

Generation Of Human CBFA2T3-GLIS2 Acute Myeloid Leukemia
By Lentivirus-Mediated Overexpression

Luc Boulianne

Department of Pathology

McGill University, Montreal

August 2021

A thesis submitted to McGill University in partial fulfilment of the
requirements of the degree of Master of Science

© Luc Boulianne 2021

Table of Contents

TABLE OF CONTENTS	1
ABSTRACT	2
INTRODUCTION.....	4
THE HEMATOPOIETIC STEM AND PROGENITOR CELLS AND THEIR PRESENCE IN CORD BLOOD	4
EPIDEMIOLOGY AND MANIFESTATIONS OF ACUTE MYELOID LEUKEMIA	6
FUSION GENES AMID THE AETIOLOGY OF ACUTE MYELOID LEUKEMIA	7
ACUTE MEGAKARYOBLASTIC LEUKEMIA	9
CBFA2T3-GLIS2 ACUTE MEGAKARYOBLASTIC LEUKEMIA	11
CBFA2T3-GLIS2 PERTURBS TRANSCRIPTIONAL ACTIVITY	13
NCAM1 (CD56) UPREGULATION IN CBFA2T3-GLIS2 ACUTE MYELOID LEUKEMIA	14
STUDY MODELS OF CBFA2T3-GLIS2 ACUTE MYELOID LEUKEMIA	15
MATERIALS AND METHODS.....	16
LENTIVIRUS PRODUCTION OF CBFA2T3-GLIS2	16
ISOLATION OF CORD BLOOD HEMATOPOIETIC STEM/PROGENITOR CELLS (CD34+LIN-) BY MAGNETIC POSITIVE SELECTION	17
LENTIVIRAL TRANSDUCTION OF CORD BLOOD HEMATOPOIETIC STEM AND PROGENITOR CELLS	17
DETECTION OF THE FUSION GENE TRANSCRIPT	18
INJECTION OF CBFA2T3-GLIS2-TRANSDUCED CELLS IN IMMUNODEFICIENT MICE	18
SACRIFICE OF MICE UPON LEUKEMIA MANIFESTATIONS	19
FLOW CYTOMETRY ANALYSIS	20
IMMUNOFLUORESCENCE IMAGING OF THE CHIMERIC PROTEIN	20
RNA SEQUENCING	21
COMPARATIVE GENOMIC HYBRIDIZATION/SINGLE-NUCLEOTIDE POLYMORPHISM ANALYSIS	22
RESULTS	22
OVEREXPRESSION OF CBFA2T3-GLIS2 IN HEMATOPOIETIC STEM AND PROGENITOR CELLS	22
CBFA2T3-GLIS2-TRANSDUCED HSPCs EXHIBITED EXTENSIVE PROLIFERATION <i>IN VITRO</i>	24
SYNTHETIC CBFA2T3-GLIS2 LEUKEMIA GENERATED IN IMMUNODEFICIENT MICE.....	27
PHENOTYPES OF SYNTHETIC CBFA2T3-GLIS2 LEUKEMIA REMAINED CONSTANT THROUGH RECIPIENT HOSTS	31
RNA EXPRESSION PROFILES FROM CBFA2T3-GLIS2 XENOGRAFTS CORRELATE WITH PEDIATRIC MEGAKARYOCYTIC LEUKEMIA	35
SYNTHETIC TUMOURS OF CBFA2T3-GLIS2 LEUKEMIA HARBORED FEW CHROMOSOMAL ABNORMALITIES.....	38
BLAST CELLS REGENERATED LEUKEMIA AFTER A SHORT PERIOD OF CELL CULTURE	40
REPOPULATING CELLS REGENERATING LEUKEMIA WERE PRESENT IN A HIGH FREQUENCY	43
DISCUSSION	45
CONCLUSION	55
ACKNOWLEDGEMENTS	56
AUTHOR'S CONTRIBUTION	57
REFERENCES.....	57
SUPPLEMENTARY DATA	74
CELL COUNTS OF BLAST CELLS IN CULTURE FOR SIX DAYS REGENERATING LEUKEMIA	74

Abstract

Pediatric acute myeloid leukemia (AML) is a genetically heterogeneous disease distinct from the adult with few chromosomal alterations and often driven by a single potent oncogene such as fusion genes. Acute megakaryoblastic leukemia (AMKL), a rare subtype of AML that accounts for 10% of cases, is associated with an adverse prognosis. Initially reported in AMKL patients, the fusion gene CBFA2T3-GLIS2 exhibits a dismal probability of overall survival as low as 14%. The chimeric protein results from the inversion of chromosome 16, thereby fusing the neryv homology region domains of CBFA2T3 to the zinc-finger domains of GLIS2. The paucity of patient samples is considerably limiting molecular and functional studies. This study reports the generation of a human CBFA2T3-GLIS2 AML model by lentiviral transduction using cord blood hematopoietic stem and progenitor cells. Indeed, blast cells harbouring the chimeric protein proliferated in immunodeficient mice and recapitulated the human disease phenotypically and molecularly. Half the primary tumours exhibited megakaryocytic surface markers in addition to the biomarker CD56 (NCAM1). Moreover, the disease could be serially regenerated in recipient animals while maintaining the original phenotype, highlighting the self-renewal capacity of leukemia-initiating cells. The tumours had very few chromosomal abnormalities and a transcriptional activity strongly expressing megakaryocytic genes. Above all, blast cells can be cultured for six days while maintaining the ability to regenerate the disease, thus paving the way to large-scale studies requiring *in vitro* phases.

La leucémie myéloïde aiguë (LMA) pédiatrique est une maladie génétiquement hétérogène distincte de chez l'adulte avec peu d'altérations chromosomiques et souvent médiée par un oncogène unique tel que les gènes de fusion. La leucémie mégakaryoblastique aiguë (LMKA), un sous-type rare de LMA qui compte pour 10% des cas, est associée à un pronostic adverse. Initialement rapporté chez des patients LMKA, le gène de fusion CBFA2T3-GLIS2 démontre une probabilité de survie globale lugubre aussi basse que 14%. La protéine chimérique résulte de l'inversion du chromosome 16 fusionnant ainsi les domaines région d'homologie nerveuse de CBFA2T3 aux domaines doigt de zinc de GLIS2. Le manque d'échantillons de patients limite considérablement les études moléculaires et fonctionnelles. Basé sur des études précédentes, il y est ici proposé que seule la surexpression du gène de fusion est suffisante afin de générer la maladie synthétiquement. Par conséquent, cette étude rapporte la génération d'un modèle humain de LMA CBFA2T3-GLIS2 par transduction lentivirale en utilisant des cellules souches et progénitrices de sang de cordon. En effet, des cellules blastiques exprimant la protéine chimérique ont proliféré dans des souris immunodéficientes, récapitulant la maladie humaine au niveau phénotypique et moléculaire. La moitié des tumeurs primaires manifestaient des marqueurs de surface mégakaryocytaires en plus du biomarqueur CD56 (NCAM1). De plus, la maladie a pu être régénérée lors de transplantations sériées dans des animaux immunodéficients tout en maintenant le phénotype original, démontrant la capacité d'autorenouvellement des cellules initiatrices de la leucémie. Les tumeurs avaient peu d'anomalies chromosomiques et avaient un profil d'expression transcriptionnelle exprimant fortement des gènes mégakaryocytaires. Par-dessus tout, des cellules blastiques peuvent être

cultivées pendant six jours tout en maintenant l'habileté de régénérer la maladie, ouvrant ainsi la voie aux études à grande échelle nécessitant des phases *in vitro*.

Introduction

The hematopoietic stem and progenitor cells and their presence in cord blood

Blood has been studied with great interest for centuries because of its all-inclusive distribution throughout the body and the dreadful outcomes of related diseases. In recent decades, researchers have made significant progress in elucidating the hierarchy defining the hematopoietic system within the bone marrow, notably by using xenograft transplantation into immunodeficient mice following sublethal irradiation (1). Serving as a common ancestor for all blood cell types, the hematopoietic stem cell (HSC) is at the apex of hematopoiesis. This ability to give rise to every cell encountered in the blood is called multipotency. Another striking characteristic is that they are able to self-renew, thereby allowing the maintenance of a pool of HSCs (1-3). HSCs mostly remain in a quiescent state and only divide on average every 157-350 days (1). However, this dormant state is rapidly reversible in hematopoietic stress to replenish all blood lineages (4). Symmetrical self-renewal division occurs when an HSC divides into two daughter cells having stemness properties. On the other hand, asymmetrical self-renewal division generates an HSC and a progenitor cell engaged in differentiation of any cell lineage (5). This heterogeneous population of hematopoietic stem and progenitor cells (HSPC) bears the surface antigen CD34, expressed on less than five percent of all blood cells (1, 6). The optimal conditions for the *in vitro* expansion of HSCs have proven to be challenging to identify and achieve (7). Hence, transplantation of HSPCs into immunodeficient mice is required to provide a supportive niche that resembles the environment normally

encountered in the human bone marrow (2). The rarity of HSCs is a major hurdle to their study in animals and their use in transplantation therapy, as only one per million cells in the human bone marrow is a transplantable HSC (8). Furthermore, as little as 30% of patients who require an allogeneic bone marrow transplantation therapy will have a human leukocyte antigen-matched sibling donor (9). Therefore, it is essential to identify a reliable and abundant source of HSPCs.

Aside from bone marrow and mobilized peripheral blood, cord blood (CB) is an excellent source of HSPCs as it is greatly available and easily banked for potential future use (10). Indeed, CB units can be cryopreserved for more than 20 years with efficient recovery of cells, thus increasing the diversity of matching donors. CB HSPCs have shown an extensive proliferation and engrafting capacity that exceeds those isolated from the bone marrow (9). Although the number of cells is lower than in the other sources, CB holds a greater proportion of long-lasting repopulating HSPCs increasing the probability of a successful transplant over time (11). Moreover, CB transplants showed a lower risk of developing graft-versus-host disease in both matched and unmatched recipients (12-14). Graft-versus-host disease is a serious adverse outcome ending in life-threatening tissue damage because of immune graft cells activated by the host's antigens. CB transplantation is executed to replenish a recipient's bone marrow in healthy HSPCs following myeloablative chemotherapy to treat hematological malignancies such as acute myeloid leukemia (AML). As with other types of acute blood cancer, AML is characterized by blast cells within the bone marrow in a proportion of at least 20%. This procedure is nowadays a well-established therapeutic strategy, especially for pediatric patients for which the required cell dose based on body weight is more readily achievable.

Epidemiology and manifestations of acute myeloid leukemia

While leukemia is the most diagnosed pediatric cancer, AML is less common and accounts for 2-5% of pediatric cancer incidence and 15-20% of pediatric leukemia (15, 16). The incidence of pediatric AML in the United States was estimated at 7.7 cases per million children aged up to 14 years between 2005 and 2009 (17). The highest incidence is found in infancy, with 1.6 cases per 100,000 infants, similar to the incidence of adults older than 40 years of age (18). AML incidence rates are higher in the first two years and decrease to reach a basal line at nine years of age. Afterwards, the numbers increase during adolescence and young adulthood and maintain this upward trend (19). Data collected throughout the last decades show a rise in pediatric cases (17, 18). AML survival rates average 60-70% in children and young adults but are lower in specific leukemia subgroups due to elevated relapse rates (20, 21).

AML is a hematological malignancy characterized by the clonal expansion of undifferentiated myeloid precursors that have acquired genetic abnormalities, thus impairing hematopoiesis, resulting in bone marrow failure. It is estimated that about 25% of all AML cases originate following exposure to cytotoxic agents or arise in the context of a pre-existing myelodysplastic or myeloproliferative syndrome (22). AML presentation includes a broad range of unspecific symptoms and clinical manifestations correlating to bone marrow failure or direct infiltration of blast cells. Fatigue, dyspnea, headache and chest pain may all result from anemia. Patients may have easy bruising or bleeding as a result of thrombocytopenia (23). Hemorrhage and coagulopathy are common complications, and up to 60% of cases may present with some form of bleeding (24).

Leukopenia may be accompanied by immune cell dysfunction and is evidenced by poorly healing skin wounds and recurrent infections. Blast cell infiltration may cause rash-like lesions, nodules and chloromas also called myeloid sarcomas (25). Similarly, presentation of pediatric acute leukemia includes fever, pallor, bleeding (petechia/purpura), bone pain, lymphadenopathy, gingival hyperplasia, chloromas and hepatosplenomegaly/splenomegaly (26).

Fusion genes amid the aetiology of acute myeloid leukemia

AML is a heterogeneous disease with the appearance of multitudinous driver mutations affecting cell components involved in self-renewal, proliferation and differentiation. Indeed, 5234 driver mutations involving 76 genes or regions have been identified, making the disease molecularly and clinically complex (27). AML has many associated genes with a low somatic mutation frequency compared to other cancer types (28). Similarly, pediatric AML is characterized by a low burden of genomic alterations, with about one-third of cases not presenting any copy number alterations (29). It was shown that the genetic lesions and their interactions are age-specific, thereby forging the concept that AML is an ontogeny-specific disease (30). Indeed, certain groups of genes are preferentially mutated in children or adults, with fusion genes being more frequent in children (31). The fusion genes, resulting from chromosomal rearrangements, produce in-frame chimeric proteins that provide blast cells with cancerous properties. As it is observed with MLL (KMT2A) rearrangements involving multiple translocation partners, fusion genes may give rise to leukemia of two different lineages (32). Nowadays, it is evident that AML classification into distinct molecular subgroups must incorporate

cytogenetics and molecular genetics in addition to cell morphology and immunophenotyping (33).

Over the years, a strong correlation has been established between recurrent fusion genes and given tumours (34), hence improving high-risk patients' stratification concerning treatment and outcome (35). Indeed, gene expression profiling of AML was shown to accurately identify the molecular subtype irrespective of the lineage of blast cells (36). Besides complementary genetic lesions, chromosomal rearrangements are considered critical events in pediatric AML (29) and involve transcription factors and epigenetic regulators. Whether the rearrangement directly includes an epigenetic regulator or not, fusion genes were demonstrated to be potent inducers of leukemogenesis through epigenetic mechanisms (37). The involvement of an epigenetic modifier is well exemplified by MLL-rearrangement leukemia that encodes for a histone 3 lysine 4 (H3K4) methyltransferase whose action induces chromatin modifications leading to transcriptional activation (32). MLL was found to be essential to sustain self-renewal in normal HSCs, thus playing a central role in cancer biology (38, 39). Moreover, MLL fusion genes caused aberrant chromatin expression profiles in regions that encode HSCs regulators and transcription factors involved in hematopoiesis and leukemia (40). MLL fusion genes induce the overexpression of the homeobox (*HOX*) cluster genes and *MEIS1*, which are deeply involved in regulating hematopoiesis through HSCs proliferation and differentiation and aberrantly expressed in AML (41, 42). Indeed, HOX genes are highly expressed in primitive HSPCs while being almost absent in differentiated CD34-cells (43). Alongside epigenetic dysregulation, fusion genes may have different mechanisms of action leading to leukemia either by directly interacting with DNA or

through protein complexes. The fusion gene BCR-ABL1 features in chronic myeloid leukemia (CML) resulting from the translocation of chromosomes 9 and 22, thereby creating a non-receptor tyrosine kinase capable of autophosphorylation due to the loss of ABL1 upstream control elements (44). Although the mechanisms whereby leukemia arises in CML are not fully elucidated, the constitutive activity of BCR-ABL1 in HSPCs results in uncontrolled proliferation, loss of differentiation and loss of response to control mechanisms and typically resembles AML (45). Interestingly, dysregulation caused by BCR-ABL1 incorporates an epigenetic component as well (46). Altogether, fusion genes in AML dysregulate expression patterns in genes primarily involved in HSC functions favouring leukemogenesis, notably in self-renewal.

Acute megakaryoblastic leukemia

Acute megakaryoblastic leukemia (AMKL) is one of the several cytological subtypes of AML for which the diagnosis is established likewise. The additional characteristic of this subtype is that at least 50% of blast cells must demonstrate immunophenotypic evidence of megakaryocytic differentiation either by immunohistochemistry or flow cytometry (47). The markers of the megakaryocytic lineage used to establish the diagnosis are the platelet glycoproteins CD41 (GPIIb/IIIa), CD61 (GPIIIa), CD42b (GPIb) (48). Establishing the diagnosis is often a difficult task due to extensive fibrosis of the bone marrow secondary to activating fibroblast factors produced by megakaryoblasts. This fibrosis causes inadequate bone marrow aspirates (dry taps) that contribute to the failure of cytogenetic analysis and difficulty in determining the accurate proportion of blast cells by microscopy. Touch preparations from bone marrow

biopsy may be performed to obtain suitable cell counting and immunohistochemistry staining (49). AMKL is a rare disease that accounts for only 1% of adults and 8-14% of children with newly diagnosed AML (50). Patients also have a bimodal age distribution with peaks in children and adults over 50 years old (51). AMKL has long been recognized with worse outcomes than the other AML subtypes for both groups of age (52). Similar to what is observed in AML, pediatric and adult AMKL are two distinct diseases with different genetic landscapes (53). Adult AMKL mainly arises in the context of evolution from myeloproliferative disorders (54, 55), while fusion genes are frequently found in pediatric AMKL, acting as powerful driver mutations (56). Of note, pediatric AMKL cases showed higher copy-number abnormalities than AML patients (57). Another genomic feature of AMKL is the high frequency of activating kinase mutations in KIT, FLT3, JAK2 and JAK3 (58, 59). Hitherto, many fusion genes have been reported in pediatric AMKL: NUP98 rearrangements (NUP98r), MLLr, HOX rearrangements, RBM15-MKL1 and CBFA2T3-GLIS2. These fusion genes are mutually exclusive and account for about 60% of cases (60). The NUP98r and CBFA2T3-GLIS2 fusion genes are cryptic and thus not detected by conventional cytogenetic techniques and are associated with dreadful 5-year overall survival rates of 22-36% and 14-28%, respectively (57, 60). Children with Down Syndrome (DS; trisomy 21) have a very high risk of developing leukemia (61), particularly AMKL (57). DS-AMKL is believed to be initiated by a GATA1 mutation and further driven by accumulated mutations (62). Despite sharing similar phenotypes, DS-AMKL and AMKL have distinct genetic features that may underlie their different clinical outcomes, which are more favourable for DS-AMKL (63). Because of the molecular heterogeneity

observed in pediatric AMKL, risk group stratification is constantly being elaborated to redefine relapse risk and tailor better treatment strategies (64).

