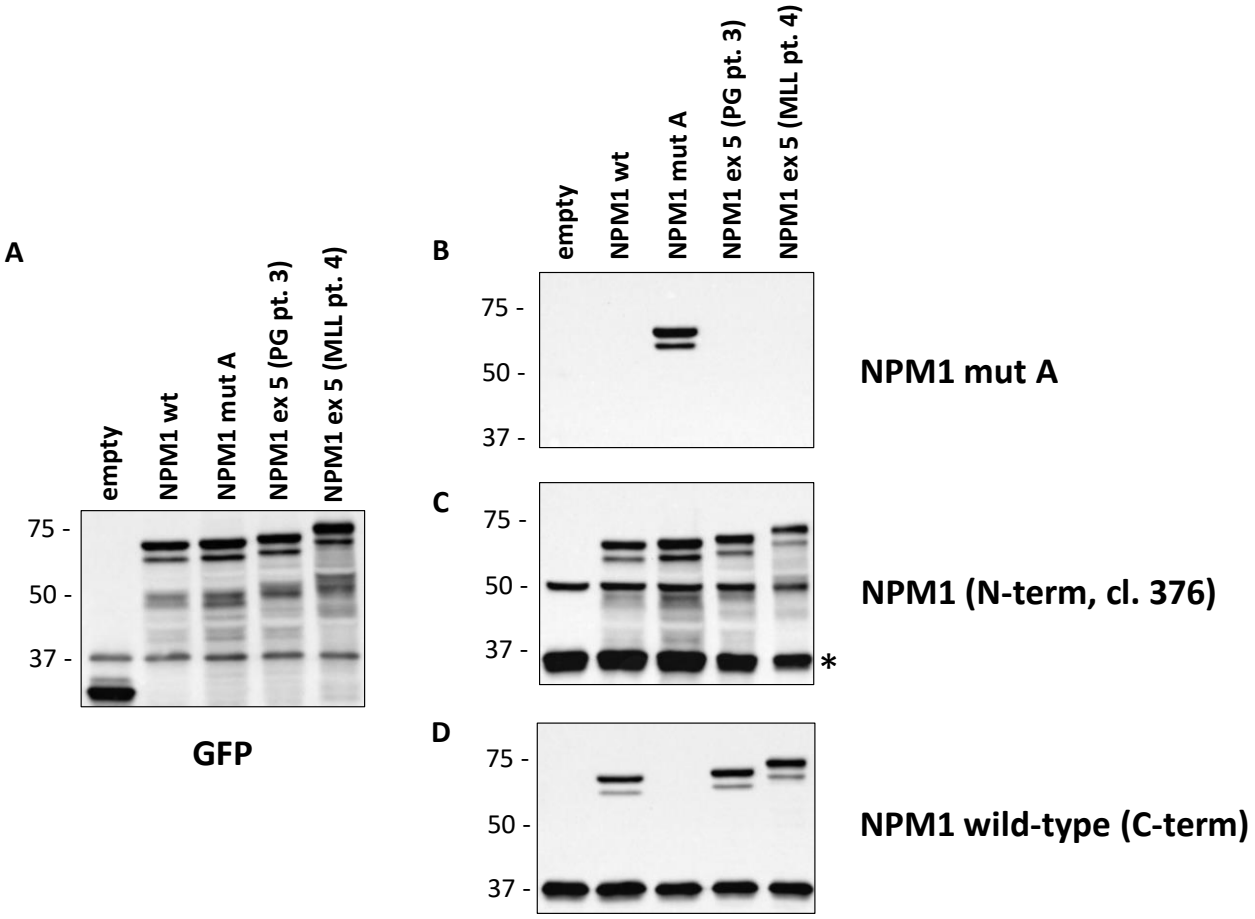
GFP: 27 kDa
NPM1 wt/mut A: 38 kDa (GFP-fusion protein: 65 kDa)
NPM1 ex 5 (PG pt. 1): predicted 38 kDa (GFP-fusion protein: 65 kDa)
NPM1 ex 5 (PG pt. 2): predicted 15 kDa (GFP-fusion protein: 42 kDa)
* endogenous NPM1



