Creating the AML-specific t(9;11) chromosomal translocation model to study the structural basis of leukemia development in mouse cells

Aleksandra Sparavier

TESI DOCTORAL UPF 2021

DIRECTORS DE LA TESI

Dr. Marc A. Marti-Renom

Dr. Luciano Di Croce

GENE REGULATION, STEM CELLS AND CANCER
CENTRE FOR GENOMIC REGULATION (CRG)

DEPARTMENT OF EXPERIMENTAL AND HEALTH SCIENCES
UNIVERSITAT POMPEU FABRA

upf. Universitat Pompeu Fabra Barcelona

cnag centre nacional d'anàlisi genòmica centro nacional de análisis genómico

CRG Centre for Genomic Regulation

ACKNOWLEDGEMENTS

First of all, I want to thank my Parents for their constant support and love. I wouldn´t be where I am without you. Thanks to my brother Michał for inspiring me to always aim higher and face challenges head on. Ania and Gosia, thanks for being the best support system I could ever imagine! A big hug to Asia, for always being there for me and the connection which does not depend on physical distance. Patryk, thanks for being the best brother-from-another-mother I never knew I needed.

Next, I would like to express my gratitude to my supervisors Marc and Luciano. Thank you for the trust you put in me and support during my PhD!

A big hug to the Marcius lab! Thanks for making me feel part of the family from the very beginning. That extends to all of the former members I had the pleasure to get to know, Yasmina and François, (I really appreciate your help when I arrived here), Mike, Silvia, Julen, Marco, Juan, it was a pleasure to interact on the professional and personal level. Irene, thanks for being "la mamma" of the group, always willing to discuss and share advice! Fra and David, thanks for all the computational-related advice, it helped me more than you can imagine☺. Maria, I appreciate all the discussions about my project, you helped me a lot. Meritxell, good luck on your PhD journey!

A huge thank you to the whole Di Croce family!! Ara, thanks for always being willing to help, in terms of experiments and beyond. Sergi, Ceci, François and Pedro, I appreciate a lot all the advice you gave me over the years and help with troubleshooting. All the former PhD students, Anna, Paul, Marc, Mar, Isa, it was a lot of fun spending time together! Ivo, thanks for being so insightful and ready to help, I benefited a lot from that. Enrique, thanks for your critical thinking and asking a lot of questions, it helped me to approach my work from a different perspective. Gianni and Livia, you always make the atmosphere in the lab lively, keep that☺ Micol, all the best for your PhD and most important enjoy it!! Pau, thanks for being effortlessly funny and always lightening the atmosphere;)

ABBREVIATIONS

AGM - aorta-gonad-mesonephros
AML - acute myeloid leukemia
ATRA - all-trans retinoic acid
bp - basepair
CBX - chromobox
CLP - common lymphoid progenitor
CMP - common myeloid progenitor
cPRC1 - canonical polycomb repressive complex 1
CRISPR - clustered regularly interspaced short palindromic repeats
CRISPR/Cas9 - CRISPR-associated protein 9
DNA - deoxyribonucleic acid
DNMT - DNA methyltransferases
dox - doxycycline
DSB - double-strand break
E - embryonic day
E14 - strain of wild-type mouse ESCs
EB - embryoid body
EHT -endothelial-to-hematopoietic transition
ESC - embryonic stem cell
GMP - granulocyte-monocyte progenitor
HAT - histone acetyltransferase
HB - hemangioblast
HDAC - histone deacetylases
HDMs - histone demetylases
HE - hemogenic endothelium
HMT - histone methyltransferase
HP - hematopoietic progenitor
HR - homologous recombination
HSC - hematopoietic stem cell
IAHC - intra-aortic hematopoietic clusters
IB - immunoblot
IF - immunofluorescence
iPSC - induced pluripotent stem cell
kDa - kilodalton
KO - knockout
MLL - mixed lineage leukemia
MLLT3 - MLLT3 Super Elongation Complex Subunit
ncPRC1 - noncanonical polycomb repressive complex 1
NHEJ - non-homologous end joining
NLS -nuclear localization signal
PcG - Polycomb group
PCR - polymerase chain reaction
PML - promyelocytic leukemia protein
PRC - Polycomb repressive complex
qPCR - quantitative PCR
RARA - retinoic acid receptor alpha

RT - reverse transcriptase
TAD - topologically associated domain
TALEN - transcription activator-like effector nuclease
TET - ten-eleven translocation
TF - transcription factor
TrxG - Trithorax group
VEGF - vascular endothelial growth factor receptor
VSM - vascular smooth muscle
WB - western blot
YEATS - (Yaf9, ENL, AF9, Taf14, Sas5)

ABSTRACT

Chromosomal translocations are large-scale genome rearrangements found in several cancers. They are formed when two (or more) non-homologous chromosomes interchange large parts as a result of erroneous DNA repair after concomitant double-strand breaks. They are especially common in leukemias, where one of the possible outcomes is the formation of fusion genes, e.g., MLL-AF9 in t(9;11)-related leukemia. Most studies so far focused solely on the function of the oncoprotein, omitting the possible effects caused by the translocation, namely the genome structure change and the heterozygous KO of the translocation partner genes. To study more in detail these mechanisms, we apply the CRISPR/Cas9 and gene-trap technology to build the equivalent of t(9;11) in mouse ESCs, namely the t(4;9) translocation, with conditional Mll-Af9 expression. We report that the induced translocation alone has only a seemingly mild effect on the hematopoietic differentiation. Strikingly however, triggering Mll-Af9 expression in ESCs abrogates the hemangioblast emergence. Collectively, our strategy allows for deciphering the different layers contributing to leukemogenesis.

RESUMEN

Las translocaciones cromosómicas son reordenamientos del genoma a gran escala observados en varios cánceres. Se forman cuando dos (o más) cromosomas no homólogos intercambian partes como resultado de una reparación errónea del ADN después de roturas concomitantes de la cromatina. Son especialmente comunes en las leucemias, donde su efecto principal es la formación de genes de fusión, como por ejemplo MLL-AF9 en la leucemia relacionada con t(9;11). La mayoría de los estudios hasta ahora se han centrado únicamente en la función de la oncoproteína, omitiendo los posibles efectos causados por la translocación, como el cambio de estructura del genoma y el KO heterocigoto de los genes asociados a la translocación. Para estudiar más a fondo estos mecanismos, hemos aplicado CRISPR/Cas9 y la tecnología gene-trap para construir el equivalente de t(9;11) en las células madre (ESCs) de ratón (t(4;9)), con expresión condicional de Mll-Af9. Descubrimos que, si la translocación inducida por sí sola tiene sólo un efecto aparentemente leve sobre la diferenciación hematopoyética, la expresión de Mll-Af9 en ESCs bloquea la emergencia del hemangioblast. En conjunto, nuestra estrategia permite descifrar los diferentes mecanismos que contribuyen a la leucemogénesis.

Table of contents

INTRODUCTION

1. Preface

The aim of this Thesis was to unravel the mechanisms involved in the leukemia related with the t(9;11) chromosomal translocation. As a model system, we used mouse embryonic stem cells that we engineered to have the equivalent translocation, t(4;9) and that we differentiated towards hematopoietic lineages. Given that leukemia is a group of hematological malignancies, a general overview of hematopoiesis will be presented first. Despite occurring predominantly in adults, the t(9;11) translocation can also arise during childhood, and even *in utero*, and for that reason both embryonic and adult hematopoietic processes will be described. We are using one of the available hematopoietic differentiation protocols, so to put the reader in context, several alternatives will be compared next. This will be followed by a general overview of leukemias, with a more detailed description of AML, in line with the focus of the Thesis. Chromosomal translocations cause large changes in the nucleus of the cells, so in order to put them in perspective, some key points on the genome structure and regulation will also be presented. Why and how these translocations occur is still not known, but several dependencies in terms of targeted regions and cell-specific processes have been reported, and they will be discussed. Lastly, the t(9;11)-related leukemia will be presented, with specific focus on the translocation partner genes (MLL and AF9), the expressed MLL-AF9 fusion protein and currently used model systems.

2. Hematopoiesis

Hematopoiesis defines a complex and tightly regulated process of blood cell types formation in the organism. There are two major types of hematopoiesis, based on the developmental timepoint when they occur, namely embryonic and adult. Although the goal of both is the same, they present some major differences, as pointed out below. In view of the focus of the thesis, hematopoiesis will be discussed taking into consideration mouse development, which however closely resembles the human one.

2.1 Embryonic hematopoiesis

In the developing embryo, two waves or phases of hematopoiesis can be observed. The first, primitive one, starts extra-embryonically, in the blood islands of the yolk sac (Figure I.1). This is followed by the so-called definitive wave, which occurs inside of the embryo, mostly in the aorta-gonad-mesonephros (AGM) region and later in the fetal liver.

The yolk sac hematopoiesis starts around E7 with the appearance of transcription factors such as Gata2 and Tal1 in the regions where the blood islands will be subsequently formed (at E7.5). At this time, the hematopoietic cell population consists mostly of primitive erythrocytes and macrophages. Paradoxically, they arise in the absence of hematopoietic stem cells (HSCs), pointing to the scenario where blood cells are produced skipping the “classical” HSC differentiation pathway. The emergence of HSCs in the yolk sac is

still under debate. Some studies report the appearance of first hematopoietic stem cells around E8-8.5 and their subsequent migration to the fetal liver, where their maturation occurs. However, other reports postulate the later, intraembryonic HSC production. At E10-11, yolk sac hematopoiesis begins to rapidly decline, in favor of other sites present inside of the embryo (Dzierzak, 1999; Ganuza et al., 2018; Keller et al., 1999; Stefanska et al., 2017; Yamane, 2018).

The definitive, intraembryonic hematopoiesis begins in the aorta-gonad-mesonephros (AGM) region at E9, with the emergence of first hematopoietic cells from the intra-aortic hematopoietic clusters (IAHC), attached to the aortic endothelium of the dorsal aorta (DA). Sometimes this first part is also referred to as the pro-definitive wave, as depicted in Figure I.1. The hematopoietic stem cells are detectable from E10.5, and they express the following markers: cKIT, CD45, CD34 and VE-cadherin. Nevertheless, it is still under debate if the AGM can generate all cell populations *de novo*, as there are contradicting studies about it. On one hand, it was reported that embryos with Ncx1 or Rac1 knockouts (KOs), deprived of proper cell migration capacities, displayed normal yolk sac functioning but no hematopoiesis in the AGM, pointing towards the first being the site of blood cell lineages production. On the other hand, however, there are studies confirming the capability of the AGM to give rise to HSCs (Medvinsky et al., 2011).

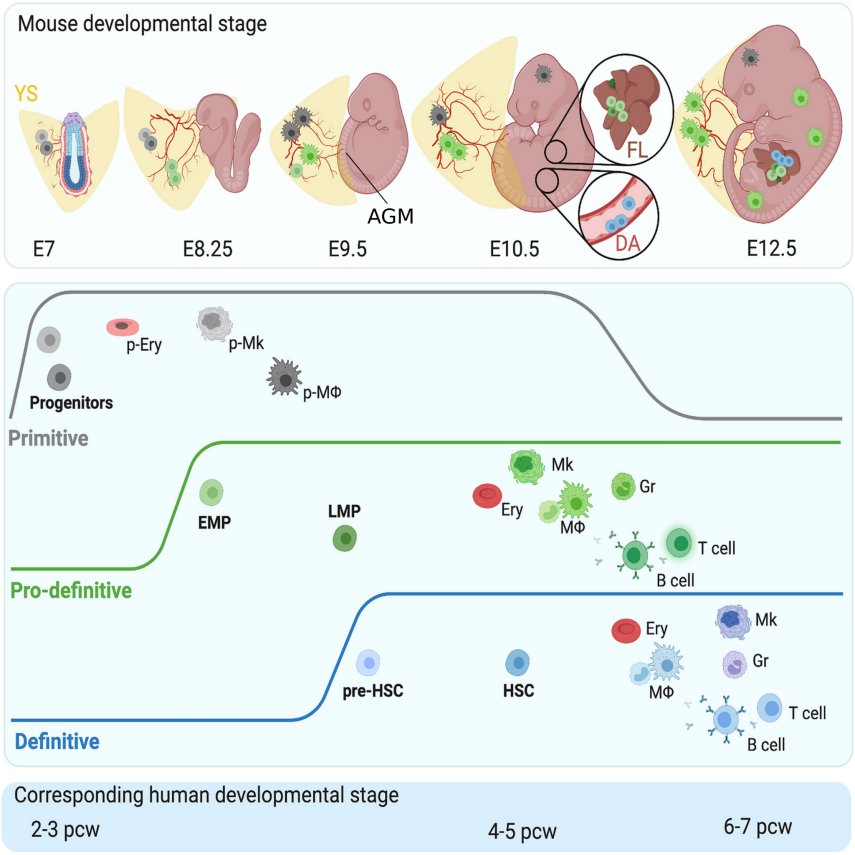


Figure I.1. Developmental stages of mouse embryonic hematopoiesis and corresponding timepoint for human (modified from Canu and Ruhrberg, 2021).

Other regions with hematopoietic activity include e.g., the placenta and the region of the embryonic head. The placenta is an organ with highly vascularized structure, and displays hematopoietic activity around E8.5-13.5, being a site of HSC generation and expansion mid-gestation. After that, the cells migrate to the fetal liver, where they mature and acquire the definitive phenotype (Sugiyama et al., 2011). In parallel, HSCs were also detected in the embryonic head, around E10.5-115, with multilineage reconstitution abilities and self-renewal capacity, as observed after transplanting them into adult recipients (Li et al., 2012).

A very important site for definitive hematopoiesis is the fetal liver. It forms at E9, with definitive erythrocytes detectable from E10 and macrophages and B cells slightly later, from E10-11. The hematopoietic stem cells are not produced here, but they migrate from the yolk sac (arguably, around E9) and the AGM region (E10), for the maturation step. The proper, definitive long-term HSCs (LT-HSCs) are visible from E11, expressing cKIT, Sca1 and CD45.

Finally, around E16.5, hematopoietic stem cells move to the bone marrow and permanently populate this niche. From this point forward, the bone marrow becomes a reservoir of hematopoietic potential in the developing organism (Sugiyama et al., 2011).

The proper functioning of the hematopoietic system in the developing embryo requires extremely precise and coordinated regulation of many processes. There are two main groups of hematopoietic regulators, controlling the primitive or subsequent definitive hematopoietic wave. In the first, primitive hematopoiesis step, the transcription factor Tal1 was shown to play a major role. Its homozygous KO completely abrogated the yolk sac hematopoiesis and caused embryonic lethality by E10.5, pointing to its role in the initial processes. A similar phenotype was observed for Rbtn2, where KO caused a complete lack of hematopoietic cells. Other factors important at this stage include Lmo2, Fli1 and Klf1. It is not yet understood how this primitive program gets shut down in favor of the definitive one, but some pluripotency transcription factors might play a role, such as Nanog. De Aja and colleagues (2019) showed that it

is responsible for blocking the erythroid lineage in the epiblast of the gastrulating embryo, by directly repressing Tal1.

The definitive hematopoiesis is more complex and therefore requires more "players". Among them, a very important one is Runx1, which was shown to be required specifically in this second wave. Its absence in the embryos did not affect the yolk sac erythropoiesis, but caused a complete lack of definitive lineages and death at E12.5. Other transcription factors playing a role exclusively in the definitive hematopoiesis are e.g., PU.1 (whose lack causes death at E17.5), C-Myb and Jumonji. Apart from all these, also signaling pathways are crucial for development of the hematopoietic system in the embryo and seem to follow this differential pattern, like the Wnt and Notch signaling pathways. With respect to the latter one, the absence of Notch1 did not affect the yolk sac, but caused a decreased HSC potential in the body of the embryo.

Although the two waves are spatially and temporally separated, they still share some regulation mechanisms and factors. Worth mentioning is Gata2, whose lack not only causes severe fetal liver anemia and death at E15, but also a reduction of primitive erythroid cells (Dzierzak, 1999; Medvinsky et al., 2011). Both Gata2 and Runx1 are important for definitive HSC development, since their haploinsufficiency was reported to cause a decreased HSC production in the AGM region (Huber et al., 2004; Yamane, 2018).

The hemangioblast hypothesis

The initial studies of embryonic hematopoiesis allowed for the observation that both in the yolk sac and the AGM region, the hematopoietic cells were present closely together with endothelial cells. Moreover, a comparative characterization showed that the two cell types share several surface markers, e.g., CD34, Flk1, Tie2, Tal1, Gata2 and Pecam1. These results led to the formation of the theory of a common precursor of hematopoietic and endothelial cell lineages, both in the primitive and the definitive waves of blood formation in the embryo (Keller et al., 1999; Medvinsky et al., 2011; Sugiyama et al., 2011).

More in-depth studies confirmed the presence of a common ancestor, which was named the hemangioblast. Using an *in vitro* model of ES cell differentiation, it was proved that this precursor was able to give rise to endothelial, hematopoietic and vascular smooth muscle cells (Figure I.2). An elegant proof of the hemangioblast potentiality was provided by Kennedy and colleagues (1997) and confirmed by Choi and colleagues (1998). They used a mix of 2 ES cell lines, carrying the neomycin or hygromycin resistance gene, respectively. The differentiation was performed in such a way that each formed colony, termed blast colony, was originating from a single cell. The colonies consisted of both adherent endothelial, and floating hematopoietic cells. Both lineages from each colony were analyzed separately, and it was shown that cells having the same origin had the same antibiotic resistance gene, meaning they were formed upon differentiation from a shared progeny (hemangioblast). Flk1, the vascular endothelial

growth factor receptor 2, is a surface marker present on the hemangioblast, and was shown to be crucial for proper hematopoietic development. Its homozygous KO resulted in embryonic death at E8.5-E9.5, with no detectable hematopoietic nor endothelial cells (Shalaby et al., 1995).

While it is widely accepted that the definitive blood cells arise from the hemangioblast, it is still under debate whether the primitive wave of hematopoiesis occurs following the same pathway. Some studies report this stepwise development of the e.g., primitive erythrocytes in the yolk sac (Gao et al., 2018; Lancrin et al., 2009), however it was shown that ES cell lines deprived of Flk-1 expression were able to give rise to (at least some) components of the primitive lineage (Drake and Fleming, 2000). More studies will be needed to shed light on these complex processes.

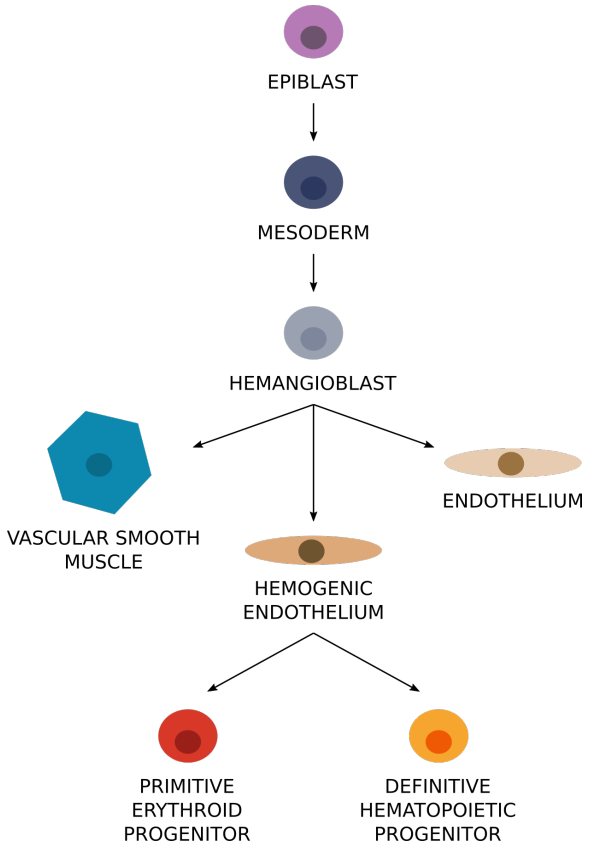


Figure I.2. Model of embryonic hematopoietic specification pathway (based on Lancrin et al., 2010)

Further experiments showed the presence of an intermediate cell population between the hemangioblast and the hematopoietic lineages. Microscopic observations of *in vitro*-derived cells and extracted tissues allowed the observation that the hemangioblast first differentiated to a specific endothelium, called hemogenic endothelium. Next, these intermediate, transient cells give rise to the hematopoietic progenitors in the endothelial-to-hematopoietic transition (EHT), a feature of the definitive (and presumably also primitive) embryonic hematopoiesis (Ema et al., 2006; Lancrin et al., 2009).

2.2 Adult hematopoiesis

After birth, the bone marrow becomes the main source of hematopoietic cells, and this event marks the start of adult hematopoiesis. In contrast to the embryonic stage, where blood cells appear before the hematopoietic stem cells (Medvinsky et al., 2011), adult hematopoiesis follows the well-known differentiation and population hierarchy, as depicted in Figure I.3. At the very top there is the (very limited in numbers) population of hematopoietic stem cells (Dzierzak and Bigas, 2018; Seita and Weissman, 2010; Wang and Wagers, 2011). They occupy the bone marrow niche, where they self-renew and, when needed, differentiate. This can happen in response to many stimuli, like lesions (bleeding, infections), chemotherapeutic agents or signals from the bone marrow niche (Fleming et al., 1993; Ikonomi et al., 2020; Passegué et al., 2005; Wilson et al., 2009).

When activated, HSCs enter the cell cycle and start to differentiate towards more specialized cell types. Historically, their specialization was thought to occur through discrete cell stages, from multipotent progenitors, through the restricted-lineage precursors, to the final step of fully differentiated blood cells (Figure I.3). With the development of more sensitive and accurate methods, it became clearer that in fact this process is much more complex and continuous. Moreover, stem and progenitor populations are shown to be heterogenous and possibly lineage-biased to some extent (Fleming et al., 1993; Laurenti and Göttgens, 2018; Wang and Wagers, 2011).

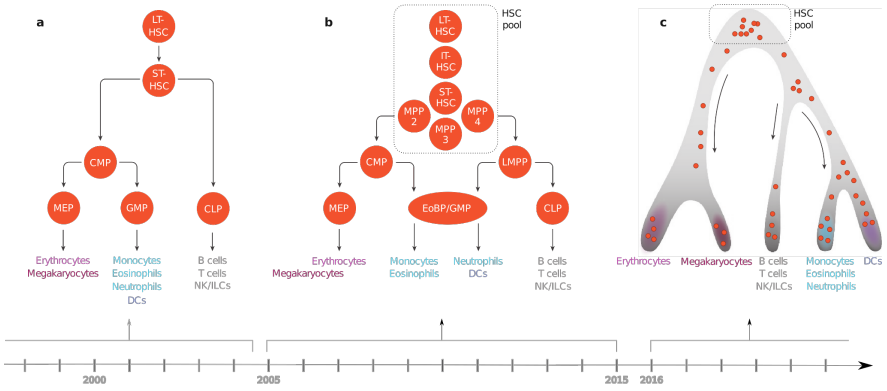


Figure I.3. Changing models of hematopoietic cell hierarchy (Modified from Laurenti and Göttgens, 2018)

Apart from different sequence of their appearance in the process of hematopoiesis, adult and embryonic hematopoietic stem cells have a number of differences. As opposed to their fetal counterparts, adult HSCs are quiescent, slowly-cycling and have less self-renewal potential. They also differ in maintenance transcription factors, as for example Gfi-1 and Tel/Etv6, crucial for adult but not fetal HSCs (Copley and Eaves, 2013; Kim et al., 2007).

2.3 Hematopoiesis-in-a-dish models

Considering the fact that the whole blood production system in the body relies on a very restricted in numbers and not easily available population of hematopoietic stem cells, it becomes evident why therapy of hematological malignancies is so challenging. The available strategy, HSC transplant, still presents major hurdles. HSCs can be isolated either from umbilical cord blood or peripheral blood upon their mobilization from the bone marrow with the recombinant

granulocyte-colony stimulating factor (G-CSF). Umbilical cord blood, however richer in HSC, provides only a limited number of cells. HSC mobilization presents an alternative, but is still not efficient in terms of recovered HSCs. In both cases, the most important limiting factor is the donor-patient immunocompatibility, crucial for a successful and long-term functioning transplant (Daniel et al., 2016; Lim et al., 2013).

In order to overcome these obstacles, much effort has been put into developing alternative solutions for obtaining transplantable hematopoietic stem cells. Currently, two main approaches are being investigated, namely *in vitro* HSC expansion and blood generation from alternative sources. So far, the efforts to culture HSCs outside of their niche (bone marrow) while maintaining their stemness have not been successful. Several small molecules, cytokines, and co-culture methods have been tested, however they allowed for only limited HSC expansion *ex vivo* (Mehta et al., 2015; Takizawa et al., 2011).

Recently, more studies aim to generate transplantable HSCs from other cell types, using (trans)differentiation. Tested starting material comprises pluripotent and differentiated cells, both human and mouse. In most cases, the process is induced through specific transcription factor (TF) overexpression, summarized in Table I.1.

Table I.1. TF-mediated HSC generation approaches.

Cell type	TFs	CFU	Engraft	Ref.
Human somatic cells				
fibroblasts	OCT4	yes	yes*	(Szabo et al., 2010)
	SOX2, miR125b	yes	yes*	(Pulecio et al., 2014)
endothelial cells	FOSB, GFI1, RUNX1, SPI1	yes	yes	(Sandler et al., 2014)
Mouse somatic cells				
fibroblasts	Gata2, Gfi1b, cFos, Etv6	yes*	no	(Pereira et al., 2013)
	Erg, Gata2, Lmo2, Runx1c, Scl	yes	yes*	(Batta et al., 2014)
	Scl, Gata2, Lmo2, Ptx2, Sox7, MycN	yes	NA	(Vereide et al., 2014)
	Scl, Lmo2, Runx1, Bmi1	yes	no	(Cheng et al., 2016)
committed blood progenitors	Runx1t1, Hlf, Lmo2, Prdm5, Pbx1, Zfp37, Mycn, Meis1	yes	yes	(Riddell et al., 2014)
endothelial cells	Fosb, Gfi1, Runx1, Spi1	yes	yes	(Lis et al., 2017)
Human pluripotent cells				
hESC	HOXB4	yes	NA	(Bowles et al., 2006)
	ERG,HOXA5, HOXA9, HOXA10, LCOR, RUNX1, SPI1	yes	yes	(Sugimura et al., 2017)
hESC-derived blood precursors	ERG, HOXA9, RORA, SOX4, MYB	yes	yes*	(Doulatov et al., 2013)
hESC, hiPSC	ETV2, GATA2/TAL1	yes	no	(Elcheva et al., 2014)
Mouse pluripotent cells				
mESC	HoxB4	yes	yes	(Kyba et al., 2002)
	Cdx4, HoxB4	yes	yes	(Wang et al., 2005b)
	Scl, Lmo2, Gata2, Pitx2, Sox7, MycN	yes	NA	(Vereide et al., 2014)

Engraft. - engraftment, * - only short-term (<6months), NA - not applicable (not tested)

Using pluripotent stem cells seems particularly tempting because of their virtually endless expansion ability and possibility to genetically engineer in a relatively straightforward way. One of the first attempts to differentiate mouse embryonic stem cells (ESCs) into hematopoietic lineages was performed by overexpressing Hoxb4 and co-culture with OP9 cells. Obtained hematopoietic cells were able to form colony-forming units and engraft irradiated mouse recipients (Kyba et al., 2002). The same strategy with the addition of Cdx4 allowed to improve the efficiency of the differentiation (Wang et al., 2005b). The hematopoiesis-promoting effect of HOXB4 expression in human embryonic stem cells is still under debate, with some studies reporting successful results (although without testing the engraftment) (Bowles et al., 2006), and others pointing otherwise (Wang et al., 2005a). Doulatov and colleagues (2013) explored a two-step human ESC (hESC) differentiation strategy, where they first isolated CD45+CD43+ myeloid progenitors from embryoid bodies, and then de-differentiated them using lentiviral-mediated overexpression (ERG, HOXA9, RORA, SOX4, MYB). Although the hematopoietic cells displayed colony forming activity, no lymphoid cells were detected and only transient engraftment was reported (Doulatov et al., 2013). An approach that yielded engraftable hematopoietic cells from hESCs was to overexpress the following TFs: ERG, HOXA5, HOXA9, HOXA10, LCOR, RUNX1 and SPI1. The differentiation consisted of isolating the population of hemogenic endothelium from EBs, which, exposed to cytokines, gave rise to hematopoietic cells through the endothelial-to-

hematopoietic transition (EHT), thus recapitulating the embryonic hematopoiesis (Sugimura et al., 2017).

In contrast to the ESC-mediated approach, mature cells, such as e.g., fibroblasts, can be transdifferentiated towards a different lineage using specific TFs. So far, experiments using mouse fibroblasts were not successful, as only in one case the obtained hematopoietic cells were able to engraft and even there, only transiently (Batta et al., 2014). In other cases, they either lacked this ability (Cheng et al., 2016; Pereira et al., 2013), or were not tested (Vereide et al., 2014). Using human fibroblasts also yielded transient engraftment (Pulecio et al., 2014; Szabo et al., 2010), pointing that more research is needed in order to develop HSCs from this cell type. Moreover, both studies took advantage of pluripotent factors (OCT4, SOX2), which are thought to promote transdifferentiation through a pluripotent intermediate, which has oncogenic potential (Daniel et al., 2016).

