

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

M.Sc. Shayda Hemmati

Born in: Shiraz, Iran

Oral examination: 11.08.2014

Identification and characterization of novel regulatory genes of post-embryonic hematopoiesis

Referees: Prof. Dr. Andreas Trumpp
Prof. Dr. Hanno Glimm

Summary

Comprehensive integration site analysis for monitoring the clonal dynamics in clinical gene therapy has revealed that the insertion of the therapeutic retroviral vector (RV) can deregulate and even substantially activate neighboring genes leading to selection advantage and clonal outgrowth. Strikingly, 7 out of 10 Wiskott Aldrich Syndrome (WAS) gene therapy patients developed acute leukemia driven by gene corrected cell clones aberrantly expressing LMO2, MDS1 or MN1. This indicates that RV-mediated activation of adjacent regions cannot only influence the fate of hematopoietic stem cells but also cause clonal dominance up to leukemia. To identify novel regulators of benign hematopoiesis we established a systematic selection strategy using the total genomic integration site dataset of normal and highly polyclonal clinical blood and bone marrow samples.

In this thesis, the unique integration site (IS) dataset within a cohort of 10 WAS gene therapy patients was systematically analyzed to select for candidate genes involved in the regulation of hematopoiesis. Initially, a total of 12.887 unique IS in vicinity of 3.267 genes were identified. Next, we selected all genes with at least 10 different IS within a 200 kb window around the gene (n=588). To enrich for genes with increased probability of transcriptional activation we then chose those genes with at least 10 IS within a 50kb window around the transcriptional start site (n=424). After stringent exclusion of all genes located within gene clusters 32 candidate genes were identified. To evaluate the hematopoietic activity of gene corrected cell clones, their contribution to blood cell formation within four years post gene therapy was monitored. We observed that these clones were detectable 15 to 93 times in a total of 102 individually analyzed patient samples, demonstrating long term activity of these hematopoietic stem cell clones. Interestingly, 20 out of the 32 highest ranked genes such as EVI1, CCND2 and LMO2 are known hematopoietic key regulators, strongly validating our selection strategy. After identification of 12 novel hematopoietic regulatory candidate genes, the top five ranked genes, ZNF217, LRRC33, PLCB4, EVL and IRF2BPL were chosen for further functional analysis in murine hematopoietic primary cells.

To evaluate the endogenous expression of candidate genes in murine hematopoietic stem and progenitor cells global transcriptome datasets from purified populations were evaluated. We observed that all five selected candidate genes were expressed in at least one out of the five analyzed hematopoietic stem and progenitor cell populations which may point to an important role in the respective cell fraction. In order to validate our selection

strategy and to further investigate whether the chosen candidate genes play roles in hematopoiesis, we performed various *in vitro* and *in vivo* assays. We tested the effect of the selected candidate genes on proliferation, differentiation, and cytokine independency as well as for their influence on long term multilineage reconstitution and self-renewal after murine bone marrow transplantation.

The first candidate gene, ZNF217, is a zinc finger protein known as a transcription factor. To analyze the impact of ZNF217 on transcriptional activity, global gene expression profiling in hematopoietic cells was performed. We observed that 337 out of 422 genes were significantly downregulated and that they are mainly involved in cellular movement indicating that ZNF217 plays a role in hematopoietic cell migration. Since ZNF217 is known as a proto-oncogene in breast and ovarian carcinoma we evaluated its effect on growth factor independency of hematopoietic cells. Interleukin 3 (IL3)-dependent cells overexpressing ZNF217 acquired the capacity to survive and form colonies in the absence of IL3 suggesting a transforming role for ZNF217 in hematopoietic cells.

The second candidate LRRC33 resembles the protein structure of Toll-like receptor (TLR) proteins which are involved in innate immunity. The overexpression of LRRC33 and TLR4 decreased the activity of NF- κ B *in vitro* when stimulated with bacterial lipopolysaccharide. This may point to an inhibitory role of LRRC33 in NF- κ B signaling, an important pathway for maintaining stem cell integrity. Transplanted LRRC33-overexpressing LSK (Lin⁻Sca-1⁺cKit⁺) cells gave rise to a 1.3-6.7-fold lower ratio of T-cells and in contrast a 1.1-7.2-fold higher amount of donor derived macrophages in secondary recipients compared to control mice.

To study the function of the third candidate gene on hematopoiesis constitutive PLCB4 knockout mice were obtained. This phospholipase has been shown to be important for brain development but has not been linked to hematopoiesis so far. Preliminary results indicate that PLCB4 deficient mice have a reduced LSK cell fraction compared to age matched littermates within the first 18 days after birth.

These results demonstrate that clinical integration site datasets can be used to identify regulatory genes of hematopoiesis. Here, we identified ZNF217 as a driver of hematopoietic transformation applying the established selection strategy. In total, we could show that four out of five candidate genes play a role in hematopoiesis and they will be further evaluated for their stem cell regulatory potential. Systematic identification of novel regulatory genes in meta-datasets derived from a larger number of gene therapy studies and subsequent validation *in vitro* and *in vivo* will allow to gain new insights into the biology of post-embryonic hematopoiesis.

Zusammenfassung

Umfassende Integrationsstellen (IS)-analysen zur Beobachtung der klonalen Zusammensetzung in Gentherapie-Studien haben gezeigt, dass die Insertion des therapeutischen retroviralen Vektors (RV) benachbarte Gene durch transkriptionelle Aktivierung deregulieren kann. Dies kann zu einem Selektionsvorteil und klonalem Auswachsen bestimmter Zellklone führen. Auffallenderweise haben 7 von 10 Gentherapie-Patienten mit Wiskott-Aldrich Syndrom eine akute Leukämie entwickelt, die von Gen-korrigierten Zellklonen mit aktivierenden Vektor-Insertionen vor den Genen LMO2, MDS1 oder MN1 getrieben wurden. Dies weist darauf hin, dass die RV initiierte Aktivierung benachbarter Genregionen nicht nur das Schicksal hämatopoetischer Stammzellen beeinflussen, sondern auch klonale Dominanz bis hin zur Leukämie-Entstehung zur Folge haben kann. Eine systematische Analyse von klinischen IS Datensätzen und die Klassifizierung von genomischen Anhäufungen dieser IS sollte zur Identifizierung bedeutender hämatopoetischer Gene führen. Deshalb wurde in der vorliegenden Arbeit mittels klinischer Daten von RV markierter polyklonaler Blutbildung eine Selektionsstrategie entwickelt, um bisher unbekannte Stammzellregulatoren zu ermitteln.

Um Kandidaten-Gene zu identifizieren, wurde der einzigartige klinische IS Datensatz von insgesamt 12887 spezifischen IS von einer Kohorte von 10 Patienten analysiert. In der Nähe dieser individuellen Insertionen wurden 3267 Protein-kodierende Gene ermittelt. Im nachfolgenden Schritt, wurden Gene gewählt, welche mindestens 10 unterschiedliche IS in einem 200 kb Fenster um ihre Loci aufwiesen (n=588). Um die Wahrscheinlichkeit zu erhöhen, dass die umliegenden Gene durch die Vektorinsertionen transkriptionell aktiviert wurden, wurden diese mit mindestens 10 IS in einem Bereich von 50 kb um den Transkriptionsstart gewählt (n=424). Nach stringentem Ausschluss von Genen, welche auf genomischer Ebene nicht einzeln, sondern in Gruppen lokalisiert vorliegen, wurden 32 eindeutig zuordenbare Kandidaten identifiziert. Um die hämatopoetische Aktivität dieser spezifischen Gen-korrigierten Klone zu untersuchen, wurde deren Beitrag zur humanen Blutbildung während einem Zeitraum von vier Jahren nach Gentherapie überprüft. Dabei haben wir festgestellt, dass diese Zellklone in 15 bis 93 Fällen von 102 individuell analysierten Blutproben detektiert werden konnten, was auf eine kontinuierliche hämatopoetische Aktivität dieser Klone hindeutet. Interessanterweise sind 20 der 32 Kandidaten-Gene mit den namhaften Vertretern LMO2, EVI1 und CCND2 bereits bekannte hämatopoetische Regulatoren, was unsere Auswahlstrategie immens stützt und validiert. Nach Identifizierung von 12 neuen hämatopoetischen Kandidaten-Genen wurden die fünf höchst Gelisteten zur detaillierten funktionalen Untersuchung in murinen hämatopoetischen