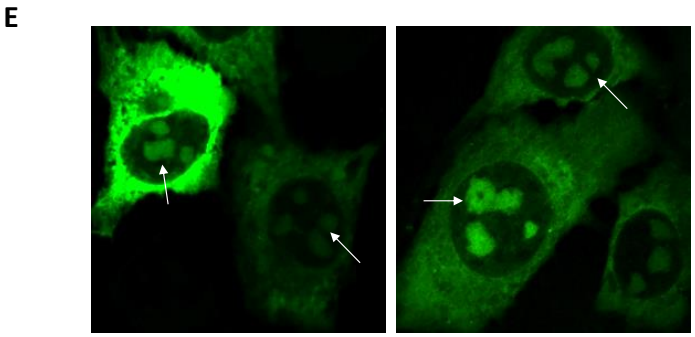
Supplemental Figure 3. Analysis of GFP-NPM1 fusion proteins (patients PG pt.1 and PG pt.2) by western blotting and confocal microscopy. A. The GFP fusion proteins are all recognized by the anti-GFP antibody at the different corresponding molecular weights (MW). **B.** As expected, based on WB data on patients samples, the anti-NPM1 mutant A recognizes the GFP-NPM1 mut A (3rd

lane), but not the new exon 5 NPM1 mutants. **C.** The antibody directed against the N-terminal portion of NPM1 (the same used in IHC, cl. 376) recognizes all forms of GFP fusion proteins and the endogenous NPM1 (asterisk). **D.** The antibody directed against the C-terminal portion of wild-type NPM1 recognizes the wt protein (2nd lane), the endogenous NPM1 (asterisk) and the NPM1 ex 5 (PG pt.1) mutant (4th lane) that retains the C-terminal portion of wt NPM1 unaltered, but not the NPM1 ex 5 (PG pt.2) mutant (5th lane) that is truncated (the faint signal at lane 5 residues after stripping and reprobing the membrane). **E.** NIH-3T3 overexpressing the new GFP-NPM1 fusion proteins showing their aberrant cytoplasmic subcellular localization. To note, the mutant protein from PG pt.1 localizes also in the nucleoli (white arrows), whilst the truncated mutant from PG pt.2 spares the nucleoli (white asterisks). Images were acquired using a Zeiss LSM 800 confocal microscope (Carl Zeiss) with a 488-nm (for eGFP) laser line for excitation, and a 63x/1.4 OIL Plan-Apochromat objective.

Supplemental Figure 4.



GFP: 27 kDa
NPM1 wt/mut A: 38 kDa (GFP-fusion protein: 65 kDa)
NPM1 ex 5 (PG pt.3): predicted 46.9 kDa (GFP-fusion protein: 73.9 kDa)
NPM1 ex 5 (MLL pt. 4): predicted 43.6 kDa (GFP-fusion protein: 70.6 kDa)
* endogenous NPM1