Using endothelial cells, both human (Sandler et al., 2014) and murine (Lis et al., 2017) turned out to be effective in terms of colony-forming activity and engraftment potential of the emerging hematopoietic cells. However, human cells displayed a myeloid bias with barely no lymphoid potential. Transdifferentiation of committed myeloid blood progenitors overcame this bias, yielding cells that showed robust multilineage potential (Riddell et al., 2014). However, the protocol had an *in vivo* step (in mice), which makes it nearly impossible to determine the exact induction signals and can affect the reproducibility.

All of the aforementioned studies are based on TF overexpression, which unfortunately limits their potential clinical applications. Most of them are virus-based, which poses a risk of oncogenesis by random insertion in the genome (e.g., through the disruption of tumor suppressor genes). Moreover, the forced TF expression does not reflect the hematopoietic events occurring physiologically, where they have a crucial role only for a restricted time. This in turn could have an effect going beyond the HSC generation in a dish, and affect the donor organism upon transplant. Lastly, some of the used factors are proto-oncogenes or pluripotent factors, creating a risk of oncogenic potential.

An alternative scenario would be to generate hematopoietic cells without TF overexpression. Long-term engraftable HSCs were obtained from both mouse and human induced pluripotent stem cells (iPSC), through *in vivo* teratoma formation (Amabile et al., 2013; Suzuki et al., 2013). However reproducible and useful in experimental work, their therapeutic use is limited due to the risk of contamination with oncogenic cells.

A different strategy stemmed from the work of the group of Gordon Keller and paved the way towards recapitulating the mouse embryonic hematopoiesis *in vitro*. They aimed to better characterize the embryonic hematopoietic differentiation pathway in order to re-create these steps *in vitro*. The first major accomplishment of the group was, in 1997, the discovery that primitive and definitive mouse embryonic hematopoietic waves share a common precursor, the

hemangioblast (HB) (Choi et al.; Kennedy et al., 1997). Many following studies of his and his alumni´s groups put great effort into deepening our knowledge in the hematopoietic specification. It was possible to determine some of the crucial regulators of differentiation, e.g., TFs, markers (Boros et al., 2011; Chen et al., 2009; Costa et al., 2012; Gandillet et al., 2009; Gilmour et al., 2014; Goode et al., 2016; Oatley et al., 2020; Serrano et al., 2010; Thambyrajah et al., 2018) as well as identify an intermediate population between the HB and hematopoietic populations. It was shown that the hemangioblast first differentiates towards a hemogenic endothelium (HE), which in turn undergoes endothelial-to-hematopoietic transition, yielding hematopoietic lineages (Lancrin et al., 2009). This sequential developmental pathway describes the events occurring in the mouse embryo (Boisset et al., 2010; Ema et al., 2006; Lancrin et al., 2010; Zovein et al., 2008), is regulated by the same TFs (Moignard et al., 2013) and is conserved amongst vertebrates (e.g., human, chick).

The protocol was adapted for specific experimental requirements, but in general it comprises forming embryoid bodies (3-4days), followed by sorting the hemangioblast (based on the Flk1 expression). After plating in hematopoiesis-promoting media (blast cultures), HB gives rise to the hemogenic endothelium first, and then, through EHT, to hematopoietic progenitors (3-4days).

In the context of leukemia, specifically childhood leukemia, it is still not known at what developmental stage chromosomal translocations

occur. This experimental strategy allows tackling this research question from a developmental perspective, and using ES cells as a starting point provides a possibility to first modify the cells (e.g., with CRISPR/Cas9 technology). Inducible expression of the RUNX1-RUNX1T1 oncogenic fusion protein allowed the dissection of the varying protein effects, depending on the differentiation stage where it was expressed. This approach allowed the observation that RUNX1-RUNX1T1 has the ability to block hematopoiesis specifically at the EHT (Regha et al., 2015).

Promising results, from the regenerative medicine point of view, came from the Kouskoff group (Pearson et al., 2015). Using specific conditions (serum-free media, temporally-restricted cytokine exposure), they generated cells able to give rise to all blood lineages and repopulate sublethally irradiated recipient mice. However, these activities did not come from differentiated hematopoietic cells, but earlier, at the hemogenic endothelium stage (expressing cKIT+ and a number of endothelial markers, like Tie-2 and ICAM2), corresponding to day 1-2 of blast culture. In fact, both Flk1+ hemangioblasts and later stages of blast culture (containing hematopoietic progenitors) failed to engraft the recipients. The successfully engrafted cells were detectable after 22 weeks, corresponding to long-term engraftment, but no secondary transplants were performed that would validate the self-renewal and repopulating activity of mESC-derived hematopoietic cells. Moreover, the study suggests the extreme sensitivity of the repopulation activity to cytokines, that goes beyond (and sometimes

opposite to) differentiation efficiency. Comparing the Bmp4+Activin A+VEGF condition with Activin A+VEGF, the absence of Bmp4 did not affect differentiation efficiency, but abrogated the engraftment ability. Moreover, differentiation yielded both primitive and definitive hematopoietic lineages, which might be problematic in terms of studying particular processes (Lacaud and Kouskoff, 2017; Sturgeon et al., 2013).

Generating *in vitro* transplantable HSCs is one of the major challenges of regenerative medicine. Substantial progress has been made since the first *in vitro* blood generation (Doetschman et al., 1985), but more research is needed to reach the goal.

3. Chromosomal translocations in leukemia

3.1 Acute myeloid leukemia

Hematopoiesis, as described above, needs tight multiple-level regulation and any threat to homeostasis can lead to hematological malignancies, such as leukemia. It is a very heterogeneous group of blood cancers, but at the base of each one of them lies the same cause, which is the uncontrolled and abnormal proliferation of one of the blood cell types (Crans and Sakamoto, 2001). Based on the affected cell lineage and the progression of the disease, 4 major leukemia subtypes have been described: acute myeloid (AML), chronic myeloid (CML), acute lymphocytic (ALL) and chronic lymphocytic leukemia (CLL). The distinction between acute and chronic is based on the differentiation stage when the developmental block occurs and the progression of the disease. In the case of acute leukemia, in contrast to the chronic one, overproliferation affects more immature progenitors and the disease progresses fairly quickly. I will focus on acute myeloid leukemia (AML), as it is the focus of the study presented here.

Recently, specific AML classifications have been described, taking into consideration the morphological properties of the patient blood cells (FAB classification), or cytogenetic features (WHO classification). The FAB classification divides AML into 8 subtypes (M0-M7), according to the maturity and origin of the leukemic cells. And so, groups M0-M5 define AML with myeloid origins, M6 -

coming from early red blood cells and M7 starts in cells that will become platelets (Zhang et al., 2021).

Alternatively, the WHO categorized AML into 6 groups, based on mutations or genomic rearrangements. They are as follows: AML with recurrent genetic abnormalities; AML with myelodysplasia-related changes (MRC); therapy-related myeloid neoplasms (t-MN); AML, not otherwise specified (NOS); myeloid sarcoma; and myeloid proliferations related to Down syndrome (DS). In most cases, leukemia appears *de novo*, but can also arise as a complication after specific treatments (t-MN), such as radiation therapy or topoisomerase II inhibitors (Hwang, 2020; Medeiros and Konoplev, 2017).

About 20–30% of patients with AML belong to the first group (AML with recurrent genetic abnormalities), with the most common aberrant feature being the presence of chromosomal translocations, such as t(8;21), t(9;11) or t(15;17) (Hwang, 2020). A large group of these rearrangements comprise the translocations involving the MLL gene, lying on the chromosome 11. MLL/KMT2A is a H3K4-specific methyltransferase, involved in crucial processes in the development and adult life of the organism. So far, more than 135 translocations involving MLL have been described, underlying its crucial role in homeostasis. Of them, the most common partner genes are AF4, AF9, ENL and ELL (Di Carlo et al., 2018).

Chromosomal translocations are events which occur upon movement of large chromatin fragments from one chromosome to another. Based on the fact whether a reciprocal translocation also occurs, they can be classified as balanced or unbalanced (loss of chromosome fragment). In order to better understand the mechanisms governing translocations and their effect, it is important to get insight into the organization and regulation of chromatin in the nucleus, which will be introduced below.

3.2 Genome structure

The information for proper cell functioning is encoded in the DNA, a string composed of two polynucleotide chains that form a double helix. In the nucleus, it is organized in a tightly regulated way, and potentially any aberrant change or misplacement can eventually lead to diseases. Genome structure has been depicted schematically in Figure I.4. At the highest resolution, the DNA is organized into nucleosomes (146bp wrapped around an octamer formed by two copies of each of the four histones: H2A, H2B, H3 and H4), which are the building blocks of chromatin. These are further organized by e.g., histone H1, which modulates their position and level of compactness. Going up with the scale, chromatin forms loops and topologically associated domains (TADs), which have been shown to be chromatin fragments with higher interaction inside compared to other genome regions. TADs are further segregated into higher-order structures, named compartments, which, depending on the transcriptional activity, are called A or B compartments. A compartment is related with regions of active gene expression,

whereas B is transcriptionally silent. The features of chromatin belonging to a specific type are also different: in A it is open and "relaxed" (euchromatin), whereas in B it is very tightly packed and closed (heterochromatin).

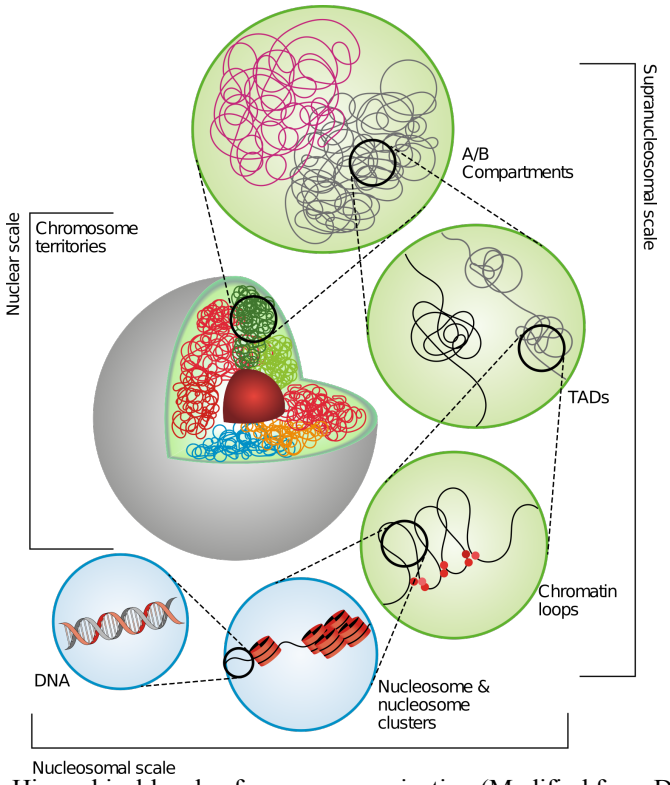


Figure I.4. Hierarchical levels of genome organization (Modified from Doğan and Liu, 2018).

On the highest organization scale, chromatin is divided into chromosomes, each occupying a non-random space in the nucleus. The specific position is often related to function and activity, with chromosomes with high numbers of active genes occupying more central positions in the nucleus and generally silent ones clustering towards the periphery and nuclear lamina (Meaburn et al., 2007).

Despite chromosomes occupying stable nuclear volumes, chromatin in a higher scale can undergo massive rearrangements. As an example, during ES cell differentiation towards cortical neurons, the compartments get bigger, interactions within B compartments increase, whereas within A compartments decrease (Zhang et al., 2019). Specific chromosome positioning also has implications in cancer. Statistical analysis showed that both the frequency of interaction between genomic loci, affected to some extent by spatial proximity, and selection pressure provided by the alteration are involved in the formation and propagation of genomic alterations in cancers (Fudenberg et al., 2011).

All the information regarding the structure of chromatin and the nucleus was discovered fairly recently, with the development of techniques such as microscopy, chromatin immunoprecipitation followed by sequencing (ChIPseq), and conformation-capture based techniques.

3.3 Regulation of gene expression

Since DNA is very highly packed in the nucleus, it needs additional factors to make the regions of interest accessible for e.g., the replication/transcription machinery. This regulation also has to be flexible and adapt to dynamic conditions, like changing gene expression during cell differentiation. One form of such regulation is provided through histones. They form the structural part of the nucleosomes, but their function goes well beyond being simple

bobbins. As part of nucleosomes, their relative position determines DNA accessibility and thus, activity status. Many other factors have the ability to rearrange this spatial conformation, like histone H1. Moreover, not all of the histone regions are “hidden” inside of the nucleosomes, their N- and C-terminal ends (tails) protrude to the outside and specific residues can be modified (Figure I.5A).

The 4 major modification types are methylation, acetylation, phosphorylation and ubiquitination. The name of the histone mark always carries the histone type, modified residue position (numbered from the N-terminus of the polypeptide chain), and the modification. And so, referring to a tri-methylated lysine in position four of the tail of histone H3, one would say H3K4me3. Histone marks carry various implications for chromatin characteristics and gene regulation (Figure I.5B). As example, H3K4me2/3, H3K9ac, H3K14ac, K3K27ac, and H4K16ac marks are a characteristic of euchromatin and regions with active transcription. In contrast, constitutive heterochromatin and silent genes are decorated with marks such as H3K9me3, H3K27me1/3, and H4K20me3 (Butler and Dent, 2013). An additional layer of complexity is provided by the DNA itself, which can also get methylated. It occurs on the cytosines in the context of CpG dinucleotides, where the methyl group is deposited in position 5 (5mC). DNA methylation is related to gene repression, and regions deprived of 5mC, called CpG islands, are found in active regions.

The presence of histone marks and DNA methylation status is modulated by chromatin-modifying enzymes. They can be divided into two groups, depending on whether they deposit (writers) or remove (erasers) the marks. Writers include histone acetyltransferases (HAT), histone methyltransferases (HMTs), DNA methyltransferases (DNMTs), and erasers comprise enzymes like histone deacetylases (HDACs) and histone demetylases (HDMs). They will be discussed in the context of hematopoiesis and their relation to leukemias.

Histone acetyltransferases are responsible for depositing the acetylation mark and thus activating transcription. The acetyl groups neutralize the positive lysine charge thus disrupting electrostatic interactions with the negatively-charged DNA backbone. As a result, chromatin becomes more open and accessible for transcriptional machinery. An example of such an enzyme is the monocytic leukemia zinc finger (MOZ) protein, which targets histone H3. MOZ is important for the emergence of hematopoietic progenitors, as homozygous KO mice displayed a lower number of such progenitors. However, even if fewer in numbers, the hematopoietic stem cells did not display any defect on lineage commitment and differentiation (Butler and Dent, 2013; Katsumoto et al., 2006).

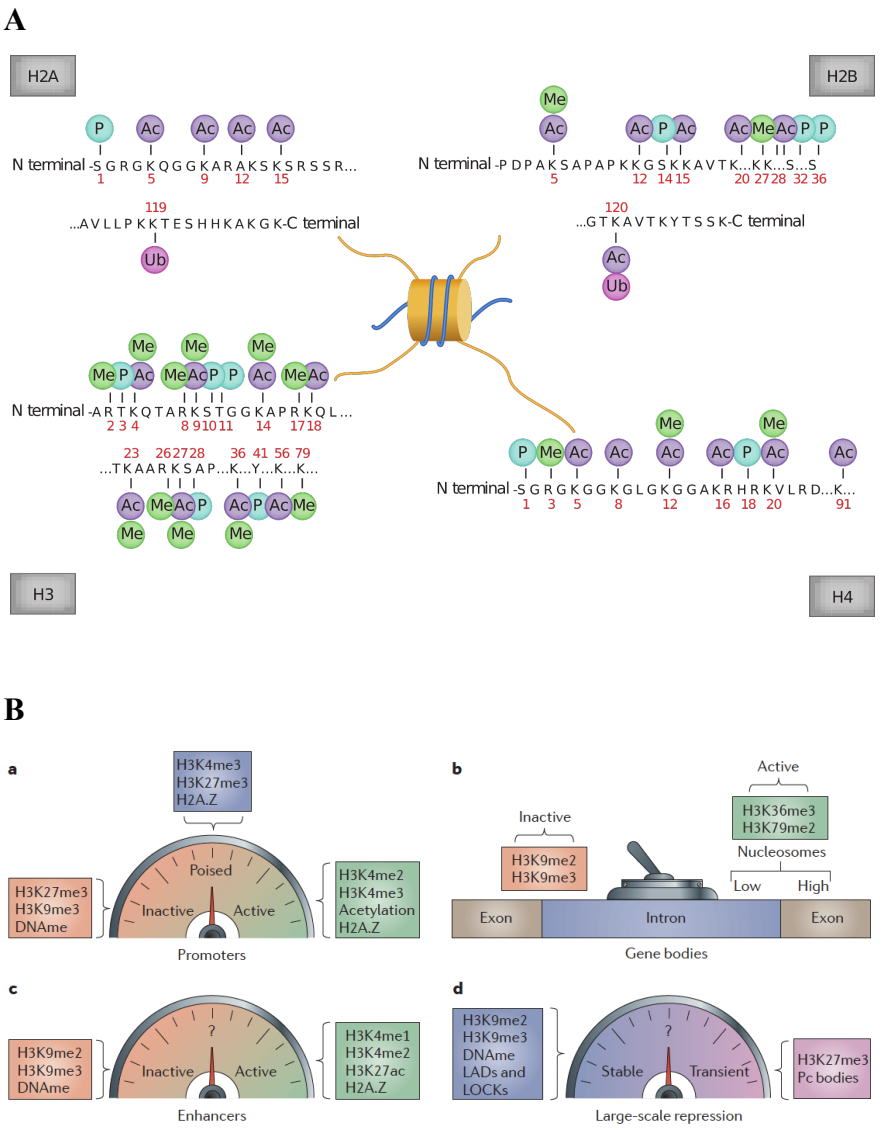


Figure I.5. **A** Histone modifications (Modified from Rodríguez-Paredes and Esteller, 2011) and **B** their effect on gene activity (Zhou et al., 2011).

Histone methyltransferases catalyze the deposition of methyl group(s) on lysines and arginines of the histone tails and, depending on the position, take part in gene activation or repression. H2K27me, catalyzed by EZH2 (discussed in more detail below), is related to

gene repression. In contrast, methylation of a different residue of the same histone, such as H3K79, implies active transcription. This mark is deposited by a disruptor of the telomeric silencing 1-like enzyme, DOT1L. It's crucial for embryo development, as mice deprived of this HMT die at E10.5-13.5. It has also been found to be involved in leukemias. In particular, in AML with t(9;11), the expressed MLL-AF9 fusion protein recruits DOT1L to its target genes and causes aberrant active transcription (Butler and Dent, 2013).

Another group of HMTs comprises the trithorax family of proteins (TrxG), also called COMPASS family. First discovered in *Drosophila*, it was shown to be crucial for embryo development. In mammals, it's composed of 6 proteins: MLL1/KMT2A (also called MLL), MLL2/KMT2B, MLL3/KMT2C, MLL4/KMT2D, SETD1A/KMT2F, SETD1B/KMT2G. There are some differences between the members in terms of structural domains (Figure I.6), but all share a conserved catalytic domain, the Su(var)3–9, Ezh2, Trithorax (SET) domain (Butler and Dent, 2013, 2013; Piunti and Shilatifard, 2016; Schuettengruber et al., 2007, 2017).

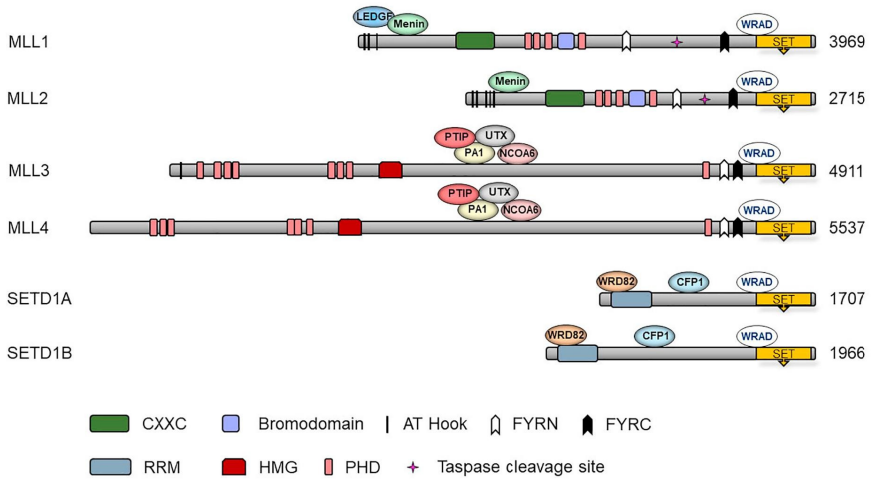


Figure I.6. Trithorax group of proteins: structures and interacting partners (from Antunes and Ottersbach, 2020).

MLL/SETD proteins are responsible for depositing the H3K3me1/2/3 marks, and are activated by binding to H2BK120ub, a shared feature with DOT1L, which also targets this chromatin mark (Antunes and Ottersbach, 2020).

TrxG proteins are indispensable for hematopoiesis, taking part in HSC self-reneval (MLL1), proper proliferation (SETD1A) and differentiation (MLL3) (Antunes and Ottersbach, 2020). Particularly important is MLL1, due to its involvement in acute leukemias, and presented in detail in the following chapters of the Thesis. MLL2 is important for embryonic development, as KO mice display retarded growth and embryonic lethality before E11.5 (Glaser et al., 2006). In adult hematopoiesis, MLL2 is needed for proper macrophage functioning (Austenaa et al., 2012). MLL3/4 are responsible for depositing H3K4me1 at enhancer regions (Antunes and Ottersbach, 2020; Hu and Shilatifard, 2016). Both of them are also crucial for

development, as null mutant mice were reported to die, either perinatally (Mll3), or at E9.5, for Mll4 (Lee et al., 2013). Moreover, MLL3 is an AML tumor suppressor, and its depletion in p53 KO HSPCs causes leukemia in transplanted mice (Chen et al., 2014).

SETD proteins are also involved in hematopoiesis. SETD1A is involved in fetal HSC proliferation, as its loss causes mice death 7-20 days post-birth and a depletion in the HSC compartment (Arndt et al., 2018). Adult SETD1A-depleted HSCs fail to reconstitute recipient mice, and are not able to differentiate properly towards erythroid lineages (Li et al., 2016). In contrast, SETD1B is dispensable for fetal, but important for adult hematopoiesis. It's required to maintain HSC and progenitor homeostasis (Antunes and Ottersbach, 2020).

A very important writer group is the polycomb (PcG) group of proteins, involved in gene silencing. Similarly to TrxG, it was discovered first in *Drosophila*, and the two groups cooperate to ensure proper embryonic development and lineage commitment. Polycomb group consists of 2 complexes, PRC1 and 2 (Figure I.7A), and catalyzes the deposition of H2AK119ub and H3K27me3, respectively. In a simplified view, PRC2, by depositing the methylation mark, both causes gene silencing and recruits PRC1 through its CBX proteins. PRC1 then deposits the ubiquitination mark. However, there is a lot more complexity and more than one pathway by which these complexes work (Butler and Dent, 2013; Di

Carlo et al., 2018; Hu and Shilatifard, 2016; Schuettengruber et al., 2007, 2017).

PRC1 can be divided into 2 subgroups, based on the mechanism of recruitment to chromatin: canonical (cPRC1) and non-canonical (ncPRC1). All of the PRC1 complexes contain RING1A/B, which confers catalytic activity, and one of the polycomb group ring finger proteins (PCGFs): NSPC1/PCGF1, MEL-18/PCGF2, PCGF3, BMI-1/PCGF4, PCGF5, or MBLR/PCGF6, modulating the catalytic activity.

cPRC1 is formed by the core RING1A/B, PCGF2/4 and, amongst others, one of the chromobox homolog proteins CBX, responsible for binding to the H3K27me3. CBX incorporation is flexible, as can be observed in ESCs, where CBX7 is predominantly included into cPRC1 and switches to CBX8 during differentiation (Vidal and Starowicz, 2017). The same was observed in hematopoiesis. CBX7-containing cPRC1 is active primarily in HSCs and is responsible for proliferation and self-renewal. Mice with CBX7 overexpression displayed aberrantly high HSC proliferation, which resulted in leukemia upon transplant (Scott et al., 2007). In the multipotent progenitor pool, the CBX8 is preferentially included, and it's crucial for lymphoid differentiation (together with PRC1.4 and PRC1.1). In fact, Pcgf1 (PRC1.1) depletion caused a myeloid bias (Di Carlo et al., 2018). Beyond its PRC1-related function, CBX8 has a role in t(9;11)-driven AML, as it directly interacts with MLL-AF9 and is involved in leukemic blast proliferation (Tan et al., 2011). cPRC1 is

recruited to chromatin by the CBXs, which recognize the H3K27me3 mark, deposited by the PRC2 complexes. Following recruitment, cPRC1 deposits the monoubiquitination mark (Conway and Bracken, 2017; Di Carlo et al., 2018).

ncPCR1 complexes lack the CBX proteins, therefore the mechanisms of their recruitment to target sites is different. Apart from the core components, they incorporate RYBP/YAF and a number of different proteins, which depend on the included PCGF, and so e.g., PCGF1-containing ncPRC1 (PRC1.1) is composed of the histone demethylase KDM2B, ubiquitin ligase SKP1, scaffolding protein BCOR etc. ncPRC1 subcomplexes bind chromatin through different subunits, e.g., the aforementioned PRC1.1 through KDM2B, which targets the unmethylated CpG islands, meaning it is PRC2-independent. Furthermore, it was reported that RYBP can also bind chromatin in the absence of the methylation mark, and actually that the deposited H2AK119ub is involved in PRC2 recruitment (Di Carlo et al., 2018).

Apart from its function in embryogenesis, PRC1 complexes have also been directly implicated in hematopoiesis (Figure I.7B). HSCs depend on PRC1.4 and PRC1.1 to maintain stemness and self-renewal potential. PRC1 containing PCGF4/BMI-1 is also important at the embryonic stage, in mesoderm commitment to hematopoiesis. It was shown that mouse ES cells overexpressing BMI-1 when differentiated, yielded primitive HSCs with higher proliferating activity compared to wild-type cells (Ding et al., 2012). This BMI-1-

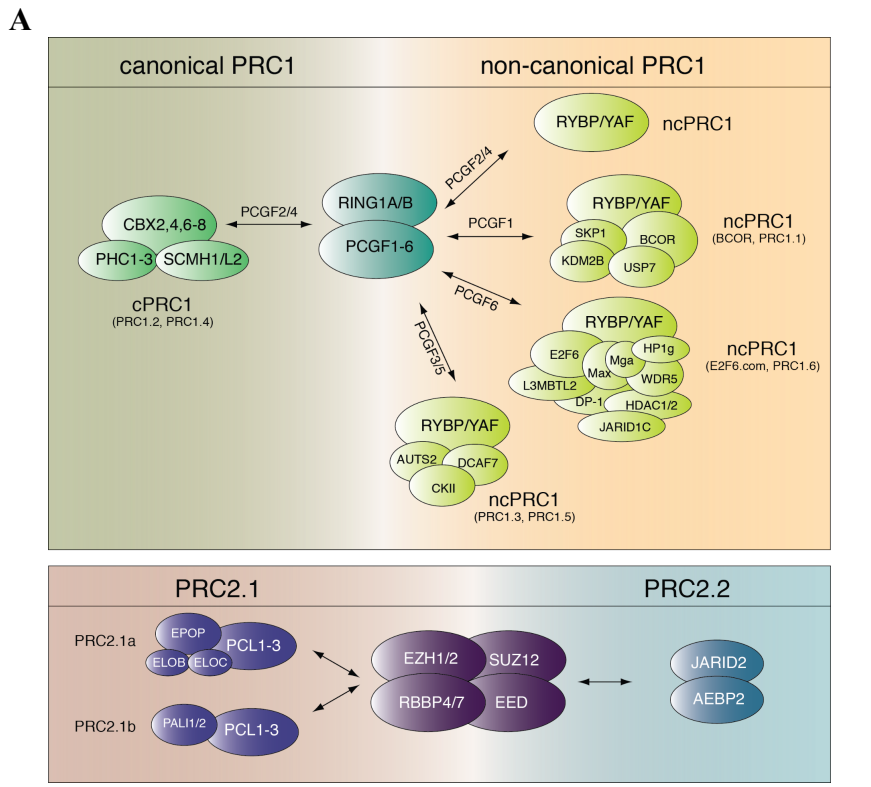
driven enhanced proliferation in HSCs was observed also in human cells (Rizo et al., 2008). In the cancer context, BMI-1 is a tumor suppressor (Di Carlo et al., 2018).

