The Functional Role of NCAM1 in Pediatric Acute Megakaryoblastic Leukemia (AMKL) Associated with the CBFA2T3::GLIS2 Gene Fusion

Emma Rose Cheetham
Department of Pathology
McGill University, Montréal

March 2023

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

Table of contents

Abstract

Acute megakaryoblastic leukemia (AMKL) is a rare subtype of acute myeloid leukemia (AML) affecting children (< 3 years old). It is characterized by abnormal megakaryoblasts and extensive myelofibrosis making bone marrow sampling difficult for diagnostics and research. AMKL is hard to treat due to the early relapse rates and resistance to chemotherapy. Recurrent chimeric gene fusions occur in AMKL such as *CBFA2T3::GLIS2* (CG2), associated with poor survival and high expression of the neural cell adhesion molecule 1 (*NCAM1*). *NCAM1* is a cell surface protein expressed within the nervous system and involved in regulation of neurogenesis, proliferation, and cell-cell adhesion. It is also expressed within the hematopoietic system on natural killer, lymphoid, and dendritic cells and its function is less understood. *NCAM1* triggers several signaling cascades including MAPK (ERK) and PI3K (AKT). *NCAM1* expression in AML subtypes is associated with promotion of leukemogenesis, reduced complete remission rates, and poor overall survival. The aim of this study was to get a global profiling of *NCAM1* expression in leukemia and normal hHSPCs and to functionally assess the role of *NCAM1* in sustaining leukemic cells *in vitro* and leukemia engraftment and progression *in vivo*. We have generated human models of AMKL carrying these genetic lesions (*CBFA2T3::GLIS2*) that recapitulate AMKL disease in our laboratory using genetic modification techniques in human cord blood cells and transplantation into immunodeficient mice. *NCAM1* knockdown (KD) was performed *in vitro*, using the M07e cell line carrying the CG2 fusion and *in vivo*, using our mCG2 models. Several functional assays were performed to assess the impact of *NCAM1* KD on proliferation, apoptosis, viability, immunophenotype, and clonogenic potential and to determine its functional role in AMKL. To further understand the function of *NCAM1* in AMKL, *NCAM1* KD was combined with inhibiting downstream MAPK and PI3K pathways. The main findings of this study were that *in vitro NCAM1* KD led to no significant impacts on cell fate outcomes of proliferation, differentiation, clonogenic potential, and cell death and viability. Combining *NCAM1* KD with pharmacological inhibition of MAPK and PI3K pathway inhibitors had no significant impacts either. Finally, it is difficult to conclude the impact of *NCAM1* KD *in vivo*, since one of the two sh*NCAM1* conditions had no leukemic burden and there were varying impacts on cell fate outcomes. These experiments will help in determining the expression patterns and functions of *NCAM1*, so that we can better understand the role it plays in AMKL development and progression. Altogether, the information gathered can further contribute to the development of therapies using *NCAM1* as a target in these high-risk AML subtypes.

Abstract (French)

La leucémie aiguë mégacaryoblastique (LAMK) est un sous-type rare de leucémie myéloïde aiguë (LMA) affectant les enfants (< 3 ans). Elle se caractérise par des mégacaryoblastes anormaux et une myélofibrose étendue rendant le prélèvement de moelle osseuse difficile pour le diagnostic et la recherche. La LAMK est difficile à traiter en raison des taux de rechute précoces et de la résistance à la chimiothérapie. Des gènes de fusion chimériques récurrents sont retrouvés dans les LAMK tels que *CBFA2T3::GLIS2* (CG2), associé à une faible survie et à une expression élevée de la molécule *NCAM1* (neural cell adhesion molecule 1). *NCAM1* est une protéine membranaire localisée à la surface cellulaire. Elle est exprimée dans le système nerveux et impliquée dans la régulation de la neurogenèse, de la prolifération et de l'adhésion cellule-cellule. Elle est également exprimée dans le système hématopoïétique sur les cellules tueuses naturelles (natural killer ou NK), lymphoïdes et dendritiques où sa fonction est moins bien comprise. *NCAM1* déclenche plusieurs cascades de signalisation, notamment MAPK (ERK) et PI3K (AKT). L'expression de *NCAM1* dans les sous-types de LMA est associée à la promotion de la leucémogenèse, à des taux de rémission complète réduits et à une faible survie globale. Le but de cette étude était d'obtenir un profil global de l'expression de *NCAM1* dans la leucémie et les cellules souches et progénitrices hématopoïétiques humaines normales et d'évaluer fonctionnellement le rôle de *NCAM1* dans le maintien des cellules leucémiques *in vitro* et la prise de greffe et la progression de la leucémie *in vivo*. Nous avons généré des modèles humains de LAMK porteurs de ces lésions génétiques (*CBFA2T3::GLIS2*) qui récapitulent la maladie dans notre laboratoire en utilisant des techniques de modification génétique dans des cellules de sang de cordon humain et une transplantation chez des souris immunodéficientes. Les études de perte de fonction de *NCAM1* (KD) ont été réalisées *in vitro*, en utilisant la lignée cellulaire M07e portant la fusion CG2 et *in vivo*, en utilisant nos modèles mCG2. Plusieurs essais cellulaires ont été effectués pour évaluer l'impact de *NCAM1* KD sur la prolifération, l'apoptose, la viabilité, le profil immunophénotypique et le potentiel clonogénique afin de déterminer son rôle fonctionnel dans LAMK. Pour mieux comprendre la fonction de *NCAM1* dans les LAMK, *NCAM1* KD a été combiné avec l'inhibition des voies MAPK et PI3K en aval. Les principales conclusions de cette étude étaient que l'inhibition de *NCAM1 in vitro* n'avait aucun impact significatif sur le destin cellulaire en terme de prolifération, différenciation, potentiel clonogénique, mort et viabilité cellulaire. La combinaison de *NCAM1* KD avec l'inhibition pharmacologique des inhibiteurs des voies MAPK et PI3K n'a pas non plus eu d'impact significatif. Enfin, il est difficile de conclure sur l'impact de *NCAM1* KD *in vivo*, car l'une des deux conditions de sh*NCAM1* n'avait

pas de charge leucémique et il y avait des impacts variables sur les résultats du destin cellulaire. Ces expériences aideront à déterminer les patrons d'expression et les fonctions de *NCAM1*, afin que nous puissions mieux comprendre le rôle qu'il joue dans le développement et la progression de la LAMK. Dans l'ensemble, les informations recueillies pourront contribuer davantage au développement de thérapies utilisant *NCAM1* comme cible dans ces sous-types de LMA à haut risque.

Introduction

Hematopoiesis

The circulatory system within the human body contains around 5L of blood, which makes up around 7% of total body weight [1]. Blood is considered one of the most regenerative tissues within the body, with around 1 trillion new blood cells being formed each day [2]. Blood functions in transporting essential nutrients and oxygen to different cells of the body and eliminating waste products [1]. Hematopoiesis is a process that occurs in the bone marrow (BM) by which all mature blood cells are produced in a hierarchical, stepwise fashion. The blood cells produced belong to either the myeloid cell lineage which includes erythrocytes, monocytes, granulocytes, platelets and megakaryocytes or to the lymphoid cell lineage which includes T- and B- cells and natural killer cells [1] (Figure 1).

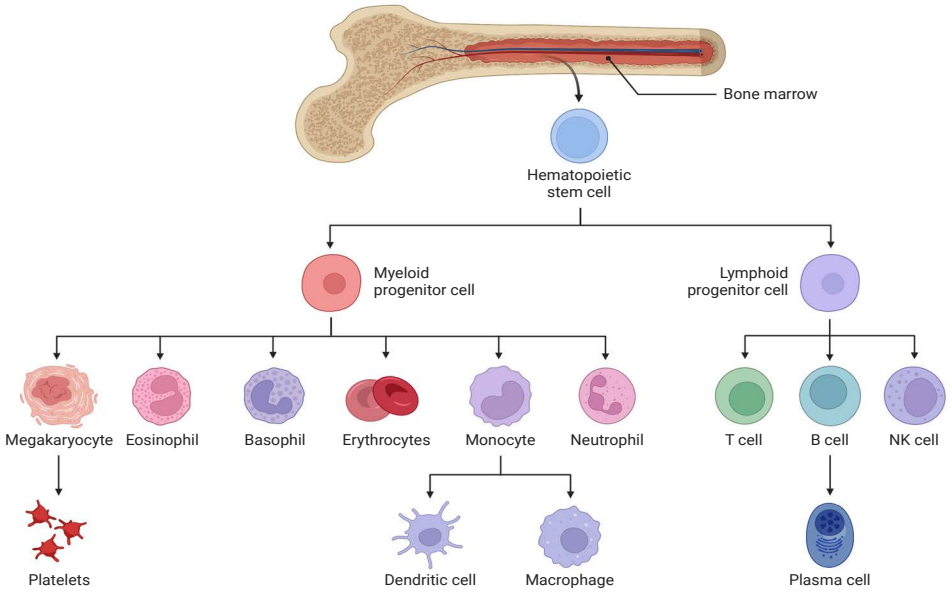


Figure 1: Simplified overview of hematopoiesis in humans. Hematopoiesis or the production of blood cells mainly occurs in the bone marrow and begins with a common multipotent hematopoietic stem cell (HSC). The common HSC can produce all blood progenitor cells belonging to the myeloid and lymphoid cell lineages. Cells belonging to the myeloid cell lineage include megakaryocytes (platelets), eosinophils, basophils, erythrocytes, monocytes (dendritic cells and macrophages) and neutrophils. Cells belonging to the lymphoid cell lineage include T- cells, B- cells (plasma cells) and natural killer cells. Reprinted from "Stem Cell Differentiation from Bone Marrow", by BioRender.com (2022). Retrieved from https://app.biorender.com/biorender-templates.

During embryonic and fetal development, hematopoiesis occurs in the yolk sac and the fetal liver with the bone marrow becoming the principal site of hematopoiesis just before birth [3]. Hematopoiesis does not occur in isolation and is regulated by many components of and in-

teractions with the surrounding microenvironment (or niche), such as lineage specific growth factors, adhesion molecules and transcription factors [4, 5]. All mature blood cells originate from a common multipotent precursor cell which is the hematopoietic stem cell or HSC. HSCs are found in the bone marrow, peripheral blood and cord blood (CB) [6] and can undergo further maturation and proliferation in secondary lymphoid organs such as the spleen and thymus [7]. An HSC has the unique ability to self-renew and is multipotent meaning it can produce identical daughter cells or it can differentiate and mature into all types of mature blood cells from all lineages [4]. HSCs can undergo both symmetric and asymmetrical division. Asymmetrical division occurs when an HSC divides and one daughter cell remains an HSC and one daughter cell differentiates into other mature blood cells. This allows for the maintenance of the HSC pool and prevents depletion. Whereas symmetric division occurs when both daughter cells remain an HSC (expansion) or both undergo differentiation and maturation into any of the blood cells. During division HSCs can also experience cell death which can lead to HSC depletion [7, 8].

HSCs are defined by their ability to fully reconstitute and repopulate bone marrow myeloid and lymphoid cells and this was first demonstrated after HSC BM transplants in mice following irradiation [4, 8]. Cord blood is a well-defined alternative source of HSCs that can be used for transplant. CB HSCs have a higher frequency of long lasting repopulating hematopoietic stem and progenitor cells (HSPCs) compared to BM. The higher frequency of long lasting HSPCs results in increased proliferative capacity, increased engraftment potential and ability to fully reconstitute bone marrow cells following transplant [9]. Compared to other blood cells, HSCs do not have distinct morphological features and can be enriched by colony-forming units or specific antigens. One way to identify HSCs is by enrichment for the cell surface antigen CD34, which is found to be expressed on the majority of HSCs and only on 5% of all blood cells [2, 5]. Disorders of the hematopoietic system include anemias, myeloproliferative and myelodysplastic syndromes and blood cancers such as lymphoma, myeloma, and leukemia [2].

Leukemia

Leukemia is a cancer of the blood forming cells or the bone marrow and is characterized by an early block in progenitor cell differentiation. The initial block in differentiation is caused by an accumulation of genetic alterations and mutations in normal HSCs. When leukemia develops a large number of abnormally functioning cells are produced, termed leukemic cells. These immature leukemic cells experience uncontrolled growth and division and can lead to overcrowding of the bone marrow causing bone marrow failure and normal HSCs to lose the ability for self-renewal, prolif-

eration and survival [10]. Leukemic cells or HSCs can alter the bone marrow microenvironment to promote their survival, and resist cell death and promote the disease progression [4]. Interactions between leukemic cells and the BM niche is also a cause for chemotherapy and drug resistance [11]. The risk factors for developing leukemia include both environmental and inherited factors that can trigger the initial DNA mutations. Environmental risk factors include exposure to carcinogens (e.g. benzene), exposure to tobacco or smoking, exposure to radiation or previous chemotherapy and cancer treatments. Inherited risk factors include certain inherited syndromes or genetic disorders, such as down syndrome, a family history of blood cancers and a compromised immune system [12, 13].

In 2020, leukemia made up around 3.4% of all new cancer diagnoses and 3.8% of all cancer deaths according to Surveillance, Epidemiology and End Results (SEER) program [12]. Different forms of leukemia can be classified based on the 1). lineage commitment and cell type affected (myelogenous or lymphocytic), 2). the speed of disease development (acute or chronic) and 3). the type of genetic alterations present. The hematopoietic system can be divided into two lineages, the myelogenous or myeloid lineage and the lymphocytic or lymphoid lineage. Depending on where the initial block in differentiation occurs in leukemia development this will determine the type of leukemic cells present. Acute leukemias progress quickly and are composed of more immature cells and are more common in children. Chronic leukemias develop more slowly, with leukemic cells retaining the ability to differentiate and are more common in adults [14]. Leukemias are made up of genetic alterations and mutations, with over 50% of leukemias having major chromosomal disruptions such as gene fusions and translocations, making leukemia a very heterogeneous disease [14]. Conventional leukemia treatments are broad and include chemotherapy, stem cell and bone marrow transplants, immunotherapy, and radiation treatments. Treatment success can vary significantly from one patient to another, and targeted therapies are limited, and more are needed to improve remission and survival rates [15].

Acute Myeloid Leukemia (AML)

Leukemia is the most common form of pediatric cancer (ages 0-15) making up around 30% of diagnoses, with the two main subtypes being acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML). Acute myeloid leukemia makes up around 18% of pediatric leukemia cases. Acute myeloid leukemia is defined as a fast-growing leukemia of progenitor/precursor cells belonging to the myeloid cell lineage. It is a very heterogeneous disease with many genetic alterations, with around 90% of cases presenting with at least one known genetic alteration that impacts normal hematopoiesis and cell function leading to bone marrow failure [12]. AMLs generally do

not respond well to conventional leukemia therapies and require more targeted therapies due to the higher rates of genetic mutations and aberrantly expressed markers [15]. Symptoms of hematopoietic disruptions and bone marrow failure include anemia, thrombocytopenia which can cause bleeding and bruising easily, bone pain, fatigue, organomegaly or enlargement of organs, and extra medullary symptoms such as skin rash and lesions (petechia/purpura) and nervous system (NS) involvement. Acute myeloid leukemias further present with high white blood cell counts or leukocytosis and myelofibrosis of the bone marrow [16].

Based on the French American British (FAB) classification of AML subtypes, which looks at morphological and cytochemical characteristics and maturity of leukemic cells, there are several different subtypes of AML (M0 through M7) (Table 1). The M0 to M5 subtypes originate from immature white blood cells, the M6 subtype originates from immature red blood cells and the M7 subtype originates from immature platelet producing cells called megakaryoblasts [17].

FAB subtype	Name
M0	Undifferentiated acute myeloblastic leukemia
M1	Acute myeloblastic leukemia with minimal maturation
M2	Acute myeloblastic leukemia with maturation
M3	Acute promyelocytic leukemia (APL)
M4	Acute myelomonocytic leukemia
M4 eos	Acute myelomonocytic leukemia with eosinophilia
M5	Acute monocytic leukemia
M6	Acute erythroid leukemia
M7	Acute megakaryoblastic leukemia

Table 1: French American British classification of AML subtypes.

Depending on the subtype of AML there are varying survival rates ranging from 22-90% [12, 13]. The M7 subtype or Acute megakaryoblastic leukemia (AMKL) is a very rare subtype of acute myeloid leukemia (AML) predominantly occurring in infants (less than 3 years of age) and makes up around 4-15% of pediatric AML cases and 1% of adult AML cases [18-20].

Acute Megakaryoblastic Leukemia (AMKL)

AMKL is characterized by abnormal megakaryoblasts and extensive myelofibrosis making bone marrow sampling difficult for diagnostics and research [18, 21]. Immunophenotyping has found that AMKL blasts express CD33+/-, CD36+/-, CD41/*ITGA2B*, CD42b/*GPIBA* and CD61/*ITGB3* markers which are commonly used for diagnosis [20, 22]. AMKL is hard to treat due to the early relapse rates and resistance to chemotherapy [18, 21, 23]. AMKL occurs almost 500-times more frequently in children with trisomy 21 or Down syndrome (DS) [13]. Compared to DS-AMKL, Non-DS AMKL is characterized by significantly poorer outcomes, lower event free survival and more frequent chimeric oncogenes impacting normal hematopoiesis [18]. Pediatric and adult

AMKL are very different diseases with distinct genetic landscapes [21]. Compared to pediatric AMKL, adult AMKL usually arises as a result of pre-existing myeloproliferative disorders. Whereas pediatric AMKL usually arises from genetic mutations that drive the leukemia development [24]. Recent large scale genomic sequencing approaches have uncovered the most frequent chimeric oncogene fusions that occur in pediatric patients such as; *CBFA2T3::GLIS2, NUP98::KDM5A, RBM15::MKL1, MLLr* and others [25] (Figure 2A). These chimeric gene fusions and rearrangements occur in around 60% of AMKL patients and are mutually exclusive to each other [23, 26].

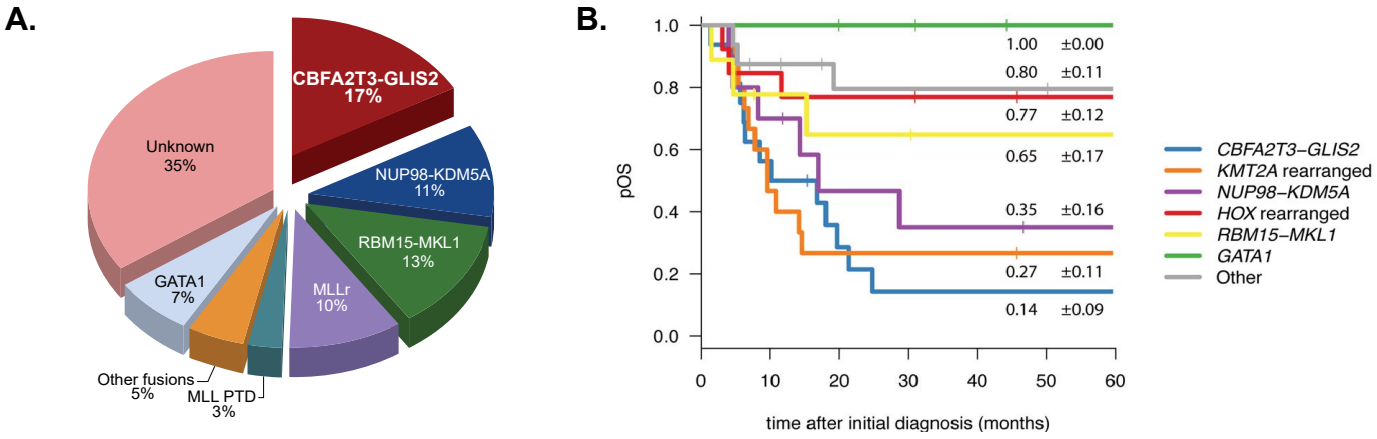


Figure 2: Frequently occurring chimeric gene fusions and probability of overall survival in pediatric non-Down syndrome AMKL. A). The most frequently occurring gene fusions in pediatric non-Down syndrome (non-DS) AMKL was assessed in 142 pediatric patients with transcriptome sequencing, reverse-transcription chain reaction (RT-PCR) and split signal fluorescence *in situ* hybridization. The results demonstrate that 17% of AMKL carried *CBFA2T3::GLIS2*, 11% had *NUP98::KDM5A*, 13% had *RBM15::MKL1*, 10% had *MLLr (KMT2Ar*), 3% had *MLL* PTD, 5% had other fusions, 7% had *GATA1* and around 35% had unknown fusions. Other fusions include single cases of each of the following: *GATA2::HOXA9, NIPBL::HOXB9, MN1::FLI1, HLX-B9::ETV6, FUS::ERG,* and *RUNX1::CBFA2T3*. Adapted from copyright © 2015, Gruber and Downing, License number: 5438381412709 [18]. **B).** Probability of overall survival (pOS) measured as time after initial diagnosis (months) in non-DS AMKL patients harboring different gene fusions; *CBFA2T3::GLIS2* (N=16); *KMT2A* rearranged (N=15); *NUP98::KDM5A* (N=10); *HOX* rearranged (N=13); *RBM15::MKL1* (N=9); *GATA1* (N=8); Other (N=16). Reprinted from copyright © 2017, de Rooj et al., License number: 5438381106045 [21].

CBFA2T3::GLIS2 fusion

The *CBFA2T3::GLIS2* fusion is the most common fusion identified in non-DS AMKL and has not been identified in adults yet [25]. *CBFA2T3::GLIS2* is a cryptic fusion, not nor-

mally identified by standard cytogenetic techniques and it is associated as an aggressive form of leukemia with poor 5-year survival ranging between 15-30% and reduced overall survival (Figure 2B) [21, 23, 26]. *CBFA2T3::GLIS2* gene fusion is caused by an inversion at chromosome 16, which leads to the fusion of the five prime nervy homology region of *CBFA2T3* (exon 11) to the three prime zinc finger domain of *GLIS2* (exon 3) (Figure 3) [25].

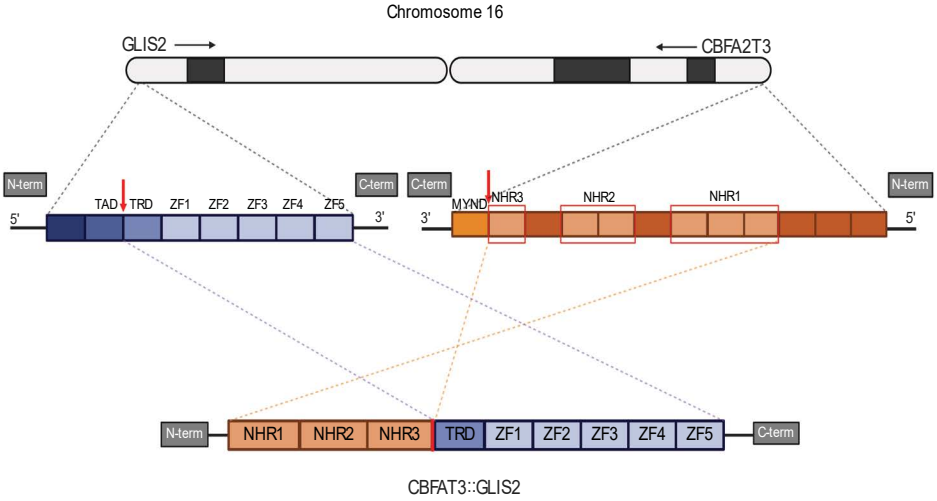


Figure 3: Structure of the *CBFA2T3::GLIS2* gene fusion. The *CBFA2T3::GLIS2* gene fusion is caused by an inversion at chromosome 16, which leads to the fusion of the five prime nervy homology region (NHR) of *CBFA2T3* to the three prime zinc finger domains (ZF1) of *GLIS2*. Red arrows depict common breakpoints at exon 11 in *CBFA2T3* and exon 3 in *GLIS2*. MYND, myeloid, nervy, and DEAF-1; NHR, nervy homology regions; TAD, topologically associating domain; TRD, trans-Repression Domain; ZF, zinc finger; N-Term, N-terminus; C-Term, C-Terminus. Figure Created with BioRender.com and adapted from 2019, Masetti et al. under the license CC BY-NC 4.0 [25].

CBFA2T3 is a member of the RUNX1T1 or previously called ETO/CBFA2T1/MTG8 family of nuclear transcriptional corepressors. It plays a role in hematopoietic stem cell quiescence, self renewal and differentiation and the development of megakaryocyte and erythrocyte progenitor cells [27-30]. *GLIS2* is a member of the GLI-similar subfamily of Krüppel-like zinc finger transcription factors and functions in kidney development and neurogenesis, where its role in leukemogenesis is not very well understood [25, 28]. *GLIS2* is not normally found in immature/differentiating progenitor cells, meaning that *CBFA2T3* may lead to its ectopic expression and activity [25, 31]. When the fusion between *CBFA2T3* and *GLIS2* occurs, the DNA binding domain of *GLIS2* is increased [28], a disruption in an allele within the genes occurs leading to a disruption in stem cell quiescence, a block in differentiation and increased self-renewal of hematopoietic cells.

Altogether, these can drive the development and the maintenance of leukemia. The C*BFA2T3::GLIS2* gene fusion is also associated with the overexpression of the cell surface marker CD56 also known as the neural cell adhesion molecule 1 (*NCAM1*) [25, 32].

Neural cell adhesion molecule 1 (*NCAM1*) or CD56

NCAM1 is a cell surface glycoprotein belonging to the immunoglobulin superfamily (IgSF) of cellular adhesion molecules (CAMs) [33]. *NCAM1* was first identified and is predominantly expressed in the nervous system on neurons and glial cells but can also be found in most tissues of the body such as the lung, kidneys [34], skeletal muscles [35], and the heart [36]. More recently, *NCAM1* was found within the hematopoietic system predominantly on natural killer cells and also on certain T- and B- cells and dendritic cells [33, 37].

***NCAM1* structure, isoforms, and function**

NCAM1 is a single gene encoded by chromosome 11 (band q23) in humans and chromosome 9 in mice and it is made up of at least 25 exons [38, 39]. The *NCAM1* gene is made up of five immunoglobulin (Ig) domains (exons 1-10), two fibronectin type III (F3I) domains (exons 11-14), a glycosylphosphatidylinositol (GPI) anchor (exon 15), a transmembrane (TM) domain (exon 16) and the cytoplasmic tail (IC) domain (exons 17-19) (Figure 4). *NCAM1* can undergo alternative RNA splicing to give rise to at least 27 distinct isoforms [40, 41]. The three main groups of isoforms are the *NCAM1*-120, *NCAM1*-140 and *NCAM1*-180 which are separated based on their molecular weights (120kDa, 140kDa and 180kDa, respectively). The Ig and F3I domains or exons 1-14 remain untouched by alternative splicing and are common to all isoforms. The *NCAM1*-120 isoforms are anchored to the plasma membrane by a GPI anchor and are missing exons 16-19 corresponding to the TM and IC domains. The *NCAM1*-140 and *NCAM1*-180 isoforms are transmembrane proteins containing intracellular regions [42-44]. *NCAM1*-140 isoforms are missing exon 15 and 18, whereas *NCAM1*-180 isoforms are only missing exon 15 ([42, 43] (Figure 4). There exists a close relative to *NCAM1*, called *NCAM2* which is encoded by a separate paralog gene. *NCAM2* shares a similar structure to *NCAM1* and around 45% overall similarity. *NCAM2* has not been well studied and not much is known about its expression and function. Like *NCAM1*, *NCAM2* is found within the nervous system primarily in the olfactory bulb and it is not expressed in AML subtypes [44-46] (our unpublished data).

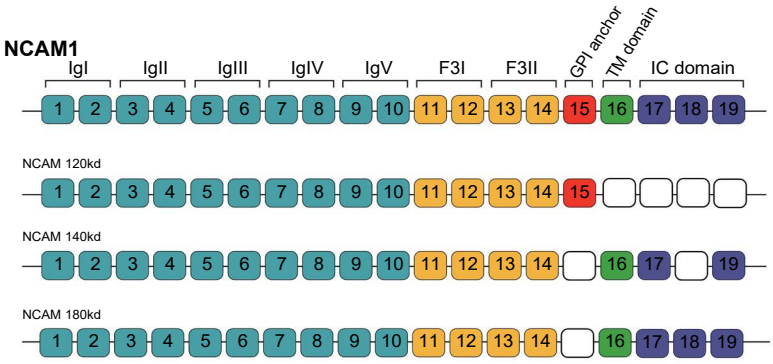


Figure 4: *NCAM1* alternative splicing isoforms. The standard *NCAM1* structure (exons 1-19) includes the immunoglobulin homology domains (Ig; IgI - IgV) (exons 1-10), fibronectin type-III domains (F3I & F3II)) (exons 11-14), glycosylphosphatidylinositol (GPI) anchor (exon 15), transmembrane domain (TM) (exon 16) and intracellular (IC) domains (exons 17-19) (Top). The exon structures for the three main isoforms of *NCAM1*; *NCAM1*-120kd, *NCAM1*-140kd and *NCAM1*-180kd (Bottom). Figure created with BioRender.com and adapted from Copyright © 2009, Gattenlöhner et al., License number: 5438390025641 [43].

NCAM1 Function

The functional role of *NCAM1* has been well characterized and studied within the nervous system where it functions in neuron differentiation and neurogenesis, synaptic plasticity, learning and memory consolidation. *NCAM1* also functions in cell-cell and cell-matrix interactions, cell-cell and cell-membrane adhesion and as a signaling receptor, transmitting signals across a cell membrane [43, 47, 48]. Its function is not well understood or studied within the hematopoietic system.

NCAM1 and signaling pathways

NCAM1 has been reported as a signaling receptor in various cell types and through different binding interactions is involved in activating intracellular signaling cascades and downstream signaling pathways that control apoptosis, cell growth and proliferation, differentiation, and survival [43]. *NCAM1* can participate in homophilic binding with other *NCAM1*s or in heterophilic dimerization with other molecules such as other adhesion molecules (e.g. L1), extracellular matrix proteins (e.g. HSPGs), growth factor receptors (e.g. FGFR), and the GDNF family receptor [43]. These interactions can occur on the same surface of the cell (cis-interactions) or between two different cell surfaces (trans-interactions) [44]. Downstream signaling cascades such as the NF-KB, PI3K, PLCγ and MAPK pathways can be activated by either homophilic binding or heterophilic dimerization [44]. The MAPK and PI3K pathways can be activated by either homophilic *NCAM1* interactions by the recruitment of non-receptor tyrosine kinases

FAK and FYN [45, 49] or by heterophilic dimerization with the fibroblast growth factor receptor (FGFR) [43, 50, 51] (Figure 5). The NF-KB pathway can be activated by *NCAM1* homophilic binding and prevent apoptosis [52, 53]. The constitutive activation of the MAPK and PI3K pathways is associated with poor survival and has been found in several malignancies and in 50-80% of AML patients [54, 55]. Protein kinase inhibitors (PKIs) are novel forms of targeted therapies under development and on clinical trials for the treatment of AML subtypes [56].

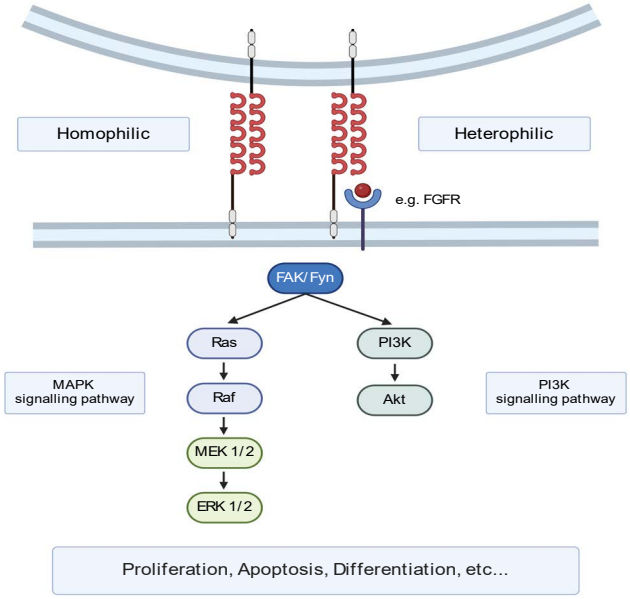


Figure 5: Signalling pathways downstream of *NCAM1*. *NCAM1* can engage signalling pathways through homophilic or heterophilic interactions. These interactions can lead to the recruitment of FYN/FAK and trigger a variety of signalling cascades such as MAPK and PI3K. These signalling cascades can lead to effects on cell proliferation, apoptosis, and differentiation. Figure created with BioRender.com. FGFR, fibroblast growth factor receptor**.**

NCAM1 expression and function regulation

Based on studies in other malignancies or the NS, and limited studies in AML, *NCAM1* seems to be regulated in various forms. *NCAM1* expression can be regulated by certain gene fusions, post translational modifications and certain transcription factors, but it remains unclear how *NCAM1* is regulated in acute myeloid leukemias (AMLs) leading to its progression and development. Certain gene fusions such as *CBFA2T3::GLIS2* can have a direct impact on *NCAM1* expression. A study has demonstrated through chromatin immunoprecipitation that *CBFA2T3* binds directly to the *NCAM1* promoter, making it a direct transcriptional target [32] (Figure 6).

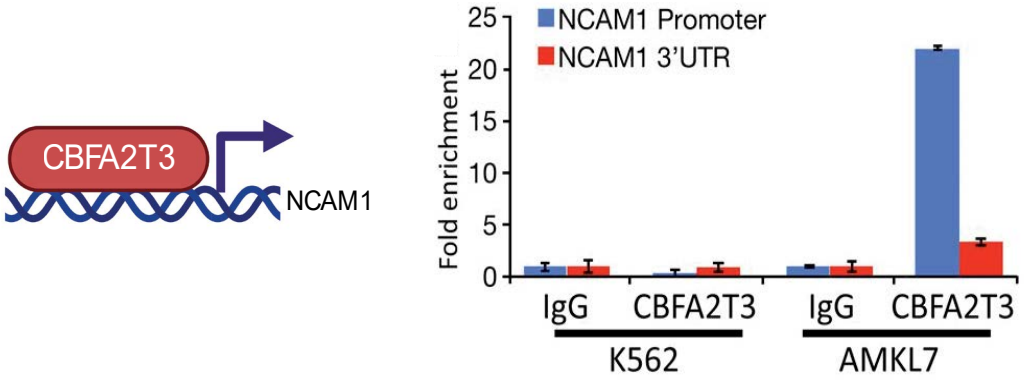


Figure 6: *NCAM1* is a direct transcriptional target of *CBFA2T3*. Chromatin immunoprecipitation (ChIP) studies in AMKL7 cells or the K562 cell line (*NCAM1* negative) demonstrated *NCAM1* is a direct transcriptional target of *CBFA2T3* (Left). ChIP was performed with a *CBFA2T3*-specific antibody and a nonspecific antibody (IgG) followed by qPCR analysis of immunoprecipitated DNA using primer pairs located in the *NCAM1* promoter or in the 3'UTR region. Error bars indicated Mean +/- SEM. Figure adapted and reprinted from 2012, Thiollier et al. under the license CC BY-NC-SA 3.0 [32].

Post translational modifications such as phosphorylation, glycosylation and polysialylation can occur and dramatically alter the function of *NCAM1* in cancer cells and tumorigenesis [45]. Polysialylation of *NCAM1* introduces a negatively charged polysialic acid and polsialylated *NCAM1* (PSA-*NCAM1*) is significantly reduced in adults and occurs more frequently in children, reducing the amount of homophilic interactions that occur [45, 57, 58]. Polysialylation of *NCAM1* is becoming an important prognostic factor in malignancies such as neuroblastoma and the presence of PSA-*NCAM1* confers a poor prognosis and overall survival [59].

Several transcription factors can modulate *NCAM1* expression and activity. For example, in multiple myeloma, SOX4 regulates *NCAM1* expression and when inhibited can lead to downregulation of *NCAM1* mRNA [60]. Certain isoforms of the runt-related transcription factor 1 (RUNX1) transcription factor are overexpressed in CD56 positive AML. The forced expression of RUNX1 induced *NCAM1* expression in CD56 negative cells [61]. Finally, DNAse hypersensitivity assays have uncovered other potential upstream regulators of *NCAM1* such as of MEIS1, MEF2C and STAT1 finding that these bind to open chromatin binding sites on *NCAM1* promoter [62]. All these transcription factors, when inhibited lead to the downregulation of *NCAM1* mRNA expression [60-62].

