DERIVING AN INDUCED PLURIPOTENT STEM CELL MODEL OF ACUTE MYELOID LEUKEMIA

DERIVING AN INDUCED PLURIPOTENT STEM CELL MODEL OF ACUTE MYELOID LEUKEMIA TO CAPTURE AND ISOLATE DISEASE HETEROGENEITY

By
DIANA GOLUBEVA, H.B.Sc.

A Thesis
Submitted to the School of Graduate Studies
In Partial Fulfillment of the Requirements
for the Degree
Master of Science

McMaster University

MASTER OF SCIENCE (2019)	McMaster Univeristy
(Biochemistry and Biomedical Sciences)	Hamilton, Ontario

TITLE	Deriving an induced pluripotent stem cell model of acute myeloid leukemia to capture and isolate disease heterogeneity
AUTHOR:	Diana Golubeva, H.B.Sc. (McMaster University)
SUPERVIOSOR:	Dr. Mick Bhatia
NUMBER OF PAGES	x, 86

LAY ABSTRACT

Leukemia is the cancer of the blood system, in which leukemic cells divide too quickly and fail to mature into functional blood cells. These cells overcrowd healthy blood cells and prevent their normal function. Acute myeloid leukemia (AML) is a subtype of leukemia where only 5-15% of patients over 60 achieve remission and the median survival for elderly patients is less than a year. Elderly patients succumb to the disease due to their inability to tolerate aggressive and non-selective treatments. The lack of specific treatments is in part, due to a wide variety of genomic mutations in AML leading to variable responses to standardized therapy. In addition, there are no therapies to exclusively eliminate leukemic but not healthy blood cells. In the laboratory, many cancers are studied by using cell or animal models of the disease. Unfortunately, in the case of AML, these models are flawed as they capture less than half of genomic variations present in patients. When using patient samples directly it is often impossible to discern healthy from leukemic cells making it problematic to derive therapies specific to cancer cells. Reprogramming cancer cells into induced pluripotent stem cells (iPSCs) is a way to capture patient specific genetic mutations and physically separate healthy and mutated cells to be able to study them separately. This model has the potential to dissect variable patient mutations to identify which are root causes of the disease. It has been shown that AML is difficult to reprogram, and here we dissect reprogramming strategies used to the past to for the first time. We then demonstrate the most successful to date capture of three AML patients in an iPSC model and interrogate the model's ability to capture and recapitulate AML disease.

ABSTRACT

Leukemia is a cancer of the blood and bone marrow where acute myeloid leukemia (AML) accounts for the most leukemic deaths annually. Only 5-10% of patients over 60 reach complete remission and the median survival for elderly patients is 8.5 months. Elderly patients succumb to the disease due to inability to tolerate conventional chemotherapy regimens. Despite advances in sequencing technologies supporting mutation identification in AML patients, it is still unclear which mutations are disease drivers and specific treatments for almost all patients are not available. This is in part because there are no known targets that distinguishing leukemic cells from their healthy counterparts. Using patient samples to derive new therapies has been challenging due to the technical inability to identify healthy from leukemic cells in a sample. Reprogramming of somatic cells into induced pluripotent stem cells (iPSCs) uniquely enables the separation of healthy and leukemic cells from AML patients through the derivation of AML and healthy isogenic iPSC cell lines. This holds potential for disease modelling and elucidation of functional driver mutation in AML. Leukemic iPSC lines have not yet been studied comprehensively due to the lack of a robust method to reprogram AML. Deriving iPSCs from leukemia patient samples has proven to be challenging, where all combined studies are limited by capturing a total of four patients spanning only two of fourteen genetic classes. Here, we for the first time interrogate strategies to selectively reprogram AML and healthy cells from patients. Using fluorescently activated cell sorting, we show that AML and healthy reprograming is restricted by the differentiation state. Further, we demonstrate that the use of myeloid

markers in deriving AML-iPSCs is disadvantageous as AML clones reside in a variety of compartments. We present, the most successful to date, derivation of mutant and healthy variety of AML-iPSCs from three patients, capturing a wide variety of genetic aberrations. AML-iPSCs demonstrate pluripotency features with the ability to model AML through hematopoietic differentiation. The derived AML-iPSC model has further potential to use gene expression and chromatic availability signatures compared to isogenic controls to identify novel genetic and epigenetic targets in AML.

ACKNOWLEDGEMENTS

I would like to thank Dr. Mick Bhatia for sharing your scientific inspirations which sparked my interest for stem cell research. Your openness to novel ideas and creativity has been a great driver for my work and will always remain to be. Thank you for providing me with the opportunity to be a part of the world-class team of scientists and students where I was also to expand my critical thinking, problem solving and communication skills.

To my committee members Dr. Karun Singh and Dr. Ray Truant, thank you for always taking a keen interest in the project, your comments have helped keep my science at an excellent level. And thank you for keeping my personal goals and aspirations in mind.

To the members of the Bhatia lab, thank you for sharing your scientific expertise, your ideas and your daily support. Thank you to all who have taken their time to teach me new techniques and how to be a better scientist. You have all had a role in shaping me into the interesting individual I am today. I look forward to working with, and keeping friendships with all of you in the future. To Jenn and Allison, thank you for providing me with your mentorship. And thank you to Deanna and Meaghan for picking up where I left off.

Thank you to Brian for always keeping my dreams alive. I of course could have never done this without my parents, Alex and Natalia. And lastly, thank you to my best friend who probably knows too much.

TABLE OF CONTENTES

LAY ABSTRACT iii

ABSTRACT iv

ACKNOWLEDGEMENTS vi

TABLE OF CONTENTS vii

LIST OF TABLES AND FIGURES ix

LIST OF ABBREVIATIONS x

DECLARATION OF ACADEMIC ACHIEVEMENT xi

1.0 INTRODUCTION

1.1 Human Pluripotent Stem Cells (hPSCs) 1

1.1.1 The Beginning

1.1.2 Human Embryonic Stem Cells (hESCs)

1.1.3 PSCs as Research Tools

1.2 Stemness in Cancer: A Rationale for Pluripotent Cancer Models 6

1.2.1 Overlaps in Pluripotency and Cancer Cellular Programs

1.2.2 Cancer Epigenetics Specific to Tissue of Origin

1.3 Acute Myeloid Leukemia 13

1.3.1 Perspective and Etiology

1.3.2 Cancer Stem Cells in AML

1.3.3 AML Genetic Heterogeneity and Targeted Therapies

1.3.4 Epigenetic Dysregulation in AML

1.3.5 Myeloid neoplasm disease models

1.4. The Potential of an iPSC Model of AML 23

1.4.1 Cellular Reprogramming and Induced Pluripotent Stem Cells (iPSCs)

1.4.2 Factors Affecting Cellular Reprogramming

1.4.3 Challenges of Reprogramming Primary Cancers

1.4.4 Reprogramming of AML

1.5 Study Rationale 33

1.6 Study Hypothesis 34

1.7 Study Objectives 35

2.0 MATERIALS AND METHODS 36
2.1 Primary Patient Samples
2.2 Reprogramming of Primary Samples
2.3 Patient Derived Xenografts
2.4 Flow Cytometry
2.5 Fluorescent in Situ Hybridization (FISH)
2.6 Cytogenomics
2.7 ddPCR
2.8 Teratoma Assay
2.9 Hematopoietic Differentiation
2.11 Cardiac Differentiation
2.12 Statistical analysis

3.0 RESULTS
3.1 Reprogramming primary AML 42
3.1.1 Determination of Primary AML Samples
3.1.2 Primary AML Reprogramming Strategy
3.1.3 Determination of Mutation Status of Derived iPSCs
3.1.4 LSC+ xenografts give rise to unique AML-iPSC clones
3.1.5 Reprogramming of AML and Healthy is Enriched in CD34+ Population
3.1.6 Reprogramming of Healthy, but not AML, Within AML Samples can be Selected for Using CD33
3.1.7AML and healthy iPSCs from AML samples all arose from blast gate
3.2 AML-iPSC Characterization 61
3.2.1 AML-iPSCs Demonstrate Pluripotency Potential
3.2.2 AML-iPSCs have Clone Dependent Myelo-erythro Differentiation Potential
3.2.3 AML-iPSCs Possess Ability to Differentiate to Other Mesoderm Tissues

4.0 DISCUSSION
4.1 AML Reprogramming 67
4.1.1 AML Reprogramming Captures Abnormal karyotype and Intra- Patient Heterogeneity
4.1.2 Reprogramming of AML Xenografts Captures Exclusive Disease Clones
4.1.3 Reprogramming Potential of AML Cells is Dependent on Cell Differentiation Stage

4.2 AML-iPSC Model Characterization: Pluripotency 71
4.2.1 AML-iPSCs have Clone Dependent Hematopoietic Differentiation Capacity
4.3 Challenges of an iPSC derived model of AML 73

5.0 CONCLUSION 74

6.0 REFERENCES 77

LIST OF TABLES AND FIGURES

LIST OF TABLES

1.1 Reprogramming of human cancers into induced pluripotent stem cells (iPSCs) 30
3.1 AML samples for reprogramming 44

LIST OF FIGURES

1.1 Applications of tumour derived cancer iPSCs 26
3.1 LSC determinations of AML samples 44
3.2 Reprogramming Strategy of Primary AML 47
3.3 AML 1 and 2 gave rise to AML and normal iPSCs 50
3.4 Summary of aberrations in AML and normal iPSCs form AML samples 1-3 51
3.5 AML 2 gave rise to AML-iPSCs 52
3.6 AML 4, 5 and 11 iPSCs determined to be normal 53
3.7 AML 6 gave rise to exclusively normal iPSCs 54
3.8 AML 7 gave rise to normal iPSCs 54
3.9 AML reprogramming of CD34 and CD33 sorted populations. 57
3.10 AML and healthy reprogramming dependence on CD34 and CD33. 58
3.11 AML iPSC and normal reprogramming originates from within AML blast population 60
3.12 AML-iPSCs possess equivalent pluripotency to normal iPSCs 62
3.13 Hematopoietic differentiation potential of AML-iPSCs 65
3.14 AML and normal iPSCs cardiomyocyte differentiation potential 66

LIST OF ABBREVIATIONS

ALS	Amyotrophic Lateral Sclerosis
AML	Acute Myeloid Leukemia
B-ALL	B Cell Acute Lymphocytic Leukaemia
BM	Bone Marrow
cDC	conventional Dendritic Cell
CFU	Colony Forming Unit
CML	Chronic Myeloid Leukaemia
FACS	Fluorescently Activated Cell Sorting
FISH	Fluorescent in situ Hybridization
GMP	Granulocyte-Monocyte Progenitor
HSC	Hematopoietic Stem Cell
iMEF	irradiated Mouse Embryonic Fibroblast
iPSC	induced Pluripotent Stem Cell
IT	Intra Testicular
IV	Inta Venous
LTC-IC	Long-Term Culture Initiating Cell
MDS	Myelodysplastic Syndromes
MNC	Mono-nuclear Cell
MPD	Myeloproliferative Disease
NK	Natural Killer
NOD	Non-obese Diabetic
NSG	NOD Scid Gamma
PDAC	Pancreatic Ductal Adenocarcinoma
PSC	Pluripotent Stem cell
SCID	Severe Combined Immune-deficient
SeV	Sendai Virus
SRC	SCID Repopulating Cell
TF	Transcription Factor
VPA	Valproic Acid

DECLARATION OF ACADEMIC ACHIEVEMENT

This thesis is completed mainly with the work of Diana Golubeva with the following contributions from the members of Dr. Mick Bhatia's laboratory. Deanna Porras conducted all cardiomyocyte differentiation of AML and normal iPSCs. Deanna Porras conducted hematopoietic differentiation of some AML clones from AML 1 and 2. Allison Boyd provided consult on AML sample section and xenografting.

1.0 INTRODUCTION

1.1 HUMAN PLURIPOTENT STEM CELLS (hPSCs)

1.1.1 The Beginning

hPSCs are defined by their ability to simultaneously carry out two opposing functions; self-renewal and differentiation, to maintain and produce all required functional cell types [1,2]. The principles of asymmetric division and pluripotency have implications across the fields of developmental biology, regenerative medicine and cancer. The first coining of the term "stem cell" is traced back to 1906 when Alexander Maximov used it to describe specific cells within the human hematopoietic system; which is used as a stem model to this day[3].

More concrete findings were published in 1932, in which observations by Sabin *et al.,* suggested that cells in the mouse bone marrow (BM) are susceptible to DNA damage by radiation, which results in reduction of mature hematopoietic cells leading to anemia[4]. The same was observed in the human body as a result of atomic bombings of Hiroshima and Nagasaki where deaths due to exposure to the lowest dose of lethal irradiation occurred as a result of hematopoietic failure. In an attempt to devise a method that could protect an otherwise very sensitive hematopoietic system from irradiation, Jacobson *et al.,* uncovered that protecting the spleens of mice allowed them to recover from an otherwise lethal dose of irradiation[5]. Subsequent studies demonstrated that the otherwise irreversibly damaged hematopoietic system could be replaced by grafts in mice, guinea pigs, and dogs[6-8].

Shortly thereafter, Till and McCulloch published a series of experiments spanning from 1961 to 1968, which for the first time, allowed hematopoietic stem and progenitor cells to be quantified[9-13]. Their experiments demonstrated that cells in BM and spleens of mice were capable of self-renewal and clonal differentiation to myelo-erythroid and lymphocyte progeny. Simultaneously Thomas *et al.,* reported the first BM transplants for acute myeloid leukemia (AML) patients[14]. Although survival of transplanted individuals was limited in the late 50's and early 60's due to graft versus host disease and graft rejection, by 1968 patients were successfully transplanted with HSCs from an HLA-identical sibling donors[15]. Ten years later, hematopoietic stem cells (HSCs) were discovered in umbilical cord blood [16].To date, the hematopoietic system continues to serve as a major discovery engine of stem cell properties, development, regenerative medicine and oncogenesis.

1.1.2 Human Embryonic Stem Cells (hESCs)

The first demonstration of *in vitro* culture of an ESC line was established in the murine system, where ESCs were derived from mouse embryos and maintained in culture[16,17]. It was not until 1998, that through the use of *in vitro* fertilization the first hESC line was derived from the inner cell mass of the pre or peri-implantation embryo by Thomson *et al.*, providing the field with an incomparable tool that allowed human stem cells to be cultured in the laboratory[18,19]. Herein, hESC hallmarks are defined by their capacity for prolonged propagation in the undifferentiated state and pluripotency; the capability to produce all three embryonic germ layers (endoderm, mesoderm, and

ectoderm). Despite significant scientific challenges and the ethical debates on the use of human embryos that were poised after Thomson's discovery, 59 cell lines with of varying features were derived and characterized[20,21].

1.1.3 PSCs as Research Tools

Once derived, and shown to be compatible with *in vitro* culture, PSCs were thought to be the holy-grail for the future of regenerative medicine for the purpose of transplantation, and a vastly superior alternative to animal models in disease modelling and drug screens.

Using PSC derived products for transplantation has been a goal for neurodegenerative disease, diabetes, heart disease, vision disorders and hematopoietic malignancies[22]. Regenerative medicine is faced with multiple challenges, such as: need for generation of mature adult tissues *in vitro* and immune-rejection of derived tissues. Despite these challenges, certain tissues, especially in immune privilege sites such as the eye, have been successfully produced from PSCs. Clinical trials have been able to reverse blindness of patients suffering from age-related macular degeneration by successfully treating patients with hESC derived retinal pigmented epithelium patches[23]. The generation of other tissues for transplantation has been a challenge. For example, there is a high need for the derivation of HSCs from pluripotent cells for transplantation as it would revolutionize the treatment of cancers and immune disease alleviating the need for human BM donors. However, despite continuous progress achieved by world class research groups, a *bona fide* HSC capable producing long-term engraftment in an

immune-deficient mouse model is still yet to be derived from a PSC source[24-31]. The only study that was able to demonstrate long term, multi-lineage reconstitution was using transcription factors (TFs) to mediate this cell fate, making it challenging to transition the generated cells into clinical applications[32]. The HSC field is still seeking to develop a process to make transgene free HSCs from pluripotent cells. Nearest to clinical trials for transplantation of blood products, is the derivation of specific cell types, red blood and platelet, cells to supplement clinical transfusions[33]. These cells would be unable to repair the whole hematopoietic system and would not treat the root cause of hematologic disease but would be an extremely valuable resource for the unmet demand of blood transfusions. Importantly, even these products require further investigation, to work out kinks related to cell maturation, and development of stroma-free mass scale production.

On the other hand, disease modelling using PSCs has been widely used to better understand genetic diseases, host pathogen interactions and cancer. PSCs allowed for the modelling of otherwise hard to obtain tissues and made modelling of known genetic alternations feasible in a variety of differentiated tissues. An example of this is the use of organoids. Organoids are *in vitro* generated 3D cell cultures that allow for the development of organ-like structure which resemble organs by their architectural and functional components. They allow investigation of human tissue and organ biology without concerns of accessibility or ethics. Pioneering work on organoids demonstrated that PSCs are able to undergo self-patterning[34]. The generation of a variety of tissues demonstrated the PSC intrinsic potential to self-organize, a homogeneous population of

cells within a uniform extrinsic environment, into 3D structures[35,36]. Organoids have thus been useful in studying tissue development, disease and PSC biology.

Modelling of disease using PSCs allows for elucidation of molecular, physiological phenotypes of disease and goes one step further to allow performance of large scale drug screens in disease specific tissues which are otherwise difficult to obtain. This has allowed for significant advancements in drug discovery for neurological disorders such as Familial dysautonomia and amyotrophic lateral sclerosis (ALS) [37,38]. PSC derived motor neurons were successfully used to model and to identify a drug candidate for ALS patients[39]. The identified drug target was FDA approved for the use of patients with epilepsy and was moved into Phase II clinical trials less than two years after initial *in vitro* experiments.

The above examples demonstrate the utility of PSC in modelling disease with drug screening capacity and the potential of an expanding variety of PSC derived tissues to be used for transplantation. The stem nature of PSCs provides another avenue of discovery, as stem features are commonly observed among many cancer types. PSCs have the potential to be used as a discovery engine to identify cancer therapies spanning many cancer subtypes which has not previously been possible.

1.2 STEMNESS IN CANCER: A REATIONALE FOR PLUTIPOTENT CANCER MODELS

1.2.1 Overlaps in Pluripotency and Cancer Cellular Programs

A vast majority of cancers undeniably possess features which overlap with pluripotency[40]. It is however unclear whether the stem cell resemblance is a result of cancers arising from more primitive cells already possessing stem properties or cellular programs in mature cells reverting to a pluripotent or more stem-like state, or both. Healthy adult stem cells possess some stem properties similar to ESCs[19]. Both are able maintain self-renewal for long periods of time while retaining the capacity to differentiate into specialized cells. Adult stem cell differentiation is, however, restricted to tissue specific specialized cells. Cancer stemness is an emerging concept and refers to similarities of cancer and stem cells measured by gene expression and epigenetic patters. Elevated stemness in cancers is associated with accelerated progression and metastases, and thereby may represent a potential target for broadly applied anti-cancer therapy development[41].

A series of reports support the correlation of stem cell feature acquisition in cancer and disease progression; as mechanisms employed by multiple poorly differentiated cancers overlap tightly with stem cell signatures at the gene expression level[41-43] . Malta *et al*., conducted a comprehensive analysis of gene expression and epigenetic data sets from 33 cancers containing a total of 12,000 human patient samples[42]. A machine learning algorithm was developed to compare these cancer samples to human

stem cells. The machine learning algorithm was trained on a set of ESCs and their counterparts in the form of lineage committed cells, to be able to rank sample stemness and determine if a test sample contains a stemness signal[44]. The algorithm was then used to detect and measure stemness in cancers[42]. As proof of principal, cancers that were observed to be poorly differentiated by histological analysis could be distinguished by higher stemness indices at both gene expression and DNA-methylation levels[42]. Even prior to this extensive analysis, studies have shown that human ESC signatures appear within signatures of aggressive tumours[41,43].

Using patient sample gene expression data from independent studies (including breast cancer, glioblastoma, and bladder carcinoma), Ben-Porath *et al.,* showed substantial evidence that samples with elevated stem gene expression levels resulted in the highest mortality in patients[41]. Impressively, it was possible to rule out that this observation was simply an artefact of rapid proliferation commonly observed in both PSCs and cancers. In order to rule-out this uncertainty, Ben-Porath and colleagues noted that immortalized cancer cell lines demonstrate elevated levels of proliferation (including cell cycle genes) but lack enrichment in stemness gene expression signatures that they used to probe patient cancer samples. This suggests that cell lines, despite features of transformation, may not capture a key feature of cancer physiology and progression which significantly affects patient survival. This could potentially explain the difficulty of transitioning cell line-based cancer therapies to the clinic and serve as rationale to use pluripotent cell models to identify broadly effective cancer therapies.

Earlier studies drew conclusions that association of stemness and cancer though the use of bulk cell transcriptome analyses, which averages expression patterns across a heterogeneous population[41]. This indicates that in some cancers, namely breast cancer, gliomas and bladder carcinomas, the stem cell signature is not restricted to exceedingly rare populations of cells. Thanks to rapid advancements in library preparation and sequencing technologies single cell transcriptome analysis allowed for the analysis of what were already known to be heterogeneous cell populations within tumors. In cancers including glioblastoma and leukaemia, elevated stemness has indeed been identified to originate from sub-populations of cancer cells[42]. It would be valuable to further investigate and functionally demonstrate whether these subpopulations play a key role in tumor maintenance and progression. If these populations in fact play a role in disease initiation, progression or relapse, the genetic and epigenetic drivers of disease can be identified to develop novel therapies targeting these populations.

