

STEM CELL SURVIVAL
IN THE FACE OF GENOMIC INSTABILITY

DISSERTATION

ANJA GEISELHART

2015

DISSERTATION

SUBMITTED TO THE
COMBINED FACULTIES FOR THE NATURAL SCIENCES AND FOR MATHEMATICS OF THE
RUPERTO-CAROLA UNIVERSITY OF HEIDELBERG, GERMANY

FOR THE DEGREE OF
DOCTOR OF NATURAL SCIENCES

PRESENTED BY
ANJA GEISELHART, M.Sc.
BORN IN RIEDLINGEN, GERMANY

ORAL EXAMINATION: 20.05.2015

STEM CELL SURVIVAL
IN THE FACE OF GENOMIC INSTABILITY

REFEREES:

PD DR. KARIN MÜLLER-DECKER

DR. MICHAEL MILSOM

Table of contents

Summary.....	VI
Zusammenfassung	VIII
1 Introduction.....	1
1.1 Hematopoiesis and the hematopoietic system.....	1
1.1.1 Hematopoietic stem cells.....	1
1.1.1.1 <i>Functional hematopoietic stem cell assay</i>	2
1.1.1.2 <i>Purification of hematopoietic stem and progenitor cells from the bone marrow</i>	3
1.1.1.3 <i>Investigating HSC fate decisions</i>	4
1.2 The hematopoietic disorder Fanconi Anemia (FA).....	5
1.2.1 The FA signaling pathway mediates DNA damage repair	6
1.2.2 Defective DNA repair causes genomic instability in FA.....	7
1.2.3 Bone marrow failure arises from a defect in the FA HSC.....	8
1.2.4 FA knockout mouse models to study FA HSC biology.....	9
1.2.5 HSC transplantation and gene therapy for FA patients.....	10
1.3 Retroviral insertional mutagenesis	12
1.3.1 Insertional mutagenesis in gene therapy clinical trials.....	13
1.3.2 Insertional mutagenesis as a powerful screening method.....	14
1.3.3 Identification of retroviral integration sites and nearby deregulated genes.....	15
2 Aim of the thesis.....	17
3 Results	19
3.1 Retroviral insertional mutagenesis screen	19
3.2 Analysis of insertion sites in transduced FA and WT HSCs.....	20
3.3 Engraftment rate of transduced FA and WT HSCs	22
3.4 Clonal dominance of transduced FA and WT HSCs	26
3.5 Identification of retroviral integration sites in dominant FA and WT HSC clones..	29
3.6 Genomic location of the retroviral integrations in mice #A1, #A2 and #G1	30
3.7 Identification of genes that drive clonal dominance	32
3.8 Characterization of genes proximal to the integration site in mouse #A1	32
3.8.1 TATA box binding protein (TBP)-associated factor, RNA polymerase I, B (Taf1b).....	32
3.8.1.1 <i>Taf1b expression levels under homeostatic conditions</i>	33
3.8.1.2 <i>Epigenetic regulation of Taf1b expression</i>	35
3.8.1.3 <i>Taf1b expression under conditions of physiologic stress</i>	36
3.8.1.4 <i>Taf1b expression during development</i>	36

3.8.2	Grainyhead-like 1 (<i>Grhl1</i>).....	37
3.8.2.1	<i>Grhl1</i> expression levels under homeostatic conditions.....	38
3.8.2.2	Epigenetic regulation of <i>Grhl1</i> expression.....	39
3.8.2.3	<i>Grhl1</i> expression under conditions of physiologic stress.....	40
3.8.2.4	<i>Grhl1</i> expression during development.....	41
3.9	Identification and characterization of deregulated genes in mouse #G1.....	42
3.9.1	The ecotropic virus integration site 1 (<i>Evi1</i>).....	42
3.9.1.1	<i>Evi1</i> expression levels under homeostatic conditions.....	43
3.9.1.2	Epigenetic regulation of <i>Evi1</i> expression.....	45
3.9.1.3	<i>Evi1</i> expression under conditions of physiologic stress.....	46
3.9.1.4	<i>Evi1</i> expression during development.....	46
3.9.2	Gene expression profiling of bone marrow sample #G1.....	47
3.10	Identification and characterization of deregulated genes in mouse #A2.....	51
3.10.1	Oxidative stress-induced growth inhibitor 1 (<i>Osgin-1</i>).....	51
3.10.1.1	<i>Osgin1</i> expression levels under homeostatic conditions.....	52
3.10.1.2	Epigenetic regulation of <i>Osgin1</i> expression.....	53
3.10.1.3	<i>Osgin1</i> expression under conditions of physiologic stress.....	55
3.10.1.4	<i>Osgin1</i> expression during development.....	55
3.10.2	Gene expression profiling of bone marrow sample #A2.....	56
3.11	Target gene validation studies.....	60
3.11.1	Generation of lentiviral vectors for target gene validation studies.....	60
3.11.1.1	Subcloning of <i>Osgin1</i>	62
3.11.1.2	Subcloning of <i>Taf1b</i>	63
3.11.2	Production of lentiviral <i>Osgin1</i> overexpression vectors.....	63
3.11.3	Lentiviral overexpression of <i>Osgin1</i> in LSK cells.....	66
3.11.4	Evaluation of <i>Osgin1</i> overexpression in a competitive transplantation setting.....	67
3.12	FA HSC growth defect <i>in vitro</i>.....	71
3.12.1	Outlook of the FA HSC <i>in vitro</i> growth defect.....	74
4	Discussion.....	75
4.1	Retroviral insertional mutagenesis screen in the FA HSC murine transplantation model.....	75
4.2	Dominant FA HSC clones are able to overcome the inherent FA HSC defect upon retroviral insertional mutagenesis.....	76
4.2.1	Identification of retroviral integration sites in the genome of dominant FA HSC clones.....	76
4.2.2	Dominant FA HSC clones allow the identification of deregulated genes in the vicinity of the proviral integration site.....	77
4.2.3	The dominant FA clones provide insight into disease-specific biology.....	78

4.3	Candidate target genes with the potential to drive clonal dominance in the face of genomic instability	78
4.3.1	Taf1b.....	79
4.3.1.1	<i>Relevance of Taf1b expression for normal and stress hematopoiesis as well as during development</i>	79
4.3.2	Grhl1.....	79
4.3.2.1	<i>Relevance of Grhl1 expression for normal and stress hematopoiesis as well as during development</i>	79
4.3.3	Evi1.....	80
4.3.3.1	<i>Relevance of Evi1 expression for normal and stress hematopoiesis as well as during development..</i>	80
4.3.3.2	<i>Evi1 as a driver of clonal dominance in the face of genome instability.....</i>	81
4.3.4	Osgin1.....	81
4.3.4.1	<i>Relevance of Osgin1 expression for normal and stress hematopoiesis as well as during development.....</i>	81
4.3.4.2	<i>Biological impact of Osgin1 overexpression in WT and FA HSCs during hematopoietic reconstitution.....</i>	82
4.3.4.3	<i>Osgin1 as a driver of clonal dominance in the face of genome instability.....</i>	83
4.3.4.4	<i>Future studies on Osgin1.....</i>	84
4.4	Development of an <i>in vitro</i> assay with the potential to assess the impact of candidate target genes on HSC biology	84
4.5	Summary and outlook.....	85
5	Material and methods	88
5.1	Animals and animal experiments	88
5.1.1	Mouse lines.....	88
5.1.2	Collection of murine primary material.....	89
5.1.2.1	<i>Peripheral blood sample collection.....</i>	89
5.1.2.2	<i>Removal of hind legs, vertebrae and sternum.....</i>	89
5.1.2.3	<i>Removal of mouse lymphoid organs.....</i>	89
5.1.2.4	<i>Collection of tail biopsies.....</i>	89
5.1.3	Transplantation of myeloablated mice.....	89
5.2	Analysis and conservation of murine primary material	90
5.2.1	Analysis of the peripheral blood.....	90
5.2.1.1	<i>Evaluation of peripheral blood cell numbers.....</i>	90
5.2.1.2	<i>Peripheral blood smear.....</i>	90
5.2.2	Purification and analysis of murine bone marrow cells.....	90
5.2.2.1	<i>Crushing of murine bone marrow cells.....</i>	90
5.2.2.2	<i>Flushing of murine bone marrow cells.....</i>	91
5.2.3	Analysis of lymphoid organs and isolation of spleenocytes.....	91
5.2.3.1	<i>Evaluation of spleen weight.....</i>	91

5.2.3.2	<i>Histology of thymus and spleen</i>	91
5.2.3.3	<i>Isolation of spleenocytes</i>	91
5.2.4	Freezing of peripheral blood and bone marrow cells.....	91
5.2.4.1	<i>Snap freezing of peripheral blood and bone marrow cells</i>	91
5.2.4.2	<i>Freezing of viable bone marrow cells</i>	92
5.3	Analysis of cell morphology using cytopins	92
5.4	Identification and isolation of stem cell populations	92
5.4.1	Isolation of low-density mononuclear cells (LDMNCs).....	92
5.4.2	Lineage depletion of murine bone marrow cells	93
5.4.3	Cell surface stainings.....	93
5.4.3.1	<i>Stainings for the identification of stem cell populations</i>	93
5.4.3.2	<i>Stainings for the identification of mature cells</i>	94
5.5	Flow cytometry and flow cytometric cell sorting	95
5.5.1	Flow cytometric analysis of cells.....	95
5.5.2	Flow cytometric cell sorting.....	95
5.5.3	Analysis of flow cytometric data.....	95
5.6	Cultivation and manipulation of cells	96
5.6.1	<i>In vitro</i> culture of sorted bone marrow cells	96
5.6.2	<i>In vitro</i> culture of adherent cell lines	96
5.6.3	Transient calcium phosphate transfection of cells for virus production.....	96
5.6.3.1	<i>Production of retroviral particles</i>	97
5.6.3.2	<i>Production of lentiviral particles</i>	97
5.6.4	Viral limiting dilution transduction of murine 3T3 cells	97
5.6.5	Viral transduction of LSK cells	98
5.6.5.1	<i>Transduction of LSK cells with retroviral vectors</i>	98
5.6.5.2	<i>Transduction of LSK cells with lentiviral vectors</i>	99
5.7	<i>In vitro</i> assays	99
5.7.1	<i>In vitro</i> proliferation of HSCs.....	99
5.7.2	<i>In vitro</i> cell fate tracking of LT-HSCs	99
5.8	Molecular cell biology methods	100
5.8.1	DNA and RNA isolation.....	100
5.8.1.1	<i>DNA isolation from primary cells</i>	100
5.8.1.2	<i>DNA isolation from mouse tail biopsies</i>	100
5.8.1.3	<i>DNA isolation from bacterial cultures</i>	100
5.8.1.4	<i>DNA isolation from agarose gels</i>	100
5.8.1.5	<i>RNA isolation from primary cells</i>	101
5.8.2	PCR.....	101
5.8.2.1	<i>Genotyping PCRs of FA knockout mice</i>	101

5.9	Retroviral integration site analysis	102
5.9.1	LM-PCR.....	102
5.9.2	nrLAM-PCR.....	104
5.10	Agarose gel electrophoresis	105
5.11	Molecular cloning procedures.....	106
5.11.1	Cloning PCRs	106
5.11.2	TOPO TA cloning	106
5.11.3	Restriction digest	107
5.11.4	Assembly of oligonucleotides	107
5.11.5	Dephosphorylation of linearized vector DNA.....	108
5.11.6	Ligation	108
5.11.7	Sequencing of DNA fragments	108
5.11.8	Transformation of competent E.coli bacteria.....	108
5.11.9	Bacterial mini and maxi cultures.....	109
5.12	mRNA expression analysis.....	109
5.12.1	Quantitative real-time PCR.....	109
5.12.2	Gene expression profiling (MicroArray)	109
6	Bibliography.....	111
7	Appendix	i
7.1	Abbreviations	i
7.2	List of figures	v
7.3	List of tables.....	vi
	Contributions.....	vii
	Acknowledgements	x

Summary

Hematopoietic stem cells (HSCs) sustain the life-long production of blood and maintain the integrity of the hematopoietic system. Therefore, they possess self-renewal capacity and are able to differentiate into all the mature blood cell lineages. These are finely tuned processes, which constantly involve dynamic HSC fate decisions and an imbalance may result in misregulation associated with cellular transformation or other types of diseases.

In the hematopoietic disorder Fanconi Anemia (FA), such an imbalance in the HSC pool results in bone marrow failure, a collapse of the entire hematopoietic system. In addition, FA patients are susceptible to developing hematologic malignancies such as acute myeloid leukemia (AML) or myelodysplastic syndrome (MDS). This is the result of a loss of function in the FA DNA repair machinery, which renders HSCs genetically unstable as they are unable to repair certain forms of DNA damage. Murine knockout models for individual FA pathway members fully recapitulate the FA HSC DNA repair defect seen in patients and demonstrate a severe engraftment defect in competitive transplantation experiments. We could recently show, that in response to proliferative stress, such as during transplantation and reconstitution, FA HSCs accumulate DNA damage and are quickly lost due to apoptotic cell death, which is a key driver of bone marrow failure. FA HSCs are also subject to cellular transformation, which may result in the outgrowth of a leukemic founder clone. The identification of novel therapeutic targets would likely provide insight into the underlying mechanism of the FA HSC defect and the associated pathology of FA. In this respect, forward genetic screens using insertional mutagenesis have proven to be a powerful screening method for the identification of genes with the potential to influence stem cell kinetics when upregulated or disrupted.

We have employed retroviral insertional mutagenesis in the context of FA HSC biology using a murine transplantation model in order to try and identify novel factors with the potential to rescue the inherent FA HSC transplantation defect. In this respect, we show that retroviral vector integrations trigger the expansion of nonmalignant dominant FA clones in transplanted mice and allow the retrospective identification of nearby genes whose deregulation had caused clonal expansion in the face of genomic instability. We identified four candidate drivers of clonal dominance in the FA HSCs: *Osgin1*, *Evi1*, *Taf1b* and *Grhl1*, which we further characterized in the context of HSC biology under homeostatic and stress conditions as well as during development. In this respect, we identified the oxidative stress-induced growth inhibitor (*Osgin1*) as a promising candidate target gene, which is expressed in FA and WT HSCs under homeostatic conditions as well as in hemogenic endothelium/HSPCs during developmental specification from normal embryonic stem cells. Furthermore, we provide insight into the hematopoietic role of *Osgin1*, which displays differential

expression levels during the individual steps of HSC commitment and their response to physiologic stress. We conclude that *Osgin1* is an essential factor for hematopoiesis and we will further characterize *Osgin1* using the tools and assays that we have developed to assess whether it impacts on FA HSC fate decisions *in vitro* and during hematopoietic reconstitution *in vivo*. Furthermore, we provide some insight into how the FA HSCs may become a dominant clone and suggest likely candidate mechanisms with the potential to compensate the inherent FA HSC defect. According to our hypothesis, the identified candidate target genes may either prevent the acquisition of DNA damage or repair it or, alternatively, block the apoptotic cell death, which is a likely cell fate outcome for FA HSCs. Our findings hold great potential for HSC research with implications for normal as well as FA HSC biology.

Zusammenfassung

Blutstammzellen bilden die Basis des blutbildenden Systems und gewährleisten dessen Integrität. Zu diesem Zweck besitzen Blutstammzellen zum einen Selbsterneuerungspotenzial und können sich darüber hinaus in alle reifen Blutzellen entwickeln. Diese schicksalhaften Entscheidungen zwischen Selbsterneuerung und Differenzierung sind stark regulierte Prozesse, deren Ungleichgewicht schließlich die Transformation betroffener Zellen begünstigt und zu Krankheiten des Blutsystems führen kann.

In der Fanconi Anämie (FA), einer hämatopoetischen Funktionsstörung, führt solch ein Ungleichgewicht zum Versagen des kompletten Blutsystems. Außerdem sind FA Patienten anfällig für Tumore des blutbildenden Systems, wie zum Beispiel der akuten myeloischen Leukämie (AML) und des myelodysplastischen Syndroms (MDS). Dies ist auf einen Funktionsverlust des FA DNA-Reparaturmechanismus zurückzuführen, welcher verhindert, dass FA Blutstammzellen DNA-Schäden reparieren können und dadurch eine genetische Instabilität hervorruft. FA Mausmodelle, die einzelne Komponenten des FA Signalweges ausgeschaltet haben, weisen ebenfalls diesen FA Stammzelldefekt auf. Dieser kommt vor allem bei Transplantationsversuchen zum Tragen, in denen die FA Stammzellen deutlich schlechter anwachsen, als dies beispielsweise bei WT Stammzellen der Fall ist. Wir konnten kürzlich zeigen, dass Stammzellen durch den Proliferationsstress, dem sie während der Transplantation ausgesetzt sind, Schäden in ihrem Erbgut akkumulieren, was schließlich den programmierten Zelltod zur Folge hat. Dieser spielt eine Schlüsselrolle in der Pathologie dieser Bluterkrankung. FA Stammzellen fallen außerdem der zellulären Transformation zum Opfer, welche die Entstehung einer leukämischen Stammzelle begünstigen kann. Die Erforschung neuartiger Therapeutika verspricht bisher unbekannte Einblicke in die Mechanismen, die dem FA Stammzelldefekt unterliegen und mit der Pathologie der FA verbunden sind. In diesem Zusammenhang haben sich genetische Screens, die das Konzept der retroviralen Insertionsmutagenese nutzen, als überaus hilfreiche Methode erwiesen, welche die Identifikation stammzellrelevanter Gene ermöglicht.

Im Rahmen dieser Arbeit haben wir einen retroviralen Insertionsmutagenese-Screen mit FA Stammzellen durchgeführt, bei dem wir ein murines Transplantationsmodell verwenden, um neuartige Faktoren mit dem Potential den angeborenen FA Stammzelldefekt zu kompensieren, ausfindig zu machen. In diesem Zusammenhang konnten wir zeigen, dass retrovirale Vektorintegrationen im Erbgut der FA Stammzellen die Expansion gutartiger dominanter FA Stammzellklone in transplantierten Mäusen fördern. Dies erlaubt die Identifikation der durch die Integrationsstelle deregulierten Gene, welche trotz der genetischen Instabilität der FA Stammzelle die Expansion der Klone begünstigen konnten. Zusammenfassend konnten wir vier Faktoren ausfindig

machen, die sehr wahrscheinlich einen Einfluss auf die klonale Dominanz der FA Stammzellen haben: *Osgin1*, *Evi1*, *Taf1b* und *Grhl1*. Diese haben wir im Hinblick auf ihre Rolle während der Hämatopoese sowohl unter homöostatischen als auch unter Stressbedingungen näher betrachtet. In diesem Zusammenhang konnten wir den oxidativen Stress-induzierten Wachstumsinhibitor *Osgin1* als einen vielversprechenden Kandidaten identifizieren, welcher nicht nur in FA, sondern auch in normalen Blutstammzellen und während der Spezifikation embryonaler Stammzellen *in vitro* exprimiert ist. Während der einzelnen Differenzierungsschritte von Stammzellen zu Vorläuferzellen, aber auch unter physiologischen Stressbedingungen, weist *Osgin1* differentielle Expressionsmuster auf, woraus wir folgern, dass *Osgin1* ein essentieller Faktor für die Hämatopoese ist. Ausblickend werden wir *Osgin1* anhand bereits etablierter Methoden genauer charakterisieren um zu verstehen, ob *Osgin1* tatsächlich Einfluss auf die Schicksalsentscheidungen der FA Blutstammzelle nimmt und darüber hinaus das Anwachsen transplanteder FA Stammzellen begünstigen kann. Weiterhin gewähren wir Einblicke in die Mechanismen, die es der FA Stammzelle erlauben, den angeborenen Defekt zu überwinden und im Rahmen der genetischen Instabilität zu einem dominanten Klon zu werden. Unserer Hypothese zu Folge könnten die identifizierten Zielgene entweder das Auftreten des DNA-Schadens verhindern, dessen Reparatur veranlassen oder den programmierten Zelltod blockieren, welcher ein sehr wahrscheinliches Schicksal für die FA Stammzelle darstellt. Unsere Erkenntnisse beinhalten großes Potential für die Stammzellforschung und könnten möglicherweise Auswirkung auf die normale als auch auf die Biologie der FA Stammzelle haben.

1 Introduction

1.1 Hematopoiesis and the hematopoietic system

Hematopoiesis is the formation of the cellular blood components, which are hierarchically organized according to their state of maturation in the hematopoietic system. Hematopoiesis occurs in the bone marrow and the lymphatic system in order to generate the mature blood cells, such as red blood cells (erythrocytes), white blood cells (leukocytes) and platelets (thrombocytes), which elicit vital functions in order to control and maintain the integrity of the hematopoietic system. In this process, erythrocytes mediate the exchange of respiratory oxygen and carbon dioxide across all tissues, the primary function of leukocytes is to protect against infections and thrombocytes facilitate coagulation and wound repair. Some leukocytes are also involved in tissue and bone remodeling and the removal of dead cells. Many blood cells are short-lived and, due to their high turnover rates, need to be constantly replenished. The average human requires approximately one hundred billion new blood cells each day. Therefore, specialized multipotent cells at the top of the hierarchy, known as hematopoietic stem cells (HSCs), allow for the continued production of so called transit amplifying or progenitor cells, which then further differentiate into all the mature blood cell lineages. Since these cells represent only a small fraction of all bone marrow cells, amplification steps at the progenitor level allows for the production of sufficient numbers of mature effector cells (Reya et al., 2001).

1.1.1 Hematopoietic stem cells

HSCs comprise the top of the hematopoietic system hierarchy and are responsible for the life-long production of all mature blood cells as well as for the maintenance of the HSC pool in blood and bone marrow. In order to fulfill these functions, HSCs have the capacity to self-renew and give rise to differentiated cells (Morrison et al., 1997; Till and McCulloch, 1980; Weissman, 2000a). Self-renewal describes the ability of HSCs to produce daughter cells that have the same stem cell properties as the parent cell, which is essential for the maintenance and expansion of HSC numbers. During differentiation, HSCs successively lose their self-renewal capacity, but gain increased proliferation potential while going through multiple progenitor stages towards committed precursors (Osawa M, 1996).

The first hint towards the existence of a common precursor of the various blood cell types arose shortly after the atomic bombings of Hiroshima and Nagasaki in 1945, with victims exposed to radiation subsequently suffering from hematopoietic failure and the inability to regenerate sufficient amounts of mature blood cells in order to fulfill vital functions. Later studies provided definitive evidence for the presence of the HSC when irradiated mice could be rescued from otherwise fatal

hematopoietic failure by bone marrow transplantation, as the injected bone marrow cells directly regenerated the hematopoietic system (Ford et al., 1956; Lorenz et al., 1951; Nowell et al., 1956). This led to the development of bone marrow transplantation as a curative therapy for hematopoietic failure following whole body irradiation, which uses bone marrow transplantation in order to transfer HSCs that possess regenerative capacity to the host organism (Kondo et al., 2003; Weissman, 2000b). In 1990, E. Donnall Thomas received the Nobel Prize for Medicine for these pioneering discoveries, which initially set the basis for the cure of leukemias and other fatal blood diseases. When physicians started to evaluate the irradiation sensitivity of normal mouse bone marrow cells in order to formulate an irradiation therapy against cancer in the 1960s, researchers could demonstrate that transplanted cells were able to develop colonies in mouse spleens, which directly correlated to the number of initially injected bone marrow cells (Till and McCulloch, 1961; Till et al., 1964). Such colonies were derived from a single precursor cell, which was able to reconstitute the hematopoietic system of irradiated secondary recipients by both self-renewal and multilineage differentiation (Becker et al., 1963; Wu et al., 1968). These findings shaped the definition of the multipotent and self-renewing HSC and provided the basis for a lot of speculation and prediction of HSC biology. Later on, however, it was demonstrated that these spleen colonies were rather primitive progenitors than bona fide HSCs as they could only give rise to cells of the myeloid lineage. Definite proof for the existence of a stem cell, which was able to generate cells of the myeloid and the lymphoid lineage, was provided another decade later (Abramson et al., 1977).

These initial experiments set the stage for modern HSC research, which focuses on the prospective isolation of HSCs in order to characterize their function and evaluate potential applications for the development of novel cell-based therapies. HSCs hold great potential for regenerative medicine due to their proliferation and differentiation capacities (Daley et al., 2003; Keller, 2005). However, the dysregulation of stem cell properties may also be the cause of certain types of diseases and cancer (Dalerba et al., 2007; Reya et al., 2001).

1.1.1.1 Functional hematopoietic stem cell assay

In order to formally address stem cell function, a plethora of *in vitro* and *in vivo* assays have been developed. However, the only method that fully allows assessing the capacity of HSCs to be able to differentiate into all mature lineages while also being capable of self-renewal is bone marrow transplantation, which serves as the “gold standard” test for functional stem cells. In this context, the mouse transplant system is well established and by far the most frequently used model system to study mammalian hematology.

When HSCs are transplanted into irradiated mice intravenously, the cells set out to find their way from the peripheral blood to their bone marrow niches in a process called homing. During this process, HSCs rely on a variety of chemokines, cytokines and adhesion molecules in order to roll and tether to the blood vessel walls and extravasate through the bone marrow endothelium. After the HSCs have reached the perivascular region of the bone marrow, they start to repopulate the blood system. (Lapidot et al., 2005). Ideally, this assay is carried out with at least one serial step of transplantation into a secondary recipient. The ability of the transplant to repopulate a secondary recipient and give rise to all the mature blood cell lineages formally proves its self-renewal and long-term reconstitution potential.

However, recent studies have cast some doubt on whether the classically defined transplantable HSCs are really responsible for the long-term maintenance of the hematopoietic system under homeostatic conditions (Busch et al., 2015; Sun et al., 2014). Nonetheless, the transplantation assay remains the most well characterized experimental procedure for HSC function that is widely available.

1.1.1.2 Purification of hematopoietic stem and progenitor cells from the bone marrow

The basis for all stem cell assays is to prospectively enrich and isolate hematopoietic stem and progenitor cells from the bone marrow. This requires cell surface markers that uniquely define HSCs and their progenitor cells. Since the initial identification of HSCs in the 1960s, it took several decades until the first immunophenotypic marker set of murine HSCs and their committed progeny could be established. In this quest, the development of fluorescence-activated cell sorting (FACS) (Hulett et al., 1969) and monoclonal antibodies (Kohler and Milstein, 1975) has been crucial. In addition, transplantation assays were used to evaluate correlations between marker expression and the functional HSC potential in hematopoiesis.

The first attempt towards enriching HSCs demonstrated that they reside in a fraction of the bone marrow, which was depleted for the expression of markers that are expressed on the surface of mature lineage committed cells: so called “lineage markers” (Muller-Sieburg et al., 1988; Muller-Sieburg et al., 1986). Addition of the markers stem cell antigen 1 (Sca-1) and c-Kit, the receptor for stem cell factor/steel factor, further refined this population to the so-called lineage⁻, Sca-1⁺, c-Kit⁺ (LSK) population (Ikuta and Weissman, 1992; Ogawa et al., 1991; Spangrude et al., 1988), which is highly enriched for stem cells but also contains multi-potent and lineage biased progenitor cells that possess more restricted self-renewal properties than the HSC compartment. Although the LSK population is commonly used as the basis for stem cell transplantation assays, other assays often require a much higher purity. Therefore, the signaling lymphocytic activation molecule (SLAM) marker set, defined according to the cluster of differentiation (CD) marker set CD150⁺ and CD48⁻, can be exploited in

order to further enrich HSCs within the LSK population to almost 50% purity (Kiel et al., 2005). Gain of CD48 with differentiation can be used to identify the LSK progenitor compartment. Within this population the stepwise loss of CD150 and the gain of CD135 allows for the distinction of the individual multipotent progenitor populations MPP2 (CD34⁺ CD48⁺ CD150⁺ CD135⁻), MPP3 (CD34⁺ CD48⁺ CD150⁻ CD135⁻) and MPP4 (CD34⁺ CD48⁺ CD150⁻ CD135⁺).

HSCs can be further distinguished into long-term (LT-) and short-term HSCs (ST-HSCs) based on their CD34 expression, two HSC populations which possess different reconstituting capacity (Wilson et al., 2008). While long-term reconstituting LT-HSCs are negative for CD34, the ST-HSCs have limited self-renewal capacity and are CD34⁺.

1.1.1.3 Investigating HSC fate decisions

The generation of mature blood cells during hematopoiesis is a finely tuned process, which constantly involves HSC fate decisions in order to decide for self-renewal, differentiation, proliferation, quiescence, survival or apoptosis. These fates are tightly regulated by a complex interplay between cell intrinsic and extrinsic mechanisms in order to maintain blood homeostasis (Zhu and Emerson, 2002). Among the intrinsic mechanisms, transcription factors such as Runx1, Lmo2, Bmi1, Gata1 and Gata1 play a role (Orkin and Zon, 2008). Extrinsic mechanisms influence the fate of stem cells via cytokines such as stem cell factor (SCF), thrombopoietin (Tpo), erythropoietin (Epo) and interleukins as well as cell-cell contacts or extracellular matrix contacts in the bone marrow niche (Metcalf, 2008; Trumpp et al., 2010). An imbalance in these highly controlled cellular programs can result in cellular transformation of HSCs into disease-initiating leukemic stem cells. Since similar signaling pathways may regulate self-renewal in both normal and deregulated HSCs, a further understanding of these regulatory mechanisms holds the potential to identify novel therapeutic targets for the treatment of diseases such as hematologic malignancies. (Reya et al., 2001). Therefore, investigating HSC fate decisions can reveal novel insights into HSC biology as well as its implications for processes such as ageing, disease and leukemic transformation.

However, the identity of single cells in culture experiments is quickly lost and their morphology does not allow distinguishing developmental stages. Therefore, the analysis of HSCs in heterogeneous populations requires innovative single cell analysis tools in order to elucidate stem cell behavior. Michael Rieger and Tim Schroeder have developed video time-lapse microscopy in order to observe hematopoiesis in cell culture and follow the fate decisions of hundreds of cells individually (Rieger and Schroeder, 2008). With this methodology, fate decisions of cells and their progeny can be investigated and displayed in pedigrees representing the exact history of the merging colonies. Using this methodology, the group of Michael Rieger could recently identify a novel instructor of HSC

differentiation, which is inducible by cytokine signals (Thalheimer et al., 2014). Furthermore, single cell transplantation studies demonstrated the potential to revise established models of the hematopoietic hierarchy and reveal novel insights into the largely heterogeneous stem cell population (Babovic and Eaves, 2014; Yamamoto et al., 2013).

1.2 The hematopoietic disorder Fanconi Anemia (FA)

As discussed above, HSC self-renewal and differentiation are tightly regulated processes, which cause disease in humans if deregulated. For example, unrestricted self-renewal is a feature of leukemic transformation while insufficient self-renewal would lead to a depletion of HSC reserves and may cause bone marrow failure: a collapse of the entire hematopoietic system.

Fanconi anemia (FA) is a hereditary disorder and the most common inherited bone marrow failure syndrome, which impacts upon the production of all hematopoietic lineages. While FA patients usually display normal blood cell counts at birth, the hematologic complications arise later on, mainly within the first decade of life. Initially, this manifests as abnormally low platelet counts in the peripheral blood (thrombocytopenia) followed by low neutrophil counts (neutropenia) before eventually developing into aplastic anemia, the insufficient production of red blood cells, leukocytes and platelets in the bone marrow. In addition, FA patients possess severely compromised hematopoietic progenitor compartments (Bagnara et al., 1992; Daneshbod-Skibba et al., 1980). Unless treated, the hematologic manifestations of FA remain the primary cause of morbidity and mortality, with patients suffering from a markedly increased risk of myelodysplastic syndrome (MDS) and acute myeloid leukemia (AML). Approximately half of all FA patients present with MDS and/or AML before they are 40 years old (Auerbach and Allen, 1991; Butturini et al., 1994; Kutler et al., 2003).

In addition to the hematologic abnormalities, FA patients suffer to varying degrees from a range of congenital developmental defects, such as skin pigmentation abnormalities, short stature, microphthalmia or thumb abnormalities. Other organ systems involved include the cardiac, renal and auditory systems (Auerbach, 2009). In addition, FA patients are susceptible to various forms of solid tumors such as squamous cell carcinoma of the head, neck and gynecologic area as well as esophageal carcinoma and tumors of the liver, brain, skin and kidney (Alter, 2003; Kutler et al., 2003).

From the genetic point of view, FA is a heterogeneous disorder caused by inactivating mutations in 16 different genes, which have been identified to date: FANCA, FANCB, FANCC (Strathdee et al., 1992), FANCD1/BRCA2, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCIJ/BRIP1, FANCL/PHF9/POG, FANCM, FANCN/PALB2, FANCO/Rad51C (Vaz et al., 2010),

FANCP/SLX4/BTBD12 (Crossan et al., 2011; Kim et al., 2011; Stoepker et al., 2011) and FANCC/XPF (Bogliolo et al., 2013; Kashiwama et al., 2013). These genes are thought to function in an epistatic signaling pathway, a common biochemical ubiquitin-phosphorylation network, which is involved in controlling multiple functions related to deoxyribonucleic acid (DNA) repair and the cellular response to stress (**Figure 1**) (Garcia-Higuera et al., 2001).

1.2.1 The FA signaling pathway mediates DNA damage repair

Intensive research has been conducted in order to reveal the molecular function of the FA signaling pathway and its individual members. In summary, several findings demonstrated that all 16 FA proteins are critical for the repair of cross-linked DNA in order to maintain genomic stability.

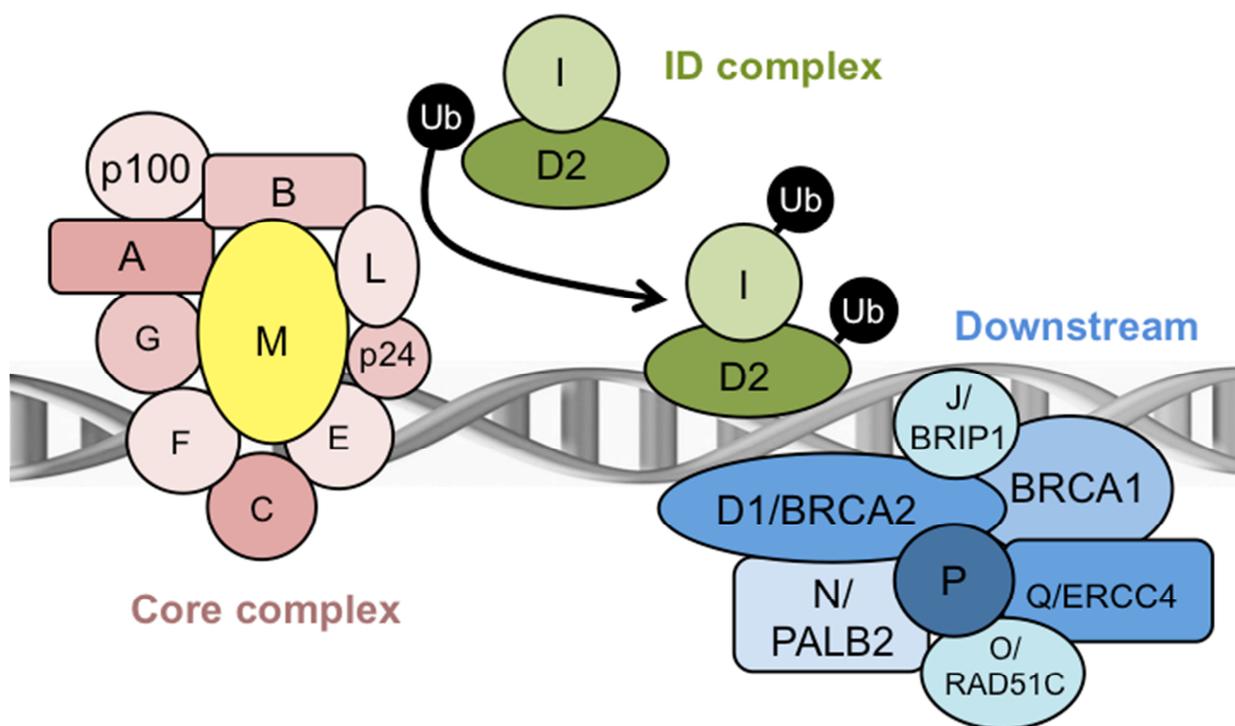


Figure 1: Schematic representation of the FA signaling pathway. FA proteins function in an epistatic signaling pathway and assemble to form the FA core complex and the ID complex, which bind to DNA and interact with classical downstream effectors to contribute to DNA damage repair. (Provided by Amelie Lier and Michael Milsom)

Upon DNA damage, FA proteins are recruited to the site of damage and assemble to form a large nuclear complex termed the FA core complex (**Figure 1**). This nuclear multi-protein complex consisting of FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL, and FANCM functions as an E3 ubiquitin ligase and, via FANCL, mediates the activation of a heterodimer composed of FANCD2 and FANCI termed the ID complex (Alpi et al., 2008; Bogliolo et al., 2013;

Garcia-Higuera et al., 2001; Howlett et al., 2002; Kashiyama et al., 2013; Kennedy and D'Andrea, 2005; Kim et al., 2011). Once mono-ubiquitinated, the ID complex can bind to single- and double-stranded DNA structures in order to repair DNA interstrand cross-links (Joo et al., 2011). This task is fulfilled by components of the excision nuclease machinery such as FANCP/SLX4/BTBD12 and FANCO/XPF, which are activated in order to excise the cross-link (Crossan et al., 2011; Kim et al., 2011; Stoepker et al., 2011). Furthermore, the ID complex interacts with classical tumor suppressors downstream of the FA pathway including FANCD1/BRCA2, FANCN/PALB2, FANCI/BRIP1 and FANCO/Rad51C thereby contributing to DNA double-strand break repair via homologous recombination (HR) (Hirano et al., 2005; Kee and D'Andrea, 2010; Niedzwiedz et al., 2004). Loss of function of any of the FA family members results in inefficient repair of DNA damage and deregulation of signaling pathways controlling cell proliferation and apoptosis.

1.2.2 Defective DNA repair causes genomic instability in FA

Fanconi anemia is a genomic instability syndrome caused by defects in the DNA repair machinery, which would usually resolve DNA interstrand cross-links, the most deleterious type of DNA lesions, which cause replication arrest and lead to DNA double strand breaks. DNA interstrand cross-links originate from endogenous sources such as nitrous acid and aldehydes or can be caused by exogenous agents such as cisplatin and its derivatives. The individual FA pathway members are critical for the repair of these cross-linked DNA in order to protect the human genome from DNA damage and maintain genomic stability (Kee and D'Andrea, 2010). Therefore, the FA pathway is an essential tumor-suppressive pathway, which prevents the development of cancer. In FA patients, the mutation in a single FA gene results in the inactivation of the whole FA signaling pathway and, ultimately, in defective DNA repair in every cell of the patients' body. For this reason, FA patients are unable to resolve certain forms of DNA damage and as a consequence of genetic instability, they are predisposed towards the development of cancer. In the hematopoietic system, genetic instability may promote malignant transformation of hematopoietic stem and/or progenitor cells and the outgrowth of an abnormal clone leading to blood malignancies with gross chromosomal changes such as MDS and/or AML (Welch et al., 2012).

FA is often referred to as a chromosome instability syndrome since FA patients display increased chromosomal aberrations, particularly radial chromosomes, at the cellular level. These aberrations can be explained by the fact that FA cells are hypersensitive to DNA interstrand cross-linking agents such as mitomycin C, cisplatin, diepoxybutane, and melphalan (Auerbach, 1993; Cervenka et al., 1981; Chen et al., 2007; Kondo et al., 2011; Sasaki and Tonomura, 1973). While mitomycin C causes radial chromosomes, diepoxybutane mainly functions as a bifunctional cross-linking agent inducing chromosomal breakage or rearrangements (Auerbach, 1993; Cervenka et al., 1981). The

increased susceptibility of FA cells to these compounds results in an exaggerated arrest in the G2/M phase of the cell cycle and apoptosis (Kaiser et al., 1982; Kubbies et al., 1985).

The genomic instability phenotype of FA cells is further promoted by the sensitivity to the inhibitory action of proinflammatory cytokines such as tumor necrosis factor alpha (TNF- α), interferon gamma (IFN- γ), and macrophage inflammatory protein-1 alpha (MIP-1 α) (Haneline et al., 1998; Rathbun et al., 1997). Both TNF- α and IFN- γ are produced at abnormally high levels in the serum and bone marrow of FA patients and have been implicated in the pathogenesis of hematopoietic failure as FA cells are prone to apoptosis when exposed to these cytokines (Dufour et al., 2003; Rosselli et al., 1994; Wang et al., 1998). For several years, it was not clear whether the abnormal response to cytokines was a cause or consequence of the bone marrow failure in FA and whether this could be linked to the DNA damage phenotype at all. However, recent findings have provided evidence for a direct link between proinflammatory cytokines and the defective DNA damage response seen in FA cells. In response to treatment with TNF- α , murine bone marrow cells defective in the FA DNA damage response demonstrated an excessive production of reactive oxygen species (Li et al., 2007). Reactive oxygen species are highly reactive and readily cause oxidative modifications to biomolecules like DNA, proteins, and lipids. Our group has recently demonstrated, that these reactive oxygen species are a metabolic byproduct of activating stem cells out of their homeostatic quiescent status into cell cycle such as in response to hematopoietic stress and comprise a physiologic source of DNA damage in HSCs (Walter et al., 2015). In addition, these findings can explain the highly penetrant bone marrow failure seen in FA patients since the inability to repair DNA damage results in accelerated cellular depletion in response to physiologic stress and leads to a collapse of the entire hematopoietic system. In normal individuals, this damage can be repaired to some extent by the coordinated action of the DNA repair machinery.

1.2.3 Bone marrow failure arises from a defect in the FA HSC

In addition to the various hematologic complications that start to arise early on, virtually all FA patients develop bone marrow failure within their first decade of life. As a result, all hematopoietic lineages are severely compromised, which strongly implies that FA patients have a defect at the level of the stem cell. Indeed, FA patients demonstrate a significant decrease in the CD34⁺ compartment, which is comprised of HSCs and progenitor cells, even well before the onset of bone marrow failure (Kelly et al., 2007). Importantly, this correlates with the fact that reduced stem and progenitor numbers can be found in umbilical cord blood taken from newborn FA patients (Auerbach et al., 1990). Thus, the depletion of hematopoietic stem and progenitor cells appears to precede the decreased production of mature hematopoietic cells and may hint towards a prenatal origin of the stem cell defect.

Moreover, bone marrow transplantation studies, in which a transplanted normal donor HSC is able to permanently reconstitute the failing hematopoietic system of a FA patient, provide evidence that the hematologic problems associated with the pathology of FA are driven by a defect at the level of the stem cell (Farzin et al., 2007; Tan et al., 2006). In addition, successful correction of the FA defect with a single functionally normal HSC was also demonstrated by the phenomenon of reverse mosaicism, which occurs as a consequence of the correction of one of the patient's nonfunctional FA alleles within a somatic cell clone either as a result of genetic recombination of two heterozygote mutations in a FA gene or from spontaneous point or frame shift mutations that restore the function of the inactivated gene. These rare genetic events have been found in several individual patients, in which a single corrected hematopoietic clone had been able to expand to sufficient numbers in order to support the hematopoietic function and survival of the patient. Importantly, cells within other somatic tissues did not contain the correcting mutations and patients presented with nonhematologic symptoms of FA (Gross et al., 2002; Mankad et al., 2006; Waisfisz et al., 1999).

1.2.4 FA knockout mouse models to study FA HSC biology

In order to study the etiology of bone marrow failure in FA, a number of murine knockout models for the components of the FA core complex (FANCA, FANCC, FANCG, FANCF, and FANCM), FANCD2, and the downstream effectors FANCD1 and FANCP have been developed (Bakker et al., 2009; Bakker et al., 2012; Crossan et al., 2011; Friedberg and Meira, 2006). Since all of these mice lack a functional FA DNA repair pathway and are unable to impact upon the acquisition of DNA damage, these models provide a platform for studying FA HSC biology in the context of genome instability. In addition they can serve as alternative transplant models in order to circumvent the problems associated with the lack of patient material due to low cell numbers (Parmar et al., 2009). In FA patients, the FA core complex genes *Fanca*, *Fancc* and *Fancg* are most commonly mutated with approximately 80% of all patients presenting with this mutations. Therefore *Fanca*, *Fancc* and *Fancg* knockouts for FA core complex loss of function are the most commonly used in FA HSC research.

These mouse models exhibit some but not all of the developmental and hematologic manifestations of human FA patients. However, the phenotype common to all of these models includes the defective regulation of cell cycle and apoptosis, spontaneous genomic instability including chromosome breakage and radial chromosomes and increased sensitivity towards DNA interstrand crosslinking agents such as diepoxybutane, mitomycin C, and cisplatin (Neveling et al., 2009). In addition to these common features of FA, the mice display some mild hematologic defects, however not all of them can be recapitulated in each individual model. While young *Fanca*^{-/-} mice present with signs of mild thrombocytopenia and impaired proliferation rates of bone marrow-derived megakaryocyte progenitors, the peripheral blood cell counts in *Fancc*^{-/-} and *Fancg*^{-/-} mice appear normal (Hadjur et

al., 2001). However, hematopoietic progenitor cells isolated from adult *Fancc*^{-/-} mice have impaired function *in vitro* (Whitney et al., 1996). Most interestingly, HSCs isolated from the bone marrow of these mouse models reveal a severe defect, which becomes evident in both, *in vitro* and *in vivo* studies. Under growth stimulatory culture conditions, *Fanca*^{-/-} and *Fancc*^{-/-} HSCs demonstrate reduced proliferation rates due to apoptosis and dramatically reduced repopulation ability in subsequent transplantation assays (Habi et al., 2005; Li et al., 2005). The *in vitro* culture provides a certain stress situation to the cells as they are forced to proliferate. Therefore, the above-mentioned findings may indicate impaired HSC maintenance during stress. Although Murine models of FA have the same overall number of immunophenotypically defined HSCs, all three mouse lines demonstrate lower frequencies of functionally defined HSCs compared to wild-type (WT) HSCs. In competitive transplantation experiments, *Fanca*^{-/-} and *Fancc*^{-/-} HSCs largely failed to engraft irradiated recipient mice, which illustrates their severe engraftment defect (Geiselhart et al., 2012).

Despite the fact that murine models of FA demonstrate a functional HSC defect during bone marrow transplantation assays, none of the FA mouse models spontaneously develops the highly penetrant bone marrow failure phenotype seen in FA patients. In addition, most of them have relatively normal hematologic function. However, these mice serve as a great model to interrogate FA pathogenesis as anemia can be elicited experimentally by *in vivo* exposure to crosslinking agents, endogenous aldehydes or conditions that model physiologic stress (Carreau et al., 1998; Hadjur et al., 2001; Langevin et al., 2011; Walter et al., 2015). Therefore, these mice allow investigating the drivers of the inherent FA HSC defect and provide a platform for the identification of potential mechanisms that may rescue it, thereby highlighting novel targets for therapeutic intervention. Furthermore, FA HSCs can serve as a model system to interrogate the biology of normal HSCs and can potentially impact upon our knowledge of more common disorders such as leukemias, myeloproliferative disorders and other bone marrow failure syndromes such as aplastic anemia.

1.2.5 HSC transplantation and gene therapy for FA patients

In light of the recent insights gained from FA mouse models, HSC transplantation currently remains the only curative treatment option for the hematologic disorders in FA patients (MacMillan et al., 2011). In fact, FA was the first disease that was successfully treated by HSC transplantation using cord blood from an unaffected human-leukocyte-antigen (HLA)-identical sibling as a starting material (Gluckman et al., 1990). Nowadays, HSCs are derived from the bone marrow, mobilized peripheral blood or umbilical cord blood of a HLA-matched donor and transplanted in order to prevent the hematological disease and malignancies, which are the most common cause of death in FA (Auerbach, 2009). However there are some major limitations to successful transplantation of FA patients such as that patients are hypersensitive to cytotoxic agents and radiotherapy used to

condition the host for bone marrow engraftment as well as the lack of availability of an HLA-matched disease-free donor (Auerbach et al., 1983; Farzin et al., 2007; Gluckman et al., 1984).

One attractive novel therapeutic modality is the genetic correction of autologous patient HSCs as it prevents the immunologic rejection of the graft as well as a source of transplantable HSCs. In this context, FA appears to be an ideal candidate for treatment via gene therapy since individual corrected HSCs, which experienced a random genetic event of reverse mosaicism, have already been demonstrated to result in sustained reversal of bone marrow failure (Mankad et al., 2006). In addition, this would suggest that host bone marrow ablation with its associated toxicity may not be necessary in FA or at least could be reduced. Initially, the feasibility of gene therapy in the context of FA had been demonstrated in FA mouse models, in which the replacement of the defective FA gene by the respective human copy DNA (cDNA) was sufficient to correct the hypersensitivity of the FA HSCs to crosslink damage and their *ex vivo* proliferation defect (Rio et al., 2002). Recent major advances have been made in the field of gene therapy, which have allowed the correction of a range of different inherited genetic disorders with a hematologic basis using a delivery system such as a retroviral vector to deliver the correcting cDNAs into patient HSCs (Aiuti et al., 2013; Biffi et al., 2013; Kohn, 2010; Watts et al., 2011).

