



(51) International Patent Classification:

C12Q 1/68 (2006.01) *G01N* 33/48 (2006.01)
C12Q 1/00 (2006.01) *G06F* 19/20 (2011.01)

(21) International Application Number:

PCT/CA2017/050020

(22) International Filing Date:

9 January 2017 (09.01.2017)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/276,496 8 January 2016 (08.01.2016) US

(71) Applicant: MCMMASTER UNIVERSITY [CA/CA]; 1280
Main Street West, Hamilton, Ontario L8S 4L8 (CA).

(72) Inventors: BHATIA, Mickie; 7 Gabel Crt, Ancaster,
Ontario L9G 4T2 (CA). DOBLE, Bradley Wayne; 605-
1276 Maple Crossing Blvd., Burlington, Ontario L7S 2J9
(CA). GUEZGUEZ, Borhane; 19 Homewood Ave,
Hamilton, Ontario L8P 2M1 (CA). ALMAKADI, Mo-
hammed Saleh; 72 Gregorio Ave, Ancaster, Ontario L9E
0C2 (CA). GARCIA-RODRIQUEZ, Juan Luis; 157
Pearl Street South, Unit 608, Hamilton, Ontario L8P 3X7
(CA).

(74) Agent: BERESKIN & PARR LLP/S.E.N.C.R.L., S.R.L.;
Scotia Plaza, 40 King Street West, 40th Floor, Toronto,
Ontario M5H 3Y2 (CA).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,
BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN,
KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA,
MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG,
NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS,
RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY,
TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN,
ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ,
TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU,
TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE,
DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU,
LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,
SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: DISEASE MODELS AND PROGNOSTIC TOOLS FOR PRE-LEUKEMIA AND LEUKEMIA

(57) Abstract: The present disclosure describes methods for creating a disease model of myelodysplastic syndrome and acute myeloid leukemia syndrome by creating gene knockouts of GSK-3 β or double knockouts of GSK-3 α and GSK-3 β , respectively. Also described are uses of these models for classifying human myeloid pre-neoplastic and neoplastic subtypes in reference to a GSK-3 β gene signature and determining risk stratification of a subject with myelodysplastic syndrome and prognosis based on their gene expression profile. The models and the GSK-3 β signature are useful for guiding treatment decisions or for monitoring treatment response. The disease models and the cells derived from them as described herein are also useful for screening and evaluating the effectiveness of test agents for treatment of MDS or leukemia.



**TITLE: DISEASE MODELS AND PROGNOSTIC TOOLS FOR
PRE-LEUKEMIA AND LEUKEMIA**

RELATED APPLICATIONS

[0001] This application claims the benefit of priority to United States Provisional application No. 62/276,496 filed January 8, 2016, the contents of which are incorporated herein by reference in their entirety.

FIELD

[0002] The disclosure relates to disease models of myelodysplastic syndromes (MDS) and acute myeloid leukemia (AML) and uses thereof. The disclosure also relates to methods and kits for detecting and/or screening for progression to AML, or an increased likelihood of progression to AML, in a subject with MDS.

BACKGROUND

[0003] Advances in divergent fields of hematology research have led to the collective appreciation that many, if not all, chronic and acute leukemias arise from proceeding, latent and/or preleukemic phases characterized by impaired hematopoiesis and defective differentiation (Meacham and Morrison, 2013; Shlush and Minden, 2015). Myelodysplastic syndromes (MDS) are considered to be preleukemic clonal diseases and are believed to originate from hematopoietic stem/progenitors cells (HSPCs) (Pandolfi et al., 2013; Woll et al., 2014), thereby affecting the development of one or more myeloid lineages causing cytopenias and dysplastic bone marrow (BM) (Garcia-Manero, 2010; Vardiman et al., 2009). These features are associated with morbid and fatal complications leading to overt acute myeloid leukemia (AML) in 30% of cases (Ades et al., 2014) with an increased risk of treatment resistance (Garcia-Manero, 2014; Jonas and Greenberg, 2015). Thus, the term MDS/AML is often used to describe a biological continuum in leukemia pathogenesis (Cazzola et al., 2013; Kreso and Dick, 2014). However, the

molecular basis of MDS to AML transition remains poorly understood (Bejar and Steensma, 2014; Larsson et al., 2013; Zhang et al., 2015), underscoring the need to better understand the specific pathways responsible for disease initiation and progression.

[0004] Although developmental pathways (e.g. Wnt/ β -catenin, Hedgehog, and Notch) were previously associated with self-renewal, deletion mutations or overexpression in these same signal mediators alone are insufficient to induce leukemic development (Takebe et al., 2011; Wang et al., 2010). Glycogen synthase kinase-3 (GSK-3) is an essential regulator of hematopoiesis (Huang et al., 2009; Trowbridge et al., 2006) and represents a signaling hub for Hedgehog (Briscoe and Therond, 2013), Wnt (Wu and Pan, 2010), and Notch (Andersson et al., 2011) that has recently been reported to play a pivotal role in various leukemias (Abrahamsson et al., 2009; Banerji et al., 2012; Wang et al., 2008). Accordingly, GSK-3 mediated networks may represent a more universal initiator of numerous pathways in cancers (Takahashi-Yanaga, 2013). Previous studies have shown functional redundancy between the two homologues of GSK-3 (α and β) (Doble et al., 2007; Gillespie et al., 2011; Itoh et al., 2012), which are encoded by separate genes on different chromosomes (Wada, 2009), but exhibit tissue-specific physiologically-important functions that may have varying impacts on human cancer etiology (Banerji et al., 2012; Takahashi-Yanaga, 2013). This suggests that the overriding tissue specific functions of GSK-3 and the interplay of α and β homologues are likely cell context specific.

SUMMARY

[0005] The roles of GSK-3 homologues have not been studied with respect to their individual impact on disease progression in MDS or AML. Here it is demonstrated that specific allelic targeting of GSK-3 α and GSK-3 β in hematopoietic tissue induces stepwise progression to AML state via a pre-neoplastic myeloid disorder that resembles human MDS. In particular, the present disclosure shows that the deletion of GSK-3 β in a transgenic mouse

model leads to a MDS-like disease *in vivo* and the combined deletion of GSK-3 α and GSK-3 β in a transgenic mouse model leads to an AML-like disease *in vivo*. These models are useful for screening for test agents to treat or prevent MDS, AML and/or progression from MDS to AML.

[0006] The present disclosure also defines a specific “GSK-3 β molecular signature” using a systematic approach based on the transcriptome comparison between WT versus GSK-3 β KO LSK cells of those genes with greater than 1.5-fold change in whole gene expression (with $p \leq 0.05$). This subset of genes was further filtered to yield a 63-gene GSK-3 β signature based on a fold change ≥ 5 (with $p \leq 0.05$). This signature was able to predict clinical outcomes of MDS patients.

[0007] Accordingly, in an aspect, there is provided a method of detecting and/or screening for progression to AML or an increased likelihood of progression to AML, in a human subject having MDS, comprising:

determining a sample gene expression profile from a sample comprising DNA from said subject, said sample profile comprising:

(a) the gene expression level of at least 3, optionally at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, or all of the human orthologues of the genes listed in Table 5; and

(b) determining the level of similarity of said sample profile to one or more control profiles.

[0008] In one embodiment, (i) a high level of similarity of the sample profile to a high risk AML progression control profile; (ii) a low level of similarity to low risk AML progression control profile; and/or (iii) a higher level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile indicates progression to AML or an increased likelihood of progression to AML.

[0009] In another embodiment, a higher level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile is indicated by hierarchical clustering.

[0010] In another embodiment, a higher level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile is indicated by a higher correlation value computed between the sample profile and the high risk AML progression control profile than an equivalent correlation value computed between the sample profile and the low risk AML progression control profile, optionally wherein the correlation value is a correlation coefficient.

[0011] In another embodiment, (i) a low level of similarity of the sample profile to a high risk AML progression control profile; (ii) a high level of similarity to low risk AML progression control profile; and/or (iii) a lower level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile indicates no progression to AML or an decreased likelihood of progression to AML.

[0012] In another embodiment, a lower level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile is indicated by hierarchical clustering.

[0013] In another embodiment, a lower level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile is indicated by a lower correlation value computed between the sample profile and the high risk AML progression control profile than an equivalent correlation value computed between the sample profile and the low risk AML progression control profile, optionally wherein the correlation value is a correlation coefficient.

[0014] In another embodiment, the method further comprises identifying the WHO MDS subtype of the subject, wherein

(a) (i) a high level of similarity of the sample profile to a high risk AML progression control profile; (ii) a low level of similarity to low risk AML progression control profile and/or (iii) a higher level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile; and

(b) a high risk WHO MDS subtype,

indicates a high risk of progression to AML, progression to AML, or an increased likelihood of progression to AML.

[0015] In another embodiment, the method further comprises identifying the WHO MDS subtype of the subject, wherein

(a) (i) a low level of similarity of the sample profile to a high risk AML progression control profile; (ii) a high level of similarity to low risk AML progression control profile; and/or (iii) a lower level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile indicates no progression to AML or an decreased likelihood of progression to AML; and

(b) a low risk WHO MDS subtype,

indicates a low risk of progression to AML, no progression to AML, or a decreased likelihood of progression to AML.

[0016] In another embodiment, the method further comprises identifying the WHO MDS subtype of the subject, wherein

(a) (i) a high level of similarity of the sample profile to a high risk AML progression control profile; (ii) a low level of similarity to low risk AML progression control profile and/or (iii) a higher level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile; and

(b) a low risk WHO MDS subtype,

or

(a) (i) a low level of similarity of the sample profile to a high risk AML progression control profile; (ii) a high level of similarity to low risk AML

progression control profile; and/or (iii) a lower level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile indicates no progression to AML or an decreased likelihood of progression to AML; and

(b) a high risk WHO MDS subtype,

indicates an intermediate risk of progression to AML.

[0017] In another embodiment, the subject identified as having progression to AML, an increased likelihood of progression to AML or a high risk of progression to AML is treated with chemotherapy, radiation, stem cell transplantation and/or bone marrow transplantation.

[0018] The disclosure also provides a method of selecting a treatment for a subject with progression to AML or an increased likelihood of progression to AML, comprising:

(a) identifying the subject with progression to AML, an increased likelihood of progression to AML or a high risk of progression to AML, according to the methods described herein; and

b) selecting a treatment for the subject, wherein the treatment comprises chemotherapy, radiation, stem cell transplantation and/or bone marrow transplantation.

[0019] The disclosure also provides a method of detecting and/or screening for myelodysplastic syndrome (MDS) or an increased likelihood of MDS, in a human subject, comprising:

determining a sample gene expression profile from a sample comprising DNA from said subject, said sample profile comprising, consisting of, or consisting essentially of:

(a) the gene expression level of at least 3, optionally at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, or all of the human orthologues of the genes listed in Table 5; and

(b) determining the level of similarity of said sample profile to one or more control profiles, wherein (i) a high level of similarity of the sample profile to an MDS specific control profile; (ii) a low level of similarity to a non-MDS control profile; and/or (iii) a higher level of similarity to an MDS specific control profile than to a non-MDS control profile indicates the presence of, or an increased likelihood of, MDS.

[0020] In one embodiment, a higher level of similarity to an MDS specific control profile than to a non-MDS specific control profile is indicated by hierarchical clustering.

[0021] In another embodiment, a higher level of similarity to an MDS specific control profile than to a non-MDS specific control profile is indicated by a higher correlation value computed between the sample profile and the MDS specific control profile than an equivalent correlation value computed between the sample profile and non-MDS specific control profile, optionally wherein the correlation value is a correlation coefficient.

[0022] In another embodiment, the subject is a pediatric subject.

[0023] In another embodiment, determining the sample gene expression profile comprises the steps:

- a) providing the sample comprising mRNA from the subject;
- b) isolating mRNA from the sample;
- c) amplifying the mRNA; and
- d) determining the gene expression level of the at least 3 genes by means of RT-PCR.

[0024] In another embodiment, the sample is derived from a clinical specimen.

[0025] In another embodiment, the sample is derived from blood or bone marrow.

[0026] The disclosure also provides a kit for detecting and/or screening for progression to AML or an increased likelihood of progression to AML, in a human subject having MDS, in a sample, comprising:

a) at least one detection agent for determining the gene expression level of at least 3, optionally at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, or all of the human orthologues of the genes listed in Table 5; and

b) instructions for use.

[0027] The disclosure also provides a

kit for detecting and/or screening for MDS or an increased likelihood of MDS in a sample from a human subject, comprising:

c) at least one detection agent for determining the gene expression level of at least 3, optionally at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, or all of the human orthologues of the genes listed in Table 5; and

d) instructions for use.

[0028] In one embodiment, the detection agents are probes and/or primers.

[0029] In another embodiment, the kit further comprises PCR reagents.

[0030] In another embodiment, the kit further comprises a computer-readable medium that causes a computer to compare gene expression from a sample at the selected genes to one or more control profiles and compute a correlation value between the sample and control profile.

[0031] The disclosure also provides a method of screening a test agent for treatment or prevention of myelodysplastic syndrome (MDS), comprising a) administering a test agent to i) a transgenic mouse model of MDS whose genome comprises a disruption of GSK-3 β , ii) a tissue from said transgenic mouse model, or iii) a cell derived from said transgenic mouse model; and b) determining the effect of said test agent on the mouse model, tissue or cell,

wherein a reduction of signs and/or symptoms of MDS in the mouse, tissue or cell identifies a test agent for the treatment or prevention of MDS.

[0032] In one embodiment, the disruption is a homozygous GSK-3 β knockout.

[0033] In another embodiment, the GSK-3 β knock out is a conditional knockout of GSK-3 β .

[0034] The disclosure also provides a method of screening a test agent for treatment or prevention of acute myeloid leukemia (AML), comprising a) administering a test agent to i) a transgenic mouse model of AML whose genome comprises a disruption of GSK-3 α and a disruption of GSK-3 β , ii) a tissue from said transgenic mouse model, or iii) a cell derived from said transgenic mouse model; and b) determining the effect of said agent on the mouse model, tissue or cell, wherein a reduction of signs and/or symptoms of AML in the mouse, tissue or cell identifies a test agent for the treatment or prevention of AML.

[0035] In one embodiment, the disruption of GSK-3 α is a homozygous GSK-3 α knockout and/or the disruption of GSK-3 β is a homozygous GSK-3 β knockout.

[0036] In another embodiment, the GSK-3 α knockout is a conditional knockout of GSK-3 α and/or the GSK-3 β knock out is a conditional knockout of GSK-3 β .

[0037] The disclosure also provides a method screening a test agent for treatment or prevention of myelodysplastic syndrome (MDS), comprising a) administering a test agent to a cell comprising a disruption of GSK-3 β and b) determining the effect of said agent on the cell.

[0038] In one embodiment, the disruption is a homozygous GSK-3 β knockout.

[0039] In another embodiment, the GSK-3 β knock out is a conditional knockout of GSK-3 β .

[0040] The disclosure also provides a method of screening a test agent for treatment or prevention of acute myeloid leukemia (AML) comprising a) administering a test agent to a cell comprising a disruption of GSK-3 α and a disruption of GSK-3 β and b) determining the effect of said agent on the cell.

[0041] In one embodiment, the disruption of GSK-3 α is a homozygous GSK-3 α knockout and/or the disruption of GSK-3 β is a homozygous GSK-3 β knockout.

[0042] In another embodiment, the GSK-3 α knockout is a conditional knockout of GSK-3 α and/or the GSK-3 β knock out is a conditional knockout of GSK-3 β .

[0043] In another embodiment, the cell is a human cell.

[0044] The present disclosure also provides a method for creating a transgenic model for myelodysplastic syndrome by creating gene knockouts of GSK-3 α .

[0045] Also provided is a method for creating a transgenic model for acute myeloid leukemia by creating gene knockouts of GSK-3 α and GSK-3 β .

[0046] Further provided is a method for classifying human myeloid pre-neoplasm subtypes comprising: determining in a subject the gene expression level for each of a set of genes selected from a GSK-3 β gene signature listed in Table 5 to obtain a subject expression profile.

[0047] Further provided is a use of the GSK-3 β signature alone or in combination with other methods for risk stratification of a subject with myelodysplastic syndrome.

[0048] Further provided is a use of the GSK-3 β signature alone or in combination with other methods for determining a prognosis of a subject based on their expression profile.

[0049] Further provided is a use of the GSK-3 β signature alone or in combination with other methods for guiding treatment based on gene expression profile.

[0050] Further provided is a use of the GSK-3 β signature alone or in combination with other methods for monitoring treatment response.

[0051] Further provided is a use of models described above for screening test agents that are in development as treatments for MDS or leukemia.

[0052] Other features and advantages of the present disclosure will become apparent from the following detailed description. It should be understood, however, that the detailed description and the specific examples while indicating embodiments of the disclosure are given by way of illustration only, since various changes and modifications within the spirit and scope of the disclosure will become apparent to those skilled in the art from this detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

[0053] Embodiments are described below in relation to the drawings in which:

[0054] Figure 1 shows that the combined deletion of GSK-3 α and β alleles leads to an AML-like disease *in vivo*. (a) Experimental design to generate conditional GSK-3 α/β mutant mice ($Gsk3\alpha^{-/-} \beta^{+/flx}$ and $Gsk3\alpha^{-/-} \beta^{flx/flx}$) and transplant strategy. (b) Transplanted recipients with $Gsk3\alpha^{-/-} \beta^{flx/flx}$ cells (Lin⁻ BM) display marked skin lesions with rapid mortality (n=15 mice for each condition) and show (c) BM hypercellularity and splenomegaly in mice transplanted with $Gsk3\alpha^{-/-} \beta^{flx/flx}$ cells (mean $\pm$ S.D., n=6). (d) $Gsk3\alpha^{-/-} \beta^{flx/flx}$ mice exhibit elevated leukocytes counts in PB and lymph nodes (LN) (mean $\pm$ S.D., n=6). (e-f) Presence of immature myeloid cells and leukemic blasts were found at high-levels in (g) BM, spleen, (h) PB and lymph node (LN) in $Gsk3\alpha^{-/-} \beta^{flx/flx}$ mice (mean $\pm$ S.D., n=6), *p<0.05 and **p<0.01. (i) Effect of GSK-3 α/β

combined deletion on HSCs and myeloid progenitors increased LSK and GMP subsets (mean±S.D., n=10) and higher cell proliferation rate (mean±S.D., n=5). (j) High mortality and similar AML features are conserved in secondary transplants of primary engrafting $Gsk3\alpha^{-/-} \beta^{flx/flx}$ in comparison to both $Gsk3\alpha^{-/-}$ and $Gsk3\alpha^{-/-} \beta^{+/flx}$ (n=12 for each condition), **p<0.01.

[0055] Figure 2 shows the GSK-3 β deficiency alone leads to myelodysplasia and impaired hematopoiesis. (a) BM transplant strategy. (b) Incremental mortality of transplanted and tamoxifen- treated recipient mice with $Gsk3\beta^{flx/flx} lin^{-}$ BM cells compared to the indicated control animals (n=15 mice for each condition). (c) Despite that the BM of $Gsk3\beta^{flx/flx}$ mice displays normal to increased cellularity (n=10), (d) existence of myelodysplasia was evident by presence of lobated or hypersegmented neutrophils (arrowheads) (n=10) and (e) increased frequency of blasts (n=10). (f) Extramedullary hematopoiesis was also observed in 11/15 $Gsk3\beta^{flx/flx}$ mice examined for accumulation of monocytes and granulocytes in the spleen and accretion of hypersegmented neutrophils in PB and micro-megakaryocytes with hyperchromatic nuclei in the spleen (n=11). (g) BM serial transplantation strategy. (h) Sustained MDS disease occurrence in secondary recipients is dependent from transplanted cell dose; (n=10 mice for each condition). (i) High mortality and similar MDS features are conserved in secondary transplants of primary engrafting $Gsk3\beta^{flx/flx}$ cells eg. myeloid blasts (n=10 for each condition), *p<0.05 and **p<0.01.

[0056] Figure 3 shows GSK-3 β -MDS BM arises from HSCs and displays MDS-IC activity. (a) Experimental design for competitive reconstitution assays using sorted HSCs (LSK) and progenitors (non-LSK) from both $Gsk3\beta^{+/flx}$ and $Gsk3\beta^{flx/flx}$ mice (n=15 for each condition). (b) Despite mortality of non-LSK transplanted mice due to lack of hematopoietic regenerative capacity, only mice transplanted with $Gsk3\beta^{flx/flx}$ LSK display high mortality and ineffective hematopoiesis (n=15 for each condition), as indicated by (c) dual features of hypercellular BM and splenomegaly (10/15 mice) (d)

with diffuse infiltration by myeloid and granulocytic cells and presence of dysplastic neutrophils (arrowheads) (n=10). (e-f) Both transplanted non-LSK $^{+/flx}$ and non-LSK $^{flx/flx}$ mice exhibit normal blood counts and myeloid lineage (n=4). (g) $Gsk3\beta^{flx/flx}$ HSCs display a higher hematopoietic stem cell function as determined by competitive reconstitution capacity (n=5), (h) along with the acquisition of higher proliferative cell rate 6 weeks post-transplant (n=5). (i) Progressive granulocytic accumulation in the BM of diseased $Gsk3\beta^{flx/flx}$ mice (n=5). (j-k) Similar myelodysplastic features and worse survival are also conserved in secondary transplants of primary engrafting $Gsk3\beta^{flx/flx}$ HSCs (n=8 for each), *p<0.05, **p<0.01 and ***p<0.001.

[0057] Figure 4 shows HSCs GSK-3 β -null and α/β -null acquires dysregulated stem cell and myeloid differentiation gene programs driven by progressive activation of the Wnt/Akt/mTOR signaling axis. (a) Principle component analysis (PCA) on all entities from global gene expression profiling of LSK HSCs from $Gsk3\alpha^{+/+} \beta^{+/flx}$ (WT), $Gsk3\alpha^{-/-} \beta^{+/+}$ (A KO), $Gsk3\alpha^{+/+} \beta^{flx/flx}$ (B KO) and $Gsk3\alpha^{-/-} \beta^{flx/flx}$ (DKO). (b) Gene set enrichment analysis (GSEA) showing the highest ranked gene signatures differentially expressed between GSK-3 genotypes. (c) Enrichment of Wnt canonical signaling signature (KEGG) in both GSK-3 B KO and DKO LSK cells correlates with (d) up-regulation of Wnt gene targets (e) significantly activated form and (f) progressive nuclear stabilization of β -catenin levels. (g) Enrichment in survival mTOR signaling (KEGG and BioCarta signatures) is also observed for both GSK-3 B KO and DKO LSK cells as well as (h) up-regulation of mTOR gene targets. (i) Except for DKO cells, mTOR complex inhibition reduces significantly total GSK-3 B KO derived-CFU-C at similar levels as WT control (mean $\pm$ S.D., n=5). (j) Inhibition of PI3K/AKT/mTOR pathway at different levels shows marked and specific reduction of CFU-Cs derived from both GSK-3 B KO and DKO LSK cells (mean $\pm$ S.D., n=5). (k) Oncogenic SANSOM_APC_MYC and WIERENGA_STAT5 targets gene signatures are also enriched in both GSK-3 β KO and DKO LSK cells. (l) Progressive

activation of AKT/PI3K pathway (REACTOME signature) correlates with (m) the increase of phosphorylated AKT levels (active form) notably in DKO LSK cells. (n) Both glycolysis and oxidative phosphorylation gene signatures (KEGG) are gradually dysregulated in both Gsk3 β KO and $\alpha\beta$ KO, (o) correlating with distinct resting metabolic activity and (p-q) differential responses to mitochondrial stress with respect to WT (mean $\pm$ S.D., n=5), **p<0.01.

[0058] Figure 5 shows the GSK-3 β -null BM expresses a unique gene expression profile and predicts pediatric MDS disease outcome. (a) Hierarchical cluster analysis of global gene profile shows similarity between mouse GSK-3 β deleted BM cells and human MDS samples (MDS-LR: n=60, MDS-HR: n=44, 5q-: n=11), CMML (n=20) and AML (n=30). Profiles of human samples were generated using human orthologues to the mouse genes listed in Table 5. (b) Unique GSK-3 β -deletion signature (63 genes, triangles) is determined between GSK-3 WT and β KO ($p \leq 0.05$, fold-change ≥ 5). (c) GSK-3 β gene signature discriminates human myeloid neoplasm subtypes. (d) Both GSK-3 β -deletion differential transcriptome ($p \leq 0.05$, fold-change ≥ 1.5) and 63-gene signatures discriminate different disease diagnosis states in both adult (normal: n=6, stable-MDS: n=23 and transformed-MDS: n=12) and (e) pediatric myelodysplasia (RAEB: n=8, RAEB-t: n=8 and RC: n=16). (f-h) Correlation of GSK-3 β -deletion gene signature with disease risk status and overall disease progression to AML in a cohort of pediatric MDS samples.

[0059] Figure 6 depicts the model of MDS onset and AML progression by allelic dosage of GSK-3 α/β .

[0060] Figure 7 shows the effect on hematopoiesis is dependent on the allelic dosage of GSK-3 α/β . (a) Experimental design outlining the strategy to generate conditional GSK-3 β knockout mice. (b-c) GSK3 β deletion was demonstrated by genomic PCR (b) mRNA gene expression and (c) Western Blot analysis of total Lin⁻ bone marrow cells from wild-type (+/+), heterozygous (+/flx) and homozygous (flx/flx) GSK-3 β mice 2 weeks after Tamoxifen

injections (n=5). (d) Increased hematopoietic CFU progenitor output of *Gsk3 β ^{flx/flx}* Lineage-depleted bone marrow cells (Lin⁻ BM) upon tamoxifen treatment (n=5, performed in triplicate). (e-f) Representative chimerism in bone marrow and peripheral blood between donor cells (CD45.2) and host cells (CD45.1) in primary recipients (n=5 for each condition). (g) Presence of severe anemia and (h) leukocytosis are detected in *Gsk3 α ^{-/-} β ^{flx/flx}* transplanted mice, (i) along with neutrophilic dermatosis in skin lesions and granulocytic infiltration in the liver (n=5 for each condition). (j) Transplanted recipients with *Gsk3 α ^{-/-}* LSK cells display cellularities similar to those observed in controls (*Gsk3 α ^{+/+}*) in both bone marrow and spleen as well similar leukocyte counts in PB and LN (n=8 mice for each condition). (k) A normal proportion of immature myeloid cells is present in BM, Spleen, PB and LN (mean±S.D., n=6), *p<0.05 and **p<0.01.

[0061] Figure 8 shows GSK-3 β -null MDS properties are accelerated in transplanted recipients. (a) Anemia, (b) granulocytosis and cytopenias were present in mice engrafted with *Gsk3 β ^{flx/flx}* cells (n=10). (c) Higher frequency of immature myeloid cells in *Gsk3 β ^{flx/flx}* recipients (n=10). (d-g) Similar myelodysplastic features are also present in peripheral blood and spleens of recipients transplanted with *Gsk3 β ^{flx/flx}* cells (n=10 for each condition), **p<0.01.

[0062] Figure 9 shows the GSK-3 β deletion leads to impaired hematopoiesis and MDS properties that are sustained in secondary recipients. (a-b) Representative bone marrow and spleen sections denoting progressive granulocytosis (MPO⁺) without monocytosis (Mac-2⁺) in GSK-3 β -null (B KO) and α/β -null (DKO) (n=4 for each genotype), scale bar 200 μ m. (c) MDS features are also conserved in secondary transplants of primary engrafting *Gsk3 β ^{flx/flx}* cells, displaying superior regenerative capacity with (d) dramatically increased cellularity of BM recipients and (e) presence of immature blasts (n=10 for each condition), **p<0.01.

[0063] Figure 10 shows the hematological disorder related to GSK-3 β deficiency is initiated solely by HSCs. (a) Effect of GSK-3 β deletion on HSCs subset demonstrating a specific increase of the LSK and CD150⁺48⁺ subsets (mean $\pm$ S.D., n=10). (b) Absolute numbers of each progenitor subpopulation in the bone marrow showing enlargement of the GMP population (mean $\pm$ S.D., n=10). (c) Presence of increased proliferation of hematopoietic stem/progenitor (HSPC) compartment (mean $\pm$ S.D., n=5) was partially compensated by (d) increased apoptosis of hematopoietic progenitor cells in *Gsk3 β ^{flx/flx}* mice (mean $\pm$ S.D., n=5). (e) Bone marrow transplant strategy in immunodeficient NOD/SCID mice. (f) Similar to congenic transplants, rapid mortality due to hematopoietic failure is observed in mice transplanted with *Gsk3 β ^{flx/flx}* BM cells (n=15 mice for LSK cells, n=6 mice for Non-LSK cells), **p<0.01. (g) Similar features of cellular cytopenias are also conserved in peripheral blood of NOD/SCID recipients transplanted with *Gsk3 β ^{flx/flx}* BM (n=10 for each condition), *p<0.05, **p<0.01. (h-i) Flow cytometric analyses revealed that most of the engrafted *Gsk3 β ^{flx/flx}* LSK BM cells generate high levels of Gr1⁺/CD11b⁺ myeloid granulocytic cells, accumulating both in bone marrow and the spleen along with the presence of anemia and cytopenia (n=10). (j-k) Both transplanted non-LSK ^{+/flx} and non-LSK ^{flx/flx} NOD/SCID recipient mice exhibit normal blood counts and myeloid lineage distribution (n=5 for each condition). (l) Representative flow gating plots demonstrating similar cell composition of myeloid progenitors subsets within the non-LSK population isolated from both *Gsk3 β ^{+/flx}* and *Gsk3 β ^{flx/flx}* BM cells (n=3). (m) Similar myelodysplastic features are also conserved in secondary transplants of primary engrafting *GSK-3 β ^{flx/flx}* HSCs (n=8 for each condition), *p<0.05, **p<0.01.

[0064] Figure 11 shows the GSK-3 α/β deletion displays specific dysregulated gene program that is not allelic dose dependent. (a-c) Most representative gene signatures present in LSK cells from GSK-3 α KO, β KO and DKO LSK. (d-e) Most representative gene signatures discriminating GSK-

3 A KO from B KO LSK cells. (f) Protein levels of GSK-3 α/β in Lin⁻ BM cells of different knockout genotypes. (g) Rapamycin treatment was inefficient to override the abnormal distribution of erythro-myeloid colony types for both GSK-3 B KO and DKO (mean $\pm$ S.D., n=5).

[0065] Figure 12 shows the GSK-3 β gene signatures predict MDS outcomes in humans. (a-b) GSEA analysis showing dysregulated epigenetic signatures previously associated with MDS/AML malignancies. (c) Mapping view of relevant functional association network of GSK-3 β 63-gene signature generated using STRING protein interaction database; stronger associations are represented by multiple lines and gene products involved in immune response are highlighted. (d) Kaplan-Meier curves demonstrate that WHO criteria as well as (e) both GSK-3 β differential transcriptome and 63- gene signatures predict faithfully patient's prognosis and (f) AML transformation risk of pediatric MDS patients, as denoted by p-value.

[0066] Figure 13 shows *in vitro* deletion of the Gsk3 β KO and Gsk3 α + β DKO in highly enriched LSK BM cell lines allow the study of the molecular features for disease onset. a) An experimental design for *in vitro* analysis, b) a representative flow cytometry histograms showing the percentage of c-Kit⁺ and Sca-1⁺ from the Lin⁻ BM analyzed at Day 0 and Day 7 after incubation, c) quantitative PCR validation of Gsk3 allele deletion after *in vitro* Tamoxifen treatment and d) representative western blots showing the decreased protein levels of GSK3 α and/or β in the respective cell line after *in vitro* Tamoxifen treatment.

DETAILED DESCRIPTION

Method of detecting and/or screening for progression to AML

[0067] Myelodysplastic syndrome (MDS) is a pre-leukemic clonal disease which may progress to lethal acute myeloid leukemia (AML). As set out in Table 5, the disclosure identifies 63 distinct genes, each of which show a statistically significant (p-value $\leq$ 0.05) difference in gene expression levels

between wild type cells and GSK-3 β KO LSK cells. As described in the Examples, the gene expression levels of the disclosed genes and human orthologues thereof, or a subset thereof, may be used in diagnostic and prognostic testing for MDS and AML and in particular, in detecting and/or screening for the progression of MDS to AML.

[0068] Accordingly, the disclosure provides a method of detecting and/or screening for progression to AML or an increased likelihood of progression to AML, in a human subject having MDS, comprising:

determining a sample gene expression profile from a sample comprising DNA from said subject, said sample profile comprising, consisting of, or consisting essentially of:

(a) the gene expression level of at least 3, optionally at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, or all of the human orthologues of the genes listed in Table 5; and

(b) determining the level of similarity of said sample profile to one or more control profiles.

[0069] In one embodiment, (i) a high level of similarity of the sample profile to a high risk AML progression control profile; (ii) a low level of similarity to low risk AML progression control profile; and/or (iii) a higher level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile indicates progression to AML or an increased likelihood of progression to AML.

[0070] Another aspect of the disclosure provides a method of detecting and/or screening for myelodysplastic syndrome (MDS) or an increased likelihood of MDS, in a human subject, comprising:

determining a sample gene expression profile from a sample comprising DNA from said subject, said sample profile comprising, consisting of, or consisting essentially of:

(a) the gene expression level of at least 3, optionally at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, or all of the human orthologues of the genes listed in Table 5; and

(b) determining the level of similarity of said sample profile to one or more control profiles, wherein (i) a high level of similarity of the sample profile to an MDS specific control profile; (ii) a low level of similarity to a non-MDS control profile; and/or (iii) a higher level of similarity to an MDS specific control profile than to a non-MDS control profile indicates the presence of, or an increased likelihood of, MDS.

[0071] Another aspect of the disclosure provides a method of detecting and/or screening for acute myeloid leukemia (AML) or an increased likelihood of AML, in a human subject, comprising:

determining a sample gene expression profile from a sample comprising DNA from said subject, said sample profile comprising, consisting of, or consisting essentially of:

(a) the gene expression level of at least 3, optionally at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, or all of the human orthologues of the genes listed in Table 5; and

(b) determining the level of similarity of said sample profile to one or more control profiles, wherein (i) a high level of similarity of the sample profile to an AML specific control profile; (ii) a low level of similarity to a non-AML control profile; and/or (iii) a higher level of similarity to an AML specific control profile than to a non-AML control profile indicates the presence of, or an increased likelihood of AML.

[0072] As used herein, the term “gene” refers to a genomic DNA sequence that comprises a coding sequence associated with the production of a polypeptide or polynucleotide product (e.g., rRNA, tRNA).

[0073] The term "expression level" of a gene as used herein refers to the measurable quantity of a gene product produced by the gene in a sample of a patient, such as a blood sample or a bone marrow sample, wherein the gene product can be a transcriptional product or a translated transcriptional product. Accordingly, the expression level can pertain to a nucleic acid gene product such as RNA or cDNA or a polypeptide gene product. The expression level is derived from a patient sample or cell and/or a control sample, and can for example be detected *de novo* or correspond to a previous determination.

[0074] As used herein, the term "orthologue" refers to genes in different species that evolved from a common ancestral gene by speciation. Normally, orthologues retain the same function in the course of evolution. Accordingly, the term "human orthologues of the genes in Table 5" refers to human orthologues of the mouse genes set out in Table 5. Determining human orthologues of mouse genes can be determined, for example, through the use of online databases such as GenBank. Table 5 provides accession numbers of human orthologues of the mouse genes listed therein.

[0075] Methods of determining gene expression levels can be determined by assaying nucleic acid expression products, for example mRNA or cDNA and/or by assaying polypeptide products. The level of gene expression can also be determined by assaying the presence and/or amount of a specific sequence in the gene, also known as a target sequence, optionally an mRNA target sequence. The level of gene expression can be determined or measured using a detection agent, wherein the agent detects a gene expression product of a gene or target sequence described herein. In an embodiment, the detection agent is an antibody, receptor protein or a nucleic acid such as a probe or primer set, capable of amplifying the gene or gene expression product.

[0076] In an embodiment, the methods comprise determining nucleic acid levels.

[0077] Wherein the gene expression level being determined is a nucleic acid, the gene expression levels can be determined using a number of methods for example a microarray chip or PCR, optionally multiplex PCR, northern blotting, or other methods and techniques designed to produce quantitative or relative data for the levels of mRNA species corresponding to specified nucleotide sequences present in a sample. Accordingly, in an embodiment, the gene expression level is determined using a microarray chip and/or PCR, optionally multiplex PCR.

[0078] Expression levels of the genes described herein can also be determined using primers (also known as target specific primers). In one embodiment, primers are attached to universal primer sequences that produce an amplification product (see for Example, US Patent No. 6,618,679). The multiplex amplification reaction used in the methods includes, but is not limited to PCR and the use of reverse transcription to generate cDNA from mRNA targets followed by nucleic acid sequence based amplification. These methods include single reaction RT-PCR, 2 step reverse transcription and PCR amplification, ligase chain reaction, cyclic probe amplification, an invader assay, bridge amplification or rolling circle amplification or a combination.

[0079] In another embodiment, the primers are designed to be intron spanning to minimize or eliminate genomic DNA signal interference. This can be accomplished using several primer design software products such as Primer6 (The Whitehead Institute, MA), OLIGO (Molecular Biology Insights, Inc, CO).

[0080] In another embodiment, amplified PCR products are separated by fragment sizes corresponding to the different target genes and the peak areas represent the mRNA level of expression. Amplified PCR products are separated by capillary electrophoresis or other separation techniques such as agarose and/or acrylamide gel, chromatography such as HPLC, or FPLC as well as microfluidic techniques. Detection is performed by, but not limited, to measuring the emission of laser induced fluorescence labels. Other methods

of detection include light absorptions and electrochemical signals. Amplified PCR products can also be detected by microarray, bead based analysis, and multiplex qPCR.