Currently the only functional demonstration of the overlapping properties between induced PSCs and cancer, is the ability of iPSCs to generate immunity against multiple cancer cell lines[45]. Immunity was induced in T and B cells and could be transferred to a non-immunized host to prevent tumor growth. Despite exciting findings, this report is exclusive to the use of the murine system and is thus limited to the use of cancer cell lines making it impossible to discern if the immune response is in fact elicited to the stem, or simply proliferative properties of the tumors.

Altogether, these studies demonstrate that a wide array of cancers share similarities with PSCs, the presence of which is correlative to clinical outcomes of

patients. This should be implicated when selecting cancer models in order to identify selective and effective therapies, however it is clear that further functional affirmation is still required[43].

1.2.2 Cancer Epigenetics Specific to Tissue of Origin

In order to perform tissue-specific functions, cells have a complex and not yet fully understood regulatory system that is able to modulate genome behaviour through alterations in chromatic structure referred to as epigenetics[46]. Epigenetic memory refers to heritable changes in chromatin structure that can be passed on through cell divisions without any alterations to DNA sequence, from the cell-of-origin to daughter cells. DNA methylation is one of the most-studied epigenetic modifications in mammalian cells and is widely known to be aberrantly regulated in cancers[47]. Methylation most frequently occurs in cytosine/guanine rich sites called CpG islands. CpG islands are found in many promoters, and when methylated, the promoter region is not available to enhancer elements thus reducing transcription of associated genes[48].

Interestingly, the cell-of-origin memory encoded in the epigenome is preserved even after cancerous transformation[40]. Global analysis of CpG sites in 10,814 tumours in a DNA-methylation based clustering analysis showed that tumours cluster according to the tissue of origin and co-cluster within organ systems. As an example, cancers of the gastrointestinal tract are more similar to each other than cancers of the kidney. Impressively, Hoadley *et al.,* showed that the cell-of-origin clustering was maintained even after tissue specific methylation sites were stringently excluded. This data suggests

that dysregulation of DNA methylation may be similar in cancers originating from tissues of the same lineage, while being distinct from the tissues of origin themselves. The epigenetic landscapes of cancers also carry features of stemness, demonstrated by Malta *et al*., using a machine learning algorithm that was used to identify stemness in DNA methylation signatures of cancers[42]. Presumably these are markers that are acquired by cells in the progress of cancer transformation and if it can be functionally shown that cancers of the same lineage possess similar epigenetic markers unique from the healthy tissue of origin, these have the potential to become selective and broadly effective cancer targets.

The presence of tissue-of-origin and stem-like features in the molecular and epigenetic landscape of cancer cells poses an interesting question about the type of cell that undergoes initial transformation. The observed primitive cancer cells can potentially arise in multiple ways: 1) a non-stem cell can become dysregulated through re-activation of stemness programs and 2) a stem or progenitor cell which already possess stem cell potential fails to balance differentiation and renewal programs leading to transformation[42,49]. It remains unclear which of these is more prominent and whether this differs across cancer subtypes. Insights into the cell-of-origin would be invaluable for the targeted treatment of classes of cancers that share a common ancestor.

1.2.3 The Potential of Capturing Cancer Epigenetics in a Pluripotent Model

Interestingly epigenetic memory has been shown to be preserved through the generation of induced pluripotent stem cells (iPSCs) via factor based reprogramming

from a variety of cell types[50,51]. Lee *et al.,* show that epigenetic memory may be implicated in iPSCs. iPSCs derived from variable tissues were shown to have indistinguishable pluripotency based on functional assays and gene expression patterns, with no lineage partialness in gene expression while in the pluripotent state. And yet, they showed that iPSCs possess a lineage biased propensity for differentiation to their tissue of origin at the expense of other lineages. This phenomenon is also captured when looking into methylation patterns in mouse iPSCs, where it was shown that iPSC DNA harbours residual methylation resembling the tissue of origin[52]. In a comprehensive study, iPSCs have been derived from a representative mouse tissue of all three lineages: hepatocytes, skin fibroblasts and melanocytes representing endoderm, mesoderm and ectoderm respectively[53]. All derived iPSCs were shown to retain a transcriptional memory of the tissue of origin at an early passage. The study by Lee, *et al.,* highlights that in human iPSCs, the epigenetic marks may not be detectable at the transcriptomic level they must be present as the iPSCs are able to functionally favour differentiation to the lineage and tissue of origin, at the expense of differentiation to other lineages and tissues. Thanks to advancements in sequencing technology, we now know that pluripotent cells are not necessarily a homogeneous population and thus it is worth noting that these earlier studies had been working with bulk cell populations making it difficult to identify sub-populations[54]. Single cell analysis may be required to identify distinct cell populations that are responsible for this preferred lineage propensity in human iPSCs.

The epigenetic memory found in iPSCs is already actively being used to study disease where pathogenesis is tightly associated with epigenetic dysregulation. Multiple

groups used iPSCs derived from fibroblasts of patients with immunodeficiency, centromeric instability and facial anomalies type I (ICF1) with mutations in DNMT3B gene, responsible for *de novo* methylation, resulting in genomic hypomethylation[55,56]. The derived iPSCs were shown to contain disease specific transcriptome and methylome in the pluripotent state and when differentiated to mesenchymal stem cells. The iPSCs contained disease specific methylation unique from wild-type iPSCs and healthy BM isolated mesenchymal stem cells. These studies used the iPSC models to elucidate disease molecular mechanisms related to telomere loss and premature senescence.

The above studies serve as proof of principle that iPSCs are able to retain memory of methylome of their tissue of origin and of a disease. This provides a promising avenue to study methylome and other epigenetic changes associated with disease using iPSCs. Since this is feasible for other disease, and abnormal epigenetics are implicated in cancers, iPSCs derived from cancer tissues have the potential to serve as cancer epigenetics models.

1.3 ACUTE MYELOID LEUKEMIA

1.3.1 Perspective and Etiology

The healthy hematopoietic system resembles a hierarchy in which rare hematopoietic stem and progenitor cells reside at the top and possess the ability to asymmetrically divide; to simultaneously maintain a tightly regulated hematopoietic stem and progenitor cell population and differentiate to all mature lymphoid and myeloid blood cells[57-59]. Leukemia is the cancer of the hematopoietic system, in which the asymmetric division of a cell in the hematopoietic hierarchy becomes dysregulated, and leads to a rapid overproduction of dysfunctional cells which fail to mature into functional blood cells. These aberrantly differentiated cells outcompete healthy hematopoietic cells in BM further disrupting the hematopoietic balance[60]. AML is a subtype of leukemia characterized by an accumulation of dysfunctional cells in the myeloid branch of the hematopoietic system. This results in the reduction of mature myeloid cells such as: red blood cells, neutrophils, and macrophages resulting in anemia, thrombocytopenia and neutropenia in patients[61].

AML is the most common acute leukemia and accounts for the highest percentage of leukemic deaths with a severely low survival rate of only 25% [62]. Where prognosis for younger patients is promising with survival of 60-70% at 5 years, the prognosis for elderly patients is bleak. This is of high importance because the median age of diagnosis of AML is 70 and despite treatment advancements, the median survival for elderly patients is 8.5 months, and 24% at 5 years. Overall, only 5-15% of patients over 60

achieve complete remission [63]. The high mortality rate results from the inability of older patients to tolerate conventional chemotherapy regimens. Despite recent targeted therapy advances, it remains a challenge to derive specific therapies as AML is vastly heterogeneous among patients and there are no identified markers distinguishing leukemic stem cells from their healthy counterparts.

1.3.2 Cancer Stem Cells in AML

Since the bulk of AML cells from patients were determined to have limited proliferation capacity, it was speculated that leukemia is likely maintained by a rare stem cell population[64]. The leukemic initiating cell (LIC), or the cell responsible for initiation of leukemia in a mouse xenograft, was first demonstrated by Lapidot *et al.,* through disease reconstitution of the severe combined immune-deficient (SCID) mice by a fraction of cells[65]. The fractionation was done by phenotype, where $CD34^{+}CD38^{-}$ cells possessed LIC capacity in AML samples. Interestingly the same population in healthy samples, referred to as SCID repopulating cells (SRCs) was shown to be responsible for hematopoietic reconstitution *in vivo* [66]. It became apparent that not only is AML constituted of a heterogeneous cell population maintained by a small subset of cells, but that it recapitulates the hierarchy distribution of the healthy hematopoietic system, implicating that cancer stem cells are responsible for the disease[67]. Whether this cell has stem properties because it resulted from a transformation event in an HSC or if it was a progenitor cell which has acquired stem like properties through transformation is still unclear.

Recent work by van Galen *et al.,* impressively combines single cell RNA sequencing and genotyping of AML samples[68]. This study overcame a hurdle that was previously faced by transcriptome analysis of AML, which was that it was impossible to distinguish signatures of healthy from leukemic cells with explicit accuracy. This study demonstrated that aberrant AML cells from multiple patients resemble multiple cell types along the myeloid lineage axis supporting the notion that AML resembles healthy hematopoietic structure. The AML cells resembled 6 healthy compartments: HSCs, progenitors, granulocyte-monocyte progenitor (GMPs), pro-monocytes, monocytes and conventional dendritic cells (cDCs). The composition of these compartments in diseased cells varied among patients, where some patients carried aberrant mutations in all compartments, primitive, differentiated, or primitive and differentiated compartments. Primitive AML cells were demonstrated to co-express opposing transcriptional programs (HSC/progenitor genes were co-expressed with GMP genes), where in normal hematopoietic cells these expression profiles were found to be exclusive. Consistent with other reports, patients that had higher HSC/progenitor gene expression in aberrant population had worse outcomes[41,68]. Other studies used next generation sequencing to provide evidence that most AML cases contain clonal heterogeneity composed of a founding clone, as well as a sub-clonal hierarchy including clonal evolution observed after relapse [69].

This data supports the notion that stem like properties of AML should serve as targets to treat disease initiating cells in aggressive cancers. There is a pressing need to dissect the clonal heterogeneity observed in AML using a functional model and understand the

contribution of leukemia-associated mutations to the transformation process, and whether HSCs are more likely to be transformed by these mutations than other subpopulations of the hematopoietic hierarchy.

1.3.3 AML Genetic Heterogeneity and Targeted Therapies

The genetic heterogeneity of AML has been studied for many decades, and the prognostic value of genetic abnormalities was recognized very early on [70]. AML can be sub-divided by chromosomal (cytogenetic) and gene specific (molecular) abnormalities where 50-60% of all cases are associated with a genetic abnormality [71]. These were first organized into a diagnostic criteria by the French-American-British (FAB) classification which relied heavily on the differentiation status of the leukemia for classification [72]. Since, more abnormalities were discovered and added onto the list "The World Health Organization classification of the myeloid neoplasms" was most recently established to classify AML based on cytogenetic and molecular abnormalities [73].

Beyond their use in diagnostics, some genetic abnormalities are being used as targets for precision treatments to work alone or in supplement with the "7+3" treatment. The typically used "7+3" chemotherapy regimen, consists of 7 days of cytarabine and 3 days of anthracycline of varying doses to maximize efficacy while minimizing treatment related mortality which is especially of concern in the elderly. Currently a few targeted therapies are being used in clinic to treat selective molecular abnormalities. The most successful of these and the only therapy used as a sole treatment of AML with no chemotherapy, is based on a molecular abnormality found in a subset of AML, acute

promyelocytic leukemia (APL). APL was first described as a highly aggressive form of hematologic disorder with survival ranging from 1 day to 1 month[74-76]. All-trans retinoic acid (ATRA) and arsenic trioxide (ATO), a differentiation therapy targeting AML disease demarcated by PML-RARα fusion protein product has changed the prognosis drastically with remission and cure rates of 80 and 90 percent respectively [77,78]. This example provides evidence that targeted differentiation therapies could be key in treatment of AML.

Since then, targeted therapies have being developed for FLT3 mutations and IDH1/2 mutations. Midostaurin, a FLT3 kinase and other kinase inhibitor, was evaluated in patients with mutated FLT3 and showed an improved overall survival (OS)[79]. This trial was limited by only assessing patients 18-60 years old, which is problematic considering that it is patients over 60 that are unable to withstand higher doses of traditional treatments and are in need of better therapies. Despite the lack of data regarding older patients, midostaurin has been approved for use in combination with chemotherapy for patients of all ages. Trails are ongoing to test IDH2 inhibitors[80]. It is important to note that despite many driver mutations identified by cytogenetic analysis and gene sequencing, only few, when introduced in healthy human cells, have been implicated to cause initiation of disease in humanized mouse model recipients[81-85]. Most of these, with the exception of MLL-AF9 fusion are shown to work synergistically with other mutations to cause disease, making it difficult to pin point pivotal driver mutations. Due to this, further work is required in order to define functional disease drivers and devise molecular targets which could effectively eliminate the root cause of the disease.

An attempt at a more broadly acting approach at targeted therapy is an antibody conjugated toxin targeting CD33, gentuzumab ozogamicin (GO). The idea behind this target, is that it is heavily expressed in the myeloid lineage of healthy and some AMLs. Trials have shown variable responses ranging from, "failed to show survival advantage" and "found to have survival benefit"[86-88]. This therapy, despite the traction it is receiving, should be heavily reconsidered not only because CD33 is found on healthy myeloid cells but also on healthy HSCs making it a dangerous guide for a toxin therapy[89,90]. This example illustrates the importance of isogenic healthy controls which must be used when generating targeted therapies. Overall, and despite advances in targeted therapies for AML are still far from ideal, and most patients who have identifiable mutations are not able to receive a specific therapy. In addition, it is still unclear which mutations are truly disease drivers and need to be treated to prevent disease relapse. Ideally, a broad acting therapy could be derived that would target multiple AML subtypes, spare healthy cells, and minimize disease evasion.

1.3.4 Epigenetic Dysregulation in AML

In addition to cytogenetic abnormalities, it is well established that the epigenome, which regulates gene expression at the level of chromatin conformation, is aberrant in AML[91]. An extensively studied example of such epigenetic aberration found in AML is DNA methylation[92]. DNA methylation is normally carried out by DNA methyltransferases (DNMT) 1, 3A, and 3B, which catalyze the conversion of cytosine to 5-methylcytosine. This is tightly linked with gene silencing, as methylation most

frequently occurs in cytosine/guanine rich sites called CpG islands. CpG islands are found in many promoters, and when methylated, the promoter region is not available to enhancer elements[47]. On the other hand, ten eleven translocation (TET) enzymes reverse DNA methylation and are associated with gene de-repression[93]. In a 2010 publication Figueroa and colleagues analyzed global DNA methylation 344 patient AML samples[94]. They concluded that the samples could be segregated into 16 groups, based on methylation signature, which were reflective of distinct cytogenetic and molecular abnormalities. How epigenetics are associated with disease establishment and progression is not fully elucidated, but because studies have shown clear involvement of epigenetics in AML, efforts have been put forth to develop epigenetics targeted therapies for the treatment of AML.

Because epigenetic modifications are reversible, they provide a rapid response and broadly acting avenue for treatment of AML. However responses to treatments thus far have been mixed. Despite approval of azacytidine and decitabine to be used for AML in combination with chemoterapy, results from clinical studies in elderly AML patients has raised controversial opinions of the treatments' efficacy[95,96]. Both, azacytidine and decitabine are DNMT inhibitors and are thought to act by reversing overall genomic hypermethylation observed in AML. A major limitation of the aforementioned clinical studies is that patient cohorts were not selected based on known genomic methylation status or methylation regulator mutations. Even apart from this limitation, because epigenetic regulation is so complex, it is difficult to imagine that an overall hypomethylation of the genome will resolve the disease. However, mutations in

epigenetic modifiers are associated with at least a quarter of all AMLs and implicated in multiple subgroups among other mutations highlighting their importance in the disease [62,80]. Because they are almost never found exclusively, it is unclear whether mutations in epigenetic modifiers such as DNMT3 have driver potential. In the mouse, *dnmt3a* loss confers a predisposition to malignant transformation, that has not advanced to leukemia which suggests that other mutations are required for disease progression[97].

Overall, epigenetics are clearly heavily implicated in AML and should be considered when developing new therapies. Because existing therapies target the epigenome broadly have limited efficacy, novel therapies should be developed to target specific epigenetic loci which are yet to be identified. Because AML epigenetics are just as inter and intra-patient heterogeneous as genetic abnormalities, single cell analyses are required to identify targetable regions specific to AML and not dysregulated in healthy cells. In addition, AML models are required which are able to verify which epigenetically dysregulated loci are required for disease initiation and progression.

1.3.5 Myeloid neoplasm disease models

AML exists as a hierarchy similar to that of a healthy hematopoietic system, where primitive cells are at the apex[67]. In order to study the hierarchy in functional terms, the appropriate assay must be selected which depends on which part of the hierarchy is to be assessed[65,67,98,99]. Currently, the repertoire of tools used to study AML includes: *in vitro* cell cultures, immortalized cell lines, colony forming unit (CFU) assays and mouse xenotransplant assays. While easy to culture, immortalized cell lines are insufficient to

capture the heterogeneity of the disease and fail to recapitulate disease features of primary AML failing to capture AML stem like properties[41]. Similarly, the *in vitro* cell culture of primary blood samples is notoriously difficult [100,101].

Hematopoietic spleen CFUs were first described by Till and McCulloch in 1961, which they soon after demonstrated to be of clonal nature; composed of cells derived from a single cell [9,10]. The first *in vitro* adaptation of this assay that allowed the observation of hematopoietic progenitor cells was developed by Ray Bradley and Donald Metcalf, where cells were plated in semi-solid medium containing methylcellulose [102]. The viscosity of the medium causes proliferating progenitors to remain in close proximity of their progeny leading to the formation of a colony unit. They were later able to employ the same assay for quantification of myelomonocytic leukemia progenitors, or AML CFU [103]. The semi-solid media has since been made more sophisticated in order to observe healthy and AML derived monocytic, granulocytic, erythrocytic, megakaryocytic, and blast colonies with the use of human hematopoietic cytokines [64,104]. The semi-solid media assay was the first *in vitro* method that allowed for the comparison of differentiation potential of healthy and AML cells and demonstrated that AML differentiation is obstructed, referred to as the AML differentiation blockade [105]. A drawback of this assay is that the absolute numbers of progenitors measured from healthy and AML are not consistent among laboratories. It is important to note that colony forming cells are not required to possess extensive self-renewal to produce colonies, and therefore the most primitive cells in the healthy or leukemic hierarchy cannot be functionally quantified and require a more sophisticated assay to test for self-renewal; the human-mouse xenograft.

Humanized mouse models have been developed in order to best mimic human biology in an *in vivo* setting and have been used extensively in hematopoietic research[106-108]. In order to support human hematopoietic xenograft growth and differentiation, highly immuno-deficient mouse strains must be used. The development SCID mice which lack functional B and T cells, was key for the aforementioned discoveries[109]. Currently even more immuno-deficient, NOD *scid* gamma (NSG) stain has been developed with a non-obese diabetic (NOD) background, containing the *scid* mutation rendering mice B and T cell deficient, and interleukin 2 receptor γ (IL-2Rγ) deletion leading to impaired cytokine signaling causing deficiency in functional natural killer (NK) cells.

In the context of leukemia, xenografting was used to estimate that approximately 1 in 250,000 AML cells is an LSC, which is capable of this reconstitution in the SCID mouse recipient[65]. Interestingly LSCs were later characterized to have the same surface phenotype ($CD34^+CD38^-CD45^+$) as healthy HSPCs [67]. To date, the LSC assay, or the ability to initiate leukemia in a mouse xenograft, is the gold standard for characterizing self-renewal and disease initiation capacity of AML and is used pre-clinically to assess novel therapeutics[110,111].

In addition to xenotransplantation assays, *in vivo* models can be used to study specific cytogenetic aberrations. Previously it was shown that retroviral expression of a leukemic fusion gene MLL-AF9 transforms human HSPCs into LSCs [112]. This data indicates that the expression of the fusion gene noted in leukemia patients is sufficient for the leukemic transformation. Most recently, the same was shown to be possible with healthy iPSCs transduced with the MLL-AF4 fusion gene. iPSC- derived blood cells were made to

engraft using 5 TFs and the expression of the MLL-AF4 was shown to cause leukemic transformation [113].

Other *in vitro* models of AML attempt to culture *ex vivo* primitive AML cells in cytokine and/or stromal cell supplemented cultures while maintaining the cell's self-renewal capacity [114]. These assays led to the identification of the leukemic long-term culture initiating cell (LTC-IC) which was able to persist in culture over 8 weeks while maintaining CFU capacity. While this identification is useful, the cultures were not successful in expanding primitive leukemic cells, which makes them an impractical candidate for high-throughput drug screens and limits their potential for clinically relevant discoveries. Overall, there is a lack of *in vitro* models which are able to entirely recapitulate the disease and capture AML genetic and epigenetic heterogeneity, while being practical enough for high through-put assays.

1.4. THE POTENTIAL OF AN IPSC MODEL OF AML

1.4.1 Cellular Reprogramming and Induced Pluripotent Stem Cells (iPSCs)

According to a theory proposed by Conrad Waddington, developmental cell fate was initially considered to be a unidirectional process in which cells starting in a pluripotent state or as he refers to it, "on top of a hill", can only travel down the hill into a differentiated state [115]. Today, it is known that differentiated cells actually have the ability to be rejuvenated back into a pluripotent state, through a process referred to as reprogramming. The first ever example of cellular reprogramming was achieved through

somatic cell nuclear transfer (SCNT), where the nucleus of a somatic cell was injected into an oocyte, which generated an embryo genetically identically to the nuclear donor cell[116,117]. This finding is important because it disproved a previous belief that stated that in order to differentiate, a cell had to permanently disable or even excise factors, making differentiation an irreversible process [118].