However, besides all of the advantages that gene therapy would bring for FA patients, it is also associated with some limitations such as the requirement for a sufficient cell number for gene modification. While these input numbers are routinely achieved in other patient groups subjected to gene therapy, CD34⁺ cells isolated from FA patients are limited in number and are also extremely sensitive to *ex vivo* culture conditions necessary for expansion (Kelly et al., 2007; Li et al., 2005). To date, the clinical gene therapy trials for FA have all failed to achieve robust engraftment of corrected patient HSCs, although advances have been made in the ability of clinicians to transduce FA CD34⁺ cells with retroviral vectors (Kelly et al., 2007; Liu et al., 1999; Walsh et al., 1994). Fortunately, some of the model systems that have been developed for FA have been able to assist in the formulation of new strategies that may help to overcome the barriers to effective gene therapy of FA. For example, novel protocols such as the “rapid” lentiviral-mediated transduction of *Fanca*^{-/-} mouse hematopoietic stem cells have recently demonstrated to minimize *ex vivo* manipulation of cells and result in engraftment levels equivalent to WT cells (Muller et al., 2008). In addition to the minimized *in vitro* manipulation period, lentiviral and foamy retroviral vectors hold the distinct advantage over γ -retroviral vectors in that they are able to efficiently transduce nondividing HSCs. (Muller et al., 2008; Si et al., 2008). In addition, alternate mobilization protocols using specific Rac GTPase small molecule inhibitors (NSC23766) or CXCR4 antagonists (AMD3100) were able to overcome the FA HSC mobilization defect towards granulocyte colony stimulating factor (G-CSF) (Milsom et al., 2009; Pulliam et al., 2008).

All these advances highlight the therapeutic benefit of gene therapy for FA patients. In addition, they reveal novel targets for therapeutic intervention and provide proof of principle for the hypothesis that the FA HSC defect can be rescued or at least compensated. In this respect, novel therapeutic targets as well as biological insight in the FA HSC defect might be obtained by looking at genes with the potential to impact on the FA HSC biology. In this setting, a forward genetic screen might be employed to identify such genes.

1.3 Retroviral insertional mutagenesis

Retroviral insertional mutagenesis is referred to as the deregulation of gene expression resulting from proviral integration into the host genome during viral replication. In this context, the integration of the retrovirus genome in the vicinity of proto-oncogenes or within tumor suppressor genes may result in the deregulated expression of these genes, thereby eliciting cellular transformation. Retroviruses can induce cellular transformation in two ways. In contrast to acute transforming retroviruses, which contain viral versions of normal cellular proto-oncogenes such as Myc and Src and are able to rapidly induce cellular transformation (Beug et al., 1979; Gonda et al., 1982; Roussel et al., 1979), retroviruses that cause mutations by proviral integration do not carry oncogenes and cause transformation with a longer latency. These viruses contain very strong promoter/enhancer elements in their genome that drive and regulate viral transcription and are able to alter the expression of cellular genes proximal to the integration site upon insertion (**Figure 2 A+B**). In addition, retroviral integration may also result in the disruption of cellular gene expression (**Figure 2 C**). One example of a retrovirus with a semi-random integration profile is the murine leukemia virus, which preferentially targets the promoter regions of actively transcribed genes (Wu et al., 2003).

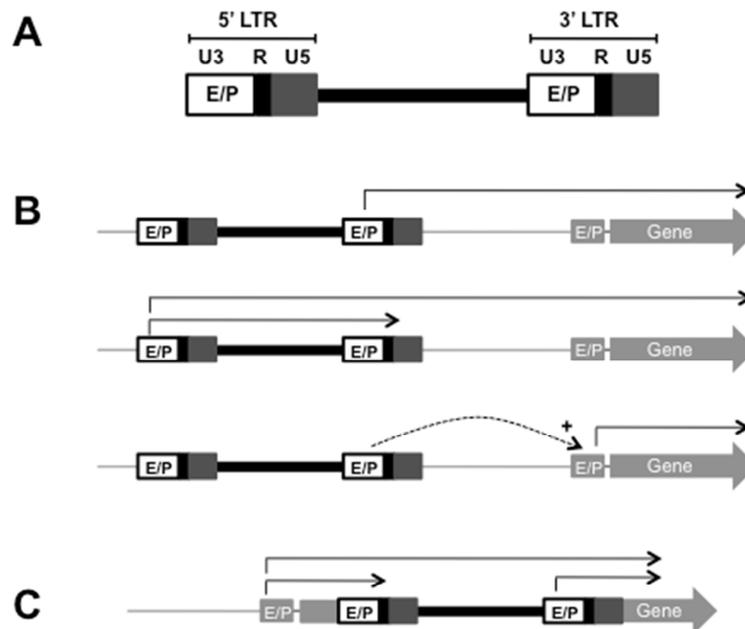


Figure 2: Mechanism of retroviral insertional mutagenesis. (A) Schematic representation of the integrated form of a recombinant retroviral vector. The long terminal repeats (LTRs), flanking the viral genome, consist of the U3 region, which contains the transcriptional enhancer/promoter (E/P) element, the R region containing a polyadenylation signal and the U5 region. These features may contribute towards insertional mutagenesis. (B) Potential mechanisms through which the integrated provirus may promote the deregulation of genes proximal to the integration site. (C) Integration within a cellular gene resulting in disrupted gene expression. (Modified from Mueller, 2011)

1.3.1 Insertional mutagenesis in gene therapy clinical trials

Besides the integration effects that occur naturally during viral infection, insertional mutagenesis has been observed as an adverse side effect in gene therapy clinical trials (Baum et al., 2003). Although replication-deficient retrovirus-based vectors have been used successfully to deliver correcting genes for the treatment of several genetic diseases including severe combined immunodeficiencies (Aiuti et al., 2002; Boztug et al., 2010; Cavazzana-Calvo et al., 2000; Gaspar et al., 2004), chronic granulomatous disease (CGD) (Ott et al., 2006) and adrenoleukodystrophy (ALD) (Cartier et al., 2009) follow up clinical studies have revealed adverse events associated with clonal genotoxicity resulting in the development of leukemia (Braun et al., 2014; Hacein-Bey-Abina et al., 2003a; Howe et al., 2008; Ott et al., 2006). Interestingly, transplanted gene-modified HSCs that had led to the leukemic development in these trials often had very similar integration sites in the genome suggesting that the nearby genes were able to promote the expansion of malignant clones. Indeed, subsequent analyses revealed several genes that are usually involved in normal hematopoiesis such as LMO2 and EVI1/MDS1 and PRDM16 but had been deregulated as a result of the retroviral integration (Aguilo et al., 2011; Chuikov et al., 2010; Deichmann et al., 2011; Gratzinger et al., 2009; Oram et al.,

2010; Ott et al., 2006). These studies have not only revealed several key players of leukemogenesis but have also led to advances in the design of viral vectors in order to improve their safety features for the successful use in gene therapy (Aiuti et al., 2013; Biffi et al., 2013; Coci et al., 2015; Suerth et al., 2014).

1.3.2 Insertional mutagenesis as a powerful screening method

The unwanted side effect of insertional mutagenesis observed in gene therapy clinical trials has been employed in genetic screens in order to identify mutations contributing to tumor formation. Retroviral vectors able to induce multiple mutations in the same cell have been associated with leukemic complications and revealed proto-oncogenes that contributed to tumorigenesis in model organisms such as the mouse. Therefore, retroviral vectors allow for the investigation of oncogenic candidate genes as well as signaling pathways that collaborate in cancer (Akagi et al., 2004; Mikkers and Berns, 2003; Uren et al., 2005). These studies have led to the identification of several key regulators of oncogenesis (Akagi et al., 2004). Probably the most prominent example is the ecotropic viral integration site 1 (Evi1) proto-oncogene, a transcription factor with a role in both self-renewal and transformation of HSCs, as it has been found repeatedly deregulated in different insertional mutagenesis screens (Buonamici et al., 2004; Li et al., 2002). However, the oncogenic potential of vector integrations is largely dependent on a high vector dose and involves the selection of clones with combinatorial hits near proto-oncogenes or other signaling genes (Du et al., 2005; Fehse et al., 2004; Modlich et al., 2005). Therefore, most of the screens for oncogenes have used replication competent viruses to increase the likelihood of a single cell acquiring the combinatorial hits required for transformation and to identify cooperating cancer mutations (Du et al., 2005).

Moreover, Kustikova and colleagues could demonstrate that retroviral vector integrations can also trigger the expansion of nonmalignant clones in murine long-term hematopoiesis (Kustikova et al., 2005). In these dominant clones, the vast majority of genes whose deregulation had caused clonal expansion were retrospectively identified as genes that are already expressed in hematopoietic stem and progenitor cells (Kustikova et al., 2007). Therefore, insertional mutagenesis screens may also be used in the absence of adverse events in order to gain insight into genes and processes underlying normal hematopoiesis such as stem cell engraftment. HSCs that harbor proviral integrations next to genes involved in stem cell self-renewal or expansion may possess an engraftment advantage upon transplantation. A very prominent example is HOXB4, a homeodomain transcription factor that promotes HSC expansion without malignant transformation (Antonchuk et al., 2002). Altogether, these studies have demonstrated, that insertional mutagenesis is a powerful genetic screening method that can be applied in an *in vivo* setting to identify genes with the potential to influence stem cell kinetics when upregulated or disrupted, which is of great interest for regenerative medicine.

1.3.3 Identification of retroviral integration sites and nearby deregulated genes

In order to identify the genes near proviral integration sites in the genome of individuals that have received gene-modified stem cells, it is necessary to recover the integration sites from the highly complex DNA background of the peripheral blood cell or bone marrow sample. For this purpose, southern blotting with the flanking sequences of proviral insertions as probes was originally used. To date, this laborious approach, which requires extensive input material has been substituted by much more sensitive polymerase chain reaction (PCR)-based technologies, which combine long terminal repeat (LTR)-chromosome junction amplification with next-generation sequencing and bioinformatics.

The amplification of provirus LTR-chromosome junctions in the genomic DNA is commonly conducted using ligation-mediated (LM)-PCR (Schmidt et al., 2001) or the more sensitive linear-amplification-mediated (LAM)-PCR (Schmidt et al., 2007; Schmidt et al., 2002). LM-PCR allows for the direct genomic sequencing of multiple unknown DNA flanking sequences at LTR-chromosome junctions by restriction digest of the genomic DNA and PCR-based extension of primers specific to the LTRs of the virus. As these primers are biotinylated, primer extension products can be enriched and ligated to an oligo linker cassette in a solid-phase LM-PCR in order to amplify the DNA by two exponential seminested PCR steps. Amplification products can then be sequenced in order to identify the genomic integration sequence. LM-PCR has been demonstrated to generate reliable and reproducible results for the investigation of clonal samples *in vivo* and *in vitro* and has the potential to recover a single dominant clone within a cellular population of 1000 other cells (Kustikova et al., 2007; Schmidt et al., 2001). Therefore, LM-PCR is very well suited for the identification of genes in the vicinity of vector integration with the potential to influence competitive fitness.

In contrast, LAM-PCR has mainly been used for high throughput analysis of viral integration sites in preclinical and clinical settings (Biasco et al., 2011; Boztug et al., 2010; Cartier et al., 2009; Gabriel et al., 2009; Hacein-Bey-Abina et al., 2003b; Montini et al., 2006; Ott et al., 2006; Paruzynski et al., 2010), in which the greater sensitivity of the method allowed for detecting and sequencing of unknown DNA flanking sequences down to a single event (Schmidt et al., 2002). LAM-PCR employs linear amplification of LTR-chromosome junctions followed by double-stranded DNA synthesis. The double-stranded DNA sequences are then digested with restriction enzymes, followed by linker ligation to the sequence and nested PCR. An advancement of this technology, nonrestrictive linear-amplification-mediated (nrLAM)-PCR, avoids the often-criticized restriction bias of the recovered integration sites, allowing for a comprehensive approach for unbiased restriction site retrieval independent of amplification efficiency during the PCR step (Gabriel et al., 2009; Paruzynski et al., 2010). nrLAM-PCR has been successfully coupled with next generation

pyrosequencing allowing for high-throughput analysis of the LTR-chromosome junctions in preclinical and clinical settings.

To reveal the specific genomic sites of integration and allow for the identification of nearby genes and other genomic features of the amplified products, bioinformatical data mining is used to align the sequence reads to their respective genomes using the UCSC basic local alignment search tool (BLAST)-like alignment tool (BLAT). In addition, bioinformatics analyses have identified clusters of insertion sites across different insertional mutagenesis studies in mice as well as in clinical gene therapy trials and summarized in a publicly available database (Akagi et al., 2004). These so called common insertion sites (CIS) reveal areas of the same genomic location in different individuals, in which the proviral insertions had clustered more frequently than it would be expected by chance and can therefore assist in identifying insertions with a selective advantage for the cell.

Altogether, these methods provide a valuable toolbox for the isolation of vector integrations and the identification of genes that might have been deregulated in response to proviral integration. Since each individual method is associated with different advantages and disadvantages, their applicability for the insertion site retrieval may vary depending on the context and the purpose of the experimental setting.

2 Aim of the thesis

Effective hematopoiesis is reliant upon dynamic HSC fate decisions in order to maintain blood homeostasis. These cell fate decisions are tightly regulated and an imbalance may result in cellular transformation or other types of diseases. In the hematopoietic disorder FA, a loss of function in the FA DNA repair machinery renders HSCs genetically unstable as they are unable to cope with certain forms of DNA damage. In response to stress, FA HSCs are either lost due to apoptosis or subject to transformation. As a result, most of the FA patients develop bone marrow failure, a collapse of the entire hematopoietic system, and have an increased likelihood of developing hematologic malignancies such as AML or MDS. But what actually determines the decision between apoptosis and transformation in individual FA HSCs? As the cancer stem cell must have evolved mechanisms to evade apoptosis in order to propagate itself, this phenomenon provides a first hint towards the potential rescue of FA HSCs despite the DNA repair defect. We are interested in the identification of these rescue mechanisms that have the potential to compensate the inherent FA HSC defect and prevent stem cell loss in the face of genome instability. For this purpose, the FA mouse model serves as an important research tool. FA HSCs isolated from the FA knockout mouse models fully recapitulate the inherent FA HSC DNA repair defect seen in patients and demonstrate reduced proliferation rates *in vitro* as well as a profound engraftment defect in competitive transplantation experiments. However, the drivers of the defect remain to be investigated since no FA knockout mouse model spontaneously develops the severe aplastic anemia that is observed in almost all FA patients. We have recently demonstrated that proliferative stress is a key driver of bone marrow failure in FA mice, leading to increased DNA damage in FA HSCs, depletion of functional FA HSCs and if FA HSCs are chronically induced into cycle, eventual collapse of the hematopoietic system leading to severe aplastic anemia. Importantly, this probably relates to the FA HSC engraftment defect since donor HSCs are subject to prolonged replicative stress during transplantation.

Here, we aim to rescue the inherent FA HSC transplantation defect in order to identify novel factors and mechanisms with the potential to prevent HSC loss in the face of genomic instability.

Specific aims:

Specific aim 1:

To conduct a retroviral insertional mutagenesis screen using FA and WT HSCs in a murine transplantation model.

Specific aim 2:

To perform a molecular analysis on dominant FA HSC clones that are able to overcome the inherent FA HSC transplantation defect.

- 2.1: The identification of the retroviral integrations sites in the genome of the genetically modified FA HSCs.
- 2.2: The analysis of whether normal gene expression is disrupted in the vicinity of the proviral integration site.

Specific aim 3:

To characterize the candidate target genes that were identified in the screen.

- 3.1: The development of assays to assess whether candidate genes impact on FA and WT HSC biology.
- 3.2: The evaluation of target gene expression and epigenetic status in HSCs under homeostatic and stress conditions as well as during development, in order to determine whether these genes may play a role in normal hematopoiesis.
- 3.3: The assessment of the biological impact of target gene overexpression in WT and FA HSCs during hematopoietic reconstitution following transplantation.

3 Results

3.1 Retroviral insertional mutagenesis screen

In order to identify genes that might be able to compensate for the inherent FA stem cell defect when upregulated or disrupted, we have set up a gain-of-function screen using retroviral insertional mutagenesis in the FA mouse model. Proof of concept for the experimental approach has already been demonstrated in murine WT cells, in which retroviral-marked HSCs were selectively expanded upon proviral integration when the selective pressure of bone marrow transplantation was applied (Geiger et al., 2012; Kustikova et al., 2005). In addition, Hartmut Geiger and colleagues have used insertional mutagenesis in the context of identifying target genes that are sufficient to confer a competitive selection advantage to HSCs in mice following irradiation (Geiger et al., 2012).

For the insertional mutagenesis screen in the FA mouse model, we have adopted the experimental design from Kustikova and colleagues (Kustikova et al., 2007) using the SF91 recombinant retroviral vector backbone, which is based on the murine leukemia virus. This replication-deficient virus preferentially integrates into the promoter regions of actively transcribed genes (Wu et al., 2003) and contains strong enhancer-promoters in the LTRs capable of mutating the expression of genes proximal to the integration site. In addition, it also includes the fluorescent protein Venus, which allows tracking of transduced cells upon transplantation.

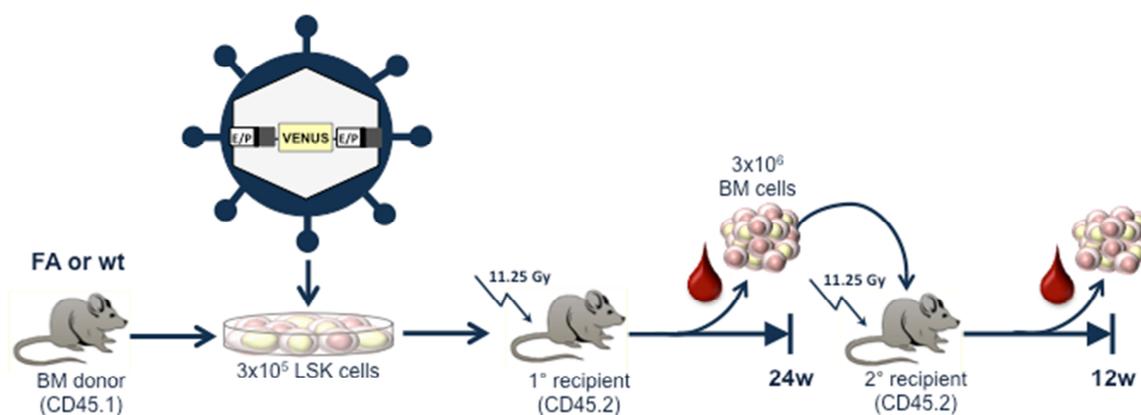


Figure 3: Experimental design of the retroviral insertional mutagenesis screen. LSK cells isolated from either FA or WT mice (CD45.1⁺ donors) were transduced with the S91 retroviral vector expressing the fluorescent marker Venus and transplanted into lethally irradiated WT mice (CD45.2⁺ primary recipients) at 3×10^5 cells per mouse. After 24 weeks, peripheral blood and bone marrow cells were isolated and 3×10^6 whole bone marrow cells were transplanted into myeloablated secondary recipients (CD45.2⁺). After 12 weeks, mice were sacrificed and peripheral blood and bone marrow cells were analyzed by FACS.

Figure 3 depicts a scheme of the retroviral insertional mutagenesis screening procedure. Murine bone marrow cells were isolated from either FA knockout mouse models (*Fanca*^{-/-}, *Fancc*^{-/-}, *Fancg*^{-/-}) or WT mice. Mature lineage-committed cells were depleted and FACS-isolated for the LSK fraction, which is enriched for stem cells and multipotent progenitors. All leukocytes isolated from these mice, including stem and progenitor cells, express the CD45.1 version of the common leukocyte antigen. This means that donor leukocytes in the peripheral blood, bone marrow and spleen can additionally be distinguished from recipient cells by antibody staining as well as Venus-fluorescence, since recipient mice express the CD45.2 epitope. These LSK cells were then transduced with the SF91 replication-deficient retroviral vector at a multiplicity of infection (MOI) in the range of 1 to 15, resulting in an average transduction frequency of 40.35% and an estimated average copy number of one to two proviral integrants per transduced cell. Irradiated WT recipient mice were subsequently transplanted with 3×10^5 unsorted retroviral-marked LSK cells per recipient. 24 weeks after the primary transplant, recipient mice were sacrificed and the hematologic parameters were assessed including the percentage of Venus-positive cells in the blood and bone marrow of the mice. 3×10^6 whole bone marrow cells were then transplanted into lethally irradiated secondary recipient mice in order to create a selection pressure for dominant stem cell clones. In this scenario a dominant stem cell clone would transplant and engraft again in the secondary recipient in order to give rise to differentiated cells, whereas all progenitor cells would be lost independent of whether they had gained a growth advantage upon retroviral integration. Therefore, we ensured the selection for stem-cell relevant genes in our analysis. 12 weeks after the secondary transplant all secondary recipient mice were sacrificed and analyzed by flow cytometry.

Primary as well as secondary recipient mice were thoroughly analyzed for the expansion of dominant stem cell clones by flow cytometry. In this context we looked at the percentage of transduced cells in bone marrow and peripheral blood as well as at the ratio of B, T and myeloid cells. The surface marker CD45.1/2 was used to distinguish donor cells from host cells in order to exclude cells that repopulated the recipient due to inefficient irradiation. Primary material, comprising cells from bone marrow and peripheral blood, was collected and stored for downstream analyses such as DNA sequencing and expression analyses. In addition, hematologic parameters such as cell counts and organ size of spleen and thymus were monitored in order to be able to distinguish malignant from non-malignant clonal expansion retrospectively.

3.2 Analysis of insertion sites in transduced FA and WT HSCs

To date, different methods to recover insertion sites from the highly complex DNA background of retroviral transduced cells have been used with constant development leading to increased sensitivity of detection. We initially performed a global analysis of the retroviral integrations present in clones

dominating in serially transplanted recipients using LM-PCR, which introduces a bias for dominant clones while neglecting integrations derived from minor clones (Schmidt et al., 2001). The reliability of this method results from enrichment of target DNA via magnetic extension primer tag selection, in which a reverse primer specific for the LTR U3 region of the virus is extended in a PCR step and subsequently coupled to magnetic beads (see **Figure 38** in 5.9.1). This is followed by a solid-phase LM-PCR, in which two exponential PCR steps amplify the target DNA. The amplification products corresponding to dominant clones recovered from the peripheral blood of transplanted primary and secondary mice were separated on an agarose gel and the prevalent bands were extracted from the gel, ignoring weak bands that might reflect insertion sites in minor clones (see **Figure 5** in 3.4). The DNA was then sequenced and blasted against the mouse genome in order to identify the genomic integration sites of the retroviral integrations.

An overview of the insertion sites that were recovered from the primary and secondary recipient mice can be seen in Table 1. In total, 174 primary recipient mice have been analyzed, of which 134 were FA (47 *Fanca*^{-/-}, 62 *Fancc*^{-/-}, 25 *Fancg*^{-/-}) and 40 were WT transplanted mice. From these primary recipients, a total of 395 insertion sites were recovered, 96 corresponding to *Fanca*^{-/-}, 130 to *Fancc*^{-/-}, 59 to *Fancg*^{-/-} and 110 to WT transplants. In the second transplantation round, 20 mice were lost due to graft failure and were therefore excluded from the analysis. In total, we could analyze 154 secondary recipient mice, of which 117 were FA (37 *Fanca*^{-/-}, 56 *Fancc*^{-/-}, 24 *Fancg*^{-/-}) and 37 were WT transplanted recipient mice. From these secondary recipient mice, we were able to recover 205 insertion sites with 43 corresponding to *Fanca*^{-/-}, 71 to *Fancc*^{-/-}, 24 to *Fancg*^{-/-} and 67 to WT transplants. These numbers demonstrate the reduction in the total amount of integration sites from 395 integrations in the primary to 205 insertion sites in the secondary recipient during an additional round of transplantation.

Table 1: Total number of insertion sites recovered from primary and secondary recipient mice

experimental group	1° recipient mice		2° recipient mice		total n° of insertion sites
	n° of mice	n° of insertion sites	n° of mice	n° of insertion sites	
<i>Fanca</i> ^{-/-}	47	96	37	43	139
<i>Fancc</i> ^{-/-}	62	130	56	71	201
<i>Fancg</i> ^{-/-}	25	59	24	24	83
WT	40	110	37	67	177
total	174	395	154	205	600

Initially, we wished to analyze this data by interrogation of all of the integration sites in FA HSCs in order to identify common insertion sites (CIS). However, CIS could not be detected with this approach since, despite the large number of experimental mice in each cohort, not enough integration sites could be recovered to analyze patterns of commonly targeted regions of the genome. We therefore sought another route of data analysis that would allow us to meaningfully compare the relatively low number of integrations sites recovered in this study.

3.3 Engraftment rate of transduced FA and WT HSCs

To find an alternative way of analyzing the data, we defined a selection strategy based on a phenomenon initially observed in human clinical gene therapy trials using retroviral vectors for the correction of certain gene-defects. These studies have provided clear evidence that, in individual patients, dominant clones can arise and take over the hematopoietic system of that patient (Braun et al., 2014; Cavazzana-Calvo and Fischer, 2007; Ott et al., 2006). Informed by these studies, we decided to next focus our analysis on individual mice with clear evidence of clonal dominance in order to select for clones in which the proviral integrant had likely impacted upon HSC biology.

For this purpose, we followed the engraftment rate of transduced cells over the time course of serial transplantations using flow cytometry (**Figure 4 A**). By plotting CD45.1 versus Venus, we could identify the total contribution of the gene-modified graft to hematopoiesis in the secondary recipient. In some cases it was clear that the donor graft was almost completely depleted over the course of two rounds of transplantations (**Figure 4 B**). In contrast, in some mice the donor cells were responsible for the product of almost all peripheral blood cells in the secondary recipients (**Figure 4 C**). Such an expansion might be the result of an individual clone having a selection advantage over the residual recipient cells due to insertionally deregulated growth-promoting genes.

If we analyzed all mice and first looked at the contribution of original donor cells to the hematopoietic system in the secondary recipients (**Figure 4 D**), we saw that in the majority of mice receiving FA donor cells, the transplanted CD45.1⁺ cells had been depleted. This phenotype was especially pronounced for the *Fanca*^{-/-} and *Fancg*^{-/-} group, in which only a few mice revealed a high CD45.1 chimerism. In contrast, mice of the *Fancc*^{-/-} group revealed an engraftment rate of CD45.1 cells almost comparable to WT. This finding correlates with the milder HSC defect that has been previously observed in this knockout mouse strain (Geiselhart et al., 2012). On average, WT donor cells were able to reconstitute secondary recipients better than FA.

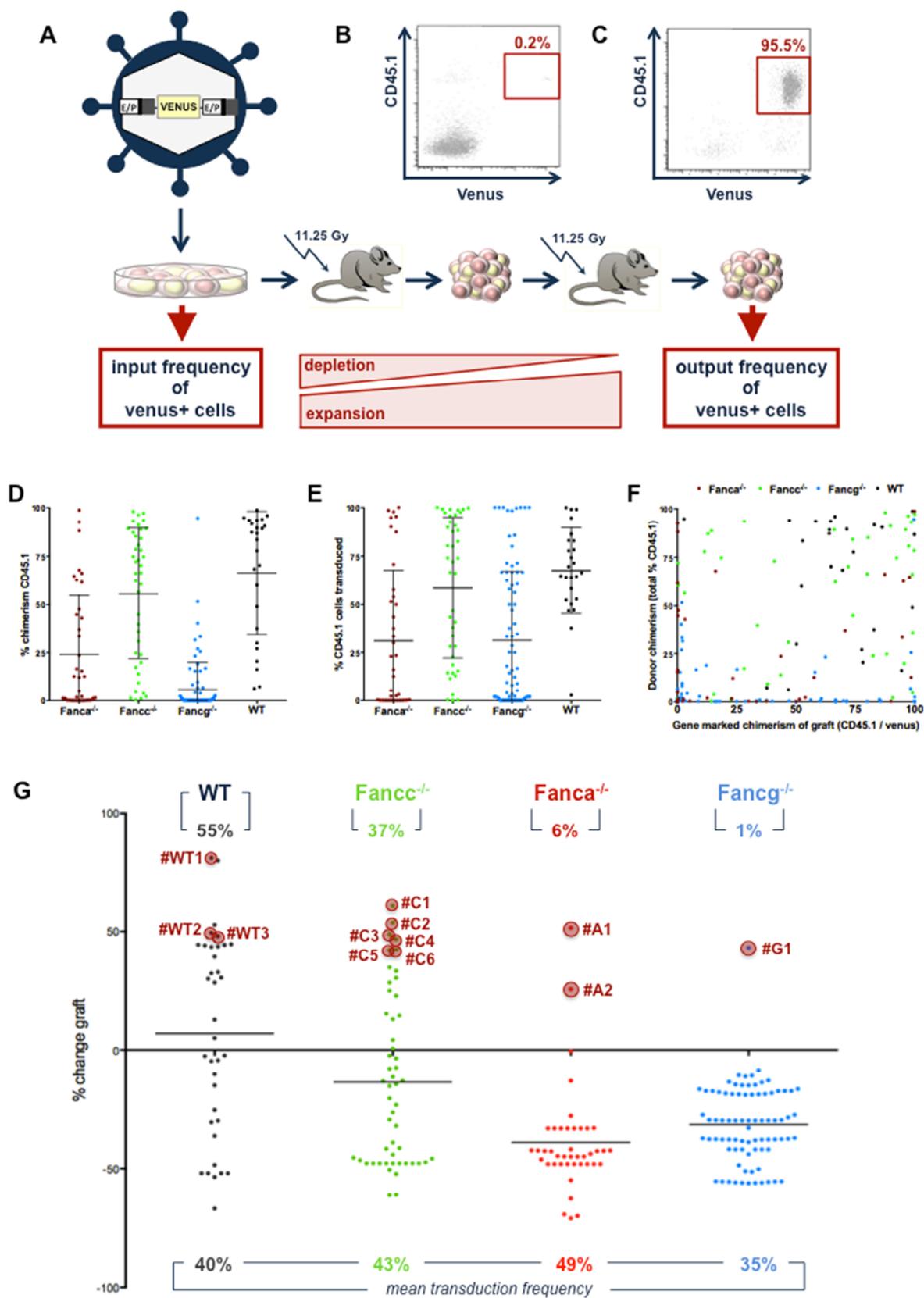


Figure 4: Engraftment rate of transduced FA and WT HSCs. (A) Schematic outline of the analysis of secondary recipient mice for the percentage of Venus-positive cells. The input frequency of SF91-transduced Venus-positive bone marrow cells was compared to the output frequency of Venus-positive bone marrow cells after serial transplantation. Transduced cells were either lost (depletion) or have expanded (expansion). (B) Representative FACS plot of a FA bone marrow sample with no engraftment (0.2%) of the transplanted (CD45.1⁺) and transduced (Venus⁺) cells. (C) Representative FACS plot of a FA bone marrow sample with strong engraftment (95.5%) of the transplanted (CD45.1⁺) and transduced (Venus⁺) cells. (D) Percentage of original donor cells (CD45.1⁺) in each individual *Fanca*^{-/-} (red), *Fancc*^{-/-} (green), *Fancg*^{-/-} (blue) and WT control (black) mouse. (E) Percentage of transduced (Venus⁺) original donor cells (CD45.1⁺) in each individual *Fanca*^{-/-} (red), *Fancc*^{-/-} (green), *Fancg*^{-/-} (blue) and WT control (black) mouse. (F) Relationship between donor chimerism (CD45.1⁺) and gene-marked chimerism (CD45.1⁺/Venus⁺) of the graft. (G) Change in percentage of transduced original donor cells (Venus⁺ CD45.1⁺) after two rounds of transplantation with individual mice that demonstrated a clear expansion of the gene-modified graft (#).

Next, we wanted to see how gene modification with the virus impacted on clonal dominance, so we looked at the percentage of CD45.1 cells that were expressing Venus (**Figure 4 E**). While the majority of mice in the WT group revealed a gene-marked chimerism between 50 - 100%, the average frequency of Venus-positive donor cells was much lower in the FA groups with around 30% for the *Fanca*^{-/-} and the *Fancc*^{-/-}. Although the *Fancc*^{-/-} group demonstrated higher transduction rates on average, only half of the mice had Venus-levels that were comparable to WT. However, some FA mice that had high level of chimerism were gene-modified, which suggests that the vector had integrated into the genome of the cells conferring an engraftment advantage.

Finally, we analyzed the relationship between the donor chimerism and the gene-marked chimerism of the graft. If we plotted these two variables together, we could observe a low level of chimerism for most of the FA mice. Especially in the *Fanca*^{-/-} and the *Fancg*^{-/-} groups, the majority of original CD45.1⁺ donor cells were depleted and very few mice displayed high Venus-positivity (**Figure 4 F**). Therefore it seems likely that in some mice, the integration of the retrovirus was able to impact upon FA HSC biology and a high gene-marked chimerism would therefore be a sign for cells that have expanded due to the insertional deregulation of growth-promoting genes.

Taken together, we could identify individual mice with high chimerism, in which retroviral-transduced Venus-positive cells have expanded and clearly dominated over the non-transduced FA cells and the residual recipient hematopoiesis. Even though we initially transplanted the same number of input cells, this phenomenon could have been an artifact relating to varying transduction efficiencies. In order to account for this, we looked at the change in percentage of gene-marked cells over two rounds of transplantation compared to the input transduction efficiency. We therefore subtracted the percentage of Venus-positive donor cells that were initially transplanted into recipient mice (input frequency) away from the final percentage of Venus-positive donor cells (output frequency) that were identified in the secondary recipient. We considered an increase in the

percentage of transduced cells over the course of serial transplantation to comprise evidence of clonal expansion, while mice in which there was no difference in the percentage of Venus-positive cells or a depletion of the transduced cells were eliminated from further analysis. The results obtained from this analysis are summarized in **Figure 4 G**. In the control group, which was composed of WT mice that were transplanted with retroviral-transduced WT cells, 56% of all mice displayed an increase in the number of transduced cells, which could be indicative of clonal dominance. In the *Fancc*^{-/-} group, which was composed of mice that had received retroviral-transduced *Fancc*^{-/-} cells, the situation was slightly different as more mice appeared in the negative area of the graph. Only 37% of all mice showed an expansion of transduced cells. In the *Fanca*^{-/-} and *Fancg*^{-/-} group, which was composed of WT mice that had received retroviral-transduced *Fanca*^{-/-} and *Fancg*^{-/-} cells respectively, we only found an expansion of transduced cells in 6% of the *Fanca*^{-/-} and 1% of the *Fancg*^{-/-} mice.

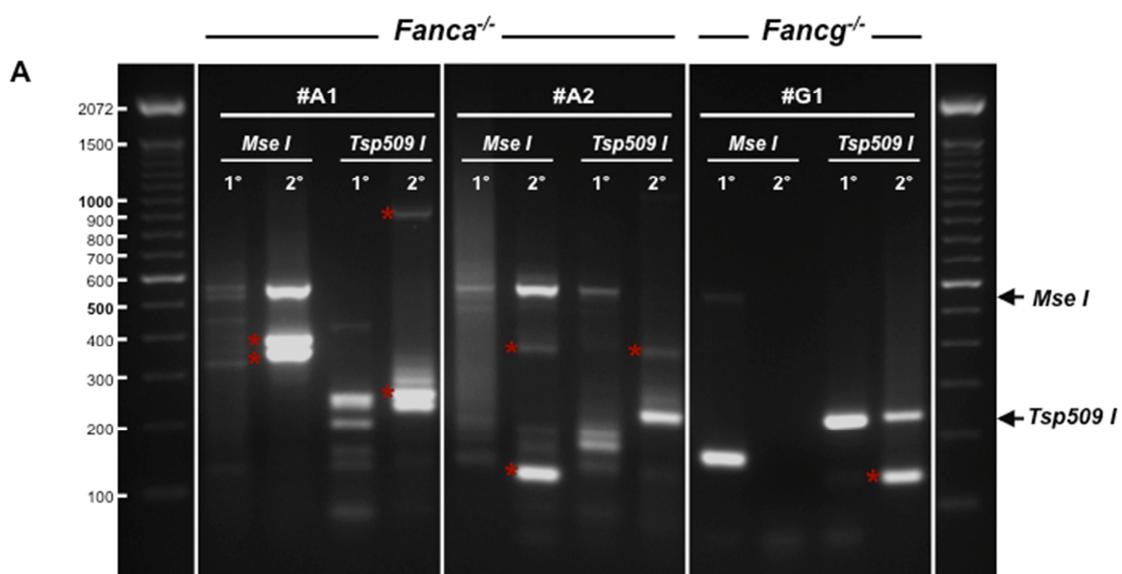
Overall, it appears that the majority of mice that had received transduced FA cells, demonstrated a severe depletion of the transplanted gene modified cells suggesting that the selection pressure for dominant clones is higher in the Fanconi model than in WT cells. Therefore, the expansion of the gene-modified graft in some of the mice suggests deregulated expression of genes that allow the HSC to overcome the FA-specific selection pressure and to become a dominant stem cell clone. As these mice would reveal interesting target genes with the potential to alter the stem cell kinetics in the context of FA HSC biology, the mice with the strongest change in graft were selected in each experimental group. In the *Fancc*^{-/-} group, mice #C1 (60.94%), #C2 (53.84%), #C3 (48.8%), #C4 (42.3%), #C5 (42.14%) and #C6 (34.9%) were selected. In *Fanca*^{-/-} and *Fancg*^{-/-} HSCs, it seems that the selection pressure for dominant clones was extremely high as only three of the mice in these groups, termed #A1 (51.6%), #A2 (25.8%) and #G1 (43.2%), demonstrated such signs of clonal dominance.

Moreover, we also considered the group of mice receiving WT transduced HSCs, as some individuals demonstrated a strong expansion of the gene-modified graft, which outnumbered the expansion rates in all FA groups. This was the case for mouse #WT1, which had a change in graft of 81.2%. Although this is probably due to a comparably low selection pressure for dominant clones in the WT setting, the analysis of retroviral integrations in the dominant clones would still provide some insight into how these cells mediated clonal dominance in a competitive setting with normally engrafting stem cells. In addition, this would allow us to reveal some aspects of normal HSC biology. Therefore we have selected three of the top WT candidate mice in order to characterize the retroviral integrations. In addition to mouse #WT1 (81.2%) we have selected mice #WT2 and #WT3 with a change in graft of 49.5 and 48.1 respectively.

3.4 Clonal dominance of transduced FA and WT HSCs

To further analyze the individual mice that had dominated in the retroviral insertional mutagenesis screen, we determined the clonal composition of their hematopoietic tissues. For this purpose LM-PCR analysis was performed on bone marrow samples from the secondary and corresponding primary recipients (**Figure 5**). The analytical LM-PCR gels show the insertions recovered from the mice after analysis with two distinct restriction enzymes (*Tsp509I* and *MseI*), which increases the likelihood of recovering all insertions. In addition, the restriction digest generated a vector-specific internal control band, which indicates a successful analysis. For the SF91 retroviral vector used in this study, the control band corresponded to 217 bp for *Tsp509I* and to 544 bp for *MseI* restriction as indicated by the arrows. These vector-specific internal control bands could be detected in almost all of the analyzed mice, however absence of the control in some of the mice could be the result of problems during primer annealing due to secondary stem loop structures. In addition, sequences with a high GC-content or a larger DNA fragment could have led to problems during the denaturation or the amplification step respectively.

All other individual bands depicted retroviral insertions corresponding to dominant clones. Primary recipients revealed polyclonal hematopoiesis, which was a result of multiple cells carrying retroviral insertions as represented by a smear or several distinct bands. After an additional round of transplantation, secondary recipients only revealed one or two dominant clones (marked by an asterisk), which are indicative of mono or biclonal hematopoiesis in these animals. This demonstrates a dramatic reduction in the complexity of integration sites in the secondary recipient despite the increase in the gene marked chimerism, which was observed in the analysis of Venus-positive cells (see **Figure 4** in 3.3).



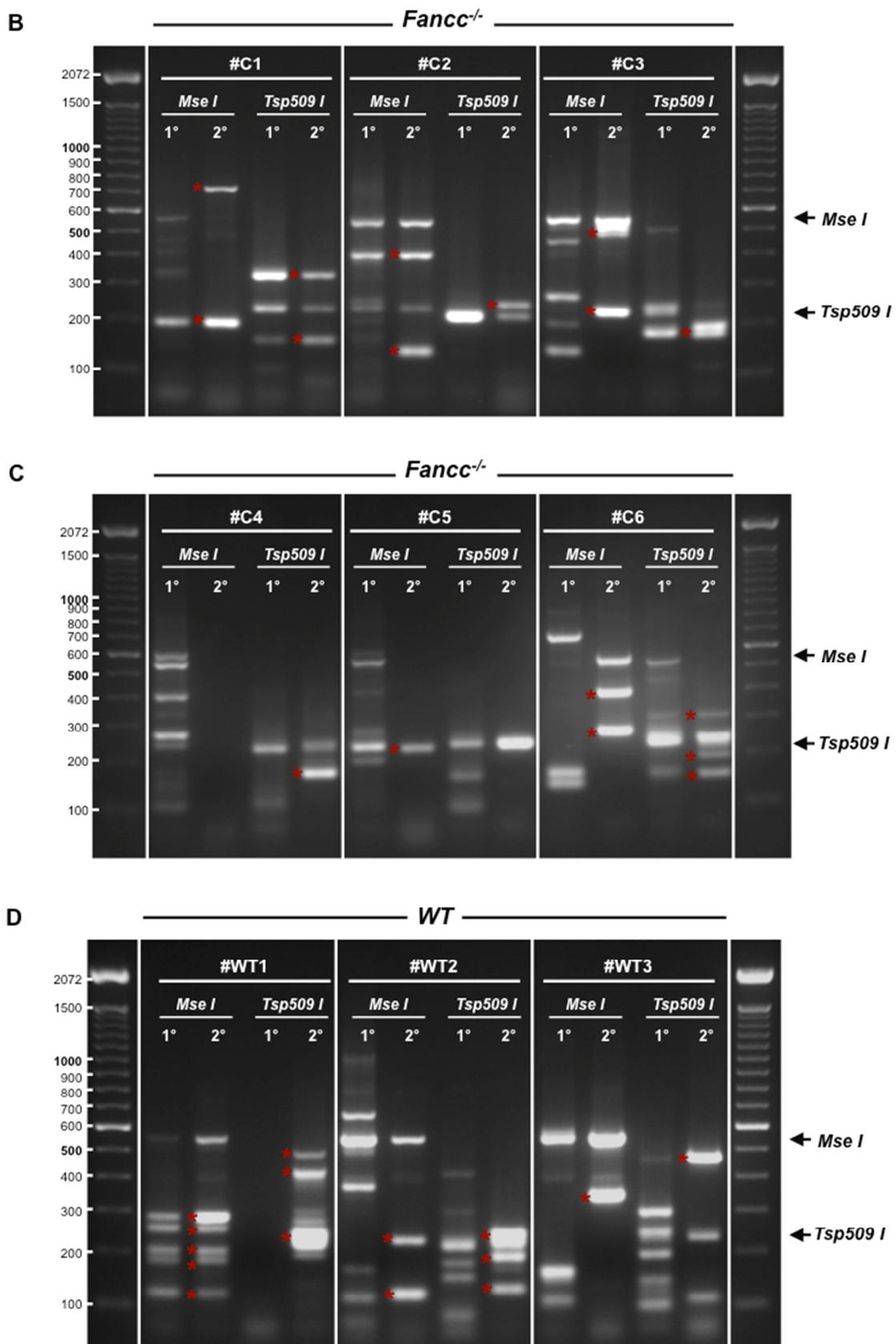


Figure 5: Clonality of FA and WT bone marrow samples. LM-PCR gels showing insertions recovered from bone marrow of primary (1°) and secondary (2°) recipients after MseI and Tsp509I restriction digest. **(A)** Analysis of *Fancc*⁺ (#A1, #A2) and *Fancc*⁻ bone marrow (#G1) **(B)+(C)** Analysis of *Fancc*⁻ bone marrow (#C1, #C2, #C3, #C4, #C5, #C6) **(D)** WT bone marrow #WT1, #WT2 and #WT3. Arrows indicate internal vector control bands for MseI (544 bp) and Tsp509I (217 bp); asterisks indicate retroviral insertions corresponding to dominant clones.

Altogether these data suggest the selective expansion of one or two dominant clones after two rounds of transplantation demonstrating the dominance of some of the retroviral-transduced FA HSC clones. These mice revealed dominant stem cell clones that had gained a growth advantage as a result of the genomic integration of the retroviral vector. In addition, we could demonstrate non-malignant clonal expansion of the dominant stem cell clones as we could not observe any signs of leukemogenesis (splenomegaly and increased peripheral blood cell counts) in these mice. Although we could detect abnormal hematopoiesis in clone #A1, which demonstrated a profound myeloid differentiation bias (**Figure 6 A+B**), we could exclude the existence of increased blasts or myelodysplasia as assessed by analysis of cellular morphology in peripheral blood smears (**Figure 6 C**) and bone marrow cytopsin (**Figure 6 D**).

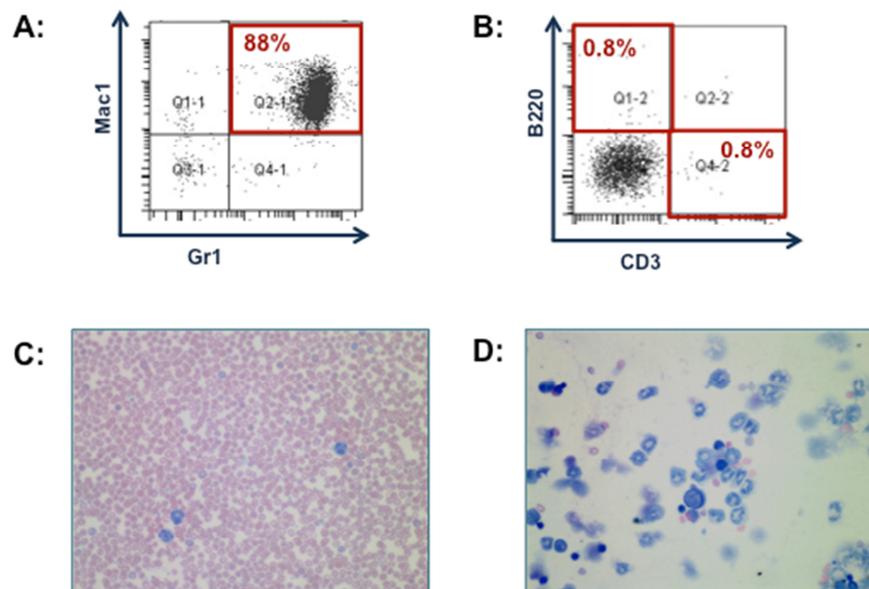


Figure 6: Abnormal hematopoiesis in mouse #A1 with a profound myeloid differentiation bias. **(A)** FACS analysis of the peripheral blood in mouse #A1 demonstrating an overrepresentation of Mac1⁺/Gr1⁺ myeloid cells (88%). **(B)** FACS analysis of the lymphoid lineage revealing decreased numbers of B (B220⁺; 0.8%) and T (CD3⁺; 0.8%) cells. **(C)** Peripheral blood smear revealing normal blood morphology. **(D)** Bone marrow cytopsin representing all stages of differentiating myeloid cells.

3.5 Identification of retroviral integration sites in dominant FA and WT HSC clones

Since the peripheral blood samples of serially transplanted mice had demonstrated dominance of one or two hematopoietic FA HSC clones, we next analyzing the retroviral vector integration sites present in these clones. For this purpose we have used nrLAM-PCR (Gabriel et al., 2009; Paruzynski et al., 2010) in collaboration with the group of Manfred Schmidt at the National Center for Tumor Diseases in Heidelberg. Genomic DNA was isolated from the bone marrow of these mice in our laboratory and the samples were handed over to Cynthia Bartholomä, who performed the nrLAM-PCR analysis together with the technical support of Stefanie Laier.

The nrLAM-PCR is an advanced form of integration site analysis, which has been used in more recent studies as it has proven to be the most comprehensive technology available for the recovery of the genomic location of proviral integrants (Schmidt et al., 2007). The nrLAM-PCR exhibits lower sensitivity due to the less efficient ligation reaction of single stranded DNA but circumvents the often-criticized restriction bias introduced by LM-PCR, as it allows for the identification of vector integrations independent of restriction enzymes (see **Figure 39** in 5.9.2).

The results from the unbiased insertion site characterization of primary bone marrow samples recovered from mice that have dominated in serial transplantations are depicted in **Figure 7**. Each bar displays the retrieval frequencies of the most prominent integration sites in each sample illustrating the relative contribution of each individual integration site to the whole genome in each mouse (**Figure 7 A**). In mice #A1, #G1, C2, #C5 and #WT3, we only found one dominant integration site, in which the integration of the retrovirus had led to a growth advantage in a single HSC clone. However, in mouse #A2 we identified two dominant integration sites, which equally contributed to hematopoiesis. This could correspond to a single dominant HSC clone that harbored two independent proviral integration sites or two dominant clones, each with a single integration site and each contributing an approximately equal amount to ongoing hematopoiesis. In contrast, all other mice revealed oligoclonal hematopoietic reconstitution. Altogether, this analysis clearly demonstrates a mono- to oligoclonal hematopoietic reconstitution in mice, fully recapitulating what was already observed in the LM-PCR analysis (see **Figure 5** in 3.4).