[0081] As used herein, a “gene expression profile” refers to the gene expression levels of one or more target genes. An expression profile can for example be detected by measuring RNA expression using methods described above such as microarray analysis, RT-PCR, multiplex PCR, directly quantitating RNA levels using for example RNA sequencing and/or by measuring polypeptide expression using methods such as flow cytometry and Western blotting.

[0082] As used herein, the term “sample gene expression profile” or “sample profile” refers to the gene expression levels of one or more target genes in a subject's genomic DNA. The gene expression profile of a sample tested according to the methods disclosed herein is referred to as a sample profile.

[0083] In some embodiments, the sample gene expression profile is compared to one or more control profiles. The control profile may be a reference value and/or may be derived from one or more samples, optionally from historical gene expression data for a patient or pool of patients who are known to have, or not have, MDS and/or AML and/or progression from MDS to AML. In such cases, the historical gene expression data can be a value that is continually updated as further samples are collected and individuals are identified as having MDS, AML and/or progression from MDS to AML as described herein. It will be understood that the control profile represents an average of the gene expression levels for selected genes as described herein. Average gene expression levels may, for example, be the mean levels or median levels.

[0084] For example, a “MDS specific control profile” or “MDS control profile” may be generated by measuring the gene expression levels of

specified genes in genomic DNA from an individual subject, or population of subjects, who are known to have MDS. Similarly, a “non-MDS control profile” may be generated by measuring the gene expression levels of specified genes in genomic DNA from an individual subject or population of subjects who are known to not have MDS.

[0085] In another example, an “AML specific control profile” or “AML control profile” may be generated by measuring the gene expression levels of specified genes in genomic DNA from an individual subject, or population of subjects, who are known to have AML. Similarly, a “non-AML control profile” may be generated by measuring the gene expression levels of specified genes in genomic DNA from an individual subject or population of subjects who are known to not have AML.

[0086] In another example, a “high risk AML progression control profile” may be generated by measuring the gene expression levels of specified genes in genomic DNA from an individual subject, or population of subjects, who are known to have a high risk of progressing from MDS to AML.

[0087] In another example, a “low risk AML progression control profile” may be generated by measuring the gene expression levels of specified genes in genomic DNA from an individual subject, or population of subjects, who are known to have a low risk of progressing from MDS to AML.

[0088] In certain embodiments, the tissue source from which the sample profile and control profile are derived is matched, so that they are both derived from the same or similar tissue.

[0089] As used herein “sample” includes but is not limited to a body fluid sample, tissue sample or tumor sample, which can be assayed for gene expression levels. The sample includes for example a blood sample, a biopsy, a frozen tissue sample, a fresh tissue specimen, a cell sample, and/or a paraffin embedded section, material from which, for example, RNA or DNA can be extracted in sufficient quantities and with adequate quality to permit

measurement of relative RNA levels or material from which polypeptides can be extracted in sufficient quantities and with adequate quality to permit measurement of relative polypeptide levels.

[0090] As used herein, the phrase “detecting and/or screening” for a condition refers to a method or process of determining if a subject has or does not have said condition. Where the condition is a likelihood or risk for a disease or disorder, the phrase “detecting and/or screening” will be understood to refer to a method or process of determining if a subject is at an increased or decreased likelihood for the disease or disorder.

[0091] As used herein, the term “progression to AML” refers to the process by which pre-leukemic syndromes such as MDS develop and/or transform into acute myeloid leukemia. It includes any and all steps along the continuum by which MDS proceeds to AML. The term “likelihood of progression to AML” refers to the likelihood that a subject diagnosed with, or suspected of having, MDS will develop AML.

[0092] A subject with an “increased likelihood of progression” or “an increased likelihood of MDS and/or AML” is also referred to having a high risk of progression or a high risk of MDS and/or AML. Likewise, a subject with a “decreased likelihood of progression” or “decreased likelihood of MDS and/or AML” is also referred to having a low risk of progression or a low risk of MDS and/or AML. An “intermediate” risk falls between high risk and low risk and a subject having a “moderate likelihood of progression” or a “moderate likelihood of MDS and/or AML” is also referred to having an intermediate risk of progression or an intermediate risk of MDS and/or AML.

[0093] The term “subject” as used herein refers to a human subject and includes, for example, a pediatric subject. In some embodiments, the subject is a patient diagnosed with, or suspected to have, MDS.

[0094] This disclosure also provides a process for normalizing target gene expression data or signal against at least one or against a geometric

means of multiple housekeeping gene target signals for improved gene expression analysis accuracy. Housekeeping genes are genes that are known to express at consistent mRNA level.

[0095] In an embodiment, the expression level is determined by one or more probes and/or one or more probe sets. In another embodiment, the one or more probes and/or the one or more probe sets, for example, include probes for the genes described herein comprised on a microarray.

[0096] The term "probe" as used herein refers to a nucleic acid molecule that comprises a sequence of nucleotides that will hybridize specifically to a target nucleic acid sequence such as the genes described herein. For example the probe comprises at least 10 or more bases or nucleotides that are complementary and hybridize to contiguous bases and/or nucleotides in the target nucleic acid sequence. The length of probe depends on the hybridization conditions and the sequences of the probe and nucleic acid target sequence and can for example be 10-20, 21-70, 71-100, 101-500 or more bases or nucleotides in length. The probes can optionally be fixed to a solid support such as an array chip, a microarray chip or bead based solid support.

[0097] In another embodiment, the expression is determined using one or more primer sets. In another embodiment, the nucleic acid targets are mRNA and are assayed using one or more target specific probes or one or more target specific primer pairs complimentary to each target sequence.

[0098] The term "primer" as used herein refers to a nucleic acid molecule, whether occurring naturally as in a purified restriction digest or produced synthetically (for example, a synthetic peptide nucleic acid (PNA) or a locked nucleic acid (LNA)), which is capable of acting as a point of synthesis when placed under conditions in which synthesis of a primer extension product, which is complementary to a nucleic acid strand is induced (e.g. in the presence of nucleotides and an inducing agent such as DNA polymerase

and at a suitable temperature and pH). The primer must be sufficiently long to prime the synthesis of the desired extension product in the presence of the inducing agent. The exact length of the primer will depend upon factors, including temperature, sequences of the primer and the methods used. A primer typically contains 15-25 or more nucleotides, although it can contain less. The factors involved in determining the appropriate length of primer are readily known to one of ordinary skill in the art.

[0099] Accordingly, in an embodiment, gene expression levels are measured using an agent that provides for determination of gene expression levels of at least one, optionally all, of the selected genes, wherein the agent comprises an oligonucleotide-immobilized substrate comprising a plurality of oligonucleotide probes or primers corresponding to the selected genes. A non-limiting example of such an agent includes a "microarray", comprising an ordered set of probes fixed to a solid surface that permits analysis such as gene expression analysis of a plurality of genes or gene sequences.

[00100] In one embodiment, a method described herein also comprises first obtaining a sample from the subject. The sample, in an embodiment, comprises a clinical specimen. As used herein, the term "clinical specimen" includes any sample from a subject which comprises nucleic acids, for example, DNA. Examples of clinical specimens useful in the present methods include, but are not limited to, blood and bone marrow samples. Other examples of samples useful in the present methods include a bodily fluid sample, optionally plasma, serum, blood, bone marrow, saliva, urine or CSF or a tissue or tumor sample. In an embodiment, the sample is submerged in a RNA preservation solution, for example to allow for storage. In an embodiment, the RNA preservation solution comprises Trizol®. The sample is in an embodiment, treated with a RNase inhibitor to prevent RNA degradation.

[00101] According to the methods described herein, similarity of the gene expression profile from a sample to one or more control profiles, may be

used to identify individuals having myelodysplastic syndrome (MDS) and/or acute myeloid leukemia (AML) or an increased likelihood of MDS and/or AML. For example, in an embodiment, the method comprises determining the level of similarity of a sample profile to one or more control profiles, wherein (i) a high level of similarity of the sample profile to an MDS and/or AML specific control profile; (ii) a low level of similarity to a non-MDS and/or non-AML control profile; and/or (iii) a higher level of similarity to an MDS and/or AML specific control profile than to a non-MDS and/or non-AML control profile indicates the presence of, or an increased likelihood of, MDS and/or AML.

[00102] Similarity of the gene expression profile from a sample to one or more control profiles, may also be used to identify individuals with MDS who have progressed to AML or have an increased likelihood of progression to AML. For example, in an embodiment, the method comprises determining the level of similarity of a sample profile to one or more control profiles, (i) a high level of similarity of the sample profile to a high risk AML progression control profile; (ii) a low level of similarity to low risk AML progression control profile; and/or (iii) a higher level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile indicates progression to AML or an increased likelihood of progression to AML.

[00103] In an embodiment, the “MDS specific control profile” is generated by measuring the gene expression levels of specified genes in genomic DNA from an individual subject, or population of subjects, who are known to have MDS. Similarly, in an embodiment, the “non-MDS control profile” is generated by measuring the gene expression levels of specified genes in genomic DNA from an individual subject, or population of subjects, who are known to not have MDS. In certain embodiments, the sample source from which the sample profile and control profile are derived is matched, so that they are both derived from the same or similar tissue. In other embodiments, the sample profile and control profile are derived from different tissues. In certain other embodiments, the MDS specific control profile and the

non-MDS control profile are derived from historical data and can indicate similarity of a sample to either the MDS or non-MDS profiles.

[00104] In another embodiment, the “AML specific control profile” is generated by measuring the gene expression levels of specified genes in genomic DNA from an individual subject, or population of subjects, who are known to have AML. Similarly, in an embodiment, the “non-AML control profile” is generated by measuring the gene expression levels of specified genes in genomic DNA from an individual subject, or population of subjects, who are known to not have AML. In certain embodiments, the sample source from which the sample profile and control profile are derived is matched, so that they are both derived from the same or similar tissue. In other embodiments, the sample profile and control profile are derived from different tissues. In certain other embodiments, the AML specific control profile and the non-AML control profile are derived from historical data and can indicate similarity of a sample to either the AML or non-AML profiles.

[00105] In an embodiment, the “high risk AML progression control profile” is generated by measuring the gene expression levels of specified genes in genomic DNA from an individual subject, or population of subjects, who are known to have a high risk of progression from MDS to AML. Similarly, in an embodiment, the “low risk AML progression control profile” is generated by measuring the gene expression levels of specified genes in genomic DNA from an individual subject, or population of subjects, who are known to have a low risk of progression from MDS to AML. In certain embodiments, the sample source from which the sample profile and control profile are derived is matched, so that they are both derived from the same or similar tissue. In other embodiments, the sample profile and control profile are derived from different tissues. In certain other embodiments, the high risk AML progression control profile and the low risk AML progression control profile are derived from historical data and can indicate similarity of a sample to either the high

risk AML progression control profile or the low risk AML progression control profiles.

[00106] Methods of determining the similarity between gene expression profiles are well known in the art. Methods of determining similarity may in some embodiments provide a non-quantitative measure of similarity, for example, using visual clustering and hierarchical clustering. In another embodiment, similarity may be determined using methods which provide a quantitative measure of similarity.

[00107] For example, in an embodiment, similarity may be measured using hierarchical clustering, optionally using Manhattan distance. The Manhattan distance function computes the distance that would be traveled to get from one data point to the other if a grid-like path is followed. The Manhattan distance between two items is the sum of the differences of their corresponding components.

[00108] The formula for this distance between a point $X=(X_1, X_2, \text{etc.})$ and a point $Y=(Y_1, Y_2, \text{etc.})$ is:

$$d = \sum_{i=1}^n |x_i - y_i|$$

Where n is the number of variables, and X_i and Y_i are the values of the variable, at points X and Y respectively.

[00109] In another embodiment, similarity may be measured by computing a "correlation coefficient", which is a measure of the interdependence of random variables that ranges in value from -1 to +1, indicating perfect negative correlation at -1, absence of correlation at zero, and perfect positive correlation at +1. In an embodiment, the correlation coefficient may be a linear correlation coefficient, for example, a Pearson product-moment correlation coefficient.

[00110] A Pearson correlation coefficient (r) is calculated using the following formula:

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}}$$

[00111] In one embodiment, x and y are the beta values for various genes in a sample profile and a control profile, respectively.

[00112] In an embodiment, a correlation coefficient calculated between the sample profile and the control profile indicates a high level of similarity to the control profile when the correlation coefficient has an absolute value between 0.5 to 1, optionally between 0.75 to 1, and a low level of similarity to the control profile when the correlation coefficient has an absolute value between 0 to 0.5, optionally between 0 to 0.25.

[00113] It will be appreciated that any “correlation value” which provides a quantitative scaling measure of similarity between gene expression profiles may be used to measure similarity.

[00114] The methods described herein may be used in combination with additional methods for classifying and/or prognosing MDS, AML and/or the progression from MDS to AML. In one embodiment, the methods further comprise identifying the WHO subtype of MDS. As used herein, the term “WHO subtype of MDS” refers to the WHO classification of myeloid neoplasms and acute leukemia (Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. Blood. 2016;127: 2391-2405). In particular, patients may be classified according to the WHO classification as having either a “low-risk MDS” or a “high-risk MDS”.

[00115] Accordingly, in one embodiment, the methods further comprise identifying the WHO subtype of the subject, where a high risk WHO subtype

indicates progression to AML or an increased likelihood of progression to AML and a low risk WHO subtype indicates no progression to AML or a decreased likelihood of progression to AML.

[00116] In another embodiment, the methods further comprise identifying the WHO MDS subtype of the subject, wherein

(a) (i) a high level of similarity of the sample profile to a high risk AML progression control profile; (ii) a low level of similarity to low risk AML progression control profile and/or (iii) a higher level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile; and

(b) a high risk WHO MDS subtype,

indicates a high risk of progression to AML, progression to AML, or an increased likelihood of progression to AML.

[00117] In another embodiment, the methods further comprise identifying the WHO MDS subtype of the subject, wherein

(a) (i) a low level of similarity of the sample profile to a high risk AML progression control profile; (ii) a high level of similarity to low risk AML progression control profile; and/or (iii) a lower level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile indicates no progression to AML or an decreased likelihood of progression to AML; and

(b) a low risk WHO MDS subtype,

indicates a low risk of progression to AML, no progression to AML, or a decreased likelihood of progression to AML.

[00118] In another embodiment, the methods further comprise identifying the WHO MDS subtype of the subject, wherein

(a) (i) a high level of similarity of the sample profile to a high risk AML progression control profile; (ii) a low level of similarity to low risk AML

progression control profile and/or (iii) a higher level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile; and

(b) a low risk WHO MDS subtype,

or

(a) (i) a low level of similarity of the sample profile to a high risk AML progression control profile; (ii) a high level of similarity to low risk AML progression control profile; and/or (iii) a lower level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile indicates no progression to AML or an decreased likelihood of progression to AML; and

(b) a high risk WHO MDS subtype,

indicates an intermediate risk of progression to AML.

[00119] Confirmation of a diagnosis of MDS or AML, or an indication of an increased likelihood of progression from MDS to AML aids in medical management of these conditions by optimizing the opportunity for early intervention and management.

[00120] Accordingly, an aspect of the disclosure provides a method of determining a course of management for a subject with MDS, AML, an increased likelihood of MDS and/or AML and/or an increased likelihood of progression from MDS to AML, comprising:

a) identifying an individual with MDS, AML, an increased likelihood of MDS and/or AML and/or an increased likelihood of progression from MDS to AML, according to the methods described herein; and

b) assigning a course of management for the MDS, AML, increased likelihood of MDS and/or AML and/or an increased likelihood of progression from MDS to AML and/or symptoms of MDS and/or AML, comprising applying

an appropriate medical intervention and/or treatment based on the results of the identification.

[00121] Methods of treating MDS, AML and progression from MDS to AML include, but are not limited to chemotherapy, radiation, stem cell transplantation and/or bone marrow transplantation.

[00122] In one embodiment, a subject identified as having progression to AML or an increased likelihood of progression to AML is treated with chemotherapy, radiation, stem cell transplantation and/or bone marrow transplantation. In another embodiment, a subject identified as having progression to AML or an increased likelihood of progression to AML is treated with increased vigilance including regular physical exams, blood tests and/or imaging in order to identify early signs and symptoms of AML. In yet another embodiment, a subject identified as having an intermediate risk of progression to AML or a moderate likelihood of progression to AML is treated with increased vigilance including regular physical exams, blood tests and/or imaging in order to identify early signs and symptoms of AML. Physical exams, blood tests and/or imaging to identify early signs and symptoms of AML may also be referred to as “watchful waiting”.

[00123] In another embodiment, the subject identified as not having progression to AML or a decreased likelihood of progression to AML is not treated with chemotherapy, radiation, stem cell transplantation and/or bone marrow transplantation.

[00124] The present disclosure also provides a method of selecting a treatment for a subject with progression to AML or an increased likelihood of progression to AML, comprising:

a) identifying the subject with progression to AML or an increased likelihood of progression to AML, according to the methods described herein; and

b) selecting a treatment for the subject, wherein the treatment comprises chemotherapy, radiation, stem cell transplantation, bone marrow transplantation and/or watchful waiting.

Kits

[00125] The disclosure also provides a kit for detecting and/or screening for MDS and/or acute myeloid leukemia (AML) or an increased likelihood of MDS and/or AML, in a sample, comprising, consisting of, or consisting essentially of:

a) at least one detection agent for determining the gene expression level of at least 3, optionally at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, or all of the human orthologues of the genes listed in Table 5; and optionally,

b) instructions for use.

[00126] The disclosure further provides a kit for detecting and/or screening for detecting and/or screening for progression to AML or an increased likelihood of progression to AML, in a human subject having MDS, in a sample, comprising, consisting of, or consisting essentially of:

a) at least one detection agent for determining the gene expression level of at least 3, optionally at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, or all of the human orthologues of the genes listed in Table 5; and optionally,

b) instructions for use.

[00127] In an embodiment, the detection agents comprise and/or are nucleic acid molecules. In a further embodiment, the detection agents are a set of probes or primers for determining the expression of at least 3, optionally at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, all of the human orthologues of the genes listed in Table 5.

[00128] In another embodiment, the detection agents comprise and/or are antibodies for determining the expression (e.g. polypeptide levels) of at least 3, optionally at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, or all of the human orthologues of the genes listed in Table 5.

[00129] In an embodiment, the kit comprises detection agents described herein, and optionally one or more specimen collectors, RNA preservation solutions and/or control samples.

[00130] In an embodiment, the specimen collector comprises a sterile vial or tube suitable for receiving a sample. In an embodiment, the specimen collector comprises RNA preservation solution.

[00131] In an embodiment the RNA preservation solution comprises one or more inhibitors of RNase. In another embodiment, the RNA preservation solution comprises Trizol® or other reagents designed to improve stability of RNA.

[00132] In another embodiment, the kit further comprises control samples. In one embodiment, the control sample is a positive control. In another embodiment, the control sample is a negative control.

[00133] In an embodiment, the antibody or probe is labeled. The label is optionally capable of producing, either directly or indirectly, a detectable signal. For example, the label may be radio-opaque or a radioisotope, such as ³H, ¹⁴C, ³²P, ³⁵S, ¹²³I, ¹²⁵I, ¹³¹I; a fluorescent (fluorophore) or chemiluminescent (chromophore) compound, such as fluorescein isothiocyanate, rhodamine or luciferin; an enzyme, such as alkaline phosphatase, beta-galactosidase or horseradish peroxidase; an imaging agent; or a metal ion.

[00134] In another embodiment, the kit comprises a computer-readable medium that causes a computer to compare gene expression from a sample at the selected genes to one or more control profiles and compute a correlation value between the sample and control profile.

[00135] In an embodiment, the kit is for use in a method described herein.

Disease models

[00136] The present disclosure shows that the deletion of GSK-3 β in a transgenic mouse model leads to a MDS-like disease *in vivo* and the combined deletion of GSK-3 α and GSK-3 β in a transgenic mouse model leads to an AML-like disease *in vivo*. Accordingly, the present disclosure provides methods for creating a disease model of myelodysplastic syndrome by disrupting or deleting GSK-3 β in a non-human transgenic model. The present disclosure provides methods for creating a disease model of acute myeloid leukemia by disrupting or deleting GSK-3 α and GSK-3 β in a non-human transgenic model. These models are useful for screening for test agents to treat and/or prevent MDS, AML and/or the progression from MDS to AML.

[00137] In one aspect of the disclosure, a transgenic non-human model of MDS and/or AML is provided. In one embodiment, the transgenic non-human model is a transgenic mouse.

[00138] The term “transgenic non-human animal” includes any member of the animal kingdom, except humans, in which one or more cells contain a genomic alteration introduced by way of human intervention such as by transgenic techniques known in the art. In one example of a genomic alteration, a transgene or nucleic acid is introduced into the cell, directly or indirectly, by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding but rather is directed to the introduction of a recombinant DNA molecule. This molecule may be integrated within a chromosome, or it may be extrachromosomally replicating DNA. In another example of a genetic alteration, a specific gene or section of genomic sequence is disrupted or deleted.

[00139] Non-limiting examples of animals that may be used in the present disclosure include mice, rats, squirrels, hamsters, guinea pigs, rabbits, pigs, sheep, baboons, monkeys, chimpanzees, birds and amphibians. In one embodiment, the animal is a mouse.

[00140] In one embodiment of the present disclosure, the genomic alteration comprises a disruption of the gene GSK-3 β and/or the gene GSK-3 α . In one embodiment, the disruption is a gene deletion or gene knock-out.

[00141] As used herein "GSK-3 α " refers to the gene encoding glycogen synthase kinase 3 alpha and includes, without limitation, GSK-3 α from any source such as those with sequences as shown in Genbank Accessions NM_019884 (human) and NM_001031667.1 (mouse), incorporated herein by reference in their entirety. In one embodiment, GSK-3 α is mouse GSK-3 α . In another embodiment, GSK-3 α is human GSK-3 α .

[00142] As used herein "GSK-3 β " refers to the gene encoding glycogen synthase kinase 3 beta and includes, without limitation, GSK-3 β from any source such as those with sequences as shown in Genbank Accessions NM_002093 (human), NM_001347232.1 (mouse) and NM_019827.6 (mouse), incorporated herein by reference in their entirety. In one embodiment, GSK-3 β is mouse GSK-3 β . In another embodiment, GSK-3 β is human GSK-3 β .

[00143] As used herein, the term "gene disruption" is used synonymously with "gene mutation" to refer to an alteration to the gene such the expression of the gene is decreased or suppressed compared to an animal where the gene is not disrupted or that the protein encoded by the gene is non-functional or of reduced function compared to the protein encoded by the wild-type gene. One example of a gene disruption is a gene deletion or gene knockout.

[00144] As used herein, the term "gene deletion" is used synonymously with "gene knockout" to specify that the entire nucleic acid sequence

corresponding to the gene of interest is deleted from the genome in at least one allele.

[00145] In one embodiment, one allele of GSK-3 α is disrupted. This is also referred to as a heterozygous disruption. In another embodiment, both alleles of GSK-3 α are disrupted. This is referred to as a homozygous disruption. Similarly, in another embodiment, one allele of GSK-3 β is disrupted. This is also referred to as a heterozygous disruption. In another embodiment, both alleles of GSK-3 β are disrupted. This is referred to as a homozygous disruption.

[00146] In one embodiment, GSK-3 α and/or GSK-3 β is conditionally disrupted. As is well known in the art, a conditional disruption such as a conditional deletion or knockout enables the disruption or deletion of a gene of interest in a specific tissue and/or at a specific time point (inducible), while in all other tissues or at all other time points the gene retains its wildtype function. For example, the Cre-lox system may be utilized for conditional deletions or disruptions of a gene. According, in one embodiment of the present disclosure, the transgenic mouse comprises a conditional disruption of one or both alleles of GSK-3 β and/or GSK-3 α , for example a conditional knockout of one or both alleles of GSK-3 β and/or GSK-3 α .

[00147] Methods of preparing transgenic animals having a deleted or disrupted gene are well known in the art. For example, embryonic stem cells containing a nucleic acid construct or vector of interest can be used to prepare a transgenic mouse. In particular, the embryonic stem cells are inserted into an early embryo for example using microinjection. For microinjection, approximately 10-20 embryonic stem cells are collected into a micropipette and injected into 3-5 day old blastocysts, preferably 3½ day old blastocysts, recovered from female mice. The injected blastocysts are re-implanted into a foster mother. When the pups are born, typically 20-21 days later, they are screened for the presence of the nucleic acid construct of the disclosure. For example, the tail tissue of the pups may be screened using

Southern blots and/or PCR. The heterozygotes are identified and can then be crossed with each other to generate homozygous animals with the gene deletion or disruption of interest. In another embodiment, as is known in art, the CRISPR/Cas9 system is used to generate a transgenic animal as described herein.

Uses of the disease models

[00148] In one embodiment, the transgenic animals described herein are useful models in studying MDS, AML and progression from MDS to AML. The animals can assist in studying the roles of GSK-3 α and/or GSK-3 β in these diseases. As shown in Figure 1, the combined deletion of GSK-3 α and GSK-3 β in a transgenic mouse leads to an AML-like disease in vivo. As shown in Figure 2, GSK-3 β deficiency alone leads to myelodysplasia and impaired hematopoiesis.

[00149] Accordingly, in another embodiment, the transgenic animals of the disclosure are useful as animal models for testing potential agents that can modulate the effect of disruption of GSK-3 β and/or GSK-3 α and treat conditions involving deletion and/or disruption of GSK-3 β and/or GSK-3 α . The transgenic animals of the disclosure are also useful as animal models for testing potential agents that can treat or prevent MDS, AML and/or the progression of MDS to AML.

[00150] In particular, in one embodiment a method of screening a test agent for treatment or prevention of myelodysplastic syndrome (MDS) is provided, the method comprising a) administering a test agent to i) a transgenic mouse model of MDS whose genome comprises a disruption of GSK-3 β , ii) a tissue from said transgenic mouse model, or iii) or a cell derived from said transgenic mouse model; and b) determining the effect of said test agent on the mouse model, tissue or cell, wherein a reduction of signs and/or symptoms of MDS in the mouse, tissue or cell identifies a test agent for the treatment or prevention of MDS.

[00151] Signs and/or symptoms of MDS in a mouse, cell or tissue model include, but are not limited to, impaired hematopoiesis, lobated or hypersegmented neutrophils, an increased frequency of blasts, extramedullary hematopoiesis, accumulation of monocytes and/or granulocytes in the spleen, accretion of hypersegmented neutrophils in the peripheral blood, micromegakaryocytes with hyperchromatic nuclei in the spleen. Other signs and/or symptoms of MDS include increased mortality.

[00152] In one embodiment, a test agent for the treatment of MDS is identified when at least one of the signs and/or symptoms of MDS in a mouse, cell or tissue model of MDS is reduced or decreased by at least 5, 10, 25, 50, 100% compared to a mouse model of MDS to which the test agent has not been administered.

[00153] In another embodiment, a method of screening a test agent for treatment or prevention of acute myeloid leukemia (AML) is provided, comprising a) administering a test agent to i) a transgenic mouse model of AML whose genome comprises a disruption of GSK-3 α and a disruption of GSK-3 β , ii) a tissue from said transgenic mouse model or iii) a cell derived from said transgenic mouse model; and b) determining the effect of said agent on mouse model, tissue or cell, wherein a reduction of signs and/or symptoms of AML in the mouse, tissue or cell identifies a test agent for the treatment or prevention of AML.

[00154] Signs and/or symptoms of AML in a mouse, cell or tissue model include, but are not limited to, increased mortality, skin lesions, bone marrow hypercellularity, splenomegaly, elevated leukocyte counts in peripheral blood and/or lymph nodes, increased presence of immature myeloid cells and leukemic blasts in the bone marrow, spleen, peripheral blood and/or lymph nodes. Other signs and/or symptoms of AML include increased mortality and/or cell death and an accumulation of blasts.

[00155] In one embodiment, a test agent for the treatment of AML is identified when at least one of the signs/and or symptoms of AML in a mouse, cell or tissue model of AML is reduced or decreased by at least 5, 10, 25, 50, 100% compared to a mouse model of AML to which the test agent has not been administered.

[00156] In embodiments where the GSK-3 β and/or GSK-3 α is a conditional disruption, the disruption is optionally induced prior to administering the test agent. The effects of the test agent on the signs and/or symptoms of MDS and/or AML can be compared to a control animal such as a transgenic animal not receiving the test agent following induction of the disruption.

[00157] The above screening methods of the disclosure can advantageously also use tissue, organs and/or cells isolated from the transgenic animal models described above.

[00158] The test agents contemplated herein include, but are not limited to chemical or biological molecules such as simple or complex organic molecules, biomolecules, metal-containing compounds, carbohydrates, peptides, proteins, peptidomimetics, glycoproteins, lipoproteins, sugars, polysaccharides, nucleic acids, antibodies, cells or combinations thereof. Such a test agent may be a naturally-occurring product or a synthetic product.

[00159] The test agents in the screening assays can be generated by methods well known to those skilled in the art, for example, well known methods for producing pluralities of compounds including without limitation: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection (see, e.g. Lam, 1997). Libraries containing large numbers of natural and synthetic compounds, including antibodies, also can be obtained from commercial sources. Combinatorial libraries of molecules

can be prepared using well known combinatorial chemistry methods (Gordon et al., J. Med. Chem. 37: 1233–1251 (1994); Gordon et al., J. Med. Chem. 37: 1385–1401 (1994); Gordon et al., Acc. Chem. Res. 29:144–154 (1996); Wilson and Czarnik, eds.

Cells, cell lines and uses thereof

[00160] The present inventors have also shown that mouse cells comprising a GSK-3 β knockout provides a MDS-like disease model and mouse cells comprising a GSK-3 α and GSK-3 β double knockout provides an AML-like disease model.

[00161] Mouse cells are commonly used as model mammalian cells. Accordingly, it is expected that human cells comprising a GSK-3 β knockout would provide a MDS-like disease model. Similarly, it is expected that human cells comprising a GSK-3 α and GSK-3 β double knockout would provide an AML-like disease model.

[00162] Accordingly, a method of screening a test agent for treatment or prevention of myelodysplastic syndrome (MDS) is provided, comprising a) administering a test agent to a cell comprising a disruption of GSK-3 β and b) detecting the effect of said agent on the cell. In one embodiment, the cell is a human cell.

[00163] In another embodiment, a method of screening a test agent for treatment or prevention of acute myeloid leukemia comprising a) administering a test agent to a cell comprising a disruption of GSK-3 α and a disruption of GSK-3 β and b) detecting the effect of said agent on the cell. In one embodiment, the cell is a human cell.

[00164] As used herein, the term cell refers to a single cell as well as a plurality of cells and also includes a cell line.

[00165] In one aspect, the methods described herein comprise screening a test agent for treatment or prevention of MDS or AML by

detecting an effect of the test agent on the cell. In one embodiment, the effect is indicative of the activity of the test agent to treat or prevent of MDS or AML.

[00166] In one embodiment, “detecting an effect” comprises monitoring or determining cell size or morphology, expression of cell markers, the emergence of cell types or the biochemical make-up of the cell. For example, in one embodiment “detecting an effect” includes, but is not limited to, using methods such as immunohistochemistry (IF1C), ELISA, reporter genes, PCR or RT-PCR, fluorescent labels, cytometric bead arrays, DNA arrays, flow cytometry or optical analysis to detect the effect of a test agent on cells comprising a disruption of GSK-3 β and/or GSK-3 α or normal (wild-type) cells.

[00167] In one embodiment, the effect of said agent on the cell is compared to the effect of said agent on a cell which does not comprise a disruption of GSK-3 β and/or GSK-3 α . In one embodiment, a test agent which has a differential effect on a cell which comprises a disruption of GSK-3 β and/or GSK-3 α compared to a cell that does not comprise a disruption of GSK-3 β and/or GSK-3 α is identified as an agent for treating or preventing MDS and/or AML.

[00168] Cells derived from the transgenic model animals described above and cells wherein the GSK-3 β and/or GSK-3 α disruption is directly introduced into the cell or cell line (for example, using techniques of genetic engineering that are well known within the art) are both contemplated within the present disclosure.

[00169] Optionally, the methods described herein further comprise testing the agents identified using the screening methods described herein for toxicity.

[00170] In one embodiment, the method of screening a test agent is a high throughput screening assay. The term “high throughput screening” as used herein refers to automated *in vitro* testing of the effect of test agents on

cells and such screening is typically performed with the aid of computer or robot-controlled processes.

[00171] In one embodiment, there is provided a composition comprising microtitre plates with a plurality of receptacles wherein one or more of the receptacles contain cells comprising a disruption of GSK-3 β and/or GSK-3 α as described herein. In one embodiment, the microtiter plates are high-throughput format microtitre plates. In one embodiment, the plates are high-density plates suitable for cell-based assays or culture. In some embodiments, the plates have 2 or more, 96, 384, or 1536 individual receptacles or wells and are suitable for use in high-throughput screening such as in automated systems and/or robotic systems.

[00172] As used herein, the term “receptacle” refers to a container suitable for the maintenance and culture of cells. In some embodiments, the receptacle may be a well on a plate such as a microtitre plate, which optionally contains a plurality of wells or receptacles. Optionally, the receptacle is designed so as to prevent contamination from adjacent receptacles. Typically, the receptacle will also contain media to provide nutrients to the one or more cells and allow for cell growth.

[00173] Terms of degree such as “substantially”, “about” and “approximately” as used herein mean a reasonable amount of deviation of the modified term such that the end result is not significantly changed. These terms of degree should be construed as including a deviation of at least $\pm 5\%$ of the modified term if this deviation would not negate the meaning of the word it modifies or unless the context suggests otherwise to a person skilled in the art.

[00174] The above disclosure generally describes the present disclosure. A more complete understanding can be obtained by reference to the following specific examples. These examples are described solely for the purpose of illustration and are not intended to limit the scope of the disclosure.

Changes in form and substitution of equivalents are contemplated as circumstances might suggest or render expedient. Although specific terms have been employed herein, such terms are intended in a descriptive sense and not for purposes of limitation.

[00175] The following non-limiting examples are illustrative of the present disclosure:

EXAMPLES

[00176] The following examples illustrate the scope of the disclosure. Specific elements of the example are for descriptive purposes only and are not intended to limit the scope of the disclosure. Those skilled in the art could develop equivalent methods and utilize comparable materials that are within the scope of the disclosure.

Example 1 Creation of disease models and gene signatures for MDS and AML

EXPERIMENTAL PROCEDURES AND MATERIALS

GSK-3 Mouse Models and *In vivo* Assays

[00177] The generation of conditional knockout mice for Gsk3 α and Gsk3 β has been described previously (MacAulay et al., 2007; Patel et al., 2008), and mice were kindly provided by Dr. Brad Doble (Stem Cell and Cancer Research Institute, Hamilton, Ontario, Canada). Tamoxifen-inducible Rosa26-CreERTM mice carrying the Cre recombinase and B6.SJL-Ptprca Pep3b/BoyJ (B16/SJL) mice were purchased from the Jackson Laboratory. To obtain conditional GSK-3 β -knockout (KO), Gsk3 $\beta^{fix/fix}$ mice were crossed with Rosa26-CreERTM mice. Subsequently these mice were crossed with Gsk3 $\alpha^{-/-}$ mice (obtained from germline deletion of GSK-3 α in conditional knockout animals, as previously described (MacAulay et al., 2007) to obtain GSK-3-double mutant mice). Both, conditional GSK-3 β -KO and GSK-3 α /GSK-3 β -knockout mice were used for experiments. Genotyping was performed on

genomic DNA isolated from ear notches with the Extract-NAMP™ Tissue PCR Kit (Sigma Aldrich). PCR analysis was performed with primers for GSK-3 α , GSK-3 β and Cre (Table 1).

[00178] For syngeneic transplantation, 1×10^6 of freshly isolated or Lin⁻ bone marrow (BM) cells from *Gsk3 α ^{-/-}*, *Gsk3 β ^{+/*flx*}*, *Gsk3 β ^{flx/flx}*, double *Gsk3 α ^{-/-} Gsk3 β ^{flx/flx}*, double *Gsk3 α ^{-/-} Gsk3 β ^{+/*flx*}* or wild-type mice were injected into the tail vein of lethally irradiated congenic recipients. Three weeks after injection, conditional knockout was induced by 3 sequential injections of 250 μ l tamoxifen (diluted in 1% captisol) into the tail vein, every 2-3 days. For competitive transplantation, 1×10^4 Lin⁻ Sca1⁺ c-Kit⁺ (LSK) or 1×10^5 Lin⁻ Sca1⁻ c-Kit⁻ (non-LSK) BM cells from *Gsk3 β ^{+/*flx*}* or *Gsk3 β ^{flx/flx}* CD45.2⁺ mice were admixed with 1×10^6 BM cells from CD45.1⁺ mice and injected into lethally irradiated BL6/SJL mice (CD45.1⁺) mice. GSK-3 β -deletion was induced by tamoxifen. Similar transplantation settings were also applied in NOD/SCID mice with tamoxifen induction 2 weeks post-transplant. Prior to secondary transplants, CD45.1⁺ cells were removed from total BM of primary recipients using the Mouse Hematopoietic Progenitor Cell Enrichment Kit containing antibodies directed against CD45.1, CD11b, B220, GR-1 and Ter119 (Stem Cell Technologies). Mice were bred, transplanted and maintained in the SCC-RI animal barrier facility and all procedures received the approval of the animal ethics board of McMaster University.