The Polycomb repressive complex 2 (PRC2) comes in two types, PRC2.1 and PRC2.2, depending on the different partners binding to the core components. There are 4 such components, namely EZH1/2, SUZ12, EED and RBBP4/7. As mentioned above, EZH1/2 belongs to the family of methyltransferases and is responsible for depositing the H3K27me3 mark. SUZ12, EED and RBBP4/7 bind DNA/RNA, H3K27me3 and histones, respectively. In the hematopoiesis context, PRC2.2 (which is formed by the core components together with JARID2 and AEBP2) functions mostly in HSCs, ensuring self-renewal and proliferation and allowing differentiation. PRC2.1 (with PCL and either EPOP or PALI1/2) takes part in proper myeloid lineage commitment. Recently, PCL3/PHF19-containing PRC2.1 was also implicated in proper maintenance of HSCs. Our lab showed that Phf19 depletion led to, amongst other effects, anterior-to-posterior homeotic transformation and increased HSC numbers at the embryonic stage. However, when mice got older, they displayed the opposite, a decrease in HSC pool, and splenomegaly was reported. HSCs were more quiescent and failed in serial transplant assay, and transcriptomic analysis revealed that Phf19 KO induced HSCs aging (they resembled HSCs coming from old animals) (Vizán et al., 2020).

The PRC2 core components also have distinct roles and have been implicated in malignancies. Although ubiquitously expressed, EZH2

is crucial for fetal hematopoiesis. It was proven by mouse studies, where it was shown that lack of Ezh2 caused embryonic lethality. In adults, EZH2 is dispensable and replaced by EZH1 in the PRC2 complexes (Hidalgo et al., 2012; Mochizuki-Kashio et al., 2011). From the cancer perspective, SUZ12 was reported to interact with PML-RARA, a fusion protein expressed in t(15;17) acute promyelocytic leukemia, resulting in PRC2-mediated silencing of hematopoietic differentiation genes (Villa et al., 2007). In the case of t(9;11) AML, the PRC2 core component, EED is necessary for the initiation and progression of the disease (Di Carlo et al., 2018).

The last out of the chromatin-modifying group of writers are DNA methyltransferases, which catalyze gene repression and the formation of heterochromatin. They comprise a family of 3 members: DNMT1, DNMT3A, DNMT3B. While DNMT3A/B can methylate DNA de-novo, DNMT1 is involved in the maintenance of this mark (Hu and Shilatifard, 2016). They are required to maintain the self-renewing ability of HSCs (Butler and Dent, 2013). DNA methylation is often deregulated in AML, and DNMT3B acts as a tumor suppressor in MLL-AF9-rearranged AML (Zheng et al., 2016).



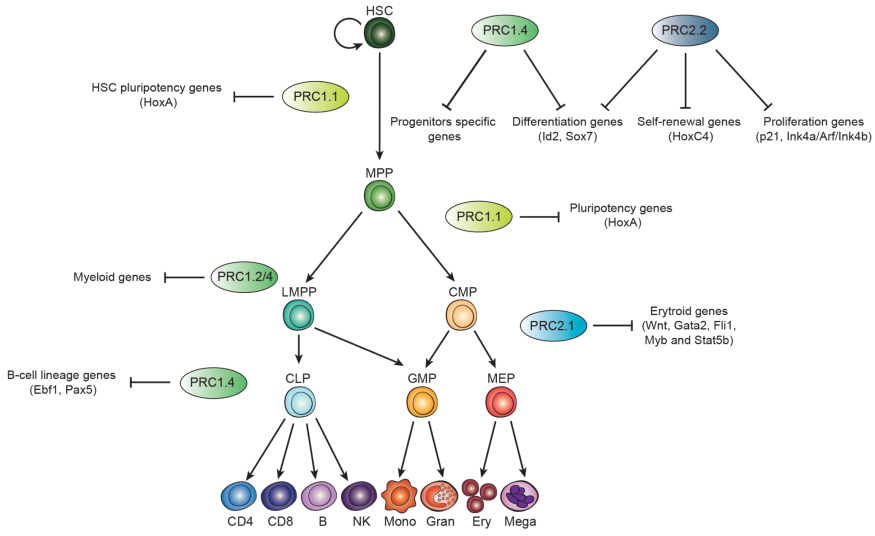


Figure I.7. **A** PcG protein complexes, **B** PcG role in adult hematopoiesis (from Di Carlo et al., 2018)

On the contrary to the previous group, the erasers are responsible for removing specific DNA/histone marks from the chromatin. LSD1 (lysine-specific demethylase 1) was the first histone demethylase to be identified. It has the ability to remove both H3K4 and H3K9 methylation marks, thus causing gene silencing. It regulates fetal hematopoiesis, as Lsd1 KO mice displayed HSC proliferation and aberrant erythroid differentiation (Sprüssel et al., 2012).

Other erasers are the DNA demethylation-catalyzing enzymes, namely the TET (ten-eleven translocation) proteins: TET1/2/3. They oxidize 5mC to 5hmC, which through subsequent steps is removed by DNA repair mechanisms. TETs are crucial for hematopoiesis, as Tet2-depleted mice develop hematological malignancies within 2-6 months after birth, and isolated HSC displayed a myeloid bias (Ko et al., 2011; Li et al., 2011).

HDACs are involved in gene repression by removing the acetylation mark from histones. They take part in embryonic hematopoiesis, specifically at the endothelial-to-hematopoietic transition (Thambyrajah et al., 2018). However, their expression must be tightly regulated, as shown in the example of HDAC1, where overexpression abrogated the ability to differentiate towards the myeloid lineages (Wada et al., 2009). HDACs are frequently involved in leukemia, e.g., the t(8;21)-related AML fusion protein RUNX1-RUNX1T1 binds and recruits them to repress differentiation genes (Glozak and Seto, 2007).

3.4 Mechanisms and effects of chromosomal translocations

Chromosomal translocations are structural genome variants that happen when the DNA repair mechanisms fail to properly correct two co-occurring double-strand DNA breaks (DSBs) on non-homologous chromosomes. As a result, part of one chromosome gets "stitched" to a different one. If the two chromosomes acquire the additional part from the other ones, the translocation is classified as balanced, whereas if some chromosome fragment is lost in the process, unbalanced. Normally, DSBs occur rarely, and when they do, they get quickly and efficiently repaired. For the translocation to occur, a second break has to happen before the first gets corrected, which *per se* is extremely improbable. Moreover, the trans-repair (joining different chromosome parts) has to "win" over the usual cis-repair, adding to the complexity. That is why chromosomal translocations are very rare (Ramsden and Nussenzweig, 2021).

The exact cause and mechanism of these rearrangements is still not known, but there are many factors that might, under special circumstances, contribute to the event. In general, there are two main models of how translocations occur. The "breakage-first" model assumes that, upon DSB, a DNA piece gets detached from the chromosome and then scans the nuclear volume until it finds a suitable partner. On the other side, the "contact-first" model postulates the necessity of close physical proximity between chromosomes prior to the occurrence of the breaks.

Regardless of the model, the process of chromosomal translocation can be assessed considering the two events summing up to the rearrangement: generation of double-strand breaks and improper repair.

Double-strand DNA breaks

Many agents have been shown to induce DSBs in the genome, both exogenous and endogenous. Extrinsic factors include e.g., ionizing radiation (IR) and chemotherapeutic agents, that cause spontaneous DNA damage (Cannan and Pederson, 2016; Jeggo and Löbrich, 2007; Mehta and Haber, 2014; Roukos and Misteli, 2014). Radiation can damage DNA either directly, by destroying the phosphodiester backbone, or indirectly, by splitting the water molecules in its close proximity. This leads to the formation of reactive oxygen species (ROS), highly reactive molecules that cause single-strand DNA breaks (SSBs). Accumulation of SSBs in a particular chromatin locus can eventually lead to the DSB formation (Cannan and Pederson, 2016). The translocation-inducing effect of IR becomes very evident if we consider that a high number of secondary leukemias and related chromosomal translocations happen as a side-effect of irradiation therapies.

Alternatively, ruptures induced by intrinsic factors can occur either spontaneously or in a targeted manner. DSBs occurring as a result of replication stress or oxidative metabolism (ROS) can happen virtually anywhere in the genome, but the targeted ones are restricted to specific loci and actually are induced as a part of normal cellular

processes (Cannan and Pederson, 2016; Roukos and Misteli, 2014). Most endogenously-generated DSBs occur during replication. They are formed as a result of replication fork stalling due to e.g., abnormal DNA secondary structures, pre-existing SSBs, collision with the transcriptional machinery or other lesions (Cannan and Pederson, 2016; Mehta and Haber, 2014).

An example of inducing targeted DSBs is provided by the RAG complex, which takes part in the V(D)J recombination during lymphocyte maturation. By mediating rounds of cleavage and differential DNA repair in the antigen-binding regions of lymphocytic immunoglobulins and T-cell receptors (TCRs), RAG forces more sequence variability. The more variable the lymphocytic antibodies, the more antigens they are able to recognize, which makes cellular defense mechanisms robust (Mehta and Haber, 2014; Nambiar and Raghavan, 2011). Cellular processes, such as replication and transcription require the relaxation of chromatin, which must be a well-coordinated process, as DNA molecule expansion in one locus must be compensated by higher packing in another (due to limited nuclear volume). One of the enzymes taking part here is topoisomerase II (TOPII), which helps resolve chromatin by inducing DSBs (Nitiss, 2009). The lesions are transient and the formed DNA ends are protected by direct interaction with TOPII. The final step is the correct reassembly of the two parts, ensuring proper chromatin continuity. However, some agents can stabilize the transient DNA-TOPII complexes, thus making the DSB permanent

and thus more prone to translocations. Such agents include anticancer drugs, e.g., etoposide (Canela et al., 2019).

DSB repair mechanisms

In order for the translocation to occur, the next step after co-occurring DSBs is the aberrant DNA repair involving joining fragments coming from different chromosomes. Independently of the model (breakage-first, contact-first), at the time of repair, the free DNA ends must be close in the nuclear space.

It is estimated that every nucleated human cell suffers from 10-50 DSBs daily (Cannan and Pederson, 2016). Therefore, in order to maintain homeostasis, it is crucial that those breaks get efficiently repaired. Any DNA lesions activate a global DNA damage response (DDR), with the aim to correct them. There are two main repair pathways concerning DSBs in the DNA, non-homologous end joining (NHEJ) and homologous recombination (HR). The pathway choice is a complex decision and depends on a number of factors, such as cell cycle stage, type of created DNA ends and availability of specific enzymes (Mehta and Haber, 2014; Roukos and Misteli, 2014).

Although being error-prone, the most widely-used repair is non-homologous end joining. It involves direct joining of the two ends created after a DSB, with possible small deletions and/or insertions at the breakpoint junction. It is functional during the whole cell cycle,

and crucially important during G0/G1, where HR is blocked (Mehta and Haber, 2014).

Contrarily to NHEJ, homologous recombination is functional during S/G2 phase and needs a template DNA for repair, in most cases provided by the sister chromatid. For that, it is thought to be error-free. In the context of chromosomal translocations, NHEJ-mediated repair has been reported to play a major role.

Specificity of chromosomal translocations

Considering that chromosomal translocations are not unlimited in terms of affected genes and cell types, many studies tried to find any possible sequence/structure/functional similarity between translocation partners. However, since they are still a very broad and heterogeneous group, it's difficult to draw conclusions applicable to all. More probably, there is a limited number of facilitating factors, but each translocation requires the cooperation between several of them.

A number of rearrangements are related to aberrant function of cellular enzymes, like the aforementioned RAG complex. RAG has been implicated in some cancer-related chromosomal translocations, like involving the IgH locus on chromosome 14. It was shown that the breaks in partner genes (e.g., LMO2, TTG1, SIL and SCL) are not random, but contain cryptic RAG recognition sites. This results in aberrant targeting of these loci by the RAG complex, followed by cleavage, eventually leading to translocations (Nambiar and

Raghavan, 2011). Interestingly, evidence exists pointing to the ability of the RAG complex to recognize, apart from the sequence, also specific DNA structures. *In vitro* studies showed that RAG can catalyze the cleavage of overhangs and flap DNA structures (Grawunder, 1997). This could also suggest an involvement of this complex in the formation of chromosomal translocations, e.g., in the t(14;18) rearrangement, found in lymphomas. The DNA fragment comprising the major breakpoint region of the partner gene, BCL2, was shown to form a non-B DNA structure. Such a domain could be cleaved by RAG, and in case of DNA repair machinery failure, take part in inter-chromosomal repair-mediated joining (Nambiar and Raghavan, 2011).

Another enzyme with possible implications in translocations is the activation-induced cytidine deaminase (AID), involved in lymphocyte specification. AID deaminates the cytosines in ssDNA regions (e.g., during transcription), which leads to mutations and eventually might lead to a DSB. It was demonstrated that AID is responsible for the breakpoints in the c-MYC gene, involved in the t(8;14) translocation (Burkitt's lymphoma). Moreover, statistical correlation studies revealed the significant presence of CpGs near some translocation breakpoint regions, which could be targeted by AID (Nambiar and Raghavan, 2011; Robbiani et al., 2008).

Many studies found a link between specific genomic sequences and rearrangements, specifically repetitive Alu sequences. Upon suffering DSBs, they can undergo chromosome arm swapping with

other Alu or non-Alu fragments. This type of repeats has been found near the breakpoint regions of e.g., BCR and ABL genes, whose translocation leads to chronic myeloid leukemia, or MLL gene, taking part in numerous translocations (Hess, 2004; Jeffs et al., 2001; Nambiar and Raghavan, 2011).

Apart from that, studies have shown that translocations occur more frequently in the transcriptionally active regions, compared with the silent ones. This can occur because euchromatin is more prone to DSBs that heterochromatin (Zheng, 2013), and additionally, genes sharing transcriptional machineries translocate more often that non-translocating genes (Roukos and Misteli, 2014; Zheng, 2013). Here also emerges the factor of spatial proximity, as in general chromosomes and genes occupying spatially proximal volumes are thought to undergo translocations significantly more often than expected (Fudenberg et al., 2011; Roukos and Misteli, 2014; Zhang et al., 2010; Zheng, 2013). However, the proximity does not seem to be necessary in all of the cases. Roukos and colleagues (2013) showed that around 10% of the translocations don't rely on initial close positioning, pointing to other mechanisms governing these specific events. The same conclusion emerges from a 3D-FISH experiment in myeloid and lymphoid cell lines. The relative position of the MLL gene and its translocation partners, ENL and AF4 was assessed. It was shown that ENL is significantly closer to MLL than AF4, which should theoretically result in higher translocation frequency. On the contrary, epidemiological data point to the MLL-

AF4 translocation as occurring on a much higher scale (Gué et al., 2006).

Another possible mechanism comes from the microscopic observation of DSB movement living cells. It was shown that the DNA ends are motile, and the translocating breaks were moving faster than others, giving them more probability to meet (Roukos and Misteli, 2014; Roukos et al., 2013). Other possible features include gene length and evolutionary conservation (the longer and more conserved genes, the more prone to translocate), fragile sites, specifically transcription start sites, chromatin loop anchors etc., (Zheng, 2013).

A number of studies found a correlation between the proximity between translocating loci and the rearrangement rate. In hematopoietic cells, where translocations are particularly common, translocation partner genes appear to be closer in space than in other cell types. This makes the "contact-first" model to be more plausible, but evidence for translocations originating from DSBs occurring far from each other also exists (Meaburn et al., 2007; Zheng, 2013).

Taking into consideration all of the above processes of creating DSBs and erroneous repair leading to chromosomal translocations, it becomes apparent why these rearrangements are extremely rare. They require a number of events to not only happen at the same time, but also occur in a skewed way. Leukemia seems to create such conditions, as translocations are a common hallmark of this cancer

type, being present in almost 60% of the cases. However, they are still rare there, with e.g., the IgH/c-Myc translocation occurring in less than 1 in a million activated mouse B cells (Zhang et al., 2010). It might point to the hematological system as permissive, on one hand, for the translocation to happen, and on the other, for the translocation to provide the cell an extremely strong growth advantage.

The big unknown remains whether translocations occur only in determined cells or are more ubiquitous but lead to cancer only in some cases. Whereas the first seems more probable, some striking evidence might point otherwise. In the t(9;11) AML mouse model, loxP sites were inserted into the introns of Mll and Af9 genes. Crossing with a Lmo2:Cre mouse, where Cre was expressed in hematopoietic stem and progenitor cells, caused the Mll-Af9 translocation in this compartment and resulted in leukemia phenotype. Interestingly, when Cre was expressed under the Lck promoter, driving its expression in T cells, the translocation was also detectable but no malignant phenotype was observed (Drynan et al., 2005), suggesting the importance of the cellular environment.

Effects of chromosomal translocations

Chromosomal translocations result mainly in two event types: displacement of gene regulatory elements and fusion of coding regions (Figure I.8). An example of the first mechanism is the effect of the aforementioned t(8;14) translocation, found in Burkitt's lymphoma. As a result, the MYC oncogene is placed under the

transcriptional control of IGH and expressed at abnormally high levels. Another such example is the t(11;14) translocation, present in T-ALL, which leads to LMO2 overexpression driven by the TCRD locus (Maddalo and Ventura, 2016; Zheng, 2013).

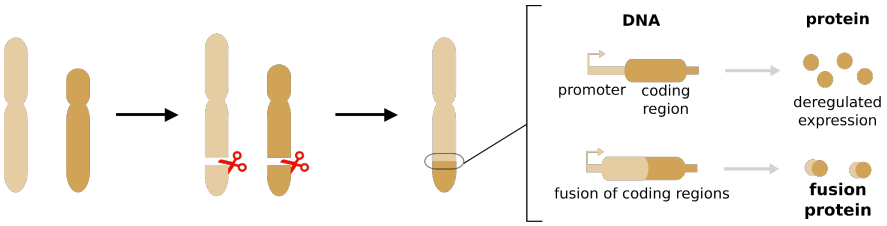


Figure I.8. Effects of chromosomal translocations (based on Roukos and Misteli, 2014)

Fusion of coding sequences upon translocation happens when breakpoints occur in either the introns of the two genes, or inside of the exons, in both cases compatible in terms of maintaining the open reading frame (ORF) of the gene. The expressed fusion protein retains some (or almost all, depending on the case) functional domains coming from the translocation partner genes, but its action might be (and actually usually is) different from them. There are two functional types of the emerging oncogenic fusion proteins. They can act as aberrant transcription factors, either stimulating, like TMPRSS2-ERG, t(9;11)/MLL-AF9, or repressing (e.g., t(15;17)/PML-RARA, t(8;21)/RUNX1-RUNX1T1) gene expression. Although with opposing activity, the effect of both groups is the same, namely a block in the hematopoietic differentiation and accumulation of leukemic blasts.

Apart from TFs, chromosomal translocation-mediated fusion genes can result in expression of aberrant tyrosine kinases (TKs). Examples include t(9;22)/BCRL-ABL (CML, ALL) and t(2;5)/NPM-ALK (Non-Hodgkin lymphoma). Usually, the N-terminal part provides the oligomerization activity, which allows for the fusion activation independently of the physiological signals (Zheng, 2013). The t(9;11) translocation and related MLL-AF9 protein will be discussed in more detail below.

4. AML-related t(9;11) chromosomal translocation

4.1 Translocation partner genes

MLL

Mixed lineage leukemia gene (MLL), the human homolog of the *Drosophila* trithorax gene, encodes a histone H3 lysine 4 (H3K4) methyltransferase. It has a crucial function in proper embryo development and hematopoiesis, by regulating the expression of HOX genes. It is a 3969 amino-acid protein, with many functional domains, such as AT-hook motifs, a DNA methyltransferase homology domain (DNMT), four PHD fingers, a transactivation domain, and a highly conserved SET domain (Figure I.9). After translation, MLL is proteolytically cleaved in the cytoplasm by taspase 1. The resulting fragments, the N-terminal (N320) and C-terminal (C180), form a non-covalent heterodimer through the FY-rich N-terminal (FYRN) and FYR C-terminal (FYRC) domain, and together with other factors (like e.g., menin), form an active complex. Interestingly, while in most cases MLL needs menin to bind to target genes (e.g., HOXA9, MECOM), in other cases the cofactor is dispensable, like in the case of PBX1, PRDM16 (Artinger et al., 2013).

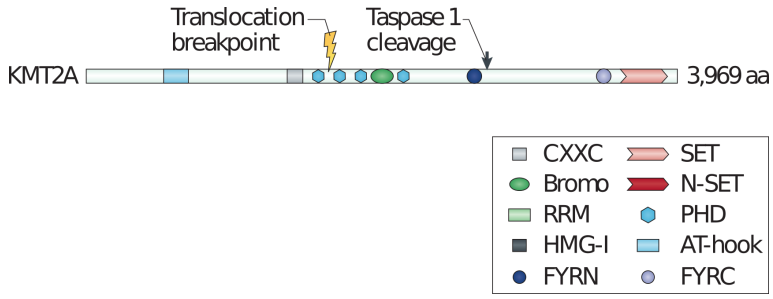


Figure I.9. MLL protein domains and breakpoint region (Modified from Rao and Dou, 2015).

The role of MLL was studied through numerous knockout and loss-of-function mutant studies. When disrupted in the germline, MLL homozygous KO caused embryonic lethality at E10.5-16.5 lack of Hox genes expression, whereas haploinsufficient mice, although survived birth, displayed homeotic transformations and hematopoietic defects, as well as Hox misregulation. This proves the crucial MLL role in HOX-dependent embryonic patterning and development. When the KO was restricted to the blood lineages (using e.g., Vav-Cre), the mice were born, but exhibited rapid bone marrow failures shortly after (3weeks). Adult HSCs depleted of MLL displayed loss of engraftment ability and loss of quiescence. Strikingly, differentiated cells (lymphocytes, macrophages, neutrophils) were not sensitive to MLL loss, pointing to its role in hematopoiesis specifically in the embryonic phase and adult HSC compartment (Antunes and Ottersbach, 2020; Jude et al., 2007).

MLL was shown to be frequently targeted in leukemia-related chromosomal translocations, with around 130 rearrangements described up to date. The most common are: AF4, AF9, ENL, ELL and AF10, present in around 80% of the cases. The MLL gene

contains two breakpoint regions, localized between intron 7 to exon 14, and between intron 17 and 28. However, in most (around 90%) cases, the breakpoint occurs between exon 9 and intron 11 (Meyer et al., 2018). As a result, the translocated MLL-fusion proteins retain the N-terminal part up to (but without) the PHD domains, and lose the C-terminal part.

AF9

The MLLT3 (AF9) gene was first discovered as a translocation partner of the MLL gene in acute myeloid leukemias. It encodes a 568-aa protein, belonging to the YEATS family. As all members, AF9 contains a highly conserved (from yeast to human) N-terminal YEATS domain, frequently present in proteins forming transcription-activation complexes. Moreover, it was shown that the YEATS domain of AF9 is necessary for histone H3 binding, suggesting its possible chromatin-modifying and chromatin-recruiting abilities. Downstream of the YEATS domain, AF9 contains a nuclear localization signal (NLS) and a serine-proline rich region. The C-terminus comprises a transcriptional activation domain, containing an ANC1 (actin non-complementing gene 1, budding yeast protein) homology domain (AHD), involved in protein-protein interactions. The functional AF9 domains are presented schematically in Figure I.10.

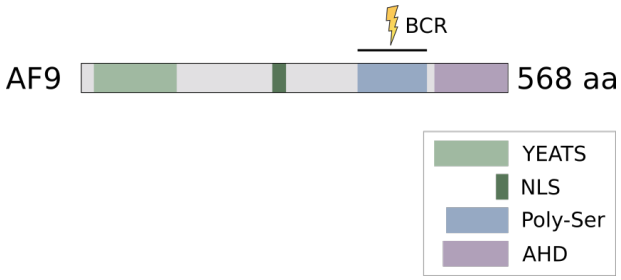


Figure I.10. AF9 protein functional domains.

AF9, through its transactivation domain, is involved in 4 major complex types with the following proteins: DOT1L methyltransferase, super-elongation (SEC) complex, CBX8 and BCOR. Surprisingly, the different groups have opposing effects in the regulation of gene expression. By interacting with either the components of SEC complex or DOT1L, AF9 mediates gene activation. On the other hand, through association with CBX8, and thus Polycomb repressive complex 1 (PRC1) or BCL6 corepressor BCOR, MLTT3 is involved in transcription repression. It is now yet known how AF9 is recruited to different complexes. However, it has been shown that, although the AHD domain is intrinsically disordered, it acquires a specific and unique conformation depending on the protein it associates with. Moreover, binding one of its interaction partners inhibits the association with a different one, probably due to structural differences (Leach et al., 2013).

Despite the broad presence in transcription-regulating complexes, pointing towards a transcription factor role, the precise function of AF9 is not yet clear. Interestingly, mouse knock-out studies revealed its role in embryonic patterning and segmentation. It was shown that

heterozygous mice were phenotypically normal and fertile, but homozygous Af9 KO caused early postnatal lethality and homeotic transformations. These malformations resemble closely the defects observed in Mll heterozygous mice knockouts, suggesting a possible AF9 role as HOX gene master regulator.

Moreover, recent studies also point to the role of AF9 in hematopoiesis, as a regulator of hematopoietic stem cell maintenance and differentiation. It is highly expressed in human HSCs and is downregulated with differentiation. In the stem cell compartment, AF9 is necessary to maintain the "stemness", as the depletion abrogated the transplantable HSC population, and overexpression caused a contrary effect, namely a 12-fold expansion of transplantable hematopoietic stem cells (Cavallari, 2019).

With respect to more differentiated blood cells, AF9 is crucial for early erythropoiesis and megakaryocytopoiesis. Its knock-down in cord blood CD34+CD38- progenitors caused complete loss of erythrocyte colony formation capacity, but had no effect on myeloid progenitor activity or cell survival. In contrast, the overexpression induced increased levels of genes associated with erythrocyte and megakaryocyte development, such as GATA1, GATA2, SCL and GFI1B. The expression of PU.1, CEBPA, and GFI, genes involved in granulocyte and monocyte development were decreased, and the apoptosis in the myeloid cultures was increased (Pina et al., 2008).

In the context of leukemia, AF9 was found to take part, together with MLL, in the t(9;11) translocation. The breakpoint region lies between exon 4 and 9, resulting in the inclusion of the trans-activating domain and (for the breakpoints upstream of exon 6) the serine-rich domain. This causes aberrant recruitment of AF9-interacting partners to MLL-target genes (e.g., HOXA9) and leads to leukemia (Aplan, 2006; Ottersbach et al., 2018).

4.2 Leukemia with the t(9;11) rearrangement

t(9;11)-induced leukemia is present both in children and adults, and depending on the age can cause predominantly ALL or exclusively AML, respectively. In AML, this translocation is detected in 3-5% of patients (Zhang et al., 2021). It involves the aforementioned MLL and AF9 genes, and results in the expression of the MLL-AF9 fusion gene. In general, in infants it is characterized by poor (ALL) or intermediate (AML) prognosis, whereas in adults it is classified as being of intermediate-high risk (Ottersbach et al., 2018). Studies on monozygotic twins showed the presence of MLL translocation with 100% concordance and rapid leukemia development, pointing that the fusion is sufficient for transformation (Antunes and Ottersbach, 2020; Ford et al., 1993; Gale et al., 1997).

The effect of the translocation in this case of leukemia can be discussed from different perspectives. As mentioned earlier, the expressed MLL-AF9 fusion protein leads to leukemia by causing aberrant expression of MLL target genes (such as HOXA9, MEIS1). This in turn leads to a differentiation block and accumulation of

aberrant hematopoietic cells, termed blasts. Moreover, blasts also become heterozygous KOs for both of the translocation genes (MLL and AF9), and, as shown before, haploinsufficiency of those factors affects hematopoiesis. The formation of the reciprocal translocation and AF9-MLL fusion gene is a rare event and was not studied thoroughly in the context of leukemia, however it might also contribute to the disease (Meyer et al., 2006, 2018). Lastly, the translocation results in the formation of new chromosomes, which could result in large-scale changes in the genome positioning, and influence gene expression. All this indicates the complexity of the translocation-related leukemias, and the crosstalk between different processes that contribute to leukemogenesis.

4.3 Models of t(9;11)-related leukemia

Since the discovery of chromosomal translocations and related changes in the nucleus as a driving force of many leukemia subtypes, much effort has been put into creating model systems recapitulating these events. The experimental models follow two main strategies: introduce malignant factors (protein/mutation) into healthy cells to assess their effect, or start from the already transformed leukemic environment with the aim to revert the phenotype. The first approach can be further subdivided into 3 groups: retroviral-based protein overexpression, transgenic mice models and genetically engineered cells. In the case of the second strategy, a number of patient-derived AML cell lines have been established. Examples of t(9;11)-related leukemia models have been summarized in Table I.2.

The earliest-developed and most used strategy is to overexpress the MLL-AF9 protein with the use of retroviruses in either murine or human hematopoietic stem and progenitor cells. With this approach, it was possible to prove that MLL-AF9 alone is able to immortalize bone marrow cells and induce leukemia in primary and secondary recipient mice (Bindels et al., 2012; Kotani et al., 2019; Krivtsov et al., 2006; Muntean et al., 2008). Moreover, this approach allowed the discovery of crucial target genes e.g., HOXA genes and MEIS1. Also, molecular assessment of the fusion protein function revealed that, despite sharing a large (approx. 1,400 aa) N-terminal part, MLL and MLL-AF9 are functionally distinct and bind to different genome regions (Xu et al., 2016).