Deep-sequencing analysis allowed for the identification of the unknown genomic sequences flanking the integrated vector DNA and revealed the closest gene for each integration site as shown in **Figure 7 B**. While this revealed up to 15 and 14 vector integrations in mice #C6 and #WT2 respectively, we could only detect a single integration in mouse #C5 with all the others ranging from three to ten. However, most of these integrations corresponded to minor clones as confirmed by the

analytical LM-PCR gels (see **Figure 5** in 3.4). The integration sites on top of the list are the genes proximal to the integration site in the dominant clones.

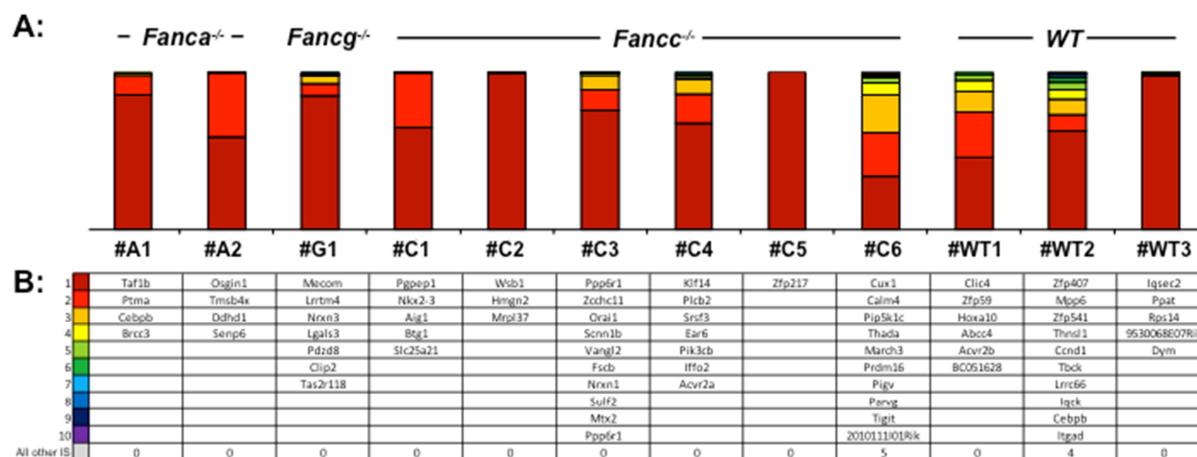


Figure 7: Insertion site characterization of FA and WT HSCs. (A) Retrieval frequencies of the most prominent integration sites in the bone marrow sample of selected mice (#) as analyzed by nrLAM-PCR. (B) Table listing the closest gene to each integration site in each mouse.

Although one could interrogate all of these mice we decided to focus our subsequent analysis on the FA clones termed #A1, #A2 and #G1, which demonstrated an extremely strong selection pressure. We hypothesized that the expansion of the gene-modified graft in these mice was due to the deregulated expression of genes that allowed the HSC to overcome the FA-specific selection pressure in order to become a dominant stem cell clone. Therefore the characterization of these dominant stem cell clones is of major interest as they would reveal target genes with the potential to compensate for the inherent FA HSC defect.

3.6 Genomic location of the retroviral integrations in mice #A1, #A2 and #G1

To further characterize the dominant FA HSC clones of mice #A1, #A2 and #G1, which demonstrated an extremely strong selection pressure, and to identify the genes proximal to the retroviral vector integrations, we next investigated their genomic locations. **Figure 8** depicts an overview of the genomic location of the retroviral integrations in each individual clone, demonstrating the complexity of the surroundings of the retroviral integrations. Since some of the integrations have occurred in gene-dense regions and retroviral enhancers can elicit their effects over large distances, we considered windows of 300 kb for the analysis of genes that are proximal to the integration site.

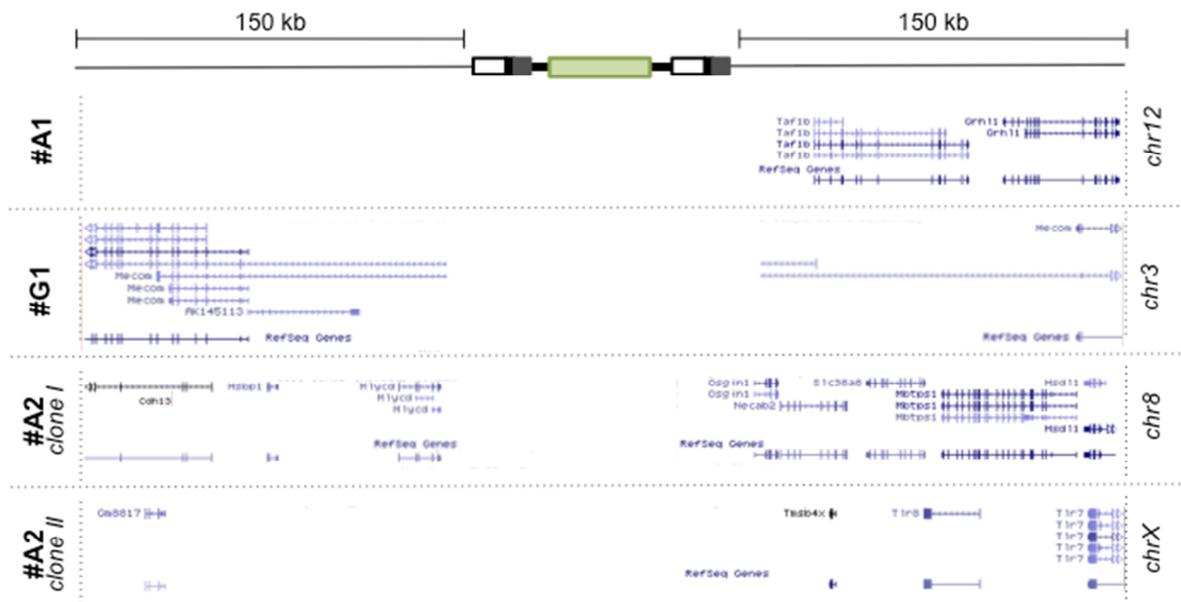


Figure 8: Genomic location of the retroviral vector integration in each individual stem cell clone. Integrated version of the retroviral vector with flanking genes 150 kb downstream and upstream of the vector integration for the dominant stem cell clones #A1, #G1, #A2 I and II.

In clone #A1, nrLAM-PCR revealed an integration of the retrovirus into chromosome 12, upstream of the TATA box binding protein-associated factor, ribonucleic acid (RNA) polymerase I, B (*Taf1b*) and the grainyhead-like 1 (*Grhl1*), a gene named after the “grainyhead-like” drosophila mutant. In clone #G1, the retroviral vector had integrated into the Evi1/MDS complex locus (*Mecom*), which is located on chromosome 3. In the two clones of mouse #A2, we found a total of 12 potential drivers of clonal dominance upstream and downstream of the two retroviral integration sites. One integration had occurred on chromosome 8, into a gene locus with eight genes in the window of 300 kb and the other was identified on the X-chromosome with four additional genes in the surroundings of the provirus. All genes, which were found proximal to the retroviral integration, are summarized in Table 2.

Table 2: Potential candidate target genes

Clone ID	Candidate target gene	
#A1	Taf1b	TATA box binding protein (Tbp) associated factor, RNA polymerase I, B
#A1	Grhl1	grainyhead-like 1
#A2	Cdh13	cadherin 13
#A2	Hsdp1	heat shock factor binding protein 1
#A2	Mlycd	malonyl-CoA decarboxylase
#A2	Osgin1	oxidative stress induced growth inhibitor 1
#A2	Necab2	N-terminal EF-hand calcium binding protein 2
#A2	Slc38a8	solute carrier family 38, member 8
#A2	Mbtps1	membrane-bound transcription factor peptidase, site 1
#A2	Hsd11	hydroxysteroid dehydrogenase like 1
#A2	Gm8817	predicted gene 8817
#A2	Tmsb4x	thymosin, beta 4, X chromosome
#A2	Tlr8	toll-like receptor 8
#A2	Tlr7	toll-like receptor 7
#G1	Mecom	MDS /Evi1 complex locus

3.7 Identification of genes that drive clonal dominance

To identify genes with the potential to alter FA stem cell kinetics and drive clonal dominance among the genes located 150 kb upstream and downstream of the retroviral integration as depicted in Table 2, we sought to identify genes that demonstrated deregulated expression in whole bone marrow samples from the individual mice termed #G1 and #A2 using quantitative RT-PCR. The results from this analysis are depicted in **Figure 15** and **Figure 20** and described in more detail in the subsequent chapters (refer to 3.9 and 3.10). For mouse #A1, we were not able to analyze the expression profiles as we did not have backup cell pellets available for RNA isolation. Fortunately, only two candidate genes, Taf1b and Grhl1, were located downstream of the integration site in mouse #A1 and therefore were immediately selected for further functional characterization.

3.8 Characterization of genes proximal to the integration site in mouse #A1

In mouse #A1 we have identified two genes, Taf1b and Grhl1, downstream of the retroviral integration on chromosome 12. Due to the lack of sufficient material to perform expression analysis, both genes were further characterized and are described in more detail in the subsequent chapters.

3.8.1 TATA box binding protein (TBP)-associated factor, RNA polymerase I, B (Taf1b)

In bone marrow of mouse #A1, we found Taf1b, the TATA box binding protein-associated factor, RNA polymerase I, B in close proximity to the retroviral vector integration on chromosome 12.

Taf1b is a component of the transcription machinery for RNA polymerase I, which initiates ribosome biogenesis and regulates eukaryotic cell growth and proliferation (Laferte et al., 2006). Among the three types of eukaryotic nuclear RNA polymerases that transcribe DNA into different forms of RNA, Polymerase I is responsible for the synthesis of most ribosomal RNAs. For the formation of the pre-initiation complex at the promoter, RNA polymerase I requires general transcription factors that assist in positioning of the RNA polymerase I to ribosomal DNA in order to initiate transcription: the upstream binding factor (UBF) and the promoter selectivity factor SL1 in humans and TIF-IB in mice (Grummt, 2003). TIF-IB/SL1 is a complex that consists of the TATA binding protein (TBP) and five TATA binding protein-associated factors (Taf1s), of which one is the 68kDa protein Taf1b (also termed Taf1-68) in mice. In general, Taf1 proteins play an integral role in the assembly of the transcription complex by mediating the recruitment of polymerase I to ribosomal DNA thereby contributing to promoter specificity and gene regulation (Heix et al., 1997; Zomerdijk et al., 1994).

Besides the generally established link between ribosomal (r)RNA synthesis and cancer (Drygin et al., 2010), Taf1b has been implicated in the pathology of colorectal carcinoma as somatic frameshift mutations in Taf1b were identified in 82% of all analyzed mismatch repair-deficient tumors (Kim et al., 2002). In addition, a Taf1b truncating mutation has been associated with lung cancer risk. However, the germ line DNA of only one patient was analyzed in this study (Renieri et al., 2014).

The importance of Taf1b for stem cell biology has just recently been demonstrated as the disruption of Taf1b by RNA interference, leading to decreased rRNA synthesis, resulted in reduced proliferation rates of ovarian germ line stem cells in *Drosophila* (Zhang et al., 2014). Therefore, Taf1b is an important component of the intrinsic transcription machinery that has the potential to modulate the cell fate, growth and proliferation of *Drosophila* stem cells. However, the role of Taf1b in the hematopoietic stem and progenitor compartment so far remains elusive. Therefore, we aimed at characterizing Taf1b in the context of HSC biology by summarizing different expression data of hematopoietic stem and progenitor cells.

3.8.1.1 Taf1b expression levels under homeostatic conditions

In order to get an initial idea about the relative expression levels of Taf1b in different tissues, we explored the BioGPS gene annotation portal (www.biogps.org), which provides a huge collection of gene expression data on a variety of biological samples. These data identified murine Taf1b to be highest expressed in testis and the mouse pro B cell line Baf3 (**Figure 9**). In addition, also the hematopoietic compartment seemed to express Taf1b.

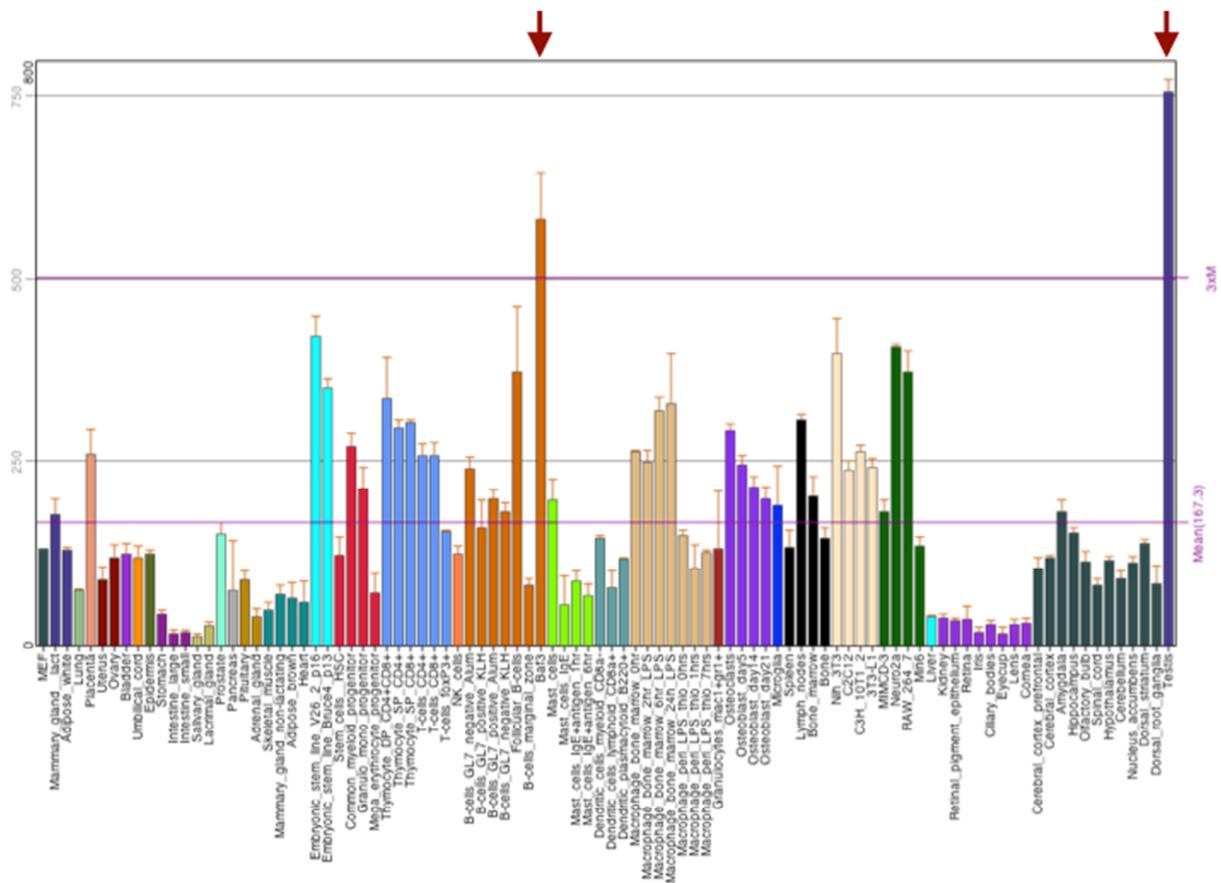


Figure 9: Taf1b expression across different tissues and cell lines. Red arrows indicate pronounced expression in the mouse pro B cell line BaF3 cells and in testis.

The role of Taf1b in the hematopoietic stem and progenitor compartment has not been characterized yet. In order to provide some insight into the relevance of Taf1b for HSC biology, we made use of a comprehensive gene expression dataset on the hematopoietic populations that was generated by high-throughput RNA sequencing through the collaborative effort of the groups of Andreas Trumpp, Michael Milsom and Daniel Lipka at the German Cancer Research Center and the members of the group of Jerome Krijgsfeld at the European Molecular Biology Laboratories in Heidelberg (Cabezas-Wallscheid et al., 2014). These data allowed assessing the Taf1b expression in HSC and MPP populations, which revealed that Taf1b was equally expressed across HSCs and a series of multipotent progenitor cells (**Figure 10 A**).

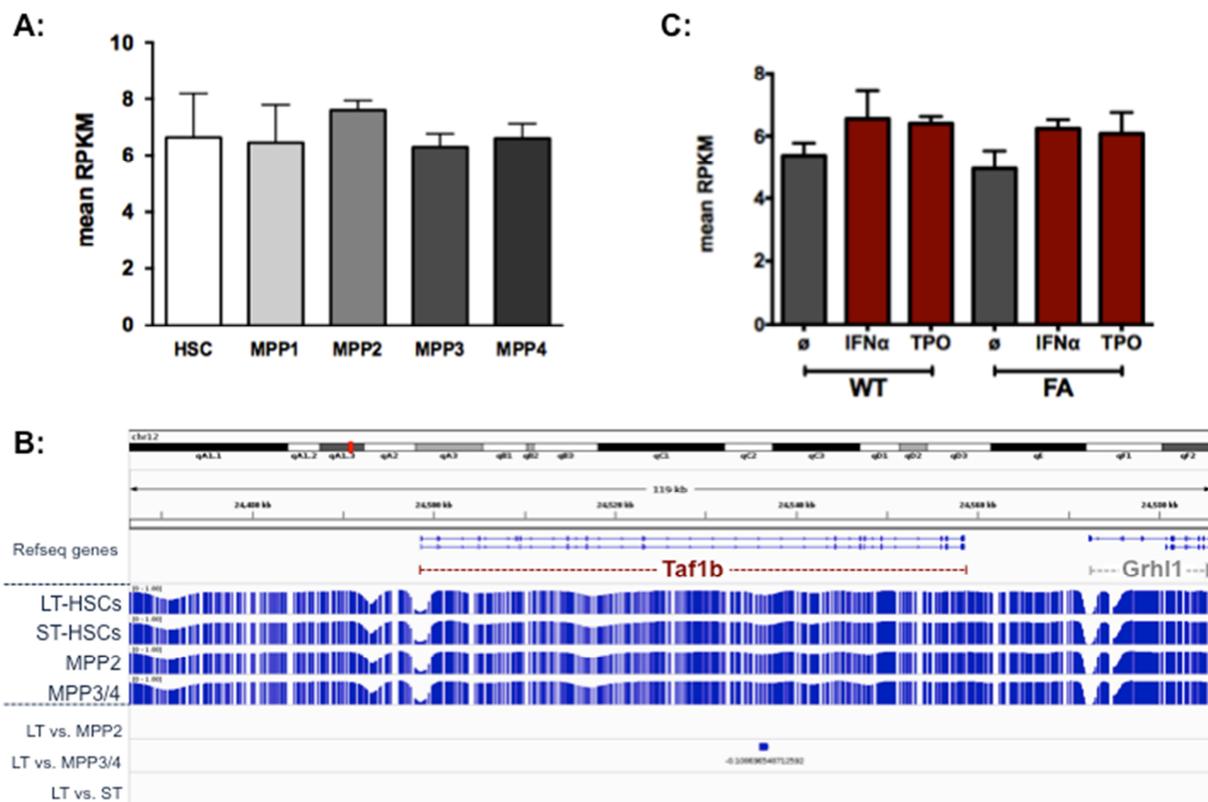


Figure 10: Taf1b expression levels and methylation pattern in HSC and progenitor populations. (A) Mean Taf1b expression levels in LT-HSCs and progenitor populations MPP1 – 4 as obtained by high-throughput RNA sequencing. (B) Taf1b methylation pattern in LT-HSCs and MPP1 - 4 generated by tagmentation-based whole genome bisulfite sequencing. (C) Mean Taf1b expression levels in FA and WT LT-HSCs under physiologic stress conditions as obtained by high-throughput RNA sequencing of IFN- α , Tpo- and control-treated LT-HSCs.

3.8.1.2 Epigenetic regulation of Taf1b expression

Gene expression can be regulated through various mechanisms. One important mechanism involves the epigenetic regulation of chromatin structure through DNA and histone modifications. DNA methylation is a stable epigenetic mark, which in mammalian cells preferentially occurs at cytosine-residues adjacent to guanine nucleotides, so called CpG-dinucleotides. CpG methylation is mediated by DNA methyltransferases. The genome-wide analysis of DNA methylation patterns provides an important measure of transcriptional regulation during differentiation (Cabezas-Wallscheid et al., 2014; Lipka et al., 2014). Therefore we were interested in the analysis of the methylation patterns at the Taf1b locus, on the long arm of chromosome 12 in the sub band A1.3, across the individual hematopoietic stem and progenitor cell populations as it was addressed by a dataset that was generated by tagmentation-based whole genome bisulfite sequencing of these hematopoietic cells by Daniel Lipka and colleagues (Lipka et al., 2014). Using this data, we found a distinct differentially

methyated region among the LT-HSC and MPP3/4 populations, which was located in an intragenic region of *Taf1b* (**Figure 10 B**). This reveals a functional region, which may be involved in the epigenetic regulation of HSC biology during hematopoiesis.

3.8.1.3 *Taf1b* expression under conditions of physiologic stress

Since we could demonstrate that *Taf1b* was expressed in HSC and MPPs under steady state conditions, we also investigated the *Taf1b* expression under conditions of physiologic stress such as in response to bleeding, infections or inflammation. HSCs are rapidly activated in response to stress and proliferate in order to replenish the system (Cheshier et al., 2007; Wilson et al., 2008). Such an activation of the HSC compartment can also be induced experimentally with agonists that mimic a stress situation like interferons, lipopolysaccharide (LPS), G-CSF or Tpo (Baldrige et al., 2010; Essers et al., 2009; Takizawa et al., 2011; Wright et al., 2001; Yoshihara et al., 2007; Zhang et al., 2008). Recently, we could show that treatment of HSCs with these agonists resulted in an increased production of intracellular reactive oxygen species and the accumulation of DNA damage, which may ultimately lead to cellular attrition (Walter et al., 2015). This demonstrates the extent of the stress the HSC has to cope with in order to maintain the integrity of the hematopoietic system. Here, IFN- α and Tpo were chosen to activate the HSCs on two consecutive days for the purpose of modeling HSC stress biology *in vivo*. In order to analyze the gene expression of highly purified LT-HSCs (Lin^- , c-Kit^+ , Sca-1^+ , CD150^+ , CD48^+ , CD34^-) under stress conditions, Amelie Lier performed high-throughput RNA sequencing of Tpo- and IFN- α -treated LT-HSC populations and compared normal HSC biology as well as FA HSC biology to the respective stress conditions (Lier and Milsom, unpublished data). Under stress conditions *Taf1b* was slightly upregulated in both WT and FA HSCs, independent of whether Tpo or IFN- α treatment induced the stress (**Figure 10 C**). This suggests that *Taf1b* might be required for the maintenance of hematopoiesis in response to stress, both in WT and FA mice.

3.8.1.4 *Taf1b* expression during development

In order to reveal whether *Taf1b* could also play a role during embryonic development of the hematopoietic stem and progenitor cell compartment, we investigated the expression levels of *Taf1b*. For this purpose, we made use of a comprehensive gene expression dataset on HoxB4-yellow fluorescent protein (YFP)-positive embryonic stem (ES) reporter cell lines, which was generated by our colleague Paul Kaschutnig. These HoxB4-YFP-positive ES cells were isolated from the *Hoxb4*-YFP reporter mouse model (Hills et al., 2011), in which the YFP-labeled HoxB4 expression is restricted to the definitive HSCs. During ES cell culture, YFP expression marks bona fide HSCs and therefore facilitates the purification and analysis of this population. In the microarray, three

populations from primitive (HoxB4⁻ CD41⁻ Flk1⁺ c-Kit⁺) to definitive/HSC-enriched (HoxB4⁺ CD41⁺ CD93⁺ c-Kit⁺) to more progenitor-like cells (HoxB4⁻ CD41⁺ c-Kit⁺) were analyzed (Kaschutnig and Milsom, unpublished data) and revealed no significant differences in the mean expression levels of Taf1b expression (Figure 11) This indicates that differential Taf1b levels might not be relevant during these individual steps of specification. However, whether uniformly expressed Taf1b is required in these rapidly proliferating cell populations or does not play a role during developmental processes at all remains to be investigated.

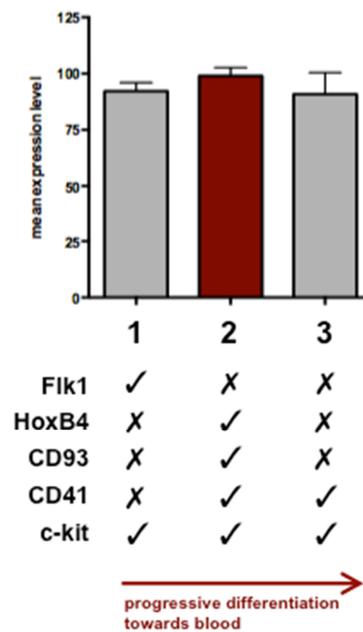


Figure 11: Taf1b expression during embryonic development. Mean expression levels of Taf1b in HoxB4-YFP-expressing embryonic stem cells during developmental specification from (1) primitive (HoxB4⁻ CD41⁻ Flk1⁺ c-Kit⁺) to (2) definitive/HSC-enriched (HoxB4⁺ CD41⁺ CD93⁺ c-Kit⁺) to more (3) progenitor-like cells (HoxB4⁻ CD41⁺ c-Kit⁺).

3.8.2 Grainyhead-like 1 (Grhl1)

We also found Grhl1 as a potential target of the vector integration in the dominant clone from mouse #A1, approximately 80 kb downstream of the retroviral integration on chromosome 12. The transcription start site of Grhl1 is separated from the vector integration by Taf1b. Despite the distant location of the transcription start site of Grhl1, it is still possible that the integration site has impacted upon Grhl1 expression due to the very strong enhancer elements of the SF91 vector, which may impact upon the gene expression of genes much further apart. Therefore, we aimed at characterizing Grhl1 in the context of HSC biology.

Grhl1 is one of three mammalian homologues of the drosophila transcription factor grainyhead (grh), which is expressed in ectoderm-derived tissues during drosophila development (Bray et al., 1989; Bray and Kafatos, 1991). Mammalian Grhl1 shares high sequence homology with Grhl2 and -3 functional domains and is expressed in the epidermis as well as in heart, lung and kidney of the developing mouse embryo (Auden et al., 2006). Mice lacking Grhl1 have demonstrated an important regulatory function of Grhl1 for desmosomal cadherin expression in the epidermis, which appears crucial for hair anchoring and epidermal differentiation (Wilanowski et al., 2008). Therefore, Grhl1 knockout mice exhibit skin barrier defects and are more susceptible to the development of skin tumors such as squamous cell carcinoma (Mlacki et al., 2014).

In neuroblastoma, Grhl1 was found to act as a tumor suppressor with high-level expression correlating with favorable patient survival. In addition, Grhl1 expression was induced upon HDAC inhibitor treatment in neuroblastoma cells, in which MYCN binding to the Grhl1 promoter had repressed its expression. This suggests a therapeutic benefit of HDAC inhibitor treatment provoking Grhl1 expression for neuroblastomagenesis at the molecular level (Fabian et al., 2014).

A role for Grhl1 in HSC biology has not been demonstrated yet. Therefore, we aimed at summarizing comprehensive expression data on Grhl1 in different compartments of the hematopoietic system in order to investigate whether Grhl1 might play a role in hematopoiesis.

3.8.2.1 Grhl1 expression levels under homeostatic conditions

To elucidate the relative expression levels of Grhl1 among a variety of biological samples including cell lines and tissues, we explored the collection of gene expression data published by the BioGPS gene annotation portal (www.biogps.org). Grhl1 was found essentially expressed in tissue from epidermis and prostate, and to a somewhat lower extent in cornea and mammary gland (**Figure 12**). In contrast, Grhl1 expression was undetectable in bone marrow as well as in individual hematopoietic populations such as mature T-, B-, myeloid and progenitor cells.

In order to investigate the Grhl1 expression levels in the hematopoietic stem and progenitor compartment we made use of high-throughput sequencing data that was generated by our colleagues and collaborators (Cabezas-Wallscheid et al., 2014). In general, the overall expression levels for Grhl1 in the HSC and MPP compartment appeared to be below 1 read per kb of transcript per million reads mapped (RPKM), which might indicate that only very few transcripts of Grhl1 were present. Independent of this finding, Grhl1 expression was increased in MPP1 cells when compared to HSCs, and decreased upon hematopoiesis from MPP2 to MPP3 and MPP4 (**Figure 13 A**). However, whether this supports a regulatory role of Grhl1 during hematopoiesis remains to be clarified.

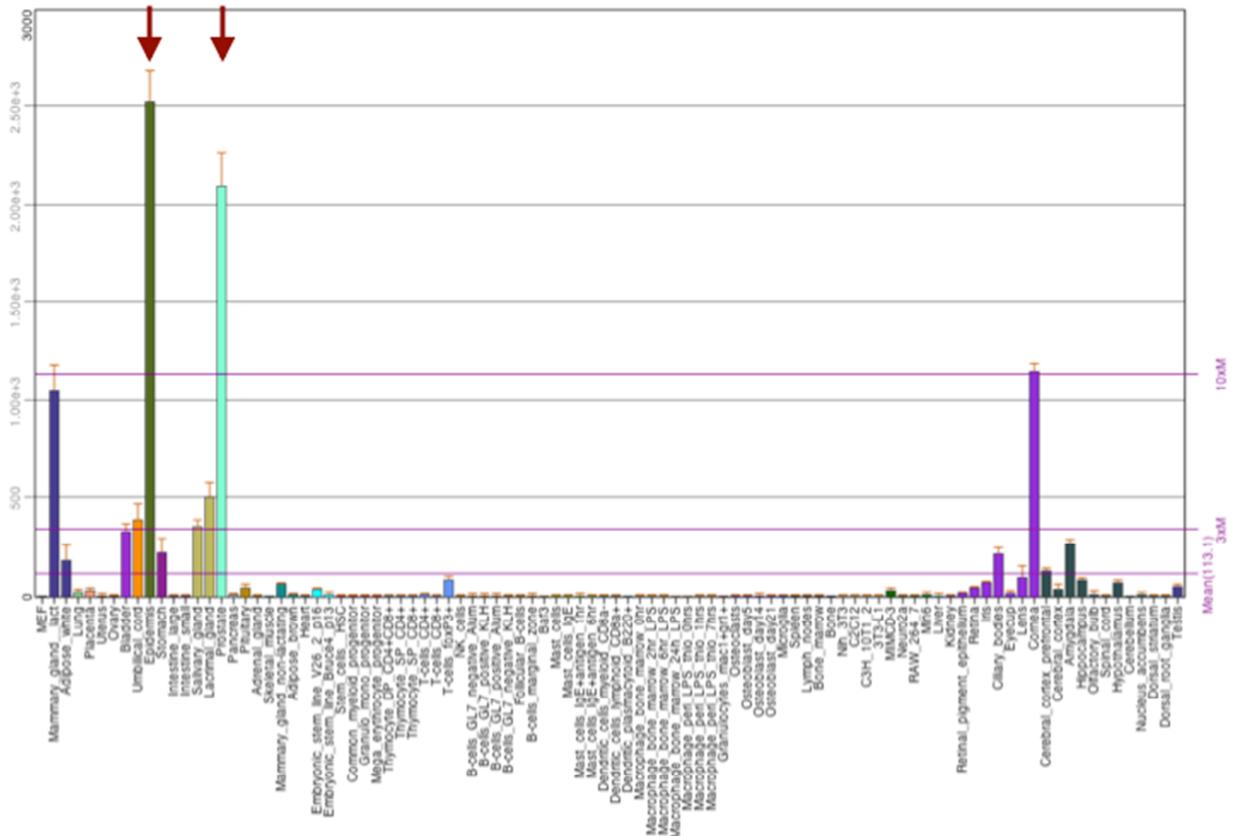


Figure 12: Grhl1 expression across different tissues and cell lines. Red arrows indicate pronounced expression of Grhl1 in epidermis and prostate.

3.8.2.2 Epigenetic regulation of Grhl1 expression

Next, we aimed at investigating the transcriptional regulation of Grhl1 during HSC differentiation and looked at the DNA methylation pattern of the Grhl1 locus as it was addressed by a dataset that was generated by tagmentation-based whole genome bisulfite sequencing by the collaborative effort of Daniel Lipka and members of our group (Cabezas-Wallscheid et al., 2014; Lipka et al., 2014). Using this data, we were able to assess the methylation status of the Grhl1 locus, on the long arm of chromosome 12 in the sub band qA1.3, which revealed no difference across the individual hematopoietic stem and progenitor cells (**Figure 13 B**). Therefore, Grhl1 expression during hematopoiesis seems to be independent of epigenetic regulation. However, the permissive methylation state for gene expression at the hypomethylated upstream promoter in HSCs and progenitors, led us to speculate that Grhl1 expression can be turned on under certain circumstances.

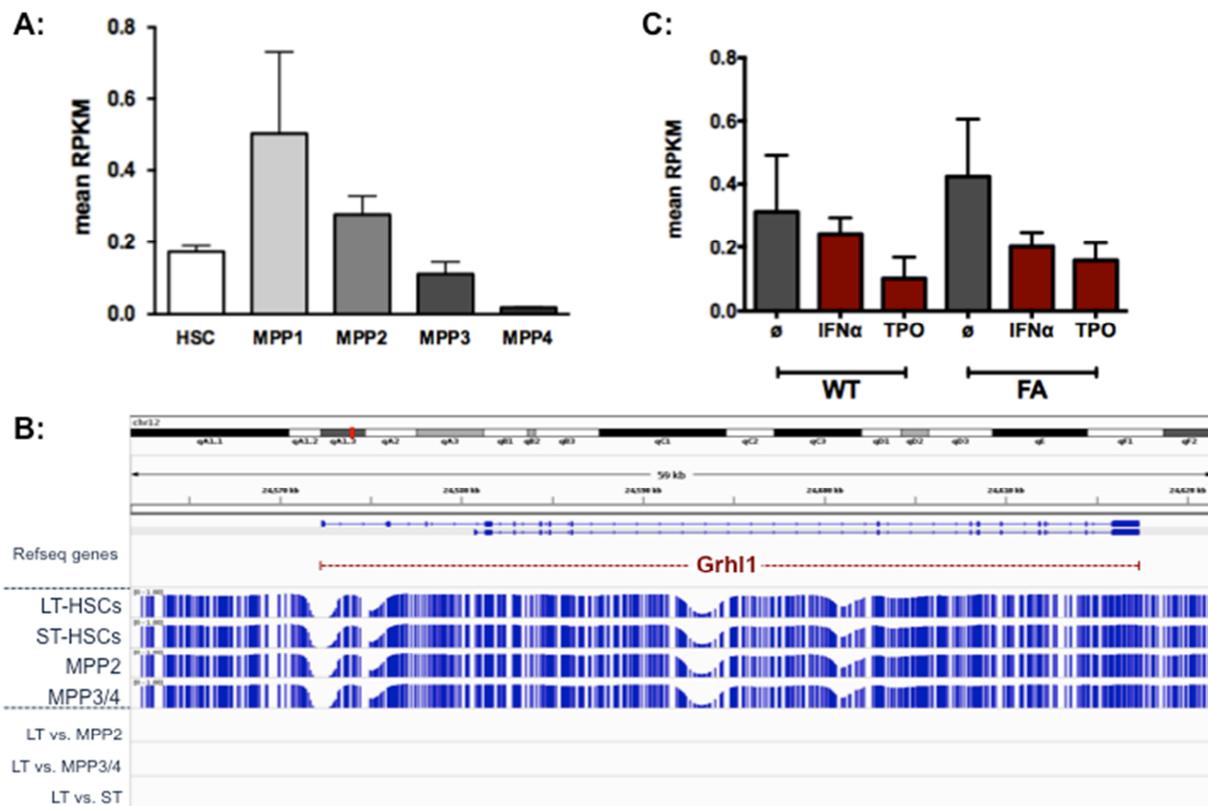


Figure 13: Grhl1 expression levels and methylation pattern in HSC and progenitor populations. (A) Mean Grhl1 expression levels in LT-HSCs and progenitor populations MPP1 – 4 as obtained by high-throughput RNA sequencing. (B) Grhl1 methylation pattern in LT-HSCs and MPP1 - 4 generated by tagmentation-based whole genome bisulfite sequencing. (C) Mean Grhl1 expression levels in FA and WT LT-HSCs under physiologic stress conditions as obtained by high-throughput RNA sequencing of IFN- α , Tpo- and control-treated LT-HSCs.

3.8.2.3 Grhl1 expression under conditions of physiologic stress

As outlined in chapter X, conditions of physiologic stress such as in response to bleeding, infections or inflammation may result in the loss of mature blood cells and in the activation of HSCs. To investigate Grhl1 expression of highly purified LT-HSCs (Lin⁻, c-Kit⁺, Sca-1⁺, CD150⁺, CD48⁻, CD34⁻) in response to stress, we experimentally induced HSC activation *in vivo* using Tpo and IFN- α treatment (Baldrige et al., 2010; Essers et al., 2009; Yoshihara et al., 2007) and performed high-throughput RNA sequencing comparing normal HSC biology as well as FA HSC biology to the respective stress conditions (Lier and Milsom, unpublished data).

This revealed that under conditions of physiologic stress as induced by Tpo- or IFN- α treatment, Grhl1 expression was slightly downregulated in WT and FA HSCs (Figure 13 C), which indicates a possible role for Grhl1 in the stress biology of normal as well as FA HSCs. In WT HSCs Tpo-treatment induced a slightly stronger downregulation of Grhl1 expression than IFN- α treatment,

which might hint towards a different role for Grhl1 in response to Tpo-induced stress. However, this remains to be investigated.

3.8.2.4 *Grhl1* expression during development

In order to investigate the *Grhl1* expression levels during HSC development, we made use of HoxB4-YFP-positive ES reporter cell lines, which were derived from the *Hoxb4*-YFP reporter mouse model (Hills et al., 2011) and analyzed for their gene expression profiles using microarray (Kaschutnig and Milsom, unpublished data). As the YFP-labeled HoxB4 expression is restricted to the definitive HSCs in these mice, during ES cell culture YFP expression marks developing HSCs. As before, three populations corresponding to mesodermal precursors of hematopoietic stem and progenitor hemogenic endothelium, HOXB4⁺ hematopoietic stem and progenitor hemogenic endothelium and downstream hematopoietic progeny of HOXB4⁺ cells were analyzed in the microarray. All three populations revealed no major differences in the mean expression levels of *Grhl1* (Figure 14), which could have two implications. It could either mean that *Grhl1* does not play a role during developmental processes of hematopoietic populations or that hematopoietic development requires *Grhl1* expression independent of differential expression levels.

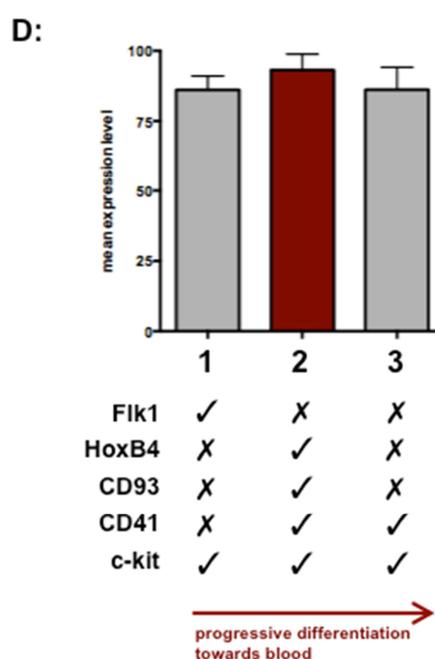


Figure 14: *Grhl1* expression during embryonic development. Mean expression levels of *Grhl1* in HoxB4-YFP-expressing embryonic stem cells during developmental specification from (1) primitive (HoxB4⁻ CD41⁻ Flk1⁺ c-Kit⁺) to (2) definitive/HSC-enriched (HoxB4⁺ CD41⁺ CD93⁺ c-Kit⁺) to more (3) progenitor-like cells (HoxB4⁻ CD41⁺ c-Kit⁺).

3.9 Identification and characterization of deregulated genes in mouse #G1

In clone #G1, we have identified the retroviral vector integration within the MDS1-EVI1 complex (Mecom), which is located on chromosome 3. This genetic locus harbors two distinct transcription start sites located almost 500 kb apart giving rise to the two transcript variants EVI1 (downstream) and MDS1 (upstream) (Nucifora et al., 1994).

To confirm that deregulated expression of one of the Mecom transcript variants was responsible for the emergence of the dominant stem cell clone #G1, we performed gene expression analysis in whole bone marrow samples from mouse #G1 using quantitative RT-PCR (Figure 15). In this respect, quantitative RT-PCR analysis confirmed the aberrant expression of the Mecom transcript variant I (EVI1) by 60-fold upregulation in comparison to unmodified *Fancg*^{-/-} bone marrow as a result of the integration into the MDS1-EVI1 complex locus. Therefore, we have identified *Evi1* as a potential driver of clonal dominance in the context of FA stem cell biology.

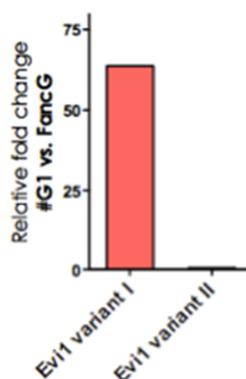


Figure 15: Aberrant expression of *Evi1* transcript variant I in the dominant stem cell clone #G1. Relative mRNA expression levels of *Evi1* transcript variants I and II in bone marrow isolated from mouse #G1. *Fancg*^{-/-} bone marrow served as control.

3.9.1 The ecotropic virus integration site 1 (*Evi1*)

EVI1, also known as the ecotropic virus integration site 1, is a transcriptional regulator that was initially identified as a site of proviral insertion in murine myeloid leukemias (Morishita et al., 1988) and is one of the most commonly identified drivers of clonal dominance, both in mouse experiments and in patients taking part in clinical gene therapy trials (Kustikova et al., 2005; Ott et al., 2006). Therefore, integration of the retroviral vector into this genetic locus can be interpreted as a positive control, which confirms the reliability of our study design.

Evi1 is a master regulator of hematopoiesis with a suggested function in myeloid differentiation, apoptosis, HSC quiescence and proliferation (Glass et al., 2014). However, the underlying molecular mechanisms still remain unclear and additional studies will be required.

The oncogenic function of Evi1 is well established and chromosomal rearrangements involving the human gene on chromosome 3q26 that may activate Evi1 expression, have been associated with MDS, a slowly developing clonal disorder that can progress to hematopoietic cancer of poor prognosis (Groschel et al., 2010; Lugthart et al., 2010). The role of Evi1 in the pathogenesis of leukemia was mostly addressed using mouse models transplanted with retroviral-transduced Evi1-expressing bone marrow cells. Overexpression of the gene product in murine models led to the development of MDS but was not sufficient to induce leukemia on its own (Buonamici et al., 2004; Cuenco and Ren, 2004) suggesting that additional cooperating hits are required for leukemic transformation (Haferlach et al., 2011; Maetzig et al., 2011; Modlich et al., 2005; Wolf et al., 2013). Since Evi1 was also frequently overexpressed in AMLs from FA patients, the genetic instability phenotype of FA patients may provide the necessary environment for leukemic transformation (Meyer et al., 2011; Meyer et al., 2007; Raynaud et al., 1996b). This suggests that EVI1 might confer a competitive advantage to FA cell clones. In addition, the pathologic functions of Evi1 are also relevant to the development of solid tumors as Evi1 expression was identified in epithelial cancers such as ovarian cancer, lung cancer and esophageal cancer (Brooks et al., 1996; Imoto et al., 2001; Jazaeri et al., 2010; Yokoi et al., 2003)

3.9.1.1 Evi1 expression levels under homeostatic conditions

In order to look at the pattern of Evi1 expression across different tissues and cell lines, we consulted the BioGPS gene annotation portal (www.biogps.org), which provides a huge collection of gene expression data on a variety of biological samples. In this data, we saw highest expression in the pluripotent mouse embryonic fibroblast cell line C3H10T1/2 (**Figure 16**). In addition, Evi1 expression was detectable in pancreas and bladder.

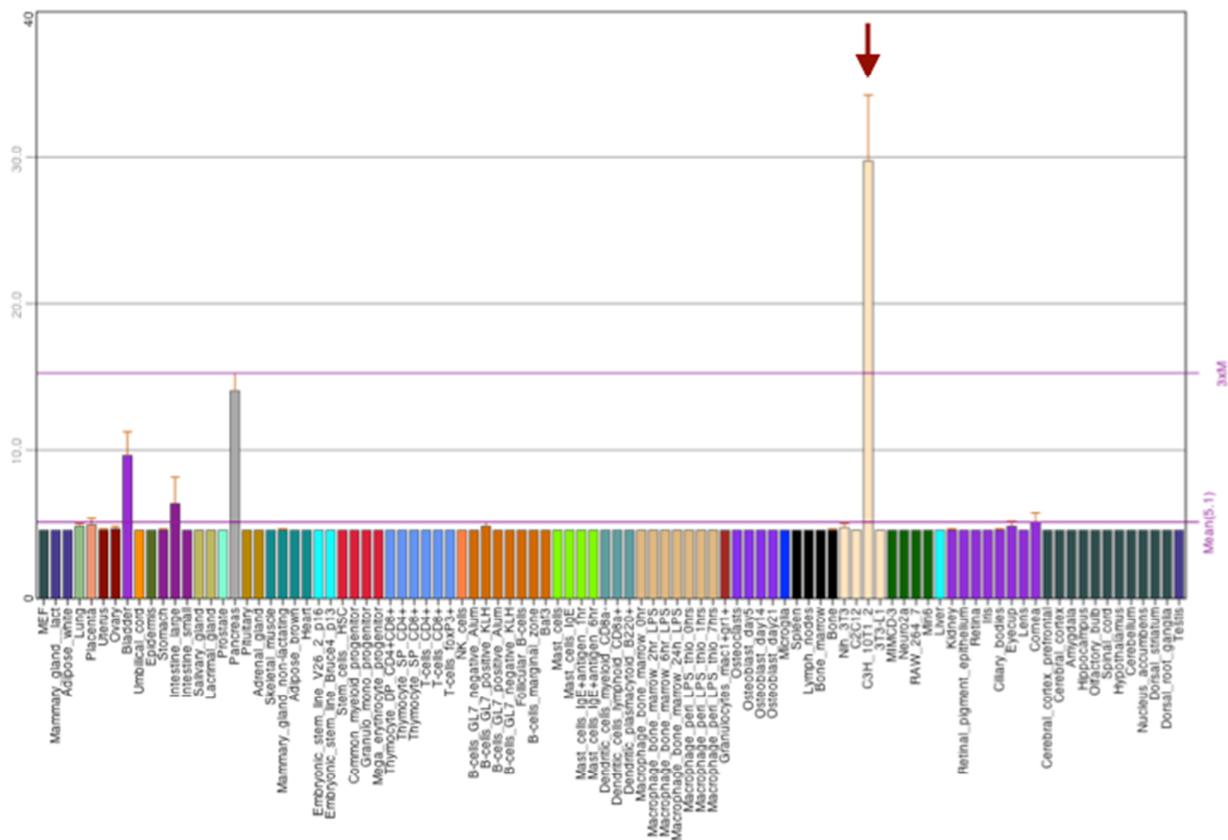


Figure 16: Evi1 expression across different tissues and cell lines. Red arrow indicates pronounced expression of Evi1 in the pluripotent mouse embryonic fibroblast cell line C3H10T1/2.

In the hematopoietic system, Evi1 is differentially expressed between HSCs and downstream MPPs as identified in gene expression data of hematopoietic populations, which was generated by high-throughput RNA sequencing through the collaborative effort of the groups of Andreas Trumpp, Michael Milsom and Daniel Lipka at the German Cancer Research Center and the members of the group of Jerome Krijgsfeld at the European Molecular Biology Laboratories in Heidelberg (Cabezas-Wallscheid et al., 2014). As illustrated in **Figure 17 A**, Evi1 presented a comparably high expression in the HSC compartment at the RNA level when compared to the progenitor cell populations. This suggests a functional role for Evi1 in HSC biology. In comparison to HSCs, Evi1 expression is about 2.1-fold decreased in the MPP1 population and about 4-fold in the in MPP2/3/4 populations. According to this expression pattern, Evi1 is likely to be involved in regulating HSC-specific processes such as HSC self-renewal, quiescence and differentiation.