CBFA2T3-GLIS2 acute megakaryoblastic leukemia

One of the AMKL fusion genes results from a cryptic inversion of chromosome 16 (p13.3q24.3) that fuses CBFA2T3 to GLIS2. The first section contains the N-terminal myristoylation and homology region domains of CBFA2T3, whereas the following part consists of GLIS2 C-terminal zinc-finger domains that bind DNA (65). CBFA2T3-GLIS2 is seen in about 20% of pediatric AMKL cases and is associated with the poorest prognosis, with overall survival of 14-28% and event-free survival of 8-35% five years following initial diagnosis (60). Extramedullary involvement in the central nervous system was found in one out of 20 cases in a cohort of CBFA2T3-GLIS2-positive patients (66). This finding aligns with another study wherein a patient sample was xenotransplanted into immunodeficient mice that consistently developed hind leg paralysis in serial transplantation. A cancerous mass was found near the spinal cord in some animals, and the flow cytometry analysis confirmed megakaryocytic blast cells. Moreover, the same initial sample generated recipient mice presenting with splenomegaly and striking nodular infiltration visible on the organ. RNA sequencing of blast cells led to the discovery of an upregulation of NCAM1 (CD56), for which the cell surface expression was confirmed by flow cytometry afterwards (67). It was determined by chromatin immunoprecipitation analysis that NCAM1 is regulated and a direct transcriptional target of CBFA2T3-GLIS2 since there was an enrichment of the chimeric protein at the promoter region (67, 68). Several studies demonstrated that the gene expression profile of CBFA2T3-GLIS2

patients clusters them apart from other AMKL subgroups (65, 67, 69). It was found an upregulated expression of genes pertaining to the JAK-STAT, sonic hedgehog, WNT and bone morphogenic protein pathways such as BMP2, BMP4, WNT3, WNT8B, PTCH1 and HHIP (65, 67). A study revealed an activation of the Hippo and TNF pathways, both involved in cancer. There was also an upregulation of several receptor tyrosine kinases as well as two of the TAM family kinases. The former group of tyrosine kinases consisted of ROR1, MET and NTRK1, while TYRO3 and AXL represented the latter family of TAM kinases (69). CBFA2T3-GLIS2-positive patients have a significantly lower burden of somatic mutations when compared with other subgroups of AMKL. The genes commonly mutated are KIT, RAS, FLT3, GATA1 and members of the JAK-STAT pathway (53, 65, 70). A third of CBFA2T3-GLIS2 cases present with a normal karyotype, whereas it represents only 9% in other subgroups of AMKL where complex karyotypes and hyperdiploidy are frequently found (70). Blast cells expressing an AML immunophenotype were reported in CBFA2T3-GLIS2-positive patients, thereby indicating that this fusion gene might not be restricted to the AMKL subtype. Interestingly, CBFA2T3-GLIS2-expressing AML was observed in patients belonging to childhood and adolescence, whereas AMKL was solely in children (66, 71). A study using an inducible mouse model of CBFA2T3-GLIS2 leukemia could show that the immunophenotype observed was AMKL when the fusion gene was expressed in fetal liver cells, whereas it was AML with adult bone marrow cells. Also, survival curves displayed a more aggressive disease in AMKL mice as seen in patients. More importantly, leukemia was not generated without the induction of CBFA2T3-GLIS2, thus strengthening the assumption that the fusion gene is the initial event leading to malignancy from alterations in the transcriptional profile (71).

CBFA2T3-GLIS2 perturbs transcriptional activity

CBFA2T3, also known as ETO2 or MTG16, is part of the myeloid translocation gene family whose members are thought to act as transcriptional co-repressors through interactions with DNA-bound transcription factors. Its expression in normal HSPCs suggested that CBFA2T3 may play a role in the homeostasis between self-renewal and differentiation (72). Indeed, the inactivation of CBFA2T3 reduced the number of HSCs as a result of the loss of self-renewal (73). CBFA2T3 was required for the proliferation of progenitors towards the erythroid-megakaryocytic lineage (74) while also repressing the developmentally primed erythroid genes later lost during erythroid differentiation (75). GATA1 is one of the genes that regulate the erythroid-megakaryocytic differentiation with a marked expression during the terminal maturation of megakaryocytes. It was demonstrated that CBFA2T3 binds to GATA1 to repress its activity in immature megakaryocytes, wherein CBFA2T3 expression is restricted (76). On the other hand, GLIS2 was not detected in normal megakaryoblasts, and its function in HSPCs biology remains unknown (67). Nevertheless, zinc-finger proteins are involved in tumorigenesis, cancer progression and metastasis formation. Therefore, it might be the case in leukemia as well (77). GLIS2 ectopic activity would likely be mediated directly or within protein complexes through DNA interactions.

Regarding the DNA-interacting property of both fusion partners, a mechanism of action of CBFA2T3-GLIS2 was proposed based on observations made in the expression level of the genes GATA1 and ERG (78, 79). It was found that the presence of CBFA2T3-GLIS2 in a cell line downregulated GATA1 while upregulating ERG that was required to

maintain leukemia (78). Associated with poor outcomes in AML, ERG expression is involved in HSPCs maintenance and stemness (80). Alterations in the expression of GATA1, together with overexpression of ERG, led to an increase in the number of immature megakaryocytes showing immortalization (81). Altogether, it seems that CBFA2T3-GLIS2, through interactions with transcriptional complexes, drives the acquisition of leukemia properties. These encompass self-renewal due to the upregulation of ERG and the block of megakaryocytic differentiation as a consequence of the downregulation of GATA1 (79).

NCAM1 (CD56) upregulation in CBFA2T3-GLIS2 acute myeloid leukemia

As previously mentioned, evidence suggests that CD56 is a direct transcriptional target of CBFA2T3-GLIS2 (67, 68). Surface expression of CD56 on CBFA2T3-GLIS2-positive blast cells and its upregulated transcription have been recurrently observed throughout studies (82). The neural cell adhesion molecule 1 (NCAM1) is a prototypic immunoglobulin superfamily member (83). Its expression in the nervous system is abundant and crucial during embryonic development (84). Nonetheless, NCAM1 can be found in various tissues after birth but also in neoplasms (85). NCAM1 is a homophilic adhesion molecule whose expression and activity are regulated by several forms of cell surface modulation (86). These cell-to-cell interactions potentially affect cell homing and dissemination, and therefore CD56 may be a key player in cancer malignancy (87). In AML, about 25% of cases are CD56-positive, and its presence on blast cells indicates a poor prognosis as it was associated with unfavourable cytogenetic abnormalities, a reduced probability of achieving complete remission and shorter overall survival (88, 89).

Thus, it makes the assessment of CD56 by flow cytometry a valuable prognostic risk factor (87). Furthermore, CD56-positive AML cell lines exhibited drug resistance by activating the mitogen-activated protein kinase (MAPK) signalling pathway (90). Consequently, CD56 might be used as a putative biomarker for guiding treatment decisions, especially for the MAPK inhibitor-based therapies that might be beneficial (91).

Study models of CBFA2T3-GLIS2 acute myeloid leukemia

Pediatric CBFA2T3-GLIS2 AMKL is a sporadic and challenging disease. The paucity of samples compelled researchers to attempt, without success, to generate an animal model to precisely study the disease (56, 65). Indeed, *in vivo* models of leukemia are critical to the maintenance of the niche interactions that contribute to the disease sustainability (2, 3). Nevertheless, *in vitro* data demonstrated increased self-renewal together with megakaryocytic differentiation in CBFA2T3-GLIS2-transduced murine HSPCs when serially replated onto methylcellulose (65, 78). In another study, transduced human CB HSPCs could be cultured for at least 12 weeks while showing immature megakaryocytic differentiation (69). Previous failures to generate CBFA2T3-GLIS2 AMKL in animal models were ascribed to the requirement of a cooperative mutation (56, 65). However, one group has generated a transgenic mouse model that successfully phenocopied the disease *in vivo* solely upon the induction of CBFA2T3-GLIS2 in fetal liver HSCs, suggesting ontogeny-related susceptibility to transformation (71).

Our laboratory has contributed to the development of synthetic human models of AMKL by lentiviral-mediated overexpression of NUP98-KDM5A in CB HSPCs (92). The cells were maintained in an optimized medium containing agonists of cell self-renewal

SR-1 and UM171 and transplanted into immunodeficient mice (93, 94). Thus, we hypothesized that overexpression of CBFA2T3-GLIS2 in developmentally relevant human CB HSPCs could lead to leukemic transformation phenocopying AMKL in animal models. It is herein reported the generation of human models of CBFA2T3-GLIS2 leukemia in immunodeficient mice.

Materials and methods

Even though this study used experimental procedures implemented in the laboratory and comprehensively described in a recent paper (92), most underwent adjustment to reflect subtle changes of a slightly different approach. Statistical analysis was done using Prism 8 (GraphPad Software, San Diego, CA, USA).

Lentivirus production of CBFA2T3-GLIS2

It was demonstrated in the literature that the MO7e cell line (ACC 104, DSMZ) harbours the fusion gene CBFA2T3-GLIS2 (78). Hence, the coding sequence was obtained from the MO7e cell line by reverse transcriptase-PCR (RT-PCR) using reported primers (65) and cloned by standard procedures into a lentiviral vector containing an N-terminal-FLAG sequence along with the reporter gene GFP (figure 1B). According to standard procedures, pseudotype VSV-G vectors were produced with HEK293T cells then titered by flow cytometry.

Isolation of cord blood hematopoietic stem/progenitor cells (CD34+Lin-) by magnetic positive selection

The protocol for the isolation of human CB HSPCs was reviewed and approved by the research ethics board of CHU Sainte-Justine. CB units were collected with the mothers' consent and distributed by Héma-Québec through its public CB bank (Montreal, Canada). HSPCs were isolated from CB newly collected within the past 30 hours using the EasySep™ Human Cord Blood CD34 Positive Selection Kit (Cat. No. 17896; StemCell Technologies, Vancouver, BC, Canada) and the “The Big Easy” EasySep™ Magnet (Cat. No. 18001; StemCell Technologies, Vancouver, BC, Canada). Preparations having at least a concentration of 90% in CD34+ cells following flow cytometry analysis were used in experiments.

Lentiviral transduction of cord blood hematopoietic stem and progenitor cells

CB HSPCs (CD34+Lin-) were thawed and pooled in a medium containing SR-1 and UM171, prompting cell expansion (93, 94). Cells were cultured for 12 hours at 37°C in humidified air having a 5% concentration in CO₂. Lentiviral particles either encoding CBFA2T3-GLIS2 or the empty vector were added to a 96-well plate coated with RetroNectin® (Cat. No. T100B; Takara Bio USA Inc., Mountain View, CA, USA) in an amount to obtain a multiplicity of infection of 50. HSPCs were equally distributed for a total number of 20,000 cells per well in the cell expansion medium supplemented with 1 µg/mL of hexadimethrine bromide. They were kept for 16 hours at 37°C in humidified air having a 5% concentration of CO₂.

Detection of the fusion gene transcript

Following the transduction of CB HSPCs and 84 days of culture ending up in only GFP-positive cells, RNA was isolated using the Direct-zol™ RNA MiniPrep Kit (Cat. No. R2050, Zymo Research, Irvine, CA, USA). Coding DNA was obtained from isolated RNA with the Applied Biosystems™ High-Capacity cDNA Reverse Transcription Kit (Cat. No. 4368814, Thermo Fisher Scientific, Waltham, MA, USA). The fusion gene transcript was amplified with the DreamTaq DNA Polymerase (Cat. No. EP0702, Thermo Fisher Scientific, Waltham, MA, USA) using reported primers flanking the breakpoint (65). The thermocycler program had an initial stage of denaturation at 95 °C for one minute. The second stage was repeated 15 times with denaturation at 95 °C for 30 seconds, followed by a decreasing annealing temperature of 1 °C for every cycle beginning at 66 °C. The final step of extension was done at 72 °C for one minute. The third stage had identical denaturation and extension steps but repeated 20 times with a constant annealing temperature of 51 °C for 30 seconds. Finally, the fourth stage consisted of a terminal extension at 72 °C for seven minutes. In order to reveal the PCR products, electrophoresis with a 2% gel was performed at 85V for one hour.

Injection of CBFA2T3-GLIS2-transduced cells in immunodeficient mice

The CHU Sainte-Justine Institutional Animal Ethics Committee reviewed and approved the animal procedures. Female NOD-SCID IL2Rg^{null} (NSG) immunodeficient mice (Cat. No. 005557; The Jackson Laboratory, USA) were used for xenotransplantation at least 6 hours after sublethal irradiation of the whole body with 2 Gy of X-rays. Two third of each well was intravenously injected into the tail vein. Serial transplantation was fulfilled

likewise on the same day that sacrifice occurred. The mice were housed in pathogen-free conditions at the animal care facility of CHU Sainte-Justine. Air-filtered cages were used, wherein sterilized food and acidified water were provided.

Sacrifice of mice upon leukemia manifestations

Transplanted mice were kept alive up to 57 weeks or until the appearance of advanced manifestations of leukemia such as reduced mobility, paleness, hunchback and dyspnea. The bone marrow was extracted upon sacrifice by gently flushing 1% FBS-RPMI 1640 medium with a 3 mL syringe through the femur, tibia and pelvic bones. The spleen cells were dissociated by placing the tissue within a sterile 100 μm cell strainer mesh in a Petri dish filled with 1% FBS-RPMI 1640 medium and using the plunger of a syringe to mash it. Live cells in tissue suspensions were counted using a hemocytometer and the Trypan blue dye. About 1×10^6 live cells were taken for immunophenotyping which was realized with the LSRFortessa™ flow cytometer (Cat. No. 649225; BD Biosciences, Franklin Lakes, NJ, USA). Other 1×10^5 live cells were used for cyto centrifugation using the Cytospin™ 4 (Cat. No. A78300003; Thermo Fisher Scientific, Waltham, MA, USA) at 1500 RPM for 5 minutes. According to the conventional procedure, a Giemsa staining was applied to the obtained stained preparations. Cells were observed with the Axioscope A.1 upright microscope (Zeiss, Oberkochen, Germany) equipped with the AxioCam 105 color with 0,5X magnification (Zeiss, Oberkochen, Germany). The oil immersion objective lens had a 100X magnification and a numerical aperture of 1,4 (Zeiss, Oberkochen, Germany). Image acquisition and processing were achieved with the Zen 2.6 software (Zeiss, Oberkochen, Germany).

Flow cytometry analysis

Flow cytometry staining was performed in line with the standard procedure using 1×10^6 live cells. The staining buffer was composed of 1 mM EDTA and 2% FBS in a solution of PBS. Non-specific and Fc-receptor binding was blocked with a 1:1000 dilution of mouse gamma globulin (Cat. No. 015-000-002; Jackson ImmunoResearch Laboratories Inc., West Grove, PA, USA). Data were first analyzed with FACSDiva (BD Biosciences, Franklin Lakes, NJ, USA) then elaborated with FlowJo (BD Biosciences, Franklin Lakes, NJ, USA).

Immunofluorescence imaging of the chimeric protein

About 1×10^5 cells from an AMKL tertiary mouse were cytocentrifuged with the Cytospin™ 4 at 1500 RPM for 5 minutes. Preparations were fixed in 2% formaldehyde for 10 minutes then permeabilized with 0.2% Triton-X-100 for 10 minutes. Non-specific binding was blocked with 10% donkey serum for 20 minutes. The mouse monoclonal anti-FLAG® M2 antibody (Cat. No. F1804; Sigma-Aldrich Corp., St. Louis, MO, USA) in a dilution 1:1000 was added, and preparations were incubated for one hour. The donkey anti-mouse secondary antibody Alexa Fluor® 647 (Cat. No. A31571; Thermo Fisher Scientific, Waltham, MA, USA) in a dilution 1:1000 was added, and preparations were incubated for one hour in the dark. DNA was stained with 300 nM DAPI for 5 minutes. Finally, the preparations were mounted with No.1 coverslips and an antifading reagent.

Fluorescence microscopy was realized with the upright microscope Axioscope.A1 equipped with the AxioCam 305 mono with 0,63X magnification (Zeiss, Oberkochen,

Germany) and the X-Cite® 120LEDmini (Excelitas Technologies Corp., Waltham, MA, USA) as the light source. Filter cubes for DAPI and Cy5 (Zeiss, Oberkochen, Germany) were used with an objective lens having a magnification of 40X and a numerical aperture of 0,75 (Zeiss, Oberkochen, Germany). Image acquisition and processing were achieved with the Zen 2.6 software using an exposure time of 0.1 and 5.0 seconds together with a light intensity of 5 and 15% for the DAPI and Cy5 channels, respectively.

RNA sequencing

Library preparation, sequencing, and data processing were performed at the genomics platform of the Institute for Research in Immunology and Cancer (Montreal, Canada). Libraries were prepared using the KAPA Stranded mRNA-Seq Kit (Roche Life Science, Basel, Switzerland), and the paired-end sequencing of 150 cycles was performed on the NextSeq® 500 system (Illumina, San Diego, CA, USA). Sequences were trimmed to remove sequencing adapters and low-quality 3' bases using Trimmomatic version 0.35 (95). Subsequently, they were aligned to the reference human genome version GRCh38 using STAR version 2.5.1b (96) with gene annotation from GENCODE version 26. Gene expression was directly estimated as read counts from the STAR mapping and computed using RSEM version 1.2.28 (97) to obtain transcript level expression and fragments per kilobase of transcript per million mapped reads (FPKM). The fusion gene sequence was confirmed with FusionCatcher (98).

Comparative genomic hybridization/single-nucleotide polymorphism analysis

Genome integrity and copy-neutral genomic aberrations were assessed in some leukemia samples with the CGX™ Onco Arrays (Cat. No. 4120-0010; PerkinElmer Inc., Waltham, MA, USA) covering over 2410 cancer-relevant regions associated with hematological disorders as well as with solid tumours. The cytogenetics laboratory of CHU Sainte-Justine performed the microarray analysis.