Although retrovirus-based models proved to be extremely useful and informative, they carry a number of limitations, like the MLL-AF9 expression levels, which are much higher compared to AML patients. In order to recreate more physiological leukemic conditions, transgenic mouse models were developed. The first one, developed in 1996, comprised the modification of mouse ES cells by homologous recombination, with the aim of inserting the corresponding AF9 cDNA part (exons 9-11) to the endogenous mouse Mll locus (inside Mll exon 10). The modified cells were then injected into the blastocyst and resulting chimeric and heterozygous mice were analyzed. Expression of Mll-AF9 from the endogenous Mll promoter allowed the achievement of physiological protein levels (Corral et al., 1996). This and more studies showed that both the chimeric and heterozygous mice developed acute myeloid

leukemia, preceded by non-malignant myeloproliferation (Dobson, 1999; Johnson et al., 2003). Of interest, it was demonstrated that only the more primitive hematopoietic lineages isolated from these mice were able to induce AML in secondary recipients (HSCs, CMPs and CLP), whereas more committed (GMPs) failed to do so (Chen et al., 2008).

The same approach of gene targeting was applied to engineer human hematopoietic stem cells (CD34+ from umbilical cord blood), thus expressing MLL-AF9 from the endogenous locus (note that here the protein sequence is fully human, as opposed to the earlier, Mll-AF9 mouse-human chimera). Transplanted modified cells were able to induce leukemia in recipient mice with an average latency of 16 weeks (Buechele et al., 2015).

An alternative was provided by Stavropoulou et al. (2016). They created transgenic mice by inserting a doxycycline-inducible MLL-AF9 expression cassette in ESCs. Here, subsequently isolated LT-HSCs and GMPs gave rise to AML (after dox treatment), but with different latencies. Also, they observed that the more primitive hematopoietic cells were transplanted, the more aggressive the phenotype was.

With the era of efficient and (relatively) easy genome engineering that came with TALENs and CRISPR/Cas9 systems, more sophisticated models of MLL-AF9-induced leukemia are being created. They are able to recapitulate not only the MLL-AF9

expression, but also e.g., chromosomal translocation, expression of the reciprocal (AF9-MLL) gene and the haploinsufficiency of wild-type MLL and AF9.

The first attempt in recapitulating the event of chromosomal translocation was undertaken by Collins et al. (2000). Crossing mice containing loxP sites in specific introns of Mll and Af9 genes with Cre-expressing mice yielded an *in vivo* translocation. However, the efficiency was extremely low, as the translocation was only detectable in the brain. A more sophisticated approach, using a conditional Cre, expressed either in HSCs (under Lmo2 promoter) or T cells (Lck promoter), yielded myeloid leukemia with 300day latency, but only in the case of HSCs (Drynan et al., 2005).

TALENs and CRISPR/Cas9 have also been used to induce the corresponding chromosomal translocation in human primary cells (CD34+ from umbilical cord blood). Modified cells displayed a growth advantage in culture and clonal expansion and induced acute myelomonocytic leukemia in recipient mice. Interestingly, it was shown that the duration of the culture prior to transplantation strongly affected the phenotype. The longer such time of *in vitro* expansion, the longer the leukemia latency and higher the percentage of myeloid leukemias in the recipient mice, pointing to the importance of the pre-treatment of the cells in the evaluation of the studied phenotype (Breese et al., 2015; Jeong et al., 2019; Schneidawind et al., 2018).

Table I.2. Experimental models of t(9;11)-related AML

cells	phenotype/findings	Ref.
Retrovirus-mediated fusion protein overexpression		
GMP	AML in primary recipients within 80days	(Krivtsov et al., 2006)
WBM	MLL-AF9 induces serial replating capacity	(Muntean et al., 2008)
	Mll wt necessary for leukemogenesis	(Thiel et al., 2010)
	Mll and MLL-Af9 have different binding sites	(Xu et al., 2016)
GMP	New mutations appeared during AML progression and relapse	(Kotani et al., 2019)
Transgenic mice		
ES cells	AML in chimeric mice with 7-month latency	(Corral et al., 1996)
ES cells	AML in heterozygous mice with 5-month latency	(Dobson, 1999)
ES cells	Myeloproliferation at the fetal stage is followed by AML in adult	(Johnson et al., 2003)
ES cells	Mll-AF9 causes leukemia in HSC, CMP, CLPs, but not GMP	(Chen et al., 2008)
ES cells	*dox-inducible MLL-AF9 AML with various latencies depending on the cell of origin	(Stavropoulou et al., 2016)
hCD34+ UCB	Leukemia with average latency of 4 months	(Buechele et al., 2015)
Translocation-based models		
ES cells (Cre-lox)	Low efficiency, detectable only in the brain	(Collins et al., 2000)
ES cells (Cre-lox)	Lmo2Cre: MPD-like myeloid leukemia, latency 10 months LckCre: translocation, Mll-Af9 expression, but no leukemia	(Drynan et al., 2005)
hCD34+ UCB	AML with mean latency of 12 months low efficiency -> long culture before the transplant-> only myeloid leukemia	(Schneidawind et al., 2018)
hCD34+ UCB	*In vitro* culture time before the transplant affects AML (latency 110 for 5d, 251 for 11d) and more aml (for 5d 6/8 mice have both aml and all)	(Jeong et al., 2019)
hCD34+ UCB	Survival advantage in extended culture, clonal expansion in colony-forming cell assays	(Breese et al., 2015)
cKIT+ BM	Increased cell and colony numbers in CFU assays (methocult)	(Sarrou et al., 2020)

WBM - whole bone marrow, UCB - umbilical cord blood, BM - bone marrow

Alternatively, cell lines isolated from t(9;11) AML patients were also used as models, such as THP-1, Mono-Mac-6, MOLM-13, UG3, KOPB-26. They provided an extremely useful model in, e.g., drug and therapeutics studies.

As shown above, a huge effort has been put into uncovering the mechanisms of t(9;11)-related leukemia, with many different cell models. Both the top-down (introducing some features of AML into healthy cells) and bottom-up (study cells coming from leukemia patients) approaches allowed us to gain tremendous knowledge about this disease and develop some drugs. However, despite being extremely useful, they are still far from recapitulating faithfully AML in the laboratory and suffer from limitations. In the case of retroviral transduction, apart from abnormally high protein levels, an issue is the random viral integration that can possibly affect e.g., tumour suppressors, as well as a transduction bias (not uniform efficiency amongst different hematopoietic progenitor cell types) (Maddalo and Ventura, 2016; Schwaller, 2020). Knock-in into the endogenous locus provides a solution to non-physiological protein levels, and actually can dramatically change the phenotype and drawn conclusions, as clearly can be seen in the example of GMPs. When transduced with MLL-AF9, these cells caused AML in mice (Krivtsov et al., 2006), but in the case where the expression was driven from the Mll locus, the modified GMPs lacked leukemogenic potential (Chen et al., 2008). Inducing chromosomal translocations provides the closest scenario to the disease, but an important limitation is the extremely low efficiency of such events. Lastly,

patient-derived cell lines, although physiologically closest to AML, over time accumulated a number of secondary mutations and rearrangements, making it hard to discriminate the effects.

Apart from the problems concerning the cells *per-se*, recent studies point out to the crucial effect of the microenvironment on the development of leukemia. Interactions with the bone marrow cells and immune system have an important impact on the disease phenotype, and many of the aforementioned models do not take it into consideration (*in vitro* approaches, immunocompromised or lethally irradiated mice, etc.) (Schwaller, 2020).

AIMS

Leukemia-related chromosomal translocations are complex rearrangements, affecting the targeted cell on different levels. Currently, the majority of studies addressed only the effect of the expressed fusion protein. In order to expand this knowledge, we aimed to separate the changes driven by the translocations from the effect of the protein, using as a model the t(9;11) rearrangement that we reconstructed in mouse ESCs.

Therefore, we proposed the following objectives:

1. Assess the suitability of mouse fusion proteins as models of human leukemic chimeric proteins.

2. Build an equivalent of the t(9;11) translocation in mouse ESCs with conditional Mll-Af9 expression.

3. Determine the effect of the translocation-driven changes on hematopoietic differentiation of ESCs.

4. Investigate how the presence of the Mll-Af9 fusion protein alters these changes.

RESULTS

I. Mouse fusion proteins can be used to model leukemias

The aim of the current study is to investigate whether the cellular background caused by a chromosomal translocation creates a permissive environment for the oncogenic fusion protein to develop leukemia. Moreover, it is still not known when, during development, these rearrangements occur, which is especially relevant in the case of childhood leukemias. In order to assess that, we applied CRISPR/Cas9 technology to induce the chromosomal translocation in mouse embryonic stem cells, which corresponds to t(9;11) rearrangement in AML patients. As a result, the expressed fusion protein (Mll-Af9) is the mouse equivalent of the human leukemic oncoprotein (MLL-AF9).

Considering that our model is based on the murine fusion protein, as a first step, we had to determine whether mouse fusion proteins could be used to model human leukemias. As discussed in the introduction, a number of mouse models took advantage of the knock-in strategy, inserting the human coding sequence of the 3' translocation partner gene into the endogenous locus of the 3' partner, resulting in the expression of a mouse-human Mll-AF9 chimera (Chen et al., 2008, Dobson, 1999; Johnson et al., 2003). However, so far there have been few studies concerning fully mouse proteins. A rare example is the murine Mll-Af9 protein, expressed upon Cre-mediated translocation in mouse ES cells (Collins et al, 2000, Dobson et al, 2005). Mll-Af9

was shown to cause leukemia in transplanted mice with a latency period of 10 months. This prompted us to investigate whether mouse leukemic fusion proteins could faithfully phenocopy their human counterparts.

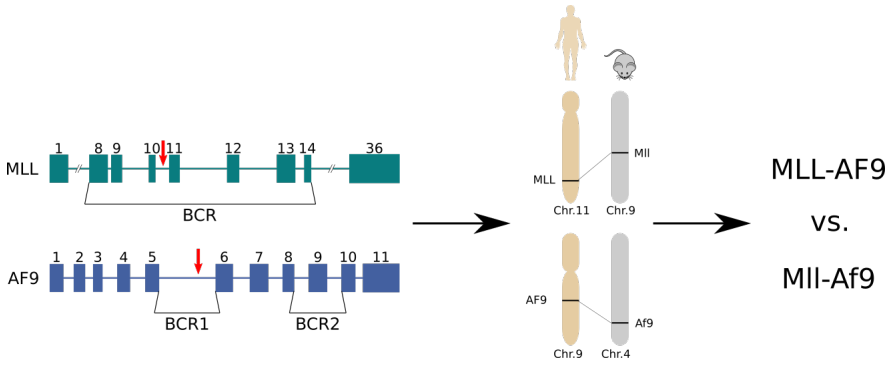


Figure III.1. Workflow of the validation of similarity between leukemic fusion proteins and their mouse equivalents. Red arrows determine the chosen breakpoint locations, BCR – breakpoint cluster region. Mouse protein sequences where then determined and compared to the human counterparts.

We started the validation by assessing the similarity of the human leukemic fusion proteins to their murine equivalents, as schematically depicted in Figure III.1. We determined the sequence of several leukemia-related fusions, taking into consideration the most frequent isoform in case of multiple or large breakpoint regions, as in the case of MLL-AF9. Next, we identified their equivalents in corresponding mouse genes (laying on different chromosomes than in humans), and assembled artificial mouse leukemic fusion proteins (like Mll-Af9). We then compared the similarity of both proteins, based on calculated sequence identity (Table III.1).

Table III.1. Degree of identity between human leukemia-associated fusion proteins and their mouse equivalents.

Translocation	Fusion protein	Human-mouse identity [%]
t(15;17)	PML-RARA	79
t(9;11)	MLL-AF9	98
t(4;11)	MLL-AF4	85
t(11;19)	MLL-ENL	91
t(8;21)	AML-ETO	92
t(9;22)	BCR-ABL	64

The comparison revealed a various degree of identity between the two species, being exceptionally high for MLL-fusions (e.g., 98% in the case of MLL-AF9). In the case of the APL-related PML-RARA, the identity between the mouse and human sequences was slightly lower, but still relatively high (79%). Based on this, we concluded that mouse fusion proteins resemble their human counterparts in terms of sequence identity, which could be indicative of their similar function. We then proceeded with the functional validation of the chosen representatives.

We chose 2 fusion proteins, Pml-Rara and Mll-Af9, for further tests. In order to reflect our experimental strategy, which will be introduced later, we included a small tag (TY1) in two copies (21 aminoacids in total) in both proteins, placed in the middle of the two parts, resulting in Pml-2xTY1-Rara and Mll-2xTY1-Af9. For simplicity, they are referred to as Pml-Rara and Mll-Af9, respectively.

We tested their functionality in a standard set of experiments used for this purpose. The proteins were overexpressed in mouse

hematopoietic stem and progenitor cells (lineage negative, lin-), using viral particles. After sorting of infected cells (based on the GFP reporter), they were plated in methylcellulose-based media supplemented with cytokines to promote cell differentiation (Methocult). Every 7 days, colonies were scored and cells replated. As expected, control cells (expressing empty vector) stopped growing after two weeks of culture (Figure III.2A). In contrast, cells expressing either Mll-Af9 or Pml-Rara kept growing longer, as a high number of colonies were detectable even after 4 weeks (e.g., around 350 colonies for Pml-Rara), pointing towards their immortalization. Not only did the cells grow longer, they also kept their "stemness" properties, as assessed by the flow cytometric analysis at the experimental endpoint (Figure III.2C-D). After 4 weeks in culture, 60-80% of Mll-Af9 expressing cells belonged to the hematopoietic stem population (lin-Sca1+cKIT+, LSK), while in the case of Pml-Rara we observed a high variability in the LSK numbers (6 and 40%), which might be caused by the differences in the initial condition (sorted cell number).

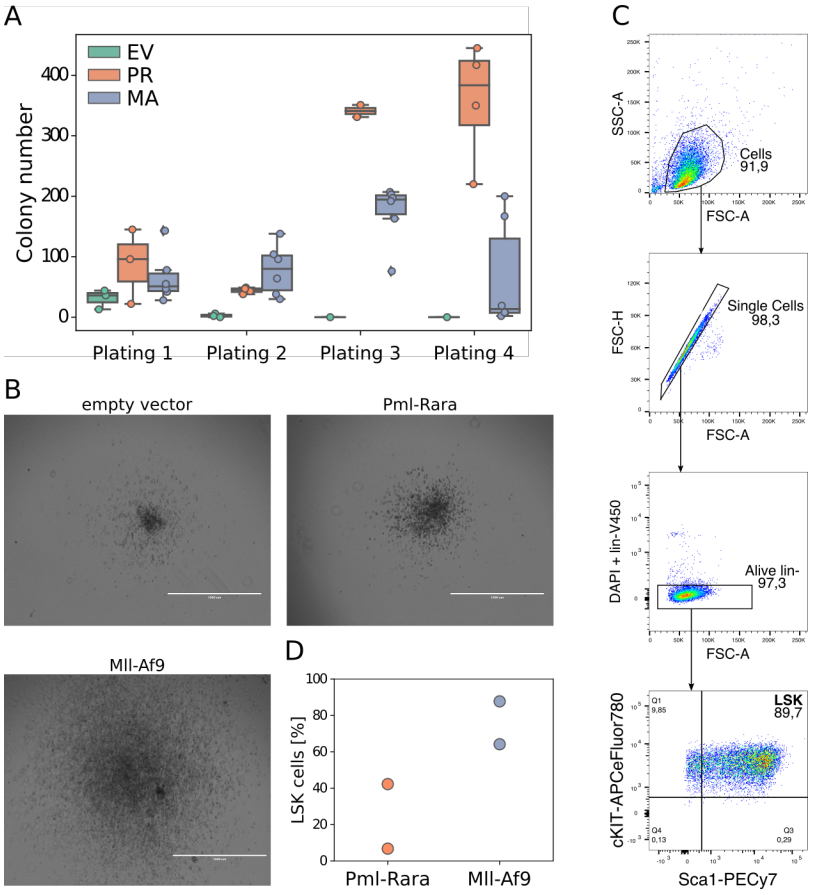


Figure II.2. Mouse fusion proteins immortalize HSPCs. (A) Methocult colony numbers at replatings in the 3 conditions. Data from 2 independent experiments. (B) Representative colonies after 1 week of culture (Plating 1). (C) Gating strategy for the LSK population, presented on the Mll-Af9-expressing cells after 4 weeks of culture. Cells were separated from debris using side/forward scatter. Next, single cells were separated from cell clusters with forward scatter height/area. Dead and lineage positive cells were sorted out with DAPI and lin-V450, respectively. Finally, hematopoietic stem and progenitor cells were gated based on the expression of two HSC markers, cKIT and Sca1. (D) Quantification of LSK cells after 4 weeks of culture. Data coming from 2 independent experiments for each condition.

We next validated the responsiveness of the mouse fusion protein to known agents. Patients with acute promyelocytic leukemia (APL), which carry the t(15;17) translocation and express PML-RARA, are treated with all-trans retinoic acid (ATRA), which causes the fusion

protein degradation and cell differentiation. Moreover, a hallmark of APL is the disruption of PML nuclear bodies and spread PML distribution (the so-called PML microspeckles), which is restored upon exposure to ATRA.

To assess whether mouse Pml-Rara will phenocopy this response, in the Methocult we included a Pml-Rara condition with the addition of clinical doses of ATRA (1μM). As a result, we observed that, at the time of replating (after 7 days), the cells treated with the drug stopped proliferating and were no longer able to form colonies, indicating terminal differentiation (Figure 3A). Loss of stemness was proved by flow cytometry and Wright-Giemsa staining (Figure 3B-C, respectively). The LSK population was almost completely abolished (from 25 to 1.55%) in the ATRA-treated samples, and the cells displayed differentiated morphology, resembling megakaryocytes. Finally, as observed by immunofluorescence, cells expressing Pml-Rara displayed a microspeckled Pml distribution, which was restored to normal (nuclear bodies) upon treatment.

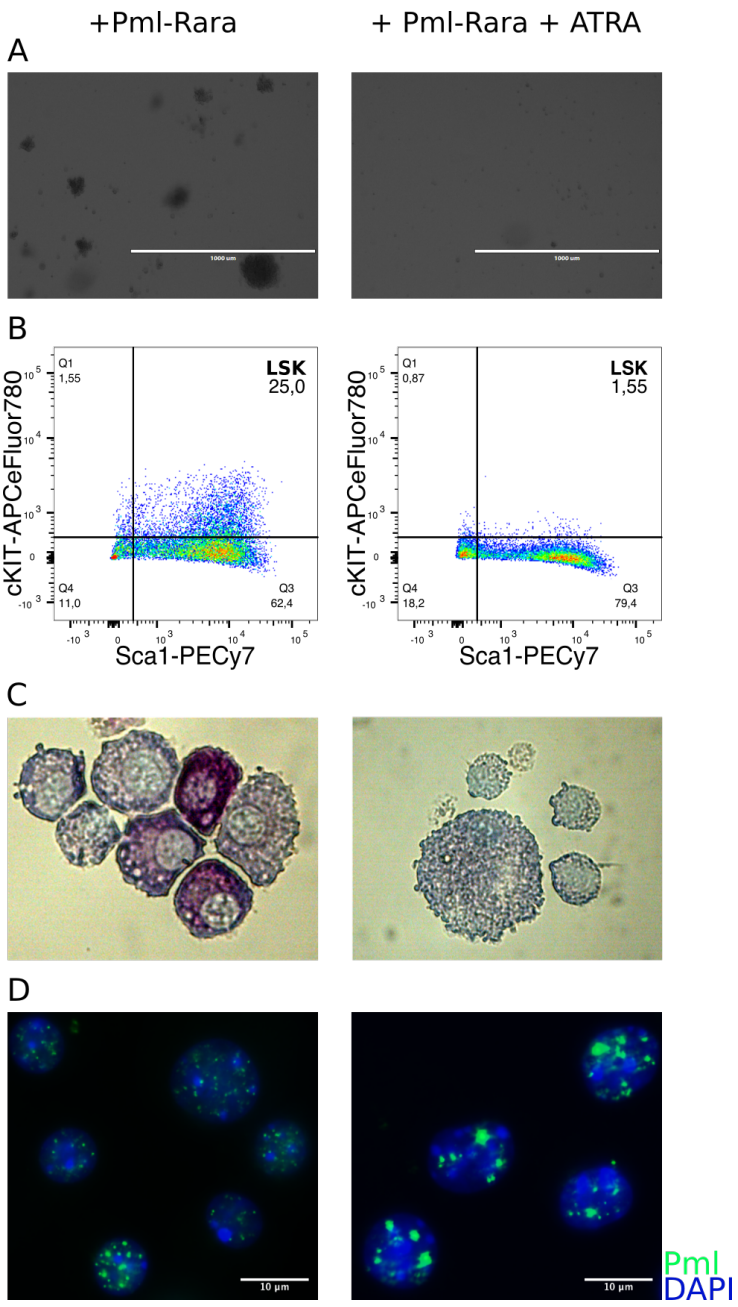


Figure III.3. Mouse Pml-Rara phenocopies the human PML-RARA responsiveness to ATRA treatment. In all cases, ATRA was used at 1µM. (A) Colonies after 7 days in Methocult culture. (B) Flow cytometric analysis of the LSK numbers. (C) Wright-Giemsa staining of the cells expressing Pml-Rara. (D) Immunostaining of the Pml distribution in the cell nuclei (Pml - green, DAPI - blue).

Considering both the high sequence identity, and the effect of the mouse Pml-Rara and Mll-Af9 proteins on the purified hematopoietic stem and progenitor cells (immortalization, retaining stemness, drug response), we concluded that mouse fusion proteins can be used as a model for their human counterparts. Of note, the TY1 tag inserted in the middle of the proteins did not block their activity, providing a strategy to precisely and accurately study the function of chimeric proteins.

II. Recapitulating embryonic hematopoiesis *in vitro*: optimization of the protocol for the hematopoietic differentiation of mouse ESCs

In order to obtain hematopoietic populations from our engineered mouse embryonic stem cell lines, we applied and adapted a protocol for ESC differentiation, which has been shown to recapitulate the embryonic hematopoiesis (Figure III.4). It closely resembles all the differentiation events occurring in the embryo, the emergence of the hemangioblast, then hemogenic endothelium and lastly, hematopoietic progenitors. We also chose this particular protocol because it does not rely on forced virus-mediated TF overexpression, which allows us to resemble as closely as possible the physiological conditions, and is relatively straightforward. As the t(9;11) translocation (with the MLL-AF9 fusion protein expressed) is the most common rearrangement found in the case of infant translocation-related acute myeloid leukemias, this differentiation system provides us with a perfect model for investigation.

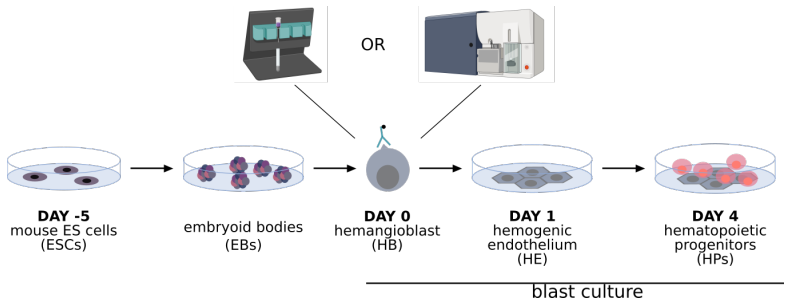


Figure III.4. Hematopoietic differentiation of mouse ES cells towards hematopoietic progenitors.

In order to use the differentiation protocol, we had to optimize it for our ES cells. In the trial experiments, we followed the specified incubation times (3 days for embryoid body formation, 3 days of blast culture), but failed to detect Flk-1+ cells (the efficiency was around 0,1%, as opposed to the reported 20-30%). In order to adapt the protocol for our cell lines, first, we analyzed the EB differentiation to define the optimal temporal window for hemangioblast selection (Figure III.5).

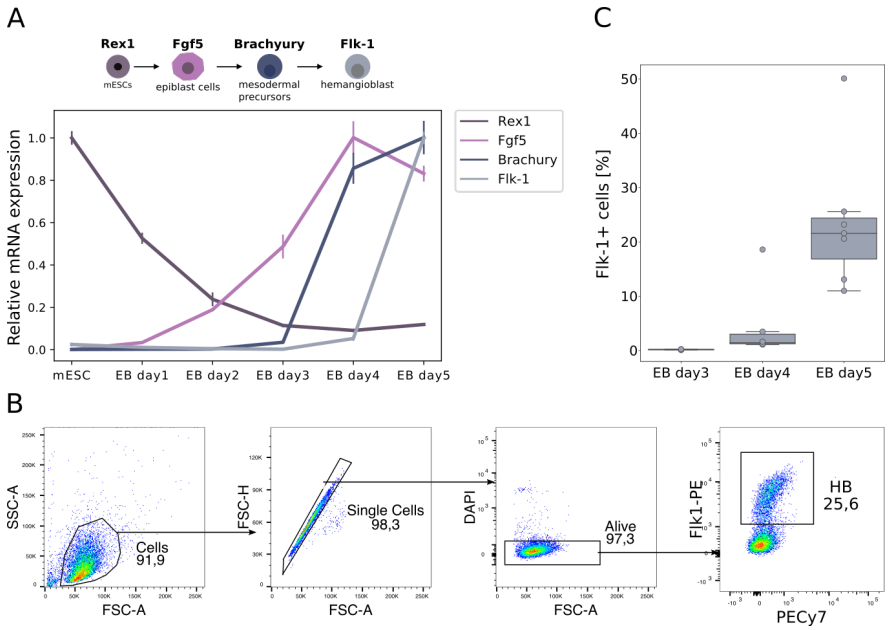


Figure III.5. Optimization of the EB differentiation culture. (A) qPCR of marker gene expression on differentiation days with the corresponding cell stages, normalized by Rplp0. Data from 3 technical replicates for each timepoint. For each gene separately, relative expression was calculated over the timepoint with maximum expression. Bars indicate standard deviation. (B) Representative plot reflecting the gating strategy of Flk-1 hemangioblasts. (C) Quantification of Flk-1+ cells on subsequent EB differentiation days. Each point represents an independent experiment.

We monitored daily the marker gene expression corresponding to subsequent differentiation cell populations from mouse ES cells to the hemangioblast, and extended the EB culture to 5 days (Figure III.5A). As expected, we observed a rapid downregulation of the pluripotency gene, Rex1, after already 1 day of EB culture. This was followed by the expression of Fgf, Brachury and Flk-1, which correspond to the appearance of epiblast cells, early mesodermal precursors and hemangioblasts, respectively. Of note, our EBs expressed a very low level of Flk-1 at day 3, in agreement with the low efficiency of detected positive cells. Flow cytometry analysis showed that day 5 is the optimal timepoint for hemangioblast selection, with around 30% of cells positive for Flk-1, as compared with the 3.5% efficiency at day 4 and 0.1% at day 3 (Figure III.5B-C). This temporal differentiation delay compared to the published work might be caused by the intrinsic differences between ES cell lines.

When plated in hematopoiesis-promoting media, the hemangioblast differentiated first into the hemogenic endothelium, and then hematopoietic precursors. According to published work, at this step both freshly isolated Flk-1+ cells and cryopreserved ones can be used. In order to make the differentiation protocol more flexible, we tested both conditions. The blast cultures are characterized by rapid growth, with clusters of hematopoietic progenitors visible after 2 days of culture (Fig. III.6A, pink arrowhead), and cell expansion clearly seen by day 4. The adherent cells comprised the hemogenic endothelium, formed by specific cell clusters, non-hemogenic

endothelium, and additionally vascular smooth muscle cells also known to be formed from the hemangioblast (Figure III.6A, blue, green and white arrowheads, respectively).

At day 4, the blast cultures were stained with hematopoietic (cKIT and CD41) and endothelial (Tie2) markers to perform the quantification of different cell lineages, as presented exemplarily in Figure III.6B. To select for hematopoiesis-related tissues, cells were first gated based on cKIT expression. Inside this cKIT+ population, we were able to track the differentiation trajectory, from the endothelial cell to hematopoietic progenitors. We were able to distinguish two populations of the hemogenic endothelium: the more primitive type 1 HE (HE1, cKIT+Tie2+CD41-), and the one more mature and undergoing the endothelial-to-hematopoietic transition (HE2, cKIT+Tie2+CD41+). Successful EHT and emergence of floating clusters of cells that lost the Tie2 cell marker was indicative of the hematopoietic progenitor population (HP, cKIT+Tie2-CD41+). We quantified the HP population coming from both fresh and cryopreserved Flk-1+ cells. Freezing these precursor cells did not abrogate their ability to give rise to hematopoietic progenitors, although it lowered the efficiency of the process, from 10 to approx. 6% of the population being classified as HPs (Figure III.6C).