The use of sequence-specific DNA-binding TFs as a tool to alter cell fate was first introduced in 1987, when the expression of a single factor was sufficient to convert mouse fibroblasts into myoblasts[119]. This finding was powerful in suggesting that cells could be manipulated into a desired fate by a finite number of factors, sending scientists in a pursuit of rolling differentiated cells back up Waddington's hill towards pluripotency. Three decades later, work by Takahashi *et al.,* described an alteration of cells fate (reprogramming) of human somatic fibroblasts into induced iPSCs through ectopic expression of pluripotency associated TFs[120]. These discoveries merited the Nobel Prize, awarded in 2010, to Drs. Gurdon and Yamanaka. Generation of iPSC is now well established in the mouse and human systems through the use of retroviral expression of four key factors — octamer-binding protein 3/4 (OCT3/4), sex determining region Y-box 2 (SOX2), Krüppel-like factor 4 (KLF4) and MYC (all four together referred to as OSKM)[120-123]. Takahashi and Yamanaka's findings overturned the regenerative medicine field. Patient-derived iPSCs had the potential to circumvent ethical dilemmas associated with hESCs as well as issues of immunological incompatibility [124]. Furthermore, iPSC technology could also provide a platform to study diseases in a patient-specific manner and provide an opportunity for development of personalized medicine.

The mechanism of cellular reprogramming using OSKM has been intensely studied but not completely elucidated. Multiple models have been proposed to decipher why reprogramming is a rare event [125]. Initial models proposed that among heterogeneous somatic cells, existed a small fraction of "elite" cells capable of cell fate alteration [126]. This model had to be reconsidered when it was shown that terminally differentiated cells in hematopoietic system and the liver could be made into iPSCs [127,128]. Second, came the stochastic and deterministic models, both of which are implemented in the two-stem reprogramming process. OSKM are referred to as pioneering factors as they are able to bind chromatin which is simultaneously not accessible to other factors. The pioneering factors are able to bind and initiate a cascade of chromatin re-modelling leading to silencing of somatic and activation of pluripotency associated genes [129]. This initial step is thought to be stochastic and inefficient, occurring rarely, due to repressive epigenetic marks preserving the chromatic in a closed conformation [130]. Later reprogramming stages are viewed as deterministic and hierarchically organized, as individual cells express late reprogramming genes in order to establish and sustain their own pluripotency pathways. These models and findings aim to make sense of the extended reprogramming time and inefficiency in order to find ways to improve the process.

Parallels exist between reprogramming of somatic cells into iPSCs and acquisition of pluripotency features during cancer progression[131]. In both scenarios, global gene expression and chromatin organization patterns shift from tissue specific towards stem and early developmental programs[41,42]. iPSCs derived from specific tissues have also been shown to inherit epigenetic modifications from the tissue-of-origin[43]. These residual marks then play a role in enhanced differentiation potential back towards the tissue-type and the cell-of-origin[51]. Reprogramming and oncogenic transformation ultimately achieve the same goal of cell fate alteration, where the cells become more primitive and lose functionalities inherent to somatic differentiated cells.

iPSCs provide a unique platform that allows researchers to isolate genetic disease

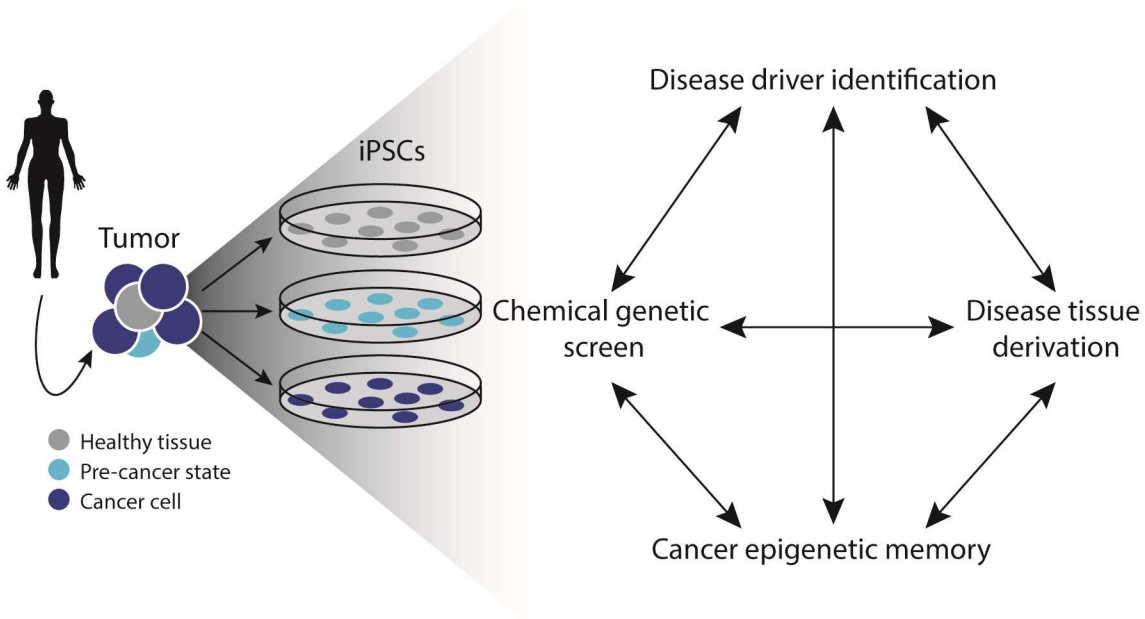

Figure 1.1 Applications of tumour derived cancer iPSCs. Through reprogramming cancer into iPSCs is uniquely possible to separate cancerous, pre-cancerous, and normal cells. These cells can then be tested side by side for genetic mutations, epigenetic components, differentiation into tissue of origin, and used in chemical genetic screens.

driver mutations in multiple ways. First, iPSCs allow for the separation of healthy cells from cancerous cells which is currently unattainable in many cancers due to a lack of definitive markers; i.e. AML and healthy blood stem cells cannot be definitively isolated due to tightly overlapping expression of potential markers[132,133]. By individually deriving iPSCs from a mixture of healthy and cancer cells from a patient, it is possible to derive healthy iPSCs as well as cancerous iPSCs which harbour genetic mutations in the same genetic background. This allows for a direct comparison to identify true driver mutations and potential cancer-specific targets. The side-by-side comparisons of healthy and cancerous iPSCs from the same individual have successfully been used in the past to identify driver mutations in oesophageal squamous carcinoma[132,133].

iPSCs derived from cancer patients enable the resources and cell line stability required, to apply genetic modifications. The iPSCs can then be subjected to genetic editing to repair or introduce mutations allows researchers to further elucidate biological and phenotypic consequences from true disease drivers, which can be isolated from the cancer associated passenger mutations as well as the patient's genetic background. Further, gene edited iPSCs can be differentiation to the tissue of origin providing a phenotypic read-out of cancer progression/regression caused by the genetic modification. Recently, gene editing technology has demonstrated that a dCas9 fusion with DNMT3 or TET1 can be used as a method to edit the methylation status of specific enhancer and promoter regions, which could potentially be used to interrogate AML relevant methylation in iPSCs[134,135].

Critically, iPSCs retain the ability to differentiate to somatic tissues when directed with appropriate cues. This allows for the step-by-step modelling of cancer progression and diseased tissues, as well as drug screening approaches. These are some of the current applications of cancer iPSCs summarized in Figure 1, which are only limited by the breadth of iPSCs developed and directed differentiation protocols to achieve disease-relevant somatic tissues.

1.4.2 Factors Affecting Cellular Reprogramming

Although the ability of cells to reprogram holds great potential, this process is inefficient [136]. Many efforts have been put forth to uncover enhancers of reprogramming in order to maximize reprogramming efficacy. The most common methods used in combination with OSKM are cell cycle regulators or epigenetic modifiers. Using cell-cycle regulators as a strategy to improve reprogramming is easily justified as one of the original reprogramming factors, MYC, itself is a well-established promoter of cell proliferation. The most studied and drastically effective regulator is tumor suppressor p53, which is often referred to as the guardian of the genome. It appears to have a binary role in reprogramming where p53 expression inhibits reprogramming, whereas p53 suppression enhances reprogramming efficiency[137-140].

Since reprogramming is known to remodel the epigenetic landscape to promote transcription of pluripotency associated genes, and post-translational histone modification enzymes are implicated in the process of reprogramming, there has been interest in using epigenetic modifiers to enhance reprogramming efficiency[141,142]. This includes epigenetic

histone marks as well as modification of the DNA itself. Modifiers can work directly or by promoting intrinsic epigenetic enzyme activity. It has been reported that DNA methyltransferase and histone deacetylase (HDAC) inhibitors are able to tremendously enhance reprogramming efficiency [143]. In particular, valproic acid (VPA) is able to improve the reprogramming efficiency of mouse embryonic fibroblasts (MEFs) by more than 100-fold. Overall, reconfiguration of the chromatin state by small molecule compounds has been shown to have positive and negative effects of reprogramming efficiency and their long term effects are not well defined.

Patient-derived iPSCs have been successfully used as disease models and are of particular interest for perusing personalized medicine [144]. However, reports have demonstrated that primary cancer patient samples are especially difficult to reprogram[145,146]. The derivation of these lines is paramount as early success has been achieved, in particular, understanding the pathogenesis of human hematopoietic malignancies and identifying novel drugs. Therefore, efforts to better understand the reprogramming process in healthy and disease contexts should continue in parallel.

1.4.3 Challenges of Reprogramming Primary Cancers

Despite the plethora of potential applications that iPSCs have in the study of cancer, there have been few reports of successful generation of iPSCs from primary human tumours[145,147]. There are multiple examples of the reprogramming of already immortalized cancer cell lines[140,148]. But as discussed previously, cancer cell lines do not

possess stemness signatures whereas primary samples do, and they also do not contain tumour heterogeneity found in primary samples.

Before TF-based reprogramming became available, researchers attempted to utilize SCNT for reprogramming cancer cells[145]. It was shown that very few cancer epigenomes can be reprogrammed into pluripotency. In a series of robust experiments, where nuclei from melanoma, leukaemia, lymphoma, medulloblastoma and breast cancer cells were transplanted into enucleated oocytes, iPSC lines could only be derived from melanoma cells[149]. Interestingly, oocytes transplanted by all cancer nuclei gave rise to viable pre-implantation embryo and normal blastocysts that all failed to produce viable embryos. This work indicates that the oocyte environment transiently dominates over cancer phenotypes and it is only later in development that cancer genes become active. ESC *in vitro* culture mimics features of the oocyte. This may explain why iPSCs derived from primary cancers display a normal phenotype in the pluripotent state – until subjected to differentiation[150-152].

Table 1.1 Reprogramming of human blood cancers into iPSCs

Primary cancer	Cancer abbrev.	Patient sample iPSC of patient samples attemplted	Reference
Acute myeloid leukemia	AML	1 of 16	Lee et al. 2017
Acute myeloid leukemia	AML	2 of ?	Chao et al. 2017
Acute myeloid leukamia	AML	1 of 3	Kotini et al. 2017
B-cell acuty lymphoid leukemia	B-CLL	0 of 3	Munoz-Lopez et al. 2016
Myeloproliferative disorders	MPDs	2 of 2	Ye et al. 2009
Myelolydplastic syndrome	MDS	3 of 5	Kotini et al. 2017
Myelolydplastic syndrome	MDS	8 of 8	Hsu et al. 2019
Chronic myeloid leukemia	CML	2 of 2	Bedel et al. 2013
Chronic myeloid leukemia	CML	2 of 2	Kumano et al. 2012
Chronic myeloid leukemia	CML	1 of 16	Hu et al. 2011
Pancreatic duct adenocarcinoma	PDAC	4 of 9 iPSC-like	Kim et al. 2013

iPSCs are valuable in studying complex hematologic disorders[153,154], and so several attempts have been made to reprogram a variety of leukaemia samples into iPSCs often with very limited success (Table 1). There are three examples of successfully reprogrammed chronic myeloid leukaemia (CML) samples[154]. Others were able to derive iPSC from myeloproliferative and myelodysplastic syndromes (MPD and MDS)[145,154,155]. MDS samples were able to reprogram, potentially signifying incomplete transformation of the cells and the presence of a reversible state which was permissive to reprogramming[156]. Despite these successes, attempts to derive iPSCs from acute leukaemia patients have been largely unsuccessful (Table 1). Two studies that reported how many AML patient samples were attempted and how many were successfully reprogrammed, were able to derive iPSCs lines from two patient samples out of nineteen samples attempted[146]. Another study reported the reprogramming and characterization of AML-iPSCs from 2 patients without disclosing if more samples were attempted[147]. A thorough study trying to reprogram samples from 3 patients with B cell acute lymphocytic leukaemia (B-ALL), used and a wide variety of reprogramming approaches. Unfortunately, they were not able to derive a single iPSC line containing leukemic aberrations, deriving iPSCs only from healthy cells within the samples.

In another tissue system, Kim *et al.,* reinforced the observation of the difficulty of obtaining iPSCs from patient cancer, while also exemplifying the utility that cancer-iPSCs offer. “iPSC-like” lines were generated from pancreatic ductal adenocarcinoma (PDAC) of a single patient that were able to model an early pancreatic cancer which was never done before, and identify a gene network activated by intermediate stage lesions.

They are referred to as iPSC-like because they require continuous low dose expression of TFs through a doxycycline-inducible system to remain viable iPSCs. This study is limited to only one cancer-iPSC line because of the difficulty of reprogramming primary pancreatic ductal adenocarcinoma.

As far as the author is aware, the studies exemplified here and summarized in Table 1 are the only known reports of primary human cancer reprogramming attempts. Identifying whether the refractory nature of cancer cells to reprogramming is a technical or a biological phenomenon remains an area of great interest.

1.4.4 Reprogramming of AML

To date, although primary AML reprogramming has been implicated in multiple studies, it is apparent that deriving AML-iPSCs is challenging[145,154,156]. Within the three aforementioned studies, only 4 patient AML cases have been reprogrammed to iPSCs (Table 1), 3 of which harbor MLL gene family network alternations which represent only 2% of adult AML patients [63]. Interestingly, AML-iPSCs have been shown to reacquire leukemic properties, including AML reconstitution in xenograft model, as well as methylation and gene expression patterns indication that AML-iPSC can be used to successfully model genetic, epigenetic and functional components of AML samples[156]. The study by Kotini and colleagues used AML iPSCs to model myelodysplastic syndrome (MDS) to AML transition, however, a limitation of this work was that the progression of disease stages iPSCs were derived from different patients[154]. Due to patient heterogeneity, it is highly unlikely that each patient progresses into AML in the

same manner, or even due to the same mutations and dysregulation patterns. Hurdles faced by reprogramming of cancers may be due to aberrant methylation states of cancer cells as reprogramming relies on the ability of pioneering factors to bind to DNA[157]. On the other hand, genetic abnormalities, cannot be taken out of consideration in regard to lack of reprogramming, but it non-cancerous genetic abnormalities are compatible with iPSC generations implicating that genetic mutations are compatible with reprogramming[123]. These studies exemplify that the hurdle of deriving AML-iPSC, holds a promising avenue to study AML *in vitro*, with the ability to capture the broad spectra of the disease with multiple AML-iPSCs lines.

1.5 STUDY RATIONALE

As discussed in great detail above, iPSCs have long proven to be useful tools that allow researchers to elucidate molecular and physiological phenotypes of disease while also providing a drug screening avenue[37,38]. In many cases, iPSCs are advantageous in their ability to model hard to obtain tissues. At glance this may not appear to be the case for blood malignancies. But, although reasonably available, AML samples are heterogeneous mixtures of healthy cells, and variable leukemic clones. This makes it difficult to draw definitive conclusions about leukemic clone function independent of other clones or healthy cells. Due to these challenges, despite many identified mutations, a very small percentage of known leukemic mutations have been functional proven to be disease drivers[81-85]. AML-iPSCs would allow for the separation of disease into clones *in vitro* allowing for functional interrogation of disease drivers.

Due to these challenges that primary samples pose, treatment of AML has been stagnant for over 50 years and new approaches are required to identify selective therapies. iPSCs are known to retain an epigenetic memory and have been used in the past to study disease where pathogenesis is tightly associated with epigenetic dysregulation[50,55,56]. In AML, iPSCs provide a unique opportunity to compare the epigenetic memory inherited from leukemic cells and compare them to isogenic healthy signatures. This memory may also be able to capture tissue-of-origin signatures from cancers of the same lineage and provide viable targets for more than one cancer type[40-42]. Functional studies are still required to validate the fairly new concept that stemness and its role in cancer, but nonetheless it would be captured by the iPSC platform.

Reprogramming cancers has been challenging in the past (Table 1.1). Identifying a way to circumvent the refraction to reprogramming in AML would be beneficial in multiple ways. Firstly it would provide a novel and multi-dimensional model of AML, which would allow to distinguish driver mutations, and provide an avenue to identifying selective therapies. Secondly, because many cancers seem refractory to reprogramming, elucidating the mechanism of unresponsiveness to pioneering factors may help answer fundamental questions about pluripotency and cancer.

1.6 STUDY HYPOTHESIS

It is hypothesized that a combination of FACS and xenografting strategy will allow for enhanced reprogramming of AML-iPSCs capturing inter and intra heterogeneity of AML patients. Further, AML-iPSCs will allow *in vitro* culture in the pluripotent state and recapitulate AML through hematopoietic differentiation.

1.7 STUDY OBJECTIVES

This thesis aim to:

1. Identify whether xenografting of AML patient samples is able to select for specific clones and enhance reprogramming of AML within AML patient samples. This was determined by performing reprogramming of AML samples *de novo* and post xenografting into NSG mice. Xenografts were hypothesized to give rise to a subset of AML clones which have LSC capacity.

2. Determine if reprogramming potential of AML is dependent of differentiation status using FACS sorting based on CD34 and CD33. Because healthy reprogramming is known to be enhanced in blood stem and progenitor cells, we hypothesized that CD34+ AML cells will be more susceptible to reprogramming. Because previous studies attempted to use myeloid CD33+ cells to reprogram AML, we hypothesized that CD33+ reprogramming will allow for the selection of AML cells for reprogramming resulting in derivation of AML and not normal iPSCs.

To test this, AML samples were be fractionated into four populations based on expression of CD34 and CD33 prior to reprogramming to identify which compartments give rise to AML versus normal iPSCs.

3. Identify if derived AML-iPSCs are able to model disease heterogeneity and AML pathophysiology thorough hematopoietic differentiation. To determine this, all derived clones will be assessed for leukemic aberrations using approaches that capture the whole

genome when possible. This allowed us to identify if AML-iPSCs were generated and if a variety of clones from patient samples can be derived. In order to assess the differentiation capacity of AML-iPSCs, cells and their normal counterparts were subjected to a differentiation assay, reading out hematopoietic phenotype markers by flow cytometry and progenitor function by CFU assay. Because previous studies have shown that AML-iPSCs can be refractory to hematopoietic differentiation, we subjected AML and normal-iPSCs to cardiomyocyte differentiation to identify if the differentiation block is hematopoietic or mesoderm specific.

2.0 MATERIALS AND METHODS

2.1 Primary Patient Samples

Primary BM peripheral blood (PB) and leukapheresis (Leuk) samples were obtained from human AML patients, and mobilized PB (M-PB) and healthy donors, in accordance with the Research Ethics Board-approved protocols at McMaster University. Mononuclear cells were isolated using density gradient centrifugation Ficoll-Paque Premium (GE Healthcare, Piscataway) followed by ammonium chloride red blood cell lysis (STEMCELL Technologies). Cells were cryopreserved in 30% fetal bovine serum with 10% dimethyl sulfoxide until use.

2.2 Reprogramming of Primary Samples

Reprogramming of AML and MPB samples was carried out in reprogramming media consisting of StemSpan SFEM II (STEMCELL Technologies) supplemented with

100 ng/mL stem cell factor (SCF), 100 ng/mL Fms-related tyrosine kinase 3 ligand (FLT3-L), and 20 ng/mL thrombopoietin (TPO), all from R&D Systems, 8µg/mL polybrene (Sigma-Aldrich) and 0.75 µM StemRegenin 1 (SR1) (STEMCELL Technologies). CytoTune –iPSC 2.0 Sendai Reprogramming kit (ThermoFisher) was used to reprogram $2.0x10^5$ cells/24-well of an ultra-low attachment plate, in 200 µL reprogramming media. Cells were transduced over 48 hours with three vectors containing KOS, c-MYC, and KLF4 at recommended MOI of 5:5:3, after which the virus and reprogramming media were removed, and each well was plated into a 6-well coated with 0.5% gelatin (Sigma-Aldrich) and pre-seeded with $2.0x10^5$ irradiated mouse embryonic fibroblasts (iMEFs). Cells were plated onto iMEFs in LIF2i media, which is PSC media: DMEM/F12 (1:1) (Gibco), 20% KOSR (Gibco), 1x NEAA (Gibco), 1 mM L-glutamine (Gibco), 0.1 mM β-mercaptoethanol, supplemented with 1 µM MEK inhibitor PD0325901 (Stemgent), 3µM GSK3 inhibitor CHIR 99021 (Sigma-Aldrich) and 10 ng/mL leukemia inhibitor factor (LIF) (Millipore). Cells were fed every other day, and colonies emerged 18-21 days post transduction. Colonies were stained with TRA-1-60 live FITC (Millipore) and individually picked into separate wells and expanded. At passage 2, all clones were transitioned into PSC media supplemented with 8 ng/mL basic fibroblastic growth factor (bFGF) (Corning). Cells were cryopreserved in DMEM:F12 (1:1), 30% KOSR, and 10% DMSO.