Results

Overexpression of CBFA2T3-GLIS2 in hematopoietic stem and progenitor cells

As reported in the literature, the pediatric AMKL cell line M07e harbours the fusion gene CBFA2T3-GLIS2 (65). Hence, RNA was isolated from M07e cells then the coding sequence was obtained by RT-PCR and confirmed by RNA sequencing. The latter revealed that 11 exons encoding two NHR domains of CBFA3T2 were fused to the zinc-finger domains of GLIS2 from exon 4 (figure 1A). The resulting sequence creates a chimeric protein of 962 amino acids. Subsequently, the sequence was cloned into a lentiviral vector inducing overexpression of the gene of interest and the reporter gene GFP (figure 1B). The empty vector solely coding GFP serves as a negative control in forthcoming experiments. In order to verify whether the newly created sequence could be correctly inserted into the genome and soundly transcribed within CB HSPCs, the CBFA2T3-GLIS2 vector was introduced into 10,000 cells previously isolated from a single donor. RT-PCR was performed using primers flanking the breakpoint after 84 days of culture that was entirely composed of GFP+ cells (figure 1B). An expected PCR product of 627 bp was observed in both transduced cell samples and the M07e cells, whereas the

band was absent with untransduced CB HSPCs (figure 1C). Immunofluorescence imaging was performed as a next step in demonstrating the vector capability to produce the CBFA2T3-GLIS2 chimeric protein localized to the nucleus (figure 1D). The chimeric protein was detected with an anti-FLAG antibody, and it colocalized with DAPI-stained nuclei, suggesting an active transcription and translation of the inserted sequence.

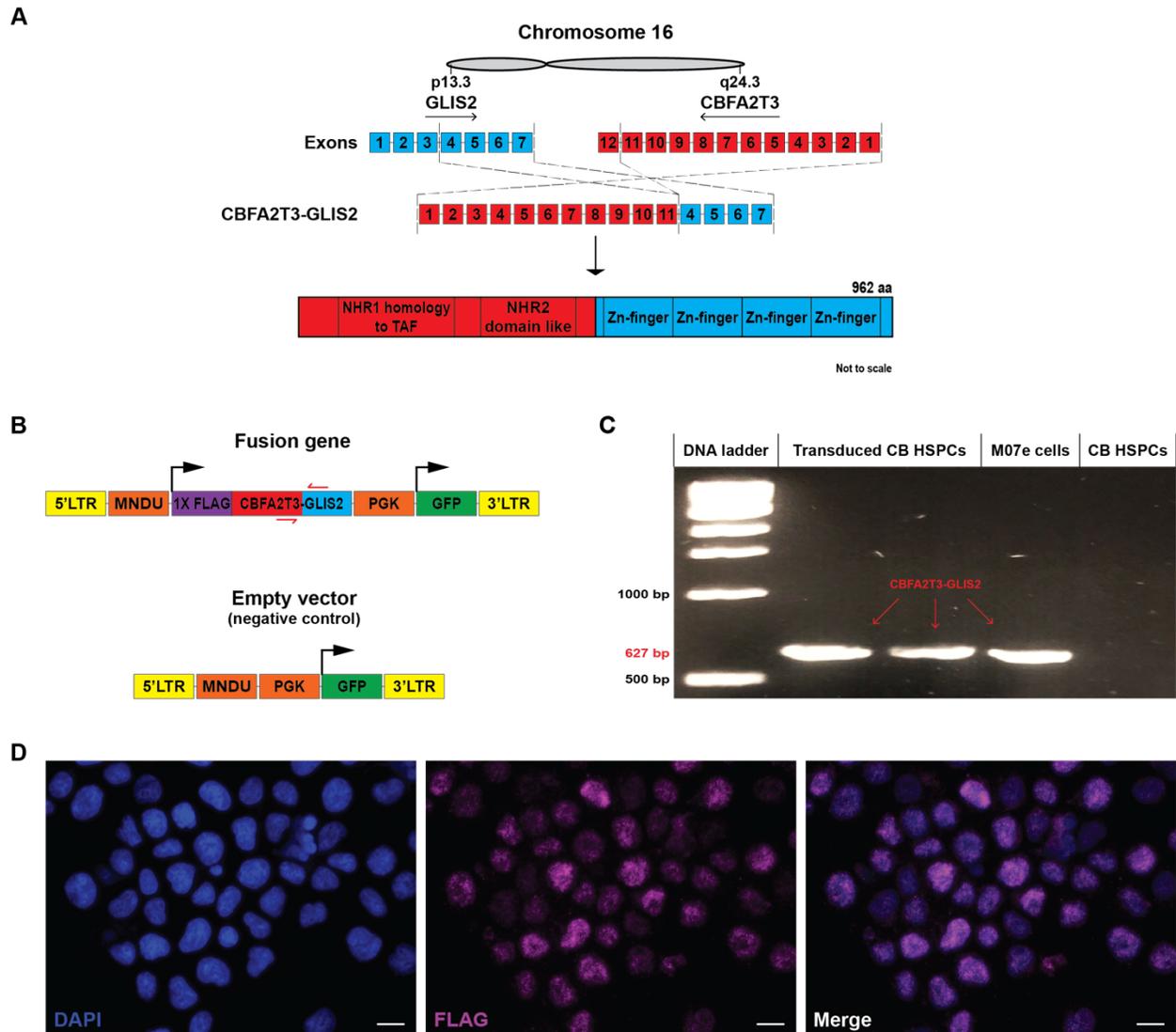


Figure 1: Overexpression of CBFA2T3-GLIS2 was induced in CB HSPCs

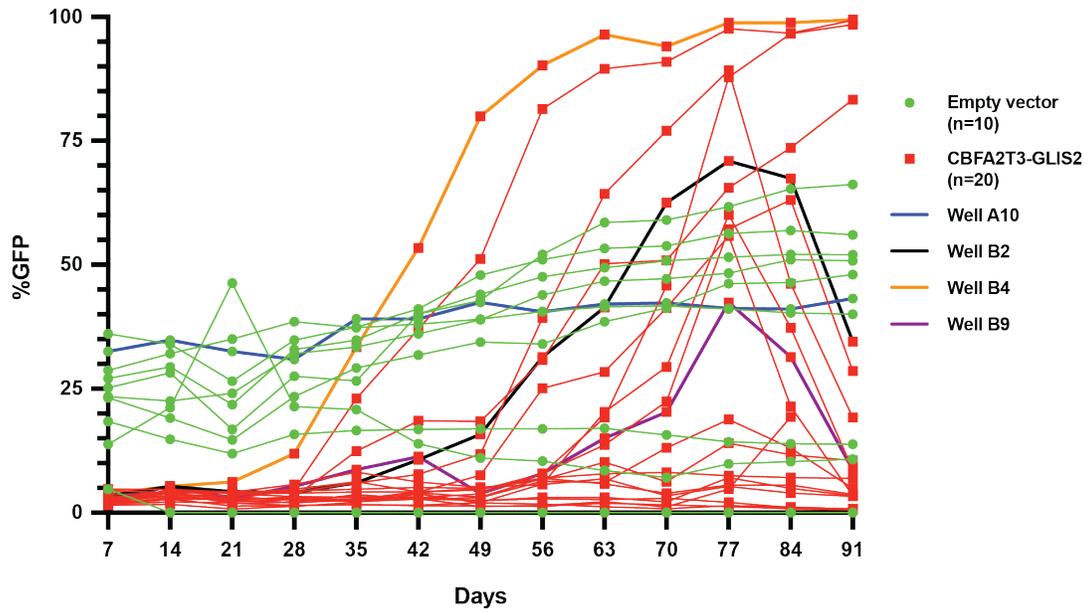
A. Schema of the CBFA2T3-GLIS2 sequence that was obtained from M07e cells by RT-PCR. RNA sequencing reveals two NHR domains of CBFA2T3 fused to the zinc-finger

domains of GLIS2. The resulting chimeric protein has 962 amino acids. **B.** The sequence was cloned into a lentiviral vector inducing overexpression of CBFA2T3-GLIS2 together with GFP. **C.** RT-PCR of transduced CB HSPCs using primers flanking the breakpoint confirmed the fusion gene transcription. **D.** Immunofluorescence imaging of the FLAG tag revealed the chimeric protein presence within the nucleus (scale=10 μm).

CBFA2T3-GLIS2-transduced HSPCs exhibited extensive proliferation *in vitro*

In order to study the proliferation capacity induced by CBFA2T3-GLIS2, CB HSPCs were isolated from four different donors. Upon experiment day, CB HSPCs were pooled then equally distributed into a 96-well plate for a number of 10,000 cells per well. The cells were transduced with either the empty or fusion-expressing lentiviral vector for a total of ten and 20 wells, respectively. The proportion of cells expressing GFP, hence indicating they had been transduced, was assessed by flow cytometry every seven days (figure 2A). Steep increases in GFP percentage were observed from day 28 for some wells having received the vector CBFA2T3-GLIS2. As a result, several CBFA2T3-GLIS2 wells surpassed the control wells and even reached 100% of cells being GFP+ for three of them. However, half of the CBFA2T3-GLIS2 wells held a constant proportion in GFP+ cells near zero from day seven until the end, and almost all CBFA2T3-GLIS2 wells that exhibited a rapid increase in GFP began to diminish at day 77. The majority of the control wells maintained a proportion of transduced cells around 50% until the experiment was terminated at day 91. The presence of steep increases rather than constant proportions in GFP+ cells suggests that the purportedly transformed cells considerably proliferated over the non-transduced cells.

A



B

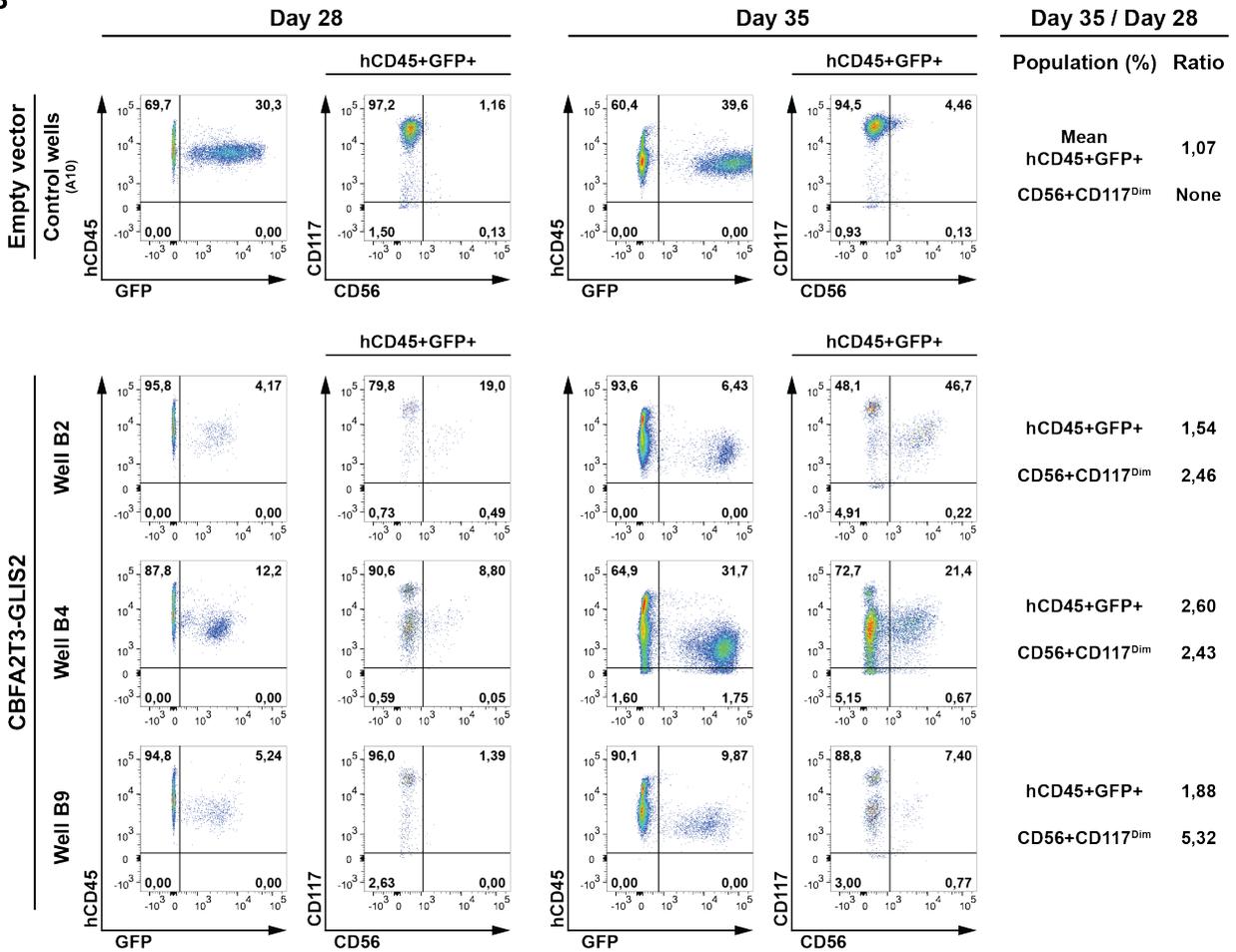


Figure 2: Extensive proliferation of CBFA2T3-GLIS2-transduced HSPCs following 28 days of culture. A. Half of the CBFA2T3-GLIS2 wells demonstrated a steep increase in the proportion of GFP+ cells beginning from day 28 to reaching 100% for three of them eventually. The control wells maintained a proportion in GFP+ cells around 50%. **B.** The proportions of hCD45+GFP+ cells and CD56+CD117^{Dim} cell population were determined by flow cytometry at day 28 and day 35 for all the control wells (represented by well A10) and three CBFA2T3-GLIS2 wells. There was a constant mean proportion in hCD45+GFP+ cells for all the control wells, whereas the proportions increased in CBFA2T3-GLIS2 wells. No CD56+CD117^{Dim} population was found in the control wells, while marked increases were observed in CBFA2T3-GLIS2 wells.

The proportions of hCD45+GFP+ and CD56+CD117^{Dim} cells at day 28 and day 35 were determined by flow cytometry for all the control wells and three CBFA2T3-GLIS2 wells (figure 2B). The CD56+CD117^{Dim} population, issued from hCD45+GFP+ cells, was studied because of the CD56 association with CBFA2T3-GLIS2 AML, while CD117 is an HSPC state marker and is expressed at a dim intensity in AML (99). The proportion of hCD45+GFP+ cells observed at day 35 versus day 28 for the control wells had a mean ratio of 1,07, indicating a constant proportion over this period. As expected, no CD56+CD117^{Dim} population was detected in the control wells. On the contrary, noteworthy increases in both hCD45+GFP+ and CD56+CD117^{Dim} populations were observed in the three shown CBFA2T3-GLIS2 wells. Although the proportions of hCD45+GFP+ cells increased by at least 1,54-fold, the greatest expansions were of the CD56+CD117^{Dim} cells ranging from 2,43-fold to 5,32-fold. Whether CBFA2T3-GLIS2

insertion into cultured HSPCs exerts immediate or delayed effects following transduction remains unknown. The phenotype and frequency of initiating cells susceptible to CBFA2T3-GLIS2-induced transformation will require more investigation.

Synthetic CBFA2T3-GLIS2 leukemia generated in immunodeficient mice

Similar to the *in vitro* study, CB HSPCs were previously isolated from six different donors. On experiment day, CB HSPCs were pooled then equally distributed into a 96-well plate for a number of 20,000 cells per well (figure 3A). Later the same day, the cells were transduced using lentiviruses by either the empty vector or the fusion gene vector for a total of five and ten wells, respectively. The entire cell content of each well was injected into a female immunodeficient mouse the following day.

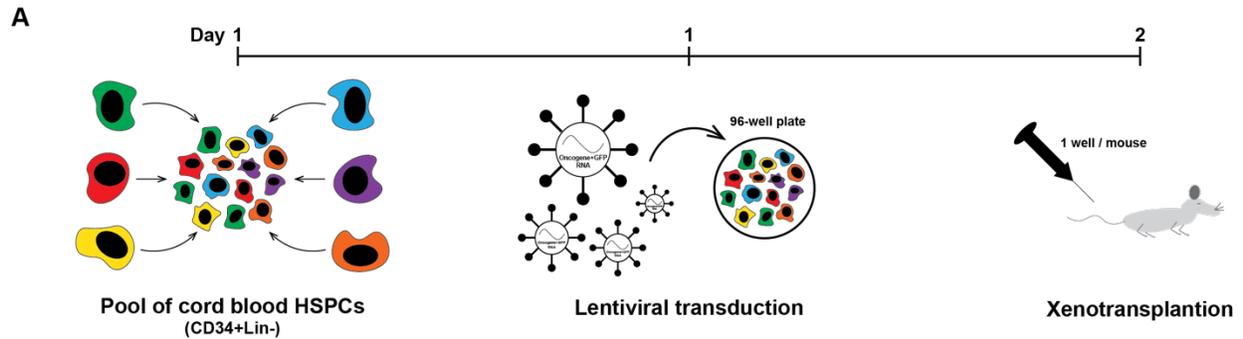
Eventually, six out of the ten mice that received cells with the CBFA2T3-GLIS2 vector ended up with leukemia following a period ranging from 9,9 to 36,1 weeks (figure 3B). The five control mice were sacrificed at week 57, thereby terminating the experiment. While the other five ill animals had a spleen of about 0,100 g, the mouse I749^{AMKL} was the sole animal presenting with an evidently enlarged spleen weighing 0,385 g. Blast cells, identified as the hCD45+GFP+ population by flow cytometry, were detected in the bone marrow of every animal in proportions varying from 40% to complete infiltration in the case of the I749^{AMKL}. Except for the I730^{AML}, the spleen was also infiltrated, though to a lesser extent than the bone marrow. Only the I749^{AMKL} that presented with splenomegaly had a spleen as infiltrated as the bone marrow. Half of the tumours strongly expressed an AMKL phenotype, while the three others were deemed AML based on the immunophenotype and cell morphology. Indeed, no CD19 nor CD3 population was

detected in any animal (data not shown), which could have indicated lymphoblastic leukemia. Furthermore, all the tumours expressed the markers CD71 and CD117 that may be encountered in AML, especially the latter. Nonetheless, more specific myeloid markers would be required to ascertain the leukemia lineage undoubtedly. Although the I747^{AML} tumour had a minor CD41 population, only three expressed the megakaryocytic markers CD41 and CD61 at a proportion larger or equal to 50% as sought to be clearly classified AMKL. The stem cell marker CD34 was detected on I750^{AML}, I731^{AMKL} and I749^{AMKL} leukemic cells, with a greater proportion for the last. CD56 was expressed on the surface of the primary mice except for the I750^{AML} and at a lower proportion for the I730^{AML}. The mouse I747^{AML} expressed CD56 strongly and exclusively among the AML tumours. Therefore, the CD56 expression was definitively more predominant in the AMKL subtype since it was observed on every tumour by more than half the blast cells. Ultimately, the primary tumours were successfully passed into secondary recipients to confirm the generation of synthetic leukemia. Moreover, the mice I749^{AMKL}, I732^{AMKL} and I750^{AML} produced 10, 8 and 7 tertiary recipients, respectively.