Flow Cytometry Analysis and Sorting

[00179] PB, spleen or BM single cell suspensions were washed in PBS, incubated for 15 min. at 4°C with a blocking solution and then stained for 20 min. at 4°C with antibodies. To check engraftment of freshly harvested BM, cells were stained with conjugated anti-mouse antibodies against CD45, CD45.1, and CD45.2 (BD Biosciences) and analyzed using a LSR II or FACS Calibur flow cytometer (BD Biosciences).

[00180] Engrafted cells were further analyzed on a LSR II after staining with conjugated anti-mouse antibodies against CD45, CD45.1, CD45.2, Granulocytes (CD11b, Gr-1), Monocytes/Macrophages (CD14, F4/80), Erythrocytes (CD71, Ter119), Megakaryocytes (CD61,CD41), B lymphocytes (B220, CD19, IgM), T lymphocytes (CD3, CD4, CD8) and HSCs/progenitors (Lineage, c-Kit, Sca-1, CD150, CD48, CD34, FCγII/III, IL-7Rα) (BD Biosciences). For β-catenin activity, bone marrow cells were first stained with lineage-antibodies and then stained with either hematopoietic stem cell or progenitor cocktails of antibodies followed by fixation with 0.8% paraformaldehyde for 10 min. Cells were then washed, permeabilized and stained with DAPI and an antibody recognizing the active (de-phosphorylated) form of β-catenin (ABC, clone 8E7, Millipore) or isotype controls directly labeled (Xenon kit, Invitrogen) with Alexa Fluor-488, as described previously (Abrahamsson et al., 2009). Subsequent analyses were done with the FlowJo software (TreeStar). For sorting Lin⁻ bone marrow, cells were stained with APC-conjugated c-Kit, PE-conjugated Sca-1, PB-conjugated CD45.2 and 7AAD (all BD Biosciences). Cells were sorted on a FACS Aria II (BD Biosciences).

Morphological Analysis

[00181] Cytospins from bone marrow, spleen or liver samples were prepared by diluting 1×10^5 cells in 100 μl PBS and spinning them on glass microscope slides using the Shandon Cytospin 3 (Block Scientific, Inc.). Blood smears from peripheral blood were prepared by pipetting 50 μl of fresh blood mixed with EDTA on a glass microscope slide and spreading it by pulling the edge of another microscope slide over the surface of the first slide starting from the blood drop. Differential staining using Wright's Giemsa stain was performed with the Shandon Kwik-Diff Stain Kit (Thermo Scientific). Tissues were fixed in 10% buffered Formalin (Fisher Scientific) and bones were decalcified using Immunocal (Decal Chemical). Sections were stained with hematoxylin and eosin (H&E) stain for morphologic and cytologic assessment.

For immunohistochemistry, MPO (A0398; Dako) and Mac2 (CL8942AP; Cedarlane) antibodies were used to identify myeloid cells. Images were acquired using either a Scanscope Slide Scanner with the Image Scope software (Aperio) or an Olympus IX 81 Inverted Microscope.

Cell cycle and apoptosis analysis

[00182] Mice were administered 200 μ l of 10 mg/ml BrdU (BD) by intraperitoneal injection at 12 and 24 hrs prior to sacrifice. Isolated bone marrow cells were collected for flow cytometric detection of BrdU-FITC uptake according to the manufacturer's instructions (FITC-BrdU flow kit; BD Biosciences). For cell cycle analysis on live cells, stained cells were incubated with 5 μ g/mL Hoechst 33342 (Invitrogen) for 10 min. prior to flow cytometric analysis. For the apoptosis assay, 7-AAD and Annexin V-APC (BD Biosciences) was used in combination with either hematopoietic stem cell or progenitor cocktails of antibodies. All procedures were analyzed on LSRII and FlowJo.

Hematopoietic Colony-Forming Cell Assay and Rapamycin treatment

[00183] Freshly isolated CD45.2 positive bone marrow cells from transplanted mice were counted by trypan blue exclusion and 3×10^3 Lin-cells or 2×10^4 to 5×10^4 total BM cells were plated in triplicate in methylcellulose medium containing rm SCF, rm IL-3, rh IL-6 and rh EPO (Methocult M3434, StemCell Technologies). After 8 to 12 days the dishes were scored for different hematopoietic colonies. To block PI3K/Akt/mTOR signaling, either vehicle, dissolved 100nM Rapamycin (Calbiochem, Millipore), dissolved 10 μ M Wortmannin (Selleck Chemicals LLC) or dissolved 175nM LY2584702 (Selleck Chemicals LLC) was added to the methylcellulose mixture prior to plating as described previously (Janes et al., 2010).

Quantitative RT-PCR

[00184] Total RNA was extracted from mouse CD45.2-positive BM and Spleen cell or FACS-sorted LSK cells using the Norgen Total RNA Purification

Kit (Norgen Biotek Corp.). 100 to 200 ng of total RNA was reverse transcribed into cDNA with qScriptTM cDNA SuperMix (Quanta Biosciences) according to the manufacturer's protocol. cDNA was diluted to a concentration of 0.5 ng/ μ l (based on the RNA concentration) and 1 or 2 μ l of the dilution was mixed with 12.5 μ l GoTaq[®] qPCR Master Mix (Promega), 2.5 μ l of 10 μ M primer mix and nuclease-free water, to a final volume of 25 μ l per reaction. Quantitative real time PCR was performed with the MX3000P QPCR System (Stratagene, CA). Each reaction was performed in triplicate and the mean Ct value was used to calculate the relative fold changes using the $\Delta\Delta$ Ct method. The primers used are listed in Table 1.

Western blot analysis

[00185] Total protein fraction from mouse Lin- bone marrow cells was extracted in 1% SDS Laemmli Sample buffer (Bio-Rad), separated by SDS-PAGE, and transferred onto nitrocellulose membranes as previously described (Benoit et al., 2010). Nuclear proteins were extracted using the NE-PER[®] Nuclear and Cytoplasmic Extraction kit (Pierce Thermo Scientific) according to the manufacturer's recommendations. Purity of the isolated nuclear fractions was validated by monitoring levels of pan histone H3 (nuclear) vs. GAPDH (cytoplasmic) by western blot. Membranes were blocked in PBS containing 5% skim milk and 0.1% TWEEN 20 (Bio-Rad). Primary antibodies anti- β -Catenin (1:1000; mouse monoclonal, 610154; BD Transduction Laboratories), anti-histone H3 (1:200 000; rabbit polyclonal, ab1791; Abcam), anti-phospho- S473-AKT (1:1000; mouse monoclonal, 4151; Cell Signaling Technology), anti-AKT (1:1000; rabbit monoclonal, clone C67E7; Cell Signaling Technology), anti-GAPDH (1:80 000; mouse monoclonal, clone 6C5; Abcam), and anti-GSK-3 α/β (1:2000; mouse monoclonal, clone 21A; Life Technologies), as well as associated HRP-conjugated secondary antibodies (1:10 000, Bio-Rad) were used as described previously (Benoit et al., 2010). Images were acquired using a ChemiDocTM XRS+ System with Image LabTM Software (Bio Rad). Quantitative optical

densitometry analysis of bands was performed using Image J software (National Institutes of Health).

Metabolic characterization

[00186] Cells were washed and placed in XF96 V3 PS cell culture microplates (Seahorse Bioscience) on the XF Base Medium (Seahorse Bioscience) supplemented for final concentrations of 0.5 mM sodium pyruvate, 1 mM L-glutamine and 3.2 g/L glucose and adjusted to pH 7.4 at room temperature. Cells were incubated for 1 hour at 37°C in an XF Prep Station (Seahorse Bioscience). XFe96 Flux Pak hydrated cartridges were placed on top of the cells for data acquisition using an XFe96 analyzer (Seahorse Bioscience). The baseline levels of oxygen consumption rates (OCR) and extra-cellular acidification rates (ECAR) were used to establish the resting cellular metabolic activity (Zhang et al., 2012). To analyze mitochondrial function, pharmacological inhibitors properly optimized were added via port injection as follows: Oligomycin (1 mM), Carbonyl cyanide-4 (trifluoromethoxy)phenylhydrazone (FCCP, 1 mM), rotenone (0.5 mM) and antimycin A (0.5 mM). ATP-linked respiration was calculated as the difference in oxygen consumption rate between baseline and after oligomycin addition. Spare respiratory capacity was calculated as the difference between baseline and after addition of FCCP.

Global gene expression profiling and clinical outcome analysis

[00187] Freshly isolated LSK BM cells from $Gsk3\alpha^{+/+}\beta^{+/flx}$ (WT), $Gsk3\alpha^{+/+}\beta^{flx/flx}$ (B KO), $Gsk3\alpha^{-/-}\beta^{+/+}$ (AKO) and $Gsk3\alpha^{-/-}\beta^{flx/flx}$ (DKO) mice were sorted based on surface marker expression (CD45.1⁺) and total RNA was extracted and amplified from each purified population, as described previously (Wang et al., 2004). Amplified-labeled RNA was hybridized to Mouse 430 2.0 genechip microarrays (Affymetrix, USA) using standard protocols at the London Regional Genomics Center (Ontario, Canada). Arrays are available at GEO database under accession number GSE53352. The data analysis,

including RMA normalization and clustering was performed using Partek Genomics Suite 6.6 software. A list of over 1000 genes significantly dysregulated between GSK-3 WT and BKO ($p < 0.05$, fold-change 2 or < 1.5) was used as a transcriptome differential gene list as well as a differential list of 63 genes with $p < 0.05$, fold-change 2 or < 5 was used as a GSK-3 β gene signature list. Gene Ontology analysis was performed using DAVID bioinformatics resource 6.7 Functional Annotation Tool Suite. The top three GOTERM_BP_FAT annotations for a given set of gene inputs were listed as annotation terms (Huang da et al., 2009). Both GSK-3 β transcriptome and 63-gene GSK-3 β molecular signature were applied to GEO data set of adult (Sridhar et al., 2009) (GSE18366) and pediatric MDS (Bresolin et al., 2012) (GSE29326) using clustering analysis to discriminate disease subtypes. Human array GEO dataset of non-diseased, MDS, CMML and AML patients (GSE15061) (Mills et al., 2009) was compared to GSK-3 β transcriptional profile using common gene entities between human and mouse array technologies as previously described (Schnerch et al., 2010). In particular, profiling of human samples was performed using arrays of human orthologues to the mouse genes. Detection of molecular signatures was performed using Gene Set Enrichment Analysis software (<http://www.broadinstitute.org/gsea>) using the MSigDB database of the Broad institute. Clinical prognostic analysis was performed as described previously (Eppert et al., 2011). Briefly, array expression data from pediatric MDS samples (GSE29326) was filtered based on either a 63 gene signature or full differential GSK-3 β gene list, multiple probe sets were averaged per gene. The log₂ expression values were centered to zero median. The transformed values for genes that had been identified as being down-regulated in GSK-3 β KO had their sign switched. Subsequently, the sum of the median-centered log₂ expression values was used as the risk score for each patient, and 32 patients were split into high- or low-risk groups depending if the score was higher or lower than the mean risk score. These groups were assessed for MDS prognosis and AML

transformation risk using progression-free survival test or progression-time to AML test respectively (univariate Kaplan-Meier analysis with log-rank test statistics).

Statistical analysis

[00188] Statistical analyses were performed by using two-tailed *t*-test comparisons with GraphPad Prism (GraphPad Software, Inc., San Diego, CA). Values were compared as means $\pm$ standard error of the mean, and $p < 0.05$ and <0.01 was considered statistically significant and is indicated by asterisks in the figures. PFS and time to AML transformation was demonstrated using Kaplan-Meier curves. The p-values between the curves were generated using Log-Rank (Mantel-Cox) test.

RESULTS

Combined GSK-3 α and β allelic deficiency leads to AML

[00189] Despite structural redundancy, *Gsk3 β ^(-/-)* knockout mice die perinatally, whereas *Gsk3 α ^(-/-)* knockout mice are viable with no noticeable phenotype in hematopoiesis (Doble et al., 2007). However, deletion of both GSK-3 α and β was required in order to manifest a block in stem cell differentiation and pluripotency retention in mouse embryonic stem cells (Kelly et al., 2011), suggestive of a redundant role of GSK-3 α and β in mouse embryonic stem cells. To investigate the direct effect of GSK-3-deletion in adult hematopoietic stem/progenitor cells (HSPC) while circumventing allelic compensatory effects and early development mortality (Rossi et al., 2012; Wirth et al., 2007), a regulated GSK-3 β -null genetic system was created by crossing floxed GSK-3 β mice with tamoxifen-inducible ROSA26-Cre transgenic mice (Rosa26-CreERTM) (Fig.1a-c) that were then crossed with *Gsk3 α ^(-/-)* to generate both *Gsk3 α ^(-/-) β ^{flx/flx} CreERTM* (α ^(-/-) β ^{flx/flx}) and *Gsk3 α ^(-/-) β ^{+flx} CreERTM* (α ^(-/-) β ^{+flx}). Having established faithful targeting and effects on the HSPC compartment (Fig.1d), lineage-depleted bone marrow (Lin⁻BM, CD45.2) was transplanted into lethally irradiated congenic recipient mice

(CD45.1) to evaluate the effects of combinatorial GSK-3 allele deletion on *in vivo* hematopoietic regenerative capacity (Fig.2a). Five weeks after transplantation, peripheral blood from mice was analyzed to determine robust donor (CD45.2) chimerism (Fig. 1e), and were subsequently administered tamoxifen to induce allele deletion (Gan et al., 2008; Lobry et al., 2013), then followed for 6 weeks, at which time a significant number of mice were analyzed (n=15). The remaining mice were followed for longer-term survival and changes in hematopoiesis (Fig.2b).

[00190] In contrast to recipient mice transplanted with either *Gsk3α*^{-/-} or *α*^{-/-}*β*^{+/*flx*} BM cells, all *Gsk3α*^{-/-}*β*^{*flx/flx*} BM recipients displayed severe illness, eruption of skin lesions, and rapid morbidity starting as early as 2 weeks post-tamoxifen treatment (Fig.2b). Since *Gsk3α*^{-/-}*β*^{+/*flx*} were equivalent to *Gsk3α*^{-/-}*β*^{+/+} BM recipients, this demonstrated that a single 13 allele is haplosufficient to sustain normal hematopoiesis even in the absence of *Gsk3α*. Further analysis of recipient hematopoietic tissue revealed competitive chimerism measured by serial transplantation (Fig. 1e and f) accompanied by BM hypercellularity, severe splenomegaly (Fig. 2c), and more importantly, displayed a greater than 5-fold increase in total leukocytes in both lymph nodes (LN) and peripheral blood (Fig. 2d). Along with anemia, leukocytosis and granulocytic infiltration in the liver (Fig. 1g and h), morphological assessment of *Gsk3α*^{-/-}*β*^{*flx/flx*} skin lesions was consistent with neutrophilic dermatosis (Fig. 1i), similar to previous reports describing human AML patients with acute febrile neutrophilic dermatosis or Sweet's syndrome (Anzalone and Cohen, 2013). By contrast, both groups of recipient mice transplanted with either *Gsk3α*^{-/-}*β*^{+/*flx*} or *Gsk3α*^{-/-} BM demonstrated normal hematopoiesis (Fig. 1j and k). Furthermore, in both BM and spleen compartments, immature myeloid cells, defined as Gr-1⁺ CD11b⁺ ckit⁺ F4/80⁻ CD115⁻ blast cell population (Fig. 2e and f), were detected at high levels in *Gsk3α*^{-/-}*β*^{*flx/flx*} recipients (Fig. 2g), along with evidence of granulocytosis in the blood (>20%) and lymph node (>10%) (Fig. 2h), when compared to both

$Gsk3\alpha^{-/-}\beta^{+/flx}$ and $Gsk3\alpha^{-/-}$ recipients (Fig. 1k). Consistent with AML blast cell morphology, pathological features of both GSK-3 α and β deficient hematopoietic cells are in accordance with both Bethesda (Kogan et al., 2002) and WHO criteria for AML (Vardiman, 2010). Strikingly, expansion of HSPC (LSK) as well as GMPs in $Gsk3\alpha^{-/-}\beta^{flx/flx}$ primary recipients (Fig. 2i), carried the AML disease to secondary recipients that experienced high mortality within a few weeks (Fig. 2j) owing to the presence of a leukemia initiating cell (L-IC) subset driven by combined $Gsk3\alpha/\beta$ deficiency.

[00191] These *in vivo* results demonstrate that induction of murine AML is not uniquely dependent on allelic dosage deficiency between GSK-3 homologues but is interdependent on α vs. β GSK-3, where disruption of GSK-3 β is the main driver for the establishment of hematopoietic disorder leading to leukemic progression that is dependent on the absence of GSK-3 α .

GSK-3 β deficiency impairs hematopoiesis by inducing myelodysplasia.

[00192] Since $Gsk3\alpha^{-/-}$ mice have no phenotype, the effect of GSK-3 β deletion in exclusion on hematopoiesis was tested. HSC enriched Lin⁻BM from both donor $Gsk3\beta^{+/flx}$ and $^{flx/flx}$ mutant mice (Fig. 1a) were assessed for their regenerative capacity in the transplantation settings (Fig. 3a). The majority of $Gsk3\beta^{flx/flx}$ recipients developed symptoms of cachexia, abnormal gait and labored breathing, and had to be sacrificed by 8-10 weeks post-tamoxifen treatment compared to $Gsk3\beta^{+/flx}$ recipients or untreated mice (Fig. 3b). Loss of mice mainly occurred due to cachexia or BM failure, leading to severe anemia (Fig. 4a). BM analysis showed significant difference in cellularity in $Gsk3\beta^{flx/flx}$ recipients (Fig. 3c) with notable increases in granulocytes along with reduced erythrocyte, monocyte and lymphocyte frequencies but no change in platelet levels (Fig. 4b). In contrast, no abnormalities were observed in any controls (e.g. untreated mice or treated mice transplanted with $Gsk3\beta^{+/flx}$ or wild-type Lin⁻BM cells) (Fig. 3b-c). Among granulocytic subsets, an increase in dysplastic neutrophils frequency (Fig. 3d), as well as undifferentiated myeloid CD11b^{high}/Gr-1⁺/c-Kit⁺ levels, (Fig. 3e and Fig. 4c)

were also observed in BM of *Gsk3 β ^{flx/flx}* recipients. This indicated abnormal myeloid differentiation and marked nuclear lobation with hypersegmentation in neutrophils (Fig. 4d) consistent with myelodysplasia (Kogan et al., 2002). Erythrocyte and platelet counts varied in some recipients illustrating features of cytopenia (Fig. 4e). In addition, examination of splenic composition of *Gsk3 β ^{flx/flx}* recipients showed a high incidence of splenomegaly (Fig. 4f) and revealed extensive growth of the granulocytic-myeloid lineage, but lack of erythrocytes (Fig. 4g). Histologically, marked expansion of the red pulp with diffuse infiltration by myeloid and granulocytic cells was observed, resulting in the destruction of splenic architecture (Fig. 3f). Dysplastic megakaryocytes in the spleen, such as the presence of micro-megakaryocytes with hypolobulated and hyperchromatic nuclei consistent with myelodysplasia was also observed (Fig. 3f). This expansion of the granulocytic compartment was accompanied by suppressed monocytic generation in *Gsk3 β ^{flx/flx}* recipient mice featuring myeloid proliferation without monocytosis in contrast to AML *Gsk3 α/β -null* phenotype (Fig. 5a-b). Based on the Bethesda criteria (Kogan et al., 2002), features from deletion of GSK-3 β alleles are indicative of severe myelodysplasia that are morphologically similar to human MDS with presence of some degree of granulocytosis (Cazzola, 2011; Natelson and Pyatt, 2013; Tiu and Sekeres, 2014).

[00193] To functionally address whether the myelodysplastic phenotype (Beurlet et al., 2013) observed contains cells with disease initiating properties, BM cells from primary recipients were serially transplanted into secondary mice (Fig. 3g). Secondary recipients developed an identical hematopoietic disorder as primary recipients, and the level of stable disease was dependent on the number of injected cells (Fig. 3h and Fig. 5c). Similar dysplastic features and lethality were also consistent with primary transplanted mice (Fig. 3i and Fig. 5d-e). Based on these results, and without being bound by theory, it is suggested that deletion of GSK-3 β allows for the generation of self-renewing cells that can be functionally defined as MDS-Initiating Cells

(MDS-ICs) capable of sustaining MDS *in vivo*. Accordingly, the present disclosure also provides self-renewing cells comprises a deletion of GSK-3 β . These cells are also useful in the methods described herein for screening for test agents for treating or preventing MDS.

HSCs are vulnerable to GSK-3 β disruption for disease onset

[00194] Since MDS is thought to be a group of clonal disorders likely to originate from HSPCs (Corey et al., 2007; Nimer, 2008), the composition of the HSPC compartment of *Gsk3 β ^{flx/flx}* recipients was examined. The BM LSK cells were increased, specifically the CD150⁺CD48⁺ subset, shown to have a myeloid commitment bias (Challen et al., 2010). However, this increase neither affected LT-HSCs (defined as LSK CD150⁺CD48⁻) nor late multipotent progenitors (MPP, defined as LSK CD150⁻CD48⁺) (Fig. 6a). Observed differentiation bias linked to *Gsk3 β* deficiency was coupled with a significant increase in the absolute numbers of granulocytic/monocytic progenitors (GMP) together with a decrease of the megakaryocyte–erythrocyte progenitor population (MEP) with no effects on common myeloid and lymphoid progenitors (CMP and CLP) (Fig. 6b). This unbalanced progenitor phenotype is exacerbated by an increase of proliferation of the LSK and LT-HSC fraction in *Gsk3 β ^{flx/flx}* recipients (Fig. 6c) and is mitigated by apoptosis of LSK CD150⁺CD48⁺ progenitors (Fig. 6d).

[00195] Beyond phenotypic analysis, HSCs (LSK) and progenitors (non-LSK) were isolated from the BM (Chao et al., 2008; Passegue et al., 2003) of both donor *Gsk3 β ^{+flx}* and *β ^{flx/flx}* mice and evaluated by *in vivo* competitive transplantation assays (Fig. 7a). As observed with Lin⁻BM recipients (Fig. 2), only mice transplanted with *Gsk3 β ^{flx/flx}* LSK cells displayed MDS features (Fig. 7b), with significant BM and splenic hypercellularity (Fig. 7c-d). These MDS features do not evolve into leukemic transformation even when tested in permissive immune-deficient NOD/SCID recipients (Bhattacharya et al., 2006; Qing et al., 2012) (Fig. 6e), and only a more rapid death rate (Fig. 6f) and marked cytopenia (Fig. 6g) was noted compared to recipients receiving WT or

GSK-3 β ^{+/*flx*} BM. Additionally, granulocyte accumulation with dysplastic neutrophils were also detected in hypercellular BM and enlarged spleens, when compared to syngeneic recipients with normal donor hematopoiesis transplanted with Gsk3 β ^{+/*flx*} LSK cells (Fig. 6h-i). In contrast, and despite the weaker short-term regenerative capacity from non-LSKs, both syngeneic and NOD/SCID engrafted recipients transplanted non-LSK cells showed no signs of disease and displayed normal hematopoietic cellularity and differentiation (Fig. 7e and Fig. 6j) (Fig. 7f and Fig. 6k), thereby indicating that Gsk3 β ^{*flx/flx*} engrafting progenitors, that include GMP and CMP fractions within the non-LSK population (Fig. 6l and Table 2), are devoid of similar disease initiation capacity as compared to LSK subsets deleted for GSK-3 β .

[00196] To characterize the kinetics of dysplasia emergence from HSCs, time course analyses revealed an acceleration of Gsk3 β ^{*flx/flx*} LSK capacity within 4 weeks post-tamoxifen treatment (Fig. 7e-f), assessing the celerity of dysplasia (Fig. 7g). Subsequent serial transplantations demonstrated a mere 55% survival of secondary recipients transplanted with primary Gsk3 β ^{*flx/flx*} LSK-engrafted BM cells with no evidence of disease in mice transplanted with primary GSK-3 β ^{+/*flx*} LSK cells (Fig. 7h). Secondary Gsk3 β ^{*flx/flx*} (LSK) recipients retained a higher regenerative capacity compared to Gsk3 β ^{+/*flx*} (LSK) recipients, thereby underscoring MDS-IC properties (Fig. 7i). Furthermore, secondary recipients exhibited disease features similar to primary recipients that included a marked increase in granulocytes frequency (Fig. 6m). Collectively, the results indicate that the HSC compartment represents the origin of MDS-IC that is dependent on GSK-3 β deletion to induce transplantable and durable hematological disorder *in vivo*.

GSK-3 β deletion drives Wnt/Akt/mTOR signaling and induces AML when combined with metabolic changes exacerbated by the absence of GSK-3 α .

[00197] To better understand the molecular nature and collaborative interplay between GSK-3 α and GSK-3 β deficiencies responsible for MDS (β -deficient) and subsequent AML (combined $\alpha\beta$ -deficient) states, the regulatory

gene networks existing in HSCs from isolated $Gsk3\alpha^{+/+}\beta^{+/flx}$ (WT), $Gsk3\alpha^{+/+}\beta^{flx/flx}$ (β KO), $Gsk3\alpha^{-/-}\beta^{+/+}$ (α KO) and $Gsk3\alpha^{-/-}\beta^{flx/flx}$ (β + α DKO) were studied.

[00198] The LSK fractions were analyzed using whole genome expression profiling (Bowman et al., 2007) where each demonstrated a distinct transcriptome phenotype (Fig. 8a). Extended gene set enrichment analysis (GSEA) performed on LSKs revealed unique gene signatures for each GSK-3 α / β allelic deletion combination (Fig. 8b). Enrichment in cell proliferation, DNA replication and mitochondria programs were observed in GSK-3 α KO LSK cells (Fig. 9a) but had no measurable effect on the biology of HSC in transplanted hosts (Fig. 2). In contrast, enrichment in stemness and uncommitted progenitors gene expression signatures in GSK-3 β KO LSK cells was observed (Fig. 9b) and reflected a potent dysregulation of myeloid cell differentiation associated MDS (Fig. 3). In combined GSK-3 β + GSK-3 α DKO LSK cells, stemness gene signatures were conjointly enriched with metabolic and cancer gene clusters (Fig. 9c) associated with AML prognosis (Valk et al., 2004; Wong et al., 2008). Without being bound by theory, this suggests that GSK-3 α deficiency contributes additively to metabolic alterations not seen in GSK-3 β deletion alone. Although, these patterns of variation in gene signatures are shared between GSK-3 α KO and GSK-3 β KO (Fig. 9d-e) with presence of compensatory GSK-3 protein homologue levels (Fig. 9f), the biological impact of each GSK-3 α and β deletion affects molecular networks in HSCs distinctively.

[00199] The canonical Wnt pathway was the most prevalent activated in GSK-3 β KO and accordingly continued to be active in combined GSK-3 β + α DKO cells (Fig. 8c) (Lane et al., 2010; Zhao et al., 2007). Up-regulation of β -catenin/Wnt gene targets, Axin 2, Lef-1, c-Myc, CyclinD1 (Fig. 8d) (Clevers and Nusse, 2012) correlated with enhanced β -catenin activity at protein level (Fig. 8e), along with progressive nuclear stabilization of β -catenin (Fig. 8f) in HSCs deleted for GSK-3 β as well as GSK-3 β + α DKO. Since Wnt/ β -catenin

signaling manipulation alone does not alter dysplastic hematopoiesis (Kirstetter et al., 2006; Scheller et al., 2006), the transcriptome changes were further analyzed for additional cooperative pathways to Wnt signaling that may account for MDS to AML transition. Compared to WT or GSK-3 α KO these analyses revealed that mTOR signaling, previously associated to myeloid disorders (Kalaitzidis et al., 2012; Magee et al., 2012; Yilmaz et al., 2006), is enriched in GSK-3 β KO, as well as GSK-3 β + α DKO cells (Fig. 8g). This was accompanied with up-regulation of oncogenic gene targets associated to hematological diseases such as Jak2, Tcf4, Cebpa (Fig. 8h). To broadly test the relevance of these transcriptome changes, mTOR kinase activity (Recher et al., 2005; Yilmaz et al., 2006) was inhibited by rapamycin in GSK-3 β KO and GSK-3 $\alpha\beta$ DKO cells that resulted in a significant reduction, albeit mild, of total CFU-C for GSK-3 β KO cells, but had no effect on progenitor output of GSK-3 $\alpha\beta$ DKO BM cells (Fig. 8i). The aberrant distribution of erythro-myeloid colony types derived from either GSK-3 β KO or $\alpha\beta$ DKO cells, and observed in untreated conditions, was not altered by rapamycin treatment (Fig. 9g). This suggests that mTOR complex inhibition cannot solely override GSK3 BM deletion phenotype. To further characterize mTOR signaling along the PI3K/AKT pathway, additional pharmacological inhibitors were employed that perturb various stages of the pathway (illustrated in Fig. 8j). Inhibition of the PI3K/AKT/mTOR signaling effectively reduced progenitors from either GSK-3 β KO or GSK-3 $\alpha\beta$ DKO in presence of pan-PI3K inhibitor (Wortmannin), which could be exacerbated with the inhibition of the S6K (LY2584702), a critical downstream effector of mTOR signal transduction (Fig. 8j). This involvement of PI3K/AKT/mTOR axis is also corroborated by the stimulation of downstream gene targets by signal transducers such APC/MYC in GSK-3 β KO and STAT-5 in $\alpha\beta$ DKO cells (Fig. 8k), previously associated with impaired hematopoietic differentiation (Bar-Natan et al., 2012; Lane et al., 2010) and in parallel with upregulation of Akt/PI3K pathway (Fig. 8l) due to the progressive accumulation of activated

P^{S47}-AKT in both GSK-3 β KO and $\alpha\beta$ DKO (Fig. 8m). These results underscore the importance of the direct role of mTOR and PI3K/AKT arising from GSK-3 α/β deletion that contributes to MDS and AML disease.

[00200] Consistent with Wnt/AKT/mTOR pathway involvement in metabolic processes (Cheng et al., 2014; Esen et al., 2013), distinct patterns of glycolysis and oxidative phosphorylation gene signatures were observed that might explain the functional discrepancy between the onset of MDS and AML based on the contribution of *Gsk3* α/β allelic dosage (Fig. 8n). An analysis of the energetic status of Lin⁻ BM cells show a gradual decrease of *in vitro* resting metabolic activity, measured by oxygen consumption (OCR) and extra-cellular acidification rates (ECAR), by loss of GSK-3 β and $\alpha\beta$ (Fig. 8o) that might be explained by mitochondrial dysregulation as suggested by gene signatures driven by *Gsk3* α KO (Fig. 9a). GSK-3 β KO and GSK-3 β + α DKO cells show differential responses to mitochondrial stress with respect to WT, with the greatest effect when GSK-3 α is deleted in GSK-3 β deficient HSCs (Fig. 8p). Spare respiratory capacity, a measure of the mitochondria's ability to provide additional ATP-derived energy in response to stress, is significantly lost in both MDS and AML phenotypes represented by *Gsk3* β KO and *Gsk3* β + α DKO cells (Fig.8q). These results are consistent with observations reporting that inactivation of GSK-3 activity decreases the contribution of OCRs to metabolic activity (Kim et al., 2010). While little is understood about the role of mitochondrial function towards AML initiation and progression, the observations in GSK-3 β + α DKO BM cells are similar to human AML patients that present with decreased spare respiratory capacity (Sriskanthadevan et al., 2015). The findings indicate that GSK-3 α has no biological impact on hematopoiesis, but deletion of this GSK-3 isoform contributes and is necessary to an altered state of metabolism and related cell cycle for evolution of MDS to AML disease in the absence GSK3 β in HSCs.

Molecular Signature of GSK-3 β deficient HSCs predicts outcomes of human MDS

[00201] Human patients with MDS encompass a heterogeneous group with distinct hematological features evident in the multiple revisions of their WHO classification into different MDS subtypes, and in the challenges presented by efforts towards early diagnosis (Cazzola et al., 2013; Garcia-Manero, 2014; Natelson and Pyatt, 2013). Understanding and applying results from early molecular pathogenesis is now critical to improve prediction of MDS patients' prognosis (Jonas and Greenberg, 2015). Based on hematopathologic resemblance of mouse GSK-3 β -null MDS phenotype to clinically classified human patients, comparative human-mouse genome expression profiling was applied among murine *Gsk3 β* deficient and normal human (non-MDS and non-AML), along with profiles of diseased patients hematopoietic cells (Mills et al., 2009). These analyses uncovered molecular similarity of GSK-3 β -null transcriptome to human MDS and MDS/myeloproliferative neoplasm (MDS/MPN) patients' BM, which included human MDS subclasses (MDS-HR, high-risk: RAEB-1 and -2, MDS-LR, low-risk: RA and RARS and 5q deletion syndrome) as well as chronic myelomonocytic leukemia (CMML), an MDS/MPN overlap syndrome with potential risk to transform into AML (Itzykson and Solary, 2013; Parikh and Tefferi, 2012; Tiu and Sekeres, 2014) (Fig. 10a). This similarity was also supported by a dysregulated epigenomic profile, where both DNA methylation (DNMT3A) (Walter et al., 2011) and chromatin modification (EZH2) (Nikoloski et al., 2010) gene targets were disrupted upon deletion of GSK-3 β (Fig. 11a), and found to be sustained by the lack of transcriptionally repressive marks on histone H3 (H3K27me3) (Fig. 11b). These same defective gene signatures are similar to epigenomic alterations recently observed in human MDS patients (Cheng et al., 2013; Huang et al., 2011; Issa, 2013; Jeong et al., 2014; Lindsley and Ebert, 2013), suggesting molecular surrogacy of this mouse model to human MDS disease.

[00202] Differential transcriptome analysis was next performed to patient specific human MDS gene expression and their AML/survival outcomes. Despite the recent attempts to stratify AML patients by identifying specific molecular gene signatures (Eppert et al., 2011; Shivarov and Bullinger, 2014), similar approaches have yet to be extended to MDS. A specific GSK-3 β molecular signature was defined using a systematic approach based on the transcriptome comparison between WT vs. *Gsk3 β* KO LSK cells from those genes with greater than 1.5 fold-change (FC) in whole gene expression with p value ≤ 0.05 (~ 1000 genes). This subset of genes was further filtered to yield a “GSK-3 β signature” based on an FC ≥ 5 (with p-value ≤ 0.05) that would allow a reasonable number of genes, totaling 63 (Fig.10b and Table 5) to be credibly applied to transcriptome comparisons and survival outcomes in patients. Gene ontology (GO) analysis of these 63 genes comprising the GSK-3 β signature revealed an enrichment for cell adhesion molecules involved in immune response and chemotaxis (Table 6). Mapping functional protein association networks through the STRING protein interactions database revealed 33 significant interactions (p-value = 3.76×10^{-8}) within the 63 gene products including those involved in immune response (Fig. 11c and Table 3). These observations are consistent with previous clinical reports proposing that immune dysregulation acts as an early sign of pathophysiologic abnormality in MDS (Billstrom et al., 1995; Marisavljevic et al., 2006; Warlick and Miller, 2011). Importantly, GSK-3 β signature allowed for molecular discrimination among MDS, CMML and AML patient transcriptomes (Fig. 10c). Based on these findings, the clinical relevance of the GSK-3 β signature was tested using two clinically annotated profiles from adult (Sridhar et al., 2009) and pediatric MDS (p-MDS) (Bresolin et al., 2012) cohort studies with retrieved patient information and clinical annotations, together with published matching gene expression profiles. Hierarchical clustering analysis allowed for unique segregation of adult and pediatric patients based on their MDS subtype using the GSK-3 β differential transcriptome and signature (Fig.

10d and e), suggesting the GSK-3 β signature may have the potential to serve as a diagnostic biomarker.

[00203] Since p-MDS compared to adult MDS offers a unique opportunity to retrospectively study long-term clinical survival outcomes (Hasle et al., 2004; Hasle and Niemeyer, 2011), the prognostic aptitude of GSK-3 β differential transcriptome and molecular signature on pediatric MDS patients followed-up over an average period of 4.4 years (0.4-9.5 years) was further tested. Independent of the molecular results, using WHO subtyping (Hasle et al., 2003), this cohort of 32 p-MDS patients could be divided into a low risk-group that involved refractory cytopenia of childhood (RCC) subtype, and a high-risk group that comprised the other WHO subtypes of p-MDS. Expectedly, the low-risk group had a better progression free survival (PFS) ($p=0.00198$) (Fig. 11d). When applying the GSK-3 β molecular signature as a prognostic marker, irrespective of the WHO subtype for the PFS of these patients, this cohort could be divided into two additionally distinct risk groups ($p=0.00634$) and ($p=0.0224$), respectively (Fig. 11e). Emphasizing the most detrimental end-point of this disease, AML evolution, the two risk groups based on WHO subtype continued to show two different outcomes ($p=0.00246$) (Fig. 10f). Two patients were excluded due to lack of detailed follow-up data (Table 4). Applying the same stratification strategy with GSK-3 β molecular signature and differential transcriptome, results continued to show two distinctive groups with statistically significant results for AML transformation prediction ($p=0.0044$) and ($p=0.00541$), respectively (Fig. 10g and Fig. 11f). The application of the GSK-3 β signature had nearly identical ability to risk-stratify human p-MDS patients, validating the prognostic capacity of the murine model of MDS.