2.3 Patient Derived Xenografts

All experimental protocols were approved by the Animal Research Ethics Board of McMaster University. AML samples were thawed and CD3 depleted using EasySep Human CD3 Positive Selection Kit II (STEMCELL Technologies) and EasySep Magnet (STEMCELL Technologies). NSG mice were irradiated at 315 cGy 24-hours prior to transplant. 5-15x10^6 cells were IV injected. BM aspirates were performed to identify human chimerism at time points prior to harvesting. BM was harvested from legs and spines 6-12 weeks post engraftment.

2.4 Flow Cytometry

All antibodies used for flow cytometry were titrated in order to generate signal based populations consistent with those demonstrated by the antibody manufacturer. All extracellular staining was performed in PBS, 3%FBS, 0.5mM EDTA (PEF) where 100k/100µL cells of interest were incubated with antibodies for 30 min at 4°, washed with 10 volumes of PEF and then stained with 7-amino actinomycin D (7-AAD, Becton Dickinson) at 1:50 to exclude nonviable cells. The following antibodies were used at 1:100 unless otherwise specified as follows: CD45-v450, SSEA3-PE, TRA-1-60-AF647 at 1:1000, CD34-FITC or APC (1:200), mCD45-FITC, CD33-PE or APC, CD3-PE (Beckman Coulter), all from BD Biosciences unless otherwise specified. Live cells were stained for 30 min at 4°C in PBS with 3% FBS and 1 mM EDTA, washed, and analyzed with an LSRII or Aria II flow cytometers (BD Biosciences).

Intracellular staining was achieved using Cytofix/Cytoperm (BD Biosciences) and Perm/Wash buffer (BD Biosciences) by following the manufacturer's instructions. Prior to performing fixation and permeabilization, cells were stained with LiveDead discrimination dye (Life Technologies) for 30 minutes at 4° in the dark. Intracellular staining was performed in Perm/Wash buffer overnight at 4° in the dark and washed with 10 volumes of PBS prior to analysis. Flow cytometry analysis was performed with the FlowJo 10.2 (FlowJo, LLC).

2.5 Fluorescent in Situ Hybridization (FISH)

Primary AML cells and iPSCs were synchronized by adding 0.1 µg/mL KaryoMAX Colcemid (ThermoFisher) to cell media for 3-4 hours. Cells were collected as single cell suspension. iPSCs were dissociated using cell dissociation buffer (Gibco) for 5-10 minutes at 37°C. Cells were incubated in a hypotonic solution of 0.075 M KCl for 15 minutes at 37°C. Cell were then fixed using 3:1 methanol: glacial acetic acid and pipetted onto a glass slide. Slides were dehydrated using a sequence of 70,80, and 100% EtOH and pre-warmed at 37°C. Probe was also pre-warmed to 37°C. Sample and probe were simultaneously denatured on a hotplate at 75°C for 2 minutes and then placed at 37°C overnight. Slides were then washed, stained with DAPI, visualized and scored manually. MYH11/CBFB probe (Empire Genomics), and PML/RARα translocation, dual fusion probe (Cytocell) were used and scored as recommended by manufacturer.

2.6 Cytogenomics

Karyotyping and comparative genomic hybridization using Cytoscan HD Array (Thermo Fisher) were performed by The Centre for Applied Genomics (TCAG), at The Hospital for Sick Children (SickKids), Toronto, Ontario, Canada. Cytoscan analysis was performed using Chromosome Analysis Suite (NetAffx 33.1, h19) and default settings.

2.7 ddPCR

DNA was isolated from iPSCs using DNeasy Blood and Tissue Kit (Quigen). Primers were designed using RefSeq transcripts and https://mutalyzer.nl/position-converter. Droplets were generated and PCR performed by TCAG, Genetic Analysis Facility, Toronto, Ontario, Canada.

2.8 Teratoma Assay

AML iPSCs were collected and injected as clumps into NOD/SCID mice intratesticular injection (IT). At 8 weeks, teratomas were harvested, fixed with 4% paraformaldehyde, embedded in paraffin, and processed for H&E staining. Sections were imaged using Aperio ImageScope (Leica).

2.9 Hematopoietic Differentiation

PCSs were differentiated using suspension embryoid bodies (EBs) as previously described [158]. Briefly, iPSCs were rolled and cultured in ultra-low wells in suspension in EB media: KO-DMEM, 20% FBS (Hyclone), 1% NEAA, 1 mM L-glutamine, 0.1 mM β-

mercaptoethanol. EB media was supplemented with 300 ng/mL SCF (R&D Systems), 300 ng/mL FLT-3L (R&D Systmes), 10 ng/mL IL-3 (Miltenyi), 10 ng/mL IL-6 (R&D Systems), 50 ng/mL G-CSF and 25 ng/mL BMP-4 and changed on days 1, and every 5th day. EBs were harvested and dissociated on day 15 and day 21 using Collagenase B (Sigma-Aldrich) at 37°C for 2 hours, then in cell dissociation buffer (Gibco) at 37°C for 10 minutes, then manually titrated and filtered to obtain a single cell suspension.

2.10 CFU Assay

Primary AML and differentiated iPSC cells were plated at $5.0x10^3$-$2.0x10^4$/0.5 mL into Methocult H4334 (STEMCELL Technologies). Cells were incubated at 37°C for 14 d and manually scored. Each CFU well represents an independent biological assay, as input cells and MethoCult formulations were individually prepared for testing in single wells.

2.11 Cardiac Differentiation

Cardiomyocyte differentiation was performed as previously reported [159]. Briefly, 10,000 cells were plated per U-bottom well in MEF-conditioned medium containing 8ng/mL bFGF and 10 μM Y-27632 ROCK1 and ROCK2 inhibitor. At day 3 media was replaced with DMEM/F12 (1:1), 20% FBS (Hyclone), 1mM L-glutamine, 1mM NEAA, 0.1 mM β-mercaptoethanol, and 50 μg/mL ascorbic acid. Aggregates were seeded on gelatin-coated wells at day 7 and then differentiation medium was changed every 2 days. Beating foci were counted and staining for alpha-smooth muscle actin (ThermoFisher) was done from 21 days.

2.12 Statistical analysis

Data are represented as means ± SEM. Prism (6.0c, GraphPad) software was used for all statistical analyses, and the criterion for statistical significance was $P<0.05$. Statistics are described in figure legends.

3.0 RESULTS

3.1 Reprogramming primary AML

3.1.1 Determination of Primary AML Samples

Primary AML samples were selected with the goals of 1) being able to test derived iPSCs for the presence or absence of a cytogenetic or molecular abnormality found in the primary sample, 2) being able to identify whether a relationship exists between leukemic stem cells (LSCs) and reprogramming potential, 3) being able to test whether reprogramming potential depends on the differentiation stage of leukemic cells and 4) capturing AML heterogeneity. To meet the first goal, the Stem Cell and Cancer Research Institute (SCC-RI) live stem cell bank was filtered to select samples which had known molecular or cytogenetic abnormalities that can be detected by fluorescent *in situ* hybridization (FISH), karyotyping, copy number array or PCR methods available. Samples were selected from all three risk-stratification groups to capture a wide variety of patient heterogeneity including complex karyotypes, isolate -7, isolate inv(3), CBFβ-MYH11 and PML-RARα gene fusions and others shown in Table 3.1.

Samples with measurable aberrations were then binned as LSC positive or negative, determined by engraftment into immune deficient mice. Some AML samples had existing annotation on their LSC status performed by Drs Allison Boyd, Lili Aslostovar. Samples without LSC status information were tested by engrafting ten million, T-cell (CD3) depleted, mononuclear cells (MNCs) into at least five sub-lethally irradiated NSG mice by intravenous (IV) injection. BM and spleens were assessed for human engraftment using pan-hematopoietic marker CD45, myeloid marker CD33, B-cell marker CD19 and stem and progenitor cell marker CD34 after 8 weeks post injection. Human chimerism was assessed by human pan-hematopoietic marker, CD45. Myeloid grafts were identified by exclusive expression of CD33 and considered LSC+. Multi-lineage grafts displayed expression of CD33 and CD19 and considered to be healthy and LSC- (Figure 3.1). Samples that were LSC positive were candidates for reprogramming both, the d*e novo* and xenografted sample with the goal of potentially identifying clones responsible for leukemic repopulation.

Because reprogramming is enhanced in the primitive compartments of healthy hematopoiesis, we thought to test whether the same applied to AML. In order to test this hypothesis we selected samples with substantial CD34 positive and negative populations that could be directly compared. Additional characteristics such as tissue source, cytogenetic and molecular abnormalities, presence of progenitors tested by CFU, percentage of CD34+ cells are summarized in Table 3.1.

Table 3.1 AML samples for reprogramming. AML samples 1-3 were later determined to give rise to AML-iPSCs. Samples 4-7 and APL sample 10 gave rise to exclusively healthy iPSCs and samples 8-10 did not give rise to any iPSCs.

								iPSC clones	
AML #	ID	Source	Genetic/molecular abnormality	LSC	CFU	CD34+ (%)	iPSC	normal	AML
1	A374.1	PB	46,XX,add(3),der(3),del(5), del(7), add(18), -20, -22,+mar1,+mar3	Y-myeloid	Y	80.3	13	0	13
2	A422	PB	45,XX,-7[19] /46,XX[1]	Y-myeloid	Y -myelo & erythro	88.0	6	0	5
3	13814.1	BM	43~46,XY,del(15)(q11.2q15)[cp5]/ 46,XY,+mar1[cp2]/46,XY[11]	Y-multiligeage	N	60.1	4	2	2
4	16150	BM	CBFbeta/MYH11 inv(16)	Y-myeloid	Y -myelo & erythro	72.8	12	12	0
	16150	PB	CBFbeta/MYH11 inv(16)	Y-myeloid	Y	77.0	0	0	0
5	A151.1	PB	46,XY,inv(16)(p13.1q22)[20]/47,sl,+8[5]	Y-myeloid	Y	30.0	3	3	0
6	15328	BM	45~46,XX,der(X)?t(X;11)(q28;q12)[2],-2[3],del(3)(q11.2)[3],-5[2], add(7)(p11.2)[3],-17[3],+3~5mar[3][cp3]	Y-multiligeage	N/A	85.8	6	6	0
7	A295.1	PB	trisomy i21	Y-multiligeage	Y	67.0	1	1	0
8	16158	BM	45, XX, del(5), -7	Y-multiligeage	Y	92.8	0	0	0
9	16626	BM	inv (3)	Y-myeloid	Y	45.0	0	0	0
10	19447	PB	complex, isolated +8	Y-multiligeage	N/A	64.0	0	0	0
11	A320	BM	PML, RARA translacation (APL)	N/A	N/A	0.1	5	5	0

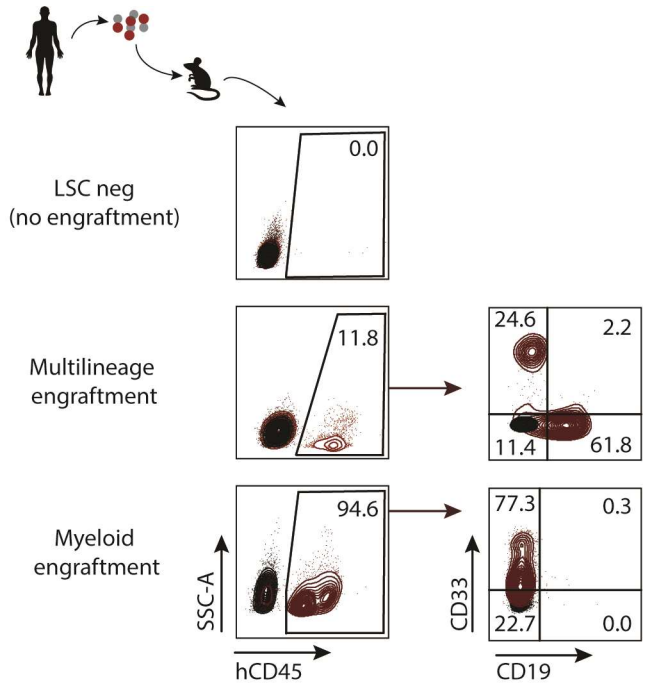

Figure 3.1 LSC determinations of AML samples. AML samples were engrafted into NSG mice which were assessed for human chimerism by human CD45. Exclusive myeloid grafts were identified by exclusive expression of CD33 and considered leukemic. Multi-lineage grafts displayed expression of CD33 and CD19 and considered to be healthy.

3.1.2 Primary AML Reprogramming Strategy

In previous reports from our and other groups, AML has been notoriously difficult to reprogram (Table 1.1). From past reports, AML either does not produce any reprogrammed colonies or it produces colonies that do not contain leukemic aberrations and likely arose from healthy progenitors. In order to increase the chances of reprogramming AML, sample reprogramming was carried out in sorted populations. Samples were fluorescently sorted into CD45+CD34+CD33+, CD45+CD34+CD33-, CD45+CD34-CD33+, and CD45+CD34-CD33- populations when present in the sample (Figure 3.3). CD45 is a pan hematopoietic marker used to exclude stromal cells in BM samples and was used across all AML samples. Because healthy hematopoietic cells are known to have reduced reprogramming potential with increased differentiation, CD34 was used to separate cells into more primitive (CD34+) and terminally differentiated (CD34-) cell fractions [160]. It was hypothesized that the AML hierarchy would behave similarly to healthy, with CD34+ cells having increased reprogramming potential. Cells were also fractionated by CD33 in order to select myeloid cells, which were hypothesized to be more likely to give rise to AML-iPSCs.

In addition to cell sorting, four select samples were engrafted into NSG mice to assess reprogramming of d*e novo* versus patient derived xenografts. This was done because the xenograft is reconstituted from rare LSCs found in the sample. This step would allow for the filtration of not just exclusively leukemic cells, but also the selection of more aggressive clones that are able to initiate the disease in the mouse model. In

addition, some reports speculated that AML cells do not reprogram due to the lack of division, and xenografts are a strategy to increase cell leukemic cell proliferation[146] .

Sorted cells from xenografts and d*e novo* samples were reprogrammed using three Sendai virus (SeV) reprogramming vectors containing KOS (KLF4, OCT3/4 and SOX2), c-MYC and KLF4, as it has been shown to have increased success in healthy and leukemic reprogramming by previous studies [145]. Findings by Lee *et al.,* also suggested that reprogramming was enhanced in naïve media, supplemented with Leukemia Inhibitory Factor (LIF), and 2 small molecule inhibitors (2i) of MEK and GSK3 [145,161]. Thus, all samples were seeded onto irradiated mouse embryonic fibroblasts (iMEFs) were maintained in LIF2i media post-transduction.

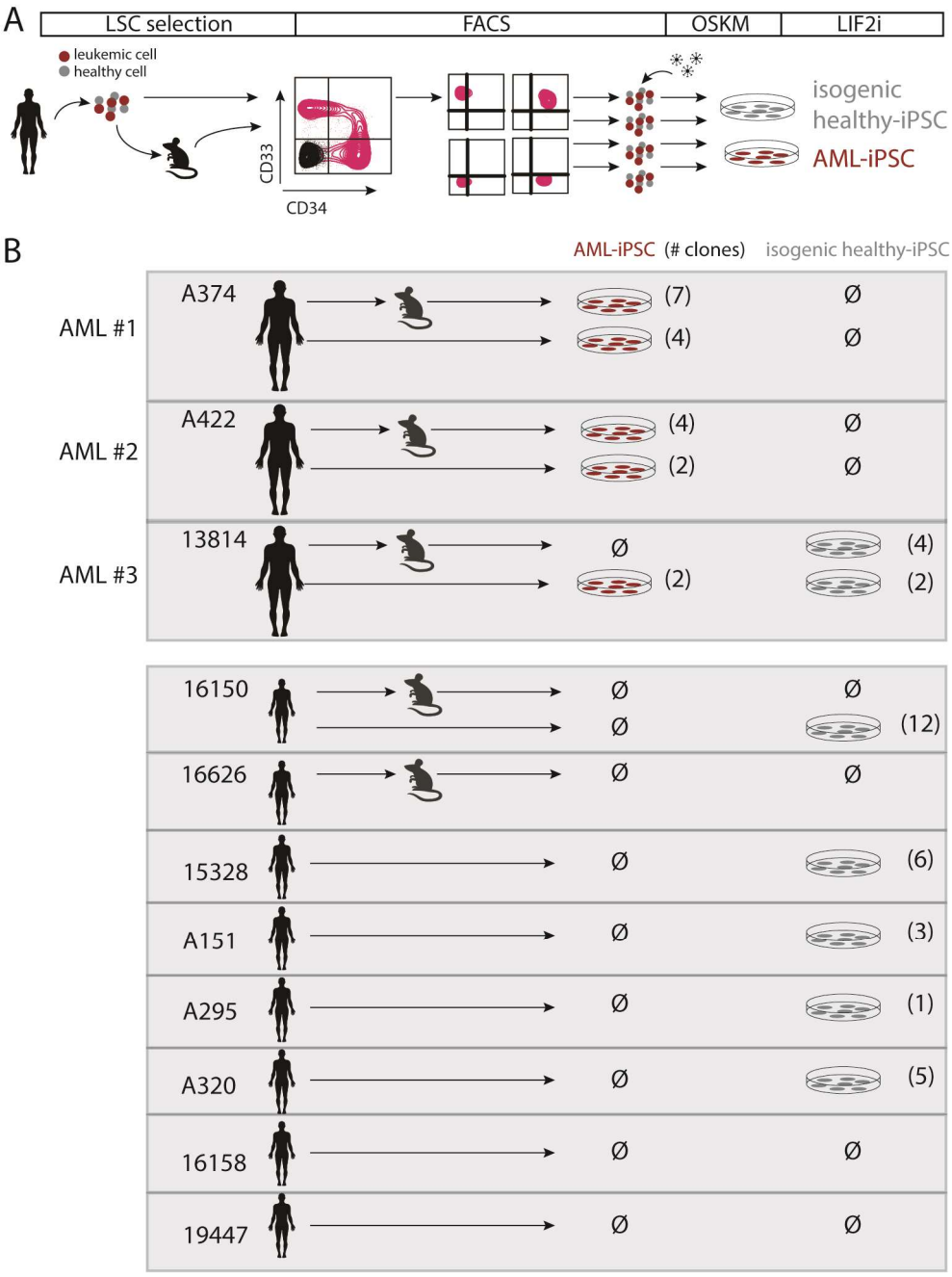

Figure 3.2 Reprogramming Strategy of Primary AML. (A) Primary AML samples were sorted directly or from xenograft on CD45, CD34 and CD33. Sorted populations were transduced with SeV virus and maintained in LIF2i for 18-21 days until iPSCs arose. (B) Summary of AML reprogramming and the leukemic status of derived iPSCs.

3.1.3 Determination of Genetic Status of Derived iPSCs

Isolated iPSCs colonies were treated as independent clones, expanded and cryopreserved at multiple early passages. To identify if leukemic or healthy clones were derived, the genetic status of each clone was tested for the presence of known aberrations.

From eleven reprogrammed AML samples, clones from three were found to capture aberrations. Thirteen AML 1 clones were tested by G-band karyotyping and determined to all have inv(3) and del(7), along with other aberrations detected in the primary sample (Figure 3.3). Some clones had identical karyotyping patterns and thus likely originated from the same leukemic clone, for example clones AML1-2 and AML1-4 (Figures 3.3 and 3.4). Thus, seven unique clones were captured from AML-1 (Figure 3.4). Five AML 2 clones were derived and tested using HD CytoScan Array and determined to capture del(7) and other aberrations from the primary sample (Figure 3.5). Overall, 3 unique AML-2 clones were derived. Four AML-3 clones were derived and tested using G-band karyotyping. Two clones were found to have aberrations, and two clones were found to have a normal karyotype Figure 3.3). The normal clones derived from AML 3 are a healthy isogenic control to compare to the leukemic clones.

FISH was used to test AML samples 5, 4 and 11. All twelve clones derived from AML 5 and three clones derived from AML 5 did not contain inv(16) tested using the CBFβ-MYH11 fusion probe (Figure 3.6A). Nine clones derived from AML 11 did not contain PML-RARα gene fusion (Figure 3.6B). AML 6 clones were tested by karyotyping and all six clones were found to have a normal karyotype (Figure 3.7). One

AML 7 clone was tested using digital droplet PCR (ddPCR) for the presence of an isodicentric chromosome 21 and was found to be normal (Figure 3.8).