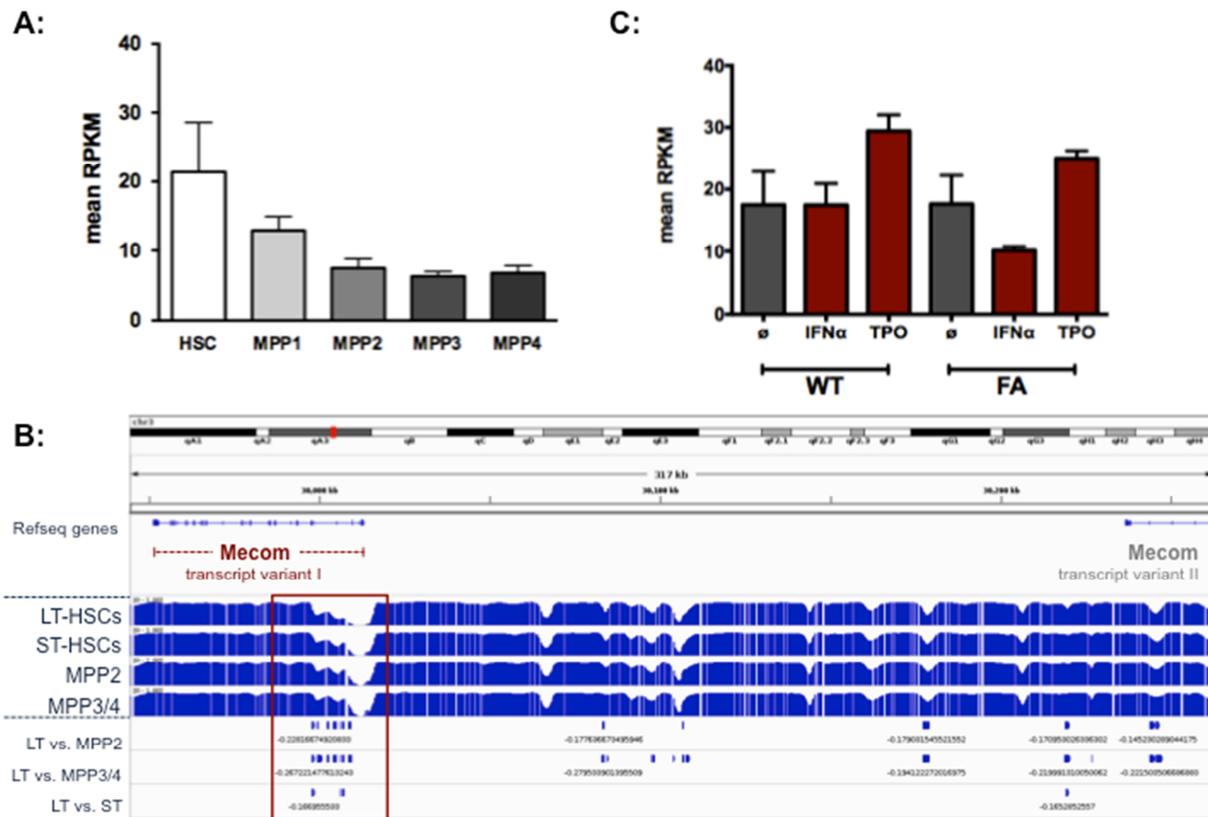


Figure 17: Evi1 expression levels and methylation pattern in HSC and progenitor populations. (A) Mean Evi1 expression levels in LT-HSCs and progenitor populations MPP1 – 4 as obtained by high-throughput RNA sequencing. (B) Evi1 methylation pattern in LT-HSCs and MPP1 - 4 generated by tagmentation-based whole genome bisulfite sequencing. Red box highlights the transcription start site of Evi1 transcript variant I with differential methylation pattern across HSC and MPP populations (C) Mean Evi1 expression levels in FA and WT LT-HSCs under physiologic stress conditions as obtained by high-throughput RNA sequencing of IFN- α , Tpo-and control-treated LT-HSCs.

3.9.1.2 Epigenetic regulation of Evi1 expression

In order to add further evidence for a role of Evi1 in regulating HSC differentiation, we investigated the epigenetic regulation of the chromatin structure at the Mecom locus, on the long arm of chromosome 3 in sub band A3, across the individual hematopoietic stem and progenitor cell populations. Therefore, we used a dataset on the methylation pattern of hematopoietic stem and progenitor cells, which was generated by Daniel Lipka using tagmentation-based whole genome bisulfite sequencing (Cabezas-Wallscheid et al., 2014; Lipka et al., 2014). This data revealed a relatively broad unmethylated region around the transcription start site of the Mecom transcript variant I (Figure 17 B), which might be relevant in regulating the transcription of Evi1 (highlighted by a red box). In addition, this area displayed several differentially methylated regions across the individual hematopoietic stem and progenitor cell populations (highlighted by a red box), which

strongly indicates that *Evi1* undergoes changes in DNA methylation during the individual differentiation steps from HSCs to the MPP populations. Furthermore, we found several additional distinct differentially methylated regions distributed across the whole *Mecom* locus.

3.9.1.3 *Evi1* expression under conditions of physiologic stress

We were also interested in the *Evi1* expression under conditions of physiologic stress such as in response to bleeding, infections or inflammation, which rapidly activates HSCs in order to ensure the continued production of mature blood cells (Cheshier et al., 2007; Wilson et al., 2008). Here, we induced HSC activation experimentally with two agonists, IFN- α and Tpo, which mimic a stress situation (Essers et al., 2009; Yoshihara et al., 2007). In order to analyze the gene expression of highly purified LT-HSCs (Lin⁻, c-Kit⁺, Sca-1⁺, CD150⁺, CD48⁻, CD34⁻) under stress conditions, Amelie Lier performed high-throughput RNA sequencing of Tpo- and IFN- α -treated LT-HSC populations and compared normal HSC biology as well as FA HSC biology to the respective stress conditions (Lier and Milsom, unpublished data). This revealed two distinct findings for Tpo and IFN- α . While *Evi1* expression was upregulated in WT and FA HSCs in response to Tpo-treatment, IFN- α treatment resulted in the opposite finding, e.g. a downregulation of *Evi1* expression, in FA HSCs (Figure 17 C). In addition, IFN- α did not impact upon the expression levels of *EVI1* in WT HSCs. With this data, we could show that *Evi1* might be involved in the regulation of HSC maintenance in response to stress, in WT and FA mice since both demonstrated deregulated *Evi1* expression levels in response to a physiologic stress stimulus. However, the underlying mechanisms may be different, depending on the type of stress the HSC has to face.

3.9.1.4 *Evi1* expression during development

Next, we wanted to evaluate *Evi1* expression levels during HSC development. Therefore, we explored a gene expression dataset on HoxB4-YFP-positive ES reporter cell lines generated by Paul Kaschutnig (Kaschutnig and Milsom, unpublished data). These HoxB4-YFP-positive ES cells were initially isolated from the *Hoxb4*-YFP reporter mouse model (Hills et al., 2011), therefore allowing for the identification of YFP-labeled HoxB4-expressing bona fide HSCs. In the microarray, three populations from primitive (HoxB4⁻ CD41⁻ Flk1⁺ c-Kit⁺) to HSC-enriched (HoxB4⁺ CD41⁺ CD93⁺ c-Kit⁺) to more progenitor-like cells (HoxB4⁻ CD41⁺ c-Kit⁺) were analyzed (Figure 18). While *Evi1* expression levels were slightly increased in the primitive cell population, the HSC-enriched and the progenitor-like fraction remained comparable. This might indicate an importance for *Evi1* during the early step of ES cell differentiation.

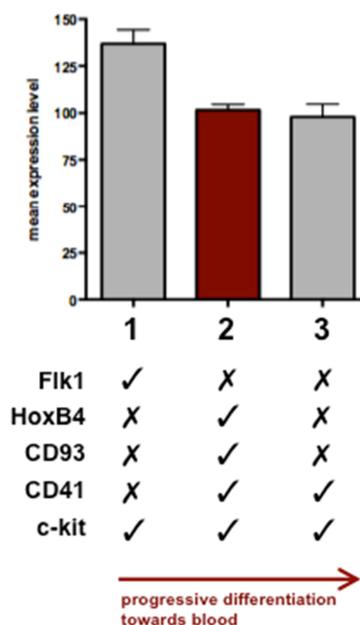


Figure 18: Evi1 expression during embryonic development. Mean expression levels of Evi1 in HoxB4-YFP-expressing embryonic stem cells during developmental specification from (1) primitive (HoxB4⁻ CD41⁻ Flk1⁺ c-Kit⁺) to (2) definitive/HSC-enriched (HoxB4⁺ CD41⁺ CD93⁺ c-Kit⁺) to more (3) progenitor-like cells (HoxB4⁻ CD41⁺ c-Kit⁺).

3.9.2 Gene expression profiling of bone marrow sample #G1

In addition to the genes that are located in close proximity to the retroviral integration site and within the defined 300 kb window, also genes outside this genomic location might have been deregulated by the strong enhancer activity of the retroviral LTRs. Indeed, it has been shown, that three-dimensional structures such as DNA supercoiling allow enhancer action over considerable genomic distances (Plank and Dean, 2014; Schoenfelder et al., 2010). One such example is the sonic hedgehog (Shh) enhancer, which displays with one megabase the so far greatest distance to its promoter (Lettice et al., 2003). Just recently, the c-myc locus was shown to be controlled by an even more distant hematopoietic stem and progenitor-specific enhancer region, which is located almost 1.7 Mb away from its promoter region (*von Paleske, doctoral thesis*). Therefore, it is very likely that the retroviral vector integration might control gene expression of genes located at distal sites by a long-range chromosomal interaction and we have to take into consideration that the emergence of the dominant stem cell clone #G1, might not exclusively result from the insertionally upregulated expression of Evi1 transcript variant I.

To address this aspect, we next aimed at investigating the global gene expression pattern in the bone marrow of mouse #G1 and performed comprehensive genome-wide expression analysis on RNA

isolated from total bone marrow cells of mouse #G1 using the Illumina Mouse Sentrix-6_v2_r3 BeadChip Array. RNA isolated from total bone marrow of three individual *Fancg*^{-/-} mice served as controls. Reverse transcription, as well as the expression analysis was performed by the Genomics and Proteomics Core Facility of the German Cancer Research Center. After normalization, which removes systematic variation from the data, we analyzed the gene expression pattern in #G1 and compared it to the *Fancg*^{-/-} control groups. This revealed 687 transcripts that were differentially expressed with a fold-change greater than 2 including 243 up-regulated transcripts and 484 down-regulated transcripts.

Using the differential gene expression data, we could confirm the transcriptional up-regulation of eGFP (2.5-fold up), which results from the expression of the fluorescent marker Venus in the retroviral vector (Table 3). In contrast to the quantitative RT-PCR analysis, which identified Evi transcript variant I, aberrant expression of Evi1 could not be confirmed with the differential gene expression data. This is most probably due to the fact that the Evi1 transcript variants are not distinguishable with the Illumina BeadChip array.

In order to obtain a more comprehensive picture of the signaling networks that are changed in the dominant stem cell clone #G1, we performed a functional analysis of the 687 most variable genes (fold change > 2) using the integrated software suite MetaCore (Thomson Reuters). Gene ontology analysis of the upregulated genes highlighted the cellular response to stress as the most significantly changed pathway in stem cell clone #G1 (**Figure 19 A**). In addition, genes involved in the regulation of apoptosis, nucleosome organization and chromosome assembly were significantly enriched in the list of upregulated genes. These processes seem to play a role in the rescue of the inherent FA HSC defect.

Under normal circumstances, the FA HSC defect would lead to cell depletion as a result of their inability to repair certain forms of DNA damage. As oxidative stress has recently been demonstrated as a source of DNA damage in HSCs, which ultimately leads to apoptosis in FA HSCs (Walter et al., 2015) it might be possible that the dominant stem cell clone #G1 was able to overcome the defect by the upregulation of a mechanisms that protects the cell from oxidative DNA damage, for example by the upregulation of heat-shock proteins such as Hsbpa1a (13-fold up), Dusp1/MKP-1 (11.7-fold up) and Dnajc6 (2.8-fold) (Table 3). In addition, the cell might have found a mechanism to evade programmed cell death, for example by the upregulation of anti-apoptotic genes such as fosb/AP-1 (20-fold), fos/c-fos (10-fold up) and Bcl-2 (2.2-fold up). In addition, Il-1 β (5.3-fold up), which has been shown to activate HSCs into cell cycle and to induce HSC proliferation *in vivo* (Sujer, *doctoral thesis*), could have influenced the survival of the #G1 stem cell clone.

Gene sets that were significantly represented in the downregulated genes of stem cell clone #G1 were associated with the immune response, defense response and cell activation (**Figure 19 B**).

A

#	Processes	0	2.5	5	7.5	10	12.5	15	17.5	-log(pValue)	pValue ↑	FDR	Ratio
1	response to stress										3.405e-24	1.638e-20	111/4295
2	response to wounding										4.592e-23	1.105e-19	63/1485
3	immune system process										1.517e-19	2.433e-16	81/2787
4	defense response										8.463e-18	1.018e-14	63/1882
5	response to external stimulus										3.277e-17	3.154e-14	75/2657
6	regulation of biological quality										1.203e-16	9.646e-14	90/2733
7	nucleosome assembly										8.905e-16	5.798e-13	19/148
8	response to cytokine										9.640e-16	5.798e-13	41/914
9	immune response										1.252e-15	6.694e-13	55/1628
10	nucleosome organization										2.351e-15	1.131e-12	20/179
11	leukocyte differentiation										4.076e-15	1.778e-12	29/457
12	chromatin assembly										4.879e-15	1.778e-12	19/162
13	myeloid leukocyte differentiation										4.879e-15	1.778e-12	19/162
14	cell activation										5.173e-15	1.778e-12	40/914
15	wound healing										6.361e-15	1.956e-12	39/874
16	regulation of apoptotic process										6.502e-15	1.956e-12	58/1864
17	negative regulation of biological process										7.206e-15	2.006e-12	103/4971
18	chromatin assembly or disassembly										7.504e-15	2.006e-12	20/190
19	cellular response to chemical stimulus										8.402e-15	2.128e-12	77/3075
20	regulation of programmed cell death										9.904e-15	2.383e-12	58/1882
21	response to abiotic stimulus										1.233e-14	2.825e-12	53/1606
22	regulation of body fluid levels										1.953e-14	4.273e-12	38/859
23	regulation of cell death										2.197e-14	4.596e-12	59/1976
24	leukocyte migration										4.401e-14	8.824e-12	24/326
25	protein-DNA complex assembly										4.670e-14	8.989e-12	19/183
26	response to biotic stimulus										6.099e-14	1.129e-11	44/1184
27	response to other organism										7.035e-14	1.202e-11	43/1138

B

#	Processes	0	3	6	9	12	15	18	21	-log(pValue)	pValue ↑	FDR	Ratio
1	immune system process										8.161e-30	4.162e-26	144/2787
2	immune response										6.484e-25	1.653e-21	99/1628
3	regulation of locomotion										1.748e-24	2.971e-21	69/838
4	regulation of immune system process										3.324e-23	3.842e-20	97/1655
5	regulation of cell motility										3.767e-23	3.842e-20	64/763
6	regulation of cellular component movement										7.829e-23	6.655e-20	67/846
7	defense response										1.947e-22	1.419e-19	103/1882
8	cell activation										2.478e-22	1.580e-19	69/914
9	response to stress										4.719e-21	2.674e-18	167/4295
10	cellular response to chemical stimulus										5.869e-21	2.993e-18	135/3075
11	regulation of cell migration										7.218e-21	3.347e-18	59/720
12	positive regulation of biological process										2.574e-20	1.094e-17	196/5578
13	positive regulation of response to stimulus										2.001e-19	7.850e-17	102/2034
14	cellular response to organic substance										6.165e-19	2.246e-16	115/2512
15	regulation of localization										7.053e-19	2.398e-16	117/2587
16	regulation of multicellular organismal process										8.477e-19	2.702e-16	128/2991
17	regulation of response to stimulus										1.783e-18	5.348e-16	156/4098
18	positive regulation of cellular process										5.387e-18	1.526e-15	176/4971
19	leukocyte activation										1.825e-17	4.897e-15	50/621
20	response to stimulus										2.294e-17	5.850e-15	282/10065
21	innate immune response										2.802e-17	6.805e-15	67/1076
22	cellular developmental process										5.952e-17	1.380e-14	158/4333
23	response to organic substance										1.240e-16	2.750e-14	137/3525
24	developmental process										6.395e-16	1.359e-13	210/6723
25	regulation of cell activation										1.064e-15	2.170e-13	48/636
26	single-organism developmental process										1.268e-15	2.487e-13	208/6670
27	regulation of immune response										1.637e-15	3.092e-13	64/1079

Figure 19: Most significantly changed gene ontology processes in stem cell clone #G1. (A) Gene ontology analysis of the upregulated genes in stem cell clone #G1. **(B)** Gene ontology analysis of the downregulated genes in stem cell clone #G1.

Table 3: Selection of differentially expressed genes in stem cell clone #G1

Fold change	Symbol	Description	Synonyms
20,07	Fosb	FBJ osteosarcoma oncogene B	
15,88	Eif2s3y	eukar. transl. initiation factor 2, subunit 3, struct. gene Y-linked	Tfy; Spy; Eif-2gy
12,97	Hspa1a	heat shock protein 1A	Hsp70-3; Hsp72
12,52	Egr1	early growth response 1	Zif268; NGFI-A
12,46	Mylc2pl	myosin light chain 2, prec. lymphocyte-specific, transcr. variant 2	PLRLC-A, -B, -C
11,69	Dusp1	dual specificity phosphatase 1	erp; mkp-1; Ptpn16
10,06	Fos	FBJ osteosarcoma oncogene	c-fos; D12Rfj1
9,73	Tcrg-V4	T-cell receptor gamma, variable 4	
8,06	Gm525	gene model 525	
7,26	Ighg	Immunogl. heavy chain (gamma polypeptide), transcript variant 1	
7,04	Plk3	Polo-like kinase 3	
6,90	Osm	oncostatin M	
6,66	Cxcl2	chemokine (C-X-C motif) ligand 2	MIP-2a; Gro2
5,99	Zfp36	zinc finger protein 36	Gos24; Nup475
5,99	Bcl11b	B-cell leukemia/lymphoma 11B, transcript variant 2	Rit1; Ctjp2
5,92	Adam8	a disintegrin and metallopeptidase domain 8	MS2; CD156
5,80	Rgs1	regulator of G-protein signaling 1	BL34
5,78	Ybx3	Y box binding protein 3	
5,36	Junb	Jun-B oncogene	
5,35	Jun	Jun oncogene	c-jun; Junc; AP-1
5,26	Gm525	gene model 525	
5,26	Il1b	Interleukin-1 beta	catabolin
⋮	⋮	⋮	⋮
2,85	Dnajc6	DnaJ (Hsp40) homolog, subfamily C, member 6	mKIAA0473
⋮	⋮	⋮	⋮
2,50	eGFP	enhanced green fluorescent protein	
⋮	⋮	⋮	⋮
2,18	Bcl2	B-cell leukemia/lymphoma 2, transcript variant 2	Bcl-2
⋮	⋮	⋮	⋮
1,30	Evi1	ecotropic viral integration site 1	Evi-1; Jbo

3.10 Identification and characterization of deregulated genes in mouse #A2

In clone #A2, we have identified two retroviral vector integrations, one which was located on chromosome 8 and one on the X-chromosome. The genetic locus on chromosome 8 contained eight genes (Cdh13, Hspb1, Mlycd, Osgin1, Necab2, Slc38a8, Mbtps1 and Hsd11) whereas the X-chromosome revealed four additional genes (Gm8817, Tmsb4x, Tlr7 and Tlr8).

To identify the genes that drive clonal dominance in mouse #A2 among the twelve potential candidates, we sought to investigate deregulated expression of genes upstream and downstream of the retroviral integrations in whole bone marrow samples from this particular mouse using quantitative RT-PCR (**Figure 20**). The results from this analysis revealed that the Oxidative Stress-induced Growth Inhibitor 1 (Osgin1) was about 3-fold upregulated and therefore the likely driver of clonal dominance in this mouse.

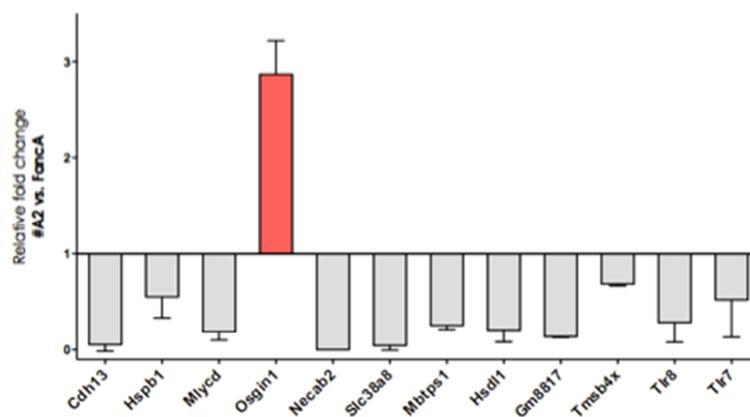


Figure 20: Aberrant expression of Osgin1 in the dominant stem cell clone #A2. Relative mRNA expression levels of genes proximal to the retroviral integration site in bone marrow isolated from mouse #A2. Red bar indicates upregulated expression of Osgin1; grey bars represent expression levels of genes that were comparable to *Fanca*^{-/-} control bone marrow.

3.10.1 Oxidative stress-induced growth inhibitor 1 (Osgin-1)

Osgin1, also referred to as OKL38 or bone marrow stromal cell-derived growth inhibitor (BDGI), was initially identified as a pregnancy-induced growth inhibitor with a major role in breast epithelial cell growth regulation and differentiation during pregnancy and tumorigenesis (Huynh et al., 2001). In addition, it was demonstrated to have an antiproliferative effect on human breast cancer cells by inducing cellular apoptosis and cell cycle arrest (Wang et al., 2005). These findings support Osgin1 as a tumor suppressor gene, however the mechanisms underlying this effect remain largely elusive.

Furthermore, Osgin1 is an oxidative stress response protein with a pathophysiological role in chronic inflammatory diseases such as atherosclerosis. In this context, Osgin1 expression is induced by high superoxide levels and was suggested to protect cells against oxidative stress (Li R, 2007; Yan et al., 2014). In response to oxidative stress or DNA damage, Osgin1 was shown to interact with p53 and translocate to the mitochondria in order to induce a cytochrome c release during apoptosis (Hu et al., 2012; Yao et al., 2008).

Loss or decreased levels of Osgin1 expression have been associated with the carcinogenesis of liver and kidney and could be correlated to an adverse outcome with shorter survival times of patients (Ong et al., 2007; Ong et al., 2004). Just recently, a nucleotide variation in the Osgin1 gene resulting in an amino acid substitution from arginine to histidine on codon 438, was identified in tumor tissue from patients with hepatocellular carcinoma. This Osgin1 variant was less effective in the translocation from nucleus and mitochondria, which reduced its apoptotic effect (Liu et al., 2014).

The role of Osgin1 in the hematopoietic system and in stem cell biology remains largely elusive. In addition, Osgin1 expression in the context of FA has not been studied yet. However, due to its suggested function as an oxidative stress response gene with a role in atherosclerosis, it may also be involved in the pathology of FA, in which oxidative stress signals can promote a collapse of the entire hematopoietic system leading to bone marrow failure.

3.10.1.1 Osgin1 expression levels under homeostatic conditions

For an overview on the relative expression levels of Osgin1, we explored the BioGPS gene annotation portal (www.biogps.org), which provides a huge collection of gene expression data on a variety of biological samples. Comparing different tissues, Osgin1 expression was found highest in liver, testis and adrenal gland (**Figure 21**).

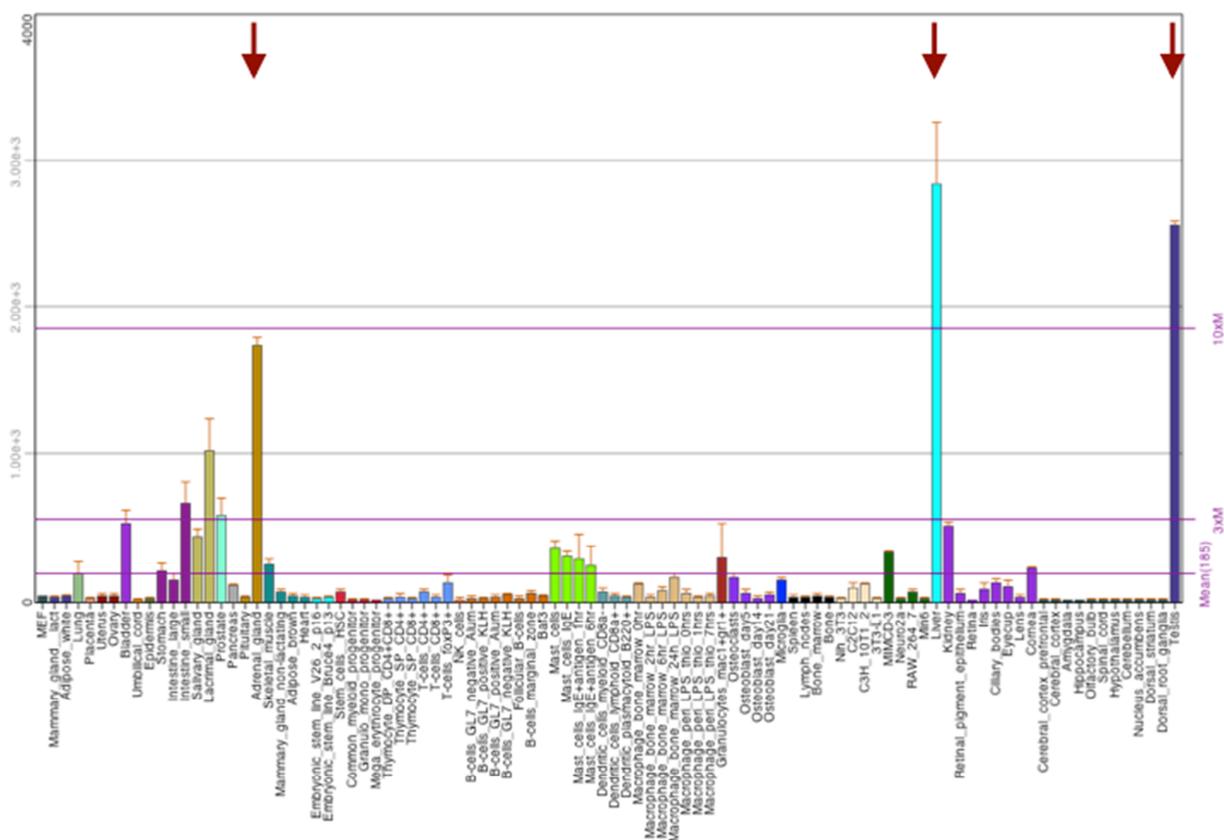


Figure 21: Osgin1 expression across different tissues and cell lines. Red arrows indicate pronounced expression of Osgin1 in adrenal gland, liver and testis.

As the role of Osgin1 in the hematopoietic stem and progenitor compartment has not been addressed to date, we were interested in the expression levels of Osgin1 in HSCs and MPPs. Therefore, we made use of a comprehensive gene expression dataset that was generated by high-throughput RNA sequencing through the collaborative effort of the groups of Andreas Trumpp, Michael Milsom and Daniel Lipka at the German Cancer Research Center and the members of the group of Jerome Krijgsfeld at the European Molecular Biology Laboratories in Heidelberg (Cabezas-Wallscheid et al., 2014). These data revealed about 3-fold higher levels of Osgin1 in HSCs and MPP4 than in MPP2 and MPP3 (**Figure 22 A**).

3.10.1.2 Epigenetic regulation of Osgin1 expression

Next, we investigated the methylation pattern of the Osgin1 locus, on the long arm of chromosome 8 in the sub band E1, using the tagmentation-based whole genome bisulfite sequencing dataset on the hematopoietic stem and progenitor cells generated by Daniel Lipka (Cabezas-Wallscheid et al., 2014; Lipka et al., 2014). This analysis identified a relatively broad hypomethylated region of approximately 10 kb upstream of the Osgin1 transcription start site, which was conserved

throughout from LT-HSCs to MPPs (Figure 22 B; highlighted by a grey box). In addition, the *Osgin1* transcription start site also revealed a distinct differentially methylated region across the HSC and MPP3/4 populations (highlighted by a red box), which might indicate some relevance for *Osgin1* in the regulation of HSC biology as HSCs seem to undergo changes in DNA methylation during differentiation towards multipotent progenitors.

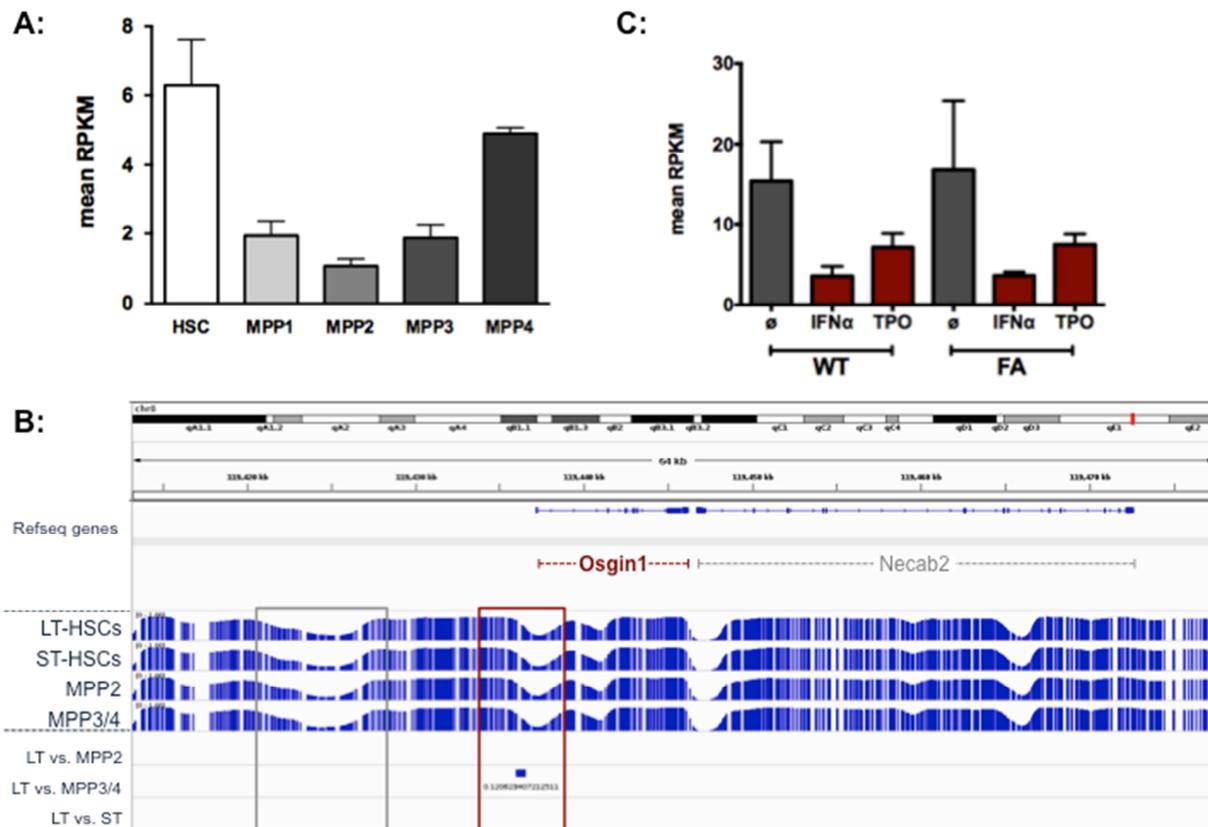


Figure 22: *Osgin1* expression levels and methylation pattern in HSC and progenitor populations. (A) Mean *Osgin1* expression levels in LT-HSCs and progenitor populations MPP1 – 4 as obtained by high-throughput RNA sequencing. (B) *Osgin1* methylation pattern in LT-HSCs and MPP1 - 4 generated by tagmentation-based whole genome bisulfite sequencing. Red box highlights the transcription start site of *Osgin1* transcript variant I with differential methylation across HSC and MPP3/4 populations; grey box highlights a broad hypomethylated region, which was conserved throughout from LT-HSCs to MPPs (C) Mean *Osgin1* expression levels in FA and WT LT-HSCs under physiologic stress conditions as obtained by high-throughput RNA sequencing of IFN- α , Tpo-and control-treated LT-HSCs.

3.10.1.3 Osgin1 expression under conditions of physiologic stress

Next, we wanted to evaluate Osgin1 expression under conditions of physiologic stress such as in response to severe bleeding or infections, which rapidly activates HSCs in order to ensure the continued production of mature blood cells. Such a stress situation can be modeled using interferons or Tpo, which can induce HSC proliferation upon administration (Baldrige et al., 2010; Essers et al., 2009; Yoshihara et al., 2007). In order to analyze the gene expression of highly purified LT-HSCs (Lin⁻, c-Kit⁺, Sca-1⁺, CD150⁺, CD48⁻, CD34⁻) under stress conditions, Amelie Lier performed high-throughput RNA sequencing of IFN- α - and Tpo-treated LT-HSC populations and compared normal HSC biology as well as FA HSC biology to the respective stress conditions (*Lier and Milsom, unpublished data*). This revealed that in response to both treatments, Osgin1 was down regulated both in WT and FA HSCs (**Figure 22 C**). This strongly suggests a role for Osgin1 in the stress biology of WT as well as FA HSCs. Although we observed comparable levels of downregulation in stressed FA and WT HSCs, the effect was much stronger in response to IFN- α when compared to Tpo treatment suggesting a more pronounced role for Osgin1 during the proinflammatory stress response.

3.10.1.4 Osgin1 expression during development

In order to investigate the Osgin1 expression levels during HSC development, we made use of HoxB4-YFP-positive ES reporter cell lines, which were derived from the *Hoxb4*-YFP reporter mouse model (Hills et al., 2011) and analyzed for their gene expression profiles using microarray by Paul Kaschutnig (*Kaschutnig and Milsom, unpublished data*). As the YFP-labeled HoxB4 expression is restricted to the definitive HSCs in these mice, during ES cell culture YFP expression marks bona fide HSCs. In the microarray, three populations were analyzed: A primitive population, which also contains mesodermal cells (HoxB4⁻ CD41⁻ Flk1⁺ c-Kit⁺), a more definitive population, which is enriched for HSCs (HoxB4⁺ CD41⁺ CD93⁺ c-Kit⁺) and a population of more progenitor-like cells (HoxB4⁻ CD41⁺ c-Kit⁺). In this dataset, a major difference in the mean expression levels of Osgin1 expression could be detected (**Figure 23**). While the expression levels of Osgin1 remain comparable in the primitive and in the progenitor-like cell population, Osgin1 is almost 4-fold upregulated in the HSC-enriched fraction. This suggests an HSC-specific role for Osgin1 during developmental processes.

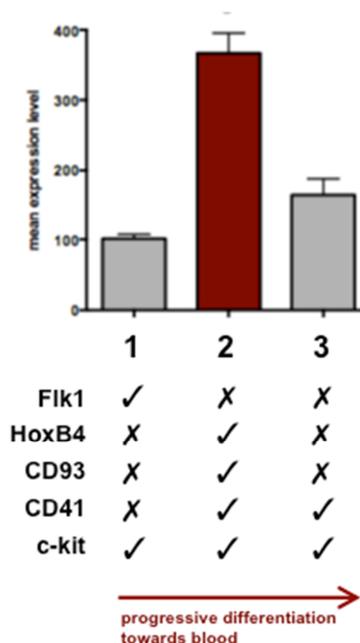


Figure 23: Osgin1 expression during embryonic development. Mean expression levels of Osgin1 in HoxB4-YFP-expressing embryonic stem cells during developmental specification from (1) primitive (HoxB4⁻ CD41⁻ Flk1⁺ c-Kit⁺) to (2) definitive/HSC-enriched (HoxB4⁺ CD41⁺ CD93⁺ c-Kit⁺) to more (3) progenitor-like cells (HoxB4⁻ CD41⁺ c-Kit⁺).

3.10.2 Gene expression profiling of bone marrow sample #A2

As outlined in chapter 3.9.2, enhancer elements as contained in the LTRs of the retroviral SF91 vector can elicit their effects over large genomic distances mediated by long-range chromosomal interactions. Therefore the emergence of the dominant stem cell clone #A2 might not exclusively result from the insertionally upregulated expression of Osgin1, but could also be the result of a change in gene expression of additional genes located at sites outside the 300 kb window that we have assessed.

In order to address this aspect, we performed gene expression analysis on RNA isolated from total bone marrow cells of mouse #A2 was using the Illumina Mouse Sentrix-6_v2_r3 BeadChip Array. RNA isolates from total bone marrow of three individual *Fanca*^{-/-} mice served as control samples. The Genomics and Proteomics Core Facility of the German Cancer Research Center performed reverse transcription, expression analysis as well as the removal of systematic variation in the microarray data by normalization. We analyzed the gene expression pattern in #A2 and compared it to the three *Fanca*^{-/-} control groups in order to identify the differentially regulated genes. This revealed 1551 transcripts that were differentially expressed with a fold-change greater than 2 including 706 up-regulated transcripts and 851 down-regulated transcripts.

Using the differential gene expression data, we could confirm the transcriptional up-regulation of eGFP (28.2-fold up), which results from the expression of the fluorescent marker Venus in the retroviral vector (Table 4). In addition, also the 3-fold upregulated expression of Osgin1 as identified by quantitative RT-PCR could be confirmed with the differential gene expression data (2.9-fold up). Despite the slight difference in the level of the fold changes between the array and the quantitative RT-PCR data, this result reliably reflects the deregulated expression of Osgin1 in stem cell clone #A2.

In order to identify the enriched biological processes in the dominant stem cell clone #A2, we performed a functional analysis on the 773 most variable genes with a fold change of more than 2.5 using the MetaCore enrichment analysis tool (Thomson Reuters). For this analysis, we chose a more stringent threshold as opposed to the analysis of stem cell clone #G1 (refer to 3.9.2), where we used a threshold of 2.0, as many more differentially regulated genes were identified in stem cell clone #A1 and we wanted to keep the overall number of genes comparable. Gene ontology analysis of the 382 upregulated genes highlighted the response to stress as the most significantly changed process in stem cell clone #A2 (**Figure 24 A**), which together with the increased expression of Osgin1 (2.9-fold up; Table 4) indicates a defense mechanism towards oxidative stress. Since oxidative stress has recently been demonstrated to cause DNA damage and induce apoptosis in FA HSCs (Walter et al., 2015) the dominant stem cell clone might have upregulated a defense mechanism against oxidative DNA damage. Furthermore, we found increased levels of anti-stress genes such as Cxcl2 (14-fold up) and Dusp1 (4.2-fold up) as well as the heat-shock proteins Dnajb2 (4.2-fold up), Dusp1 (4.2-fold up) and Serpinh1 (3.8-fold up) among the top upregulated candidates (Table 4). In addition, cell cycle regulators such as Cdkn1a (8.2-fold up), Plk3 (7-fold up), Gas6 (5.7-fold up), Cdkn2a/p16 (3.9-fold up) and Il-1 beta (3.6-fold up) as well as anti-apoptotic genes such as fosb/AP-1 (7.2-fold up) might indicate a mechanism that promotes cell cycle progression and proliferation while suppressing programmed cell death.

Gene sets that were significantly represented among the 393 downregulated genes of stem cell clone #G1 were associated with the regulation of the immune system (**Figure 24 B**).

Table 4 Selection of differentially expressed genes in stem cell clone #A2

Fold change	Symbol	Description	Synonyms
28,20	eGFP	enhanced green fluorescent protein	
24,51	Cfd	complement factor D (adipsin)	Adn; DF; factor D
14,05	Cxcl2	chemokine (C-X-C motif) ligand 2	Mgsa-b; Scyb; MIP-2a; Gro2
12,10	Xist	inactive X specific transcripts on chromosome X.	
9,65	Adipoq	adiponectin, C1Q and collagen domain containing	adipo; Acrp30; GBP28
8,96	Mmp13	matrix metalloproteinase 13	MMP-13; Mmp1; Clg
8,76	Cma2	chymase 2, mast cell	Mcp10
8,23	Cma1	chymase 1, mast cell	Mcpt5; MMCP-5; Mcp-5
8,18	Cdkn1a	cyclin-dependent kinase inhibitor 1A (P21)	p21Cip1; P21; Waf1; CDKI
7,72	Rgs1	regulator of G-protein signaling 1	BL34
7,68	Mcpt4	mast cell protease 4	Mcp-4; MMCP-4
7,19	Fosb	FBJ osteosarcoma oncogene B	
7,00	Plk3	polo-like kinase 3 (Drosophila)	CNK, FNK, PRK
6,45	Igfbp5	insulin-like growth factor binding protein 5	IGFBP-5
6,44	Esm1	endothelial cell-specific molecule 1	ESM-1
6,34	Igfbp7	insulin-like growth factor binding protein 7	Fstl2; mac25
6,17	Mcpt6	mast cell protease 6	MMCP-6; Mcp-6
6,07	Egr1	early growth response 1	Zif268; NGFI-A
6,04	Bglap2	bone gamma-carboxyglutamate protein 2	OG2; mOC-B
5,90	Ctsk	cathepsin K	MMS10-Q; catK; Ms10q
5,66	Gas6	growth arrest specific 6	Gas-6
5,57	Cox6b2	cytochrome c oxidase subunit VIb polypeptide 2	
5,56	Mgp	matrix Gla protein	Mglap
5,53	Bgn	biglycan	
5,28	Bglap1	bone gamma carboxyglut. protein 1, transcript variant 2	OC; mOC-A; Bglap; OG1
5,28	Sparc	secreted acidic cysteine rich glycoprotein	Osteonectin, BM-40
5,28	Ctla4	cytotoxic T-lymphocyte-associated protein 4	Ctla-4; Cd152; Ly-56
4,60	Cxcl14	chemokine (C-X-C motif) ligand 14	
4,51	Tnfaip3	tumor necrosis factor, alpha-induced protein 3	A20; Tnfaip3
4,23	Dnajb2	DnaJ homolog, subfamily B, member 2, transcr. var. 2	
4,19	Dusp1	dual specificity phosphatase 1	erp; mkp-1; Ptpn16
4,14	Ccl4	chemokine (C-C motif) ligand 4	Scya4; Mip1b; Act-2; MIP-1B
3,93	Aldh1a1	aldehyde dehydrogenase family 1, subfamily A1	Raldh1; E1; Ahd2; Aldh1
3,85	Cdkn2a	cyclin-dependent kinase inhibitor 2A, transcr. var. 1	p16; MTS1; Arf; Pctr1
3,82	Serpinh1	serpin peptidase inhibitor, clade H, member 1	HSP47
3,60	Il1b	Interleukin-1 beta	catabolin
3,41	Bbc3	BCL2 binding component 3	PUMA; PUMA/JFY1
2,90	Osgin1	oxidative stress induced growth inhibitor 1	Ok138

A



B

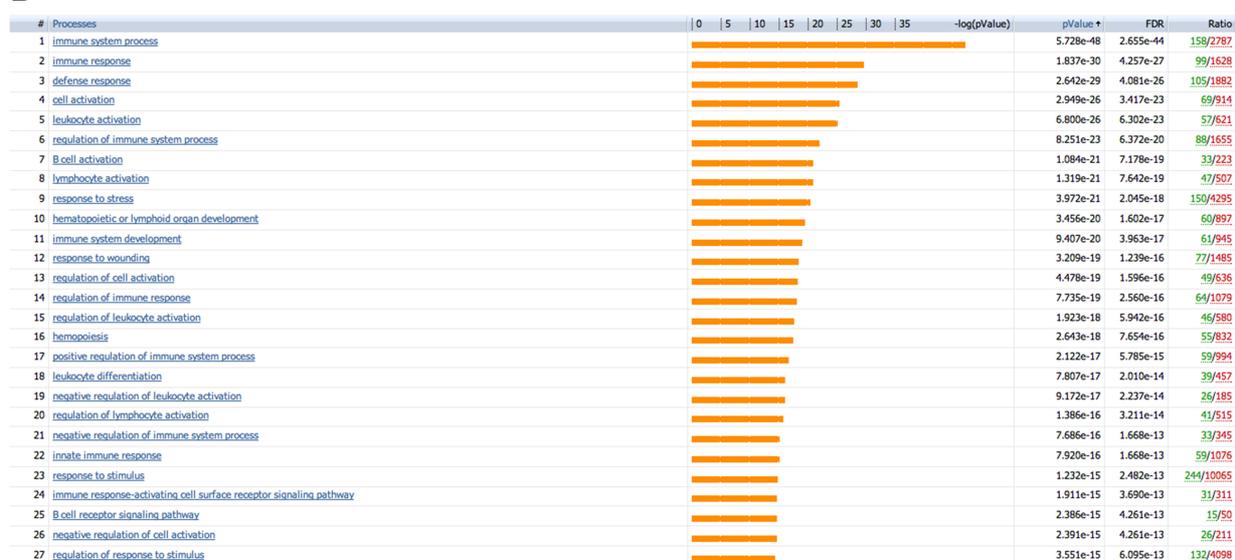


Figure 24: Most significantly changed gene ontology processes in stem cell clone #A2. (A) Gene ontology analysis of the upregulated genes in stem cell clone #A2. **(B)** Gene ontology analysis of the downregulated genes in stem cell clone #A2.

3.11 Target gene validation studies

To investigate whether the genes, which were identified in the retroviral insertional mutagenesis screen, were able to confer a selection advantage to FA HSCs upon transplantation, we next performed target gene validation studies. For this purpose lentiviral expression constructs were generated to be able to perform gain-of-function or loss-of-function studies by constitutive overexpression or knockdown of target genes *in vivo*.

3.11.1 Generation of lentiviral vectors for target gene validation studies

For the purpose of target gene validation, we decided to use lentiviral vectors as they provide several advantages over other types of viral vectors. In general, lentiviral vectors allow for functional gene analysis by stable long-term expression of the transgene as they integrate their genome into the host cell and are potent in transducing both actively proliferating and non-dividing cells (Naldini et al., 1996; Weinberg et al., 1991) while for example γ -retroviruses only transduce cells that divide shortly after infection (Miller et al., 1990; Roe et al., 1993). Therefore, they are particularly useful for studies in primary hematopoietic stem and progenitor cells, which do not divide very frequently *in vitro* during short term culture. In addition, the design of the lentiviral vectors allows for integration of relatively large genetic sequences of up to 9 kb. However the major advantage of using lentiviruses is their safety profile. In comparison to retroviral vectors, which have been found to convey mutagenic side effects in several human gene therapy trials and animal models (Cavazzana-Calvo et al., 2000; Hacein-Bey-Abina et al., 2003a; Li et al., 2002; Modlich et al., 2005; Nienhuis et al., 2006) lentiviral vectors show a different integration profile not associated with insertional mutagenesis (Montini et al., 2006). Since we aimed at investigating a genetic effect that was identified in response to insertional mutagenesis, we wanted to get around the integrational upregulation of random genes and only look at the effects of the overexpressed transgenes of interest. For this reason, lentiviral vectors are very well suited for the purpose of target gene validation studies in this context.

The lentiviral vectors used in this study are based on the lentiviral “gene ontology” (LeGO) vectors, which are human immunodeficiency (HIV)-1 derived third generation lentiviral vectors that allow for stable transgene expression under the control of the spleen focus-forming virus (SFFV) promoter/enhancer (Weber et al., 2008). LeGOiV2, the vector backbone for the lentiviral vectors generated in this study, contains an encephalomyocarditis virus internal ribosomal entry site (IRES) for expression of the fluorescent marker gene Venus and a multiple cloning site (MCS) with the five restriction sites BamHI, EcoRI, SbfI, StuI, NotI (**Figure 25 A**).

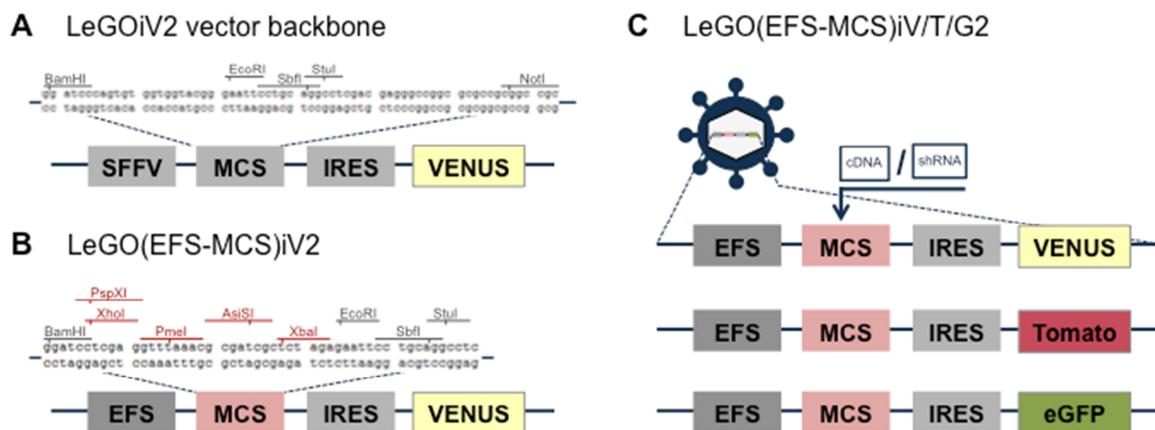


Figure 25: Generation of lentiviral vectors for target gene validation. (A) Schematic representation of the LeGOiV₂ vector backbone with the SFFV promoter, the multiple cloning site (MCS) with detailed representation of the restriction sites, the internal ribosomal entry sequence (IRES) and the Venus fluorescent reporter protein (B) Schematic representation of the LeGO(EFS-MCS)iV₂ after insertion of the elongation factor 1 (EFS) promoter and the optimized MCS containing five additional restriction sites (represented in red). (C) Schematic representation of the three final lentiviral vector constructs LeGO(EFS-MCS)iV/T/G₂ containing the different marker proteins Venus, tdTomato and eGFP respectively.

