Molecular synergy underlies the co-occurrence patterns and phenotype of NPM1-mutant acute myeloid leukemia.

Oliver M Dovey[1], Jonathan L Cooper[1], Annalisa Mupo[1], Carolyn S Grove[1,3,4], Claire Lynn[5], Paul Green[1], Nathalie Conte[6], Robert M Andrews[7], Kevin Shannon[8], Tyler Jacks[9], Roland Rad[10], Mark Arends[11], Penny Wright[2], Allan Bradley[1], Ignacio Varela[12], and George S Vassiliou[1-2†].

[1]The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK;

[2]Department of Haematology, Cambridge University Hospitals NHS Trust, Cambridge, UK

[3]School of Pathology and Laboratory Medicine, University of Western Australia, Crawley, Western Australia, Australia

[4]PathWest Division of Clinical Pathology, Queen Elizabeth II Medical Centre, Nedlands, Western Australia, Australia

[5]Leukemia and Stem Cell Biology Group, Division of Cancer Studies, Department of Haematological Medicine, King's College London, Denmark Hill Campus, London SE5 9NU, UK

[6]Mouse Informatics, European Bioinformatics Institute (EMBL-EBI), Wellcome Trust Genome Campus, Hinxton, Cambridge, UK

[7]Institute of Translation, Innovation, Methodology and Engagement, Cardiff University School of Medicine, Heath Park, Cardiff, UK

[8]Department of Pediatrics, University of California San Francisco, San Francisco, CA; Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, California, USA

[9]David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA; Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

[10]Department of Medicine II, Klinikum Rechts der Isar, Technische Universität München, 81675 Munich, Germany; German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany

[11]Edinburgh Cancer Research Centre, University of Edinburgh, Western General Hospital, Edinburgh, UK

[12]Instituto de Biomedicina y Biotecnología de Cantabria, Santander 39011, Spain.

†Correspondence: gsv20@sanger.ac.uk

Running Title: Molecular synergy in NPM1 mutant AML

Include article word count here: 3992

Abstract Word Count: 248

Scientific Category: Myeloid Neoplasia

Number of figures: 5

Number of references: 33

Key Points

1 or 2 key points no more than 140 characters

NPM1c and NRAS-G12D cooperate to cause high penetrance AML in mice.

Aggressive onset of mutant NPM1-FLT3 AML is underpinned by distinctive molecular and cellular synergies.

Abstract

Mutations affecting *NPM1* define the commonest subgroup of acute myeloid leukemia (AML). They frequently co-occur with mutations of *FLT3*, usually internal tandem duplications (ITD), and less commonly of *NRAS* or *KRAS*. Co-occurrence of mutant *NPM1* with *FLT3-ITD* carries a significantly worse prognosis than the *NPM1-RAS* combination. To understand the molecular basis of these observations we compare the effects of the two combinations on hematopoiesis and leukemogenesis in knock-in mouse models. Early effects of these mutations on hematopoiesis show that compound *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ or *Npm1*^{cA};*Flt3*^{ITD} share a number of features: *Hox* gene over-expression, enhanced self-renewal, expansion of hematopoietic progenitors and a bias towards myeloid differentiation. The most notable differences were that *Npm1*^{cA};*Flt3*^{ITD} mutants, displayed significantly higher peripheral leucocyte counts, early depletion of common lymphoid progenitors and a monocytic bias compared to the granulocytic bias observed in *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ mutants. Underlying this was a striking molecular synergy manifested as a dramatically altered gene expression profile in *Npm1*^{cA};*Flt3*^{ITD}, but not *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$, progenitors compared to wild type. Both compound models developed high penetrance AML although latency in *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ mutants was significantly longer (median survival 138 days post-pIpC in *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ vs 52.5 days in *Npm1*^{cA};*Flt3*^{ITD} mice). During AML evolution, both models acquired additional copies of the mutant *Flt3* or *Nras* alleles, but only *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ mice showed acquisition of other mutations observed in human AML, including *IDH1* R132Q. Our results show that molecular complementarity underlies the frequent co-occurrence of mutant *NPM1* and *FLT3-ITD*, and the poorer AML prognosis associated with this mutation combination compared to *NPM1-NRAS*.

Introduction

Advances in genomics have defined the somatic mutational landscape of acute myeloid leukemia (AML), leading to a detailed characterisation of their prognostic significance and patterns of mutual co-occurrence or exclusivity.[1, 2] Mutations in *NPM1*, the gene for Nucleophosmin, characterise the most common subgroup of AML representing 25-30% of all cases, result in cytoplasmic dislocation of the protein (*NPM1c*) and are mutually exclusive of leukemogenic fusion genes.[1-3] As is often the case for fusion genes, progression to AML after the acquisition of mutant *NPM1* is contingent upon the gain of additional somatic mutations such as those that activate STAT and/or RAS signalling.[4] For reasons that are not clear, this transforming step favours acquisition of internal tandem duplications in *FLT3* (*FLT3*-ITD) over other somatic mutations with similar effects such as those involving *NRAS* or *KRAS*.[1-4] Furthermore, the *NPM1c/FLT3-ITD* combination is associated with a significantly worse prognosis compared to combinations of *NPM1c* with mutant *NRAS*, *KRAS* or other mutations.[2]

Whilst the adverse prognostic impact of *NPM1/FLT3-ITD* vs *NPM1/RAS* co-mutation influences clinical decisions in AML, its molecular basis and that of the frequent co-occurrence of *NPM1c* and *FLT3-ITD* in AML are unknown. Here, in order to investigate these phenomena, we compare the interaction of *Npm1c* with *Flt3-ITD* to its interaction with $Nras^{G12D}$ in knock-in mice. Individually, knock-in models of *NPM1c*, *FLT3-ITD* and *NRAS-G12D* display enhanced myelopoiesis and progression to myeloproliferative disorders or AML in a significant proportion of animals.[5-7] Also, we and others have previously shown that *Npm1c* and *Flt3-ITD* synergise to drive rapid-onset AML, but the interaction between *Npm1c* and mutant $Nras^{G12D}$ has not, to our knowledge, been previously investigated in knock-in mice.[8] Our findings reveal that the combination of *Npm1c* and *Flt3-ITD* has an early profound effect on gene expression and hemopoiesis, whilst *Npm1c* and *Nras-G12D* display only modest molecular synergy and subtler cellular changes. Also, whilst both types of co-mutation drove AML in the majority of mice, the leukemias in *Npm1c;Flt3-ITD* mice were more aggressive and undifferentiated than those which developed in *Npm1c;Nras-G12D animals*. At the genomic level, there was frequent amplification in both models of the mutant *Flt3-ITD* or *Nras-G12D* allele, however additional somatic mutations in AML driver genes (e.g. Idh1 and Ptpn11) were seen only in *Npm1c;Nras-G12D* AMLs. Our findings propose that the molecular synergy between *Npm1c* and *Flt3-ITD* underpin the co-occurrence patterns, phenotype and prognosis of NPM1-mutant AML.

Materials and methods

Animal husbandry

Mx1-Cre+;*Npm1*$^{flox\text{-}cA/+}$ were crossed with *Nras*$^{LSL\text{-}G12D}$ or *Flt3*^{ITD} mice, to generate triple transgenic animals (*Mx1-Cre*;*Npm1*$^{flox\text{-}cA/+}$;*Nras*$^{LSL\text{-}G12D/+}$ *and Mx1-Cre*;*Npm1*$^{flox\text{-}cA/+}$;*Flt3*$^{ITD/+}$).[9] To activate conditional alleles (*Npm1*^{cA} and *Nras*^{G12D}) in approximately 12-14 week old *Mx1-Cre*;*Npm1*$^{flox\text{-}cA/+}$;*Nras*$^{LSL\text{-}G12D/+}$ mice, *Mx1-Cre* was induced by intraperitoneal administration of 5 doses of 200 ㄴ g pIpC over a 10 day period. As described recently, *Mx-1 Cre;Npm1*$^{Flox\text{-}cA/+}$;*Flt3*$^{ITD/+}$ mutants do not require pIpC induction of *Mx1-Cre* and recombination of the *Npm1*$^{flox\text{-}cA}$ allele[8]. For pre-leukemic analyses *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ were sacrificed 4-5 weeks post pIpC and *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ were sacrificed at 5 weeks of age. Genotyping for mutant alleles was performed as previously.[5, 10]

Hematological measurements

Blood counts were performed on a VetABC analyzer (Horiba ABX).

Histopathology

Formalin fixed, paraffin embedded (FFPE) sections were stained with hematoxylin and eosin. Samples from leukemic mice were also stained with anti-CD3, anti-B220 and anti-myeloperoxidase, and detected using immunoperoxidase. All material was examined by two experienced histopathologists (P.W. and M.A.) blinded to mouse genotypes. Selected samples were also studied for total ERK1/2 (p44/42 MAPK, clone 137F5, Cell Signalling) and pERK1/2 (phosphor-p44/42 MAPK, clone 197G2, Cell Signalling).

Colony-forming assays and serial re-plating

Nucleated cells (3 x10^4) from bone marrow aspirates of mutant and wild-type mice were suspended in cytokine-containing methylcellulose-based media (M3434, Stem Cell Technologies) and plated in duplicate wells of 6-well plates. Colony-forming units (CFUs) were counted 7 days later. For serial re-plating, 3 x10^4 cells were re-seeded and colonies counted after 7 days.

Flow cytometry and cell sorting

For flow cytometry, single cell suspensions of bone marrow cells or splenocytes were passed through a 0.4 ㄴ m nylon filter and suspended in 0.85% NH_4Cl for 5 minutes to lyse erythrocytes. Cells were then suspended in Hank's Balanced Salt Solution (HBSS) supplemented with 2% FCS and 10 ㄴ M HEPES. Progenitor populations were defined and stained as described in supplementary methods.

Gated cellularity was calculated by multiplying the percentage of gated cells by the total number of nucleated cells from bone marrow samples after erythrocyte depletion.

Retroviral transduction of bone marrow progenitors

Lineage depleted bone marrow aspirates, isolated from wildtype and *Flt3*$^{ITD/+}$ mice, were transduced with MSCV-*Hoxa9*-GFP and/or MSCV-*Nkx2-3*-CFP retroviruses and expanded for 7 days in liquid culture (X-Vivo, Lonza, supplemented with 10ng/ml IL-3, 10ng/ml IL-6 and 50ng/ml SCF, Peprotech). CFP, GFP or double positive cells were FACS sorted and 2.5 x10^4 cells re-plated in semi-solid media as previously described. For cloning strategy, see supplemental methods.

Microarray and comparative genomic hybridization analysis

Mouse gene expression profiles were generated using the Illumina MouseWG-6 v2 Expression BeadChip platform (Illumina). DNA copy number variation in leukemic samples was assessed using the Mouse Genome Comparative Genomic Hybridization 244K Microarray (Agilent Technologies). Full details of analysis are provided in supplemental methods. For mouse gene expression profiling, n=4-10 (Lin$^-$) or n=3-5 (MPP).

AML exome sequencing and mutation calling

Whole exome sequencing (WES) of AML bone marrow and control C57BL/6N or 129Sv tail DNA was performed using the Agilent SureSelect Mouse Exon Kit (Agilent Technologies) and paired-end sequencing on a HiSeq2000 sequencer (Illumina). Validation of mutations was performed using MiSeq sequencing (Illumina) of amplicon libraries as described before (See Supplemental Methods Figure S1 and Supplemental Tables 6 and 7 for primer sequences).[11] Full details of analysis are provided in supplemental methods.

Statistics

Student *t* test or one-way analysis of variance (ANOVA, Bonferroni adjusted) were used for statistical comparisons as appropriate and unless stated. Error bars represent standard error of the mean (SEM). Significant values are reported as: * P<0.05 Vs wildtype, ** P<0.01 Vs wildtype, *** P<0.001 Vs wildtype, ⚌ P<0.05 Vs *Flt3*$^{ITD/+}$, ⚌⚌ P<0.01 Vs *Flt3*$^{ITD/+}$, ⚌⚌⚌ P<0.001 Vs *Flt3*$^{ITD/+}$), ♣P<0.05 Vs *Nras*$^{G12D/+}$, ♣♣P<0.01 Vs *Nras*$^{G12D/+}$, ♣♣♣P<0.001 Vs *Nras*$^{G12D/+}$, † P<0.05 *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ Vs *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$, †† P<0.01 *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ Vs *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$, ††† P<0.001 *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ Vs *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$.

Results

Mutant *Npm1* co-operates with *Nras-G12D* and *Flt3-ITD* to increase self-renewal of hematopoietic progenitors and expand myelopoiesis

To understand the impact of the studied mutations alone and in combination, we studied hemopoietic cell compartments of *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$, *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$, *Npm1*$^{cA/+}$;-*Nras*$^{G12D/+}$, *Flt3*$^{ITD/+}$ and wild type (WT) control mice sacrificed 4-6 weeks after activation of conditional mutations (Figure 1). Compared to *Flt3*$^{ITD/+}$ single mutants, *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$ mice showed a marked increase of peripheral WCCs (56±13.4 vs 6.5±0.5 x10^{6} g/L, p<0.001) and spleen weights (0.63g vs 0.16g, p<0.001), but not of total bone marrow cellularity (Figure 1B). By contrast, both *Nras*$^{G12D/+}$ and *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ mutants exhibited only subtle increases in spleen size (WT: 0.12g, *Nras*$^{G12D/+}$: 0.18g, *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$: 0.19g, p<0.01 and p<0.001 respectively) and bone marrow cellularity (WT: 28.1±1.9 x 10^{6}, *Nras*$^{G12D/+}$: 43.7±2.6 x 10^{6} and *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$: 41.3±3.2 x 10^{6}, p<0.01 compared to WT) (Figure 1B).

Expanded myelopoiesis and myeloproliferation were previously documented in single *Nras*$^{G12D/+}$ and *Flt3*$^{ITD/+}$ mutant mice.[6, 7] The addition of mutant *Npm1* augmented these phenotypes (Supplemental Figure S1A). In particular, total Mac-1^{+} splenocytes increased in number (27%-50% for *Nras*$^{G12D/+}$; and from 57-73% for *Flt3*$^{ITD/+}$). Notably, these cells were predominantly granulocytic (Mac-1^{+}/Gr-1^{+}) in *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ mice in contrast to cells from *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$ mice, which were predominantly monocytic Mac-1^{+}/Gr-1^{-} (Supplemental Figure S1A).