Figure 3.3 AML 1 and 2 gave rise to AML and normal iPSCs. Using G-banding, it was found that AML1 clones 1-12 were found to have chromosomal aberration seen in primary AML 1. From primary AML 3, clones 1-2 were found to have leukemic aberration, whereas clones 3-8 were found to be normal by G-band karyotyping

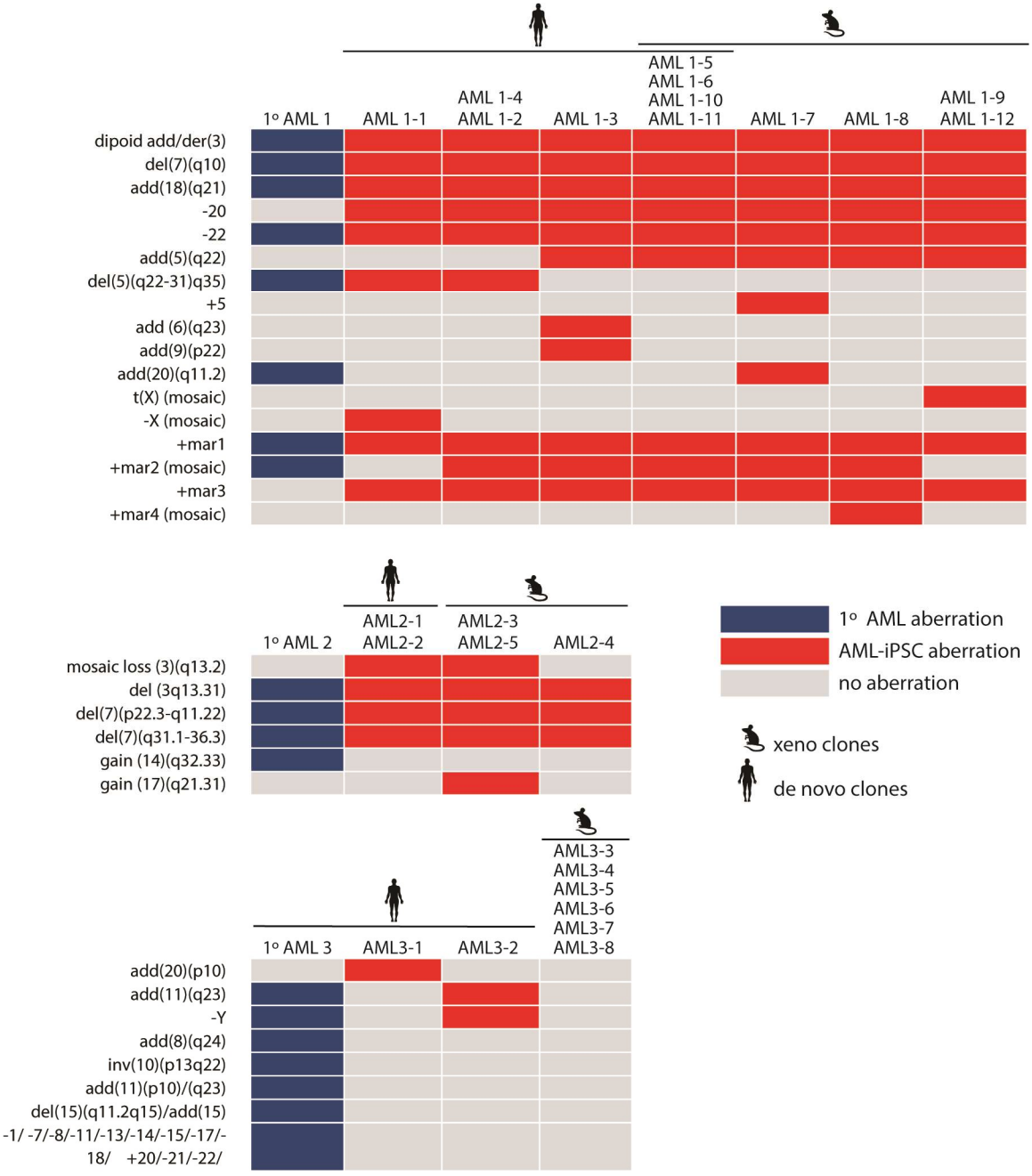

Figure 3.4 Summary of aberrations in AML and normal iPSCs form AML samples 1-3. Primary AML aberrations are shown in blue, AML-iPSC aberrations are shown in red. Clones that showed identical aberration are stacked in columns. Mouse and human silhouettes identify which clones were derived from xenografted and *de novo* samples respectively. Some identical clones in sample 1 were derived from xenografted and *de novo* sample shown by overlap.

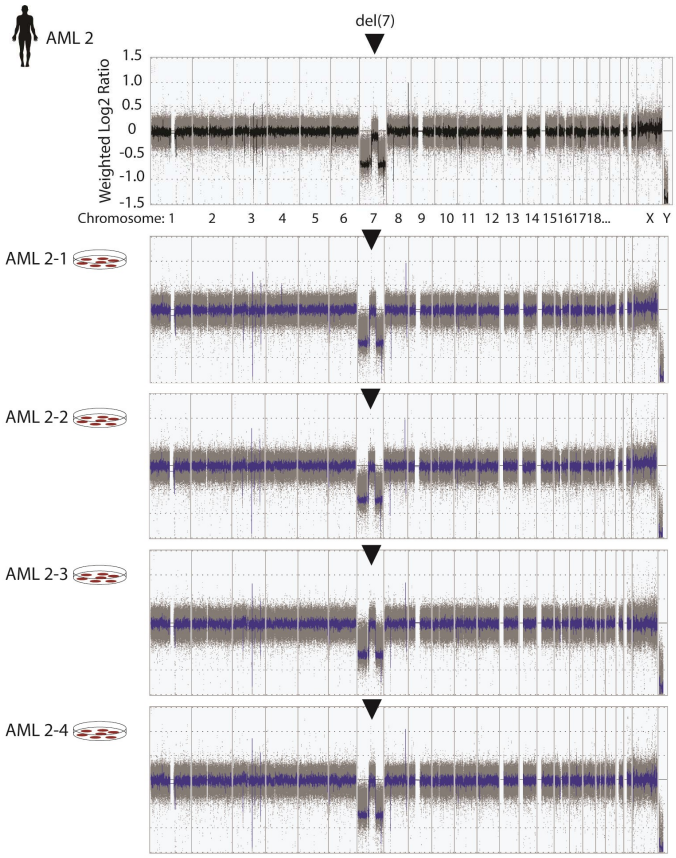

Figure 3.5 AML 2 gave rise to AML-iPSCs. Gene copy number was assessed in AML 2 and derived iPSCs by Affymetrix HD CytoScan microarray and analyzed using ChAS software. All 4 clones contain del(7) present in primary AML 2.

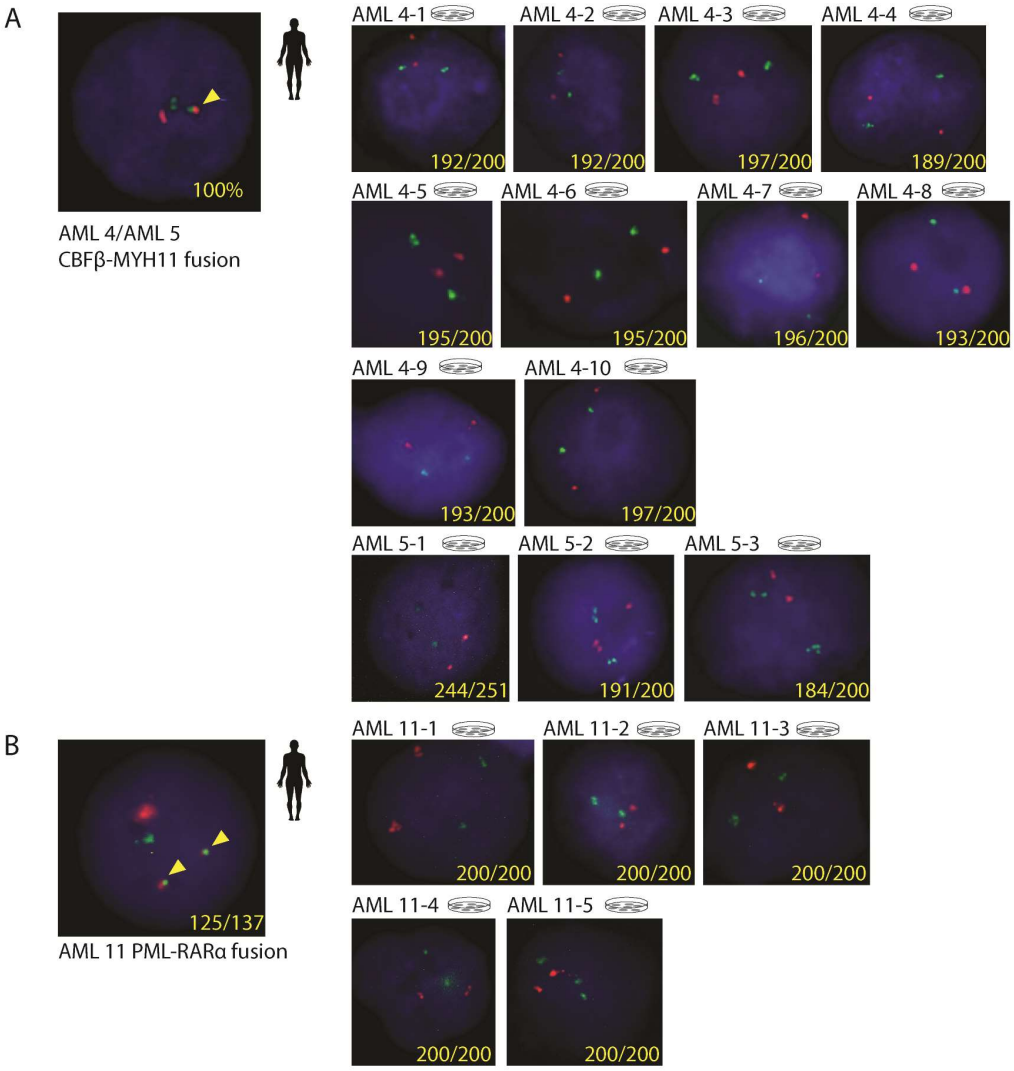

Figure 3.6 AML 4, 5 and 11 iPSCs determined to be normal. (A) FISH was performed on AML samples 4 and 5 using CBFβ-MYH11 hybridization probe. When the gene fusion is present the foci pattern is 1 green, 1 red, 1 red-green. All clones derived from AML 4 and 5 show no gene fusion with a 2 green, 2 red pattern. (B) FISH was performed on AML 11 and derived iPSCs using PML-RARα dual hybridization probe. Primary AML demonstrates gene fusion foci pattern of1 green, 1 red and 2 green-red. All clones demonstrate no gene fusion with the normal, 2 green, 2 red foci pattern. Yellow arrows point togreen-red foci. Yellow numbers indicate the number of nuclei with the shown pattern out of the nuclei scored.

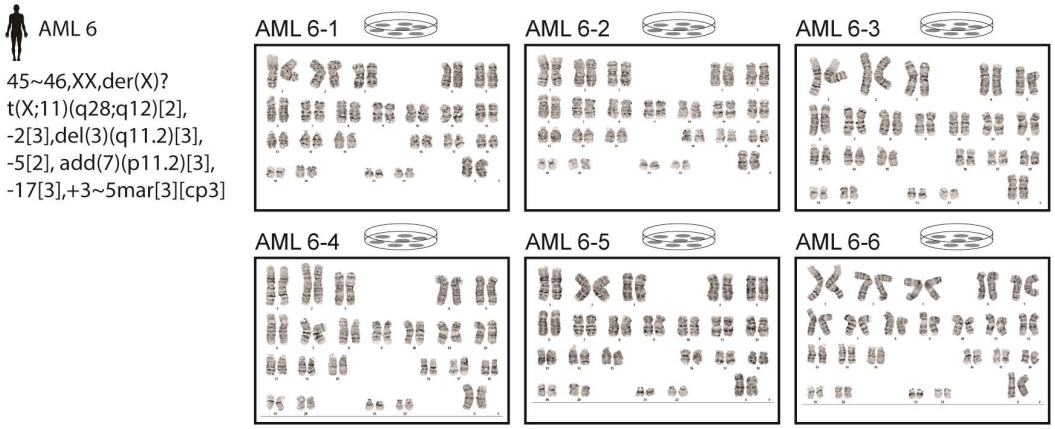

Figure 3.7 AML 6 gave rise to exclusively normal iPSCs. Using G-band karyotyping, it was found that all 6 clones did not contain aberrations found in AML6.

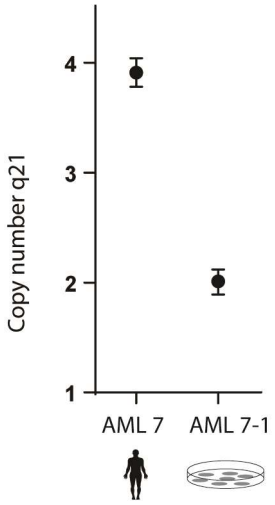

Figure 3.8 AML 7 gave rise to iPSCs devoid of isodicentric chromosome 21. AML 7 and AML7-iPSCs assessed by ddPCR for long arm (q) of chromosome 21. Copy number determined relative to control probe. Error bars represent Poisson 95% confidence intervals.

3.1.4 LSC+ xenografts give rise to unique AML-iPSC clones

Four AML samples were selected to be reprogrammed *de novo* and post xenograft into immune-deficient mice. Remaining samples were reprogrammed *de novo* only (Figure 3.2). AML samples were separated using FACS to achieve >99.5% pure

populations for reprogramming (Figure 3.2). Colonies were identified by TRA-1-60 staining and manually isolated on day 18-21 post transduction and cultured in naïve conditions to increase reprogramming efficiency [145].

Based on the aberrations found in the iPSC clones, reprogramming of xenograft samples did not select for a single LSC clone in AML 1 or 2 (Figure 3.4). It appears that in the case of AML 1, reprogramming gave rise to 3 clones unique to xenograft samples and 3 clones unique to *de novo* samples. Only 1 clone overlapped and was derived from both the *de novo* and xenograft sample. In the case of AML 2, reprogramming gave rise to 2 clones unique to xenograft samples and 1 clone unique to d*e novo* samples. There were no overlapping clones derived from xenograft and *de novo* in AML 2.

In the case of AML 3, AML-iPSCs were derived only form *de novo* sample (Figure 3.4). Unlike the other two samples, the two AML-iPSCs clones from sample 3 did not capture all of the mutations found in the primary sample and thus represent a subset of mutations that was selected for by reprogramming. Two iPSC clones derived from *de novo* and all iPSCs derived from exclusively myeloid the xenografts were determined to be normal by karyotype (Figure 3.3).

3.1.5 Reprogramming of AML and Healthy is Enriched in CD34+ Population

AML is known to recapitulate the healthy hematopoietic hierarchy where reprogramming potential is dependent on the differentiation status[67,160] . Thus, in order to identify if AML reprogramming is also enhanced in the more primitive stem and progenitor compartment AML samples were fractionated by FACS using CD34 prior to

reprogramming outlined in (Figure 3.9). In order to increase probability of capturing a leukemic stem or progenitor cell for reprogramming, samples were also sorted on CD33. AML samples which gave rise to AML-iPSCs were sorted into all four populations when the population existed within the sample. In the AML samples 1, 2 and 3, which gave rise to AML-iPSCs, colonies arose almost exclusively from the CD34+ population in all three samples (Figure 3.9). There was only a single occurrence in which a colony arose from the CD34 negative population in AML 1.

In order to more clearly discern the relationship between leukemic reprogramming and CD34, Fisher's exact test was used. In order to assess AML reprogramming, AML samples 1, 2 and 3 were tested and reprogramming was shown to be dependent on expression of CD34 in a total of 71 reprogrammed wells (Figure 3.10). AML samples 3, 4, 5, 6, and 7 gave rise to healthy iPSCs and were used to assess the relationship between healthy reprogramming and CD34. As has been previously shown in healthy samples, within AML patient samples reprogramming was shown to depend on CD34 expression in a total of 111 reprogrammed wells.

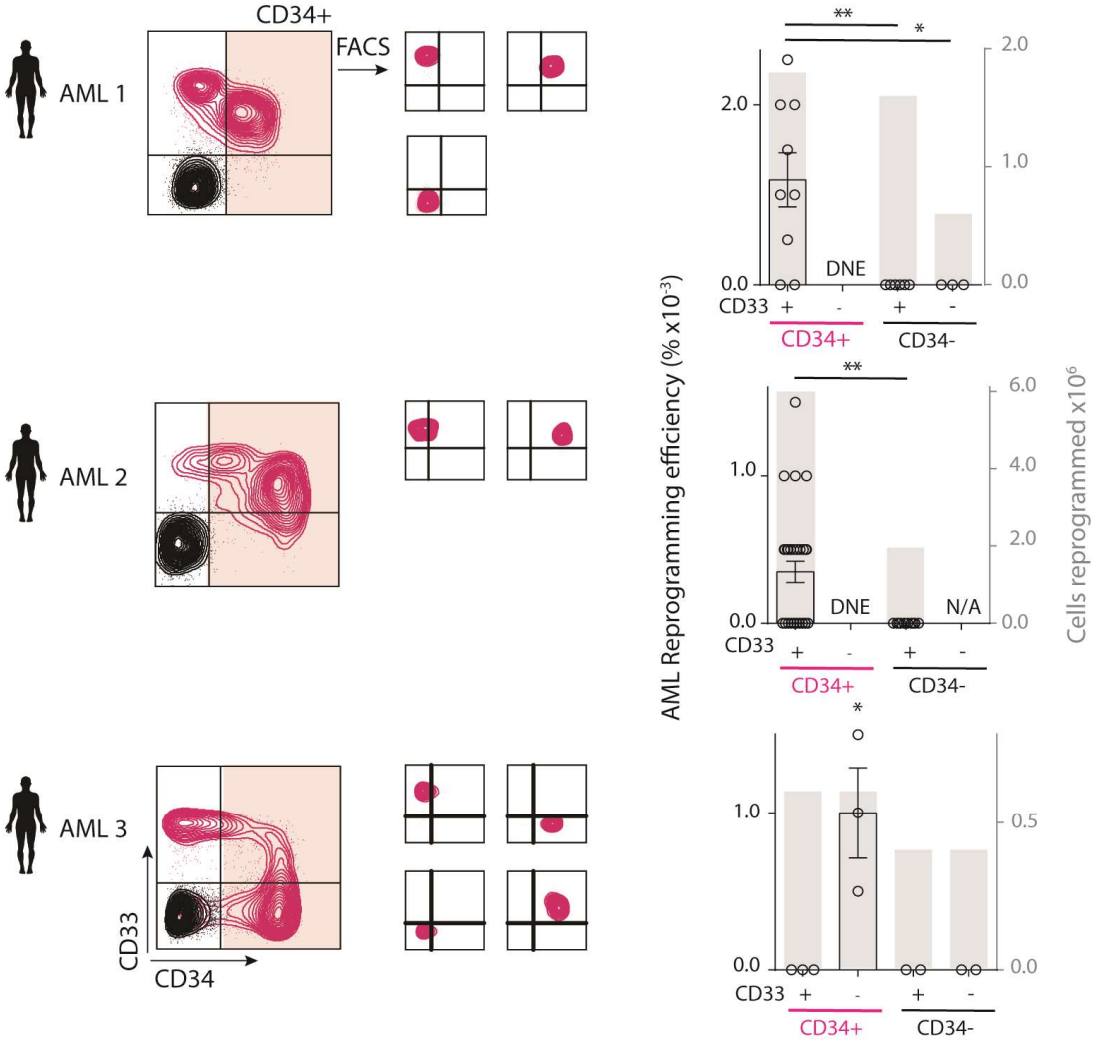

Figure 3.9 AML reprogramming of CD34 and CD33 sorted populations. AML samples were FACS separated by CD34 and CD33 and populations were reprogrammed. Leukemic reprogramming efficiency of populations is in black, plotted on left axis. Total cells that were attempted to be reprogrammed per population are shown on right axis in grey. Error bars represent ∓ SEM. Two-way ANOVA with Tukey's multiple comparison test was performed for statistical analysis.

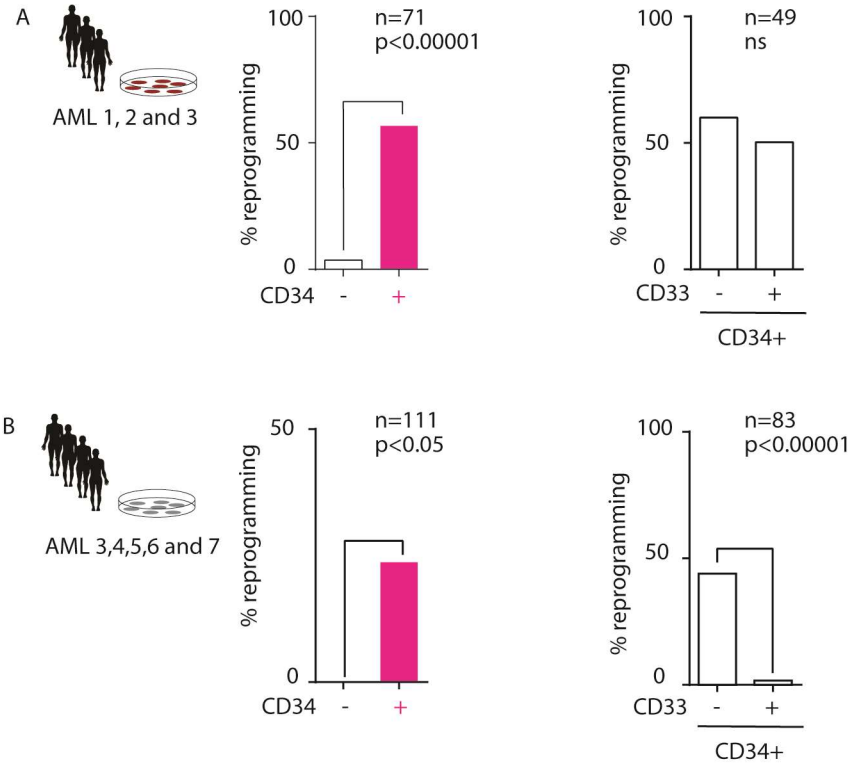

Figure 3.10 AML and healthy reprogramming dependence on CD34 and CD33. (A) Leukemic reprogramming dependence on CD34 (left) and on CD33 within CD34+ (right) was measured using Fisher's exact test. (B) Normal reprogramming from AML sample dependence on CD34 (left) and on CD33 within CD34+ (right) was measured using Fisher's exact test.

3.1.6 Reprogramming of Healthy, but not AML, Within AML Samples can be Selected for Using CD33

Myeloid marker, CD33 was used because it was hypothesized to aid in selection of leukemic cells for reprogramming. Based on AML samples 1, 2 and 3, leukemic reprogramming in not enhanced in CD33+ or CD33- compartments within CD34+ cells (Figure 3.10A). In fact the AML-iPSCs arose from exclusively CD34+CD33+ population in AML samples 1 and 2. But, in AML sample 3, AML-iPSCs arose exclusively from CD34+CD33- population (Figure 3.9).