The immunophenotype and cell morphology were examined for comprehending the contrasts between the presentations of CBFA2T3-GLIS2 leukemia. The best representative of each group is shown for comparison (figure 3C). The immunophenotype was determined by flow cytometry and cell morphology evidenced by cytocentrifugation followed by a Giemsa staining upon sacrifice. The blast cells, identified by a single double-positive population expressing hCD45 and GFP that is not seen in healthy animals, are shown in the upper right quadrant of the first column. In the next column, blast cells from the I749^{AMKL} strongly expressed CD41 and CD61 as displayed in the upper right quadrant,

whereas they were absent in both tissues of the I750^{AML}. Another striking difference of AMKL animals shown in the third column is the dual expression of CD56 and CD34 in both tissues forming distinct populations once again in the upper right quadrant. Nevertheless, a small population of about 10% seemed to be present in the bone marrow of the I750^{AML} though it is challenging to ascertain. Also, a quarter of the I749^{AMKL} blast cells solely expressed CD56 without CD34 in the lower right quadrant. The I750^{AML} was the unique AML tumour with a small proportion of CD34+ cells in the bone marrow, as shown in the upper left quadrant. It is possible to appreciate the common positivity of the markers CD71 and CD177 across generated tumours in the upper right quadrant of the last but one column. Lastly, cells from the I749^{AMKL} exhibited blebs on the plasma membrane (indicated by arrows), a putative characteristic of megakaryocytic differentiation. Also, smaller cells with dark coloration are seen, principally in the I750^{AML}.

In light of these observations, it is evident that synthetic CBFA2T3-GLIS2 leukemia has a heterogeneous expression of cell surface markers allowing its classification into the AMKL subtype and a less precise presentation of AML. Furthermore, CD41+CD61+ cells had blebs on the surface, which were absent on double-negative cells. Astonishingly, the phenotype associated with a primary mouse is also reflected on secondary and tertiary recipients.



B

Primary leukemia	Latency (weeks)	Spleen weight (g)	Blast cells (%)		Immunophenotype	Recipient mice	
			Bone marrow	Spleen		Sec.	Ter.
1731 ^{AMKL}	28.1	0.068	40	14	CD34+ CD41+ CD45+ CD56+ CD61+ CD71+ CD117+	1	0
1732 ^{AMKL}	20.3	0.090	47	38	CD34- CD41+ CD45+ CD56+ CD61+ CD71+ CD117+	2	8
1749 ^{AMKL}	29.4	0.385	97	98	CD34+ CD41+ CD45+ CD56+ CD61+ CD71+ CD117+	1	10
1730 ^{AML}	36.1	0.097	77	0	CD34- CD41- CD45+ CD56 ^{low} CD61- CD71+ CD117+	1	0
1747 ^{AML}	9.9	0.045	89	20	CD34- CD41 ^{Low} CD45+ CD56+ CD61- CD71+ CD117+	2	0
1750 ^{AML}	28.1	0.118	67	21	CD34 ^{Low} CD41- CD45+ CD56- CD61- CD71+ CD117+	2	7

(-) ≤ 20% ; 20% < (Low) > 50% ; (+) ≥ 50%

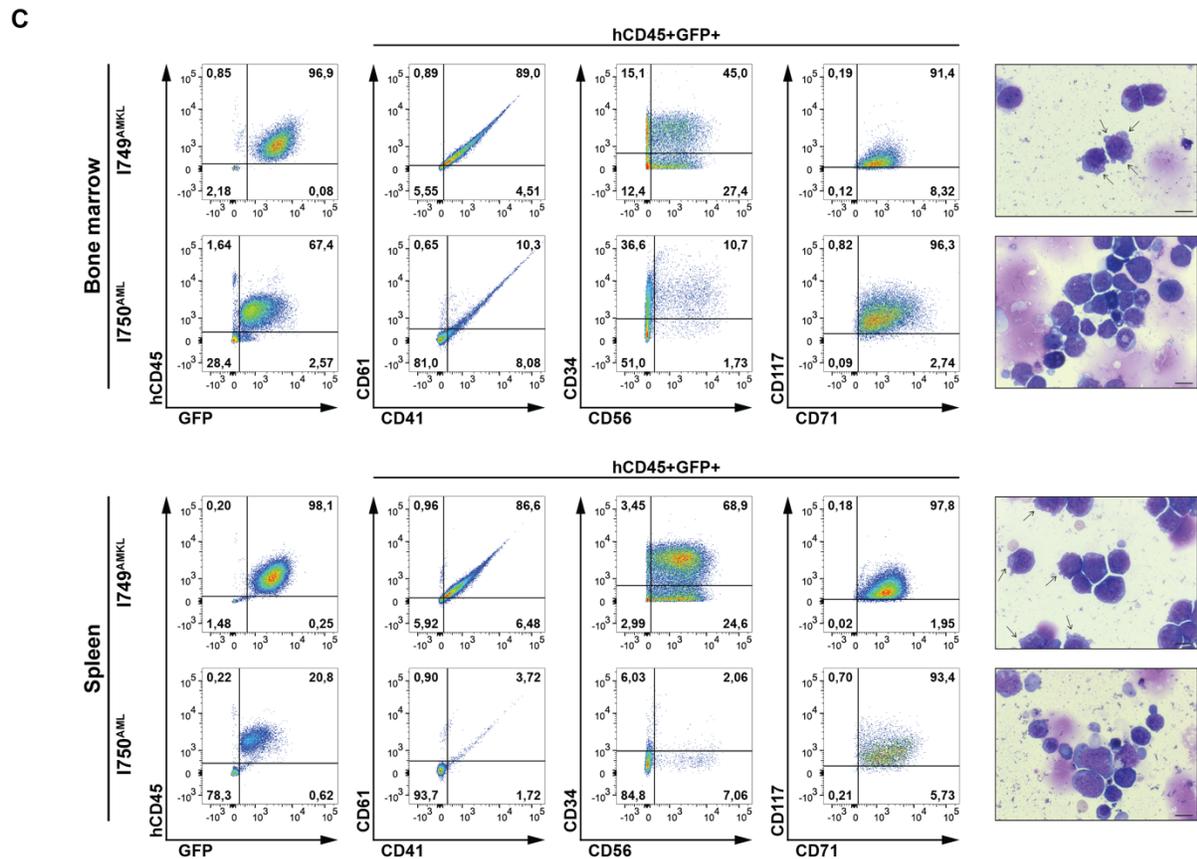


Figure 3: CBFA2T3-GLIS2-transduced HSPCs generated AML with a dual phenotype in immunodeficient mice. **A.** CB HSPCs (CD34+Lin-) were transduced by lentiviruses carrying the fusion gene together with GFP. The cells were injected into immunodeficient mice the next day. **B.** Six animals had leukemia, of which three clearly expressed the megakaryocytic markers CD41 and CD61. CD56 was observed across the tumours with the exception of one. The generation of leukemia was confirmed since the disease was successfully passed into recipient mice. **C.** A tumour representing each phenotype was displayed to contrast the immunophenotype and cell morphology of cells harvested from the bone marrow and the spleen. Blebs indicated by arrows were present on the cells from both tissues of the I749^{AMKL}, as evidenced by cytocentrifugation preparations stained by Giemsa. (scale=10 μ m).

Phenotypes of synthetic CBFA2T3-GLIS2 leukemia remained constant through recipient hosts

According to the leukemia passage, the latency of generated recipient mice was depicted on a survival graph (figure 4A). A significant diminution in latency over the number of passages ($p < 0,05$) was observed. However, it is essential to mention that the number of injected cells was 50 times greater for the secondary and tertiary groups. Also, the tertiary group was misrepresented since only half the primary tumours generated tertiary recipients. Nevertheless, the tertiary recipients had similar latencies among them with respect to the initial tumour of which they were issued.

The spleen weight of the recipients issued from the tumours I732^{AMKL}, I749^{AMKL} and I750^{AML} perdured through the passages, although the I749^{AMKL} descendants showed

a greater variation than the controls and the other lineages (figure 4B). A significant difference was found between the I749^{AMKL} descendants and the other lineages and the controls ($p < 0,05$). The blast cells infiltration within the bone marrow and the spleen were identically determined by identifying the hCD45+GFP+ population using flow cytometry (figure 4C). The infiltration level mainly remained constant in both tissues for all lineages. The bone marrow was significantly more infiltrated than the spleen ($p < 0,05$), and the spleens of the I749^{AMKL} descendants were significantly more infiltrated than the other lineages ($p < 0,05$). As it was with the primary tumours, the spleens of the AML tumours seem to be less infiltrated than the AMKL ones. Nonetheless, further studies will be required to validate this observation.

Bone marrow cells from recipients of the I749^{AMKL} and I750^{AML} were examined regarding the immunophenotype and cell morphology using identical means (figure 4D). The I749^{AMKL} descendants had a total infiltration and conspicuously expressed the megakaryocytic markers CD41 and CD61. The proportion of cells displaying CD34+CD117+, which are undifferentiation markers, increased from secondary to tertiary recipients. However, it is unclear whether the same phenomenon occurred with the CD71+CD56+ cells since the population in the secondary recipient is ill-defined. Blebs were also visible on the surface of cells from the descendants of the I749^{AMKL} (indicated by arrows). The bone marrow of the I750^{AML} descendants was entirely infiltrated, which represents an impressive increase compared to the 67% seen in the primary tumour. There was a continual absence of the markers CD41, CD61 and CD56, while blebs were still not observed on cells. Furthermore, the low proportion of CD34+ cells seen in the bone marrow of the I750^{AML} vanished in the secondary passage.

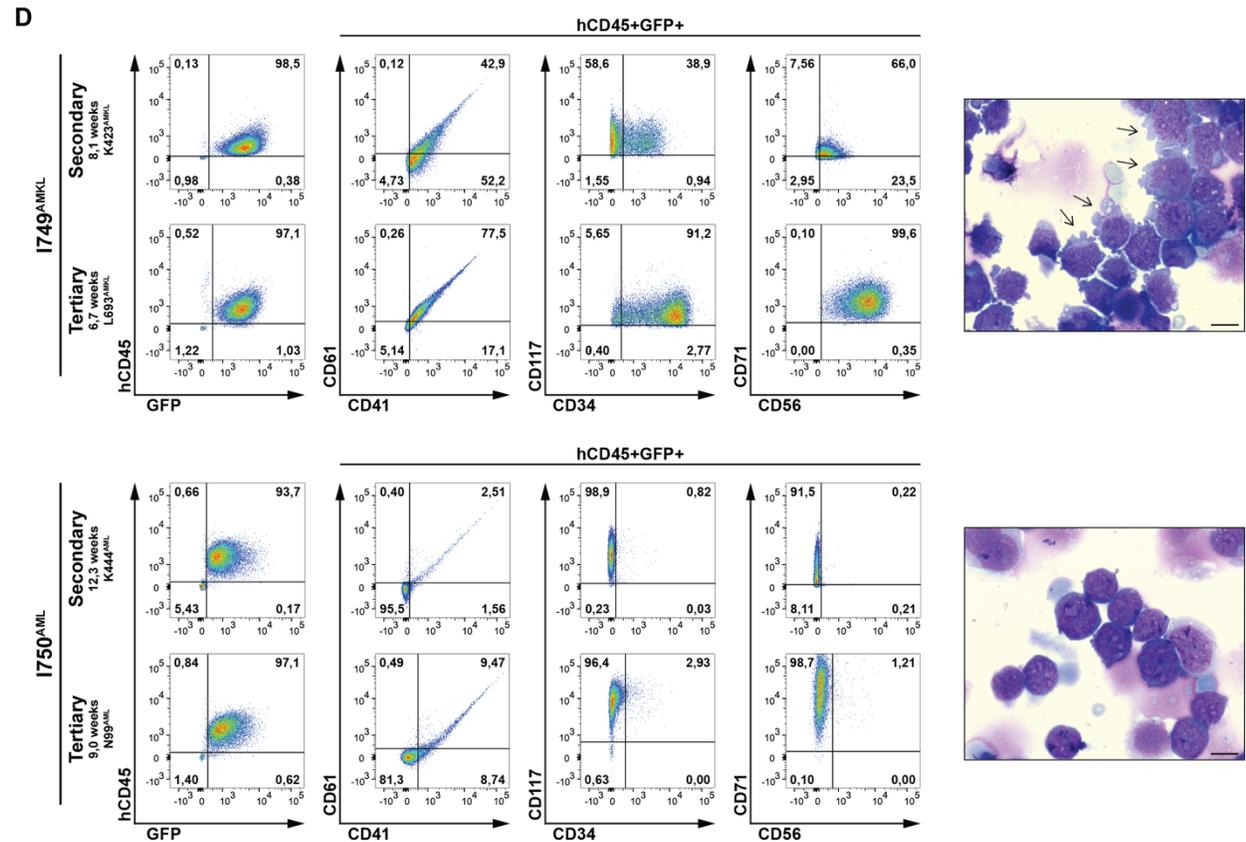
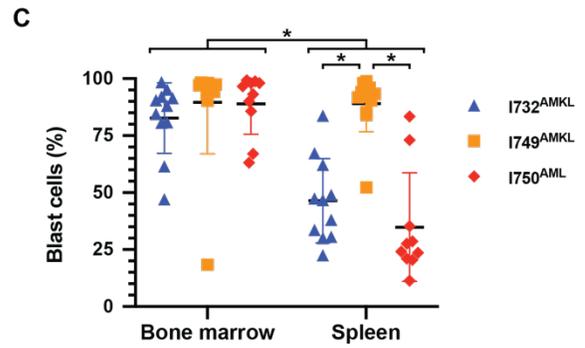
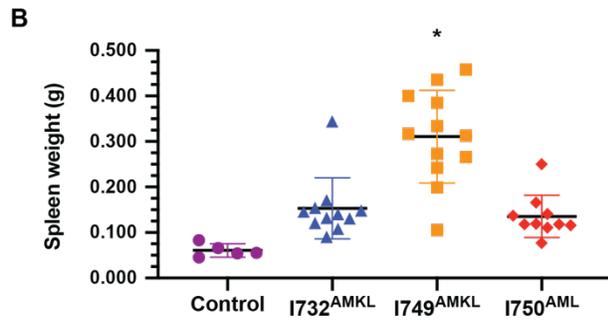
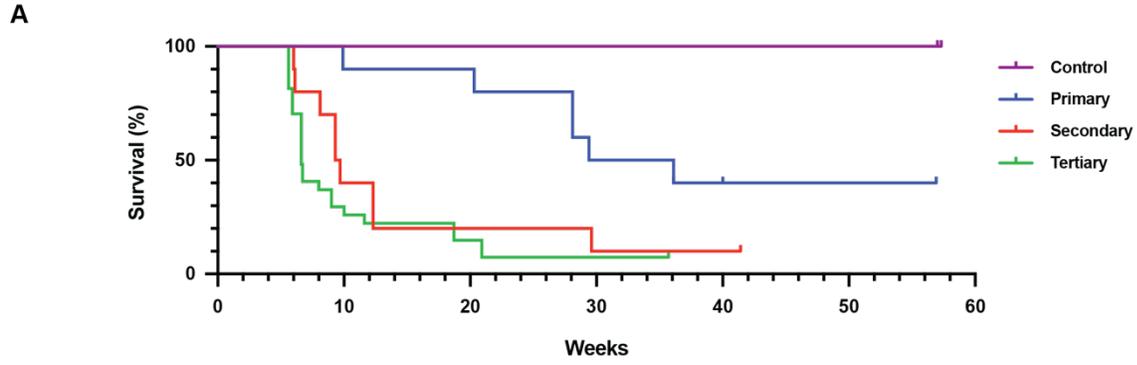


Figure 4: Self-renewal of primary tumours assessed by serial transplantation. A.

Survival curves of recipient mice serially transplanted with synthetic CBFA2T3-GLIS2 AMKL cells depicted significant latency diminution over the number of passages ($p < 0,05$). Also, it is possible to appreciate that the latency was inversely correlated to the number of passages. **B. C.** The spleen weights and infiltration levels were consistent in the recipients of the primary tumours 1732^{AMKL}, 1749^{AMKL} and 1750^{AML}. Only the descendants of the 1749^{AMKL} presented with significantly greater spleen weights than the controls and the other lineages ($p < 0,05$). Furthermore, the bone marrow was significantly more infiltrated than the spleen ($p < 0,05$), while the spleens of the descendants of the 1749^{AMKL} were significantly more infiltrated than the ones of the other lineages ($p < 0,05$). **D.** The immunophenotype and cell morphology were thoroughly similar to the corresponding primary tumour, with blebs present on CD41+CD61+ cells (scale=10 μm).

Overall, recipient mice exhibited a latency that was inversely correlated to the number of passages. The spleen weights and infiltration levels were consistent in recipient hosts with reference to the primary tumour they were issued. Likewise, the immunophenotype and cell morphology were, with blebs present on CD41+CD61+ cells. Therefore, synthetic CBFA2T3-GLIS2 leukemia shows a self-renewal capacity as assessed by serial transplantation and maintenance of phenotypic characteristics.