Nras$^{G12D/+}$ mice have been shown to have increased hematopoietic stem and progenitor cell (HSCP)numbers, due to increased proliferation and self-renewal of the HSC and MPP compartments.[12, 13] Our results confirm these data demonstrating significant increases in total myeloid progenitors, GMPs and CMPs, as well as the total number of early progenitors (LSK, and MPP) in both *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ and *Nras*$^{G12D/+}$ bone marrow cells when compared to WT controls (Figure 1C and Supplemental Figure S2A). Our data also reveal that *Nras*$^{G12D/+}$ stem and progenitor cell composition is largely unaltered by the addition of mutant NPM1. Also in agreement with previous studies, hematopoiesis in *Flt3*$^{ITD/+}$ mice was characterised by increased numbers of total myeloid progenitors (LK p<0.05 and GMPs p<0.01) and early progenitor populations (LSK, MPP and LMPP, p<0.01, p<0.01 and p<0.05 respectively) (Figure 1C and Supplemental Figure S2A). [14, 15] Of note, there was a substantial decrease in the size of the common lymphoid progenitor (CLP) population in *Flt3*$^{ITD/+}$ and *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ mice (Figure 1C) but not in single or compound *Nras*$^{G12D/+}$ mutants. This was in part due to the reduction in Il-7Rα-positive cells (Figure S2B). *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$

mice also exhibited robust increases in numbers of LK, GMP, LSK, MPP and LMPP populations (over what was observed with *Flt3*$^{ITD/+}$) when compared to WT. In direct comparison with *Flt3*$^{ITD/+}$ mutants, numbers of CMP and MEP progenitors in *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$ mice were reduced (from 55x10^3 to 16x10^3, $p<0.05$ and from 61x10^3 to 17x10^3, $p<0.05$), yet GMPs (proposed as direct descendants of CMPs [16]) remain expanded when compared to *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ mutants. This demonstrates that *Flt3*$^{ITD/+}$ mutant myelopoiesis is dramatically altered by the addition of *Npm1*$^{cA/+}$. Also, when compared to *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ mice, *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ mice showed an increase in LMPPs, reduction in lymphoid progenitors (CLP) and increase in GMPs (Figure 1E).

In order to assess the effects on the earliest detectable hematopoietic stem cell compartment (HSC) we opted to perform E-SLAM staining (CD45$^+$/EPCR$^+$/CD48$^-$/CD150$^+$).[17] Importantly, and unlike many other HSC FACS strategies, this does not rely on cell surface expression of FLT3, and reveals the percentage of E-SLAM detectable HSCs is decreased in *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ mice further so in *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$mutants (Figure 1D). Finally, using serial re-plating of bone marrow cells in semi-solid media we show that *Npm1*$^{cA/+}$ co-mutation markedly increased self-renewal of *Flt3*$^{ITD/+}$ cells (as shown previously [8]) and also of *Nras*$^{G12D/+}$ (Figure 1F).

The *Npm1*$^{cA/+}$ transcriptional signature persists in double mutant hemopoietic progenitors

To examine their combined effects on transcription we performed comparative global gene expression profiling of lineage negative (Lin$^-$) bone marrow cells using microarrays. *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ and *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ cells displayed a dramatically altered gene expression profile compared to single *Nras*$^{G12D/+}$ or *Flt3*$^{ITD/+}$ mutants (Figure 2A and Supplemental Figure S3B). Previously, we showed that mouse *Npm1*$^{cA/+}$ Lin$^-$ cells overexpressed several homeobox *(Hox)* genes (in particular overexpression of *Hoxa5*, *Hoxa7*, *Hoxa9* and two other homeobox genes, *Hopx* and *Nkx2-3*).[18] Here, we show that this signature, absent from *Nras*$^{G12D/+}$ or *Flt3*$^{ITD/+}$ singular mutant mice, persists in compound *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ and *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ Lin$^-$ progenitors. (Figure 2A, Supplemental Figure S3A-C). Gene Set Enrichment Analysis (GSEA) of *Npm1*$^{cA/+}$ single and compound mutant cell expression profiles, showed significant enrichment for genes up-regulated in NPM1-mutant and *MLL*-fusion gene positive human leukemias (Figure 2A).

Overexpression of the homeobox gene NKX2.3 in NPM1-mutant AML

Using the human TCGA AML dataset, we compared gene expression profiles of NPM1 mutant (NPM1c^{+ve}) to NPM1 wildtype (NPM1^{wt}) AML.[1] In agreement with previously published analyses, both

HOXA and *HOXB* genes were significantly overexpressed in NPM1c^{+ve} AML (Figure 2B).[19] We also noted that another homeobox gene, *NKX2-3*, was also overexpressed in keeping with our findings in *Npm1*$^{cA/+}$ mice (Figure 2A). Recently, *NKX2-3* overexpression was shown to be the most effective discriminant of *MLL-MLLT4* (*MLL-AF6*) driven AML to AML driven by other *MLL*-fusion genes. [20] Whilst overexpression of *Hox* genes such as *Hoxa9* has been shown to impart increased self-renewal and proliferation of hematopoietic progenitors, the effects of *Nkx2-3* overexpression are unknown. [21] To study this we performed retroviral gene transfer of fluorescently tagged *Nkx2-3*-CFP and *Hoxa9*-GFP into wildtype and *Flt3*$^{ITD/+}$ Lin^{-} bone marrow cells. Cells were subsequently sorted and plated in semi-solid methylcellulose for colony formation assays (Figure 2Ci). We find that overexpression of *Nkx2-3* increases clonogenic potential, albeit to a lesser extent compared to *Hoxa9* overexpression, in both wildtype and *Flt3*$^{ITD/+}$ progenitors. Notably, this is not augmented in combined transfected cells. (Figure 2Cii).

In order to mitigate the impact of the studied driver mutations on cell surface phenotypes, we performed transcriptome analysis on a homogeneous population of purified LSK-multipotent progenitor cells (MPPs, cell surface profile: Lin^{-}/CD34^{+}/Flt3^{+}/CD48^{+}/CD150^{-}) (Figure 2D). Transcription profiles of MPPs from single *Nras*$^{G12D/+}$ or *Flt3*$^{ITD/+}$ and the respective *Npm1*$^{cA/+}$ compound mutant MPPs revealed distinct transcriptional changes. In particular, compared to WT, both *Nras*$^{G12D/+}$ and *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ MPPs displayed similarly small numbers of differentially expressed genes yet only ~20% of these were shared (Figure 2Di). Also, gene/pathway enrichment analyses did not uncover significant overlap with a pre-established expression signature (data not shown). In contrast, the "addition" of *Npm1*$^{cA/+}$ to *Flt3*$^{ITD/+}$ in MPPs led to differential expression of a large number of additional genes, whilst also retaining most of the transcriptional changes attributable to *Flt3*$^{ITD/+}$ (Figure 2Dii). This demonstrates the powerful synergy between *Npm1*$^{cA/+}$ and *Flt3*$^{ITD/+}$ (Figure 2Dii, Tables S2a-b). Pathway analysis of genes differentially expressed in *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$ MPPs revealed enrichment of genes in the JAK-STAT pathway (Supplemental Figure 2E, Supplemental Tables S4a-b), including the negative regulators *Cish* and *Socs2* (Figure 2F). A number genes, encoding proteins involved in MAPK signalling are also deregulated (i.e. down-regulation of the MAPK pathway inhibitor, *Dusp6* and up-regulation of activators, *Rasgrp1*, *Rasgrf2* and *RasIIIb)* (Figure 2F). Other notable dysregulated genes included those involved in chromatin organisation (down-regulated, *Hdac10*, *Hdac11*, *Cbx7*, *Fbxl10/Kdm2b*, *Chd3*, *Satb1* and *H1F0*) and hematopoietic lineage or myeloid cell differentiation (*Bcl6*, *Bmp1*, *Lmo2*, *Ldb1* all down-regulated with *Cd74* and *Thy1* up-regulated) (Figure 2F, Supplemental Figure 3D). Of particular note is the down regulation of two positive regulators of murine lymphoid hemopoiesis, *Bcl11a* and *Kdm2a/Fbxl10* (knock-out mice of either gene are devoid of detectable CLPs[22, 23]). Many of the genes in our *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$ data set are also

present as deregulated in a recently published Tet2$^{-/-}$;*Flt3*$^{ITD/+}$ mouse model that also develops AML (76/418 genes, Supplemental Figure 3F and Supplemental Table 6,) which serves to verify our dataset but also reveals a distinguishing expression signature of FLT3-ITD which includes *Socs2*, *Id1*, *Csfr3r* and *Bcl11a*. In contrast a lack of correlation between deregulated gene sets of *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ and *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ MPPs (Supplemental Figure S3E) emphasises the molecular distinction between these compound mutants. [24]

Notably, *Hox* gene expression was not significantly altered in MPP populations from any of the *Npm1*$^{cA/+}$ models when compared to wildtype or to single *Nras*$^{G12D/+}$ and *Flt3*$^{ITD/+}$ mutants (Figure 2E and Figure S3C). These results are in agreement with observations that *Hox* gene expression in human NPM1c^{+ve} AML is comparable to that seen in normal HSCs and myeloid progenitors.[19] However, although we observe expanded myeloid progenitor populations, these data propose that the observed pattern of homeobox gene dysregulation is a direct molecular consequence of NPM1c rather than a change in cellular composition.

Npm1$^{cA/+}$ and _Nras_^{G12D} collaborate to promote high penetrance AML

To understand the leukemogenic potential of combined *Npm1*$^{cA/+}$ and *Nras*^{G12D} mutations, we aged cohorts of *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ mice along with *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$, single *Npm1*$^{cA/+}$, *Nras*^{G12D} and *Flt3*$^{ITD/+}$ mutant and wildtype control mice. Compound *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ and *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ mice had significantly reduced survival (median 138 and 52.5 days respectively) when compared to wildtype (618 days), *Npm1*$^{cA/+}$ (427 days), *Nras*$^{G12D/+}$ (315 days) and *Flt3*$^{ITD/+}$ (also 315 days) (Figure 3A, Supplemental Figure S4A). No difference in the survival of *Nras*$^{G12D/+}$ and *Flt3*$^{ITD/+}$ mutant mice was observed (p value = 0.85, see Supplemental Figure S4A for all comparisons). Moribund *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ and *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ mice displayed increased spleen and liver weights compared to moribund mice of other genotypes (Figure 3B). At time of sacrifice they also had significantly increased blood leukocyte (32.6±12, *Nras*$^{G12D/+}$ compared to 359±62 x10^6/L, *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$; and 151±34, *Flt3*$^{ITD/+}$ compared 250±33 x10^6/L, *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$) and reduced platelet counts (1046±227, *Nras*$^{G12D/+}$ compared to 504.9±209 x10^6/L, *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$; and 607±99, *Flt3*$^{ITD/+}$ compared to 225±25.8 x10^6/L, *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$). Infiltration of spleen tissue with myeloid cells was confirmed by FACS (Mac-1/Gr-1) (Supplemental Figure S4B). Independent histopathological analysis of formalin fixed paraffin embedded tissues from moribund mice (Figure 3C) revealed an increase in AML incidence from 41% (*Flt3*$^{ITD/+}$) to 100% in *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ samples and from 13% (*Nras*$^{G12D/+}$) to 85% in *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ samples (45% AML with maturation, AML$^+$ and 40% AML without maturation, AML$^-$ as defined by Bethesda classification) (Figure 3C).

Additional somatic mutations are required for progression to AML in *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ mice.

Npm1$^{cA/+}$;*Flt3*$^{ITD/+}$ mice succumb to AML very rapidly, compared to *Npm1*$^{cA/+}$ and *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$mice. We hypothesised that the slower onset of AML in the latter two genotypes -may be due to the requirement for additional cooperating somatic mutations. To test this, we performed whole exome sequencing and array comparative hybridisation (aCGH) of AMLs from *Npm1*$^{cA/+}$, *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$and *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ mice. We first confirmed the frequent development of loss-of-heterozygosity (LOH) at the *Flt3* locus in *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ AMLs [8, 25] and used quantitative amplicon sequencing (MiSeq; Methods, Supplemental Methods Figure S1) to quantify the *Flt3*^{ITD} variant allele fractions (VAFs), which were greater than 0.5 in 5/5 (range: 0.55-0.95, Figure 4Ai). Results of aCGH showed that LOH was copy-neutral (i.e. acquired uniparental disomy, Supplemental Figure 4Aii), with duplication of the *Flt3*^{ITD} allele. Using another quantitative amplicon sequencing assay we also show that recombination of the *Npm1*$^{flox\text{-}cA}$ allele was complete or near-complete (Figure S5A). Interestingly, aCGH of *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ samples revealed amplification of chr3 in 5/10 samples tested (Figure 4Bi). This was exclusively seen in *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ AMLs and mapped to a minimally amplified region (chr3: 102743581-103470550) which contained the genes *Nr1h5*, *Sike1*, *Csde1*, *Ampd1*, *Dennd2c*, *Bcas*, *Trim33* and *Nras* (Supplementary Table S10). This was confirmed by amplicon specific MiSeq PCR of the *Nras*^{G12D} allele and demonstrates gains of mutant *Nras* in these samples. This assay also identified gains of mutant *Nras* in 3 of 5 copy neutral samples by aCGH at the *Nras* locus (and 3 out of 4 further samples not assayed by aCGH). In summary, increased *Nras*^{G12D} dosage was detected in 11/14 *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ AMLs (Figure 4Bii). Staining of FFPE tissues from *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ AMLs for pERK1/2 (activated downstream of mutant RAS), demonstrated the level of RAS pathway activation approximately correlated with *Nras*^{G12D} gene dosage (Figure 4C).