Primärzellen ausgewählt. Um die endogenen Expressionslevel dieser Kandidaten in definierten hämatopoetischen Stamm- und Progenitorzell-Populationen zu bewerten, wurden globale Transkriptom-Daten ausgewertet. Alle fünf Kandidaten wurden in mindestens einer der fünf aufgereinigten Zellpopulationen exprimiert, was auf eine wichtige Rolle in der entsprechenden Fraktion hindeuten könnte. Um umfassender zu untersuchen, ob die ausgewählten Kandidaten-Gene Stammzell-regulatorisches Potential besitzen, wurden verschiedene *in vitro* und *in vivo* Experimente verwendet. Es wurde einerseits der Einfluss der Kandidaten auf Proliferation, Differenzierungsfähigkeit und Zytokin-Abhängigkeit, andererseits deren Auswirkung auf die Langzeitblutbildung und Selbsterneuerungsfähigkeit im lebenden Organismus untersucht. Das erste Kandidatengen ist ein Zinkfingerprotein, ZNF217, und als Transkriptionsfaktor bekannt. Um das Potential als transkriptioneller Regulator in hämatopoetischen Zellen zu untersuchen, wurden nach dessen Überexpression globale Genexpressionsanalysen durchgeführt. 337 von 422 deregulierten Genen waren signifikant herunter-reguliert und der Großteil kann in den Prozess der zellulären Bewegung eingruppiert werden, was auf eine Rolle von ZNF217 bei der Migration von hämatopoetischen Zellen hindeuten könnte. Da ZNF217 in Brust- und Ovarialkarzinomen als Proto-Onkogen beschrieben wurde, wurde sein Einfluss auf die Wachstumsfaktorunabhängigkeit in hämatopoetischen Zellen untersucht. Die Überexpression von ZNF217 führte unter Mangel von Interleukin 3 (IL3) zum Überleben und Wachstum von Zytokin-unabhängigen Kolonien, was auf eine transformierende Eigenschaft von ZNF217 in hämatopoetischen Zellen schließen lässt. Der zweite Kandidat, LRRC33 ähnelt in seiner Proteinstruktur Toll-like-Rezeptoren (TLR), welche eine Rolle in der angeborenen Immunität spielen. Es konnte gezeigt werden, dass nach Stimulierung mit bakteriellem Lipopolysaccharid *in vitro* die Überexpression von LRRC33 und TLR4 zu einer reduzierten Aktivität von NFκB führt. Dies deutet auf einen inhibitorischen Einfluss von LRRC33 auf das NFκB *Signaling* hin, was ein wichtiger Signalweg zur Erhaltung hämatopoetischer Stammzellintegrität darstellt. Die Transplantation von LRRC33 exprimierenden LSK (lin-Sca1+ckit+) Zellen ergab im Vergleich zu Kontrollmäusen eine 1,3-6,7 fach geringere Menge an T-Zellen und einen 1,1-7,2 fach höheren Anteil an Makrophagen in sekundär transplantierten Rezipienten. Um die Funktion des dritten Kandidatengens, PLCB4 zu untersuchen, wurden Knockout-Mäuse eingesetzt. Es wurde bereits gezeigt, dass diese Phospholipase eine große Bedeutung während der Gehirnentwicklung besitzt, jedoch ist bisher noch kein Zusammenhang zur Blutentwicklung beschrieben worden. Erste Ergebnisse deuten darauf hin, dass PLCB4 defiziente Mäuse einen im Vergleich zu altersgemäßen Wildtyp-Wurfgeschwistern geringeren Anteil an hämatopoetischen LSK Zellen während den ersten 18 Tagen nach Geburt aufweisen.

Die Ergebnisse der hier vorliegenden Arbeit zeigen, dass klinische Integrationsstellen-Datensätze verwendet werden können, um bisher unbekannte Stammzell-regulatorische Gene der Blutbildung zu identifizieren. Durch die Anwendung der hiermit entwickelten Selektionsstrategie konnte ZNF217 als Faktor hämatopoetischer Transformierung ermittelt werden. Zusammenfassend kann gesagt werden, dass vier von fünf ausgewählten Kandidaten eine Rolle während der Blutzellentwicklung spielen und weiterführende Untersuchungen das Stammzell-regulatorische Potential auf molekularer Ebene erfassen werden. Die systematische Identifizierung bisher unbekannter Stammzell-regulatorischer Gene mittels übergreifender Datensätze einer größeren Anzahl von Gentherapiestudien sowie deren schrittweise Validierung in *in vitro* und *in vivo* Ansätzen werden dazu beitragen, neue Einblicke in die Biologie adulter Hämatopoese zu gewinnen.

Table of Contents

Summary	I
Zusammenfassung.....	III
List of Abbreviations.....	X
List of Figures.....	XII
List of Tables.....	XIV
1 Introduction.....	1
1.1 Hematopoiesis and hematopoietic stem cells	1
1.1.1 Hematopoietic stem cell heterogeneity	2
1.1.2 Purification of hematopoietic stem and progenitor cells	3
1.1.3 Regulation of hematopoietic stem cell fate.....	5
1.1.4 Process of leukemogenesis	7
1.2 Wiskott-Aldrich Syndrome	8
1.2.1 Treatments of Wiskott–Aldrich Syndrome	9
1.2.1.1 Hematopoietic stem cell transplantation.....	9
1.2.1.2 Gene therapy.....	10
1.3 Retroviral vector system.....	11
1.3.1 Insertional mutagenesis: the side effect of retroviral gene transfer.....	12
1.3.2 Monitoring clonal dynamics of marked hematopoietic stem cells by LAM-PCR14	
2 Specific aims	15
3 Material and methods.....	18
3.1 Materials.....	18
3.1.1 Equipment and devices	18
3.1.2 Plastic and disposables.....	20
3.1.3 Bacteria culture	21
3.1.4 Cell culture.....	22
3.1.5 Freezing and thawing medium.....	23
3.1.6 Cytokines, medium supplements and solutions for cell culture.....	24

3.1.7	Antibodies	25
3.1.8	PCR primers and reagents	26
3.1.9	Plasmids.....	28
3.1.10	Chemicals and reagents	28
3.1.11	Enzymes and reaction buffers for cloning.....	29
3.1.12	Commercial kits.....	29
3.1.13	Western blot buffers and reagents	30
3.1.14	Bacteria strains.....	33
3.1.15	Cell lines.....	33
3.1.16	Mouse strains	34
3.1.17	Surgical instruments.....	34
3.2	Methods.....	35
3.2.1	Molecular biology methods	35
3.2.2	Cell biology methods.....	42
3.2.3	Mouse experiments.....	49
4	Results	50
4.1	Establishment of a selection strategy for identification of novel candidate stem cell regulatory genes.....	50
4.1.1	Assessing the contribution of candidate-gene-associated clones in hematopoiesis	55
4.1.2	Analyzing long-term multilineage contribution of hematopoietic stem and progenitor cells <i>in vivo</i>	57
4.1.3	Evaluating the expression of candidate genes in hematopoietic stem and progenitor cells.....	60
4.2	Investigating the function of candidate hematopoietic regulatory genes.....	61
4.2.1	Examining proliferative activity and survival of single LSK cells <i>in vitro</i>	62
4.2.1.1	Evaluating growth kinetics of LSK cells cultured <i>in vitro</i>	62
4.2.1.2	Analyzing survival and growth kinetics of CCND2-transduced LSK cells.....	64

4.2.2	Producing functional lentiviral overexpression vectors encoding for the candidate genes	65
4.2.3	Studying the impact of candidate genes on hematopoiesis <i>in vitro</i>	66
4.2.3.1.	Studying the viability of hematopoietic cells overexpressing candidate genes	66
4.2.3.2.	Analyzing cytokine dependency in hematopoietic cells overexpressing candidate genes	67
4.2.3.3.	Assessing gene expression profiling of ZNF217-transduced cells	71
4.2.3.4.	Evaluating the role of EVL in colony forming activity.....	73
4.2.3.5.	Analyzing the impact of LRRC33 on NF- κ B activity.....	74
4.2.4.	Investigating the function of candidate regulatory genes <i>in vivo</i>	74
4.2.4.1.	Analyzing hematopoiesis in mice transplanted with LRRC33-transduced LSK cells.....	76
4.2.4.2.	Characterizing BM cells from PLCB4 knockout mice <i>in vitro</i> and <i>in vivo</i>	78
4.2.4.2.1.	Examining BM cells from PLCB4 knockout mice in pre-B cell CFC assay.....	79
4.2.4.2.2.	Investigating multilineage engraftment of BM cells from PLCB4 knockout mice.....	82
5.	Discussion.....	84
5.1.	The integration site repertoire from gene therapy patients can be used for selecting candidate hematopoietic regulatory genes	84
5.2.	Evaluation of the candidate genes' expression in defined hematopoietic stem and progenitor cell populations	86
5.2.1.	Efficient purification of hematopoietic stem population using CD34 ⁻ CD48 ⁻ CD135 ⁻ CD150 ⁺ LSK surface markers.....	86
5.3.	Investigating the role of candidate genes <i>in vitro</i> and <i>in vivo</i>	89
5.3.1.	Single LSK (Lin ⁻ Sca-1 ⁺ c-Kit ⁺) cells show different growth kinetics under stem cell condition <i>in vitro</i>	89