NCAM1 and leukemogenesis

NCAM1 expression is deregulated not only in AML subtypes but also in other malignancies including neuroblastoma, multiple myeloma, small cell lung cancer, ovarian cancer,

and pancreatic cancer [63] and may play a role in disease development and progression [64]. There is limited research on the functional role of *NCAM1* in AML subtypes. The few studies that have looked into this found that the overexpression of *NCAM1* is associated with promotion of leukemogenesis and disease progression. In certain leukemias that carry a chromosomal translocation (t(8;21) (q22;q22)) considered lower risk of relapse have significantly reduced remission and response to therapies when *NCAM1* is highly expressed [64, 65]. In AML, *NCAM1* is found to contribute to resistance to chemotherapy, reduced complete remission, higher relapse rates and poor overall survival. Reduced survival curves and complete remission rates occurred in patients with high *NCAM1* expression (6 months) compared to those without *NCAM1* (12 months), demonstrating a more aggressive disease in *NCAM1*+ AMKL [66] (Figure 7).

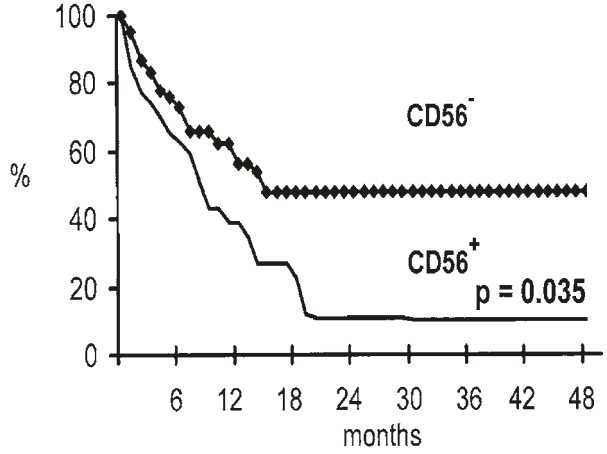


Figure 7: Percent overall survival is significantly reduced in CD56 positive AML patients. Survival analysis in CD56+ and CD56- AML patients demonstrated reduced mean overall survival (%) (6 months) in CD56+ patients compared to CD56- patients (12 months). Reprinted from Copyright © 2001, Raspadori et al. License number: 5438380536100 [66].

Studies found that *NCAM1* knockdown led to reduced cell growth and proliferation and increased cell death demonstrating that *NCAM1* functions in sustaining leukemic cells and promoting their survival [62] (Figure 8A). Another *in vivo* study found that *NCAM1* KD in mice treated with a chemotherapy drug lived significantly longer compared to *NCAM1* positive mice, demonstrating that *NCAM1* functions in chemotherapy resistance in AML [62] (Figure 8B).

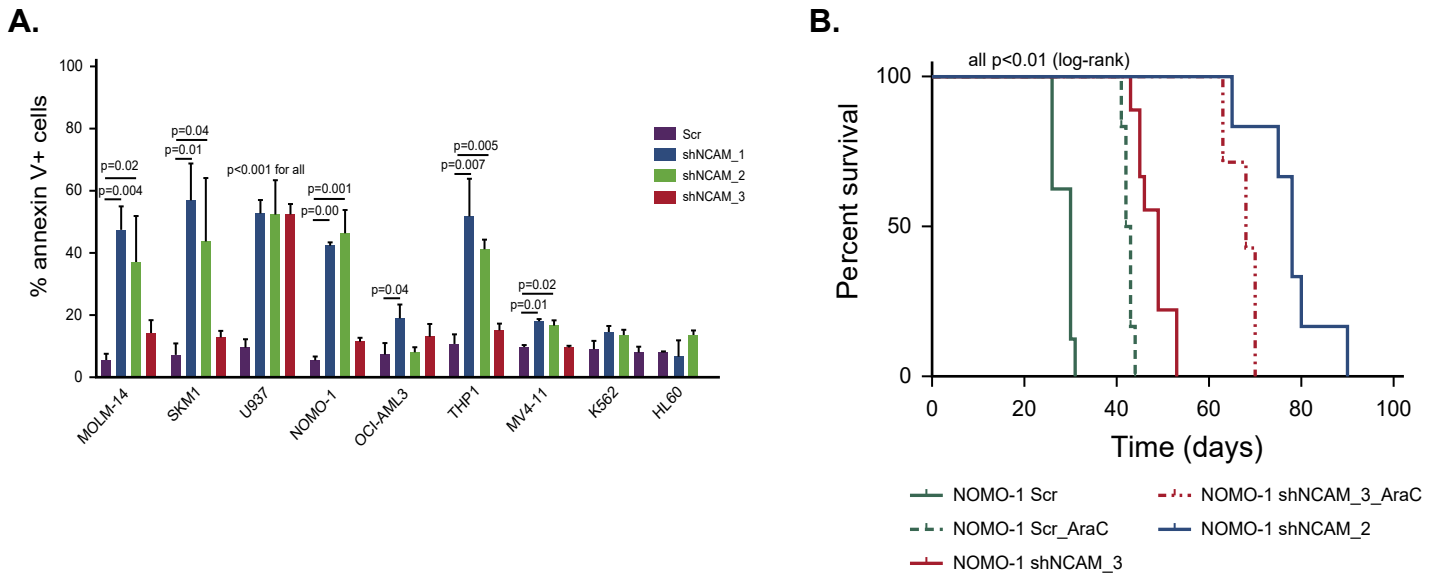


Figure 8: *NCAM1* KD leads to an increase in cell death and when combined with chemotherapy prolongs survival in mice. A). Cell death (% annexinV+ cells) was assessed 5 days after the addition of doxycycline induction of different shRNA targeting *NCAM1* (sh*NCAM1*_1, _2, _3) compared to a scrambled (Scr) control in different cell lines (e.g. U937). Statistical analyses were performed by the 2-tailed Student t-test. Error bars represent mean of 3 experiments for each condition +/- SD of mean. Reprinted from Copyright © 2019, Sasca et al., License number: 5438381282800 [62]. **B).** Percent survival of NSG mice (N=5 per group) transplanted with NOMO-1 cells carrying a scrambled control (Scr) or shRNA targeting *NCAM1* (sh*NCAM1*_2 and sh*NCAM1*_3). All mice received doxycycline starting on day 15 post transplant. Between day 20-30 post transplant, NOMO-1_scr and sh*NCAM1*_3 mice were treated with either a vehicle or Ara-C (10mg/kg) once daily and survival of mice are shown as Kaplan Meier curves. Reprinted from Copyright © 2019, Sasca et al., License number: 5438381282800 [62].

Synthetic AML models

As mentioned earlier acquiring samples of AMKL can be limited and difficult due to myelofibrosis which makes bone marrow aspirations difficult and low blast cell counts [67]. AML cell lines are available but are not always fully representative of the disease. Having *in vivo* models is essential to be able to study the disease and to have the essential niche interactions that contribute to leukemic cell maintenance and disease progression and development. Our laboratory has developed several synthetic models of AML including *NUP98::KDM5A* AMKL and *CBFA2T3::GLIS2* AMKL. These models were generated by overexpression of the fusion (e.g. *CBFA2T3::GLIS2*) in cord blood-derived human hematopoietic stem and progenitor cells (hHSPCs) and transplantation into immunodeficient mice [68] (unpublished) (Figure 9A). Patient derived xenograft models have also been generated in our lab through transplantation of patient leukemic cells carrying gene fusions

(e.g. *NUP98::BPTF*) into NSG mice (Figure 9B). Our CG2 PDX and synthetic models recapitulate AMKL disease both molecularly and phenotypically, with the expression of AMKL cell surface markers (CD34+CD41+CD61+) and high *NCAM1* expression and can be used in functional studies.

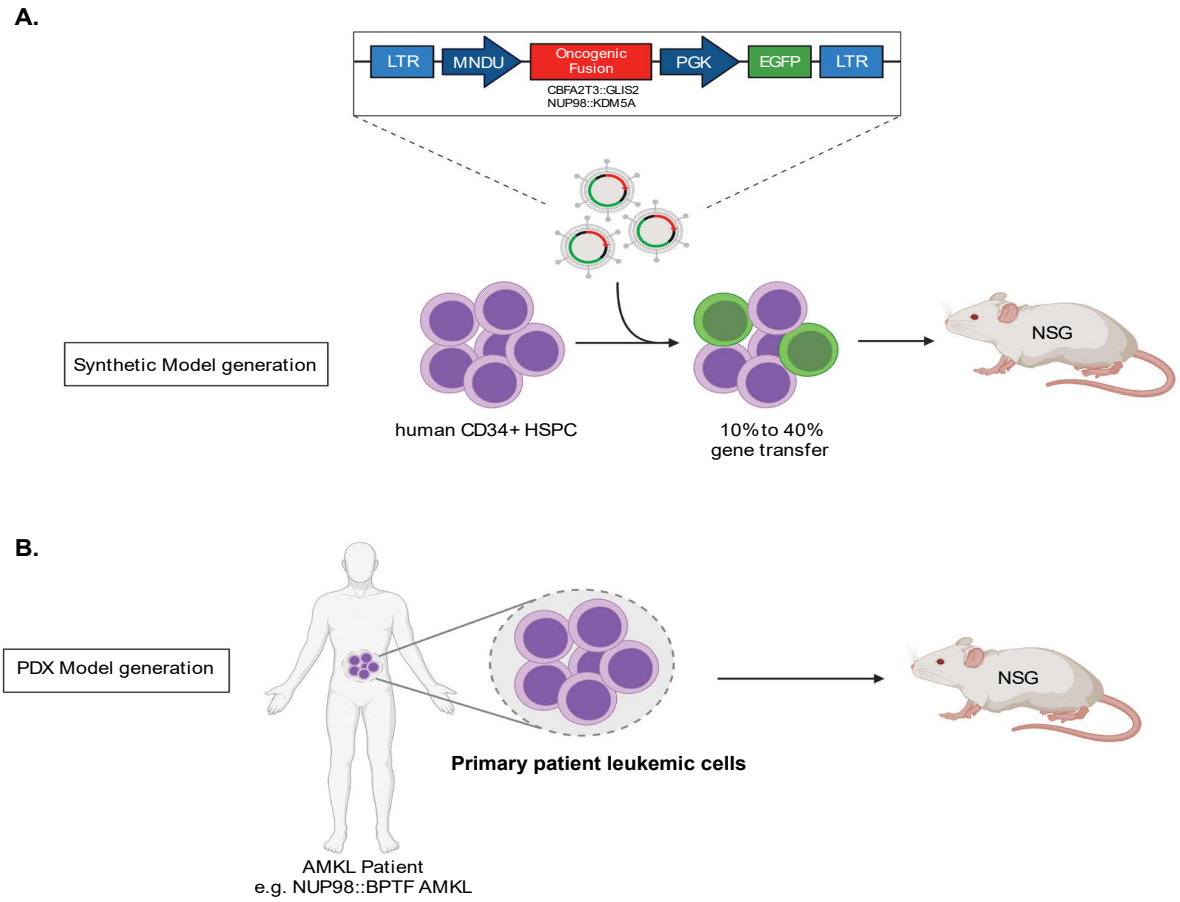


Figure 9: Generation of synthetic and patient derived xenograft models. Synthetic and patient derived xenograft (PDX) murine models were previously generated in our lab. **A).** Synthetic models were generated using a lentiviral-based approach of overexpression of the fusion (e.g. *CBFA2T3::GLIS2* or *NUP98::KDM5A* fusion) in human CD34+ cord blood hematopoietic stem and progenitor cells (CD34+ HSPC) and transplantation into NOD-scid IL2Rgnull (NSG) mice the next day. Around 30-40% gene transfer (%GFP) was achieved. **B).** PDX models were generated by first acquiring primary patient leukemic cells and transplanting them in NSG mice. Figure created with BioRender.com.

Hypothesis and Objectives

Since AML is a very heterogeneous disease with varying genetic and immunophenotypic landscapes, it can be difficult determining the type of therapeutic approach to use and determining prognosis. Treatments for AMKL have not evolved for years and there are currently not many widely available targeted therapies. Being able to identify unique and specific biomarkers, such as

NCAM1, is important for creating targeted and patient specific therapies that could help to improve the dismal AMKL survival rates. My hypothesis is that *NCAM1* is a functional dependency in CG2 AMKL. My objectives were to get a global profiling of *NCAM1* expression in leukemia and normal hHSPCs and using lentiviral shRNA-mediated knockdown of *NCAM1* to functionally assess the role of *NCAM1* in sustaining leukemic cells *in vitro* and leukemia engraftment and progression *in vivo*.

Materials and Methods

All experimental procedures have undergone optimization for specific use in the following experiments.

Cell lines and primary cells

M07e (ACC 104) with the *CBFA2T3::GLIS2* fusion, as described in the literature [31], ML-2 (ACC 15), U937 (ACC 5), CMK (ACC 392) and HEL (ACC 11) cell lines were purchased from DSMZ (Germany). HEK293T, K562, and HL-60 cell lines were purchased from American Type Culture Collection (ATCC). CHRF-288-11 was a gift from Micheal A. Lieberman (University of Cincinnati, Cincinnati, OH). Primary xenograft cells were isolated from bone marrow or spleen from our synthetic models (described below) and express the fluorescent reporter gene GFP. The cell lines were cultured as recommended by the manufacturer. M07e was grown in RPMI 1640 medium supplemented with 10% fetal bovine serum (FBS, Wisent) and 10ng/ml recombination human Interleukin-3 (IL-3, Gibco). U937, CHRF-288-11, HEL and ML-2 were maintained in RPMI 1640 medium supplemented with 10% FBS and CMK was maintained in RPMI supplemented with 20% FBS. K562 and HL-60 were maintained in IMDM supplemented with 10% FBS and 20% FBS, respectively. HEK293T was grown in DMEM supplemented with 10% FBS. Cells were expanded in culture and used at a low passage (≤ 20) for each experiment and were routinely checked for mycoplasma contamination. Primary xenograft cells were grown and cultured in optimized serum free media supplemented with cytokines, 500 nM SR1 [69] (Cedarlane) and 750 nM UM729 [70] (StemCell Technologies) at 37 °C and 5% CO_2.

Cord blood isolation

The isolation of human CD34+Lin- cord blood hematopoietic stem and progenitor cells (HSPCs) was performed on human umbilical cord blood (CB) units (≤ 24 hours old) provided by Héma-Quebec with the mother's consent through its public CB bank (Montréal, Canada). The protocol and handling of CB units was approved by the Research Ethics Board of CHU Sainte-Jus-

tine (approval number 3453). RosetteSep™ (StemCell Technologies) combined with Ficoll-Paque (Fisher Scientific) layering was used to pellet differentiated cells such as T cells, B cells, myeloid cells, platelets, and red blood cells during density gradient centrifugation, leaving a pre-enriched layer of CD34+ cells. This was followed by magnet positive selection of CD34+ HSPCS using the EasySep™ Human cord blood CD34 positive selection kit (StemCell Technologies) in order to get a higher purity of CD34+ HSPCs. The purity of the sample was verified with CD34 antibody surface staining and flow cytometry and only HSPC populations with a purity above 70% were used in experiments. In culture, CD34+ cord blood cells were maintained in optimized culture conditions [69, 71] at 37 °C and 5% CO_2.

Lentivirus production and transduction

shRNA for knockdown vectors

For short hairpin (sh) RNA-mediated knockdown (KD) of *NCAM1*, a lentiviral approach was used. The goal of gene knockdown is to reduce the expression level of a target gene, in this case *NCAM1*. The lentiviral vector for performing gene knockdown contained an mCherry fluorescent reporter gene used for assessing transduction and gene transfer efficiencies, a miR-E backbone that functions to improve mature shRNA levels and knockdown efficiencies [72] and a short hairpin (shRNA) against the genes of interest for knocking down the gene (Figure 13A). The shRNA molecules are processed into the cell where, through RNA interference via small interfering RNA (siRNA), knockdown of the gene and reduction in expression occurs [73]. The mir-E backbone for shRNAs has been shown to improve knockdown efficiency, increase accuracy and precision of dicer cleavage and reduces off target effects [72]. This backbone was implemented into the design of the shRNA structures targeting *NCAM1* to have higher levels of knockdown and to be able to more accurately study the functional effects of knocking down *NCAM1*. The following vectors were used as controls for shRNA mediated knockdown: MNDU-mCherry-miRE-shRenilla (583) and MNDU-mCherry-miRE-shNon-targeting (774). ShRenilla is from Renilla reniformis luciferases and the MNDU plasmid was a gift from Dr. Guy Sauvageau's laboratory with permission from Dr. R. Keith Humpries (Terry Fox Laboratory). The shNon-targeting control was purchased from Sigma and modified to fit the miR-E MNDU design. Both of these control target sequences do not target mammalian genes and are thus non-silencing and non-targeting. The following vectors were used for targeting *NCAM1*: MNDU-mCherry-miRE-sh*NCAM1*(599), MNDU-mCherry-miRE-sh*NCAM1*(601), and MNDU-mCherry-miRE-sh*NCAM1*(603). The shRNA sequences used to target *NCAM1* were taken from Fellman et al. [72] and subcloned into an MNDU3-miR-E

lentiviral expression vectors (Supplemental Table S1).

Transfection

Prior to lentiviral particle production, midipreps and maxipreps were performed to isolate and purify plasmid DNA for use in transfections using The PureLink™ HiPure Plasmid Filter Midiprep and Maxiprep Kits (Invitrogen). The lentiviral particles were produced in HEK293T cells by transfection with a lentiviral vector carrying the shRNAs targeting the genes of interest, as described above and with the packaging vectors containing *gag*, *pol*, and *rev* genes (pLp1 and pLp2) and VSV-G envelope plasmid using Lipofectamine 2000 transfection reagent (Thermo Fisher Scientific) and suspended in OptiMem media (Gibco) following standard procedures. The viral supernatant was collected 48 hours post transduction, concentrated using 5x PEG-It virus precipitation solution (Systems Bioscience) following manufacturer's instructions and tittered by flow cytometry.

Lentiviral transduction for knockdown

Lentivirus transduction was performed using 5µg/ml polybrene transduction compound (MilliporeSigma) for *in vivo* studies and 8µg/ml protamine sulfate transduction compounds for *in vitro* studies, following standard procedures. The toxicity of these compounds and transduction efficiencies on both cell lines and primary xenograft cells was tested prior to performing experiments to determine optimal concentrations. Cell lines or primary xenograft cells were transduced at a multiplicity of infection (MOI) of 5 and 15 respectively, with lentiviral particles carrying shRNA targeting *NCAM1*, a Renilla control and a Non-targeting control for 24-, 48- or 72-hours. Post-transduction, gene transfer efficiencies or % mCherry was assessed by flow cytometry.

Synthetic mice model generation and husbandry

All mouse experimental procedures were performed in accordance with the Canadian Council of Animal Care and with approval from the CHU Sainte-Justine Research ethic boards and Animal Institutional ethics committee (approval number: MP-21-2017-1566). Previously in our lab, synthetic xenograft murine models were generated using a lentiviral-based approach of overexpression of the fusion (e.g. *CBFA2T3::GLIS2* or *NUP98::KDM5A* fusion) in human cord blood hematopoietic stem and progenitor cells and transplantation the next day into NOD-scid IL2Rg-null (NSG) mice from Jackson Laboratories (Bar Harbor, ME, USA) [68] (Figure 9A). The following vector was used for this approach: MNDU-*CBFA2T3::GLIS2*-GFP and MNDU-*NUP98::KDM5A*-GFP. Patient derived xenograft (PDX) models were also generated in our lab. AMKL patient

leukemic cells carrying fusions such as *NUP98::BPTF* (NTF) were harvested, underwent ficoll layering and were then transplanted into NSG mice (Figure 9B). These approaches created models that phenocopy human disease both molecularly and phenotypically. The mice were bred in pathogen-free conditions at the animal care facility of CHU Sainte-Justine. Mice were kept in ventilated microinsulator cages and provided with sterilized food and acidified water. Several generations of models were generated and used in the lab. The models used throughout these experiments are *CBFA2T3::GLIS2* AMKL models, referred to as mCG2-1 (I749 branch) and mCG2-2 (I732 branch), as well as *NUP98::KDM5A* AMKL models referred to as N5A and *NUP98:BPTF* PDXs referred to as NTF.

Transplantation and *in vivo* studies

Immunodeficient NSG mice (N=6 per shRNA condition, N=3 for Non-transduced conditions) were obtained from Jackson laboratories (Bar harbor, ME, USA). Cells were transplanted via intravenous injection via tail-vein in mice recipients that underwent sub-lethal irradiation (whole-body irradiation with 2Gy X-rays, CP160 irradiator, Faxitron X-Ray Corporation, USA) 24 hours prior to xenotransplantation. Primary CG2 xenograft spleen cells were either Non-transduced or transduced with shRenilla (sh583) and shNon-targeting (sh774) controls and three sh*NCAM1*s (sh599, sh601, sh603) and a fixed volume of cells, 80% of the well, were transplanted 24 hours post transduction into each recipient mouse (around 400,000 cells, 1 well/1 mouse).

After xenotransplantation, the mice were monitored through serial blood sampling for human CD45+ and GFP+ cells via flow cytometry analysis to assess leukemic cell infiltration and to select an endpoint for the experiment. hCD45 staining is used to differentiate mice cells from human cells. Standard blood lysis procedures were followed. At 9 weeks post-transplantation or earlier if mice were showing advanced signs of leukemia (reduced mobility, paleness, hunchback and/or dyspnea), the mice were sacrificed following standard necropsy procedures. Following sacrifice, spleens were weighed, and bone marrow (femur) and spleen cells were isolated and harvested in RPMI 1640 media supplemented with 1% FBS for flow cytometry analysis. Harvested cells underwent protein surface staining and flow cytometry, clonogenic progenitor assays, cell death (AnnexinV/DAPI) assays, and were directly lysed in Trizol reagent for RNA extraction and qPCR analysis. One tibia and portion of spleen per mouse was fixed in 10% neutral buffered formalin for histopathology.

Immunohistochemistry

Femurs and spleen tissues were fixed in 10% neutral buffered formalin and were sent to

the OPTILAB pathology laboratory at CHU Sainte-Justine (Montréal, Quebec) for immunohistochemical staining. Formalin-fixed femurs were decalcified for 48 hours prior to embedding in paraffin. Paraffin embedded samples were cut longitudinally and transversally into 4μm sections and stained with an antibody against human *NCAM1*/CD56 (Supplemental Table S2) and mounted onto slides. Images were acquired with a Zeiss Axio.Scope.A1 microscope (Zeiss) equipped with a color camera (Axiocam 105 color, Zeiss) and operated with the Zen 2.6 blue edition software (Zeiss).

Flow cytometry

Flow cytometry was used to categorize cells of heterogeneous samples based on the presence or absence of specific cell surface markers, to measure protein expression patterns and to determine subpopulations expressing *NCAM1*. Flow cytometry was used for cell sorting, automated cell counting, titrating antibodies, titrating lentivirus, assessing cord blood purity, measuring gene transfer (%mCherry), measuring surface protein expression (MFI), and assessing cell death and viability with the AnnexinV/DAPI assay. The LSR Fortessa cytometer (BD Bioscience) was used and passed single cells within the sample by a laser within the flow cytometer to detect the fluorophore, count and/or sort the cells. Excitation spectrum from the flow cytometer are ranges of wavelengths that add energy to the fluorochromes present on antibody marked cells. This energy excites the fluorochromes and causes them to emit light at a different wavelength, termed the emission spectrum. The emitted spectrum fluorescence measured directly reflects the amount and type of cell present within the sample and the lasers also measure forward and side scatter of the laser beam. Forward scatter indicates the size and volume of the cells and side scatter indicates the complexity and granularity (cytoplasmic granule content /nuclear components) [74]. The staining buffer used was composed of sterile PBS, 2%FBS and 1mM EDTA (Thermo Fisher Scientific). Prior to staining, primary xenograft cells were blocked for non-specific binding with a mouse gamma globulin (dilution 1/1000, Jackson ImmunoResearch Laboratories). Compensation beads (BD Bioscience) were used to set up voltages on the flow cytometer and to be able to differentiate between antibodies/markers. Compensation is performed because there can be emission spectrum overlaps, termed spectral overlap, when target cells are stained with multiple fluorochromes at once. Spectral overlap between different fluorochromes can make it difficult to differentiate between markers. Compensation is used to correct this by subtracting partial fractions of the signals emitted from each fluorochrome [74]. Data were first analyzed with FACSDiva (BD Biosciences) and a more detailed analysis was performed with FlowJo software (BD Biosciences).

Cell counting and sorting

Automated cell counting was performed with the LSRFortessa cytometer equipped with a high throughput sampler (HTS) (BD Biosciences). Cell counts were measured using a fixed and known sample volume and sample flow rate. Cells were sorted based on the presence of fluorescent reporter genes (mCherry and/or GFP) with a FACSAria II flow cytometer (BD Biosciences) at the flow cytometry platform of CHU Sainte-Justine (Montréal, Canada) or at the Flow cytometry core facility of the Institute for Research in Immunology and Cancer (IRIC, Montréal, Canada).

Immunophenotyping, surface staining and gene transfer

Cells either expressed fluorescent reporter genes (mCherry and/or GFP) or were stained with fluorophore conjugated antibodies against surface proteins of interest which allowed for the expression and mean fluorescence intensity (MFI) of the marker to be quantified using fluorescence. All antibodies were titrated prior to use by performing serial dilutions and assessment by flow cytometry to determine optimal dilution to use. Cells were stained with a variety of markers for stem cell progenitors, for megakaryocytic, myeloid, erythroid and lymphoid lineage cells, for differentiation and for *NCAM1* (Supplemental Table S2) following standard protocols.

Apoptosis assay

To understand the functional role of *NCAM1* on cell death and viability, the AnnexinV/DAPI assay was performed. For this assay, cells were stained with an Annexin V-APC conjugated antibody (Supplemental Table S2) and DAPI (500ng/ml) (BioLegend) and suspended in binding buffer, according to standard protocols. This assay was used to determine the amount of viable, apoptotic and necrotic cells within a sample. Between 20,000-60,000 cells were used and stained with AnnexinV to detect apoptotic cells and DAPI to detect necrotic cells. AnnexinV functions by binding to phosphatidyl serine residues on the outside/plasma membrane of dying cells and DAPI migrates to the inside of dying cells/broken/ruptured membrane to bind to DNA content [75, 76]. Viable cell populations are double negative, early apoptotic cells are $DAPI^{neg}/AnnexinV^{+}$, late apoptotic cells are double positive and necrotic cells are $DAPI^{+}/AnnexinV^{neg}$.

Colony formation assays

Colony formation assays (CFAs) are used to measure clonogenic potential of cells or the ability of single cells to produce colonies of daughter cells, to proliferate (colony size) as well as assess cell differentiation and types of hematopoietic progenitor cells present (e.g. CFU-GM, CFU-E, CFU-GEMM, BFU-E, etc…). In this case CFAs were used to understand the functional

role of *NCAM1* and if the knockdown of *NCAM1* impacts a cells ability to produce colonies, proliferate, survive and differentiate. Non-transduced or shRNA transduced cells that were either sorted (for mCherry and/or GFP) or unsorted, were seeded in Methocult H4034 optimum (StemCell Technologies) at a density of 500 cells/ml (*in vitro*) or 1000 cells/ml (*in vivo*) in 35 mm dishes in duplicates. After 14 days of incubation at 37°C, colony number, size and morphology were manually assessed with the EVOS FL Auto imaging system (LifeTechnologies). Filters for mCherry (PE-Texas Red) and GFP were used to differentiate colonies with 10x and 20x objectives.

RNA purification, qRT-PCR, qPCR, and ddPCR

In order to determine the knockdown efficiency of *NCAM1*, transduced cells were lysed in Trizol reagent, RNA was purified and converted to cDNA and relative gene expression was determined with quantitative polymerase chain reaction (qPCR) or digital droplet polymerase chain reaction (ddPCR). Both ddPCR and qPCR function using Taq polymerases for amplification but the main difference between the two is that ddPCR partitions the PCR reactions into thousands of individual droplets which allow for direct and individual quantification of DNA (end-point PCR in each droplet) without the need for standard calibration curves. Standard calibration curves are used in qPCR to assess PCR efficiency and the PCR product is measured in real time after each cycle (exponential phase). ddPCR allows for more precise and reproducible results and the detection of low copy number samples and smaller fold changes in gene expression [77, 78]. Between 40,000-200,000 of Non-transduced or shRNA transduced cells were set aside in 100µl of Trizol lysis reagent and placed at –80°C until RNA purification could be performed. Total RNA was purified and isolated from samples using the Direct-zol RNA microprep kit (Zymo Research) following manufacturer's instructions. Following purification, RNA concentration was measured using a Nanodrop (Thermo Scientific) and RNA quality was assessed using TapeStation systems of the Nanopore sequencing platform at CHUSJ (Montréal, Canada). The complementary DNA (cDNA) was obtained by performing qRT-PCR on 30-70ng of the isolated RNA using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Thermo Fisher Scientific). The thermal cycling conditions were as follows: 25°C for 10min, followed by 37°C for 120min and ending with 85°C for 5min with a holding stage at 4°C.

Prior to qPCR analysis and to confirm the qRT-PCR was successful, cDNA product amplification was performed by PCR using primer pairs specific to *NCAM1* (Forward 5'-CTGCATTGCTGAGAACAAGG-3' and Reverse 5'- AATTCCATGGCAGTCTGGTT-3') and DreamTaq DNA Polymerase (Thermo Fisher Scientific). PCR cycling conditions were as follows: initial denaturation stage at 94°C for 3min for 1 cycle; denaturation stage at 95°C for 45s for 35 cycles,

annealing stage at 60°C for 30s for 35 cycles, extension stage at 72°C for 60s for 35 cycles; final extension stage at 72°C for 5min for 1 cycle, and holding stage 4°C. PCR products were visualized by running a 2% gel electrophoresis at 400amp and 80V for 45min and gels were visualized using the Kodak Gel Logic 2200 Imaging System (Kodak) and ChemiDoc Imaging systems (Bio-Rad) (Expected product size: 105 bp).

Relative gene expression was determined by qPCR analysis performed by the Genomics platform at the Institute of Immunology and Cancer (IRIC) using assays designed with the Universal Probe Library (Roche) and TaqMan™ Fast Advanced Master Mix (Thermo Fisher Scientific) or by ddPCR. For qPCR, cDNA Samples were diluted ½ and plated in triplicates in 96-well plates. The QuantStudio qPCR instrument (Thermo Fisher Scientific) was used to detect the amplification level and relative *NCAM1* expression comparison (2^-ΔΔCT) was calculated using *PSMF1* and *ABL* as endogenous controls and tested samples were compared to control shRenilla sample (calibrator for comparative Ct method). For ddPCR, primers for *NCAM1* were validated prior to experiments using an eight temperature gradient with two different dilutions, to determine the best annealing temperature and specificity of assay. Samples were first converted into droplets using QX200 droplet generator (Bio-Rad) and a primer based reaction setup with 2X QX200 EvaGreen ddPCR supermix (Bio-Rad) was used. ddPCR cycling conditions were as follows: enzyme activation stage at 95°C for 5min for 1 cycle; denaturation stage at 95°C for 30s for 50 cycles, annealing and extension stage at 57°C for 60s for 50 cycles, signal stabilization stages at 4°C for 5min for 1 cycle and 90°C for 5min for 1 cycle and a final holding stage at 12°C. Samples were analyzed with the QX200 Droplet reader (Bio-Rad) and data analysis using QuantaSoft software (Bio-Rad).

In vitro pharmacological assays

Drugs

Pharmacological inhibition was performed by adding MAPK inhibitors or PI3K inhibitors to primary CG2 xenograft cells for a 48 hour period. Trametinib (HY-10999), Cobimetinib (HY-13064), AZD8330 (HY-12058), A-443654 (HY-10425), BGT226 (HY-13334), PIK-75 hydrochloride (HY-13281) were purchased from MedChemExpress LLC and used for pharmacological inhibition of the MAPK and PI3K pathways. All drugs were prepared as 10mM stocks in DMSO.

Dose response curves

Dose response assays were performed to better understand the efficacy of the drugs and to determine the half maximal inhibitory concentrations or IC50s. IC50 is the drug concentration

at which half the biological process is inhibited [79]. Dose response curves were generated at the High-throughput screening core facility of the IRIC institute. The curves were generated by seeding 2000-5000 cells per well in a 384-well plate in quadruplicates with either inhibitor drugs or vehicle (≤ 0.1% DMSO). The drugs were provided as either 1mM or 10mM stocks and 10 doses of the drugs were tested, beginning at 1μM or 10μM, 3-fold dilutions down to generate 10 data points on the curve. The read-out was performed after a 6-day incubation at 37°C and 5% CO_2 with Cell titer-Glo© luminescent cell viability assay measuring ATP in metabolically active cells (Promega). The percentage of viability was calculated as follows: 100 x (mean luminescence compound/mean luminescence DMSO vehicle) and the IC50s were calculated using nonlinear regression analysis (four parameters) in GraphPad Prism 8 software.

Drug synergies

Synergy assays were performed to understand the *in vitro* interactions between inhibitory drugs on inhibiting biological processes and if the interactions are additive, antagonistic, synergistic or have no effect. Synergy assays were performed at the High-throughput screening core facility of the IRIC institute. Synergy was interpreted by taking into consideration four reference models, the Bliss independence model (BLISS), the Loewe additivity model (Loewe), the Highest Single Agent model (HSA) and the Zero interaction potency model (ZIP). Around 2000-5000 cells were seeded per well in quadruplicates in a 384-well plate and combined with either two inhibitors or vehicle (≤ 0.1% DMSO). The drugs were provided as either 1mM or 10mM stocks and 8 doses of the drugs were tested, beginning at 1μM or 10μM, 3-fold dilutions down to generate 8x8 matrices and single agent dose-response curves. The read-out was performed after a 6 day incubation at 37°C and 5% CO_2 with Cell titer-Glo© luminescent cell viability assay (Promega). The average synergy scores were calculated for each model using R (v3.6.1) package Synergyfinder (v2.0.12) software [80] and visualized in 2D and 3D synergy maps using the python package synergy (v.0.3.6) [81]. Usually a score ≤ –10 can be interpreted as an antagonistic response, between –10 to 10 as additive and ≥10 as synergistic. These values give a general idea of the type of interaction occurring but require further testing to confirm.

***In vitro* drug studies**

To further understand the function of *NCAM1* in AMKL, *NCAM1* KD was combined with inhibiting downstream MAPK and PI3K pathways. Around 6 million cells for each Non-transduced or shRNA transduced conditions were seeded and transduced at an MOI of 12-15 (Day 0) and 72 hours later (Day 3) the cells were assessed for gene transfer efficiency and knockdown efficien-

cy by flow cytometry and qPCR. The remaining cells were sorted at the Flow Cytometry Core facility at IRIC for mCherry+ cell population. The sorted cells were either 1). sent to High-throughput screening core facility of the IRIC institute for dose responses and 2). seeded in triplicates with either Trametinib (50nM), A-443654 (10nM) or DMSO (≤ 0.1%). The drug treated and DMSO treated cells were incubated for an additional 48 hours and on Day 5 were assessed for knockdown efficiency by flow cytometry and qPCR and several functional assays looking at different cell fate outcomes such as cell proliferation, cell death and cell viability (Figure 21A).

Statistics

The data is represented as Mean +/- SEM using Graphpad Prism 8 software. Multiple t-tests were performed, normalization, and nonlinear regression (four parameters) with GraphPad prism as well. Significance depicted as P-values: * <0.05, ** <0.01, *** <0.001, **** <0.0001.

Results

Analysis of *CBFA2T3::GLIS2* AMKL bulk sample datasets uncovered a defined megakaryocytic and erythroid signature with high and differential *NCAM1* expression.

Transcriptomic and differential gene expression analyses were performed on bulk RNA sequencing datasets which included an institutional CHUSJ dataset with *CBFA2T3::GLIS2* AMKL models (N=10) & patient samples (N=2), *NUP98::KDM5A* AMKL models (N=5) & patient samples (N=2) and normal CB-CD34+ cells (N=4) and a validation cohort dataset from St-Judes which includes both *CBFA2T3::GLIS2* AMKL (N=12) and other genetic subtypes of AMKL from pediatric patients at diagnosis (N=61) (Figure 10A). This data depicts that our mCG2 models correlate closely with CG2 patient samples and have high expression of *NCAM1* and AMKL-associated markers CD44, *SPN* (CD43) and CD34 [32]. A common gene signature of upregulated (N=399) and downregulated (N=330) genes was defined for CG2 AMKL. *NCAM1* is both common to and highly expressed in AMKL and when compared to all the other genes expressed, it is the most differentially expressed between all samples within the datasets (Figure 10A). Furthermore, our mCG2 models cluster with CG2 patient samples (N=14) by expression profiling PCA analysis and they cluster apart and differentially than other AMKL genotypes (NUP98) and normal CB-CD34+ cells (Figure 10B).