RNA expression profiles from CBFA2T3-GLIS2 xenografts correlate with pediatric megakaryocytic leukemia

RNA sequencing was performed with RNA extracted from CBFA2T3-GLIS2 xenografts (five primary and I730^{AML} secondary tumours) and a CBFA2T3-GLIS2 AMKL patient to assess whether the expression profile of the models correlates with the human disease. Principal component analysis (PCA) was performed with the RNAseq datasets from the CBFA2T3-GLIS2 xenografts and a patient sample and compared to previously sequenced NUP98-KDM5A AMKL and normal CB-CD34+ cells from our local cohort (figure 5). The PCA analysis showed model and patient tumours clustering according to their genetic subtype.

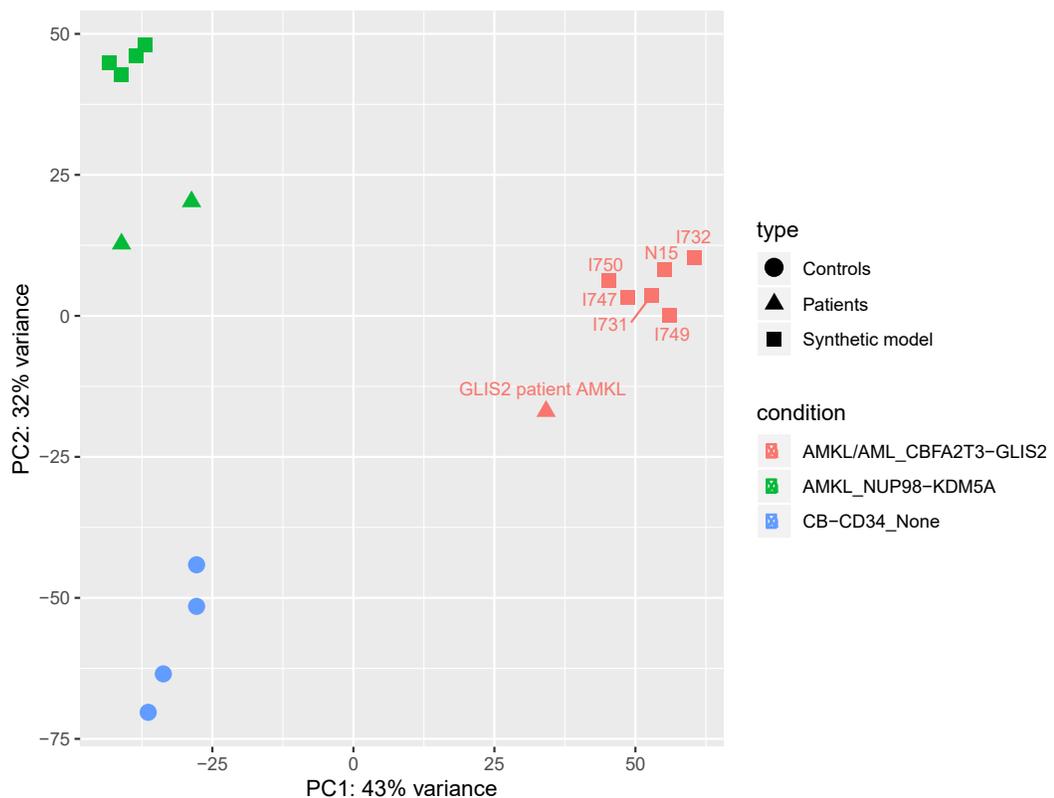


Figure 5: Principal component analysis (PCA) plot of the first and second principal components of CBFA2T3-GLIS2 models. NUP8-KDM5A models, patient and control cord blood (CB)-CD34⁺ cells are shown for comparison. The analysis was calculated using the 500 most variable genes between all conditions from the RNAseq dataset.

Hierarchical clustering performed with the RNAseq datasets from the synthetic CBFA2T3-GLIS2 xenografts and an AMKL patient cohort (53) showed genotype-matched grouping of model and patient tumours (figure 6). Overall, global analysis of the RNAseq datasets suggests that the engineered CBFA2T3-GLIS2 leukemia xenografts correlate with the human disease molecularly.

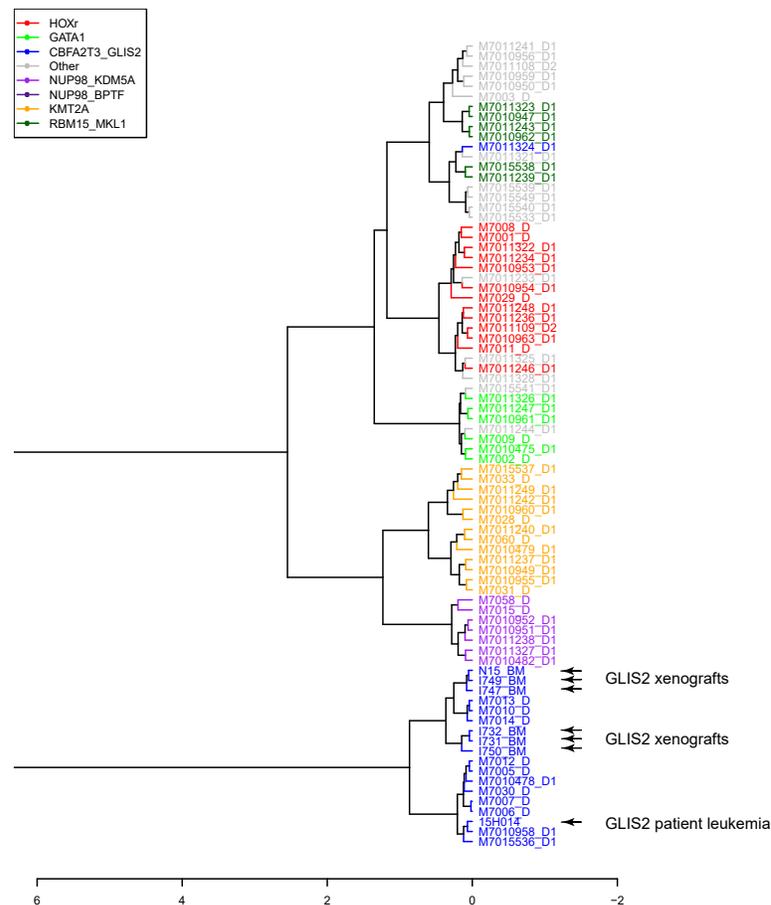


Figure 6: Hierarchical clustering was performed with the RNAseq datasets from the CBFA2T3-GLIS2 xenografts and an AMKL patient cohort (53), using the top 100 most variant genes. The Expression profiles from the CBFA2T3-GLIS2 xenografts cluster with the genotype-matched patient tumours.

Differential expression of select HSPC, megakaryocytic and myeloid markers were evaluated in CBFA2T3-GLIS2 models and a patient tumour, normal CB HSPCs (n=4 different donors) and two AML patients (fusion genes NUP98-NSD1 and MLL-AF9) (table I). The megakaryocytic markers, notably CD41 and CD61, were unexpectedly upregulated in all the synthetic tumours, including the AML ones. Indeed, the expression levels were relatively superior compared to both normal CB HSPCs and AML patients. Furthermore, most of the myeloid surface markers as well as the transcription factor CEBPA, which is essential for myeloid lineage commitment, were less expressed in the synthetic models. As previously reported in the literature, NCAM1 was expressed in all CBFA2T3-GLIS2 leukemia models. Altogether, the upregulation of megakaryocytic genes and NCAM1 in every model suggests a heterogeneous disease with a megakaryocytic component, even in the absence/low expression of CD41 and CD61 on the cell surface.

		CBFA2T3-GLIS2							AML		CB	
Gene		Patient	Cell line	I731 ^{AMKL}	I732 ^{AMKL}	I749 ^{AMKL}	N15 ^{AML} (I730 ^{AML})	I747 ^{AML}	I750 ^{AML}	NUP98- NSD1 patient	MLL-AF9 patient	HSPCs
HSPC	CD34	104	1	25	5	143	122	3	117	194	0	119
	KIT (CD117)	29	115	29	60	59	45	54	37	89	5	47
	PROM1 (CD133)	0	9	0	0	1	0	0	0	21	0	49
	THY1 (CD90)	0	107	1	0	2	0	0	1	0	0	0
Megakaryocytic	ITGA2B (CD41)	408	98	426	246	244	262	447	384	46	1	22
	GP9 (CD42a)	11	6	24	30	19	8	50	14	4	0	2
	GP1BA (CD42b)	212	38	96	107	94	112	148	134	3	0	8
	GP1BB (CD42c)	134	9	88	93	10	31	162	80	5	7	6
	GP5 (CD42d)	0	0	0	0	0	0	0	0	0	0	0
	ITGB3 (CD61)	18	13	21	44	29	23	46	29	8	0	10
	MPL (CD110)	4	0	41	28	7	14	32	19	3	0	4
Myeloid	LYZ	210	18	35	92	21	20	20	19	649	1922	378
	MPO	10	1	12	8	11	3	1	4	401	4	284
	ITGAM (CD11b)	8	7	9	4	8	5	2	6	20	70	9
	ANPEP (CD13)	8	32	6	3	20	3	2	3	36	2	52
	CD14	9	0	0	0	0	0	0	0	1	1	41
	FUT4 (CD15)	2	5	4	6	4	4	5	4	29	22	7
	CD33	103	22	31	6	161	94	8	77	79	61	27
	CD68	131	9	5	5	7	2	6	4	35	245	65
	IL3RA (CD123)	3	3	9	9	15	7	7	9	28	9	17
Others	CD44	1954	417	74	248	380	388	399	337	129	336	387
	PTPRC (CD45)	72	25	97	73	61	68	66	61	185	116	136
	TFRC (CD71)	19	99	36	38	78	79	21	66	53	55	129
	NCAM1 (CD56)	652	36	40	115	80	49	153	36	0	9	1
	GATA1	1	1	29	20	1	5	25	11	6	0	21
	ERG	97	27	89	81	40	89	80	93	36	3	64
	CEBPA	1	1	0	0	1	0	0	0	64	70	23
	HOXA9	0	0	0	0	0	0	0	0	106	133	45
	CBFA2T3	66	53	58	47	35	39	58	67	37	3	29
	GLIS2	17	15	8	4	7	6	7	9	0	0	1

Values are in FPKM

0	100
---	-----

Table I: Gene expression of the CBFA2T3-GLIS2 synthetic models. Megakaryocytic markers were upregulated in every synthetic CBFA2T3-GLIS2 tumour compared to the normal CB HSPCs and two AML patients.

Synthetic tumours of CBFA2T3-GLIS2 leukemia harbored few chromosomal abnormalities

Once the immunophenotype of the generated tumours was characterized, attention was given to the possibility that blast cells may carry chromosomal

abnormalities. Thus, comparative genomic hybridization and single nucleotide polymorphism arrays (CGH/SNP) were performed on tumours of different passages (table II). Only the primary mouse I731^{AMKL} was not represented. The analysis revealed that the I732^{AMKL} descendants carried a del(11)(p14.3-p12) in proportions from 50 to 80%. Interestingly, the number of carrying blast cells increased from the secondary to the tertiary recipient. Although this chromosomal region includes cancer-related genes such as APIP, CAPRIN1, CD44, ELF5, FANCF, LMO2, PAX6, RAG1, RAG2, TRAF6, WT1 and PRRG4, the clinical significance of this deletion remains unknown. Surprisingly, the I749^{AMKL} did not present chromosomal abnormalities even though this leukemia seemed quite aggressive because of the splenomegaly and complete infiltration in the bone marrow and spleen. Also, nothing was discovered in the tumour I747^{AML} despite having the shortest latency of only 9,9 weeks. However, the secondary recipient of the I747^{AML} had a duplication of the terminal region of chromosome 1 in 90% of the cells. Unfortunately, little is known on how the genes of this area might impact leukemia. A chromosomal abnormality was observed in the secondary recipient of the I750^{AML} as 85% of the cells harboured a del(20)(q11.22-q13.2). This latter deletion has a diagnostic significance in secondary AML (100). The CGH analysis of parental cord blood cells was normal and did not reveal any cancer-predisposing chromosomal anomalies. Briefly, more investigation will be required to determine whether the altered genetic regions actually impact leukemia generation and progression.

Leukemia	Chromosomal abnormality	Clinical significance
K465 ^{AMKL} (1732 ^{AMKL} Sec.)	del(11)(p14.3-p12) ~50% cells	Unknown
L691 ^{AMKL} (1732 ^{AMKL} Ter.)	del(11)(p14.3-p12) ~80% cells	Unknown
I749 ^{AMKL}	None	None
L693 ^{AMKL} (1749 ^{AMKL} Ter.)	None	None
N15 ^{AML} (1730 ^{AML} Sec.)	dup(1)(q25.3-qter) ~90% cells	Unknown
N877 ^{AML} (1747 ^{AML} Sec.)	None	None
K444 ^{AML} (1750 ^{AML} Sec.)	del(20)(q11.22-q13.2) ~85% cells	Diagnostic of secondary AML

Table II: Chromosomal abnormalities were found in recipients from three primary tumours. Although two other abnormalities were observed, only the del(20q) seen in the I750^{AML} secondary recipient has a diagnostic significance of secondary AML.

Blast cells regenerated leukemia after a short period of cell culture

In order to determine whether leukemia could be generated after a period of cell culture, splenic cells of the K423^{AMKL}, the unique I749^{AMKL} secondary recipient, were cultured for six days then transplanted (figure 7). There was a surface expression of the megakaryocyte markers (CD41; CD61) together with the CD34+CD117+ and CD56+CD71+ populations from day one. This expression profile remained for the next five days, with the fluorescence intensity becoming much brighter as soon as three days of culture. Furthermore, the cell culture was performed with six independent wells, wherein the number of cells had uniformly expanded by about ten folds to ultimately a mean of 2,9 million cells (supplementary table I). The entire cell content of each well was transplanted after six days into one recipient mouse that generated leukemia ensuing a

mean latency of 11,2 weeks. Harvested blast cells showed a megakaryocytic immunophenotype, with a greater proportion of cells concurrently expressing CD34 and CD117 at sacrifice. Consequently, this elevated expression of HSPCs markers might indicate a superior stemness state than observed in culture. In the end, the repopulating cells, which are capable of regenerating leukemia, were either maintained or injected in a sufficient number ensuing six days of cell culture.

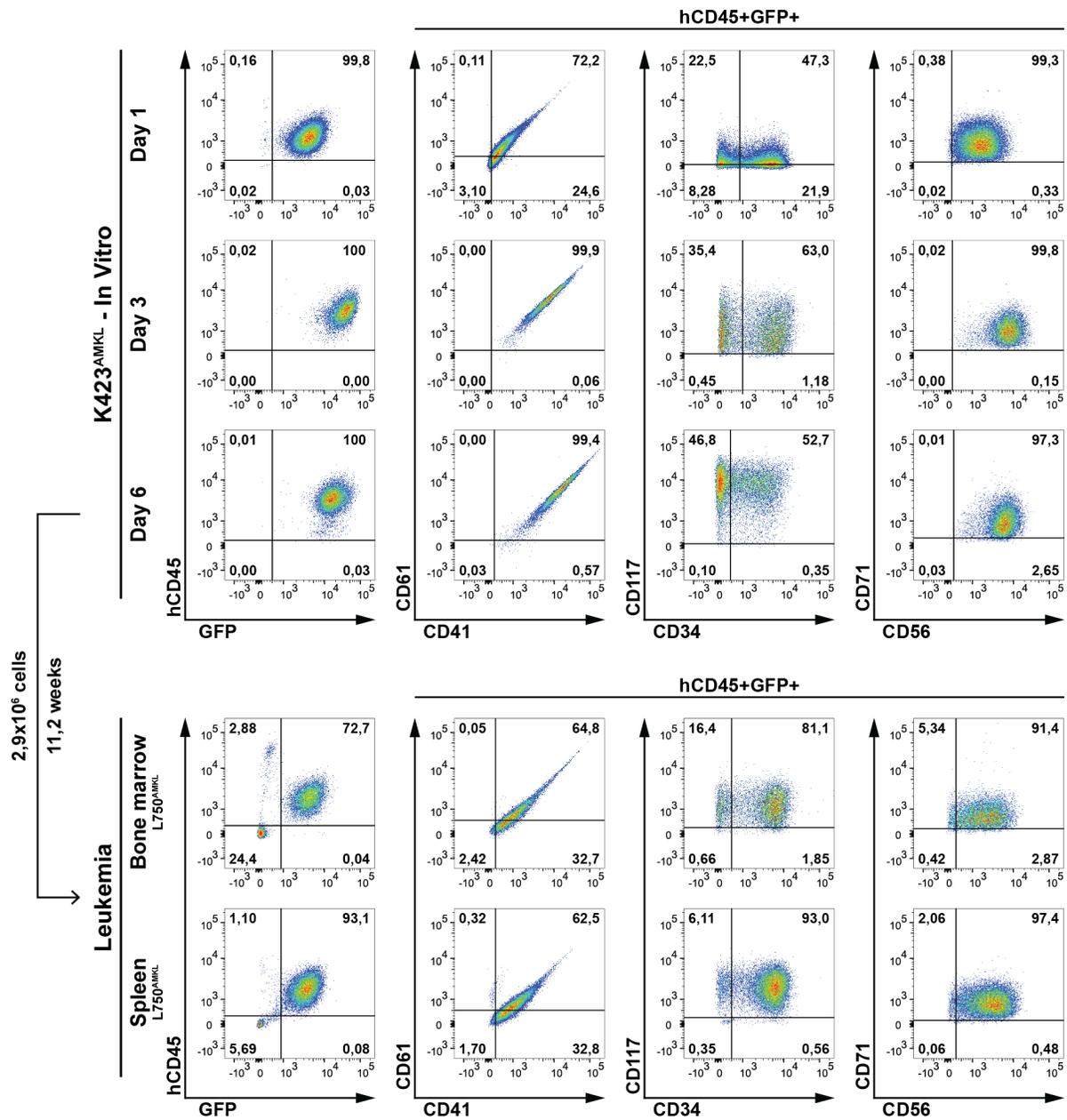


Figure 7: Blast cells had the potential to regenerate leukemia after six days of cell culture. Blast cells were transplanted into a recipient mouse following six days of cell culture and regenerated AMKL. Also, a more significant proportion of cells concurrently expressed CD34 and CD117 at sacrifice.