5.3.2. ZNF217 overexpression increased survival and proliferation of hematopoietic cells <i>in vitro</i>	90
5.3.2.1. Hematopoietic cells overexpressing ZNF217 show increased growth activity	90
5.3.2.2. Stable ZNF217 expression induces IL3 independency of FDCP-Mix and 32D cells	91
5.3.2.3. Expression profiling of ZNF217 overexpressing hematopoietic cells indicated deregulation of pathways involved in cellular movement.....	91
5.3.3. EVL overexpression increases cell viability and colony forming ability	91
5.3.4. Analyzing the effect of LRRC33 <i>in vivo</i> and <i>in vitro</i>	92
5.3.4.1. LRRC33 inhibits NF-kB activity through toll-like receptor 4	92
5.3.4.2. LRRC33 influences macrophage and T-cell differentiation <i>in vivo</i>	92
5.3.5. Analyzing PLCB4 function <i>in vitro</i> and <i>in vivo</i> using PLCB4 knockout mice	93
5.3.5.1. Eighteen-day old PLCB4 knockout mice show decreased level of LSK cell count.....	93
Conclusion	96
References.....	97
Publications and Conferences	108
Declaration	109
Acknowledgments.....	110

List of Abbreviations

ALL	Acute lymphoblastic leukemia
AML	Acute myeloid syndrome
ATP	Adenosine-5'-triphosphate
BM	Bone marrow
BMT	Bone marrow transplantation
BSA	Bovine serum albumin
CCND2	Cyclin D2
CDK	Cyclin dependent kinase
cDNA	Complementary DNA
CFC	Colony forming cell
CFU	Colony forming unit-spleen
CIS	Common integration site
CLP	Common lymphoid progenitors
CMP	Common myeloid progenitors
ddNTP	Dideoxynucleotides
DMSO	Dimethylsulfoxide
DNA	Desoxyribonucleic acid
dNTP	Deoxynucleosidetriphosphates
dTTP	Deoxythymidine triphosphate
EMT	Epithelial-Mesenchymal Transition
EVL	Ena/VASP-like
FACS	Fluorescence activated cell sorting
GMP	Granulocyte-monocyte progenitors
HPC	Hematopoietic stem cells
HSC	Hematopoietic stem cells
HSCT	Hematopoietic stem cell transplantation
IL3	Interleukin 3
IRF2BPL	Interferon regulatory factor 2 binding protein like

IS	Integration site
KO mice	knockout mice
LAM-PCR	Linear-amplification mediated PCR
LPS	Lipopolysaccharides
LRRC33	Leucine rich repeat 33
LTR	Long terminal repeats
LV	lentiviral vector
MDS	Myelodysplastic syndrome
MEP	Megakaryocyte-erythroid progenitors
MOI	Multiplicity of infection
MPP	Multipotent progenitors
NK	Natural killer cells
PCR	Polymerase Chain Reaction
PEI	Polyethylenimine
PLCB4	Phospholipase C beta4
PPT	Polypurine tract
PVDF	Polyvinylidene difluoride
RNA	Ribonucleic acid
RT	Room temperature
RV	Retroviral vector
SIN	Self-inactivating
SNP	Single nucleotide polymorphism
ss	Single stranded
T-ALL	T-cell acute lymphoblastic leukemia
TLR	Toll like receptor
TSS	Transcription start site
VSV-G	Vesicular stomatitis virus-glycoprotein g
WAS	Wiskott Aldrich Syndrome
X-CGD	X-linked chronic granulomatous disease
X-SCID	X-linked severe combined immunodeficiency
ZNF217	Zinc finger protein 217

List of Figures

Figure 1: Model of the hematopoietic developmental hierarchy.....	2
Figure 2: LSK cell population contains 5 subpopulations.....	5
Figure 3: Fate decisions of hematopoietic stem cells..	6
Figure 4. Hypothetical model for the progression of acute myeloid leukemia (AML) involving 5q loss..	7
Figure 5. Retroviral gene transfer into hematopoietic stem cells in gene therapy.	10
Figure 6: Packaging and transduction process of retroviral vectors.	12
Figure 7: Insertional mutagenesis following retroviral integration.	13
Figure 8: LAM-PCR detects the precise position of viral IS in the genome.	14
Figure 9: An overview of sequential steps of selecting candidate stem cell regulatory genes..	51
Figure 10: The position of the IS and their flanking genes were observed along the chromosome. ...	52
Figure 11: Schematic view of genes in vicinity of IS excluded or included as candidate hematopoietic regulatory genes.....	52
Figure 12: Selecting candidate hematopoietic regulatory genes by analyzing the IS repertoire.	53
Figure 13: The distribution of 32 highly-ranked candidate genes in individual WAS patients in four years.	54
Figure 14: Long term clonal activity of the candidate genes..	56
Figure 15: The engraftment of transplanted HSC and HPC with various combinations of stem cell markers in recipient mice.....	58
Figure 16: Differentiation capability of HSC_1 and HPC_1 cells into recipient mice hematopoiesis. ..	59
Figure 17: RNA expression of the selected candidate genes in HSC and HPC populations.	
Figure 18: Experimental layout to investigate the role of selected candidate genes <i>in vitro</i> and <i>in vivo</i>	62
Figure 19: Proliferation kinetics of single-cell-derived LSK clones.	63
Figure 20: Survival of CCND2-transduced-LSK cells compared to mock transduced control.....	64
Figure 21: LV constructs encoding for individual candidate transgenes: pCCL-PGK-Transgene-IRES-eGFP.	65
Figure 22: Overexpression of the candidate genes in transduced cells at RNA and protein level.	66
Figure 23: Metabolic activity of transduced hematopoietic C1498 cells.....	67
Figure 24: Growth kinetics of FDCP-Mix cells transduced with the LV of candidate genes in the presence (left) or absence (right) of IL3.....	68
Figure 25: CFC assay on transduced 32D cells transduced with the LV of four candidate genes in absence of IL3 cytokine.	69
Figure 26: c-Myc and ZNF217 colonies in cytokine free CFC assay.	70
Figure 27: Global gene expression profiling of hematopoietic 32D cells overexpressing ZNF217..	71
Figure 28: Top deregulated biological networks identified in ZNF217 overexpressing cells.....	71
Figure 29: Categories and subcategories of ZNF217 deregulated genes.....	72
Figure 30: CFC assay with transduced C1498 hematopoietic cells overexpressing EVL..	73
Figure 31: NF- κ B activity in transfected 293T cells.	74
Figure 32: Investigating the function of the candidate genes <i>in vivo</i> ..	75
Figure 33 : Number of mice transplanted and engrafted with transduced LSK cells.....	76
Figure 34: Multilineage engraftment in primary LRRC33 and control GFP transplants.....	77

Figure 35: Multilineage engraftment in secondary LRRC33 and control GFP transplants.....	78
Figure 36: PLCB4 knockout mice.	79
Figure 37: Analyzing LSK population in BM from PLCB4 mice.....	80
Figure 38: Myeloid and B-cell CFC assay with the whole BM cells from -/-, +/- and +/+ PLCB4 mice at different ages.	81
Figure 39: Transplantation of BM cells obtained from -/-, +/- and +/+ 40-day old PLCB4 mice into lethally irradiated recipients..	82
Figure 40: Long-term multilineage reconstitution of transplanted PLCB4-knockout BM cells in primary recipient mice.....	83

List of Tables

Table 1: Summary of single cell BM reconstitution data	4
Table 2: The 5 candidate hematopoietic regulatory genes selected for further <i>in vitro</i> and <i>in vivo</i> analyses.....	55
Table 3: HSC and HPC marker combinations used for isolating various hematopoietic populations	57

1 Introduction

1.1 Hematopoiesis and hematopoietic stem cells

Hematopoietic stem and progenitor cells (HSC and HPC) normally reside in the BM and represent about 0.01% of nucleated marrow cells [1]. These cells make a heterogeneous population in various aspects such as their degree of self-renewal, differentiation [2] and life span [3]. Since most mature blood stem cells have a limited life span, the ability of HSC to preserve themselves by self-renewal is critical [4]. Through hematopoietic differentiation, these cells lose their multipotency and increasingly become mature. HPC are able to give rise to large numbers of both myeloid (e.g. macrophages, granulocytes, erythrocytes, and megakaryocytes or platelets) and lymphoid lineages (e.g. T-cells and B-cells) [1, 5].