Figure 10

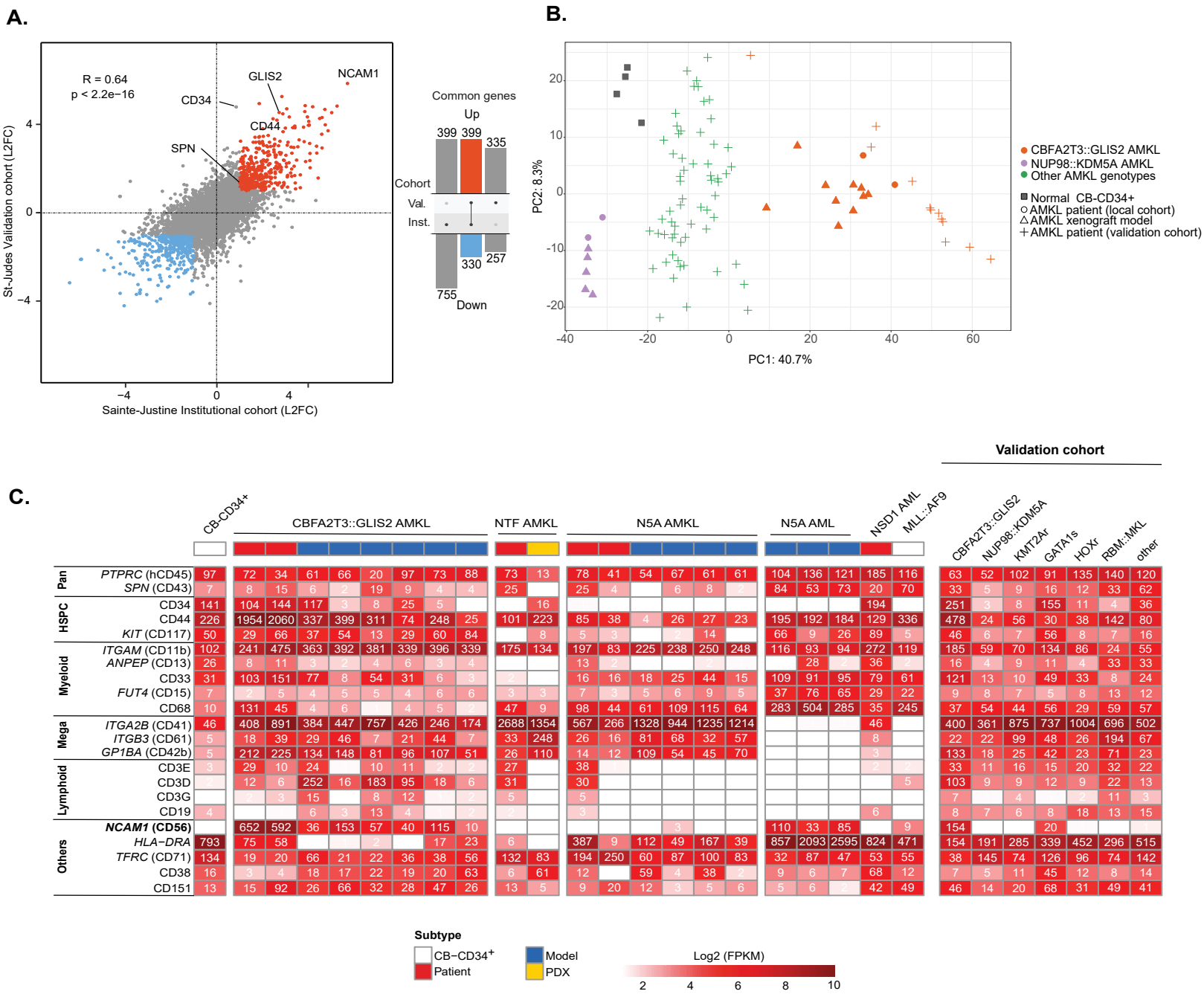


Figure 10: Transcriptomic profiling and differential gene analyses of *CBFA2T3::GLIS2* AMKL model samples correlate with the patient disease and demonstrate mega-erythroid signature with high *NCAM1* expression. A). Left panel, correlation of differential gene expression (log2 fold-change (L2FC)) in *CBFA2T3::GLIS2* (CG2) acute megakaryoblastic leukemia (AMKL) models and patients from our institutional cohort (CHUSJ) compared with a validation cohort of pediatric CG2 AMKL (St. Jude). Differentially expressed genes, defined as |L2FC>1| and FDR q-value < 0.05, common to both datasets are indicated in blue or orange and define the CG2 signature. Institutional (Inst.) dataset: CG2 AMKL models (N=10) and CG2 patient samples (N=2) compared with *NUP98::KDM5A* (N5A) AMKL models (N=5), N5A patient samples (N=2)

and normal CB-CD34+ cells (N=4). Validation (Val.) dataset: CG2 AMKL (N=12) vs other genetic subtypes of AMKL (N=61) from pediatric patients at diagnosis. Right panel, upset plots showing differentially expressed genes that are jointly over- (N=399) or under-expressed (N=330) in CG2 leukemias, corresponding to blue and orange dots in Left panel. **B).** Principal component analysis (PCA) plot of *CBFA2T3::GLIS2* (CG2, orange) and *NUP98::KDM5A* (N5A, purple) model (▲) and patient-derived (●) leukemias, normal cord blood (CB)-CD34+ cells (gray, ■) and a published validation dataset of pediatric AMKL [21] (green, +) calculated with the 729 differentially expressed genes associated with CG2 signature (common up/down genes described in Fig. A). **C).** Heatmap depicting RNAseq gene expression (FPKM) of selected hematopoietic lineage markers in leukemic cells derived from *CBFA2T3::GLIS2* (CG2) leukemia, as compared with normal cord blood cells (CB-CD34+), *NUP98* rearranged AMKL (*NUP98::BPTF*, NTF; *NUP98::KDM5A*, N5A) and *NUP98::NSD1* and *MLL-AF9* AML from our institutional cohort. Gene expression in a validation cohort [21] is represented as mean expression per indicated AMKL genotype (right panels). Pan, Pan-hematopoietic; HSPC, hematopoietic stem and progenitor cells; Mega, megakaryocytic; *KMT2Ar* or *HOXr, KMT2A* or *HOX* rearranged; *GATA1*s, *GATA1* truncating-mutation; *RBMKL*, *RBM15::MKL1*; PDX, patient derived xenograft.

A further look into selected gene expression by profiling of bulk RNAseq datasets (same institutional and validation cohorts) shows that CG2 leukemic cells have an early megakaryocytic and erythroid lineage signature. Both CG2 model and patient samples have remarkably high *NCAM1* as well as high HSPC marker CD34, Pan-hematopoietic-associated hCD45, *SPN*, AMKL-associated CD41, CD61, CD42 and AML-marker CD44 expression compared to other AMKL and AML subtypes [28, 32, 82, 83] with the absence/low expression of myelo-monocytic markers (*FUT4*/CD15, *ITGAM*/CD11b and CD68)) [84] (Figure 10C). This megakaryocytic and erythroid marker signature is further validated by immunophenotyping (Figure 11 and Supplemental Figures S1 and S2) and assessment of surface protein expression. This confirmed that CG2 models express highly *NCAM1* as well as megakaryocytic markers compared to other AMKL and AML subtypes at both a protein and gene level.

Figure 11

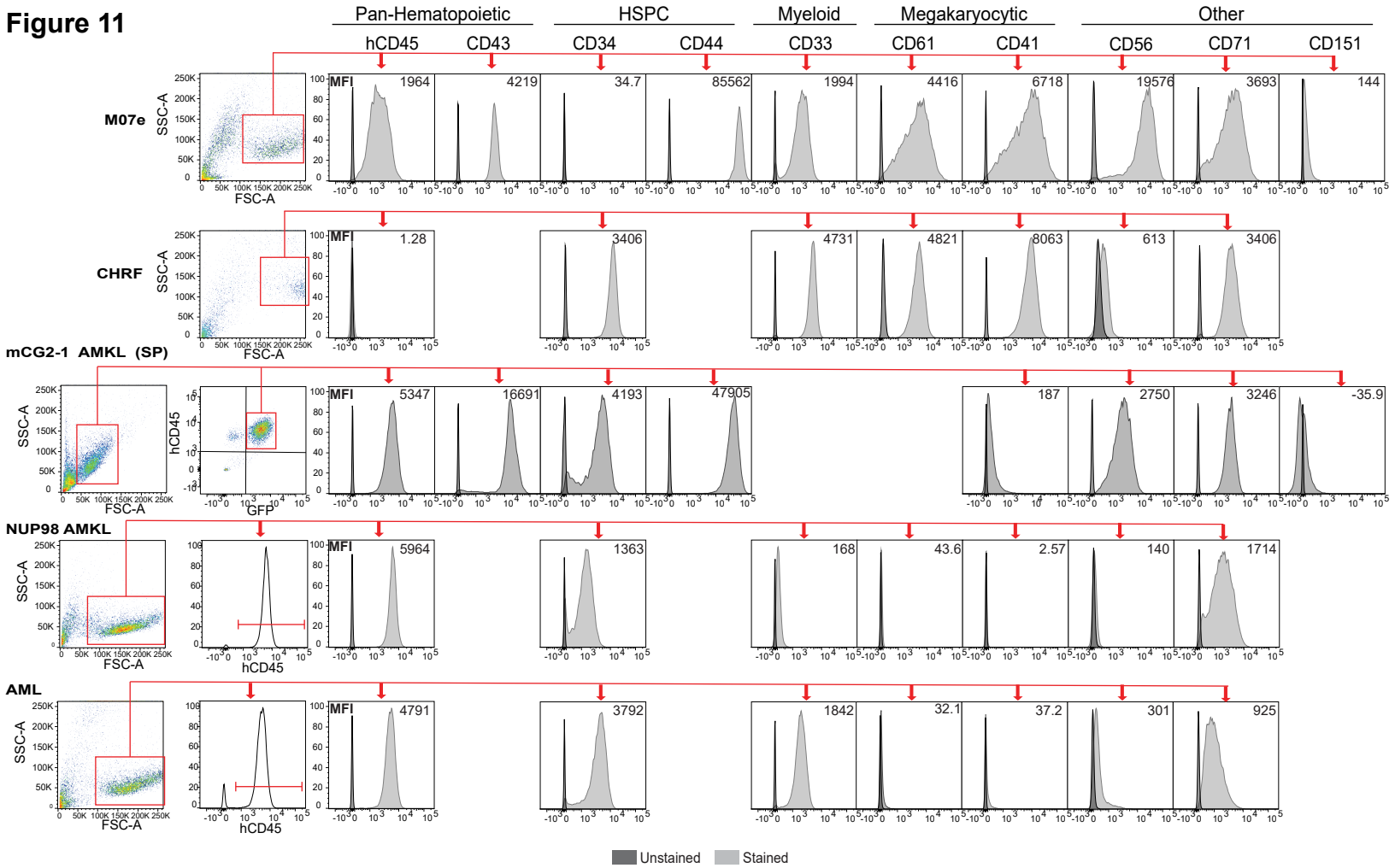


Figure 11: *CBFA2T3::GLIS2* AMKL cell line and model cells express megakaryocytic and erythroid markers with high *NCAM1* expression at the cell surface. Flow cytometry pseudo color dot plots depicting gating strategy and flow cytometry plots depicting mean fluorescence intensity (MFI) of selected hematopoietic lineage markers (Pan-hematopoietic, HSPC, Myeloid, megakaryocytic, and other markers) in the M07e cell line, CHRF cell line, Primary mCG2-1 and *NUP98* AMKL model (O969) spleen cells and primary AML model (O330) spleen cells. M07e and CHRF are gated on viable cells. mCG2-1 AMKL spleen cells first gated on viable cells, then gated on hCD45 and GFP double positive cells. *NUP98* AMKL and AML cells first gated on viable cells and then on hCD45 positive cells. Dark grey plots are unstained, light gray plots are stained. SP, spleen; HSPC, hematopoietic stem and progenitor cells

Analysis of single cell CG2 AMKL sample datasets defined megakaryocytic and erythroid lineage signature and heterogeneity between samples.

Single cell RNA sequencing (scRNAseq) was performed in order to determine at a single cell level differences in gene expression and to uncover populations of cells within heterogeneous samples that have common gene signatures and lineage compositions. A comparison of

normal bipotent megakaryocyte-erythroid progenitors (MEP), megakaryocyte progenitors (MKP) and platelets to all other hematopoietic lineages from normal BM cell populations datasets was performed (datasets (Human Cell Atlas, HCA58)) (Figure 12A). Using these identification labels, lineage and cellular composition was further assessed by scRNAseq in CG2 AMKL from models (N=4) and pediatric patients (N=2) compared to a diverse phenotypic and genetic subtype of adult AML. This demonstrates that CG2 AMKL from models or patients overall favors a more megakaryocytic progenitor signature with varying erythroid committed cells (Figure 12B Left, orange/green/blue) compared to other samples of AML (grey shaded samples in Figure 12B Left). AML has a more heterogenous lineage composition (Figure 12B Right) compared to AMKL which has a more specific MKP/MEP/platelet lineage composition. Even within AMKL samples (mCG2-1, mCG2-2, mCG2-5 and mCG2-6), there is MEP proportion heterogeneity seen with the mCG2-6 model partially projecting into the erythroid space (blue shading radar plot, Figure 12B Left).

Heterogeneity is further seen with UMAPS of *NCAM1*, *ITGA2B*/CD41 and CD34 in model and patient CG2 samples. Here, *NCAM1* is homogeneously expressed within model and patient samples but mCG2-6 has a more heterogenous expression with some cell populations not expressing *NCAM1*. The MEP/MKP/platelet specific genes such as *ITGA2B* (CD41) [85] is homogeneously expressed across all CG2 AMKL cells. Stem cell marker CD34, is homogeneously expressed between some but not all CG2 samples (Figure 12C).

Several AML cell lines (ML-2, U937, HEL, HL-60), AMKL cell lines (CHRF, CMK, M07e) and one chronic myelogenous leukemia cell lines (K562) and several synthetic mice models harboring CG2, *NUP98* and *T21* fusions were assessed for *NCAM1* protein surface expression (MFI) with flow cytometry. Heterogeneous expression levels of *NCAM1* was found between cell lines, synthetic models and different AML subtypes (ranging between 4%-90% *NCAM1* positive cells). The M07e cell line which harbors the *CBFA2T3::GLIS2* fusion [31] had significantly higher *NCAM1* expression compared to the U937 cell line and ML-2 cell lines which do not have the CG2 fusion and were also positive for *NCAM1*. There was low/absent *NCAM1* expression in the CHRF, CMK, HEL, HL-60, and K562 cell lines (Figure 12D). All synthetic models harboring the *CBFAT3::GLIS2* fusion (M894, M902, N783 and M893) were positive for *NCAM1* and had higher expression compared to the other AML subtype models (Figure 12E). Aberrant *NCAM1* expression can also be seen in CG2 model bone marrow (femur) and spleen samples that underwent immunohistochemical *NCAM1* staining. Both the spleen and bone marrow from a CG2 leukemic mice compared to the healthy tissues from a Non-leukemic mouse show higher level of *NCAM1* protein expression (Figure 12F, brown staining).

Figure 12

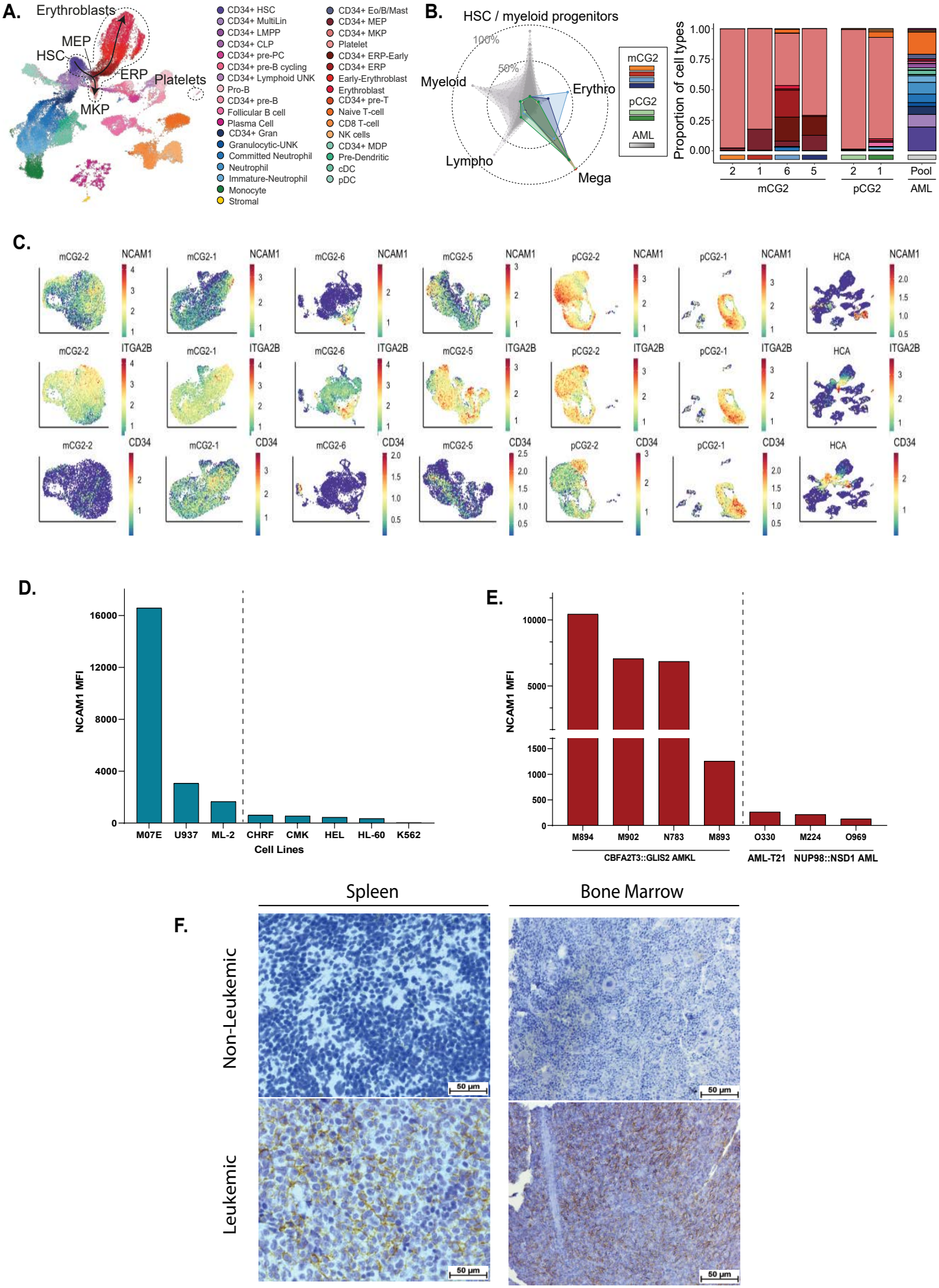

A.
Erythroblasts
MEP
HSC
ERP
Platelets
MKP
CD34+ HSC
CD34+ MultiLin
CD34+ LMPP
CD34+ CLP
CD34+ pre-PC
CD34+ pre-B cycling
CD34+ Lymphoid UNK
Pro-B
CD34+ pre-B
Follicular B cell
Plasma Cell
CD34+ Gran
Granulocytic-UNK
Committed Neutrophil
Neutrophil
Immature-Neutrophil
Monocyte
Stromal
CD34+ Eo/B/Mast
CD34+ MEP
CD34+ MKP
Platelet
CD34+ ERP-Early
CD34+ ERP
Early-Erythroblast
Erythroblast
CD34+ pre-T
Naive T-cell
CD8 T-cell
NK cells
CD34+ MDP
Pre-Dendritic
cDC
pDC
B.
HSC / myeloid progenitors
100%
50%
Myeloid
Erythro
Lympho
Mega
mCG2
pCG2
AML
Proportion of cell types
1.00
0.75
0.50
0.25
0.00
2
1
6
5
2
1
Pool
mCG2
pCG2
AML
C.
mCG2-2
mCG2-1
mCG2-6
mCG2-5
pCG2-2
pCG2-1
HCA
NCAM1
ITGA2B
CD34
D.
NCAM1 MFI
16000
12000
8000
4000
0
M07E
U937
ML-2
CHRF
CMK
HEL
HL-60
K562
Cell Lines
E.
NCAM1 MFI
10000
5000
1500
1000
500
0
M894
M902
N783
M893
O330
M224
O969
CBFA2T3::GLIS2 AMKL
AML-T21
NUP98::NSD1 AML
F.
Spleen
Bone Marrow
Non-Leukemic
Leukemic
50 µm

Figure 12: *CBFA2T3::GLIS2* AMKL samples demonstrate a megakaryocytic-erythroid lineage signature with high and homogeneous *NCAM1* expression whereas expression levels vary between AML subtypes. A). Single cell analysis (UMAP) of normal bone marrow lineages. The stem-mega-erythroid compartment is circled, relevant populations are labeled, and differentiation trajectories are highlighted with arrows. **B).** Radar plot (Left) of lineage composition assessed by scRNAseq in CG2 models and pediatric patients of AMKL, as compared to diverse phenotypic and genetic subtypes of adult AML. The detailed proportion of each cell type is presented as stacked bar plots (Right). Color coding of populations is as depicted in (A). **C).** UMAPS embedded gene expression of selected markers (CD56/*NCAM1*, CD41/*ITGA2B* and CD34) in AMKL mice models (mCG2-2, mCG2-1, mCG2-6 and mCG2-5) and CG2 AMKL patient samples (pCG2-1 and pCG2-2). CG2, *CBFA2T3::GLIS2*; HCA, Human Cell Atlas. **D).** Bar graph of *NCAM1* mean fluorescence intensity (MFI) in AML cell lines (ML-2, U937, HEL, HL-60), AMKL cell lines (CHRF, CMK, M07e) and one chronic myelogenous leukemia cell lines (K562). The dotted line separates *NCAM1* positive (Left) and *NCAM1* negative (Right) cell lines. **E).** Bar graph of *NCAM1* mean fluorescence intensity (MFI) in several synthetic AMKL mice models harboring the *CBFA2T3::GLIS2* fusion (M894, M902, N783 and M893) and AML models harboring the *NUP98::NSD1* fusion (M224 and O969) and the *T21* fusion. The dotted line separates *NCAM1* positive (Left) and *NCAM1* negative (Right) cell lines. **F).** Immunohistochemical *NCAM1* stained portions of mice spleen and femur (bone marrow). Upper Left is the stained spleen portion of a Healthy non-leukemic mouse and Bottom Left is the spleen of a leukemic mCG2 AMKL mouse. Upper right is the stained femur portion of a healthy non-leukemic mouse and Bottom right is the femur of a leukemic mCG2 AMKL mouse.

Overall, our mCG2 models cluster with CG2 patient samples and apart from other AML subtypes harboring different fusions, meaning our mCG2 models are representative of disease. *CBFA2T3::GLIS2* AMKL has a defined megakaryocytic and erythroid marker signature with less lineage heterogeneity compared to AML subtypes. CG2 AMKL has remarkably high and differential and homogeneous expression of *NCAM1* compared to other AML subtypes, making *NCAM1* an interesting potential therapeutic biomarker.

Functional role of *NCAM1* in CG2 AMKL

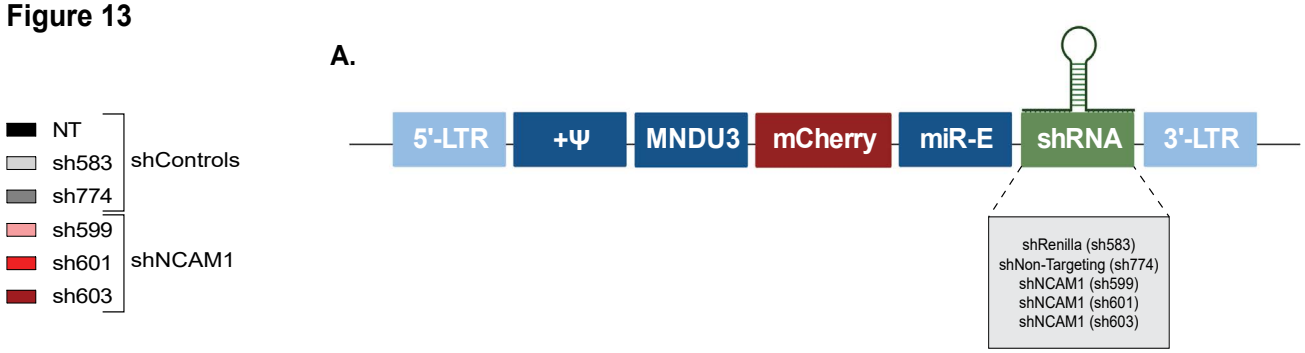
The aim of this study is to functionally assess the role of *NCAM1* in sustaining leukemic cells *in vitro* and leukemia engraftment and progression *in vivo*. *NCAM1* knockdown (KD) was performed both *in vitro*, using the M07e cell line carrying the CG2 fusion and our model cells and

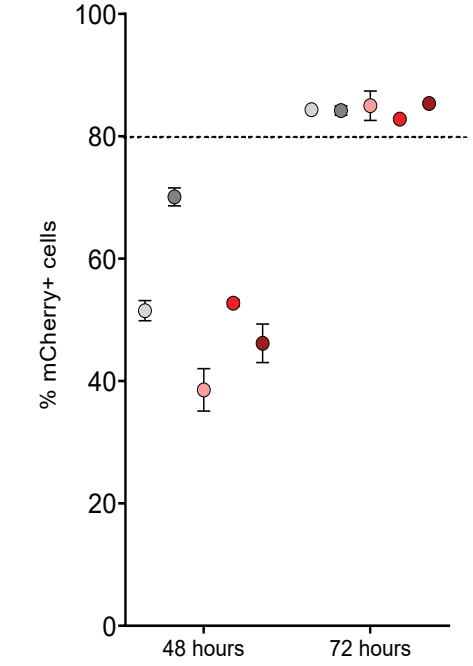
in vivo, using our mCG2 models. Several functional assays were performed to assess the impact of *NCAM1* KD on different cell fate outcomes of proliferation, apoptosis, viability, immunophenotype, and clonogenic potential. Pharmacological inhibition was combined with *NCAM1* KD to further understand the role of *NCAM1* in downstream signaling pathways.

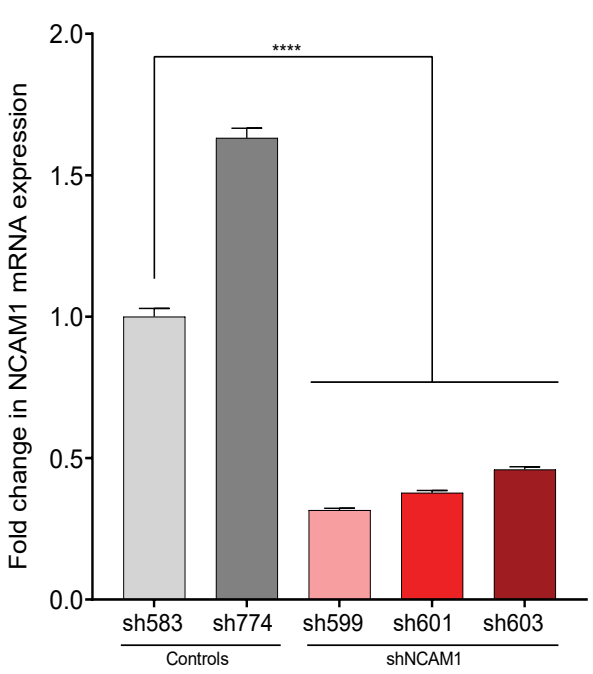
Lentiviral shRNA targeting control genes and *NCAM1* have high gene transfer efficiency and gene knockdown capabilities.

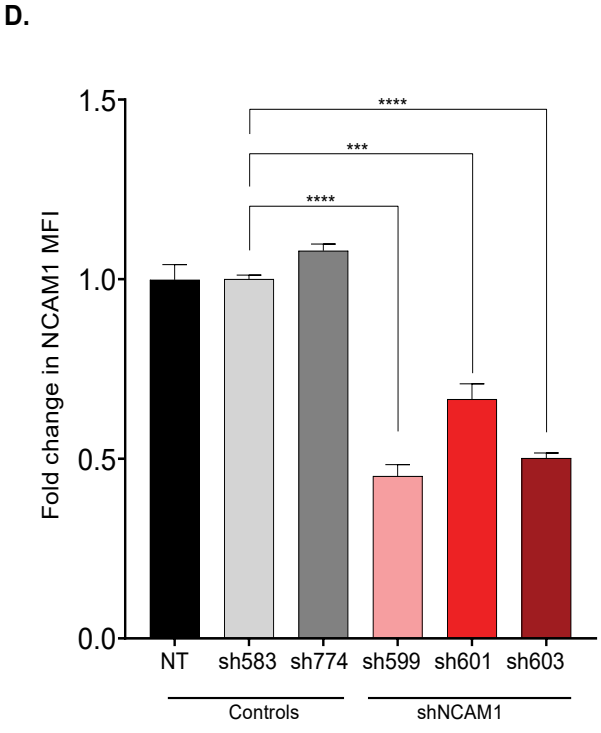
Prior to performing *in vitro* and *in vivo* lentiviral shRNA mediated *NCAM1* knockdown and assessing the function of *NCAM1*, the knockdown efficiency of the shRNAs generated had to be validated. The M07e cell line (NCAM+) was used to validate the following shRNAs; shRenilla (583) control, shNon-targeting (774) control, sh*NCAM1* (599), sh*NCAM1* (601), and sh*NCAM1* (603). Around 250,000 cells were seeded and transduced at an MOI of 5 in triplicates for each shRNA lentiviral vector. Gene transfer efficiency was assessed by measuring the percentage of viable cells expressing the reporter gene mCherry by flow cytometry. Knockdown efficiency was assessed by measuring *NCAM1* protein surface expression (MFI) by flow cytometry and gene expression by ddpcr, respectively. Both gene transfer (e.g. transduction) and knockdown efficiencies were assessed 48- and 72-hours post transduction. The gene transfer efficiency or % mCherry+ cells for shRNA transduced cell conditions ranged between 30-70% at 48 hours which was significantly lower than 72 hours post transduction which ranged between 80-85% (Figure 13B). Flow cytometry analysis demonstrated that gene transfer efficiency was less variable between shRNA transduced cell conditions and greater and more efficient at 72 hours compared to 48 hours for all shRNA transduced conditions. Therefore, 72 hours post transduction was selected as the timepoint for all functional analyses. At 72hrs, all three sh*NCAM1*-transduced cell populations (sh599, 601 and 603) showed significant decreases in *NCAM1* mRNA expression compared to the shRenilla control cells, as assessed by ddPCR (Figure 13C and 48hr shown in Supplemental Figure S3A). Transcript levels relative to control ranged between 30-50% (levels of knockdown 50-70%). *NCAM1* protein surface expression or mean fluorescence intensity (MFI) measured by flow cytometry at 72 hours post transduction showed 1.7-fold, 2.2-fold and 2.0-fold decreases (around 50% reduction in protein levels) of *NCAM1* expression in the three sh*NCAM1*-transduced cell populations (599, 601 and 603) in comparison to shRenilla-transduced cells, respectively (Figure 13D-E and 48hr shown in supplemental Figure S3B-C). Overall, the five shRNA transduced conditions had high gene transfer efficiencies and the three sh*NCAM1*transduced conditions had over 50% knockdown of *NCAM1*. The two shControl transduced conditions acted similarly and will be good controls for future experiments.

Figure 13


A.
NT
sh583
sh774
shControls
sh599
sh601
sh603
shNCAM1
5'-LTR
+Ψ
MNDU3
mCherry
miR-E
shRNA
3'-LTR
shRenilla (sh583)
shNon-Targeting (sh774)
shNCAM1 (sh599)
shNCAM1 (sh601)
shNCAM1 (sh603)


B.
% mCherry+ cells
100
80
60
40
20
0
48 hours
72 hours


C.
Fold change in NCAM1 mRNA expression
2.0
1.5
1.0
0.5
0.0
****
sh583
sh774
sh599
sh601
sh603
Controls
shNCAM1


D.
Fold change in NCAM1 MFI
1.5
1.0
0.5
0.0
****
***
****
NT
sh583
sh774
sh599
sh601
sh603
Controls
shNCAM1

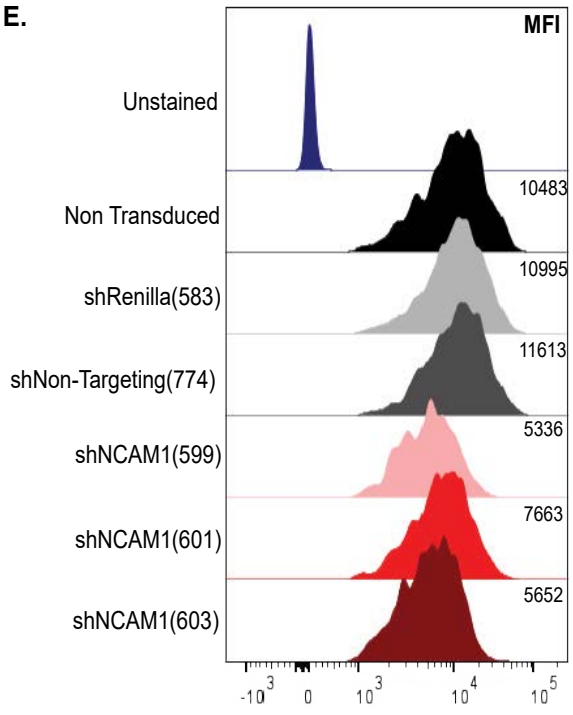

E.
MFI
Unstained
Non Transduced
10483
shRenilla(583)
10995
shNon-Targeting(774)
11613
shNCAM1(599)
5336
shNCAM1(601)
7663
shNCAM1(603)
5652
-10^3
0
10^3
10^4
10^5

Figure 13: Validation of lentiviral shRNA targeting control genes and *NCAM1* demonstrated high gene transfer and gene knockdown efficiencies. A). Structure of the lentiviral vectors used for short hairpin (sh) mediated gene knockdown. Between the 5'- and 3'- LTR regions all vectors contain the retroviral Ψ packaging sequence, an MNDU3 promoter, followed by an mCherry fluorescent reporter gene, a miR-E backbone, and the insert of the short hairpin (shRNA) against the gene of interest (e.g sh*NCAM1* and shRenilla and shNon-targeting as controls). Figure created with BioRender.com. **B).** Percentage gene transfer rates (%mCherry+ cells) 48- and 72- hours post transduction with indicated shRNA against *NCAM1* in comparison to control vectors (shRenilla (583) and Non-targeting (774)) or Non-transduced (NT) cells in the M07e cell line. Depicted as Mean +/- SEM, N=3/sh. **C).** Relative *NCAM1* gene expression detected by ddPCR normalised against *ABL* endogene and calibrated to control condition shRenilla (sh583). Depicts fold change in *NCAM1* mRNA expression between shControl conditions and sh*NCAM1* conditions 72 hours post transduction in the M07e cell line. Multiple t-tests performed, p less than 0.05. Depicted as Mean +/- SEM, N=3/sh. **D).** *NCAM1* surface protein fold change in MFI quantified by flow cytometry 72 hours post transduction with indicated shRNA against *NCAM1* in comparison to control vectors (shRenilla (583) and Non-targeting (774)) or Non-transduced (NT) cells in the M07e cell line. Normalised against shRenilla, multiple t-tests performed, depicts Mean +/- SEM, N=3/sh. **E).** Flow cytometry histograms showing visual shifts in *NCAM1* surface protein expression 72 hours post transduction in the M07e cell line. Depicts shifts in *NCAM1* MFI for indicated shRNA against *NCAM1* (sh599, 601, 603) in comparison to control vectors (shRenilla and Non-targeting) or Non-transduced cells. P-values: * <0.05, ** <0.01, *** <0.001, **** <0.0001. MFI, mean fluorescence intensity. SEM, standard error of the mean.

The gene transfer and knockdown efficiencies for functional analyses *in vitro* were efficient

To study the functional role of *NCAM1 in vitro*, lentiviral shRNA mediated *NCAM1* gene knockdown was performed on the M07e (*NCAM1*+), U937 (*NCAM1*+) and K562 (*NCAM1*neg) cell lines. Around 400,000 cells of each cell line were transduced at an MOI of 4 with two shRNA-control vectors (shRenilla (sh583) and shNon-targeting (sh774)) and three shRNA-*NCAM1* vectors (sh599, sh601 and sh603) in triplicates. Gene transfer efficiency was assessed by measuring the percentage of viable cells expressing the reporter gene mCherry by flow cytometry. Knockdown efficiency was assessed by measuring *NCAM1* protein surface expression (MFI) and gene expression by flow cytometry and qPCR respectively. Both gene transfer and knockdown efficiencies were assessed 24-, 48- and 72- hours post transduction and only the results at 72hrs are presented for clarity (Figure 14A).