To facilitate the cloning procedure when integrating transgenes and to ensure stable but non-mutagenic expression levels, LeGOiV₂ was modified in order to obtain the LeGO(EFS-MCS)iV₂ vector as follows (Figure 25 B). In a first step, the SFFV promoter in the original construct was replaced with the elongation factor 1 (EFS) promoter (Kim et al., 1990), which allows for stable transgene expression, especially in cells of the hematopoietic compartment. In contrast to the retroviral enhancer-promoter SFFV, which has been shown to have significantly greater potential of activating neighboring promoters, the cellular EFS promoter derived from a human gene reduces the genotoxic risk and does not get silenced (Zychlinski et al., 2008). The EFS promoter was amplified from the SRS11-EFS-MGMT(P140K)-IRES-Venus plasmid and subcloned into the LeGOiV₂ vector.

Next, five unique restriction sites (PspXI, XhoI, PmeI, AsiSI and XbaI) were added to the MCS of the LeGOiV₂ to generate an optimized MCS with a total of nine unique restriction sites of which four are eight-base-cutters, five are six-base-cutters and two are blunt-end-cutters. This provides some flexibility for the integration of the selected target genes in a later cloning step.

Furthermore, we aimed at developing a panel of vectors, each expressing a different fluorescent marker gene as we wanted to have the option of using several vectors simultaneously as well as choosing the color of the vector according to the experimental conditions such as antibody combinations in stainings for flow cytometry. For this purpose, the newly generated EFS-MCS cassette was subcloned into the LeGOiV₂ and LeGOiG₂ vector backbones in order to generate two

additional versions of the vector, LeGO(EFS-MCS)iT₂ with the fluorescent tdTomato-cassette and LeGO(EFS-MCS)iG₂ with the fluorescent eGFP-cassette (**Figure 25 C**). This provides some flexibility in terms of colors as the expression of the inserted transgenes can be monitored by either Venus-, tdTomato- or eGFP expression according to the type of vector used in the experiment.

3.11.1.1 Subcloning of Osgin1

Out of the target genes, which were identified in the insertional mutagenesis screen, we decided to subclone Osgin1 as the most interesting candidate into the LeGO(EFS-MCS)iV/T/G₂ vectors in order to obtain expression vectors for overexpression of Osgin1 *in vivo*.

The full-length cDNA sequence for murine Osgin1 was purchased from the Mammalian Gene Collection (MGC). The MGC clone with the ID 37844 was obtained in the pCMV-SPORT6 vector and correctness of the sequence was confirmed using sanger sequencing with the M13 forward and reverse primers. Blasting of the sequencing against the murine genome confirmed that this cDNA clone picked up all five splice variants of Osgin1. In order to subclone Osgin1 into the LeGO(EFS-MCS)iV/T/G₂ vector, EcoRI and XhoI restriction enzymes were used to cut insert and vector (**Figure 26 A**). Presence of the insert resulting in the LeGO(EFS-Osgin1)iV/T/G₂ vector was confirmed with a test digest using EcoRI and XhoI. Gel electrophoresis identified a 1.5 kb fragment corresponding to the Osgin1 insert and the LeGO(EFS-MCS)iV2/iT₂ and iG₂ vectors of approximately 7.7 and 8.4 kb respectively.

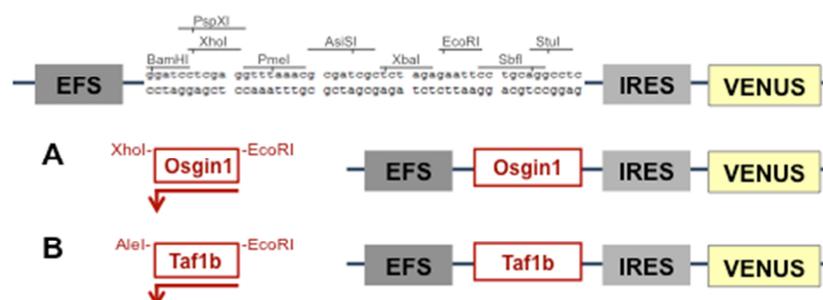


Figure 26: Subcloning of Osgin1 and Taf1b into the lentiviral vector backbone LeGO(EFS-MCS)iV₂. (A) The full-length cDNA for Osgin1 was subcloned into the LeGO(EFS-MCS)iV₂ vector using XhoI and EcoRI restriction enzymes (B) The restriction digest for the generation of the LeGO(EFS-Taf1b)iV₂ vector was set up using AclI and EcoRI restriction enzymes.

3.11.1.2 Subcloning of Taf1b

In addition to Osgin1, we also generated a lentiviral overexpression construct for Taf1b. For this purpose, the group of Ingrid Grummt at the German Cancer Research Center provided the full-length cDNA sequence for murine Taf1b, which was contained in the pT β -FLAG-mTaf168 vector (ID 1059). The identity of the construct was verified by DNA sequencing using a set of customized primers that allow generation of a full-length sequencing product. In order to subclone Taf1b into the LeGO(EFS-MCS)iV/T/G₂ vector backbone, a restriction digest was set up with *AleI* and *EcoRI* (**Figure 26 B**) to isolate the 2 kb fragment corresponding to Taf1b from the pT β -FLAG-mTaf168. To linearize the LeGO(EFS-MCS)iV/T/G₂ vectors, *PmeI* and *EcoRI* enzymes were used. However, so far we were not yet successful in generating the ligation product LeGO(EFS-Taf1b)iV/T/G₂.

3.11.2 Production of lentiviral Osgin1 overexpression vectors

The generated replication-deficient HIV-derived lentiviral vectors, which contained the candidate target gene Osgin1 (LeGO(EFS-transgene)iV₂) and the corresponding empty control vectors (LeGO(EFS-MCS)iT₂) were produced by transient co-transfection of the vector plasmids, envelope and packaging constructs into human embryonic kidney 293T producer cells using calcium phosphate co-precipitation (**Figure 27**). Due to safety concerns, the viral components were delivered on three separate plasmids (Naldini et al., 1996): the lentiviral expression construct, the Gag/Pol packaging plasmid pSPAX2 and the envelope plasmid pMD2.G. The lentiviral expression plasmid encodes the LeGO(EFS-transgene)iV₂ construct while the pSPAX2 expresses the structural genes Gag and Pol and the regulatory Tat and Rev components, which are necessary for the packaging and replication process of the lentiviral vectors respectively (Zufferey et al., 1997). To obtain a broad tropism as well as vector particle stability, the viral particles were pseudotyped with the ecotropic envelope plasmid pMD2.G, which is derived from the G-protein of the vesicular stomatitis virus (VSV-G) (Burns et al., 1993). All virus stocks were harvested, concentrated by ultracentrifugation and re-suspended in serum-free medium for long-term storage.

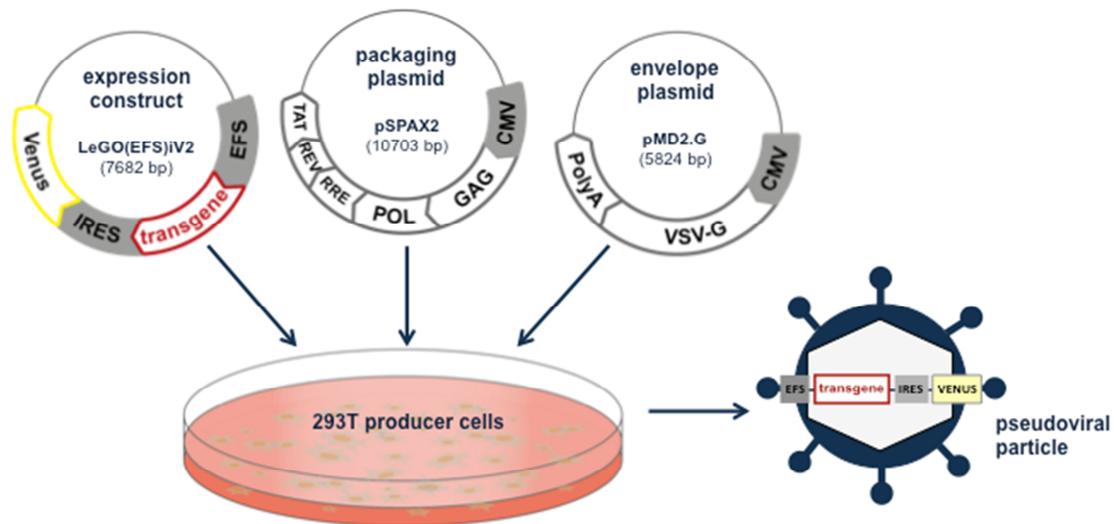


Figure 27: Production of lentiviral overexpression vectors. Lentiviral particles were produced by transient co-transfection of the human embryonic kidney producer cell line 293T with the desired expression construct (LeGo(EFS)iV₂), the packaging plasmid (pSPAX2) and the envelope plasmid (pMD2.G).

Since we found that the initial protocol, which was routinely used for the production of lentiviral vectors for *in vitro* applications in our laboratory, was not sufficient to generate high viral titers that efficiently transduce primary bone marrow cells *in vivo*, we aimed at optimizing the original protocol. In order to obtain high lentiviral titers, we first of all increased the overall virus yield by increasing the number of transduced human embryonic kidney 293T cells. For a better handling during transfection and also afterwards during medium changes and virus harvest we switched from 10 cm cell culture dishes to T175 cell culture flasks by scaling up the individual components used in the transfection protocol. In addition, the optimal ratio of the components in the transfection mixture was evaluated by adjusting the amounts of the lentiviral plasmid, the packaging plasmid and the envelope plasmid. LeGO(EFS-MCS)iT₂ was used as a sample vector to investigate the potency of the pseudoviral particles generated according to the three different transfection protocols A, B and C (**Figure 28**). In contrast to the above mentioned standard protocol for lentiviral production, LeGO(EFS-MCS)iT₂ pseudoviral particles, which were generated to test the efficiency of the different transduction mixtures, were not concentrated using ultracentrifugation and frozen immediately upon harvest on day three. The efficiency of the viral particles was evaluated following limiting dilution transduction of murine 3T3 cells using flow cytometry according to the percentage of transduced cells (**Figure 29 A**)

As shown in **Figure 28**, the highest transduction efficiency was obtained with viral particles that were produced using transfection mixture B with a titer of 9.6×10^5 transducing units/ml. Therefore, the

production of all lentiviral overexpression vectors and corresponding control constructs was based on protocol B.

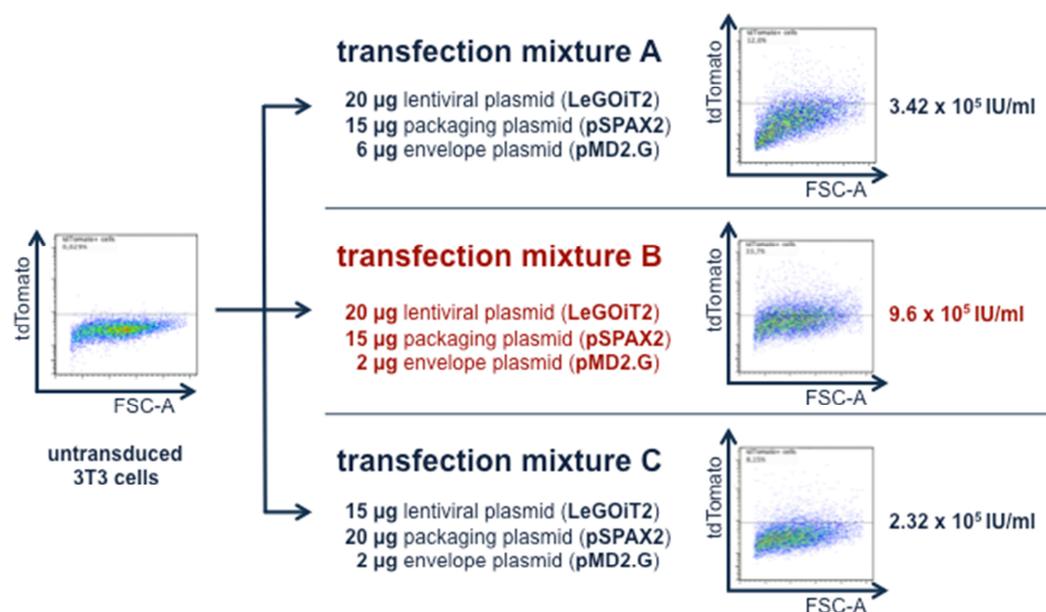


Figure 28: Evaluation of transfection mixtures for the generation of lentiviral vectors. Murine 3T3 cells were transduced with a limiting dilution of the pseudoviral particles that were generated according to the transfection mixtures A, B and C and evaluated using flow cytometry. Viral particles generated according to transfection mixture B (shown in red) resulted in the highest transduction efficiency.

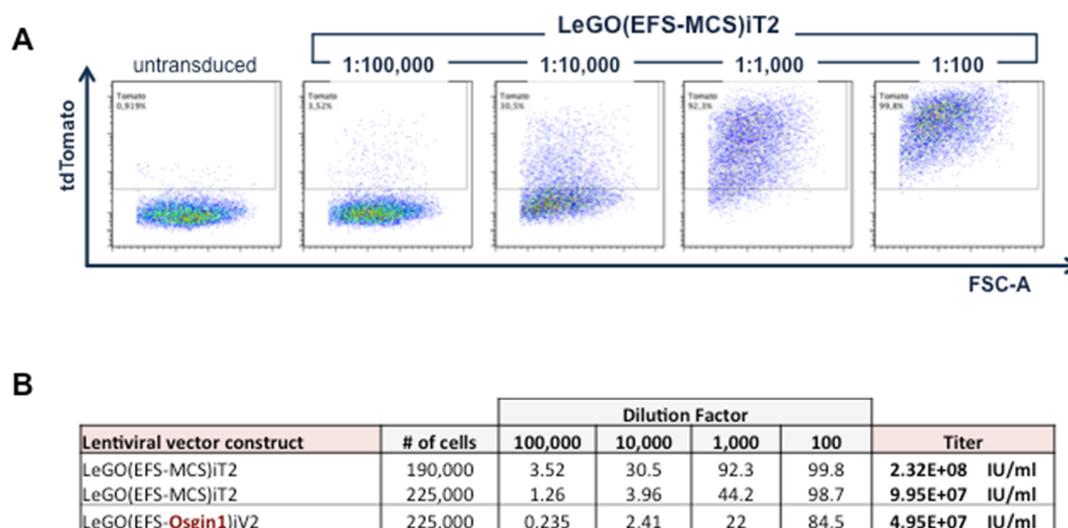


Figure 29: Analysis of lentiviral vector titers. (A) Representative FACS plots of the limiting dilution transduction of murine 3T3 cells with LeGO(EFS-MCS)iT2. Each FACS plots depicts the transduction efficiency at the indicated dilution of the virus (B) Range of viral titers obtained using limiting dilution analysis.

Furthermore, the virus titers for all lentiviral vectors packaged with the G protein of VSV-G and generated according to protocol B were determined following limiting dilution transduction of murine 3T3 cells using flow cytometry (**Figure 29 A**).

As shown in **Figure 29 B**, high titers were readily obtained for LeGO(EFS-MCS)iT₂ and the corresponding overexpression construct for Osgin1, LeGO(EFS-Osgin)iV₂, ranging from 5×10^7 to 2.3×10^8 transducing units/ml. These titers were used to evaluate the relative amount of virus needed for a certain infection rate in order to be able to transduce primary bone marrow LSK cells at comparable levels.

3.11.3 Lentiviral overexpression of Osgin1 in LSK cells

In order to transduce primary bone marrow cells with the lentiviral vectors generated in this study, LSK cells were isolated from either FA or WT mice and cultured on retronectin-coated plates in the presence of serum-free medium supplemented with 50 ng/ml mSCF, mTpo and hFlt3-L. For an efficient transduction it is necessary to keep the cells to be transduced in a relatively small volume of medium to increase the likelihood of cells coming into contact with the pseudoviral particles. LSK cells were subsequently transduced with the Osgin1 overexpression construct or the vector control at a MOI ranging from 10 to 50 aiming for approximately 5×10^6 transducing units/well. In order to enhance the transduction efficiency of the lentiviral particles, 4 µg/ml polybrene (Sigma-Aldrich) was added to the culture for a maximum of 12 hours.

To test whether the LeGO(EFS-MCS)iT₂ vector and the LeGO(EFS-Osgin1)iV₂ constructs allowed for efficient marking of transduced primary LSK cells, the transduction efficiency of lentiviral vectors was evaluated 36 – 48 hours post-transduction by flow cytometry. The results from this analysis are shown in **Figure 30**. We could confirm that the EFS promoter provided high expression levels of the marker gene, particularly in LSK cells independent of whether the Osgin1 transgene was incorporated in the vector or not. Empty vector controls LeGO(EFS-MCS)iT₂ as well as LeGO(EFS-Osgin1)iV₂ overexpression constructs achieved comparable levels of transduction ranging from 32 – 90% of transduced WT as well as FA cells. As expected, increasing MOIs also resulted in higher expression levels as a direct consequence of the number of vector integrations. However, above an MOI of 34 the transduction efficiency was not increased anymore and the high viral load was also found to be cytotoxic as more dead cells were observed in the well 36 – 48h post transduction. Therefore, we chose to keep the MOIs below 34 and aim for a transduction efficiency of approximately 70%.

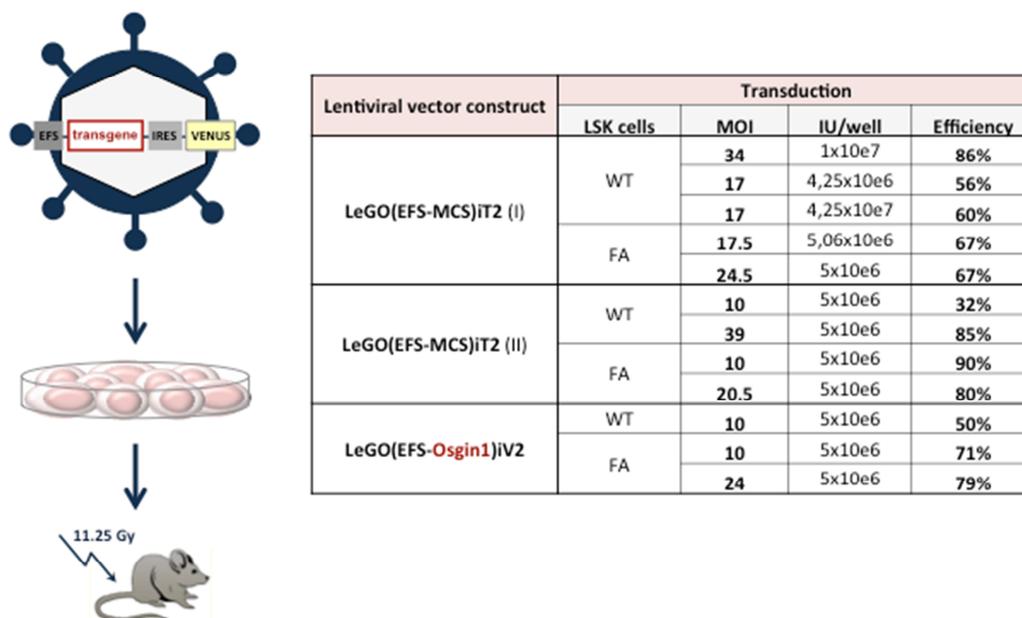


Figure 30: Transduction of primary bone marrow cells. Purified and FACS-isolated bone marrow LSK cells were transduced with the lentiviral vector constructs and evaluated for their transduction efficiencies 36 – 48 hours post transduction. Table depicts the transduced cell types (WT/FA), the multiplicity of infection (MOI), the infectious units per well (IU/well) as well as the evaluated transduction efficiency for each lentiviral vector construct. Transduced cells were transplanted into myeloablated WT recipient mice.

3.11.4 Evaluation of Osgin1 overexpression in a competitive transplantation setting

In order to provide proof-of-principle for the potential of Osgin1 to confer an engraftment advantage to FA or WT HSCs *in vivo*, the lentiviral overexpression construct for Osgin1 was utilized to assess the biological consequences of Osgin1 expression in a competitive transplantation experiment (Figure 31 A). For this purpose, CD45.1⁺ LSK bone marrow cells were isolated from either FA or WT mice and transduced *in vitro* with the overexpression construct for Osgin1 or the empty control vector respectively at a MOI ranging from 10 to 39. The details for the transduction are provided in Figure 30. Afterwards, the transduction rates of the CD45.1⁺ progeny cells were evaluated using flow cytometry and the cells were mixed according to the schematic in Figure 31 B in order to obtain a 1:1 ratio of Osgin1-overexpressing and control-transduced cells. To investigate the biological consequences of the transgene overexpression *in vivo* and to assess whether Osgin1 was able to confer a competitive advantage to HSCs, Osgin1-overexpressing and control-transduced cells were harvested from the cell culture plates and injected into the tail vein of lethally irradiated CD45.2⁺ WT mice according to the schematic in Figure 31 A.

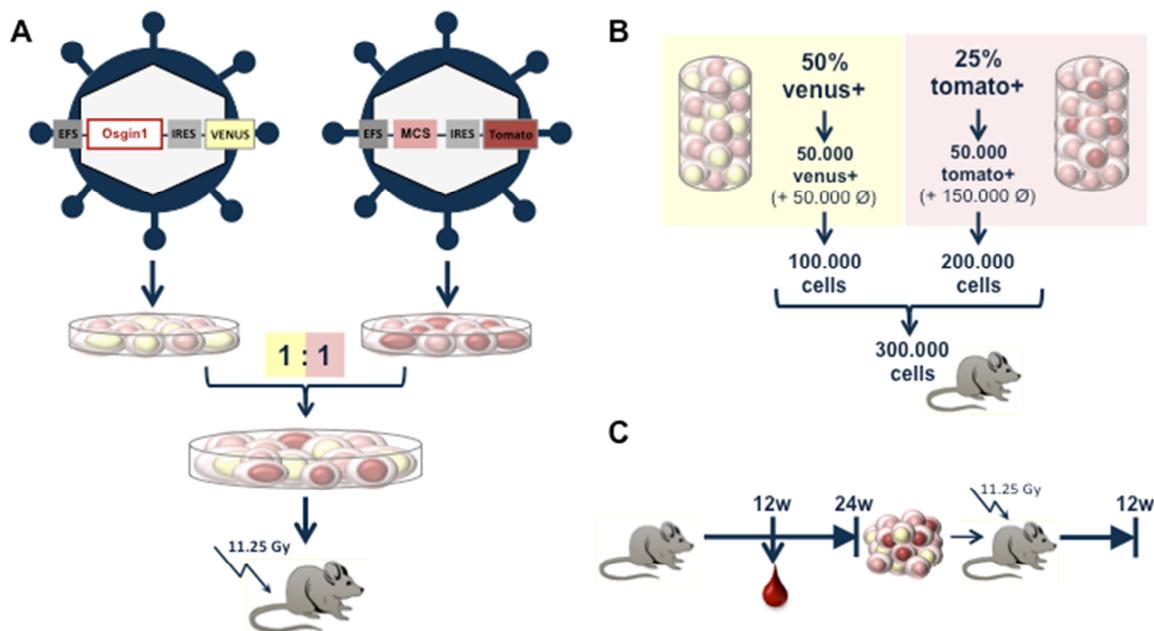


Figure 31: Overexpression of Osgin1 in a competitive transplantation setting. (A) Lentiviral vector constructs LeGO(EFS-Osgin1) iV_2 and LeGO(EFS-MCS) iT_2 were used to transduce CD45.1⁺ LSK cells, which were then mixed in a 1:1 ratio and transplanted into myeloablated CD45.2⁺ WT recipient mice. (B) Schematic example of how the transduced cells were mixed in order to normalize for unequal transduction efficiencies. (C) Time scale of the analysis of transplanted mice. The peripheral blood chimerism was evaluated at 12 and 24 weeks post transplantation and whole bone marrow cells were transplanted into secondary CD45.2⁺ recipients. These were analyzed 12 weeks post transplant.

Figure 31 C depicts the time scale of the analysis. At 12 weeks post transplant, mice were subjected to an intermediate bleed and the percentage contribution of FA and WT cells to the peripheral blood was determined by flow cytometric analysis, taking advantage of the differential expression of CD45 subtypes on the surface of transplanted CD45.1⁺ donor FA and WT leukocytes in contrast to the CD45.2⁺ recipient WT leukocytes (Figure 32 A). This analysis revealed a high peripheral blood chimerism for WT (Figure 32 B) donor cells in all six transplanted mice. In contrast, FA donor cells had engrafted the recipients to a much lower extent and only one mouse revealed a donor-graft comparable to WT levels (Figure 32 C). However, since we initially transplanted a mixture of transduced and untransduced cells, we next analyzed the ratio of Venus-positive Osgin1-overexpressing donor cells and tdTomato-positive control-transduced donor cells and compared it to the bulk of untransduced donor cells in these mice (Figure 32 D). This analysis revealed neither Venus- nor tdTomato-expressing donor cells in the peripheral blood of WT and FA mice at 12 weeks post transplant (Figure 32 E+F), which clearly indicated that the transduced cells had been outcompeted by non-transduced cells during the initial weeks after transplantation. This is likely the

result of a very low input number of transduced cells. However, we still wanted to analyze the mice at a later time point in order to check for the gene-marked chimerism in the bone marrow.

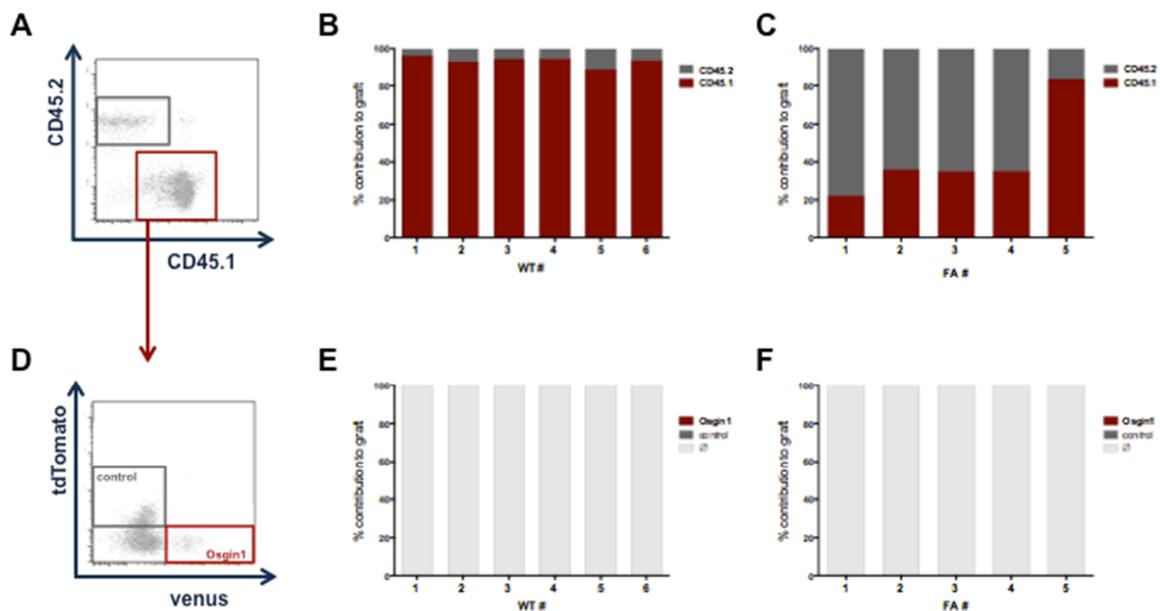


Figure 32: Analysis of the peripheral blood chimerism of FA and WT mice transplanted with a mixture of Osgin1-overexpressing and control-transduced cells at 12 weeks post transplant. (A) Gating scheme for the distinction of CD45.1⁺ donor and CD45.2⁺ recipient cells in transplanted animals. (B) Relative contribution of CD45.1⁺ WT donor cells (red) to the peripheral blood chimerism in transplanted animals. (C) Relative contribution of CD45.1⁺ FA donor cells (red) to the peripheral blood chimerism in transplanted animals. (D) Gating scheme for the distinction of Osgin1-overexpressing (Venus⁺) and control-transduced (tdTomato⁺) CD45.1⁺ donor cells in transplanted animals. (E) Relative contribution of CD45.1⁺ WT donor cell populations to the peripheral blood chimerism in transplanted animals. (F) Relative contribution of CD45.1⁺ FA donor cell populations to the peripheral blood chimerism in transplanted animals.

At 24 weeks post transplant the mice were sacrificed and bone marrow cells were harvested in order to analyze the donor chimerism (Figure 33). Similar to the analysis of the peripheral blood chimerism at 12 weeks post transplant, the analysis of the bone marrow at 24 weeks post transplant revealed different donor chimerism for WT and FA (Figure 33 A). While almost all WT mice exhibited a donor chimerism above 90% (Figure 33 B) with one exception, which showed a lower chimerism of around 40%, FA donor cells had only engrafted two individuals. However when we looked at the ratio of Venus-positive Osgin1-overexpressing donor cells and tdTomato-positive control-transduced donor cells in the bone marrow of the mice (Figure 33 D), we saw the same picture as in the analysis of the peripheral blood chimerism at 12 weeks post transplant, for both WT and FA cells (Figure 33 E+F).

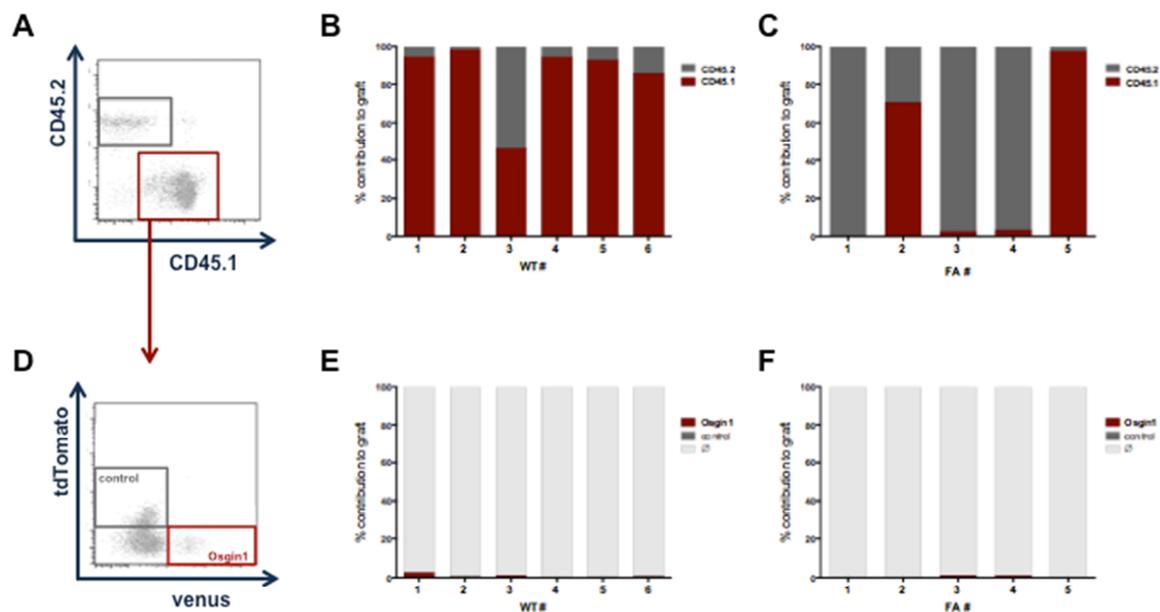


Figure 33: Analysis of the peripheral blood chimerism of FA and WT mice transplanted with a mixture of Osgin1-overexpressing and control-transduced cells at 24 weeks post transplant. (A) Gating scheme for the distinction of CD45.1⁺ donor and CD45.2⁺ recipient cells in transplanted animals. **(B)** Relative contribution of CD45.1⁺ WT donor cells (red) to the peripheral blood chimerism in transplanted animals. **(C)** Relative contribution of CD45.1⁺ FA donor cells (red) to the peripheral blood chimerism in transplanted animals. **(D)** Gating scheme for the distinction of Osgin1-overexpressing (Venus⁺) and control-transduced (tdTomato⁺) CD45.1⁺ donor cells in transplanted animals. **(E)** Relative contribution of CD45.1⁺ WT donor cell populations to the peripheral blood chimerism in transplanted animals. **(F)** Relative contribution of CD45.1⁺ FA donor cell populations to the peripheral blood chimerism in transplanted animals.

In addition, 300.000 total bone marrow cells were transplanted into irradiated secondary CD45⁺ WT recipient mice in order to restrict the analysis to a stem cell readout. With these mice, we are still waiting for the final observation endpoints at 12 weeks post transplant.

3.12 FA HSC growth defect *in vitro*

FA is a fatal hereditary disorder with hematologic abnormalities such as aplastic anemia and thrombocytopenia, which originate from severely defective HSCs at the top of the hierarchy of the hematopoietic system. This HSC defect has been studied intensively using the various FA knockout mouse models that are available to date (Geiselhart et al., 2012). HSCs isolated from these mouse knockout models are severely compromised in numbers and demonstrate a severe HSC engraftment defect, when transplanted into lethally irradiated recipient mice. However, very little is known about the mechanistic basis of this defect and the consequences on HSC fate decisions.

In order to investigate the behavior of FA HSCs and their progeny in a more simplified system, we aimed at assessing the inherent FA HSC defect *ex vivo*. In this respect, we first sought to examine the growth properties of highly purified *Fancc*^{-/-} HSCs (Lin⁻, c-Kit⁺, Sca-1⁺, CD150⁺, CD48⁻) *in vitro* over 7 days under conditions of minimal cytokine stimulation with mSCF and mTpo. Under physiologic conditions SCF is produced by niche cells, such as fibroblasts and endothelial cells and acts as the ligand for the receptor tyrosine kinase c-Kit, which is highly expressed on hematopoietic stem and progenitor cells (Ashman, 1999). Tpo, which is produced by the liver, is the primary regulator of platelet generation, but also plays an important role in controlling the HSC compartment (Hitchcock and Kaushansky, 2014). As both, SCF and Tpo, are crucial for the maintenance of HSCs *in vivo* (Kimura et al., 1998; McCarthy et al., 1977; Qian et al., 2007; Thoren et al., 2008), these two cytokines provide the basis for the survival and expansion of primitive hematopoietic cells *in vitro* (Broudy, 1997; Kaushansky, 2006). In addition, we kept the cells under serum-free culture conditions using StemSpan SFEM culture medium (Stem Cell Technologies), which is recommended for the culture and expansion of mouse hematopoietic cells.

Immediately after sorting of Lin⁻, c-Kit⁺, Sca-1⁺, CD150⁺, CD48⁻ from *Fancc*^{-/-} as well as from WT bone marrow, HSCs were plated into 96-well plates with 100 cells per well. In order to be able to follow their proliferation rates *in vitro* over the time course of 7 days, the initial input number of cells as well as the output number was determined using a flow cytometry-based technology suited for low input cell numbers. Using a defined number of fluorescence-coupled beads, which was added to the test sample and analyzed along with the cells, we were able to enumerate the relative cell counts of each sample by relating the number of cells counted to the total number of fluorescent beads analyzed. This analysis revealed that the proliferation rates of FA HSCs (36.31-fold increase) were drastically reduced in comparison to WT HSCs (83.73-fold increase), demonstrating that FA HSCs have a profound growth defect under these culture conditions (**Figure 34**).

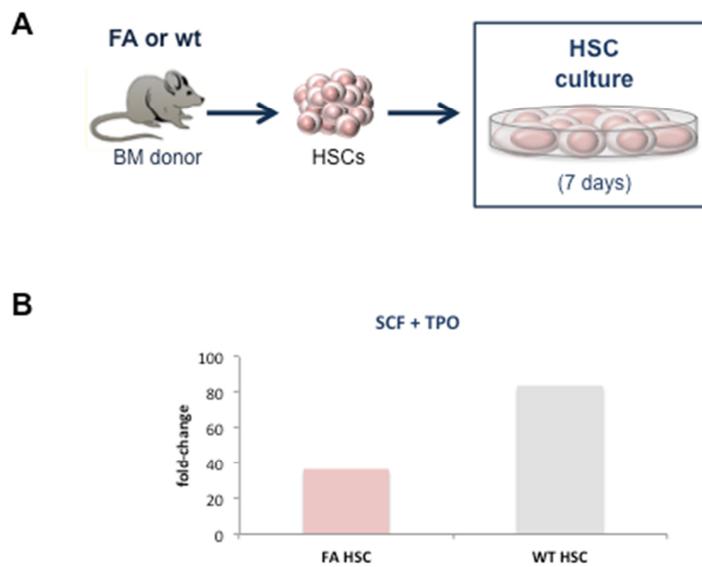


Figure 34: FA HSCs demonstrate a severe growth defect *in vitro*. (A) Schematic representation of the experimental procedure. FA or WT LT-HSCs were isolated and subjected to *in vitro* culture for 7 days. (B) Fold-change of FA (red) and WT HSC (grey) proliferation *in vitro* over 7 days.

Based on this data and in collaboration with the group of Michael Rieger at the LOEWE Center for Cell and Gene Therapy in Frankfurt we then went on to analyze the basis for this growth defect using *in vitro* cell fate tracking as developed by the group of Tim Schroeder (Schroeder, 2008; Schroeder, 2011). This method has already demonstrated great value in analyzing cell fate decisions such as the balance between self-renewal and differentiation or the instructive function of cytokines during hematopoietic lineage choice as it allows for continuous long-term tracking of individual single cells and their progeny over several days *in vitro* (Rieger, 2009; Thalheimer et al., 2014).

In order to look at the growth defect of individual FA HSCs and their progeny we have isolated *Fanca*^{-/-} LT-HSCs (Lin⁻, c-Kit⁺, Sca-1⁺, CD150⁺, CD48⁻, CD34⁻) together with the FACS core facility of the German Cancer Research Center in Heidelberg. Immediately afterwards the cells were shipped to Frankfurt and Frederic Thalheimer put them into *in vitro* culture supplemented with SCF and Tpo in order to follow their cell fate outcome using time-lapse microscopy (cell observer).

Looking at the cell fate and lineage decisions of individual FA HSCs and their progeny, we found some significant differences between WT and FA HSCs already during the first two days of culture (Figure 35 A). While WT cells were happily growing and divided after about 1.5 days (1d14h) in culture and divided even a second (2d1h) and third (2d2h) time, the FA LT-HSCs, divided only once. In comparison to WT, FA HSCs divided about 4 hours earlier (1d10h), but then underwent cell death about 10 hours after their first cell division. This striking difference is further depicted in the representative dendrogram plot (Figure 35 B), which demonstrated a single cell division event for

the FA HSCs while WT HSCs were still proliferating. Overall we have observed very high rates of cell death in FA LT-HSCs, after the first and after the second *in vitro* division (**Figure 35 C+D**), which again revealed some drastic aspects of the HSC defect. This data demonstrates that programmed cell death is a frequent cell fate decision for *Fanca*^{-/-} LT-HSCs as they are induced into cycle and is likely a major component of the strong negative selective pressure we observed when *Fanca*^{-/-} LT-HSCs are asked to reconstitute lethally irradiated recipient mice. This novel insight into the cell biology of stem cells deficient in the FA DNA repair pathway was recently published as part of a peer-reviewed manuscript, which identifies stress hematopoiesis as a major driver of bone marrow failure in this disease (Walter et al., 2015).

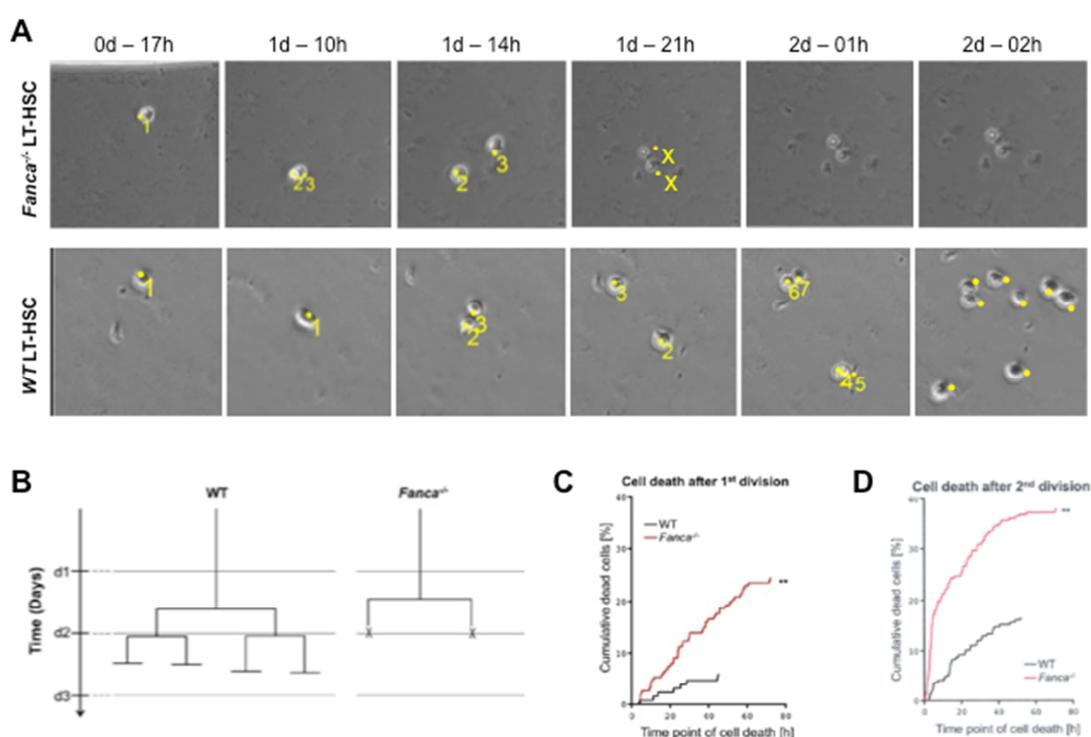


Figure 35: Cell fate of FA and WT LT-HSCs *in vitro*. (A) Representative images of the time-lapse video microscopy analysis of FACS-isolated FA and WT LT-HSCs. X indicates cell death. (B) Representative dendrogram plot for a single cell division event comparing WT and FA LT-HSCs. X indicates cell death. (C) Cumulative incidence of LT-HSC death after the first *in vitro* division. n= 126 (WT) and 187 (FA) cells tracked. (D) Cumulative incidence of LT-HSC death after the second *in vitro* division. n= 218 (WT) and 263 (FA) cells tracked. **P<0.01

3.12.1 Outlook of the FA HSC *in vitro* growth defect

Altogether the *in vitro* cell culture assay that we have developed allowed us to perform some initial analyses on the processes underlying the FA HSC defect. Initially, we found that the FA HSCs had reduced proliferation rates in the *in vitro* culture system. However, using time-lapse video microscopy, we could demonstrate that they were not as actively proliferating as their WT counterparts but instead were more rapidly dying. These data provide the basis for further analyses, which are intended to decipher the processes underlying the FA HSC defect. In this respect, this technique could serve as a potential surrogate assay to very quickly and efficiently screen for factors that are able to rescue or change the inherent FA HSC cell fate choice. Therefore, future studies evolve around overexpressing the genes, which were identified in the insertional mutagenesis screen, in the respective murine FA knockout background in order to analyze them for their potential to rescue the FA HSC defect. In addition, this will eventually allow us to add some mechanistic data to the question on how the target genes can compensate for the FA HSC defect in the insertional mutagenesis setting.

4 Discussion

4.1 Retroviral insertional mutagenesis screen in the FA HSC murine transplantation model

Retroviral insertional mutagenesis has the potential to randomly deregulate gene expression within the genome of transduced HSCs, thereby promoting clonal outgrowth of genetically modified cells that have gained a proliferative advantage over non-transduced cells (Kustikova et al., 2005) and provides the basis for genetic screens aiming to identify candidates with the potential to impact upon certain cellular processes relevant for HSC maintenance (Geiger et al., 2012; Kustikova et al., 2007). In this respect, we were the first to successfully conduct such a forward genetic screen in the context of FA HSC biology.

Although insertional mutagenesis can facilitate clonal dominance in WT HSCs, it was not clear whether it could impact upon the biology of FA knockout HSCs in a murine transplantation model. Nonetheless, we successfully transduced purified FA HSCs *in vitro* with the replication-deficient SF91 murine γ -retrovirus. In light of the findings that murine FA HSCs are prone to excessive apoptosis during *in vitro* culture conditions, which involve cytokine stimulation to induce HSC proliferation required for γ -retroviral transduction (Li et al., 2005; Walter et al., 2015), the generation of transplantable gene-modified cells was a first step towards conducting a successful insertional mutagenesis screen in the FA HSC transplant model.

Although we were able to recover the retroviral integrations from the transplanted mice, we could not perform a bioinformatics analysis for clusters of insertion sites across all mice as it was carried out in previous studies (Akagi et al., 2004). Although we had set up the screen with a relatively large number of mice, not enough integration sites could be recovered to reveal areas of the same genomic locations with a selective advantage for the cell. This is in part due to the FA HSC defect, which works as an additional selection pressure towards the successful engraftment of FA HSCs.

On the other hand, this engraftment defect allowed us to base our analysis on the selection of individual mice in which gene-modification with the virus had resulted in the outgrowth of dominant stem cell clones despite the severe FA HSC engraftment defect. We were thus able to demonstrate that retroviral insertional mutagenesis had impacted upon the FA HSC biology in the murine transplantation model.

4.2 Dominant FA HSC clones are able to overcome the inherent FA HSC defect upon retroviral insertional mutagenesis

As observed in patients undergoing clinical gene therapy trials for other hematologic diseases, dominant FA HSC clones arose in individual mice and took over the hematopoietic system as a result of the retroviral integration. In these mice, the proviral integrant had impacted upon the FA HSC biology resulting in the outgrowth of the genetically modified cell.

We were able to select twelve individual mice and evaluate the clonality of the bone marrow samples. We could demonstrate that over the course of serial transplantations, polyclonal hematopoiesis in the primary recipient was restricted towards oligo- or monoclonal hematopoiesis with the outgrowth of dominant clones. These dominant clones appear to have been able to compensate for the inherent FA HSC engraftment defect and gain a growth advantage despite their presumed genomic instability.

Although FA patients are predisposed towards malignant transformation, we could not detect any signs of leukemogenesis in primary nor secondary transplants, which demonstrates the non-malignant expansion of the dominant FA HSC clones in all selected mice.

4.2.1 Identification of retroviral integration sites in the genome of dominant FA HSC clones

The molecular analysis of the dominant clones using nrLAM-PCR helped us to identify the genomic location of the retroviral integration sites and revealed the genes that are located in the vicinity of the integrant with the potential to influence FA HSC biology. Moreover, this method provided additional proof for the mono- to oligoclonal hematopoietic reconstitution of these mice and revealed the relative contribution of each individual integration site to the genome in each mouse.

Due to the high selection pressure that we have observed in *Fanca*^{-/-} and *Fancg*^{-/-}, we focused our subsequent analysis on the three mice with evidence of clonal dominance in these transplant groups (#A1, #A2 and #G1). In these mice, we could demonstrate retroviral vector integration into the region of transcriptionally active or epigenetically accessible genes, as it has been suggested in the literature already (Bushman et al., 2005; Mitchell et al., 2004; Wu et al., 2003). In all four integration sites recovered from mice #A1, #A2 and #G1, we identified the vector integrations in genomic regions that show lower methylation than the genome-wide average in all subpopulations from HSCs to MPPs (refer to 3.8.1.2, 3.8.2.2, 3.9.1.2, 0), which may facilitate vector integration due to the transcriptionally permissive chromatin state. It is likely that these sites might be prone to the integration of retroviral vectors.

4.2.2 Dominant FA HSC clones allow the identification of deregulated genes in the vicinity of the proviral integration site

Using quantitative RT-PCR and gene expression profiling, we could demonstrate that the normal gene expression was disrupted in the vicinity of the proviral integration in mice #A2 and #G1. In this respect, we could identify *Osgin1* and *Evi1* as the likely drivers of clonal dominance in these mice.

Although, quantitative RT-PCR was able to clearly confirm the aberrant expression of *Evi1* using specific primers to distinguish transcript variant I and II, we could not recapitulate this finding with the global gene expression data, which was probably due to the fact that the analysis of the *Mecom* locus considered both *Evi1* isoforms and might not reflect what was observed using specific primers.

For mouse #A1, we were not able to confirm deregulated expression of *Taf1b* and/or *Grhl1* due to the lack of backup cell pellets for RNA isolation, however we could still select both for further functional characterization.

Dominant FA HSC clones have upregulated stress response genes to compensate the FA HSC engraftment defect

Gene expression profiling on the bone marrow of mice #G1 and #A2 provided insight into the potential mechanism underlying the compensation of the FA HSC engraftment defect. To these ends, we were able to show that both dominant FA HSC clones had upregulated genes that were part of the cellular stress response. We hypothesize that these mechanisms provided the dominant FA HSCs with an advantage in order to proliferate and engraft in the face of genome instability. In light of the recent finding, that FA HSCs suffer DNA damage as a result of their inability to tolerate and impact upon the production of intracellular reactive oxygen species and undergo apoptosis (Walter et al., 2015), the oxidative stress response would be a possible explanation of how the dominant FA HSC clones were able to overcome the inherent defect. In addition, gene ontology analysis confirmed the regulation of processes such as apoptosis, cell cycle regulation and chromosome assembly in the dominant stem cell clones of mice #G1 and #A2.