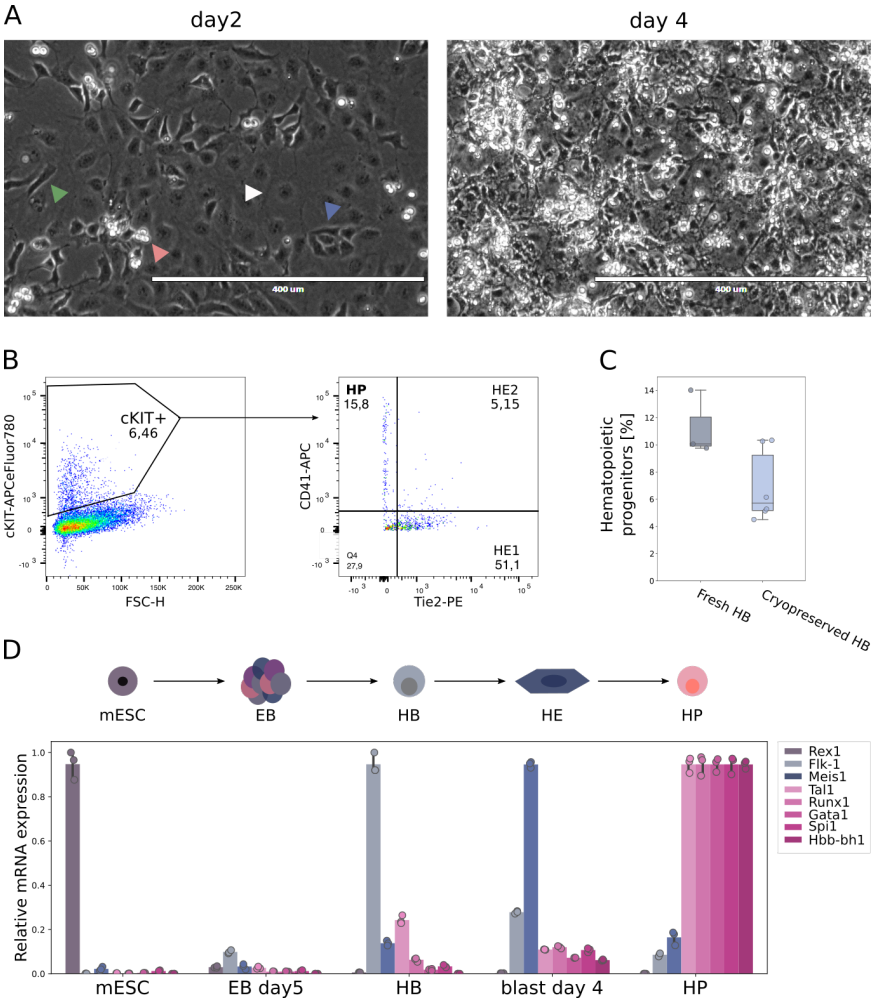


Figure II.6. Hematopoietic progenitors can be successfully isolated from blast cultures. (A) Representative microscopic images taken at indicated timepoints of the blast culture. Distinct cell populations are indicated with arrowheads (hemogenic endothelium, endothelium, hematopoietic progenitors and vascular smooth muscle cells as blue, green, pink and white arrow, respectively). (B) Gating strategy for the identification of cell populations based on marker gene expression by flow cytometry. Plots shown for selected alive single cells. (C) Quantification of hematopoietic progenitors coming from both freshly-isolated and cryopreserved hemangioblasts. Data from 2 independent experiments. (D) Gene expression analysis by qPCR (normalized with Rplp0), with the corresponding cell stages depicted above. For each gene separately, the relative expression was calculated relatively to the timepoint with highest expression. Data from 3 technical replicates.

Sorted hematopoietic progenitors were successfully validated by qPCR of marker genes essential for embryonic hematopoiesis, like Runx1, Tal1 and Spi1. Including previous differentiation steps (mES cells, EBs, hemangioblast, hemogenic endothelium) allowed to clearly distinguish all the developmental stages based on marker gene expression patterns. As seen in the EB differentiation, mouse ES cells displayed high expression of the pluripotency gene Rex1, and hemangioblasts were enriched for Flk-1. Total blast cultures, considering that HPs were a minority, displayed a high expression of the endothelial marker Meis1. All the tested hematopoietic markers (Tal1, Runx1, Gata1, Spi1, Hbb-bh1), were highly expressed specifically in the sorted HP population, validating the nature of these cells. Interestingly, Tal1 showed a slight upregulation at the hemangioblast stage, consistent with its crucial role for the differentiation of the hemangioblast towards hemogenic endothelium.

As a "proof-of-concept" experiment to evaluate the differentiation protocol in terms of assessing the effect of leukemia-related events, we aimed to test the effect of the mouse Mll-2xTY1-Af9 protein. To achieve that, using CRISPR/Cas9 technology, we developed ES cell lines carrying the inducible expression of Mll-2xTY1-Af9 in the Rosa26 locus (named "MAD"). The inserted cassette allowed for protein expression upon cell exposure to doxycycline (Figure III.7A).

Positive clones were validated in terms of the presence of the expression cassette in the Rosa26 locus, karyotyped, and

subsequently differentiated towards hematopoietic progenitors, where fusion protein expression was induced by adding doxycycline at the time of plating hemangioblasts in blast media (Figure III.7B). Inducible cell lines readily expressed Mll-Af9 exclusively after dox treatment, as shown by qPCR and WB (Figure III.7C-D).

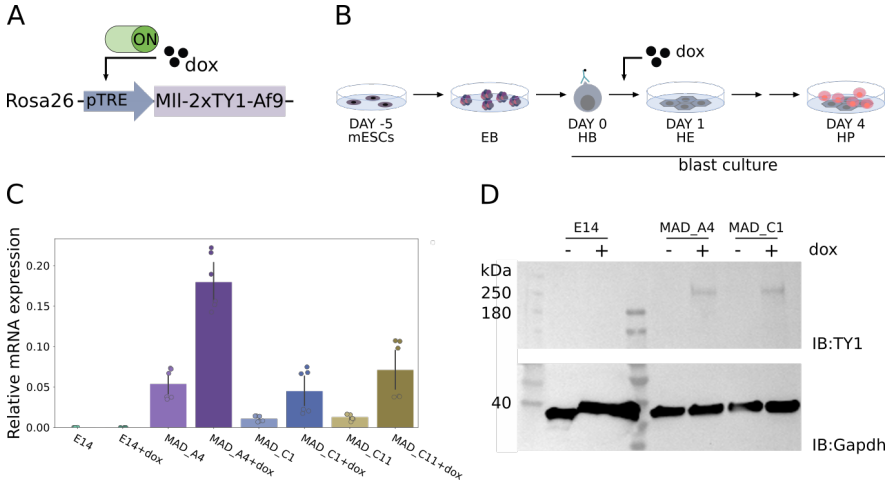


Figure III.7. Developed cell lines express Mll-Af9 in blast cultures upon treatment with doxycycline (A) Scheme of the inducible Mll-Af9 expression from the Rosa26 locus. Promoter is activated by adding doxycycline to the cell medium. (B) Experimental scheme of fusion gene expression induction. Doxycycline was added at the moment of hemangioblast plating in the blast media, at 1µg/ml. Mll-Af9 expression at day 4 of blast culture was measured by qPCR (C, data from 2 independent experiments) and western blot (D).

Having validated the fusion protein expression in differentiated cells, we proceeded with assessing possible differences in hematopoietic populations. The cultures displayed a normal morphology (Figure III.8A). Analysis of the marker gene expression by flow cytometry revealed a slightly higher number of hematopoietic progenitor cells in the cultures treated with dox, meaning expressing Mll-Af9 (Figure III.8b). Of note, all the cell lines displayed a lower number of

progenitors as compared to the wild-type, which can be caused by CRISPR/Cas9 manipulation and expansion from a single cell, possibly causing a genetic drift. Another explanation could possibly be experimental variability.

Marker gene expression in blast culture was analyzed by qPCR (Figure III.8C). No difference in Mll expression levels compared to the wild-type cells were observed upon dox treatment. This is expected, considering our experimental design. In the inducible cell lines Mll-Af9 is introduced as an additional copy, leaving the Mll and Af9 alleles intact. Moreover, we specifically designed the Mll and Af9 primers to span the breakpoint junction, which allows us to detect the expression of only the wild-type alleles and not the fusion. Interestingly however, in all the 3 tested cell lines, Af9 was upregulated specifically upon Mll-Af9 expression. The expression of hematopoietic genes, such as Tal1, Gata2 and Spi1 was highly variable depending on the cell line. The clearest example of that was Spi1, whose expression was more different between cell lines than between the non-treated and treated condition. Known MLL-AF9 target genes were also assessed and several of them showed a marked upregulation, like the Hoxa9 gene, expressed 4-9-fold more in the +dox condition compared to -dox and wild-type cells. Hoxa5 also was upregulated, but wasn't as consistent as Hoxa9 and clone-to-clone variability was observed. However, some targets, like Meis1 and Evi1 were not affected by Mll-Af9 expression. We are currently performing an in-depth characterization of the hematopoietic progenitors expressing Mll-Af9.

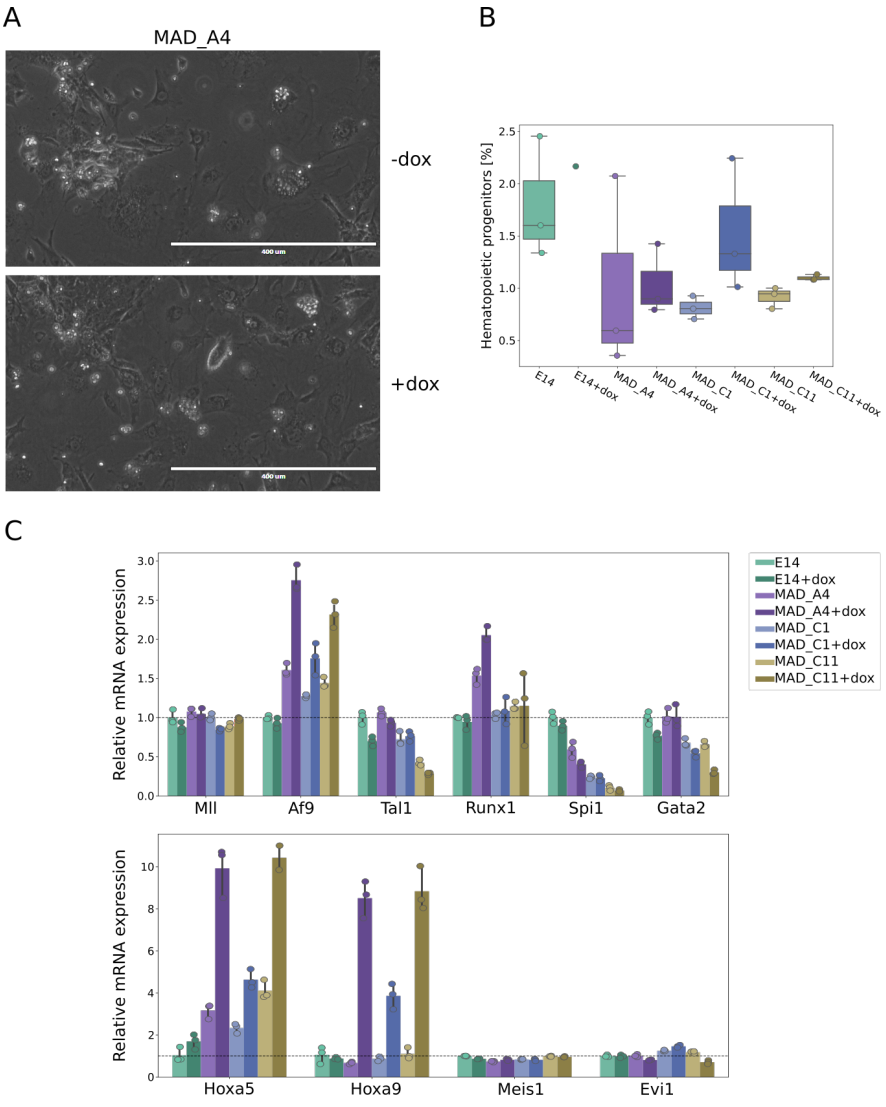


Figure III.8. Mll-Af9 expressed at the post-hemangioblast stage allows for differentiation into hematopoietic progenitors. (A) Representative images of blast cultures after 4 days -/+dox, used at 1µg/ml. (B) Quantification of hematopoietic progenitors at day 4 (cKIT+Tie2-CD41+). Data from 2 independent experiments. (C) qPCR of marker genes in total blast samples. Data from 3 technical replicates.

In summary, the hematopoietic differentiation protocol together with inducible protein expression is a well-functioning experimental

model. Our preliminary results sum up to the previous validation of the Mll-Af9 activity, performed in mouse purified HSPCs, proving further that it affects hematopoietic cells and can be used as a tool to model leukemia.

III. Recreating the t(9;11) translocation in mouse ESCs

Prompted by our positive results, we aimed to recreate not only the Mll-Af9 expression, but in fact the whole t(9;11)-related leukemia environment in mouse ES cells. In mice, Mll and Af9 lay on chromosome 9 and 4, respectively, meaning that the equivalent mouse translocation would be t(4;9). Furthermore, we wanted to have control over the expression of the fusion Mll-Af9 protein in order to distinguish between the different layers (changes driven by the chromosomal translocation vs. expressed fusion protein). Apart from that, we aimed to apply a strategy which would allow us to specifically target the expressed Mll-Af9 protein. Currently, fusion proteins are studied using antibodies targeting the wild-type alleles, and analyzed either subtracting the common signals between the two, or e.g., performing subsequent immunoprecipitations. However, this is not a direct approach and can lead to biases and loss of information. From previously shown experiments (Chapter I), we already knew that the tag could be inserted in the middle of the protein, and we applied it in the model design.

To fulfil all our experimental aims, we took advantage of the CRISPR/Cas9 system and combined it with gene-trap and endogenous tagging strategies in mouse ES cells, as depicted schematically in Figure III.9. We designed gRNAs to target the introns of choice of the wild-type alleles (intron 10 for Mll and intron 5 for Af9), and a donor template in order to favor homology-mediated repair. It's composed of a promoter-less mCherry cassette flanked by

2 loxP sites and homology arms of various lengths. By including splicing signals (donor and acceptor splice sites), after a successful rearrangement, we force the expression of Mll-mCherry fusion. The presence of loxP sites allows for regulation of the Mll-Af9 expression, which is triggered after cell exposure to the Cre recombinase. This approach gives us a number of advantages. First of all, we can control the "switching on" of the oncogenic Mll-Af9 fusion protein. Secondly, we can take advantage of the mCherry expression to select for the positive clones. Lastly, by inserting a small tag (TY1) permanently in between the two moieties (fused to the last exon of Mll, namely exon 10), we allow for precise targeting of the oncogenic fusion protein in any antibody-based assay, such as protein immunoprecipitation followed by mass spectrometry or chromatin immunoprecipitation followed by sequencing (ChIPseq).

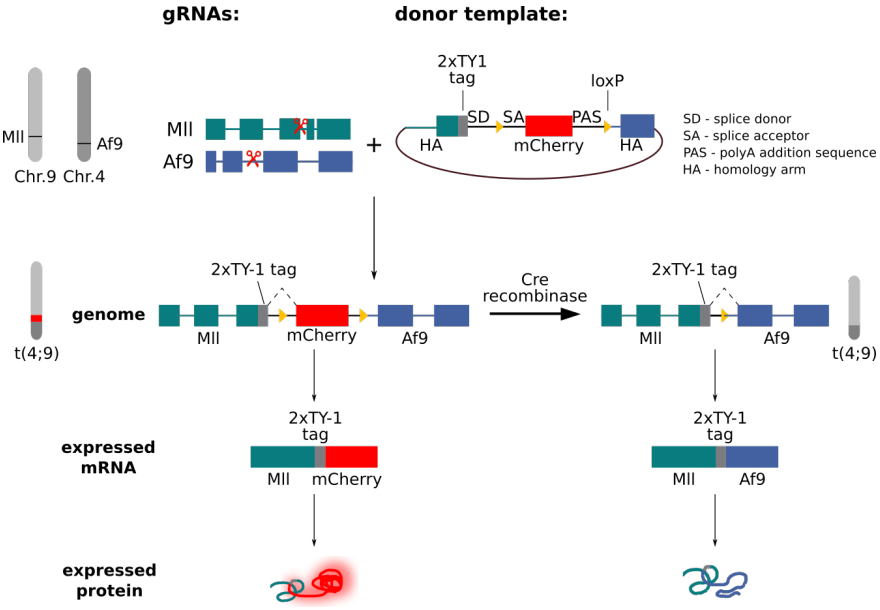


Figure III.9. Strategy of the CRISPR/Cas9-induced combined chromosomal translocation and gene trap.

The screening for positive ES cell lines was performed as depicted in Figure III.10. After transfection of the CRISPR/Cas9 components into ES cells, the mCherry+ were sorted, expanded either in single clones or bulk population first, and subsequently genotyped by PCR on the genomic DNA (in the agarose gels each lane represents a single clone). Considering that mCherry expression was driven from the endogenous Mll promoter, whose expression level in mouse ES cells is very low, the efficiency of sorting was also low (less than 1%). Considering the big size of the inserted fragment together with the homology arms (3164 bp), to facilitate genotyping, the whole product was split in 2 parts, “start” and “end” part. As can be seen from the representative agarose gels, we were able to detect clones positive for either the start (clone C) or end part (clones A and B), reflecting the insertion of the mCherry cassette in the Mll or Af9 locus, respectively.

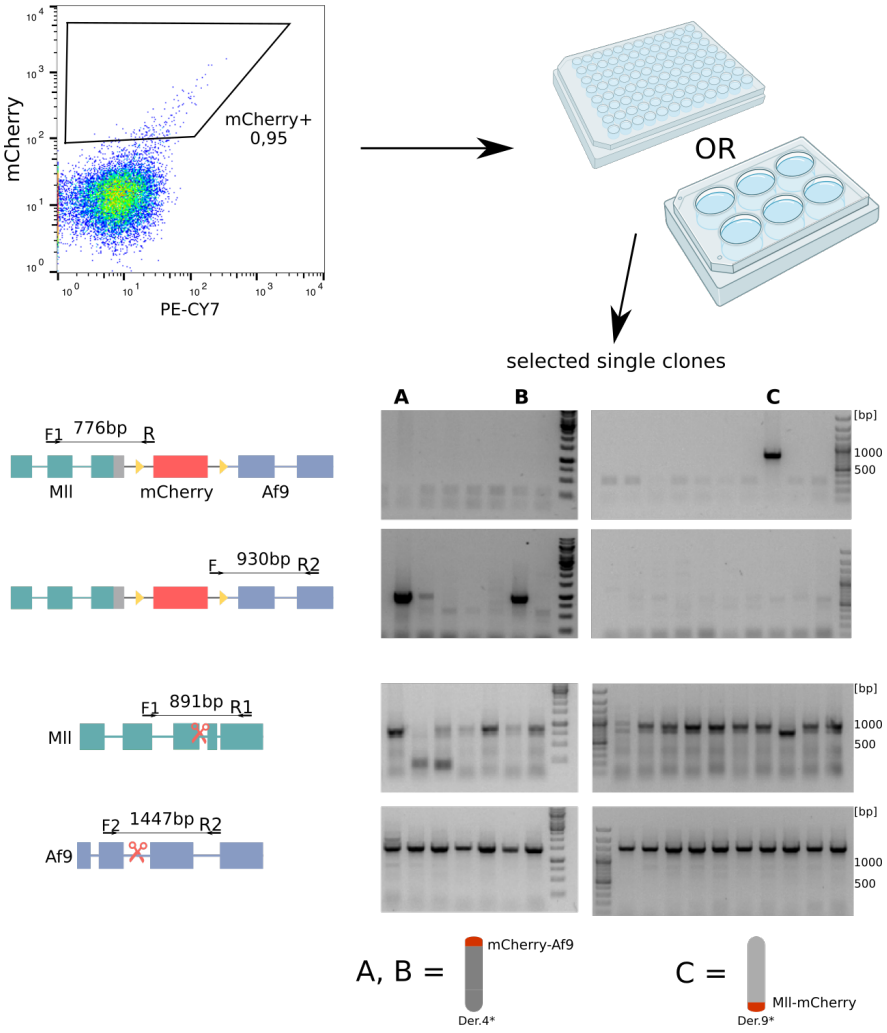


Figure III.10. Screening for positive ES cell lines with the chromosomal translocation. After transfection, cells were sorted based on mCherry expression, expanded directly in 96well plates or first in bulk, and genotyped with PCR on genomic DNA as depicted. Red scissors indicate sites targeted with gRNAs.

Obtaining clones such as A and B was not expected, since in our design mCherry expression is driven by the endogenous promoter of Mll. Here, the cassette was not integrated on the 5´ but still clones were sorted as mCherry+. This could be due to i) aberrant mCherry transcription from an internal cryptic start site, ii) not efficient

separation of cells during sorting due to low mCherry expression levels in general. We did not observe clones positive for both start and end fragments, indicating lack of cell lines with the full t(4;9) translocation.

Leukemic cells with translocations retain one copy of the wild-type allele of both MLL and AF9, so in order to mirror that, we also monitored their integrity in our clones. In the case of the 3 representative clones both wild-type Mll and Af9 were present in all cases. However, in the case of clone C, the band corresponding to Mll was migrating at a lower size than the rest of clones, which indicated a deletion. Sequencing of the fragment confirmed a 186bp-deletion in the intron 10 of Mll, which however did not affect Mll splicing and transcription. We also identified a number of aberrant clones that e.g., lacked wild-type alleles. This was most probably caused by the cassette insertion in both Mll alleles, meaning they were homozygous Mll knock-outs.

In order to achieve full translocation, we performed an extensive optimization study, where we tested and assessed a number of factors, summarized in Table 2. They were categorized in 3 groups, depending on the protocol part they were involved in. We tested, eg. different homology arm lengths (from 700bp to 2kbp), additional selection for the plasmid encoding gRNAs and Cas9 (fluorescence-based sorting or antibiotic resistance), a number of transfection reagents, various additives to favor homologous recombination and block NHEJ (RS-1, NU7441). All of the above conditions were tested

separately as well as in combinations, but unfortunately, they all failed to yield a successful translocation.

Table III.2. Optimization of the CRISPR/Cas9-mediated translocation in mouse ESCs.

pre-transfection	at transfection	post-transfection
increasing length of AF9 homology arms in the donor plasmid (700bp, 2kb)	transfection conditions (cell number, plasmid amounts, transfection reagents, incubation times)	selection cominations (cell sorting/antibiotic resistance)
selection method for the gRNA-encoding plasmid (GFP, puromycin)	number of transfections (single/serial)	sorting strategies (directly single-cell, bulk first, etc.)
	additives (NHEJ inhibitor NU7441, HR activator RS-1)	sorting timing after transfection (1-3 days)

Eventually, after numerous unsuccessful trials, we decided to apply a two-step strategy: isolate clones with cassette insertion on either ens, and induce the full translocation by using gRNAs targeting the other allele. We selected and validated a cell line that inserted the mCherry cassette into one Mll allele (Representative clone C from Figure 8). We designed and tested gRNAs targeting Cas9 to the Af9 locus and also, other ones, mediating the cleavage of the donor plasmid downstream of the Af9 homology arm. This was to ensure that this part was “free” and accessible for the subsequent recombination. After optimization rounds, we finally obtained a cell line with the full chromosomal translocation, and successfully validated the gene trap strategy.

After the aforementioned start/end PCR screening, we were able to identify a potentially positive clone, named TRL_mCherry (=containing the translocation between the Mll and Af9 genes, and the mCherry gene trap). We then confirmed the presence of the

translocation by amplifying the whole fragment of genomic DNA, and confirmed it with sequencing (Figure III.11).

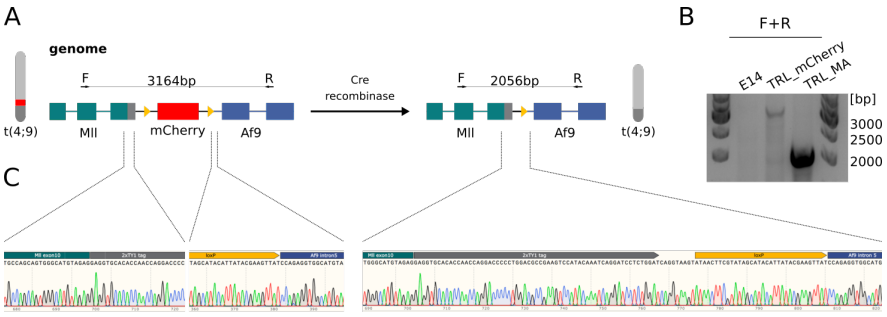


Figure III.11. Isolated positive cell line contains the induced t(4;9) chromosomal translocation. (A) Position of the primers and the amplified gDNA fragment sizes. mCherry cassette excision was induced by transiently expressing Cre at the ESC stage. (B) PCR reactions using the indicated primers for wild-type (E14), t(4;9) cell line (TRL_mCherry), and Cre-treated cell line (TRL_MA). (C) Sequencing of the amplified fragments.

The amplification and sequencing of the fragment corresponding to the translocation in the ES cells transiently exposed to the Cre recombinase showed proper excision of the mCherry cassette, validating the gene trap strategy. The Cre-treated cells were named TRL_MA, indicating that they contain the translocation and express Mll-Af9 fusion (shown below).

The cells containing the t(4;9) translocation expressed only Mll-mCherry fusion protein, whereas Mll-Af9 could only be detected after exposing the cells to Cre recombinase, which led to the loss of mCherry. We were able to show this specificity on RT-PCR, qPCR and Western blot (Figure III.12B, D and E, respectively). Moreover, sequencing of both fusion mRNAs showed proper splicing and inclusion of the TY1 tag in both fusion proteins (Figure III.12C).

Additionally, we observed a decreased expression of the wild-type Mll and Af9 genes in cells with the translocation, corresponding to their heterozygous knock-out induced as a result of the rearrangement.

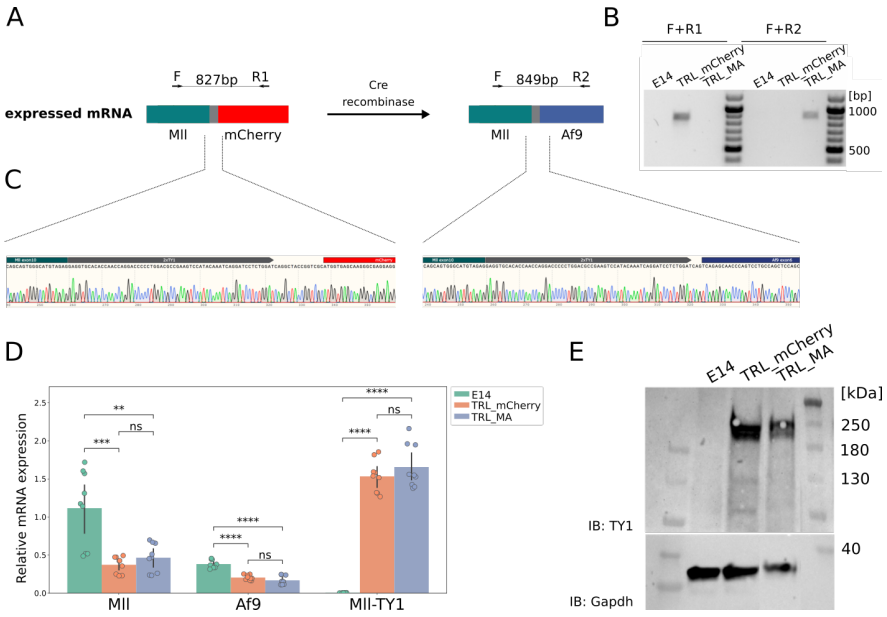


Figure III.12. Gene-trap strategy allows precise control over t(4;9)-induced Mll-Af9 expression (A) Scheme of the primers used for the specific amplification of the fusion Mll-mCherry and Mll-Af9 cDNAs together with the expected band sizes. (B) RT-PCR reactions run on agarose gels. (C) Sequencing of the purified fragments. (D) qPCR of wild-type alleles and fusion proteins in ES cell lines, normalized with Gapdh. Data from 3 independent experiments. Significance was determined with the t-Student test. (E) Western blot analysis of the expression of fusion proteins, with the indicated antibodies.

Finally, fusion protein expression was evaluated by Western blot, using the TY1 antibody. Despite being positioned in the middle of the proteins, it allowed for a successful Mll-mCherry and Mll-Af9 detection. Although the size of both proteins is around 180 kDa (185kDa for Mll-mCherry and 179kDa for Mll-Af9), they are visible

around the band corresponding to 250kDa, most probably due to post-translational modifications.

Taken together, we have achieved a mouse ES cell line with the t(4;9) translocation, which is the equivalent of the human t(9;11) leukemia-related rearrangement. Moreover, our gene-trap strategy proved to be successful, as Mll-Af9 protein expression was detected only after exposing cells to the Cre recombinase. Finally, the fusion proteins could be efficiently targeted using an antibody against the TY1 tag.

IV. At first glance, murine t(4;9) translocation in the absence of the oncogene expression only slightly affects hematopoiesis, but incorporating Mll-Af9 expression from the ESC stage blocks it.

Once the cell line was validated, in parallel with screening for more clones, we moved forward with the assessment of the effect of the induced chromosomal translocation on the hematopoietic differentiation. In order to do that, we took wild type ESCs together with the TRL_mCherry cell line and differentiated them towards hematopoietic progenitors.

At the stage of EB culture we did not observe any notable differences between cell lines, and the hemangioblast differentiation efficiency was comparable, as seen by the percentage of Flk-1+ cells (Figure III.13A). Blast cultures, however, showed various results. Out of 3 performed experiments, in 2 we noticed differences in the amount of differentiated cell types. By visual inspection, we observed more endothelial cells in cultures coming from the translocated cell line, compared to the wild-type cells, as shown in Figure III.13B. Flow cytometry analysis confirmed the difference, and pinpointed the showed differences in the hemogenic endothelial cells compartment. Cells with the translocation displayed a lower number of the HE1 population and a higher number of the more mature hemogenic endothelium (HE2), compared to the wild-type cells. The number of hematopoietic progenitors was comparable (Figure III.13C).