Figure 14

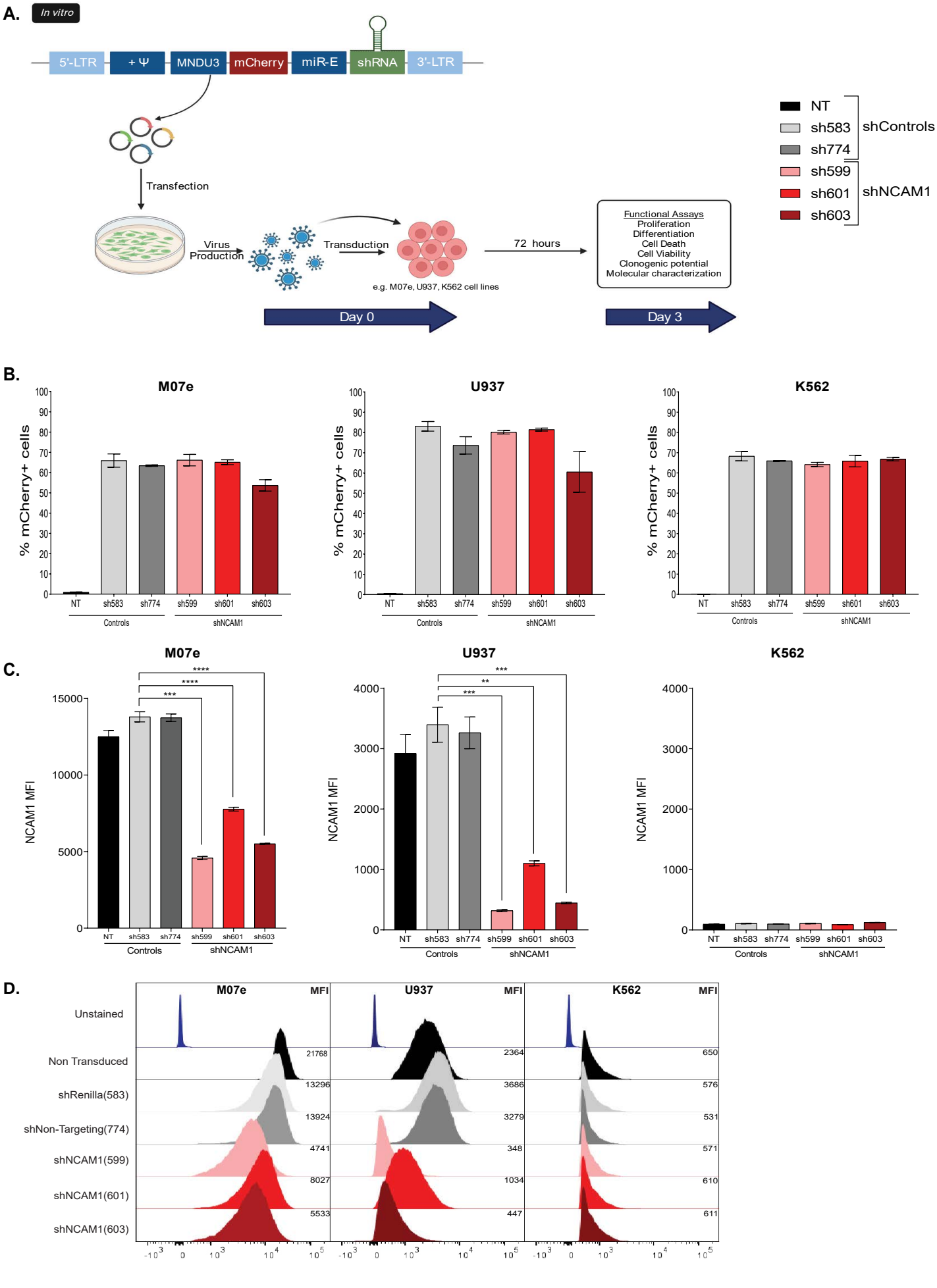

A.
In vitro
5'-LTR
+ Ψ
MNDU3
mCherry
miR-E
shRNA
3'-LTR
Transfection
Virus Production
Transduction
e.g. M07e, U937, K562 cell lines
72 hours
Functional Assays
Proliferation
Differentiation
Cell Death
Cell Viability
Clonogenic potential
Molecular characterization
Day 0
Day 3
NT
sh583
sh774
shControls
sh599
sh601
sh603
shNCAM1
B.
M07e
U937
K562
% mCherry+ cells
Controls
shNCAM1
C.
NCAM1 MFI
D.
MFI
Unstained
Non Transduced
shRenilla(583)
shNon-Targeting(774)
shNCAM1(599)
shNCAM1(601)
shNCAM1(603)
21768
13296
13924
4741
8027
5533
2364
3686
3279
348
1034
447
650
576
531
571
610
611

Figure 14 continued.

E.

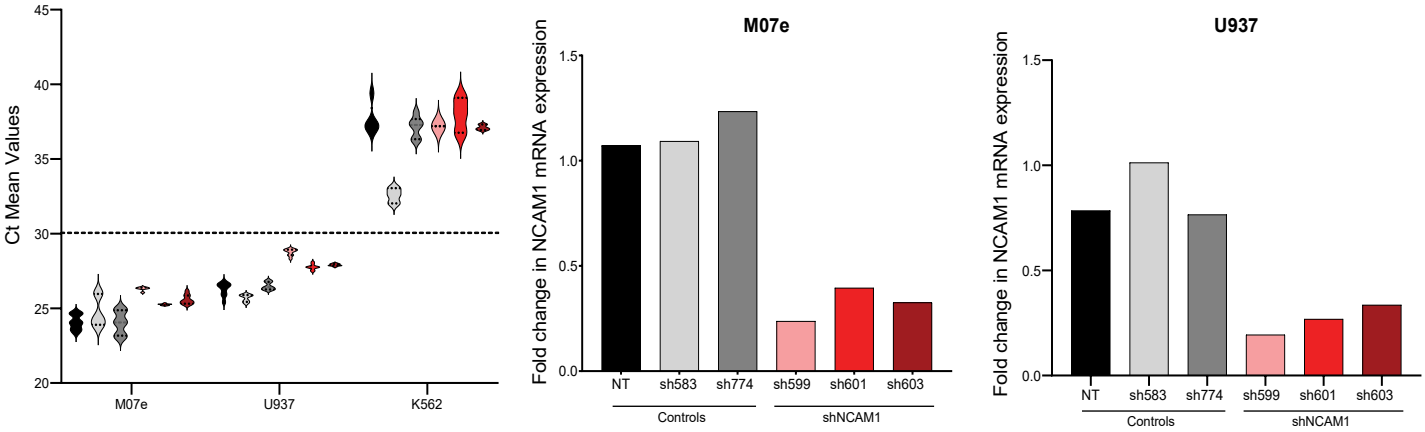


Figure 14: Gene transfer and gene knockdown were efficient in the M07e, U937 and K562 cell lines. A). Outline of *in vitro NCAM1* KD experiment. Top: The lentiviral vector used for short-hairpin (sh) mediated gene knockdown. Between the 5'- and 3'- LTR regions all vectors contain the retroviral Ψ packaging sequence, an MNDU3 promoter, an mCherry fluorescent reporter gene, a miR-E backbone, and the cDNA which contains the insert of the short hairpin (shRNA) against the gene of interest (e.g sh*NCAM1* and shRenilla and shNon-targeting as controls). Bottom: A summary of the methodologies. HEK293T cells were transfected with lentiviral packaging vectors and each shRNA vector to produce the lentivirus used for shRNA mediated gene knockdown. On Day 0, the M07e, U937 and K562 cell lines were transduced. In vitro functional analysis was performed at 72-hours post transduction (Day 3), to assess different cell fates. Figure created with BioRender.com. **B).** Percentage gene transfer rates (%mCherry+ cells) 72 hours post transduction with indicated shRNA against *NCAM1* in comparison to control vectors (shRenilla (583) and Non-targeting (774)) or Non-transduced (NT)) cells in the M07e (Left), U937 (Middle) and K562 (Right) cell lines. Depicted as Mean +/- SEM, N=3/sh. **C).** *NCAM1* surface protein expression (MFI) quantified by flow cytometry 72 hours post transduction with indicated shRNA against *NCAM1* in comparison to control vectors (shRenilla (583) and Non-targeting (774)) or Non-transduced (NT)) cells in the M07e (Left), U937 (Middle) and K562 (Right) cell lines. Multiple t-tests performed, depicts Mean +/- SEM, N=3/sh. **D).** Flow cytometry histograms showing visual shifts in *NCAM1* surface protein expression with MFI values 72 hours post transduction in the M07e (Left), U937 (Middle) and K562 (Right) cell lines. Depicts visual shifts in *NCAM1* MFI for indicated shRNA against *NCAM1* (sh599, 601, 603) in comparison to control vectors (shRenilla and Non-targeting) or Non-transduced cells. N=3/sh. **E).** Mean cycle threshold (Ct) values of mRNA expression in the M07e, U937 and K562 cell lines (Left). Relative *NCAM1* gene expression detected by qPCR normalised against *ABL* and *PSFM1* endogenous reference genes and calibrated to control condition

shRenilla (sh583). Depicts fold change in *NCAM1* expression between shControl conditions and sh*NCAM1* conditions 72 hours post transduction in the M07e (Middle), U937 (Right) cell lines. Depicted as Mean +/- SEM, N=3/sh. P-values: * <0.05, ** <0.01, *** <0.001, **** <0.0001. MFI, mean fluorescence intensity. SEM, standard error of the mean.

NCAM1 KD *in vitro* had no effect on cell proliferation, death, viability, and clonogenic potential

Several functional assays were performed after *NCAM1* shRNA-mediated KD *in vitro* in the M07e, U937 and K562 cell lines to assess different cell fate outcomes such as cell proliferation, cell death and viability and clonogenic potential (Figure 14A).

Cell proliferation *in vitro*

Cell proliferative capacity was measured by performing cell counts using a fixed volume of cells and flow cytometry HTS. Around 400,000 cells were transduced in triplicates for each cell line and cell counts were performed on the day of transduction (Day 0) and 24-, 48- and 72- hours post transduction. The number of cells increased with increasing timepoints being the highest at 72 hours post transduction for all three cell lines. The cell number between Non-transduced (NT) conditions and shRNA transduced (sh583, sh774 and sh*NCAM1*s) conditions did not vary significantly. The cell number for the two shControl transduced conditions (sh583 and sh774) compared to the three sh*NCAM1* transduced (sh599, sh601 and sh603) conditions also did not vary significantly. The M07e cell line appears to proliferate less (cell counts ≤ 1000000) compared to the U237 and K562 cell lines (cell counts ≤ 2500000). The cell number values were corrected for the volume removal at earlier timepoints (Figure 15A). Overall, *NCAM1* KD did not significantly increase or decrease cell proliferation.

Cell death and viability *in vitro*

Cell death and cell viability were assessed using the AnnexinV/DAPI assay and flow cytometry. Cell viability and cell death were measured 24-, 48- and 72- hours post transduction with only the 72hour data represented. Percentages of viable, apoptotic and necrotic cell populations were determined through gating strategies on FlowJo (Figure 15B). The percentage of viable cells 72 hours post transduction was found to be between 80-95% for all three cell lines and all Non-transduced and transduced conditions. The percentage of viable cells did not vary significantly between NT and shRNA transduced conditions and between shControl transduced (sh583 and sh774) and sh*NCAM1* transduced (sh599, 601 and 603) conditions for all three cell lines (Figure

15C). The percentage of necrotic cells was below 12% for all cell lines and all conditions, with no significant differences (Figure 15C). The percentage of early and late apoptotic cells (AnnexinV+ cells) 72 hours post transduction, was found to be below 10% for the M07e and U937 cells lines and below 20% for the K562 cell line (Figure 15C-D). The percentage of annexin positive cells did not vary significantly between NT and shRNA transduced conditions and between shControl transduced and sh*NCAM1* transduced conditions for all three cell lines (Figure 15C-D). Overall, *NCAM1* KD did not impact cell viability and cell death.

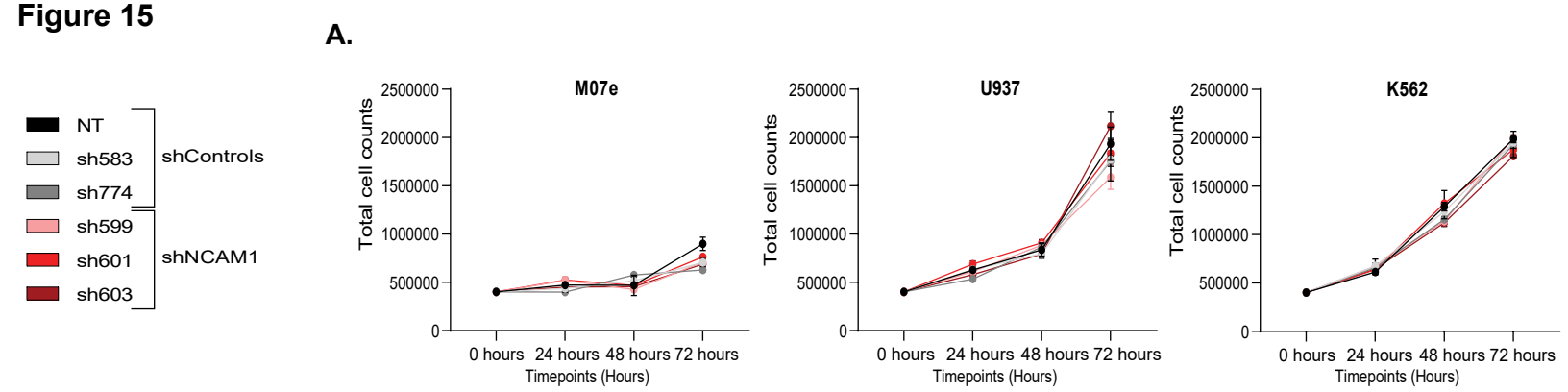
Clonogenic potential *in vitro*

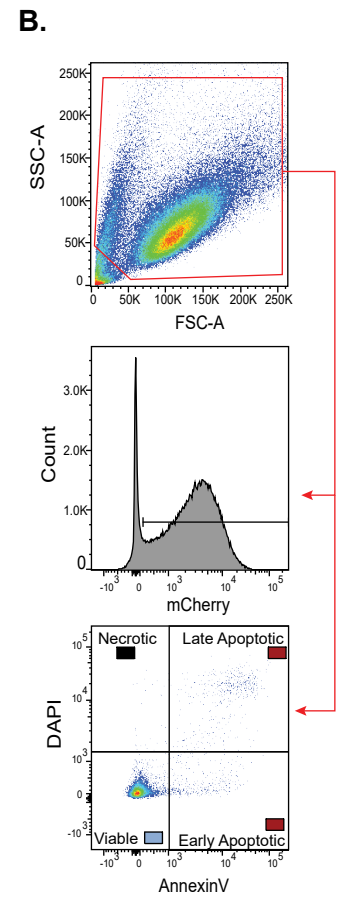
Cell clonogenic potential was assessed by performing colony formation assays for the M07e cell line only 72 hours post transduction. Around 500 cells were plated in duplicates for each condition and allowed to incubate for 14 days. Colony counts were performed using the EVOS cell imaging system for both mCherry positive and negative colonies. There were no significant differences in the ratio of mCherry+ colonies to total cells plates for all shRNA transduced cell conditions compared to the shRenilla transduced control. There were significant decreases between the NT and all shRNA transduced conditions (shControls or sh*NCAM1*, Figure 15E Left). Every plate had similar colony sizes, distribution and progenitor cell makeup for all conditions (Figure 15E Right). Overall, *NCAM1* knockdown *in vitro* does not impact cell clonogenic potential.

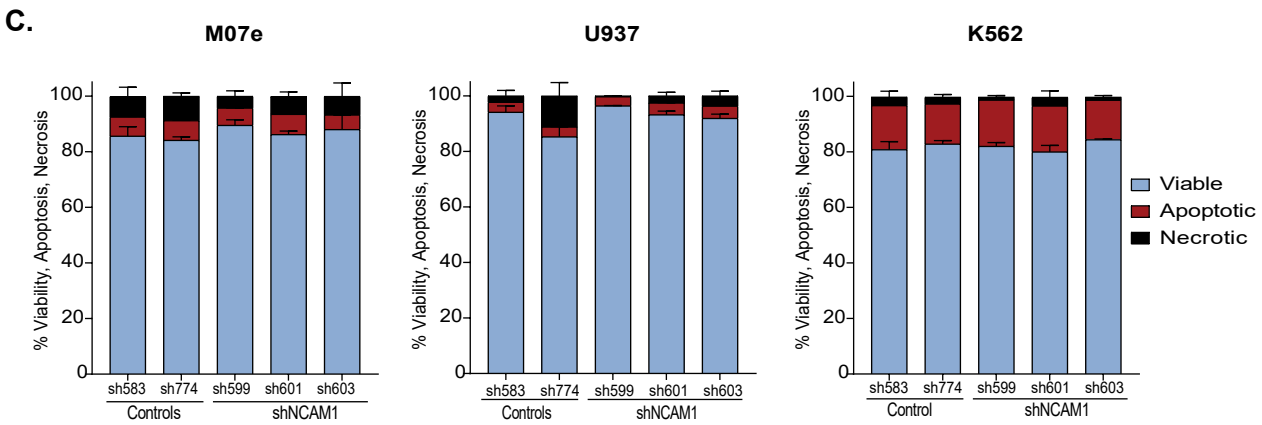
The leukemic cell infiltration, and gene transfer and knockdown efficiencies for functional analyses *in vivo.*

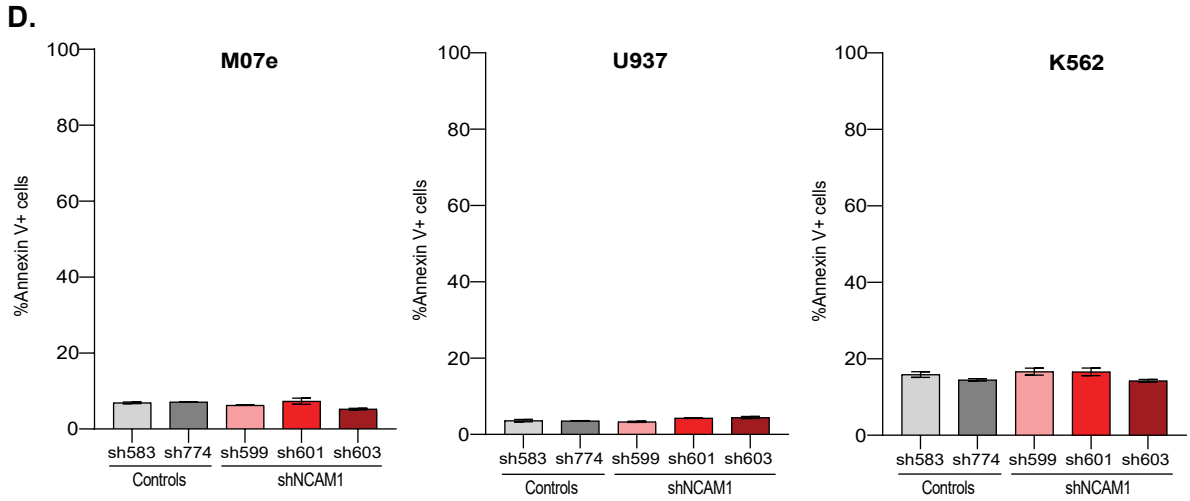
Prior to performing and assessing the functional role of *NCAM1 in vivo*, gene transfer, gene knockdown efficiencies and leukemic cell infiltration were assessed. Primary *CBFA2T3::GLIS2* cells from our mCG2 models were either Non-transduced (N=3) or transduced with the five aforementioned shRNAs (N=6). Around 1 million cells were transduced at an MOI of 15 and 80% of the wells (around 400,000 cells) were transplanted into each mouse (1 mouse = 1 well) (Figure 16A). Not all the mice made it to the endpoint, some unexpectedly died ≤ 5 days post-transplant due to unknown reasons. This included one shRenilla mouse and three sh*NCAM1*-601 mice. Later on, closer to endpoint, two more mice died due to possible femoral dislocations, this included two sh*NCAM1*-603 mice. One sh*NCAM1*-603 mouse was sacrificed 4 days prior to the endpoint because of obvious signs of leukemia development and was included in the functional analysis (Supplemental Table S3).

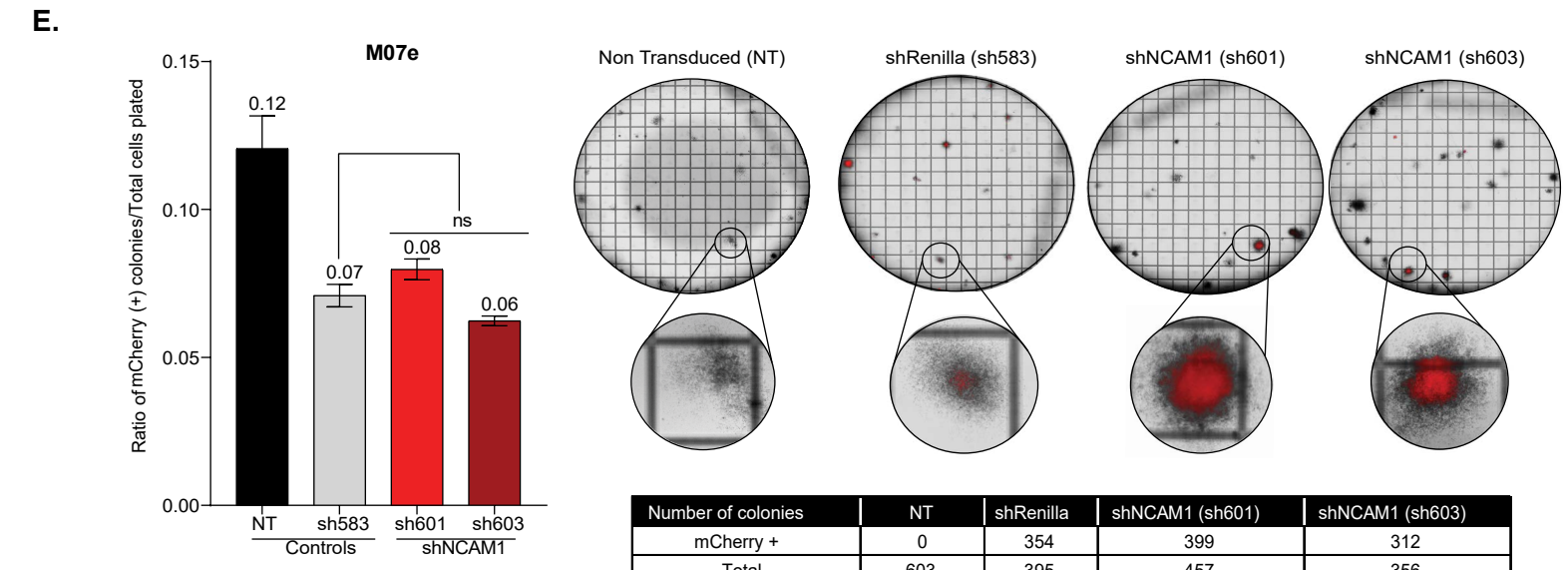
Figure 15











Number of colonies	NT	shRenilla	shNCAM1 (sh601)	shNCAM1 (sh603)
mCherry +	0	354	399	312
Total	603	395	457	356

Figure 15: No impact of *NCAM1* KD *in vitro* on cell proliferation, death, viability, and clonogenic potential in the M07e, U937 and K562 cell lines. A). Total cell counts to assess cell proliferation for the M07e (Left), U937 (Middle) and K562 (Right) cell lines over 24-, 48- and 72-hours. The cells were either Non-transduced (NT) or transduced with shControls or sh*NCAM1*s. Depicted as Mean +/- SEM, N=6/sh. **B).** Flow cytometry plots depicting the gating strategy used for the AnnexinV/DAPI assay to determine viable (blue), early apoptotic (dark red), late apoptotic (dark red) and necrotic (black) cell populations using FlowJo. Viable cells were first gated plotting SSC-A against FSC-A, then mCherry+ cells were selected and then AnnexinV-APC was plotted against DAPI to determine viable and dead cells. **C).** Percentages of viable, apoptotic and necrotic cells as depicted by stacked bar graphs for the M07e (Left), U937 (Middle) and K562 (Right) cell lines 72 hours post transduction. The cells were either Non-transduced (NT) or transduced with shControls or sh*NCAM1*s. Depicted as Mean +/- SEM, N=3/sh. Blue = Viable, Dark Red= early/late apoptotic, Black=necrotic. **D).** Percentages of apoptotic cells (AnnexinV+ cells) for the M07e (Left), U937 (Middle) and K562 (Right) cell lines 72 hours post transduction. The cells were either Non-transduced (NT) or transduced with shControls or sh*NCAM1*s. Depicted as Mean +/- SEM, N=3/sh. **E).** The ratio of mCherry+ colonies to total cells plated for the M07e cell line 72 hours post transduction either Non-transduced (NT) or transduced with shRenilla control and sh*NCAM1* (sh601 and sh603) (Left). Representative images of the plates and colonies for each condition with overlapping mCherry+ expression (Top Right) and a table of mCherry+ colonies and total colonies counted in the clonogenic progenitor cell assay (Bottom Right). Depicted as Mean +/- SEM, N=6/sh. P-values: * <0.05, ** <0.01, *** <0.001, **** <0.0001. SEM, standard error of the mean.

The transduction efficiency (%mCherry+ cells) and *NCAM1* protein surface expression (MFI) was assessed the day after transplant (Day 2), during week 6- and 9- post-transplant, and on the day of sacrifice/endpoint. The gene transfer efficiency (%mCherry+ cells) the day after transplant was above 80% for all shRNA transduced conditions (shRenilla and sh*NCAM1*s) (Figure 16B). There were significant decreases in *NCAM1* protein surface expression (MFI) the day after transplant for both sh*NCAM1* transduced (sh601 and sh603) conditions compared to the shRenilla transduced control condition (25-45% level of knockdown) (Figure 16C).

Figure 16

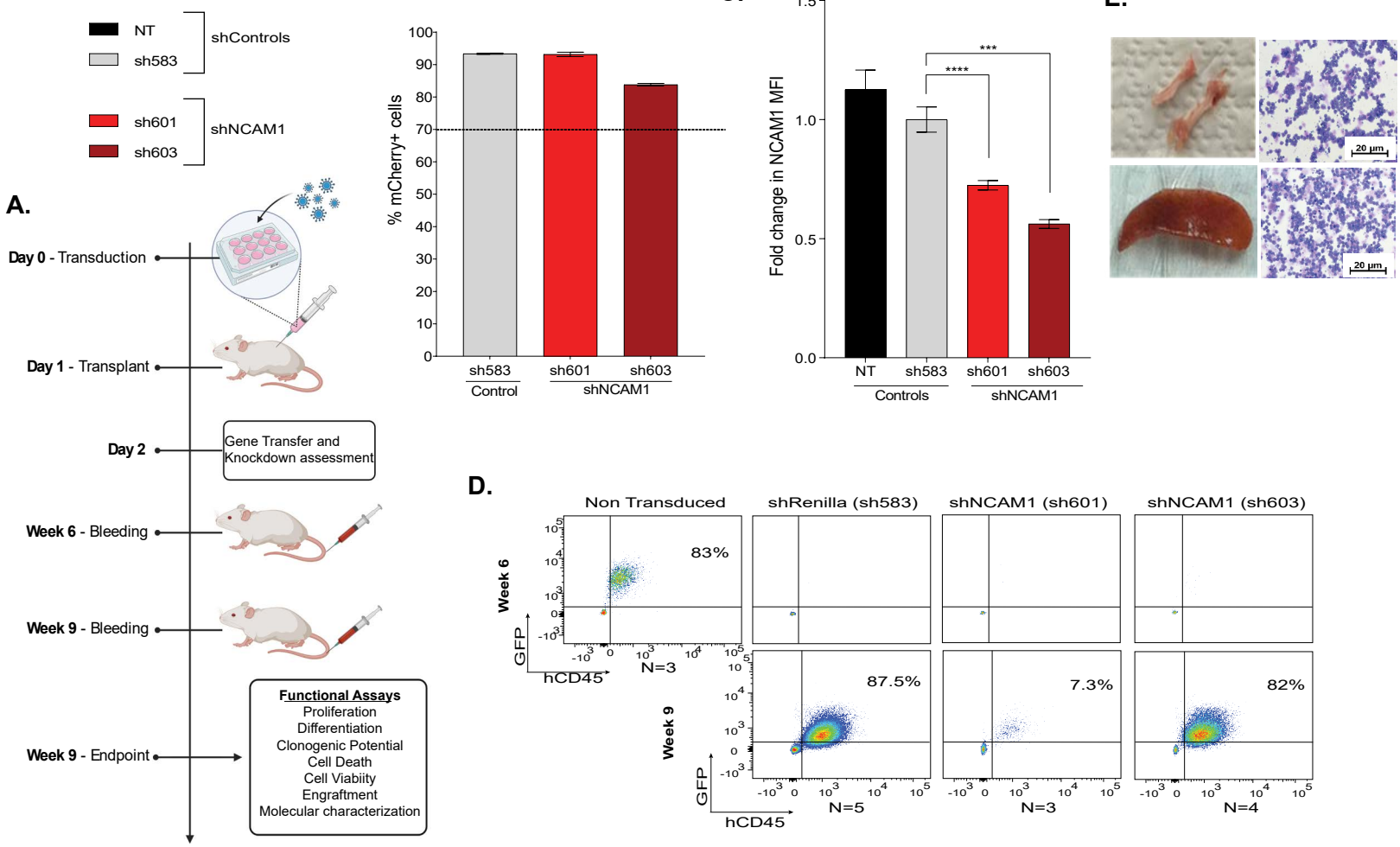

NT
sh583
shControls
sh601
sh603
shNCAM1
A.
Day 0 - Transduction
Day 1 - Transplant
Day 2
Gene Transfer and Knockdown assessment
Week 6 - Bleeding
Week 9 - Bleeding
Week 9 - Endpoint
Functional Assays
Proliferation
Differentiation
Clonogenic Potential
Cell Death
Cell Viabiity
Engraftment
Molecular characterization
B.
% mCherry+ cells
sh583
sh601
sh603
Control
shNCAM1
C.
Fold change in NCAM1 MFI
NT
sh583
sh601
sh603
Controls
shNCAM1
E.
20 µm
20 µm
D.
Non Transduced
shRenilla (sh583)
shNCAM1 (sh601)
shNCAM1 (sh603)
Week 6
Week 9
GFP
hCD45
83%
87.5%
7.3%
82%
N=3
N=5
N=3
N=4
F.
Spleen Weight (mg)
Non-Leukemic
NT
sh583
sh601
sh603
Controls
shNCAM1
G.
% GFP infiltration
73.6
81.7
88.8
90.7
29.1
32.5
86.1
84.0
Bone Marrow
Spleen
Controls
shNCAM1
H.
% mcherry+ cells
Day 2
Blood
Bone Marrow
Spleen
Control
shNCAM1
I.
Fold change in NCAM1 MFI
Day 2
Bone Marrow
Spleen
Controls
shNCAM1
J.
MFI
Non-Leukemic
Non Transduced
shRenilla(583)
shNCAM1(601)
shNCAM1(603)
2568
3518
2377
1856
2681
3201
2083
1712
Spleen
Bone Marrow

Figure 16: Assessment of leukemic cell infiltration, gene transfer, and gene knockdown efficiencies following *NCAM1* KD *in vivo*. A). A summary of the methodologies for *NCAM1* KD *in vivo*. Primary mCG2 leukemia model cells were either Non-transduced (NT) or transduced with an shControl (shRenilla (583)) and shRNA against *NCAM1* (sh*NCAM1*-601 and -603). The cells were transplanted into immunodeficient mice and gene transfer and KD efficiency was assessed Day 2 post transduction. Engraftment and disease progression were monitored using serial blood sampling (week 6 and week 9 post transplant). The mice were sacrificed when moribund (endpoint) following standard necropsy procedures. Several functional assays, to assess different cell fates, were performed on isolated bone marrow and spleen cells. Figure created with BioRender.com. **B).** Percentage gene transfer rates (%mCherry+ cells) 2 days post transduction with indicated shRNA against *NCAM1* in comparison to control vector (shRenilla (583)) cells in primary mCG2 leukemia model cells. Depicted as Mean +/- SEM, N=3 for NT and N=6/sh. **C).** Fold change in *NCAM1* surface protein expression (MFI) normalised to the shRenilla control, quantified by flow cytometry 2 days post transduction with indicated shRNA against *NCAM1* in comparison to control vector (shRenilla (583)) or Non-transduced (NT)) cells in primary mCG2 leukemia model cells. Multiple t-tests performed, depicts Mean +/- SEM, N=3 for NT and N=6/sh. **D).** Images of leukemic mouse femur (Top Left) and spleen (Bottom Left). Images of giemsa stained isolated bone marrow (femur) cells (Top Right) and isolated spleen cells (Bottom Right) from mice models. **E).** Flow cytometry pseudocolor plots depicting blood samples taken week 6 and week 9 post transplant to assess leukemic cell infiltration. The plots depict hCD45 stained cells against GFP+ cells in cells isolated from peripheral blood in CG2 mice that were transplanted with either Non-transduced (NT) cells, shRenilla-control (583) cells or sh*NCAM1* (601 or 603) cells. N=3 for NT, N=5 for sh583, N=3 for sh601 and N=4 for sh603. **F).** Spleen weights (mg) of a healthy (Non-leukemic) (N=1), Non-transduced mice (NT, N=3), shRenilla-control mice (sh583, N=5), and sh*NCAM1* mice (N=3 for sh601 and N=4 for sh603) on the day of endpoint (week 9). Multiple t-tests performed, depicts Mean +/- SEM. **G).** Percentage leukemic cell infiltration (%hCD45/GFP+ cells) on the day of endpoint (week 9) in isolated bone marrow and spleen cells from mice transplanted with Non-transduced (NT) cells, shRenilla-control cells and sh*NCAM1* (601 and 603) cells. Depicted as Mean +/- SEM, N=3 for NT, N=5 for sh583, N=3 for sh601 and N=4 for sh603. **H).** Percentage gene transfer rates (%mCherry+ cells) on the day of endpoint (week 9) in isolated blood, bone marrow and spleen cells from mice transplanted with Non-transduced (NT) cells, shRenilla (583) control cells and sh*NCAM1* (601 and 603) cells. Data from Figure B of Percentage gene transfer rates (%mCherry+ cells) 2 days post transduction were added for easy comparison.

Depicted as Mean +/- SEM, N=3 for NT, N=5 for sh583, N=3 for sh601 and N=4 for sh603. **I).** Fold change in *NCAM1* surface protein expression (MFI) normalised to shRenilla control, quantified by flow cytometry on the day of endpoint (week 9) in isolated bone marrow and spleen cells from mice transplanted with Non-transduced (NT) cells, shRenilla-control cells and sh*NCAM1* (601 and 603) cells. Data from Figure C) of *NCAM1* surface protein expression (MFI) 2 days post transduction were added for easy comparison. Depicted as Mean +/- SEM, N=3 for NT, N=5 for sh583, N=3 for sh601 and N=4 for sh603. **J).** Flow cytometry histograms showing visual shifts in *NCAM1* surface protein expression with MFI depicted on the day of endpoint (week 9) in isolated bone marrow and spleen cells from healthy Non-leukemic mice and mice transplanted with Non-transduced (NT) cells, shRenilla (583) control cells and sh*NCAM1* (601 and 603) cells, N=3 for NT, N=5 for sh583, N=3 for sh601 and N=4 for sh603. MFI, mean fluorescence intensity. SEM, standard error of the mean.