In this context, the analysis of more dominant FA HSC clones would be of interest in order to add additional evidence to our hypothesis. For this purpose, back-up cell pellets from the mice dominating in the insertional mutagenesis screen (e.g. mice #C1 – #C6) could be analyzed and compared to non-transduced bone marrow cells from FA knockout mice. Moreover, also the cellular mechanism that allowed normal HSC clones #WT1 - #WT3 to expand would be of interest in order to see whether this was distinct from the mechanism that dominant FA clones employed.

4.2.3 The dominant FA clones provide insight into disease-specific biology

Dominant FA HSC clones are able to overcome the FA HSC selection pressure and allow the identification of genes that are likely drivers of clonal dominance. However, the question arises whether the observed effects are truly disease specific and can, therefore, provide insight into the pathology of FA HSCs rather than also contributing towards clonal expansion in an insertional mutagenesis screen using WT cells.

If we evaluate the outcome of the screen globally and look at all mice that have received transduced FA HSCs, we could only identify a few mice with dominant clones due to the extremely strong selection pressure in the FA model. This clearly argues for a disease-specific outcome of the screen, as such a selection pressure has not been observed in any of the previously conducted insertional mutagenesis screens.

However, what has been observed in many WT screens already was the fact that Evi1, which led to clonal expansion in the dominant stem cell clone #G1, reliably drives clonal dominance upon retroviral insertional mutagenesis. This might mean that the mechanism of acquisition of clonal dominance in FA cells is not that different from WT cells. However, Evi1 has been found to be frequently overexpressed in AMLs from FA patients harboring 3q amplifications, suggesting that Evi1 may promote the survival and transformation of the FA stem cells (Meyer et al., 2011; Meyer et al., 2007; Raynaud et al., 1996a; Raynaud et al., 1996b; White et al., 2013). This suggests that Evi1 would indeed be able to confer a disease-specific advantage to FA HSCs despite its established role in WT hematopoiesis and might therefore be interesting to follow up on.

Moreover, we identified 3 novel hits, Taf1b, Osgin1, Grhl1, which have not been demonstrated to come up as frequent hits in WT screens (Akagi et al., 2004) and, therefore, hold the potential to provide insight into the pathology of FA or even represent targets for therapeutic intervention.

4.3 Candidate target genes with the potential to drive clonal dominance in the face of genomic instability

The candidate targets, which were identified in the screen, have demonstrated the potential to drive clonal dominance in the face of genome instability. This might have implications for therapeutic intervention in FA, but may also contribute towards the understanding of normal versus leukemic HSC biology. Therefore, we characterized the candidate target genes in order to evaluate whether these genes are relevant for normal hematopoiesis. Along those lines we were able to summarize gene expression data on HSCs under homeostatic and stress conditions as well as during development and provide epigenetic information on HSC differentiation for all four candidate target genes Osgin1,

Evi1, Taf1b and Grhl1. To functionally address the biological impact of the target genes on hematopoietic reconstitution *in vivo*, we generated lentiviral vector constructs for overexpression of the target genes in FA and WT HSCs. These could be employed for the overexpression of Osgin1 in a competitive transplantation setting.

4.3.1 Taf1b

4.3.1.1 Relevance of Taf1b expression for normal and stress hematopoiesis as well as during development

Taf1b, a component of the transcription machinery for RNA polymerase I, which was shown to contribute to promoter specificity and gene regulation in cells (Heix et al., 1997; Zomerdijk et al., 1994), has not been characterized in the context of HSC biology before. For the first time, to our knowledge, we provide a summary of different expression data for Taf1b on hematopoietic stem and progenitor cells in normal and stress hematopoiesis as well as during development.

The pronounced expression of Taf1b across all hematopoietic cell types, including HSCs, might suggest a general role for Taf1b in the maintenance of the hematopoietic compartment. However, this precludes a specific function during hematopoiesis with relevance for a certain hematopoietic population. In contrast, we could identify a functional region among LT-HSCs and MPP3/4, which might be involved in the epigenetic regulation of HSC biology during hematopoiesis as well as slightly upregulated expression levels of Taf1b in response to a stress stimulus, which might indicate some relevance for HSC stress biology, both in WT and FA HSCs.

In summary, Taf1b appears to have a broad function among various types of tissues. Although, the gene expression data of HSCs during different conditions led us to speculate that Taf1b might be equally relevant for all components of the hematopoietic compartment, functional data on its role during hematopoietic reconstitution will be needed to further support this initial speculation, especially since we lack proof for its deregulated expression in the dominant stem cell clone #A1.

4.3.2 Grhl1

4.3.2.1 Relevance of Grhl1 expression for normal and stress hematopoiesis as well as during development

Grhl1 expression, which is important for murine embryonic development as well as in the pathology of skin tumors and neuroblastoma, has not been addressed in the context of hematopoiesis yet. Here,

we provide a first summary of *Grhl1* expression in hematopoietic stem and progenitor cells under normal and stress conditions as well as during development.

Grhl1 expression is very specific to certain types of tissue such as epidermis and prostate whereas it seems absent in most other tissues. Although expressed at very low levels in the hematopoietic stem and progenitor compartment, we could demonstrate differential expression levels with a peak in MPP2 cells and a stepwise decrease along with differentiation. In addition, *Grhl1* expression was downregulated under conditions that mimic physiologic stress. This suggests that *Grhl1* might be relevant for HSC differentiation as well as during HSC stress. However, the very low expression levels have to be taken into consideration since the differential expression levels, which were detected with RNA sequencing, could also be a result of insensitivity when reaching the detection limit. Although essential for murine development, *Grhl1* is not differentially expressed during hematopoietic specification.

Taken together, we are hesitant to suggest a role for *Grhl1* during hematopoiesis but consider generating some functional data on its role during hematopoietic reconstitution in order to further address this issue.

4.3.3 *Evi1*

4.3.3.1 Relevance of *Evi1* expression for normal and stress hematopoiesis as well as during development

In the only dominant stem cell clone of the *Fancg*^{-/-} experimental group (#G1), we could identify *Evi1*. *Evi1* is one of the most commonly identified drivers of clonal dominance in WT HSCs, which serves as a positive control for our screen. The importance of *Evi1* for the regulation of hematopoiesis has already been demonstrated. Here we add additional evidence for the importance of *Evi1* during hematopoiesis as we could demonstrate differential expression of *Evi1* between HSCs and downstream progenitor populations as well as during development. In addition, several regulatory regions in the epigenome of the *Mecom* locus may have implications for the regulation of *Evi1* transcription during HSC differentiation.

Furthermore, we could provide novel insight into the role of *Evi1* in regulating the response towards physiologic stress in both WT and FA HSCs as *Evi1* expression levels were differentially regulated upon treatment with IFN- α or Tpo. In this context, *Evi1* seems to be crucial for the stress response in FA HSCs. While it is known that FA HSCs are hypersensitive towards the inhibitory action of proinflammatory cytokines (Haneline et al., 1998; Rathbun et al., 1997) and readily undergo apoptosis (Dufour et al., 2003; Rosselli et al., 1994; Wang et al., 1998), WT HSCs can tolerate the

inflammatory stress much better than FA HSCs. Under such conditions, Evi1 levels are decreased in comparison to homeostatic conditions whereas they remain stable in WT HSCs. This suggests, that Evi1 is integral to the inflammatory stress response in HSCs, however FA HSCs fail to maintain Evi1 levels. Likewise, FA HSCs demonstrate a similar defect in response to Tpo-treatment. Although Evi1 expression is increased with Tpo, FA HSCs do not reach comparable levels of Evi1 as WT HSCs. Moreover, since we could demonstrate that increased Evi1 levels could provide a clonal advantage for the dominant stem cell clone #G1 during transplantation, we conclude that Evi1 is an important player of the HSC response towards physiologic stress.

4.3.3.2 Evi1 as a driver of clonal dominance in the face of genome instability

The role of Evi1 in the pathogenesis of leukemia and MDS is well established. In addition, Evi1 has implications for FA HSC biology as it was frequently overexpressed in AMLs from FA patients (Meyer et al., 2011; Meyer et al., 2007; Raynaud et al., 1996a). Although we could verify Evi1 overexpression in the bone marrow of mouse #G1, we could not detect any signs of leukemogenesis in this mouse. This is consistent with the finding that overexpression of Evi1 is not sufficient to induce leukemia in murine models (Buonamici et al., 2004; Cuenco and Ren, 2004) and requires a collaborating hit (Haferlach et al., 2011; Maetzig et al., 2011; Modlich et al., 2005; Wolf et al., 2013). However, we were surprised that the genomic instability phenotype of the *Fancg*^{-/-} HSC, which would provide a perfect environment for oncogenic transformation, did not result in leukemogenesis in this particular mouse. In this respect it is not surprising that Evi1 confers a competitive advantage to the FA HSC clone #G1 as also cancer cells seem to exploit this mechanistic route, however it might be interesting to investigate how non-malignant expansion has been possible and which factors make the difference towards the expansion of cancer cells as it is not really known how Evi1 contributes to the transformation of cells. This finding suggests that Evi1 may impact on cell survival as a contributing component of leukemic transformation. However this remains to be investigated and we therefore interested in overexpressing Evi1 in the context of FA HSC biology using lentiviral vectors.

4.3.4 **Osgin1**

4.3.4.1 Relevance of Osgin1 expression for normal and stress hematopoiesis as well as during development

We could demonstrate that among all the genes located in the vicinity of the two retroviral integrations in mouse #A2, Osgin1 was the only one that was overexpressed. As Osgin1 has already been implicated in the protection of cells against oxidative stress (Li R, 2007; Yan et al., 2014),

which has been demonstrated to cause DNA damage and HSC attrition in FA (Walter et al., 2015), the upregulation of *Osgin1* would provide a likely explanation for the growth advantage of the dominant FA HSC clone #A2. Here we provide novel insight into the role of *Osgin1* in the hematopoietic system under homeostatic and stress conditions as well as during different developmental steps, which, to our knowledge, has not been addressed yet.

We could demonstrate that *Osgin1* is expressed in HSCs as well as in hemogenic endothelium/HSCPs during developmental specification from embryonic stem cells. This strongly indicates that *Osgin1* is relevant for the HSC compartment during homeostasis as well as during development. Furthermore, *Osgin1* is differentially regulated during the individual steps of HSC commitment, which correlates with the identification of a relatively broad and conserved regulatory region in the epigenome of the *Osgin1* locus and a differentially methylated region between the HSC and MPP3/4 compartment. These findings strongly indicate that the *Osgin1* expression in HSCs is regulated through DNA methylation and provides evidence for a role of *Osgin1* in HSC maintenance and differentiation. *Osgin1* is also expressed in FA HSCs, approximately to the same level as in WT HSCs. In response to stress, *Osgin1* levels are downregulated to comparable levels in both WT and FA HSCs, indicating that *Osgin1* is involved in the regulation of the stress response in both WT and FA HSCs. Altogether, we conclude that *Osgin1* is an essential factor for hematopoiesis as it is involved in HSC maintenance, differentiation as well as during HSC development and stress.

4.3.4.2 Biological impact of *Osgin1* overexpression in WT and FA HSCs during hematopoietic reconstitution

To functionally address the biological impact of target gene overexpression in WT and FA HSCs during hematopoietic reconstitution following transplantation, we have overexpressed *Osgin1* in a competitive transplantation experiment. However, so far we were not successful in obtaining functional data on the role of *Osgin1* during this process. Although, we were successful in generating lentiviral vectors, which allowed us to overexpress *Osgin1* in murine LSK cells, transplantation of transduced LSK cells did not contribute to our understanding of *Osgin1* function for HSC engraftment and expansion yet. This is due to several reasons, which all result from the inadequate experimental setup of the transplantation assay. First of all, the input transduction frequency of *Osgin1* as well as control-transduced cells was very low. Although WT donor cells could be detected in transplanted animals at 12 weeks post transplant, we could not detect any transduced cells in the peripheral blood regardless of whether input cells were transduced with the *Osgin1*-overexpressing or the control vector. The same held true for FA donor cells, which still contributed to the graft at 12 weeks post transplant but were negative for *Osgin1* or the control virus. This strongly implicates that the input frequency of transduced cells was indeed too low as these cells were overgrown and

depleted by the non-transduced cells in the graft. This is not surprising since in relation to the transduced cells, the unmodified cells were overrepresented, which demonstrates another reason of why this experiment failed to generate sufficient results. The competitive setting of the experiment certainly adds another layer of selection pressure, which in retrospect appears too harsh in this context. Therefore, we suggest repeating this experiment in a less-competitive manner and with higher input transduction efficiency. With this revised set-up, we would aim to analyze the effects of Osgin1 overexpression on the engraftment of transplanted LSK cells and hope to be able to obtain insight into the role of Osgin1 in this context.

4.3.4.3 Osgin1 as a driver of clonal dominance in the face of genome instability

Osgin1 has been implicated in the biology of human breast cancer (Wang et al., 2005) as well as in the carcinogenesis of liver and kidney cells (Ong et al., 2007; Ong et al., 2004). In this context, loss of Osgin1 or decreased levels was associated with an adverse outcome whereas high levels of Osgin1 were beneficial as they had an antiproliferative effect on the cancer cells. In the case of the dominant stem cell clone #A1, increased levels of Osgin1 also seemed beneficial as the normally defective cell had been able to expand without any evidence of malignant transformation. It is surprising that Osgin1 has an apoptotic effect on cancer cells while conferring an engraftment advantage to the FA HSC, which would normally undergo apoptosis and fail to engraft. As nothing is known about the effect of supraphysiologic Osgin1 levels on the hematopoietic system of both WT and FA mice, the above suggested overexpression experiments would provide an insight to this aspect. Moreover, Osgin1 was implicated in the pathophysiology of chronic inflammatory diseases such as atherosclerosis, in which Osgin1 expression is induced in response to oxidative stress in order to protect the cells from the oxidative damage (Li R, 2007; Yan et al., 2014). This aspect could help explaining the survival of the dominant FA HSC clone #A1, which could survive in the face of genomic instability, although this was recently shown to predispose the cells towards oxidative DNA damage (Walter et al., 2015). In addition, the results from the gene ontology analysis of clone #A1, which demonstrate the upregulation of cellular stress response mechanisms, add further evidence to the hypothesis on how the FA HSC clone #A1 could survive. According to these findings, it is very likely that the regulation of oxidative stress can provide a potential mechanism towards the compensation of the FA HSC defect. However, in order to fully proof this hypothesis further insight into the mechanistic regulation as well as the identification of additional key regulators will be needed.

4.3.4.4 Future studies on Osgin1

We are especially interested in the analysis of Osgin1 as our most promising candidate target gene, which seems to play a major role in the maintenance of the hematopoietic system and, in addition, indicates some relevance to FA HSC biology. The mechanistic insight obtained from previous studies on Osgin1, which demonstrate its potential as an oxidative stress response gene make us confident to further assess the role of Osgin1 in the pathology of FA. Therefore, in addition to the overexpression experiments suggested above, it would also be interesting to obtain a genetic model for Osgin1 in order to be able to analyze the hematopoietic compartment in Osgin1-overexpressing mice. However, such a mouse model is, to our knowledge, not available to date and would have to be generated in a time-consuming effort. Crossing of Osgin1-overexpressing mice with the FA knockout mouse model would generate a valuable research tool that would allow us to study the effect of Osgin1 overexpression in the context of FA HSC biology. Likewise, the effects of perturbed Osgin1 expression would be of interest to us. In this respect, a nucleotide variation in the Osgin1 gene, which was recently identified in tumor tissue from patients with hepatocellular carcinoma (Liu et al., 2014), would be of help. This less effective variant of Osgin1, which results in an amino acid substitution from arginine to histidine on codon 438, could be overexpressed in the context of WT HSC biology and analyzed for its effect on the cell fate using the *in vitro* assay that we have established. Alternatively, a short-hairpin (sh)RNA targeting Osgin1 could be cloned into the lentiviral vector construct in order to knock down Osgin1 expression in WT HSCs

4.4 **Development of an *in vitro* assay with the potential to assess the impact of candidate target genes on HSC biology**

We were successful in developing an *in vitro* assay, which not only allowed us to reveal the cell fate of FA HSCs in culture, but also provides the basis for assessing the impact of target gene overexpression on HSC fate decisions.

The insights gained from the *in vitro* cell fate tracking experiments, which we have conducted in collaboration with the group of Michael Rieger at the LOEWE Center for Cell and Gene Therapy in Frankfurt, have contributed to our understanding of the cell biology of stem cells deficient in the Fanconi anemia DNA repair pathway as recently published in a peer-reviewed manuscript (Walter et al., 2015). We could show that FA HSCs demonstrate a proliferative disadvantage when compared to WT HSCs and readily undergo apoptosis already after one day in culture. This led us to conclude that programmed cell death is a frequent cell fate decision for FA HSCs and a major driver of bone marrow failure in this disease. Moreover, the increased rates of moribund FA HSCs account for the strong negative selection pressure in the transplant setting, which we have applied in the insertional

mutagenesis screen. In this respect, the dominant FA HSC clones must have gained the ability to overcome this selection pressure by choosing an alternate and more advantageous cell fate as opposed to apoptosis in order to expand and reconstitute the hematopoietic system of the transplanted host mice. Therefore, the identified candidate target genes promise to reveal insight into the underlying mechanisms of the apoptosis escape in combination with the ability to compensate for the inherent FA HSC defect. These processes are of great interest to regenerative medicine as well as for the general understanding of HSC biology, especially in the context of hematologic disorders and ageing. Therefore, we are very much interested in investigating the potential of the candidate target genes to rescue or change the inherent FA HSC cell fate choice. In this respect, the *in vitro* cell culture assay that we have developed serves as a potential surrogate assay to overexpress the candidate target genes, which were identified in the insertional mutagenesis screen, in the respective murine FA knockout background using the lentiviral vectors that we have developed. With this we aim at providing some mechanistic insight into the compensation of the FA HSC defect in the insertional mutagenesis setting. Future work includes the inhibition of apoptosis to see how this impacts on FA cell survival in order to formally demonstrate *in vitro* and *in vivo* that apoptosis is the major route via which FA deficient cells are lost. Some of the tools, which would allow us to interrogate this are already in place. Together with Julia Detzer, a bachelor student in our lab, we have validated retroviral overexpression constructs for the anti-apoptotic Bcl2 family members Bcl2, Bcl-XL and Mcl1. In addition, conditional Bak/Bax double knockout mice, which are resistant to cell-intrinsic and extrinsic apoptosis have been obtained (<http://jaxmice.jax.org/strain/006329.html>).

4.5 Summary and outlook

The data presented in this thesis contributes to the understanding of the severe FA HSC defect, which results in hematologic abnormalities and bone marrow failure in FA patients. Using a murine FA transplantation model, we have demonstrated the possibility to compensate for this severe engraftment defect, which would normally result in the death of the FA HSCs upon transplantation. This was, to our knowledge, the first proof that FA HSCs can be rescued as a result of genomic provirus integration. Furthermore, we suggest four candidate target genes (Taf1b, Grhl1, Evi1 and Osgin1) with the potential to rescue the defective FA HSCs in the murine FA transplant setting and provide data on their relevance for HSC biology. In addition, we propose a likely mechanistic explanation of how the rescue of stem cells that would normally undergo apoptosis might have happened and provide the basis for future experiments aiming at further elucidating the mechanistic basis of FA HSC survival.

Taken together, these findings have allowed us to build a model of how the FA HSCs might have gained a growth advantage in the insertional mutagenesis screen despite the genomic instability

phenotype (**Figure 36**). As we have recently shown, during transplantation and reconstitution, stem cells suffer DNA damage as a result of the replicative stress. Normal HSCs can deal with the stress and repair the damage whereas it remains largely unresolved in FA HSCs leading to decreased HSC function and cellular depletion (Walter et al., 2015). Since we have been able to detect dominant stem cell clones in some of the retroviral-transduced FA HSCs, we hypothesize that we have rescued the FA HSC defect by deregulation of cellular genes that might play a role in the pathology of FA HSCs. According to their likely mechanism of action, we postulate that these target genes might fall into different categories depending on which mechanistic aspect of FA HSC depletion they impact upon.

A candidate target gene rescuing the phenotype by acting at the upstream part of our model indicated as target gene A (**Figure 36**), would act by limiting the amount of DNA damage that the cell was exposed to. For example, if metabolic reactive oxygen species were driving DNA damage in this setting, then the upregulation of detoxifying enzymes that eliminate the metabolic reactive oxygen species would rescue the HSC depletion phenotype.

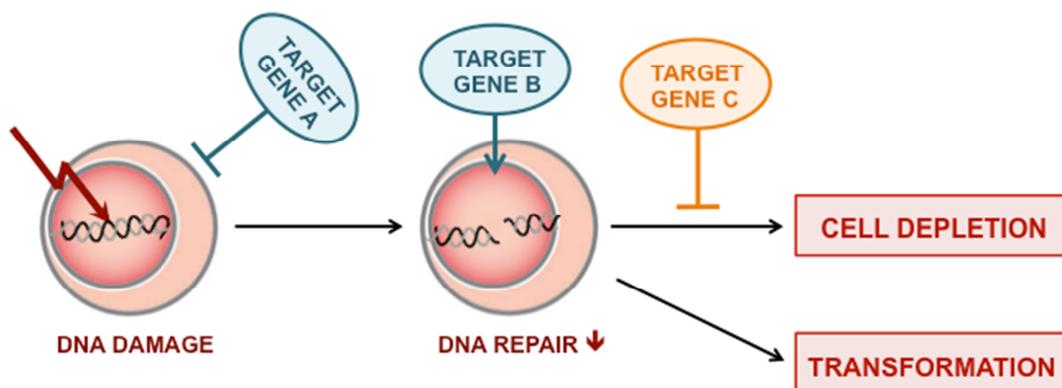


Figure 36: Schematic model of how the potential candidate target genes might rescue the FA HSC during transplantation and reconstitution. Upon transplantation, the FA HSC suffers DNA damage, which remains unrepaired due to the inactive FA DNA repair. The likely consequences are cell depletion or transformation. A potential candidate target gene may either (A) inhibit the cause of the DNA damage, (B) compensate for the DNA damage or (C) block the mechanism through which the cell is lost.

Genes acting at position B would act by enhancing the mechanism through which the DNA damage is repaired in the absence of a functional FA pathway, for example by the upregulation of a component of an alternate DNA repair pathway.

Finally candidates acting at position C would inhibit the mechanism through which the damaged cell is lost, for example by upregulation of anti-apoptotic genes such as Bcl-2 or inhibitors of p53. However, this mechanism has the likely consequence of cellular transformation as the DNA damage can be accumulated if the apoptotic escape mechanism is blocked.

The identification of such candidate target genes might have broader implications for FA patients as they could impact upon the pathology of FA HSCs. In this respect, the first two classes of target genes (A+B) may be of interest in terms of preventing the genetic instability that can lead to cell depletion or transformation. In contrast, the third class of target genes (C) may be useful to inhibit genes that promote the survival of genetically unstable cells. In this context, this would provide a suitable model for the investigation of how and why HSCs are lost.

Moreover, similar mechanisms may be in play in tumor cells, in which the FA pathway is inactivated. Indeed, in several tumor types of non-FA patients, the FA pathway has been found to be inactivated in somatic cells as opposed to the germ line loss of function seen in FA patients. This is the case in breast and ovarian cancer patients, which present with a loss of heterozygosity in the downstream pathway members like FANCD1/BRCA2, FANCN/PALB2 or FANCO/RAD51C. Therefore, the identified candidates might also have some relevance to cancer in general as the FA pathway is potentially a common tumor suppressor pathway suggesting that the tumor cells acquire genetic instability through the inactivation of the FA pathway.

5 Material and methods

5.1 Animals and animal experiments

All animal experiments were carried out in accordance with the *Guide for the Care and Use of Laboratory Animals* published by the US National Institutes of Health (NIH Publication No. 85-23, revised 1996) and the *European Community guidelines for the use of experimental animals*. All procedures involving animals were registered with and approved by the Animal Care and Use Committees of the German Regierungspräsidium Karlsruhe für Tierschutz und Arzneimittelüberwachung (TVA no. G-89/13).

5.1.1 Mouse lines

The mouse lines were bred and kept under specific pathogen-free conditions (SPF) in individually ventilated cages (IVC) at the German Cancer Research Center (DKFZ, Heidelberg) animal facility.

C57Bl/6 J (CD45.2⁺ WT mice)

C57Bl/6 J mice are referred to as CD45.2⁺ WT mice. These mice were either purchased from Harlan Laboratories or were bred at the DKFZ in-house facility.

B6.SJL-Ptprc^a Pepc^b/BoyJ (CD45.1⁺ WT mice)

B6.SJL-Ptprc^a Pepc^b/BoyJ mice are referred to as CD45.1⁺ WT mice. These mice were purchased from Charles River Laboratories, Italy or were bred at the DKFZ in-house facility.

FA knockout mouse lines

FA knockout mice are on a C57Bl/6 background and carry a knockout allele, which was generated by homologous recombination of the respective FA locus in embryonic stem cells. Therefore, the expression of the respective FA protein is repressed.

In the course of this work, knockout mice for the *Fanca*^{-/-} (Cheng et al., 2000; Wong et al., 2003), *Fancc*^{-/-} (Chen et al., 1996) and *Fancg*^{-/-} (Yang et al., 2001) proteins were obtained by breeding of heterozygous individuals. Homozygous FA knockout mice in the litter were identified by genotyping of the littermates. Therefore, total DNA isolated from mouse tail biopsies (5.8.1.2) was amplified by PCR (5.8.2.1) and analyzed by gel electrophoresis (5.10) on a 1 - 2% agarose gel.

5.1.2 Collection of murine primary material

5.1.2.1 Peripheral blood sample collection

Peripheral blood was obtained by puncturing the craniofacial capillary bed of the mice. Approximately 100 µl of peripheral blood were collected into a tube containing EDTA for blood cell assessment (5.2.1, 5.4.3.2). Approximately 0.5 – 1 ml of peripheral blood was collected when a sample of the blood cells was snap frozen (5.2.4.1) for later isolation of DNA (5.8.1.1) and analysis of vector integrations (5.9). In this case, the mice were sacrificed by cervical dislocation immediately after they had been bled.

5.1.2.2 Removal of hind legs, vertebrae and sternum

Mice were sacrificed by cervical dislocation and hind legs (femora, tibiae and ilia), vertebrae and sternum were removed and collected in ice-cold Iscove's Modified Dulbecco's Medium (IMDM; Life Technologies).

5.1.2.3 Removal of mouse lymphoid organs

From each mouse that was sacrificed during the course of this work, lymphoid organs such as spleen and thymus were checked in order to exclude hematologic abnormalities. In case of abnormal organ size, lymphoid organs were removed and kept on ice-cold IMDM (Life Technologies) for further analysis of weight and histology (5.2.3).

5.1.2.4 Collection of tail biopsies

Tail biopsies were collected from the offspring of FA knockout animals and the genomic DNA was isolated (5.8.1.2) for the purpose of genotyping (5.8.2.1).

5.1.3 Transplantation of myeloablated mice

For the transplantation of purified murine bone marrow cells (5.4) modified by viral transduction (5.6.5), host mice were lethally irradiated by whole-body irradiation with 10 Gy (2 x 550 rad) using the irradiator from the DKFZ animal facility (Buchler GmbH, Braunschweig, Cesium source). The transplantation of the respective number of cells contained in 200 µl of PBS was performed by injection into the tail vein of the myeloablated mice. Mice were kept on antibiotic (trimethoprim/sulfamethoxazole) containing water for 3 weeks post transplantation. In addition, the skin of the mouse was treated with biafine cream for three days to prevent burns.

5.2 Analysis and conservation of murine primary material

5.2.1 Analysis of the peripheral blood

The peripheral blood of mouse lines (5.1.1) and transplanted animals (5.1.3) was collected (5.1.2.1) and analyzed as described below.

5.2.1.1 Evaluation of peripheral blood cell numbers

Peripheral blood cell counts were obtained in order to detect hematologic abnormalities using a Hemavet 950 FS veterinary blood cell counting machine (Drew Scientific). Therefore, an aliquot of 30 μ l peripheral blood was analyzed. In case of abnormal white blood cell counts, peripheral blood smears (5.2.1.2) and bone marrow cytopins (5.3) were performed.

5.2.1.2 Peripheral blood smear

Peripheral blood smears were generated in order to evaluate the cellular composition as well as the morphology of the peripheral blood sample and identify abnormal cells such as blasts. Therefore, a drop of approximately 5 μ l of blood was placed on the surface of a glass slide, and spread along the surface of the glass slide in order to obtain a thin layer of blood. Glass slides were air-dried, fixed in Methanol and histologically stained with Wright's stain (Sigma-Aldrich) according to the manufacturer's instructions.

5.2.2 Purification and analysis of murine bone marrow cells

In order to purify murine bone marrow cells, hind legs (femora, tibiae and iliac crests), vertebrae and sternum were dissected by cleaning the bones from adherent soft tissue using a scalpel (Feather Safety Razor) and by removing the spinal cord. Bones were either crushed (5.2.2.1) or flushed (5.2.2.2) and the bone marrow cells were kept in 2% FCS/PBS (PAA Laboratories/Sigma-Aldrich) on ice. Murine bone marrow cells were further processed (5.4) in order to be analyzed by flow cytometry (5.5) and for their cell morphology (5.3) or transplanted into myeloablated mice (5.1.3).

5.2.2.1 Crushing of murine bone marrow cells

When bone marrow cells were pooled from multiple mice, bones were crushed in a mortar with IMDM (Life Technologies) and the cell suspension was filtered through a 40 μ m cell strainer (Becton Dickinson) and washed by spin down and resuspension in 2% FCS/PBS (PAA Laboratories/Sigma-Aldrich).

5.2.2.2 Flushing of murine bone marrow cells

When only a small sample of bone marrow cells was analyzed from an individual mouse, bones were harvested from one or two femurs by flushing the cells out of the femur into 1 or 2 ml 2% FCS/PBS (PAA Laboratories/Sigma-Aldrich) respectively using a 1 ml syringe (Terumo). The cell suspension was filtered through a 40 µm cell strainer (Becton Dickinson).

5.2.3 **Analysis of lymphoid organs and isolation of spleenocytes**

5.2.3.1 Evaluation of spleen weight

In case of abnormal size, PBS storage buffer was removed from the outside surface of the spleen by gentle dabbing on tissue paper and the weight was determined. A weight of over 200 mg, which corresponds to approximately twice its average weight, was considered as abnormal. In this case, detailed analysis of the tissue histology (5.2.3.2) was carried out and spleenocytes were isolated (5.2.3.3) for further analysis.

5.2.3.2 Histology of thymus and spleen

In case of abnormal thymus size and/or spleen weight, organs were embedded in paraffin and sections were histologically stained for analysis of tissue morphology using Wright's stain (Sigma-Aldrich) according to the manufacturer's instructions.

5.2.3.3 Isolation of spleenocytes

To isolate spleenocytes from murine spleens (5.1.2.3) and get rid of the capsule, a piece of tissue was put into a 40 µm cell strainer (Becton Dickinson) on top of a cell culture plate with PBS and meshed up with the back of a syringe (Terumo). The mesh was washed with 2% FCS/PBS (PAA Laboratories/Sigma-Aldrich) and the cell suspension was thoroughly resuspended and kept on ice for cell counts using a cell counting machine such as the Hemavet 950 FS (Drew Scientific), cell morphology (5.3) and flow cytometry (5.5.1).

5.2.4 **Freezing of peripheral blood and bone marrow cells**

5.2.4.1 Snap freezing of peripheral blood and bone marrow cells

To freeze down blood cells for later molecular analysis, aliquots of peripheral blood (5.1.2.1) and purified murine bone marrow cells (5.2.2) were lysed with ACK lysis buffer (Lonza) for 5 min at room temperature, spun down at 450 x g 4°C and resuspended in 2% FCS/PBS (PAA

Laboratories/Sigma-Aldrich). After an additional spin, the supernatant was discarded and the cell pellets were snap frozen in liquid nitrogen in order to be stored at -80°C for further use.

5.2.4.2 Freezing of viable bone marrow cells

Viable bone marrow cells were purified from the bones (5.2.2) and frozen using dimethylsulfoxide (DMSO; Sigma-Aldrich) as a cryoprotective agent. Therefore, bone marrow cells were resuspended in 2% DMSO/FCS (Sigma-Aldrich) and aliquoted into cryotubes (Thermo Scientific). Cryotubes were placed into cryogenic freezing containers (Nalgene) with 250 ml isopropanol (Sigma-Aldrich), which provide a cooling rate of 1°C per min in order to freeze down the cells slowly and reduce ice crystal formation. Cryotubes were kept in the freezing containers for approximately 24 hours at -80°C and subsequently transferred to the vapor phase of liquid nitrogen for long-term storage.

5.3 **Analysis of cell morphology using cytopins**

To analyze the cell morphology of purified bone marrow cells (5.2.2) or spleenocytes (5.2.3.3), cytopins were performed to concentrate the cells on object slides. Therefore, cytopin columns (Thermo Scientific) were set up on object slides and the cell suspension was prepared with 1×10^5 cells/ml in 10% FCS/PBS (PAA Laboratories/Sigma-Aldrich). For two final cell concentrations on the slide, 100 μ l of cell suspension corresponding to 1×10^4 cells was applied to the upper column and 200 μ l corresponding to 2×10^4 to the lower column. Cytopins were performed at 28 x g for 5 min and low acceleration in a cytopin cytocentrifuge (Thermo Scientific). Coverslips were air-dried, fixed in Methanol and histologically stained with Wright's stain (Sigma-Aldrich) according to the manufacturer's instructions.

5.4 **Identification and isolation of stem cell populations**

5.4.1 **Isolation of low-density mononuclear cells (LDMNCs)**

The LDMNCs were isolated from the purified murine bone marrow cells (5.2.2) by two steps of density gradient centrifugation using Histopaque 1083 (Sigma-Aldrich).

Therefore, an equal volume of the bone marrow cell suspension was layered onto the surface of an equal volume of the ficoll in a falcon (Greiner; Sarstedt) or polystyrene tube (Becton Dickinson). After centrifugation at room temperature and 450 x g for 20 min without break, the opaque interphase and the first parts of the ficoll containing the LDMNCs were harvested and washed with ice-cold PBS.

5.4.2 Lineage depletion of murine bone marrow cells

To further enrich the isolated LDMNC fraction (5.4.1) for the lineage-negative cells, the mature lineage-positive cells were excluded by lineage-depletion as described below.

The samples were stained with a panel of rat anti-mouse biotin-conjugated lineage markers (see table Table 5) for a maximum of 1 hour at 4°C. After two washing steps, the sample was incubated with magnetic Biotin Binder Dynabeads (Life Technologies) for 45 – 60 min at 4°C. Subsequently, the lineage-positive cells bound to these magnetic polystyrene beads, were depleted using a Dynamag-15 magnet (Life Technologies). All centrifugation steps were conducted at 450 x g for 5 min at 4°C. All wash steps were performed with ice-cold PBS supplemented with 2% FCS (PAA Laboratories).

Table 5: Antibodies for lineage depletion

Antigen	Label	Clone	Company
CD5	biotin	53-7.3	BioLegend
B220	biotin	RA3-6B2	BioLegend
CD11b	biotin	M1/70	BioLegend
CD8a	biotin	53-6.7	BioLegend
Gr-1	biotin	RB6-8C5	BioLegend
TER-119	biotin	TER-119	BioLegend

5.4.3 Cell surface stainings

In order to stain the cell surface of peripheral blood and bone marrow cells, specific monoclonal antibody stainings were performed. Therefore, the cells were incubated with the respective panel of antibodies in a 2% FCS/PBS (PAA Laboratories/Sigma-Aldrich) solution at 4°C for 1 hour. Subsequently, the cells were washed two times with 2% FCS/PBS (PAA Laboratories/Sigma-Aldrich) and incubated with the dead cell marker 7-amino actinomycin D (7AAD, Life Technologies) in order to be able to exclude the dead cells from the analysis using flow cytometry (5.5).

5.4.3.1 Stainings for the identification of stem cell populations

To prospectively isolate LSK (lin⁻, c-Kit⁺, Sca-1⁺), HSC (lin⁻, c-Kit⁺, Sca-1⁺, CD48⁻, CD150⁺) or LT-HSC (lin⁻, c-Kit⁺, Sca-1⁺, CD48⁻, CD150⁺, CD34⁻) fractions by flow sorting (5.5.2), lineage-depleted bone marrow cells (5.4.2) were stained for the respective cell surface markers using the antibody combinations outlined in Table 6.

Table 6: Antibodies for flow cytometric cell sorting of stem cell populations

Antigen	Label	Clone	Company
Streptavidin	PE	-	BD Pharming
Sca-1	APC-Cy7	D7	eBioscience
CD117 (c-Kit)	APC	2B8	eBioscience
CD48	PB	HM48-1	BioLegend
CD150	PE-Cy5	TC15-12F12.2	BioLegend
CD34	FITC	RAM34	eBioscience

Abbreviations: allophycocyanin (APC), allophycocyanin-Cyanine 7 (APC-Cy7), fluorescein isothiocyanate (FITC), pacific blue (PB), phycoerythrin (PE), phycoerythrin-Cyanine 5 (PE-Cy5)

5.4.3.2 Stainings for the identification of mature cells

For the assessment of mature peripheral blood and bone marrow cells using flow cytometric analysis (5.5.1), samples were stained for lineage markers (CD4, CD8a, CD11b and B220). In order to distinguish host from donor cells after bone marrow transplantation using flow cytometry (5.5), peripheral blood and bone marrow samples were stained for the CD45.1 and CD45.2 cell surface markers. Therefore, the combination of monoclonal antibodies as listed in Table 7 was used.

Antibody-stained peripheral blood and bone marrow samples were treated with 1 ml ACK lysing buffer (Lonza) at room temperature for 10 min to lyse erythrocytes. Subsequently, samples were washed with ice-cold PBS (Sigma-Aldrich) supplemented with 2% FCS (PAA Laboratories) and analyzed by flow cytometry (5.5).

Table 7: Antibodies for flow cytometric analysis of mature cells

Antigen	Label	Clone	Company
CD45.1	APC-Cy7	A20.1	eBioscience
CD45.2	PB	104	BioLegend
CD11b	APC	M1/70	eBioscience
B220	APC	RA3-6B2	eBioscience
CD4	PE	GK1.5	eBioscience
CD8a	PE	53-6.7	eBioscience
B220	PE	RA3-6B2	eBioscience

Abbreviations: allophycocyanin (APC), allophycocyanin-Cyanine 7 (APC-Cy7), pacific blue (PB), phycoerythrin (PE)

5.5 Flow cytometry and flow cytometric cell sorting

5.5.1 Flow cytometric analysis of cells

Cells were analyzed by flow cytometry after cell surface staining (5.4.3) and resuspension in 2% FCS/PBS (PAA Laboratories/Sigma-Aldrich) on either a LSR II (Becton Dickinson) or LSR Fortessa (Becton Dickinson) cell analyzer equipped with 350 nm, 405 nm, 488 nm, 561 nm, and 640 nm excitation lasers. Prior to analysis of multicolor samples, compensation was manually adjusted according to signals from OneComp eBeads (eBioscience) stained with single antibodies. For all experiments, dead cells were excluded by using 7AAD (Life Technologies). The analysis of the flow cytometric data was performed as described in 5.5.3.

5.5.2 Flow cytometric cell sorting

All sorting experiments were performed using a BD FACS Aria I, II or III flow cytometer (BD Biosciences) at the DKFZ Flow Cytometry Service Unit according to the gating strategy shown in **Figure 37**.

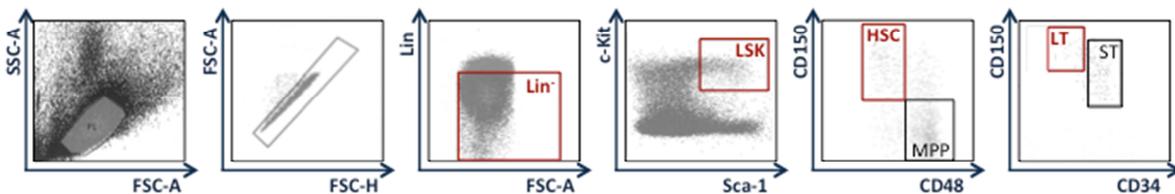


Figure 37: Schematic of the flow cytometric cell sorting to prospectively purify the respective stem cell populations from murine bone marrow. Cells were gated according to forward and side scatter characteristics and doublets were excluded in order to enrich for lineage negative (lin⁻), LSK (lin⁻, c-Kit⁺, Sca-1⁺), HSC (lin⁻, c-Kit⁺, Sca-1⁺, CD48⁻, CD150⁺) or long-term (LT)-HSC (lin⁻, c-Kit⁺, Sca-1⁺, CD48⁻, CD150⁺, CD34⁻) fractions by flow sorting. In addition, gates for multipotent progenitor cells (MPPs) and short-term (ST)-HSCs are displayed.

5.5.3 Analysis of flow cytometric data

Analysis of flow cytometric data as obtained in 5.5.1 and 5.5.2 was performed using the FlowJo software version X (TreeStar). Analysis was performed on live cells gated according to forward and side scatter characteristics.

5.6 Cultivation and manipulation of cells

5.6.1 *In vitro* culture of sorted bone marrow cells

LSK cells, HSCs or LT-HSCs were purified from murine bone marrow (5.4), resuspended in media containing 1% Pen/Strep (Sigma-Aldrich), 2% L-glutamine (Gibco) and cytokines (Peprotech) as described individually and cultivated in low attachment suspension culture plates (Cellstar) at 37°C and 5% CO₂ in a standard cell culture incubator.

5.6.2 *In vitro* culture of adherent cell lines

The adherent human embryonic kidney cell line 293T stably expressing the SV40 T-antigen and the 293T-derived phoenix-GP cell line (Nolan Laboratory) were used for virus production (5.6.3). For virus titration (5.6.4), the murine embryonic fibroblast cell line 3T3 was used.

All cell lines were cultured in DMEM growth media (Sigma-Aldrich) supplemented with 10% FCS (PAA Laboratories), 1% Pen/Strep (Sigma-Aldrich) and 1% L-glutamine (Gibco) in a standard cell culture incubator at 37°C and 5% CO₂. Confluent plates of cells were split using trypsin-EDTA (Life Technologies) and further expanded with fresh media. 293T and phoenix-GP cells were cultured on 0.1% gelatine-coated (Sigma-Aldrich) tissue culture dishes.

5.6.3 Transient calcium phosphate transfection of cells for virus production

Retroviral (5.6.3.1) and lentiviral (5.6.3.2) supernatant was generated by transient calcium phosphate transfection of the respective cell lines according to the manufacturer's protocol of the Calcium Phosphate kit (Life Technologies). Therefore, the respective cell lines were cultured as described (5.6.2). On the day before transfection, confluent plates of cells were split and the cells were seeded on plates in order to reach 70 – 80% confluency on the day of transfection. The respective transfection mixtures were prepared as individually described (5.6.3.1; 5.6.3.2). To obtain a fine precipitate of the transfection components, hepes buffered saline (contained in the kit) was thoroughly mixed with the transfection mixture and incubated for 45 min at room temperature. Meanwhile, the culture media was replaced with fresh media supplemented with 25 µM chloroquine (AppliChem) to improve the transfection efficiency. The transfection mixture was added dropwise to the cells and the transduced cells were incubated at 37°C and 5% CO₂. In the morning of the next day, the culture media was changed to fresh culture media by carefully aspirating the transfection mix and adding fresh media. In the evening, the culture media was changed to collection media, which in addition to the components of the growth media contained 1% HEPES buffer (Life Technologies) and was adjusted to pH 7.8 – 7.9. From the morning of the next day onwards, the viral supernatant

of each plate was collected and filtered using a 0.45 µm filter (Merck Millipore). Viral particles were concentrated by ultracentrifugation, resuspended in the respective media and stored at -80°C. One aliquot was used for virus titration (5.6.4)

5.6.3.1 Production of retroviral particles

Retroviral supernatant was generated by transient calcium phosphate transfection of the phoenix-GP cell line (Nolan Laboratory) with the retroviral transfection mixture for each 10 cm diameter plate of cells receiving 8 µg SF91 γ-retroviral plasmid, 10 µg Gag/Pol packaging plasmid (M57DAW), 3 µg envelope plasmid (Eco) and 2M CaCl₂. For details see 5.6.3.

In the morning after the transfection, the retroviral supernatant of each plate was collected, stored at 4°C and fresh collection media was gently added to the cells in order to be collected again in the evening. The morning as well as the evening harvests were combined, filtered using a 0.45 µm filter (Merck Millipore) and viral particles were concentrated by ultracentrifugation at 17.000 x g at 4°C overnight. In the morning of the next day, the supernatant was aspirated and the virus pellet was resuspended in IMDM (Life Technologies).

5.6.3.2 Production of lentiviral particles

Lentiviral supernatant was generated by transient calcium phosphate transfection of the 293T cells (5.6.3) with the transfection mixture containing 50 µg SF91 lentiviral plasmid, 37.5 µg Gag/Pol packaging plasmid (pSPAX2), 5 µg envelope plasmid (pMD2.G) and 2M CaCl₂. For details see 5.6.3.

In the morning after the transfection, the lentiviral supernatant of each plate was collected, centrifuged for 10 min at 450 x g and 4°C to get rid of the cell debris and filtered with the 0.45 µm filter (Merck Millipore). Lentiviral particles were concentrated by ultracentrifugation at 7.520.000 x g for 2 hours at 4°C and subsequently resuspended in StemSpan SFEM (Stem Cell Technologies) for long-term storage at -80°C.

5.6.4 Viral limiting dilution transduction of murine 3T3 cells

To evaluate the efficiency of the viral particles as generated in 5.6.3, a limiting dilution transduction of murine 3T3 cells was performed. 3T3 cells were cultured as described (5.6.2) and seeded on 6-well plates (Corning) with 100.000 cells per well. On the next day, the cells of two wells were harvested, combined and counted in order to note the average number of cells to be transduced. A limiting dilution of the virus to be tested was prepared with culture media and 8 µg/ml polybrene (Sigma-Aldrich) and each virus dilution was added to the corresponding well. In the morning of the next

day, the media of the transfected cells was changed to fresh growth media without polybrene (Sigma-Aldrich). Approximately 36 – 48 hours later, cells were harvested and resuspended in 2% FCS/PBS (PAA Laboratories/Sigma-Aldrich) in order to be analyzed using flow cytometry (5.5.1). The infectious units (IU) of the virus were calculated according to the dilution factor, the frequency of transduced cells, the count noted on the day of transduction and the total volume:

$$IU/ml = \frac{(DF \times \text{Frequency of positive cells} \times \# \text{ of cells})}{Volume}$$

5.6.5 Viral transduction of LSK cells

For the transduction of LSK cells with retroviral or lentiviral vectors as generated in 5.6.3.1 and 5.6.3.2 respectively, low-density mononuclear cells were isolated from the bone marrow (5.4.1) and sorted for the LSK fraction as described (5.5.2). The efficiency of transduced cells was assessed using flow cytometry (5.5.1) at 36 – 72 hours post transduction. In this respect, untransduced control cells were used to adjust the flow cytometry settings.

5.6.5.1 Transduction of LSK cells with retroviral vectors

For the insertional mutagenesis screen murine LSK cells were transduced with the SF91 retroviral vector. Therefore, LSK cells isolated as described (5.5.2) were pre-stimulated by *in vitro* culture in IMDM (Life Technologies) culture medium containing 1% Pen/Strep (Sigma-Aldrich), 1% L-glutamine (Gibco), 10% FCS (PAA Laboratories), 100 ng/ml recombinant murine SCF (Peprotech), 100 ng/ml recombinant murine Tpo (Peprotech) and 100 ng/ml G-CSF (Amgen), for at least 36 hours on low attachment suspension culture plates (Cellstar) at 37°C and 5% CO₂. On day 3, cells were counted using the Neubauer chamber (Braun) and 300.000 cells were transduced with the retroviral vector at a MOI in the range of 1 - 15. Transduction of the indicated cell numbers was performed in 12-well cell culture plates (Corning) pre-coated with 4 µg/cm² retronectin (TaKaRa) and with the addition of fresh cytokine-supplemented media. In detail, the virus was immobilized on the retronectin-coated plates by spin-down for 45 min at 880 x g and 24°C and the appropriate volume of thoroughly resuspended cells was added onto each well of the virus-coated plate. In addition, one well of cells was kept untransduced as a control. About 8 hours later, the same amount of virus was added to each well, except for the control well, in order to perform a second infection of LSK cells. Transduced cells were incubated until the next day at 37°C and 5% CO₂. On day 4, the cells were harvested.

5.6.5.2 Transduction of LSK cells with lentiviral vectors

For the transduction of murine LSK cells with lentiviral LeGO vectors, sorted LSK cells (5.5.2) were plated on Retronectin-coated low attachment suspension culture plates (Cellstar) and cultured with StemSpan SFEM (Stem Cell Technologies) culture medium containing 1% Pen/Strep (Sigma-Aldrich), 1% L-glutamine (Gibco) and 50 ng/ml of each recombinant murine SCF (Peprotech), recombinant murine Tpo (Peprotech) and recombinant human Flt3-L (Peprotech). Subsequently, LSK cells were transduced by adding the virus solution to each well aiming at 5×10^6 IU/ml. In addition 4 µg/ml polybrene (Sigma-Aldrich) was added to each well in order to enhance transduction and the cells were incubated at 37°C and 5% CO₂. In the morning of day 2 the culture medium was changed by replacing about half the volume of polybrene-containing medium, which is toxic to the cells, with fresh cytokine-supplemented media. In the afternoon of day 3, approximately 36 – 48 hours post transduction the cells were harvested.