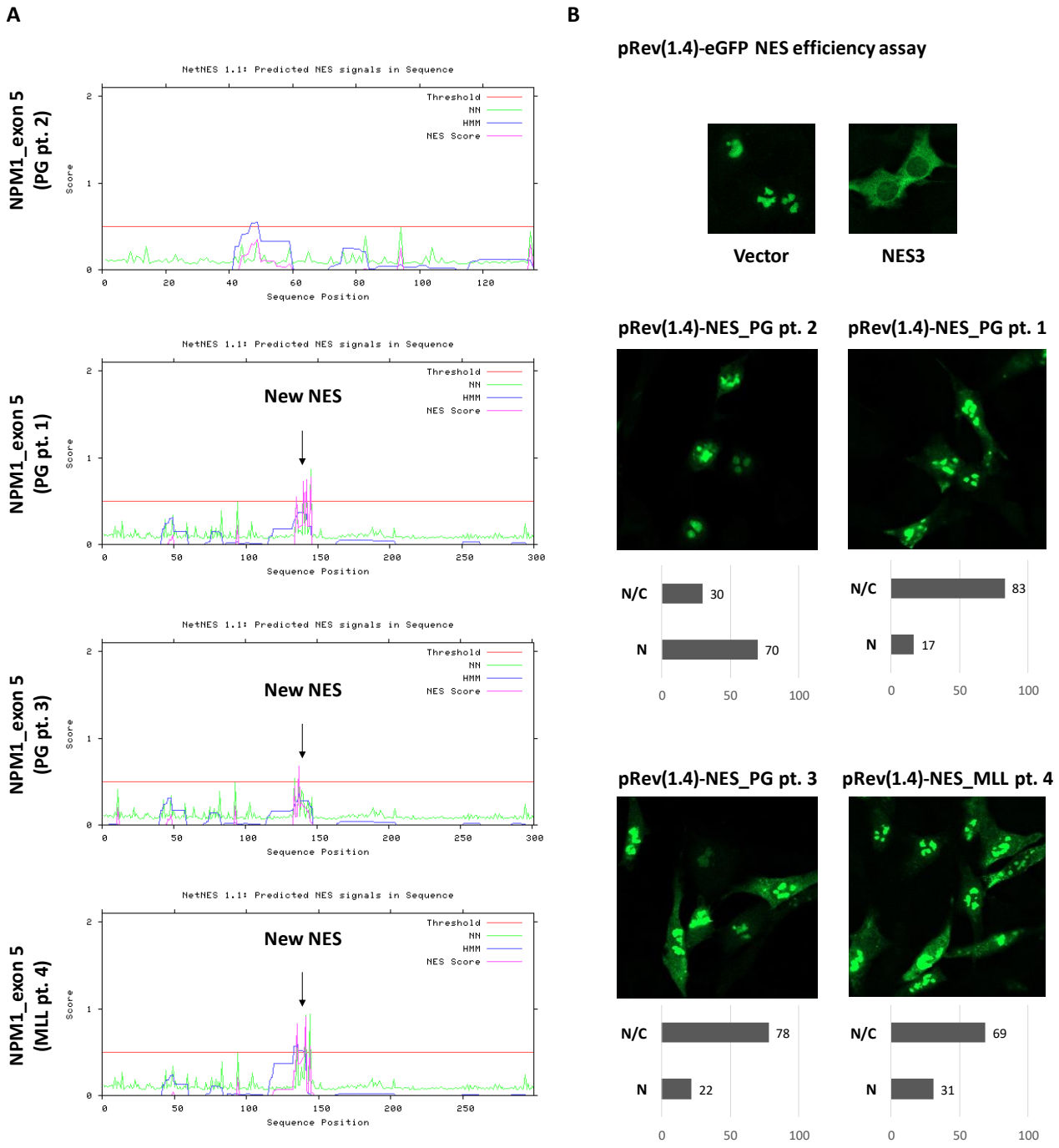
pEGFP-NPM1_exon 5 (PG pt. 3)

pEGFP-NPM1_exon 5 (MLL pt. 4)

Supplemental Figure 4. Analysis of GFP-NPM1 fusion proteins (patients PG pt.3 and MLL pt.4) by western blotting and confocal microscopy. A. The GFP fusion proteins are all recognized by the anti-GFP antibody at the different corresponding MW. **B.** As expected, the anti-NPM1 mutant A

recognizes the GFP-NPM1 mut A (3rd lane), but not the new exon 5 NPM1 mutants. **C.** The antibody directed against the N-terminal portion of NPM1 (the same used in IHC, cl. 376) recognizes all forms of GFP fusion proteins and the endogenous NPM1 (asterisk). **D.** The antibody directed against the C-terminal portion of wild-type NPM1 recognizes the wt protein (2nd lane), the endogenous NPM1 (asterisk) and both the NPM1 ex 5 mutant (PG pt. 1) (4th lane) and NPM1 ex 5 mutant (MLL pt. 4) (5th lane) that retain the C-terminal portion of wt NPM1 unaltered. **E.** NIH-3T3 overexpressing the new GFP-NPM1 fusion proteins showing their aberrant cytoplasmic subcellular localization. Also in these cases, the mutant proteins localizes in the nucleoli (white arrows), due to the conservation of the C-terminal NoLS of the wt NPM1. Images were acquired using a Zeiss LSM 800 confocal microscope (Carl Zeiss) with a 488-nm (for eGFP) laser line for excitation, and a 63x/1.4 OIL Plan-Apochromat objective.