After transplantation of the cells into recipient mice, engraftment occurred over several weeks (total 9 weeks) and serial blood sampling was performed to assess the level of leukemic cell infiltration (%GFP/hCD45+ cells) and to select an endpoint. Blood was sampled at week 6- and 9- post-transplant (Figure 16A). The %GFP/hCD45+ cells or infiltration of leukemic cells was around 83% for NT conditions and between 0-10% for the shRNA transduced conditions during week 6 of post-transplant (Figure 16D Top). Whereas leukemic cell infiltration was above 80% for all shRNA transduced conditions except for the sh601 transduced condition which was around 7% during week 9 post-transplant (Figure 16D Bottom). The endpoint was selected as week 9 due to the infiltration data and mice beginning to look and act leukemic, with palpable enlarged spleens. Mice spleens were weighed and assessed for amount of leukemic cell infiltration (Figure 16F). A healthy, Non-leukemic mouse was used as a control and its spleen weight was found to be under 100mg, whereas all the NT, shRenilla control and sh*NCAM1*-603 transplanted mice had spleen weights above 300mg. The sh*NCAM1*-601 transplanted mice had spleen weights closer to that of the healthy mouse, around 100mg and significantly decreased to the weights of the NT and shRenilla transplanted control mice (Figure 16F). Leukemic cell infiltration (%GFP/hCD45+ viable cells) was assessed on isolated spleen and bone marrow cells (Figure 16E) and found to be between 70-90% for all NT, shRenilla control and sh*NCAM1*-603 transplanted mice. Organ infiltration for the sh*NCAM1*-601 transplanted mice ranged between 20-35% and was significantly lower than the NT and shRenilla transplanted control mice (Figure 16G).

Gene transfer efficiency (%mCherry/GFP+ viable cells) was measured on isolated blood,

spleen and bone marrow cells and found to be above 70% for all NT, shRenilla control and sh*NCAM1*-603 organs. The sh*NCAM1*-601 transplanted mice blood, bone marrow and spleen cells had a gene transfer rate below 50% and were significantly reduced compared to that of the shRenilla conditions (Figure 16H). Gene knockdown efficiency was assessed on isolated bone marrow and spleen cells by measuring *NCAM1* protein surface expression (MFI) with flow cytometry and measuring mRNA expression with qPCR on sorted cells (mCherry/GFP+ only). *NCAM1* protein surface expression was significantly decreased in both the sh*NCAM1* (sh601 and sh603) spleen cells and only the sh*NCAM1*-603 bone marrow cells compared to the shRenilla control conditions (between 25-64% knockdown) (Figure 16I-J). Relative *NCAM1* mRNA expression was significantly decreased for the sh*NCAM1*-603 conditions spleen and bone marrow cells and not for the sh*NCAM1*-601 compared to the shRenilla control conditions (overall knockdown 20-60%) (Supplemental Figure S4).

NCAM1 KD *in vivo* had varying impacts on cell death, viability, and proliferation.

Several functional assays were performed *in vivo* to assess the functional impact of *NCAM1* KD. The functional assays were performed on cells isolated from the bone marrow and spleen and assessed different cell fate outcomes such as cell viability and cell death, and clonogenic potential (Figure 16A).

Cell death and viability *in vivo*

Cell viability and death was assessed using the AnnexinV/DAPI assay and flow cytometry the day after transplant (Day 2) and on isolated spleen and bone marrow cells week 9 post-transplant (endpoint). Using a specific gating strategy on FlowJo the percentage of viable, apoptotic and necrotic cells was determined (Figure 17A). The percentage of viable cells was found to be between 60-80% with around 20% of necrosis occurring the day after transplant (Day 2) for shRenilla control and both sh*NCAM1* transduced conditions (Figure 17B). Whereas the percentage of viable cells week 9 post-transplant (endpoint) was above 90% for all conditions with negligible amounts of necrosis (below 4%) (Figure 17B). The percentage of apoptotic cells the day after transplant (Day 2) was under 5% for shRenilla control and sh*NCAM1*-601 transduced conditions and under 10% for sh*NCAM1*-603 conditions (Figure 17C). The percentage of apoptotic cells week 9 post transplant (endpoint) was lower and below 2% for all conditions (Figure 17C).

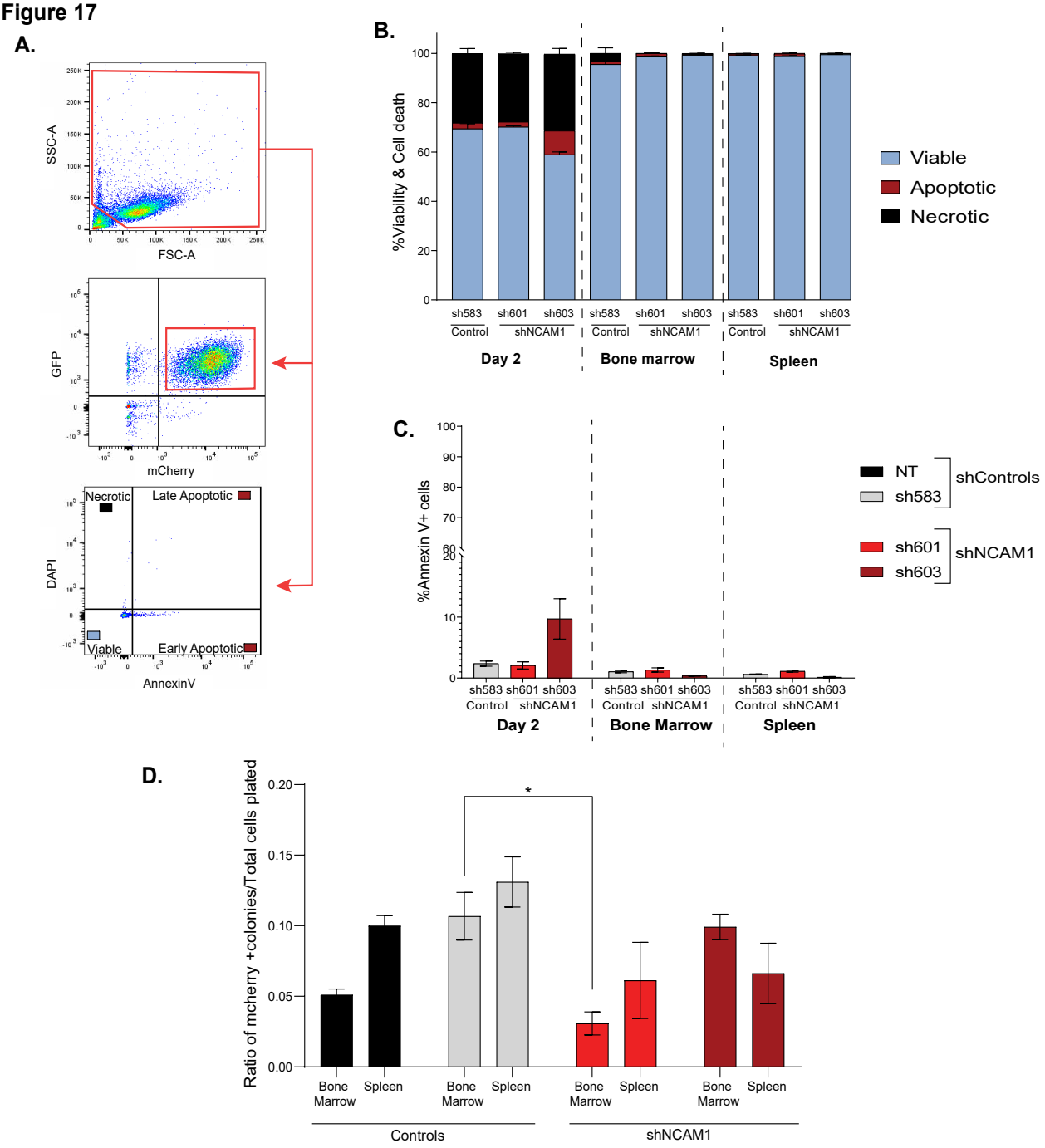
Clonogenic potential *in vivo*

Clonogenic potential was assessed on sorted bone marrow and spleen cells week 9 post

transplant (endpoint). Around 1000 sorted mCherry+/GFP+ cells were plated in duplicates for each condition and allowed to incubate for 14 days. Colony counts were performed for mCherry positive colonies. There was a significant decrease in the ratio of mCherry+ colonies to total cells plated for the bone marrow cells of the sh*NCAM1*-601 transduced condition compared to the shRenilla control bone marrow cells (Figure 17D). The ratios for the spleen cells of both sh*NCAM1* and the bone marrow cells of sh*NCAM1*-603 transduced conditions were not significantly decreased. All colonies looked similar in size, makeup and distribution.

Overall, *NCAM1* knockdown *in vivo* may potentially have an impact on cell fate outcomes but it is difficult to conclude without further testing since the two sh*NCAM1* gave opposing results.

Figure 17



Number of colonies	Organ	NT	shRenilla	shNCAM1 (sh601)	shNCAM1 (sh603)
mCherry+	Spleen	255	131	61	66
	Bone Marrow	603	395	31	99

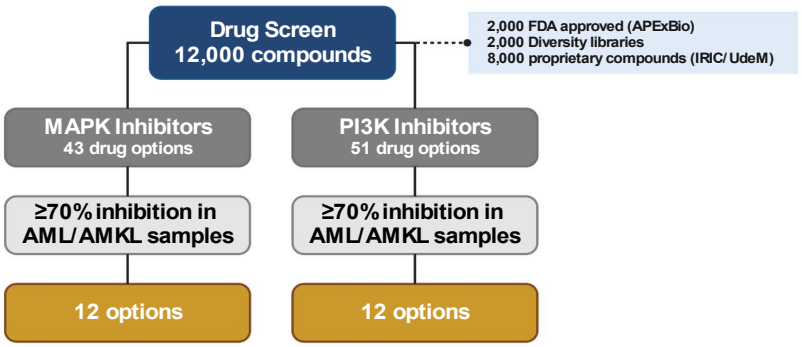
Figure 17: Impact of *NCAM1* KD *in vivo* on cell death, viability, and clonogenic potential. **A).** Flow cytometry plots depicting the gating strategy used to determine viable (blue), early apoptotic (dark red), late apoptotic (dark red) and necrotic (black) primary mCG2 model cell populations using FlowJo. Viable cells were first gated plotting SSC-A against FSC-A, then mCherry+ cells were gated against GFP to select for the double positive cells and finally AnnexinV was plotted against DAPI to determine viable and dead cells. **B).** Percentages of viable, apoptotic and necrotic primary mCG2 model cells as depicted by stacked bar graphs in cells 2 days post transplant and isolated bone marrow and spleen cells on the day of endpoint (week 9). The cells were either transduced with shRenilla controls or sh*NCAM1* (601 and 603). Depicted as Mean +/- SEM, N=3 for NT, N=5 for sh583, N=3 for sh601 and N=4 for sh603. Blue = Viable, Dark Red= early/late apoptotic, Black=necrotic. **C).** Percentages of apoptotic (AnnexinV+ cells) cells of primary mCG2 models 2 days post transduction and cells isolated from the bone marrow and spleen on the day of endpoint (week 9). The cells were either transduced with shRenilla controls or sh*NCAM1* (601 and 603). Depicted as Mean +/- SEM, N=3 for NT, N=5 for sh583, N=3 for sh601 and N=4 for sh603. **D).** The ratio of mCherry+ colonies to total cells plated of cells isolated from the bone marrow and spleen on the day of endpoint (week 9). The cells were either Non-transduced (NT) or transduced with shRenilla controls (583) or sh*NCAM1* (601 and 603) (left). Table of mCherry+ colonies and total colonies counted in a clonogenic progenitor cell assay (Bottom). Depicted as Mean +/- SEM, N=3 for NT, N=5 for sh583, N=3 for sh601 and N=4 for sh603. SEM, standard error of the mean.

CG2 AMKL cells are sensitive to MAPK and PI3K pathway inhibition *in vitro*

To study the role of signaling pathways downstream of *NCAM1* potentially sustaining leukemic cells *in vitro*, the sensitivity of primary leukemic cells to MAPK and PI3K pathway inhibitors combined with KD were tested. Drug candidates were first selected from a wide scale drug screen previously performed in the lab, consisting of 12,000 compounds (2,000 FDA approved APExBio), 2,000 diversity libraries and 8,000 proprietary compounds from IRIC/UDEM) (unpublished results from the Cellot lab). Of these 12,000 compounds, drugs inhibiting components of the MAPK and PI3K pathways were selected (43 drug options for MAPK and 51 drug options for PI3K). Of these 94 options, only drugs that had over 70% inhibition levels in our target cells AMKL/AML of interest were selected, narrowing it down to 12 options per pathway (Figure 18A.). For MAPK pathway inhibitors, Trametinib (MEK1/2 inhibitor), Cobimetinib (selective MEK inhibitor) and AZD8330 (MEK1/2 inhibitor) were selected. For the PI3K pathway, A-443654 (AKT inhibitor), PIK-75 (PI3K inhibitor) and BGT-226 (PI3K/mTOR inhibitor) were selected (Figure 18B).

Figure 18

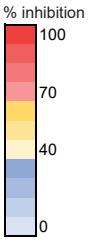
A.



B.

MAPK Pathway Drugs

Drug name	CAS number	Information	GLIS2-AMKL		Normal HSPCs (CB-CD34+)			NUP98 AMKL		PDX
			mCG2-1	mCG2-2	CD34_P1	CD34_P2	CD34_P3	NTF	N5A	AML
Trametinib (GSK1120212)	871700-17-3	MEK1 and MEK2 inhibitor, potent and selective	81	81	85	61	76	56	39	48
Cobimetinib	934660-93-2	Selective MEK inhibitor	83	86	89	75	88	65	34	64
AZD8330	869357-68-6	MEK 1/2 inhibitor	87	87	87	80	87	68	36	72
GDC-0623	1168091-68-6	MEK1 inhibitor, potent and ATP-uncompetitive	84	83	79	31	17	55	44	50
TAK-733	1035555-63-5	MEK allosteric site inhibitor	84	86	84	62	79	62	37	64
PD0325901	391210-10-9	MEK inhibitor	72	77	69	62	62	36	35	44
Pimasertib (AS-703026)	1236699-92-5	MEK1/2 inhibitor	74	76	72	67	58	36	42	44
TAE226 (NVP-TAE226)	761437-28-9	FAK inhibitor,potent and ATP-competitive	71	83	74	49	54	67	33	16
SCH772984 HCl		ERK1/2 inhibitor	76	78	80	67	68	61	34	49
SCH772984 TFA		ERK1/2 inhibitor	74	76	76	64	56	59	33	51
SCH772984	942183-80-4	ERK1 and ERK2 inhibitor	67	74	68	49	69	62	32	48



PI3K Pathway Drugs

Drug name	CAS number	Information	GLIS2-AMKL		Normal HSPCs (CB-CD34+)			NUP98 AMKL		PDX
			mCG2-1	mCG2-2	CD34_P1	CD34_P2	CD34_P3	NTF	N5A	AML
A-443654	552325-16-3	Akt inhibitor,potent and selective	99	99	98	98	99	98	98	98
PIK-75	372196-77-5	Inhibitor of PI3K isoform p110	98	98	98	98	98	98	98	98
CUDC-907	1339928-25-4	Potent PI3K/HDAC inhibitor	99	99	99	99	99	99	98	99
BEZ235 (NVP-BEZ235)	915019-65-7	PI3K/mTOR inhibitor,ATP-competitve	89	89	95	89	72	76	63	75
NVP-BGT226	1245537-68-1	PI3K/mTOR inhibitor,selective and novel	99	99	99	99	99	99	98	99
PF-04691502	1013101-36-4	PI3K/mTOR (FRAP) inhibitor	85	84	89	84	80	74	77	71
GSK2126458	1086062-66-9	PI3K/mTOR inhibitor	89	88	94	94	90	83	93	77
PF-05212384 (PKI-587)	1197160-78-3	PI3K?, PI3K? and mTOR inhibitor	92	89	59	85	40	86	83	79
LY3023414	1386874-06-1	class I PI3K isoforms, mTOR and DNA-PK inhibitor	90	88	93	92	94	80	67	75
ZSTK474	475110-96-4	Potent PI3K inhibitor	73	69	73	50	27	54	64	54
PKI-402	1173204-81-3	PI3K inhibitor,selective, reversible and ATP-competitive	74	70	32	28	10	55	35	56
GDC-0980 (RG7422)	1032754-93-0	Class I PI3K inhibitor for PI3K	78	79	78	73	44	65	70	65

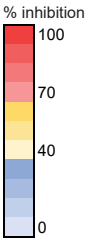
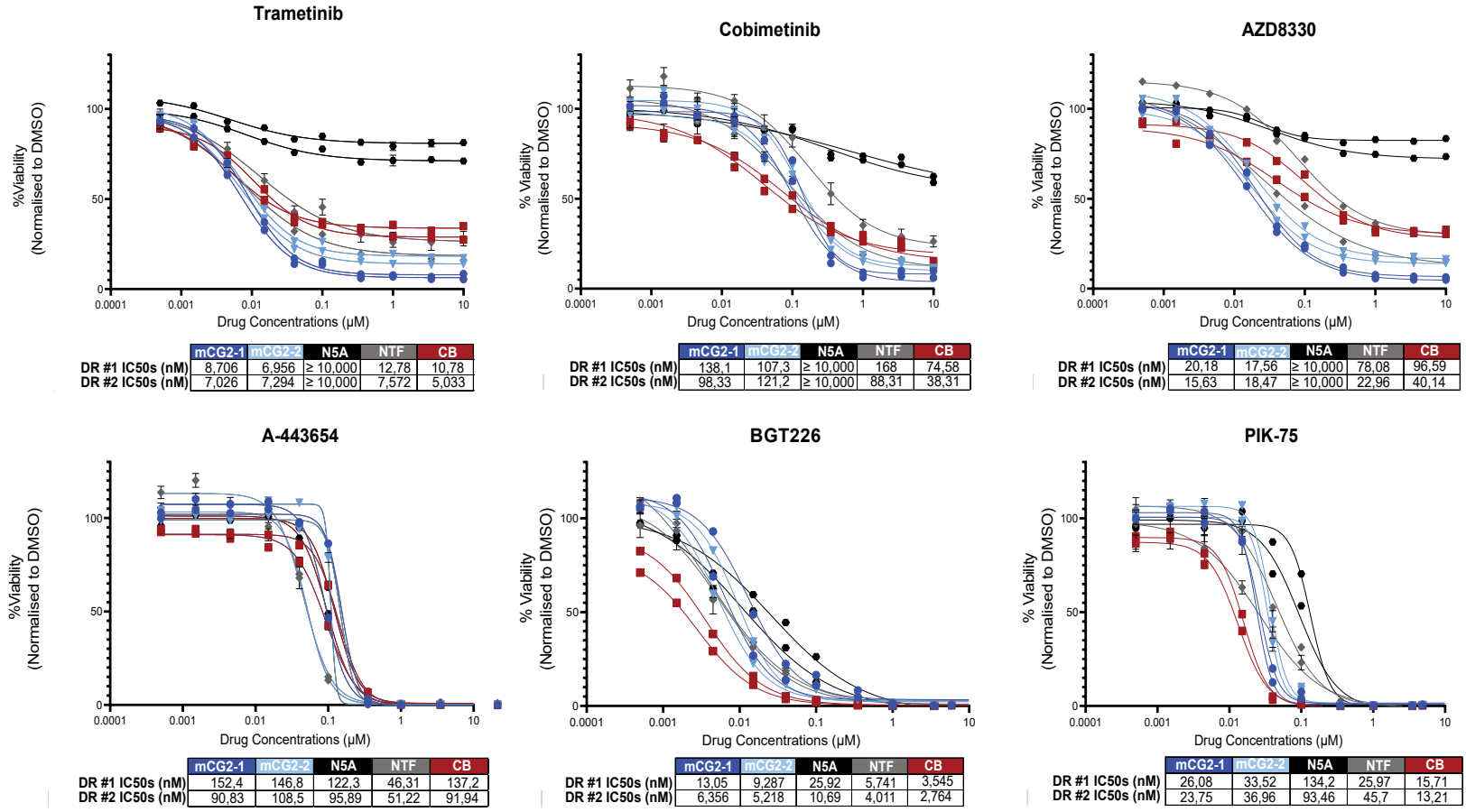


Figure 18: MAPK and PI3K pathway drug selection process and AMKL cell sensitivities. A). Flow chart depicted selection process of pathway inhibitor drugs. Drug candidates were first selected from a wide scale drug screen, consisting of 12,000 compounds (2,000 FDA approved APExBio), 2,000 diversity libraries and 8,000 proprietary compounds from IRIC/UDEM) (unpublished results from the Cellot lab). Of these 12,000 compounds, drugs inhibiting components of the MAPK and PI3K pathways were selected (43 drug options for MAPK and 51 drug options for PI3K). Of these 94 options, only drugs that had over 70% inhibition levels in our target cells AMKL/AML of interest were selected, narrowing it down to 12 options per pathway. Figure created with BioRender.com. **B).** Tables of the 12 MAPK inhibitors (Upper table) and 12 PI3K inhibitors (Bottom table) with over 70% inhibition levels in our target cells. Depicts the Drug name, CAS numbers, brief information about each drug and the percentage inhibition score in *CBFA2T3::GLIS2*-AMKL cells (mCG2-1 and mCG2-2), Normal CB-CD34+ cells (CD34_P1, CD34_P2, CD34_P3), *NUP98* AMKL cells (NTF and N5A) and Patient derived xenografts (PDX) cells (AML-*T21*). Between 70-100% inhibition in red, 40-70% inhibition in yellow and 0-40% inhibition in blue.

Prior to performing experiments and functional assays combining an inhibitor with knock-down of *NCAM1*, *in vitro* sensitivity and cell death and viability of primary leukemic cells was tested. Dose responses (DR) (N=2) and IC50 values were determined after incubation of AMKL (*CBFA2T3::GLIS2*), NTF and N5A xenografts or normal CB-CD34+ cells with Trametinib, Cobimetinib, AZD8330, BGT226, PIK-75, A-443654 or DMSO for 6 days (10 concentrations, 4 replicates), followed by a viability readout with Cell-Titer Glo. CG2 cells, NTF and CB cells were sensitive to MAPK inhibitors (Trametinib, Cobimetinib and AZD8330) with IC50s ranging between 6nm-130nM (Figure 19 Top). N5A cells were not sensitive to MAPK inhibitors with percentage viability remaining high. All cell types were sensitive to the PI3K inhibitors (BGT226, PIK-75 and A-443654) with IC50s ranging between 3-130nM (Figure 19 Bottom).

Figure 19



	mCG2-1	mCG2-2	N5A	NTF	CB
DR #1 IC50s (nM)	8,706	6,956	≥ 10,000	12,78	10,78
DR #2 IC50s (nM)	7,026	7,294	≥ 10,000	7,572	5,033

	mCG2-1	mCG2-2	N5A	NTF	CB
DR #1 IC50s (nM)	138,1	107,3	≥ 10,000	168	74,58
DR #2 IC50s (nM)	98,33	121,2	≥ 10,000	88,31	38,31

	mCG2-1	mCG2-2	N5A	NTF	CB
DR #1 IC50s (nM)	20,18	17,56	≥ 10,000	78,08	96,59
DR #2 IC50s (nM)	15,63	18,47	≥ 10,000	22,96	40,14

	mCG2-1	mCG2-2	N5A	NTF	CB
DR #1 IC50s (nM)	152,4	146,8	122,3	46,31	137,2
DR #2 IC50s (nM)	90,83	108,5	95,89	51,22	91,94

	mCG2-1	mCG2-2	N5A	NTF	CB
DR #1 IC50s (nM)	13,05	9,287	25,92	5,741	3,545
DR #2 IC50s (nM)	6,356	5,218	10,69	4,011	2,764

	mCG2-1	mCG2-2	N5A	NTF	CB
DR #1 IC50s (nM)	26,08	33,52	134,2	25,97	15,71
DR #2 IC50s (nM)	23,75	36,96	93,46	45,7	13,21

Figure 19: AMKL and normal cells are sensitive to MAPK and PI3K inhibitors. Dose response (DR) curves (N=2) for three MAPK pathway inhibitor drugs (Trametinib, Cobimetinib and AZD8330) and three PI3K pathway inhibitor drugs (A-443654, BGT226 and PIK-75). Dose responses for each drug were performed on mCG2-1 (I749) cells, mCG2-2 (I732) cells, *NUP98::KDM5A* (N5A) cells, *NUP98::BPTF* (NTF) cells and normal CB-CD34+ cells. The IC50 values (in nM) are normalised to DMSO and are depicted in tables below each graph.

CG2 AMKL xenograft cells (N=1) were exposed to Trametinib, AZD8330, Cobimetinib, A-443654, BGT226 and PIK-75 for 24-, 48-, and 72- hours at concentrations surrounding the leukemia specific IC50 values. Induction of apoptosis was more significative at 48h with higher levels of cell death compared to DMSO controls, as measured by AnnexinV/DAPI staining (Figure 20, data for 24hr and 72hr not shown). Concentrations that induced between 10-30% apoptosis were selected for Trametinib (50nM) and A-443654 (10nM) for functional assays combining KD with drug in order to see any synergistic, additive and/or combinatorial effects.

Figure 20

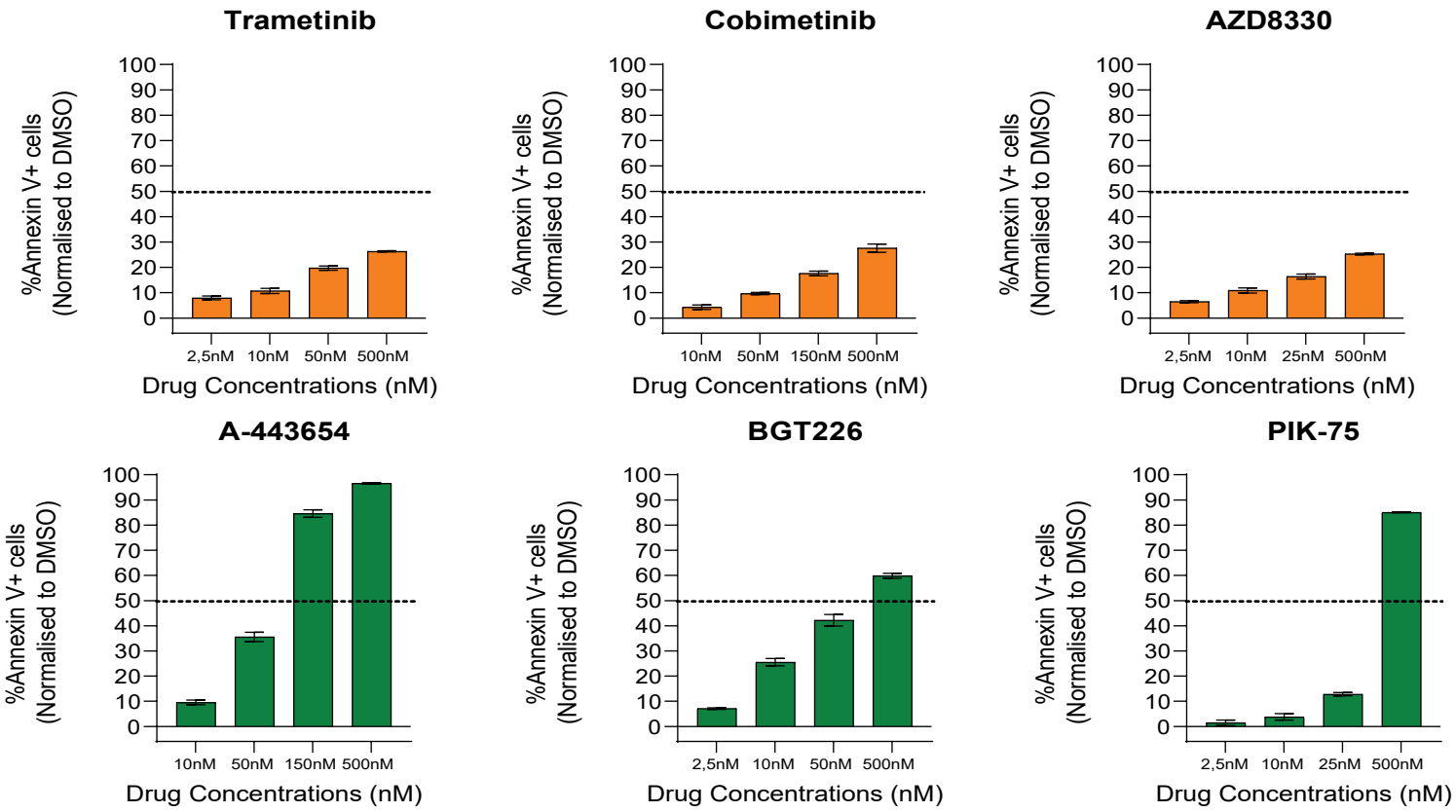


Figure 20: MAPK and PI3K drugs lead to significant impacts on cell death and viability at 48 hours in CG2 AMKL cells. Percentages of apoptotic cells (AnnexinV+ cells) normalised to DMSO in primary mCG2-1 cells 48 hours post transduction after treatment with MAPK and PI3K inhibitors at varying concentrations. N=3/sh.

Sensitivity to MAPK or PI3K pathway inhibition is not impacted by *NCAM1* knockdown in CG2 AMKL cells

To evaluate if *NCAM1* knockdown sensitize CG2 cells to MAPK (Trametinib) or PI3K (A-443654) pathway inhibition, dose responses (N=1) and IC50 values were determined after incu-

bation of sorted AMKL (*CBFA2T3::GLIS2*) (N=2) (mCherry/GFP+) cells either Non-transduced (NT) or transduced with shRenilla and sh774 as controls and three sh*NCAM1* (sh599, 601 and 603) with Trametinib, A-443654 or DMSO for 6 days (10 concentrations, 4 replicates), followed by a viability readout with Cell-Titer Glo. All Non-transduced and transduced conditions were sensitive to both Trametinib and A-443654 with IC50s ranging between 1-6nM and 40-110nM respectively (Figure 21B-C).

***NCAM1* KD combined with pathway inhibitor treatment did not impact cell proliferation, death, and viability in CG2 AMKL.**

Based on the dose response curves and apoptosis data, 50nM Trametinib and 10nM A-443654 were selected to be used for functional assays combining drug with knockdown. To assess the role of *NCAM1* in downstream signaling pathways, *NCAM1* shRNA-mediated knockdown was combined with either a MAPK or PI3K pathway inhibitor. Around 6 million cells from two primary mCG2 model cells (mCG2-1; I749 and mCG2-2; I732) were transduced at an MOI of 15 and 72-hours post transduction the cells were sorted for mCherry/GFP+ cells only and 80,000-250,000 cells (80,000 for sh599) were seeded in triplicates with either 50nM of Trametinib or 10nM of A-443654 for 48 hours. The cells were either Non-transduced (NT) or transduced with shRenilla and sh774 as controls and three sh*NCAM1* (sh599,601 and 603). Several functional assays were performed to assess knockdown efficiency and different cell fates such as cell death and viability, and cell proliferation (Figure 21A).

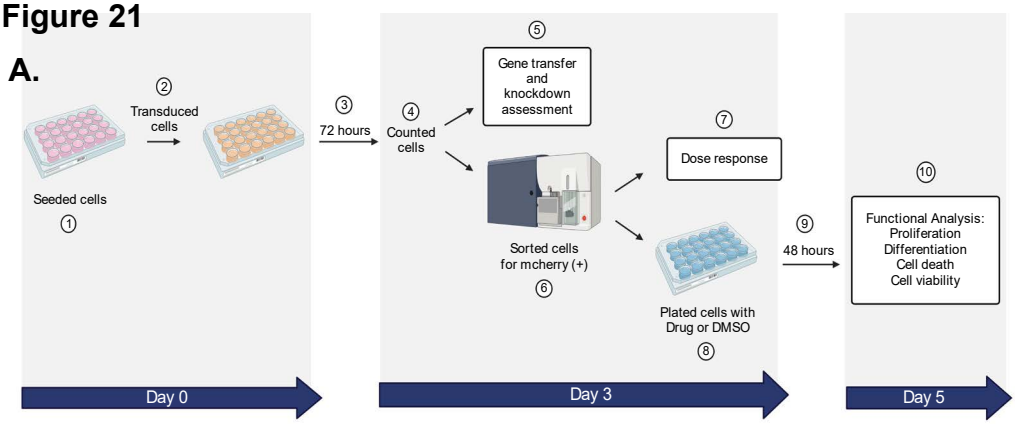
The knockdown efficiency was determined by measuring *NCAM1* protein surface expression (MFI) with flow cytometry and *NCAM1* relative mRNA expression with qPCR (not shown, RNA degraded) on day 3 and day 5. Relative mRNA expression could not be quantified because the mRNA was degraded. This was demonstrated with high CT values (above 25-30) and through bioanalyzer with RINe values below 10. There were significant decreases in *NCAM1* protein expression day 3 and 5 post transduction with Trametinib and A-443654 in all three sh*NCAM1* transduced conditions (sh599, 601 and 603) compared to the shRenilla transduced control as assessed by flow cytometry. There was around 40-75% knockdown of *NCAM1* for all three sh*NCAM1* transduced cell conditions on day 3 in both cell types. The average fold changes in sh*NCAM1* transduced cell MFI for sh599, sh601 and sh603 on day 5 for both cell types were 10.7-fold, 4.8-fold and 6.9-fold for Trametinib (between 50-90% knockdown) respectively and 11.3-fold, 2.5-fold and 7.8-fold for A-443654 respectively (between 50-90% knockdown) (Figure 21 D-E).

There were no significant differences in the percent of AnnexinV+ cells or apoptosis for cells treated with Trametinib or A-443654 for all shRNA conditions. Cell death remained below 10% for A-443654 treated cells and below 50% for Trametinib treated cells in both cell lines for all conditions. There were significant increases in apoptosis between Non-transduced and shRNA transduced conditions for Trametinib and A-443654 treated cells (Figure 21F).

Cell proliferative capacity was measured by performing cell counts using a fixed volume of cells and flow cytometry HTS. Around 80,000-250,000 cells were seeded in triplicates for each cell line and drug condition and cell counts were performed on day 3 and day 5 post-transduction. Compared to DMSO controls, all conditions in both cell types treated with Trametinib were significantly different whereas the conditions treated with A-443654 did not significantly differ from DMSO. The cell number between shControl and sh*NCAM1* conditions did vary significantly and the cell number between Non-transduced conditions and sh*NCAM1* transduced conditions did not vary significantly. *NCAM1* KD did not reduce cell proliferation for all three sh*NCAM1* transduced conditions as the cells either continued to proliferate (sh601/603) or remained stable (sh599). All shControl transduced conditions experienced decreases in cell proliferation after treatment with both drugs in both cell types compared to the sh*NCAM1* transduced conditions (Figure 21G).

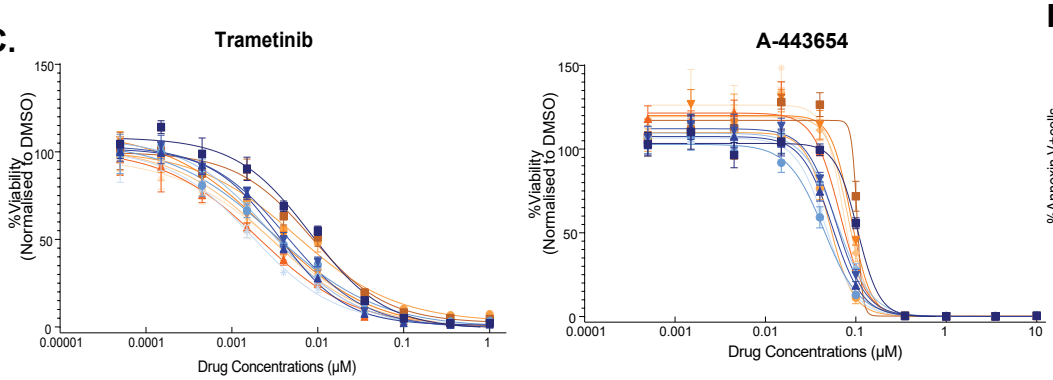
Overall, combining *NCAM1* knockdown (over 50% KD) with inhibitors of MAPK or PI3K did not impact cell viability, cell death and cell proliferation.