[00204] Although prognostic systems like the International Prognostic Scoring System (IPSS) have been validated in adult MDS patients (Greenberg et al., 1997), the IPSS has failed to show prognostic value in p-MDS (Hasle et al., 2004). Accordingly, the GSK-3 β signature was combined

with the current WHO classification of p-MDS. Using this combined modeling, the patients could be divided into three risk groups: low, intermediate and high-risk with regards to AML transformation probability ($p=0.00212$) (Fig. 10h). Patients were labeled low-risk if they had low cut-off GSK-3 β signature and low-risk WHO subtype. High-risk patients are the ones who had high cutoff GSK-3 β signature and were defined as high-risk based on WHO classification. If a patient had either a significant GSK-3 β signature or met the high-risk definition of WHO classification, he or she was labeled an intermediate-risk. This combined model was able to further risk-stratify our p-MDS cohort and provided valuable prognostic information by suggesting a third risk group, which is the intermediate-risk. These definitions for each risk group and especially the low-risk allowed an estimation of a lower probability for the low-risk group to evolve to AML using the GSK-3 β -WHO combined model (1/12) in comparison to only using the WHO classification (3/16) by around 10% (Fig. 10f and h). It is proposed that this increased precision of risk groups provides valuable information for more refined treatment planning in human MDS. An expanded use of this combined WHO and GSK-3 β signature for larger cohorts in both pediatric and adult MDS patients is warranted towards improved patient management and risk stratification.

DISCUSSION

[00205] Although it is widely accepted that genetic lesions identified in cancers play a major role in neoplasias (Alexandrov et al., 2013), evidence for mutations that initiate, drive or sustain human cancer is uncertain. For human MDS, the majority of genetic alterations have not been causally correlated to disease initiation, and seem to require dysregulated proliferative and survival pathways to perpetuate pre-leukemic disease for subsequent transition into AML (Zhang et al., 2015). It is proposed that stepwise allelic deficiency of GSK-3, a central integrative mediator of mitogenic/morphogenic signals, is sufficient for MDS disease initiation (β allele) and AML transformation (combined α and β alleles) as depicted in Fig. 6. Without being bound by

theory, the findings described herein suggest the pathways tightly regulated by GSK-3 β are critical for evolution of MDS from healthy HSCs that allow a state for additional insults, such as loss of GSK-3 α , to lead to AML. Upon GSK3 deficiency, the study implicates an inter-relationship with the mTOR/AKT/PI3K pathways during MDS/AML progression not reported to date. Effects of GSK-3 β deficiency to induce MDS was unique to HSCs and not progenitors and allowed emergence of MDS-IC capacity. These findings are consistent with recent reports indicating that a specific pre-leukemic stem cell subset of human MDS from low to intermediate risk patients originates from the HSC compartment (Pang et al., 2013; Schanz et al., 2012; Tehranchi et al., 2010; Woll et al., 2014).

[00206] Despite sharing similar sequences, it is found that the role of GSK-3 β is molecularly different from *Gsk3 α* in hematopoiesis, correlating with distinct HSC transcriptomes. Notably, increase of GSK-3 α protein levels in GSK-3 β -deficient BM cells did not compensate or “rescue” the effects responsible of the observed MDS state. This is in contrast to previous reports showing functional redundancy between GSK-3 α and β alleles in embryonic and hematopoietic development (Doble et al., 2007; Huang et al., 2009; Itoh et al., 2012). This discrepancy is likely due to fundamental differences between fetal and adult somatic HSCs revealed in this study, and is consistent with the idea that properties of HSCs change during ontogeny (Babovic and Eaves, 2014). Recent evidence suggests that this switch is orchestrated by a concordant change in the level of expression in HSCs of *Pten* and *Cebpa* (Ye et al., 2013; Zhang et al., 2006). Both *Pten* and *Cebpa*, which play a crucial role in regulating the Akt/mTOR pathway, survival and quiescence of HSCs (Kharas et al., 2010), were dysregulated in GSK3 β -deficient adult LSK fractions that develop MDS. Consequently, GSK3 β deletion in the fetal LSK fraction (Huang et al., 2009) might not affect fetal HSCs due to the lack of involvement of Akt/mTOR/Wnt signaling axis. More importantly, it is uniquely revealed herein that the GSK3 β -null myelodysplastic

state acquires stemness properties and dysregulated myeloid differentiation programs that are exacerbated upon additional Gsk3 α deletion inducing a metabolic alteration specific to AML onset driven by GSK3 β -null HSCs.

[00207] Although previous mouse models have been found to be similar to MDS, transition of mouse MDS to AML, or any relationship to human MDS to AML evolution, has not been demonstrated. Using the GSK3 β -null pre-leukemic gene signature, clinical outcomes of p-MDS patients were predicted. Without being bound by theory, it is suggested that a “combined GSK-3-WHO” risk stratification model will improve the probability to detect AML development (~10%) by discriminating the p-MDS cohort into three risk categories: Low, intermediate and high-risk. Although this analysis was performed on a small (n=32) but rare number of patients where survival and molecular gene expression profiles have been applied, it is believed that implementing this novel disease risk stratification approach will be crucial for better risk stratification of larger MDS patient cohorts. This is the first demonstration that a transcriptional phenotype of such pre-neoplastic state can allow risk stratification of MDS patients and may aid in early intervention by targeting molecular networks involved in the earliest stages of AML onset.

Example 2 *In vitro* deletion of the GSK-3 β KO and GSK-3 α + β DKO in highly enriched LSK BM cell lines allow the study of the molecular features for disease onset

[00208] In order to understand the transition steps driving MDS and/or AML onset, an *in vitro* deletion system was developed (Figure 13). Bone marrow cells were extracted from WT, Gsk3 β (flx/flx) and Gsk3 α (-/-) + β (flx/flx) mice, homogenized and Lineage depleted. Then, freshly isolated 1.5×10^6 cells were incubated in 6-well-plate culture dishes with Stem Span medium containing rm SCF (100ng/mL), rm IL-3 (10ng/mL) and rh IL-6 (10ng/mL) (StemCell Technologies) for 6 days. Fresh medium was added every 2 days. At day 7, cells were treated with 4-OH-Tamoxifen

(SigmaAldrich) at 5 μ M for 24 hours. Media were replaced with fresh one and incubated cells for molecular analysis for another 5-7 days post-treatment (A).

[00209] At Day 7, the percentage of LSK cells is robustly increased, allowing the study and targeting of this purified population along the progression to MDS or AML (B). As observed by PCR or western blot analysis, effective silencing took place after *in vitro* Tamoxifen treatment (C and D). Interestingly, Cyclin D1 expression (*Ccnd1*) is increased in Gsk3 β KO and Gsk3 α + β DKO cells which resembles the *in vivo* data. Collectively, the *in vitro* model presented shows a suitable way to study the HSC transition from healthy state to MDS to AML.

[00210] Different embodiments of the disclosure have been shown by the above examples. Those skilled in the art could develop alternatives to the methods mentioned above that are within the scope of the disclosure and defined claims.

TABLE 1**A**

Primer Name	Primer Sequence	SEQ ID No.
Exon2_GSK3b_Fwd	5'- GTGTCTGTATAACTGACTTCCTGTGGC -3'	SEQ ID No.1
Neo_GSK3b_Fwd	5'- GGGGCAACCTTAATTTTCATT -3'	SEQ ID No.2
Intron2_GSK3b_Rev	5'- TCTGGGCTATAGCTATCTAGTAACG -3'	SEQ ID No.3
Intron1_2_GSK3a_Fwd	5'-CCCCACCAAGTGATTTCACTGCTA-3'	SEQ ID No.4
Intron2_3_GSK3a_Rev	5'-AACATGAAATTCCGGGCTCCAACCTCTAT-3'	SEQ ID No.5
R26R-univ_Fwd	5'- AAAGTCGCTCTGAGTTGTTAT -3'	SEQ ID No.6
R26R-wt_Rev	5'- GGAGCGGGAGAAATGGATATG -3'	SEQ ID No.7
CreER Rev	5'- CCTGATCCTGGCAATTTTCG -3'	SEQ ID No.8

B

Gene Name	Forward Primer Sequence	SEQ ID No.
<i>Axin2</i>	5'- TGACTCTCCTTCCAGATCCCA -3'	SEQ ID No.9
<i>Ccndi</i>	5'-TCCTCTCCAAATGCCAGAG-3'	SEQ ID No.11
<i>Cebpa</i>	5'-CAAGAACAGCAACGAGTACCG-3'	SEQ ID No.13
<i>Jak2</i>	5'-TTGTGGTATTACGCCTGTGTATC-3'	SEQ ID No.15
<i>Lef1</i>	5'-TGTTTATCCCATCACGGGTGG-3'	SEQ ID No.17
<i>Mtor</i>	5'-ACCGGCACACATTTGAAGAAG-3'	SEQ ID No.19
<i>Myc</i>	5'-TCCTGTACCTCGTCCGATTTC-3'	SEQ ID No.21
<i>Tcf4</i>	5'-TCATATCCCACAGTCCAGCA-3'	SEQ ID No.23
<i>β-actin</i>	5'-ATGAACGACGTAGCCATTGTG-3'	SEQ ID No.25

Gene Name	Reverse Primer Sequence	SEQ ID No.
<i>Axin2</i>	5'- TGCCCACTAGGCTGACA -3'	SEQ ID No.10
<i>Ccndi</i>	5'-GCAGGAGAGGAAGTTGTTGG-3'	SEQ ID No.12
<i>Cebpa</i>	5'-GTCACTGGTCAACTCCAGCAC-3'	SEQ ID No.14
<i>Jak2</i>	5'-ATGCCTGGTTGACTCGTCTAT-3'	SEQ ID No.16
<i>Lef1</i>	5'-CATGGAAGTGTGCGCTGACAG-3'	SEQ ID No.18
<i>Mtor</i>	5'-CTCGTTGAGGATCAGCAAGG-3'	SEQ ID No.20
<i>Myc</i>	5'-GGTTTGCCCTCTTCTCCACAG-3'	SEQ ID No.22
<i>Tcf4</i>	5'-TTTGTGCCACTACGATGGAA-3'	SEQ ID No.24
<i>β-actin</i>	5'-GGAATCCTTCTGACCCATGC-3'	SEQ ID No.26

TABLE 2. Cell composition in progenitors within NON-LSK population

Genotype of injected cells	Total number of injected cells	Freq. of progenitors (%)	Estimated total number of progenitors	Freq. of CMP (%)	Estimated total number of CMP	Freq. of MEP (%)	Estimated total number of MEP
Non-LSK Gsk3 β ^{+/f/y}	10,000	6.95 $\pm$ 1.24	6950 $\pm$ 1240	20.3 $\pm$ 0.8	1409 $\pm$ 12	13.6 $\pm$ 0.6	945 $\pm$ 6
Non-LSK Gsk3 β ^{0/y/f/y}	10,000	8.66 $\pm$ 1.45	8660 $\pm$ 1450	30.5 $\pm$ 1.05	2641 $\pm$ 27	9.7 $\pm$ 0.7	840 $\pm$ 7

The proportion of cells for each subset are estimated and expressed as the means $\pm$ SD.

TABLE 3. STRING annotations process of GSK3 β signature

P-value	3.76E-08
Interactions observed	33
Interactions expected	10.7
Nb total of proteins	62

TABLE 4**Disease Characteristics and Clinical Outcomes of Pediatric MDS Cases**

Patient No.	WHO Subtype	Risk group	GSK3 β signature	GSK3 β full transcriptome	AML evolution	Survival
P1	RCC	Low	LOW	LOW	NO	Alive
P2	RCC	Low	LOW	LOW	NO	Alive
P3	RCC	Low	LOW	LOW	NO	Alive
P4	RCC*	Low	LOW	LOW	YES	Dead
P5	RAEB	High	LOW	LOW	NO	Alive
P6	RAEB-t	High	HIGH	HIGH	YES	Alive
P7	RAEB	High	HIGH	HIGH	YES	Alive
P8	RAEB*	High	LOW	HIGH	YES	Alive
P9	RCC	Low	HIGH	HIGH	NO	Alive
P10	RAEB-t	High	HIGH	LOW	NO	Alive
P11	RAEB	High	LOW	HIGH	YES	Alive
P12	RAEB	High	HIGH	HIGH	YES	Dead
P13	RAEB	High	HIGH	HIGH	YES	Alive
P14	RAEB	High	LOW	HIGH	NO	Dead
P15	RCC	Low	LOW	LOW	NO	Alive
P16	RAEB-t	High	HIGH	HIGH	YES	Dead
P17	RCC	Low	LOW	LOW	NO	Alive
P18	RCC	Low	LOW	LOW	NO	Alive
P19	RAEB*	High	LOW	LOW	Unknown	Dead
P20	RCC	N/A	LOW	LOW	NO	Dead
P21	RCC	Low	HIGH	HIGH	NO	Alive
P22	RCC	Low	LOW	HIGH	NO	Alive
P23	RAEB-t	High	HIGH	HIGH	NO	Alive
P24	RCC	Low	LOW	LOW	NO	Alive
P25	RAEB-t	High	LOW	LOW	YES	Alive
P26	RAEB-t	High	HIGH	HIGH	YES	Alive
P27	RAEB-t	High	HIGH	HIGH	YES	Alive
P28	RAEB	High	LOW	LOW	YES [¶]	Alive
P29	RCC	Low	HIGH	LOW	YES	Alive
P30	RCC	Low	LOW	LOW	NO	Alive
P31	RCC	Low	HIGH	HIGH	YES	Alive
P32	RCC	Low	LOW	LOW	NO	Alive

* Therapy related MDS

[¶] Unknown time of progression

RCC: Refractory cytopenia of childhood

RAEB: Refractory anemia with excess blasts

RAEB-t: Refractory anemia with excess blasts in transformation

TABLE 5 - GSK3 β signature (63 entities)

	Gene Symbol	Column ID	Entrez Gene	Gene Title	RefSeq Transcript ID	p-value (GSK3del vs. wt)	Fold-Change (GSK3del vs. wt)	Fold-Change (GSK3del vs. wt) (Description)	Human Orthologue Accession Number
1	Cd177	1424509_at	68891	CD177 antigen matrix metalloproteinase 9	NM_026862	0.0125607	-69.8839	GSK3del down vs wt	NM_020406
2	Mmp9	1416298_at	17395		NM_013599	0.00250717	-43.7735	GSK3del down vs wt	NM_004994
3	Ltf	1450009_at	17002	lactotransferrin	NM_008522	0.0240287	-42.9909	GSK3del down vs wt	NM_002343 /// NM_001199149 /// NM_001321121 /// NM_001321122
4	Fpr2	1422953_at	14289	formyl peptide receptor 2	NM_008039	0.0259369	-26.1706	GSK3del down vs wt	NM_001005738 /// NM_001462
5	Mmp8	1449366_at	17394	matrix metalloproteinase 8	NM_008611	0.0405542	-24.1443	GSK3del down vs wt	NM_001304441 /// NM_001304442 /// NM_002424
6	Lrg1	1417290_at	76905	leucine-rich alpha-2-glycoprotein 1	NM_029796	0.00895133	-23.672	GSK3del down vs wt	NM_052972
7	Slco4c1	1437870_at	227394	solute carrier organic anion transporter family, member 4C1	NM_172658	0.0339975	-13.9839	GSK3del down vs wt	NM_180991
8	Sgms2	1428663_at	74442	sphingomyelin synthase 2	NM_028943	0.0102935	-13.7437	GSK3del down vs wt	NM_001136257
9	Crispld2	1434758_at	78892	cysteine-rich secretory protein LCCL domain containing 2	NM_030209	0.0121872	-12.975	GSK3del down vs wt	NM_031476
10	Il1r2	1419532_at	16178	interleukin 1 receptor, type II	NM_010555	0.000629529	-12.2028	GSK3del down vs wt	NM_001261419 /// NM_004633 /// NM_173343
11	Olfm4	1437060_at	380924	olfactomedin 4	NM_001030294	0.00124201	-11.2403	GSK3del down vs wt	NM_006418

	Gene Symbol	Column ID	Entrez Gene	Gene Title	RefSeq Transcript ID	p-value (GSK3del vs. wt)	Fold-Change (GSK3del vs. wt)	Fold-Change (GSK3del vs. wt) (Description)	Human Orthologue Accession Number
12	Pglyrp1	1449184_at	21946	peptidoglycan recognition protein 1	NM_009402	0.0165793	-10.5091	GSK3del down vs wt	NM_005091
13	Prom1	1419700_a_at	19126	prominin 1	NM_001163577 /// NM_001163578 /// NM_001163581 /// NM_001163582 /// NM_001163583	0.0450329	-9.77284	GSK3del down vs wt	NM_001145847 /// NM_001145848 /// NM_001145849 /// NM_001145850 /// NM_001145851
14	Abca13	1446715_at	268379	ATP-binding cassette, sub-family A (ABC1), member 13	NM_178259	0.00293415	-9.67912	GSK3del down vs wt	NM_152701
15	Cd300lf	1427994_at	246746	CD300 antigen like family member F	NM_001169153 /// NM_145634	0.0450968	-9.26463	GSK3del down vs wt	NM_001289082 NM_001289083 /// NM_001289084 /// NM_001289085 /// NM_001289086
16	Ptprc	1440165_at	19264	Protein tyrosine phosphatase, receptor type, C	NM_001111316 /// NM_001268286 /// NM_011210	0.0312012	-9.24411	GSK3del down vs wt	NM_001267798 /// NM_002838 /// NM_080921 /// NM_080922
17	Asprv1	1452732_at	67855	aspartic peptidase, retroviral-like 1	NM_026414	0.025127	-7.69176	GSK3del down vs wt	NM_181718 /// NM_020437 /// NM_018136 /// NM_001206846
18	Ccr1	1419609_at	12768	chemokine (C-C motif) receptor 1	NM_009912	0.0214601	-7.42381	GSK3del down vs wt	NM_001295
19	Trem1	1447284_at	58217	triggering receptor expressed on myeloid cells 1	NM_021406	0.0265773	-7.1401	GSK3del down vs wt	NM_001242589 /// NM_001242590 /// NM_018643
20	S100a6	1421375_a_at	20200	S100 calcium binding protein A6 (calcyclin)	NM_011313	0.0296629	-6.57213	GSK3del down vs wt	NM_014624

	Gene Symbol	Column ID	Entrez Gene	Gene Title	RefSeq Transcript ID	p-value (GSK3del vs. wt)	Fold-Change (GSK3del vs. wt)	Fold-Change (GSK3del vs. wt) (Description)	Human Orthologue Accession Number
21	Cybb	1436779_at	13058	cytochrome b-245, beta polypeptide	NM_007807	0.0203902	-6.5296	GSK3del down vs wt	NM_000397
22	Anxa1	1448213_at	16952	annexin A1	NM_010730	0.00276651	-6.03677	GSK3del down vs wt	NM_000700
23	Rhou	1449027_at	69581	ras homolog gene family, member U	NM_133955	0.0455354	-5.97568	GSK3del down vs wt	NM_021205
24	Cdh1	1448261_at	12550	cadherin 1	NM_009864	0.00899184	-5.86039	GSK3del down vs wt	NM_004360 /// NM_001317184 /// NM_001317185 /// NM_001317186
25	Gsr	1421817_at	14782	glutathione reductase	NM_010344	0.0101014	-5.84119	GSK3del down vs wt	NM_001195104 /// NM_000637 /// NM_001195102 /// NM_001195103
26	Lyst	1434674_at	17101	lysosomal trafficking regulator	NM_010748	0.0310648	-5.82347	GSK3del down vs wt	NM_000081
27	Scrg1	1420764_at	20284	scrapie responsive gene 1	NM_009136	0.0380157	-5.81111	GSK3del down vs wt	NM_001329597 /// NM_007281
28	Pygl	1417741_at	110095	liver glycogen phosphorylase	NM_133198	0.0352145	-5.7345	GSK3del down vs wt	NM_002863 /// NM_001163940
29	Hsd11b1	1449038_at	15483	hydroxysteroid 11-beta dehydrogenase 1	NM_001044751 /// NM_008288	0.0158629	-5.7238	GSK3del down vs wt	NM_181755 /// NM_001206741 /// NM_005525
30	Ch131	1451537_at	12654	chitinase 3-like 1	NM_007695	0.0273668	-5.70256	GSK3del down vs wt	NM_001276
31	Stk17b	1430165_at	98267	serine/threonine kinase 17b (apoptosis-inducing)	NM_133810	0.044122	-5.64855	GSK3del down vs wt	NM_004226
32	Itgb2	1450678_at	16414	integrin beta 2 family with sequence similarity 101, member B	NM_008404	0.0223336	-5.51468	GSK3del down vs wt	NM_000211 /// NM_001127491 /// NM_001303238
33	Fam101b	1456603_at	76566		NM_029658	0.00468747	-5.3396	GSK3del down vs wt	NM_182705

	Gene Symbol	Column ID	Entrez Gene	Gene Title	RefSeq Transcript ID	p-value (GSK3del vs. wt)	Fold-Change (GSK3del vs. wt)	Fold-Change (GSK3del vs. wt) (Description)	Human Orthologue Accession Number
34	Bmx	1422084_at	12169	BMX non-receptor tyrosine kinase	NM_009759	0.0166326	-5.25095	GSK3del down vs wt	NM_001721 /// NM_203281 /// NM_001320866
35	Lmna	1421654_a_at	16905	lamin A	NM_001002011 /// NM_001111102 /// NM_019390	0.0256853	5.13549	GSK3del up vs wt	NM_001257374 /// NM_001282624 /// NM_001282625 /// NM_001282626 /// NM_005572
36	Nr4a2	1447863_s_at	18227	nuclear receptor subfamily 4, group A, member 2	NM_001139509 /// NM_013613	0.0490427	5.15859	GSK3del up vs wt	NM_006186 /// NM 173171 /// NM 173172 /// NM 173173 /// NM 001139509 /// NM 013613
37	Tmem158	1428074_at	72309	transmembrane protein 158	NM_001002267	0.0208402	5.33134	GSK3del up vs wt	NM_015444
38	Slc22a3	1420444_at	20519	solute carrier family 22 (organic cation transporter), member 3	NM_011395	0.0128353	5.39371	GSK3del up vs wt	NM_021977
39	Bcl11a	1419406_a_at	14025	B cell CLL/lymphoma 11A (zinc finger protein)	NM_001159289 /// NM_001159290 /// NM_001242934 /// NM_016707	0.0135952	5.39611	GSK3del up vs wt	NM_018014 /// NM_022893 /// NM_138553 /// NM_138559
40	Pla2g16	1445597_s_at	225845	phospholipase A2, group XVI	NM_139269	0.00803401	5.51137	GSK3del up vs wt	NM_001128203 /// NM_007069
41	Eltf1	1418058_at	170757	EGF, latrophilin seven transmembrane domain containing 1	NM_133222	0.0239181	5.60039	GSK3del up vs wt	NM_022159

	Gene Symbol	Column ID	Entrez Gene	Gene Title	RefSeq Transcript ID	p-value(GSK3del vs. wt)	Fold-Change(GSK3del vs. wt)	Fold-Change(GSK3del vs. wt) (Description)	Human Orthologue Accession Number
42	Mrpl52	1415761_at	68836	mitochondrial ribosomal protein L52	NM_026851	0.0153308	5.77285	GSK3del up vs wt	NM_178336 /// NM_180981 /// NM_180982 /// NM_181304 /// NM_181305 /// NM_181306 /// NM_181307
43	Uba7	1426971_at	74153	ubiquitin-like modifier activating enzyme 7	NM_023738	0.0372808	5.83357	GSK3del up vs wt	NM_003335
44	Pdgfrb	1436970_a_at	18596	platelet derived growth factor receptor, beta polypeptide	NM_001146268 /// NM_008809	0.016662	6.11288	GSK3del up vs wt	NM_002609
45	Gpsm1	1436917_s_at	67839	G-protein signalling modulator 1 (AGS3-like, C. elegans)	NM_001199146 /// NM_001199147 /// NM_153410	0.0248096	6.22353	GSK3del up vs wt	NM_001145638 /// NM_001145639 /// NM_001200003 /// NM_015597
46	Alad	1424877_a_at	17025	aminolevulinatase, delta-, dehydratase	NM_008525	0.027027	6.47613	GSK3del up vs wt	NM_000031 /// NM_001003945 /// NM_001317745
47	Grasp	1441894_s_at	56149	GRP1 (general receptor for phosphoinositides 1)-associated scaffold protein	NM_019518	0.0148299	6.56348	GSK3del up vs wt	NM_001271856 /// NM_181711
48	Tfr2	1459994_x_at	50765	transferrin receptor 2	NM_015799	0.0402812	6.72018	GSK3del up vs wt	NM_001206855 /// NM_003227
49	Jakmip1	1441317_x_at	76071	janus kinase and microtubule interacting protein 1	NM_178394	0.0372636	6.9886	GSK3del up vs wt	NM_144720 /// NM_001099433 /// NM_001306134 /// NM_001306133

	Gene Symbol	Column ID	Entrez Gene	Gene Title	RefSeq Transcript ID	p-value(GSK3del vs. wt)	Fold-Change(GSK3del vs. wt)	Fold-Change(GSK3del vs. wt) (Description)	Human Orthologue Accession Number
50	Cisd1	1443822_s_at	52637	CDGSH iron sulfur domain 1 protein phosphatase 1, regulatory (inhibitor) subunit 13B	NM_134007	0.0367244	7.11296	GSK3del up vs wt	NM_018464
51	Ppp1r13b	1434895_s_at	21981	FERM domain containing 4B	NM_011625	0.0237935	7.62797	GSK3del up vs wt	NM_015316
52	Frmd4b	1438169_a_at	232288	immunoglobulin joining chain	NM_145148	0.0100937	7.99232	GSK3del up vs wt	NM_015123
53	Igj	1424305_at	16069	coiled-coil-helix-coiled-coil-helix domain containing 6	NM_152839	0.0475385	8.28519	GSK3del up vs wt	NM_144646
54	Chchd6	1438659_x_at	66098	spinster homolog 2 (Drosophila)	NM_001167736 /// NM_025351	0.00947016	8.32688	GSK3del up vs wt	NM_001320610 /// NM_032343
55	Spns2	1451601_a_at	216892	zinc finger protein, multitype 1	NM_153060	0.0455561	8.54671	GSK3del up vs wt	NM_001124758
56	Zfpm1	1451046_at	22761	tetraspanin 33	NM_009569	0.028258	8.99192	GSK3del up vs wt	NM_153813
57	Tspan33	1455618_x_at	232670	coagulation factor II (thrombin) receptor	NM_146173	0.014267	9.48573	GSK3del up vs wt	NM_178562
58	F2r	1437308_s_at	14062	POU domain, class 2, associating factor 1	NM_010169	0.0436825	9.75891	GSK3del up vs wt	NM_001992 /// NM_001311313
59	Pou2af1	1416957_at	18985	aldehyde dehydrogenase family 1, subfamily A1	NM_011136	0.0036546	10.5672	GSK3del up vs wt	NM_006235
60	Aldh1a1	1416468_at	11668	TSC22 domain family, member 1	NM_013467	0.0316714	12.0742	GSK3del up vs wt	NM_000689
61	Tsc22d1	1454971_x_at	21807		NM_001177751 /// NM_009366 /// NM_207652	0.0248376	12.4207	GSK3del up vs wt	NM_001243797 /// NM_001243798 /// NM_001243799 /// NM_006022 /// NM_183422

	Gene Symbol	Column ID	Entrez Gene	Gene Title	RefSeq Transcript ID	p-value (GSK3del vs. wt)	Fold-Change (GSK3del vs. wt)	Fold-Change (GSK3del vs. wt) (Description)	Human Orthologue Accession Number
62	Ighm	1427329_a_at	16019	immunoglobulin heavy constant mu	---	0.0117056	14.2817	GSK3del up vs wt	---
63	Ctnn	1433908_a_at	13043	cortactin	NM_001252572 /// NM_007803	0.0370396	17.9394	GSK3del up vs wt	NM_001184740 /// NM_005231 /// NM_138565

TABLE 6

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0006955-immune response	12	19.04761905	6.00E-05	IL1R2, PTPRC, CYBB, POU2AF1, LYST, CCR1, IGJ, PGLYRP1, LTF, CD300LF, TREM1, IGHM	54	690	13528	4.3568438	0.042651052	0.042651052	0.090941023
GOTERM_BP_FAT	GO:0006952-defense response	10	15.87301587	5.77E-04	ANXA1, LTF, ITGB2, FPR2, F2R	54	615	13528	4.073471846	0.342171175	0.188933526	0.87039784
GOTERM_BP_FAT	GO:0006935-chemotaxis	5	7.936507937	0.003517813	LYST, CCR1, PDGFRB, ITGB2, FPR2	54	160	13528	7.828703704	0.92257419	0.573785142	5.200279553
GOTERM_BP_FAT	GO:0042330-taxi	5	7.936507937	0.003517813	LYST, CCR1, PDGFRB, ITGB2, FPR2	54	160	13528	7.828703704	0.92257419	0.573785142	5.200279553
GOTERM_BP_FAT	GO:0007626-locomotory behavior	6	9.523809524	0.004266118	LYST, CCR1, NR4A2, PDGFRB, ITGB2, FPR2	54	274	13528	5.485806975	0.955122876	0.539736799	6.273387126
GOTERM_BP_FAT	GO:0000272-polysaccharide catabolic process	3	4.761904762	0.005333467	PYGL, PGLYRP1, CHI3L1	54	28	13528	26.84126984	0.979399911	0.539983065	7.784400353
GOTERM_BP_FAT	GO:0006954-inflammatory response	6	9.523809524	0.006674176	CYBB, CCR1, ANXA1, ITGB2, FPR2, F2R	54	325	13528	4.624957265	0.998208896	0.651512828	12.36799084
GOTERM_BP_FAT	GO:0007610-behavior	7	11.11111111	0.009725882	LYST, CCR1, PGLYRP1, NR4A2, PDGFRB, ITGB2, FPR2	54	469	13528	3.739082366	0.999171217	0.637108098	13.76634398
GOTERM_BP_FAT	GO:0060326-cell chemotaxis	3	4.761904762	0.010172102	LYST, PDGFRB, ITGB2	54	39	13528	19.27065527	0.999402504	0.604595245	14.35331852
GOTERM_BP_FAT	GO:0030005-cellular di-, tri-valent inorganic cation homeostasis	5	7.936507937	0.011896539	PTPRC, CCR1, TFR2, LTF, F2R	54	227	13528	5.518029042	0.999831517	0.619172241	16.58681483
GOTERM_BP_FAT	GO:0042981-regulation of apoptosis	9	14.28571429	0.012364115	PTPRC, LYST, MMP9, ANXA1, NR4A2, STK17B, CDH1, PPP1R13B, F2R	54	804	13528	2.804311774	0.999880513	0.594741508	17.18297769
GOTERM_BP_FAT	GO:0043067-regulation of programmed cell death	9	14.28571429	0.013069925	PTPRC, LYST, MMP9, ANXA1, NR4A2, STK17B, CDH1, PPP1R13B, F2R	54	812	13528	2.776683087	0.999928893	0.580336379	18.07535567
GOTERM_BP_FAT	GO:0010941-regulation of cell death	9	14.28571429	0.013342163	PTPRC, LYST, MMP9, ANXA1, NR4A2, STK17B, CDH1, PPP1R13B, F2R	54	815	13528	2.766462168	0.999941799	0.55631221	18.41714885
GOTERM_BP_FAT	GO:005066-di-, tri-valent inorganic cation homeostasis	5	7.936507937	0.01414534	PTPRC, CCR1, TFR2, LTF, F2R	54	239	13528	5.240973191	0.999967777	0.548691096	19.41778739
GOTERM_BP_FAT	GO:0030003-cellular cation homeostasis	5	7.936507937	0.017315372	PTPRC, CCR1, TFR2, LTF, F2R	54	254	13528	4.9314669	0.999996891	0.595777823	23.2563414
GOTERM_BP_FAT	GO:0050508-cation homeostasis	5	7.936507937	0.025478696	PTPRC, CCR1, TFR2, LTF, F2R	54	286	13528	4.37969438	0.999999993	0.713251768	32.36968739
GOTERM_BP_FAT	GO:0043062-extracellular structure organization	4	6.349206349	0.025827702	CRISPLD2, MMP9, CDH1, F2R	54	163	13528	6.147693706	0.999999995	0.696383512	32.84036223
GOTERM_BP_FAT	GO:0042592-homeostatic process	8	12.6984127	0.026355491	GSR, PTPRC, PYGL, CCR1, TFR2, PDGFRB, LTF, F2R	54	751	13528	2.668639345	0.999999996	0.680383157	33.2859475
GOTERM_BP_FAT	GO:0050730-regulation of peptidyl-tyrosine phosphorylation	3	4.761904762	0.029098976	PTPRC, PDGFRB, ITGB2	54	68	13528	11.05228758	1	0.696105139	36.07857791
GOTERM_BP_FAT	GO:0015674-di-, tri-valent inorganic cation transport	4	6.349206349	0.031551828	PTPRC, TFR2, LTF, F2R	54	176	13528	5.693602694	1	0.706254448	38.48254395
GOTERM_BP_FAT	GO:0019725-cellular homeostasis	6	9.523809524	0.035158709	GSR, PTPRC, CCR1, TFR2, LTF, F2R	54	466	13528	3.225560324	1	0.72726022	41.86456809
GOTERM_BP_FAT	GO:0006928-cell motion	6	9.523809524	0.037703908	LYST, ANXA1, NR4A2, PDGFRB, ITGB2, FPR2	54	475	13528	3.164444444	1	0.735175182	44.14570211
GOTERM_BP_FAT	GO:0006959-humoral immune response	3	4.761904762	0.038301963	POU2AF1, LTF, TREM1	54	79	13528	9.513361463	1	0.72440072	44.66943964
GOTERM_BP_FAT	GO:0009617-response to bacterium	4	6.349206349	0.039796257	LYST, PGLYRP1, LTF, F2R	54	193	13528	5.192093648	1	0.722478449	45.95806547
GOTERM_BP_FAT	GO:0030335-positive regulation of cell migration	3	4.761904762	0.047498671	MMP9, PDGFRB, F2R	54	89	13528	8.444444444	1	0.770552088	52.16735777
GOTERM_BP_FAT	GO:0048878-chemical homeostasis	6	9.523809524	0.049350087	PTPRC, PYGL, CCR1, TFR2, LTF, F2R	54	512	13528	2.935763889	1	0.770004894	53.55709374

Category	Term	Count	%	pValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_BP_FAT	GO:0051209--release of sequestered calcium ion into cytosol	2	3.174603175	0.053499143	PTPRC, F2R	54	14	13528	35.78835979	1	0.784610182	56.53572873
GOTERM_BP_FAT	GO:0051283--negative regulation of sequestering of calcium ion	2	3.174603175	0.053499143	PTPRC, F2R	54	14	13528	35.78835979	1	0.784610182	56.53572873
GOTERM_BP_FAT	GO:0051282--regulation of sequestering of calcium ion	2	3.174603175	0.053499143	PTPRC, F2R	54	14	13528	35.78835979	1	0.784610182	56.53572873
GOTERM_BP_FAT	GO:0009811--response to wounding	6	9.523809524	0.055712968	CYBB, CCR1, ANXA1, ITGB2, FPR2, F2R	54	530	13528	2.8360587	1	0.785920587	58.05108203
GOTERM_BP_FAT	GO:0040017--positive regulation of locomotion	3	4.761904762	0.056391187	MMP9, PDGFRB, F2R	54	98	13528	7.66893424	1	0.777979909	58.50535173
GOTERM_BP_FAT	GO:0051272--positive regulation of cell motion	3	4.761904762	0.056391187	MMP9, PDGFRB, F2R	54	98	13528	7.66893424	1	0.777979909	58.50535173
GOTERM_BP_FAT	GO:0006873--cellular ion homeostasis	5	7.936507937	0.058378201	PTPRC, CCR1, TFR2, LTF, F2R	54	374	13528	3.349178055	1	0.778174333	59.80994224
GOTERM_BP_FAT	GO:0055082--cellular chemical homeostasis	5	7.936507937	0.061189099	PTPRC, CCR1, TFR2, LTF, F2R	54	380	13528	3.296296296	1	0.783034541	61.59014382
GOTERM_BP_FAT	GO:0016052--carbohydrate catabolic process	3	4.761904762	0.067978243	PYGL, PGLYRP1, CHI3L1	54	109	13528	6.895005097	1	0.807700642	65.59073067
GOTERM_BP_FAT	GO:0007204--elevation of cytosolic calcium ion concentration	3	4.761904762	0.069068123	PTPRC, CCR1, F2R	54	110	13528	6.832323232	1	0.802837238	66.19547462
GOTERM_BP_FAT	GO:0005976--polysaccharide metabolic process	3	4.761904762	0.070163822	PYGL, PGLYRP1, CHI3L1	54	111	13528	6.770770771	1	0.79819047	66.793429
GOTERM_BP_FAT	GO:0042742--defense response to bacterium	3	4.761904762	0.071265279	LYST, PGLYRP1, LTF	54	112	13528	6.71031746	1	0.793750468	67.38455734
GOTERM_BP_FAT	GO:0060402--calcium ion transport into cytosol	2	3.174603175	0.07191705	PTPRC, F2R	54	19	13528	26.37037037	1	0.787353859	67.72970456
GOTERM_BP_FAT	GO:0060401--cytosolic calcium ion transport	2	3.174603175	0.075558207	PTPRC, F2R	54	20	13528	25.05185185	1	0.794928795	69.59596008
GOTERM_BP_FAT	GO:0051607--defense response to virus	2	3.174603175	0.075558207	PTPRC, LYST	54	20	13528	25.05185185	1	0.794928795	69.59596008
GOTERM_BP_FAT	GO:0030574--collagen catabolic process	2	3.174603175	0.075558207	MMP9, MMP8	54	20	13528	25.05185185	1	0.794928795	69.59596008
GOTERM_BP_FAT	GO:0050801--ion homeostasis	5	7.936507937	0.075776692	PTPRC, CCR1, TFR2, LTF, F2R	54	409	13528	3.062573576	1	0.78694702	69.70467269
GOTERM_BP_FAT	GO:0051480--cytosolic calcium ion homeostasis	3	4.761904762	0.077991613	PTPRC, CCR1, F2R	54	118	13528	6.369114878	1	0.788041691	70.78645494
GOTERM_BP_FAT	GO:0006811--ion transport	7	11.11111111	0.078482295	PTPRC, CYBB, SLC04C1, TFR2, SLC22A3, LTF, F2R	54	768	13528	2.283371914	1	0.781615196	71.0211744
GOTERM_BP_FAT	GO:0006026--aminoglycan catabolic process	2	3.174603175	0.079185348	PGLYRP1, CHI3L1	54	21	13528	23.85890653	1	0.771626918	71.3544129
GOTERM_BP_FAT	GO:0043065--positive regulation of apoptosis	5	7.936507937	0.08735267	PTPRC, LYST, MMP9, STK17B, PPP1R13B	54	430	13528	2.913006029	1	0.801814254	74.97221091
GOTERM_BP_FAT	GO:0043068--positive regulation of programmed cell death	5	7.936507937	0.089073961	PTPRC, LYST, MMP9, STK17B, PPP1R13B	54	433	13528	2.892823539	1	0.800640593	75.67807319
GOTERM_BP_FAT	GO:0010942--positive regulation of cell death	5	7.936507937	0.090230729	PTPRC, LYST, MMP9, STK17B, PPP1R13B	54	435	13528	2.879523201	1	0.797415171	76.141943
GOTERM_BP_FAT	GO:0002521--leukocyte differentiation	3	4.761904762	0.093209461	PTPRC, MMP9, BCL11A	54	131	13528	5.73706531	1	0.800994456	77.29867755
GOTERM_BP_FAT	GO:0016477--cell migration	4	6.349206349	0.093590597	LYST, NRA2, PDGFRB, ITGB2	54	276	13528	3.630703167	1	0.795119626	77.44284469
GOTERM_BP_FAT	GO:0044243--multicellular organismal catabolic process	2	3.174603175	0.097112643	MMP9, MMP8	54	26	13528	19.27065527	1	0.800573888	78.73520993

REFERENCES:

Abrahamsson, A. E., Geron, I., Gotlib, J., Dao, K. H., Barroga, C. F., Newton, I. G., Giles, F. J., Durocher, J., Creusot, R. S., Karimi, M., et al. (2009). Glycogen synthase kinase 3beta missplicing contributes to leukemia stem cell generation. *Proc Natl Acad Sci U S A* 106, 3925-3929.