Hierarchical relationships exist among HSC, their progenies, and mature blood cells. In hierarchical differentiation models, multipotent HSC with self-renewal ability [16-18]; initially give rise to multipotent progenitors (MPP). MPP give rise to lymphoid-primed multipotent progenitors (LMPP) [6]. MPP and LMPP are considered as multipotent progenitor cells with limited or no self-renewal activity, which allows only short-term (up to 8 weeks) multilineage repopulation. These multipotent progenitors generate lineage-committed progenitors such as common myeloid progenitors (CMP)[7] which generate granulocyte-macrophage progenitors (GMP)[8], common lymphoid progenitors (CLP) [9], or megakaryocyte-erythroid progenitors (MEP), with loss of self-renewal ability [10] (Figure 1). Current models accept that HSC self-renewal capability is associated with multipotency and that HSC lose differentiation potential in the process of blood maturation. According to the present “classical” model, hierarchical differentiation proceeds with the first lineage commitment at the MPP stage, giving rise to all mature blood lineages. Several recent studies have questioned this scheme of lineage commitment largely because LMPPs with both neutrophil-monocyte and lymphoid potentials, but little or no megakaryocyte-erythroid potential (MkE), have been detected [3, 11] suggesting that an alternative MkE differentiation pathway may exist [10, 12].

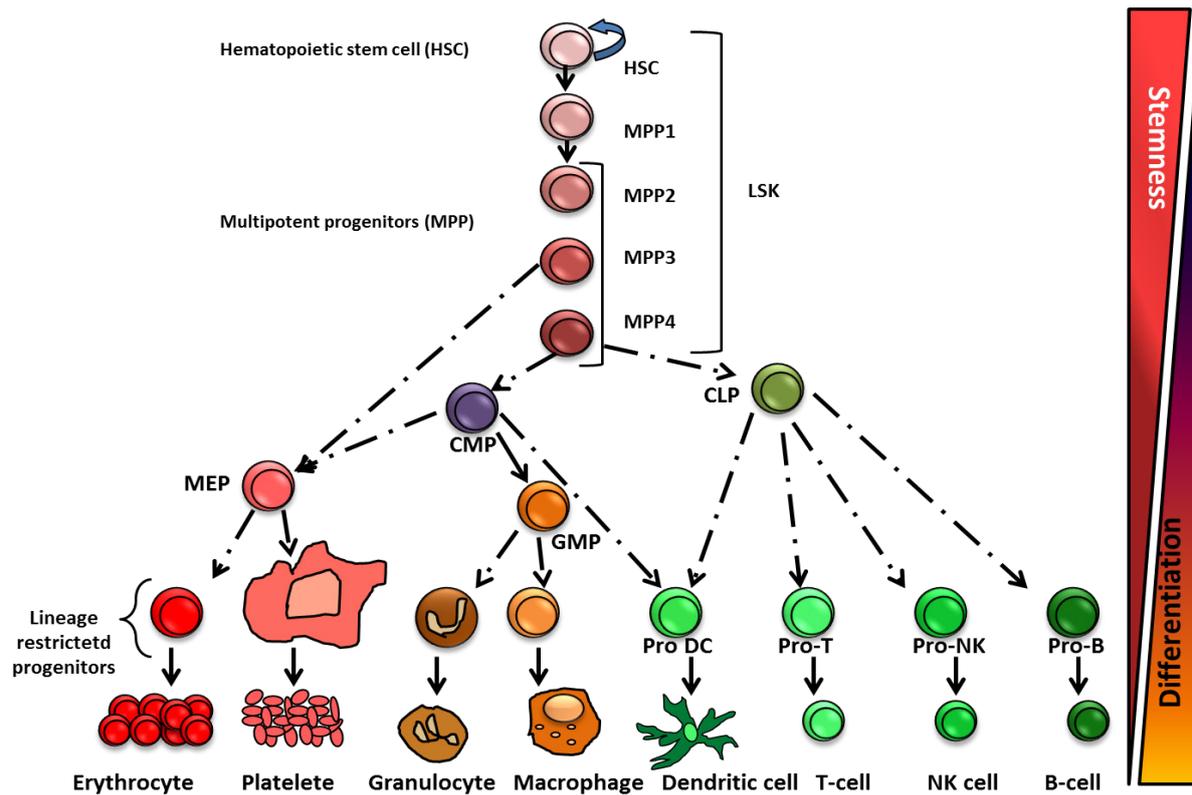


Figure 1: Model of the hematopoietic developmental hierarchy. HSC, hematopoietic stem cell; MPP, multipotent progenitor; LSK, Lineage⁻ Sca-1⁺ cKit⁺ cells; CLP, common lymphoid progenitor; CMP, common myeloid progenitor; GMP, granulocyte-monocyte progenitor; MEP, megakaryocyte-erythroid progenitors.

1.1.1 Hematopoietic stem cell heterogeneity

The fact that a single HSC can generate millions of mature blood cells gave the first evidence to use clonal analyses for HSC detection [5, 13]. Fifty years ago the first HSC were defined as colony forming units spleen (CFU-S) since they were able to form visible colonies in spleen in a 1-2 week period, however they were found to contain very few longer-term hematopoietic repopulating cells [14-16]. CFU-S colonies contained cells which were able to generate secondary spleen colonies, and could reconstitute lethally irradiated mice [5, 13, 17].

A significant observation from the CFU-S study was the highly variable types of daughter cells generated in transplant experiments [17, 18]. This finding was later supported by *in vitro* analyses of multilineage colonies, where external stimuli could be kept the same [19, 20], and later strengthened by retroviral vectors (RV) used to mark cells which can be tracked over longer period of time *in vivo* [21, 22].

Various HSC subtypes have been defined using several criteria to distinguish the mature blood cell outputs. Mueller-Sieburg et al. have defined different myeloid-biased, lymphoid-

biased and balanced subtypes based on the dominant lineage in the total number of donor-derived blood cells (excluding the recipient's contribution) [4].

Dykstra et al. studied different differentiation behavior of HSC by their relative contribution to the total number of circulating myeloid and lymphoid cells (including both donor and recipient HSC). In the latter method HSC subtypes are defined as lymphoid-deficient (α), balanced (β) and myeloid-deficient (γ and δ). Both α and β -HSC contribute equally to the circulating pool of myeloid cells [3]. Serial transplantation experiments have shown that only α and β -HSC remain active during serial transplantation and their unique differentiation behavior is maintained in primary transplants over years [2, 3]. In addition, their initial expression pattern of differentiation is expressed by their clonally generated daughter cells when they are transplanted into secondary recipient mice [2-4, 11].

Yamamoto et al. defined an intermediate-term-HSC (IT-HS) in addition to previously defined short and long-term-HSC (ST- and LT-HSC) [10, 23, 24]. They proposed using ST-, IT-, and LT-HSC to classify HSC based on their reconstitution time period [10]. Since granulocytes are extremely short-lived, their reconstitution directly reflects HSC activity. Therefore the reconstitution of granulocytes (and not other lineages such as erythrocytes, B-cells and T-cells) was considered as a parameter of this classification [3, 25, 26].

1.1.2 Purification of hematopoietic stem and progenitor cells

Numerous studies have been conducted to purify HSC and HPC populations using flow cytometry. This technique is primarily based on fluorochrome-conjugated antibodies against cell surface markers of HSC and HPC and the physiological properties of these cells such as high-efflux activity of certain fluorescent dyes. The surface marker combinations to purify HSC and HPC are primarily based on Lineage markers, Sca-1, and c-Kit resulting in the so-called "LSK" (Lineage⁻ Sca-1⁺ c-Kit⁺) population [27]. LSK cells comprise a heterogeneous population including self-renewing long term HSC and various HPC. Although compared to whole BM, LSK cells are about 1000 fold enriched for HSC activity, the majority of these cells are multipotent progenitors and only about 1 out of 30 LSK cells is a long-term multilineage repopulating cell [28].

Markers in LSK scheme can be used to further subfractionate this population to distinguish HSC from differentiating progenitors. Additional markers included in the LSK scheme, such as Tie2 [29], CD34[23], and EPCR[30], can be used to further purify this population. The "signaling lymphocyte activation molecule" (SLAM) family of proteins, including CD150,

CD48, and CD244 can also be used to enhance the LSK enrichment scheme. CD48 and CD244 are used to identify and eliminate differentiating progenitor lineage cells, while CD150 identifies cells with potent stem cell activity [31].

The other strategy is to identify HSC based on their high efflux ability of fluorescent dyes such as a DNA-binding dye, Hoechst 33342 or a mitochondrial-binding dye, Rhodamine123 (Table 1)[32-34].

Table 1: Summary of single cell BM reconstitution data

Phenotype of cell	Frequency of reconstitution (%)	Reference
LSK CD150 ⁺ CD48 ⁻	47	[31]
CD41 ⁻ CD150 ⁺ CD48 ⁻	45	[31]
LSK SP ^{lo} Thy1.1 ^{lo} CD34 ⁻ CD135 ⁻	35	[35]
Lin ⁻ SP Rho ^{lo}	33	[36]
LSK CD34 ^{lo/-}	22	[23]
CD150 ⁺ CD48 ⁻	21	[31]
LSK Thy1.1 ^{lo}	18	[37]

LSK: Lin⁻Sca1⁺c-Kit⁺; SP: side population (can efflux Hoechst 33342); Rho: Rhodamine [38].