However, due to low sample size the differences were not significant, and more replicates will be needed to undoubtedly determine the effect. Regardless, the differences could not be attributed to increased apoptosis or cell death, as both cell lines displayed a similar mortality rate (Figure III.13D).

Marker gene expression in total blast cultures showed a similar expression of genes crucial for early stages of hematopoiesis, like Tal1 and Runx1 (Figure 10E). The markers of more specialized cell types were either comparable (Gata2), or downregulated, like Spi1, a marker crucial for myelopoiesis. Interestingly, we observed an upregulation of some known MLL-AF9 target genes (Hoxa5, Dnmt3b), while others were not affected (Hoxa9, Meis1). This was not expected, yet very interesting, as in those cells Mll-Af9 is not yet expressed. Again, sample size was not sufficient for the differences to be statistically significant.

In order to assess specifically the hematopoietic progenitor populations, they were sorted based on marker genes, and subjected to qPCR analysis (Figure III.13F). As expected, both wild-type Mll and Af9 were downregulated in the cells containing the translocation. The tested hematopoietic genes showed comparable expression in the two conditions. Interestingly, Spi1, which showed a notable expression difference in the total blast cultures, in the progenitors was expressed at a similar level, suggesting that it might be e.g., deregulated in the endothelium. Taken together, we observed a possible difference in hematopoietic differentiation of mouse ES

cells caused by changes in the genome due to the induced translocation, but more experiments and replicates are needed in order to draw definite conclusions.

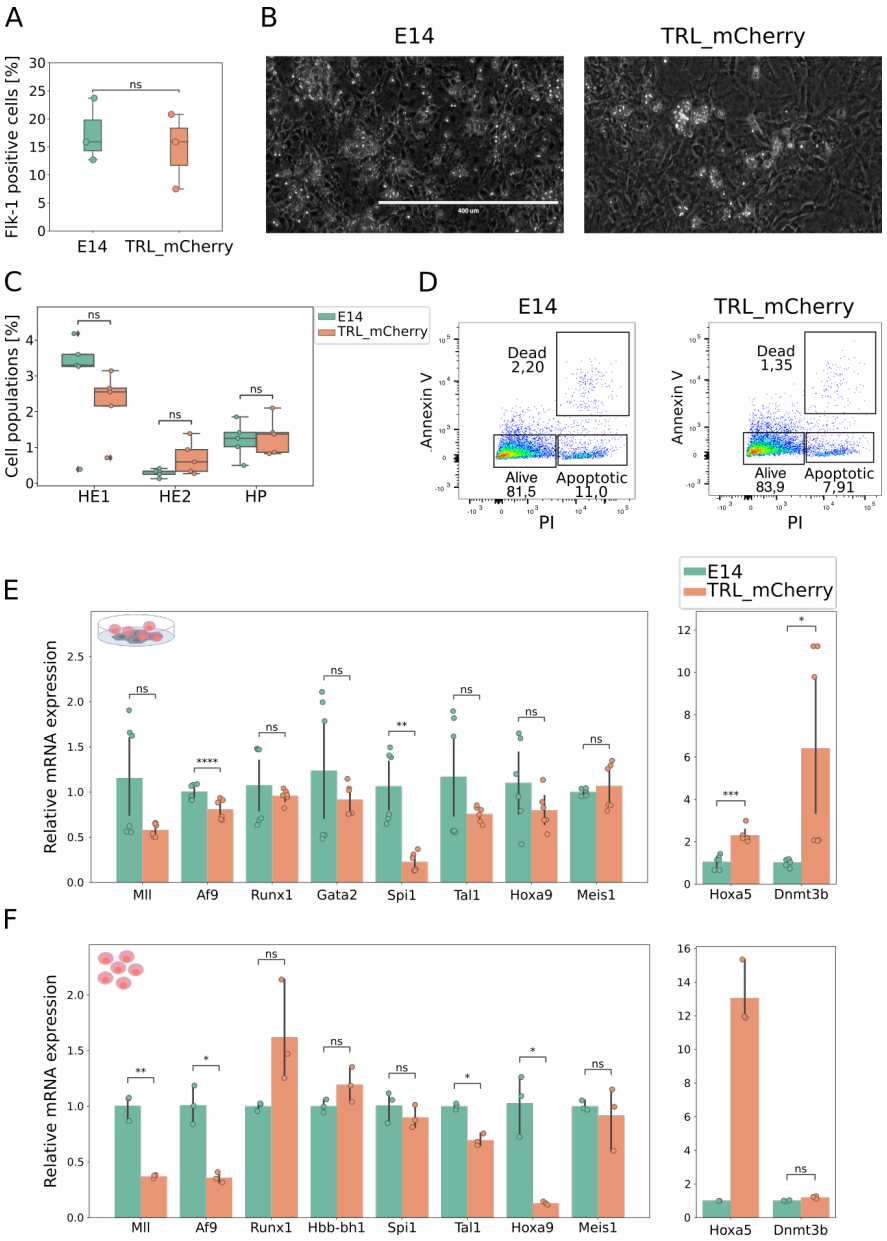


Figure II.13. Chromosomal translocation-induced background does not strongly affect hematopoietic differentiation. (A) Flk-1+ cells at d5 of EB differentiation. (B) Blast cultures at day 4 of culture. (C-E) Assessment of blast cultures at day 4 of differentiation: Flow cytometry analysis of the cell populations (C, data from 3 independent experiments), Annexin V staining (D), and qPCRs illustrating marker gene expression (E, data from two independent experiments). (F) Marker gene expression in sorted HPs. Data from 3 technical replicates. (E, F) Relative expression was calculated over the wild-type cells. (A,C,E,F) Significance was calculated using the t-Student test.

Our preliminary, yet promising results prompted us to investigate further our translocation model and include the expression of the Mll-Af9 protein. To achieve it, we unlocked its expression at the stage of ES cells by transient transfection with the Cre-expressing plasmid and selection based on the reporter GFP expression. Thus, the obtained mouse ES cells contained both the chromosomal translocation and expressed Mll-Af9.

After plating in the differentiation media, the TRL_MA cells formed EBs comparable to both wild-type and TRL_mCherry cells. Strikingly, hemangioblast sorting at day 5 showed almost complete abrogation of this population specifically in the cells expressing the Mll-Af9 fusion protein (Figure III.14A). Whereas both wild-type cells and cells with the translocation yielded a similar number of hemangioblasts (approx. 15%), in the case of TRL_MA the number was around 0.1-0.2%. When plated in blast-promoting media, the cells yielded endothelial-like cells with abnormal morphology. Moreover, “cobblestone-like” adherent cells and floating hematopoietic progenitors were completely absent (Figure III.14B). Compared to wild-type and translocation, blast cultures showed a high number of dead cells. Flow cytometry confirmed the lack of

hemogenic endothelium and progenitors, but interestingly a significantly higher number of cells expressed cKIT, as compared to wild-type and TRL_mCherry conditions (Figure III.14C D).

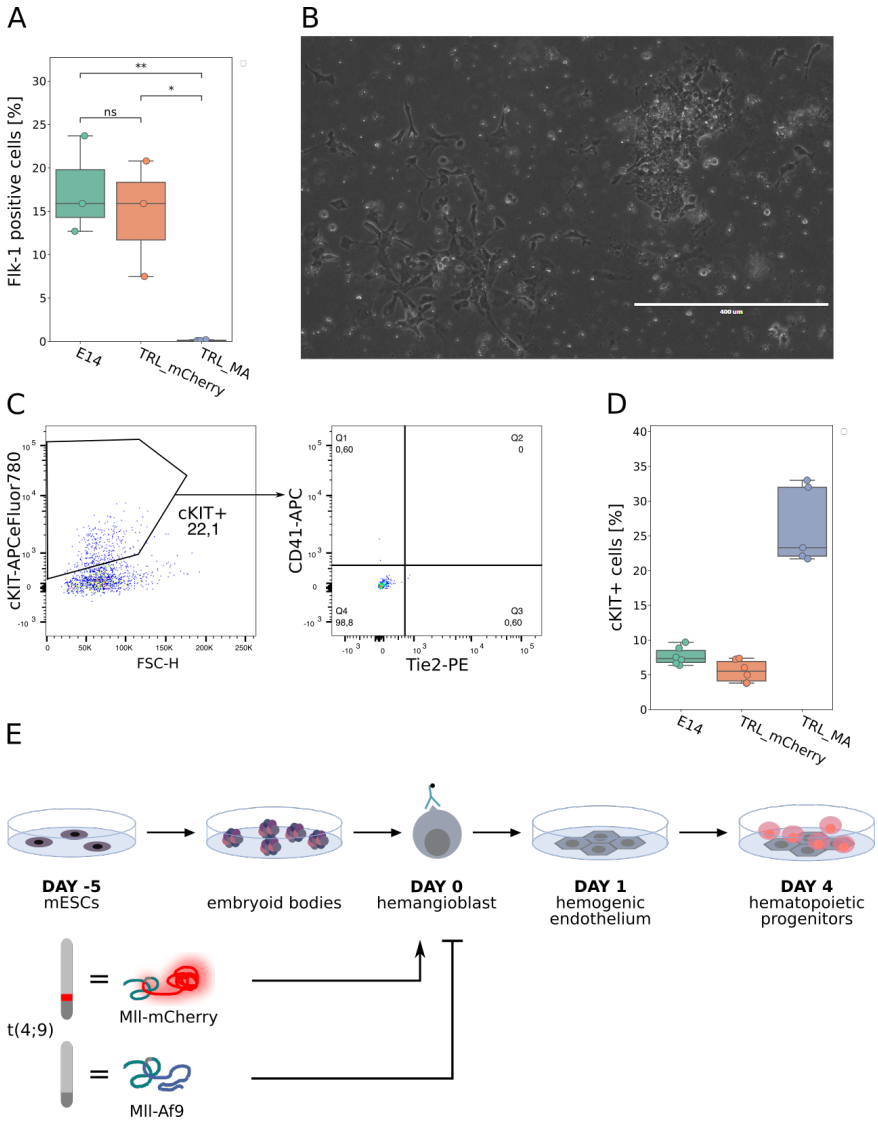


Figure III.14. Mll-Af9 expression in ES cells strongly impairs hematopoietic differentiation. (A) Quantification of Flk1+ cells at fay 5 of EB culture. Data coming from 2 independent experiments. Significance was calculated using the t-Student test (B) Representative image of TRL_MA blast culture at day 4 of blast culture. (C) Flow cytometry analysis of the blast culture showing complete lack of

CD41 and Tie2 markers. (D) Quantification of cKIT+ cells for wild-type cells, TRL_mCherry and TRL_MA. Data coming from 2 independent experiments. (E) Proposed model of the effect of the chromosomal translocation and expression of the Mll-Af9 fusion protein on hematopoietic differentiation of ES cells.

Based on the obtained results, we were able to determine that, whereas the translocation-induced change in the genome of mouse ES cells, seemingly only slightly affects hematopoietic differentiation. However, if the protein expression is "switched-on" at the ES cell stage, it blocks hematopoietic differentiation and results in the abrogation of the hemangioblast population Figure III.14E).

Taken together, we applied the CRISPR/Cas9 system to build leukemia models in mouse ES cells. We achieved the equivalent translocation to the human leukemia-related t(9;11). Apart from that, we developed cell lines inducibly expressing mouse Mll-Af9. Based on our preliminary results, we can conclude that, in embryonic development, the timing of the Mll-Af9 expression has an important effect. When expressed in ES cells, it blocks almost completely hematopoietic differentiation upstream of the hemangioblast, and the few cells that manage to "escape" this block are not able to yield hematopoietic colonies. When expressed later in differentiation, after the hemangioblast differentiation stage, Mll-Af9 has a mild effect on hematopoietic progenitors. On top of all that, the changes caused due to the induced translocation (genome structure change, heterozygous knockout of Mll and Af9) seem to affect the differentiation in a fusion protein-independent manner.

DISCUSSION

Chromosomal translocations are a hallmark of approximately 10% of cancers, found predominantly in leukemias and lymphomas. The two main effects include placing an oncogene under a strong promoter or, when breakpoints occur in coding regions of two genes, formation of a fusion gene and expression of a chimeric protein (Figure I.8). The latter affects most of the translocation-related leukemias, with the expressed protein being the driver of the disease. One example of such rearrangement is the t(9;11) translocation, affecting the histone methyltransferase MLL and the super-elongation complex member, AF9 gene. Upon joining of the two gene fragments, the fusion MLL-AF9 protein is expressed.

In this Thesis, we aimed to further characterize the mechanistic events leading to malignant transformation in the t(9;11)-related leukemia. As a model system, we chose mouse ESCs for their virtually unlimited self-renewal and expansion ability and experimental flexibility. Engineered cell lines can be used to create mouse models or test in differentiations to the three lineages (endoderm, mesoderm, ectoderm). With this approach, we could also tackle the question of whether chromosomal translocations occur only in defined cell types, or are ubiquitous but offer proliferative advantage only in some cases. Furthermore, it was already shown that inducing the Mll-Af9 translocation in mice leads to leukemia, which suggested the oncogenic potential of mouse fusion proteins. We validated that possibility by first, determining the high sequence identity for several human leukemic fusion proteins and their mouse equivalents (Table III.1). Secondly, using two chosen

representatives, Pml-Rara and Mll-Af9, we showed their ability to immortalize primary mouse hematopoietic stem cells, and in the case of Pml-Rara, a drug-dependent response (Figures III.2-3). Based on these results, we proceeded with our Mll-Af9 translocation model in mouse ESCs.

We chose mouse pluripotent cells over the human because of their easier manipulation *in vitro*, perspective of creating a mouse model and a better-established protocol for differentiation towards hematopoietic stem and progenitor cells. We applied the differentiation protocol of mouse ESCs towards hematopoietic progenitors and determined that this system could efficiently express the Mll-Af9 fusion protein (Figure III.5-6 and III.7, respectively). This was achieved by developing cell lines containing a cassette for doxycycline-inducible Mll-Af9 protein in the Rosa26 locus of ESCs. We validated the activity of the murine fusion protein by detecting marked upregulation of Hoxa5 and Hoxa9, some of the well-known MLL-AF9 target genes (Figure III.8).

Going further, we took advantage of the CRISPR/Cas9 technology to obtain the chromosomal translocation between genes Mll and Af9, resulting in t(4;9) rearrangement. Of note, because the translocation partner genes in mouse and human lie on different chromosomes, the structural changes in the nuclear space will not be fully equivalent to the patients with t(9;11) translocation, urging for cautious assessment of the obtained results. We combined the translocation with the gene-trap strategy to precisely control the timing of Mll-Af9 expression

(Figure III.9). This was achieved by inserting a mCherry cassette equipped with splice donor (SD, fused to the last exon of Mll) and splice acceptor (SA, fused to mCherry) sites between the parts coming from the two genes. Additionally, loxP sites were placed flanking the SA site and mCherry. The chromosomal translocation and correct insertion of the cassette were confirmed by amplification of the whole genomic region and sequencing (Figure III.11). We showed on both mRNA and protein level that, upon translocation, only Mll-mCherry fusion was expressed. Mll-Af9 was detectable only when the engineered mouse ESCs were exposed to the Cre recombinase, here achieved by transient transfection of an expression plasmid (Figure III.12A-C, E). We also assessed the expression of the wild-type Mll and Af9 alleles and showed their decreased expression after the translocation, consistent with their heterozygous KO as a result of the rearrangement (Figure III.12D). Having validated the system, we could conclude that we created a *bona fide* model of human t(9;11) leukemia in mouse ESCs, which will allow us (and other laboratories world-wide) to assess the different factors contributing to this disease.

CRISPR/Cas9 is a very powerful tool for building cell models. It allows to induce e.g., gene KOs, knock-ins (KIs), inversions and translocations. However, in order to obtain a homogeneous cell population, expansion from a single cell is necessary. This in turn can cause genetic drifts and introduce clone-to-clone variability. Apart from the desirable modification, CRISPR/Cas9 was reported to cause, in some cases, important secondary off-target effects. As an

example, large, megabase-scale deletions and loss-of-heterozygosity was observed in both HEK293T and human CD34+ hematopoietic stem cells treated with Cas9 and gRNAs targeting the beta globin locus (Boutin et al., 2021). All of these observations urge for including in the studies several single-clone derived cell lines to ensure that the assessed effect can be undoubtedly attributed to the introduced modification.

Based on this, our main focus was to obtain replicates of the cell line with translocation, which proved to be extremely difficult. So far, the developed models with *de novo* induced MLL-AF9 translocation comprised either isolated hematopoietic cells, both human and murine, and whole mice, where the rearrangement was induced *in vivo*. However, in most cases a bulk population without selection was used, thus containing wild-type cells, cells with other types of modifications and cells containing the translocation (Breese et al., 2015; Sarrou et al., 2020; Schneidawind et al., 2018). Breese and colleagues induced in parallel the MLL-AF9 and MLL-AF4 translocations, and estimated the frequency, but only for MLL-AF4. Based on their calculations, they were able to achieve the translocation approximately in 1 every 40.000 cells, which gives an idea about the efficiency of inducing translocations in human hematopoietic stem and progenitor cells. High efficiency was achieved by Jeong and colleagues, by using CRISPR/Cas9 technology in human CD34+ hematopoietic stem and progenitor cells (Jeong et al., 2019). After transduction, they serially diluted the population and analyzed the genomic DNA after 2 weeks of culture.

They detected amplicons corresponding to the MLL-AF9 translocation in samples coming from starting concentrations of 200cells/sample, which translates to a 0.5% efficiency. However, in this case also single-cell selection was not performed. The aforementioned studies reflect the difficulty of inducing chromosomal translocation, and so in cells that are thought to be targets of these rearrangements in leukemias. This is even more challenging in mouse ESC, where such translocations do not occur naturally.

In contrast to the discussed reports, we aimed to perform an in-depth study on how the MLL-AF9 chromosomal translocation affects the cell state. In order to do that, we needed pure and homogeneous single cell-derived populations. In order to improve screening efficiency, we took advantage of a mCherry-mediated gene trap, where the expression of the reporter gene depends on the endogenous Mll promoter and allows for fluorescence-based selection (Figure III.10). However, this approach only forced for the donor plasmid integration at the 5´-end, into the Mll gene. Mll-mCherry would be expressed regardless of the recombination happening also at the 3´-end, namely the Af9 gene, and this is what we observed. We only detected clones with one-sided insertion of the donor plasmid, either in the Mll or Af9 locus. The latter one was unexpected as in this case the mCherry, despite lacking an endogenous promoter, was expressed. This could occur as a result of erroneous integration in a different genomic locus or presence of a cryptic promoter in the Mll homology arm of the donor plasmid, among other possibilities. With

respect to the first hypothesis, we were able to confirm that the mCherry cassette integrated in the Af9 gene retained almost 200bp of the backbone of the donor plasmid upstream of the Mll homology arm, indicating a possibility of this end being "free". However, we cannot exclude that the integration happened even more upstream. It was reported that Mll contains a cryptic promoter, placed in the intron 10, which resulted in expression of a N-truncated Mll mRNA isoform and translation into protein (Scharf et al., 2007). Our 5´-homology arm in the donor plasmid comprises 672bp spanning part of exon 9, intron 9 and exon 10 of Mll, thus it is positioned upstream of the postulated internal promoter. However, we cannot rule out the possibility of the presence of more of such promoters. Nevertheless, if it was true, we could speculate that this should result in comparable frequencies of both types of selected clones, and in fact we consistently isolated a lot more clones with the donor plasmid insertion in the Mll gene, compared to the Af9.

Despite our extensive effort, we were not able to identify cell lines positive for the full t(4;9) translocation. To address this limitation, we applied a two-step strategy, where we first isolated a clone with the mCherry cassette insertion in the Mll locus, and afterwards forced the full translocation by targeting the Af9 locus. This approach proved to be successful, but still the frequency was low (1 in 288 screened clones). Currently, we are performing a massive-scale screen in order to identify more positive cells. Additionally, we are currently setting up the whole-chromosome painting technique

(FISH), to definitely demonstrate the rearrangement on the chromosome level.

We aimed to investigate how the t(4;9) chromosomal translocation affects the cell in the absence of the expressed Mll-Af9 fusion protein. To achieve that, performed the hematopoietic differentiation of the translocation-containing cell line (TRL_mCherry) and used wild-type ESCs (E14) as a control (Figure III.13). The first part of the protocol, namely embryoid body culture, did not show any notable differences between conditions, suggesting that the presence of the translocation *per se* (in the absence of the expression of the chimeric oncogene) and the heterozygous KO of Mll and Af9 did not strongly affect the ability of mouse ESCs to differentiate. No significant differences were also noted at the time of hemangioblast selection, indicating that differentiating ESCs could progress towards the hematopoietic lineages. When further differentiated, hemangioblasts from both conditions yielded the cell lineages in the blast cultures: hemogenic endothelium (HE), non-hemogenic endothelium, hematopoietic progenitors (HPs) and vascular smooth muscle cells (VSM). This suggested further that the t(4;9) translocation was permissive in the context of hematopoiesis. Interestingly, in 2 out of 3 performed experiments, we observed a higher number of endothelial cells in the TRL_mCherry sample. Flow cytometry analysis revealed further differences in the HE compartment, with concomitant loss of cells belonging to the type 1 and increase of cell number in type 2 (more mature) HE, in the case of cells with the chromosomal translocation. We speculated that this

shift in cell populations could reflect the faster maturation of the HE, which however did not translate to more efficient endothelial-to-hematopoietic transition and larger population of HPs. Gene expression analysis on the total blast population and purified HPs revealed no major differences in the expression of genes crucial for early steps of hematopoiesis, such as Tal1, Runx1 or Gata2. Interestingly, we observed a significant downregulation of the Spi1 marker in the TRL_mCherry cell line compared to wild-type, but only in total blast samples. Purified HPs showed no differences in Spi1 expression levels. The same dependency could be observed for Dnmt3b. This could suggest that the effect of both Spi1 and Dnmt3b could concern other cell types in the blast cultures, like the (non)hemogenic endothelium or smooth muscle cells.

Furthermore, we wanted to assess the expression of Hox genes. They were identified as targets of MLL, but also AF9, and reported to be upregulated in t(9;11) leukemias, as a consequence of MLL-AF9 activity. We wondered whether they would be affected only by the translocation-driven changes and decreased Mll and Af9 expression levels. We hypothesized that Mll and Af9 downregulation might negatively affect their expression. In line with that, we observed a decrease in Hoxa9 expression, especially marked in the HP population. This was already observed in Mll KO mouse ESCs (Ernst et al.). Gene expression analysis during EB differentiation revealed a broad Hox gene downregulation. Same as in our experiments, Flk-1 expression pattern was not disturbed, pointing towards a non-essential role of Mll for hemangioblast specification. Strikingly, we

detected a marked upregulation of the Hoxa5 expression, both in blast cultures and purified HPs containing the chromosomal translocation. It was reported that overexpressing another one of the HOX proteins, the human HOXB4 caused a significant increase in the HE population, without perturbing the non-hemogenic endothelium development. However, a similar role for Hoxa5 was not reported, which could provide a possible explanation for enhanced endothelial specification in TRL_mCherry blast cultures. In contrast, it has been shown that increasing its expression levels in endothelial cells induced senescence and blocked vasculogenesis (Cuevas et al., 2015). However, whereas Mll-Af9 expression induced a marked Hoxa5 upregulation in the blast cultures (MAD cell lines, Figure III.8), we did not observe any difference in the endothelial population there. This would argue against the role of Hoxa5 as a unique regulator of the development of endothelial cells. In order to determine whether Hoxa5 is involved in the increased endothelial lineages in our blast cultures, we will need to assess gene expression patterns in each of the cell lineages purified from the blast cultures separately and also include the Flk-1+ hemangioblasts for registering any possible changes occurring at the pre-blast stage. Af9 was also reported to be a master regulator of the Hox genes, based on homozygous mutant mice studies, where the animals showed broad homeotic transformations. We will also need to examine a possible effect of Af9 expression reduction in TRL_mCherry cells.

We are currently investigating the observed effect of enhanced endothelial specification caused by the t(4;9) translocation. A similar

phenotype was observed in the context of Sox7 overexpression, a marker crucial for the establishment of the hemogenic endothelium (Costa et al., 2012). Additionally, Sox7 overexpression led to Spi1 downregulation, consistent with decreased Spi1 levels in our TRL_mCherry blast cultures. However, the authors report that the enhanced HE formation had a negative effect on the HP population. In contrast, t(4;9) translocation did not impair HP emergence. The difference could be due to i) different Sox7 expression levels in both models, ii) involvement of other TFs and mechanisms in the observed TRL_ mCherry phenotype.

Interestingly, it was reported that fusing an oligomerization module to the Mll part was sufficient for inducing leukemia. Insertion of the LacZ gene yielded both myeloid and lymphoid leukemias, although with long latency periods (Dobson et al., 2000). A similar leukemia-inducing effect was reported upon fusing the FK506 binding protein (FKBP12) to the Mll portion (Eguchi et al., 2004). This approach allowed for inducible dimerization upon adding the AP20187 compound. Mll-FKBP12 dimerization induced cell immortalization, blocked differentiation, increase in several target genes in Mll-rearranged leukemia (Hoxa7, Hoxa9, Meis1). In our model, after the chromosomal translocation the Mll-mCherry chimera is expressed. mCherry is monomeric, suggesting that this fusion protein does not oligomerize, concomitant with the opposite effect on the Hoxa9 expression from the reported above. However, to confirm or deny any role of Mll-mCherry in hematopoiesis, we will need to assess its

possible interaction partners and/or involvement in signaling pathways.

Whereas the chromosomal translocation alone did not grossly affect the hematopoietic differentiation, releasing the Mll-Af9 expression in ESCs blocked severely impaired it. The oncogenic fusion expression was “switched on” by transient Cre expression. We proved that these cells (TRL_MA) lost the Mll-mCherry expression in favour of Mll-Af9 (Figure III.12). When differentiated into EBs, TRL_MA cells did not display any noticeable differences neither in size nor number of EBs. Strikingly, Flk-1 staining, marking the hemangioblast, revealed an almost total abrogation of this precursor in the case with the cells containing both the t(4;9) translocation and Mll-Af9, compared to E14 and TRL_mCherry cells (0.1% vs. median 15%, Figure III.14). Considering that neither the cells with the chromosomal translocation, nor the un-induced MAD cell lines displayed such defect, we hypothesize that Mll-Af9, if expressed in ESCs, severely impairs hemangioblast differentiation. When plated in blast-promoting media, the (very limited in numbers) TRL_MA Flk-1+ cells yielded only endothelial-like cells, with lack of HE, HPs or vascular smooth muscle cells, confirmed by flow cytometry analysis of marker genes. Interestingly, we observed a significant upregulation of the cKIT expression, together with loss of CD41 and Tie2 markers. This further supports our initial observation of improper hematopoietic differentiation.

cKIT is a ubiquitous cell marker, expressed on several cell lineages, including ESCs. We will need to examine the embryoid bodies differentiated from TRL_MA cells and track the mesodermal specification steps upstream of the Flk-1 marker. This will allow us to determine the effect of Mll-Af9, which could occur through: i) block of mesoderm specification in favor of other lineages, ii) block of mesodermal differentiation at a specific precursor stage, iii) other possible mechanisms. Hox genes are crucial for hematopoietic differentiation and need to be tightly regulated. It is possible that their sustained activation from the ESC stage by Mll-Af9 impairs this process.

Another study used a different MLL fusion protein, MLL-AF4. It was virally overexpressed in human ESCs (Bueno, 2012). Upon EB differentiation, the fusion protein enhanced endothelial specification, increasing 3-fold the population of hemogenic precursors. Furthermore, these precursors were skewed towards the endothelial lineage, with impairment in the hematopoietic cells compartment. This seemingly goes in contrast to our results, although several factors have to be accounted for here: differences between species (human vs. mouse), expression levels (overexpression vs. physiological levels), and difference between proteins (MLL-AF4 vs. Mll-Af9). It was shown that expression levels are a crucial factor for leukemogenesis, illustrated when transforming GMPs. Retrovirus-mediated MLL-AF9 overexpression readily yielded leukemia in transplanted mice, whereas knocking of the AF9 sequence into Mll locus and physiological levels failed to do so (Chen et al., 2008;

Krivtsov et al., 2006). Apart from that, it was shown that, despite having a large part of the protein in common, MLL-AF4 and MLL-AF9 affect the transcriptional program in different ways (Prange, 2017). Determining the scale of the effect on Mll-Af9 on ESC differentiation will allow us to further understand the differences between the two MLL fusion proteins.