5.7 *In vitro* assays

5.7.1 *In vitro* proliferation of HSCs

For the analysis of the HSC proliferation rates, FACS-isolated HSCs (5.5.2) were resuspended in StemSpan SFEM (Stem Cell Technologies) medium supplemented with 1% Pen/Strep (Sigma-Aldrich), 2% L-glutamine (Gibco) and 100 ng/ml SCF (Peprotech) and TPO (Peprotech) and counted using Greiner Terasaki plates (Sigma-Aldrich). The cell suspension was set up in triplicates with 100 cells per 96-well and 100 µl medium and the cells were cultured for 7 days at 37°C.

To determine the proliferation rates of the cells in culture over 7 days, the input number of cultured cells was compared to the output cell number using a bead-based method that allows counting of low cell numbers. Therefore, a known volume of cell suspension was analyzed along with a known concentration of beads using flow cytometry (5.5.1). By relating the number of cells counted to the total number of beads analyzed, the number of cells/µl could be determined for each sample.

5.7.2 *In vitro* cell fate tracking of LT-HSCs

For the *in vitro* cell fate tracking experiments, FACS-isolated LT-HSCs (5.5.2) were resuspended in StemSpan SFEM (Stem Cell Technologies) containing 100 ng/ml SCF (Peprotech) and TPO (Peprotech) and seeded in 24-well plates (Corning) equipped with silicon culture inserts (IBIDI). Plates were gas-tight sealed with adhesive tape (Baiersdorf) after 5% CO₂ saturation by pre-incubation in a standard cell culture incubator for 2 hours. Time-lapse microscopy was performed using a CellObserver (Zeiss) at 37°C. Phase contrast images were acquired every 2-3 min using a 10x

phase contrast objective (Zeiss) and an AxioCamHRm camera (at 1388x1040 pixel resolution) with a self-written VBA module remote controlling Zeiss AxioVision 4.8 software. Movies were assembled using QuickTime 7.1.6 software. Cell tracking was performed using a self-written computer program (TTT) (Rieger, 2009). Individual cells were observed and tracked manually until the fate of all progeny in the third generation was determined.

5.8 Molecular cell biology methods

5.8.1 DNA and RNA isolation

DNA and RNA were extracted and purified as specified below and the concentration and purity was assessed using the NanoDrop spectrophotometer (Peqlab Biotechnologie). To assess the purity of DNA and RNA, the ratio of absorbance at 260 nm and 280 nm was used. A ratio of 1.8 is generally accepted for pure DNA and a ratio of 2.0 for pure RNA. Lower ratios in either case may indicate contaminations.

5.8.1.1 DNA isolation from primary cells

DNA from snap frozen mouse peripheral blood or bone marrow cells (5.2.4.1) was isolated using the DNeasy blood & tissue kit (Qiagen) according to the manufacturer's guidelines.

5.8.1.2 DNA isolation from mouse tail biopsies

DNA from mouse tail biopsies (5.1.2.4) was isolated by digestion of the tail tissue in 500 μ l 50 nM NaOH (Sigma-Aldrich) at 95°C for 1 hour. Afterwards, 100 μ l 1M Tris (Sigma-Aldrich) was added and DNA samples were stored at 4°C.

5.8.1.3 DNA isolation from bacterial cultures

DNA from bacterial cultures (5.11.9) was isolated using the mini or maxi plasmid purification kit (Qiagen) according to the manufacturer's guidelines.

5.8.1.4 DNA isolation from agarose gels

To isolate DNA from agarose gels, which were run using gel electrophoresis (5.10), the respective gel band was cut using a scalpel (Feather Safety Razor) and extracted and purified using the QIAquick gel extraction kit (Qiagen) according to the manufacturer's guidelines.

5.8.1.5 RNA isolation from primary cells

The extraction and purification of total RNA from freshly isolated (5.2.2) or frozen bone marrow samples (5.2.4.1) was performed using the Arcturus PicoPure RNA Isolation Kit (Applied Biosystems) combined with RNase-Free DNase (Quiagen) treatment according to the manufacturer's protocol. Samples were stored at -80°C for further use.

5.8.2 PCR

PCR was used to amplify specific fragments of DNA to determine the genotype of FA knockout mice (5.1.1) or for molecular cloning purposes (5.11) as well as for the analysis of vector integrations in murine blood and bone marrow cells (5.9). For the reaction, the template DNA, a product-specific oligonucleotide pair, dNTPs (Qiagen), polymerase buffer and the polymerase were mixed and placed into a standard thermal cycler to denature, anneal primers to and elongate the respective DNA fragments using the indicated PCR conditions as described (5.8.2.1 and 5.11.1). The PCR products were analyzed on agarose gels using gel electrophoresis (5.10) or purified from the gels (5.8.1.4) for further use.

5.8.2.1 Genotyping PCRs of FA knockout mice

In order to determine the genotype of the respective FA knockout mice, genotyping PCRs were performed on genomic DNA isolated from tail biopsies (5.8.1.2) using specific primers (Table 8). PCR reactions were set up in a final volume of 20 µl using the polymerases Apex taq (Sigma-Aldrich) for the amplification of *Fanca*^{-/-} and *Fancc*^{-/-} specific products and Red taq (Sigma-Aldrich) for *Fancg*^{-/-} according to the manufacturer's guidelines. Genotyping of *Fancg*^{-/-} animals involved two separate PCRs, one for the knockout and one for the WT gene product. All PCRs were performed at an initial denaturation of 95°C for 5 min and 30 cycles of 95°C for 45 sec, 55°C for 40 sec, 72°C for 40 sec and a final extension at 72°C for 7 min in a standard thermal cycler. Amplified PCR products were analyzed by agarose gel electrophoresis (5.10).

Table 8: Oligonucleotides for genotyping of FA knockout mouse lines (5' to 3')

primer	sequence (5'-3')	Amplicon
Fanca-I	AGCCGATGTTCCAGACGCTATGC	KO = 320bp WT = 500bp
Fanca-II_rev	GGTATCTCAGGAGTTTCAGAGCAGAATCC	
Fanca-III_fwd	GCTTCCAGAGGAAGTCTTCCTTCACG	
Fancc-KO	TTGAATGGAAGGATTGGAGC	KO= 614bp WT= 819bp
Fancc-Ex8	CCTGCCATCTTCAGAATTGT	
Fancc-Buc	GAGCAACACAAATGGTAAGG	
Fancg-KO_fwd	GCATCATCGAAATTGCCGTCAACCAAGCTC	KO = 227bp
Fancg-KO_rev	TCGTGCACGCGGATTTCCGGCTCCAACAATG	
Fancg-WT_fwd	CTTGTAGAGTGAGGAGGAGTTCCCTAAGCC	WT = 400bp
Fancg-WT_rev	GGCGACAATGTCCAGCCAGGTCATTCCAGC	

5.9 Retroviral integration site analysis

5.9.1 LM-PCR

LM-PCR was used to assess the clonality of SF91-transduced murine peripheral blood and bone marrow cells and was performed essentially as described (Schmidt et al., 2001). **Figure 38** depicts a schematic outline of the procedure.

200 ng of genomic DNA isolated from murine peripheral blood and bone marrow cell pellets (5.8.1.1) were digested in two separate restriction digests with 4 U of restriction enzymes Tsp509I (New England Biolabs) or MseI (New England Biolabs) for 2 hours at 37°C and purified by standard ethanol precipitation. For primer extension the restriction-digested DNA was added to the reaction mixture containing 2U of native Pfu polymerase (Stratagene), 10X native Pfu reaction buffer (Stratagene), 200 mM dNTPs (Qiagen) and 0.25 pmol of biotinylated retroviral primer LTR I (Table 9) and the reaction was run at 95°C for 5 min, 64°C for 30 min and 72°C for 15 min in a standard PCR machine. The extension product was diluted in 450 µl H₂O and concentrated in a Microcon30 centrifugal filter device (Millipore) according to the manufacturer's guidelines. The concentrated sample was mixed with 200 µg of streptavidin-coated magnetic beads, which were prepared according to the manufacturer's instructions of the Dynabeads kilobase binder kit (Dyna) and incubated shaking at room temperature overnight. The captured DNA was retrieved using a magnetic particle concentrator, washed with H₂O, resuspended in 100 µl H₂O and ligated to the oligo linker cassette, which was generated by annealing of the denatured OC1 and OC2 (Table 9) as described (5.11.4). The ligation was performed in the presence of 2U T4 DNA ligase using the fast-link ligation kit (Cambio) for 15 min at room temperature. The sample was washed with H₂O and denatured in 0.1M NaOH for 10 min shaking at room temperature to recover the ligation product

and get rid of the biotinylation. The DNA was stored at -20°C or immediately processed in two exponential PCRs. In the first exponential PCR, 2 μl of the ligation product was amplified in a total volume of 50 μl using the Taq PCR kit (Qiagen) and 25 pmol of the retroviral primer LTRII and the linker-specific primer OCI (Table 9) at an initial denaturation of 95°C for 5 min and 30 cycles of 95°C for 1 min, 58°C for 45 sec, 72°C for 90 sec and a final extension at 72°C for 10 min. A second exponential PCR was performed to amplify 1 μl of the amplified product from the first PCR using 25 pmol of LTRIII and OCII nested primers under the reaction and cycling conditions described above. The PCR products were isolated from a 2% agarose gels (5.8.1.4) after electrophoresis (5.10) and sequenced (5.11.7) using the LTRIII and OCII primers (Table 9).

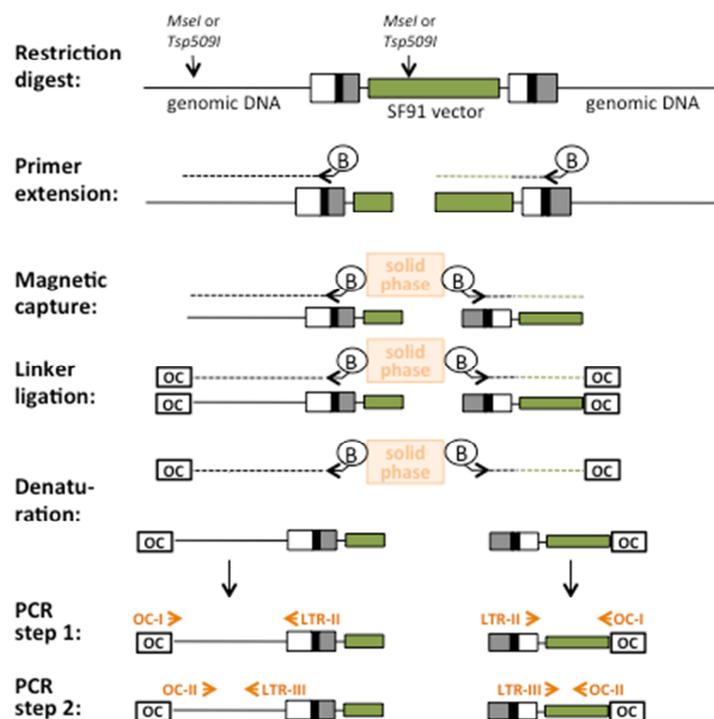


Figure 38: Schematic outline of LM-PCR to amplify SF91 vector flanking sequences in genomic DNA isolated from murine peripheral blood and bone marrow cells. After restriction digest with *MseI* or *Tsp509I*, primer extension is performed with biotinylated LTR-specific primers (B). The fragments are immobilized on paramagnetic beads that can be captured with a magnet (solid phase) and a oligonucleotide linker cassette (OC) is ligated to the unknown end of the fragments. Denaturation recovers the free nonbiotinylated DNA strand, which is exponentially amplified using LTR-specific primers (LTR-II and LTR-III) and oligo cassette primers (OC-I and OC-II) for the two exponential PCRs. (Based on Schmidt, 2001)

Table 9: Retroviral LTR oligo sequences and oligo cassette (5' to 3')

primer name	sequence (5'-3')
LTRI (bio)	GAGCTCTCTGGCTAACTAGG
LTRII	AGCTTGCCTTGAGTGCTTCA
LTRIII	AGTAGTGTGTGCCCGTCTGT
OC1	GACCCGGGAGATCTGAATTCAGTGGCACAGCAGTTAGG
OC2	CCTAACTGCTGTGCCACTGAATTCAGATC
OCI	GACCCGGGAGATCTGAATTC
OCII	AGTGGCACAGCAGTTAGG

5.9.2 nrLAM-PCR

nrLAM-PCR to identify SF91 vector flanking genomic sequences in genomic DNA isolated from murine bone marrow cell pellets (5.8.1.1) was conducted in collaboration with the group of Manfred Schmidt at the German Cancer Research Center as previously described (Paruzynski et al., 2010). **Figure 39** depicts a schematic outline of the procedure.

Briefly, 1000 ng of genomic DNA were used to amplify the vector flanking sequences by two 50-cycles linear PCR amplification steps with a biotinylated primer LTR SFFV (Table 10) hybridizing to the 5-prime region of the LTRs of the vector. Subsequent steps involved magnetic capture of the biotinylated PCR-products and ligation of the single stranded linker cassette (LC). 2 μ l of the eluate was used as template in an exponential PCR using a biotinylated vector- (LTRII SFFV) and adaptor-specific primer (LCI) (Table 10). Magnetic capture of the biotinylated PCR-products was performed and 1/5 of the eluate was used for a re-amplification by a second exponential PCR step with nested vector- and adaptor-specific primers. The PCR products were prepared for 454 pyrosequencing (Roche) as previously described (Paruzynski et al., 2010). An additional PCR with special fusion-primers carrying 454 sequencing adaptors was performed for both nrLAM-PCR products, which were pooled afterwards for the individual samples. DNA barcoding was used to allow parallel sequencing of multiple samples in a single sequencing run and the obtained data was analyzed using the semi-automated bioinformatical data mining pipeline as described (Arens et al., 2012). Sequences were trimmed by removing vector- and linker cassette-specific parts and aligned to the mouse genome using UCSC BLAT (mm10), while nearby genes and other integrating features were annotated as previously described (Cartier et al., 2009; Paruzynski et al., 2010).

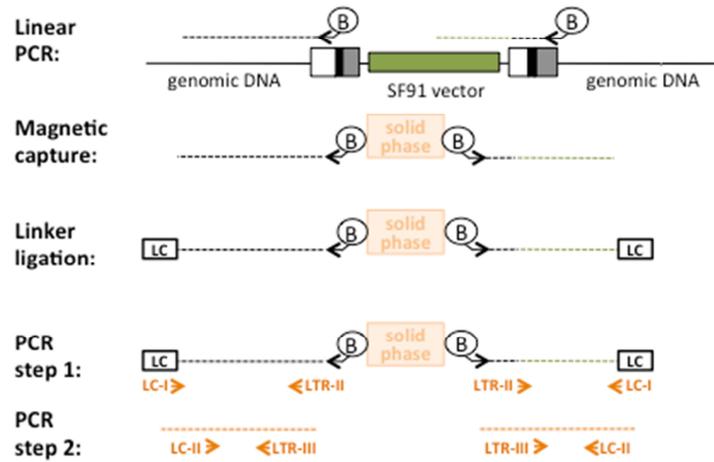


Figure 39: Schematic outline of nrLAM-PCR to amplify SF91 vector flanking sequences in genomic DNA isolated from murine bone marrow cells. After linear amplification of the vector-genome junctions with biotinylated primers (B), the fragments are immobilized on paramagnetic beads that can be captured with a magnet (solid phase) and a single-stranded linker cassette (LC) is ligated to the unknown end of the fragments. The fragments are exponentially amplified using LTR primers (LTR-II and LTR-III) and linker cassette primers (LC-I and LC-II) for the two exponential PCRs. (Based on (Gabriel et al., 2009))

Table 10: Retroviral LTR oligo sequences and linker cassette (5' to 3')

primer name	sequence (5'-3')
LTR (bio)	GTTTGGCCCAACGTTAGCTATT
LTRII (bio)	GCCCTTGATCTGAACTTCTC
LTRIII	TTCCATGCCTTGCAAATGGC
LC	P-CCTAACTGCTGTGCCACTGAATTCAGATCTCCCGGGTddC
LCI	GACCCGGGAGATCTGAATTC
LCII	GATCTGAATTCAGTGGCACAG

5.10 Agarose gel electrophoresis

PCR products were analyzed for their molecular weights on agarose gels by separation of the DNA fragments in an electric field. The gels were composed of 1 - 2% agarose (Roth) diluted in 1X TAE buffer (AppliChem) and contained ethidiumbromide (Roth). The PCR products were mixed with loading buffer containing bromphenol blue and separated at 130 V in a chamber containing TAE buffer. DNA fragments were visualized using an UV transilluminator and compared to a DNA ladder control (Invitrogen) with a band pattern of known size.

5.11 Molecular cloning procedures

Molecular cloning includes all the procedures used to produce recombinant DNA molecules and involves reactions comprising of the DNA fragment of interest as well as of a vector/plasmid backbone that allows replication in bacteria. In detail, the gene of interest was copied using PCR (5.11.1), excised out of the source DNA using restriction digest (5.11.3) or assembled from individual oligonucleotides (5.11.4). This so-called insert was ligated (5.11.6) into the linearized form of the plasmid vector and transformed into competent bacteria (5.11.8), which were grown in mini and maxi cultures (5.11.9) in order to amplify the cloned DNA. The amplified DNA was isolated from bacterial cultures (5.8.1.3) for further use. Molecular cloning requires several intermediate steps of DNA analysis on agarose gels (5.10), gel purification of the desired DNA fragment (5.8.1.4) as well as sequencing (5.11.7).

5.11.1 Cloning PCRs

PCR (5.8.2) was used to amplify a specific gene out of plasmid DNA isolated from bacterial cultures (5.8.1.3) adding specific restriction sites to the 3' and 5' ends of the DNA fragments compatible to the restriction sites in the target vector. Therefore, 25 pmol of each specific primer pair (Table 11) that was designed complementary to the respective genes was used in the reaction mix, which was set up in a final volume of 100 μ l using 0.2 mM dNTPs, 1X reaction buffer and 1U taq polymerase of the Taq PCR core kit (Qiagen). All PCRs were performed at an initial denaturation of 94°C for 5 min and 25 cycles of 94°C for 1 min, 60°C for 1 min, 70°C for 1 min in a standard thermal cycler. Amplified PCR products were analyzed by agarose gel electrophoresis (5.10), gel-purified (5.8.1.4) and TA cloned (5.11.2).

Table 11: Oligonucleotides for amplification of the EFS promoter

primer name	primer	sequence (5'-3')
EFS + NheI	fwd	AAGCTAGCATCGATTGGCTCCGGTG
EFS + BamHI	rev	GGATCCCGCGTCACGACAC

5.11.2 TOPO TA cloning

TOPO TA cloning was used to quickly and efficiently clone the gel-purified (5.8.1.4) PCR-amplified (5.11.1) products into the PCR2.1-TOPO backbone vector using the TOPO TA subcloning kit (Life Technologies). The reaction was set up according to the manufacturer's guidelines with 3 μ l of the PCR product, 1 μ l of the TOPO vector, 1 μ l salt solution and 1 μ l H₂O. All components were mixed and incubated for 30 min at room temperature. Subsequently, 5 μ l of the cloning reaction mix was transformed into competent E.coli bacteria as described (5.11.8) and

plated on pre-warmed selective plates containing 0.8 mg X-Gal and 0.8 mg IPTG required for blue/white screening of bacterial colonies. In this respect, white colonies contained plasmids with successful insertion of the PCR-amplified product and were picked for inoculation of bacterial mini cultures (5.11.9). The amplified DNA was isolated from bacterial cultures (5.8.1.3) and sent for sequencing analysis (5.11.7) in order to be able to decide for a perfect clone from which the desired DNA fragment could be excised (5.11.3).

5.11.3 Restriction digest

Restriction digests were performed to excise the genes of interest from the source DNA and linearize the self-replicating DNA plasmids prior to ligation (5.11.6) as well as for the analysis of newly ligated DNA fragments. The enzymatic digestion of the DNA fragments was performed using specific restriction endonucleases (New England Biolabs) that flank the DNA sequence and are present at the site of insertion. To prepare vector and insert for ligation, the digest was set up with 1 µg of DNA in a total volume of 100 µl. For analytical purposes 10 – 20 ng of DNA were digested in a total volume of 20 µl. In addition to the DNA, the mix containing the desired restriction enzymes (New England Biolabs) and the respective buffer (diluted to 1x) was set up according to the manufacturer's guidelines. Digestion was carried out overnight or for at least 1h for analytical purposes at the enzyme-specific reaction temperature in a heating block. Agarose gel-electrophoresis was used to analyze the products of the restriction digest (5.11.3) and desired products were isolated by gel purification (5.8.1.4).

5.11.4 Assembly of oligonucleotides

To assemble two oligonucleotides (Table 12), an equimolar mixture of both was denatured in the presence of MgCl₂ and Tris HCl for 5 min in a hot block at 95°C followed by a slow cool-down phase to room temperature and subsequent concentration using the Microcon30 centrifugal filter device (Millipore) according to the manufacturer's guidelines.

Table 12: Oligonucleotides for manufacture of the multiple cloning site (MCS)

name	sequence (5'-3')
MCS 1	GATCCTCGAGGTTTAAACGCGATCGCTCTAGAG
MCS 2	AATTCTCTAGAGCGATCGCGTTTAAACCTCGAG

5.11.5 Dephosphorylation of linearized vector DNA

To prevent self-ligation of the linearized vector in the ligation reaction (5.11.6), the vector DNA was de-phosphorylated using calf-intestinal alkaline phosphatase (Roche) treatment at 37°C for 30 min directly after the restriction digest of the vector DNA (5.11.3). To inactivate the phosphatase, the DNA was load straight onto an agarose gel and gel purified (5.8.1.4).

5.11.6 Ligation

To join the phosphodiester bonds of two DNA fragments with specific complementary end sequences as generated by restriction digest (5.11.3), a ligation reaction using T4 DNA ligase (Roche) was performed. Therefore, DNA insert and vector were mixed in an approximate 3:1 ratio and the 10X ligation buffer (Roche) was diluted to 1X with H₂O. The reaction was incubated in a PCR thermal cycler, which provides stable temperatures at 16°C for 16 hours. To amplify the ligation product, it was transformed into competent bacteria (5.11.8).

5.11.7 Sequencing of DNA fragments

Newly PCR-cloned DNA fragments (5.11.1) subcloned into the TOPO-TA vector (5.11.2) as well as LM-PCR products (5.9.1) were analyzed by sanger sequencing to confirm the correctness of the sequence and exclude errors, which may have been mistakenly introduced by the polymerase. For this purpose, 20 µl of DNA at 100 ng/µl were sent to GATC Biotech in Konstanz, who performed the sequencing reactions.

5.11.8 Transformation of competent E.coli bacteria

To amplify DNA plasmids, chemically competent E.coli bacteria, either One Shot TOP10 (Life Technologies) or Stbl3 (Life Technologies), which were specifically used for lentiviral vector backbones, were transformed by heat-shock. Therefore, competent bacteria were defrosted on ice for approximately 3 min, mixed with a small volume of DNA and incubated for 30 – 40 min on ice. The heat-shock was performed in a water bath at 42°C for 45 sec and the bacteria were put on ice for 2 min immediately afterwards. After adding 200 µl of S.O.C medium (Life Technologies), the bacteria were incubated shaking for 1 hour at 37°C and subsequently spread on pre-warmed LB-agar (Merck) plates to obtain bacterial colonies at 37°C overnight. Due to the antibiotic resistance of the plasmid, the respective antibiotic in the LB agar plate allows for the selection of plasmid-containing bacteria whereas non-transformed bacteria won't grow.

5.11.9 Bacterial mini and maxi cultures

Bacterial mini cultures were obtained by inoculation of 2 ml LB medium (Merck) with a single colony picked from the LB agar plate that was obtained after bacterial transformation (5.11.8). The bacterial culture was grown in the presence of the respective antibiotic shaking at 37°C for a maximum of 16 hours. To further expand the bacteria, the mini culture was transferred to an Erlenmeyer containing 200 ml of LB medium (Merck) and the respective antibiotic and grown overnight for a maximum of 16 hours at 37°C.

5.12 mRNA expression analysis

5.12.1 Quantitative real-time PCR

For quantitative real-time PCR, RNA extracted and purified from frozen murine bone marrow cell pellets (5.8.1.5) was reverse transcribed into cDNA according to the manufacturer's protocol of the SuperScript VILO cDNA Synthesis Kit (Life Technologies). cDNA was diluted in RNase-free water and stored at -20°C or immediately used to set up the quantitative PCR reactions.

PCR reactions were set up with 2.5 ng of cDNA in 384-well plates using primers (Table 13) at concentrations of 0.5 µM. qRT-PCR was performed and analyzed on the ViiA7 Real-Time PCR System (Applied Biosystems) using Power SYBR Green PCR Master Mix (Applied Biosystems) and amplification conditions of an initial denaturation step of 95°C for 10 min and 40 cycles of 15 sec at 95°C and 60 sec at 60°C. Expression values for each target gene were normalized against both *Oaz1* and *Sdha* and the relative gene expression was analyzed applying relative quantification methods.

5.12.2 Gene expression profiling (MicroArray)

For gene expression profiling, frozen RNA samples (5.8.1.5) were submitted to the DKFZ Genomics and Proteomics Service Unit and their personnel further processed the samples. Quality control of the RNA was performed using an Agilent Bioanalyzer 2100 in order to exclude samples with a RIN score below 7 from the analysis. Sample amplification was performed using the Ovation PicoSL WTA system and labeled with the BiotinIL Module from NuGEN. For expression analysis the samples were hybridized to the Illumina Mouse Sentrix-6_v2_r3 BeadChip Array.

The gene expression raw data was normalized by the core facility personnel and further analyzed using Microsoft Office™ Excel. The deregulated genes were listed according to their fold-changes and genes with a fold-change above a certain threshold were analyzed using the integrated software suite MetaCore (Thomson Reuters) to assess gene ontology.

Table 13: qRT-PCR primer

gene target	primer	sequence (5'-3')
Cdh13	fwd	CTGTGGGGGTCATTGTCAACT
	rev	GTTGGTCTGTGGGTTGGTGT
Hsbp1	fwd	TGCAGCAGATGCAAGACAAGT
	rev	TCCAGGTCGTCAATCCGACT
Mlycd	fwd	GCACGTCCGGGAAATGAAC
	rev	GCCTCACACTCGCTGATCTT
Osgin1	fwd	CTCTCTGGACACATCCCCTAC
	rev	GAAAGGTACTCTAGGTCCTGGT
Necab2	fwd	CGCCGCGCCGATAAAAATG
	rev	GTCACACAGTTCCTTGGTGTC
Slc38a8	fwd	TCCTCAGAGTGATCGGGGAC
	rev	GACAGGGGAAAGATGACCAGC
Mbtps1	fwd	CTGGTGGTTTTGCTCTGTGG
	rev	GGCTGTGAAGTATCCGTTGAAAG
Hsd11	fwd	CAGCCCGAAAGAGCATCAC
	rev	GCCCATCTTCCATACTGCTTGA
Gm8817	fwd	AAGGCGGACAGGAACATCAG
	rev	CAGCATGAATACAGTGGAGTCTC
Tmsb4x	fwd	ATGTCTGACAAACCCGATATGGC
	rev	CCAGCTTGCTTCTCTTGTTCA
Tlr8	fwd	GAAAACATGCCCCCTCAGTCA
	rev	CGTCACAAGGATAGCTTCTGGAA
Tlr7	fwd	ATGTGGACACGGAAGAGACAA
	rev	GGTAAGGGTAAGATTGGTGGTG
Evi1 variant I	fwd	CTTTGAATCCAAGGCAGAGC
	rev	GACAGCATGTGCTTCTCCAA
Evi1 variant II	fwd	AAGTAATGAGTGTGCCTATGGC
	rev	AGTTGACTCTCGAAGCTCAAAC
Sdha	fwd	AAGTTGAGATTTGCCGATGG
	rev	TGGTTCTGCATCGACTTCTG
Oaz1	fwd	TTTCAGCTAGCATCCTGTACTCC
	rev	GACCCTGGTCTTGTCGTTAGA

6 Bibliography

- Abramson, S., R.G. Miller, and R.A. Phillips. 1977. The identification in adult bone marrow of pluripotent and restricted stem cells of the myeloid and lymphoid systems. *J Exp Med* 145:1567-1579.
- Aguilo, F., S. Avagyan, A. Labar, A. Sevilla, D.F. Lee, P. Kumar, . . . H.W. Snoeck. 2011. Prdm16 is a physiologic regulator of hematopoietic stem cells. *Blood* 117:5057-5066.
- Aiuti, A., L. Biasco, S. Scaramuzza, F. Ferrua, M.P. Cicalese, C. Baricordi, . . . L. Naldini. 2013. Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. *Science* 341:1233-151.
- Aiuti, A., S. Slavin, M. Aker, F. Ficara, S. Deola, A. Mortellaro, . . . C. Bordignon. 2002. Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning. *Science* 296:2410-2413.
- Akagi, K., T. Suzuki, R.M. Stephens, N.A. Jenkins, and N.G. Copeland. 2004. RTCGD: retroviral tagged cancer gene database. *Nucleic Acids Res* 32:D523-527.
- Alpi, A.F., P.E. Pace, M.M. Babu, and K.J. Patel. 2008. Mechanistic insight into site-restricted monoubiquitination of FANCD2 by Ube2t, FANCL, and FANCI. *Mol Cell* 32:767-777.
- Alter, B.P. 2003. Cancer in Fanconi anemia, 1927-2001. *Cancer* 97:425-440.
- Antonchuk, J., G. Sauvageau, and R.K. Humphries. 2002. HOXB4-induced expansion of adult hematopoietic stem cells ex vivo. *Cell* 109:39-45.
- Arens, A., J.U. Appelt, C.C. Bartholomae, R. Gabriel, A. Paruzynski, D. Gustafson, . . . M. Schmidt. 2012. Bioinformatic clonality analysis of next-generation sequencing-derived viral vector integration sites. *Hum Gene Ther Methods* 23:111-118.
- Ashman, L.K. 1999. The biology of stem cell factor and its receptor C-kit. *Int J Biochem Cell Biol* 31:1037-1051.
- Auden, A., J. Caddy, T. Wilanowski, S.B. Ting, J.M. Cunningham, and S.M. Jane. 2006. Spatial and temporal expression of the Grainyhead-like transcription factor family during murine development. *Gene Expr Patterns* 6:964-970.
- Auerbach, A.D. 1993. Fanconi anemia diagnosis and the diepoxybutane (DEB) test. *Exp Hematol* 21:731-733.
- Auerbach, A.D. 2009. Fanconi anemia and its diagnosis. *Mutat Res* 668:4-10.
- Auerbach, A.D., B. Adler, R.J. O'Reilly, D. Kirkpatrick, and R.S. Chaganti. 1983. Effect of procarbazine and cyclophosphamide on chromosome breakage in Fanconi anemia cells: relevance to bone marrow transplantation. *Cancer Genet Cytogenet* 9:25-36.
- Auerbach, A.D., and R.G. Allen. 1991. Leukemia and preleukemia in Fanconi anemia patients. A review of the literature and report of the International Fanconi Anemia Registry. *Cancer Genet Cytogenet* 51:1-12.
- Auerbach, A.D., Q. Liu, R. Ghosh, M.S. Pollack, G.W. Douglas, and H.E. Broxmeyer. 1990. Prenatal identification of potential donors for umbilical cord blood transplantation for Fanconi anemia. *Transfusion* 30:682-687.
- Babovic, S., and C.J. Eaves. 2014. Hierarchical organization of fetal and adult hematopoietic stem cells. *Exp Cell Res* 329:185-191.
- Bagnara, G.P., P. Strippoli, L. Bonsi, M.F. Brizzi, G.C. Avanzi, F. Timeus, . . . et al. 1992. Effect of stem cell factor on colony growth from acquired and constitutional (Fanconi) aplastic anemia. *Blood* 80:382-387.
- Bakker, S.T., H.J. van de Vrugt, M.A. Rooimans, A.B. Oostra, J. Steltenpool, E. Delzenne-Goette, . . . J.P. de Winter. 2009. Fancm-deficient mice reveal unique features of Fanconi anemia complementation group M. *Hum Mol Genet* 18:3484-3495.
- Bakker, S.T., H.J. van de Vrugt, J.A. Visser, E. Delzenne-Goette, A. van der Wal, M.A. Berns, . . . H. te Riele. 2012. Fancf-deficient mice are prone to develop ovarian tumours. *J Pathol* 226:28-39.

- Baldrige, M.T., K.Y. King, N.C. Boles, D.C. Weksberg, and M.A. Goodell. 2010. Quiescent haematopoietic stem cells are activated by IFN-gamma in response to chronic infection. *Nature* 465:793-797.
- Baum, C., J. Dullmann, Z. Li, B. Fehse, J. Meyer, D.A. Williams, and C. von Kalle. 2003. Side effects of retroviral gene transfer into hematopoietic stem cells. *Blood* 101:2099-2114.
- Becker, A.J., C.E. Mc, and J.E. Till. 1963. Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells. *Nature* 197:452-454.
- Beug, H., A. von Kirchbach, G. Doderlein, J.F. Conscience, and T. Graf. 1979. Chicken hematopoietic cells transformed by seven strains of defective avian leukemia viruses display three distinct phenotypes of differentiation. *Cell* 18:375-390.
- Biasco, L., A. Ambrosi, D. Pellin, C. Bartholomae, I. Brigida, M.G. Roncarolo, . . . A. Aiuti. 2011. Integration profile of retroviral vector in gene therapy treated patients is cell-specific according to gene expression and chromatin conformation of target cell. *EMBO Mol Med* 3:89-101.
- Biffi, A., E. Montini, L. Lorioli, M. Cesani, F. Fumagalli, T. Plati, . . . L. Naldini. 2013. Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. *Science* 341:1233-1238.
- Bogliolo, M., B. Schuster, C. Stoepler, B. Derkunt, Y. Su, A. Raams, . . . J. Surrallés. 2013. Mutations in ERCC4, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia. *Am J Hum Genet* 92:800-806.
- Boztug, K., M. Schmidt, A. Schwarzer, P.P. Banerjee, I.A. Diez, R.A. Dewey, . . . C. Klein. 2010. Stem-cell gene therapy for the Wiskott-Aldrich syndrome. *N Engl J Med* 363:1918-1927.
- Braun, C.J., K. Boztug, A. Paruzynski, M. Witzel, A. Schwarzer, M. Rothe, . . . C. Klein. 2014. Gene therapy for Wiskott-Aldrich syndrome--long-term efficacy and genotoxicity. *Sci Transl Med* 6:227ra233.
- Bray, S.J., B. Burke, N.H. Brown, and J. Hirsh. 1989. Embryonic expression pattern of a family of Drosophila proteins that interact with a central nervous system regulatory element. *Genes Dev* 3:1130-1145.
- Bray, S.J., and F.C. Kafatos. 1991. Developmental function of Elf-1: an essential transcription factor during embryogenesis in Drosophila. *Genes Dev* 5:1672-1683.
- Brooks, D.J., S. Woodward, F.H. Thompson, B. Dos Santos, M. Russell, J.M. Yang, . . . R. Taetle. 1996. Expression of the zinc finger gene EVI-1 in ovarian and other cancers. *Br J Cancer* 74:1518-1525.
- Broudy, V.C. 1997. Stem cell factor and hematopoiesis. *Blood* 90:1345-1364.
- Buonamici, S., D. Li, Y. Chi, R. Zhao, X. Wang, L. Brace, . . . G. Nucifora. 2004. EVI1 induces myelodysplastic syndrome in mice. *J Clin Invest* 114:713-719.
- Burns, J.C., T. Friedmann, W. Driever, M. Burrascano, and J.K. Yee. 1993. Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells. *Proc Natl Acad Sci U S A* 90:8033-8037.
- Busch, K., K. Klapproth, M. Barile, M. Flossdorf, T. Holland-Letz, S.M. Schlenner, . . . H.R. Rodewald. 2015. Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. *Nature* 518:542-546.
- Bushman, F., M. Lewinski, A. Ciuffi, S. Barr, J. Leipzig, S. Hannenhalli, and C. Hoffmann. 2005. Genome-wide analysis of retroviral DNA integration. *Nat Rev Microbiol* 3:848-858.
- Butturini, A., R.P. Gale, P.C. Verlander, B. Adler-Brecher, A.P. Gillio, and A.D. Auerbach. 1994. Hematologic abnormalities in Fanconi anemia: an International Fanconi Anemia Registry study. *Blood* 84:1650-1655.
- Cabezas-Wallscheid, N., D. Klimmeck, J. Hansson, D.B. Lipka, A. Reyes, Q. Wang, . . . A. Trumpp. 2014. Identification of regulatory networks in HSCs and their immediate progeny via integrated proteome, transcriptome, and DNA Methylation analysis. *Cell Stem Cell* 15:507-522.
- Carreau, M., O.I. Gan, L. Liu, M. Doedens, C. McKerlie, J.E. Dick, and M. Buchwald. 1998. Bone marrow failure in the Fanconi anemia group C mouse model after DNA damage. *Blood* 91:2737-2744.

- Cartier, N., S. Hacein-Bey-Abina, C.C. Bartholomae, G. Veres, M. Schmidt, I. Kutschera, . . . P. Aubourg. 2009. Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science* 326:818-823.
- Cavazzana-Calvo, M., and A. Fischer. 2007. Gene therapy for severe combined immunodeficiency: are we there yet? *J Clin Invest* 117:1456-1465.
- Cavazzana-Calvo, M., S. Hacein-Bey, G. de Saint Basile, F. Gross, E. Yvon, P. Nusbaum, . . . A. Fischer. 2000. Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* 288:669-672.
- Cervenka, J., D. Arthur, and C. Yasis. 1981. Mitomycin C test for diagnostic differentiation of idiopathic aplastic anemia and Fanconi anemia. *Pediatrics* 67:119-127.
- Chen, C.C., T. Taniguchi, and A. D'Andrea. 2007. The Fanconi anemia (FA) pathway confers glioma resistance to DNA alkylating agents. *J Mol Med (Berl)* 85:497-509.
- Chen, M., D.J. Tomkins, W. Auerbach, C. McKerlie, H. Youssoufian, L. Liu, . . . M. Buchwald. 1996. Inactivation of Fac in mice produces inducible chromosomal instability and reduced fertility reminiscent of Fanconi anaemia. *Nat Genet* 12:448-451.
- Cheng, N.C., H.J. van de Vrugt, M.A. van der Valk, A.B. Oostra, P. Krimpenfort, Y. de Vries, . . . F. Arwert. 2000. Mice with a targeted disruption of the Fanconi anemia homolog Fanca. *Hum Mol Genet* 9:1805-1811.
- Cheshier, S.H., S.S. Prohaska, and I.L. Weissman. 2007. The effect of bleeding on hematopoietic stem cell cycling and self-renewal. *Stem Cells Dev* 16:707-717.
- Chuikov, S., B.P. Levi, M.L. Smith, and S.J. Morrison. 2010. Prdm16 promotes stem cell maintenance in multiple tissues, partly by regulating oxidative stress. *Nat Cell Biol* 12:999-1006.
- Coci, E.G., T. Maetzig, D. Zychlinski, M. Rothe, J.D. Suerth, C. Klein, and A. Schambach. 2015. Novel self-inactivating vectors for reconstitution of Wiskott-Aldrich Syndrome. *Current gene therapy*
- Crossan, G.P., L. van der Weyden, I.V. Rosado, F. Langevin, P.H. Gaillard, R.E. McIntyre, . . . K.J. Patel. 2011. Disruption of mouse Slx4, a regulator of structure-specific nucleases, phenocopies Fanconi anemia. *Nat Genet* 43:147-152.
- Cuenca, G.M., and R. Ren. 2004. Both AML1 and EVI1 oncogenic components are required for the cooperation of AML1/MDS1/EVI1 with BCR/ABL in the induction of acute myelogenous leukemia in mice. *Oncogene* 23:569-579.
- Dalerba, P., R.W. Cho, and M.F. Clarke. 2007. Cancer stem cells: models and concepts. *Annu Rev Med* 58:267-284.
- Daley, G.Q., M.A. Goodell, and E.Y. Snyder. 2003. Realistic prospects for stem cell therapeutics. *Hematology Am Soc Hematol Educ Program* 398-418.
- Daneshbod-Skibba, G., J. Martin, and N.T. Shahidi. 1980. Myeloid and erythroid colony growth in non-anaemic patients with Fanconi's anaemia. *Br J Haematol* 44:33-38.
- Deichmann, A., M.H. Brugman, C.C. Bartholomae, K. Schwarzwaelder, M.M. Verstegen, S.J. Howe, . . . G. Wagemaker. 2011. Insertion Sites in Engrafted Cells Cluster Within a Limited Repertoire of Genomic Areas After Gammaretroviral Vector Gene Therapy. *Mol Ther*
- Drygin, D., W.G. Rice, and I. Grummt. 2010. The RNA polymerase I transcription machinery: an emerging target for the treatment of cancer. *Annu Rev Pharmacol Toxicol* 50:131-156.
- Du, Y., S.E. Spence, N.A. Jenkins, and N.G. Copeland. 2005. Cooperating cancer-gene identification through oncogenic-retrovirus-induced insertional mutagenesis. *Blood* 106:2498-2505.
- Dufour, C., A. Corcione, J. Svahn, R. Haupt, V. Poggi, A.N. Beka'ssy, . . . V. Pistoia. 2003. TNF-alpha and IFN-gamma are overexpressed in the bone marrow of Fanconi anemia patients and TNF-alpha suppresses erythropoiesis in vitro. *Blood* 102:2053-2059.

- Essers, M.A., S. Offner, W.E. Blanco-Bose, Z. Waibler, U. Kalinke, M.A. Duchosal, and A. Trumpp. 2009. IFN α activates dormant haematopoietic stem cells in vivo. *Nature* 458:904-908.
- Fabian, J., M. Lodrini, I. Oehme, M.C. Schier, T.M. Thole, T. Hielscher, . . . H.E. Deubzer. 2014. GRHL1 acts as tumor suppressor in neuroblastoma and is negatively regulated by MYCN and HDAC3. *Cancer Res* 74:2604-2616.
- Farzin, A., S.M. Davies, F.O. Smith, A. Filipovich, M. Hansen, A.D. Auerbach, and R.E. Harris. 2007. Matched sibling donor haematopoietic stem cell transplantation in Fanconi anaemia: an update of the Cincinnati Children's experience. *Br J Haematol* 136:633-640.
- Fehse, B., O.S. Kustikova, M. Bubenheim, and C. Baum. 2004. Pois(s)on--it's a question of dose. *Gene Ther* 11:879-881.
- Ford, C.E., J.L. Hamerton, D.W. Barnes, and J.F. Loutit. 1956. Cytological identification of radiation-chimaeras. *Nature* 177:452-454.
- Friedberg, E.C., and L.B. Meira. 2006. Database of mouse strains carrying targeted mutations in genes affecting biological responses to DNA damage Version 7. *DNA Repair (Amst)* 5:189-209.
- Gabriel, R., R. Eckenberg, A. Paruzynski, C.C. Bartholomae, A. Nowrouzi, A. Arens, . . . M. Schmidt. 2009. Comprehensive genomic access to vector integration in clinical gene therapy. *Nat Med* 15:1431-1436.
- Garcia-Higuera, I., T. Taniguchi, S. Ganesan, M.S. Meyn, C. Timmers, J. Hejna, . . . A.D. D'Andrea. 2001. Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. *Mol Cell* 7:249-262.
- Gaspar, H.B., K.L. Parsley, S. Howe, D. King, K.C. Gilmour, J. Sinclair, . . . A.J. Thrasher. 2004. Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector. *Lancet* 364:2181-2187.
- Geiger, H., S.A. Pawar, E.J. Kerschen, K.J. Nattamai, I. Hernandez, H.P. Liang, . . . M. Hauer-Jensen. 2012. Pharmacological targeting of the thrombomodulin-activated protein C pathway mitigates radiation toxicity. *Nat Med*
- Geiselhart, A., A. Lier, D. Walter, and M.D. Milsom. 2012. Disrupted Signaling through the Fanconi Anemia Pathway Leads to Dysfunctional Hematopoietic Stem Cell Biology: Underlying Mechanisms and Potential Therapeutic Strategies. *Anemia* 2012:265790.
- Glass, C., M. Wilson, R. Gonzalez, Y. Zhang, and A.S. Perkins. 2014. The role of EVI1 in myeloid malignancies. *Blood Cells Mol Dis* 53:67-76.
- Gluckman, E., R. Berger, and J. Dutreix. 1984. Bone marrow transplantation for Fanconi anemia. *Semin Hematol* 21:20-26.
- Gluckman, E., A. Devergie, H. Bourdeau-Eperou, D. Thierry, R. Traineau, A. Auerbach, and H.E. Broxmeyer. 1990. Transplantation of umbilical cord blood in Fanconi's anemia. *Nouv Rev Fr Hematol* 32:423-425.
- Gonda, T.J., D.K. Sheiness, and J.M. Bishop. 1982. Transcripts from the cellular homologs of retroviral oncogenes: distribution among chicken tissues. *Mol Cell Biol* 2:617-624.
- Gratzinger, D., S. Zhao, R. West, R.V. Rouse, H. Vogel, E.C. Gil, . . . Y. Natkunam. 2009. The transcription factor LMO2 is a robust marker of vascular endothelium and vascular neoplasms and selected other entities. *Am J Clin Pathol* 131:264-278.
- Groschel, S., S. Lugthart, R.F. Schlenk, P.J. Valk, K. Eiwen, C. Goudswaard, . . . R. Delwel. 2010. High EVI1 expression predicts outcome in younger adult patients with acute myeloid leukemia and is associated with distinct cytogenetic abnormalities. *J Clin Oncol* 28:2101-2107.
- Gross, M., H. Hanenberg, S. Lobitz, R. Friedl, S. Herterich, R. Dietrich, . . . H. Hoehn. 2002. Reverse mosaicism in Fanconi anemia: natural gene therapy via molecular self-correction. *Cytogenetic and genome research* 98:126-135.
- Grummt, I. 2003. Life on a planet of its own: regulation of RNA polymerase I transcription in the nucleolus. *Genes Dev* 17:1691-1702.

- Habi, O., M.C. Delisle, N. Messier, and M. Carreau. 2005. Lack of self-renewal capacity in *Fancc*^{-/-} stem cells after ex vivo expansion. *Stem Cells* 23:1135-1141.
- Hacein-Bey-Abina, S., C. von Kalle, M. Schmidt, F. Le Deist, N. Wulffraat, E. McIntyre, . . . A. Fischer. 2003a. A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. *N Engl J Med* 348:255-256.
- Hacein-Bey-Abina, S., C. Von Kalle, M. Schmidt, M.P. McCormack, N. Wulffraat, P. Leboulch, . . . M. Cavazzana-Calvo. 2003b. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* 302:415-419.
- Hadjur, S., K. Ung, L. Wadsworth, J. Dimmick, E. Rajcan-Separovic, R.W. Scott, . . . F.R. Jirik. 2001. Defective hematopoiesis and hepatic steatosis in mice with combined deficiencies of the genes encoding *Fancc* and *Cu/Zn superoxide dismutase*. *Blood* 98:1003-1011.
- Haferlach, C., U. Bacher, T. Haferlach, F. Dicker, T. Alpermann, W. Kern, and S. Schnittger. 2011. The *inv(3)(q21q26)/t(3;3)(q21;q26)* is frequently accompanied by alterations of the *RUNX1*, *KRAS* and *NRAS* and *NF1* genes and mediates adverse prognosis both in MDS and in AML: a study in 39 cases of MDS or AML. *Leukemia* 25:874-877.
- Haneline, L.S., H.E. Broxmeyer, S. Cooper, G. Hangoc, M. Carreau, M. Buchwald, and D.W. Clapp. 1998. Multiple inhibitory cytokines induce deregulated progenitor growth and apoptosis in hematopoietic cells from *Fac*^{-/-} mice. *Blood* 91:4092-4098.
- Heix, J., J.C. Zomerdijk, A. Ravanpay, R. Tjian, and I. Grummt. 1997. Cloning of murine RNA polymerase I-specific TAF factors: conserved interactions between the subunits of the species-specific transcription initiation factor TIF-IB/SL1. *Proc Natl Acad Sci U S A* 94:1733-1738.
- Hills, D., R. Gribi, J. Ure, N. Buza-Vidas, S. Luc, S.E. Jacobsen, and A. Medvinsky. 2011. *Hoxb4*-YFP reporter mouse model: a novel tool for tracking HSC development and studying the role of *Hoxb4* in hematopoiesis. *Blood* 117:3521-3528.
- Hirano, S., K. Yamamoto, M. Ishiai, M. Yamazoe, M. Seki, N. Matsushita, . . . M. Takata. 2005. Functional relationships of *FANCC* to homologous recombination, translesion synthesis, and BLM. *Embo J* 24:418-427.
- Hitchcock, I.S., and K. Kaushansky. 2014. Thrombopoietin from beginning to end. *Br J Haematol* 165:259-268.
- Howe, S.J., M.R. Mansour, K. Schwarzwaelder, C. Bartholomae, M. Hubank, H. Kempfski, . . . A.J. Thrasher. 2008. Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J Clin Invest* 118:3143-3150.
- Howlett, N.G., T. Taniguchi, S. Olson, B. Cox, Q. Waisfisz, C. De Die-Smulders, . . . A.D. D'Andrea. 2002. Biallelic inactivation of *BRCA2* in Fanconi anemia. *Science* 297:606-609.
- Hu, J., H. Yao, F. Gan, A. Tokarski, and Y. Wang. 2012. Interaction of *OKL38* and *p53* in regulating mitochondrial structure and function. *PLoS One* 7:e43362.
- Hullett, H.R., W.A. Bonner, J. Barrett, and L.A. Herzenberg. 1969. Cell sorting: automated separation of mammalian cells as a function of intracellular fluorescence. *Science* 166:747-749.
- Huynh, H., C.Y. Ng, C.K. Ong, K.B. Lim, and T.W. Chan. 2001. Cloning and characterization of a novel pregnancy-induced growth inhibitor in mammary gland. *Endocrinology* 142:3607-3615.
- Ikuta, K., and I.L. Weissman. 1992. Evidence that hematopoietic stem cells express mouse *c-kit* but do not depend on steel factor for their generation. *Proc Natl Acad Sci U S A* 89:1502-1506.
- Imoto, I., A. Pimkhaokham, Y. Fukuda, Z.Q. Yang, Y. Shimada, N. Nomura, . . . J. Inazawa. 2001. *SNO* is a probable target for gene amplification at *3q26* in squamous-cell carcinomas of the esophagus. *Biochem Biophys Res Commun* 286:559-565.