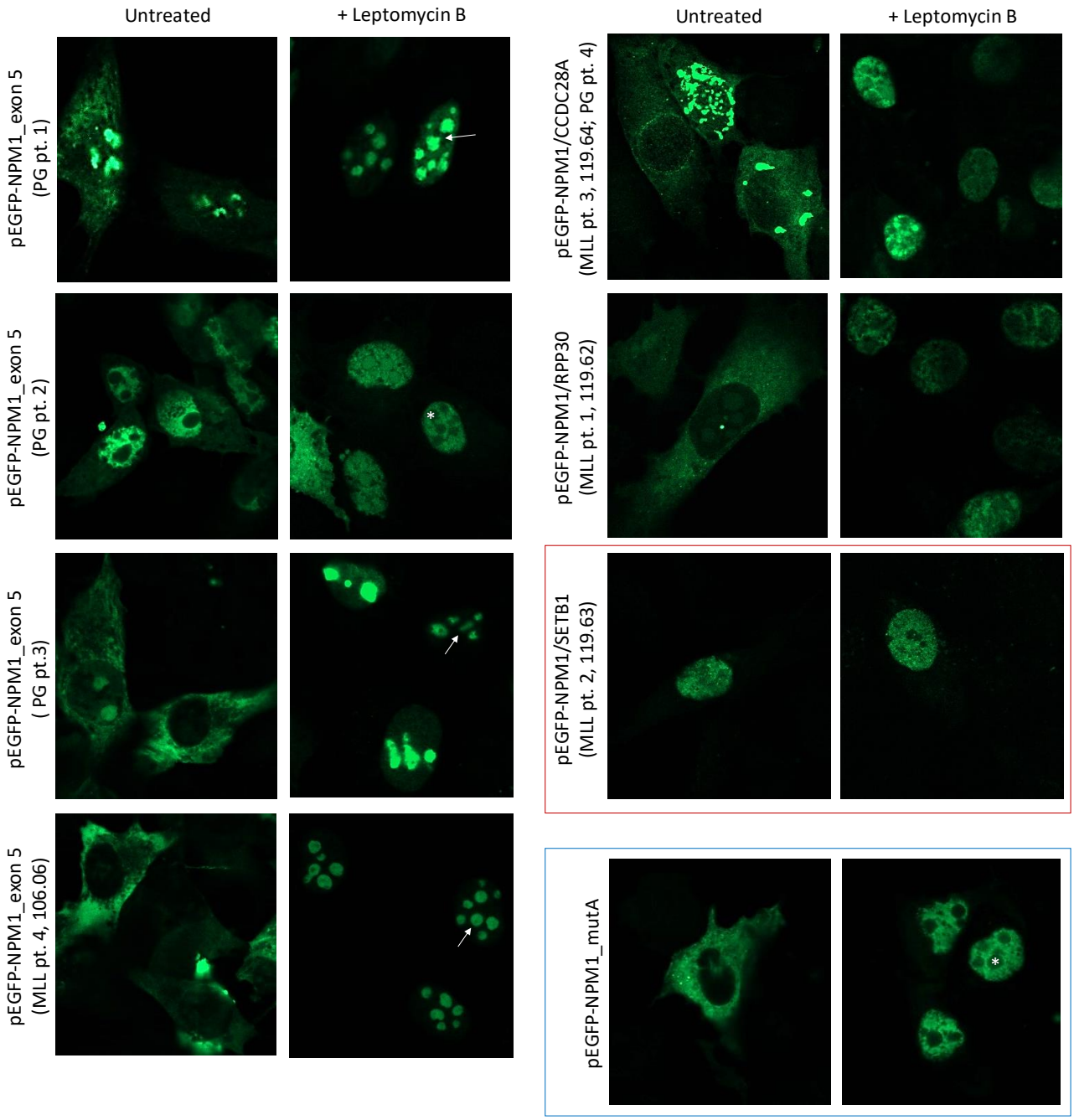
Supplemental Figure 5.



Supplemental Figure 5. New intramolecular NES are acquired by new exon-5 NPM1 mutants. **A.** Search for acquisition of a NES motif in the new exon-5 mutants protein sequence was performed by applying the NES detection NetNES 1.1 (http://www.cbs.dtu.dk/services/NetNES) and NESential server (http://seq.cbrc.jp/NESsential). A new intramolecular NES with high score was detected in PG pt.1, PG pt.3 and MLL pt.4 (arrows). **B.** The sequence of either the C-terminus of the PG pt.2 mutant or of the individual newly discovered intramolecular NES (from PG pt.1, PG pt.3 and MLL pt.4) were cloned in the fusion protein pRev(1.4)–eGFP and expressed in NIH-3T3 cells to evaluate their