Genome structure is a crucial determinant for proper cell functioning. This control has to be carried out on multiple levels, from ensuring that the chromosomes occupy their designated territories to maintaining genes separate or together in terms of interaction-mediated activation or repression. Such control over the genes is achieved by e.g., segregating them into the same of different topologically-associated domains. Any disruption of specific borders between TADs can lead to severe impairments. It was shown that disrupting the boundary between neighboring TADs segregating genes crucial for proper limb development lead to severe malformations (Lupiáñez et al., 2015). In the context of AML, it was shown that deleting a boundary dividing the HOXA genes, lying between HOXA7 and HOXA9, disrupted the long-range interactions between the posterior HOXA genes and as a consequence reduced the leukemic blast burden (Zhang et al., 2020). In MLL-AF9-driven leukemias, this fusion protein was postulated to modulate genome structure by targeting regions occupied by CTCF (Prange, 2017). It will be of great importance to investigate in our model i) how the induced chromosomal translocation affects genome structure in

ESCs and during differentiation, and ii) whether Mll-Af9 has an additional modifying role in this context.

Taken together, we combined CRISPR/Cas9, gene-trap strategy and endogenous tagging, to create the t(9;11) leukemia-related model in mouse ESCs. Despite the fact that much effort will be needed in order to pinpoint specific factors and pathways yielding the leukemic cell transformation, our approach offers an innovative and flexible model that allows for dissection of the different layers contributing to the phenotype and precise targeting of the leukemic fusion protein.

CONCLUSIONS

Based on the results presented in this Thesis, the following conclusions can be drawn:

1. Mouse fusion proteins can be used as a proxy for human leukemic chimeric oncogenes;
2. Engineered t(4;9) translocation is viable in mouse ESCs and provides a *bona fide* model of t(9;11) leukemia-related rearrangement;
3. Translocation-driven structural changes in the genome of mouse ESCs in the absence of Mll-Af9 fusion protein do not grossly affect their ability to differentiate towards hematopoietic lineages;
4. The combination of chromosomal translocation and Mll-Af9 expression in mouse ESCs severely impairs their hematopoietic differentiation;
5. Gene-trap strategy provides efficient and tight regulation of the Mll-Af9 fusion protein expression;
6. Internal tagging of the Mll-Af9 with the TY1 tag allows for efficient protein targeting.

MATERIALS AND METHODS

Cell culture

E14Tg2a mouse ES cells were grown on 0.1% gelatine-coated (Millipore, ES-006-B) plates in Glasgow's Minimum Essential Medium (GMEM) supplemented with 20% fetal bovine serum (FBS, Hyclone, 12389802), 1x Glutamax (Gibco, 25030-024), 1x non-essential aminoacids (NEAA Gibco, 11140-050), 1x sodium pyruvate (Gibco, 11360-070), 100uM 2-mercaptoethanol (Millipore, ES-007-E), 1x penicillin/streptomycin (P/S) and homemade mouse leukemia inhibitory factor (LIF), at a concentration determined on a batch-to-batch basis.

After transfection, ES cell selection was performed by adding puromycin at a concentration 1ug/ml (A1113803, Life Technologies) and daily media changes, or alternatively by FACS sorting based on the reporter fluorescence.

Mouse hematopoietic stem and progenitor cells were isolated from bones of sacrificed mice using the Lineage Depletion Kit (Miltenyi Biotec, 130-090-858). Briefly, mouse bones were crushed in a mortar, washed in PBS and filtered to isolate whole bone marrow cells. Next, they were stained with a biotinylated cocktail of lineage-specific hematopoietic markers and subsequently anti-biotin microbeads. Negative selection was performed using a magnetic separator and appropriate columns (QuadroMACS separator, 130-091-051, and LS columns, 130-042-401, both from Miltenyi Biotec). Selected negative fraction (flow-through), which contained hematopoietic stem and progenitor cells (lin-) was cultured in non-

tissue culture treated plates in StemSpan media (Stemcell technologies) supplemented with stem cell factor (SCF), Flt3 ligand and IL-11 (all from Peprotech).

Colony-forming assay was performed by plating sorted lin- cells in methylcellulose-based medium (MethoCult™ GF M3434, Stemcell technologies, 03434). After every 7 days, colonies were counted, cells were replated and subjected to analysis. For the ATRA treatment, the drug was added at the moment of plating at a final concentration of 1ug/ml.

Retrovirus production and infection

Sequences of mouse fusion proteins were determined by combining the coding regions of mouse genes corresponding to the parts of partner genes involved in leukemic chromosomal translocation. As an example, Mll-Af9 was constructed by combining exons 1-10 of Mll and 6-11 of Af9, which corresponds to the MLL-AF9 fusion protein found in patients with t(9;11)-related leukemia.

Constructing vectors for viral overexpression was performed by cloning amplified coding regions into a modified retroviral vector (PINCO) using Gibson cloning technology. Additionally, a small tag was inserted in the middle of the two parts (Mll-2xTY1-Af9 and Pml-2xTY1-Rara). The vector was modified in terms of the promoter of the GFP reporter, where pCMV was replaced with the Ef1a promoter.

Retroviral particles were produced using the Phoenix-ECO cells. Briefly, at day 0 cells were plated in 10-cm dishes at a density 2mln cells per plate. The following day, they were co-transfected using the calcium phosphate method with10ug of the retroviral plasmid of interest vector and 5ug of the packaging ECO plasmid per plate. The next day, the media were changed. 48h and 72h after infection, supernatants containing viral particles were collected, filtered through a 45um filter, mixed with PEG800 in a 1:4 ratio (10ml PEG for 40ml of supernatant) and kept overnight at 4 degrees. Following, they were spun down at 1500xg for 30 minutes, resuspended in PBS and stored at -80 degrees.

Lin- cells were spin-infected 2 days after isolation from mice, in retronectin-coated 24-well plates (RetroNectin reagent, Takara T202), by centrifugation at 2500rpm for 1h. Next, cells were collected and plated in StemSpan media for 2 days before sorting.

CRISPR/Cas9 gene editing of mouse ESCs

Donor plasmid for Rosa26 insertion was constructed by cloning the Mll-2xTY1-Af9 coding sequence under a doxycycline-inducible tight TRE promoter, the rTtA repressor and a cassette for puromycin-resistance into a vector containing homology arms for the Rosa26 locus. Both this vector and a plasmid encoding Cas9 and a sgRNA targeting the Rosa26 locus were a gift from Manuel Irimia´s lab (CRG). Cloning was performed using the Gibson technology.

The donor plasmid for translocation was constructed by the Protein Engineering Facility (CRG) by assembling synthetic blocks. The sequence of fragments corresponding to homology arms for both Mll and Af9 genes were confirmed by sequencing the corresponding amplicons from the gDNA of our E14 mouse ES cells. Plasmid variants were cloned by replacing the Af9 homology arm region with a longer fragment.

gRNAs targeting introns of Mll and Af9, as well as the donor plasmid were designed using the Genome Browser CRISPR 10K tool and CRISPOR software. They were cloned in pSpCas9(BB)-2A-GFP (PX458) and pSpCas9(BB)-2A-Puro (PX459) vectors (both were a gift from Feng Zhang, Addgene #48138 and # 62988, respectively). The gRNAs were validated using the surveyor assay and the most efficient ones were used in transfections.

Mouse ES cells were transfected with donor plasmids and CRISPR/Cas9 components using a reverse transfection method. The transfection reagents and vector amounts were optimized on a case-to-case basis. The media was changed after 18h of incubation, and cells were subjected to puromycin selection (starting 24h post-transfection), or cell sorting (48h post-transfection), depending on the reporter gene present in the plasmids. After puromycin selection, single clones were picked, expanded and validated by PCR amplification of the corresponding gDNA fragment, using primers listed below. In the case of sorting, if performed in bulk, cell population was first expanded and validated by genotyping. If the

expected PCR band was detected, cells were plated for picking and single clones were checked. If cells were sorted in a 96-well plate (single-cell sorting), clones were expanded and validated separately.

Table VI.1. Primers used for gDNA genotyping and gRNA cloning (all 5´- 3´)

Primer	Forward primer	Reverse primer
Rosa26 screening 5' end	CGCTGATTGGCTTCTTTTCCT	CCTCCCCCTGAACCTGAAACA
Rosa26 screening 3' end	GTGGATGTGGAATGTGTGCGA	GGGGGAGGAGACATCCACCT
Translocation screening 5' end	GAGAAGCCACCTCCAGTAAGTAA	CCTGATTTGTATGGACTTCGGCGT
Translocation screening 3' end	GGGATGCGGTGGGCTCTAT	GGGATGCGGTGGGCTCTAT
Translocation screening full insert	GAGAAGCCACCTCCAGTAAGTAA	CAAGTGACACTCTGACACTGTTCT
Mll wild-type allele	GTGTGGGAGATGGGAGGCTTA	GCCAGCACACCTGGACCTG
Af9 wild-type allele	GTACTCCCACTGGATTGTGCA	GGGATGCGGTGGGCTCTAT
Mll gRNA	CACCGGCACAGTGGCATCATCACAG	AAACCTGTGATGATGCCACTGTGCC
Af9 gRNA	CACCGAGATACAACCTGCTTCGTAA	AAACTCTTAAGCCACTAACACCTTC
Af9 gRNA #2	CACCGAAGGTGTTAGTGGCTTAAGA	AAACTTACGAAGCAGGTTGTATCTC
RHA gRNA	CACCGTAATGAATCGGCCAACGCGC	AAACGCGCGTTGGCCGATTCATTAC
RHA gRNA #2	CACCGTTGCGCGCTTGGCGTAATCA	AAACTGATTACGCCAAGCGCGCAAC

RHA - right homology arm (sequence in the donor plasmid downstream of the AF9 homology arm)

Fluorescence-activated cell sorting (FACS)

Analysis of marker gene expression or transfection efficiency was performed using a LSR II Analyser or Fortessa analyser. For cell sorting, either the Influx Cell Sorter or FacsAria Cell Sorter were used (all of the equipment belongs to BD Biosciences). Cell viability of differentiated blast cultures was assessed using the annexin V kit (Invitrogen), according to the manufacturer's instructions. Data was analysed using the FlowJo software.

Table VI.2. List of antibodies used for FACS analysis.

Antibody	Conjugate	Provider	Cat. number
CD41	APC	ThermoFisher	17-0411-80
Tie2	PE	eBioscience	12-5987-81
cKIT	APCeFluo780	Invitrogen	47-1171-82
cKIT	biotin	Invitrogen	13-1171-82
Flk1	biotin	eBioscience	13-5821-85
streptavidin	PE	BD Bioscience	554061
streptavidin	PE-CF594	BD Bioscience	562284
Sca1	PECy7	Invitrogen	25-5981-81

Karyotyping of mouse ES cell lines

Isolate cell lines were karyotyped by microscopic analysis of metaphase spreads. Briefly, ES cells were treated with 0.2ug/ml KaryoMax Colcemid Solution (Gibco, 15212012) for 3 hours. Following, the cells were collected and pelleted. Cells were disrupted in ice-cold hypotonic 0.56% KCl solution, and chromosomes were fixed in ice-cold fixative solution (methanol: glacial acetic acid, 3:1). Suspensions were dropped on microscopic slides, dried and stained using Fluoroshield with DAPI (Sigma, F6057). Images were taken using an inverted microscope (Leica).

Cytospin

0.1-0.2mln lin- cells were collected, washed with PBS and placed into the cytofunnels mounted on microscope slides. Cells were then centrifuged for 3 minutes at 350rpm and left in room temperature to dry completely.

Immunostaining

Cells attached to the slides were fixed in 4% paraformaldehyde, permeabilized in 0.1% triton in PBS+1%BSA, stained with the anti-Pml polyclonal antibody (Invitrogen, PAS-79836, raised in rabbit). As a secondary antibody, Alexa 488-conjugated anti-rabbit antibody was used (Molecular probes, A21206). Stainings were performed for 1 hour at room temperature in the dark. Subsequently, slides were washed, mounted with Vectashield with DAPI, and analyzed on an inverted microscope (Leica).

Wright-Giemsa staining

Slides were stained in Wright-Giemsa solution for 1 minute in Coplin glass. Subsequently, they were stained in Giemsa solution for 10 minutes. After washing and drying, slides were analyzed using an inverted microscope (Leica).

Hematopoietic differentiation of ESCs

ES cells were differentiated towards hematopoietic progenitors using a published protocol. Briefly, stem cells were first conditioned by splitting in growth media where GMEM was substituted with Iscove's Modified Dulbecco's Medium (IMDM, Lonza 12-722F). After 2 days, cells were collected and plated in non-tissue culture treated 10-cm plates bacterial plates (Stelinin, ThermoScientific) in EB differentiation media, at 0.3mln cells in 25ml per plate. The differentiation medium was as follows: IMDM, 15% FBS, 1% Glutamax, 50ug/ml acorbic acid (SigmaAldrich, 50-81-7), 0.45mM MTG and 0.18mg/ml human transferrin (Roche, 10652202001). After 5-day incubation, EBs were collected, dissociated, passed through a cell strainer and counted. Hemangioblast selection was performed using a biotynylated anti Flk-1 antibody (eBioscience 13-5821-85), at a concentration of 5uL per 10mln of cells. Depending on the sorting strategy, the stained cells were treated in two different ways. For FACS, cells were stained with a secondary PE-streptavidin antibody (BD Biosciences, 554061) at a dilution 1:400, washed and resuspended in media with DAPI. In the case of MACS, Flk-1-stained cells were incubated with anti-biotin microbeads anti-biotin

microbeads (Miltenyi Biotec: 130-090-485), 20ul per 10mln cells, washed and separated on a MACS column.

Sorted hemangioblasts were plated on gelatin-coated plates in blast media at a density of 0.05-0.1mln cells per ml (for 6-well plates, 2ml of cell suspension were plates in each well, meaning 0.1-0.2mln cells. For bigger plates, the amounts were scaled up proportionally to the surface). The blast media consisted of: IMDM, 10% FCS, 1% Glutamax, 15% D4T medium (gift from Anna Bigas lab, IMIM), 25ug/ml ascorbic acid, 0.45mM MTG, 0.18mg/ml human transferrin, 5ng/ml VEGF (Peprotech: 450-32), and IL-6 (Peprotech: 216–16). After 4 days of culture, total blast samples were collected, and cell populations were determined by staining with the antibodies against cKIT, Tie2 and CD41.

RNA extraction, RT-PCR and qPCR

RNA was extracted using the RNeasy Mini kit (74106, Qiagen), following the manufacturer's protocol. Reverse transcription was performed using the qScript cDNA Synthesis Kit (Quantabio, 95047-025).

RT-PCR was performed using corresponding primers and 2xPCR master mix (Promega). For qPCR, SYBR green (Roche) was used, and samples were run on a Lightcycler 480 (Roche). Normalization of gene expression was performed using either the Rplp0 or Gapdh housekeeping gene.

Table VI.3. List of primers for RT-PCR and qPCR (all 5´- 3´)

Gene	Forward primer	Reverse primer
Rplp0	TTCATTGTGGGAGCAGAC	CAGCAGTTTCTCCAGAGC
Gapdh	GTATGACTCCACTCACGGCAAA	TTCCCATTCTCGGCCTTG
Rex1	GAGACTGAGGAAGATGGCTTCC	CTGGCGAGAAAGGTTTTGCTCC
Fgf5	GTGTCTCAGGGGATTGTAGGAATACG	GTGAAGGAAAGTTCCGGTTGC
Brachyury	GAACCTCGGATTCACATCGT	TTCTTTGGCATCAAGGAAGG
Flk-1	CACCTGGCACTCTCCACCTTC	GATTTCATCCCACTACCGAAAG
Meis1	GCAGTTGGCACAAGATACAGGAC	ACTGCTCGGTTGGACTGGTCTA
Tal1	GCCAGCCGCTCGCCTCACTA	CCGCACTACTTTGGTGTGAGGA
Runx1	CCAGCAAGCTGAGGAGCGGCG	CGGATTTGTAAAGACGGTGA
Gata1	ACGACCACTACAACACTCTGGC	TTGCGGTTCCTCGTCTGGATTC
Gata2	CTTCAACCATCTCGACTCGCAG	GCAACAAGTGTGGTCGGCACAT
Spi1	GAGGTGTCTGATGGAGAAGCTG	ACCCACCAGATGCTGTCCTTCA
Hbb-bh1	GTTGGAGGAGAAACTCTGGGAAG	GCCCAAGGATGTCAGCACTTTC
Hoxa5	CGCAAGCTGCACATTAGTCACG	GAGAGGCAAAGGGCATGAGCTA
Hoxa9	GCCTTCTCCGAAAACAATGCCG	TTCCGAGTGGAGCGAGCATGTA
Evi1	CCTCCATGTTCAGCTTCCGAGC	AAGTGCCGTGTTAGGTTCGCAG

Dnmt3b	CGCACAACCAATGACTCTGCTG	GGTGACTTCAGAAGCCATCCGT
Mll	GTGTGGGAGATGGGAGGCTTA	GTTCACAACACACTTGGCAATACACA
Af9	CCCAATGACTCAGACGTGGA	CATGATCGACCTTAAAGGACCTTGTT
Mll-mCherry (RT-PCR)	GAGAAGCCACCTCCAGTAAGTAA	TGCTTGATCTCGCCCTTCAG
Mll-Af9 (RT-PCR)	GAGAAGCCACCTCCAGTAAGTAA	TCAGGATGTTCCAGATGTTTCCAG
Mll-mCherry	CGCCGAAGTCCATACAAATCAGGAT	GAACTCGTGGCCGTTCACGGA
Mll-Af9	CGCCGAAGTCCATACAAATCAGGAT	CATGATCGACCTTAAAGGACCTTGTT

Western blot

Protein extraction was performed by resuspending cell pellets in a SDS lysis buffer (25 mM Tris HCl pH 7.6, 1 % SDS, 1 mM EDTA, 1 mM EGTA, protease inhibitors ((**cOmplete EDTAfree tablet from** Roche)), and incubating at 95 degrees for 10 minutes. Reactions were centrifuged, and protein-containing supernatants were collected. Protein quantification was performed using the Pierce™ BCA Protein Assay Kit (ThermoFisher, 23225), according to the manufacturer's protocol.

For western blot, samples of 50ug total protein were run on a Mini-PROTEAN® Precast Gels (Bio-Rad) and subsequently transferred to polyvinylidene difluoride (PVDF) membranes (162-0177, Bio-Rad).Next, membranes were incubated with either anti-TY1 (ThermoFisher, MA5-23513) or anti-Gapdh antibody (Santa cruz,

SC-32233), and in both cases, with a secondary anti-mouse-HRP (Bethyl laboratories, A90-516P). For visualisation, Pierce ECL Western Blotting Substrate kit (32106, Thermo Fischer) was used, according to manufacturer's protocol, together with the IBright Imaging System (ThermoFisher).

BIBLIOGRAPHY

Amabile, G., Welner, R.S., Nombela-Arrieta, C., D'Alise, A.M., Di Ruscio, A., Ebralidze, A.K., Kraytsberg, Y., Ye, M., Kocher, O., Neuberg, D.S., et al. (2013). In vivo generation of transplantable human hematopoietic cells from induced pluripotent stem cells. Blood *121*, 1255–1264.

Antunes, E.T.B., and Ottersbach, K. (2020). The MLL/SET family and haematopoiesis. Biochim. Biophys. Acta BBA - Gene Regul. Mech. *1863*, 194579.

Aplan, P.D. (2006). Chromosomal translocations involving the MLL gene: Molecular mechanisms. DNA Repair *5*, 1265–1272.

Arndt, K., Kranz, A., Fohgrub, J., Jolly, A., Bledau, A.S., Di Virgilio, M., Lesche, M., Dahl, A., Höfer, T., Stewart, A.F., et al. (2018). SETD1A protects HSCs from activation-induced functional decline in vivo. Blood *131*, 1311–1324.

Artinger, E.L., Mishra, B.P., Zaffuto, K.M., Li, B.E., Chung, E.K.Y., Moore, A.W., Chen, Y., Cheng, C., and Ernst, P. (2013). An MLL-dependent network sustains hematopoiesis. Proc. Natl. Acad. Sci. *110*, 12000–12005.

Austenaa, L., Barozzi, I., Chronowska, A., Termanini, A., Ostuni, R., Prosperini, E., Stewart, A.F., Testa, G., and Natoli, G. (2012). The Histone Methyltransferase Wbp7 Controls Macrophage Function through GPI Glycolipid Anchor Synthesis. Immunity *36*, 572–585.

Batta, K., Florkowska, M., Kouskoff, V., and Lacaud, G. (2014). Direct Reprogramming of Murine Fibroblasts to Hematopoietic Progenitor Cells. Cell Rep. *9*, 1871–1884.

Bindels, E.M.J., Havermans, M., Lugthart, S., Erpelinck, C., Wocjtowicz, E., Krivtsov, A.V., Rombouts, E., Armstrong, S.A., Taskesen, E., Haanstra, J.R., et al. (2012). EVI1 is critical for the pathogenesis of a subset of MLL-AF9-rearranged AMLs. Blood *119*, 5838–5849.

Boisset, J.-C., van Cappellen, W., Andrieu-Soler, C., Galjart, N., Dzierzak, E., and Robin, C. (2010). In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. Nature *464*, 116–120.

Boros, K., Lacaud, G., and Kouskoff, V. (2011). The transcription factor Mxd4 controls the proliferation of the first blood precursors at the onset of hematopoietic development in vitro. Exp. Hematol. *39*, 1090–1100.

Boutin, J., Rosier, J., Cappellen, D., Prat, F., Toutain, J., Pennamen, P., Bouron, J., Rooryck, C., Merlio, J.P., Lamrissi-Garcia, I., et al. (2021). CRISPR-Cas9 globin editing can induce megabase-scale copy-neutral losses of heterozygosity in hematopoietic cells. Nat. Commun. *12*, 4922.

Bowles, K.M., Vallier, L., Smith, J.R., Alexander, M.R.J., and Pedersen, R.A. (2006). HOXB4 Overexpression Promotes Hematopoietic Development by Human Embryonic Stem Cells. Stem Cells *24*, 1359–1369.

Breese, E.H., Buechele, C., Dawson, C., Cleary, M.L., and Porteus, M.H. (2015). Use of Genome Engineering to Create Patient Specific MLL Translocations in Primary Human Hematopoietic Stem and Progenitor Cells. PLOS ONE *10*, e0136644.

Buechele, C., Breese, E.H., Schneidawind, D., Lin, C.-H., Jeong, J., Duque-Afonso, J., Wong, S.H.K., Smith, K.S., Negrin, R.S., Porteus, M., et al. (2015). MLL leukemia induction by genome editing of human CD34+ hematopoietic cells. Blood *126*, 1683–1694.

Bueno, C. (2012). A human ESC model for MLL-AF4 leukemic fusion gene reveals an impaired early hematopoietic-endothelial specification. Cell Res. *22*, 17.

Butler, J.S., and Dent, S.Y.R. (2013). The role of chromatin modifiers in normal and malignant hematopoiesis. Blood *121*, 3076–3084.

Canela, A., Maman, Y., Huang, S.N., Wutz, G., Tang, W., Zagnoli-Vieira, G., Callen, E., Wong, N., Day, A., Peters, J.-M., et al. (2019). Topoisomerase II-Induced Chromosome Breakage and Translocation Is Determined by Chromosome Architecture and Transcriptional Activity. Mol. Cell *75*, 252-266.e8.

Cannan, W.J., and Pederson, D.S. (2016). Mechanisms and Consequences of Double-Strand DNA Break Formation in Chromatin: DOUBLE-STRAND DNA BREAK FORMATION IN CHROMATIN. J. Cell. Physiol. *231*, 3–14.

Canu, G., and Ruhrberg, C. (2021). First blood: the endothelial origins of hematopoietic progenitors. Angiogenesis *24*, 199–211.

Chen, C., Liu, Y., Rappaport, A.R., Kitzing, T., Schultz, N., Zhao, Z., Shroff, A.S., Dickins, R.A., Vakoc, C.R., Bradner, J.E., et al. (2014). MLL3 Is a Haploinsufficient 7q Tumor Suppressor in Acute Myeloid Leukemia. Cancer Cell *25*, 652–665.

Chen, M.J., Yokomizo, T., Zeigler, B.M., Dzierzak, E., and Speck, N.A. (2009). Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. Nature *457*, 887–891.

Chen, W., Kumar, A.R., Hudson, W.A., Li, Q., Wu, B., Staggs, R.A., Lund, E.A., Sam, T.N., and Kersey, J.H. (2008). Malignant Transformation Initiated by Mll-AF9: Gene Dosage and Critical Target Cells. Cancer Cell *13*, 432–440.

Cheng, H., Ang, H.Y.-K., A. EL Farran, C., Li, P., Fang, H.T., Liu, T.M., Kong, S.L., Chin, M.L., Ling, W.Y., Lim, E.K.H., et al. (2016). Reprogramming mouse fibroblasts into engraftable myeloerythroid and lymphoid progenitors. Nat. Commun. *7*, 13396.

Choi, K., Kennedy, M., Kazarov, A., Papadimitriou, J.C., and Keller, G. A common precursor for hematopoietic and endothelial cells. 8.

Conway, E.M., and Bracken, A.P. (2017). Unraveling the Roles of Canonical and Noncanonical PRC1 Complexes. In Polycomb Group Proteins, (Elsevier), pp. 57–80.

Copley, M.R., and Eaves, C.J. (2013). Developmental changes in hematopoietic stem cell properties. Exp. Mol. Med. *45*, e55–e55.

Corral, J., Lavenir, I., Impey, H., Warren, A.J., Forster, A., Larson, T.A., Bell, S., McKenzie, A.N.J., King, G., and Rabbitts, T.H. (1996). 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.

Costa, G., Mazan, A., Gandillet, A., Pearson, S., Lacaud, G., and Kouskoff, V. (2012). SOX7 regulates the expression of VE-cadherin in the haemogenic endothelium at the onset of haematopoietic development. Development *139*, 1587–1598.

Crans, H., and Sakamoto, K. (2001). Transcription factors and translocations in lymphoid and myeloid leukemia. Leukemia *15*, 313–331.

Cuevas, I., Layman, H., Coussens, L., and Boudreau, N. (2015). Sustained Endothelial Expression of HoxA5 In Vivo Impairs Pathological Angiogenesis And Tumor Progression. PLOS ONE *10*, e0121720.

Daniel, M.G., Pereira, C.-F., Lemischka, I.R., and Moore, K.A. (2016). Making a Hematopoietic Stem Cell. Trends Cell Biol. *26*, 202–214.

Di Carlo, V., Mocavini, I., and Di Croce, L. (2018). Polycomb complexes in normal and malignant hematopoiesis. J. Cell Biol. jcb.201808028.

Ding, X., Lin, Q., Ensenat-Waser, R., Rose-John, S., and Zenke, M. (2012). Polycomb Group Protein Bmi1 Promotes Hematopoietic Cell Development from Embryonic Stem Cells. Stem Cells Dev. *21*, 121–132.

Dobson, C.L. (1999). The Mll-AF9 gene fusion in mice controls myeloproliferation and specifies acute myeloid leukaemogenesis. EMBO J. *18*, 3564–3574.

Doetschman, T.C., Eistetter, H., Katz, M., Schmidt, W., and Kemler, R. The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. 22.

Doğan, E.S., and Liu, C. (2018). Three-dimensional chromatin packing and positioning of plant genomes. Nat. Plants *4*, 521–529.

Doulatov, S., Vo, L.T., Chou, S.S., Kim, P.G., Arora, N., Li, H., Hadland, B.K., Bernstein, I.D., Collins, J.J., Zon, L.I., et al. (2013). Induction of Multipotential Hematopoietic Progenitors from Human Pluripotent Stem Cells via Respecification of Lineage-Restricted Precursors. Cell Stem Cell *13*, 459–470.

Drake, C.J., and Fleming, P.A. (2000). Vasculogenesis in the day 6.5 to 9.5 mouse embryo. Blood *95*, 1671–1679.

Drynan, L.F., Pannell, R., Forster, A., Chan, N.M.M., Cano, F., Daser, A., and Rabbitts, T.H. (2005). Mll fusions generated by Cre-loxP-mediated de novo translocations can induce lineage reassignment in tumorigenesis. EMBO J. *24*, 3136–3146.

Dzierzak, E. (1999). Embryonic Beginnings of Definitive Hematopoietic Stem Cells. Ann. N. Y. Acad. Sci. *872*, 256–264.

Dzierzak, E., and Bigas, A. (2018). Blood Development: Hematopoietic Stem Cell Dependence and Independence. Cell Stem Cell *22*, 639–651.

Eguchi, M., Eguchi-Ishimae, M., and Greaves, M. (2004). The small oligomerization domain of gephyrin converts MLL to an oncogene. Blood *103*, 3876–3882.

Elcheva, I., Brok-Volchanskaya, V., Kumar, A., Liu, P., Lee, J.-H., Tong, L., Vodyanik, M., Swanson, S., Stewart, R., Kyba, M., et al. (2014). Direct induction of haematoendothelial programs in human pluripotent stem cells by transcriptional regulators. Nat. Commun. *5*, 4372.

Ema, M., Yokomizo, T., Wakamatsu, A., Terunuma, T., Yamamoto, M., and Takahashi, S. (2006). Primitive erythropoiesis from mesodermal precursors expressing VE-cadherin, PECAM-1, Tie2, endoglin, and CD34 in the mouse embryo. Blood *108*, 4018–4024.

Ernst, P., Mabon, M., Davidson, A.J., Zon, L.I., and Korsmeyer, S.J. An Mll-Dependent Hox Program Drives Hematopoietic Progenitor Expansion. Curr. Biol. 7.