Ades, L., Itzykson, R., and Fenaux, P. (2014). Myelodysplastic syndromes. *Lancet*.

Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A., Behjati, S., Biankin, A. V., Bignell, G. R., Bolli, N., Borg, A., Borresen-Dale, A. L., et al. (2013). Signatures of mutational processes in human cancer. *Nature* 500, 415-421.

Andersson, E. R., Sandberg, R., and Lendahl, U. (2011). Notch signaling: simplicity in design, versatility in function. *Development* 138, 3593-3612.

Anzalone, C. L., and Cohen, P. R. (2013). Acute febrile neutrophilic dermatosis (Sweet's syndrome). *Curr Opin Hematol* 20, 26-35.

Babovic, S., and Eaves, C. J. (2014). Hierarchical organization of fetal and adult hematopoietic stem cells. *Exp Cell Res* 329, 185-191.

Banerji, V., Frumm, S. M., Ross, K. N., Li, L. S., Schinzel, A. C., Hahn, C. K., Kakoza, R. M., Chow, K. T., Ross, L., Alexe, G., et al. (2012). The intersection of genetic and chemical genomic screens identifies GSK-3alpha as a target in human acute myeloid leukemia. *J Clin Invest* 122, 935-947.

Bar-Natan, M., Nelson, E. A., Xiang, M., and Frank, D. A. (2012). STAT signaling in the pathogenesis and treatment of myeloid malignancies. *JAKSTAT* 1, 55-64.

Bejar, R., and Steensma, D. P. (2014). Recent developments in myelodysplastic syndromes. *Blood* 124, 2793-2803.

Beurlet, S., Chomienne, C., and Padua, R. A. (2013). Engineering mouse models with myelodysplastic syndrome human candidate genes; how relevant are they? *Haematologica* 98, 10-22.

Bhattacharya, D., Rossi, D. J., Bryder, D., and Weissman, I. L. (2006). Purified hematopoietic stem cell engraftment of rare niches corrects severe lymphoid deficiencies without host conditioning. *J Exp Med* 203, 73-85.

Billstrom, R., Johansson, H., Johansson, B., and Mitelman, F. (1995). Immune-mediated complications in patients with myelodysplastic syndromes--clinical and cytogenetic features. *Eur J Haematol* 55, 42-48.

Bowman, T. V., Merchant, A. A., and Goodell, M. A. (2007). Molecular profiling of hematopoietic stem cells. *Methods Mol Med* 134, 1-16.

Bresolin, S., Trentin, L., Zecca, M., Giordan, M., Sainati, L., Locatelli, F., Basso, G., and te Kronnie, G. (2012). Gene expression signatures of pediatric myelodysplastic syndromes are associated with risk of evolution into acute myeloid leukemia. *Leukemia* 26, 1717-1719.

Briscoe, J., and Therond, P. P. (2013). The mechanisms of Hedgehog signalling and its roles in development and disease. *Nat Rev Mol Cell Biol* 14, 416-429.

Cazzola, M. (2011). Risk assessment in myelodysplastic syndromes and myelodysplastic/myeloproliferative neoplasms. *Haematologica* 96, 349-352.

Cazzola, M., Della Porta, M. G., and Malcovati, L. (2013). The genetic basis of myelodysplasia and its clinical relevance. *Blood* 122, 4021-4034.

Challen, G. A., Boles, N. C., Chambers, S. M., and Goodell, M. A. (2010). Distinct hematopoietic stem cell subtypes are differentially regulated by TGF-beta1. *Cell Stem Cell* 6, 265-278.

Chao, M. P., Seita, J., and Weissman, I. L. (2008). Establishment of a normal hematopoietic and leukemia stem cell hierarchy. *Cold Spring Harb Symp Quant Biol* 73, 439-449.

Cheng, J. X., Anastasi, J., Watanabe, K., Kleinbrink, E. L., Grimley, E., Knibbs, R., Shen, Q. J., and Vardiman, J. W. (2013). Genome-wide profiling reveals epigenetic inactivation of the PU.1 pathway by histone H3 lysine 27 trimethylation in cytogenetically normal myelodysplastic syndrome. *Leukemia* 27, 1291-1300.

Cheng, S. C., Quintin, J., Cramer, R. A., Shepardson, K. M., Saeed, S., Kumar, V., Giamarellos-Bourboulis, E. J., Martens, J. H., Rao, N. A., Aghajani-Refah, A., et al. (2014). mTOR- and HIF-1alpha-mediated aerobic glycolysis as metabolic basis for trained immunity. *Science* 345, 1250684.

Clevers, H., and Nusse, R. (2012). Wnt/beta-catenin signaling and disease. *Cell* 149, 1192-1205.

Corey, S. J., Minden, M. D., Barber, D. L., Kantarjian, H., Wang, J. C., and Schimmer, A. D. (2007). Myelodysplastic syndromes: the complexity of stem-cell diseases. *Nat Rev Cancer* 7, 118-129.

Doble, B. W., Patel, S., Wood, G. A., Kockeritz, L. K., and Woodgett, J. R. (2007). Functional redundancy of GSK-3alpha and GSK-3beta in Wnt/beta-catenin signaling shown by using an allelic series of embryonic stem cell lines. *Dev Cell* 12, 957-971.

Eppert, K., Takenaka, K., Lechman, E. R., Waldron, L., Nilsson, B., van Galen, P., Metzeler, K. H., Poepl, A., Ling, V., Beyene, J., et al. (2011). Stem cell gene expression programs influence clinical outcome in human leukemia. *Nat Med* 17, 1086-1093.

Esen, E., Chen, J., Karner, C. M., Okunade, A. L., Patterson, B. W., and Long, F. (2013). WNT-LRP5 signaling induces Warburg effect through mTORC2 activation during osteoblast differentiation. *Cell Metab* 17, 745-755.

Gan, B., Sahin, E., Jiang, S., Sanchez-Aguilera, A., Scott, K. L., Chin, L., Williams, D. A., Kwiatkowski, D. J., and DePinho, R. A. (2008). mTORC1-dependent and -independent regulation of stem cell renewal, differentiation, and mobilization. *Proc Natl Acad Sci U S A* 105, 19384-19389.

Garcia-Manero, G. (2010). Prognosis of myelodysplastic syndromes. *Hematology Am Soc Hematol Educ Program* 2010, 330-337.

Garcia-Manero, G. (2014). Myelodysplastic syndromes: 2014 update on diagnosis, risk-stratification, and management. *Am J Hematol* 89, 97-108.

Gillespie, J. R., Ulici, V., Dupuis, H., Higgs, A., Dimattia, A., Patel, S., Woodgett, J. R., and Beier, F. (2011). Deletion of glycogen synthase kinase-3beta in cartilage results in up-regulation of glycogen synthase kinase-3alpha protein expression. *Endocrinology* 152, 1755-1766.

Greenberg, P., Cox, C., LeBeau, M. M., Fenaux, P., Morel, P., Sanz, G., Sanz, M., Vallespi, T., Hamblin, T., Oscier, D., et al. (1997). International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood* 89, 2079-2088.

Hasle, H., Baumann, I., Bergstrasser, E., Fenu, S., Fischer, A., Kardos, G., Kerndrup, G., Locatelli, F., Rogge, T., Schultz, K. R., et al. (2004). The International Prognostic Scoring System (IPSS) for childhood myelodysplastic syndrome (MDS) and juvenile myelomonocytic leukemia (JMML). *Leukemia* 18, 2008-2014.

Hasle, H., and Niemeyer, C. M. (2011). Advances in the prognostication and management of advanced MDS in children. *Br J Haematol* 154, 185-195.

Hasle, H., Niemeyer, C. M., Chessells, J. M., Baumann, I., Bennett, J. M., Kerndrup, G., and Head, D. R. (2003). A pediatric approach to the WHO classification of myelodysplastic and myeloproliferative diseases. *Leukemia* 17, 277-282.

Huang, G., Zhao, X., Wang, L., Elf, S., Xu, H., Sashida, G., Zhang, Y., Liu, Y., Lee, J., Menendez, S., et al. (2011). The ability of MLL to bind RUNX1 and

methylation H3K4 at PU.1 regulatory regions is impaired by MDS/AML-associated RUNX1/AML1 mutations. *Blood* 118, 6544-6552.

Huang, J., Zhang, Y., Bersenev, A., O'Brien, W. T., Tong, W., Emerson, S. G., and Klein, P. S. (2009). Pivotal role for glycogen synthase kinase-3 in hematopoietic stem cell homeostasis in mice. *J Clin Invest* 119, 3519-3529.

Issa, J. P. (2013). The myelodysplastic syndrome as a prototypical epigenetic disease. *Blood* 121, 3811-3817.

Itoh, S., Saito, T., Hirata, M., Ushita, M., Ikeda, T., Woodgett, J. R., Algul, H., Schmid, R. M., Chung, U. I., and Kawaguchi, H. (2012). GSK-3alpha and GSK-3beta proteins are involved in early stages of chondrocyte differentiation with functional redundancy through RelA protein phosphorylation. *J Biol Chem* 287, 29227-29236.

Itzykson, R., and Solary, E. (2013). An evolutionary perspective on chronic myelomonocytic leukemia. *Leukemia* 27, 1441-1450.

Jeong, M., Sun, D., Luo, M., Huang, Y., Challen, G. A., Rodriguez, B., Zhang, X., Chavez, L., Wang, H., Hannah, R., et al. (2014). Large conserved domains of low DNA methylation maintained by Dnmt3a. *Nat Genet* 46, 17-23.

Jonas, B. A., and Greenberg, P. L. (2015). MDS prognostic scoring systems - Past, present, and future. *Best Pract Res Clin Haematol* 28, 3-13.

Kalaitzidis, D., Sykes, S. M., Wang, Z., Punt, N., Tang, Y., Ragu, C., Sinha, A. U., Lane, S. W., Souza, A. L., Clish, C. B., et al. (2012). mTOR complex 1 plays critical roles in hematopoiesis and Pten-loss-evoked leukemogenesis. *Cell Stem Cell* 11, 429-439.

Kelly, K. F., Ng, D. Y., Jayakumaran, G., Wood, G. A., Koide, H., and Doble, B. W. (2011). beta-catenin enhances Oct-4 activity and reinforces pluripotency through a TCF-independent mechanism. *Cell Stem Cell* 8, 214-227.

Kharas, M. G., Okabe, R., Ganis, J. J., Gozo, M., Khandan, T., Paktinat, M., Gilliland, D. G., and Gritsman, K. (2010). Constitutively active AKT depletes hematopoietic stem cells and induces leukemia in mice. *Blood* 115, 1406-1415.

Kim, Y. M., Seo, Y. H., Park, C. B., Yoon, S. H., and Yoon, G. (2010). Roles of GSK3 in metabolic shift toward abnormal anabolism in cell senescence. *Ann N Y Acad Sci* 1201, 65-71.

Kirstetter, P., Anderson, K., Porse, B. T., Jacobsen, S. E., and Nerlov, C. (2006). Activation of the canonical Wnt pathway leads to loss of hematopoietic stem cell repopulation and multilineage differentiation block. *Nat Immunol* 7, 1048-1056.

Kogan, S. C., Ward, J. M., Anver, M. R., Berman, J. J., Brayton, C., Cardiff, R. D., Carter, J. S., de Coronado, S., Downing, J. R., Fredrickson, T. N., et al. (2002). Bethesda proposals for classification of nonlymphoid hematopoietic neoplasms in mice. *Blood* 100, 238-245.

Kreso, A., and Dick, J. E. (2014). Evolution of the cancer stem cell model. *Cell Stem Cell* 14, 275-291.

Lane, S. W., Sykes, S. M., Al-Shahrour, F., Shterental, S., Paktinat, M., Lo Celso, C., Jesneck, J. L., Ebert, B. L., Williams, D. A., and Gilliland, D. G.

(2010). The Apc(min) mouse has altered hematopoietic stem cell function and provides a model for MPD/MDS. *Blood* 115, 3489-3497.

Larsson, C. A., Cote, G., and Quintas-Cardama, A. (2013). The changing mutational landscape of acute myeloid leukemia and myelodysplastic syndrome. *Mol Cancer Res* 11, 815-827.

Lindsley, R. C., and Ebert, B. L. (2013). The biology and clinical impact of genetic lesions in myeloid malignancies. *Blood* 122, 3741-3748.

Lobry, C., Ntziachristos, P., Ndiaye-Lobry, D., Oh, P., Cimmino, L., Zhu, N., Araldi, E., Hu, W., Freund, J., Abdel-Wahab, O., et al. (2013). Notch pathway activation targets AML-initiating cell homeostasis and differentiation. *J Exp Med* 210, 301-319.

Magee, J. A., Ikenoue, T., Nakada, D., Lee, J. Y., Guan, K. L., and Morrison, S. J. (2012). Temporal changes in PTEN and mTORC2 regulation of hematopoietic stem cell self-renewal and leukemia suppression. *Cell Stem Cell* 11, 415-428.

Marisavljevic, D., Kraguljac, N., and Rolovic, Z. (2006). Immunologic abnormalities in myelodysplastic syndromes: clinical features and characteristics of the lymphoid population. *Med Oncol* 23, 385-391.

Meacham, C. E., and Morrison, S. J. (2013). Tumour heterogeneity and cancer cell plasticity. *Nature* 501, 328-337.

Mills, K. I., Kohlmann, A., Williams, P. M., Wieczorek, L., Liu, W. M., Li, R., Wei, W., Bowen, D. T., Loeffler, H., Hernandez, J. M., et al. (2009). Microarray-based classifiers and prognosis models identify subgroups with

distinct clinical outcomes and high risk of AML transformation of myelodysplastic syndrome. *Blood* 114, 1063-1072.

Natelson, E. A., and Pyatt, D. (2013). Acquired myelodysplasia or myelodysplastic syndrome: clearing the fog. *Adv Hematol* 2013, 309637.

Nikoloski, G., Langemeijer, S. M., Kuiper, R. P., Knops, R., Massop, M., Tonnissen, E. R., van der Heijden, A., Scheele, T. N., Vandenberghe, P., de Witte, T., et al. (2010). Somatic mutations of the histone methyltransferase gene EZH2 in myelodysplastic syndromes. *Nat Genet* 42, 665-667.

Nimer, S. D. (2008). MDS: a stem cell disorder--but what exactly is wrong with the primitive hematopoietic cells in this disease? *Hematology Am Soc Hematol Educ Program*, 43-51.

Pandolfi, A., Barreyro, L., and Steidl, U. (2013). Concise review: preleukemic stem cells: molecular biology and clinical implications of the precursors to leukemia stem cells. *Stem Cells Transl Med* 2, 143-150.

Pang, W. W., Pluvineau, J. V., Price, E. A., Sridhar, K., Arber, D. A., Greenberg, P. L., Schrier, S. L., Park, C. Y., and Weissman, I. L. (2013). Hematopoietic stem cell and progenitor cell mechanisms in myelodysplastic syndromes. *Proc Natl Acad Sci U S A* 110, 3011-3016.

Parikh, S. A., and Tefferi, A. (2012). Chronic myelomonocytic leukemia: 2012 update on diagnosis, risk stratification, and management. *Am J Hematol* 87, 610-619.

Passegue, E., Jamieson, C. H., Ailles, L. E., and Weissman, I. L. (2003). Normal and leukemic hematopoiesis: are leukemias a stem cell disorder or a

reacquisition of stem cell characteristics? Proc Natl Acad Sci U S A 100 Suppl 1, 11842-11849.

Qing, Y., Lin, Y., and Gerson, S. L. (2012). An intrinsic BM hematopoietic niche occupancy defect of HSC in scid mice facilitates exogenous HSC engraftment. Blood 119, 1768-1771.

Recher, C., Beyne-Rauzy, O., Demur, C., Chicanne, G., Dos Santos, C., Mas, V. M., Benzaquen, D., Laurent, G., Huguet, F., and Payrastre, B. (2005). Antileukemic activity of rapamycin in acute myeloid leukemia. Blood 105, 2527-2534.

Rossi, L., Lin, K. K., Boles, N. C., Yang, L., King, K. Y., Jeong, M., Mayle, A., and Goodell, M. A. (2012). Less is more: unveiling the functional core of hematopoietic stem cells through knockout mice. Cell Stem Cell 11, 302-317.

Schanz, J., Tuchler, H., Sole, F., Mallo, M., Luno, E., Cervera, J., Granada, I., Hildebrandt, B., Slovak, M. L., Ohyashiki, K., et al. (2012). New comprehensive cytogenetic scoring system for primary myelodysplastic syndromes (MDS) and oligoblastic acute myeloid leukemia after MDS derived from an international database merge. J Clin Oncol 30, 820-829.

Scheller, M., Huelsken, J., Rosenbauer, F., Taketo, M. M., Birchmeier, W., Tenen, D. G., and Leutz, A. (2006). Hematopoietic stem cell and multilineage defects generated by constitutive beta-catenin activation. Nat Immunol 7, 1037-1047.

Shivarov, V., and Bullinger, L. (2014). Expression profiling of leukemia patients: key lessons and future directions. Exp Hematol 42, 651-660.

Shlush, L. I., and Minden, M. D. (2015). Preleukemia: the normal side of cancer. *Curr Opin Hematol* 22, 77-84.

Sridhar, K., Ross, D. T., Tibshirani, R., Butte, A. J., and Greenberg, P. L. (2009). Relationship of differential gene expression profiles in CD34+ myelodysplastic syndrome marrow cells to disease subtype and progression. *Blood* 114, 4847-4858.

Sriskanthadevan, S., Jeyaraju, D. V., Chung, T. E., Prabha, S., Xu, W., Skrtic, M., Jhas, B., Hurren, R., Gronda, M., Wang, X., et al. (2015). AML cells have low spare reserve capacity in their respiratory chain that renders them susceptible to oxidative metabolic stress. *Blood*.

Takahashi-Yanaga, F. (2013). Activator or inhibitor? GSK-3 as a new drug target. *Biochem Pharmacol* 86, 191-199.

Takebe, N., Harris, P. J., Warren, R. Q., and Ivy, S. P. (2011). Targeting cancer stem cells by inhibiting Wnt, Notch, and Hedgehog pathways. *Nat Rev Clin Oncol* 8, 97-106.

Tehranchi, R., Woll, P. S., Anderson, K., Buza-Vidas, N., Mizukami, T., Mead, A. J.,

Astrand-Grundstrom, I., Strombeck, B., Horvat, A., Ferry, H., et al. (2010). Persistent malignant stem cells in del(5q) myelodysplasia in remission. *N Engl J Med* 363, 1025-1037.

Tiu, R. V., and Sekeres, M. A. (2014). Making sense of the myelodysplastic/myeloproliferative neoplasms overlap syndromes. *Curr Opin Hematol* 21, 131-140.

Trowbridge, J. J., Xenocostas, A., Moon, R. T., and Bhatia, M. (2006). Glycogen synthase kinase-3 is an in vivo regulator of hematopoietic stem cell repopulation. *Nat Med* 12, 89-98.

Valk, P. J., Verhaak, R. G., Beijen, M. A., Erpelinck, C. A., Barjesteh van Waalwijk van Doorn-Khosrovani, S., Boer, J. M., Beverloo, H. B., Moorhouse, M. J., van der Spek, P. J., Lowenberg, B., and Delwel, R. (2004). Prognostically useful gene-expression profiles in acute myeloid leukemia. *N Engl J Med* 350, 1617-1628.

Vardiman, J. W. (2010). The World Health Organization (WHO) classification of tumors of the hematopoietic and lymphoid tissues: an overview with emphasis on the myeloid neoplasms. *Chem Biol Interact* 184, 16-20.

Vardiman, J. W., Thiele, J., Arber, D. A., Brunning, R. D., Borowitz, M. J., Porwit, A., Harris, N. L., Le Beau, M. M., Hellstrom-Lindberg, E., Tefferi, A., and Bloomfield, C. D. (2009). The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood* 114, 937-951.

Wada, A. (2009). GSK-3 inhibitors and insulin receptor signaling in health, disease, and therapeutics. *Front Biosci* 14, 1558-1570.

Walter, M. J., Ding, L., Shen, D., Shao, J., Grillot, M., McLellan, M., Fulton, R., Schmidt, H.,

Kalicki-Veizer, J., O'Laughlin, M., et al. (2011). Recurrent DNMT3A mutations in patients with myelodysplastic syndromes. *Leukemia* 25, 1153-1158.

Wang, Y., Krivtsov, A. V., Sinha, A. U., North, T. E., Goessling, W., Feng, Z., Zon, L. I., and Armstrong, S. A. (2010). The Wnt/beta-catenin pathway is required for the development of leukemia stem cells in AML. *Science* 327, 1650-1653.

Wang, Z., Smith, K. S., Murphy, M., Piloto, O., Somervaille, T. C., and Cleary, M. L. (2008). Glycogen synthase kinase 3 in MLL leukaemia maintenance and targeted therapy. *Nature* 455, 1205-1209.

Warlick, E. D., and Miller, J. S. (2011). Myelodysplastic syndromes: the role of the immune system in pathogenesis. *Leuk Lymphoma* 52, 2045-2049.

Wirth, D., Gama-Norton, L., Riemer, P., Sandhu, U., Schuch, R., and Hauser, H. (2007). Road to precision: recombinase-based targeting technologies for genome engineering. *Curr Opin Biotechnol* 18, 411-419.

Woll, P. S., Kjallquist, U., Chowdhury, O., Doolittle, H., Wedge, D. C., Thongjuea, S., Erlandsson, R., Ngara, M., Anderson, K., Deng, Q., et al. (2014). Myelodysplastic syndromes are propagated by rare and distinct human cancer stem cells in vivo. *Cancer Cell* 25, 794-808.

Wong, D. J., Nuyten, D. S., Regev, A., Lin, M., Adler, A. S., Segal, E., van de Vijver, M. J., and Chang, H. Y. (2008). Revealing targeted therapy for human cancer by gene module maps. *Cancer Res* 68, 369-378.

Wu, D., and Pan, W. (2010). GSK3: a multifaceted kinase in Wnt signaling. *Trends Biochem Sci* 35, 161-168.

Ye, M., Zhang, H., Amabile, G., Yang, H., Staber, P. B., Zhang, P., Levantini, E., Alberich-Jorda, M., Zhang, J., Kawasaki, A., and Tenen, D. G. (2013).

C/EBPa controls acquisition and maintenance of adult haematopoietic stem cell quiescence. *Nat Cell Biol* 15, 385-394.

Yilmaz, O. H., Valdez, R., Theisen, B. K., Guo, W., Ferguson, D. O., Wu, H., and Morrison, S. J. (2006). Pten dependence distinguishes haematopoietic stem cells from leukaemia-initiating cells. *Nature* 441, 475-482. Zhang, J., Grindley, J. C., Yin, T., Jayasinghe, S., He, X. C., Ross, J. T., Haug, J. S., Rupp, D., Porter-Westpfahl, K. S., Wiedemann, L. M., et al. (2006). PTEN maintains haematopoietic stem cells and acts in lineage choice and leukaemia prevention. *Nature* 441, 518-522.

Zhang, L., Padron, E., and Lancet, J. (2015). The molecular basis and clinical significance of genetic mutations identified in myelodysplastic syndromes. *Leuk Res* 39, 6-17.

Zhao, C., Blum, J., Chen, A., Kwon, H. Y., Jung, S. H., Cook, J. M., Lagoo, A., and Reya, T. (2007). Loss of beta-catenin impairs the renewal of normal and CML stem cells in vivo. *Cancer Cell* 12, 528-541.

CLAIMS:

1. A method of detecting and/or screening for progression to AML or an increased likelihood of progression to AML, in a human subject having MDS, comprising:

determining a sample gene expression profile from a sample comprising DNA from said subject, said sample profile comprising:

(a) the gene expression level of at least 3, optionally at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, or all of the human orthologues of the genes listed in Table 5; and

(b) determining the level of similarity of said sample profile to one or more control profiles.

2. The method of claim 1, wherein

(i) a high level of similarity of the sample profile to a high risk AML progression control profile; (ii) a low level of similarity to low risk AML progression control profile; and/or (iii) a higher level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile indicates progression to AML or an increased likelihood of progression to AML.

3. The method of claim 2, wherein a higher level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile is indicated by hierarchical clustering.

4. The method of claim 2, wherein a higher level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile is indicated by a higher correlation value computed between the sample profile and the high risk AML progression control profile than an equivalent correlation value computed between the sample profile and the low

risk AML progression control profile, optionally wherein the correlation value is a correlation coefficient.

5. The method of claim 1, wherein

(i) a low level of similarity of the sample profile to a high risk AML progression control profile; (ii) a high level of similarity to low risk AML progression control profile; and/or (iii) a lower level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile indicates no progression to AML or an decreased likelihood of progression to AML.

6. The method of claim 5, wherein a lower level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile is indicated by hierarchical clustering.

7. The method of claim 5, wherein a lower level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile is indicated by a lower correlation value computed between the sample profile and the high risk AML progression control profile than an equivalent correlation value computed between the sample profile and the low risk AML progression control profile, optionally wherein the correlation value is a correlation coefficient.

8. The method of any one of claims 1-7, further comprising identifying the WHO MDS subtype of the subject, wherein

(a) (i) a high level of similarity of the sample profile to a high risk AML progression control profile; (ii) a low level of similarity to low risk AML progression control profile and/or (iii) a higher level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile; and

(b) a high risk WHO MDS subtype, indicates a high risk of progression to AML, progression to AML, or an increased likelihood of progression to AML.

9. The method of any one of claims 1-7, further comprising identifying the WHO MDS subtype of the subject, wherein

(a) (i) a low level of similarity of the sample profile to a high risk AML progression control profile; (ii) a high level of similarity to low risk AML progression control profile; and/or (iii) a lower level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile indicates no progression to AML or an decreased likelihood of progression to AML; and

(b) a low risk WHO MDS subtype, indicates a low risk of progression to AML, no progression to AML, or a decreased likelihood of progression to AML.

10. The method of any one of claims 1-7, further comprising identifying the WHO MDS subtype of the subject, wherein

(a) (i) a high level of similarity of the sample profile to a high risk AML progression control profile; (ii) a low level of similarity to low risk AML progression control profile and/or (iii) a higher level of similarity to a high risk AML progression control profile than to a low risk AML progression control profile; and

(b) a low risk WHO MDS subtype,
or

(a) (i) a low level of similarity of the sample profile to a high risk AML progression control profile; (ii) a high level of similarity to low risk AML progression control profile; and/or (iii) a lower level of similarity to a high risk AML progression control profile than to a low risk AML progression control

profile indicates no progression to AML or an decreased likelihood of progression to AML; and

- (b) a high risk WHO MDS subtype,
indicates an intermediate risk of progression to AML.

11. The method of any one of claims 1-10, wherein the subject identified as having progression to AML or an increased likelihood of progression to AML is treated with chemotherapy, radiation, stem cell transplantation and/or bone marrow transplantation.

12. A method of selecting a treatment for a subject with progression to AML or an increased likelihood of progression to AML, comprising:

- (a) identifying the subject with progression to AML or an increased likelihood of progression to AML, according to the method of any one of claims 1-11; and
- (b) selecting a treatment for the subject, wherein the treatment comprises chemotherapy, radiation, stem cell transplantation and/or bone marrow transplantation.

13. A method of detecting and/or screening for myelodysplastic syndrome (MDS) or an increased likelihood of MDS, in a human subject, comprising:

determining a sample gene expression profile from a sample comprising DNA from said subject, said sample profile comprising, consisting of, or consisting essentially of:

- (a) the gene expression level of at least 3, optionally at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, or all of the human orthologues of the genes listed in Table 5; and
- (b) determining the level of similarity of said sample profile to one or more control profiles, wherein (i) a high level of similarity of the sample profile

to an MDS specific control profile; (ii) a low level of similarity to a non-MDS control profile; and/or (iii) a higher level of similarity to an MDS specific control profile than to a non-MDS control profile indicates the presence of, or an increased likelihood of, MDS.

14. The method of claim 13, wherein a higher level of similarity to an MDS specific control profile than to a non-MDS control profile is indicated by hierarchical clustering.

15. The method of claim 13, wherein a higher level of similarity to an MDS specific control profile than to a non-MDS control profile is indicated by a higher correlation value computed between the sample profile and the MDS specific control profile than an equivalent correlation value computed between the sample profile and the non-MDS control profile, optionally wherein the correlation value is a correlation coefficient.

16. The method of any one of claims 1-15, wherein the subject is a pediatric subject.

17. The method of any one of claims 1-16, wherein determining the sample gene expression profile comprises the steps:

- a) providing the sample comprising mRNA from the subject;
- b) isolating mRNA from the sample;
- c) amplifying the mRNA; and
- d) determining the gene expression level of the at least 3 genes by means of RT-PCR.

18. The method of any one of claims 1-17, wherein the sample is a clinical specimen.

19. The method of any one of claims 1-17, wherein the sample is derived from blood and/or bone marrow.

20. A kit for detecting and/or screening for progression to AML or an increased likelihood of progression to AML, in a human subject having MDS, in a sample, comprising:

a) at least one detection agent for determining the gene expression level of at least 3, optionally at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, or all of the human orthologues of the genes listed in Table 5; and

b) instructions for use.

21. A kit for detecting and/or screening for MDS or an increased likelihood of MDS in a sample from a human subject, comprising:

a) at least one detection agent for determining the gene expression level of at least 3, optionally at least 5, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, or all of the human orthologues of the genes listed in Table 5; and

b) instructions for use.

22. The kit according to claim 20 or 21 wherein the detection agents are probes and/or primers.

23. The kit according to any one of claims 20 to 22, further comprising, PCR reagents.

24. The kit according to any one of claims 20-23, further comprising a computer-readable medium that causes a computer to compare gene expression from a sample at the selected genes to one or more control

profiles and compute a correlation value between the sample and control profile.

25. A method of screening a test agent for treatment or prevention of myelodysplastic syndrome (MDS), comprising a) administering a test agent to i) a transgenic mouse model of MDS whose genome comprises a disruption of GSK-3 β , ii) a tissue from said transgenic mouse model, or iii) a cell derived from said transgenic mouse model; and b) determining the effect of said test agent on the mouse model, tissue or cell, wherein a reduction of signs and/or symptoms of MDS in the mouse, tissue or cell identifies a test agent for the treatment or prevention of MDS.

26. The method of claim 25, wherein the disruption is a homozygous GSK-3 β knockout.

27. The method of claim 25 or 26, wherein the GSK-3 β knock out is a conditional knockout of GSK-3 β .

28. A method of screening a test agent for treatment or prevention of acute myeloid leukemia (AML), comprising a) administering a test agent to i) a transgenic mouse model of AML whose genome comprises a disruption of GSK-3 α and a disruption of GSK-3 β , ii) a tissue from said transgenic mouse model, or iii) a cell derived from said transgenic mouse model; and b) determining the effect of said agent on the mouse model, tissue or cell, wherein a reduction of signs and/or symptoms of AML in the mouse, tissue or cell identifies a test agent for the treatment or prevention of AML.

29. The method of claim 28, wherein the disruption of GSK-3 α is a homozygous GSK-3 α knockout and/or the disruption of GSK-3 β is a homozygous GSK-3 β knockout.

30. The method of claim 29, wherein the GSK-3 α knockout is a conditional knockout of GSK-3 α and/or the GSK-3 β knock out is a conditional knockout of GSK-3 β .

31. A method of screening a test agent for treatment or prevention of myelodysplastic syndrome (MDS), comprising a) administering a test agent to a cell comprising a disruption of GSK-3 β and b) determining the effect of said agent on the cell.

32. The method of claim 31, wherein the disruption is a homozygous GSK-3 β knockout.

33. The method of claim 32, wherein the GSK-3 β knock out is a conditional knockout of GSK-3 β .

34. A method of screening a test agent for treatment or prevention of acute myeloid leukemia comprising a) administering a test agent to a cell comprising a disruption of GSK-3 α and a disruption of GSK-3 β and b) determining the effect of said agent on the cell.

35. The method of claim 32, wherein the disruption of GSK-3 α is a homozygous GSK-3 α knockout and/or the disruption of GSK-3 β is a homozygous GSK-3 β knockout.

36. The method of claim 25, wherein the GSK-3 α knockout is a conditional knockout of GSK-3 α and/or the GSK-3 β knock out is a conditional knockout of GSK-3 β .