efficiency, as previously described.[5] Subcellular localization is shown for each condition. The Rev(1.4) (vector) is nuclear by definition, having its own NES sequence inactivated by mutation. The positive control Rev(1.4)-NES3-eGFP (NES3) is cytoplasmic in the majority of transfected cells. The C-terminal sequence of the PG pt.2 drives cytoplasmic localization only in a minority of cells, whereas all other NES motifs (i.e. NES_PG pt.1, NES_PG pt.3, and NES_MLL pt.4) result in cytoplasmic positive staining in a greater proportion of cells. Histograms show different strength of various NPM NES motifs, measured as the percentage of cytoplasmic positive cells (partial or exclusive eGFP cytoplasmic staining). N= Nucleus, nucleolar and/or nucleoplasmic eGFP positivity; N/C= partial or exclusive eGFP cytoplasmic staining.

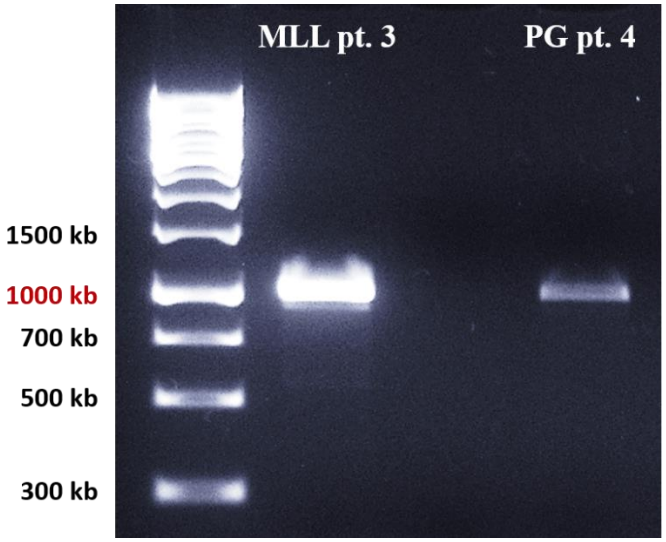
Supplemental Figure 6.



Supplemental Figure 6. Cytoplasmic localization of the new NPM1 mutants/fusion transcripts is NES-dependent. NIH-3T3 overexpressing the new GFP-NPM1 fusion proteins were either left untreated or treated with the Crm1/XPO1-inhibitor leptomycin B and the subcellular localization of the GFP-NPM1 fusion proteins was analyzed by confocal microscopy. In all cases but NPM1/SETPB1 (red box), in the untreated cells the GFP-NPM1 fusion protein is localized in the cytoplasm, whilst upon treatment with leptomycin B it relocalizes either into the nucleoplasm sparing the nucleoli (white asterisks), for those constructs lacking the nucleolar binding site (NoLS)