Fleming, W., Alpern, E., Uchida, N., Ikuta, K., Spangrude, G., and Weissman, I. (1993). Functional heterogeneity is associated with the cell cycle status of murine hematopoietic stem cells. J. Cell Biol. *122*, 897–902.

Ford, A.M., Ridge, S.A., Cabrera, M.E., Mahmoud, H., Steel, C.M., Chan, L.C., and Greaves, M. (1993). In utero rearrangements in the trithorax-related oncogene in infant leukaemias. Nature *363*, 358–360.

Fudenberg, G., Getz, G., Meyerson, M., and Mirny, L.A. (2011). High order chromatin architecture shapes the landscape of

chromosomal alterations in cancer. Nat. Biotechnol. *29*, 1109–1113.

Gale, K.B., Ford, A.M., Repp, R., Borkhardt, A., Keller, C., Eden, O.B., and Greaves, M.F. (1997). Backtracking leukemia to birth: Identification of clonotypic gene fusion sequences in neonatal blood spots. Proc. Natl. Acad. Sci. *94*, 13950–13954.

Gandillet, A., Serrano, A.G., Pearson, S., Lie-A-Ling, M., Lacaud, G., and Kouskoff, V. (2009). Sox7-sustained expression alters the balance between proliferation and differentiation of hematopoietic progenitors at the onset of blood specification. Blood *114*, 4813–4822.

Ganuza, M., Chabot, A., Tang, X., Bi, W., Natarajan, S., Carter, R., Gawad, C., Kang, G., Cheng, Y., and McKinney-Freeman, S. (2018). Murine hematopoietic stem cell activity is derived from pre-circulation embryos but not yolk sacs. Nat. Commun. *9*, 5405.

Gao, L., Tober, J., Gao, P., Chen, C., Tan, K., and Speck, N.A. (2018). RUNX1 and the endothelial origin of blood. Exp. Hematol. *68*, 2–9.

Gilmour, J., Assi, S.A., Jaegle, U., Kulu, D., van de Werken, H., Clarke, D., Westhead, D.R., Philipsen, S., and Bonifer, C. (2014). A crucial role for the ubiquitously expressed transcription factor Sp1 at early stages of hematopoietic specification. Development *141*, 2391–2401.

Glaser, S., Schaft, J., Lubitz, S., Vintersten, K., van der Hoeven, F., Tufteland, K.R., Aasland, R., Anastassiadis, K., Ang, S.-L., and Stewart, A.F. (2006). Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. Development *133*, 1423–1432.

Glozak, M.A., and Seto, E. (2007). Histone deacetylases and cancer. Oncogene *26*, 5420–5432.

Goode, D.K., Obier, N., Vijayabaskar, M.S., Lie-A-Ling, M., Lilly, A.J., Hannah, R., Lichtinger, M., Batta, K., Florkowska, M., Patel, R., et al. (2016). Dynamic Gene Regulatory Networks Drive Hematopoietic Specification and Differentiation. Dev. Cell *36*, 572–587.

Grawunder, U. (1997). A complex of RAG-1 and RAG-2 proteins persists on DNA after single- strand cleavage at V(D)J recombination signal sequences. Nucleic Acids Res. *25*, 1375–1382.

Gué, M., Sun, J.-S., and Boudier, T. (2006). Simultaneous localization of MLL, AF4 and ENL genes in interphase nuclei by 3D-FISH: MLL translocation revisited. BMC Cancer *6*, 20.

Hess, J.L. (2004). MLL: a histone methyltransferase disrupted in leukemia. Trends Mol. Med. *10*, 500–507.

Hidalgo, I., Herrera-Merchan, A., Ligos, J.M., Carramolino, L., Nuñez, J., Martinez, F., Dominguez, O., Torres, M., and Gonzalez, S. (2012). Ezh1 Is Required for Hematopoietic Stem Cell Maintenance and Prevents Senescence-like Cell Cycle Arrest. Cell Stem Cell *11*, 649–662.

Hu, D., and Shilatifard, A. (2016). Epigenetics of hematopoiesis and hematological malignancies. Genes Dev. *30*, 2021–2041.

Huber, T.L., Kouskoff, V., Joerg Fehling, H., Palis, J., and Keller, G. (2004). Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. Nature *432*, 625–630.

Hwang, S.M. (2020). Classification of acute myeloid leukemia. BLOOD Res. *55*, S1–S4.

Ikonomi, N., Kühlwein, S.D., Schwab, J.D., and Kestler, H.A. (2020). Awakening the HSC: Dynamic Modeling of HSC Maintenance Unravels Regulation of the TP53 Pathway and Quiescence. Front. Physiol. *11*, 848.

Jeffs, A.R., Wells, E., and Morris, C.M. (2001). Nonrandom distribution of interspersed repeat elements in theBCR andABL1 genes and its relation to breakpoint cluster regions. Genes. Chromosomes Cancer *32*, 144–154.

Jeggo, P.A., and Löbrich, M. (2007). DNA double-strand breaks: their cellular and clinical impact? Oncogene *26*, 7717–7719.

Jeong, J., Jager, A., Domizi, P., Pavel-Dinu, M., Gojenola, L., Iwasaki, M., Wei, M.C., Pan, F., Zehnder, J.L., Porteus, M.H., et al. (2019). High-efficiency CRISPR induction of t(9;11) chromosomal

translocations and acute leukemias in human blood stem cells. Blood Adv. *3*, 2825–2835.

Johnson, J.J., Chen, W., Hudson, W., Yao, Q., Taylor, M., Rabbitts, T.H., and Kersey, J.H. (2003). Prenatal and postnatal myeloid cells demonstrate stepwise progression in the pathogenesis of MLL fusion gene leukemia. Blood *101*, 3229–3235.

Jude, C.D., Climer, L., Xu, D., Artinger, E., Fisher, J.K., and Ernst, P. (2007). Unique and Independent Roles for MLL in Adult Hematopoietic Stem Cells and Progenitors. Cell Stem Cell *1*, 324–337.

Katsumoto, T., Aikawa, Y., Iwama, A., Ueda, S., Ichikawa, H., Ochiya, T., and Kitabayashi, I. (2006). MOZ is essential for maintenance of hematopoietic stem cells. Genes Dev. *20*, 1321–1330.

Keller, G., Lacaud, G., and Robertson, S. (1999). Development of the hematopoietic system in the mouse. Exp. Hematol. *27*, 777–787.

Kennedy, M., Firpo, M., Choi, K., Wall, C., Robertson, S., Kabrun, N., and Keller, G. (1997). A common precursor for primitive erythropoiesis and definitive haematopoiesis. Nature *386*, 488–493.

Kim, I., Saunders, T.L., and Morrison, S.J. (2007). Sox17 Dependence Distinguishes the Transcriptional Regulation of Fetal from Adult Hematopoietic Stem Cells. Cell *130*, 470–483.

Ko, M., Bandukwala, H.S., An, J., Lamperti, E.D., Thompson, E.C., Hastie, R., Tsangaratou, A., Rajewsky, K., Koralov, S.B., and Rao, A. (2011). Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. Proc. Natl. Acad. Sci. *108*, 14566–14571.

Kotani, S., Yoda, A., Kon, A., Kataoka, K., Ochi, Y., Shiozawa, Y., Hirsch, C., Takeda, J., Ueno, H., Yoshizato, T., et al. (2019). Molecular pathogenesis of disease progression in MLL-rearranged AML. Leukemia *33*, 612–624.

Krivtsov, A.V., Twomey, D., Feng, Z., Stubbs, M.C., Wang, Y., Faber, J., Levine, J.E., Wang, J., Hahn, W.C., Gilliland, D.G., et al. (2006). Transformation from committed progenitor to leukaemia

stem cell initiated by MLL–AF9. Nature *442*, 818–822.

Kyba, M., Perlingeiro, R.C.R., and Daley, G.Q. (2002). HoxB4 Confers Definitive Lymphoid-Myeloid Engraftment Potential on Embryonic Stem Cell and Yolk Sac Hematopoietic Progenitors. Cell *109*, 29–37.

Lacaud, G., and Kouskoff, V. (2017). Hemangioblast, hemogenic endothelium, and primitive versus definitive hematopoiesis. Exp. Hematol. *49*, 19–24.

Lancrin, C., Sroczynska, P., Stephenson, C., Allen, T., Kouskoff, V., and Lacaud, G. (2009). The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage. Nature *457*, 892–895.

Lancrin, C., Sroczynska, P., Serrano, A.G., Gandillet, A., Ferreras, C., Kouskoff, V., and Lacaud, G. (2010). Blood cell generation from the hemangioblast. J. Mol. Med. *88*, 167–172.

Laurenti, E., and Göttgens, B. (2018). From haematopoietic stem cells to complex differentiation landscapes. Nature *553*, 418–426.

Leach, B.I., Kuntimaddi, A., Schmidt, C.R., Cierpicki, T., Johnson, S.A., and Bushweller, J.H. (2013). Leukemia Fusion Target AF9 Is an Intrinsically Disordered Transcriptional Regulator that Recruits Multiple Partners via Coupled Folding and Binding. Structure *21*, 176–183.

Lee, J.-E., Wang, C., Xu, S., Cho, Y.-W., Wang, L., Feng, X., Baldridge, A., Sartorelli, V., Zhuang, L., Peng, W., et al. (2013). H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. ELife *2*, e01503.

Li, Y., Schulz, V.P., Deng, C., Li, G., Shen, Y., Tusi, B.K., Ma, G., Stees, J., Qiu, Y., Steiner, L.A., et al. (2016). Setd1a and NURF mediate chromatin dynamics and gene regulation during erythroid lineage commitment and differentiation. Nucleic Acids Res. gkw327.

Li, Z., Cai, X., Cai, C.-L., Wang, J., Zhang, W., Petersen, B.E., Yang, F.-C., and Xu, M. (2011). Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. Blood *118*, 4509–4518.

Li, Z., Lan, Y., He, W., Chen, D., Wang, J., Zhou, F., Wang, Y., Sun, H., Chen, X., Xu, C., et al. (2012). Mouse Embryonic Head as a Site for Hematopoietic Stem Cell Development. Cell Stem Cell *11*, 663–675.

Lim, W.F., Inoue-Yokoo, T., Tan, K.S., Lai, M.I., and Sugiyama, D. (2013). Hematopoietic cell differentiation from embryonic and induced pluripotent stem cells. Stem Cell Res. Ther. *4*, 71.

Lis, R., Karrasch, C.C., Poulos, M.G., Kunar, B., Redmond, D., Duran, J.G.B., Badwe, C.R., Schachterle, W., Ginsberg, M., Xiang, J., et al. (2017). Conversion of adult endothelium to immunocompetent haematopoietic stem cells. Nature *545*, 439–445.

Lupiáñez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R., et al. (2015). Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions. Cell *161*, 1012–1025.

Maddalo, D., and Ventura, A. (2016). Somatic Engineering of Oncogenic Chromosomal Rearrangements: A Perspective. Cancer Res. *76*, 4918–4923.

Meaburn, K.J., Misteli, T., and Soutoglou, E. (2007). Spatial genome organization in the formation of chromosomal translocations. Semin. Cancer Biol. *17*, 80–90.

Medeiros, L.J., and Konoplev, S.N. (2017). Molecular Biology of Leukemias. In The Molecular Basis of Human Cancer, W.B. Coleman, and G.J. Tsongalis, eds. (New York, NY: Springer New York), pp. 505–559.

Medvinsky, A., Rybtsov, S., and Taoudi, S. (2011). Embryonic origin of the adult hematopoietic system: advances and questions. Development *138*, 1017–1031.

Mehta, A., and Haber, J.E. (2014). Sources of DNA Double-Strand Breaks and Models of Recombinational DNA Repair. Cold Spring Harb. Perspect. Biol. *6*, a016428.

Mehta, R.S., Rezvani, K., Olson, A., Oran, B., Hosing, C., Shah, N., Parmar, S., Armitage, S., and Shpall, E.J. (2015). Novel Techniques

for Ex Vivo Expansion of Cord Blood: Clinical Trials. Front. Med. *2*.

Meyer, C., Schneider, B., Jakob, S., Strehl, S., Attarbaschi, A., Schnittger, S., Schoch, C., Jansen, M., van Dongen, J.J.M., and Den Boer, M.L. (2006). The MLL recombinome of acute leukemias. Leukemia *20*, 777–784.

Meyer, C., Burmeister, T., Gröger, D., Tsaur, G., Fechina, L., Renneville, A., Sutton, R., Venn, N.C., Emerenciano, M., Pombo-de-Oliveira, M.S., et al. (2018). The MLL recombinome of acute leukemias in 2017. Leukemia *32*, 273–284.

Mochizuki-Kashio, M., Mishima, Y., Miyagi, S., Negishi, M., Saraya, A., Konuma, T., Shinga, J., Koseki, H., and Iwama, A. (2011). Dependency on the polycomb gene Ezh2 distinguishes fetal from adult hematopoietic stem cells. Blood *118*, 6553–6561.

Moignard, V., Woodhouse, S., Fisher, J., and Göttgens, B. (2013). Transcriptional hierarchies regulating early blood cell development. Blood Cells. Mol. Dis. *51*, 239–247.

Muntean, A.G., Giannola, D., Udager, A.M., and Hess, J.L. (2008). The PHD fingers of MLL block MLL fusion protein–mediated transformation. Blood *112*, 4690–4693.

Nambiar, M., and Raghavan, S.C. (2011). How does DNA break during chromosomal translocations? Nucleic Acids Res. *39*, 5813–5825.

Nitiss, J.L. (2009). DNA topoisomerase II and its growing repertoire of biological functions. Nat. Rev. Cancer *9*, 327–337.

Oatley, M., Bölükbası, Ö.V., Svensson, V., Shvartsman, M., Ganter, K., Zirngibl, K., Pavlovich, P.V., Milchevskaya, V., Foteva, V., Natarajan, K.N., et al. (2020). Single-cell transcriptomics identifies CD44 as a marker and regulator of endothelial to haematopoietic transition. Nat. Commun. *11*, 586.

Ottersbach, K., Sanjuan-Pla, A., Torres-Ruíz, R., Bueno, C., Velasco-Hernández, T., and Menendez, P. (2018). The "Never-Ending" Mouse Models for MLL-Rearranged Acute Leukemia Are Still Teaching Us. HemaSphere *2*, e57.

Passegué, E., Wagers, A.J., Giuriato, S., Anderson, W.C., and Weissman, I.L. (2005). Global analysis of proliferation and cell cycle gene expression in the regulation of hematopoietic stem and progenitor cell fates. J. Exp. Med. *202*, 1599–1611.

Pearson, S., Cuvertino, S., Fleury, M., Lacaud, G., and Kouskoff, V. (2015). In Vivo Repopulating Activity Emerges at the Onset of Hematopoietic Specification during Embryonic Stem Cell Differentiation. Stem Cell Rep. *4*, 431–444.

Pereira, C.-F., Chang, B., Qiu, J., Niu, X., Papatsenko, D., Hendry, C.E., Clark, N.R., Nomura-Kitabayashi, A., Kovacic, J.C., Ma'ayan, A., et al. (2013). Induction of a Hemogenic Program in Mouse Fibroblasts. Cell Stem Cell *13*, 205–218.

Pina, C., May, G., Soneji, S., Hong, D., and Enver, T. (2008). MLLT3 Regulates Early Human Erythroid and Megakaryocytic Cell Fate. Cell Stem Cell *2*, 264–273.

Piunti, A., and Shilatifard, A. (2016). Epigenetic balance of gene expression by Polycomb and COMPASS families. Science *352*, aad9780.

Prange, K.H.M. (2017). MLL-AF9 and MLL-AF4 oncofusion proteins bind a distinct enhancer repertoire and target the RUNX1 program in 11q23 acute myeloid leukemia. 12.

Pulecio, J., Nivet, E., Sancho-Martinez, I., Vitaloni, M., Guenechea, G., Xia, Y., Kurian, L., Dubova, I., Bueren, J., Laricchia-Robbio, L., et al. (2014). Conversion of Human Fibroblasts Into Monocyte-Like Progenitor Cells. STEM CELLS *32*, 2923–2938.

Ramsden, D.A., and Nussenzweig, A. (2021). Mechanisms driving chromosomal translocations: lost in time and space. Oncogene *40*, 4263–4270.

Rao, R.C., and Dou, Y. (2015). Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. Nat. Rev. Cancer *15*, 334–346.

Regha, K., Assi, S.A., Tsoulaki, O., Gilmour, J., Lacaud, G., and Bonifer, C. (2015). Developmental-stage-dependent transcriptional response to leukaemic oncogene expression. Nat. Commun. *6*.

Riddell, J., Gazit, R., Garrison, B.S., Guo, G., Saadatpour, A., Mandal, P.K., Ebina, W., Volchkov, P., Yuan, G.-C., Orkin, S.H., et al. (2014). Reprogramming Committed Murine Blood Cells to Induced Hematopoietic Stem Cells with Defined Factors. Cell *157*, 549–564.

Rizo, A., Dontje, B., Vellenga, E., de Haan, G., and Schuringa, J.J. (2008). Long-term maintenance of human hematopoietic stem/progenitor cells by expression of BMI1. Blood *111*, 2621–2630.

Robbiani, D.F., Bothmer, A., Callen, E., Reina-San-Martin, B., Dorsett, Y., Difilippantonio, S., Bolland, D.J., Chen, H.T., Corcoran, A.E., Nussenzweig, A., et al. (2008). AID Is Required for the Chromosomal Breaks in c-myc that Lead to c-myc/IgH Translocations. Cell *135*, 1028–1038.

Rodríguez-Paredes, M., and Esteller, M. (2011). Cancer epigenetics reaches mainstream oncology. Nat. Med. *17*, 330–339.

Roukos, V., and Misteli, T. (2014). The biogenesis of chromosome translocations. Nat. Cell Biol. *16*, 293–300.

Roukos, V., Voss, T.C., Schmidt, C.K., Lee, S., Wangsa, D., and Misteli, T. (2013). Spatial dynamics of chromosome translocations in living cells. Science *341*, 660–664.

Sainz de Aja, J., Menchero, S., Rollan, I., Barral, A., Tiana, M., Jawaid, W., Cossio, I., Alvarez, A., Carreño-Tarragona, G., Badia-Careaga, C., et al. (2019). The pluripotency factor NANOG controls primitive hematopoiesis and directly regulates *Tal1*. EMBO J. *38*.

Sandler, V.M., Lis, R., Liu, Y., Kedem, A., James, D., Elemento, O., Butler, J.M., Scandura, J.M., and Rafii, S. (2014). Reprogramming human endothelial cells to haematopoietic cells requires vascular induction. Nature *511*, 312–318.

Sarrou, E., Richmond, L., Carmody, R.J., Gibson, B., and Keeshan, K. (2020). CRISPR Gene Editing of Murine Blood Stem and Progenitor Cells Induces MLL-AF9 Chromosomal Translocation and MLL-AF9 Leukaemogenesis. Int. J. Mol. Sci. *21*, 4266.

Scharf, S., Zech, J., Bursen, A., Schraets, D., Oliver, P.L., Kliem, S., Pfitzner, E., Gillert, E., Dingermann, T., and Marschalek, R.

(2007). Transcription linked to recombination: a gene-internal promoter coincides with the recombination hot spot II of the human MLL gene. Oncogene *26*, 1361–1371.

Schneidawind, C., Jeong, J., Schneidawind, D., Kim, I.-S., Duque-Afonso, J., Wong, S.H.K., Iwasaki, M., Breese, E.H., Zehnder, J.L., Porteus, M., et al. (2018). MLL leukemia induction by t(9;11) chromosomal translocation in human hematopoietic stem cells using genome editing. Blood Adv. *2*, 832–845.

Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., and Cavalli, G. (2007). Genome Regulation by Polycomb and Trithorax Proteins. Cell *128*, 735–745.

Schuettengruber, B., Bourbon, H.-M., Di Croce, L., and Cavalli, G. (2017). Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. Cell *171*, 34–57.

Schwaller, J. (2020). Learning from mouse models of MLL fusion gene-driven acute leukemia. Biochim. Biophys. Acta BBA - Gene Regul. Mech. *1863*, 194550.

Scott, C.L., Gil, J., Hernando, E., Teruya-Feldstein, J., Narita, M., Martinez, D., Visakorpi, T., Mu, D., Cordon-Cardo, C., Peters, G., et al. (2007). Role of the chromobox protein CBX7 in lymphomagenesis. Proc. Natl. Acad. Sci. *104*, 5389–5394.

Seita, J., and Weissman, I.L. (2010). Hematopoietic stem cell: self-renewal versus differentiation: Hematopoietic stem cell. Wiley Interdiscip. Rev. Syst. Biol. Med. *2*, 640–653.

Serrano, A.G., Gandillet, A., Pearson, S., Lacaud, G., and Kouskoff, V. (2010). Contrasting effects of Sox17- and Sox18-sustained expression at the onset of blood specification. Blood *115*, 3895–3898.

Shalaby, F., Rossant, J., Yamaguchi, T.P., Gertsenstein, M., Wu, X.-F., Breitman, M.L., and Schuh, A.C. (1995). Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice. Nature *376*, 62–66.

Sprüssel, A., Schulte, J.H., Weber, S., Necke, M., Händschke, K., Thor, T., Pajtler, K.W., Schramm, A., König, K., Diehl, L., et al.

(2012). Lysine-specific demethylase 1 restricts hematopoietic progenitor proliferation and is essential for terminal differentiation. Leukemia *26*, 2039–2051.

Stavropoulou, V., Kaspar, S., Brault, L., Sanders, M.A., Juge, S., Morettini, S., Tzankov, A., Iacovino, M., Lau, I.-J., Milne, T.A., et al. (2016). MLL-AF9 Expression in Hematopoietic Stem Cells Drives a Highly Invasive AML Expressing EMT-Related Genes Linked to Poor Outcome. Cancer Cell *30*, 43–58.

Stefanska, M., Batta, K., Patel, R., Florkowska, M., Kouskoff, V., and Lacaud, G. (2017). Primitive erythrocytes are generated from hemogenic endothelial cells. Sci. Rep. *7*.

Sturgeon, C.M., Ditadi, A., Clarke, R.L., and Keller, G. (2013). Defining the path to hematopoietic stem cells. Nat. Biotechnol. *31*, 416–418.

Sugimura, R., Jha, D.K., Han, A., Soria-Valles, C., da Rocha, E.L., Lu, Y.-F., Goettel, J.A., Serrao, E., Rowe, R.G., Malleshaiah, M., et al. (2017). Haematopoietic stem and progenitor cells from human pluripotent stem cells. Nature *545*, 432–438.

Sugiyama, D., Inoue-Yokoo, T., Fraser, S.T., Kulkeaw, K., Mizuochi, C., and Horio, Y. (2011). Embryonic Regulation of the Mouse Hematopoietic Niche. Sci. World J. *11*, 1770–1780.

Suzuki, N., Yamazaki, S., Yamaguchi, T., Okabe, M., Masaki, H., Takaki, S., Otsu, M., and Nakauchi, H. (2013). Generation of Engraftable Hematopoietic Stem Cells From Induced Pluripotent Stem Cells by Way of Teratoma Formation. Mol. Ther. *21*, 1424–1431.

Szabo, E., Rampalli, S., Risueño, R.M., Schnerch, A., Mitchell, R., Fiebig-Comyn, A., Levadoux-Martin, M., and Bhatia, M. (2010). Direct conversion of human fibroblasts to multilineage blood progenitors. Nature *468*, 521–526.

Takizawa, H., Schanz, U., and Manz, M. (2011). Ex vivo expansion of hematopoietic stem cells: mission accomplished? Swiss Med. Wkly.

Tan, J., Jones, M., Koseki, H., Nakayama, M., Muntean, A.G.,

Maillard, I., and Hess, J.L. (2011). CBX8, a Polycomb Group Protein, Is Essential for MLL-AF9-Induced Leukemogenesis. Cancer Cell *20*, 563–575.

Thambyrajah, R., Fadlullah, M.Z.H., Proffitt, M., Patel, R., Cowley, S.M., Kouskoff, V., and Lacaud, G. (2018). HDAC1 and HDAC2 Modulate TGF-β Signaling during Endothelial-to-Hematopoietic Transition. Stem Cell Rep. *10*, 1369–1383.

Thiel, A.T., Blessington, P., Zou, T., Feather, D., Wu, X., Yan, J., Zhang, H., Liu, Z., Ernst, P., Koretzky, G.A., et al. (2010). MLL-AF9-Induced Leukemogenesis Requires Coexpression of the Wild-Type Mll Allele. Cancer Cell *17*, 148–159.

Vereide, D.T., Vickerman, V., Swanson, S.A., Chu, L.-F., McIntosh, B.E., and Thomson, J.A. (2014). An Expandable, Inducible Hemangioblast State Regulated by Fibroblast Growth Factor. Stem Cell Rep. *3*, 1043–1057.

Vidal, M., and Starowicz, K. (2017). Polycomb complexes PRC1 and their function in hematopoiesis. Exp. Hematol. *48*, 12–31.

Villa, R., Pasini, D., Gutierrez, A., Morey, L., Occhionorelli, M., Viré, E., Nomdedeu, J.F., Jenuwein, T., Pelicci, P.G., Minucci, S., et al. (2007). Role of the Polycomb Repressive Complex 2 in Acute Promyelocytic Leukemia. Cancer Cell *11*, 513–525.

Vizán, P., Gutiérrez, A., Espejo, I., García-Montolio, M., Lange, M., Carretero, A., Lafzi, A., de Andrés-Aguayo, L., Blanco, E., Thambyrajah, R., et al. (2020). The Polycomb-associated factor PHF19 controls hematopoietic stem cell state and differentiation. Sci. Adv. *6*, eabb2745.

Wada, T., Kikuchi, J., Nishimura, N., Shimizu, R., Kitamura, T., and Furukawa, Y. (2009). Expression Levels of Histone Deacetylases Determine the Cell Fate of Hematopoietic Progenitors. J. Biol. Chem. *284*, 30673–30683.

Wang, L.D., and Wagers, A.J. (2011). Dynamic niches in the origination and differentiation of haematopoietic stem cells. Nat. Rev. Mol. Cell Biol. *12*, 643–655.

Wang, L., Menendez, P., Shojaei, F., Li, L., Mazurier, F., Dick, J.E.,

Cerdan, C., Levac, K., and Bhatia, M. (2005a). Generation of hematopoietic repopulating cells from human embryonic stem cells independent of ectopic HOXB4 expression. J. Exp. Med. *201*, 1603–1614.

Wang, Y., Yates, F., Naveiras, O., Ernst, P., and Daley, G.Q. (2005b). Embryonic stem cell-derived hematopoietic stem cells. Proc. Natl. Acad. Sci. *102*, 19081–19086.

Wilson, A., Laurenti, E., and Trumpp, A. (2009). Balancing dormant and self-renewing hematopoietic stem cells. Curr. Opin. Genet. Dev. *19*, 461–468.

Xu, J., Li, L., Xiong, J., denDekker, A., Ye, A., Karatas, H., Liu, L., Wang, H., Qin, Z.S., Wang, S., et al. (2016). MLL1 and MLL1 fusion proteins have distinct functions in regulating leukemic transcription program. Cell Discov. *2*.

Yamane, T. (2018). Mouse Yolk Sac Hematopoiesis. Front. Cell Dev. Biol. *6*, 80.

Zhang, H., Zhang, Y., Zhou, X., Wright, S., Hyle, J., Zhao, L., An, J., Zhao, X., Shao, Y., Xu, B., et al. (2020). Functional interrogation of HOXA9 regulome in MLLr leukemia via reporter-based CRISPR/Cas9 screen. ELife *9*, e57858.

Zhang, J., Gao, X., and Yu, L. (2021). Roles of Histone Deacetylases in Acute Myeloid Leukemia With Fusion Proteins. Front. Oncol. *11*, 741746.

Zhang, L., Kang, X., Lu, J., Zhang, Y., Wu, X., Wu, G., Zheng, J., Tuladhar, R., Shi, H., Wang, Q., et al. (2019). Installation of a cancer promoting WNT/SIX1 signaling axis by the oncofusion protein MLL-AF9. EBioMedicine *39*, 145–158.

Zhang, Y., Gostissa, M., Hildebrand, D.G., Becker, M.S., Boboila, C., Chiarle, R., Lewis, S., and Alt, F.W. (2010). The Role of Mechanistic Factors in Promoting Chromosomal Translocations Found in Lymphoid and Other Cancers. In Advances in Immunology, (Elsevier), pp. 93–133.

Zheng, J. (2013). Oncogenic chromosomal translocations and human cancer (Review). Oncol. Rep. *30*, 2011–2019.

Zheng, Y., Zhang, H., Wang, Y., Li, X., Lu, P., Dong, F., Pang, Y., Ma, S., Cheng, H., Hao, S., et al. (2016). Loss of Dnmt3b accelerates MLL-AF9 leukemia progression. Leukemia *30*, 2373–2384.

Zhou, V.W., Goren, A., and Bernstein, B.E. (2011). Charting histone modifications and the functional organization of mammalian genomes. Nat. Rev. Genet. *12*, 7–18.

Zovein, A.C., Hofmann, J.J., Lynch, M., French, W.J., Turlo, K.A., Yang, Y., Becker, M.S., Zanetta, L., Dejana, E., Gasson, J.C., et al. (2008). Fate Tracing Reveals the Endothelial Origin of Hematopoietic Stem Cells. Cell Stem Cell *3*, 625–636.