Whole exome sequencing of AML samples revealed the average cumulative number of single nucleotide variants (SNVs) and small insertions or deletions (Indels) per AML sample was positively correlated to median survival rates. That is, *Npm1*$^{cA/+}$ 6.8±0.9, *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ 3.3±0.5 and *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$2.6±0.7 (mean number of mutations ± S.E.M.) (Figure 5A). Given the short latency to AML, it is unsurprising that *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ AMLs exhibit the fewest mutations. Moreover, *Npm1*$^{cA/+}$ AMLs spontaneously acquire mutations in genes involved in RAS signalling (*Nras*^{Q61H}, *Cbl*^{S374F}, *Ptpn11*^{S502L} and *Nf1*$^{W1260*/R683*}$) confirming this genetic interaction. Likewise, detection of a spontaneous tyrosine kinase domain mutation in Flt3, (*Flt3*^{D842G}) further confirms the importance of mutant FLT3 in the progression of mutant NPM1 AML (Figure 5B, Supplemental Table 9). Interestingly, another spontaneously acquired mutation, in a single *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ AML, occurs

in the Idh1 gene (R132Q) and although R132H/R132C are the commonly detected mutations in human AML,[1] the R132Q mutation has been demonstrated as being an equivalent pathogenic variant in chondrosarcoma.[26] aCGH also reveals the presence of complete or partial gain of chr7 in 7/8 *Npm1*$^{cA/+}$ and 4/9 *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ AMLs (Figure 5C and Supplemental Figure S5B). A minimally amplified region contains a number of genes previously implicated in leukemic transformation such as *Nup98*, *Wee1* and *Eed* (Supplemental Figure S5C).[18, 28, 29] The amplified region of murine chr7 is syntenic to human chr11 (Supplemental Table S11). Trisomy 11 has been reported in human AML and MDS, often in concert with *MLL*-PTD mutations.[30, 31] Single copy loss of chromatin modifying enzymes commonly found in NK-AML, *WT1*, *Asxl1* and *Dnmt3b* occurs simultaneously in one *Npm1*$^{cA/+}$ sample (Supplemental Figure S5D) and a focal deletion of Ezh2 is detected in a single *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ AMLs (Figure 5c). Other significant chromosomal gains include a region of chr6 in RN8 (which contains the oncogene *Ret* as well as *Kras*) and two instances of chr15 amplification, RN3 and RN6 which contain the *Ghr* gene, a common insertion site in our *Npm1*$^{cA/+}$ transposon insertional mutagenesis screen (Figure 5C, Supplemental Figure S6 and Supplemental Table 10 for full results).[18]

Discussion

Whilst the mutational drivers of AML and their patterns of co-occurrence are now well understood, the molecular basis for the frequency and prognostic impact of these patterns remains unknown. Of particular clinical relevance are the co-occurrence patterns of mutant *NPM1*, the equal most common mutation type in human AML.[1,2] Co-mutation of *NPM1* with *FLT3-ITD* is both significantly more frequent and carries a worse prognosis than co-mutation of *NPM1* with *NRAS or KRAS*. To understand the basis of this observation we investigated the interactions of these mutations in bespoke experimental models (Figure 1A). We first looked at the short-term impact of these mutations on hemopoiesis in young mice and confirmed that single *Npm1*$^{cA/+}$ mutant mice have normal bone marrow cellularity, white blood cell counts (WCC) and splenic weight.[18] Also, as described before, single *Flt3*$^{ITD/+}$ and *Nras*$^{G12D/+}$ had moderate but significant increases in splenic size, whilst *Nras*$^{G12D/+}$ also had raised WCC and bone marrow cellularity.[6,7] However, whilst the introduction of *Npm1*$^{cA/+}$ into the *Nras*$^{G12D/+}$ background did not alter these parameters significantly, the *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ co-mutation led to a dramatic rise in white cell count and splenic size (Figure 1B). At the cellular level, the *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ combination did not lead to significant changes on the size of progenitor and stem cell compartments when compared to *Nras*$^{G12D/+}$ alone. In sharp contrast, when compared to *Flt3*$^{ITD/+}$ single mutants, *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ mice displayed reductions in

CMP and MEP and increases in LSK progenitors. Furthermore, *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ mice showed a profound reduction in HSCs (Figure 1 C–E).

This differential impact of *Npm1*$^{cA/+}$ *on Flt3*$^{ITD/+}$ versus *Nras*$^{G12D/+}$ was reflected in the marked differences in gene expression profiles (GEP) between these two types of double-mutant mice. In fact the *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ model displayed only minimal differences to single *Nras*$^{G12D/+}$, whilst *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ lin- progenitors had profoundly different GEPs to *Flt3*$^{ITD/+}$ reflecting the dramatic impact of this combination on gene expression and in turn on progenitor cell fate. In order to discern the effects on transcription from changes in cellular composition, we studied gene expression in purified MPPs. Interestingly, whilst the impact of "adding" *Npm1*$^{cA/+}$ was much more dramatic with *Flt3*$^{ITD/+}$ than with *Nras*$^{G12D/+}$ (Supplemental Figure S3), the characteristic *Hox* gene signature of *Npm1*$^{cA/+}$ was not detectable in MPPs (Figure 2D-F). By contrast, other differences in MPP gene expression between *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ and *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ may have a role in the observed cellular phenotypes, for example Kdm2b, a critical gene for lymphopoiesis (is down-regulated in *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ (Figure 2F and Supplemental Figure S3D), which displays a significant reduction in CLPs.[22] Our transcriptome analysis of mouse progenitors and human AML reveals overexpression of *Nkx2-3* to be a distinguishing molecular feature of mutant NPM1. *NKX2-3* has been described as a prominently expressed and a novel distinguishing marker of MLL-AF6 and MLL-ENL from other of forms of MLL leukemia (fusion or PTD). [20, 32] Here we report this phenomenon in human NPM1c^{+ve} NK-AML. A full appreciation of *Nkx2-3* over-expression and how it may contribute to the effects on leukemogenesis has yet to be discerned and may warrant closer inspection, especially as it is suggestive of a common mechanism of deregulation of homeobox gene expression in these leukemia sub-types.

To study how these differences impact on leukemogenesis we aged double mutant mice and report that, like *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ animals (Mupo et al[8]), *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ mice also develop highly penetrant AML. However, AML latency was markedly shorter in *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ (median, 52.5 days) than in *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ (median, 138 days) mice. Interestingly single mutant *Flt3*$^{ITD/+}$ and *Nras*$^{G12D/+}$ mice did not display a different survival (Figure 3A) indicating that the interaction with *Npm1*^{cA} was central to this difference. Also, *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ AMLs lacked myeloid differentiation, which was frequently seen in *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ leukemias. To understand the genetic events involved in leukemic progression, we performed copy number analysis and exome sequencing of *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ and *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ AMLs. We found that the commonest somatic event during progression was an increase in mutant allele burden (*Nras*$^{G12D/+}$ or *Flt3*$^{ITD/+}$), through copy neutral LOH

(6/6 *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ and 3/8 *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ AMLs) or copy number gain (5/8 *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ AMLs).

Using faithful mouse models, our study provides strong and comparative evidence that the co-occurrence of mutant *NPM1* with *FLT3-ITD* is more formidable in AML potentiality, led by strikingly different molecular and cellular consequences, compared to its co-occurrence with mutant *NRAS*. This is a very plausible explanation for the frequent co-occurrence and worse prognosis of double mutant *NPM1c/ FLT3-ITD* AML.

Acknowledgements

We wish to thank Prof. Gary Gilliland for the *FLt3*^{ITD} mouse. OMD, JLC and GSV are funded by a Wellcome Trust Senior Fellowship in Clinical Science (WT095663MA). AM is funded by the Kay Kendall Leukaemia Fund project grant. CG was funded by a Bloodwise Clinical Research Training Fellowship. IV is funded by Spanish Ministerio de Economía y Competitividad subprograma Ramón y Cajal. We thank Servicio Santander Supercomputación for their support. GSV is a consultant for and holds stock in Kymab Ltd, and receives an educational grant from Celgene.

Authorship

Contribution: O.M.D., J.L.C., A.M., C.S.G., C.L., P.G. and G.S.V. performed mouse experiments. O.M.D. and G.S.V. analyzed results; P.W. and M.A. performed histopathological analysis of mouse samples; O.M.D., N.C. and R.M.A. performed microarray analysis; I.V. performed analysis of next generation sequencing; O.M.D. and G.S.V. designed the study. O.M.D. and G.S.V. wrote the paper with the help of K.S., T.J., R.R., P.W., M.A. and A.B.

Conflict of interest disclosure: GSV is a consultant for and holds stock in Kymab Ltd, and receives an educational grant from Celgene. All other authors declare no competing financial interests.

Correspondence: George Vassiliou, [1]The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, UK; e-mail: gsv20@sanger.ac.uk.

References

1. Network TCGAR. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med*. 2013; 368(22):2059-74.
2. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N Engl J Med*. 2016; 374(23):2209-21.
3. Falini B, Mecucci C, Tiacci E, et al. Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. *N Engl J Med*.

4. Welch JS, Ley TJ, Link DC, et al. The origin and evolution of mutations in acute myeloid leukemia. *Cell*. 2012; 150(2):264-78.
5. Vassiliou GS, Cooper JL, Rad R, et al. Mutant nucleophosmin and cooperating pathways drive leukemia initiation and progression in mice. *Nat Genet*.
6. Lee BH, Tothova Z, Levine RL, et al. FLT3 mutations confer enhanced proliferation and survival properties to multipotent progenitors in a murine model of chronic myelomonocytic leukemia. *Cancer Cell*. 2007; 12(4):367-80.
7. Li Q, Haigis KM, McDaniel A, et al. Hematopoiesis and leukemogenesis in mice expressing oncogenic NrasG12D from the endogenous locus. *Blood*. 2011; 117(6):2022-32.
8. Mupo A, Celani L, Dovey O, et al. A powerful molecular synergy between mutant Nucleophosmin and Flt3-ITD drives acute myeloid leukemia in mice. *Leukemia*. 2013;27(9):1917-20.
9. Kuhn R, Schwenk F, Aguet M, Rajewsky K. Inducible gene targeting in mice. *Science*. 1995;269(5229):1427-9.
10. Haigis KM, Kendall KR, Wang Y, et al. Differential effects of oncogenic K-Ras and N-Ras on proliferation, differentiation and tumor progression in the colon. *Nat Genet*. 2008;40(5):600-8.
11. McKerrell T, Park N, Moreno T, et al. Leukemia-associated somatic mutations drive distinct patterns of age-related clonal hemopoiesis. *Cell Rep*. 2015;10(8):1239-45.
12. Wang J, Kong G, Liu Y, et al. Nras(G12D/+) promotes leukemogenesis by aberrantly regulating hematopoietic stem cell functions. *Blood*. 2013;121(26):5203-5207.
13. Li Q, Bohin N, Wen T, et al. Oncogenic Nras has bimodal effects on stem cells that sustainably increase competitiveness. *Nature.* 2013 Dec 5;504(7478):143-147.
14. Shih AH, Jiang Y, Meydan C, et al. Mutational cooperativity linked to combinatorial epigenetic gain of function in acute myeloid leukemia. *Cancer Cell*. 2015;27(4):502-515.
15. Mead AJ, Kharazi S, Atkinson D, et al. FLT3-ITDs instruct a myeloid differentiation and transformation bias in lymphomyeloid multipotent progenitors. *Cell Rep*. 2013 Jun 27;3(6):1766-1776.
16. Akashi K, Traver D, Miyamoto T, Weissman IL. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature*. 2000 Mar 9; 404(6774):193-197.
17. Kent DG, Copley MR, Benz C, et al. Prospective isolation and molecular characterization of hematopoietic stem cells. *Blood*. 2009 Jun 18;113(25):6342-50
18. Vassiliou GS, Cooper JL, Rad R, et al. Mutant nucleophosmin and cooperating pathways drive leukemia initiation and progression in mice. *Nat Genet*. 2011;43(5):470-475.
19. Spencer DH, Yound MA, Lamprecht TL, et al. Epigenomic analysis of the HOX gene loci reveals mechanisms that may control canonical expression patterns in AML and normal hematopoietic cells. *Leukemia*. 2015;29(6):1279-1289.
20. Lavalee VP, Baccelli I, Krosl J, et al. The transcriptomic landscape and directed chemical interrogation of MLL-rearranged acute myeloid leukemias. *Nat Genet*. 2015;47(9):1030-1037.
21. Thorsteinsdottir U, Mamo A, Kroon E, et al. Overexpression of the myeloid leukemia-associated Hoxa9 gene in bone marrow cells induces stem cell expansion. *Blood*. 2002;99(1):121-129.
22. Andricovich J, Kai Y, Peng W, et al. Histone demethylase KDM2B regulates lineage commitment in normal and malignant hematopoiesis. *J Clin Invest*. 2016; 126(3):905-920.
23. Yu Y, Wang J, Khaled W, et al. Bcl11a is essential for lymphoid development and negatively regulates p53. *J Exp Med*. 2012; 209(13):2467-2483.

24. Shih AH, Jiang Y, Meydan C, et al. Mutational cooperativity linked to combinatorial epigenetic gain of function in acute myeloid leukemia. *Cancer Cell*. 2015 Apr 13;27 (4): 502-512.

25. Stirewalt DL, Pogosova-Agadjanyan EL, Tsuchiya K, et al. Copy-neutral loss of heterozygosity is prevalent and a late event in the pathogenesis of FLT3/ITD AML. *Blood Cancer J*. 2014;4:e208

26. Hirata M, Sasaki M, Cairns RA, et al. Mutant IDH is sufficient to initiate enchondromatosis in mice. *Proc Natl Acad Sci*. 2015;112(9):2829-2834.

27. Weisberg E, Nonami A, Chen Z, et al. Identification of Wee1 as a novel therapeutic target for mutant RAS-driven acute leukemia and other malignancies. *Leukemia*. 2015 Jan;29(1):27-37.

28. Shi J, Wang E, Zuber J, et al. The Polycomb complex PRC2 supports aberrant self-renewal in a mouse model of MLL-AF9;Nras(G12D) acute myeloid leukemia. *Oncogene*. 2013;32(7):930-938.

29. Danis E, Yamauchi T, Echanique K, et al. Inactivation of Eed impedes MLL-AF9-mediated leukemogenesis through Cdkn2a-dependent and Cdkn2a-independent mechanisms in a murine model. *Exp Hematol*. 2015;43(11):930-935.

30. Abdel-Wahab O, Levine R. The spliceosome as an indicted conspirator in myeloid malignancies. *Cancer Cell*. 2011;20(4):420-423.

31. Wang SA, Jabbar K, Lu G, et al. Trisomy 11 in myelodysplastic syndromes defines a unique group of disease with aggressive clinicopathologic features. *Leukemia*. 2010 Jan 24(4); 740-747

32. Garcia-Cuellar MP, Buttner C, Bartenhagen C, et al. Leukemogenic MLL-ENL Fusions Induce Alternative Chromatin States to Drive a Functionally Dichotomous Group of Target Genes. *Cell Rep*. 2016;15(2):310-322.

33. Dovey OM, Chen B, Mupo A, et al. Identification of a germline F692L drug resistance variant in cis with Flt3-ITD in knock-in mice. *Haematologica*. 2016;101(8):e328-331.

Figure Legends

Figure 1. Mutant *Npm1* co-operates with *Nras*-G12D and *Flt3-ITD* to enhance myeloid differentiation and enhance progenitor self-renewal.