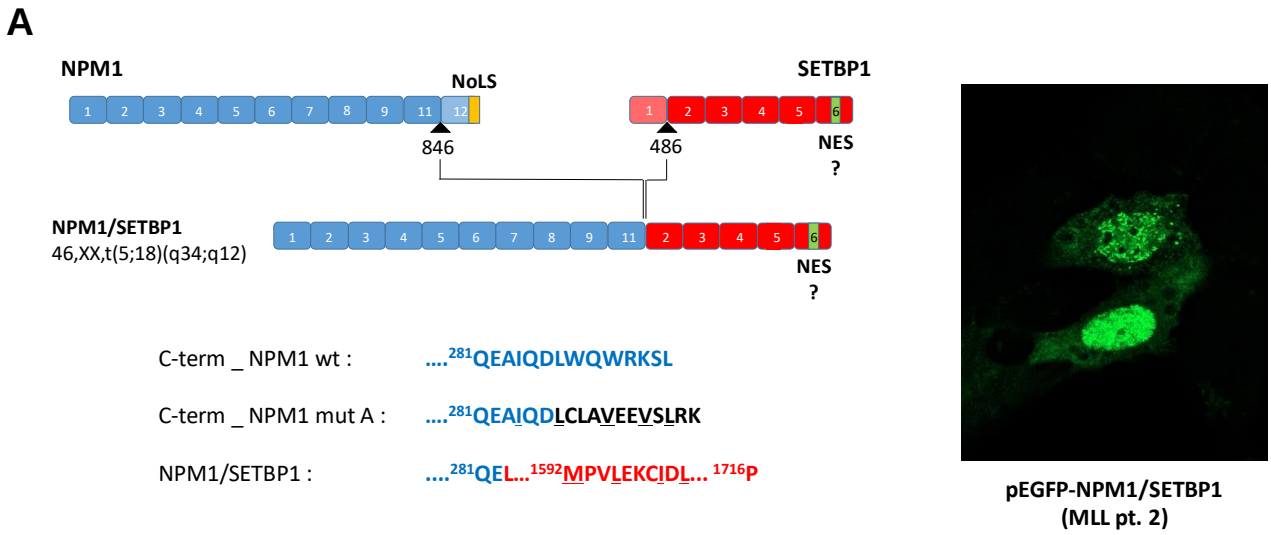
corresponding to the C-terminus of the wild-type NPM1 (i.e. NPM1 mutA; NPM1_exon 5 (PG pt.2); NPM1/RPP30 (MLL pt.1); NPM1/CCDC28A (MLL pt.3, and PG pt.4)), or into nucleoplasm/nucleoli, for those constructs retaining the C-terminus of the wild-type NPM1 (i.e. NPM1_exon 5 (PG pt. 3); NPM1_exon 5 (MLL pt. 4)) (white arrows). NPM1/SETPB1 subcellular localization does not change with leptomycin B (red box). Results are representative of 3 independent experiments. Images were acquired using a Zeiss LSM 800 confocal microscope (Carl Zeiss) with a 488-nm (for eGFP) laser line for excitation, and a 63x/1.4 OIL Plan-Apochromat objective.

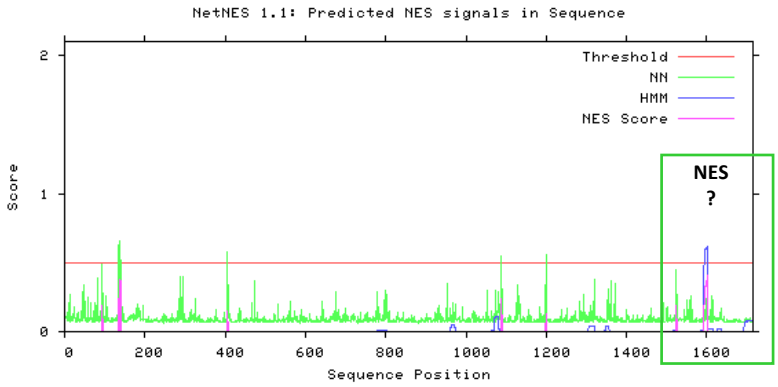
Supplemental Figure 7.



Supplemental Figure 7. PCR for *NPM1/CCDC28A* fusion transcript. PCR amplification was performed on genomic DNA from BM from either MLL pt.3 or PG pt.4 using primers complementary to *NPM1* exon 11 and *CDC28A* exon 2 (F 5′-CTGAGGCCCCAGAACTATCT-3' and R 5′-GTCGCTGTGAAGTTGACTGT-3', respectively). The PCR product from PG pt.4 (4th lane) showed the same size (about 1000 kb) as MLL pt.3 (2nd lane) when run in a 1% agarose gel in parallel with DNA size marker (1 Kb Plus Ladder, 1st lane). The 3rd lane is empty. Gel electrophoresis was analyzed by UVidoc Imager (Uvitec).