The heterogeneity of LSK cells has been shown by further purifying them with CD34 CD48 CD150 CD135 cell surface markers. Based on the expression of CD150 and CD48 cell surface markers, LSK cells can be divided to three populations. CD150⁺ CD48⁻ cells can be further subdivided based on CD34 expression. CD150⁺ CD48⁺ subset is almost entirely CD34⁺ CD135⁻ and CD150⁺CD48⁻ can be subdivided to CD135⁺ and CD135⁻ subpopulations.

The frequency of MPP1, MPP2, MPP3, and MPP4 are 3, 5, 18, and 13 times more than HSC frequency in whole murine BM respectively [39] (Figure 2).

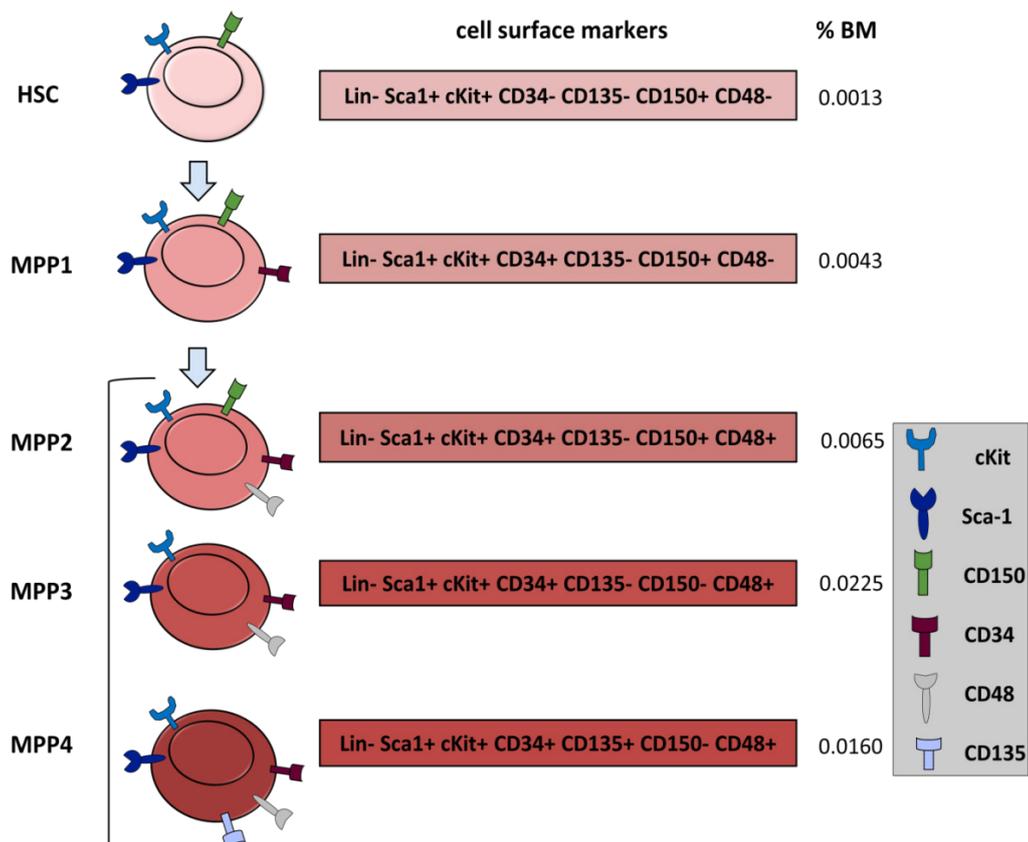


Figure 2: LSK cell population contains 5 subpopulations. HSC, MPP1, MPP2, MPP3, and MPP4 populations are depicted according to their cell surface markers. HSC: hematopoietic stem cells, MPP: multipotent progenitor cells, BM: bone marrow.

1.1.3 Regulation of hematopoietic stem cell fate

The task of HSC is to maintain a certain number of cells in each hematopoietic lineage and to adapt blood cell production to the need of the organism. This is done through the regulation of various processes such as differentiation versus self-renewal, proliferation versus quiescence and survival versus apoptosis [40]. The cellular environment (niche) in bone marrow contains various cells such as endothelium, adipocytes and osteoblasts which play a crucial role in the regulation of HSC [41] (Figure 3).

Both intrinsic and extrinsic signals such as cytokines, cell-cell contacts or contacts with the extracellular matrix in the subendosteal, influence the fate of stem cells. Two hypotheses attempt to describe action and interactions of exogenous signals and intrinsic factors [42, 43].

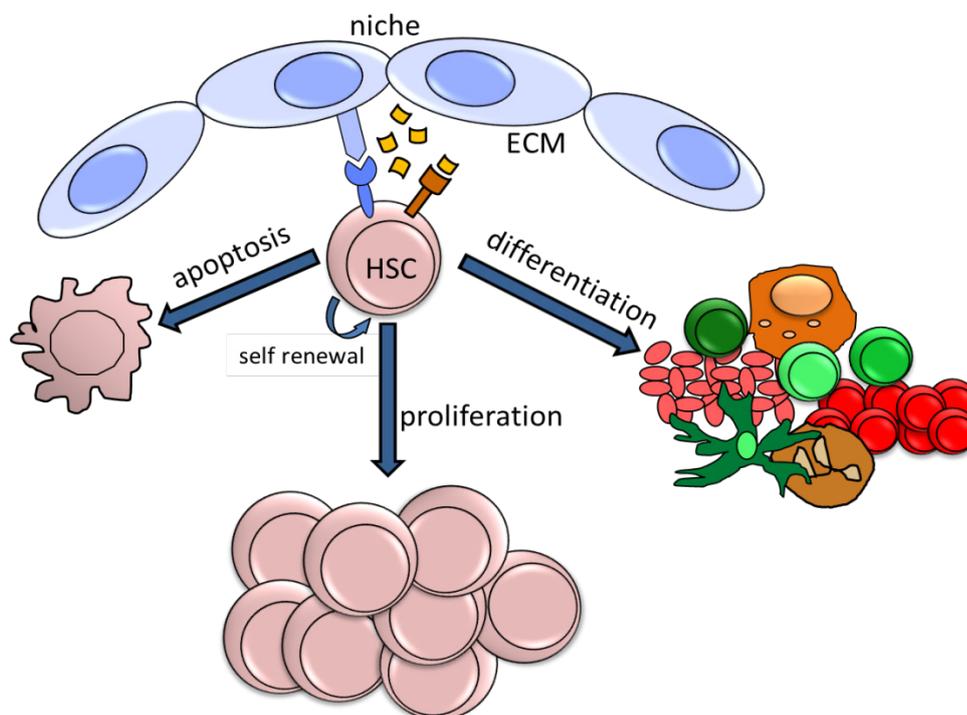


Figure 3: Fate decisions of hematopoietic stem cells. HSC have the unique ability to reconstitute hematopoiesis throughout life while keeping the number of stem cells constant (self-renewal). During the differentiation process, HSC and progenitor cells must constantly go through fate decisions. Adult hematopoiesis is largely dependent on the microenvironment in BM, which consists of a heterogeneous mixture of cells, such as stroma, osteoblasts, endothelial cells and extracellular matrix (ECM) [44].

The intrinsic, stochastic model hypothesizes that specific factors, such as transcription factors, which are important for the appropriate cell fate decisions are expressed in HSC and can be up or downregulated by random-based autoregulatory mechanisms. Transcription factors can drive the expression of relevant lineage-specific genes, thereby initiating the phenotypic change in the progenitor cell down a specific differentiation pathway [45-47]. As an example, transcription factor Pax5 controls the development of B-cells by repressing lineage-inappropriate genes and activating B-cell-specific genes [48]. According to this model, HSC compared to more differentiated progenitor cells express a wide range of genes at basal levels [49].

In another model, exogenous signals can determine the HSC fate via their cognate receptors whose cytoplasmic domains activate various pathways involved in survival, proliferation, and differentiation [50, 51]. For instance, it is well known that low level of TGF- β 1 maintains the multiple differentiation potentials of hematopoietic stem and progenitor cells, which is associated with its negative controlling effect on cells cycling. TGF- β 1 functions as a key physiological factor ensuring the maintenance of a stem cell reserve [52].

The innovative use of imaging techniques demonstrated the influence of lineage-specific differentiation of hematopoietic progenitor cells on single-cell level after the using cytokines [40].

1.1.4 Process of leukemogenesis

Accumulation of genetic alterations in HSC causes dysregulation of these cells which finally leads to neoplasia. In accordance with available models for the origin and development of neoplasia, transformation to leukemia has long been thought to be a multistep process. Foulds first proposed a multistep progression model in which a normal cell should go through a number of different intermediate stages before a malignancy develops [53-55]. This model is consistent with a dysregulation process in myelodysplastic syndrome (MDS) in which BM hyperplasia and ineffective hematopoiesis leads to dysplasia. Subsequent events involving several genetic mutations such as *p53*, *Rb* or *ras* can enable the subclone to escape from apoptotic process and develop to leukemia (Figure 4).