Repopulating cells regenerating leukemia were present in a high frequency

Hitherto, serial transplantation was achieved in different quantities since the frequency of repopulating cells capable of regenerating leukemia was unknown. Therefore, a limiting dilution assay (LDA) was done using splenic cells of the primary tumour I749^{AMKL} (figure 8) to estimate the frequency of leukemia-initiating cells (LIC) able to reconstitute leukemia in recipient mice. The LDA consists in transplanting groups of sublethally irradiated mice with a range of tumour cell doses and determining the frequency of leukemic and non-leukemic mice in each group. An estimation of the LICs frequency can be calculated using computational software and assuming a Poisson distribution and a single-hit hypothesis. Ascending quantities of cells were injected into recipient mice, with the smallest being $2,5 \times 10^2$ and the biggest $1,0 \times 10^6$ cells. As a result, mice having received at least $1,0 \times 10^5$ cells totally ended up with leukemia, which was ascertained by flow cytometry upon signs of disease. The highest fraction without leukemia was 0,38 following an injection of $2,5 \times 10^2$ cells (figure 8A). Although the lowest quantity unable to generate leukemia was not achieved, the striking repopulating cell frequency of one in 3445 cells could be calculated using the ELDA software (101) with a 95% confidence interval of 1/8511 to 1/1395 (CI) (figure 8B). The single-hit Poisson test was significant ($p = 2.23e-07$), while it was not the case with the heterogeneity test ($p = 1$). Also, a regression slope of 0.204 was calculated. The significant survival curves showed that latency before leukemia was negatively correlated with the number of injected cells (figure 8C). This estimated frequency of repopulating cells, in sum, truly reflects the appalling aggressiveness observed with the tumour I749^{AMKL} particularly.

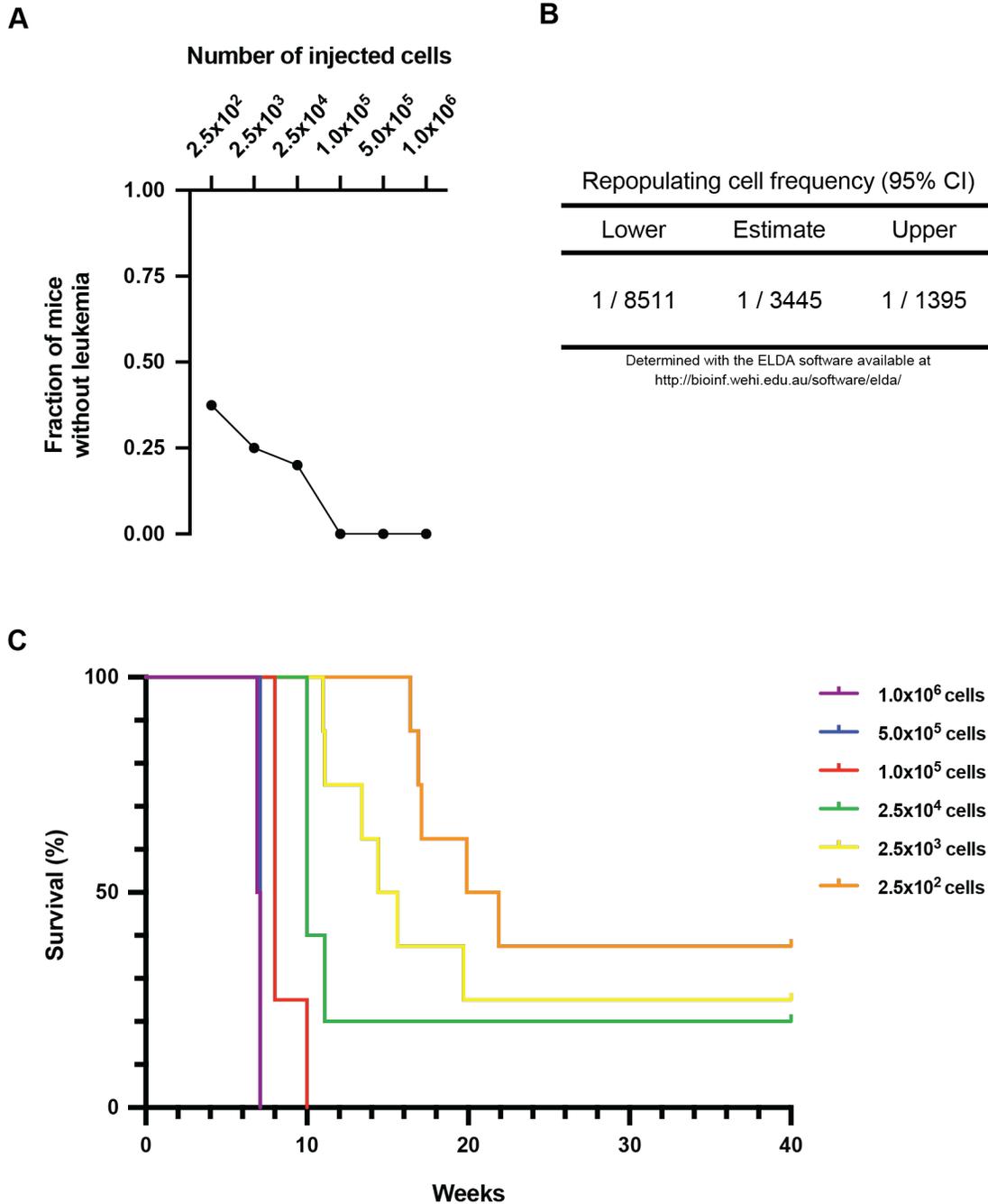


Figure 8: Limiting dilution assay (LDA) determined the frequency of repopulating cells for a single primary tumour. A. The highest fraction without leukemia was 0,38 with $2,5 \times 10^2$ injected cells. **B.** The repopulating cells frequency was estimated as being 1/3445 cells. **C.** Significant survival curves showed an inverse correlation of latency with the number of injected cells.

Discussion

The successful generation of human CBFA2T3-GLIS2 leukemia has been achieved by lentiviral transduction using CB HSPCs and transplantation in immunodeficient mice. The synthetic CBFA2T3-GLIS2 xenograft models presented myeloid/megakaryocytic features and self-renewal while recapitulating the human disease phenotypically and molecularly.

According to the CGH analysis, the donor cord blood had a normal molecular karyotype. As proposed by researchers within the field, the requirement of a cooperative mutation appeared unnecessary to drive CBFA2T3-GLIS2 leukemia (56, 65). Also, another group successfully generated the *in vitro* model using the same experimental approach and observed malignancy in CBFA2T3-GLIS2-transduced cells (69). The odds of receiving a CB unit with predispositions to leukemia in two different experiments performed in separate research centres are low. Next-generation sequencing could be helpful nonetheless in the detection of a pre-existing alteration that might be present in CB HSPCs.

The CBFA2T3-GLIS2 sequence was obtained from the M07e cell line and inserted into CB HSPCs by lentiviral transduction (figure 1). RT-PCR verified CBFA2T3-GLIS2 transcription. Immunofluorescence imaging validated the oncoprotein translation and presence within the nucleus, thus presumably modifying the genomic landscape and expression patterns by creating altered transcriptional complexes. The lentivirus genus is part of the retrovirus family, wherein members were proven to affect the expression of adjacent genes (102) upon genome insertion in regions transcriptionally active (103, 104),

particularly in the presence of a strong promoter (105). The integration site was not investigated in this study, although there was a probability of insertional mutagenesis. However, it must have happened in each independently transduced pool of cells having generated leukemia and provided it affected a tumorigenic gene. Even though insertional mutagenesis occurred, the generated tumours are proper human models of CBFA2T3-GLIS2 leukemia since they reflect the pediatric disease.

HSPCs, transformed in the experiments by the overexpression of CBFA2T3-GLIS2, are reputed to be difficult to culture with a span no longer than a few weeks (106). The *in vitro* proliferation capacity of CBFA2T3-GLIS2-transduced cells was extensively studied (figure 2). The proportion of GFP-positive cells was recurrently assessed by flow cytometry every seven days and led to the observation of a sudden steep increase for some CBFA2T3-GLIS2 wells from day 28. The proportion of GFP-positive cells in most wells that showed a rapid increase began nevertheless to reduce at day 77. Three wells were entirely composed of CBFA2T3-GLIS2-transduced cells and remained as such until the experiment was terminated at day 91. This period is similar to the 84 days reported by another group of researchers that introduced CBFA2T3-GLIS2 into human CB HSPCs by lentiviral transduction as well (69). The extensive proliferation observed in the transduced cells might sensibly be due to a gained proliferative advantage over the non-transduced ones resulting from the fusion gene insertion. Correspondingly, transduced cells having not acquired any proliferative advantage might explain the lower proliferation rate seen in the CBFA2T3-GLIS2 wells without increases in GFP. As mentioned, HSPCs are heterogeneous regarding cell type and level of differentiation. Self-renewal and differentiation potentialities diminish through differentiation and are influenced by intrinsic

and external factors such as epigenetic modulations and environmental cytokines, respectively (107). Since the medium was identical, the sudden decrease in GFP seen in some CBFA2T3-GLIS2 wells is probably linked to intrinsic factors related to the level of differentiation of the transduced cells. Indeed, the molecular landscape evolves in accordance with differentiation (108). It is then necessary that the lentiviral construct enters the cell type entirely sensitive to the fusion gene effects to overly transform with a continuous and long-term self-renewal activity. Undoubtedly, defining the cell type from which leukemia originated in the present models would increase the generation rate in future experiments.

In the end, it is possible that the cells inside the wells that reached 100% in GFP were immortal and could have been cultured for more than 91 days. Indeed, a more extended period of culture was observed with the synthetic NUP98-KDM5A AMKL model also generated in the laboratory (92).

Immunophenotyping was achieved to detail the extensive proliferation seen in culture. A comparative analysis of data from day 28 and day 35 revealed a sudden increase in GFP as well as the appearance of CD56 accompanied by CD117^{Dim}. It suggests that NCAM1 could serve as an early indicator of cancerous transformation as early as 28 days following transduction.

In another experimental setting, a pool of CB HSPCs from four donors was transduced by either the empty vector or the fusion gene vector before being injected into immunodeficient mice the next day (figure 3). As a result, six mice out of ten developed leukemia, with half presenting AMKL while the others were deemed undetermined AML. The generation of CBFA2T3-GLIS2 AML in synthetic human models will enforce

awareness in clinics that this fusion gene can be encountered in patients presenting with AML. It will require physicians to add this fusion gene in the AML panel detection carried out by either RT-PCR or RNA sequencing in order to establish the correct diagnosis. Most CBFA2T3-GLIS2 AML patients were in adolescence and young adulthood at diagnosis in two different cohorts (66, 71), strengthening the hypothesis that age dictates the leukemia subtype due to changes in cell ontogeny (71). As previously stated, the genomic landscape of driver mutations in both AML and AMKL is age-specific (30, 53). Consequently, the same oncogene such as CBFA2T3-GLIS2 may exert different phenotypes through interactions with distinctive transcriptional complexes depending on the age. Nevertheless, further epidemiological data will be required to ascertain the average age at diagnosis for both presentations of CBFA2T3-GLIS2 leukemia.

The median latency was 20,3 and 28,1 weeks for the AMKL and AML tumours, respectively (figure 3B). The inducible mouse model observed similar corresponding latencies of 23,4 and 35,9 weeks (71). There were up to 19 replicates in the latter study, suggesting that the latencies obtained might still be alike if more than ten replicates had been used in this experiment. More importantly, it bolsters the observation of a longer latency for CBFA2T3-GLIS2 AML.

Splenomegaly resulting from nodular infiltration is a feature of pediatric AML (26). In this regard, a CFA2T3-GLIS2 AMKL mouse presented with an enlarged spleen of 0,385g (figure 3B). The condition was not present in the other primary mice, even if blast cells were found within the organ. Another group observed the same phenomenon and had a single patient-derived xenograft model with splenomegaly (67). In order to understand the conditions for spleen infiltration, the gene expression profile of splenic

blast cells, which can also be found in a lower quantity in animals without splenomegaly, could be compared in a future study. Also, the comparison could include bone marrow blast cells. Indeed, spleen infiltration probably implies distinctive abilities attributed to cell evasion allowing blast cells to disseminate throughout the body from the bone marrow to cell homing to keep them at this specific organ. To this end, researchers studying multiple myeloma, a hematological neoplasm characterized by an expansion of plasma cells, compared the expression profiles of the tumour cells normally found inside the bone marrow with those circulating in peripheral blood. They found that the expression profile of the circulating tumour cells was similar to those residing within the bone marrow, except for a few genes involved in interferon and inflammatory response, hypoxia, cell cycle and migration (109). This finding suggests that circulating cells evolved from a clone within the bone marrow to form a novel clonal branch in peripheral blood. It would then not be surprising to see the same phenomenon in AML considering the concept of clonal evolution wherein the disease continuously evolves through the creation of new clones, each resulting from a novel genetic abnormality (110, 111). However, in the absence of evidence, there is a possibility that the blast cells found in the spleen do not transit by the bone marrow following injection but directly migrate to the spleen and engage in cell replication. A system allowing the *in vivo* tracking of cells such as bioluminescence imaging would provide answers regarding cell dissemination (112). This approach has already been used with immune cells and could easily be implemented in the near future.

Flow cytometry analysis determined the proportion of blast cells within the bone marrow and the spleen, but the leukemic burden was not estimated in other organs of interest, such as the nervous system and liver. The level of infiltration varies considerably

among the mice with regard to the tissue (i.e., bone marrow or spleen), leukemia subtype and even within the latter (figure 3B). As a striking example, the I730^{AML} presented without blast cells in the spleen, whereas the I749^{AMKL} had a 98% infiltration. On the contrary, the spleen infiltrations of the Syngeneic female mice were used in this study, thereby reducing the possibility that the host environment appreciably influenced disease evolution because of interactions with human grafted cells. The range of infiltration in hematopoietic organs was wide, but leukemic invasion in non-hematopoietic organs will need to be investigated with luciferase-expressing xenograft models to quantify the total tumour burden better.

The megakaryocyte markers (CD41; CD61) and the stemness markers (CD34; CD117) were assessed in the blast cells population identified as hCD45+GFP+. Also, CD56 and CD71 provided further information on the surface expression landscape (figure 3C). The difference in the expression level of megakaryocyte markers in both tissues is striking and unequivocally led to the classification of tumours into AMKL and undetermined AML. The shallow expression of CD41 and CD61 seen by flow cytometry for AML tumours could be due to artefactual adherence of platelets to blast cells or differential recycling of integrins (113). Accordingly, the proportion of megakaryocytic markers must be at least 30% not to be deemed spurious (114) and 50% for a tumour to be clearly classified as AMKL (47). CD34 was found in the I731^{AMKL} and I749^{AMKL} and to a lesser extent in the I750^{AML}. The expression of CD34 in AMKL patients harbouring blast cells at an immature stage is putative (115). Blast cells of all tumours expressed CD71 and CD117 on their surface. CD71 is usually expressed on the surface of erythroid progenitors and is encountered in AML and AMKL (116). CD56 was associated with

CBFA2T3-GLIS2 patients (82) and was broadly observed in the tumours but absent for one AML tumour. As shown, CD56 expression was more predominant in the I749^{AMKL} that had splenomegaly, hence invoking the possible role of NCAM1 in cell dissemination. Functional studies silencing NCAM1 might reveal a correlation with spleen invasion.

Cell morphology has long been used to classify AML despite the associated difficulties of identifying the proper lineage (117). Cytocentrifugation was performed with cells harvested from the bone marrow and the spleen, followed by a Giemsa staining (figure 3C). Blast cell morphology was evident in both tissues and subtypes, and cells of the AMKL immunophenotype had blebs on their surface. Cell blebs are a distinctive characteristic of megakaryocytes and were observed in CBFA2T3-GLIS2 AMKL patients (82). Blast cells in CBFA2T3-GLIS2-positive patients were reportedly pleomorphic, ranging from small cells having scanty cytoplasm and inconspicuous nucleoli to large cells with abundant cytoplasm and prominent nucleoli (82). This description matches the observations made for all tumours without consideration for leukemia classification. Also, there seem to be cells that resemble maturing erythroblasts, which are described as being smaller with condensed chromatin having a hyperchromatic “ink-dot” appearance (118). Even though suitable immunophenotyping for this lineage was not achieved, the concomitant existence of an erythroid component in AMKL is well established (119).

Additional noteworthy features of this model are that the phenotype was consistent in recipient mice (figure 4), and the latency was reduced through passages (figure 4A). As a result, it demonstrated that leukemia self-renewed since the tumours could be serially transplanted and regenerate the disease. The latency was significantly different between the groups, although only some primary tumours were passed into tertiary

recipients, and the number of injected cells was 50 times higher for the secondary and tertiary recipients. An equal number of injected cells with equal recipients would have given an unbiased statistical analysis between the groups. In this regard, the comparison between the control and primary groups is the most accurate since an equal number of cells was injected. Nevertheless, as traditionally observed (120, 121), a latency that is negatively correlated with the number of passages is unquestionably visible in this experiment. Splenomegaly was present in all the recipient mice issued from the I749^{AMKL}, with spleen weights significantly greater than the control, I732^{AMKL} and I750^{AML} groups (figure 4B). Flow cytometry revealed that the bone marrow was significantly more infiltrated than the spleen (figure 4C). Unsurprisingly, the descendants of the I749^{AMKL} had spleen infiltrations significantly higher, nearing the 100% proportion in blast cells. Surface expression and cell morphology were indistinguishable from the primary mice (figure 4D), thereby indicating that observable heterogeneity of a primary tumour can be passed along to at least a tertiary recipient.