Reprogramming of healthy progenitors in AML samples that gave rise to healthy-iPSCs (AML samples 3, 4, 5, 6 and 7) was however, significantly enhanced in the CD34+CD33- population. In fact healthy iPSCs arose from almost exclusively CD34+CD33- population.

3.1.7AML and healthy iPSCs from AML samples all arose from blast gate

Immuno-phenotyping AML samples is common practice for clinicians used for diagnosis and monitoring. In healthy samples it is possible to gate granulocyte, lymphocyte, monocyte and blast populations using flow cytometry CD45 staining, typical gating pattern shown in Figure 3.11A[162]. In AML samples, the profile is used to identify blast population shown by Vo and colleagues[163]. Populations that gave rise to reprogramming were plotted on CD45 versus side scatter area to determine where they fit on the SCC-A and CD45 profile in relation to typical blast gating. Samples 1 and 2 from which the CD34+CD33+ populations gave rise to AML-iPSCs fit into the blast gate (Figure 3.11B). Interestingly sample 3 CD34+CD33- population which gave rise to a mixture of healthy and AML iPSCs also fit into the blast gate (Figure 3.11C). And finally AML 4 and 5 sample populations CD34+CD33- which gave rise to exclusively healthy iPSCs also fit within the blast gate (Figure 3.11D).

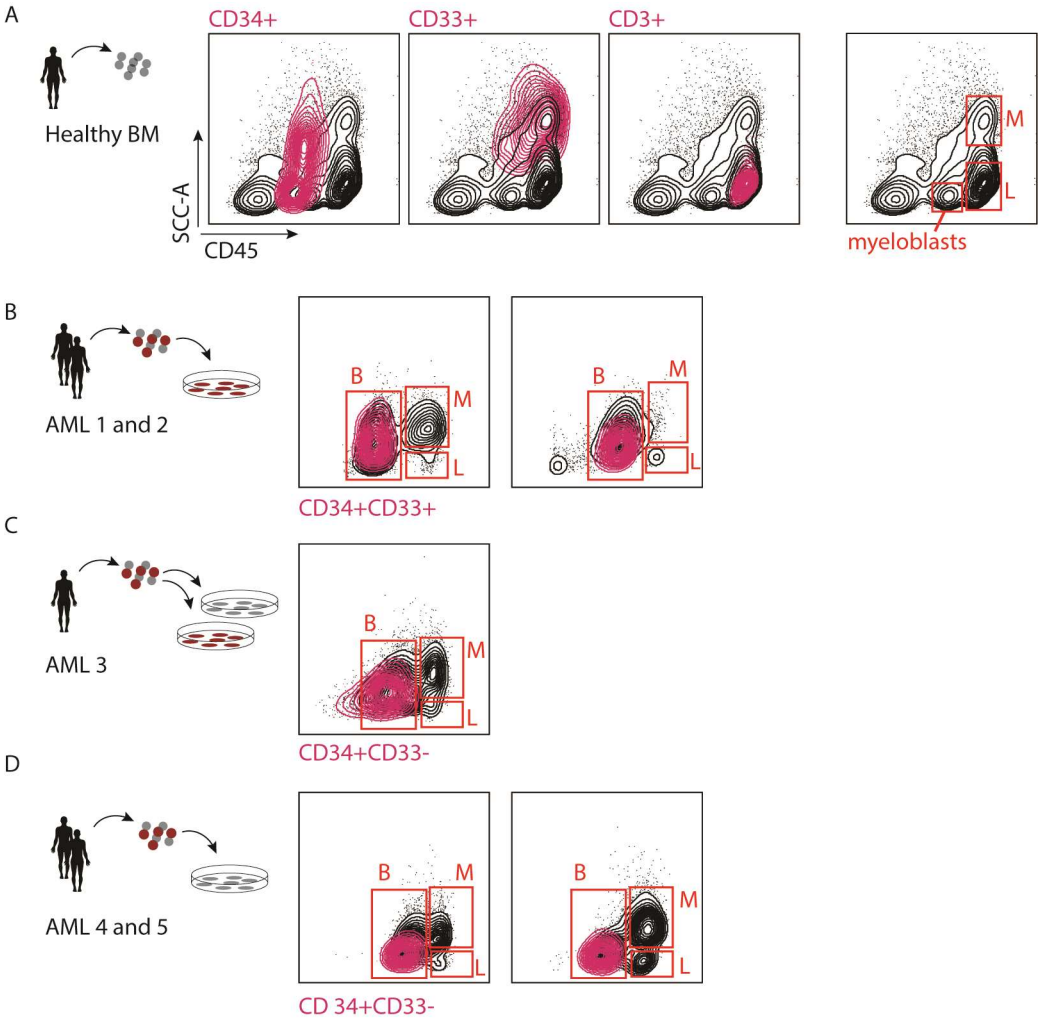

Figure 3.11 AML iPSC and normal reprogramming originates from within AML blast population. (A) Healthy BM CD45 and SSC-A flow profile. Total BM in black overlaid by CD34+, CD33+, and CD3+ populations in pink, showing how they fit into the profile. Typical gating of total sample is shown by M (monocyte), L (lymphocyte), and myeloblast gates. (B) AML samples 1 and 2 which gave rise to AML-iPSCs shown in black left to right. Overlaid in pink is the CD34+CD33+ population which gave rise to AML-iPSCs. Gates for (A), (B) and (C) are labelled as M (monocytes), L (lymphocytes, and B (AML blast). (C) AML sample 3 shown in black, overlaid by CD34+CD33- population which gave rise to AML and healthy iPSCs. (D) AML 4 and 5 shown in black from left to right, overlaid with CD34+CD33- population which gave rise to exclusively healthy iPSCs.

3.2 AML-iPSC Characterization

3.2.1 AML-iPSCs Demonstrate Pluripotency Potential

AML and normal iPSCs had indistinguishable morphology and expanded similarly when pick passaged every 6 days (Figure 3.12A). Once cultured 3-4 weeks post derivation, cells from AML-iPSCs 1,2 and 3 were dissociated on day 6 post passage and tested for their expression of external pluripotency markers TRA-1-60 and SSEA3, and internal pluripotency markers OCT3/4 and NANOG. Pluripotency markers showed expression not different from healthy iPSC and PSC controls (Figure 3.12B).

In addition, two unique clones derived from AML 1 were functionally interrogated for pluripotency using the teratoma assay which demonstrated that AM1- 2 and AML1-8 are capable of giving rise to all three germ lineages *in vivo* (Figure 3.12C).

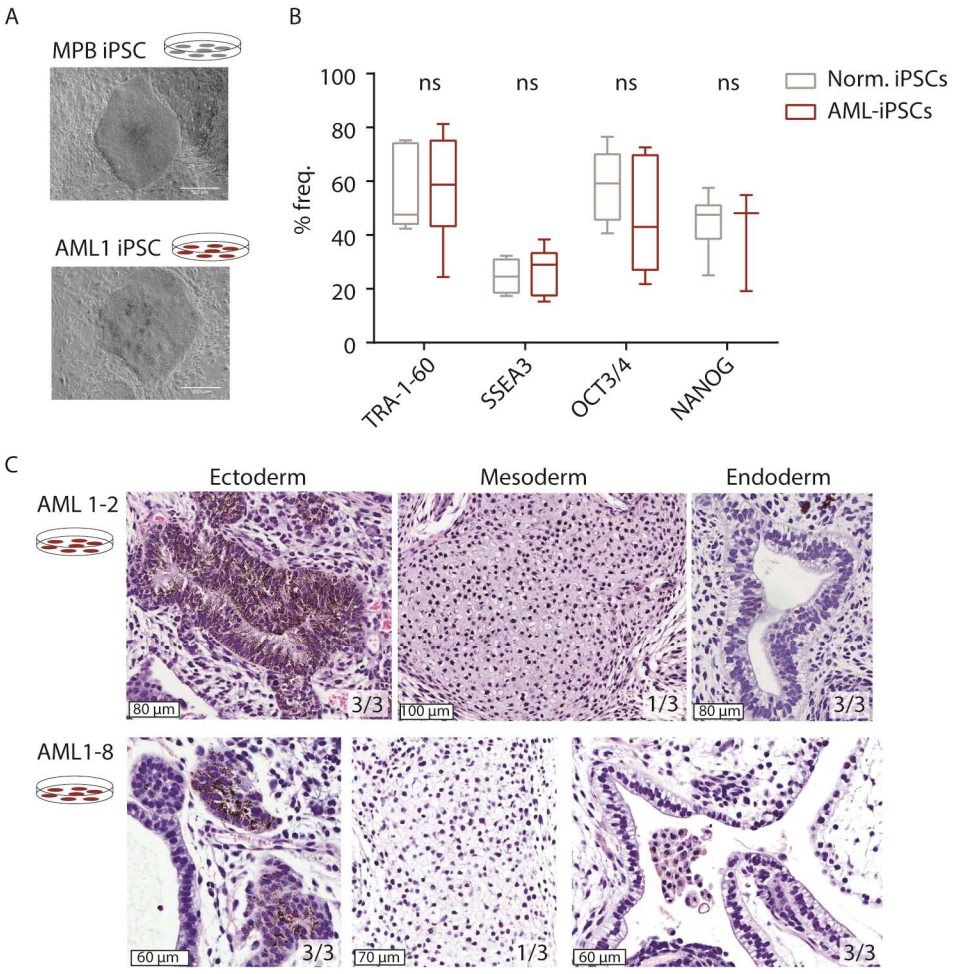

Figure 3.12 AML-iPSCs possess equivalent pluripotency to normal iPSCs. (A) AML and MPB-iPSCs imaged using phase-contrast microscopy. Scale bars are 500 µm. (B) Surface and internal pluripotency markers measured by flow cytometry. Error bars represent ∓ SEM. Statistical analysis done using two-way ANOVA and Tukey's multiple comparison test. Population distribution was considered normal using D'Agostino & Pearson omnibus normality test (C) Hematoxylin and eosin staining of teratomas sections at 8 weeks. Tissues shown are pigmented epithelium (ectoderm), columnar epithelium (endoderm) and hyaline cartilage (mesoderm). Frequency of tissue per independent mouse teratomas after 1 scored tissue section is recorded in bottom right corner.

3.2.2 AML-iPSCs have Clone Dependent Myelo-erythro Differentiation Potential

When subjected to hematopoietic differentiation, AML- iPSC clones from AML patients 1, 2 and 3 demonstrated variable differentiation capacities between different clones. Phenotypically differentiation was measured by co-expression of CD34 and CD45

marking the stem and progenitor cell populations. Expression of pan-hematopoietic marker CD45 is also shown (Figure 3.13). Further differentiation cells were plated into a CFU assay where colonies were manually counted on day 14 post to measure functional progenitor output while distinguishing myeloid and erythroid colonies.

Clones from AML 1, demonstrated variable differentiation capacity. Clones AML1-2, AML1-4, AML1-9, and AML1-11 all produced significantly lower number of phenotypic and functional progenitors than healthy iPSCs (Figure 3.13A and C). Overall expression of CD45 was also significantly reduced (Figure 3.13B). Clone AML1-12 was on par with healthy iPCSs in its ability to produce phenotypic progenitors and CD45+ hematopoietic cells (Figure 3.13 A and B). It was also the only AML 1 clone that was capable of producing erythroid colonies in the CFU, however even this clone had significantly less functional progenitors than the normal iPSCs (Figure 3.13C).

All clones from AML 2 demonstrated reduced differentiation capacity. Both, AML2-2 and AML 2-3 produced significantly less progenitors thank healthy iPSCs by phenotype and function and none of the progenitors were able to give rise to erythroid colonies (Figure 3.13A and C). Both clones also produced significantly less hematopoietic cells (Figure 3.13B).

Clones from AML3 include 2 clones containing leukemic aberrations (AML3-1 and AML3-2) and 2 clones shown to have normal karyotype in Figure 3.3 (AML3-3 and AML3-4). Both normal clones, AML 3-3 and AML 3-4, showed normal capacity to produce phenotypic and functional progenitors with the ability to produce erythroid colonies as well as hematopoietic cells (Figure 3.13A-C). Leukemic clones on the other

hand showed variable differentiation capacity. AML3-1 was on par with healthy in producing progenitors and hematopoietic cells in every capacity and was able to produce erythroid colonies (Figure 3.13A-C). Whereas, AML3-2 had significantly reduced differentiation by all measures and had no capacity to produce erythroid colonies (Figure 3.13A-C).

3.2.3 AML-iPSCs Possess Ability to Differentiate to Other Mesoderm Tissues

Because many AML-iPSC clones demonstrated diminished or reduced hematopoietic differentiation capacity, we thought to test their ability to differentiation to another mesoderm lineage cell type, such as cardiomyocytes. The ability to produce cardiomyocytes was measured by smooth muscle actin (SMO) expression as was done previously[159]. All AML-iPSC clones that were tested for hematopoietic differentiation were able to produce cardiomyocyte structures and expressed SMO as healthy iPSC control (Figure 3.14).

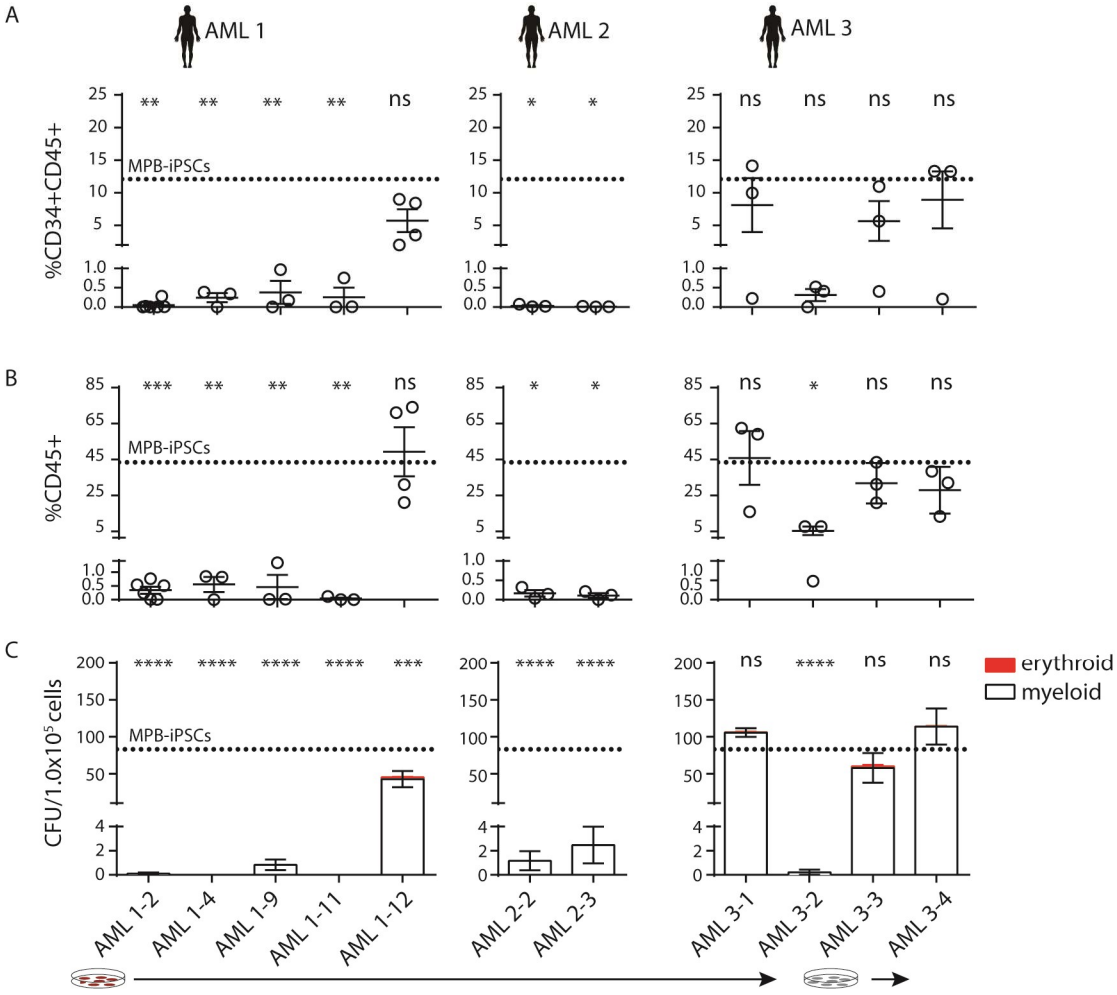

Figure 3.13 Hematopoietic differentiation potential of AML-iPSCs. AML-iPSCs derived from patients 1, 2 and 3 were differentiated to hematopoietic stem and progenitor cells. (A) On day 21 post differentiation cells were analyzed for presence of stem and progenitor cells using flow cytometry for %CD34+CD45+ expression (B) Cells were also analyzed using flow cytometry for %CD45 pan-hematopoietic marker. (C) On day 21, cells were plated into CFU assay and 14 days later colonies were manually scored as myeloid or erythroid. All reported results were performed at least as three independent experiments for each sample. Dotted line represents positive control differentiation average. Data shows average with error bars representing ∓SEM. All comparisons are made to MPB-iPSC control using one-way ANOVA and Dunnett's multiple comparison test.

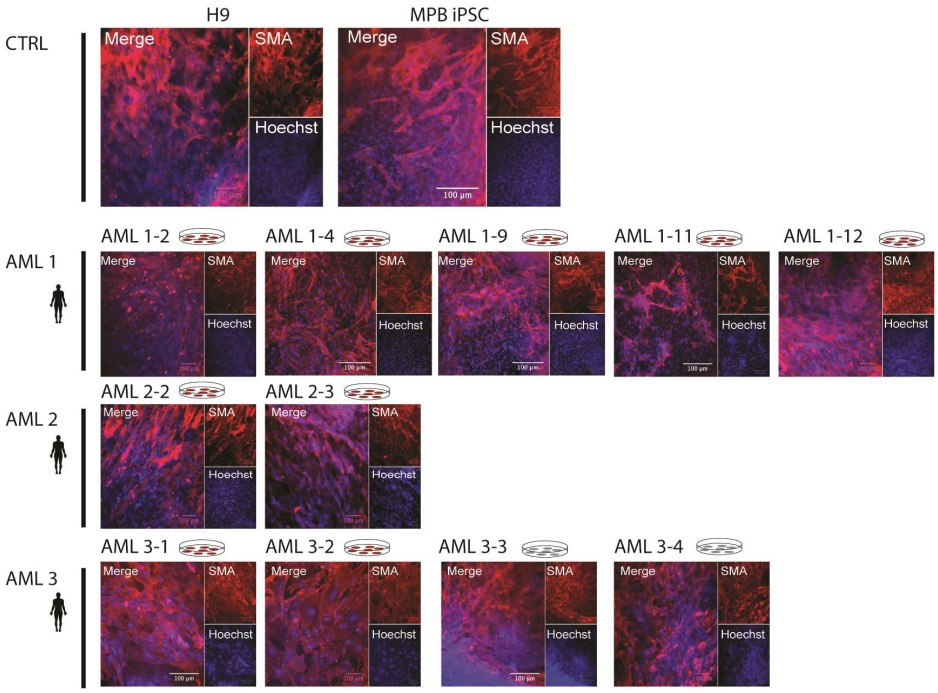

Figure 3.14 AML and normal iPSCs cardiomyocyte differentiation potential. AML and normal iPSCs were differentiated towards cardiomyocytes and stained for smooth muscle actin (SMO) and Hoechst nucleic stain. All iPSC clones were differentiated a minimum of three times, with a representative image shown. Error bars are 100µm.

4.0 DISCUSSION

4.1 AML Reprogramming

The overarching goal of this work was to generate a pluripotent AML-iPSC model which would have the potential to capture intra and inter AML patient heterogeneity *in vitro.* This model has the potential to model disease phenotype through differentiation, capture AML epigenetic abnormalities, and interrogate disease mutations.

4.1.1 AML Reprogramming Captures Abnormal karyotype and Intra- Patient Heterogeneity

Eleven AML samples were reprogrammed using FACS and xenografting to increase reprogramming efficiency of human primary AML. As noted in previous studies of AML and other cancers (Table 1.1), AML cells possess a barrier to reprogramming despite having already acquired pluripotency features through cancer transformation (Figure 4.1). Here, three of eleven samples gave rise to AML-iPSCs capturing leaukemic aberrations, and five of eleven AML samples giving rise to healthy iPSCs derived from healthy stem and progenitor cells in the sample. Samples with a variety of aberrations were selected for reprogramming, however the three that successfully gave rise to AML-iPSCs possessed gross karyotypic abnormalities spanning intermediate and adverse risk stratifications (Table 3.1).

AML-iPSCs derived from AMLs 1 and 2 were able to capture all mutations found in the primary samples detected by G-band karyotyping (AML 1) and copy number array

(AML2) thus completely capturing the clonal heterogeneity from these two patients (Figure 3.4). AML-iPSCs derived from AML 3 did not capture all karyotyping abnormalities from the primary sample. In fact del(15)(q11.2q15), which is present in over 70% of all aberrant cells in AML 3 by G-band karyotyping (Figure 3.3). This indicates that only rare clones were permissive for reprogramming in AML 3 whereas the majority of AML cells were refractory to reprogramming. In order to further evaluate this, additional reprogramming of AML 3 *de novo* should be performed to attempt to capture additional leukemic clones and the full heterogeneity of the sample for modelling.