37. The method of any one of claims 31-36, wherein the cell is a human cell.

34/47

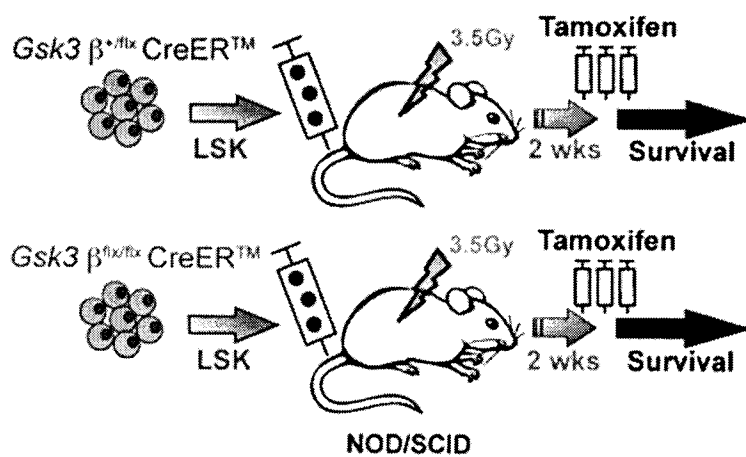
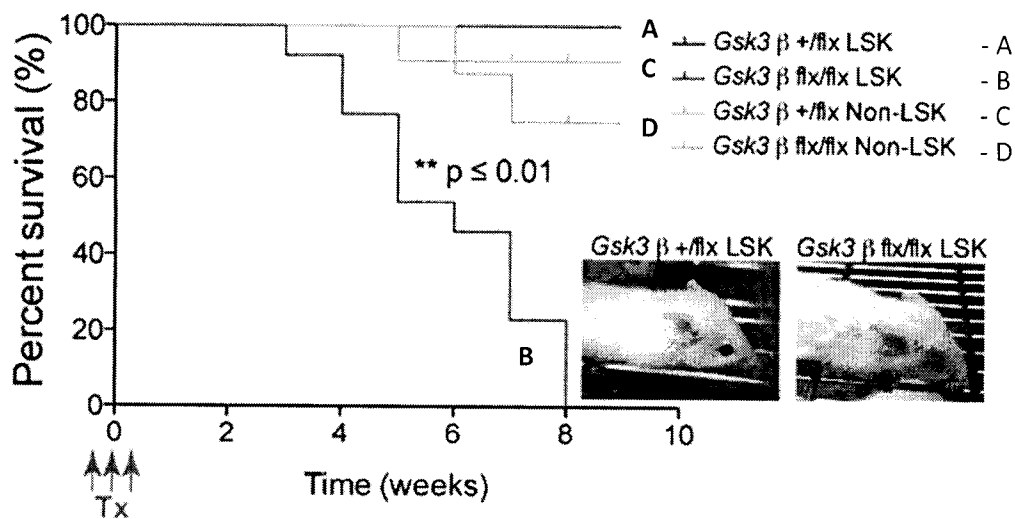
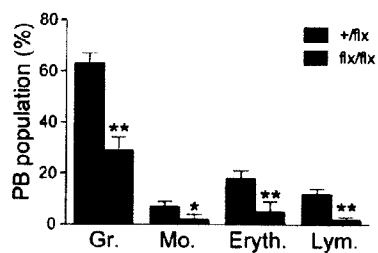
E**F**

FIGURE 10 (cont)

35/47

G



H

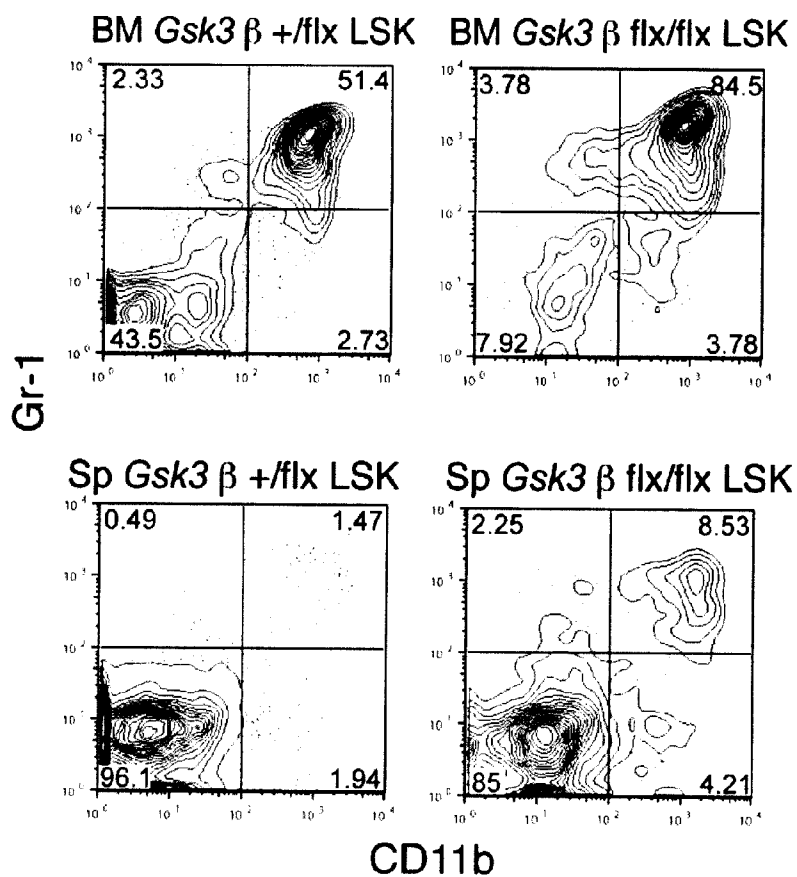


FIGURE 10 (cont)

36/47

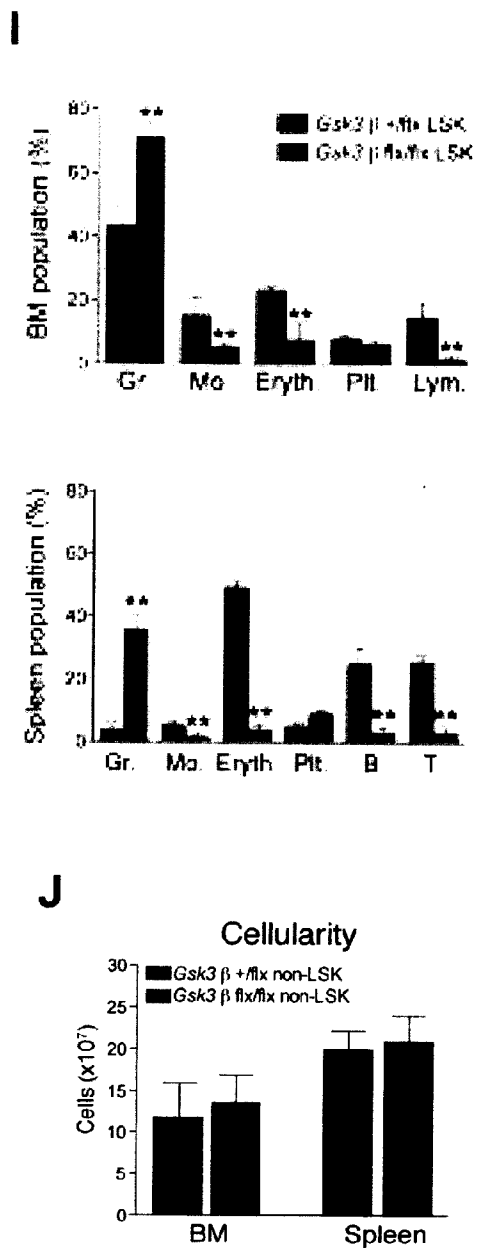


FIGURE 10 (cont)

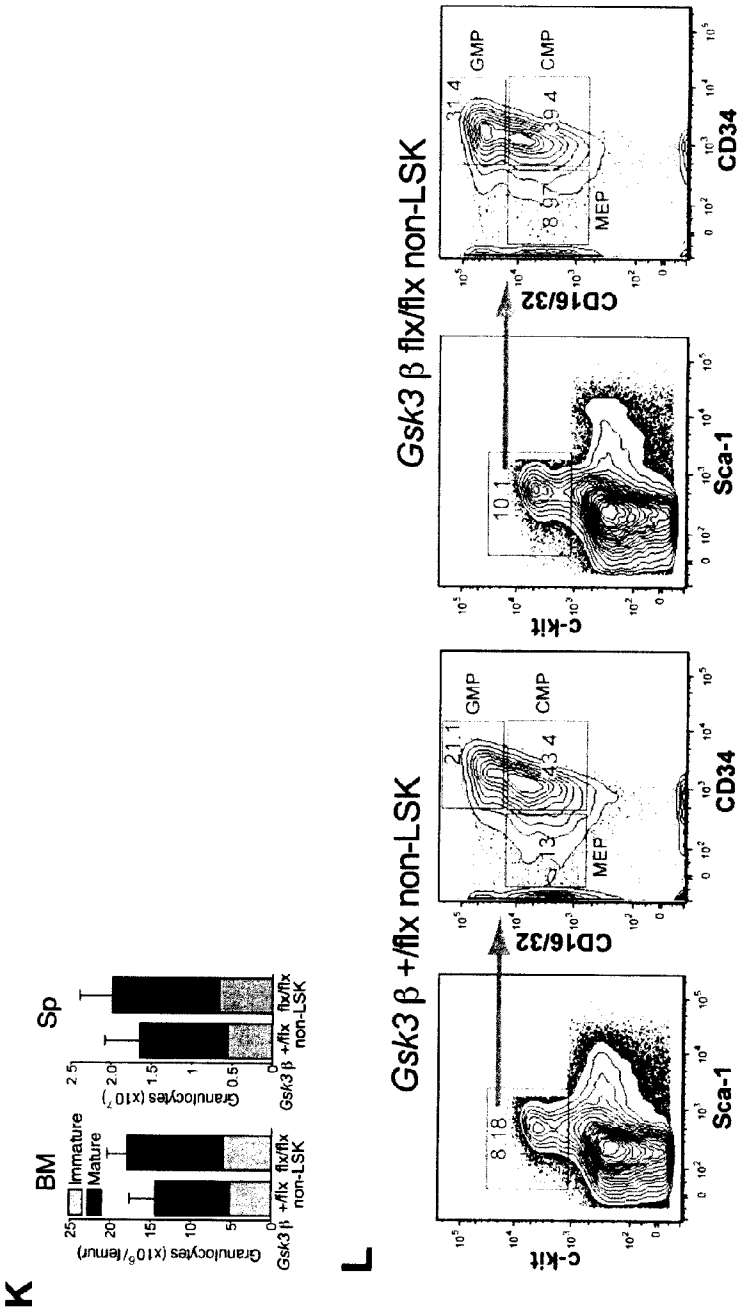


FIGURE 10 (cont)

38/47

M

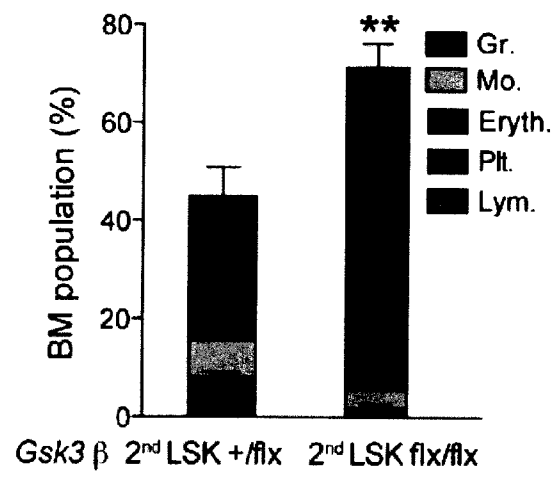


FIGURE 10 (cont)

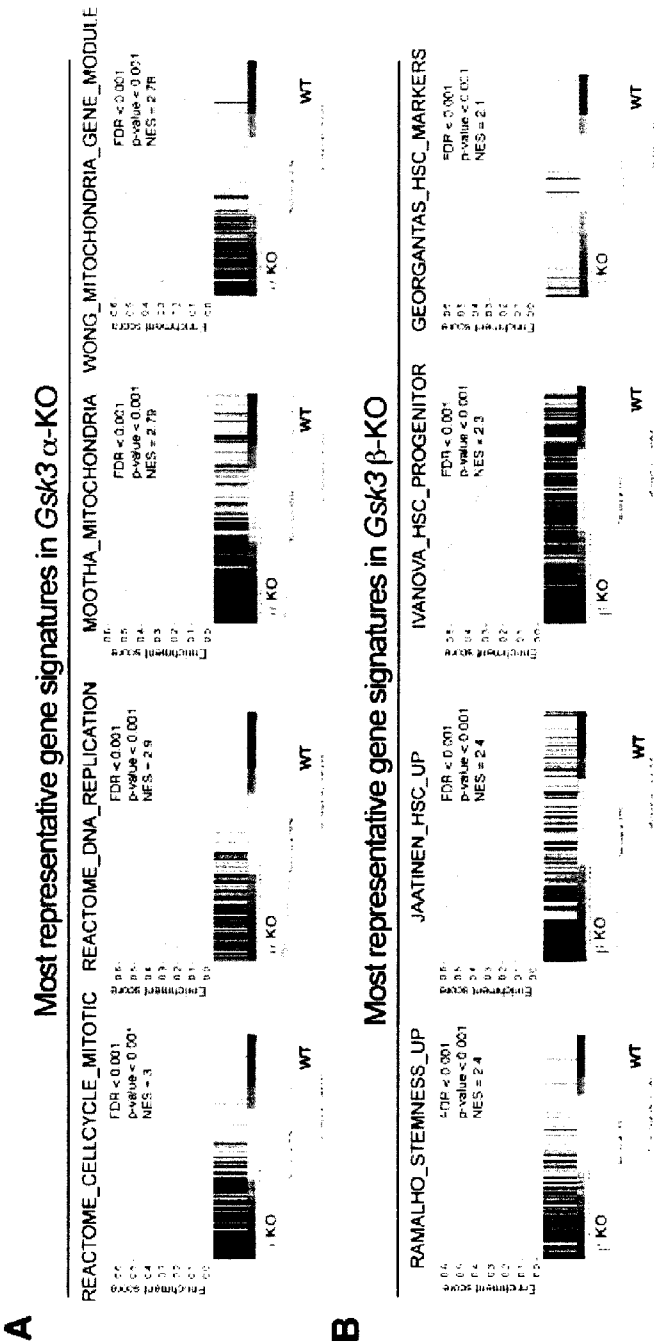


FIGURE 11

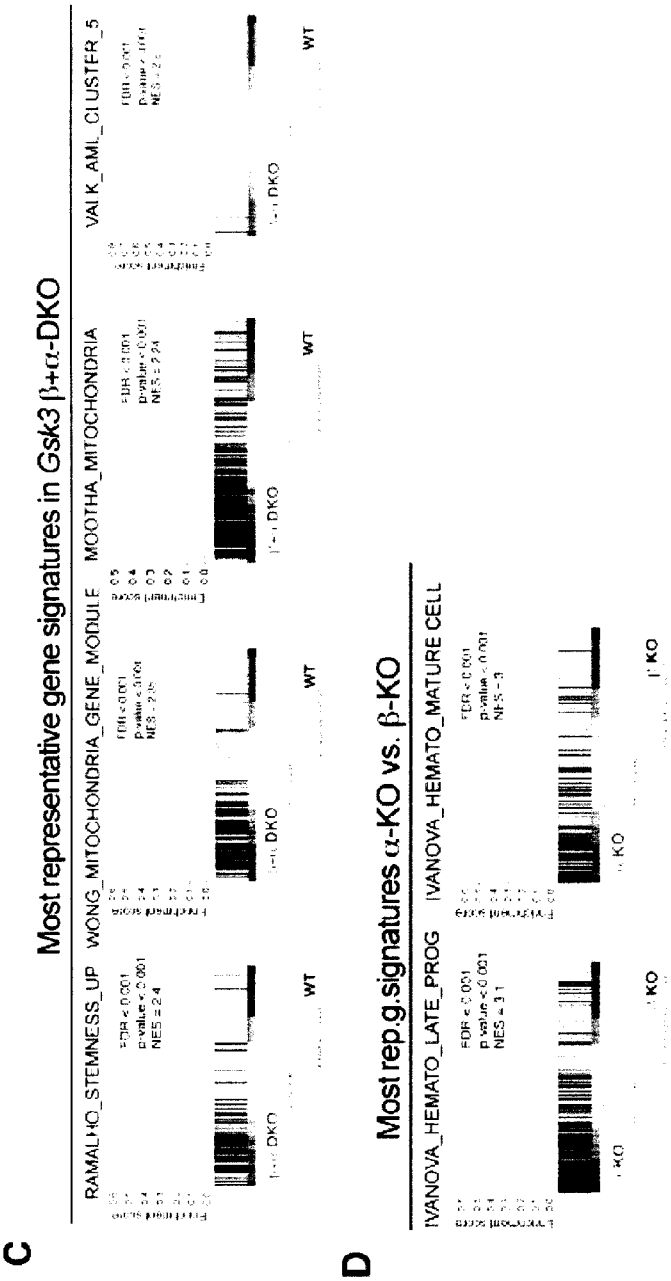


FIGURE 11 (cont)

E

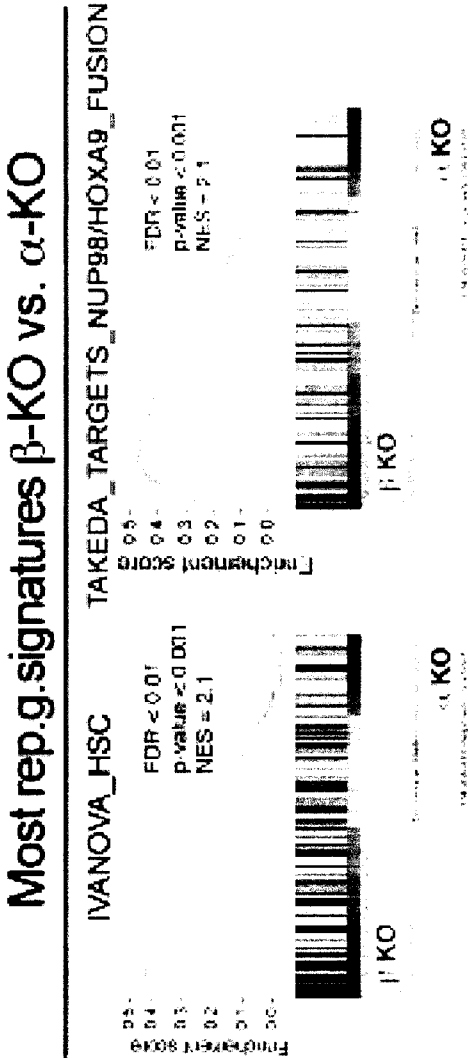


FIGURE 11 (cont)

42/47

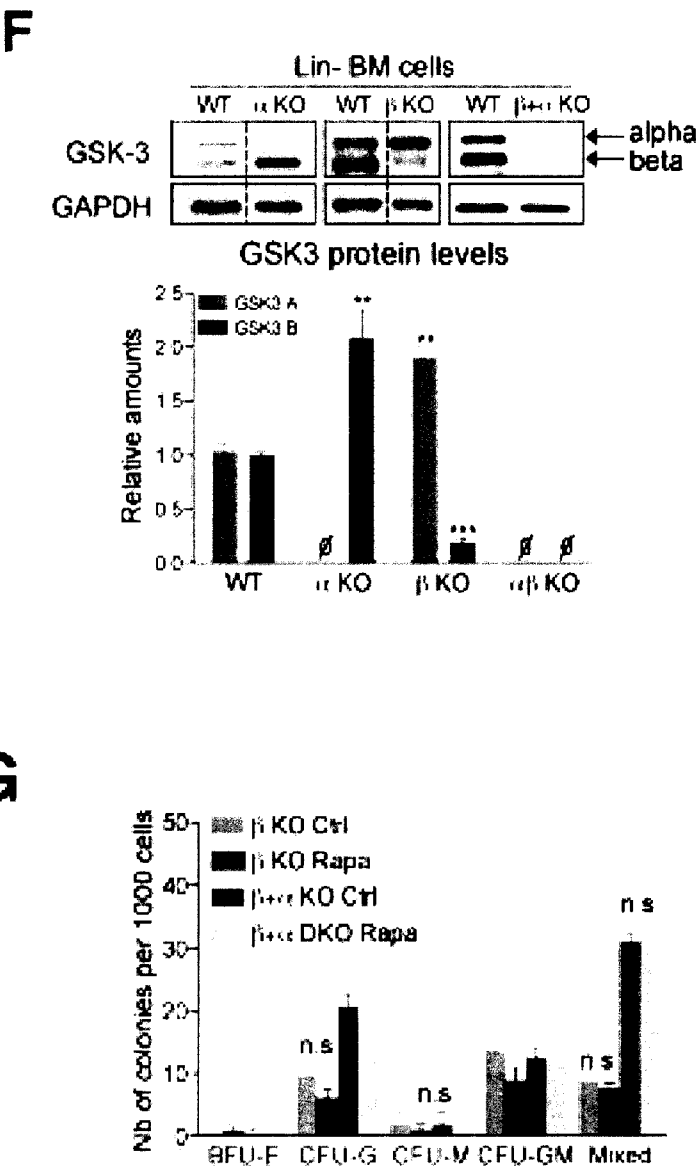


FIGURE 11 (cont)

43/47



FIGURE 12

44/47

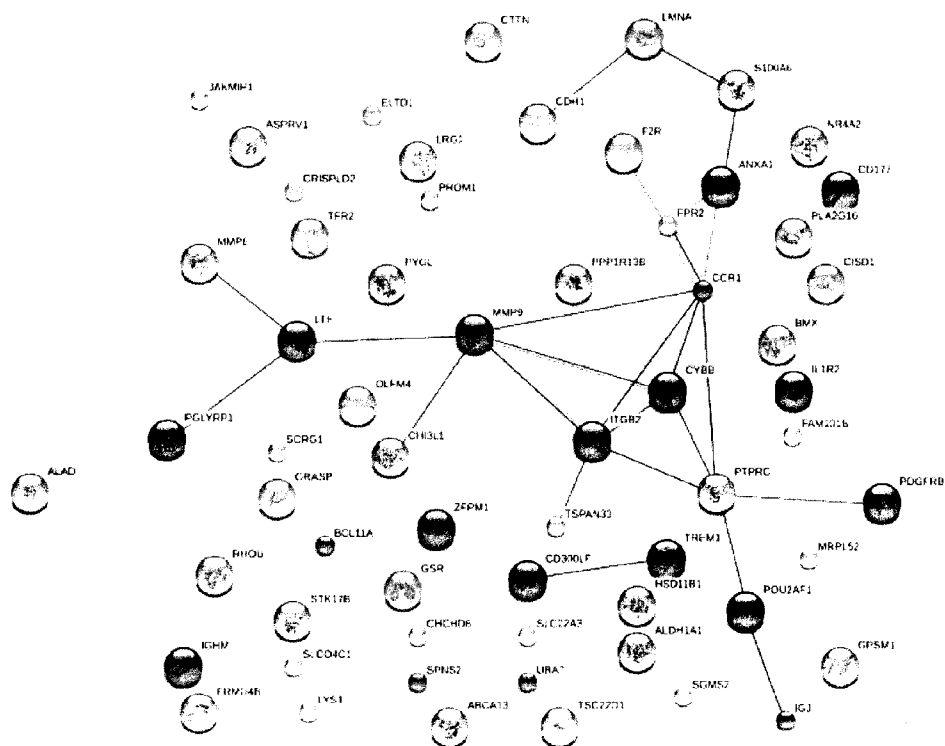
C**Known protein-protein interactions of Gsk3 β signature**

FIGURE 12 (cont)

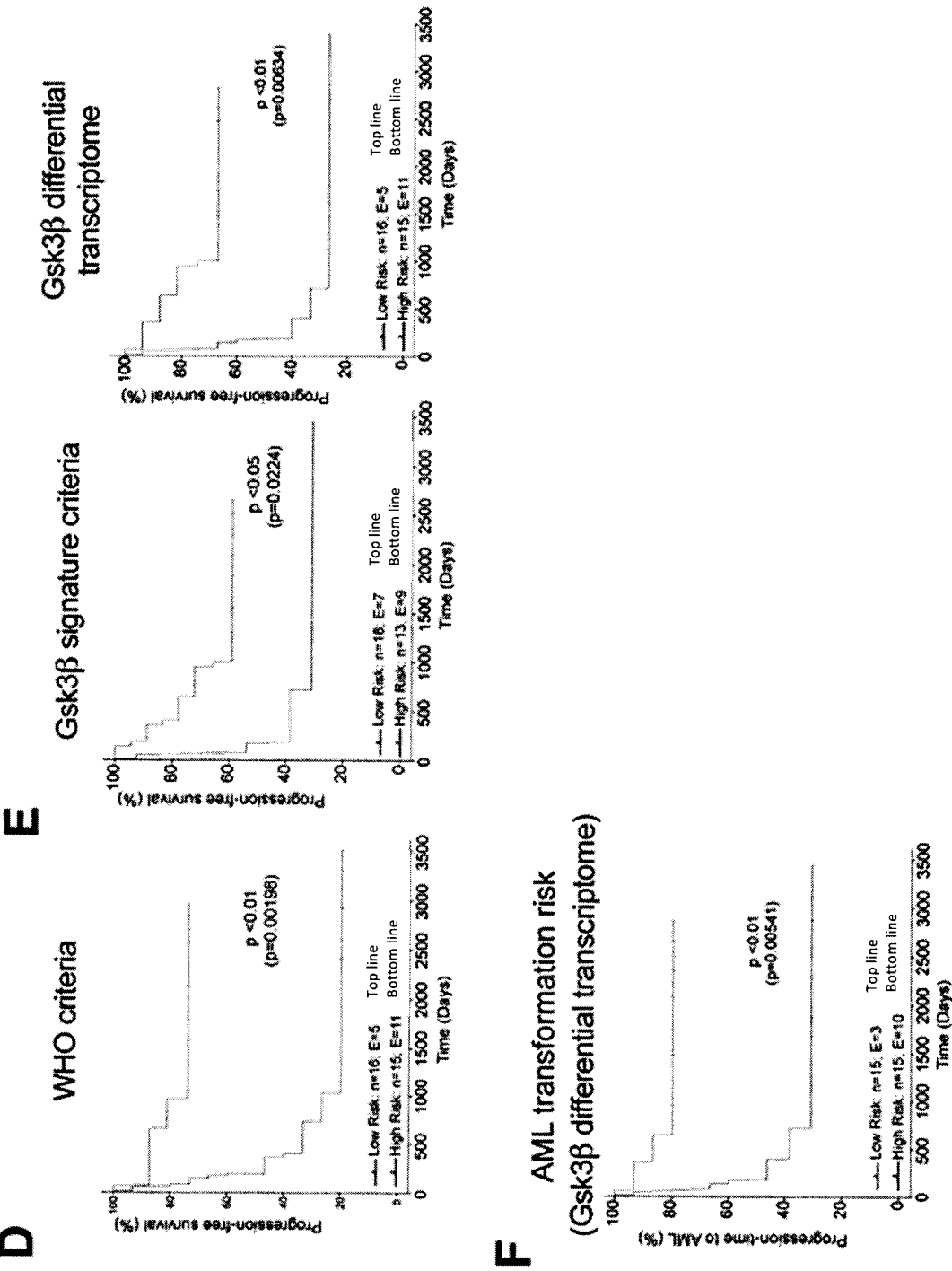


FIGURE 12 (cont)

46/47

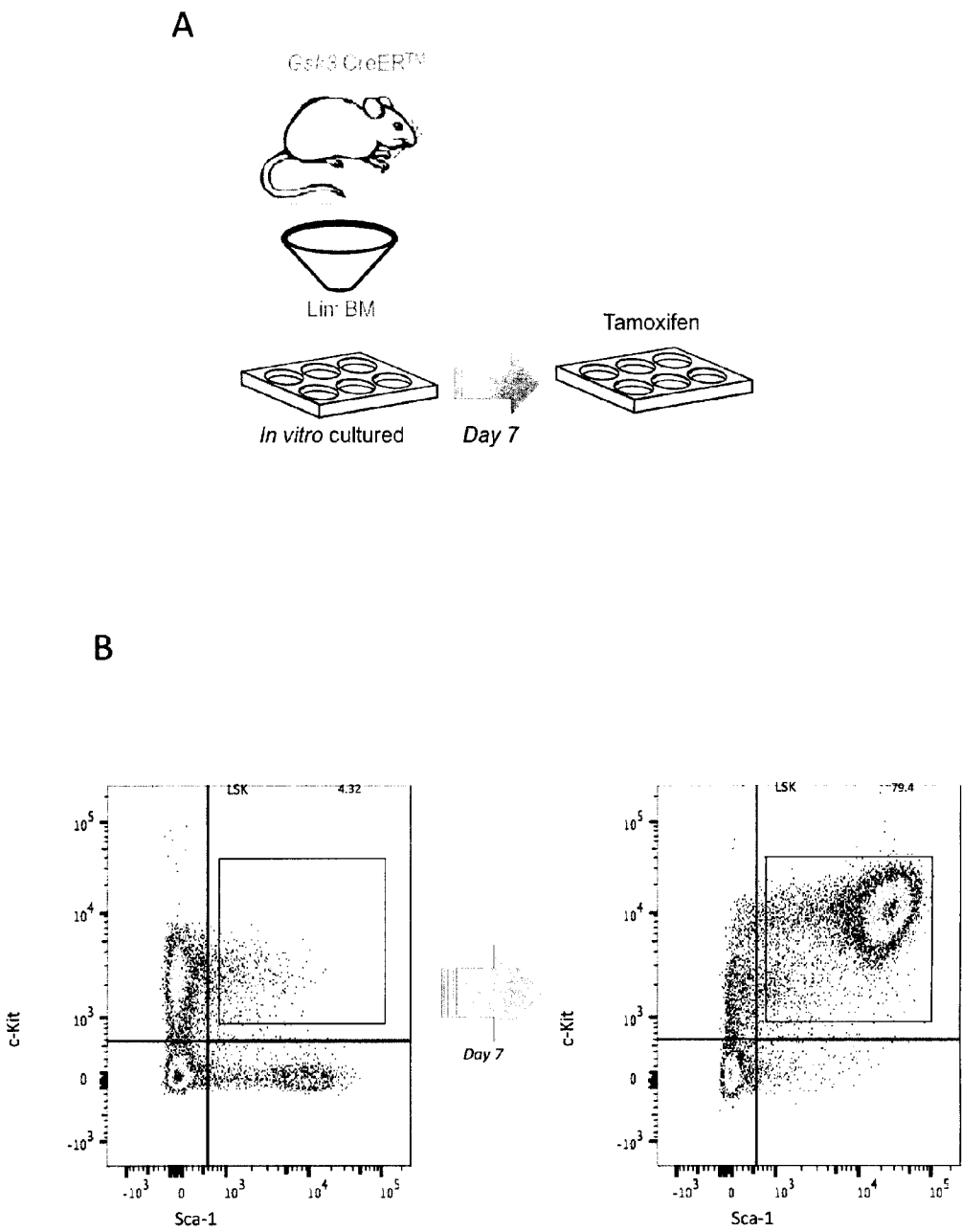


FIGURE 13

47/47

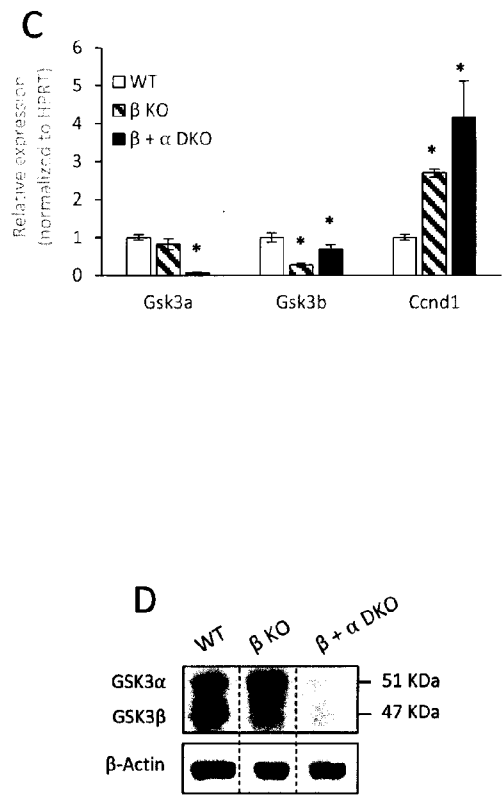


FIGURE 13 (cont)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CA2017/050020

A. CLASSIFICATION OF SUBJECT MATTER

IPC: **C12Q 1/68** (2006.01), **C12Q 1/00** (2006.01), **G01N 33/48** (2006.01), **G06F 19/20** (2011.01)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12Q 1/68 (2006.01), C12Q 1/00 (2006.01), G01N 33/48 (2006.01), G06F 19/20 (2011.01)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic database(s) consulted during the international search (name of database(s) and, where practicable, search terms used)

DWPI, BIOSIS, CABA, EMBASE, FSTA, MEDLINE, INPADOC, AUPATFULL, CANPATFULL, CNFULL, DEFULL, EPFULL, FRFULL, GBFULL, INFULL, JPFULL, RFULL, PCTFULL, USFULL, Canadian Patent Database, SCOPUS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2005/0233330 A1 (MANO) 20 October 2005 (20-10-2005)	
A	BRESOLIN, S; et al., " <i>Gene expression signatures of pediatric myelodysplastic syndromes are associated with risk of evolution into acute myeloid leukemia</i> ". <i>Leukemia</i> , July 2012 (07-2012), Vol. 26(7), pp. 1717-1719, ISSN 08876924	
A	SRIDHAR, K., " <i>Relationship of differential gene expression profiles in CD34+ myelodysplastic syndrome marrow cells to disease subtype and progression</i> ". <i>Blood</i> , 26 November 2009 (26-11-2009), Vol. 114(23), pp. 4847-4858, ISSN 00064971	
A	MILLS, K., " <i>Microarray-based classifiers and prognosis models identify subgroups with distinct clinical outcomes and high risk of AML transformation of myelodysplastic syndrome</i> ". <i>Blood</i> , 1 January 2009 (01-01-2009), Vol. 114(5), pp. 1063-1072, ISSN 00064971	

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

Date of mailing of the international search report
07 April 2017 (07-04-2017)Name and mailing address of the ISA/CA
Canadian Intellectual Property Office
Place du Portage I, C114 - 1st Floor, Box PCT
50 Victoria Street
Gatineau, Quebec K1A 0C9
Facsimile No.: 819-953-2476

Authorized officer

Claire Plumb (819) 639-7501

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CA2017/050020**Box No. I****Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)**

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

- a. ☒ forming part of the international application as filed:
- ☒ in the form of an Annex C/ST.25 text file.
 - ☐ on paper or in the form of an image file.
- b. ☒ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
- c. ☐ furnished subsequent to the international filing date for the purposes of international search only:
- ☐ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).
 - ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CA2017/050020**Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of the first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claim Nos.: 11
because they relate to subject matter not required to be searched by this Authority, namely:

Please see Continuation of Box No. II

2. ☒ Claim Nos.: 1-10 and 12-24
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

Please see Continuation of Box No. II

3. ☐ Claim Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please see Continuation of Box No. III

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claim Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim Nos.: 1-24

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CA2017/050020

Patent Document Cited in Search Report	Publication Date	Patent Family Member(s)	Publication Date
US2005233330A1	20 October 2005 (20-10-2005)	None	

Continuation of: Box No. II

Claim 11 is directed to a method for treatment of the human or animal body by surgery or therapy, which the International Searching Authority is not required to search under PCT Rule 39.1(iv). However, this Authority has carried out a search based on the alleged effect or purpose/use of the product defined in claim 11.

The International Searching Authority has not carried out a search for the entire scope of claims 1-24, under Article 17(2)(b) of the PCT. The description, claims, and/or drawings fail to comply with the prescribed requirements to such an extent that a meaningful search could not be carried out. Claims 1-24 so lack clarity and/or support that a meaningful search over the whole of the claimed scope is impossible. Consequently, the search has been established for the parts of the application which appear to be clear and supported, namely methods and kits featuring the detection of all 63 genes listed in Table 5. The description provides a set of 63 genes that, when detected at particular levels, allow a GSK-3 β signature associated with myelodysplastic syndrome (MDS) or acute myeloid leukemia (AML) to be detected (see [00202]). The description further discloses the use of the signature to divide pediatric MDS patients into a low risk-group and a high risk-group (see [00203]), and that the information produced following detection of the signature can be combined with further information to prognose transformation to AML (see [00204]). It follows that the description provides sufficient support for claims featuring the detection of the complete set of 63 genes recited in Table 5 and the subsequent analysis of expression information for the detection of a GSK-3 β signature linked to MDS or AML. The description does not, however, provide the guidance necessary to determine which genes, chosen from within the set of 63 genes, would provide sufficient useful information to assess the presence of a GSK-3 β signature linked to either MDS or AML. The international search has therefore been restricted to claims which feature the detection of the set of 63 genes recited in Table 5 of the description and the use of information related to the expression of these genes in a human subject with MDS to detect or screen for an increased likelihood of progression to AML.

Continuation of: Box No. III

The claims are directed to a plurality of inventive concepts as follows:

Group A - Claims 1-24 are directed to methods or kits featuring the genes listed in Table 5 for use in the analysis of either myelodysplastic syndrome (MDS) or acute myeloid leukemia (AML); and

Group B - Claims 25-37 are directed to methods featuring a transgenic mouse whose genome comprises either a disruption of GSK-3 β or a disruption of both GSK-3 β and GSK-3 α , for screening a test agent for treatment or prevention of either MDS or AML.

The only common technical feature of the two groups are the diseases MDS and AML, both of which are known in the art and cannot serve to unify Groups A and B. The method or kits for analysis featuring genes of Group A and the method for screening of Group B featuring a transgenic mouse therefore do not share a common, novel technical element that would serve to unify them.

The claims must be limited to one inventive concept as set out in PCT Rule 13.