The PCA and the hierarchical clustering analysis grouped the tumours according to the genetic subtype (figure 5-6). Therefore, tumour heterogeneity may be more subtle than a rough observation of cell morphology and surface expression. For instance, a clinical case reported blast cells without megakaryocytic markers following a first flow cytometry analysis. However, it turned out there was an intracytoplasmic expression of CD61 that was later on detected after cell permeabilization (122). Therefore, the classification of the tumours might be inappropriate because of an intracytoplasmic expression of megakaryocytic markers by very immature blast cells. Although available laboratory techniques allowing cell permeabilization could have been used to determine

the lineage commitment, RNA sequencing was preferred due to the possibility of efficiently obtaining a comprehensive view of the transcriptional profile (table I). Strikingly, the tumours of both designated classifications strongly expressed megakaryocytic markers, such as CD41 and CD61, when compared with normal CB HSPCs and two AML samples. Furthermore, the myeloid transcriptional landscape was substantially different from the one shown by the AML samples. Finally, it seems that the generated tumours share more similarities than thought, which go far beyond cell morphology and surface expression. It would thus be sensible to test for gene expression or intracytoplasmic protein expression, considering that post-translational modifications and protein trafficking may alter antigen presentation on the surface, especially with cancer.

CGH/SNP was performed for several tumours at different passages to find genetic lesions that blast cells may be carrying out (table II). Although a del(11q) and a dup(1p) were discovered in two different tumours, the del(20q) found in the K444^{AML}, a secondary of the I750^{AML}, was the sole chromosomal abnormality to hold a clinical significance of diagnosis (100). Indeed, the del(20q) has been associated with secondary AML cases accounting for 8% of them (123). On the other hand, the del(11p) has been reported in pediatric AML with unknown significance, albeit being more frequent than the del(20q) (124). The third of CBFA2T3-GLIS2 cases present with a normal karyotype (70). A similar observation was made in this study, though there are few primary tumours. Overall, these chromosomal abnormalities may likely influence the course of leukemia and be simply undocumented. Manifestly, functional studies on such large genomic regions involving a myriad of genes were beyond the scope of this study.

Splenic blast cells of the K423^{AMKL}, an I749^{AMKL} secondary recipient, were cultured for six days then transplanted into immunodeficient mice, thus regenerating leukemia (figure 7). Despite a slightly lower proportion of cells expressing CD34 in culture, the mouse immunophenotype profile at sacrifice was reasonably similar. These results indicate that repopulating cells having stemness properties were either maintained or injected in a sufficient number before proliferating once more inside the recipient hosts. An LDA was carried out with splenic cells of the primary tumour I749^{AMKL} to define the frequency of LICs (figure 8). It was found using the ELDA software (101) that one in 3445 cells was able to regenerate leukemia (95% confidence interval of 1/8511 to 1/1395). The LDA is based on the single-hit Poisson model that posits that one LIC is sufficient to generate leukemia (125). The frequency might be underestimated in this experiment because a minimal cell dose was not obtain in which none of the mice would have developed leukemia. In order to obtain more informative results, lower cell doses should be tested, and the number of replicates should be increased to account for technical variability (126). An interesting way to study repopulating cells that could be conducted in conjunction with an LDA is through single-cell sorting using flow cytometry (127). Indeed, this less frequent population could be specifically studied once isolated. Altogether, the ability to cultivate synthetic blast cells regenerating leukemia is a milestone in CBFA2T3-GLIS2 research and opens the door to functional studies or drug screens for which a period of culture is necessary (128). Furthermore, knowing the frequency of repopulating cells facilitates targeting these critical cells ascribed to cancer relapse and resistance. (129).

Conclusion

This study reports the successful generation of a human CBFA2T3-GLIS2 AML model in immunodeficient mice, phenocopying the basic clinical features. It was achieved by overexpression of the fusion gene inserted into the genome of human CB HSPCs by lentiviral transduction. Although the gene expression profile suggested a homogeneous disease reflecting the AMKL subtype, only half the tumours presented the associated surface markers. The immunophenotype revealed the expression of CD56, and *in vitro* evidence suggests a possible correlation with extensive cell proliferation. Leukemia could be serially transplanted into tertiary recipient mice without conspicuously altering blast cells, as evidenced by cell morphology, flow cytometry and CGH/SNP. More importantly, the CBFA2T3-GLIS2-transduced cells could be cultured for six days and regenerated leukemia when transplanted into recipient hosts. Because of the paucity of patient samples, this achievement paves the way to large-scale functional studies requiring a period of culture such as drug screens. Any discovery involving a therapeutic agent would undoubtedly be a breakthrough considering the dreadful prognosis associated with this fusion gene. Indeed, the LDA performed in this study points towards an aggressive disease harbouring a high frequency of repopulating cells. Overall, it is demonstrated for the first time that overexpression of CBFA2T3-GLIS2 in human CB HSPCs leads to aggressive *de novo* megakaryocytic leukemia in immunodeficient mice, recapitulating the human disease. This outcome is likely promoted by favourable differentiation levels of the precursor cells and culture conditions.

Acknowledgements

First and above all, I want to express my sincere gratitude to the department of pathology for the friendly and supportive environment, essentially provided by Professor Edith Zorychta and Dr. Hua Ling. I definitively enjoyed being amongst keen people with overflowing potential from all around the world. I want to thank my supervisor Dr. Sonia Cellot for having given me the chance to do research and experiments composing such a venturing study. Also, thanks to my research advisors Dr. Alan Spatz and Dr. Josée Hébert for their guidance and Dr. Kolja Eppert as a thesis examiner. I thank Mathieu Roussy for having done the cloning of CBFA2T3-GLIS2 before my arrival in the laboratory and for the isolation and transduction of CB HSPCs. Obviously, it allowed the generation of the synthetic model of human CBFA2T3-GLIS2 leukemia. Louise Laramée was also involved in the cloning of CBFA2T3-GLIS2 while providing help at diverse moments, so I thank her for that. This study could not have been brought this far without Sophie Cardin's sustained support and who took part in every experiment, namely in injection, animal handling and sacrifice. Thanks to Mélanie Bilodeau for having assisted me in the planning of experiments and the analysis of results and to Alexandre Rouette and the genomics platform of IRIC for the RNA sequencing. I want to thank the animal facility of CHU Saint-Justine for the care and housing of mice as well as the cytogenetics laboratory and France Léveillé for having performed the CGH/SNP analysis. Finally, special thanks to the Leukemia & Lymphoma Society of Canada for having funded Dr. Cellot generously. Undoubtedly, the funding received was instrumental in completing this study.

Author's contribution

The author participated in the planning of all experiments along with the other laboratory members involving Louise Laramée, Sophie Cardin, Mélanie Bilodeau, Alexandre Rouette and Mathieu Roussy. The author did the illustrations, graphs, tables and statistical analyses, including the LDA. Luc Boulianne performed the sacrifice together with the autopsy of mice helped by Sophie Cardin. The author did the whole RT-PCR experiment leading to the detection of the fusion gene transcript. Luc Boulianne developed and experimented the protocol for the immunofluorescence imaging procedure. Microscopy was entirely achieved by Luc Boulianne from the acquisition to the processing of images. Luc Boulianne did flow cytometry analyses and the associated figures presented throughout this study. Cell culture performed at several moments throughout this study was carried out by the author. The author did RNA and DNA isolation prior to sequencing and CGH/SNP, respectively.

References

1. Doulatov S, Notta F, Laurenti E, Dick JE. Hematopoiesis: a human perspective. *Cell stem cell*. 2012;10(2):120-36.
2. Pinho S, Frenette PS. Haematopoietic stem cell activity and interactions with the niche. *Nature Reviews Molecular Cell Biology*. 2019;20(5):303-20.
3. Gao X, Xu C, Asada N, Frenette PS. The hematopoietic stem cell niche: from embryo to adult. *Development (Cambridge, England)*. 2018;145(2):dev139691.
4. Wilson A, Laurenti E, Oser G, van der Wath RC, Blanco-Bose W, Jaworski M, et al. Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell*. 2008;135(6):1118-29.

5. Lessard J, Faubert A, Sauvageau G. Genetic programs regulating HSC specification, maintenance and expansion. *Oncogene*. 2004;23(43):7199-209.
6. Civin CI, Strauss LC, Brovall C, Fackler MJ, Schwartz JF, Shaper JH. Antigenic analysis of hematopoiesis. III. A hematopoietic progenitor cell surface antigen defined by a monoclonal antibody raised against KG-1a cells. *J Immunol*. 1984;133(1):157-65.
7. Sauvageau G, Iscove NN, Humphries RK. In vitro and in vivo expansion of hematopoietic stem cells. *Oncogene*. 2004;23(43):7223-32.
8. Wang JC, Doedens M, Dick JE. Primitive human hematopoietic cells are enriched in cord blood compared with adult bone marrow or mobilized peripheral blood as measured by the quantitative in vivo SCID-repopulating cell assay. *Blood*. 1997;89(11):3919-24.
9. Ballen KK, Gluckman E, Broxmeyer HE. Umbilical cord blood transplantation: the first 25 years and beyond. *Blood*. 2013;122(4):491-8.
10. Shearer WT, Lubin BH, Cairo MS, Notarangelo LD. Cord Blood Banking for Potential Future Transplantation. *Pediatrics*. 2017;140(5).
11. Miller PH, Knapp DJ, Eaves CJ. Heterogeneity in hematopoietic stem cell populations: implications for transplantation. *Current opinion in hematology*. 2013;20(4):257-64.
12. Gluckman E, Rocha V, Boyer-Chammard A, Locatelli F, Arcese W, Pasquini R, et al. Outcome of cord-blood transplantation from related and unrelated donors. Eurocord Transplant Group and the European Blood and Marrow Transplantation Group. *N Engl J Med*. 1997;337(6):373-81.

13. Rocha V, Wagner JE, Jr., Sobocinski KA, Klein JP, Zhang MJ, Horowitz MM, et al. Graft-versus-host disease in children who have received a cord-blood or bone marrow transplant from an HLA-identical sibling. Eurocord and International Bone Marrow Transplant Registry Working Committee on Alternative Donor and Stem Cell Sources. *N Engl J Med*. 2000;342(25):1846-54.
14. Wagner JE, Rosenthal J, Sweetman R, Shu XO, Davies SM, Ramsay NK, et al. Successful transplantation of HLA-matched and HLA-mismatched umbilical cord blood from unrelated donors: analysis of engraftment and acute graft-versus-host disease. *Blood*. 1996;88(3):795-802.
15. Turcotte LM, Liu Q, Yasui Y, Arnold MA, Hammond S, Howell RM, et al. Temporal Trends in Treatment and Subsequent Neoplasm Risk Among 5-Year Survivors of Childhood Cancer, 1970-2015. *Jama*. 2017;317(8):814-24.
16. Linabery AM, Ross JA. Trends in childhood cancer incidence in the U.S. (1992-2004). *Cancer*. 2008;112(2):416-32.
17. Puumala SE, Ross JA, Aplenc R, Spector LG. Epidemiology of Childhood Acute Myeloid Leukemia. *Pediatric blood & cancer*. 2013;60(5):728-33.
18. Tarlock K, Meshinchi S. Pediatric acute myeloid leukemia: biology and therapeutic implications of genomic variants. *Pediatr Clin North Am*. 2015;62(1):75-93.
19. Creutzig U, Kutny MA, Barr R, Schlenk RF, Ribeiro RC. Acute myelogenous leukemia in adolescents and young adults. *Pediatr Blood Cancer*. 2018;65(9):e27089.
20. Armstrong GT, Chen Y, Yasui Y, Leisenring W, Gibson TM, Mertens AC, et al. Reduction in Late Mortality among 5-Year Survivors of Childhood Cancer. *N Engl J Med*. 2016;374(9):833-42.

21. Pui CH, Carroll WL, Meshinchi S, Arceci RJ. Biology, risk stratification, and therapy of pediatric acute leukemias: an update. *J Clin Oncol*. 2011;29(5):551-65.
22. Granfeldt Østgård LS, Medeiros BC, Sengeløv H, Nørgaard M, Andersen MK, Dufva IH, et al. Epidemiology and Clinical Significance of Secondary and Therapy-Related Acute Myeloid Leukemia: A National Population-Based Cohort Study. *J Clin Oncol*. 2015;33(31):3641-9.
23. Rose-Inman H, Kuehl D. Acute Leukemia. *Hematology/oncology clinics of North America*. 2017;31(6):1011-28.
24. Dixit A, Chatterjee T, Mishra P, Kannan M, Choudhry DR, Mahapatra M, et al. Disseminated intravascular coagulation in acute leukemia at presentation and during induction therapy. *Clin Appl Thromb Hemost*. 2007;13(3):292-8.
25. Appelbaum FR. 98 - Acute Leukemias in Adults. In: Niederhuber JE, Armitage JO, Doroshow JH, Kastan MB, Tepper JE, editors. *Abeloff's Clinical Oncology (Fifth Edition)*. Philadelphia: Content Repository Only!; 2014. p. 1890-906.
26. Nazemi KJ, Malempati S. Emergency Department Presentation of Childhood Cancer. *Emergency Medicine Clinics of North America*. 2009;27(3):477-95.
27. Papaemmanuil E, Gerstung M, Bullinger L, Gaidzik VI, Paschka P, Roberts ND, et al. Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N Engl J Med*. 2016;374(23):2209-21.
28. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*. 2013;499(7457):214-8.

29. Radtke I, Mullighan CG, Ishii M, Su X, Cheng J, Ma J, et al. Genomic analysis reveals few genetic alterations in pediatric acute myeloid leukemia. *Proceedings of the National Academy of Sciences of the United States of America*. 2009;106(31):12944-9.
30. Bolouri H, Farrar JE, Triche T, Jr., Ries RE, Lim EL, Alonzo TA, et al. The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. *Nature medicine*. 2018;24(1):103-12.
31. Mercher T, Schwaller J. Pediatric Acute Myeloid Leukemia (AML): From Genes to Models Toward Targeted Therapeutic Intervention. *Front Pediatr*. 2019;7:401.
32. Winters AC, Bernt KM. MLL-Rearranged Leukemias-An Update on Science and Clinical Approaches. *Front Pediatr*. 2017;5:4.
33. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016;127(20):2391-405.
34. Mertens F, Johansson B, Fioretos T, Mitelman F. The emerging complexity of gene fusions in cancer. *Nat Rev Cancer*. 2015;15(6):371-81.
35. Creutzig U, van den Heuvel-Eibrink MM, Gibson B, Dworzak MN, Adachi S, de Bont E, et al. Diagnosis and management of acute myeloid leukemia in children and adolescents: recommendations from an international expert panel. *Blood*. 2012;120(16):3187-205.
36. Ross ME, Mahfouz R, Onciu M, Liu HC, Zhou X, Song G, et al. Gene expression profiling of pediatric acute myelogenous leukemia. *Blood*. 2004;104(12):3679-87.

37. Eriksson A, Lennartsson A, Lehmann S. Epigenetic aberrations in acute myeloid leukemia: Early key events during leukemogenesis. *Experimental Hematology*. 2015;43(8):609-24.
38. McMahon KA, Hiew SY, Hadjur S, Veiga-Fernandes H, Menzel U, Price AJ, et al. Mll has a critical role in fetal and adult hematopoietic stem cell self-renewal. *Cell stem cell*. 2007;1(3):338-45.
39. Jude CD, Climer L, Xu D, Artinger E, Fisher JK, Ernst P. Unique and independent roles for MLL in adult hematopoietic stem cells and progenitors. *Cell stem cell*. 2007;1(3):324-37.
40. Guenther MG, Lawton LN, Rozovskaia T, Frampton GM, Levine SS, Volkert TL, et al. Aberrant chromatin at genes encoding stem cell regulators in human mixed-lineage leukemia. *Genes Dev*. 2008;22(24):3403-8.
41. Dorrance AM, Liu S, Yuan W, Becknell B, Arnoczky KJ, Guimond M, et al. Mll partial tandem duplication induces aberrant Hox expression in vivo via specific epigenetic alterations. *J Clin Invest*. 2006;116(10):2707-16.
42. Milne TA, Martin ME, Brock HW, Slany RK, Hess JL. Leukemogenic MLL fusion proteins bind across a broad region of the Hox a9 locus, promoting transcription and multiple histone modifications. *Cancer Res*. 2005;65(24):11367-74.
43. Alharbi RA, Pettengell R, Pandha HS, Morgan R. The role of HOX genes in normal hematopoiesis and acute leukemia. *Leukemia*. 2013;27(5):1000-8.
44. Apperley JF. Chronic myeloid leukaemia. *Lancet*. 2015;385(9976):1447-59.
45. Houshmand M, Simonetti G, Circosta P, Gaidano V, Cignetti A, Martinelli G, et al. Chronic myeloid leukemia stem cells. *Leukemia*. 2019;33(7):1543-56.

46. Koschmieder S, Vetrie D. Epigenetic dysregulation in chronic myeloid leukaemia: A myriad of mechanisms and therapeutic options. *Semin Cancer Biol.* 2018;51:180-97.
47. Klairmont MM, Hoskoppal D, Yadak N, Choi JK. The Comparative Sensitivity of Immunohistochemical Markers of Megakaryocytic Differentiation in Acute Megakaryoblastic Leukemia. *American journal of clinical pathology.* 2018;150(5):461-7.
48. Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DA, Gralnick HR, et al. Criteria for the diagnosis of acute leukemia of megakaryocyte lineage (M7). A report of the French-American-British Cooperative Group. *Ann Intern Med.* 1985;103(3):460-2.
49. Hahn AW, Li B, Prouet P, Giri S, Pathak R, Martin MG. Acute megakaryocytic leukemia: What have we learned. *Blood reviews.* 2016;30(1):49-53.
50. Pagano L, Pulsoni A, Vignetti M, Mele L, Fianchi L, Petti MC, et al. Acute megakaryoblastic leukemia: experience of GIMEMA trials. *Leukemia.* 2002;16(9):1622-6.
51. Gassmann W, Löffler H. Acute megakaryoblastic leukemia. *Leuk Lymphoma.* 1995;18 Suppl 1:69-73.
52. Giri S, Pathak R, Prouet P, Li B, Martin MG. Acute megakaryocytic leukemia is associated with worse outcomes than other types of acute myeloid leukemia. *Blood.* 2014;124(25):3833-4.
53. de Rooij JD, Branstetter C, Ma J, Li Y, Walsh MP, Cheng J, et al. Pediatric non-Down syndrome acute megakaryoblastic leukemia is characterized by distinct genomic subsets with varying outcomes. *Nature genetics.* 2017;49(3):451-6.