4.1.2 Reprogramming of AML Xenografts Captures Exclusive Disease Clones

AML samples 1-4 were reprogrammed *de novo* and from xenografts (Figure 3.2). AML samples 1 and 2, both gave rise to clones that were derived exclusively from either xeno or *de novo* sample (Figure 3.4). In AML 1, one clone that was derived most frequently was derived twice from the xeno and twice from the *de novo* sample (iPSCs AML1-5,1-6,1-10 and 1-11 shown in Figure 3.4). All other clones from AML 1 and 2 were not derived from both. This suggests that the xenografts are selecting a subset of AML samples which potentially becomes expanded in the xenograft and is thus gains a higher potential for reprogramming. It is possible that specific clones that are not proliferative in the *de novo* sample, become proliferative in the xenograft which allows for their reprogramming which would be suggested by a previous study by Munoz-Lopez and colleagues[146].

In the case of AML 3, AML-iPSCs were only derived from *de novo* sample. Normal iPSCs were also derived from the *de novo* sample as expected, since more than

50% of cells in the primary sample were normal by karyotype (Figure 3.3). Surprisingly, although reprogrammed xenografts were exclusively myeloid indicating leukemic engraftment, all iPSCs derived from the xenografts were normal by karyotype (Figure 3.3). This is potentially explained by the presence of multi-lineage engraftment in few mice which were not used for reprogramming. The multi-lineage engraftment in some animals indicates that healthy stem and progenitor cells are responsible for the reconstitution and it is possible that the exclusively myeloid animals, the graft was resulting from a myeloid biased cell [164].

AML 4 gave rise to healthy iPSCs exclusively from the d*e novo* sample, with no iPSC colonies arising from the xenograft (Figure 3.2). This is indicates that leukemic cells which were refractory to reprogramming in the AML sample gave rise to the xenografts. This is an example that demonstrates that simple increase in proliferation by xenograft is not sufficient to precipitate reprogramming of leukemic cells.

All interpretations concerning clonal composition and capture of intra patient heterogeneity are limited by the resolution of G band karyotyping for AML samples 1 and 3. This method is limited by resolution of approximately five Mb and is unable to detect any genetic mutations associated with AML. Copy number array used for clonal classification of AML 2 iPSCs has a resolution of approximately 100 bp but is not effective for detecting genetic mutations or structural translocations associated with AML. In order to draw conclusions regarding capturing of clones and patient heterogeneity, the primary AML samples and iPSCs should by analyzed using whole exome sequencing (WES). This analysis would provide a basis that would allow AML-

iPSCs to be used to functionally demonstrate driver versus passenger nature of AML associated mutation in the future.

4.1.3 Reprogramming Potential of AML Cells is Dependent on Cell Differentiation Stage

It is well established that reprograming is enhanced in the more primitive healthy hematopoietic compartment[160]. FACS separation of patient samples demonstrates that AML reprogramming is completely restricted to the more primitive CD34+ compartment, with only one case of CD34- reprogramming form a total of 71 wells (Figure 3.9A). Healthy cell reprogramming from AML patient samples was also restricted to the CD34+ compartment exclusively (Figure 3.9B). This demonstrates, that even though much more rare, AML reprogramming is also limited by the differentiation state. Just as in healthy cells, additional reprogramming factors may be required for reprogramming of terminally differentiated CD34- leukemic cells.

From the fractionation strategy it is clear that selection of CD34+ cells is conducive for AML and healthy for OSKM mediated reprogramming. Interestingly, despite its use in other studies, selection for CD33+ cells did not provide a selective advantage for leukemic reprogramming (Figure 3.9A). In fact, it seems to be specific to the patient sample, where AMLs 1 and 2 yield AML reprogramming form the CD34+CD33+ compartment and AML 3 yields leukemic reprogramming from the CD34+CD33- compartment. This can be explained by aberrant AML cells residing in different compartments along the myeloid differentiation axis and implicates that selection by CD33 should not be employed in samples where the differentiation status of aberrant cells in unknown[68].

AML samples that gave rise to healthy were analyzed in the same manner and shown to reprogram from the CD34+ fraction exclusively (Figure 3.9B). Healthy reprogramming was significantly enriched in the CD34+CD33- population oppose to the CD34+CD33+. This indicates that reprogramming of healthy stem and progenitor cells is occurring from non-myeloid compartment which should be utilized to acquire healthy iPSCs from AML samples when desired.

A gating strategy that was considered, but not employed in this study was to gate on blast populations [162,163]. This strategy was not used due to the subjective nature of blast gates. When assessing sorted and reprogrammed CD34+ population on blast gating, it was observed that leukemic and healthy iPSCs were derived from blast gates (Figure 3.11) indicating that it would not be a productive strategy to select for leukemic reprogramming. This gating strategy would be interesting to use on AML samples that do not express CD34+ cells at all, in order to determine if AML reprogramming is possible.

4.2 AML-iPSC Model Characterization: Pluripotency

AML-iPSCs were tested for phonotypic expression of pluripotency markers by flow cytometry. All AML-iPSCs were not different from normal iPSCs. In order to functionally test pluripotency a teratoma assay was performed and two independent clones derived from AML 1 demonstrated potential to form all three germ lineages. Ectoderm and endoderm representatives were found in all three sections of independent teratomas. However, mesoderm representatives were only present in one of three teratomas indicating that despite the ability to produce cartilage, AML-iPSCs may be impaired in their mesoderm differentiation capacity.

4.2.1 AML-iPSCs have Clone Dependent Hematopoietic Differentiation Capacity

The ability to produce functional and phenotypic progenitors as well as hematopoietic cells was variable among the clones, with normal clones resembling normal hematopoietic differentiation (Figure 3.13). This suggests that mutation status is supported by differentiation capacity.

In AML 1, where all clones contained leukemic aberrations, most were blocked in hematopoietic differentiation. Only clone AML1-12 demonstrated hematopoietic differentiation similar to normal iPSCs. In fact AML1-9 and AML1-12 were considered to be derived from the same clone based on aberration pattern (Figure 3.4). In terms of hematopoietic differentiation, these two were the only clones able to produce appreciable functional progenitors forming colonies in the CFU assay, however clone AMl1-9 had significantly lower potential than AML1-12 and healthy iPSCs. This indicates that, since karyotyping is limited by sensitivity there are likely additional genetic or epigenetic differences between AML1-9 and AML1-12 that were not captured. In the case of AML 2, both clones showed significantly reduced differentiation potential. The impaired hematopoietic differentiation of multiple AML-iPSC clones is not surprising as it has been reported on previously[145,154]. This poses an interesting question of whether the hematopoietic differentiation block is as a result of a genetic or epigenetic component that is also responsible for the lack of complete differentiation of AML cells. The case of AML 3 was interesting because isogenic healthy iPSCs in fact resembled normal hematopoietic differentiation (Figure 3.13). The leukemic clones had drastically different

differentiation, with one being on par with normal and the second completely lacking any hematopoietic potential.

Because some clones demonstrated the potential to produce stem and progenitor cells, it will be of interest to determine if they are capable of engraftment into immuno-deficient mice. Although it is well known that true HSCs have never been produced from PSCs, it has been shown that AML-iPSCs and iPSCs with leukemic aberration are able to reconstitute immune-deficient mice[32,113]. In terms of the derivation method, there was no evidence in any of the samples that would indicate that that is a differentiation potential difference between clones derived from xenografted versus de novo samples (Figure 3.4 and 3.13).

It was of interest to determine if AML-iPSC differentiation to other tissues was affected due to the known genetic aberrations. This was necessary because AML does not arise form somatic mutations, but form mutation that occur within the hematopoietic compartment, and thus is it reasonable to assume that these mutations may cause defects in differentiation to other tissues. The teratoma assay suggested that endoderm and ectoderm structures were consistently present, other mesoderm layer tissues may have reduced differentiation potential (Figure 3.12). The closest to hematopoiesis and thus most likely impaired is cardiomyocyte differentiation. All derived clones were capable of successful cardiomyocyte differentiation suggesting that the differentiation block is specific to hematopoiesis (Figure 3.14).

Overall AML-iPSCs with varying haematopoietic differentiation potential provide the perfect platform to investigate what mechanism is responsible for the lack of hematopoietic differentiation with potential to link this to AML patients.

4.3 Challenges of an iPSC derived model of AML

As discussed previously, artificially derived model of diseases are associated with certain limitations which should be considered. In the case of AML-iPSCs model, multiple challenges should be considered and ideally addressed experimentally. Firstly, it remains unknown if the reprogramming process stochastically or selectively isolates leukemic clones. Despite allowing for the isolation and culture of disease clones with the potential to decipher the contributions of genetic abnormalities, it has not yet been shown that any of the leukemic clones selected by reprogramming were clinically relevant to disease maintenance or relapse. This remains a high bar, as there have never yet been a study which was able to identify and definitively show which leukemic clones or their progeny were clinically responsible for initiation of disease relapse.

Secondly, it remains unknown whether the reprogramming process itself may introduce additional genetic insults to an already unstable leukemic genome. This should be considered when isolating the genetic contributions to the phenotype of AML. If a mutation is found to have a relevant function it should be further identified to be present within a patient sample. Despite this limitation, if the mutation cannot be identified in the patient sample, it should still be considered as it may be occurring to due to a predisposition caused by multiple leukemia associated mutations and present in genomes of other patients in a relevant manner.

As in every model, limitations of AML-iPSCs should be considered when using the model and findings should be validated in primary AML samples and healthy human blood.

5.0 CONCLUSION

AML-iPSCs have the unique potential to identify targets that are selective for leukemic cells which has not been possible with primary sample screening and current AML models. Thus it was crucial to investigate the challenges of reprogramming AML cells. Here we have directly interrogated strategies used for reprogramming of AML and were successful in deriving AML-iPSCs from three patient samples capturing unique genetic aberrations and intra patient heterogeneity. We demonstrated that reprogramming AML samples xenograft versus d*e novo*, gives rise to unique clones, but did not have any bearing on the capacity of the clones to differentiate to hematopoietic cells. AML sample sorting allowed us to identify that AML samples reprogramming is dependent on differentiation status where the more primitive, CD34+ fraction was exclusively able to give rise to AML-iPSCs. It was also determined that CD33 selection was not effective at enhancing leukemic reprogramming, as AML-iPSCs arose from both CD33 positive and negative compartments. However, CD33 negative compartment in AML samples was significantly more likely to select for healthy reprogramming. Further, we determined that AML-iPSCs resemble normal iPSCs while in the pluripotent state. When subjected to differentiation, AML-iPSCs have either diminished or normal hematopoietic differentiation, while all retaining the ability to differentiate to other mesoderm tissue.

This provides a platform with a functional differentiation read-out which can be used to identify the genetic or epigenetic link to aberrant hematopoiesis in AML.

6.0 REFERENCES

1 Knoblich, J. A. Asymmetric cell division during animal development. *2* **1**, 11-20 (2001).

2 Morrison, S. J. & Kimble, J. Asymmetric and symmetric stem-cell divisions in development and cancer. *Nature* **441**, 1068-1074 (2006).

3 Konstantinov, I. E. In Search of Alexander A. Maximow: The Man Behind the Unitarian Theory of Hematopoiesis. *Perspectives in Biology and Medicine* **43**, 269-276, doi:10.1353/pbm.2000.0006 (2000).

4 Sabin, F. R. The Production of Osteogenic Sarcomata and the Effects on Lymph Nodes and Bone Marrow of Intravenous Injections of Radium Chloride and Mesothorium in Rabbits. *J Exp Med* **56**, 267-289, doi:10.1084/jem.56.2.267 (1932).

5 Jacobson, L. O., Marks, E. K. & Robson, M. J. Effect of spleen protection on mortality following x-irradiation. *J Lab Clin Med* **12**, 197-201 (1949).

6 Lorenz, E., Congdon, C. & Uphoff, D. Modification of Acute Irradiation Injury in Mice and Guinea-Pigs by Bone Marrow Injections. *Radiology* **58**, 863-875 (1952).

7 Ford, C. E., Hamerton, J. L., Barnes, D. W. & Loutit, J. F. Cytological identification of radiation-chimaeras. *Nature* **177**, 452-454 (1956).

8 Ferrebee, J. W., Lochte, H. L., Jaretzki, A., Sahler, O. D. & Thomas, E. D. Successful marrow homograft in the dog after radiation. *Surgery* **43**, 516-520 (1958).

9 Till, J. E. & McCulloch, E. A. A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiat Res* **178**, AV3-7, doi:10.1667/RRAV01.110.2307/3570892 (1961).

10 Becker, A. J., McCulloch, E. A. & Till, J. E. Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells. *Nature* **197**, 452-454 (1963).

11 Siminovitch, L., McCulloch, E. A. & Till, J. E. The distribution of colony-forming cells among spleen colonies. *J Cell Comp Physiol* **62**, 327-336 (1963).

12 Wu, A. M. Cytological Evidence for a Relationship between Normal Hematopoietic Colony-Forming Cells and Cells of the Lymphoid System. *J Exp Med* **127**, 455-464, doi:10.1084/jem.127.3.455 (1968).

13 Wu, A. M., Till, J. E., Siminovitch, L. & McCulloch, E. A. A cytological study of the capacity for differentiation of normal hemopoietic colony-forming cells. *J Cell Physiol* **69**, 177-184 (1967).

14 Thomas, E. D., Lochte, H. L. J., Lu, W. C. & Feffebee, J. W. Intravenous infusion of bone marrow in patients receiving radiation and chemotherapy. *N Engl J Med* **257**, 491-496 (1957).

15 Starzl, T. E. The birth of clinical organ transplantation. *J Am Coll Surg* **192**, 431-446 (2001).

16 Prindull, G., Prindull, B. & Meulen, N. Haematopoietic stem cells (CFUc) in human cord blood. *Acta Paediatr Scand* **67**, 413-416 (1978).

17 Evans, M. J. & Kaufman, M. H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-156 (1981).
18 Trounson, A., Leeton, J., Besanko, M., Wood, C. & Conti, A. Pregnancy established in an infertile patient after transfer of a donated embryo fertilised in vitro. *Br Med J* **286**, 835-838 (1983).
19 Thomson, J. A. *et al.* Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145-1147 (1998).
20 Robertson, J. A. Embryo stem cell research: ten years of controversy. *J Law Med Ethics* **38**, 191-203 (2010).
21 Adewumi, O. *et al.* Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. *Nat Biotechnol* **25**, 803-816, doi:10.1038/nbt1318 (2007).
22 Trounson, A. & DeWitt, N. D. Pluripotent stem cells progressing to the clinic. *Nat Rev Mol* **17**, 194-200, doi:10.1038/nrm.2016.10 (2016).
23 da Cruz, L. *et al.* Phase 1 clinical study of an embryonic stem cell–derived retinal pigment epithelium patch in age-related macular degeneration. *Nat Biotechnol* **36**, 328-337, doi:10.1038/nbt.4114 (2018).
24 Reid, J. C. *et al.* CXCL12/CXCR4 Signaling Enhances Human PSC-Derived Hematopoietic Progenitor Function and Overcomes Early In Vivo Transplantation Failure. *Stem Cell Reports* **10**, 1625-1641, doi:10.1016/j.stemcr.2018.04.003 (2018).
25 Doulatov, S. *et al.* Induction of Multipotential Hematopoietic Progenitors from Human Pluripotent Stem Cells via Respecification of Lineage-Restricted Precursors. *Cell Stem Cell* **13**, 459-470, doi:10.1016/j.stem.2013.09.002 (2013).
26 Gori, J. L. *et al.* Vascular niche promotes hematopoietic multipotent progenitor formation from pluripotent stem cells. *J Clin Invest* **125**, 1243-1254, doi:10.1172/jci79328 (2015).
27 Lee, J. *et al.* mRNA-mediated glycoengineering ameliorates deficient homing of human stem cell–derived hematopoietic progenitors. *J Clin Invest* **127**, 2433-2437, doi:10.1172/jci92030 (2017).
28 Ramos-Mejía, V. *et al.* HOXA9 promotes hematopoietic commitment of human embryonic stem cells. *Blood* **124**, 3065-3075, doi:10.1182/blood-2014-03558825 (2014).
29 Ran, D. *et al.* RUNX1a enhances hematopoietic lineage commitment from human embryonic stem cells and inducible pluripotent stem cells. *Blood* **121**, 2882-2890, doi:10.1182/blood-2012-08451641 (2013).
30 Risueño, R. M. *et al.* Inability of Human Induced Pluripotent Stem Cell-Hematopoietic Derivatives to Downregulate MicroRNAs In Vivo Reveals a Block in Xenograft Hematopoietic Regeneration. *Stem Cells* **30**, 131-139, doi:10.1002/stem.1684 (2012).
31 Wang, L. *et al.* Generation of hematopoietic repopulating cells from human embryonic stem cells independent of ectopicHOXB4expression. *J Exp Med* **201**, 1603-1614, doi:10.1084/jem.20041888 (2005).

32 Sugimura, R. *et al.* Haematopoietic stem and progenitor cells from human pluripotent stem cells. *Nature* **545**, 432-438, doi:10.1038/nature22370 (2017).

33 Lu, S. J., Feng, Q., Li, F., Kimbrel, E. A. & Lanza, R. in *Principles of Regenerative Medicine (Third Edition)* (eds A. Atala, R. Lanza, A. G. Mikos, & R. Nerem) (Academic Press Elsevier, 2019).

34 Sasai, Y. Cytosystems dynamics in self-organization of tissue architecture. *Nature* **493**, 318-326, doi:10.1038/nature11859 (2013).

35 Eiraku, M. *et al.* Self-Organized Formation of Polarized Cortical Tissues from ESCs and Its Active Manipulation by Extrinsic Signals. *Cell Stem Cell* **3**, 519-532, doi:10.1016/j.stem.2008.09.002 (2008).

36 Eiraku, M. *et al.* Self-organizing optic-cup morphogenesis in three-dimensional culture. *Nature* **472**, 51-56, doi:10.1038/nature09941 (2011).

37 Lee, G. *et al.* Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs. *Nature* **461**, 402-406, doi:10.1038/nature08320 (2009).

38 Sareen, D. *et al.* Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. *Sci Transl Med* **5**, 149 (2013).

39 McNeish, J., Gardner, J. P., Wainger, B. J., Woolf, C. J. & Eggan, K. From Dish to Bedside: Lessons Learned While Translating Findings from a Stem Cell Model of Disease to a Clinical Trial. *Cell Stem Cell* **17**, 8-10, doi:10.1016/j.stem.2015.06.013 (2015).

40 Hoadley, K. A. *et al.* Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. *Cell* **173**, 291-304, doi:10.1016/j.cell.2018.03.022 (2018).

41 Ben-Porath, I. *et al.* An embryonic stem cell–like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet* **40**, 499-507, doi:10.1038/ng.127 (2008).

42 Malta, T. M. *et al.* Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. *Cell* **173**, 338-353, doi:10.1016/j.cell.2018.03.034 (2018).

43 Kim, J. *et al.* A Myc Network Accounts for Similarities between Embryonic Stem and Cancer Cell Transcription Programs. *Cell* **143**, 313-324, doi:10.1016/j.cell.2010.09.010 (2010).

44 Sokolov, A., Paull, E. O. & Stuart, J. M. ONE-CLASS DETECTION OF CELL STATES IN TUMOR SUBTYPES. *Pac Symp Biocomput* **21**, 405-416 (2016).

45 Kooreman, N. G. *et al.* Autologous iPSC-Based Vaccines Elicit Anti-tumor Responses In Vivo. *Cell Stem Cell* **22**, 501-513.e507, doi:10.1016/j.stem.2018.01.016 (2018).

46 Bernstein, B. E., Meissner, A. & Lander, E. S. The Mammalian Epigenome. *Cell* **128**, 669-681, doi:10.1016/j.cell.2007.01.033 (2007).

47 Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev* **16**, 6-21, doi:10.1101/ (2002).

48 Greger, V., Passarge, E., Höpping, W., Messmer, E. & Horsthemke, B. Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. *Hum Genet* **83**, 155-158 (1989).
49 Sell, S. Cellular origin of cancer: dedifferentiation or stem cell maturation arrest? *Environ Health Perspect* **101**, 15-26 (1993).
50 Lee, J. H. *et al.* Somatic transcriptome priming gates lineage-specific differentiation potential of human-induced pluripotent stem cell states. *Nat Commun* **5**, doi:10.1038/ncomms6605|www.nature.com/naturecommunications (2014).
51 Bar-Nur, O., Russ, H. A., Efrat, S. & Benvenisty, N. Epigenetic Memory and Preferential Lineage-Specific Differentiation in Induced Pluripotent Stem Cells Derived from Human Pancreatic Islet Beta Cells. *Cell Stem Cell* **9**, 17-23, doi:10.1016/j.stem.2011.06.007 (2011).
52 Kim, K. *et al.* Epigenetic memory in induced pluripotent stem cells. *Nature* **467**, 285-290, doi:10.1038/nature09342 (2010).
53 Ohi, Y. *et al.* Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. *Nat Cell Biol* **13**, 541-549, doi:10.1038/ncb2239 (2011).
54 Nakanishi, M. *et al.* Human Pluripotency Is Initiated and Preserved by a Unique Subset of Founder Cells. *Cell* **177**, 910-924.e922, doi:10.1016/j.cell.2019.03.013 (2019).
55 Huang, K. *et al.* Selective demethylation and altered gene expression are associated with ICF syndrome in human-induced pluripotent stem cells and mesenchymal stem cells. *Hum Mol Genet* **23**, 6448-6457, doi:10.1093/hmg/ddu365 (2014).
56 Sagie, S. *et al.* Induced pluripotent stem cells as a model for telomeric abnormalities in ICF type I syndrome. *Hum Mol Genet* **23**, 3629-3640, doi:10.1093/hmg/ddu071 (2014).
57 Leary, A. G., Strauss, L. C., Civin, C. I. & Ogawa, M. Disparate differentiation in hemopoietic colonies derived from human paired progenitors. *Blood* **66** (1985).
58 Mayani, H., Dragowska, W. & Lansdorp, P. M. Lineage commitment in human hemopoiesis involves asymmetric cell division of multipotent progenitors and does not appear to be influenced by cytokines. *J Cell Physiol* **157**, 579-586 (1993).
59 Bryder, D., Rossi, D. J. & Weissman, I. L. Hematopoietic stem cells: the paradigmatic tissue-specific stem cell. *Am J Pathol* **169**, 338-346, doi:10.2353/ajpath.2006.060312 (2006).
60 Boyd, A. L. *et al.* Identification of Chemotherapy-Induced Leukemic-Regenerating Cells Reveals a Transient Vulnerability of Human AML Recurrence. *Cancer Cell* **34**, 483-498.e485, doi:10.1016/j.ccell.2018.08.007 (2018).
61 Shipley, J. L. & Butera, J. N. Acute myelogenous leukemia. *Exp Hematol* **37**, 649-658, doi:10.1016/j.exphem.2009.04.002 (2009).