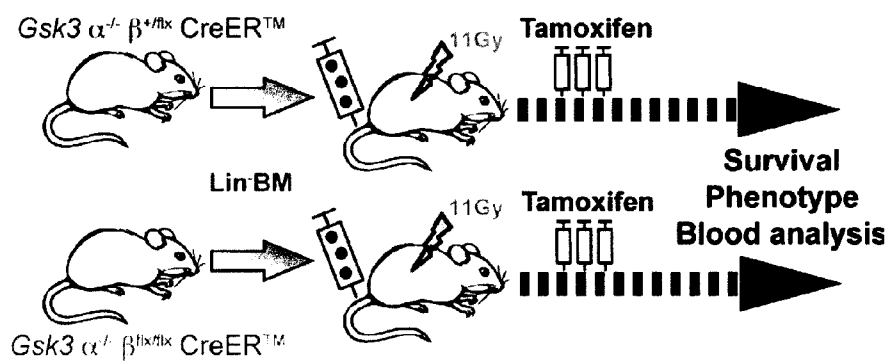
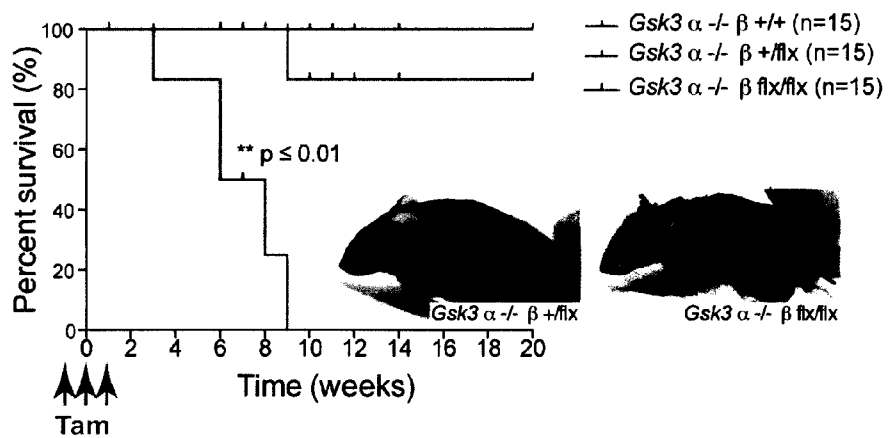
A**B**

FIGURE 1

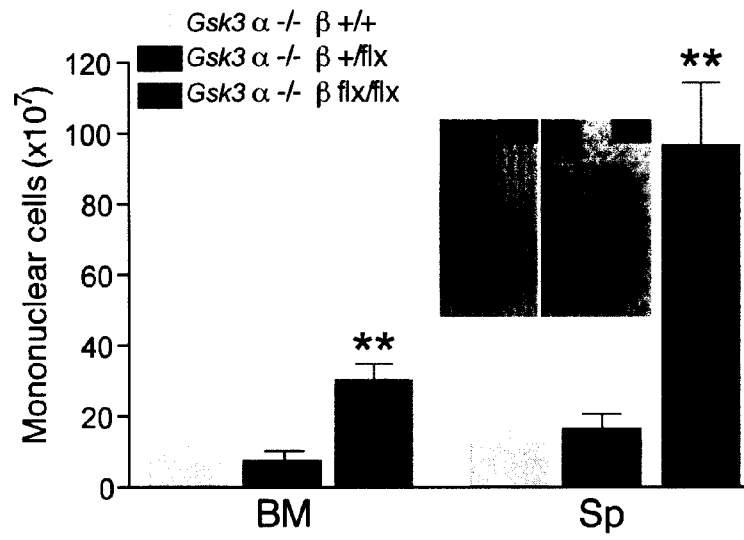
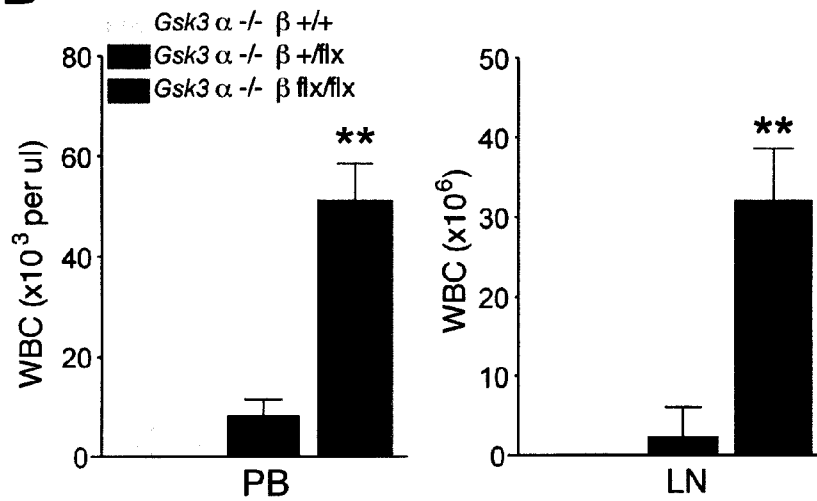
C**D**

FIGURE 1 (cont)

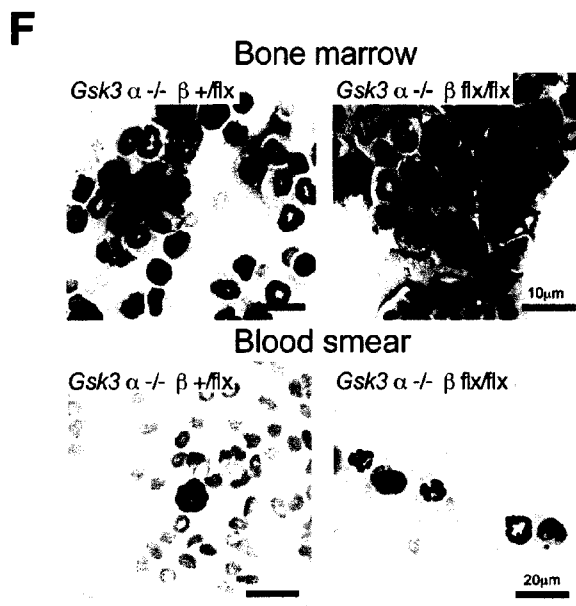
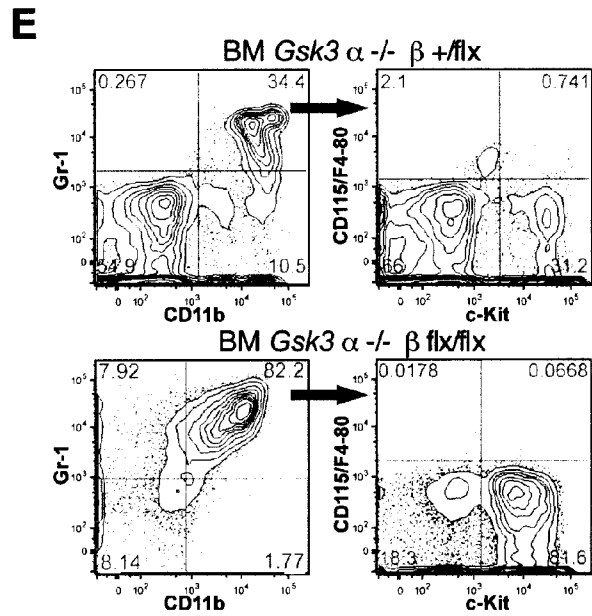


FIGURE 1 (cont)

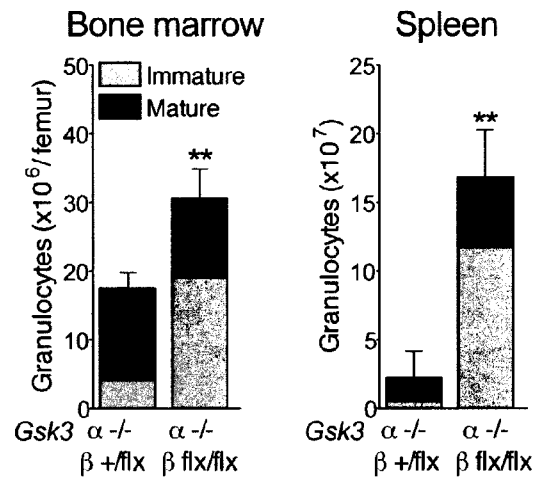
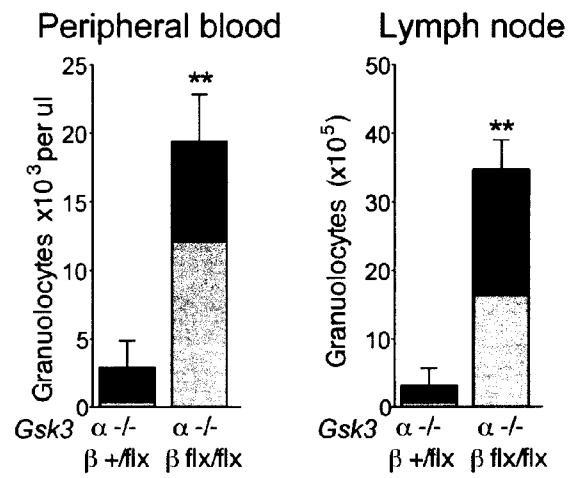
G**H**

FIGURE 1 (cont)

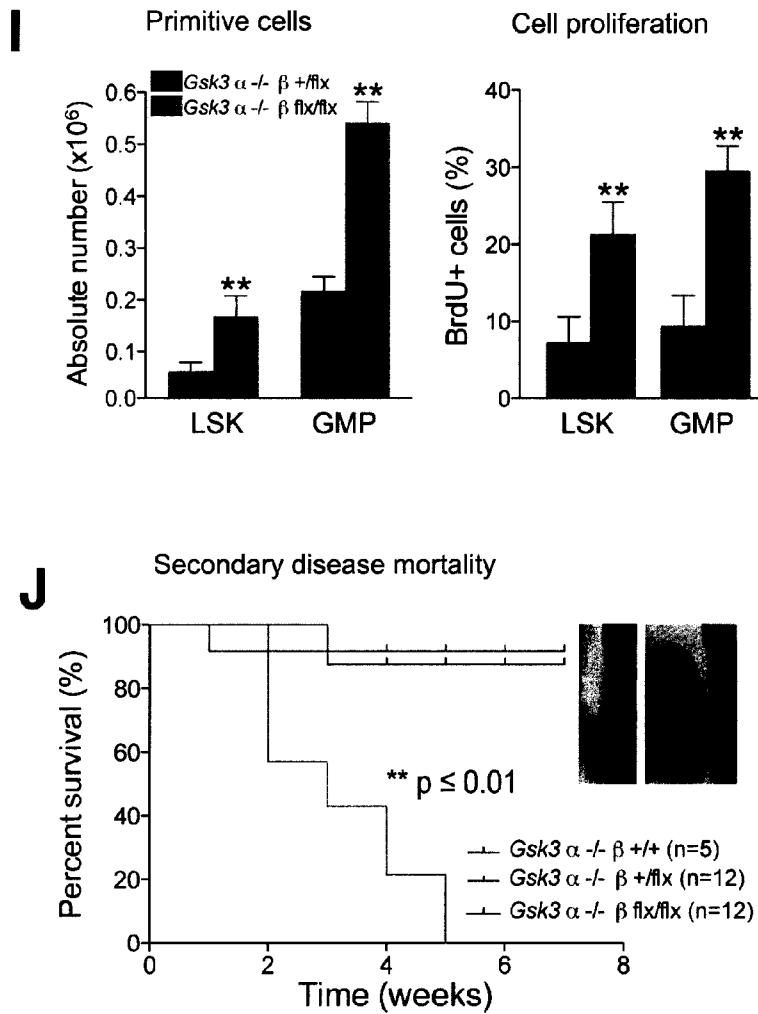


FIGURE 1 (cont)

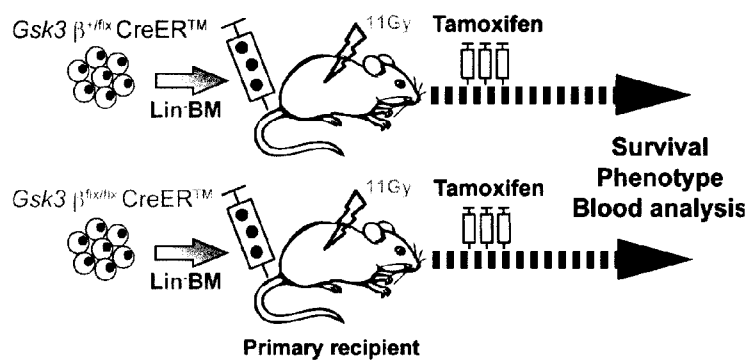
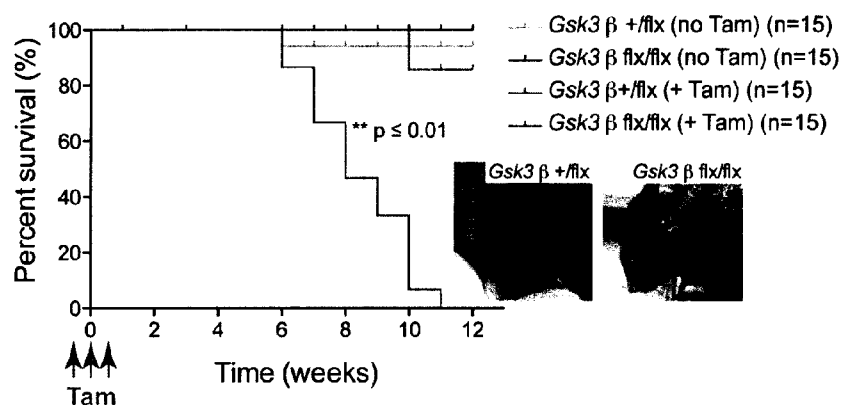
A**B**

FIGURE 2

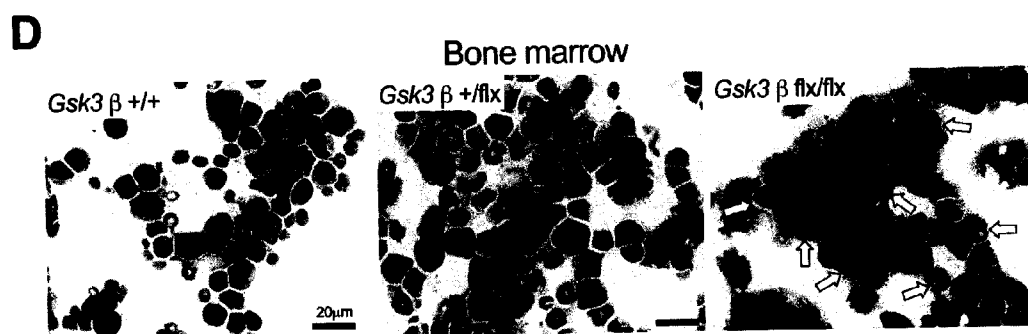
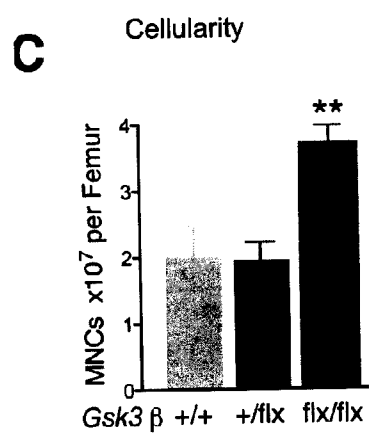


FIGURE 2 (cont)

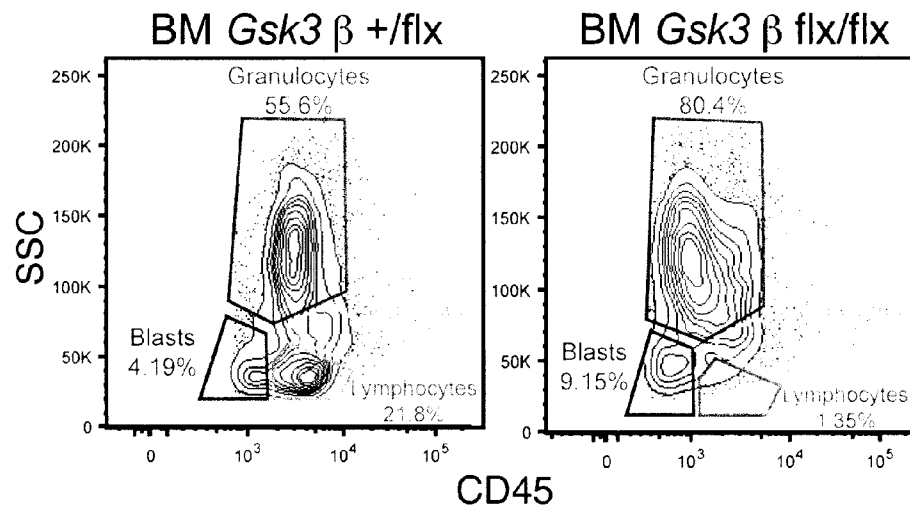
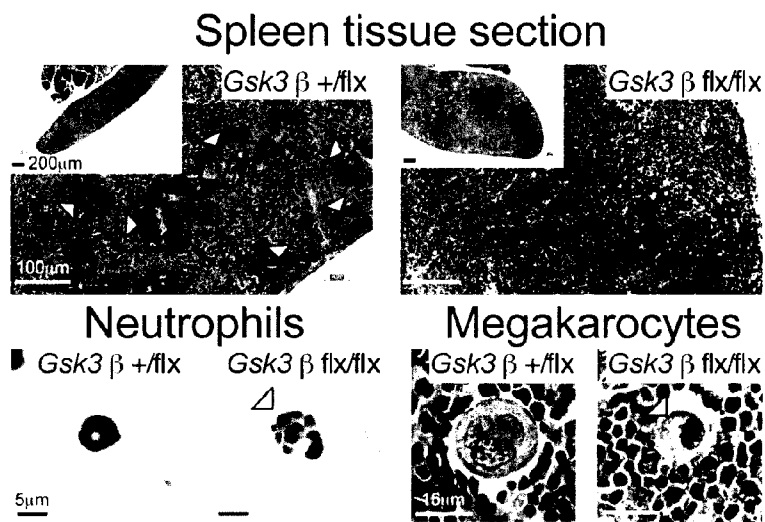
E**F**

FIGURE 2 (cont)

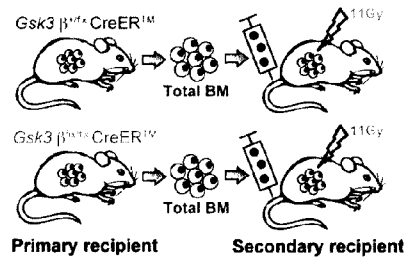
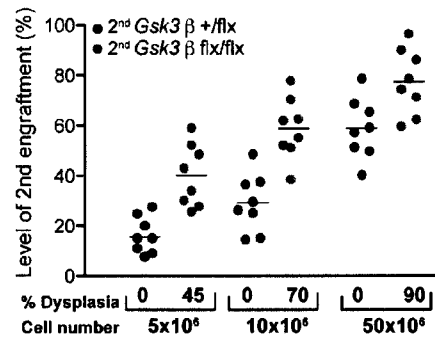
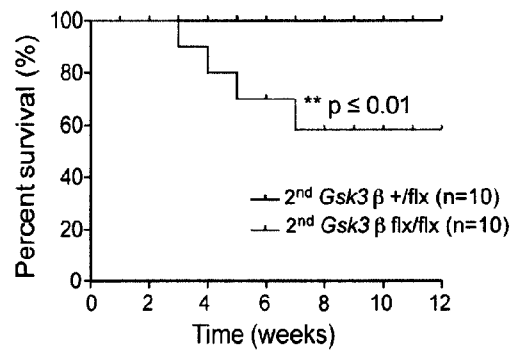
G**H** Dysplasia initiating activity**I** Disease mortality

FIGURE 2 (cont)

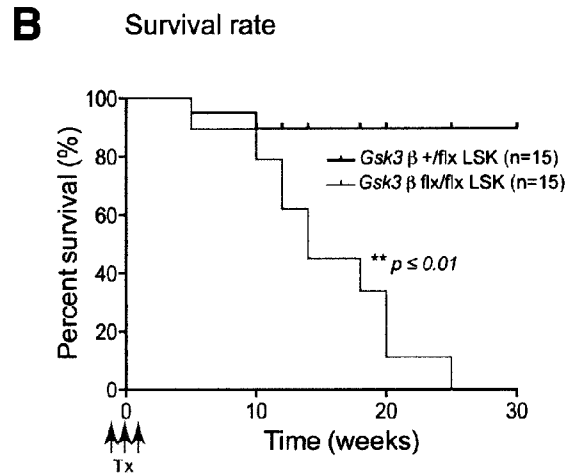
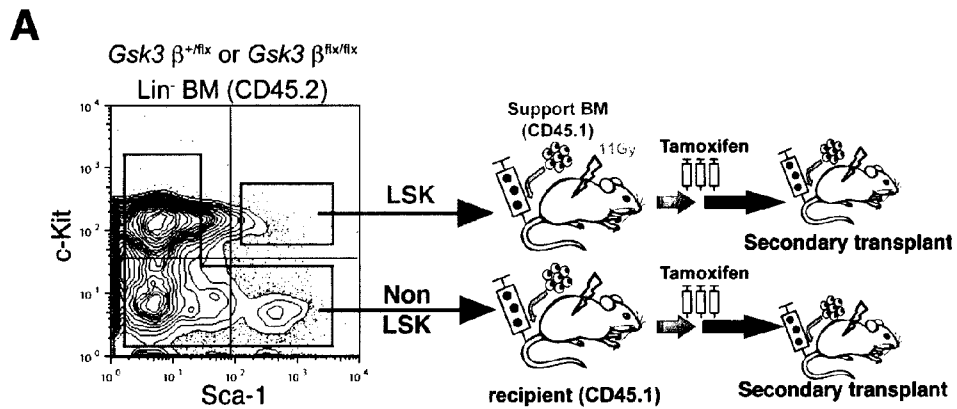


FIGURE 3

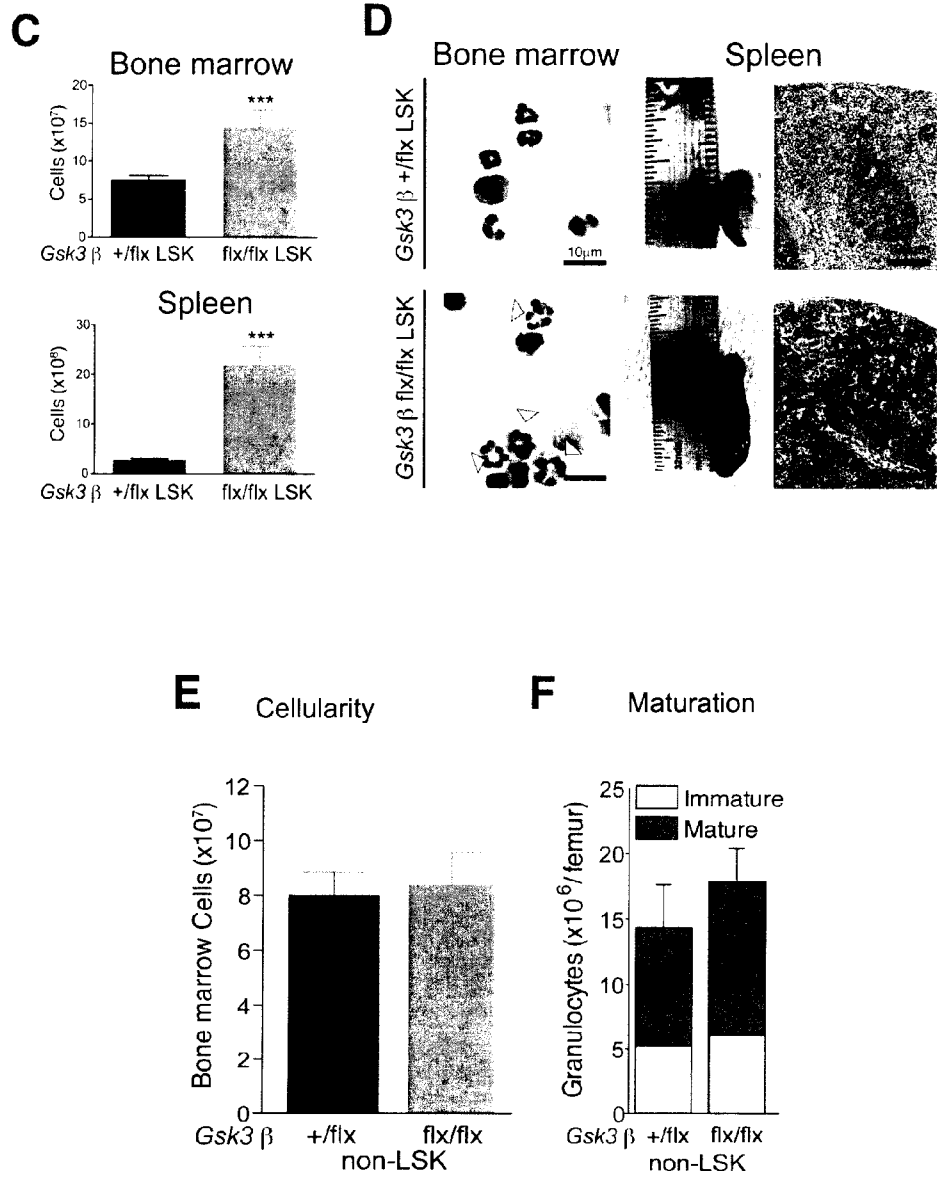


FIGURE 3 (cont)

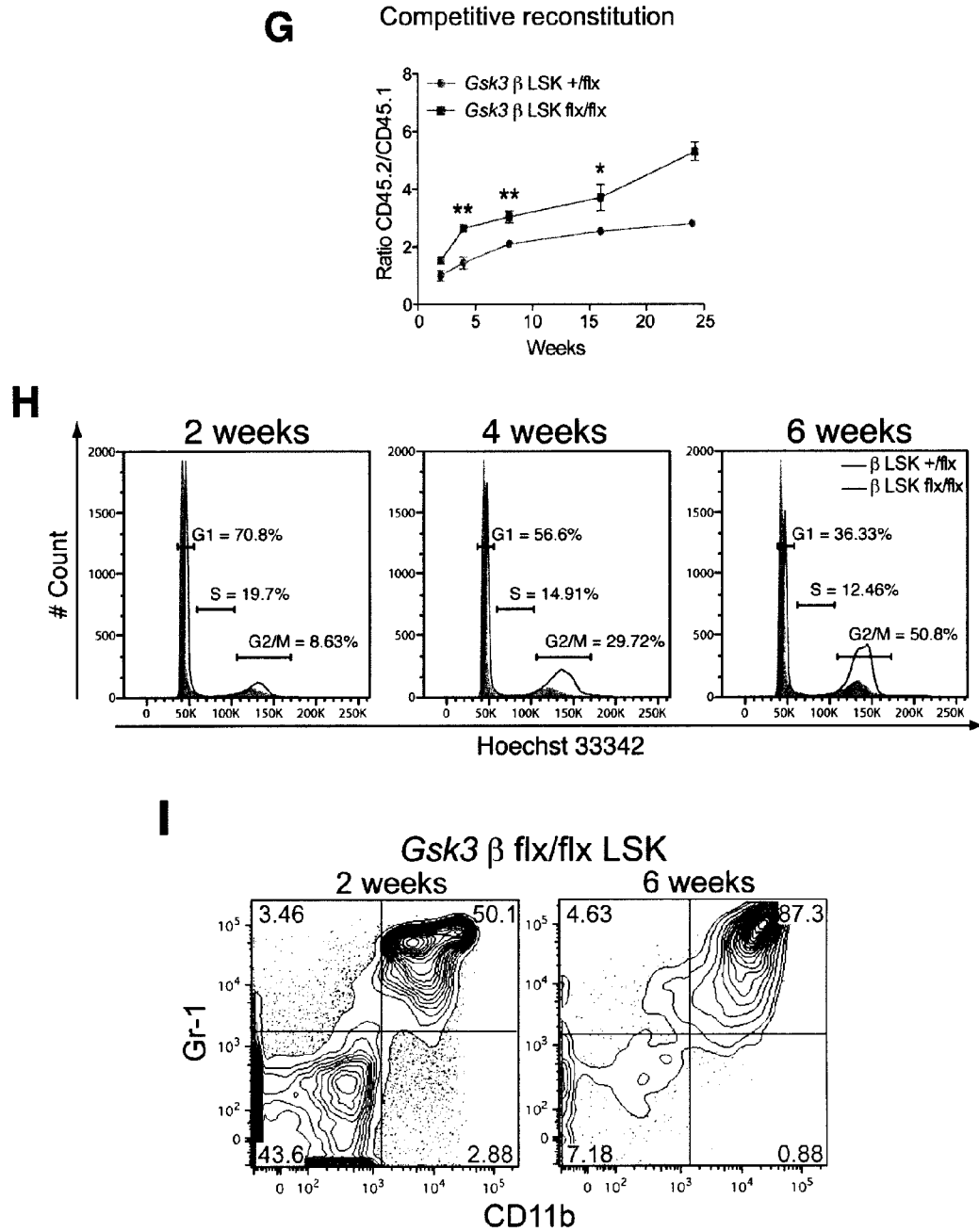
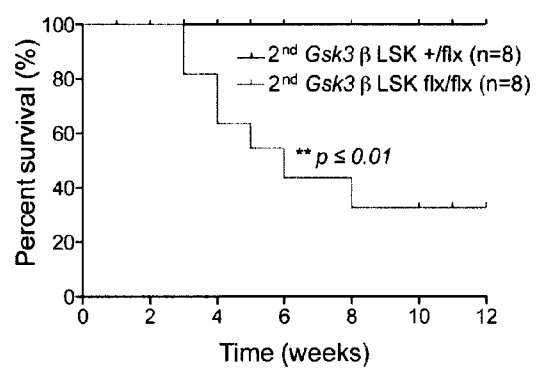


FIGURE 3 (cont)

J

Secondary transplant

**K**

Regenerative capacity

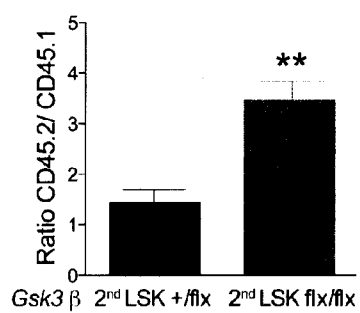
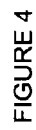


FIGURE 3 (cont)

A



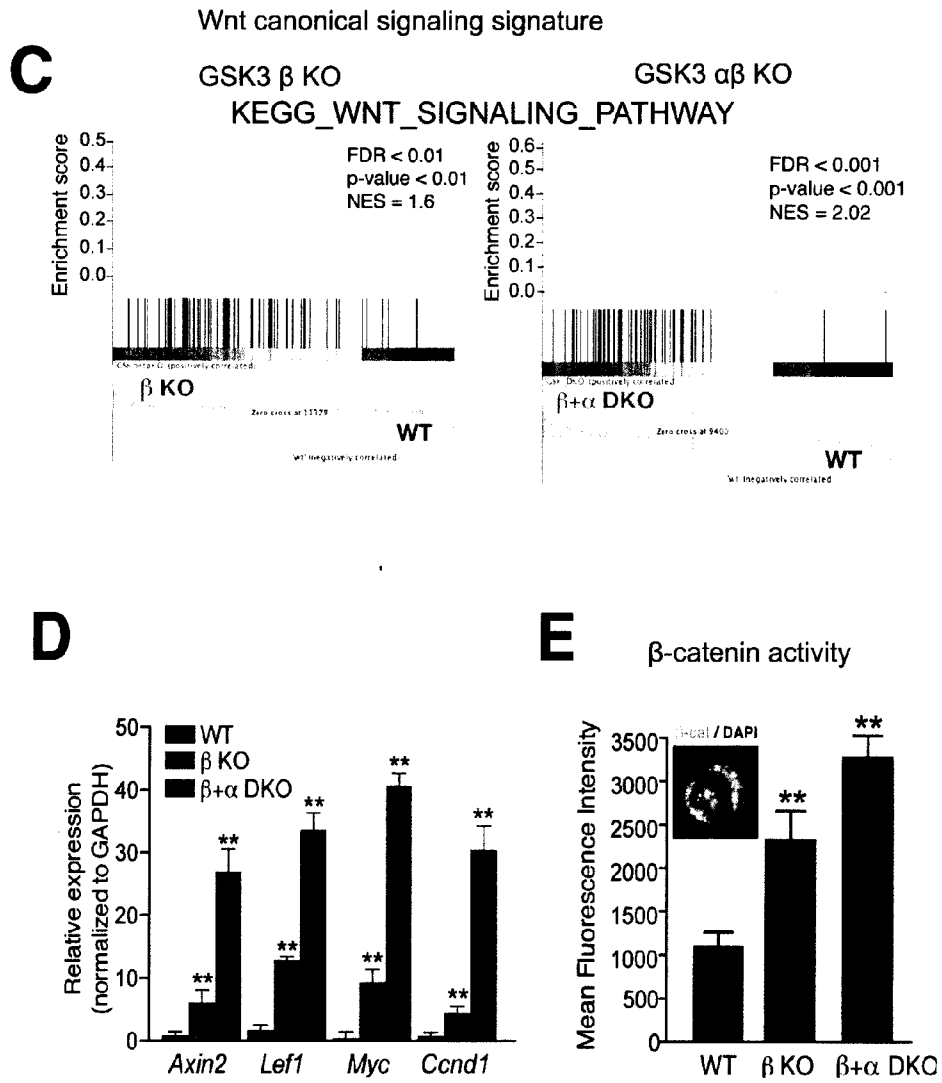


FIGURE 4 (cont)

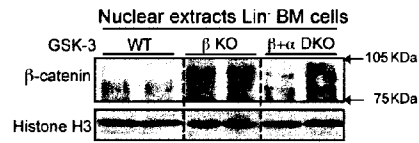
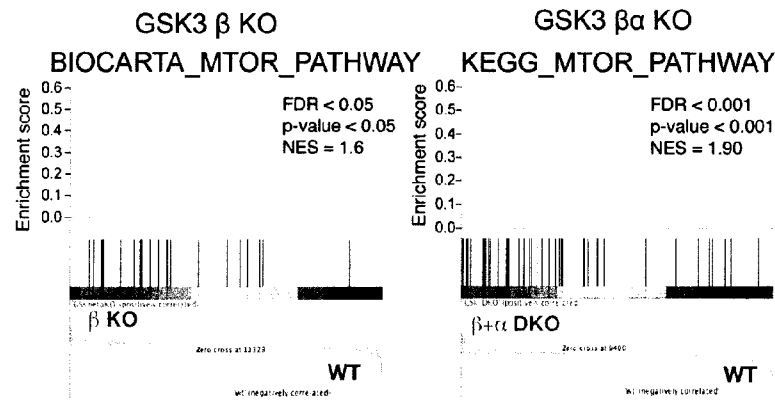
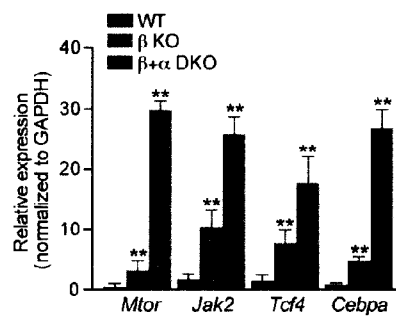
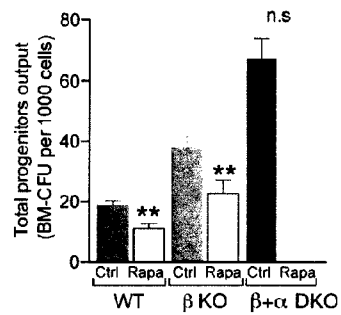
F **β -catenin stabilization****G****mTOR signaling signature****H****I****mTOR inhibition**

FIGURE 4 (cont)

Transcriptional dysregulation

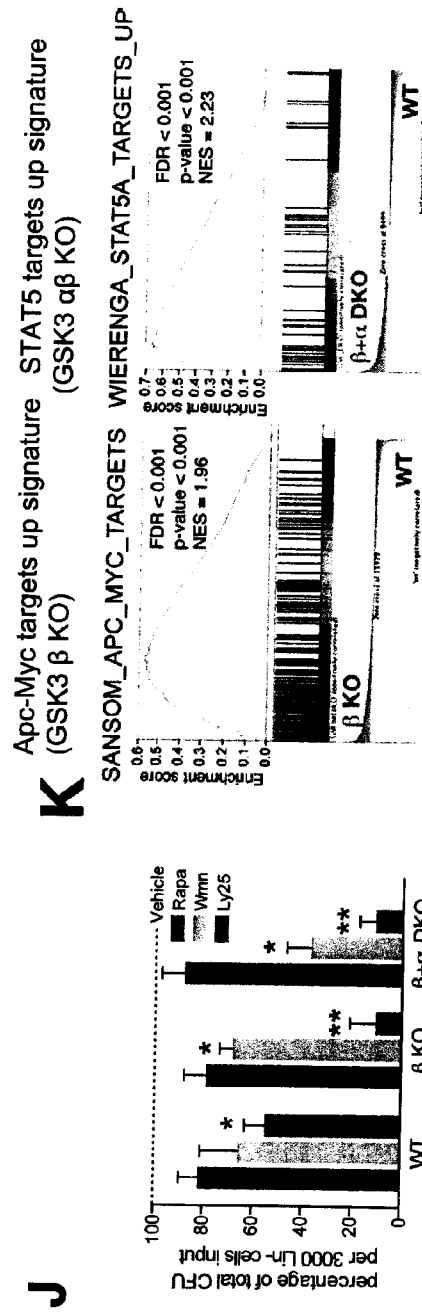


FIGURE 4 (cont)

AKT/P13K signaling signature
GSK3 β KO

L

REACTOME_P13K_AKT_ACTIVATION

Enrichment score
0.6
0.5
0.4
0.3
0.2
0.1
0.0

Enrichment score
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0

FDR < 0.001
p-value < 0.001
NES = 2.05

FDR < 0.05
p-value < 0.01
NES = 1.62



M

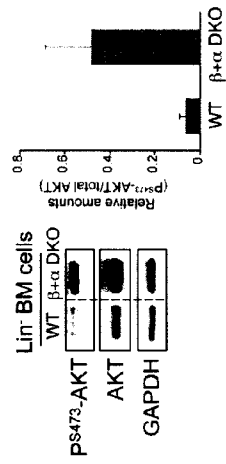
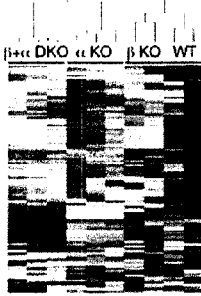
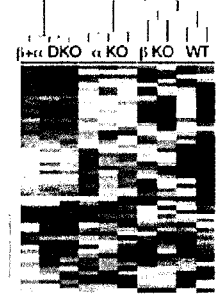
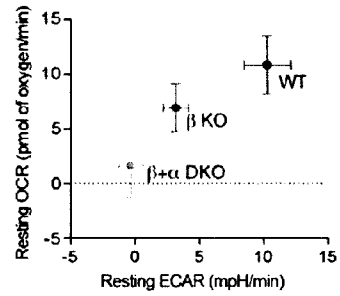


FIGURE 4 (cont)

NKEGG_GLYCOLYSIS
GLUCONEOGENESISKEGG_OXIDATIVE
PHOSPHORYLATION**O**

Metabolic homeostasis

**P**

Mitochondrial stress

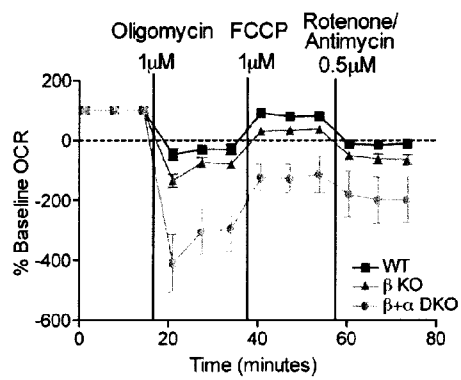
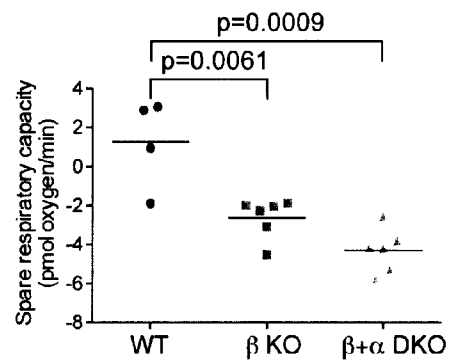
**Q**