(**A**) Schema for *Mx-1 Cre*, *Npm1*$^{flox\text{-}cA}$, *Nras*$^{LSL\text{-}G12D}$ and *Flt3*^{ITD} inter-crosses. (**B**) *Nras*$^{G12D/+}$ mice show a subtle and *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$ mice a marked increase in white cell count (WCC), compared to wildtype. Splenic sizes were significantly increased in all mutant genotypes except *Npm1*$^{cA/+}$, with *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$ showing the most striking phenotype. Bone marrow cellularity was increased only in the presence of the *Nras*$^{G12D/+}$ allele. (**C**) FACS analysis at 4-5 weeks after mutation induction. Gating strategies depicted are from wildtype mice. Significant differences in the stem and progenitor cell compartments of *Nras*$^{G12D/+}$ and *Flt3*$^{ITD/+,}$ but not *Npm1*$^{cA/+}$ single mutant mice, as previously reported. In double mutant mice, the *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ combination was not significantly different to *Nras*$^{G12D/+}$, in contrast to *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ which was markedly different to both *Flt3*$^{ITD/+}$ *and* *Npm1*$^{cA/+}$ single mutants. (**D**) Using a cell surface phenotype independent of FLT3 staining, we found that CD45+/EPCR+/CD150+/CD48- HSCs were reduced slightly in *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ and markedly in *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$mice. (**E**) Summary of hematopoietic effects of *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ and *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ double mutations in mice. LK, Lin^{-}/Kit^{+}; LSK, Lin^{-}/Sca-1^{+}/Kit^{+}; CMP, common myeloid progenitor; MEP, megakaryocyte-erythroid progenitor; GMP, granulocyte-monocyte progenitor;

MPP, multi-potent progenitor; LMPP, lymphoid primed multi-potent progenitor; CLP, common lymphoid progenitor and HSC, hematopoietic stem cell. (**F**) Single *Npm1*$^{cA/+}$ and double *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ or *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ mutant hematopoietic progenitors show increased self-renewal potential in whole bone marrow serial replating assays Mean ± SEM (n=4-8).

Figure 2. Impact of *Npm1*$^{cA/+}$ on the transcriptomes of *Nras*$^{G12D/+}$ and *Flt3*$^{ITD/+}$ mutant hematopoietic progenitors.

(**A**) Overlap of differentially expressed mRNAs reveals that *Npm1*$^{cA/+}$ has a dramatic impact on Lin- progenitor gene expression profiles when combined with *Flt3*$^{ITD/+}$, but only a modest impact when combined with *Nras*$^{G12D/+}$. Nonetheless, the characteristic hallmarks of *Npm1*$^{cA/+}$ are retained in both double mutant progenitors, namely overexpression of *Hoxa* genes and of the homeobox genes *Hopx* and *Nkx2-3* (also seen in single *Npm1*$^{cA/+}$ progenitors). Gene Set Enrichment Analysis reveals enrichment of differentially expressed genes from these models in human AMLs harboring mutant *NPM1* or *MLL* gene fusions (**B**) Comparison of human *NPM1*-mutant (*NPM1*^{c}) versus *NPM1-wildtype* (*NPM1*^{WT}) normal karyotype AML (NK-AML) also shows marked overexpression of *HOXA* and *HOXB* genes, as well as of *NKX2.3* raising the possibility that the latter may mediate some of the effect of *NPM1*^{c}. (**C**) Effects of *Nkx2-3* and *Hoxa9* over-expression on mouse hematopoietic progenitors. (i) Lin^{-} bone marrow progenitors from wildtype *and Flt3*$^{ITD/+}$ mice were transduced with MSCV-*Nkx2.3*-CFP and/or MSCV-*Hoxa9*-GFP constructs, maintained in liquid culture for 7 days, FACS sorted for CFP and GFP single and for double transfected cells and plated in semi-solid media. (ii) Colony assays of 2500 transduced cells show that both MSCV-*Hoxa9* and MSCV-*Nkx2-3* conferred an increase in self-renewal of both wildtype and *Flt3*$^{ITD/+}$ cells. However, double MSCV-*Hoxa9*/MSCV-*Nkx2-3* transfected cells showed no further changes in self-renewal when compared to MSCV-*Hoxa9* alone. Mean ± SEM (n=3); *p<0.05; **p<0.01; ***p<0.001; students t-test). (**D**) Sorting strategy for LSK/CD34^{+}/Flt3^{+}/CD48^{+} progenitor cells and overlap of differentially expressed genes (Illumina MouseWG-6 v2 Expression BeadChip) for (**i**) *Nras*$^{G12D/+}$ Vs *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ and (**ii**) *Flt3*$^{ITD/+}$ Vs *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ MPPs datasets. (**E**) Heat map of normlaised *Hox* gene expression in purified (**i**) MPP and (**ii**) Lin^{-} populations reveal that *Npm1*$^{cA/+}$ mutants (single or double) have similar patterns of *Hox* gene expression to wildtype (normalised average expression values are used to generate heat map values). (**F**) Differentially expressed genes in *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ MPPs vs wildtype controls .

Figure 3. *Npm1*^{cA} and *Nras*^{G12D} co-operate to drive high penetrance AML.

(**A**) Kaplan Meier survival curves of wildtype (n=23), *Npm1*$^{cA/+}$ (n=34), *Nras*$^{G12D/+}$ (n=40), *Flt3*$^{ITD/+}$ (n=39), *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ (n=46) and *Npm1*$^{cA/+}$;*Flt3*$^{ITD+/}$ (n=40). Double mutant (*Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ and *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$) mice had a significantly shortened survival when compared to single mutants, whilst *Npm1*$^{cA/+}$;*Flt3*^{ITD} had significantly shorter survival than *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ mice. (**B**) Spleen and liver weights, blood leukocyte (WCC) and platelet (Plts) counts of wildtype (n=13), *Npm1*$^{cA/+}$ (n=17), *Nras*$^{G12D/+}$ (n=22), *Flt3*$^{ITD/+}$ (n=30), *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ (n=15) and *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ (n=29), Mean ±SEM, one way ANOVA (Bonferroni adjusted) (*Vs wildtype, ⩵Vs *Flt3*$^{ITD/+}$, ♣Vs Nras$^{G12D/+}$). (**C**) Characteristic

histopathology from sick mice demonstrate increased incidence of AML in compound *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ and *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ mice compared to *Nras*$^{G12D/+}$ and *Flt3*$^{ITD/+}$ mice. Complete effacement of splenic tissue and infiltration of myeloid blast cells in liver tissue from *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ and *Npm1*$^{cA/+}$; *Flt3*$^{ITD+/}$ AMLs are presented. H&E, Haematoxylin and eosin; MPO, myeloperoxidase.

Figure 4. Leukemic progression in double mutant mice involves increased *Nras*^{G12D} or *Flt3*^{ITD} allele dosage

(**A**) Increase in *Flt3*^{ITD} allele burden in AMLs from *Npm1*^{cA}; *Flt3*^{ITD} mice through loss of heterozygosity for the locus. (**i**) *Flt3*^{ITD} amplicon sequencing (MiSeq) of leukemic bone marrow or spleen DNA (FN2-FN7). Tail DNA amplified from 2-week-old Flt3$^{+/+}$, *Flt3*$^{ITD/+}$, *Flt3*$^{ITD/ITD}$ mice was used as control. (**ii**) Normalised Log2 ratio plots show copy neutrality of chr5 and the Flt3 locus in 7/7 *Npm1*^{cA}; *Flt3*^{ITD} murine AMLs (FN-AMLs) tested. In-set: standard *Flt3*^{ITD} PCR genotyping of the same FN-AML samples; note reduction in the wildtype allele is visible. (**B**) (**i**) Summary of aCGH showing copy number gain at the *Nras* locus in AMLs RN6-10. (**ii**) Allele fractions for *Nras*^{wt} vs *Nras*^{G12D} show that copy number gains in RN6-10 involved *Nras*^{G12D}, and that an additional 3 cases (RN3-5) show copy-neutral loss-of-heterozygosity. In addition, two more RN AMLs show gains in mutant NRAS when measuring *Nras*^{wt} vs *Nras*^{G12D} –allele fractions (aCGH was not performed on these). Results of two *Npm1*$^{cA/+}$ samples are also shown for comparison purposes (N6, N7). (**C**) Increased gene dosage of *Nras*^{G12D} correlates with increased levels of phosphorylated RAS effectors pERK1/2. FN2,3,4,6,7= *Npm1*^{cA};*Flt3*^{ITD} mice, RN1-14= *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ mice.

Figure 5. Somatic mutations in *Npm1*$^{cA/+}$, *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ and *Npm1*^{cA}; *Flt3*^{ITD}AMLs. (**A**) Exome sequencing identifies an increased number of somatic nucleotide variants (SNVs) and small indels in *Npm1*$^{cA/+}$, compared to *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ (RN-AML) and *Npm1*^{cA}; *Flt3*^{ITD} (FN-AML) AML samples. Total AMLs sequenced: *Npm1*$^{cA/+}$ (n=12), *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ (n=14) and *Npm1*^{cA}; *Flt3*^{ITD} (n=7), mean/range, one way ANOVA (Bonferroni adjusted) (**=p<0.01 Vs Npm1$^{cA/+}$). (**B**) Summary of SNVs/Indels detected in AMLs from each genotype as indicated. Those in blue are genes mutated in the TCGA AML dataset. Those in red are exact or synonymous mutations detected in the TCGA AML dataset. (**C**) Co-occurance of SNVs and CNVs. Depicted are SNVs and focal copy number variations (CNVs) which have been formally detected in the TCGA AML[19] dataset or detected as common insertion sites (CIS) in our previously published Npm1$^{cA/+}$ Sleeping Beauty Transposon screen [25]. Mutant allele copy gains, chromosome gains and losses depicted. For copy number variation, colour coded boxes are based on log2 ratios (aCGH) and are not representative of CNV size. For a complete overview of all CNV and SNV co-occurrence see Supplemental Figure S6.

Figure 1.

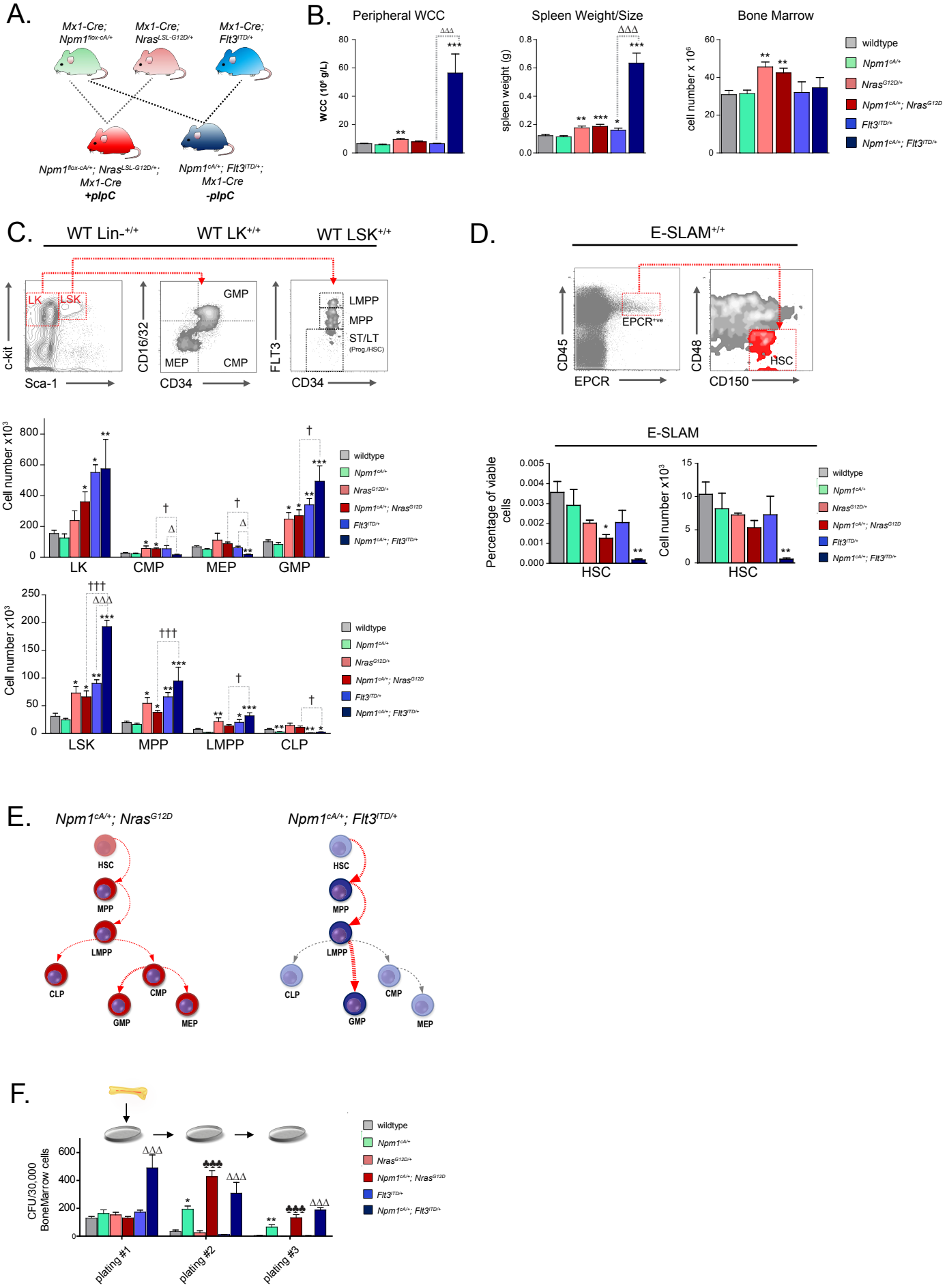

A.
Mx1-Cre; Npm1flox-cA/+
Mx1-Cre; NrasLSL-G12D/+
Mx1-Cre; Flt3ITD/+
Npm1flox-cA/+; NrasLSL-G12D/+; Mx1-Cre +pIpC
Npm1cA/+; Flt3ITD/+; Mx1-Cre -pIpC
B.
Peripheral WCC
Spleen Weight/Size
Bone Marrow
WCC (10⁶ g/L)
spleen weight (g)
cell number x 10⁶
wildtype
Npm1cA/+
NrasG12D/+
Npm1cA/+; NrasG12D
Flt3ITD/+
Npm1cA/+; Flt3ITD/+
C.
WT Lin-+/+
WT LK+/+
WT LSK+/+
LK
LSK
c-kit
Sca-1
CD16/32
CD34
GMP
MEP
CMP
FLT3
LMPP
MPP
ST/LT
(Prog./HSC)
Cell number x10³
LK
CMP
MEP
GMP
LSK
MPP
LMPP
CLP
D.
E-SLAM+/+
EPCR+ve
CD45
EPCR
CD48
CD150
HSC
E-SLAM
Percentage of viable cells
Cell number x10³
E.
Npm1cA/+; NrasG12D
Npm1cA/+; Flt3ITD/+
HSC
MPP
LMPP
CLP
CMP
GMP
MEP
F.
CFU/30,000 BoneMarrow cells
plating #1
plating #2
plating #3