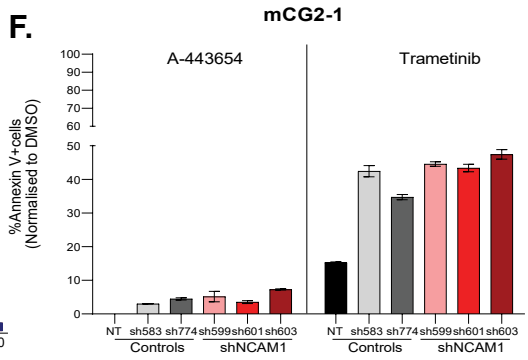
Figure 21

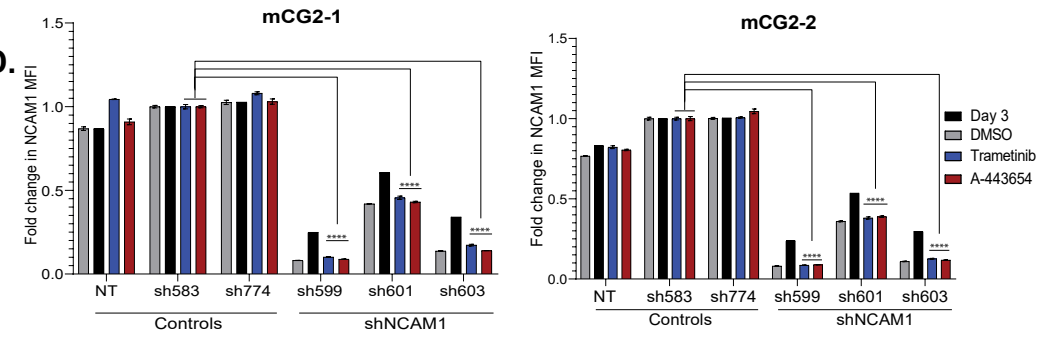


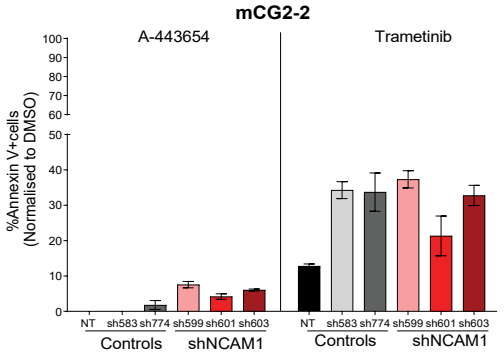
B.

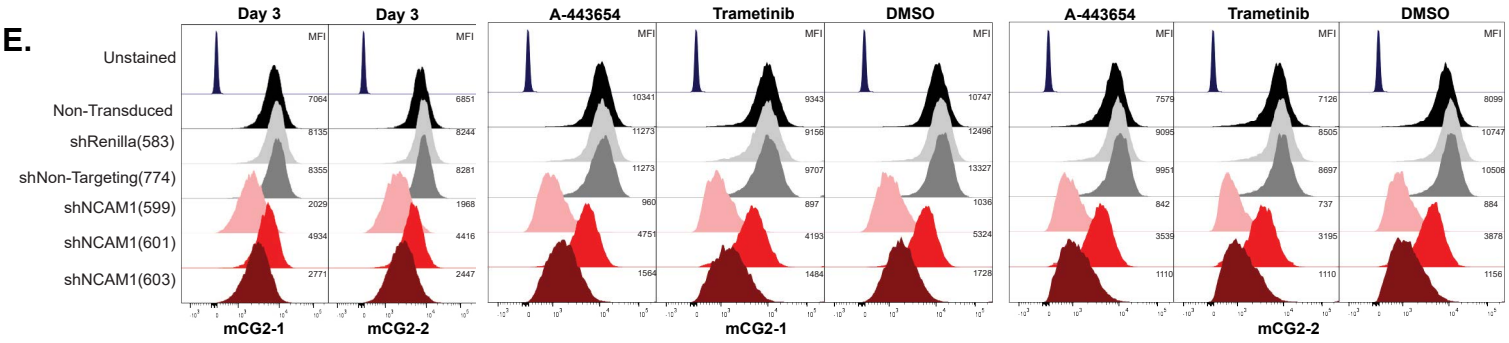
	Sample ID	Cell Type	IC50 concentration (nM) Trametinib	A-443654
	Dose Response 1& 2	mCG2-1	7.867	121.6
	Non-Transduced (NT)	mCG2-1	8.16	104.3
	shRenilla (sh583)	mCG2-1	3.545	54.74
	shNon-Targeting (sh774)	mCG2-1	4.357	60.35
	shNCAM1 (sh599)	mCG2-1	3.077	44.55
	shNCAM1 (sh601)	mCG2-1	2.894	63.03
	shNCAM1 (sh603)	mCG2-1	1.729	46.61
	Dose Response 1& 2	mCG2-2	7.139	127.7
	Non-Transduced (NT)	mCG2-2	8.431	103.2
	shRenilla (sh583)	mCG2-2	2.04	67.03
	shNon-Targeting (sh774)	mCG2-2	2.74	88.04
	shNCAM1 (sh599)	mCG2-2	5.656	51.28
	shNCAM1 (sh601)	mCG2-2	2.745	81.86
	shNCAM1 (sh603)	mCG2-2	3.293	72.19











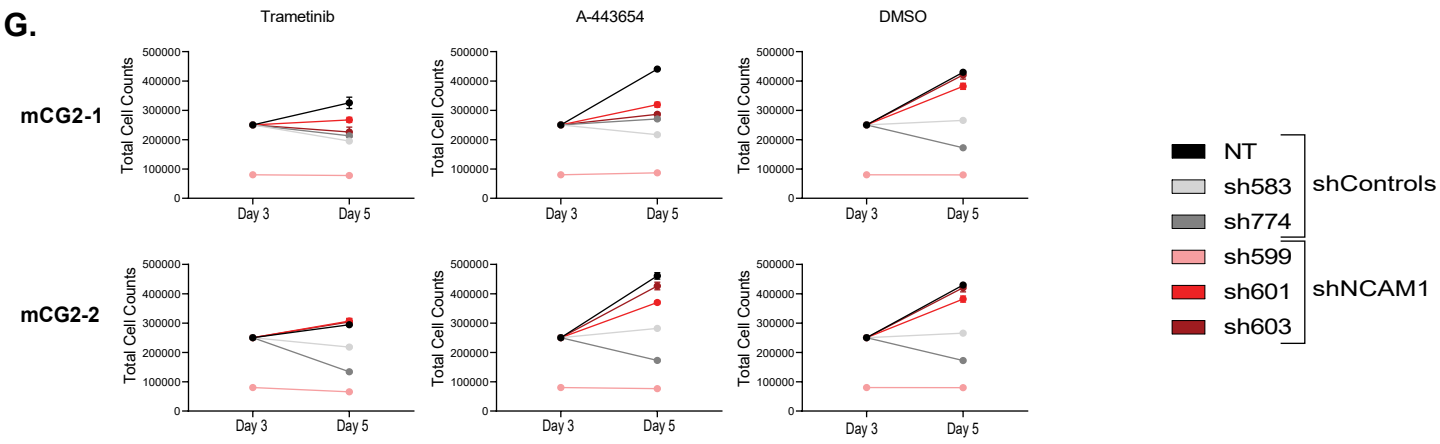


Figure 21: *NCAM1* KD combined with pathway inhibitors does increase AMKL cell sensitivity and does not significantly impact cell proliferation, viability, and death *in vitro*. A). Outline of pharmacological inhibition combined with *NCAM1* KD. On Day 0, primary mCG2 model cells were seeded and were either Non-transduced (NT) or transduced with shControls (shRenilla (583) and shNon-targeting (774)) or sh*NCAM1* (599, 601, 603). After 72 hours, the cells were counted and then 1). assessed for gene transfer efficiency and knockdown efficiency by flow cytometry and qPCR and 2). sorted for mCherry+ cell population. The sorted cells were then either 1). sent for dose responses and 2). seeded in triplicates with either Trametinib (50nM), A-443654 (10nM) or DMSO (≤ 0.1%). The drug treated and DMSO treated cells were incubated for an additional 48 hours and on day 5 several functional assays looking at different cell fate outcomes were performed. Figure created with BioRender.com. **B).** Table legend for Figure C) of colour code, sample ID, cell type, and IC50 values (nM) after treatment with Trametinib and A-443654. **C).** Dose responses (N=1) for the MAPK pathway inhibitor Trametinib and PI3K pathway inhibitor A-443654 combined with shRNA KD conditions normalised to DMSO. Dose responses for each drug were performed on mCG2-1 (I749) cells and mCG2-2 (I732) cells that were either Non-transduced (NT) or transduced with shControls (shRenilla (583) and shNon-targeting (774)) or sh*NCAM1*s (599, 601, and 603) and normalised to DMSO controls. **D).** Fold change in *NCAM1* surface protein expression (MFI) normalised to shRenilla control, quantified by flow cytometry 3 days and 5 days post transduction in mCG2-1 (I749) cells and mCG2-2 (I732) cells that were either Non-transduced (NT) or transduced with shControls (shRenilla (583) and shNon-targeting (774)) or sh*NCAM1*s (599, 601, and 603) and treated with Trametinib (50nM), A-443654 (10nM) or DMSO. Depicted as Mean +/- SEM, N=3/sh. **E).** Flow cytometry histograms showing visual shifts in *NCAM1* surface protein expression with MFI depicted 3 days and 5 days post transduction in mCG2-1 (I749) cells and mCG2-2 (I732) cells that were either Non-transduced (NT) or transduced with shControls (shRenilla (583) and shNon-targeting (774)) or sh*NCAM1*s (599, 601, and 603) and treated with Trametinib (50nM), A-443654 (10nM) or DMSO. N=3/sh. **F).** Percentages of apoptotic cells (AnnexinV+ cells) normalised to DMSO in primary mCG2-1 and mCG2-2 cells 5 days post transduction after treatment with Trametinib (50nM) and A-443654 (10nM). The cells that were either Non-transduced (NT) or transduced with shControls (shRenilla (583) and shNon-targeting (774)) or sh*NCAM1*s (599, 601, and 603). N=3/sh. **G).** Total cell counts to assess cell proliferation in primary mCG2-1 and mCG2-2 cells 3- and 5-days post transduction after treatment with Trametinib (50nM), A-443654 (10nM) or DMSO. The cells that were either Non-transduced (NT) or transduced with shControls (shRenilla (583) and shNon-targeting (774)) or sh*NCAM1*s (599, 601,

and 603). Depicted as Mean +/- SEM, N=6/sh. MFI, mean fluorescence intensity. SEM, standard error of the mean.

Trametinib and A-443654 combination is not synergistic in AMKL cells.

Finally, drug synergism was assessed with Trametinib and A-443654 to determine if a combination therapy would be more efficient to inhibit AMKL compared to single agent. Synergism was assessed after incubation with inhibitors for 6-days (10 concentrations, 4 replicates) in five cell types, two cord blood pools (CB1, CB2), two CG2 cell types (mCG2-1 and mCG2-2), one N5A and one NTF. The average synergy scores were between 0 - -10 for Bliss and ZIP models, below –10 for Loewe model and between –10 - 10 for HSA model for all cell type conditions (Figure 22A). The *CBFA2T3::GLIS2* cell types (mCG2-1; L694 and mCG2-2; N786) did not have average synergy scores above 0 for all four models. Average synergy scores are depicted in 2D and 3D models for the BLISS modelling method (Figure 22B and Supplemental Figure S5).

Overall, *NCAM1* is highly and differentially expressed in CG2 AMKL. *NCAM1* KD (around 50% KD) *in vitro* does not impact cell fate outcomes as expected. Combining *NCAM1* KD with pharmacological pathway inhibition does not sensitize leukemic cells or have any further impacts on cell fate outcomes. Results of *NCAM1* KD *in vivo* are not conclusive since the two sh*NCAM1* conditions had varying results on engraftment and cell fate outcomes.

Figure 22

A.

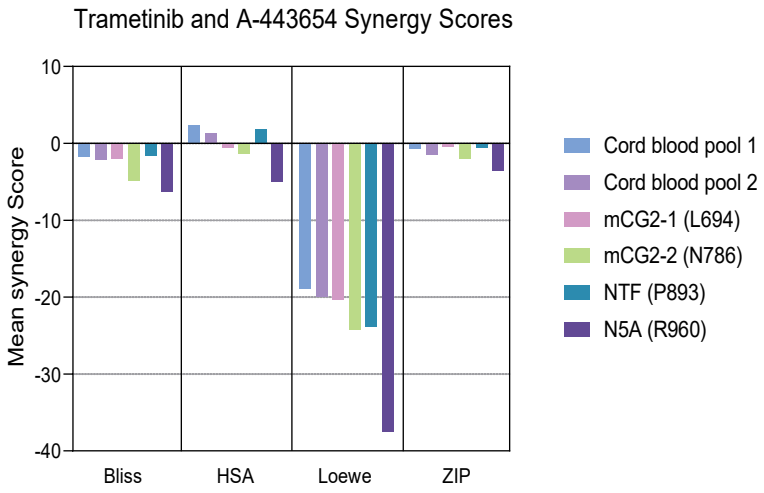


Figure 22 continued.

B.

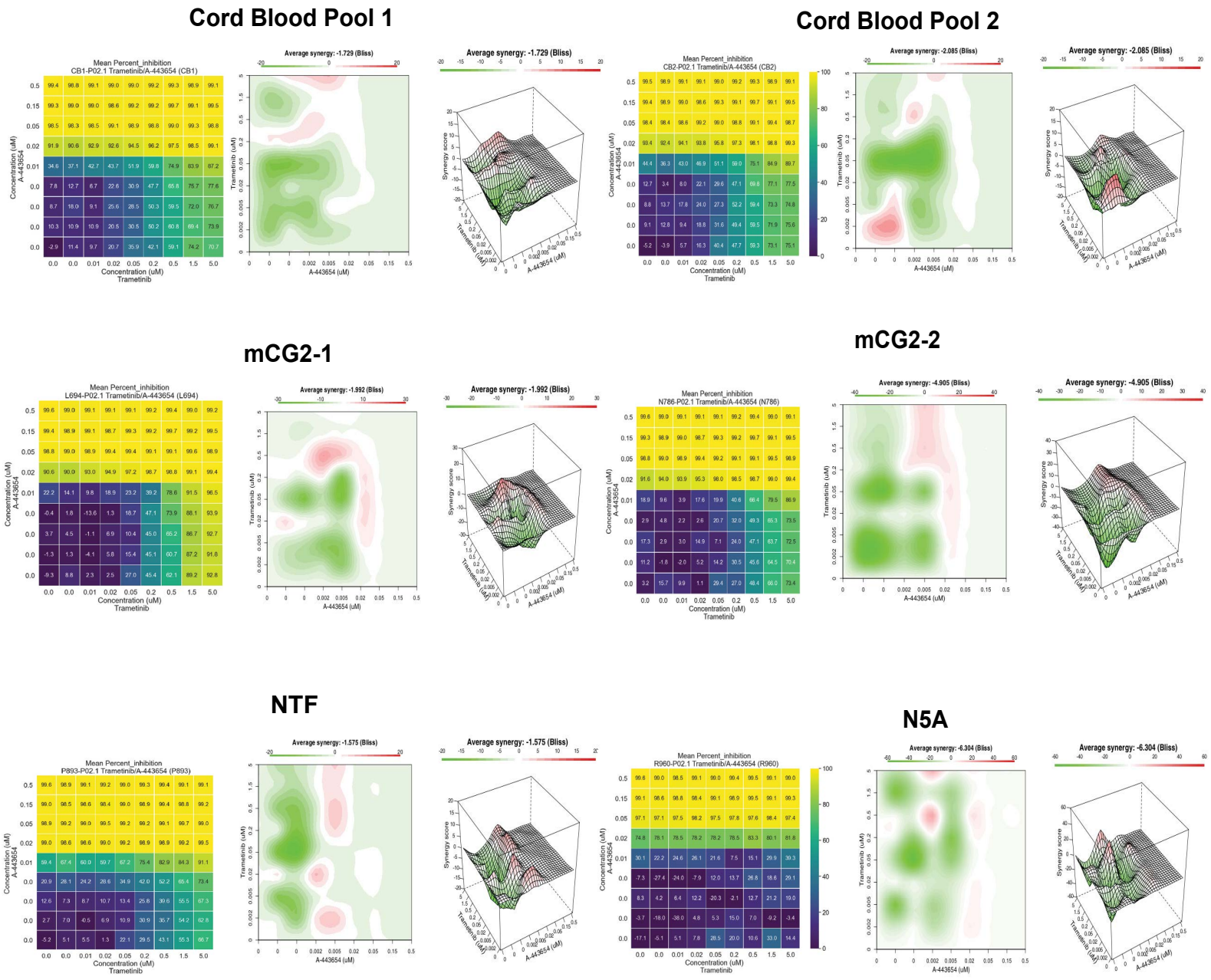


Figure 22: Trametinib and A-443654 combination is not synergistic in CG2 AMKL cells.

A). Mean synergy scores for Trametinib and A-443654 combinatorial matrices performed with two cord blood pools (CB1 and CB2), mCG2-1 (L694), mCG2-2 (N786), NUP98::BPTF (NTF, P893) and *NUP98::KDM5A* (N5A, R960) primary cells (cell titer-Glo, 6 days of culture). Mean synergy scores depicted for the Bliss, HSA, Loewe and ZIP reference models. **B).** Mean percent inhibition depicted in 8x8 matrices of Trametinib and A-443654 drug concentration combinations tested and 2D and 3D models of average synergy scores for the Bliss reference model in two cord blood pools, mCG2-1 (L694), mCG2-2 (N786), *NUP98::BPTF* (NTF, P893) and *NUP98::KDM5A* (N5A, R960) primary cells.

Discussion

NCAM1 is a remarkably highly expressed and differentially expressed in *CBFA2T3::GLIS2* AMKL compared to other subtypes of AML. Its high expression makes it a potentially unique and promising target to use for therapies. shRNA targeting *NCAM1* with high transduction and moderate knockdown efficiencies were successfully generated. Knockdown of *NCAM1* demonstrated that it does not play a significant role in cell fate outcomes such as cell viability, death, proliferation and clonogenic potential *in vitro*. Combining pharmacological inhibition of downstream pathways with *NCAM1* KD *in vitro* does not seem to have major impacts on cell fate outcomes either. It is unclear the role *NCAM1* plays *in vivo* and further studies would be required.

According to transcriptomic and PCA analysis, our CG2 models correlate with CG2 patient samples and cluster differentially from other AML subtypes. This demonstrates that our models phenocopy human disease molecularly and phenotypically and this is important because CG2 AMKL samples are difficult to acquire and are hypocellular due to myelofibrosis [67]. Commercially available AMKL cell lines, such as M07e, are not completely representative of human disease as these cell lines can undergo genetic drift by serial passage and are not fully suited for basing new targeted therapies upon [10]. Having established CG2 models with defined biomarker signatures allows us to have an accurate representation of CG2 leukemia to work with and study the functional role of *NCAM1*.

Expression profiling and immunophenotyping further demonstrate that CG2 leukemic cells follow a megakaryocytic and erythroid lineage composition with remarkably high expression of *NCAM1* compared to other AML subtypes that carry a variety of chimeric gene fusions. AMKL is characterized by having over 30% of immature blasts that express megakaryocytic markers such as CD41, CD42b or CD61 [86, 87]. The expression of these markers is seen in our CG2 models and M07e cell line. Our models and the M07e cell line also express a marker signature considered a consensus myeloid leukemia-associated immunophenotype consisting of CD33+/CD56+/CD34+/- markers [88, 89]. There are many subtypes of AML that carry different gene fusions, biomarkers and have contrasting patient outcomes. When CD56 positive leukemias that harbor different gene fusions, such as *KMT2A/MLL* and *CBFA2T3::GLIS2* were compared, the 5-year event free survival was 39% and 19% respectively [90]. Genetic alterations are not the sole explanation for differences in patient prognosis and survival. Factoring in unique multidimensional phenotypes is important in determining subsets of patients with poor prognosis [91]. The RAM phenotype for example, is a unique phenotype found in hard-to-treat leukemias with poor prognosis and is characterized by $CD56^{BRIGHT}$, $CD38^{DIM/NEG}$, $CD45^{DIM/NEG}$ and $HLA\text{-}DR^{NEG}$ cell populations. The

RAM phenotype is most frequently found in infants with CG2 AMKL [82, 90] and can be observed in our CG2 model and patient samples. One study found that CD56 positive AML without the RAM phenotype had a 3-year event free survival (EFS) of 52%, whereas CD56 positive AML with the RAM phenotype had a 3-year EFS of 16% [91]. These findings demonstrate the importance of not only looking into *NCAM1* as a potential therapeutic target, but also the surrounding molecular landscapes. Being able to have a better understanding of lineage composition and defined genetic biomarker signatures can help with diagnostics and developing targeted therapies.

Transcriptomic, immunophenotyping and immunohistochemical analyses have demonstrated that *NCAM1* is both highly and differentially expressed in CG2 AMKL model spleen and bone marrow cells, patient samples and cell lines and that heterogeneity can occur between AML subtypes. Comparing M07e and U937, two *NCAM1* positive cell lines, there is a significant increase in *NCAM1* MFI in M07e that harbors the CG2 fusion [31] compared to U937 that does not harbor this fusion [92, 93]. One can speculate that it is the binding interaction mentioned earlier between *CBFA2T3* and the *NCAM1* promoter that could lead to the overexpression of *NCAM1* in M07e. Comparing our mCG2 models (mCG2-1, -2, -5 and -6), there are differences in lineage composition and expression profiles. mCG2-6 follows a more megakaryocytic lineage with erythroid projection and heterogeneous *NCAM1* expression, with both *NCAM1*+ and *NCAM1* neg cell populations. The other mCG2 models follow a more homogeneous megakaryocytic lineage with homogeneous *NCAM1* expression. Differences in biomarker expression (e.g. CD34 and CD41) and lineage composition can be seen across CG2 patient samples and other subtypes of AML. These differences amongst our models could be from differences in the cell of origin or the surrounding microenvironment (niche).

NCAM1 is not normally expressed on megakaryocytic cells and is usually found to be expressed on cells of the nervous system or cells within the hematopoietic system belonging to the lymphoid lineage, such as natural killer cells. Our models are generated with cord blood pools, these cells are CD34+ HSPCS and globally do not express *NCAM1*. Differences in *NCAM1* expression or populations of cells expressing *NCAM1* in patient samples could also be a result of the cell of origin or malignant transformation. One can speculate that since *CBFA2T3* may lead to ectopic expression and activity of *GLIS2* in megakaryoblasts in AMKL and since *NCAM1* is a potential direct transcriptional target of *CBFA2T3*, then this malignant transformation may lead to its aberrant expression as well [32].

The bone marrow microenvironment is composed of many different cell types secreting soluble factors and expressing a variety of cell adhesion molecules. Studies have shown that direct cell-cell contact with adhesion models such as N-cadherin, CD44 and *NCAM1* are essential for he-

matopoiesis, niche development and immature HSPC maintenance and development [94]. Leukemic stem cells are known to express specific cell adhesion molecules that provide them with proliferation advantage and avoiding cell death within the BM microenvironment [95, 96]. The molecular landscape of proteins, genes/fusions, and transcription factors vary as well as their interactions with each other. *NCAM1* expression can vary based on which genetic fusion is present and the interactions they may have. The presence or absence of transcription factors can also control *NCAM1* expression. Differences in bone marrow microenvironment composition and interactions that occur between transcription factors, fusions, and other cells and molecules could explain the differences between our mCG2 models. Utilizing tools such as RNA sequencing or transcriptomic analysis to perform more single cell analyses on intracellular and extracellular BM niche components and interactions could be beneficial in understanding what else is expressed along with *NCAM1* and the types of interactions that occur. Performing this type of analysis before and after the knockdown of *NCAM1* can also be beneficial in understanding how the microenvironment and expression profiles change.

There are still many unknowns when it comes to studying the expression and function of *NCAM1* in the hematopoietic system and leukemia. One thing is clear, which is it is aberrantly and highly expressed in a unique subset of AML which makes it an interesting potential biomarker to study for generating more targeted therapies.

The shRNA (shControls & sh*NCAM1*s) used for functional analysis experiments were validated prior to use and demonstrated that an efficient gene transfer (transduction efficiency) and gene knockdown occurred in the M07e cell line for all shRNA conditions. A high gene transfer directly reflects the amount of transduction occurring, if the gene transfer is high (above 70%), this allows us to be confident that the majority of the cells were transduced and that the results of functional analyses directly reflect cells that have experienced *NCAM1* KD. Flow cytometry MFI demonstrated that there is no impact of the viral transduction on *NCAM1* KD since the Non-transduced control and shRNA transduced control cells had similar expression levels of *NCAM1*. Both flow cytometry and ddPCR analysis demonstrated that the cells transduced with the two shRNA controls, shRenilla and shNon-targeting, have similar gene transfer and *NCAM1* expression levels. This means that they are both equally good to use as controls as they both act similarly and both do not target mammalian genes. Throughout these experiments the shRenilla control was used for comparisons for simplicity. Finally, by assessing both the surface protein expression and relative gene expression of *NCAM1*, the knockdown was validated using two different methods and demonstrated an average of 50% knockdown (at both a surface protein and mRNA level). Two of the three sh*NCAM1* transduced conditions (sh599 and

sh603) had a stronger knockdown compared to the third sh*NCAM1* transduced condition (601).

High *NCAM1* expression may play a role in sustaining leukemic cells and promoting their growth and survival. *NCAM1* can play a role in several cell fate outcomes such as cell proliferation, cell death and viability, cell differentiation, clonogenic potential, cell cycle and cellular senescence. To see if *NCAM1* does indeed have a functional role in AMKL, *NCAM1* was inhibited *in vitro* by lentiviral mediated shRNA knockdown, that were previously validated to give on average a 50% knockdown of *NCAM1* in cell lines carrying CG2 and the impact on some of the cellular fates were assessed.

To assess the functional role of *NCAM1 in vitro* on different cell fate outcomes, the gene transfer and level of knockdown was assessed by flow cytometry and qPCR and found to be efficient in the M07e, U937 and K562 cell lines. *NCAM1* knockdown was greater in the U937 cell line compared to the M07e cell line. As expected, the K562 cell line had no changes in *NCAM1* expression. There was no impact of viral transduction in the two *NCAM1* positive cell lines since the Non-transduced control and both shRNA transduced controls had similar gene transfer and KD efficiencies.

Knockdown of *NCAM1* did not produce the expected results of decreasing cell proliferation, cell viability and clonogenic potential and increasing cell death *in vitro*. As expected, cells continued to proliferate in the *NCAM1*neg cell line K562. *NCAM1* KD did not increase cell death and reduce cell viability *in vitro* as expected. For both proliferation assessment and cell death/viability, there were no differences between Non-transduced and shRNA transduced conditions and between shControl and sh*NCAM1* transduced conditions meaning that the stress of viral transduction and *NCAM1* KD both did not impact cell fates. The expectation was to see an increase in cell death and a decrease in viability after *NCAM1* KD. Other studies have shown that *NCAM1* KD *in vitro* increases the amount of cell death and reduces cell viability [62].

Overall, the cells remain viable and there is very little cell death in the form of necrosis and apoptosis occurring after *NCAM1* KD. This could be a result of a weak KD or if there were differences between shControl and sh*NCAM1* transduced conditions, then another explanation for this could be cell cycle arrest or cellular senescence. Different forms of cellular stress can cause a cell to experience cell death or undergo cellular senescence. Cellular senescence is a hallmark of cancer and is a form of cell cycle arrest, where the cell remains viable and metabolically active, and can resist cell death [97]. One study demonstrates that *NCAM1* can play a role in cell cycle and cellular senescence and that partial KD of *NCAM1* led to increased number of cells arrested in G0/G1 and no changes in cell death [62]. Repeating this with a greater level of KD and looking into cellular senescence and cell cycle arrest as a possible reason for the lack of induced apoptosis

after knockdown could be beneficial in understanding the functional role of *NCAM1* in AMKL.

NCAM1 KD did not decrease clonogenic potential *in vitro*. Clonogenic assays assess the ability of single progenitor cells to survive, divide and produce more daughter cells. The expectation was to see a reduction in clonogenic potential after *NCAM1* KD, since *NCAM1* plays a role in promoting proliferation and cell survival. There is a difference between Non-transduced and shRenilla meaning that the stress of viral transduction could lead to a reduction in clonogenic potential. But since there is no difference between shRenilla and the two sh*NCAM1* conditions, then there is no impact of KD alone on clonogenic potential.

Not only the M07e cell line but also the U937 positive control cell line did not produce the expected phenotype after *NCAM1* KD as seen in other studies [62]. This could indicate that the overall level of knockdown was not sufficient. Since M07e overall has a higher *NCAM1* MFI than U937 (Figure 12D), then perhaps the M07e cell line would need to experience a greater KD compared to U937 to see functional impacts. Other studies using a doxycycline inducible shRNA knockdown of *NCAM1*, which resulted in almost full knockdown of *NCAM1*, saw a reduction in cell growth and proliferation and saw an increase in apoptosis and cell cycle arrest [62, 98].

MAPK and PI3K pathways are upregulated or overactivated in AML and *NCAM1* is known to bind receptors that can activate these pathways as well. These pathways regulate cell growth, and promote cell survival, differentiation, and proliferation [19, 20]. To study the effect of *NCAM1* on cell fates, *NCAM1* KD was combined with MAPK or PI3K inhibitors. CG2, NTF and N5A primary cell subtypes were sensitive to PI3K inhibition (similar IC50s for A-443654, BGT226 and PIK-75) but when treated with MAPK inhibitors, only the N5A subtype was resistant. Differences in genetic makeup could explain these differences. Dose responses demonstrated that normal cells (CB-CD34+) had similar IC50s to the other cell types, meaning that both leukemic and normal cells are equally as sensitive to the drugs.

Trametinib and A-443654 were selected to use for the functional assays based on the drug screening and the preliminary viability studies. Drugs that induced around 10-30% cell death after 48 hours were selected so that when KD was combined with these inhibitors a potentially additive effect would be able to be seen. If the percentage of cell death was higher than 30% than this would make any potentially additive effects more difficult to see. A-443654 is a potent pan-AKT inhibitor, preventing phosphorylation activity catalyzed by AKT. It can equally target AKT1, AKT2 and/or AKT3. Akt is the main mediator of pro-survival signals in the PI3k pathway [99]. Trametinib is a potent and selective MAPK pathway inhibitor targeting MEK1/2 preventing its phosphorylation and activation [100].

The expectation of combining a single inhibitor with *NCAM1* KD was to see a greater decrease in proliferation and cell viability and an increase in cell death in mCG2 model primary cells [49, 62, 101, 102]. Trametinib and A-443654 were not used in combination since combinatorial studies demonstrated that there was a lack of synergy in all cell types tested. Combining different PI3K and MAPK inhibitors has been shown to lead to synergistic effects [54, 103]. Perhaps the specific combination of Trametinib and A-443654 is not synergistic and other combinations would need to be tested or other forms of assessing synergy would need to be attempted.

The combination of *NCAM1* KD with Trametinib or A-443654 did not make the CG2 cells more sensitive and did not significantly change the IC50 values compared to the shControl transduced conditions. The IC50s were slightly decreased in all shRNA transduced conditions compared to the NT controls and this could indicate that viral transduction may impact cell sensitivity to drug treatments. The knockdown of *NCAM1* was efficient as assessed by flow cytometry and the combination of *NCAM1* KD with inhibitors did not have significant impacts on cell fate outcomes. The combination of inhibitor with *NCAM1* KD did not reduce cell growth and proliferation. There were no differences between the Non-transduced and sh*NCAM1* transduced conditions which demonstrates that transduction alone did not have an impact on cell proliferation. Cells treated with Trametinib (50nM) did have minimal increases and reductions in cell proliferation compared to cells treated with A-443654 (10nM) and DMSO. This could indicate that Trametinib alone may impact cell proliferation. Since A-443654 had very little differences compared to the DMSO controls this could indicate that A-443654 alone had no impact.

Compared to preliminary tests, cells that underwent *NCAM1* KD and were treated with Trametinib had almost double the amount of cell death occurring, but there still were no significant differences between shControl transduced conditions and sh*NCAM1* transduced conditions, meaning that *NCAM1* KD does not cause this increase in cell death and that viral transduction stress may be the cause. Most studies found that the induction of apoptosis was both dose- and time- dependent for both A-443654 And Trametinib [20, 21], seeing detectable effects between 48 - 96 hours [55]. When PI3K inhibitors targeting AKT or MAPK inhibitors targeting ERK are used in different forms of leukemia, such as CLL and AML, cell death is increased, proliferation and growth are reduced, cells become arrested, and cells are sensitized to chemotherapy drugs [19-21].

There are several explanations for why no impact of combining *NCAM1* KD with drug treatment was observed. There are many pathways found in AML subtypes that sustain leukemic cells and promote their survival (e.g. RTK pathways, JAK/STAT, Hedgehog, etc…) and multiple ways of activating them. *NCAM1* interactions are not the only way in which MAPK

or PI3K pathways can be activated. *NCAM1* interacting with FGFR receptors are only one of many ways in which these pathways can be activated. Both the PI3K and MAPK pathways are complex and interact and regulate each other and share many downstream effectors [19]. Another reason could be because *NCAM1* can act as a reservoir for cytokines and growth factors and provide continuous supply to nearby receptors. Since *NCAM1* was not entirely knocked down or out (around 50%), the remaining *NCAM1*s acting as reservoirs can still provide surface receptors (such as FGFR) with the factors needed to trigger and activate these downstream pathways [62]. AML is a heterogeneous disease with many genetic alterations present. The presence of other mutations, such as the RAS mutation can have an impact on cell sensitivity to MAPK inhibitors, like Trametinib. One study found that a RAS-mutant cell line experienced more apoptosis after exposure to Trametinib compared to a RAS-wildtype cell line (below 10%) [104]. This demonstrates that transcriptomic analysis and understanding what other fusions or mutations are present can help with determining which inhibitors to use in combination with *NCAM1* KD.

Finally, different concentrations and combinations of MAPK/PI3K pathway inhibitor drugs can be used to improve these results. Other studies have shown a synergistic effect on apoptosis induction when combining certain PI3K and MAPK inhibitors in AML [54, 103] and that these effects can be further potentiated with the addition of a chemotherapy drug [54]. One study found that combining Trametinib at 50nM and Ara-C chemotherapy drug led to increased apoptosis and decreased cell growth in a synergistic fashion [62]. Since *NCAM1* KD was combined with a single inhibitor and not two inhibitors, perhaps an increased dosage would be needed if no other drugs are used to see an impact on these cell fates. Altogether, testing different inhibitors, prolonging drug treatment times, testing different concentrations or combining inhibitors with a stronger KD or full knockout of *NCAM1* may be needed to see functional impacts on cell fates.

The role of high *NCAM1* expression *in vivo* has been shown to impact disease progression, leukemic cell engraftment and resistance to chemotherapy drugs. The function was assessed by performing *NCAM1* KD *in vivo* and assessing different cell fate outcomes.

The transplanted cells had high gene transfers (above 80%) and significant knockdown of *NCAM1*. This means that the cells that were transplanted into the mice had efficient gene transfer and knockdown meaning that the majority of the cells transplanted were transduced. On the day of sacrifice (endpoint), the knockdown efficiency was preserved, and mice spleen weights confirmed that splenomegaly occurred and that there was high infiltration of leukemic cells (hCD45+/GFP+). An expectation of *NCAM1* knockdown was to see an engraftment disadvantage and an increased time to leukemia onset, meaning a lower presence or infiltration of the transplanted leukemic cells.

It is possible that this was observed in the sh*NCAM1*-601 transplanted mice, since the percentage of leukemic cell infiltration in the blood and the spleen weights were significantly decreased and more similar to the healthy mouse with less leukemic burden. There was also a decrease in gene transfer and in *NCAM1* KD in the sh*NCAM1*-601 transplanted mice on the day of endpoint compared to the day of transplant. This could mean that engraftment did not occur and that there was a reduction in the number of transduced (hCD45+/GFP+/mCherry+) cells. Since the other group of sh*NCAM1*-603 transplanted mice did not exhibit a similar trend in blood infiltration, spleen weights, gene transfer and *NCAM1* KD, it is difficult to conclude if *NCAM1* KD did have an impact on leukemic cell engraftment or if other circumstances such as initial transplantation issues was the cause. Transplanted cells can be lost *in vitro* prior to transplant or *in vivo* and have issues with homing and engraftment early on. Performing bone marrow aspirations or using luciferase expressing models can help to confirm initial homing, assess leukemic burden early on and track engraftment as it occurs and could help to better understand the early impacts of *NCAM1* KD.

There were no significant differences between the Non-transduced mice and the shRenilla transplanted mice, meaning that the stress of viral transduction did not have an impact on gene transfer, leukemic cell infiltration and KD efficiency. One solution would be to improve the level of knockdown and repeat the *in vivo* studies with a third sh*NCAM1* to eliminate any off-target effects. Other studies demonstrated that *NCAM1* plays a role in sustaining leukemic cells *in vivo* and found that using a doxycycline inducible *NCAM1* knockdown system reduced engraftment and prolonged mice survival [62, 98].