FIGURE 4 (cont)

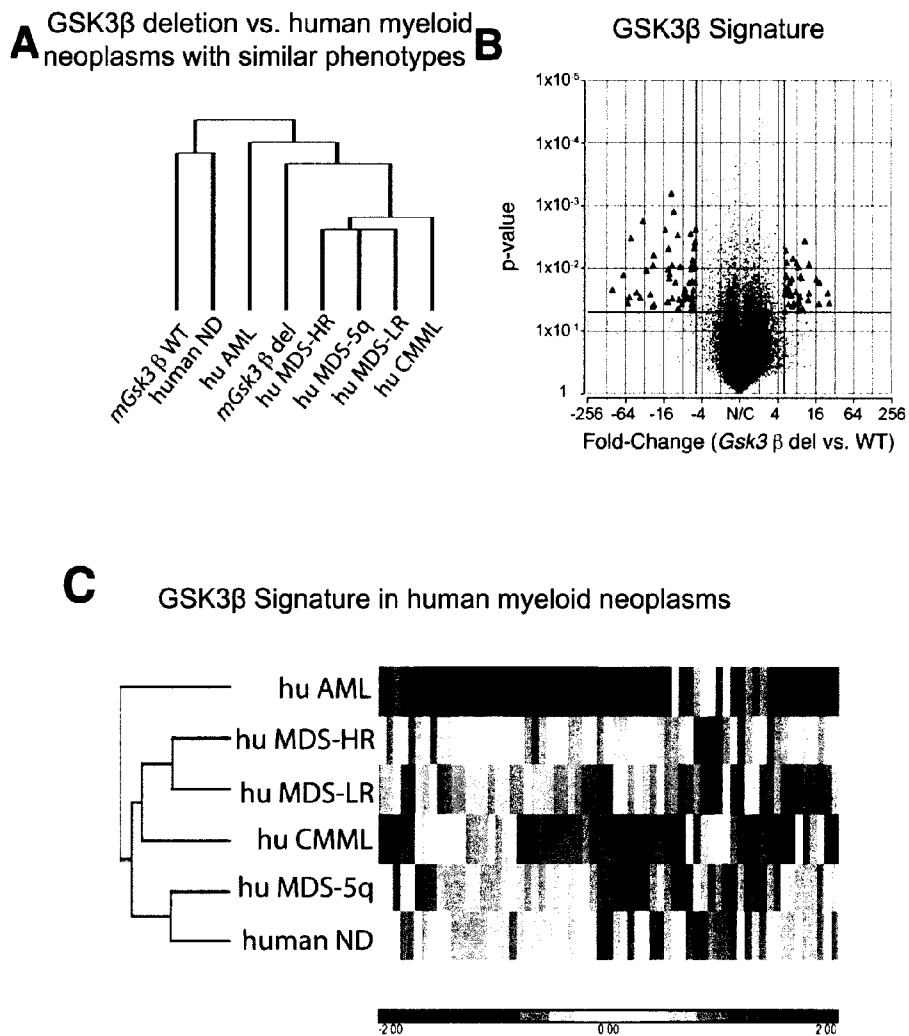


FIGURE 5

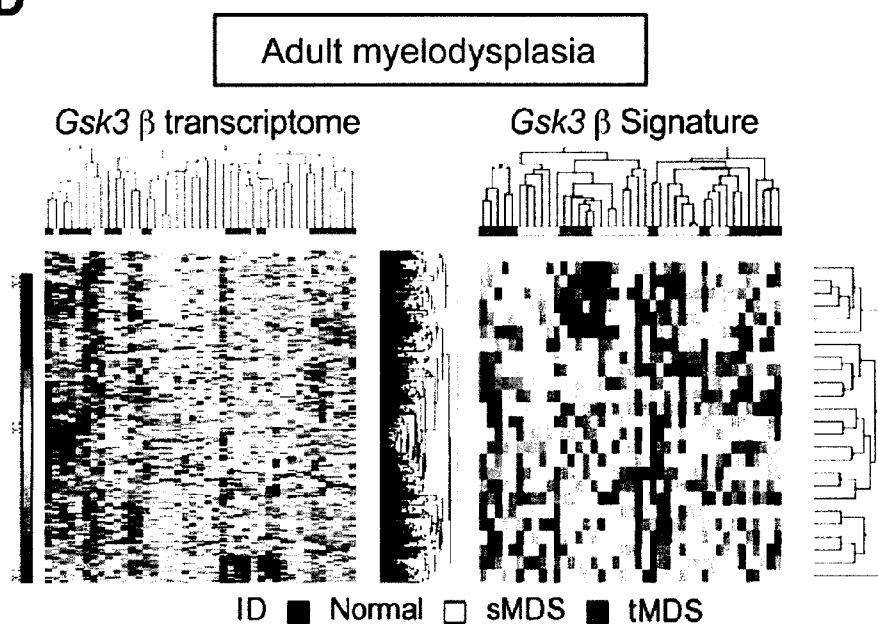
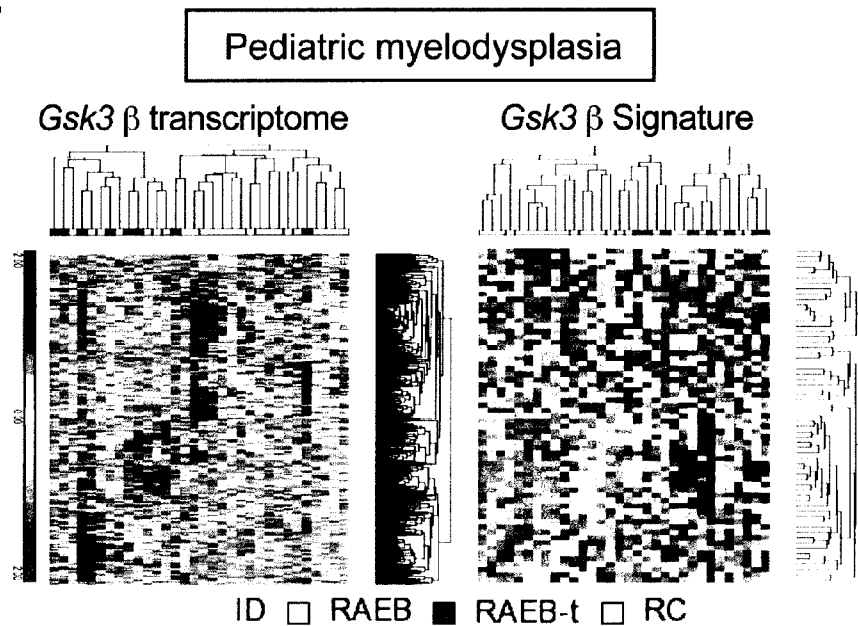
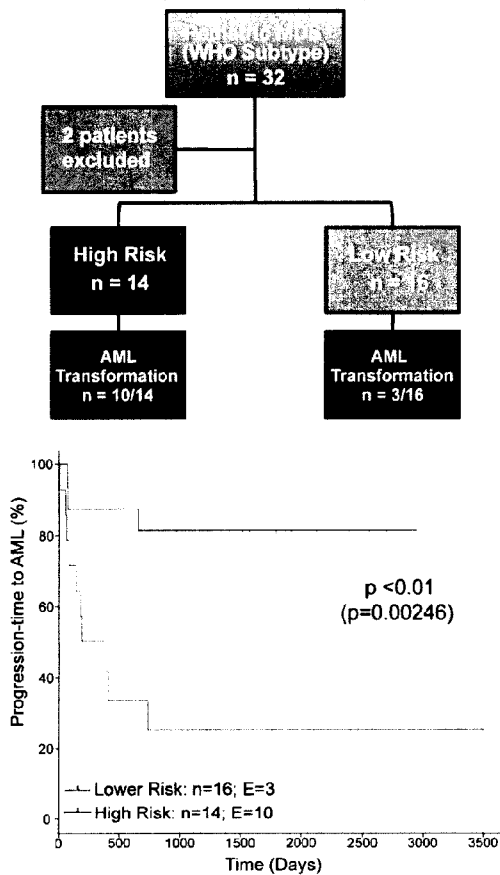
D**E**

FIGURE 5 (cont)

F

AML Transformation Risk (WHO criteria)

**G**

AML Transformation Risk (Gsk3 β signature criteria)

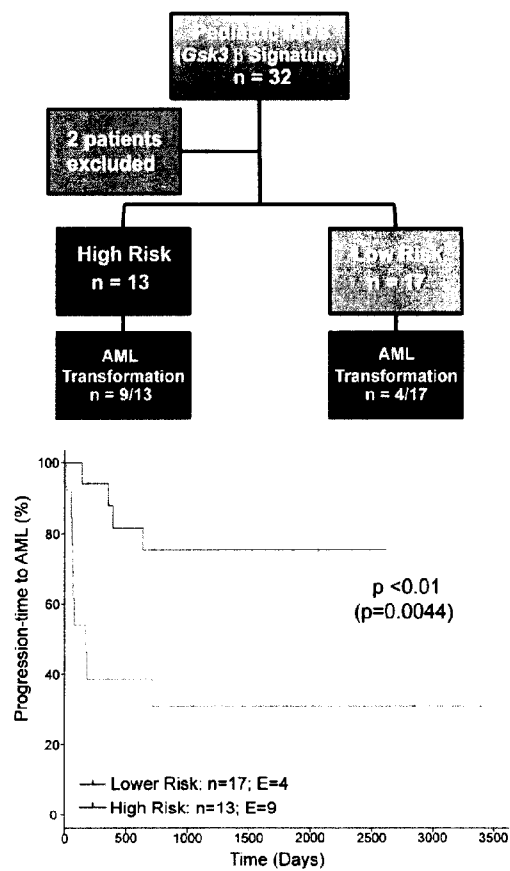


FIGURE 5 (cont)

H

AML Transformation Risk (WHO/*Gsk3* β signature criteria)

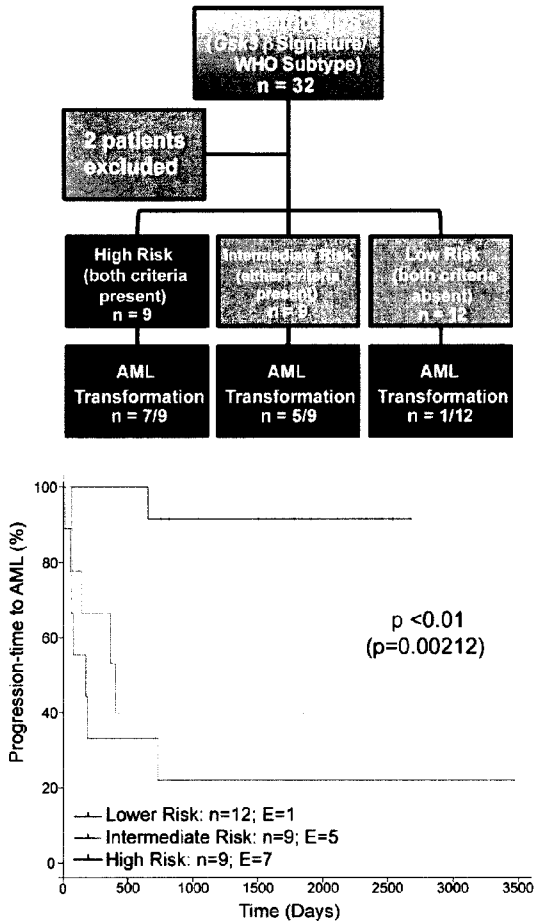


FIGURE 5 (cont)

Model of disease progression

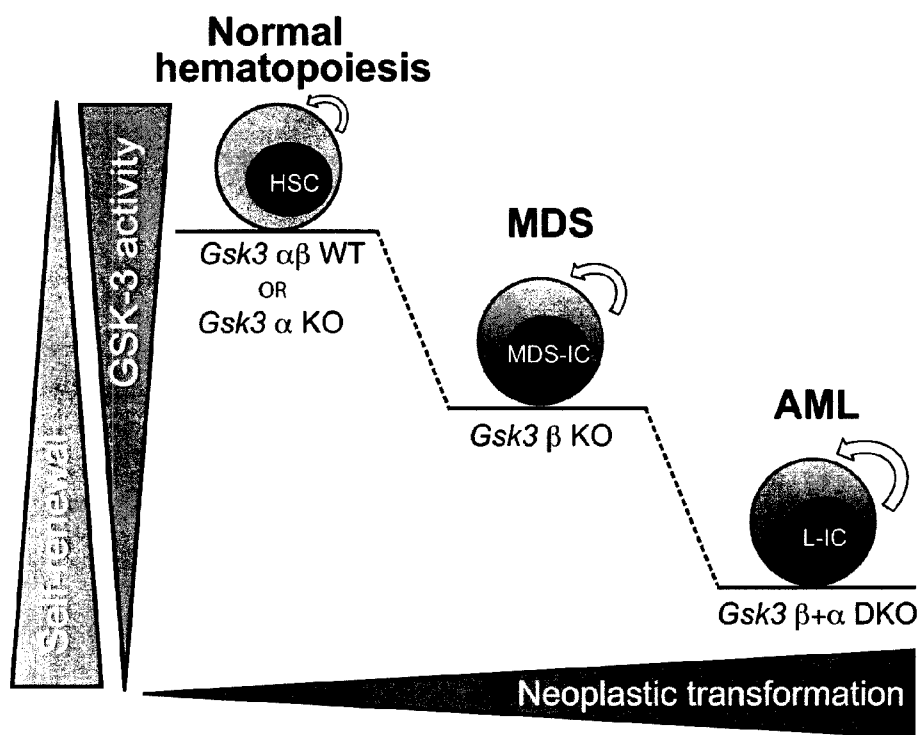


FIGURE 6

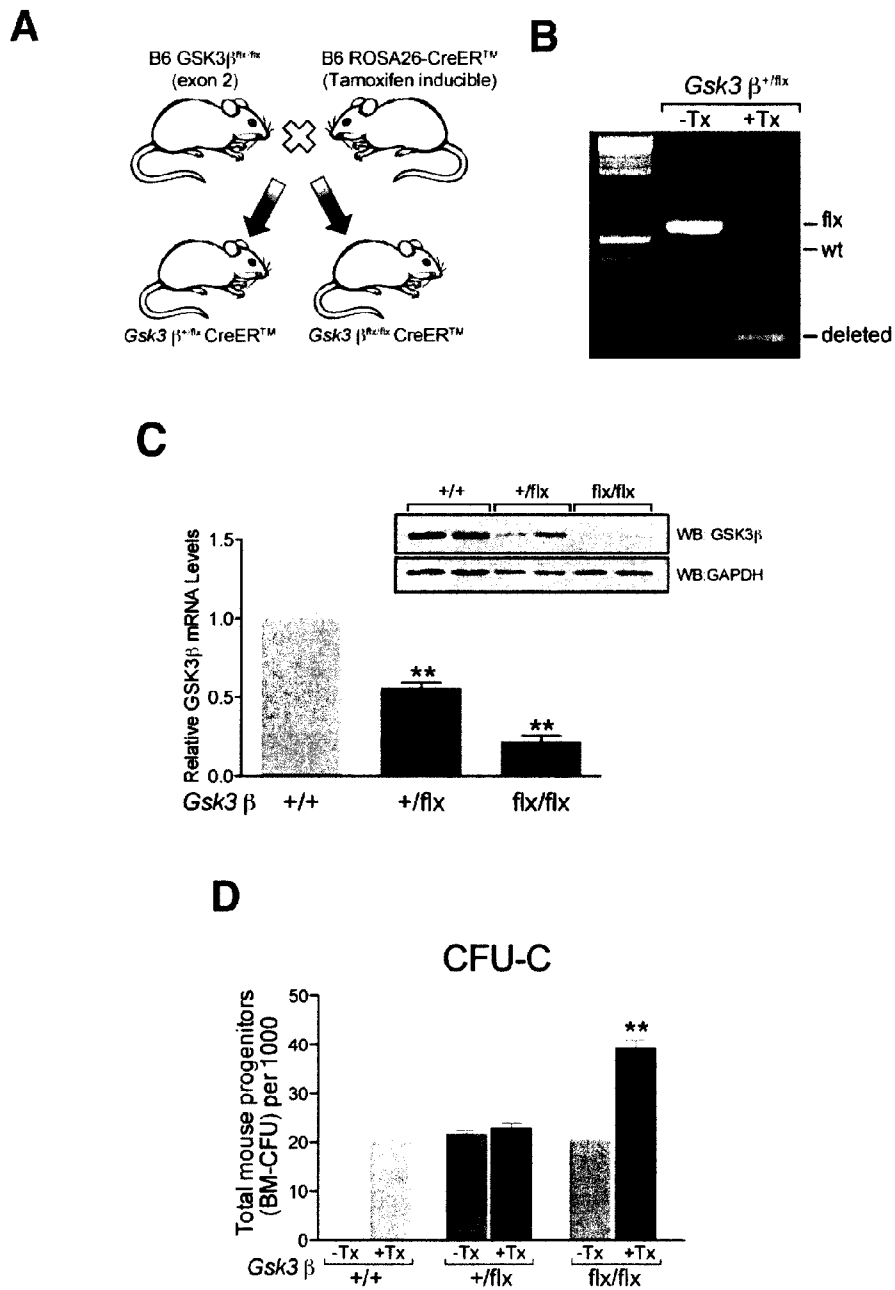


FIGURE 7

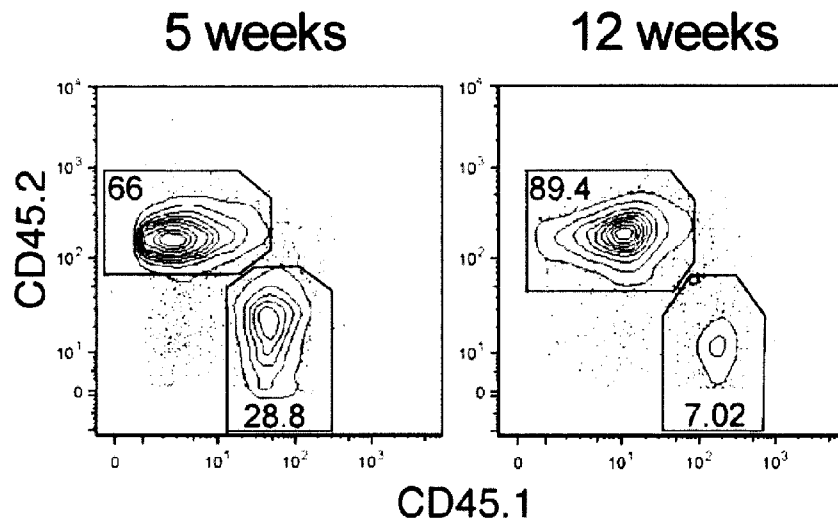
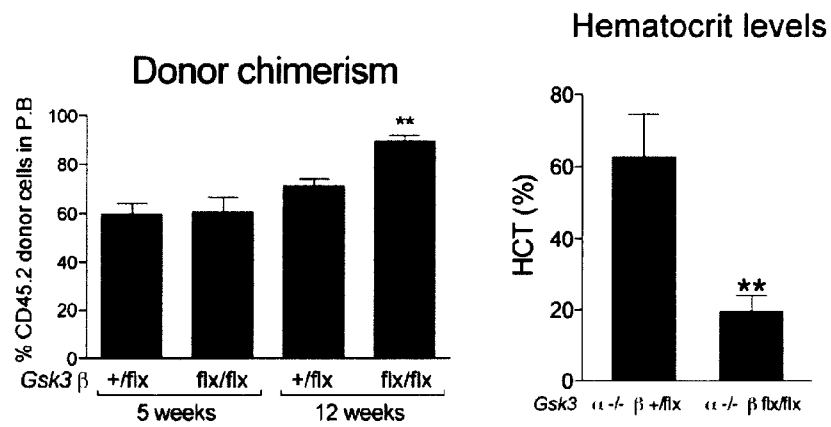
E**G****F**

FIGURE 7 (cont)

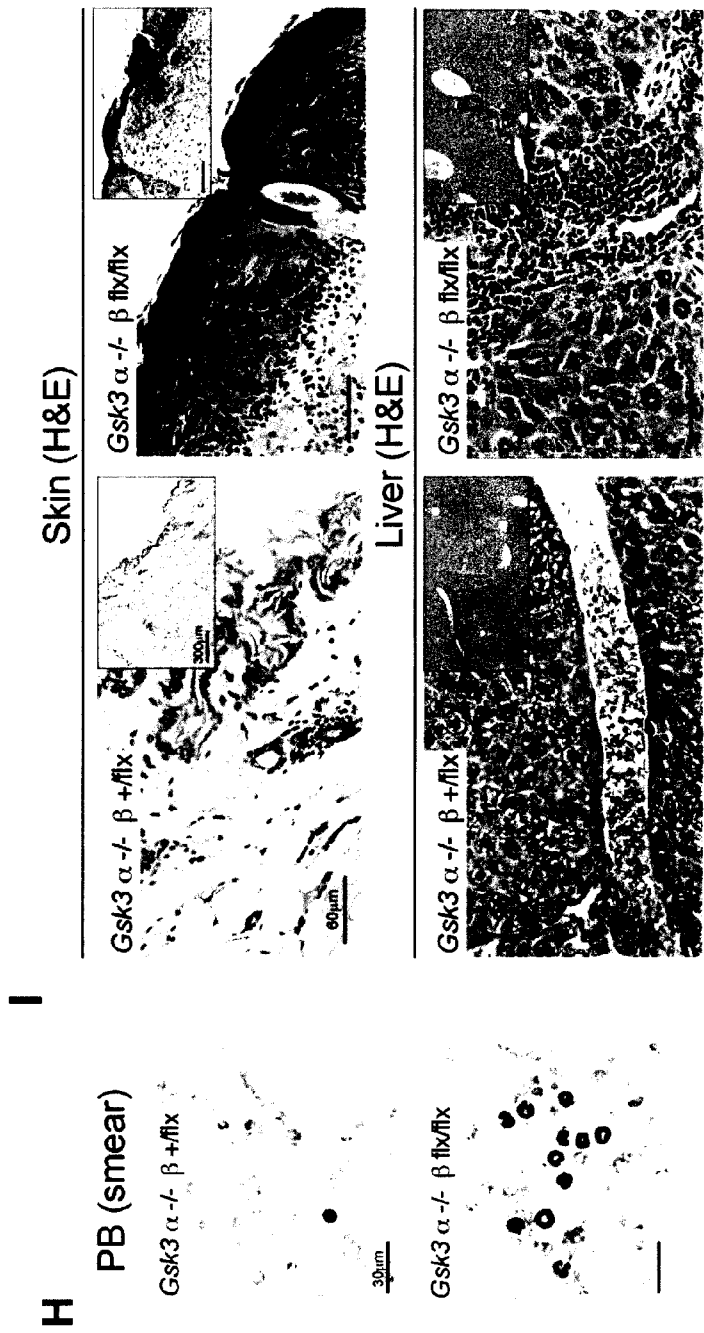


FIGURE 7 (cont)

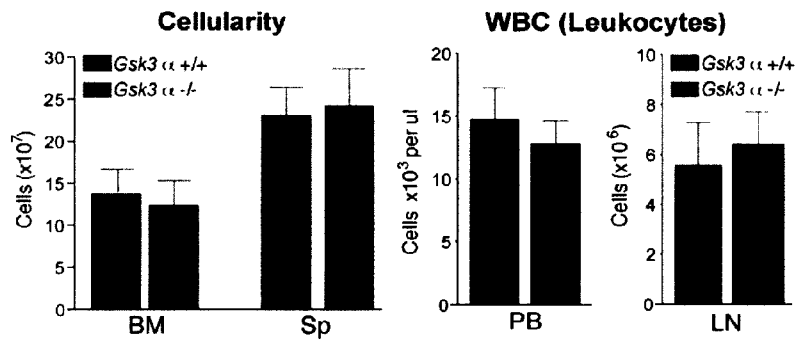
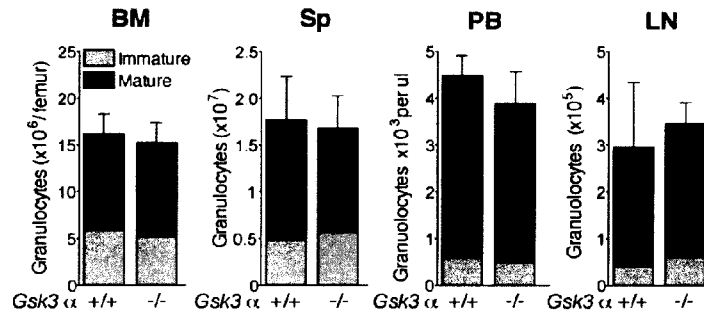
J**K**

FIGURE 7 (cont)

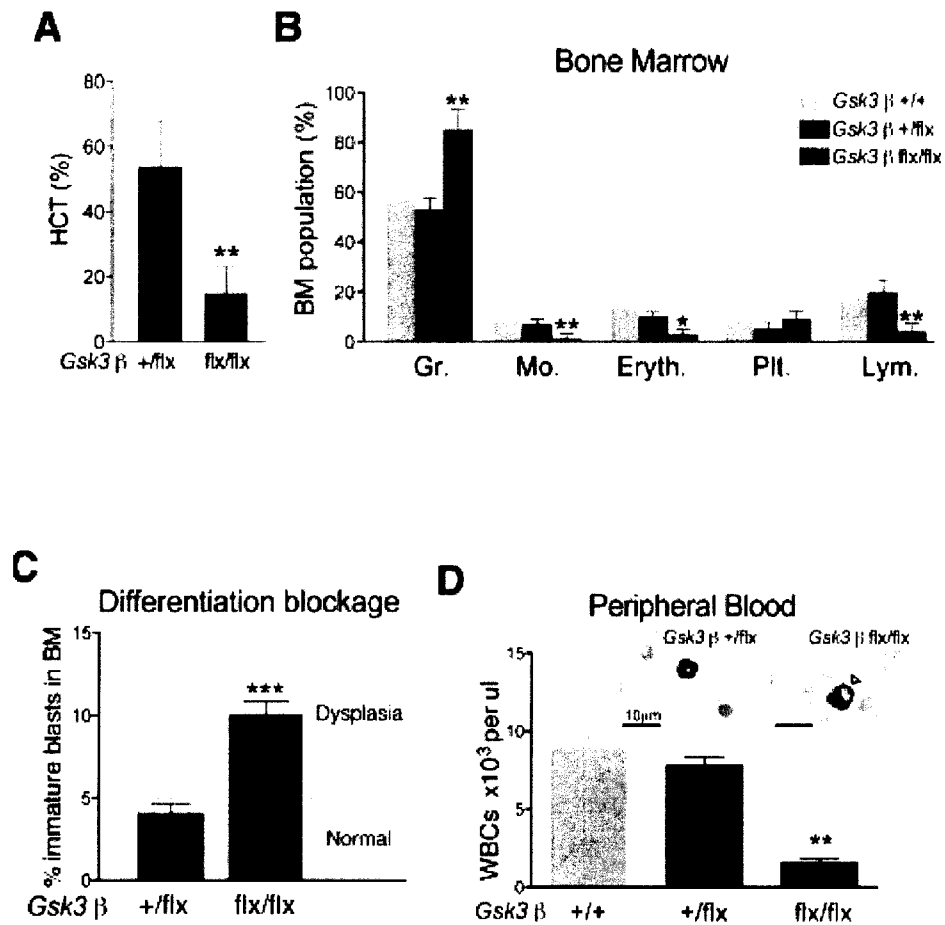


FIGURE 8

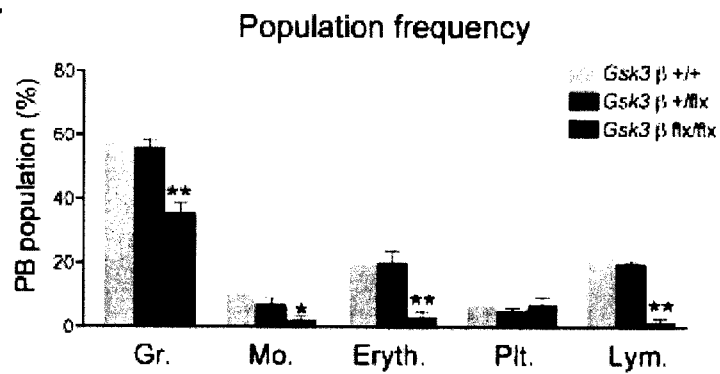
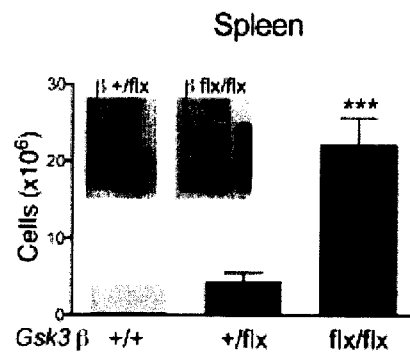
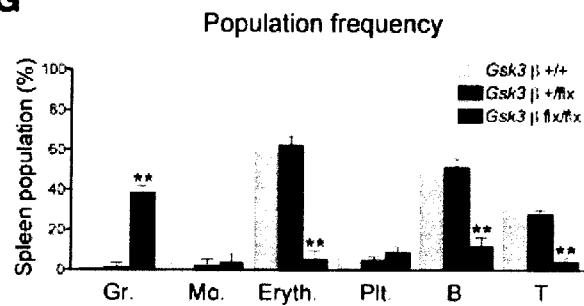
E**F****G**

FIGURE 8 (cont)

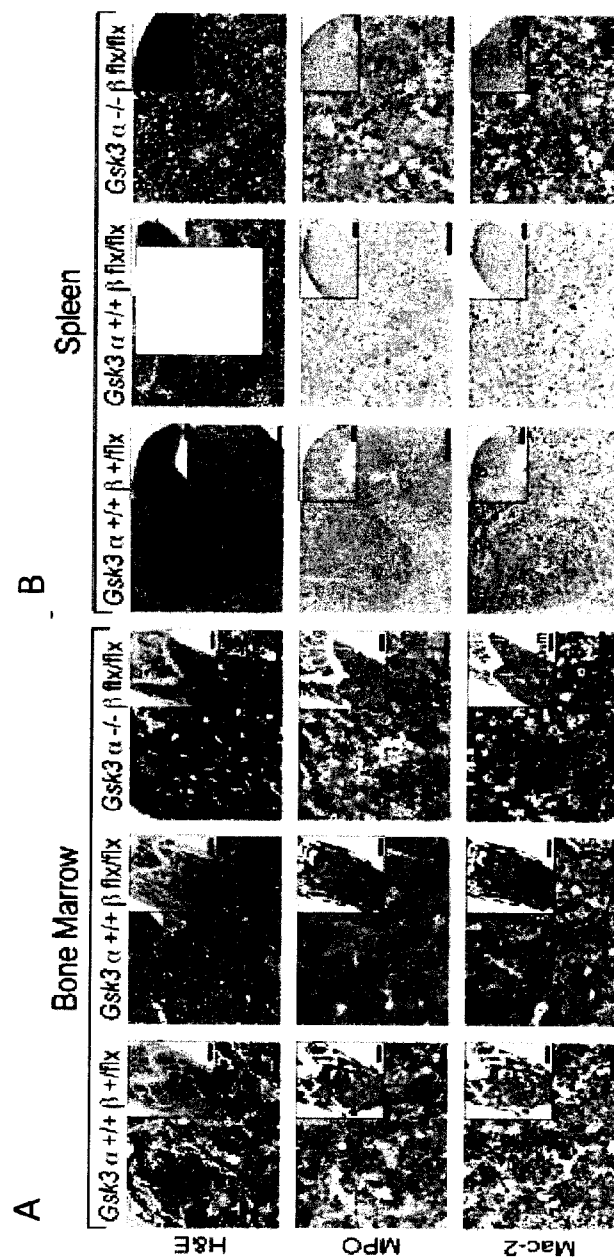


FIGURE 9

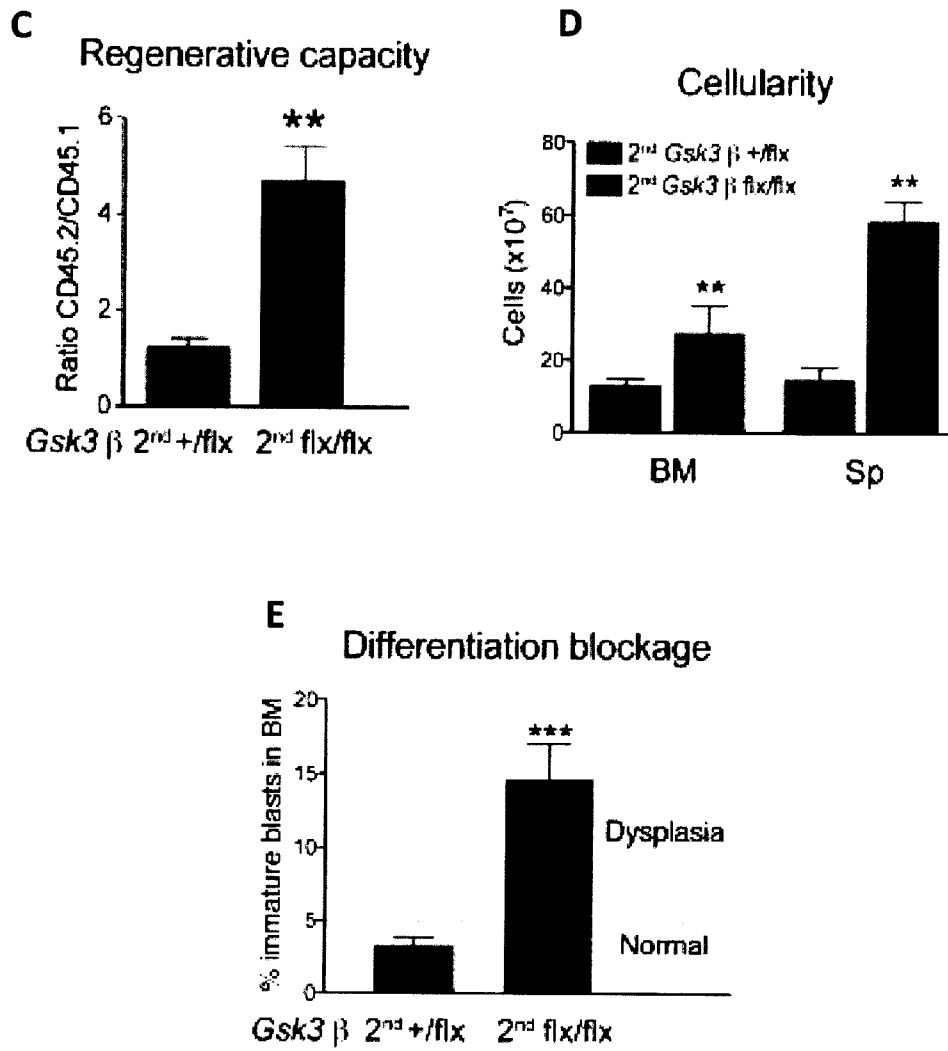


FIGURE 9 (cont)

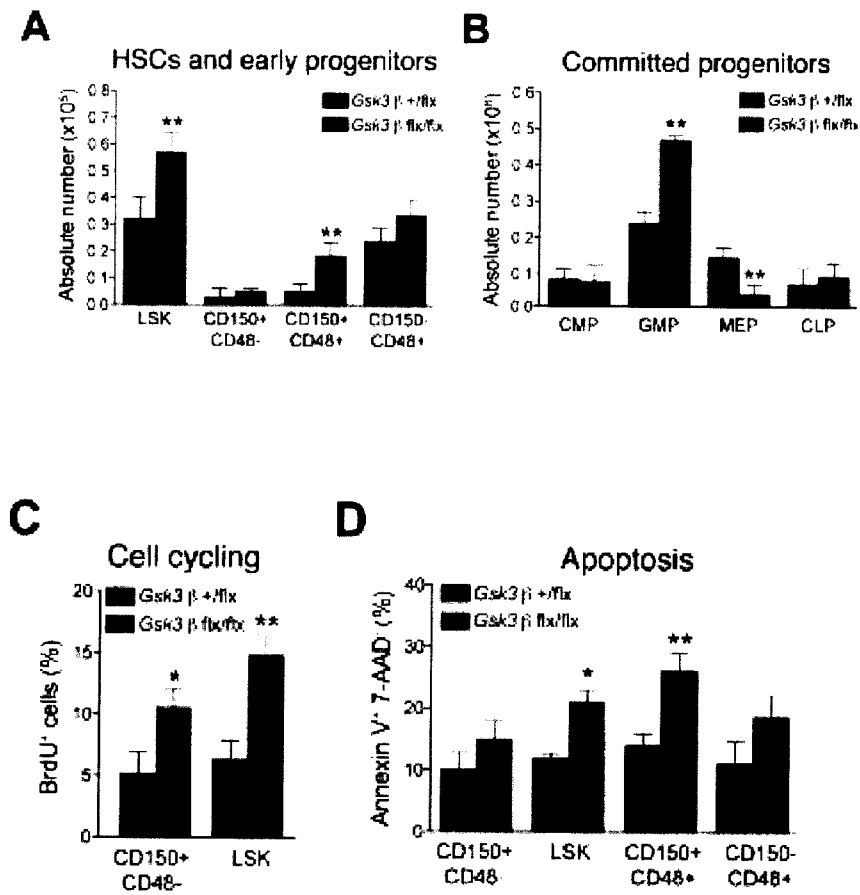


FIGURE 10