Figure 2

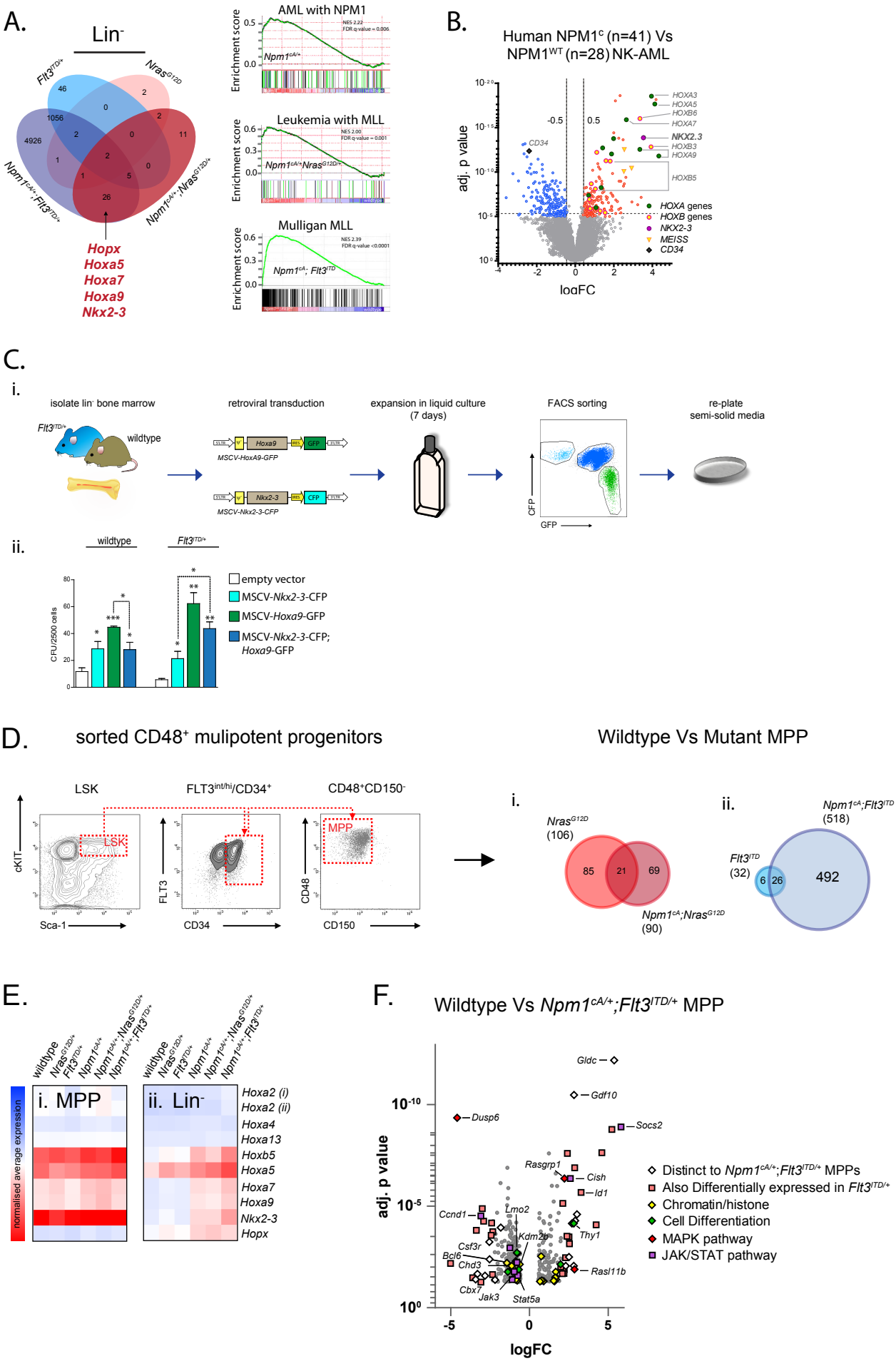
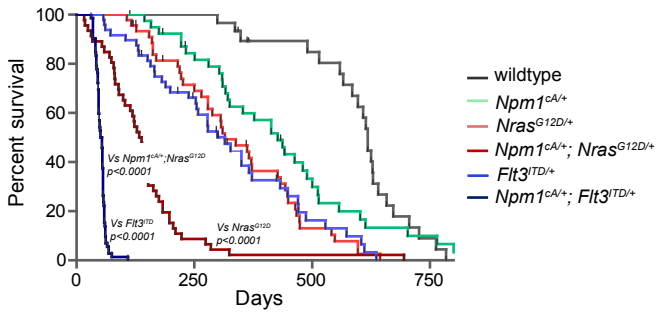
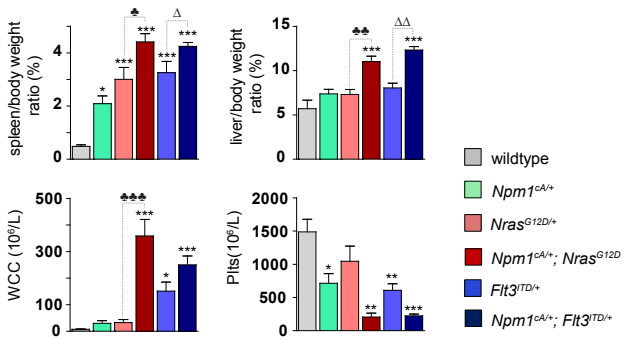

A.
Lin-
Flt3ITD/+
Nras G12D
Npm1cA/+;Flt3ITD/+
Npm1cA/+;Nras G12D/+
46
2
0
1056
2
2
4926
2
0
11
1
2
0
1
5
26
Hopx
Hoxa5
Hoxa7
Hoxa9
Nkx2-3
AML with NPM1
Leukemia with MLL
Mulligan MLL
Enrichment score
B.
Human NPM1c (n=41) Vs
NPM1WT (n=28) NK-AML
adj. p value
logFC
-0.5
0.5
CD34
HOXA3
HOXA5
HOXB6
HOXA7
NKX2.3
HOXB3
HOXA9
HOXB5
HOXA genes
HOXB genes
NKX2-3
MEISS
CD34
C.
i.
isolate lin- bone marrow
retroviral transduction
expansion in liquid culture
(7 days)
FACS sorting
re-plate
semi-solid media
wildtype
MSCV-HoxA9-GFP
MSCV-Nkx2-3-CFP
CFP
GFP
ii.
wildtype
Flt3ITD/+
CFU/2500 cells
empty vector
MSCV-Nkx2-3-CFP
MSCV-Hoxa9-GFP
MSCV-Nkx2-3-CFP;
Hoxa9-GFP
D.
sorted CD48+ mulipotent progenitors
Wildtype Vs Mutant MPP
LSK
FLT3int/hi/CD34+
CD48+CD150-
cKIT
Sca-1
FLT3
CD34
CD48
CD150
MPP
i.
Nras G12D
(106)
85
21
69
Npm1cA;Nras G12D
(90)
ii.
Npm1cA;Flt3ITD
(518)
Flt3ITD
(32)
6
26
492
E.
wildtype
Nras G12D/+
Flt3ITD/+
Npm1cA/+
Npm1cA/+;Nras G12D/+
Npm1cA/+;Flt3ITD/+
i. MPP
ii. Lin-
normalised average expression
Hoxa2 (i)
Hoxa2 (ii)
Hoxa4
Hoxa13
Hoxb5
Hoxa5
Hoxa7
Hoxa9
Nkx2-3
Hopx
F.
Wildtype Vs Npm1cA/+;Flt3ITD/+ MPP
adj. p value
logFC
Gldc
Gdf10
Dusp6
Socs2
Rasgrp1
Cish
Id1
Ccnd1
Lmo2
Kdm2b
Thy1
Csf3r
Bcl6
Chd3
Rasl11b
Cbx7
Jak3
Stat5a
Distinct to Npm1cA/+;Flt3ITD/+ MPPs
Also Differentially expressed in Flt3ITD/+
Chromatin/histone
Cell Differentiation
MAPK pathway
JAK/STAT pathway

Figure 3

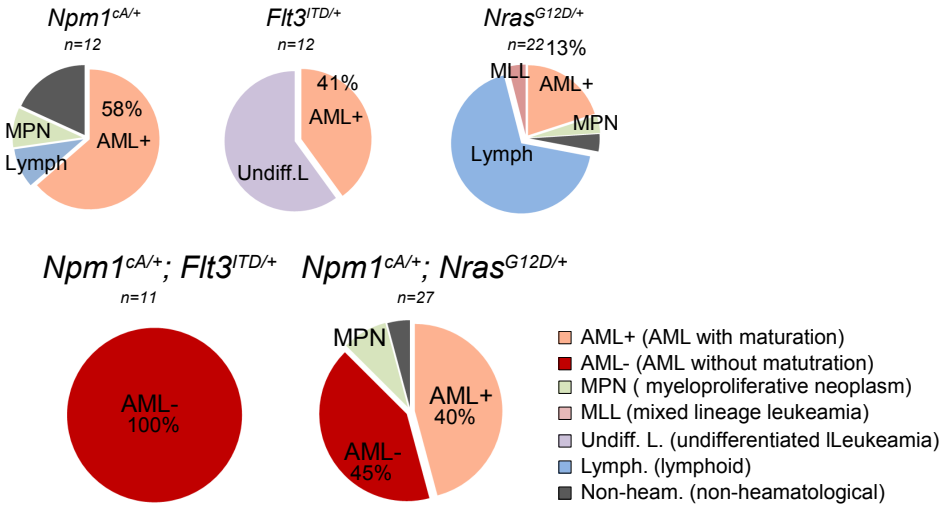
A.


Percent survival
Days
wildtype
Npm1cA/+
NrasG12D/+
Npm1cA/+; NrasG12D/+
Flt3ITD/+
Npm1cA/+; Flt3ITD/+
Vs Npm1cA/+;NrasG12D p<0.0001
Vs Flt3ITD p<0.0001
Vs NrasG12D p<0.0001

B.


spleen/body weight ratio (%)
liver/body weight ratio (%)
WCC (10⁶/L)
Plts(10⁶/L)
wildtype
Npm1cA/+
NrasG12D/+
Npm1cA/+; NrasG12D
Flt3ITD/+
Npm1cA/+; Flt3ITD/+

C.


Npm1cA/+
n=12
58%
AML+
MPN
Lymph
Flt3ITD/+
n=12
41%
AML+
Undiff.L
NrasG12D/+
n=22
13%
MLL
AML+
MPN
Lymph
Npm1cA/+; Flt3ITD/+
n=11
AML-
100%
Npm1cA/+; NrasG12D/+
n=27
MPN
AML+
40%
AML-
45%
AML+ (AML with maturation)
AML- (AML without matutration)
MPN (myeloproliferative neoplasm)
MLL (mixed lineage leukeamia)
Undiff. L. (undifferentiated lLeukeamia)
Lymph. (lymphoid)
Non-heam. (non-heamatological)

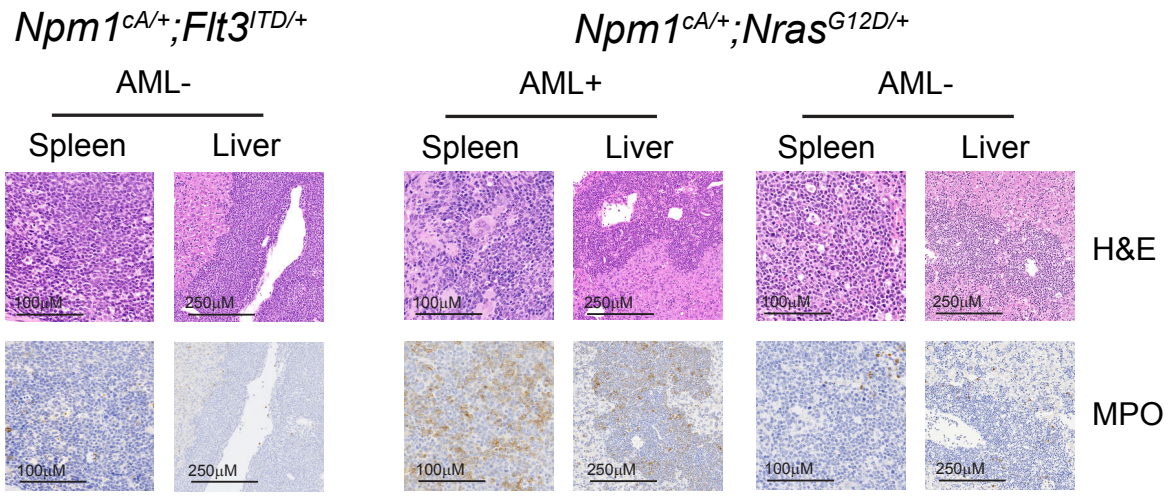
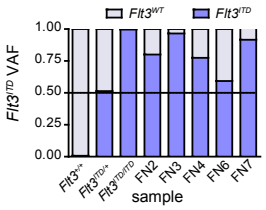

Npm1cA/+;Flt3ITD/+
AML-
Spleen
Liver
Npm1cA/+;NrasG12D/+
AML+
AML-
Spleen
Liver
Spleen
Liver
H&E
MPO
100µM
250µM

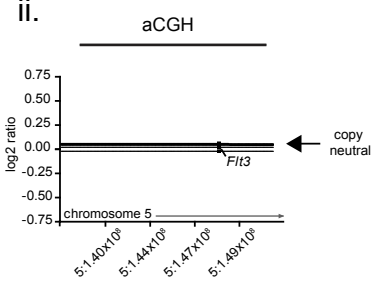
Figure 4

A. *Npm1*^{cA}; *Flt3*^{ITD}

i.

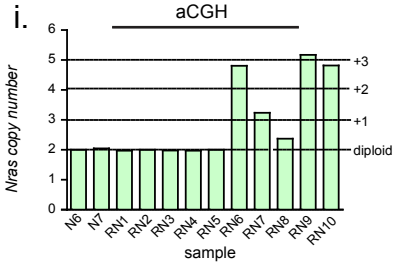


ii.

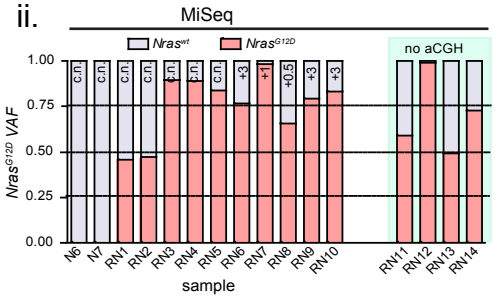


B. *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$

i.



ii.