Engraftment is measured as either the ability of an HSC to find a suitable microenvironment for self-renewal or the ability of HSC to retain stemness [105]. In AMKL the bone marrow niche is important in sustaining leukemic stem cells and promoting their growth, differentiation and survival [4, 94]. In other studies the knockdown of other adhesion molecules, such as N-cadherin, led to reduced engraftment in the bone marrow but not the spleen [106]. Perhaps, assessing engraftment in the bone marrow after *NCAM1* KD would shed more light on if engraftment is truly impacted or not. The sh*NCAM1*-601 transplanted mice did not look as visibly leukemic compared to the sh*NCAM1*-603 mice. This could mean that leukemic cell engraftment was delayed and that a survival assay would be required to see if leukemia onset was slower.

Overall, there was no impact of *NCAM1* KD *in vivo* on different cell fate outcomes as assessed by flow cytometry and colony formation assays. Cell viability was not significantly reduced and cell death was not significantly increased, which is the opposite of what was expected. No difference was observed in the level of cell death or viable cells between sh*NCAM1*-601 and

sh*NCAM1*-603 transplanted mice. This is interesting since there were differences in infiltration, gene transfer and level of KD, and one would expect to see a difference in the cell fate outcomes as well. There appears to be a higher amount of cell death in the form of necrosis occurring the day after transplant or 24hrs post transduction compared to the day of endpoint. This could mean that after the initial transduction of cells there is an increase in cellular stress that causes cells to become necrotic instead of apoptotic. This could mean that cells are not given enough time to recover from transduction and are not the best quality when transplanted into the mice. Sorting cells prior to transplant or utilizing a doxycycline-inducible knockdown may reduce the effects of premature cell death and allow for transplantation of healthy cells into the mice.

It is difficult to conclude if *NCAM1* KD impacted clonogenic potential of the xenograft cells because the two sh*NCAM1* conditions had opposing results. As mentioned earlier, the expectation was to see a reduction in clonogenic potential after *NCAM1* KD. The BM cells of sh*NCAM1*-601 transplanted mice had a significant reduction compared to the BM/SP cells of shNCAM-603 transplanted mice. Other studies found that *NCAM1* KO mice cells produced less colonies compared to *NCAM1* WT mice cells [62].

Different *NCAM1* isoforms have differing roles in certain neurological diseases and so perhaps this same concept applies to leukemia. Studies have shown that different isoforms of *NCAM1* can trigger different downstream signaling cascades. The *NCAM1*-140 isoform is responsible for activating pro-proliferation and survival pathways such as the MAPK and PI3K pathways, and anti-apoptotic pathways such as NF-KB in malignant neoplasms [43]. The *NCAM1*-180 isoform has low affinity for surface receptors responsible in the signaling cascades and is unlikely to trigger any. The *NCAM1*-120 isoform is not a transmembrane protein and does not participate in downstream signaling [43, 44]. It has also been noted that the *NCAM1*-140 and -180 isoforms decrease with age whereas the *NCAM1*-120 isoforms increase with age [44]. One can infer that in pediatric leukemia, the main isoforms of *NCAM1* present would be the -140 and -180 types. One study found that the *NCAM1*-180 isoform is not normally highly expressed in leukemia cell lines [107], which could mean that the *NCAM1*-140 isoform is the dominant isoform in leukemia. Determining more specifically which isoforms dominate our CG2 AMKL model and samples can help in designing a more targeted knockdown. The shRNA used in these experiments target regions of *NCAM1* that are common to all isoforms. Sh*NCAM1*-599 targets exon 5, sh*NCAM1*-601 targets exon 4 and sh*NCAM1*-603 targets exon 11. Since the sh*NCAM1* used is more general, a more specific shRNA may be beneficial to directly target the isoform that plays a role in leukemia development and progression and reducing any compensation by other isoforms that may be occurring.

The level of knockdown in these experiments is considered moderate (averaging 50%) and usually is good enough to have an impact but based on the results improving the level of knockdown would be beneficial. Other studies were able to achieve an average of 80% KD using similar siRNA mediated *NCAM1* KD methods [108]. Studies have found that utilizing a straight knockdown of *NCAM1*, such as the shRNA-mediated knockdown used in these experiments, can lead to rapid and almost complete cell death early on, most likely resulting from the cellular stress of viral transduction [62]. In these experiments, when cell death was assessed 24 hours post transduction, there were higher levels of necrosis occurring. When cells are presented with high levels of environmental stress, they can bypass other forms of cell death and undergo necrosis, potentially leading to off target effects. To avoid this, other forms of knocking down *NCAM1* expression or inhibiting it entirely need to be considered such as 1). Performing a full knock out of *NCAM1*, 2). performing a conditional knockdown of *NCAM1*, 3). Targeting upstream and downstream regulators of *NCAM1* or 4). Inhibiting binding interactions and post translational modifications.

Full knockout resulting in over 90% efficient genome editing using CRSIPR/Cas9 based techniques of *NCAM1* have been performed in muscle cells [109] and natural killer cells [33]. But this technique is not well defined in AML. Conditional knockdown such as using a doxycycline-inducible knockdown technique can prevent high levels of premature cell death and can allow recovery of cells before *NCAM1* is knocked down. This has proven an efficient technique with other studies finding that conditional *NCAM1* KD led to almost complete downregulation of *NCAM1* and decreased cell growth and proliferation, increased cell cycle arrest, increased cell death and sensitization to chemotherapy drugs [62, 98]. The average level of knockdown achieved in these experiments was around 50%, which could be considered moderate. Another study with moderate levels of knockdown did not induce cell death and only decreased cell growth and increased the amount of G0/G1 arrested cells [62]. Utilizing an inducible knockdown system would be advantageous in not only increasing the level of knockdown but also allowing to select the specific time of KD, allowing cells to recover post transduction, selecting transduced cells and reducing off target effects such as premature cell death and initial engraftment disadvantages. We planned on using a doxycycline inducible KD approach for *NCAM1*. We have started optimizing this doxycycline-inducible knockdown system both *in vitro* and *in vivo*. Several stages were troubleshooted such as puromycin selection and doxycycline toxicity, the leakiness of the vector and cell viability and gene transfer. *In vivo*, the impact of doxycycline administration on mice dehydration and weight was also assessed (Supplemental Figure S6). Further optimizations are needed.

Targeting upstream regulators of *NCAM1* in neural diseases, multiple myeloma, and AML has proven to help induce apoptosis, reduce proliferation, and downregulate the mRNA expression of *NCAM1* [38, 60, 98] all of which prolong disease progression. Defined upstream regulators of *NCAM1* across different diseases include MEIS1, MEF2, STAT1, MLL-AF9, RUNX1, and SOX4 transcription factors [60-62, 98, 110]. Studies found that these transcription factors bind directly to *NCAM1* promoters and upregulate *NCAM1* expression. Suppression of these transcription factors *in vitro* lead to the downregulation of *NCAM1* expression [60-62, 98, 110], reduced proliferation and increased apoptosis [61, 64]. Transcription factor RUNX1 is frequently mutated in cancers and overexpressed in CD56+ AML and plays an important role in hematopoietic stem cell differentiation, and maturation [110]. Forced overexpression of RUNX1 in CD56 negative cell lines leads to the overexpression of *NCAM1* [61, 64].

Targeting downstream effectors of *NCAM1* is well studied across neural diseases but not well studied in AML. *NCAM1* is known to bind through both homophilic and heterophilic interactions to components of the ECM and can lead to constitutive activation of signaling pathways. The interaction between *NCAM1* and the FGFR was first described in neurons and more recently defined in other malignancies [111]. *NCAM1* is a non-canonical binding partner of FGFR through predominantly heterophilic interactions and can lead to the constitutive activation of this pathway by preventing ubiquitination and degradation. The heterophilic binding of *NCAM1* to FGFR can activate the PI3K-AKT and MAPK-ERK1/2 pathways and lead to their constitutive activation which is thought to be the cause of drug resistance [49, 62, 112, 113]. Pharmaceutical inhibition of FGFR has been attempted with no significant impacts on cell viability and proliferation [62]. As mentioned earlier, *NCAM1* can act as a reservoir at the cell surface and accumulate cytokines, chemokines and growth factors for the constitutive activation of receptors [62]. Another reason for this could be that there are multiple ways in which these downstream pathways can be activated, and *NCAM1* is just one of them. Targeting other pathways involving *NCAM1* may have more effects on cell fate outcomes. A study found that both the p48 RUNX1 isoform and *NCAM1* induce the NF-KB pathway and increase anti-apoptotic effects of BLC2L [61, 64]. Targeting p48 RUNX1 isoforms with siRNA or blocking NF-KB increased the amount of apoptosis in CD56+ AML cell lines [61]. There has been some success with utilizing inhibitory *NCAM1* targeting drugs such as Cimetidine in salivary gland tumors which inhibited *NCAM1* through the NF-KB pathway [114].

NCAM1 is involved in both homophilic and heterophilic binding interactions and preventing these interactions can be another way to study the role of *NCAM1*. *NCAM1* can undergo post translational modifications such as polysialylation. Polysialylation confers a negative charge to

the extracellular domains of *NCAM1* and prevents homophilic interactions and have major impacts on downstream signaling [33, 108]. Heterophilic interactions can be prevented with the use of *NCAM1* targeting antibodies. Some studies have had success with downregulating *NCAM1* expression in neuroblastoma cell lines utilizing *NCAM1* targeting antibodies [63]. Investigating *NCAM1* function with neutralizing antibodies is limited mainly because *NCAM1* antibodies are not widely available and of the ones available not all of them function well in inhibiting *NCAM1* [37].

Conclusion

Overall, this study clearly demonstrated that *NCAM1* is both highly and differentially expressed in *CBFA2T3::GLIS2* AML subtypes and that heterogeneity does occur. Diving deeper into microenvironment composition and dynamics and the cell of origin could improve our understanding of these hard-to-treat leukemias. Functional assessments demonstrated that there was no impact on cell fate outcomes of proliferation, cell death, cell viability, and clonogenic potential after partial *NCAM1* knockdown *in vitro*. Combining *NCAM1* KD with pharmacological inhibitors of potential downstream signaling pathways (MAPK and PI3K) did not have any impact on cell fates either. It is difficult to conclude the functional role of *NCAM1 in vivo* since impacts of *NCAM1* KD varied between the two sh*NCAM1* transduced conditions. Both *in vitro* and *in vivo* studies would need to be repeated with an improved or full knockdown of *NCAM1* and looking into other functional roles such as cell cycle and senescence.

Significance

These experiments will help in determining the expression patterns and functions of *NCAM1*, so that we can better understand the role it plays in AMKL development and progression. Since AML is a very heterogeneous disease, being able to identify unique and specific biomarkers is important for creating targeted and patient specific therapies that could help to improve the dismal survival rates. There is a concensus that the BM microenvironment can play an important role in leukemic stem cell survival and leukemia development and progression. Developing treatments that target microenvironment molecules (e.g. adhesion molecules) can help to improve treatment success. *NCAM1* is an ideal potential therapeutic biomarker to use to guide the development of new therapies since it is both highly and differentially expressed in a specific subset of pediatric AMKL containing the CG2 fusion. Using *NCAM1* as a target for antibody-drug conjugates is currently under clinical trials for several malignancies including neuroblastoma [59]. Even if *NCAM1* does not have a direct functional role in CG2 AMKL, using *NCAM1* to guide the development of new therapies can help to improve the high relapse rates and poor AMKL survival rates.

Acknowledgements

I would like to begin by thanking my supervisor, Dr. Sonia Cellot, for allowing me the opportunity to be a part of her laboratory and for providing encouragement and advice over the last two years. Special thanks to my committee members Dr. François Mercier and Dr. Alexander Gregorieff for their suggestions and guidance and to Dr. Noël Raynal as thesis reviewer. I would next like to thank Mélanie Bilodeau and Louise Laramée for their endless knowledge, advice and support. Thank you for always being there to teach or to answer my questions without hesitation. I am extremely grateful to Sophie Cardin and Verena Gress. I could not have completed my thesis without their help, guidance and unwavering support. I have gained experience and learned many invaluable skills and lessons from them both that I will take with me forever and for that I am truly thankful. Many thanks to members of the Institute of Research in Immunology and Cancer (IRIC): Raphaëlle Lambert and Caroline Paradis from the Genomic core facility for qPCR analysis; Annie Gosselin and Angelique Bellemare-Pelletier from the Flow cytometry core facility for cell sorting; and Simon Mathien and Geneviève Boucher from the High-throughput screening and Bioinformatics core facility for dose response and synergy experiments and analysis, respectively. Thank you to members of CHU Sainte-Justine Research center: Karine Provencher from the OPTILAB pathology laboratory for immunohistochemical staining; Bastien Paré from the Nanopore sequencing platform for assessing RNA quality; Ines Boufaied from the Flow cytometry platform for cell sorting; and the animal care facility for helping to support my experiments. Special thanks to Véronique Lisi, Nehme El-Hachem, Alexandre Rouette, Azer Farah, under the supervision of Dr. Vincent-Philippe Lavallée, for their help with bioinformatic analysis and generation of associated figures. I'd like to recognize the assistance of former lab member Luc Boulianne for his work with the synthetic model generation. Thank you to Furat Fatima, for her kindness and willingness to always lend a helping hand. Thank you to Mathieu Roussy, Nadia Emely Chauca Torres, Julian Matias Freue and Billy Ta for their support and contributions. Many thanks to McGill University, the Pathology department and to Dr. Edith Zorychta and Dr. Hua Ling for their unwavering support. Thanks should also go to the Canadian institutes of health research (CIHR) for supporting my research through funding in the form of the Canada Graduate Scholarship Master's (CGS M). This study was supported by the Canadian Cancer Society Research Institute (CCSRI)/Cole Foundation Impact grant and an operating grant from the Richard and Edith Strauss Foundation (S. Cellot). Finally, I cannot begin to express my thanks to my parents, Helene Chouinard and Brodie Cheetham, my brothers, Kyle Cheetham and Connor Cheetham, to Abhijit Gupta and all my friends who provided endless love and support and for always believing in me every step of the way.

Author's contribution

Emma Rose Cheetham participated in the planning of all experiments along with the other laboratory members involving, Sophie Cardin, Verena Gress, Louise Laramée, and Mélanie Bilodeau. The author did figures, graphs, tables and statistical analyses, with some illustrations made by Mathieu Roussy, Nehme El-Hachem, Véronique Lisi, Alexandre Rouette and Azer Farah. Cord blood isolation was performed by the author and Sophie Cardin, Nadia Emely Chauca Torres, and Furat Fatima. Lentivirus transfection, production and transductions was performed by the author. The author did all shRNA validation and *in vitro NCAM1* KD experiments and functional assays. The author performed all functional experiments of cell staining, cell counting, Annexin V/DAPI assays, and colony formation assays and colony counting. The author performed flow cytometry, FlowJo analysis and all associated figures. Emma Rose Cheetham did RNA isolation and whole RT-PCR, PCR and gels, and ddPCR experiments. The author performed ddPCR and qPCR analyses and all the associated figures. The author did all *in vitro* pharmacological inhibition experiments and functional analyses with aid from Verena Gress and members of the IRIC platform for drug dose responses and synergies. Emma Rose Cheetham performed the sacrifice, bleeding and autopsy of mice together with the help by Sophie Cardin and Verena Gress. The author did all the *in vivo NCAM1* KD functional analyses experiments. Microscopy was entirely achieved by Emma Cheetham from the acquisition to the processing of images. Cell culture performed throughout all experiments was performed by the author. The optimization and development of the inducible doxycycline knockdown system was performed by the author and Sophie Cardin, Verena Gress and Louise Laramée. Initial generation of synthetic and PDX models of AMKL and molecular characterization were performed by many members of the Cellot lab (Sophie Cardin, Mathieu Roussy, Verena Gress, Luc Boulianne, Mélanie Bilodeau).

References:

1. Alberts, B., et al., Renewal by multipotent stem cells: Blood cell formation, in Molecular Biology of the Cell. 4th edition. 2002, Garland Science.
2. Doulatov, S., et al., Hematopoiesis: a human perspective. 2012. 10(2): p. 120-136.
3. Cumano, A. and I. Godin, Ontogeny of the Hematopoietic System. Annual Review of Immunology, 2007. 25(1): p. 745-785.
4. Hoggatt, J., Y. Kfoury, and D.T.J.A.R.o.P.M.o.D. Scadden, Hematopoietic stem cell niche in health and disease. 2016. 11: p. 555-581.
5. Maxie, M.G., Jubb, Kennedy & Palmer's pathology of domestic animals. Volume 3. 2016, Elsevier: St. Louis, Missouri.
6. Lubin, B.H. and W.T.J.P. Shearer, Cord blood banking for potential future transplantation. 2007. 119(1): p. 165-170.
7. Moore, M.A., Hematopoietic stem cells, in Principles of Tissue Engineering. 2014, Elsevier. p. 989-1040.
8. Eaves, C.J.J.B., The Journal of the American Society of Hematology, Hematopoietic stem cells: concepts, definitions, and the new reality. 2015. 125(17): p. 2605-2613.
9. Mayani, H. and P.M. Lansdorp, Biology of Human Umbilical Cord Blood-Derived Hematopoietic Stem/Progenitor Cells. STEM CELLS, 1998. 16(3): p. 153-165.
10. Milan, T., et al., Pediatric leukemia: Moving toward more accurate models. 2019. 74: p. 1-12.
11. Ladikou, E., et al., Acute myeloid leukaemia in its niche: the bone marrow microenvironment in acute myeloid leukaemia. 2020. 22(3): p. 1-9.
12. Lin, X., et al., Global, regional, and national burdens of leukemia from 1990 to 2017: a systematic analysis of the global burden of disease 2017 study. 2021.13 (7): p. 10468.
13. Puumala, S.E., et al., Epidemiology of childhood acute myeloid leukemia. 2013. 60(5): p. 728-733.
14. Mak, T.W., M.E. Saunders, and B.D. Jett, Primer to the immune response. 2013: Newnes.
15. Ci, T., et al., Delivery strategies in treatments of leukemia. 2022. 51(6): p. 2121-2144.
16. Calvo, C., et al., Infant acute myeloid leukemia: A unique clinical and biological entity. 2021. 13(4): p. 777.
17. Schiffer, C. and R.J.H.-F.C.M.t.e.H. Stone, Canada: BC Decker, Morphologic classifica-

tion and clinical and laboratory correlates. 2003.

18. Gruber, T.A. and J.R.J.B. Downing, The Journal of the American Society of Hematology, The biology of pediatric acute megakaryoblastic leukemia.2015.126(8):p. 943-949.
19. Pagano, L., et al., Acute megakaryoblastic leukemia: experience of GIMEMA trials. 2002. 16(9): p. 1622-1626.
20. Athale, U.H., et al., Biology and outcome of childhood acute megakaryoblastic leukemia: a single institution's experience. 2001. 97(12): p. 3727-3732.
21. De Rooij, J.D., et al., Pediatric non–Down syndrome acute megakaryoblastic leukemia is characterized by distinct genomic subsets with varying outcomes. 2017. 49(3): p. 451-456.
22. Inaba, H., et al., Heterogeneous cytogenetic subgroups and outcomes in childhood acute megakaryoblastic leukemia: a retrospective international study. 2015. 126 (13): p. 1575-1584.
23. Smith, J.L., et al., Comprehensive transcriptome profiling of cryptic CBFA2T3-GLIS2 fu sion-positive AML defines novel therapeutic options—a COG and target pediatric AML study. 2018. 132: p. 881.
24. Dang, J., et al., AMKL chimeric transcription factors are potent inducers of leukemia. 2017. 31(10): p. 2228-2234.
25. Masetti, R., et al., CBFA2T3-GLIS2-positive acute myeloid leukaemia. A peculiar paediatric entity. 2019. 184(3): p. 337-347.
26. Masetti, R., et al., The changing scenario of non-Down syndrome acute megakaryoblastic leukemia in children. 2019. 138: p. 132-138.
27. Fischer, M.A., et al., Myeloid translocation gene 16 is required for maintenance of haematopoietic stem cell quiescence. 2012. 31(6): p. 1494-1505.
28. Gruber, T.A., et al., An Inv (16)(p13. 3q24. 3)-encoded CBFA2T3-GLIS2 fusion protein defines an aggressive subtype of pediatric acute megakaryoblastic leukemia. 2012. 22(5): p. 683-697.
29. Leung, A., et al., Uncoupling VEGFA functions in arteriogenesis and hematopoietic stem cell specification. 2013. 24(2): p. 144-158.
30. Lopez, C.K., et al., Pediatric acute megakaryoblastic leukemia: multitasking fusion proteins and oncogenic cooperations. 2017. 3(9): p. 631-642.
31. Thirant, 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): p. 452-465.

32. Thiollier, C., et al., Characterization of novel genomic alterations and therapeutic approaches using acute megakaryoblastic leukemia xenograft models. Journal of Experimental Medicine, 2012. 209(11): p. 2017-2031.

33. Picard, L.K., et al., Human NK cells responses are enhanced by CD56 engagement. European Journal of Immunology, 2022. 52(9): p. 1441-1451.

34. Filiz, S., et al., Localization of neural cell adhesion molecule (N-CAM) immunoreactivity in adult rat tissues. 2002. 77(3): p. 127-135.

35. Fidziańska, A. and A.J.F.N. Kamińska, Neural cell adhesion molecule (N-CAM) as a marker of muscle tissue alternations. Review of the literature and own ob servations. 1995. 33(3): p. 125-128.

36. Gaardsvoll, H., et al., Age-related changes in expression of neural cell adhesion molecule (NCAM) in heart: a comparative study of newborn, adult and aged rats. 1993. 61(1): p. 100-107.

37. Van Acker, H.H., et al., CD56 in the immune system: more than a marker for cytotoxicity? 2017. 8: p. 892.

38. Damgaard, T., et al., Regulation of the CD56 promoter and its association with proliferation, anti-apoptosis and clinical factors in multiple myeloma. 2009. 50(2): p. 236-246.

39. Nguyen, C., et al., Localization of the human NCAM gene to band q23 of chromosome 11: the third gene coding for a cell interaction molecule mapped to the distal portion of the long arm of chromosome 11. 1986. 102(3): p. 711-715.

40. Cunningham, B.A., et al., Neural cell adhesion molecule: structure, immunoglobulin-like domains, cell surface modulation, and alternative RNA splicing. 1987. 236(4803): p. 799-806.

41. Reyes, A.A., et al., At least 27 alternatively spliced forms of the neural cell adhesion mol ecule mRNA are expressed during rat heart development. 1991. 11(3): p. 1654-1661.

42. Horstkorte, R., et al., *NCAM1*. 2012, Springer New York. p. 1183-1187.

43. Gattenlöhner, S., et al., Specific Detection of CD56 (NCAM) Isoforms for the Identifica tion of Aggressive Malignant Neoplasms with Progressive Development. The American Journal of Pathology, 2009. 174(4): p. 1160-1171.

44. Berezin, V., Structure and function of the neural cell adhesion molecule NCAM. Vol. 663.

2009: Springer Science & Business Media.

45. Choi, S., Encyclopedia of signaling molecules. Vol. 337. 2012: Springer New York.

46. Parcerisas, A., et al., The hidden side of ncam family: Ncam2, a key cytoskeleton organi zation molecule regulating multiple neural functions. 2021. 22(18): p. 10021.

47. Nielsen, J., N. Kulahin, and P.S. Walmod, Extracellular Protein Interactions Mediated by the Neural Cell Adhesion Molecule, NCAM: Heterophilic Interactions Between NCAM and Cell Adhesion Molecules, Extracellular Matrix Proteins, and Viruses. 2010, Springer New York. p. 23-53.

48. Walmod, P., et al., Cell adhesion molecules of the immunoglobulin superfamily in the nervous system, in Handbook of neurochemistry and molecular neurobiol ogy: Neural protein metabolism and function. 2007. p. 35-151.

49. Schmid, R.S., et al., NCAM stimulates the Ras-MAPK pathway and CREB phosphoryla tion in neuronal cells. 1999. 38(4): p. 542-558.

50. Beggs, H.E., et al., *NCAM1*40 Interacts with the Focal Adhesion Kinase p125fak and the SRC-related Tyrosine Kinase p59fyn. 1997. 272(13): p. 8310-8319.

51. Doherty, P., F.S.J.M. Walsh, and C. Neuroscience, CAM-FGF receptor interactions: a model for axonal growth. 1996. 8(2-3): p. 99-111.

52. Krushel, L.A., et al., NF-κB activity is induced by neural cell adhesion molecule binding to neurons and astrocytes. 1999. 274(4): p. 2432-2439.

53. Simpson, C.S. and B.J.J.J.o.B.C. Morris, Regulation of neuronal cell adhesion molecule expression by NF-κB. 2000. 275(22): p. 16879-16884.

54. Darici, S., et al., Targeting PI3K/Akt/mTOR in AML: rationale and clinical evidence. 2020. 9(9): p. 2934.

55. Milella, M., et al., Therapeutic targeting of the MEK/MAPK signal transduction module in acute myeloid leukemia. 2001. 108(6): p. 851-859.

56. Ling, Y., et al., Protein kinase inhibitors for acute leukemia. Biomarker Research, 2018. 6(1).

57. Acheson, A., J.L. Sunshine, and U.J.T.J.o.c.b. Rutishauser, NCAM polysialic acid can reg ulate both cell-cell and cell-substrate interactions. 1991. 114(1): p. 143-153.

58. Chuong, C.-M. and G.M.J.J.o.N. Edelman, Alterations in neural cell adhesion molecules during development of different regions of the nervous system. 1984. 4(9): p. 2354-2368.

59. Heinly, B. and C.J.F.i.O. Grant, Cell Adhesion Molecules in Neuroblastoma: Complex

Roles, Therapeutic Potential. 2022. 12.

60. Iqbal, M.S., et al., CD56 expression in human myeloma cells derived from the neurogen ic gene expression: possible role of the SRY-HMG box gene, SOX4. 2010. 91(2): p. 267-275.

61. Gattenloehner, S., et al., Novel RUNX1 isoforms determine the fate of acute myeloid leu kemia cells by controlling CD56 expression. 2007. 110(6): p. 2027-2033.

62. Sasca, D., et al., *NCAM1* (CD56) promotes leukemogenesis and confers drug resistance in AML. Blood, The Journal of the American Society of Hematology, 2019. 133(21): p. 2305-2319.

63. Feng, Y., et al. Differential killing of CD56-expressing cells by drug-conjugated human antibodies targeting membrane-distal and membrane-proximal non-overlapping epitopes. in MAbs. 2016. Taylor & Francis.

64. Zecchini, S., U.J.S. Cavallaro, and F.o.t.N.C.A.M. NCAM, Neural cell adhesion molecule in cancer: expression and mechanisms. 2010: p. 319-333.

65. Yang, D.H., et al., Predictable prognostic factor of CD56 expression in patients with acute myeloid leukemia with t (8: 21) after high dose cytarabine or alloge neic hematopoietic stem cell transplantation. 2007. 82(1): p. 1-5.

66. Raspadori, D., et al., CD56 antigenic expression in acute myeloid leukemia identifies pa tients with poor clinical prognosis. Leukemia, 2001. 15(8): p. 1161-1164.

67. Hahn, A.W., et al., Acute megakaryocytic leukemia: What have we learned. 2016. 30(1): p. 49-53.

68. Cardin, S. and et al., Human models of NUP98-KDM5A megakaryocytic leukemia in mice contribute to uncovering new biomarkers and therapeutic vulnerabili ties. Blood Advances, 2019. 3(21): p. 3307.

69. Boitano, A.E., et al., Aryl Hydrocarbon Receptor Antagonists Promote the Expansion of Human Hematopoietic Stem Cells. Science, 2010. 329(5997): p. 1345-1348.

70. Pabst, C., et al., Identification of small molecules that support human leukemia stem cell activity ex vivo. 2014. 11(4): p. 436-442.

71. Fares, I., et al., Pyrimidoindole derivatives are agonists of human hematopoietic stem cell self-renewal. 2014. 345(6203): p. 1509-1512.

72. Fellmann, C. and et al., An Optimized microRNA Backbone for Effective Single-Copy RNAi. Cell Reports, 2013. 5(6): p. 1704.

73. Han, H., RNA Interference to Knock Down Gene Expression. 2018, Springer New York.

p. 293-302.

74. Adan, A., et al., Flow cytometry: basic principles and applications. 2017. 37(2): p. 163-176.

75. Reutelingsperger, C.P.M. and W.L. Van Heerde, Annexin V, the regulator of phosphati dylserine-catalyzed inflammation and coagulation during apoptosis. Cellular and Molecular Life Sciences (CMLS), 1997. 53(6): p. 527-532.

76. Wallberg, F., T. Tenev, and P. Meier, Analysis of Apoptosis and Necroptosis by Fluores cence-Activated Cell Sorting. Cold Spring Harbor Protocols, 2016. 2016(4): p. pdb.prot087387.

77. Taylor, S.C., G. Laperriere, and H. Germain, Droplet Digital PCR versus qPCR for gene expression analysis with low abundant targets: from variable nonsense to publica tion quality data. Scientific Reports, 2017. 7(1).

78. Campomenosi, P., et al., A comparison between quantitative PCR and droplet digital PCR technologies for circulating microRNA quantification in human lung cancer. BMC Biotechnology, 2016. 16(1).

79. Zhang, H., et al., A Strategy to Model Nonmonotonic Dose-Response Curve and Estimate IC50. PLoS ONE, 2013. 8(8): p. e69301.

80. Ianevski, A., A.K. Giri, and T.J.N.a.r. Aittokallio, SynergyFinder 2.0: visual analytics of multi-drug combination synergies. 2020. 48(W1): p. W488-W493.

81. Wooten, D.J. and R. Albert, synergy: a Python library for calculating, analyzing and visu alizing drug combination synergy. Bioinformatics, 2021. 37(10): p. 1473-1474.

82. Eidenschink Brodersen, L., et al., A recurrent immunophenotype at diagnosis independent ly identifies high-risk pediatric acute myeloid leukemia: a report from Children's Oncology Group. 2016. 30(10): p. 2077-2080.

83. Vodyanik, M.A., J.A. Thomson, and I.I.J.B. Slukvin, Leukosialin (CD43) defines hemato poietic progenitors in human embryonic stem cell differentiation cultures. 2006. 108(6): p. 2095-2105.

84. Tuffaha, M.S.A., H. Guski, and G. Kristiansen, Markers and Immunoprofile of Myeloid Neoplasm. 2018, Springer International Publishing. p. 181-184.

85. Van Galen, P., et al., Single-Cell RNA-Seq Reveals AML Hierarchies Relevant to Disease Progression and Immunity. Cell, 2019. 176(6): p. 1265-1281.e24.

86. Le, Q., et al., CBFA2T3-GLIS2 model of pediatric acute megakaryoblastic leukemia iden tifies FOLR1 as a CAR T cell target. 2022.

87. Paredes-Aguilera, R., et al., Biology, clinical, and hematologic features of acute mega karyoblastic leukemia in children. 2003. 73(2): p. 71-80.
88. Feller, N., et al., Defining consensus leukemia-associated immunophenotypes for detec tion of minimal residual disease in acute myeloid leukemia in a multicenter setting. 2013. 3(8): p. e129-e129.
89. Gadgeel, M., et al., Aberrant myelomonocytic CD56 expression in Down syndrome is fre quent and not associated with leukemogenesis. 2021. 100(7): p. 1695-1700.
90. Pardo, L.M., et al., Deciphering the significance of CD56 expression in pediatric acute myeloid leukemia: a report from the Children's Oncology Group. 2020. 98(1): p. 52-56.
91. Brodersen, L.E., et al., A Recurrent Immunophenotype at Diagnosis Independently Identi fies High Risk Pediatric AML: A Report from the Children's Oncology Group Tri al AAML0531. 2015. 126(23): p. 560.
92. Strefford, J., et al., The characterisation of the lymphoma cell line U937, using compara tive genomic hybridisation and multi-plex FISH. 2001. 94(1-2): p. 9-14.
93. MacKinnon, R.N., et al., Detailed molecular cytogenetic characterisation of the myeloid cell line U937 reveals the fate of homologous chromosomes and shows that centromere capture is a feature of genome instability. 2020. 13(1): p. 1-14.
94. Stopp, S., et al., Expression of the melanoma cell adhesion molecule in human mesenchy mal stromal cells regulates proliferation, differentiation, and maintenance of hematopoietic stem and progenitor cells. 2013. 98(4): p. 505.
95. Sofia, C.A., S. Daniela, and D. Sérgio, Bone marrow malignancies as paradigms of dys functional cell adhesion mechanisms.
96. Roversi, F.M., et al., Hematopoietic Cell Kinase (HCK) Is a Player of the Crosstalk Be tween Hematopoietic Cells and Bone Marrow Niche Through CXCL12/CXCR4 Axis. 2021. 9: p. 634044.
97. Kumari, R., P.J.F.i.c. Jat, and d. biology, Mechanisms of cellular senescence: cell cycle arrest and senescence associated secretory phenotype. 2021. 9: p. 645593.
98. Sasca, D., et al., Targeting Aberrant Ncam (neural cell adhesion molecule; CD56) Expres sion in Acute Myeloid Leukemia. 2015. 126(23): p. 311.
99. Luo, Y., et al., Potent and selective inhibitors of Akt kinases slow the progress of tumors *in vivo*. 2005. 4(6): p. 977-986.
100. Gilmartin, A.G., et al., GSK1120212 (JTP-74057) is an inhibitor of MEK activity and ac

tivation with favorable pharmacokinetic properties for sustained *in vivo* pathway inhibition. 2011. 17(5): p. 989-1000.

101. Paratcha, G., F. Ledda, and C.F.J.C. Ibáñez, The neural cell adhesion molecule NCAM is an alternative signaling receptor for GDNF family ligands. 2003. 113(7): p. 867-879.

102. Ditlevsen, D.K., et al., The role of phosphatidylinositol 3-kinase in neural cell adhesion molecule-mediated neuronal differentiation and survival. 2003. 84(3): p. 546-556.

103. Sandhöfer, N., et al., Dual PI3K/mTOR inhibition shows antileukemic activity in MLL-re arranged acute myeloid leukemia. 2015. 29(4): p. 828-838.

104. Kerstjens, M., et al., MEK inhibition is a promising therapeutic strategy for MLL-rear ranged infant acute lymphoblastic leukemia patients carrying RAS muta tions. 2017. 8(9): p. 14835.

105. Grenier, J.M., et al., Adhesion Molecules Involved in Stem Cell Niche Retention During Normal Haematopoiesis and in Acute Myeloid Leukaemia. 2021. 12.

106. Hosokawa, K., et al., Knockdown of N-cadherin suppresses the long-term engraftment of hematopoietic stem cells. 2010. 116(4): p. 554-563.

107. Vander Borght, A., et al., The 180 splice variant of NCAM—containing exon 18—is spe cifically expressed in small cell lung cancer cells. 2018. 7(3): p. 376.

108. Frese, C.K., et al., Quantitative map of proteome dynamics during neuronal differentia tion. 2017. 18(6): p. 1527-1542.

109. Stadelmann, C., et al., mRNA-mediated delivery of gene editing tools to human primary muscle stem cells. Molecular Therapy - Nucleic Acids, 2022. 28: p. 47-57.

110. Okuda, T., et al., RUNX1/AML1: a central player in hematopoiesis.2001.74(3):p.252-257.

111. Williams, E.J., et al., Activation of the FGF receptor underlies neurite outgrowth stimu lated by L1, N-CAM, and N-cadherin. 1994. 13(3): p. 583-594.

112. Francavilla, C., et al., The binding of NCAM to FGFR1 induces a specific cellular re sponse mediated by receptor trafficking. 2009. 187(7): p. 1101-1116.

113. Takase, N., et al., NCAM-and FGF-2-mediated FGFR1 signaling in the tumor microen vironment of esophageal cancer regulates the survival and migration of tu mor-associated macrophages and cancer cells. 2016. 380(1): p. 47-58.

114. Fukuda, M., K. Kusama, and H. Sakashita, Cimetidine inhibits salivary gland tumor cell adhesion to neural cells and induces apoptosis by blocking NCAM expres sion. BMC Cancer, 2008. 8(1): p. 376.

Supplemental Methods

RNA sequencing and bioinformatic analysis

Library preparation, sequencing, and data processing were performed at the Institute for Research in Immunology and Cancer's Genomics Platform (IRIC, Montréal, Canada). Libraries were prepared using KAPA Stranded mRNA-Seq Kit and sequencing was performed on the Illumina Nextseq500. Gene expression were estimated directly from STAR mapping as read count values as well as computed using RSEM version 1.2.28 [1] in order to obtain transcript level expression and Fragments Per Kilobase of transcript per Million mapped fragments (FPKM), using a stranded RNA protocol. Sequences were also analyzed using FusionCatcher [2] in order to identify potential gene fusions

Differential expression analysis

All differential expression analysis (DEG) was performed using FPKM values, due to data availability, with the limma pipeline and empirical Bayes moderated t-statistics [3]. An offset of 1 was added to the FPKM values prior to log2 transformation and lowly expressed genes were removed. Both validation cohort (St. Jude, first setting) [4] and institutional cohort (CHUSJ, second setting) were analyzed separately. In the first setting, gene expression profiling from all AMKL samples with the CG2 fusion transcript were contrasted to all other AMKL samples. In the second setting, all AMKL samples with the CG2 fusion transcript were contrasted to AMKL samples involving *NUP98::KDM5A* fusion gene and human CB-CD34+ cells. A gene was considered significantly regulated if its (absolute log) fold-change is equal or higher than 1 and corresponding adjusted (False Discovery Rate or FDR) p-value <0.05.