- Jazaeri, A.A., J.S. Ferriss, J.L. Bryant, M.S. Dalton, and A. Dutta. 2010. Evaluation of EVI1 and EVI1s (Delta324) as potential therapeutic targets in ovarian cancer. *Gynecol Oncol* 118:189-195.
- Joo, W., G. Xu, N.S. Persky, A. Smogorzewska, D.G. Rudge, O. Buzovetsky, . . . N.P. Pavletich. 2011. Structure of the FANCI-FANCD2 complex: insights into the Fanconi anemia DNA repair pathway. *Science* 333:312-316.
- Kaiser, T.N., A. Lojewski, C. Dougherty, L. Juergens, E. Sahar, and S.A. Latt. 1982. Flow cytometric characterization of the response of Fanconi's anemia cells to mitomycin C treatment. *Cytometry* 2:291-297.
- Kashiyama, K., Y. Nakazawa, D.T. Pilz, C. Guo, M. Shimada, K. Sasaki, . . . T. Ogi. 2013. Malfunction of nuclease ERCC1-XPF results in diverse clinical manifestations and causes Cockayne syndrome, xeroderma pigmentosum, and Fanconi anemia. *Am J Hum Genet* 92:807-819.
- Kaushansky, K. 2006. Lineage-specific hematopoietic growth factors. *N Engl J Med* 354:2034-2045.
- Kee, Y., and A.D. D'Andrea. 2010. Expanded roles of the Fanconi anemia pathway in preserving genomic stability. *Genes Dev* 24:1680-1694.
- Keller, G. 2005. Embryonic stem cell differentiation: emergence of a new era in biology and medicine. *Genes Dev* 19:1129-1155.
- Kelly, P.F., S. Radtke, C. von Kalle, B. Balcik, K. Bohn, R. Mueller, . . . D.A. Williams. 2007. Stem cell collection and gene transfer in Fanconi anemia. *Mol Ther* 15:211-219.
- Kennedy, R.D., and A.D. D'Andrea. 2005. The Fanconi Anemia/BRCA pathway: new faces in the crowd. *Genes Dev* 19:2925-2940.
- Kiel, M.J., O.H. Yilmaz, T. Iwashita, C. Terhorst, and S.J. Morrison. 2005. SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell* 121:1109-1121.
- Kim, D.W., T. Uetsuki, Y. Kaziro, N. Yamaguchi, and S. Sugano. 1990. Use of the human elongation factor 1 alpha promoter as a versatile and efficient expression system. *Gene* 91:217-223.
- Kim, N.G., H. Rhee, L.S. Li, H. Kim, J.S. Lee, J.H. Kim, and N.K. Kim. 2002. Identification of MARCKS, FLJ11383 and TAF1B as putative novel target genes in colorectal carcinomas with microsatellite instability. *Oncogene* 21:5081-5087.
- Kim, Y., F.P. Lach, R. Desetty, H. Hanenberg, A.D. Auerbach, and A. Smogorzewska. 2011. Mutations of the SLX4 gene in Fanconi anemia. *Nat Genet* 43:142-146.
- Kimura, S., A.W. Roberts, D. Metcalf, and W.S. Alexander. 1998. Hematopoietic stem cell deficiencies in mice lacking c-Mpl, the receptor for thrombopoietin. *Proc Natl Acad Sci U S A* 95:1195-1200.
- Kohler, G., and C. Milstein. 1975. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* 256:495-497.
- Kohn, D.B. 2010. Update on gene therapy for immunodeficiencies. *Clin Immunol* 135:247-254.
- Kondo, M., A.J. Wagers, M.G. Manz, S.S. Prohaska, D.C. Scherer, G.F. Beilhack, . . . I.L. Weissman. 2003. Biology of hematopoietic stem cells and progenitors: implications for clinical application. *Annu Rev Immunol* 21:759-806.
- Kondo, N., A. Takahashi, E. Mori, T. Noda, M.Z. Zdzienicka, L.H. Thompson, . . . T. Ohnishi. 2011. FANCD1/BRCA2 plays predominant role in the repair of DNA damage induced by ACNU or TMZ. *PLoS One* 6:e19659.
- Kubbies, M., D. Schindler, H. Hoehn, A. Schinzel, and P.S. Rabinovitch. 1985. Endogenous blockage and delay of the chromosome cycle despite normal recruitment and growth phase explain poor proliferation and frequent edomitosis in Fanconi anemia cells. *Am J Hum Genet* 37:1022-1030.
- Kustikova, O., B. Fehse, U. Modlich, M. Yang, J. Dullmann, K. Kamino, . . . C. Baum. 2005. Clonal dominance of hematopoietic stem cells triggered by retroviral gene marking. *Science* 308:1171-1174.

- Kustikova, O.S., H. Geiger, Z. Li, M.H. Brugman, S.M. Chambers, C.A. Shaw, . . . C. Baum. 2007. Retroviral vector insertion sites associated with dominant hematopoietic clones mark "stemness" pathways. *Blood* 109:1897-1907.
- Kutler, D.I., B. Singh, J. Satagopan, S.D. Batish, M. Berwick, P.F. Giampietro, . . . A.D. Auerbach. 2003. A 20-year perspective on the International Fanconi Anemia Registry (IFAR). *Blood* 101:1249-1256.
- Laferte, A., E. Favry, A. Sentenac, M. Riva, C. Carles, and S. Chedin. 2006. The transcriptional activity of RNA polymerase I is a key determinant for the level of all ribosome components. *Genes Dev* 20:2030-2040.
- Langevin, F., G.P. Crossan, I.V. Rosado, M.J. Arends, and K.J. Patel. 2011. Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. *Nature* 475:53-58.
- Lapidot, T., A. Dar, and O. Kollet. 2005. How do stem cells find their way home? *Blood* 106:1901-1910.
- Lettice, L.A., S.J. Heaney, L.A. Purdie, L. Li, P. de Beer, B.A. Oostra, . . . E. de Graaff. 2003. A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum Mol Genet* 12:1725-1735.
- Li, J., D.P. Sejas, X. Zhang, Y. Qiu, K.J. Nattamai, R. Rani, . . . Q. Pang. 2007. TNF-alpha induces leukemic clonal evolution ex vivo in Fanconi anemia group C murine stem cells. *J Clin Invest* 117:3283-3295.
- Li R, C.W., Yanes R, Lee S, Berliner JA. 2007. OKL38 is an oxidative stress response gene stimulated by oxidized phospholipids. *J Lipid Res*.
- Li, X., M.M. Le Beau, S. Ciccone, F.C. Yang, B. Freie, S. Chen, . . . D.W. Clapp. 2005. Ex vivo culture of Fancd2 stem/progenitor cells predisposes cells to undergo apoptosis, and surviving stem/progenitor cells display cytogenetic abnormalities and an increased risk of malignancy. *Blood* 105:3465-3471.
- Li, Z., J. Dullmann, B. Schiedlmeier, M. Schmidt, C. von Kalle, J. Meyer, . . . C. Baum. 2002. Murine leukemia induced by retroviral gene marking. *Science* 296:497.
- Lipka, D.B., Q. Wang, N. Cabezas-Wallscheid, D. Klimmeck, D. Weichenhan, C. Herrmann, . . . C. Plass. 2014. Identification of DNA methylation changes at cis-regulatory elements during early steps of HSC differentiation using tagmentation-based whole genome bisulfite sequencing. *Cell Cycle* 13:3476-3487.
- Liu, J.M., S. Kim, E.J. Read, M. Futaki, I. Dokal, C.S. Carter, . . . C.E. Walsh. 1999. Engraftment of hematopoietic progenitor cells transduced with the Fanconi anemia group C gene (FANCC). *Hum Gene Ther* 10:2337-2346.
- Liu, M., Y. Li, L. Chen, T.H. Chan, Y. Song, L. Fu, . . . X.Y. Guan. 2014. Allele-specific imbalance of oxidative stress-induced growth inhibitor 1 associates with progression of hepatocellular carcinoma. *Gastroenterology* 146:1084-1096.
- Lorenz, E., D. Uphoff, T.R. Reid, and E. Shelton. 1951. Modification of irradiation injury in mice and guinea pigs by bone marrow injections. *J Natl Cancer Inst* 12:197-201.
- Lugthart, S., S. Groschel, H.B. Beverloo, S. Kayser, P.J. Valk, S.L. van Zelderen-Bhola, . . . H. Dohner. 2010. Clinical, molecular, and prognostic significance of WHO type inv(3)(q21q26.2)/t(3;3)(q21;q26.2) and various other 3q abnormalities in acute myeloid leukemia. *J Clin Oncol* 28:3890-3898.
- MacMillan, M.L., M.R. Hughes, S. Agarwal, and G.Q. Daley. 2011. Cellular therapy for fanconi anemia: the past, present, and future. *Biol Blood Marrow Transplant* 17:S109-114.
- Maetzig, T., M.H. Brugman, S. Bartels, N. Heinz, O.S. Kustikova, U. Modlich, . . . C. Baum. 2011. Polyclonal fluctuation of lentiviral vector-transduced and expanded murine hematopoietic stem cells. *Blood* 117:3053-3064.
- Mankad, A., T. Taniguchi, B. Cox, Y. Akkari, R.K. Rathbun, L. Lucas, . . . M. Grompe. 2006. Natural gene therapy in monozygotic twins with Fanconi anemia. *Blood* 107:3084-3090.
- McCarthy, K.F., G.D. Ledney, and R. Mitchell. 1977. A deficiency of hematopoietic stem cells in steel mice. *Cell Tissue Kinet* 10:121-126.
- Metcalf, D. 2008. Hematopoietic cytokines. *Blood* 111:485-491.

- Meyer, S., C. Bristow, M. Wappett, S. Pepper, A.D. Whetton, H. Hanenberg, . . . H. Tonnie. 2011. Fanconi anemia (FA)-associated 3q gains in leukemic transformation consistently target EVI1, but do not affect low TERC expression in FA. *Blood* 117:6047-6050.
- Meyer, S., W.D. Fergusson, A.D. Whetton, F. Moreira-Leite, S.D. Pepper, C. Miller, . . . H. Tonnie. 2007. Amplification and translocation of 3q26 with overexpression of EVI1 in Fanconi anemia-derived childhood acute myeloid leukemia with biallelic FANCD1/BRCA2 disruption. *Genes Chromosomes Cancer* 46:359-372.
- Mikkers, H., and A. Berns. 2003. Retroviral insertional mutagenesis: tagging cancer pathways. *Adv Cancer Res* 88:53-99.
- Miller, D.G., M.A. Adam, and A.D. Miller. 1990. Gene transfer by retrovirus vectors occurs only in cells that are actively replicating at the time of infection. *Mol Cell Biol* 10:4239-4242.
- Milsom, M.D., A.W. Lee, Y. Zheng, and J.A. Cancelas. 2009. Fanca^{-/-} hematopoietic stem cells demonstrate a mobilization defect which can be overcome by administration of the Rac inhibitor NSC23766. *Haematologica* 94:1011-1015.
- Mitchell, R.S., B.F. Beitzel, A.R. Schroder, P. Shinn, H. Chen, C.C. Berry, . . . F.D. Bushman. 2004. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol* 2:E234.
- Mlacki, M., C. Darido, S.M. Jane, and T. Wilanowski. 2014. Loss of Grainy head-like 1 is associated with disruption of the epidermal barrier and squamous cell carcinoma of the skin. *PLoS One* 9:e89247.
- Modlich, U., O.S. Kustikova, M. Schmidt, C. Rudolph, J. Meyer, Z. Li, . . . C. Baum. 2005. Leukemias following retroviral transfer of multidrug resistance 1 (MDR1) are driven by combinatorial insertional mutagenesis. *Blood* 105:4235-4246.
- Montini, E., D. Cesana, M. Schmidt, F. Sanvito, M. Ponzoni, C. Bartholomae, . . . L. Naldini. 2006. Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. *Nat Biotechnol* 24:687-696.
- Morishita, K., D.S. Parker, M.L. Mucenski, N.A. Jenkins, N.G. Copeland, and J.N. Ihle. 1988. Retroviral activation of a novel gene encoding a zinc finger protein in IL-3-dependent myeloid leukemia cell lines. *Cell* 54:831-840.
- Morrison, S.J., N.M. Shah, and D.J. Anderson. 1997. Regulatory mechanisms in stem cell biology. *Cell* 88:287-298.
- Mueller, L.U., Milsom, M.D., Williams, D.A. 2011. Insertional Mutagenesis Strategies in Cancer Genetics. *Springer Verlag* Chapter 6, p. 131-165.
- Muller-Sieburg, C.E., K. Townsend, I.L. Weissman, and D. Rennick. 1988. Proliferation and differentiation of highly enriched mouse hematopoietic stem cells and progenitor cells in response to defined growth factors. *J Exp Med* 167:1825-1840.
- Muller-Sieburg, C.E., C.A. Whitlock, and I.L. Weissman. 1986. Isolation of two early B lymphocyte progenitors from mouse marrow: a committed pre-pre-B cell and a clonogenic Thy-1^{lo} hematopoietic stem cell. *Cell* 44:653-662.
- Muller, L.U., M.D. Milsom, M.O. Kim, A. Schambach, T. Schuesler, and D.A. Williams. 2008. Rapid lentiviral transduction preserves the engraftment potential of Fanca^(-/-) hematopoietic stem cells. *Mol Ther* 16:1154-1160.
- Naldini, L., U. Blomer, P. Gally, D. Ory, R. Mulligan, F.H. Gage, . . . D. Trono. 1996. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* 272:263-267.
- Neveling, K., D. Endt, H. Hoehn, and D. Schindler. 2009. Genotype-phenotype correlations in Fanconi anemia. *Mutat Res* 668:73-91.
- Niedzwiedz, W., G. Mosedale, M. Johnson, C.Y. Ong, P. Pace, and K.J. Patel. 2004. The Fanconi anaemia gene FANCC promotes homologous recombination and error-prone DNA repair. *Mol Cell* 15:607-620.
- Nienhuis, A.W., C.E. Dunbar, and B.P. Sorrentino. 2006. Genotoxicity of retroviral integration in hematopoietic cells. *Mol Ther* 13:1031-1049.

- Nowell, P.C., L.J. Cole, J.G. Habermeyer, and P.L. Roan. 1956. Growth and continued function of rat marrow cells in x-radiated mice. *Cancer Res* 16:258-261.
- Nucifora, G., C.R. Begy, H. Kobayashi, D. Roulston, D. Claxton, J. Pedersen-Bjergaard, . . . J.D. Rowley. 1994. Consistent intergenic splicing and production of multiple transcripts between AML1 at 21q22 and unrelated genes at 3q26 in (3;21)(q26;q22) translocations. *Proc Natl Acad Sci U S A* 91:4004-4008.
- Ogawa, M., Y. Matsuzaki, S. Nishikawa, S. Hayashi, T. Kunisada, T. Sudo, . . . S. Nishikawa. 1991. Expression and function of c-kit in hemopoietic progenitor cells. *J Exp Med* 174:63-71.
- Ong, C.K., C. Leong, P.H. Tan, T. Van, and H. Huynh. 2007. The role of 5' untranslated region in translational suppression of OKL38 mRNA in hepatocellular carcinoma. *Oncogene* 26:1155-1165.
- Ong, C.K., C.Y. Ng, C. Leong, C.P. Ng, K.T. Foo, P.H. Tan, and H. Huynh. 2004. Genomic structure of human OKL38 gene and its differential expression in kidney carcinogenesis. *J Biol Chem* 279:743-754.
- Oram, S.H., J.A. Thoms, C. Pridans, M.E. Janes, S.J. Kinston, S. Anand, . . . B. Gottgens. 2010. A previously unrecognized promoter of LMO2 forms part of a transcriptional regulatory circuit mediating LMO2 expression in a subset of T-acute lymphoblastic leukaemia patients. *Oncogene* 29:5796-5808.
- Orkin, S.H., and L.I. Zon. 2008. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* 132:631-644.
- Osawa M, H.K., Hamada H, Nakauchi H. 1996. Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell. *Science*
- Ott, M.G., M. Schmidt, K. Schwarzwaelder, S. Stein, U. Siler, U. Koehl, . . . M. Grez. 2006. Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EV11, PRDM16 or SETBP1. *Nat Med* 12:401-409.
- Parmar, K., A. D'Andrea, and L.J. Niedernhofer. 2009. Mouse models of Fanconi anemia. *Mutat Res* 668:133-140.
- Paruzynski, A., A. Arens, R. Gabriel, C.C. Bartholomae, S. Scholz, W. Wang, . . . C. von Kalle. 2010. Genome-wide high-throughput integrome analyses by nrLAM-PCR and next-generation sequencing. *Nat Protoc* 5:1379-1395.
- Plank, J.L., and A. Dean. 2014. Enhancer function: mechanistic and genome-wide insights come together. *Mol Cell* 55:5-14.
- Pulliam, A.C., M.J. Hobson, S.L. Ciccone, Y. Li, S. Chen, E.F. Srour, . . . D.W. Clapp. 2008. AMD3100 synergizes with G-CSF to mobilize repopulating stem cells in Fanconi anemia knockout mice. *Exp Hematol* 36:1084-1090.
- Qian, H., N. Buza-Vidas, C.D. Hyland, C.T. Jensen, J. Antonchuk, R. Mansson, . . . S.E. Jacobsen. 2007. Critical role of thrombopoietin in maintaining adult quiescent hematopoietic stem cells. *Cell Stem Cell* 1:671-684.
- Rathbun, R.K., G.R. Faulkner, M.H. Ostroski, T.A. Christianson, G. Hughes, G. Jones, . . . G.C. Bagby. 1997. Inactivation of the Fanconi anemia group C gene augments interferon-gamma-induced apoptotic responses in hematopoietic cells. *Blood* 90:974-985.
- Raynaud, S., H. Cave, M. Baens, C. Bastard, V. Cacheux, J. Grosgeorge, . . . B. Grandchamp. 1996a. The 12;21 translocation involving TEL and deletion of the other TEL allele: two frequently associated alterations found in childhood acute lymphoblastic leukemia. *Blood* 87:2891-2899.
- Raynaud, S.D., M. Baens, J. Grosgeorge, K. Rodgers, C.D. Reid, M. Dainton, . . . P. Marynen. 1996b. Fluorescence in situ hybridization analysis of t(3; 12)(q26; p13): a recurring chromosomal abnormality involving the TEL gene (ETV6) in myelodysplastic syndromes. *Blood* 88:682-689.
- Renieri, A., M.A. Mencarelli, F. Cetta, M. Baldassarri, F. Mari, S. Furini, . . . E. Frullanti. 2014. Oligogenic germline mutations identified in early non-smokers lung adenocarcinoma patients. *Lung Cancer* 85:168-174.
- Reya, T., S.J. Morrison, M.F. Clarke, and I.L. Weissman. 2001. Stem cells, cancer, and cancer stem cells. *Nature* 414:105-111.

- Rieger, M.A., and T. Schroeder. 2008. Exploring hematopoiesis at single cell resolution. *Cells Tissues Organs* 188:139-149.
- Rieger, M.A.H.P.S.S., B.M.; Eitelhuber, A.C.; Schroeder, T. 2009. Hematopoietic cytokines can instruct lineage choice. *Science* 325:
- Rio, P., J.C. Segovia, H. Hanenberg, J.A. Casado, J. Martinez, K. Gottsche, . . . J.A. Bueren. 2002. In vitro phenotypic correction of hematopoietic progenitors from Fanconi anemia group A knockout mice. *Blood* 100:2032-2039.
- Roe, T., T.C. Reynolds, G. Yu, and P.O. Brown. 1993. Integration of murine leukemia virus DNA depends on mitosis. *Embo J* 12:2099-2108.
- Rosselli, F., J. Sanceau, E. Gluckman, J. Wietzerbin, and E. Moustacchi. 1994. Abnormal lymphokine production: a novel feature of the genetic disease Fanconi anemia. II. In vitro and in vivo spontaneous overproduction of tumor necrosis factor alpha. *Blood* 83:1216-1225.
- Roussel, M., S. Saule, C. Lagrou, C. Rommens, H. Beug, T. Graf, and D. Stehelin. 1979. Three new types of viral oncogene of cellular origin specific for haematopoietic cell transformation. *Nature* 281:452-455.
- Sasaki, M.S., and A. Tonomura. 1973. A high susceptibility of Fanconi's anemia to chromosome breakage by DNA cross-linking agents. *Cancer Res* 33:1829-1836.
- Schmidt, M., G. Hoffmann, M. Wissler, N. Lemke, A. Mussig, H. Glimm, . . . C. von Kalle. 2001. Detection and direct genomic sequencing of multiple rare unknown flanking DNA in highly complex samples. *Hum Gene Ther* 12:743-749.
- Schmidt, M., K. Schwarzwaelder, C. Bartholomae, K. Zaoui, C. Ball, I. Pilz, . . . C. von Kalle. 2007. High-resolution insertion-site analysis by linear amplification-mediated PCR (LAM-PCR). *Nat Methods* 4:1051-1057.
- Schmidt, M., P. Zickler, G. Hoffmann, S. Haas, M. Wissler, A. Muessig, . . . C. von Kalle. 2002. Polyclonal long-term repopulating stem cell clones in a primate model. *Blood* 100:2737-2743.
- Schoenfelder, S., I. Clay, and P. Fraser. 2010. The transcriptional interactome: gene expression in 3D. *Curr Opin Genet Dev* 20:127-133.
- Schroeder, T. 2008. Imaging stem-cell-driven regeneration in mammals. *Nature* 453:345-351.
- Schroeder, T. 2011. Long-term single-cell imaging of mammalian stem cells. *Nat Methods* 8:S30-35.
- Si, Y., A.C. Pulliam, Y. Linka, S. Ciccone, C. Leurs, J. Yuan, . . . D.W. Clapp. 2008. Overnight transduction with foamyviral vectors restores the long-term repopulating activity of Fancc^{-/-} stem cells. *Blood* 112:4458-4465.
- Spangrude, G.J., S. Heimfeld, and I.L. Weissman. 1988. Purification and characterization of mouse hematopoietic stem cells. *Science* 241:58-62.
- Stoepker, C., K. Hain, B. Schuster, Y. Hilhorst-Hofstee, M.A. Rooimans, J. Steltenpool, . . . J.P. de Winter. 2011. SLX4, a coordinator of structure-specific endonucleases, is mutated in a new Fanconi anemia subtype. *Nat Genet* 43:138-141.
- Strathdee, C.A., H. Gavish, W.R. Shannon, and M. Buchwald. 1992. Cloning of cDNAs for Fanconi's anaemia by functional complementation. *Nature* 358:434.
- Suerth, J.D., V. Labenski, and A. Schambach. 2014. Alpharetroviral vectors: from a cancer-causing agent to a useful tool for human gene therapy. *Viruses* 6:4811-4838.
- Sun, J., A. Ramos, B. Chapman, J.B. Johnnidis, L. Le, Y.J. Ho, . . . F.D. Camargo. 2014. Clonal dynamics of native haematopoiesis. *Nature* 514:322-327.
- Takizawa, H., R.R. Regoes, C.S. Boddupalli, S. Bonhoeffer, and M.G. Manz. 2011. Dynamic variation in cycling of hematopoietic stem cells in steady state and inflammation. *J Exp Med* 208:273-284.
- Tan, P.L., J.E. Wagner, A.D. Auerbach, T.E. Defor, A. Slungaard, and M.L. Macmillan. 2006. Successful engraftment without radiation after fludarabine-based regimen in Fanconi anemia patients undergoing genotypically identical donor hematopoietic cell transplantation. *Pediatric blood & cancer* 46:630-636.

- Thalheimer, F.B., S. Wingert, P. De Giacomo, N. Haetscher, M. Rehage, B. Brill, . . . M.A. Rieger. 2014. Cytokine-regulated GADD45G induces differentiation and lineage selection in hematopoietic stem cells. *Stem Cell Reports* 3:34-43.
- Thoren, L.A., K. Liuba, D. Bryder, J.M. Nygren, C.T. Jensen, H. Qian, . . . S.E. Jacobsen. 2008. Kit regulates maintenance of quiescent hematopoietic stem cells. *J Immunol* 180:2045-2053.
- Till, J.E., and E.A. McCulloch. 1961. A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiat Res* 14:213-222.
- Till, J.E., and E.A. McCulloch. 1980. Hemopoietic stem cell differentiation. *Biochim Biophys Acta* 605:431-459.
- Till, J.E., E.A. McCulloch, and L. Siminovitch. 1964. A Stochastic Model of Stem Cell Proliferation, Based on the Growth of Spleen Colony-Forming Cells. *Proc Natl Acad Sci U S A* 51:29-36.
- Trumpp, A., M. Essers, and A. Wilson. 2010. Awakening dormant haematopoietic stem cells. *Nat Rev Immunol* 10:201-209.
- Uren, A.G., J. Kool, A. Berns, and M. van Lohuizen. 2005. Retroviral insertional mutagenesis: past, present and future. *Oncogene* 24:7656-7672.
- Vaz, F., H. Hanenberg, B. Schuster, K. Barker, C. Wiek, V. Erven, . . . C.G. Mathew. 2010. Mutation of the RAD51C gene in a Fanconi anemia-like disorder. *Nat Genet* 42:406-409.
- Waisfisz, Q., N.V. Morgan, M. Savino, J.P. de Winter, C.G. van Berkel, M.E. Hoatlin, . . . H. Joenje. 1999. Spontaneous functional correction of homozygous fanconi anaemia alleles reveals novel mechanistic basis for reverse mosaicism. *Nat Genet* 22:379-383.
- Walsh, C.E., M. Grompe, E. Vanin, M. Buchwald, N.S. Young, A.W. Nienhuis, and J.M. Liu. 1994. A functionally active retrovirus vector for gene therapy in Fanconi anemia group C. *Blood* 84:453-459.
- Walter, D., A. Lier, A. Geiselhart, F.B. Thalheimer, S. Huntscha, M.C. Sobotta, . . . M.D. Milsom. 2015. Exit from dormancy provokes DNA-damage-induced attrition in haematopoietic stem cells. *Nature*
- Wang, J., T. Otsuki, H. Youssoufian, J.L. Foe, S. Kim, M. Devetten, . . . J.M. Liu. 1998. Overexpression of the fanconi anemia group C gene (FAC) protects hematopoietic progenitors from death induced by Fas-mediated apoptosis. *Cancer Res* 58:3538-3541.
- Wang, T., D. Xia, N. Li, C. Wang, T. Chen, T. Wan, . . . X. Cao. 2005. Bone marrow stromal cell-derived growth inhibitor inhibits growth and migration of breast cancer cells via induction of cell cycle arrest and apoptosis. *J Biol Chem* 280:4374-4382.
- Watts, K.L., J. Adair, and H.P. Kiem. 2011. Hematopoietic stem cell expansion and gene therapy. *Cytotherapy* 13:1164-1171.
- Weber, K., U. Bartsch, C. Stocking, and B. Fehse. 2008. A multicolor panel of novel lentiviral "gene ontology" (LeGO) vectors for functional gene analysis. *Mol Ther* 16:698-706.
- Weinberg, J.B., T.J. Matthews, B.R. Cullen, and M.H. Malim. 1991. Productive human immunodeficiency virus type 1 (HIV-1) infection of nonproliferating human monocytes. *J Exp Med* 174:1477-1482.
- Weissman, I.L. 2000a. Stem cells: units of development, units of regeneration, and units in evolution. *Cell* 100:157-168.
- Weissman, I.L. 2000b. Translating stem and progenitor cell biology to the clinic: barriers and opportunities. *Science* 287:1442-1446.
- Welch, J.S., T.J. Ley, D.C. Link, C.A. Miller, D.E. Larson, D.C. Koboldt, . . . R.K. Wilson. 2012. The origin and evolution of mutations in acute myeloid leukemia. *Cell* 150:264-278.
- White, D.J., R.D. Unwin, E. Bindels, A. Pierce, H.Y. Teng, J. Muter, . . . S. Meyer. 2013. Phosphorylation of the leukemic oncoprotein EVI1 on serine 196 modulates DNA binding, transcriptional repression and transforming ability. *PLoS One* 8:e66510.

- Whitney, M.A., G. Royle, M.J. Low, M.A. Kelly, M.K. Axthelm, C. Reifsteck, . . . M. Grompe. 1996. Germ cell defects and hematopoietic hypersensitivity to gamma-interferon in mice with a targeted disruption of the Fanconi anemia C gene. *Blood* 88:49-58.
- Wilanowski, T., J. Caddy, S.B. Ting, N.R. Hislop, L. Cerruti, A. Auden, . . . S.M. Jane. 2008. Perturbed desmosomal cadherin expression in grainy head-like 1-null mice. *Embo J* 27:886-897.
- Wilson, A., E. Laurenti, G. Oser, R.C. van der Wath, W. Blanco-Bose, M. Jaworski, . . . A. Trumpp. 2008. Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell* 135:1118-1129.
- Wolf, S., C. Rudolph, M. Morgan, G. Busche, G. Salguero, R. Stripecke, . . . U. Modlich. 2013. Selection for Evi1 activation in myelomonocytic leukemia induced by hyperactive signaling through wild-type NRas. *Oncogene* 32:3028-3038.
- Wong, J.C., N. Alon, C. McKerlie, J.R. Huang, M.S. Meyn, and M. Buchwald. 2003. Targeted disruption of exons 1 to 6 of the Fanconi Anemia group A gene leads to growth retardation, strain-specific microphthalmia, meiotic defects and primordial germ cell hypoplasia. *Hum Mol Genet* 12:2063-2076.
- Wright, D.E., S.H. Cheshier, A.J. Wagers, T.D. Randall, J.L. Christensen, and I.L. Weissman. 2001. Cyclophosphamide/granulocyte colony-stimulating factor causes selective mobilization of bone marrow hematopoietic stem cells into the blood after M phase of the cell cycle. *Blood* 97:2278-2285.
- Wu, A.M., J.E. Till, L. Siminovitch, and E.A. McCulloch. 1968. Cytological evidence for a relationship between normal hemotopoietic colony-forming cells and cells of the lymphoid system. *J Exp Med* 127:455-464.
- Wu, X., Y. Li, B. Crise, and S.M. Burgess. 2003. Transcription start regions in the human genome are favored targets for MLV integration. *Science* 300:1749-1751.
- Yamamoto, R., Y. Morita, J. Oechara, S. Hamanaka, M. Onodera, K.L. Rudolph, . . . H. Nakauchi. 2013. Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell* 154:1112-1126.
- Yan, X., S. Lee, B.G. Gugiu, L. Koroniak, M.E. Jung, J. Berliner, . . . R. Li. 2014. Fatty acid epoxyisoprostane E2 stimulates an oxidative stress response in endothelial cells. *Biochem Biophys Res Commun* 444:69-74.
- Yang, Y., Y. Kuang, R. Montes De Oca, T. Hays, L. Moreau, N. Lu, . . . A.D. D'Andrea. 2001. Targeted disruption of the murine Fanconi anemia gene, *Fancg/Xrcc9*. *Blood* 98:3435-3440.
- Yao, H., P. Li, B.J. Venters, S. Zheng, P.R. Thompson, B.F. Pugh, and Y. Wang. 2008. Histone Arg modifications and p53 regulate the expression of OKL38, a mediator of apoptosis. *J Biol Chem* 283:20060-20068.
- Yokoi, S., K. Yasui, T. Iizasa, I. Imoto, T. Fujisawa, and J. Inazawa. 2003. TERC identified as a probable target within the 3q26 amplicon that is detected frequently in non-small cell lung cancers. *Clin Cancer Res* 9:4705-4713.
- Yoshihara, H., F. Arai, K. Hosokawa, T. Hagiwara, K. Takubo, Y. Nakamura, . . . T. Suda. 2007. Thrombopoietin/MPL signaling regulates hematopoietic stem cell quiescence and interaction with the osteoblastic niche. *Cell Stem Cell* 1:685-697.
- Zhang, P., S. Nelson, G.J. Bagby, R. Siggins, 2nd, J.E. Shellito, and D.A. Welsh. 2008. The lineage-c-Kit+Sca-1+ cell response to Escherichia coli bacteremia in Balb/c mice. *Stem Cells* 26:1778-1786.
- Zhang, Q., N.A. Shalaby, and M. Buszczak. 2014. Changes in rRNA transcription influence proliferation and cell fate within a stem cell lineage. *Science* 343:298-301.
- Zhu, J., and S.G. Emerson. 2002. Hematopoietic cytokines, transcription factors and lineage commitment. *Oncogene* 21:3295-3313.
- Zomerdijk, J.C., H. Beckmann, L. Comai, and R. Tjian. 1994. Assembly of transcriptionally active RNA polymerase I initiation factor SL1 from recombinant subunits. *Science* 266:2015-2018.

Zufferey, R., D. Nagy, R.J. Mandel, L. Naldini, and D. Trono. 1997. Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nat Biotechnol* 15:871-875.

Zychlinski, D., A. Schambach, U. Modlich, T. Maetzig, J. Meyer, E. Grassman, . . . C. Baum. 2008. Physiological promoters reduce the genotoxic risk of integrating gene vectors. *Mol Ther* 16:718-725.

7 Appendix

7.1 Abbreviations

A

7-AAD	7-amino actinomycin
ALD	adrenoleukodystrophy
AML	acute myeloid leukemia
APC	allophycocyanin
APC-Cy7	allophycocyanin-Cyanine 7

B

BLAST	basic local alignment search tool
BLAT	BLAST-like alignment tool
bp	base pair

C

°C	degree Celsius
CD	cluster of differentiation
cDNA	copy deoxyribonucleic acid
CGD	chronic granulomatous disease
CIS	common insertion sites
cm	centimeter

D

DKFZ	German Cancer Research Center
DNA	deoxyribonucleic acid

E

EFS	elongation factor 1
eGFP	enhanced green fluorescent protein
Epo	erythropoietin
ES cells	embryonic stem cells
Evi1	ecotropic virus integration site 1

F

FA	Fanconi anemia
FACS	fluorescence-activated cell sorting
FCS	fetal calf serum
FITC	fluorescein isothiocyanate

G

G-CSG	granulocyte colony stimulating factor
Grhl1	grainyhead-like 1
Gy	gray

H

h	human
HCl	hydrochloric acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HLA	human-leukocyte-antigen
HSC	hematopoietic stem cell

I

IFN	interferon
IFN- α	interferon alpha
IFN- γ	interferon gamma
IMDM	Iscove's Modified Dulbecco's Medium
IRES	internal ribosomal entry site
IVC	individually ventilated cages

K

kb	kilo bases
kDA	kilo dalton
KO	knockout

L

LAM-PCR	linear-amplification-mediated polymerase chain reaction
LB	Lysogeny broth
LDMNC	low density mononuclear cell
LeGO	lentiviral "gene ontology"
Lin ⁻	lineage-negative
LM-PCR	ligation-mediated polymerase chain reaction
LSK	lineage ⁻ , Sca1 ⁺ , c-kit ⁺
LT-HSC	long-term hematopoietic stem cell
LTR	long terminal repeats

M

μ g	microgram
μ l	microliter
μ m	micrometer
m	murine
MCS	multiple cloning site
MDS	myelodysplastic syndrome
Mecom	Evi1/MDS complex locus
mg	milligram
MGC	mammalian gene collection
min	minute

MIP-1 α	macrophage inflammatory protein-1 alpha
ml	milliliter
MOI	multiplicity of infection
MPP	multipotent progenitor
mRNA	messenger ribonucleic acid

N

ng	nanogram
nrLAM-PCR	nonrestrictive (nr) linear-amplification-mediated (LAM) polymerase chain reaction (PCR)

O

Osgin1	oxidative stress-induced growth inhibitor 1
--------	---

P

%	percent
PB	pacific blue
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PE	phycoerythrin
PE-Cy5	phycoerythrin-Cyanine 5
Pen/Strep	penicillin-streptomycin

R

rad	radiation absorbed dose
RNA	ribonucleic acid
RNase	ribonuclease
RPKM	reads per kilobase of transcript per million reads mapped
rRNA	ribosomal ribonucleic acid
RT-PCR	real-time polymerase chain reaction

S

Sca-1	stem cell antigen 1
SCF	stem cell factor
sec	second
SFFV	spleen focus-forming virus
shRNA	short hairpin ribonucleic acid
SLAM	signaling lymphocytic activation molecule
SPF	specific pathogen-free conditions
ST-HSC	short-term hematopoietic stem cell

T

Taf1b	TATA box binding protein-associated factor, RNA polymerase I, B
TNF α	tumor necrosis factor alpha
Tpo	thrombopoietin

7.2 List of figures

<i>Figure 1: Schematic representation of the FA signaling pathway</i>	6
<i>Figure 2: Mechanism of retroviral insertional mutagenesis</i>	13
<i>Figure 3: Experimental design of the retroviral insertional mutagenesis screen</i>	19
<i>Figure 4: Engraftment rate of transduced FA and WT HSCs</i>	24
<i>Figure 5: Clonality of FA and WT bone marrow samples</i>	28
<i>Figure 6: Abnormal hematopoiesis in mouse #A1 with a profound myeloid differentiation bias</i>	28
<i>Figure 7: Insertion site characterization of FA and WT HSCs</i>	30
<i>Figure 8: Genomic location of the retroviral vector integration in each individual stem cell clone</i>	31
<i>Figure 9: Taf1b expression across different tissues and cell lines</i>	34
<i>Figure 10: Taf1b expression levels and methylation pattern in HSC and progenitor populations</i>	35
<i>Figure 11: Taf1b expression during embryonic development</i>	37
<i>Figure 12: Grhl1 expression across different tissues and cell lines</i>	39
<i>Figure 13: Grhl1 expression levels and methylation pattern in HSC and progenitor populations</i>	40
<i>Figure 14: Grhl1 expression during embryonic development</i>	41
<i>Figure 15: Aberrant expression of Evi1 transcript variant I in the dominant stem cell clone #G1</i>	42
<i>Figure 16: Evi1 expression across different tissues and cell lines</i>	44
<i>Figure 17: Evi1 expression levels and methylation pattern in HSC and progenitor populations</i>	45
<i>Figure 18: Evi1 expression during embryonic development</i>	47
<i>Figure 19: Most significantly changed gene ontology processes in stem cell clone #G1</i>	49
<i>Figure 20: Aberrant expression of Osgin1 in the dominant stem cell clone #A2</i>	51
<i>Figure 21: Osgin1 expression across different tissues and cell lines</i>	53
<i>Figure 22: Osgin1 expression levels and methylation pattern in HSC and progenitor populations</i>	54
<i>Figure 23: Osgin1 expression during embryonic development</i>	56
<i>Figure 24: Most significantly changed gene ontology processes in stem cell clone #A2</i>	59
<i>Figure 25: Generation of lentiviral vectors for target gene validation</i>	61
<i>Figure 26: Subcloning of Osgin1 and Taf1b into the lentiviral vector backbone LeGO(EFS-MCS)iV2</i>	62
<i>Figure 27: Production of lentiviral overexpression vectors</i>	64
<i>Figure 28: Evaluation of transfection mixtures for the generation of lentiviral vectors</i>	65
<i>Figure 29: Analysis of lentiviral vector titers</i>	65
<i>Figure 30: Transduction of primary bone marrow cells</i>	67
<i>Figure 31: Overexpression of Osgin1 in a competitive transplantation setting</i>	68
<i>Figure 32: Analysis of the peripheral blood chimerism of FA and WT mice transplanted with a mixture of Osgin1-overexpressing and control-transduced cells at 12 weeks post transplant</i>	69

<i>Figure 33: Analysis of the peripheral blood chimerism of FA and WT mice transplanted with a mixture of Osgin1-overexpressing and control-transduced cells at 24 weeks post transplant</i>	70
<i>Figure 34: FA HSCs demonstrate a severe growth defect in vitro</i>	72
<i>Figure 35: Cell fate of FA and WT LT-HSCs in vitro</i>	73
<i>Figure 36: Schematic model of how the potential candidate target genes might rescue the FA HSC during transplantation and reconstitution</i>	86
<i>Figure 37: Schematic of the flow cytometric cell sorting to prospectively purify the respective stem cell populations from murine bone marrow</i>	95
<i>Figure 38: Schematic outline of LM-PCR to amplify SF91 vector flanking sequences in genomic DNA isolated from murine peripheral blood and bone marrow cells</i>	103
<i>Figure 39: Schematic outline of nrLAM-PCR to amplify SF91 vector flanking sequences in genomic DNA isolated from murine bone marrow cells</i>	105

7.3 List of tables

<i>Table 1: Total number of insertion sites recovered from primary and secondary recipient mice</i>	21
<i>Table 2: Potential candidate target genes</i>	32
<i>Table 3: Selection of differentially expressed genes in stem cell clone #G1</i>	50
<i>Table 4 Selection of differentially expressed genes in stem cell clone #A2</i>	58
<i>Table 5: Antibodies for lineage depletion</i>	93
<i>Table 6: Antibodies for flow cytometric cell sorting of stem cell populations</i>	94
<i>Table 7: Antibodies for flow cytometric analysis of mature cells</i>	94
<i>Table 8: Oligonucleotides for genotyping of FA knockout mouse lines (5' to 3')</i>	102
<i>Table 9: Retroviral LTR oligo sequences and oligo cassette (5' to 3')</i>	104
<i>Table 10: Retroviral LTR oligo sequences and linker cassette (5' to 3')</i>	105
<i>Table 11: Oligonucleotides for amplification of the EFS promoter</i>	106
<i>Table 12: Oligonucleotides for manufacture of the multiple cloning site (MCS)</i>	107
<i>Table 13: qRT-PCR primer</i>	110

Contributions

This thesis would not have been completed without the contributions and help of many collaborators, colleagues and mentors.

Dr. Michael Milsom designed the insertional mutagenesis screen and set up and analyzed the first transplants before I took over the project in 2011. In addition to providing guidance and ideas that shaped the project, he also provided critical feedback on draft versions of this manuscript.

Sina Huntscha supported me in the analysis of the transplanted mice and conducted many of the LM-PCRs that provided the basis for the analysis of the insertion sites in transduced HSCs. Moreover, she assisted in all matters concerning animal experiments, including the genotyping and transplantation of the mice.

My Thesis Advisory Committee suggested the analysis of the engraftment rate of transduced FA and WT HSCs, which culminated in the selection of individual mice with clear signs of clonal dominance. My many fruitful discussions with its members, **Dr. Michael Milsom, Prof. Andreas Trumpp, Prof. Hanno Glimm, Prof. Anthony Ho** and **Dr. Stefan Fröhling**, and their critical evaluation of my data at committee meetings were invaluable.

Dr. Cynthia Bartholomä, with technical support from **Stefanie Laier** (both working on behalf of the Department of Molecular and Gene Therapy at the DKFZ Heidelberg), generated the nrLAM-PCR data on bone marrow isolated from mice. **Dr. Anna Paruzynski** provided scientific input on the evaluation of the data and suggested follow-up experiments.

Nina Cabezas-Wallscheid generated the comprehensive gene expression dataset that provided insight into the expression levels of the selected target genes under homeostatic conditions. It is publicly available (Cabezas-Wallscheid et al., 2014).

Dr. Daniel Lipka from the Division of Epigenomics and Cancer Risk Factors at the DKFZ contributed the data on the epigenetic regulation of the selected target genes (Cabezas-Wallscheid et al., 2014; Lipka et al., 2014), as well as his expertise on all matters concerning the interpretation of the DNA methylation data.

Amelie Lier performed the evaluation of the target gene expression under conditions of physiologic stress. Her high-throughput RNA sequencing data on purified LT-HSCs proved to be a great resource.

Paul Kaschutnig contributed the gene expression data on the HoxB4-YFP-positive ES cells, which provided a first hint of whether the selected target genes could play a role during embryonic development of the HSC and progenitor compartment.

Tatjana Schmidt and **Sabine Henze** performed gene expression profiling of the bone marrow samples #A2 and #G1 at the Genomics and Proteomics Core Facility of the DKFZ. **Paul Kaschutnig** and **Franziska Zickgraf** provided help in the analysis of the gene expression data.

The *in vitro* cell fate tracking experiments were performed in collaboration with the group of **Prof. Michael Rieger** at the LEOWE Center for Cell and Gene Therapy in Frankfurt. **Frederic Thalheimer** put the sorted LT-HSCs into culture and evaluated the data on their cell fate outcome. This data is now published to some extent (Walter et al., 2015).

All flow cytometric cell sorts were conducted at the Flow Cytometry Core Facility of the DKFZ by **Ann Atzberger, Klaus Hexel, Jens Hartwig** and **Dr. Steffen Schmidt**.

Dr. Dagmar Walter's knowledge of flow cytometry was invaluable. She designed and set up many flow cytometric assays and assisted in the evaluation of the data.

Julia Detzer contributed to the project during her time as a bachelor student by providing technical support.

Some chapters of the introduction are partially based on the review “Disrupted Signaling through the Fanconi Anemia Pathway Leads to Dysfunctional Hematopoietic Stem Cell Biology: Underlying Mechanisms and Potential Therapeutic Strategies” (Geiselhart et al., 2012). Therefore, parts of the text in these sections might contain suggestions and corrections from co-authors. **Figure 1** was kindly provided by Amelie Lier and Michael Milsom.

Declarations

Declarations according to § 8 (3) b) and c) of the doctoral degree regulations:

- b) I hereby declare that I have written the submitted dissertation myself and in this process have used no other sources or materials than those expressly indicated,
- c) I hereby declare that I have neither applied for permission to be examined at any other institution, nor have I presented the dissertation in this or any other form at any other institution, nor have I used this dissertation in any other form in another examination.

Acknowledgements

My time as a PhD student at the German Cancer Research Center (DKFZ) will always be one of my fondest memories, primarily because of the experiences and people involved.

First and foremost, I thank **Dr. Michael Milsom** for his excellent supervision and support during this project, and for generously sharing his experience, knowledge, encouragement and advice, without which, this thesis would not have been possible.

I am grateful to **PD Dr. Karin Müller-Decker** for her interest in reading and evaluating this thesis, and for agreeing to serve as first examiner. I am also grateful to **Prof. Jan Lohmann** for taking on the role of chairman of the defense of thesis.

I would also like to thank **Prof. Andreas Trumpp** for facilitating a well-equipped and thriving research environment at HI-STEM and the Division of Stem Cells and Cancer (DKFZ).

I owe a great deal of thanks to all the members of **HI-STEM** for their scientific discussions, support, technical assistance and advice, as well as for always providing a pleasant and exciting working atmosphere. In particular, I want to thank the legendary Milsom group, **Sina Huntscha**, **Amelie Lier** and **Dr. Dagmar Walter**, for spending the majority of my PhD time with me, as well as **Paul Kaschutnig** and **Dr. Ruzhica Bogeska**, who both stepped in later to carry on the traditions. Moreover, I very much enjoyed working together with post-graduate students **Irem Bayindir** and **David Brocks**, as well as with ‘my’ undergraduate student **Julia Detzer**. Thank you all for the great times we spent together.

Thanks also to the **Helmholtz International Graduate School for Cancer Research**, with **Dr. Lindsay Murrells**, **Magdalena Uhler**, **Evelyn Müller** and **Heike Riehm-Geier**, for creating the perfect environment in which to conduct a successful PhD, for providing further education possibilities, and for funding during my doctoral period. Similarly, I would like to thank the PhD Careers team, **Dr. Barbara Janssens** and **Marion Gürth**, for their excellent support and career advice.

I wish to thank the members of the **2012/13 PhD student council**, as well as the **3rd HFYLS** and **8th IPSCC crew**, some of whom have become very good friends, for everything that we achieved and for the truly memorable time that we spent together.

Even with the support of all these people, this thesis would not have been completed without my friends and family, who provided support, encouragement and telephone emergency services around the clock.

In particular, I want to thank my parents **Fine** and **Georg Geiselhart** and my sister **Jana Reck**, as well as my adopted family in Stuttgart. However, most of all, I wish to thank **Frank Burkart** who certainly contributed the most.