These genetic aberrations can be initiated by exogenous factors such as chemicals, radiation and chemotherapeutic agents or by endogenous random circumstances. Extensive damages to the genetic material are normally recognized by the cell and these cells are eliminated by apoptosis [55].

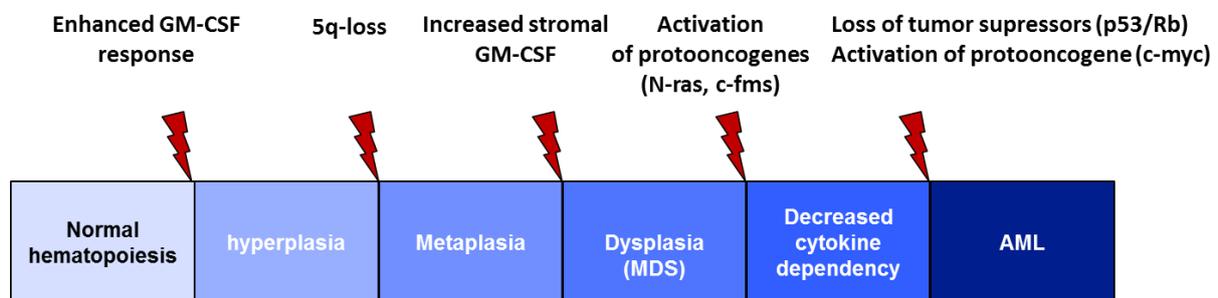


Figure 4. Hypothetical model for the progression of acute myeloid leukemia (AML) involving 5q loss. Model of abnormal myeloproliferative evolution in which early events, including alterations in cytokine response and loss of heterozygosity (5q), are followed by activation of at least one protooncogene coupled with the successive loss of a tumor suppressor gene. Adapted from [55]. MDS: myelodysplastic syndrome [55].

Due to the accessibility of tumor material and successful cytogenetic techniques which detect chromosomal alterations, acute leukemia is one of the best characterized genetic diseases. Childhood acute lymphoblastic leukemia (ALL) is characterized by recurring aneuploidies (e.g., high hyperdiploidy and hypodiploidy) and chromosomal translocations such as t(9;22) BCR-ABL1 and rearrangement of the TLX1, TLX3, LYL1 and MLL genes in T-

lineage ALL (T-ALL) [56-58]. With the exception of BCR-ABL1 these abnormalities are less common in adult ALL [59].

Development of high throughput technology has enabled to investigate the mutation status at both RNA and DNA level and detect mutations in leukemic cells with no cytogenetic abnormalities. It is shown that leukemic clones carry several genetic changes which contribute to the development of leukemia [60-62]. Genetic analyses of leukemic clones have demonstrated the occurrence of several mutations per leukemic cell and thus different subclones with specific changes. This allows tracking both the dynamics and development of individual subclones and suggests that several clones are able to give rise to leukemia [63, 64].

1.2 Wiskott-Aldrich Syndrome

Wiskott–Aldrich Syndrome (WAS) is an X-linked primary immunodeficiency disorder that is characterized by severe immunodeficiency, microthrombocytopenia, eczema and an increased incidence of autoimmunity and malignancies [65]. The incidence of this rare disorder is approximately one to four cases per 1,000,000 live male births [66] [67].

The gene responsible for WAS is located on the short arm of the X chromosome at Xp11.22–p11.23 [65]. The *WAS* gene encodes the WAS protein (WASp), which is a 502-amino acid protein. WASp is involved in actin polymerization and associated coupling of receptor engagement, signaling events, and cytoskeletal rearrangement and is exclusively expressed in hematopoietic cells [68, 69]. There is a strong correlation between the level of WASp and the severity of WAS. Complete WASp absence leads to a pronounced defect in the function of multiple hematopoietic cell lineages, resulting in thrombocytopenia with small platelets and progressive lymphopenia with abnormal lymphoid and myeloid cell function [70]. On the other hand, decreases in WASp levels have variable effects on the function of the different cell types, especially platelets [71].

WAS disease is caused by mutations in the *WAS* gene expressed exclusively in hematopoietic cells. Since actin cytoskeleton plays a prominent role in the basic mechanisms of cell adhesion and migration, the effects of *WAS* gene mutations on these processes has become more substantial [72-74]. Heterogeneous mutations spanning the entire *WAS* gene have been described. These mutations in *WAS* alter the function or expression of the WASp intracellular protein. Around 300 unique mutations spanning the *WAS* gene have been described. The most frequent ones are missense mutations, followed by splice-site mutations, deletions, nonsense mutations, insertions, and complex mutations [65, 70, 75].

Most missense mutations are located in exons one to four, splice-site mutations are predominantly in introns six to ten, whereas nonsense mutations, insertions, deletions, and complex mutations are distributed throughout the whole WAS gene [76, 77]. The effect of a given mutation on WASp expression correlates with severity of disease with some exceptions. In general, mutations that abolish WASp expression or result in the expression of a truncated protein are associated with WAS disease [76, 78, 79].

1.2.1 Treatments of Wiskott–Aldrich Syndrome

1.2.1.1 Hematopoietic stem cell transplantation

Despite advances in diagnosis and clinical care, the prognosis of WAS disease remains poor. Allogeneic hematopoietic stem cell transplantation (HSCT) can be considered, especially if a human leukocyte antigen (HLA)–identical related sibling donor is available [80]. The best transplantation outcome has been achieved with HLA-identical sibling donors and matched unrelated donors when the age of the recipient is less than five years at the time of the transplant [81, 82]. Acute or chronic graft-versus-host disease (GVHD) is observed especially in the first year after HSCT but can persist longer, and is treated with immunosuppression. Recently, autoimmunity independent of GVHD has been observed in patients with WAS following HSCT, and is associated with reduced donor chimerism [83, 84].

1.2.1.2 Gene therapy

Gene therapy is an alternative to HSCT in the treatment of WAS [85-87]. To avoid the immunological barriers imposed by GVHD in allogeneic HSCT, autologous gene-modified HSCT has been used [88]. In addition, recent advances in gene therapy make it an attractive alternative to HSCT for WAS, especially in cases where no HLA-matched donor is available.

The procedure of gene therapy consists of isolating autologous HSC (CD34⁺) from the patient, transducing the cells *ex vivo* with an RV that expresses WASp as a transgene, and then reinfusing them back into the patient (Figure 5) [89]. After hematopoietic reconstitution, the clonality, kinetics and contribution of gene - corrected cells to blood formation can be observed by means of specific PCR technologies [90-93]. The stable persistence of the corrected gene results in clinical improvement.

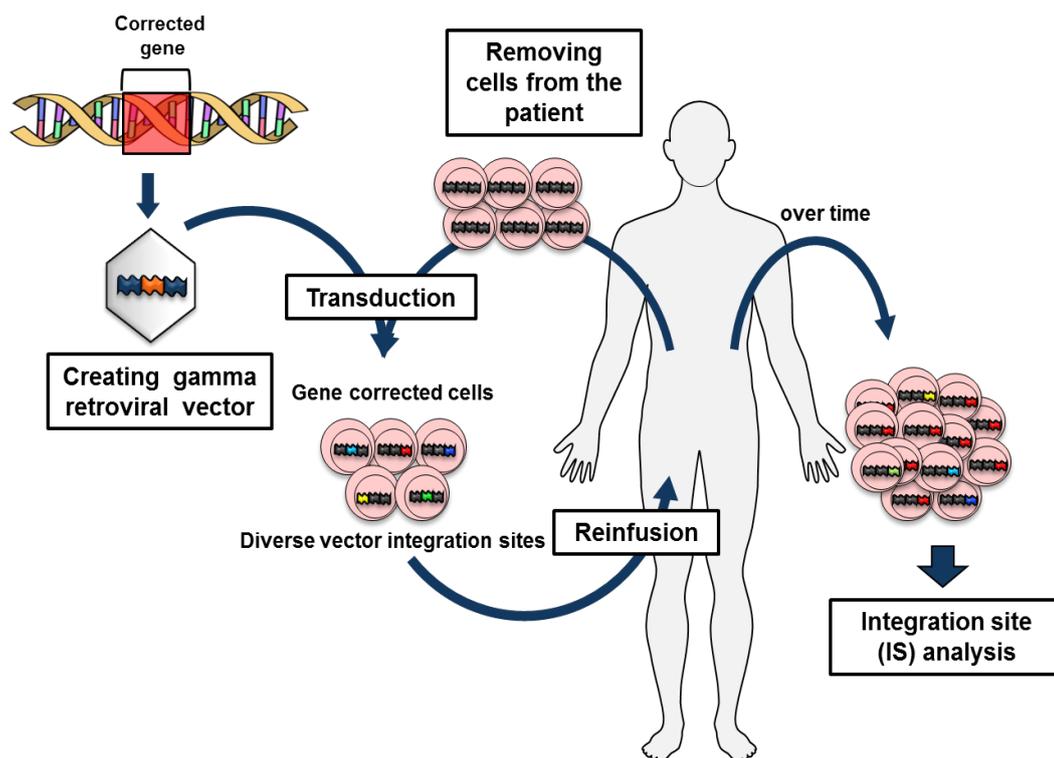


Figure 5. Retroviral gene transfer into hematopoietic stem cells in gene therapy. The corrected transgene is introduced to $CD34^+$ cells by gamma retroviral vectors. After transduction of the hematopoietic progenitor cells, they are transferred back into the patient. The gene-corrected cells can be analyzed using LAM-PCR. Adapted from [94]. IS: integration site.