C. *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$

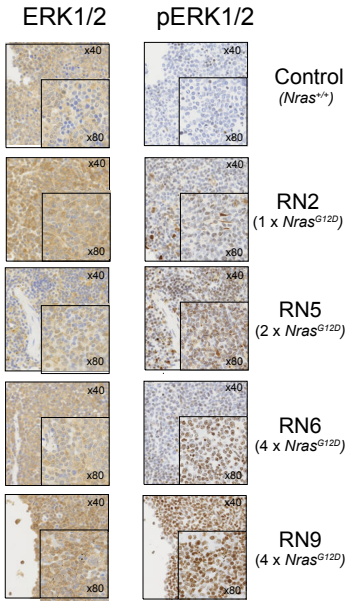
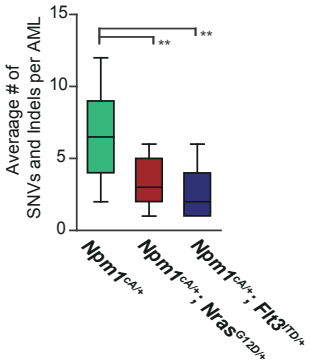
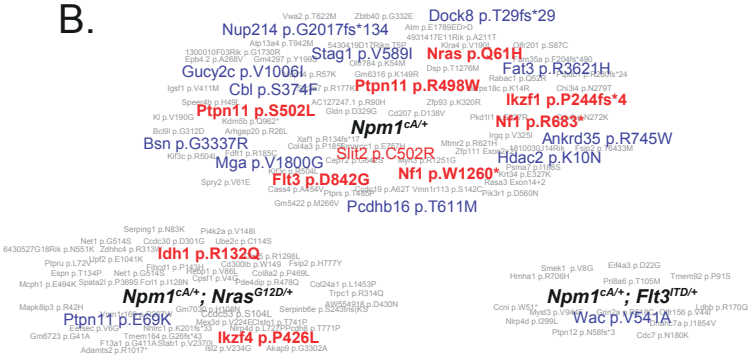


Figure 5

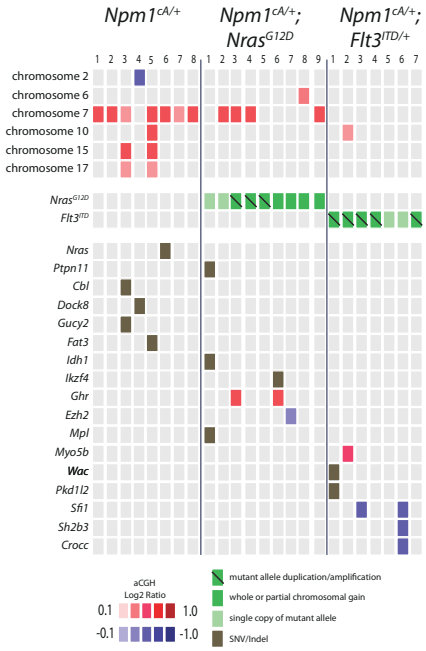
A.


Averaage # of SNVs and Indels per AML
15
10
5
0
**
**
Npm1cA/+
Npm1cA/+; NrasG12D/+
Npm1cA/+; Flt3ITD/+

B.


Dock8 p.T29fs*29
Nup214 p.G2017fs*134
Stag1 p.V589I
Nras p.Q61H
Gucy2c p.V1006I
Fat3 p.R3621H
Cbl p.S374F
Ptpn11 p.R498W
Ikzf1 p.P244fs*4
Ptpn11 p.S502L
Npm1cA/+
Nf1 p.R683*
Bsn p.G3337R
Ankrd35 p.R745W
Slit2 p.C502R
Hdac2 p.K10N
Mga p.V1800G
Nf1 p.W1260*
Flt3 p.D842G
Pcdhb16 p.T611M
Idh1 p.R132Q
Npm1cA/+; NrasG12D/+
Ptpn11 p.E69K
Ikzf4 p.P426L
Npm1cA/+; Flt3ITD/+
Wac p.V541A

C.


Npm1cA/+
Npm1cA/+; NrasG12D
Npm1cA/+; Flt3ITD/+
chromosome 2
chromosome 6
chromosome 7
chromosome 10
chromosome 15
chromosome 17
NrasG12D
Flt3ITD
Nras
Ptpn11
Cbl
Dock8
Gucy2
Fat3
Idh1
Ikzf4
Ghr
Ezh2
Mpl
Myo5b
Wac
Pkd1l2
Sfi1
Sh2b3
Crocc
aCGH
Log2 Ratio
0.1
1.0
-0.1
-1.0
mutant allele duplication/amplification
whole or partial chromosomal gain
single copy of mutant allele
SNV/Indel

Supplemental Methods

Flow cytometry and cell selection

Progenitor populations were defined as follows;

HSCs (CD45$^+$, EPCR$^+$, CD48$^-$, CD150$^+$)

Short term-long term-progenitor-HSCs (ST/LT$^{prog/HSC}$) (Lin$^-$, c-kit$^+$, Sca-1$^+$, Flt3$^-$)

Multipotent progenitors (MPP) (Lin$^-$, c-kit$^+$, Sca-1$^+$, Flt3$^+$)

Lymphoid primed multipotent progenitors (LMPP) (Lin$^-$, c-kit$^+$, Sca-1$^+$, Flt3^{hi})

Common lymphoid progenitors (CLP) (Lin$^-$, Flt3^{hi}, Il-7rα^+, c-kit^{int}, Sca-1^{int})

Granulocyte-monocyte progenitors (GMP) (Lin$^-$, Il-7rα^-, c-kit$^+$, Sca-1$^-$, CD34$^+$, CD16/32$^+$)

Common myeloid progenitors (CMP) (Lin$^-$, Il-7rα^-, c-kit$^+$, Sca-1$^-$, CD34$^+$, CD16/32$^-$)

Megakaryocyte-erythroid progenitors (MEP) (Lin$^-$, Il-7rα^-, c-kit$^+$, Sca-1$^-$, CD34$^-$, CD16/32$^-$)

Antibodies were obtained from eBiosciences or BioLegend unless stated. Markers for lineage depletion (Lin$^-$) included B220 (RA3-6B2), CD3ε (145-2C11), Mac-1 (M1/70), Gr-1 (RB6-8C5) and Ter119 (TER-119). E-PCR (RMEPCR1560, Stem Cell Technologies), CD45 (30-F11), CD150 (TC15-12F12.2), CD48 (HM48-1), CD34 (RAM34), CD16/32 (2.4G2), FLT3 (A2F10), c-kit (2B8), Sca-1 (E13-161.7) and Il7-rα (A7R34).

For MPP gene expression, sorting was defined as: Lin$^-$, c-kit$^+$, Sca-1$^+$, CD34$^+$, CD48$^+$ and CD150$^-$. Samples were flow-sorted after removal of lineage positive cells using a magnetic activated cell sorting (MACS) mouse lineage depletion kit containing antibodies for CD5, B220, CD11b, Gr-1 (Ly-6G/C), 7-4, and Ter-119 following the manufacturers' instructions (Miltenyi Biotec).

Gene expression profiling, microarrays

Gene expression profiles of lineage negative (lin$^-$) or 1,000 sorted multipotent progenitors (MPPs) from mutant and wildtype controls were compared using the Illumina MouseWG-6 v2 Expression BeadChip platform (Illumina). Lin$^-$ populations were separated from whole bone marrow using magnetic activated cell sorting (MACS, Miltenyi Biotec) and RNA isolated using a standard Trizol (Thermofisher) protocol. Flow sorted MPP populations were sorted directly into Trizol LS (Thermofisher) using a Mo-Flow™ XDP (Beckman Coulter) and RNA extracted according to the manufacturer. Extracted RNA was prepared for array hybridization using the TargetAmp™-Nano Labeling Kit (Epicentre). Global profiling was done using Illumina MouseWG-6 v2.0 Expression BeadChip. Data were quantile normalized and analyzed using the Bioconductor, *lumi* and *limma* [2]packages with *P* values adjusted for multiple testing (Bioconductor, http://www.bioconductor.org/; lumi, http://www.bioconductor.org/packages/2.0/bioc/html/lumi.html; RTCGD, http://rtcgd.ncifcrf.gov/).[1-3]

Adjusted *P* value (<0.05) was used to identify significantly differentially expressed genes. Gene set enrichment analysis was carried out using GSEA v2.1.0 (Broad Institute).[4, 5] (All data is deposited into ArrayExpress under the following accession numbers, E-MTAB-5358, E-MTAB-5359 and E-MTAB-5361.)

Comparative gene expression analysis of NPM1+ve and NPM1-ve AML samples from the Cancer Genome Atlas (TCGA) , generated using the Affymetrix Human Genome U133 Plus 2.0 Array (Affymetrix) was performed using the GCRMA (http://www.bioconductor.org/repository/devel/vignette/gcrma.pdf), *limma* and *affy* packages in Bioconductor .[2, 6, 7] (.CEL file sample IDs are listed in Supplemental Table 1.1, datasets and samples are listed in Supplemental Table 2, note that samples positive for *MLL* gene fusions were removed from the control sample set.)

Copy number variation, comparative genomic hybridisation (aCGH)

DNA copy number variation in leukemic samples was assessed using the Mouse Genome Comparative Genomic Hybridization 244K Microarray (Agilent Technologies). DNA was labeled with Cy3 or Cy5 according to Agilent aCGH genomic labeling protocol (Agilent Technologies). Raw data was extracted using Agilent Feature Extraction and normalised using R Package aCGH Spline. Subsequent data analysis was performed in R using aCGH Bioconductor packages (http://www.bioconductor.org).[7] (All data is deposited into ArrayExpress under the following accession number, E-MTAB-5356.)

Mouse AML mutation calling and validation.

Sequence reads were aligned against the reference mouse genome (GRCm38) using the Burrows-Wheeler algorithm (BWA; specifically, *aln* for HiSeq--paired-end exome sequencing data and *mem* for MiSeq-250bp-paired-end sequencing data). For the detection of the *Flt3*^{ITD} and *Npm1*$^{flox\text{-}cA}$ or *Npm1*^{cA} alleles, a fasta entry containing these sequences was appended to the reference genome. Sam/bam files were sorted and indexed using SAMTOOLS.[8] Where necessary we also performed PCR duplicate marking using PICARD tools (http://picard.sourceforge.net) and local realignment around indels using GATK.[9] The in-house software RAMSES, was used to detect somatic mutations and indels identified using PINDEL.[10] Functional consequences of mutations were predicted using an in-house script employing Ensembl Perl API.[11] All potential transcript annotations were calculated and recorded. The most deleterious of all potential annotations was reported for each mutation. All data is submitted to the European Nucleotide Archive (ENA study accession PRJEB18526, secondary study ERP020464).

Further details of the exome sequencing and amplicon specific validation workflow are outlined in Supplemental Methods Figure S1.

Retroviral transduction

Cloning of mouse *Nkx2-3 and Hoxa9* into MSCV-GFP/CFP retroviral backbones.

mRNA extracted from homozygous wildtype C57BL/6N mouse bone marrow cells was reverse transcribed using SuperScript III (Invitrogen) and the subsequent cDNA was used as template to amplify full length *Nkx2-3* or *Hoxa9* cDNA using high fidelity *taq polymerase* (KAPA HiFi HotStart ReadyMix, Kapa Biosystems) using the manufacturer's instructions and the following primers;

EcoRI-*mNkx2-3*-XhoI Fwd:

gaattcgccaccatgatgttaccaagcccggtcacctccacccctttctc

EcoRI-*mNkx2-3*-XhoI Rev:
tcgagtcacttgtcgtcatcgtctttgtagtcaatgtcatgatccttgtaatcgccgtcgtgccaagccctgatgccctgcaaagtcccctgcgtgcacg

This fragment was cloned into an EcoRI/XhoI linearized fragment obtained from the MSCV-IRES-GFP (Addgene plasmid # 20672) retroviral backbone using standard molecular biology techniques.

MluI-*Hoxa9*-XhoI Fwd:

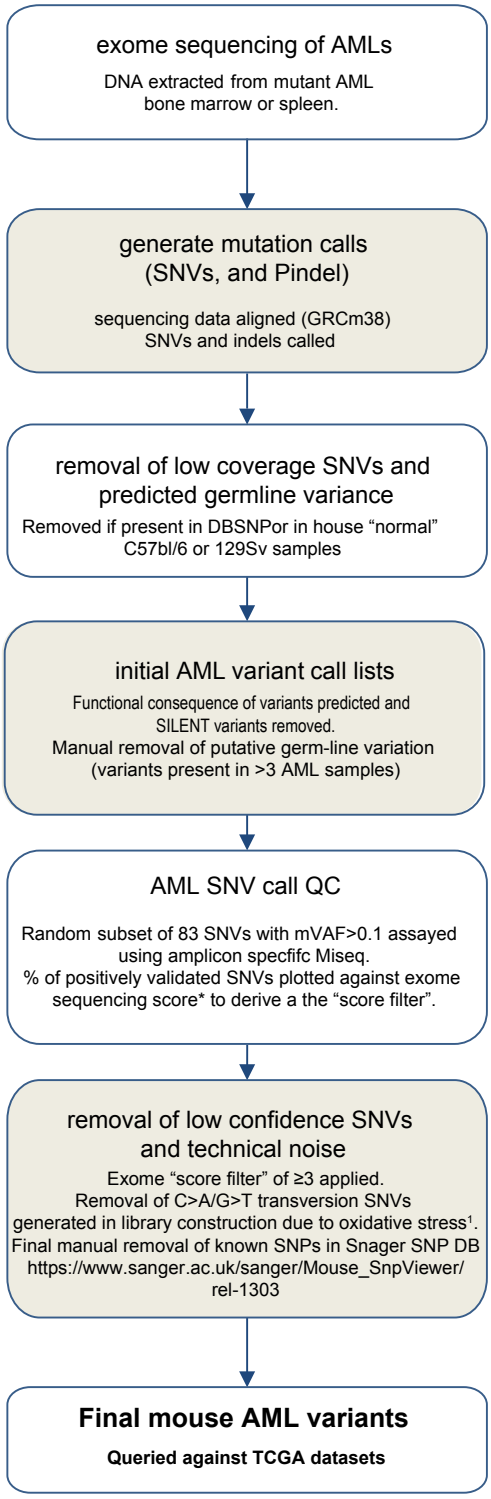
aattcacgcgtatggccaccaccggggccctgggcaactactatgtggac

MluI- *Hoxa9*-XhoI Rev:
ctcgagttaagcgtaatctggaacatcgtatgggtagccgtcgtgctcgtcttttgctcggtccttgttgattttcttcattttcatcctgcggttctgg

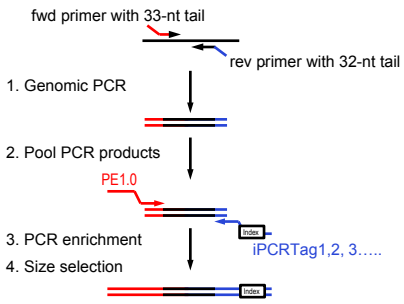
This fragment was cloned into a MluI/XhoI linearized fragment of MSCV-IRES-CFP, a kind gift from Dr Brian Huntley.