54. Mesa RA, Li CY, Ketterling RP, Schroeder GS, Knudson RA, Tefferi A. Leukemic transformation in myelofibrosis with myeloid metaplasia: a single-institution experience with 91 cases. *Blood*. 2005;105(3):973-7.
55. Radaelli F, Mazza R, Curioni E, Ciani A, Pomati M, Maiolo AT. Acute megakaryocytic leukemia in essential thrombocythemia: an unusual evolution? *Eur J Haematol*. 2002;69(2):108-11.
56. Dang J, Nance S, Ma J, Cheng J, Walsh MP, Vogel P, et al. AMKL chimeric transcription factors are potent inducers of leukemia. *Leukemia*. 2017;31(10):2228-34.
57. Gruber TA, Downing JR. The biology of pediatric acute megakaryoblastic leukemia. *Blood*. 2015;126(8):943-9.
58. Malinge S, Ragu C, Della-Valle V, Pisani D, Constantinescu SN, Perez C, et al. Activating mutations in human acute megakaryoblastic leukemia. *Blood*. 2008;112(10):4220-6.
59. Walters DK, Mercher T, Gu TL, O'Hare T, Tyner JW, Loriaux M, et al. Activating alleles of JAK3 in acute megakaryoblastic leukemia. *Cancer cell*. 2006;10(1):65-75.
60. Masetti R, Guidi V, Ronchini L, Bertuccio NS, Locatelli F, Pession A. The changing scenario of non-Down syndrome acute megakaryoblastic leukemia in children. *Critical Reviews in Oncology/Hematology*. 2019;138:132-8.
61. Hasle H, Clemmensen IH, Mikkelsen M. Risks of leukaemia and solid tumours in individuals with Down's syndrome. *Lancet*. 2000;355(9199):165-9.
62. Labuhn M, Perkins K, Matzk S, Varghese L, Garnett C, Papaemmanuil E, et al. Mechanisms of Progression of Myeloid Preleukemia to Transformed Myeloid Leukemia in Children with Down Syndrome. *Cancer cell*. 2019;36(2):123-38.e10.

63. Yoshida K, Toki T, Okuno Y, Kanezaki R, Shiraishi Y, Sato-Otsubo A, et al. The landscape of somatic mutations in Down syndrome-related myeloid disorders. *Nature genetics*. 2013;45(11):1293-9.
64. de Rooij JD, Masetti R, van den Heuvel-Eibrink MM, Cayuela JM, Trka J, Reinhardt D, et al. Recurrent abnormalities can be used for risk group stratification in pediatric AMKL: a retrospective intergroup study. *Blood*. 2016;127(26):3424-30.
65. Gruber TA, Larson Gedman A, Zhang J, Koss CS, Marada S, Ta HQ, et al. An Inv(16)(p13.3q24.3)-encoded CBFA2T3-GLIS2 fusion protein defines an aggressive subtype of pediatric acute megakaryoblastic leukemia. *Cancer cell*. 2012;22(5):683-97.
66. Masetti R, Pigazzi M, Togni M, Astolfi A, Indio V, Manara E, et al. CBFA2T3-GLIS2 fusion transcript is a novel common feature in pediatric, cytogenetically normal AML, not restricted to FAB M7 subtype. *Blood*. 2013;121(17):3469-72.
67. Thiollier C, Lopez CK, Gerby B, Ignacimoutou C, Poglio S, Duffourd Y, et al. Characterization of novel genomic alterations and therapeutic approaches using acute megakaryoblastic leukemia xenograft models. *The Journal of experimental medicine*. 2012;209(11):2017-31.
68. Masetti R, Bertuccio SN, Astolfi A, Chiarini F, Lonetti A, Indio V, et al. Hh/Gli antagonist in acute myeloid leukemia with CBFA2T3-GLIS2 fusion gene. *Journal of hematology & oncology*. 2017;10(1):26.
69. Smith JL, Ries RE, Hylkema T, Alonzo TA, Gerbing RB, Santaguida MT, et al. Comprehensive Transcriptome Profiling of Cryptic CBFA2T3-GLIS2 Fusion-Positive AML Defines Novel Therapeutic Options: A COG and TARGET Pediatric AML Study. *Clin Cancer Res*. 2019.

70. Hara Y, Shiba N, Ohki K, Tabuchi K, Yamato G, Park MJ, et al. Prognostic impact of specific molecular profiles in pediatric acute megakaryoblastic leukemia in non-Down syndrome. *Genes, chromosomes & cancer*. 2017;56(5):394-404.
71. Lopez CK, Noguera E, Stavropoulou V, Robert E, Aid Z, Ballerini P, et al. Ontogenic Changes in Hematopoietic Hierarchy Determine Pediatric Specificity and Disease Phenotype in Fusion Oncogene-Driven Myeloid Leukemia. *Cancer Discov*. 2019;9(12):1736-53.
72. Lindberg SR, Olsson A, Persson AM, Olsson I. The Leukemia-associated ETO homologues are differently expressed during hematopoietic differentiation. *Experimental hematology*. 2005;33(2):189-98.
73. Fischer MA, Moreno-Miralles I, Hunt A, Chyla BJ, Hiebert SW. Myeloid translocation gene 16 is required for maintenance of haematopoietic stem cell quiescence. *The EMBO journal*. 2012;31(6):1494-505.
74. Chyla BJ, Moreno-Miralles I, Steapleton MA, Thompson MA, Bhaskara S, Engel M, et al. Deletion of Mtg16, a target of t(16;21), alters hematopoietic progenitor cell proliferation and lineage allocation. *Molecular and cellular biology*. 2008;28(20):6234-47.
75. Stadhouders R, Cico A, Stephen T, Thongjuea S, Kolovos P, Baymaz HI, et al. Control of developmentally primed erythroid genes by combinatorial co-repressor actions. *Nature communications*. 2015;6:8893.
76. Hamlett I, Draper J, Strouboulis J, Iborra F, Porcher C, Vyas P. Characterization of megakaryocyte GATA1-interacting proteins: the corepressor ETO2 and GATA1 interact to regulate terminal megakaryocyte maturation. *Blood*. 2008;112(7):2738-49.

77. Cassandri M, Smirnov A, Novelli F, Pitolli C, Agostini M, Malewicz M, et al. Zinc-finger proteins in health and disease. *Cell Death Discovery*. 2017;3:17071.
78. Thirant C, Ignacimoutou C, Lopez CK, Diop M, Le Mouel L, Thiollier C, et al. ETO2-GLIS2 Hijacks Transcriptional Complexes to Drive Cellular Identity and Self-Renewal in Pediatric Acute Megakaryoblastic Leukemia. *Cancer cell*. 2017;31(3):452-65.
79. Thirant C, Lopez C, Malinge S, Mercher T. Molecular pathways driven by ETO2-GLIS2 in aggressive pediatric leukemia. *Molecular & cellular oncology*. 2017;4(6):e1345351.
80. Knudsen KJ, Rehn M, Hasemann MS, Rapin N, Bagger FO, Ohlsson E, et al. ERG promotes the maintenance of hematopoietic stem cells by restricting their differentiation. *Genes & development*. 2015;29(18):1915-29.
81. Stankiewicz MJ, Crispino JD. ETS2 and ERG promote megakaryopoiesis and synergize with alterations in GATA-1 to immortalize hematopoietic progenitor cells. *Blood*. 2009;113(14):3337-47.
82. Masetti R, Bertuccio SN, Pession A, Locatelli F. CBFA2T3-GLIS2-positive acute myeloid leukaemia. A peculiar paediatric entity. *Br J Haematol*. 2019;184(3):337-47.
83. Crossin KL, Krushel LA. Cellular signaling by neural cell adhesion molecules of the immunoglobulin superfamily. *Dev Dyn*. 2000;218(2):260-79.
84. Edelman GM, Jones FS. Gene regulation of cell adhesion: a key step in neural morphogenesis. *Brain Res Brain Res Rev*. 1998;26(2-3):337-52.
85. Jin L, Hemperly JJ, Lloyd RV. Expression of neural cell adhesion molecule in normal and neoplastic human neuroendocrine tissues. *Am J Pathol*. 1991;138(4):961-9.

86. Cunningham BA, Hemperly JJ, Murray BA, Prediger EA, Brackenbury R, Edelman GM. Neural cell adhesion molecule: structure, immunoglobulin-like domains, cell surface modulation, and alternative RNA splicing. *Science*. 1987;236(4803):799-806.
87. Xu S, Li X, Zhang J, Chen J. Prognostic value of CD56 in patients with acute myeloid leukemia: a meta-analysis. *J Cancer Res Clin Oncol*. 2015;141(10):1859-70.
88. Raspadori D, Damiani D, Lenoci M, Rondelli D, Testoni N, Nardi G, et al. CD56 antigenic expression in acute myeloid leukemia identifies patients with poor clinical prognosis. *Leukemia*. 2001;15(8):1161-4.
89. Djunic I, Virijevic M, Djurasinovic V, Novkovic A, Colovic N, Kraguljac-Kurtovic N, et al. Prognostic significance of CD56 antigen expression in patients with acute myeloid leukemia. *Med Oncol*. 2012;29(3):2077-82.
90. Sasca D, Szybinski J, Schuler A, Shah V, Heidelberger J, Haehnel PS, et al. Neural cell adhesion molecule 1 (NCAM1; CD56) promotes leukemogenesis and confers drug resistance in AML. *Blood*. 2019.
91. Sykes SM. NCAM1 supports therapy resistance and LSC function in AML. *Blood*. 2019;133(21):2247-8.
92. Cardin S, Bilodeau M, Roussy M, Aubert L, Milan T, Jouan L, et al. Human models of NUP98-KDM5A megakaryocytic leukemia in mice contribute to uncovering new biomarkers and therapeutic vulnerabilities. *Blood Adv*. 2019;3(21):3307-21.
93. Fares I, Chagraoui J, Gareau Y, Gingras S, Ruel R, Mayotte N, et al. Cord blood expansion. Pyrimidoindole derivatives are agonists of human hematopoietic stem cell self-renewal. *Science (New York, NY)*. 2014;345(6203):1509-12.

94. Boitano AE, Wang J, Romeo R, Bouchez LC, Parker AE, Sutton SE, et al. Aryl hydrocarbon receptor antagonists promote the expansion of human hematopoietic stem cells. *Science (New York, NY)*. 2010;329(5997):1345-8.
95. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114-20.
96. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21.
97. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011;12:323.
98. Nicorici D, Şatalan M, Edgren H, Kangaspeska S, Murumägi A, Kallioniemi O, et al. FusionCatcher – a tool for finding somatic fusion genes in paired-end RNA-sequencing data. *bioRxiv*. 2014:011650.
99. Hans CP, Finn WG, Singleton TP, Schnitzer B, Ross CW. Usefulness of anti-CD117 in the flow cytometric analysis of acute leukemia. *Am J Clin Pathol*. 2002;117(2):301-5.
100. Xu X, Bryke C, Sukhanova M, Huxley E, Dash DP, Dixon-Mciver A, et al. Assessing copy number abnormalities and copy-neutral loss-of-heterozygosity across the genome as best practice in diagnostic evaluation of acute myeloid leukemia: An evidence-based review from the cancer genomics consortium (CGC) myeloid neoplasms working group. *Cancer Genet*. 2018;228-229:218-35.
101. Hu Y, Smyth GK. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J Immunol Methods*. 2009;347(1-2):70-8.

102. Nienhuis AW, Dunbar CE, Sorrentino BP. Genotoxicity of retroviral integration in hematopoietic cells. *Mol Ther.* 2006;13(6):1031-49.
103. Beard BC, Dickerson D, Beebe K, Gooch C, Fletcher J, Okbinoglu T, et al. Comparison of HIV-derived lentiviral and MLV-based gammaretroviral vector integration sites in primate repopulating cells. *Mol Ther.* 2007;15(7):1356-65.
104. Schröder AR, Shinn P, Chen H, Berry C, Ecker JR, Bushman F. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell.* 2002;110(4):521-9.
105. Montini E, Cesana D, Schmidt M, Sanvito F, Bartholomae CC, Ranzani M, et al. The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. *J Clin Invest.* 2009;119(4):964-75.
106. Derakhshani M, Abbaszadeh H, Movassaghpour AA, Mehdizadeh A, Ebrahimi-Warkiani M, Yousefi M. Strategies for elevating hematopoietic stem cells expansion and engraftment capacity. *Life Sci.* 2019;232:116598.
107. Zon LI. Intrinsic and extrinsic control of haematopoietic stem-cell self-renewal. *Nature.* 2008;453(7193):306-13.
108. Seita J, Weissman IL. Hematopoietic stem cell: self-renewal versus differentiation. *Wiley Interdiscip Rev Syst Biol Med.* 2010;2(6):640-53.
109. Garcés JJ, Simicek M, Vicari M, Brozova L, Burgos L, Bezdekova R, et al. Transcriptional profiling of circulating tumor cells in multiple myeloma: a new model to understand disease dissemination. *Leukemia.* 2020;34(2):589-603.
110. Welch JS, Ley TJ, Link DC, Miller CA, Larson DE, Koboldt DC, et al. The origin and evolution of mutations in acute myeloid leukemia. *Cell.* 2012;150(2):264-78.

111. Vetrie D, Helgason GV, Copland M. The leukaemia stem cell: similarities, differences and clinical prospects in CML and AML. *Nat Rev Cancer*. 2020;20(3):158-73.
112. Kim JE, Kalimuthu S, Ahn BC. In vivo cell tracking with bioluminescence imaging. *Nucl Med Mol Imaging*. 2015;49(1):3-10.
113. Betz SA, Foucar K, Head DR, Chen IM, Willman CL. False-positive flow cytometric platelet glycoprotein IIb/IIIa expression in myeloid leukemias secondary to platelet adherence to blasts. *Blood*. 1992;79(9):2399-403.
114. Dunphy CH. Comprehensive review of adult acute myelogenous leukemia: cytomorphological, enzyme cytochemical, flow cytometric immunophenotypic, and cytogenetic findings. *J Clin Lab Anal*. 1999;13(1):19-26.
115. Wang SA, Hasserjian RP. Acute Erythroleukemias, Acute Megakaryoblastic Leukemias, and Reactive Mimics: A Guide to a Number of Perplexing Entities. *American journal of clinical pathology*. 2015;144(1):44-60.
116. Liu Q, Wang M, Hu Y, Xing H, Chen X, Zhang Y, et al. Significance of CD71 expression by flow cytometry in diagnosis of acute leukemia. *Leuk Lymphoma*. 2014;55(4):892-8.
117. Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DA, Gralnick HR, et al. Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group. *Br J Haematol*. 1976;33(4):451-8.
118. Moonim MT, Porwit A. CHAPTER 3 - Normal bone marrow histology. In: Porwit A, McCullough J, Erber WN, editors. *Blood and Bone Marrow Pathology (Second Edition)*. Edinburgh: Churchill Livingstone; 2011. p. 45-62.

119. San Miguel JF, Gonzalez M, Cañizo MC, Ojeda E, Orfao A, Caballero MD, et al. Leukemias with megakaryoblastic involvement: clinical, hematologic, and immunologic characteristics. *Blood*. 1988;72(2):402-7.
120. Leibovici J. Serial passage of tumors in mice in the study of tumor progression and testing of antineoplastic drugs. *Cancer Res*. 1984;44(5):1981-4.
121. Stromberg PC, Grants IS, Kociba GJ, Krakowka GS, Mezza LE. Serial syngeneic transplantation of large granular lymphocyte leukemia in F344 rats. *Vet Pathol*. 1990;27(6):404-10.
122. Käfer G, Willer A, Ludwig W, Krämer A, Hehlmann R, Hastka J. Intracellular expression of CD61 precedes surface expression. *Ann Hematol*. 1999;78(10):472-4.
123. Milosevic JD, Puda A, Malcovati L, Berg T, Hofbauer M, Stukalov A, et al. Clinical significance of genetic aberrations in secondary acute myeloid leukemia. *American Journal of Hematology*. 2012;87(11):1010-6.
124. Inaba H, Zhou Y, Abla O, Adachi S, Auvrignon A, Beverloo HB, et al. Heterogeneous cytogenetic subgroups and outcomes in childhood acute megakaryoblastic leukemia: a retrospective international study. *Blood*. 2015;126(13):1575-84.
125. Bonnefoix T, Bonnefoix P, Verdiel P, Sotto JJ. Fitting limiting dilution experiments with generalized linear models results in a test of the single-hit Poisson assumption. *J Immunol Methods*. 1996;194(2):113-9.
126. Macken C. Design and analysis of serial limiting dilution assays with small sample sizes. *J Immunol Methods*. 1999;222(1-2):13-29.

127. Penter L, Dietze K, Bullinger L, Westermann J, Rahn HP, Hansmann L. FACS single cell index sorting is highly reliable and determines immune phenotypes of clonally expanded T cells. *Eur J Immunol.* 2018;48(7):1248-50.
128. Cellot S, Wilhelm BT, Barabé F. Synthetic human leukemia models: towards precision medicine. *Oncotarget.* 2017;8(57):96480-1.
129. Kreso A, Dick JE. Evolution of the cancer stem cell model. *Cell Stem Cell.* 2014;14(3):275-91.

Supplementary data

Cell counts of blast cells in culture for six days regenerating leukemia

Well	Cell count		
	Day 1	Day 3	Day 6
1	287,500	1,137,500	2,664,500
2	287,500	1,062,500	2,963,800
3	287,500	1,112,500	2,503,400
4	287,500	1,095,000	2,788,600
5	287,500	1,080,000	3,073,300
6	287,500	1,297,500	3,533,200
Mean	287,500	1,130,833	2,921,133
Fold	-	3.93	10.16

Supplementary table I: Cell counts of splenic cells harvested from the tumour

K423^{AMKL}. The mean cell count was calculated at each time point together with the extent of the increase. The number of cells was ten times higher after six days of culture.