After WAS gene therapy trials with gammaretroviral vector, 9 out of 10 treated patients showed sustained engraftment and correction of WASp expression in lymphoid and myeloid cells and platelets. Gene therapy resulted in partial or complete resolution of immunodeficiency, autoimmunity, and bleeding diathesis [95, 96].

One of the side effects of gene therapy is the integration of gammaretroviral vectors into the host genome which might cause serious consequences (section 1.3.1). It was observed that 7 out of 10 WAS patients treated with gene therapy developed leukemia two to five years after bone marrow transplantation (BMT) [97].

Efforts at improving the efficacy and safety of autologous gene-modified HSCT have recently focused on using of lentiviral vectors. Three patients with WAS received autologous gene-modified HSCT using cells transduced with lentiviral vectors encoding for WASp [85, 86]. Stable and durable integration of the corrected gene and expression of WASp was noted [98]. No insertional oncogenesis-associated events have been reported so far; however, long-term safety and efficacy in gene therapy trial using lentiviral vectors are still being studied.

1.3 Retroviral vector system

Viral vectors contain viral elements on the transfer vector. These elements include viral packaging signal (ψ), which is involved in regulating packaging the retroviral RNA genome into the viral envelope during replication, a promoter which initiates the transcription of the transgene and a polyadenylation signal which marks the end of ectopic RNA. Primer binding site (PBS) is required for initiating reverse transcription. Polypurine tract (PPT) allows the viruses to generate double strand DNA from RNA genome. Long terminal tandem repeats (LTRs) are identical sequences at 5' and 3' ends of viral genome which are necessary for the process of DNA synthesis. Moreover, they are used by viruses to insert their genetic material into the host genomes. To reduce the risk of productive recombination and replication by the viral vectors and increase their safety level, several accessory viral genes such as vif, vpr, vpu and nef were removed [99].

Virus particles encoding transgenes can be produced in a packaging cell line such as HEK293. Packaging cells have been either engineered to stably express viral proteins such as capsid and reverse transcriptase or they get transfected by the helper plasmids encoding viral genes. Produced viral particles can transduce target cells and integrate their genetic material into the host genome after reverse transcription (Figure 6).

Viral vectors are widely used in gene therapy trials to stably integrate the coding sequences of the therapeutic genes into patient's genome [94, 99-103].

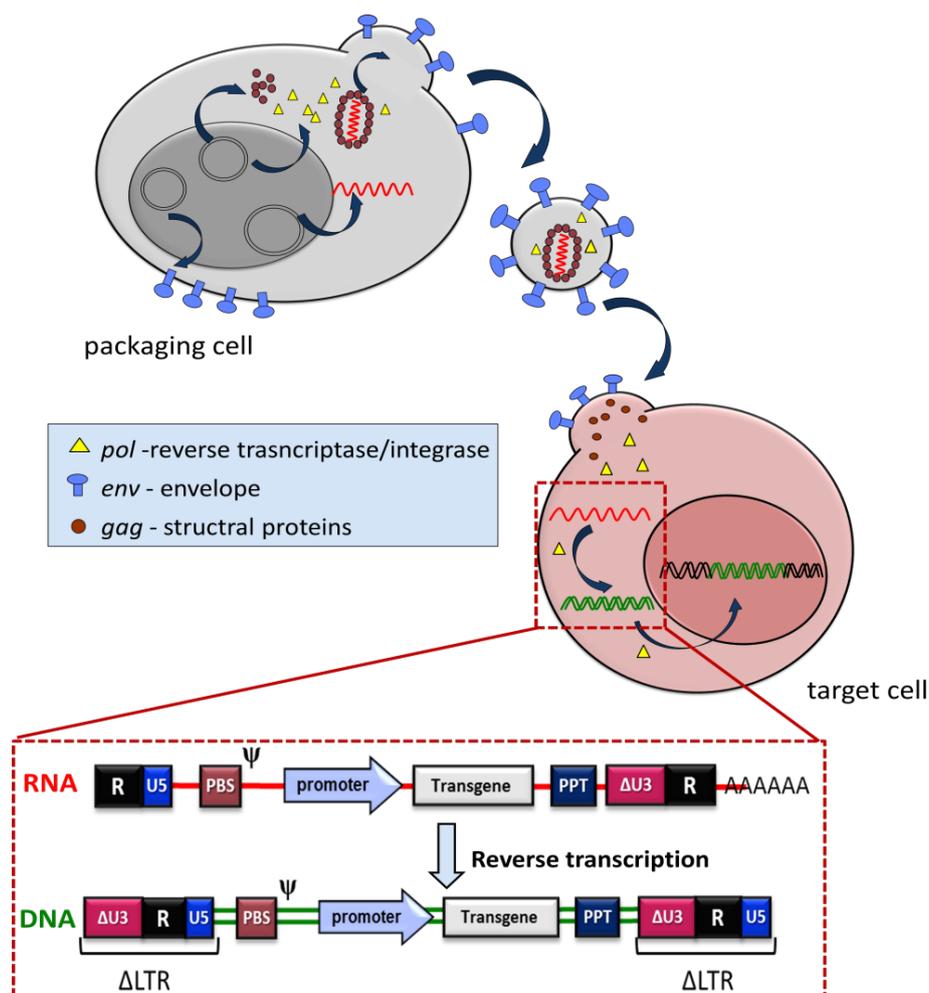


Figure 6: Packaging and transduction process of retroviral vectors. Retroviral vectors are generated by co-transfection of a packaging cell line like HEK293 cells with the cDNA expressing transfer plasmid along with two helper/packaging plasmids which encode the structural capsid (*gag*) and envelope (*env*) proteins, reverse transcriptase and integrase enzymes. Packaging cells produce viral particles, whose genome only encodes sequences from the transfer plasmid, and can transduce the target cells. Transduction starts when the viral Env protein interacts with the cellular receptor and enters the target cell. The RNA genome is reverse transcribed into dsDNA which then enters the nucleus, is integrated into the host genome and gets expressed by the host cell's machinery. Proviral mRNA contains a packing signal for the viral envelope (ψ), a binding site for primer (PBS) during reverse transcription, and a polypurine tract (PPT) which allows the synthesis of double-stranded DNA. Adapted from [104].

1.3.1 Insertional mutagenesis: the side effect of retroviral gene transfer

Insertional vectors are gene delivery tools in clinical gene therapy, providing constant transgene expression in the host cell genome [86, 105, 106]. However, severe side effects of gammaretroviral vector insertion have raised concern about the safety of gene therapy (section 1.2.1.2) [107-109]. Two to five years after gammaretroviral gene therapy trial, seven

patients developed acute leukemia (one acute myeloid leukemia (AML), four T cell acute lymphoblastic leukemia (T-ALL), and two primary T-ALL with secondary AML after chemotherapy) [97].

Insertional analysis demonstrated that vector integration occurred preferentially in the vicinity of transcription start sites and was clustered in regions proximal to proto-oncogenes. It has been shown that the integration of gammaretroviral vector used in gene therapy is semi-random with the preference of integrating in ± 5 -10 kb from the transcription start sites (TSS) [94, 99-103]. Due to the strong enhancer and promoter in the LTR region of RV, other genes in vicinity of RV integration might get transcriptionally activated because of insertional mutagenesis (Figure 7). This constant activation might lead to clonal expansion up to malignant transformation over time. Studying the IS in gene therapy patients treated for primary immune deficiency diseases revealed that years after BMT, clones carrying the IS in vicinity of specific genes such as LMO2 and MDS1/EVI1 become dominant over other clones. Analyzing the IS repertoire of WAS gene therapy patients which developed leukemia reveals that LMO2 is one of the most frequently targeted locus which gets activated due to insertional mutagenesis [97]. The gene therapy-induced leukemia and the mouse models suggest that enforced expression of LMO2 causes a clonal growth advantage in T-cell lineage which eventually leads to T-cell acute lymphoblastic leukemia (T-ALL) [110, 111].