62 Shallis, R. M., Wang, R., Davidoff, A., Ma, X. & Zeidan, A. M. Epidemiology of acute myeloid leukemia: Recent progress and enduring challenges. *Blood Rev* **36**, 70-87, doi:10.1016/j.blre.2019.04.005 (2019).
63 Döhner, H. *et al.* Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood* **115**, 453-474, doi:10.1182/blood-2009-07235358 (2010).
64 Griffin, J. D. & Löwenberg, B. Clonogenic cells in acute myeloblastic leukemia. *Blood* **68**, 1185-1195 (1986).
65 Lapidot, T. *et al.* A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* **367**, 645-648 (1994).
66 Bhatia, M. *et al.* Quantitative Analysis Reveals Expansion of Human Hematopoietic Repopulating Cells After Short-term Ex Vivo Culture. *J Exp Med* **186**, 619-624, doi:10.1084/jem.186.4.619 (1997).
67 Bonnet, D. & Dick, J. E. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med* **3**, 730-737 (1997).
68 van Galen, P. *et al.* Single-Cell RNA-Seq Reveals AML Hierarchies Relevant to Disease Progression and Immunity. *Cell*, doi:10.1016/j.cell.2019.01.031 (2019).
69 Ding, L. *et al.* Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* **481**, 506-510, doi:10.1038/nature10738 (2012).
70 Yunis, J. J. *et al.* Refined chromosome study helps define prognostic subgroups in most patients with primary myelodysplastic syndrome and acute myelogenous leukaemia. *Br J Haematol* **68**, 189-194 (1988).
71 Grimwade, D. *et al.* The importance of diagnostic cytogenetics on outcome in AML: analysis of 1,612 patients entered into the MRC AML 10 trial. The Medical Research Council Adult and Children's Leukaemia Working Parties. *Blood* **92**, 2322-2333 (1998).
72 Bennett, J. M. *et al.* Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group. *Br J Haematol* **33**, 451-458 (1976).
73 Arber, D. A. *et al.* The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood* **127**, 2391-2405, doi:10.1182/blood-2016-03-643544 (2016).
74 Larson, R. A. *et al.* Evidence for a 15;17 translocation in every patient with acute promyelocytic leukemia. *Am J Med* **76**, 827-841 (1984).
75 GHITIS, J. Acute promyelocytic leukemia? *Blood*, 237-240 (1963).
76 PISCIOTTA, A. V. & SCHULZ, E. J. Fibrinolytic purpura in acute leukemia. *Am J Med* **19**, 824-828 (1955).
77 Wang, Z. Y. & Chen, Z. Acute promyelocytic leukemia: from highly fatal to highly curable. *Blood* **111**, 2505-2515, doi:10.1182/blood-200707-102798 (2008).
78 Dos Santos, G. A., Kats, L. & Pandolfi, P. P. Synergy against PML-RARa: targeting transcription, proteolysis, differentiation, and self-renewal in acute

promyelocytic leukemia. *J Exp Med* **210**, 2793-2802, doi:10.1084/jem.20131121 (2013).

79 Döhner, H. *et al.* Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood* **129**, doi:10.1182/blood-2016-08733196 (2017).

80 Estey, E. H. Acute myeloid leukemia: 2019 update on risk-stratification and management. *Am J Hematol* **93**, 1267-1291, doi:10.1002/ajh.25214 (2018).

81 Papaemmanuil, E. *et al.* Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N Engl J Med* **374**, 2209-2221, doi:10.1056/NEJMoa1516192 (2016).

82 Corral, J. *et al.* An Mll-AF9 fusion gene made by homologous recombination causes acute leukemia in chimeric mice: a method to create fusion oncogenes. *Cell* **85**, 853-861 (1996).

83 Bursen, A. *et al.* The AF4·MLL fusion protein is capable of inducing ALL in mice without requirement of MLL·AF4. *Blood* **115**, 3570-3579, doi:10.1182/blood-2009-06-229542 (2010).

84 Watanabe-Okochi, N. *et al.* AML1 mutations induced MDS and MDS/AML in a mouse BMT model. *Blood* **111**, 4297-4308, doi:10.1182/blood-2007-01-068346 (2008).

85 Nagase, R. *et al.* Expression of mutant Asxl1 perturbs hematopoiesis and promotes susceptibility to leukemic transformation. *J Exp Med* **215**, 1729-1747, doi:10.1084/jem.20171151 (2018).

86 Petersdorf, S. H. *et al.* A phase 3 study of gemtuzumab ozogamicin during induction and postconsolidation therapy in younger patients with acute myeloid leukemia. *Blood* **121**, 4854-4860, doi:10.1182/blood-201301-466706 (2013).

87 Burnett, A. K. *et al.* Identification of Patients With Acute Myeloblastic Leukemia Who Benefit From the Addition of Gemtuzumab Ozogamicin: Results of the MRC AML15 Trial. *J Clin Oncol* **29**, 369-377, doi:10.1200/jco.2010.31.4310 (2011).

88 Hills, R. K. *et al.* Addition of gemtuzumab ozogamicin to induction chemotherapy in adult patients with acute myeloid leukaemia: a meta-analysis of individual patient data from randomised controlled trials. *Lancet Oncol* **15**, 986-996, doi:10.1016/s1470-2045(14)70281-5 (2014).

89 Knapp, D. J. H. F. *et al.* Single-cell analysis identifies a CD33+ subset of human cord blood cells with high regenerative potential. *Nat Cell Biol* **20**, 710-720, doi:10.1038/s41556-018-0104-5 (2018).

90 Godwin, C. D., Gale, R. P. & Walter, R. B. Gemtuzumab ozogamicin in acute myeloid leukemia. *Leukemia* **31**, 1855-1868, doi:10.1038/leu.2017.187 (2017).

91 Goldberg, A. D., Allis, C. D. & Bernstein, E. Epigenetics: A Landscape Takes Shape. *Cell* **128**, 635-638, doi:10.1016/j.cell.2007.02.006 (2007).

92 Figueroa, M. E. *et al.* MDS and secondary AML display unique patterns and abundance of aberrant DNA methylation. *Blood* **114**, 3448-3458, doi:10.1182/blood-2009-01-200519 (2009).

93 Rasmussen, K. D. & Helin, K. Role of TET enzymes in DNA methylation, development, and cancer. *Genes Dev* **30**, 733-750, doi:10.1101/gad.276568 (2016).
94 Figueroa, M. E. *et al.* DNA Methylation Signatures Identify Biologically Distinct Subtypes in Acute Myeloid Leukemia. *Cancer Cell* **17**, 13-27, doi:10.1016/j.ccr.2009.11.020 (2010).
95 Wouters, B. J. & Delwel, R. Epigenetics and approaches to targeted epigenetic therapy in acute myeloid leukemia. *Blood* **127**, 42-52, doi:10.1182/blood-201507-604512 (2016).
96 Sekeres, M. A. & Steensma, D. P. Boulevard of broken dreams: drug approval for older adults with acute myeloid leukemia. *J Clin Oncol* **30**, 4061-4063 (2012).
97 Mayle, A. *et al.* Dnmt3a loss predisposes murine hematopoietic stem cells to malignant transformation. *Blood*, doi:10.1182/blood2014-08-594648 (2015).
98 Hope, K. J., Jin, L. & Dick, J. E. Human acute myeloid leukemia stem cells. *Arch Med Res* **34**, 507-514, doi:10.1016/j.arcmed.2003.08.007 (2003).
99 Hope, K. J., Jin, L. & Dick, J. E. Acute myeloid leukemia originates from a hierarchy of leukemic stem cell classes that differ in self-renewal capacity. *Nat Immunol* **5**, 738-743, doi:10.1038/ni1080 (2004).
100 Brenner, S. *et al.* CXCR4-Transgene Expression Significantly Improves Marrow Engraftment of Cultured Hematopoietic Stem Cells. *Stem Cells* **22**, 1128-1133, doi:10.1634/stemcells.2003-0196 (2004).
101 Hofmeister, C. C., Zhang, J., Knight, K. L., Le, P. & Stiff, P. J. Ex vivo expansion of umbilical cord blood stem cells for transplantation: growing knowledge from the hematopoietic niche. *Bone Marrow Transplant* **39**, 11-23, doi:10.1038/sj.bmt.1705538 (2006).
102 Bradley, T. R. & Metcalf, D. The growth of mouse bone marrow cells in vitro. *Aust J Exp Biol Med Sci* **44**, 287-299 (1966).
103 Metcalf, D., Moore, M. A. & Warner, N. L. Colony formation in vitro by myelomonocytic leukemic cells. *J Natl Cancer Inst* **43**, 983-1001 (1969).
104 Metcalf, D. Hematopoietic cytokines. *Blood* **111**, 485-491 (2008).
105 Sachs, L. The differentiation of myeloid leukaemia cells: new possibilities for therapy. *Br J Haematol* **40**, 509-517 (1978).
106 Shultz, L. D., Ishikawa, F. & Greiner, D. L. Humanized mice in translational biomedical research. *Nat Rev Immunol* **7**, 118-130, doi:10.1038/nri2017 (2007).
107 Legrand, N., Weijer, K. & Spits, H. Experimental Models to Study Development and Function of the Human Immune System In Vivo. *J Immunol* **176**, 2053-2058, doi:10.4049/jimmunol.176.4.2053 (2006).
108 Zhang, B., Duan, Z. & Zhao, Y. Mouse models with human immunity and their application in biomedical research. *J Cell Mol Med* **13**, 1043-1058, doi:10.1111/j.1582-4934.2008.00347.x (2009).
109 Bosma, G. C., Custer, R. P. & Bosma, M. J. A severe combined immunodeficiency mutation in the mouse. *Nature* **301**, 527-530 (1983).

110 Sachlos, E. *et al.* Identification of Drugs Including a Dopamine Receptor Antagonist that Selectively Target Cancer Stem Cells. *Cell* **149**, 1284-1297, doi:10.1016/j.cell.2012.03.049 (2012).
111 Wunderlich, M. *et al.* AML cells are differentially sensitive to chemotherapy treatment in a human xenograft model. *Blood* **121**, e90-97, doi:10.1182/blood-2012-10464677 (2013).
112 Krivtsov, A. V. *et al.* Transformation from committed progenitor to leukaemia stem cell initiated by MLL–AF9. *Nature* **442**, 818-822, doi:10.1038/nature04980 (2006).
113 Tan, Y. T. *et al.* Respecifying human iPSC-derived blood cells into highly engraftable hematopoietic stem and progenitor cells with a single factor. *Proc Natl Acad Sci U S A* **115**, 2180-2185 (2018).
114 Sutherland, H. J., Blair, A. & Zapf, R. W. Characterization of a hierarchy in human acute myeloid leukemia progenitor cells. *Blood* **87**, 4754-4761 (1996).
115 Waddington, C. H. *Waddington, C. H.*, (George Allen & Unwin, 1957).
116 GURDON, J. B., ELSDALE, T. R. & M., F. Sexually mature individuals of Xenopus laevis from the transplantation of single somatic nuclei. *Nature* **182**, 64-65 (1958).
117 GURDON, J. B. The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. *J Embryol Exp Morphol* **10**, 622-640 (1962).
118 Weismann, A. *The Germ-Plasm: A Theory of Heredity*. (Charles Scribner's Sons, 1893).
119 Davis, R. L., Weintraub, H. & Lassar, A. B. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**, 987-1000 (1987).
120 Takahashi, K. *et al.* Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* **131**, 861-872, doi:10.1016/j.cell.2007.11.019 (2007).
121 Takahashi, K. & Yamanaka, S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* **126**, 663-676, doi:10.1016/j.cell.2006.07.024 (2006).
122 Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917-1920 (2007).
123 Park, I. H. *et al.* Disease-Specific Induced Pluripotent Stem Cells. *Cell* **134**, 877-886, doi:10.1016/j.cell.2008.07.041 (2008).
124 Hochedlinger, K. & Jaenisch, R. Nuclear reprogramming and pluripotency. *Nature* **441**, 1061-1067 (2006).
125 Takahashi, K. & Yamanaka, S. A decade of transcription factor-mediated reprogramming to pluripotency. *Nat Rev Mol Cell Biol* **17**, 183-193, doi:10.1038/nrm.2016.8 (2016).
126 Yamanaka, S. A Fresh Look at iPS Cells. *Cell* **137**, 13-17, doi:10.1016/j.cell.2009.03.034 (2009).
127 Hanna, J. *et al.* Direct Reprogramming of Terminally Differentiated Mature B Lymphocytes to Pluripotency. *Cell* **133**, 250-264, doi:10.1016/j.cell.2008.03.028 (2008).

128 Stadtfeld, M., Brennand, K. & Hochedlinger, K. Reprogramming of Pancreatic β Cells into Induced Pluripotent Stem Cells. *Current Biology* **18**, 890-894, doi:10.1016/j.cub.2008.05.010 (2008).
129 Soufi, A., Donahue, G. & Zaret, K. S. Facilitators and Impediments of the Pluripotency Reprogramming Factors' Initial Engagement with the Genome. *Cell* **151**, 994-1004, doi:10.1016/j.cell.2012.09.045 (2012).
130 Buganim, Y. *et al.* Single-Cell Expression Analyses during Cellular Reprogramming Reveal an Early Stochastic and a Late Hierarchic Phase. *Cell* **150**, 1209-1222, doi:10.1016/j.cell.2012.08.023 (2012).
131 Friedmann-Morvinski, D. & Verma, I. M. Dedifferentiation and reprogramming: origins of cancer stem cells. *EMBO Rep* **15**, 244-253, doi:10.1002/embr.201338254 (2014).
132 Hochedlinger, K. *et al.* Reprogramming of a melanoma genome by nuclear transplantation. *Genes Dev* **18**, 1875-1885, doi:10.1016/j.molmed.2012.04.001. (2004).
133 Li, L., Connelly, M. C., Wetmore, C., Curran, T. & Morgan, J. I. Mouse embryos cloned from brain tumors. *Cancer Res* **63**, 2733-2736 (2003).
134 Liu, X. S. *et al.* Editing DNA Methylation in the Mammalian Genome. *Cell* **167**, 233-247.e217, doi:10.1016/j.cell.2016.08.056 (2016).
135 Baumann, V. *et al.* Targeted removal of epigenetic barriers during transcriptional reprogramming. *Nature Communications* **10**, doi:10.1038/s41467-019-10146-8 (2019).
136 Yamanaka, S. Strategies and New Developments in the Generation of Patient-Specific Pluripotent Stem Cells. *Cell Stem Cell* **1**, 39-49, doi:10.1016/j.stem.2007.05.012 (2007).
137 Kawamura, T. *et al.* Linking the p53 tumour suppressor pathway to somatic cell reprogramming. *Nature* **460**, 1140-1144, doi:10.1038/nature08311 (2009).
138 Marión, R. M. *et al.* A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. *Nature* **460**, 1149-1153, doi:10.1038/nature08287 (2009).
139 Hong, H. *et al.* Suppression of induced pluripotent stem cell generation by the p53–p21 pathway. *Nature* **460**, 1132-1135, doi:10.1038/nature08235 (2009).
140 Utikal, J. *et al.* Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. *Nature* **460**, 1145-1148, doi:10.1038/nature08285 (2009).
141 Knaupp, A. S. *et al.* Transient and Permanent Reconfiguration of Chromatin and Transcription Factor Occupancy Drive Reprogramming. *Cell Stem Cell* **21**, 834-845.e836, doi:10.1016/j.stem.2017.11.007 (2017).
142 Buganim, Y., Faddah, D. A. & Jaenisch, R. Mechanisms and models of somatic cell reprogramming. *Nature Rev Genet* **14**, 427-439, doi:10.1038/nrg3473 (2013).
143 Huangfu, D. *et al.* Induction of pluripotent stem cells by defined factors is greatly improved by small molecule compounds. *Nat Biotechnol* **26**, 795-797 (2008).

144 Ferreira, L. M. R. & Mostajo-Radji, M. A. How induced pluripotent stem cells are redefining personalized medicine. *Gene* **520**, 1-6, doi:10.1016/j.gene.2013.02.037 (2013).
145 Lee, J. H. *et al.* Brief Report: Human Acute Myeloid Leukemia Reprogramming to Pluripotency Is a Rare Event and Selects for Patient Hematopoietic Cells Devoid of Leukemic Mutations. *Stem Cells* **35**, 2095-2102, doi:10.1002/stem.2655 (2017).
146 Muñoz-López, A. *et al.* Development Refractoriness of MLL-Rearranged Human B Cell Acute Leukemias to Reprogramming into Pluripotency. *Stem Cell Reports* **7**, 602-618, doi:10.1016/j.stemcr.2016.08.013 (2016).
147 Kim, J. *et al.* An iPSC Line from Human Pancreatic Ductal Adenocarcinoma Undergoes Early to Invasive Stages of Pancreatic Cancer Progression. *Cell Reports* **3**, 2088-2099, doi:10.1016/j.celrep.2013.05.036 (2013).
148 Carette, J. E. *et al.* Generation of iPSCs from cultured human malignant cells. *Blood* **115**, 4039-4042, doi:10.1182/blood-2009-07-231845 (2010).
149 Doulatov, S. *et al.* Drug discovery for Diamond-Blackfan anemia using reprogrammed hematopoietic progenitors. *Sci Transl Med* **9** (2017).
150 Hu, K. *et al.* Efficient generation of transgene-free induced pluripotent stem cells from normal and neoplastic bone marrow and cord blood mononuclear cells. *Blood* **117**, e109-119, doi:10.1182/blood-2010-07-298331 (2011).
151 Kumano, K. *et al.* Generation of induced pluripotent stem cells from primary chronic myelogenous leukemia patient samples. *Blood* **119**, 6234-6242, doi:10.1182/blood-2011-07-367441 (2012).
152 Bedel, A. *et al.* Variable Behavior of iPSCs Derived from CML Patients for Response to TKI and Hematopoietic Differentiation. *PLoS One* **8**, e71596, doi:10.1371/journal.pone.0071596 (2013).
153 Ye, Z. *et al.* Human-induced pluripotent stem cells from blood cells of healthy donors and patients with acquired blood disorders. *Blood* **114**, 5473-5480, doi:10.1182/blood-2009-04-217406 (2009).
154 Kotini, A. G. *et al.* Stage-Specific Human Induced Pluripotent Stem Cells Map the Progression of Myeloid Transformation to Transplantable Leukemia. *Cell Stem Cell* **20**, 315-328.e317, doi:10.1016/j.stem.2017.01.009 (2017).
155 Hsu, J. *et al.* Reprogramming identifies functionally distinct stages of clonal evolution in myelodysplastic syndromes. *Blood*, blood.2018884338, doi:10.1182/blood.2018884338 (2019).
156 Chao, M. P. *et al.* Human AML-iPSCs Reacquire Leukemic Properties after Differentiation and Model Clonal Variation of Disease. *Cell Stem Cell* **20**, 329-344.e327, doi:10.1016/j.stem.2016.11.018 (2017).
157 Blelloch, R. *et al.* Reprogramming Efficiency Following Somatic Cell Nuclear Transfer Is Influenced by the Differentiation and Methylation State of the Donor Nucleus. *Stem Cells* **24**, 2007-2013, doi:10.1634/stemcells.2006-0050 (2006).
158 Chadwick, K. Cytokines and BMP-4 promote hematopoietic differentiation of human embryonic stem cells. *Blood* **102**, 906-915, doi:10.1182/blood-2003-03-0832 (2003).

159 Takashima, Y. *et al.* Resetting Transcription Factor Control Circuitry toward Ground-State Pluripotency in Human. *Cell* **158**, 1254-1269, doi:10.1016/j.cell.2014.08.029 (2014).
160 Eminli, S. *et al.* Differentiation stage determines potential of hematopoietic cells for reprogramming into induced pluripotent stem cells. *Nat Genet* **41**, 968-976, doi:10.1038/ng.428 (2009).
161 Lee, J. H. *et al.* Lineage-Specific Differentiation Is Influenced by State of Human Pluripotency. *Cell Reports* **19**, 20-35, doi:10.1016/j.celrep.2017.03.036 (2017).
162 Lacombe, F. *et al.* Flow cytometry CD45 gating for immunophenotyping of acute myeloid leukemia. *Leukemia* **11**, 1878-1886 (1997).
163 Vo, T. T. *et al.* Relative Mitochondrial Priming of Myeloblasts and Normal HSCs Determines Chemotherapeutic Success in AML. *Cell* **151**, 344-355, doi:10.1016/j.cell.2012.08.038 (2012).
164 Cho, R. H., Sieburg, H. B. & Muller-Sieburg, C. E. A new mechanism for the aging of hematopoietic stem cells: aging changes the clonal composition of the stem cell compartment but not individual stem cells. *Blood* **111**, 5553-5561, doi:10.1182/blood-200711-123547 (2008).