Visualization

The two gene expression datasets (institutional and validation cohorts) were combined and quantile-normalized prior to use in a scaled heatmap. Heatmaps, and principal component analysis (PCA) were visualized using ggplot2 and custom scripts in R programming language.

Single cell RNA sequencing

Live single cell suspensions were processed on the Chromium Controller according to manufacturer's protocol aiming at a capture of 8,000 cells per sample. Sequencing

libraries were generated as per manufacturer's instructions and sequenced on a NovaSeq S6000 aiming at 20,000 reads per cell. Fastq files were aligned to the genome and reads counted using CellRanger V3. Count matrices were analysed using Seurat V4. Droplets expressing less than 200 genes were considered empty droplets and discarded. From the remaining cells, we removed those whose percentage of mitochondrial reads was more than 2 standard deviations above the mean mitochondrial percentage of a sample's cells. Reads were log-normalized with a scaling factor of 10,000, principal component analysis was performed using the 2,000 most variable genes and data was embedded in two dimensions using Uniform Manifold Approximation and Projection computed on the first 30 principal components.

We further removed, from each sample, clusters of empty droplets that had low RNA counts and did not express genes associated with a specific cell type. After filtering, PCA, and two-dimensions embedding was repeated. Cell type classification was performed using Seurat's Categorical data transfer method using the Human Cell Atlas dataset [5]. We obtained this annotated dataset using Bioconductor's Experiment Hub web resource with the HCA Data 1.10.0 R package. We restricted the analysis to a maximum of 3,000 cells per cell type using random subsampling to avoid the overrepresentation of abundant cell types. AML single-cell data was downloaded from Gene Expression Omnibus (accession GSE116256) [6] and processed identically to the AMKL samples, including cell type classification

Supplemental References

1. Li, B. and C.N.J.B.b. Dewey, RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. 2011. 12(1): p. 1-16.
2. Nicorici, D., et al., FusionCatcher–a tool for finding somatic fusion genes in paired-end RNA-sequencing data. 2014: p. 011650.
3. Ritchie, M.E., et al., limma powers differential expression analyses for RNA-sequencing and microarray studies. 2015. 43(7): p. e47-e47.
4. De Rooij, J.D., et al., Pediatric non–Down syndrome acute megakaryoblastic leukemia is characterized by distinct genomic subsets with varying outcomes. 2017. 49(3): p. 451-456.
5. Hay, S.B., et al., The Human Cell Atlas bone marrow single-cell interactive web portal. 2018. 68: p. 51-61.
6. Van Galen, P., et al., Single-Cell RNA-Seq Reveals AML Hierarchies Relevant to Disease Progression and Immunity. Cell, 2019. 176(6): p. 1265-1281.e24.

Supplemental Table S1: Lentiviral vector shRNA target sequences

Abbreviation	Name	Target sequence (5' --> 3')	Source
sh583	shRenilla	AGGATTATAATGCTTATCTA	Renilla reniformis luciferase, Gift from Guy Sauvageau
sh774	ShNon-Targeting	AGCGCGATAGCGCTAATAATTA	Leuconostoc or Actinomyces, Modified Sigma Non-Mammalian shRNA in pLKO.1-CMV-TurboGFP
sh599	shNCAM1	TCCGATTCATAGTCCTGTCCAA	Fellman, mRNA Sequence ID: ref NM_001242607.1
sh601	shNCAM1	TTCAGAAGCTCATGTTCAAGAA	Fellman, mRNA Sequence ID: ref NM_001242607.1
sh603	shNCAM1	GTCCATGTACCTTGAAGTGCAA	Fellman, mRNA Sequence ID: ref NM_001242607.2

Supplemental Table S2: List of antibodies.

Target protein	Application	Conjugation	Company	Catalogue Number	Dilution used
CD56/*NCAM1*	Flow cytometry	BV605	BD Biosciences	742659	3.0:100
CD56/*NCAM1*	Flow cytometry	BV605	BioLegend	362538	1.0:100
CD56/*NCAM1*	Immunohistochemistry	Unconjugated	Agilent	IR62861-2	/
hCD45/*PTPRC*	Flow cytometry	APC-CY7	BioLegend	304014	3.5:100
CD34	Flow cytometry	BV421	BD Biosciences	562577	3.0:100
CD43/*SPN*	Flow cytometry	APC	BD Biosciences	560198	2.5:100
CD44	Flow cytometry	BUV737	BD Biosciences	564941	2.0:100
CD71/*FTRC*	Flow cytometry	BUV395	BD Biosciences	743308	3.0:100
CD151	Flow cytometry	BV650	BD Biosciences	742925	2.5:100
CD41/*ITGA2B*	Flow cytometry	APC	BioLegend	303710	3.0:100
CD61/*ITGB3*	Flow cytometry	PerCP-CY5.5	BD Biosciences	564173	3.5:100
CD19	Flow cytometry	BV421	BioLegend	302234	3.0:100
CD3	Flow cytometry	PE-CY7	BD Pharmingen	557851	2.0:100
CD11b	Flow cytometry	PE	BD Biosciences	347557	2.0:100
CD117/c-KIT	Flow cytometry	APC	BD Biosciences	341096	3.5:100
CD33	Flow cytometry	PE-CY7	BD Biosciences	333946	2.0:100
CD15/*FUT4*	Flow cytometry	FITC	BD Biosciences	347423	2.0:100
HLA-DR	Flow cytometry	PE	BD Biosciences	347367	2.0:100
CD13	Flow cytometry	APC-CY7	BioLegend	301710	3.5:100
Annexin V	Flow cytometry	APC	BioLegend	640941	0.5:100

Supplemental Table S3: Mice cause of death.

Group	Mouse ID	Cell Dose	Survival (days)	Spleen weight (mg)	Cause of death
NT #1	Q-002	150ul, 80% of well	50	364	High leukemic cell infiltration, obvious signs of leukemia
NT #2	Q-004	150ul, 80% of well	50	360	High leukemic cell infiltration, obvious signs of leukemia
NT #3	Q-005	150ul, 80% of well	50	383	High leukemic cell infiltration, obvious signs of leukemia
sh583 #1	Q-003	150ul, 80% of well	68	345	Day of endpoint, high leukemic cell infiltration, obvious signs of leukemia
sh583 #2	Q-008	150ul, 80% of well	68	425	Day of endpoint, high leukemic cell infiltration, obvious signs of leukemia
sh583 #3	Q-009	150ul, 80% of well	68	456	Day of endpoint, high leukemic cell infiltration, obvious signs of leukemia
sh583 #4	Q-100	150ul, 80% of well	68	416	Day of endpoint, high leukemic cell infiltration, obvious signs of leukemia
sh583 #5	Q-101	150ul, 80% of well	68	308	Day of endpoint, high leukemic cell infiltration, obvious signs of leukemia
sh583 #6	Q-104	150ul, 80% of well	≤ 5	N/A	Unknown causes
sh601 #1	Q-102	150ul, 80% of well	≤ 5	N/A	Unknown causes
sh601 #2	Q-103	150ul, 80% of well	≤ 5	N/A	Unknown causes
sh601 #3	Q-152	150ul, 80% of well	≤ 5	N/A	Unknown causes
sh601 #4	Q-105	150ul, 80% of well	68	100	Day of endpoint, high leukemic cell infiltration, obvious signs of leukemia
sh601 #5	Q-106	150ul, 80% of well	68	112	Day of endpoint, high leukemic cell infiltration, obvious signs of leukemia
sh601 #6	Q-107	150ul, 80% of well	68	84	Day of endpoint, high leukemic cell infiltration, obvious signs of leukemia
sh603 #1	Q-109	150ul, 80% of well	62	308	Possible femoral dislocation
sh603 #2	Q-110	150ul, 80% of well	68	373	Day of endpoint, high leukemic cell infiltration, obvious signs of leukemia
sh603 #3	Q-111	150ul, 80% of well	62	253	Possible femoral dislocation
sh603 #4	Q-188	150ul, 80% of well	64	398	Obvious signs of leukemia
sh603 #5	Q-189	150ul, 80% of well	68	470	Day of endpoint, high leukemic cell infiltration, obvious signs of leukemia
sh603 #6	Q-190	150ul, 80% of well	68	468	Day of endpoint, high leukemic cell infiltration, obvious signs of leukemia

N/A, not available

Figure S1

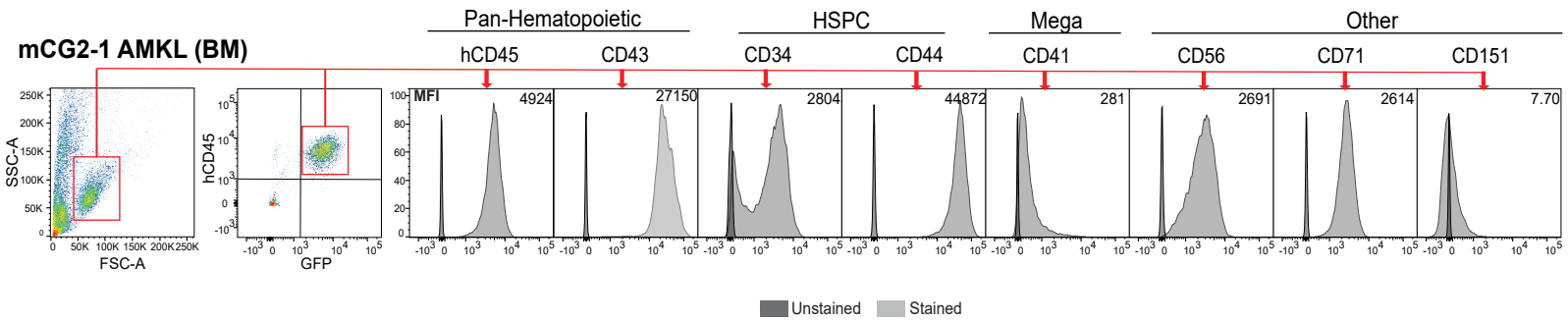


Figure S1: *CBFA2T3::GLIS2* AMKL model bone marrow cells express megakaryocytic and erythroid markers with high *NCAM1* expression at the cell surface. Flow cytometry pseudo color dot plots depicting gating strategy and flow cytometry plots depicting mean fluorescence intensity (MFI) of selected hematopoietic lineage markers (Pan-hematopoietic, HSPC, Myeloid, megakaryocytic, and other markers) in the mCG2-1 (I749) model isolated bone marrow cells (See Figure 11 for analysis of spleen cells). FSC-A plotted against SSC-A to gate for viable cells, then GFP plotted against hCD45 to gate for double positive human leukemic cells. Dark grey plots are unstained, light gray plots are stained. HSPC, hematopoietic stem and progenitor cells; BM, bone marrow.

Figure S2

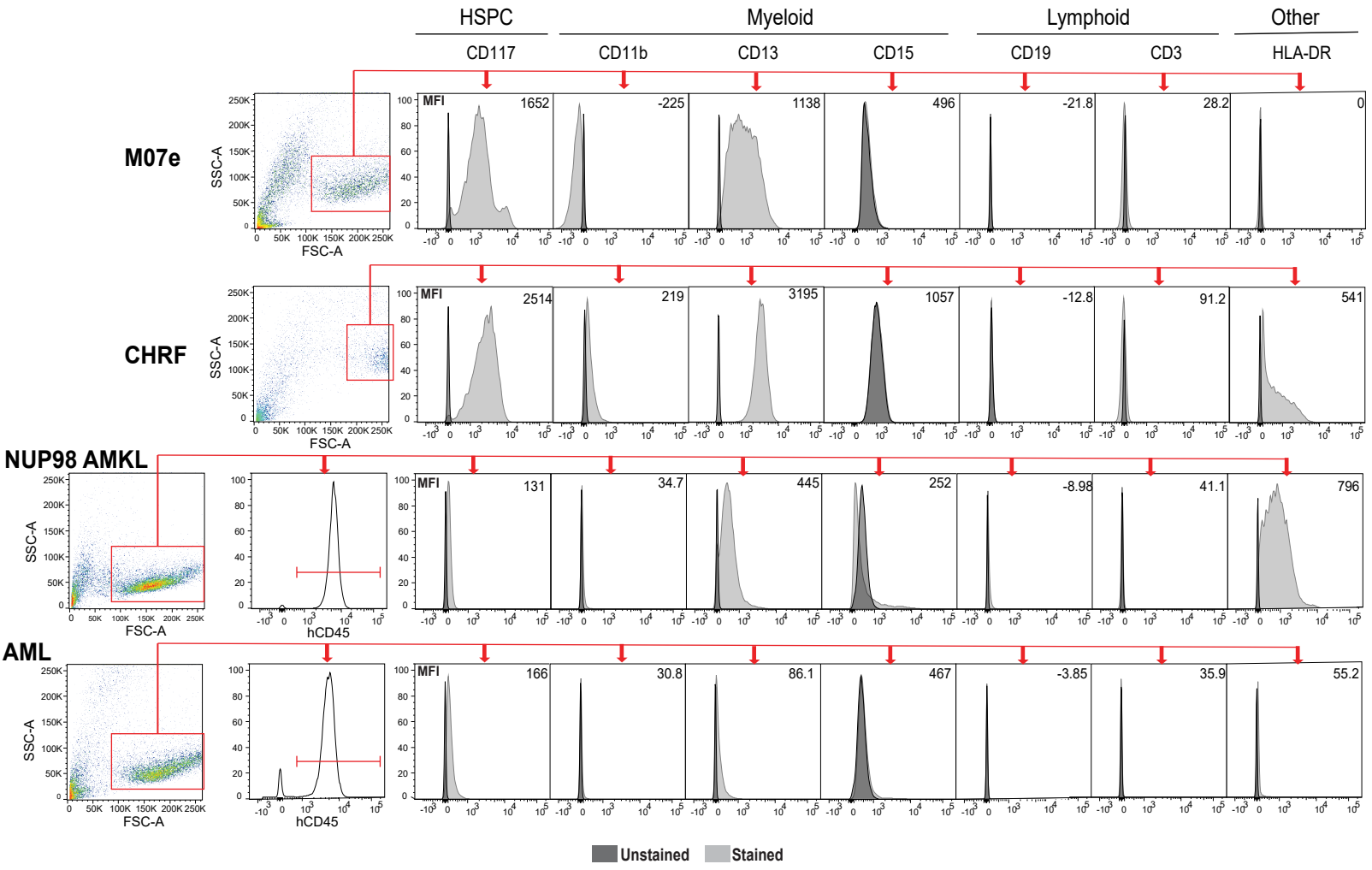


Figure S2: Expression of selected cell surface markers on different subtypes of AML.

Flow cytometry pseudocolor dot plots depicting gating strategy and flow cytometry histograms depicting mean fluorescence intensity (MFI) of selected hematopoietic lineage markers (HSPC, myeloid, lymphoid and others) in the M07e cell line, CHRF cell line, primary mCG2-1 and *NUP98* AMKL model spleen cells and primary AML model spleen cells. M07e and CHRF are gated on viable cells. *NUP98* AMKL and AML cells first gated on viable cells and then on hCD45 positive cells. Dark grey plots are unstained controls, light gray plots are stained. Dark grey plots are unstained, light gray plots are stained. HSPC, hematopoietic stem and progenitor cells.

Figure S3

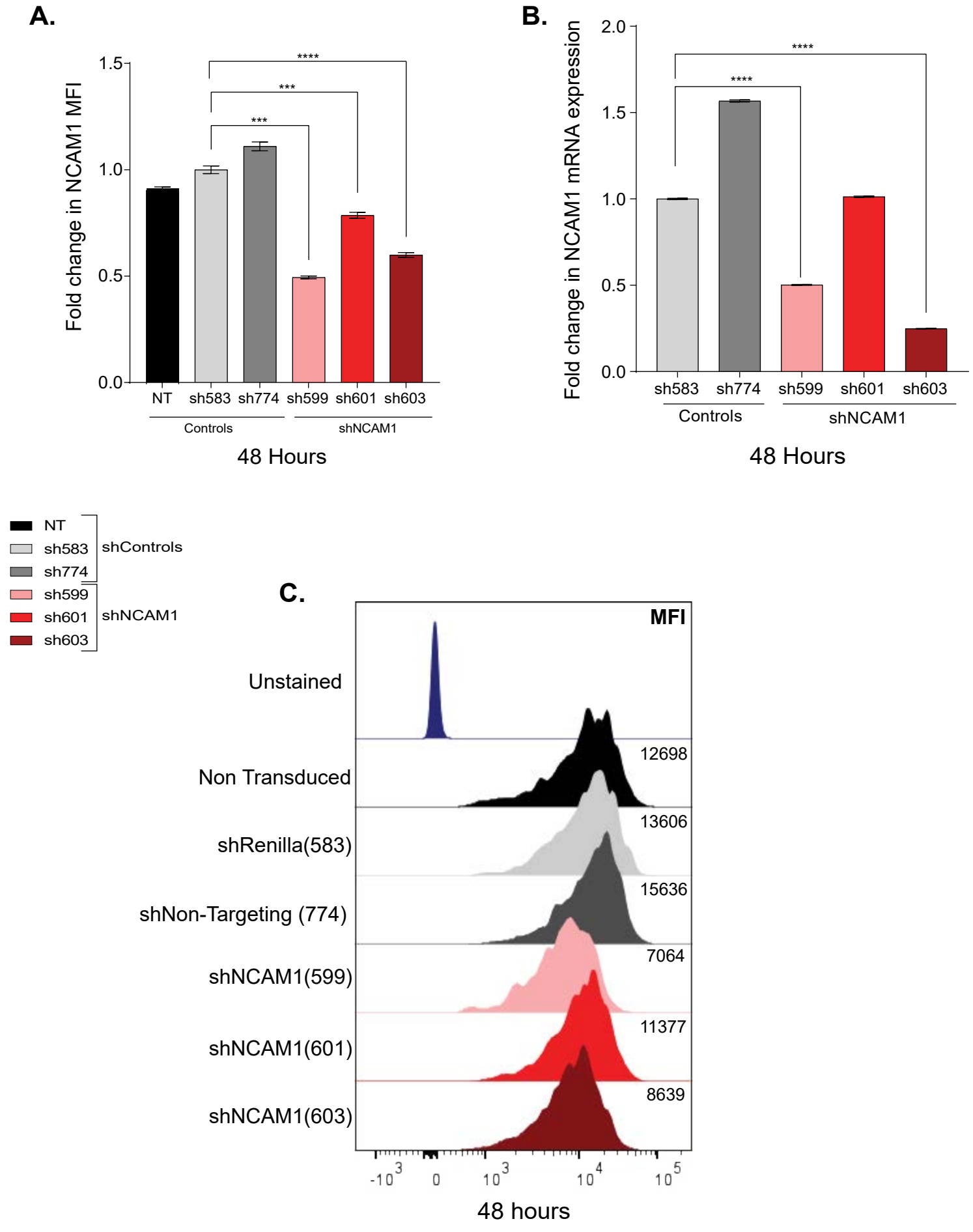

A.
Fold change in NCAM1 MFI
1.5
1.0
0.5
0.0
****
***
***
NT
sh583
sh774
sh599
sh601
sh603
Controls
shNCAM1
48 Hours
B.
Fold change in NCAM1 mRNA expression
2.0
1.5
1.0
0.5
0.0
****
****
sh583
sh774
sh599
sh601
sh603
Controls
shNCAM1
48 Hours
NT
sh583
sh774
shControls
sh599
sh601
sh603
shNCAM1
C.
MFI
Unstained
Non Transduced
12698
shRenilla(583)
13606
shNon-Targeting (774)
15636
shNCAM1(599)
7064
shNCAM1(601)
11377
shNCAM1(603)
8639
-10^3
0
10^3
10^4
10^5
48 hours

Figure S3: Validation of lentiviral shRNA, targeting control genes and *NCAM1*, gene transfer, and knockdown efficiencies at 48 hours. A). Fold change in *NCAM1* surface protein expression (MFI) quantified by flow cytometry 48 hours post transduction with indicated shRNA against *NCAM1* in comparison to control vectors (shRenilla (583) and Non-targeting (774)) or Non-transduced (NT) cells in the M07e cell line. Normalised against shRenilla, multiple t-tests performed, depicts Mean +/- SEM, N=3/sh. **B).** Relative *NCAM1* gene expression detected by ddPCR normalised against *ABL* endogene and calibrated to control condition shRenilla (sh583). Depicts change in *NCAM1* expression between shControl conditions and sh*NCAM1* conditions 48 hours post transduction in the M07e cell line. Multiple t-tests performed, p less than 0.05. Depicted as Mean +/- SEM, N=3/sh. **C).** Flow cytometry histograms showing visual shifts in *NCAM1* surface protein expression 48 hours post transduction in the M07e cell line. Depicts visual shifts in *NCAM1* MFI for indicated shRNA against *NCAM1* (sh599, 601, 603) in comparison to control vectors (shRenilla and Non-targeting) or Non-transduced cells. P-values: * <0.05, ** <0.01, *** <0.001, **** <0.0001. MFI, mean fluorescence intensity. SEM, standard error of the mean.

Figure S4

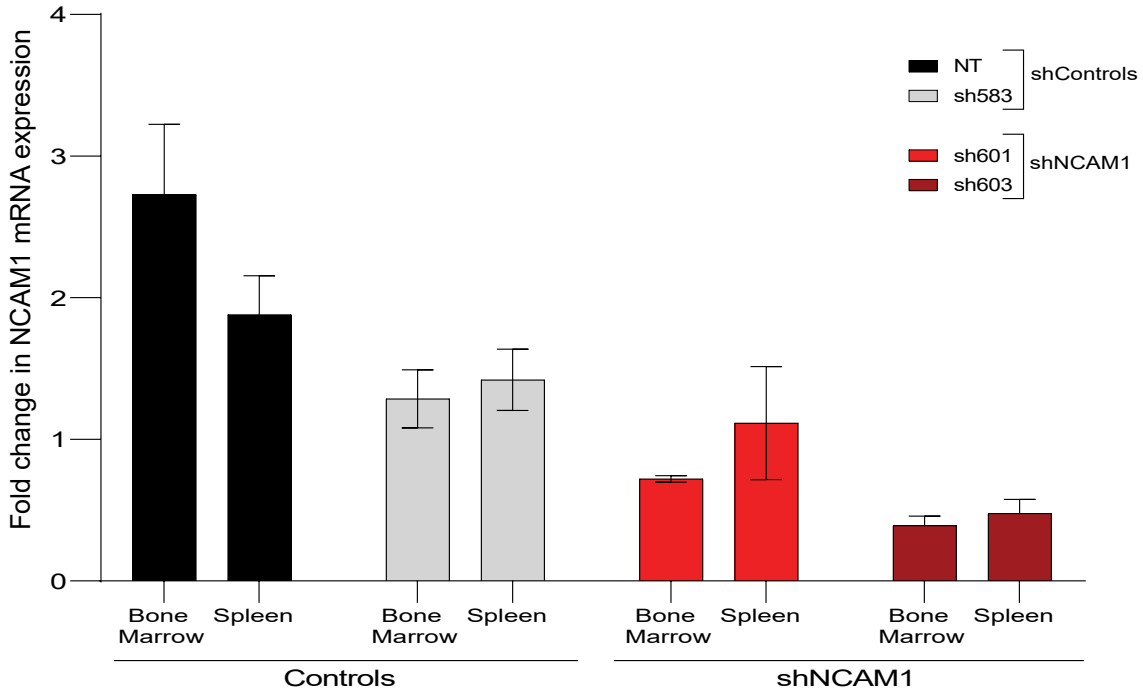


Figure S4: Relative *NCAM1* gene expression assessed by qPCR after *NCAM1* KD *in vivo.* *NCAM1* gene expression detected by qPCR normalised against *ABL* and *PSFM1* endogenous reference genes and calibrated to control condition shRenilla (sh583). Depicts fold change in *NCAM1* expression between shControl conditions and sh*NCAM1* conditions on the day of endpoint (week 9) in isolated mCG2-1 bone marrow and spleen cells. Depicted as Mean +/- SEM, N=3 for NT, N=5 for sh583, N=3 for sh601 and N=4 for sh603.

Figure S5

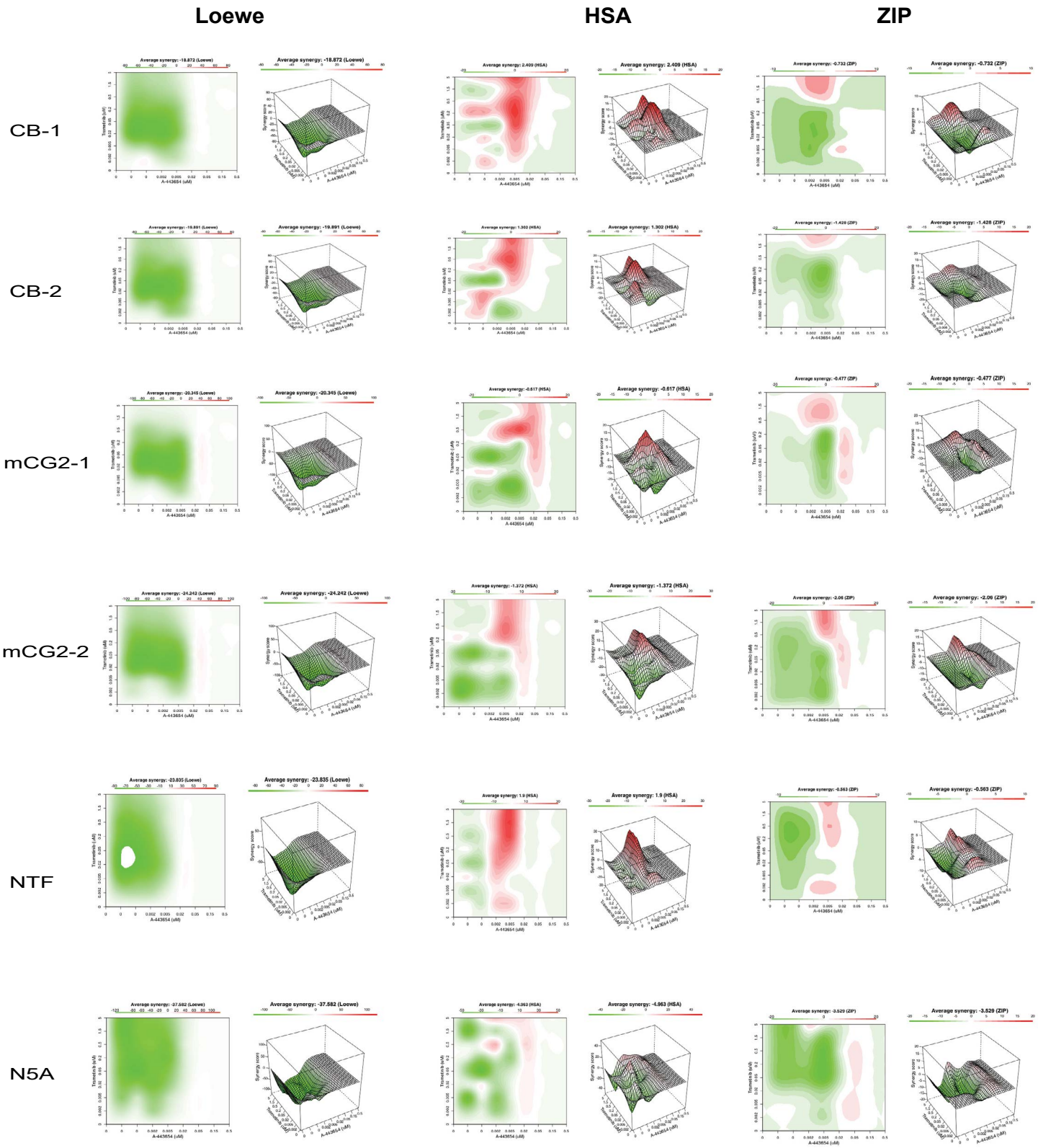
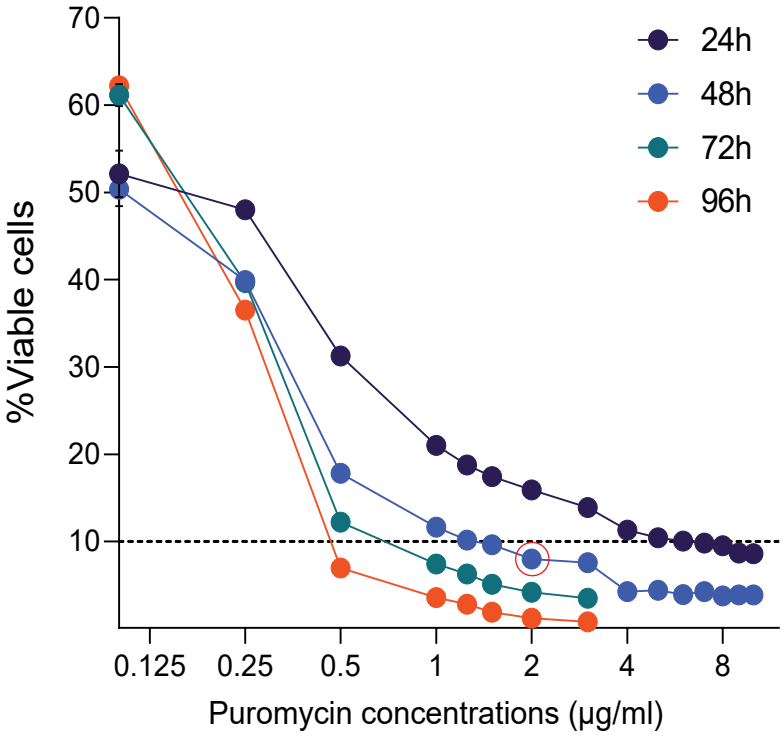


Figure S5: Other synergy reference models for Trametinib and A-443654 combination. Illustrations of 2D and 3D models of average synergy scores for the Loewe, HSA and ZIP reference models in two cord blood pools (CB1 and CB2), mCG2-1, mCG2-2, *NUP98::BPTF* (NTF) and *NUP98::KDM5A* (N5A) primary cells.

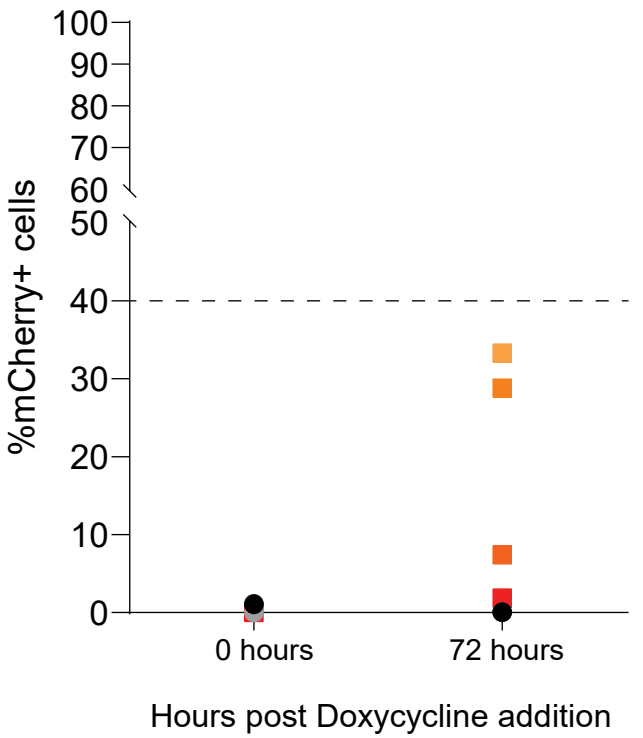
Figure S6

A.

B.

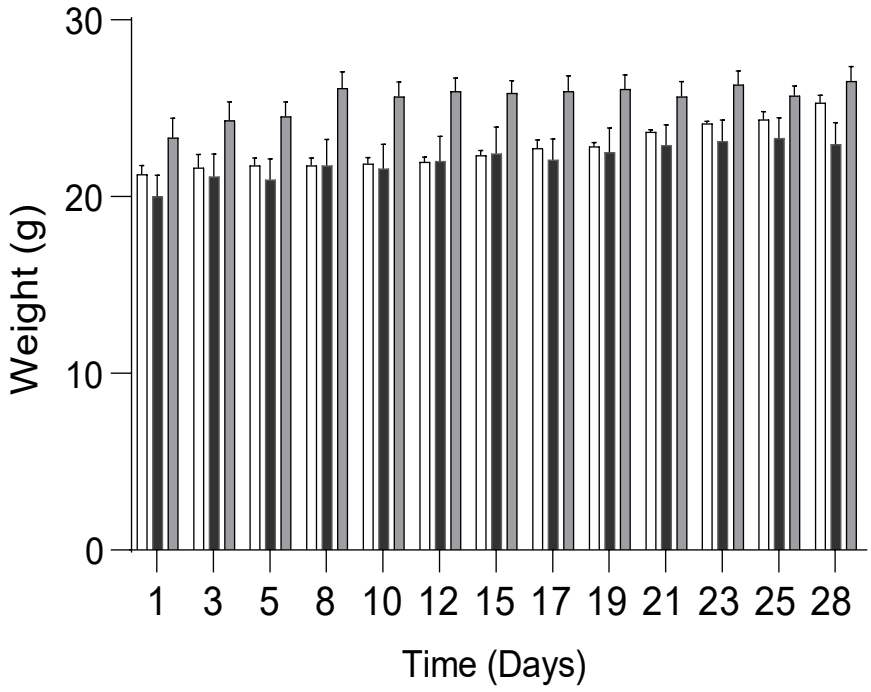


C.



	Transduced	Puromycin	Doxycycline
●	NT	No	No
●	NT	2µg/ml	No
●	NT	No	2µg/ml
■	Tet-shRenilla	2µg/ml	No
■	Tet-shRenilla	2µg/ml	100ng/ml
■	Tet-shRenilla	2µg/ml	1µg/ml
■	Tet-shRenilla	2µg/ml	2µg/ml

D.



Control (Not treated)
Doxycyclin in 2% sucrose water (Doxycyclin Hyclate: 2mg/ml)
Doxycyclin in food pellet (625 mg/kg)

Figure S6: Supplemental inducible lentiviral structure for the inducible doxycycline mediated knockdown system. A). The lentiviral structure for the inducible doxycycline dependent knockdown system. Between the 5' and 3' LTR regions the vector contains the retroviral Ψ packaging sequence, an inducible P_{TRE3G} promoter, followed by an mCherry fluorescent reporter gene, a miR-E backbone, the insert of the short hairpin (shRNA) against the gene of interest (e.g sh*N-CAM1* and shRenilla and shNon-targeting as controls), a constitutive *PGK* promoter, a puromycin resistance gene for selection, an IRES (internal ribosome entry site) and the Tet-On 3G transactivator protein gene (TA). The constitutive promoter *PGK* drives the expression of the TET-On 3G transactivator protein which is unable to bind to the inducible promoter P_{TRE3G} in the absence of doxycycline. When doxycycline is added, it will bind to the transactivator protein inducing a conformational change that will allow binding to the inducible P_{TRE3G} promoter (tet-operator sequences) and triggering of transcription and knockdown. Figure created with BioRender.com. **B).** Puromycin killing curve performed with various concentrations of puromycin (0.125ug/ml to 8ug/ml) over a 24-, 48-, 72- and 96- hour periods. The percentage of viable cells was assessed using PI staining in the M07e cell line. Puromycin selection at 2ug/ml for 48 hours (circled in red) was deemed as optimal and produced under 10% viable cells. **C).** The percentage of %mCherry+ cells on the day of doxycycline addition and 72 hours later in the M07e cell line assessed by flow cytometry. Several conditions were tested (Table) with the addition of puromycin and without. The Non-transduced conditions with no puromycin were used as negative control for mCherry expression or to assess doxycycline toxicity baseline. The other Non-transduced condition with puromycin was as a selection control. Overall, induced mCherry expression was detected in less than 40% of cells. **D).** Doxycycline was administered to mice as either a food pellet (625mg/kg) or in 2% sucrose water (2mg/ml). Body weight (g) and dehydration levels was assessed over 28 days and doxycycline was found to have no impact.