Supplemental Methods Figure S1

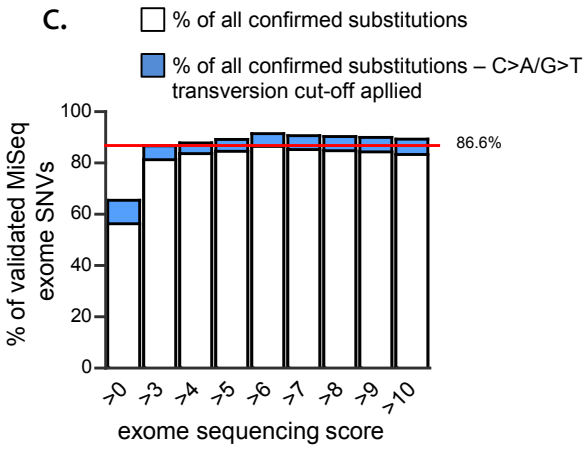
a.



b.



c.



Supplemental Table Legends

Supplemental Table S1. Comparatively altered gene expression in lineage negative bone marrow aspirates compared to wildtype (significantly differentially expressed genes, adj. $p<0.05$).

Supplemental Table S2. TCGA datasets[6] used for comparative gene expression analysis of human AML.

Supplemental Table S3. Comparatively altered gene expression in human AML based on NPM1 mutation status (for significantly differentially expressed genes adj.$p<0.05$).

Supplemental Table S4a. Comparative gene (probes) expression in multipotent progenitors, *Nras*$^{G12D/+}$ and *Npm1*$^{cA/+}$;*Nras*$^{G12D/+}$ compared to wildtype.

Supplemental Table S4b. Comparative gene (probes) expression in multipotent progenitors, *Flt3ITD*$^{/+}$ and *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$compared to wildtype.

Supplemental Table S5a. Kegg Pathways enriched in *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ multipotent progenitor cells (DAVID[12]).

Supplemental Table S5a. Functional annotational clustering of gene ontology terms (GO-term) enriched in *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ multipotent progenitor cells (DAVID[12]).

Supplemental Table S6. Overlap of differentially expressed genes in *Tet2*$^{-/-}$;*Flt3*$^{ITD/+}$ LSK cells and *Npm1*$^{cA/+}$;*Flt3*$^{ITD/+}$ MPP cells compared to wildtype.

Supplemental Table S7a. 1st Round MiSeq amplicon specific primer sequences.

Supplemental Table S7b. 2nd Round MiSeq iPCR-tag primer sequences.

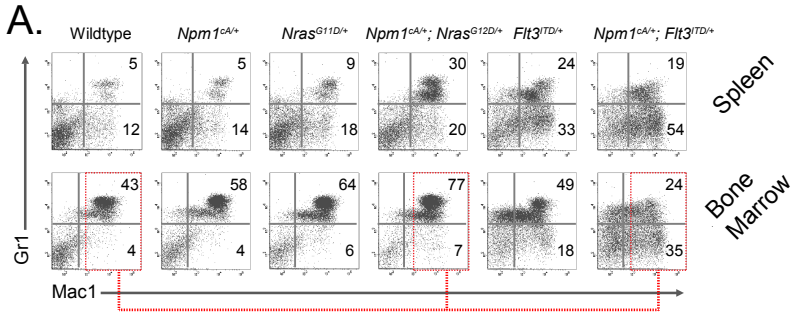
Supplemental Table S8. Validation of exome sequencing: MiSeq amplicon sequencing results (mVAF).

Supplemental Table S9. Combined Single Nucleotide Variant (SNV) and insertions/deletions (Indels) detected by the Exome sequencing pipeline (detailed in Methods and outlined in Supplemental Methods Figure S1)

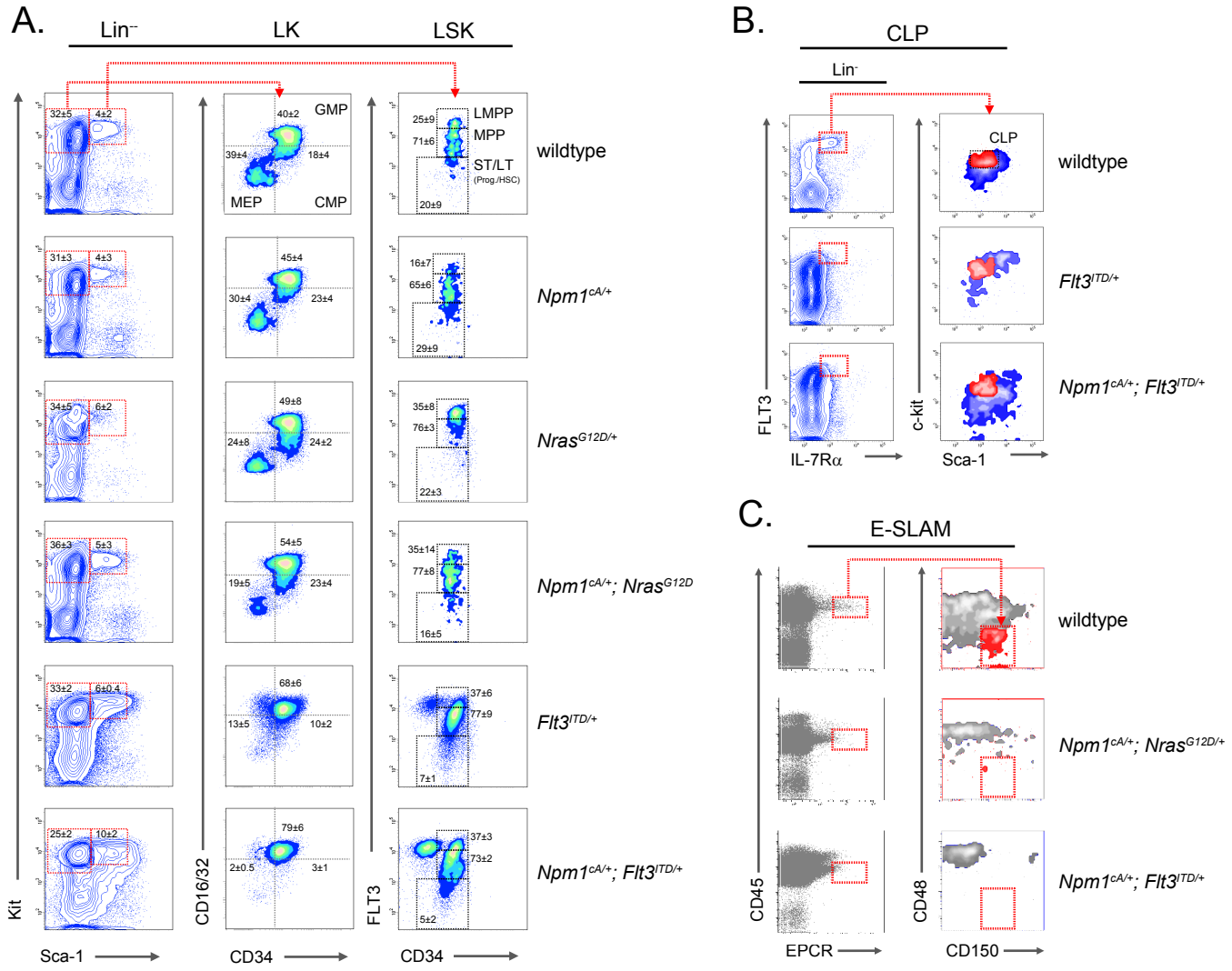
Supplemental Table S10. aCGH results summary.

Supplemental Table S11. Mouse-Human synteny of chromosome regions with altered copy number, identified by aCGH, in murine AMLs. (Only genes identified as mutated in the TCGA AML data-set or hits in our Npm1$^{cA/+}$ Sleeping Beauty insertional mutagenesis screen are included.)

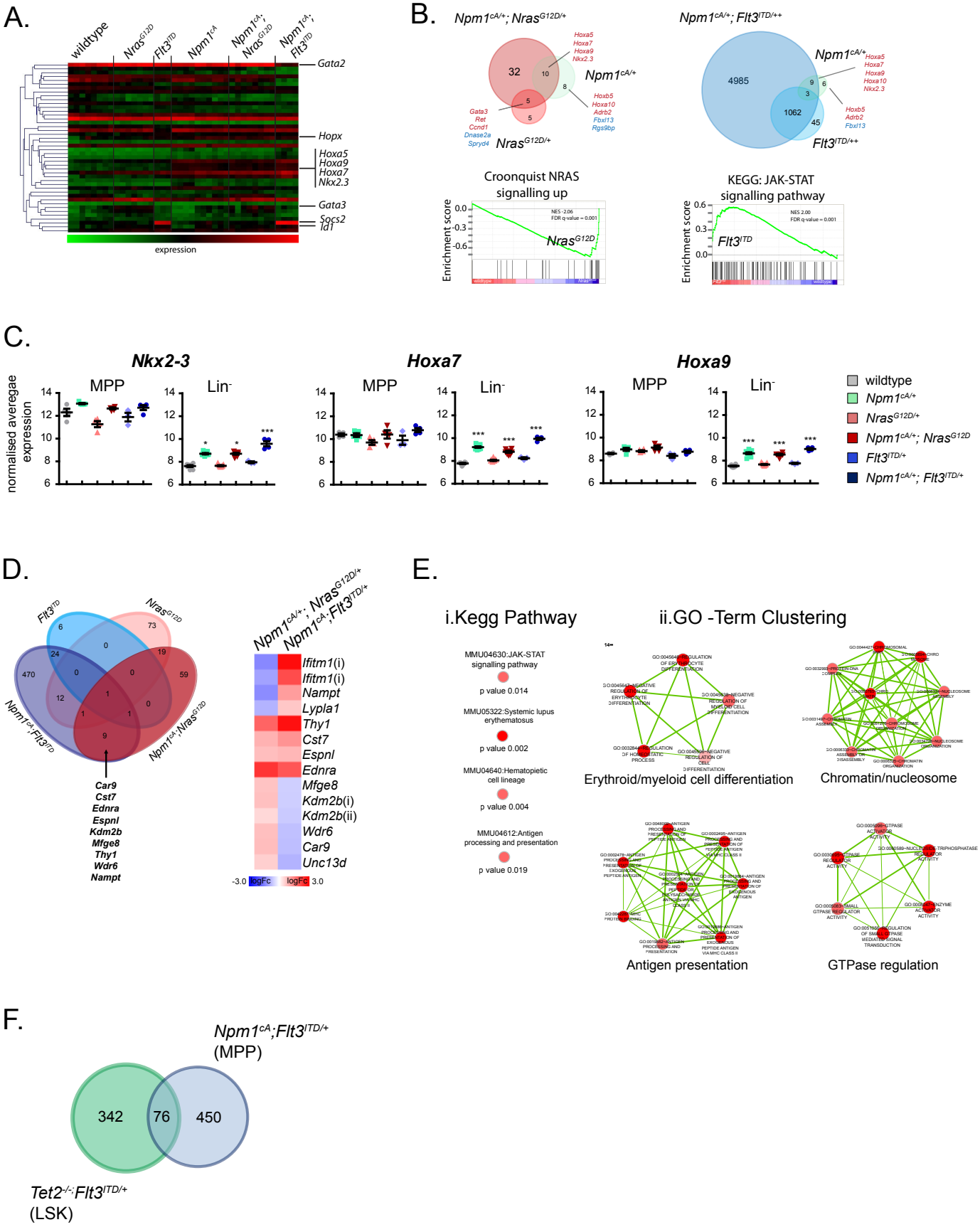
Supplemental Figure S1



Supplemental Figure S2



Supplemental Figure S3



Supplemental Figure S4

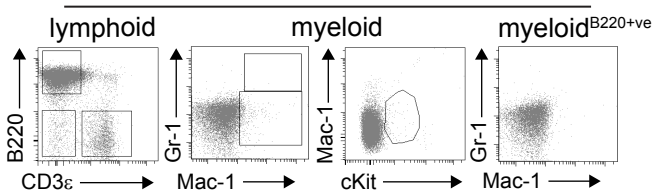
A.

Mantel-Cox Test

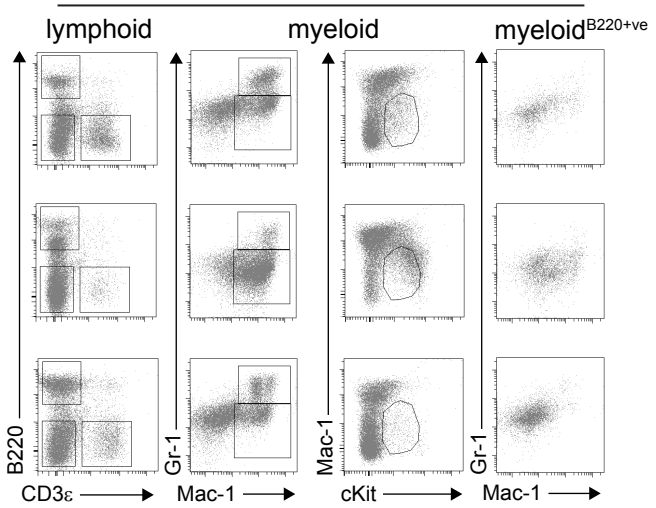
	wildtype	*Npm1*$^{cA/+}$	*Nras*$^{G12D/+}$	*Npm1*$^{cA/+}$; *Nras*^{G12D}	*Flt3*$^{ITD/+}$	*Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$
Npm1$^{cA/+}$	p<0.05	p<0.05				
Nras$^{G12D/+}$	p<0.0001	**p<0.0001**	**p<0.0001**			
Npm1$^{cA/+}$; *Nras*^{G12D}	p<0.0001	p<0.05	**n.s.**	p<0.0001		
Flt3$^{ITD/+}$	p<0.0001	**p<0.0001**	p<0.0001	**p<0.0001**	p<0.0001	
Npm1$^{cA/+}$; *Flt3*$^{ITD/+}$	p<0.0001	**p<0.0001**	p<0.0001	**p<0.0001**	**p<0.0001**	
median survival (days)	618	427	315	138	315	52.5

B.