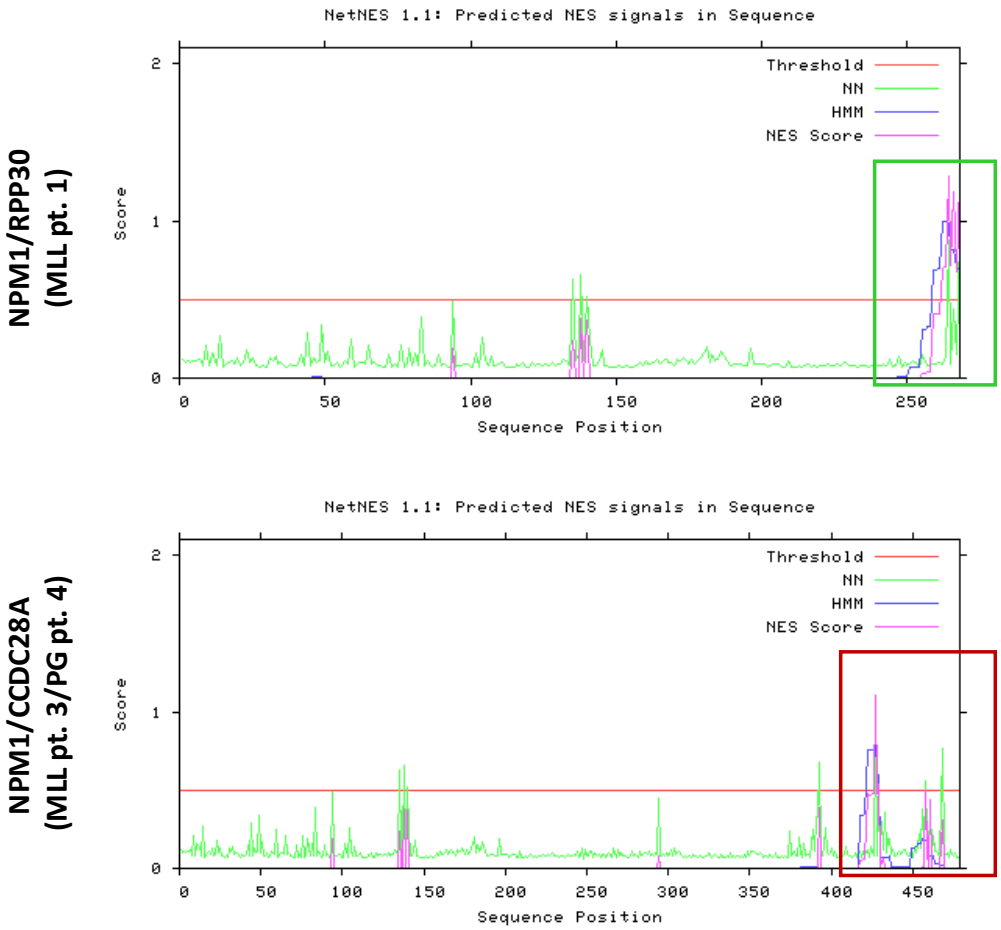
Supplemental Figure 8.





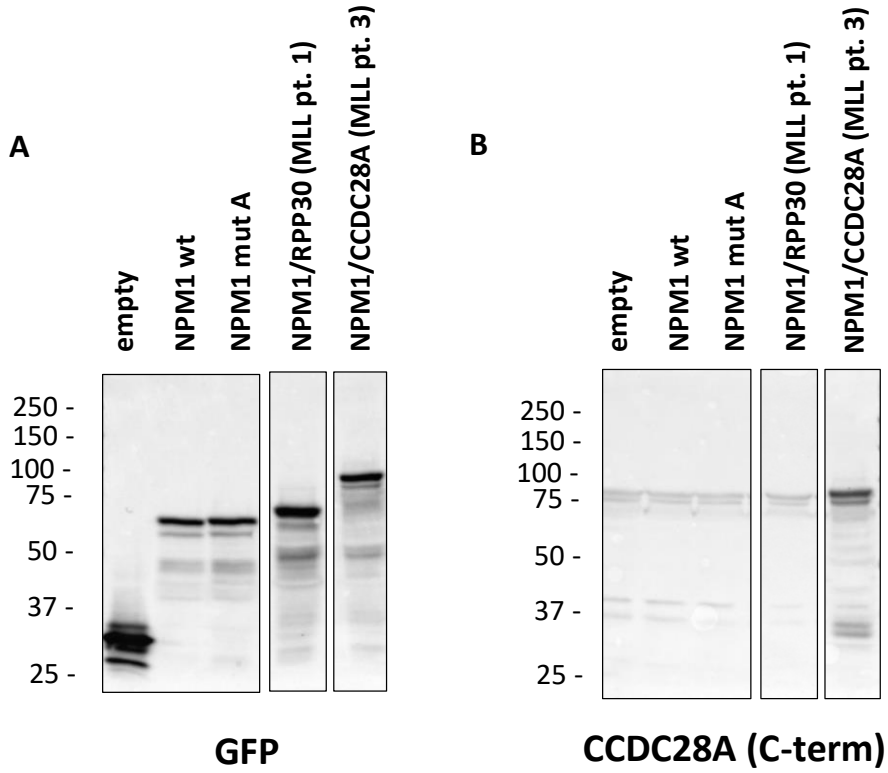
Supplemental Figure 8. *NPM1/SETBP1* fusion transcript. A. In MLL pt.2, *NPM1* was rearranged with *SETBP1* gene at the end of exon 11 (breakpoint at position 846 of *NPM1* cDNA based on the transcript ENST00000296930) and *SETBP1* was rearranged with *NPM1* at the beginning of exon 2 (breakpoint at position 486 in *SETBP1* cDNA ENST00000282030.5). The encoded fusion was predicted to be in-frame. The new fusion protein is 1716 aa long, with a predicted MW of 188.56 kDa. In the protein sequence, in correspondence of SETBP1 exon 6, a NES motif was predicted to be present (MPVLEKCIDL). Expression of the new GFP-NPM1/SETBP1 fusion protein in NIH-3T3 shows it is localized in the nucleoplasm. Images were acquired using a Zeiss LSM 800 confocal microscope (Carl Zeiss) with a 488-nm (for eGFP) laser line for excitation, and a 63x/1.4 OIL Plan-Apochromat objective. **B.** Search for acquisition of a NES motif was performed by applying the NES detection NetNES 1.1 (http://www.cbs.dtu.dk/services/NetNES) and NESential server (http://seq.cbrc.jp/NESsential).

Supplemental Figure 9



Supplemental Figure 9. NES motif prediction in the new NPM1 fusion proteins NPM1/RPP30 and NPM1/CCDC28A. Search for acquisition of a NES motif in the new NPM1 fusion protein sequence was performed by applying the NES detection NetNES 1.1 (http://www.cbs.dtu.dk/services/NetNES) and NESential server (http://seq.cbrc.jp/NESsential). A new C-terminal NES (green box) is created in the NPM1/RPP30. Two NES are present at the C-terminus of NPM1/CCDC28A (red box).

Supplemental Figure 10



GFP: 27 kDa
NPM1 wt/mut A: 38 kDa (GFP-fusion protein: 65 kDa)
NPM1/RPP30 (MLL pt. 1): predicted 29 kDa (GFP-fusion protein: 56 kDa)
NPM1/CCDC28A (MLL pt. 3 and PG pt. 4): predicted 53.1 kDa (GFP-fusion protein: 80.1 kDa)

Supplemental Figure 10. Western blotting of GFP-NPM1 fusion proteins (patients MLL pt. 1 and MLL pt. 3/PG pt. 4). A. The GFP fusion proteins are all recognized by the anti-GFP antibody at the different corresponding MW. **B.** The anti-CCDC28A antibody directed against the C-terminal part of CCDC28A recognizes only the NPM1/CCDC28A fusion protein.

Supplemental References.

1. Falini B, Mecucci C, Tiacci E, et al; GIMEMA Acute Leukemia Working Party. Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. N Engl J Med. 2005;352(3):254-266.

2. Martelli MP, Manes N, Liso A, et al. A western blot assay for detecting mutant nucleophosmin (NPM1) proteins in acute myeloid leukaemia. Leukemia. 2008 Dec;22(12):2285-8.

3. Quentmeier H, Martelli MP, Dirks WG, et al. Cell line OCI/AML3 bears exon-12 NPM gene mutation-A and cytoplasmic expression of nucleophosmin. Leukemia. 2005 Oct;19(10):1760-7.

4. Falini B, Bolli N, Shan J, et al. Both carboxy-terminus NES motif and mutated tryptophan(s) are crucial for aberrant nuclear export of nucleophosmin leukemic mutants in NPMc+ AML. Blood. 2006;107: 4514-4523.

5. Bolli N, Nicoletti I, De Marco MF, et al. Born to be exported: COOH-terminal nuclear export signals of different strength ensure cytoplasmic accumulation of nucleophosmin leukemic mutants. Cancer Res. 2007 Jul 1;67(13):6230-7.

6. Henderson BR, Eleftheriou A. A comparison of the activity, sequence specificity, and CRM1-dependence of different nuclear export signals. Exp Cell Res 2000; 256: 213–24.