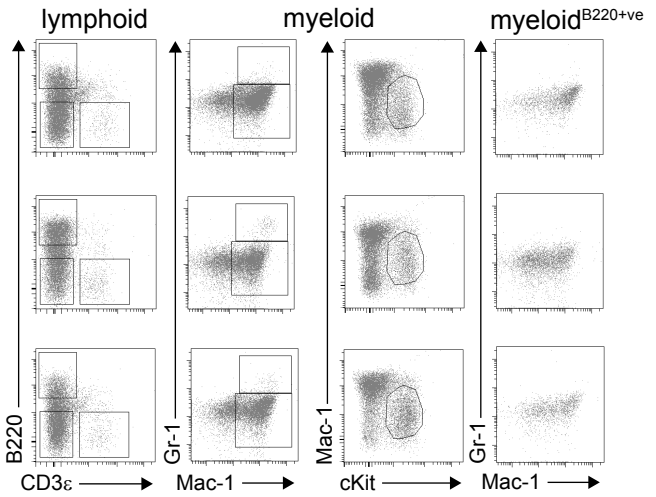
wildtype splenocytes



Npm1^{cA}; *Nras*$^{G12D/+}$ AML

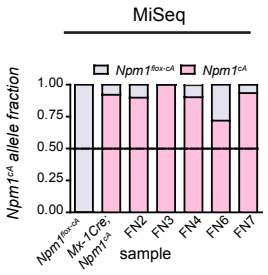


Npm1^{cA}; *Flt3*$^{ITD/+}$ AML

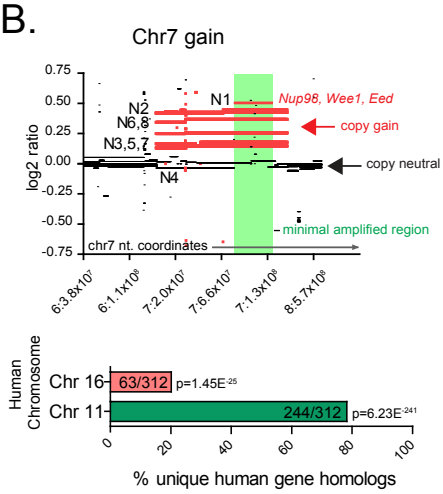


Supplemental Figure S5

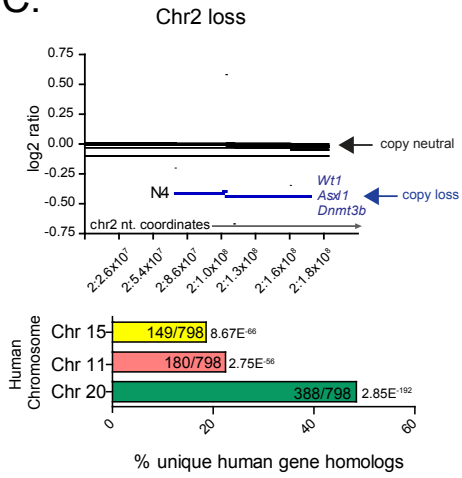
A.



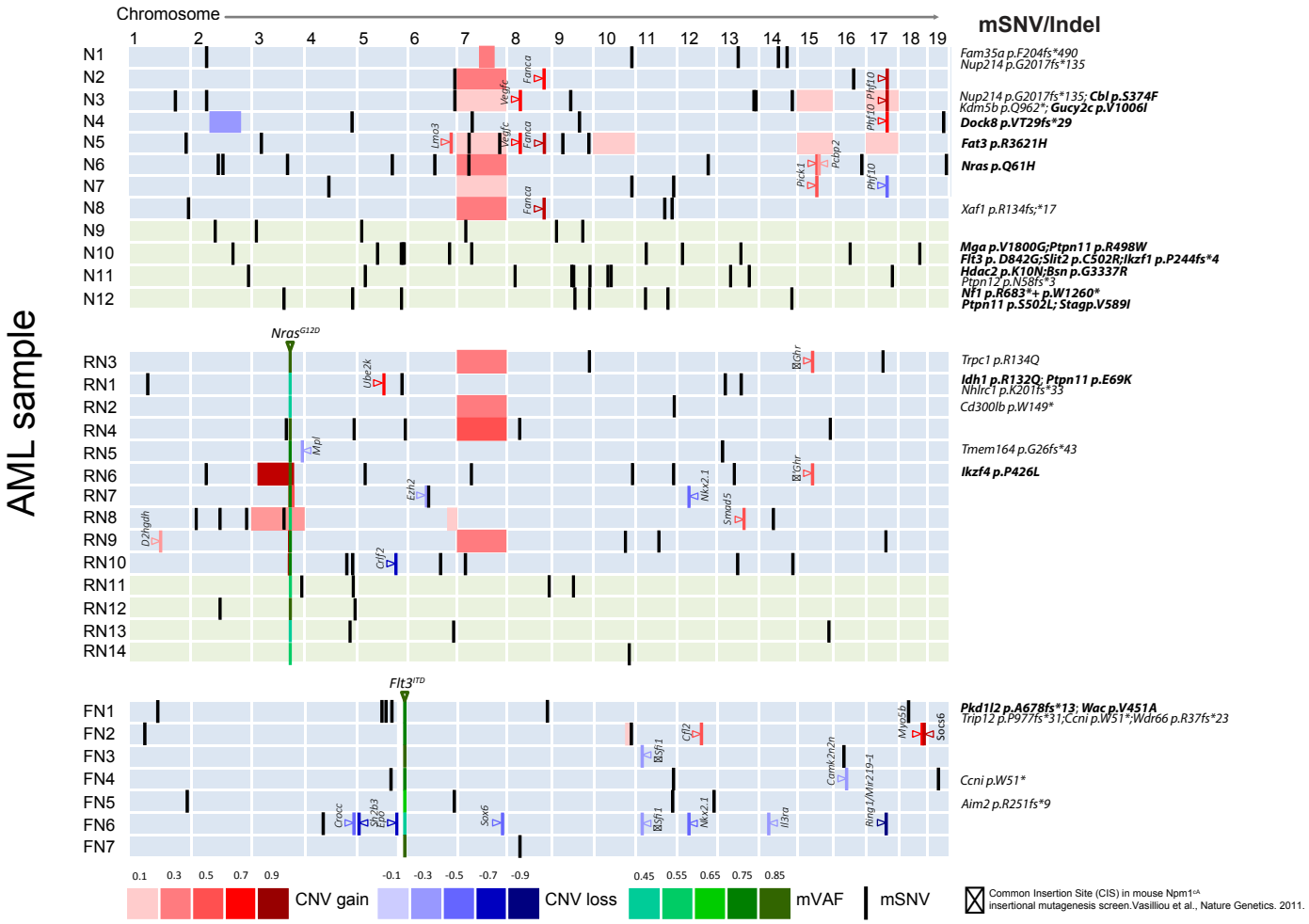
B.



C.



Supplemental Figure S6



Supplemental Figure Legends

Supplemental Figure S1. Pre-leukemic phenotypes of hematopoietic tissues. (**A**) Mac-1/Gr-1 staining of bone marrow and splenocytes from all genotypes shows an increase in myeloid commitment in *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ (predominantly Mac-1^{+}/Gr-1^{+}granulocytic) and *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$ (predominantly Mac-1^{+}/Gr-1^{-} monocytic) compared to singular mutants.

Supplemental Figure S2. Representative plots of pre-leukemic progenitor FACS of *Npm1*^{cA}, *Nras*$^{G12D/+}$, *Flt3*$^{ITD/+}$ and compound *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ or *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$ mice. (**A**) Lin^{-}, LK, LSK. Percentages of parent populations are shown for LK, LSK, GMP, MEP, CMP, LMPP, MPP and ST/LT$^{prog-HSC}$ populations, mean ± SEM (n=4-8). Representative plots of (**B**) CLP and (**C**) E-SLAM, HSC FACS plots and gates. Note an absence of the double positive FLT3/Il7-Rα in the CLP stain. LK (Lin^{-}/Kit^{+}), LSK (Lin^{-}/Kit^{+}/Sca-1^{+}), CMP (common myeloid progenitor), MEP (megakaryocyte-erythroid progenitor), GMP (granulocyte-monocyte progenitor), MPP (multi-potent progenitor), LMPP (lymphoid primed multi-potent progenitor), CLP (common lymphoid progenitor) and HSC (hematopoietic stem cell).

Supplemental Figure S3. Global gene expression analysis of lineage negative and LSK-progenitors. (**A**) Heat map of *Hox* gene expression in Lin^{-} bone marrow from singular and compound *Npm1*$^{cA/+}$, *Nras*$^{G12D/+}$, *Flt3*$^{ITD/+}$ mice (normalised average expression values are used to generate heat map values). (**B**) Venn diagrams of overlapping differentially expressed genes in *Npm1*$^{cA/+}$, *Nras*$^{G12D/+}$, *Flt3*$^{ITD/+}$, *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ and *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$ Lin^{-} bone marrow. Select over-expressed (red font) or under-expressed (blue font) are displayed. GSEA of differentially expressed genes in *Nras*$^{G12D/+}$ or *Flt3*$^{ITD/+}$ only mutants reveal enrichments for NRAS and JAK-STAT signalling pathways respectively. (**C**) Box-whisker plots of normalised average expression of *Nkx2-3*, *Hoxa7* and *Hoxa9*, as detected by microarrays in MPP and Lin^{-} populations. n=4-10 (Lin^{-}) or n=3-5 (MPP) for all genotypes (Mean ± Min-Max). (**D**) (**i**) Venn diagram and (**ii**) heat map of overlapping and distinct differentially expressed genes in sorted MPP populations from *Npm1*$^{cA/+}$, *Nras*$^{G12D/+}$, *Flt3*$^{ITD/+}$, *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ and *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$ reveals only a small sub-set of 12 deregulated genes shared in compound *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ and *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$ mice (log fold change, logFC, Adj. $p<0.05$ was used to identify significantly differentially expressed genes). (**E**) Results of gene-annotation enrichment analysis and functional annotation of differentially expressed genes in *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$ compared to wildtype MPPs (using DAVID). Statistically significant enriched Kegg pathways and enriched Gene Ontology term (GO-Term) clusters are shown (as depicted using Cytoscape 3.3.0). (**F**) A number of differentially expressed genes in *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$ multipotent progenitors (MPPs) are also deregulated in compound Tet2$^{-/-}$; *Flt3*$^{ITD/+}$ lineage negative/Sca-1^{+}/c-Kit^{+} (LSK) progenitors when compared to wildtype controls.

Supplemental Figure S4. *Npm1*^{cA} and oncogenic *Nras*^{G12D} co-operate to develop AML. (**A**) Comparative survival statistics (Median survival and Mantel-Cox Test p values) of data presented in Figure 3a, Kaplan Meier. FACS analysis of three of *Npm1*$^{cA/+}$; *Nras*$^{G12D/+}$ and *Npm1*$^{cA/+}$; *Flt3*$^{ITD/+}$ AMLs confirms myeloid infiltration in secondary lymphoid tissue (splenocytes); lymphoid (CD3ε/B220), myeloid (Mac-1/Gr-1, Mac-1/Kit) and B220^{+} myeloid (B220^{+}/Mac-1/Gr-1).

Supplemental Figure S5. Array comparative hybridisation (aCGH) of *Npm1$^{cA/+}$*, *Npm1$^{cA/+}$*; *Nras$^{G12D/+}$* and *Npm1^{cA}; Flt3^{ITD}* murine AML. (**A**) To determine the extent of recombination of the *Npm1$^{flox\text{-}cA}$* allele FN-AMLs, we quantified the fraction of *Npm1$^{flox\text{-}cA}$* and *Npm1^{cA}* allele reads using targeted amplicon specific MiSeq (see Materials and methods). As controls we used *Mx-1 Cre*;*Npm1$^{cA/+}$* and *Npm1$^{flox\text{-}cA}$* gDNA 4 months post pIpC injection. (**B**) Normalised Log2 ratio plots show gains (whole chromosome or smaller regions) of chr 3 in *Npm1$^{cA/+}$*; *Nras$^{G12D/+}$*. Green highlighted region denotes minimally mapped region of common chromosomal gain or loss (chr3: 102743581-103470550). (**C**) A commonly amplified region of chr7 (ch7: 91838150-131492236) is detected in 7/8 *Npm1$^{cA/+}$* and 4/9 *Npm1$^{cA/+}$; Nras$^{G12D/+}$* (not represented). 31 of the 312 genes in this region, syntenic to human chr11, are mutated in the TCGA AML data-set and include *Nup98*, *Wee1* and *Eed*. (**D**) A region of chr2 (chr2: 77889234-171131931) is deleted in *Npm1$^{cA/+}$* AML, N4. Of the 741 genes within this region 57 are in the TCGA AML dataset and include the commonly deleted genes in AML; *Asxl1*, *Wt1* and *Dnmt3b*. Black smoothed line indicates copy neutral regions. Red or blue smoothed line denotes gain or loss, respectively, of a chromosomal region defined on the x-axis for a particular sample. For (**C**) and (**D**) enrichment of syntenic human-mouse genes are shown (enrichment p-values as determined using DAVID[12]).

Supplemental Figure S6. Combined copy number and somatic variants for *Npm1$^{cA/+}$*, *Npm1$^{cA/+}$; Nras$^{G12D/+}$* and *Npm1$^{cA/+}$; Flt3$^{ITD/+}$* AMLs. (**A**) Combined copy number (aCGH) and somatic variants (while exome sequencing) in *Npm1$^{cA/+}$* (N-AML) compared to murine *Npm1$^{cA/+}$*; *Nras$^{G12D/+}$* (RN-AML) and *Npm1^{cA}; Flt3^{ITD}* (FN-AML) AML samples.

Supplemental Methods Figure S1. Exome sequencing and mutant somatic variant validation. (**A**) Exome sequencing and MiSeq validation "pipe-line" for detecting non-synonymous mouse AML variants. (**B**) Representation of MiSeq amplicon sequencing protocol. (1) Genomic PCR was performed with genome specific/MiSeq adapter primer sequences, Supplemental Table 2.1. (2) Pooled PCR products were then (3) amplified by PCR enrichment using a universal PE1.0 forward and a unique iPCRTag reverse primer, Supplemental Table 2.2. Samples were further purified and sequenced on an Illumina MiSeq. (**C**) The percentage of SNVs detected by exome sequencing and validated by MiSeq amplicon specific sequencing increases to 83% when using an "exome sequencing score" ≥3. This is further increased to 86% upon removal of C>A/G>T trans-version SNVs with mVAF<0.3. Note, the exome sequencing score (generated by RAMSES) is a confidence value derived from the following criteria for each SNV within a given sample; (**i**) the presence of mutations in both forward and reverse reads, (**ii**) unique or multiple genomic loci alignment (BLAT) and (**iii**) read quality and depth.

References

1. Yang YH, Dudoit S, Luu P, et al. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res*. 2002;30(4):e15.

2. Smyth, GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol*. 2004;3:Article3.

3. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc*. 1995;57(1)289–300.

4. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci*. 2005;102(43):15545-15550.

5. Mootha VK, Lindgren CM, Eriksson KF, et al. PGC-1alpha responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet*. 2003;34(3):267-273.

6. Network TCGAR. Genomic and epigenomc landscape of adult de novo acute lyeloid leukemia. *N Engl J Med*. 2013;368(22):2059-2074.

7. Gentleman RC, Carey VJ, Bates DM, et al. Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biol*. 2004;5(10):R80.

8. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-2079.

9. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010 Sep;20(9):1297-1303.

10. Martinez N, Almaraz C, Varela I, et al. Whole-exome sequencing in splenic marginal zone lymphoma reveals mutations in genes involved in marginal zone differentiation. *Leukemia*. 2014;28(6):1334-1340.

11. Ye K, Scgulz MH, Long Q, Apweiler R, Ning Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics*. 2009;25(21):2865-2871.

12. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nature Protoc*. 2009;4(1):44-57.