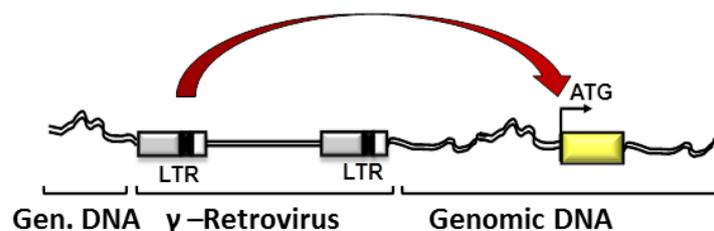


Figure 7: Insertional mutagenesis following retroviral integration. Genes in vicinity of retroviral integration might get activated by proviral enhancer/promoter. LTR: long tandem repeat.

The genotoxicity risk of gammaretroviral vectors can be reduced by replacing the strong promoter-enhancer of vector LTR with a relatively weak internal promoter derived from endogenous genes to produce a self-inactivating (SIN) vector [112]. It was shown that SIN lentiviral vectors with a promoter derived from the human phosphoglycerate kinase gene (PGK) did not accelerate tumor induction in lymphoma-prone mice, in contrast to gammaretroviral vectors with strong enhancer–promoter sequences located in the LTRs [113].

1.3.2 Monitoring clonal dynamics of marked hematopoietic stem cells by LAM-PCR

Several PCR-based methods are available for the identification and characterization of unknown flanking DNA sequences [91, 92, 103, 114].

The linear amplification-mediated PCR (LAM-PCR) proved to exhibit the highest sensitivity, allowing the detection of various vector integration sites (IS) in one sample. LAM-PCR developed by our group characterizes unknown DNA regions flanking viral IS with a sensitivity down to single-cell level [101]. The integration of retroviral vectors into distinct chromosomal sites creates unique sequences between the genomic DNA and 5' and 3' LTRs of the vector that makes vector IS a clonal marker of each transduced cell and all of its progeny [101].

The broad application spectrum of LAM-PCR has been approved by its application as a tool for the molecular follow up of gene-modified cells [109, 115-117]. The combination of this sensitive PCR method with high-throughput sequencing allows detecting the IS repertoire, clonality and the viral integration patterns in pre-clinical and clinical studies [87, 116, 118-120].

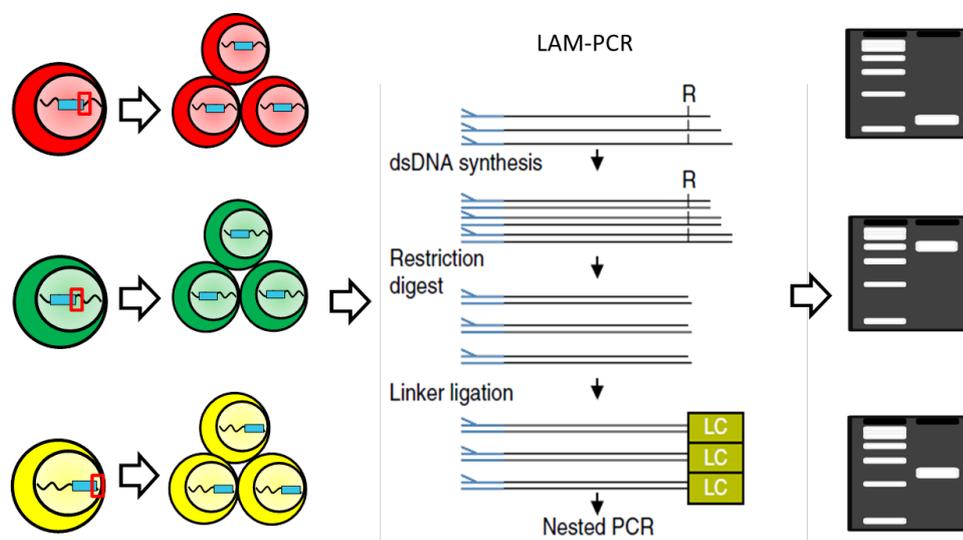


Figure 8: LAM-PCR detects the precise position of viral IS in the genome. A PCR product which starts from the known sequence of the vector (LTR) extends through the unknown neighboring genome. After performing LAM-PCR and sequencing the PCR-amplified DNA fragments the precise integration position within the genome will be revealed. LAM-PCR: linear amplification-mediated PCR; LTR: long tandem repeat.

LAM-PCR is a multiple step procedure. LTR-specific, biotinylated primers are used to amplify the genome-RV junctions. Single-stranded DNA fragments composed of the known primer-vector-sequence and unknown adjacent part of the host cells genome were immobilized on streptavidin-coated magnetic beads. Double-stranded DNA was synthesized and subsequently digested with restriction enzymes. A restriction-site-specific linker of a known sequence was coupled to the remaining double-stranded DNA. The DNA fragments were denatured and removed from the solid phase. Single-stranded DNA was further amplified using linker-specific and nested primers. Finally, amplified DNA was purified, concentrated and submitted for high resolution gel electrophoresis (Figure 8). To detect the precise location of the vector IS in the genome, amplified DNA was sequenced by high-through-put next generation sequencing platforms [101, 121].

Specific aims

In clinical gene therapy studies, retroviral vectors (RV) are used to stably integrate a therapeutic transgene into the host cell genome promoting its long term expression in all progenies of transduced blood cells. To monitor clonal selection and vector-associated genotoxicity in individual subjects the complete integration site (IS) repertoire has been analyzed in Wiskott Aldrich Syndrome (WAS) gene therapy patients using highly sensitive Linear Amplification Mediated (LAM) PCR. Detailed IS analyses proved that RV integrations are semi-randomly distributed and that common integration sites (CIS) cluster close to transcription start sites of genes possibly pointing to an *in vivo* selection. Moreover, we have observed that therapeutic-vector-mediated transcriptional activation of neighboring genes can lead to leukemia. Analyzing the IS repertoire in normal individuals enabled us to monitor the dynamics of each individual RV-marked clone in highly polyclonal human post-transplant hematopoiesis. We hypothesize that systematic re-analysis of this large and unique clinical IS dataset can help to identify candidate genes involved in the regulation of hematopoietic stem cells. Towards a precise selection and a further molecular understanding of the regulatory role of candidate genes in hematopoiesis determined by gammaretroviral CIS we specifically aim:

Aim 1: To select candidate hematopoietic regulatory genes

We aim to establish a strategy for selecting novel hematopoietic regulatory genes based on systematically investigating the genome-wide RV integration site pattern obtained from 10 clinical WAS gene therapy patients. Moreover, we aim to assess the contribution of human hematopoietic cell clones harboring IS close to the candidate genes to blood formation in individual patients for up to four years. To further select for stem cell relevant genes we seek to determine the endogenous expression level of each candidate in well-defined murine hematopoietic stem and progenitor cell populations.

Aim 2: To examine the impact of selected candidate genes on hematopoietic stem cells *in vitro* and *in vivo*

To validate and characterize the molecular mechanism of the selected candidate genes in hematopoietic cells lentiviral based overexpression vectors will be used. We aim to investigate the influence of the selected genes on proliferation, viability and cytokine dependency *in vitro* as well as long term multilineage reconstitution and self-renewal of hematopoietic primary cells *in vivo*.

Systematic identification of novel candidate regulatory genes in meta-datasets derived from a large number of gene therapy studies and subsequent validation *in vitro* and *in vivo* will help us to gain new insights into the biology of post-embryonic hematopoiesis.

3 Material and methods

3.1 Materials

3.1.1 Equipment and devices

Equipment	manufacturer
Analytical Balance TE 124S	Sartorius, Göttingen
Biofuge® pico	Heraeus, Hanau
Camera Lumix DMC-FZ50	Panasonic, Hamburg
Centrifuge inserts	Kendro, Langenselbold
Centrifuge Multifuge® 3SR	Heraeus, Hanau
Cryobox Nalgene	Thermo Fisher Scientific, Schwerte
Electrophoreses Power Supply 200/2000	Elchrom Scientific, Cham
Flow Cytometer BD™ LSRII	Becton, Dickinson and Company, Heidelberg
Flow Cytometer FACS Aria™ Cell Sorter	Becton, Dickinson and Company, Heidelberg
Fluorescence Microscope Axiovert 200	Zeiss, Oberkochen
Freezer -20°C	Liebherr, Biberach an der Riss
Freezer -80°C	Sanyo (Panasonic), Hamburg
Fridge 4°C	Liebherr, Biberach an der Riss
Gel Documentation	Peqlab, Erlangen
Gel Electrophoreses Chamber	Biometra, Göttingen
Incubator Heracell® 150	Thermo Fisher Scientific, Schwerte
Isoflurane Vaporizer Vapor 19.3	Dräger, Lübeck
LightCycler® 480 Real-Time PCR	Roche, Mannheim

Light Microscope	Zeiss, Oberkochen
Microplate Reader	Tecan, Männedorf
Microwave	Bartscher, Salzkotten
Mini Protean® Tetra Cell	BioRad, Munich
Molecular Imager® ChemiDoc™ XRS+	BioRad, Munich
Multipette Plus	Eppendorf, Hamburg
NanoDrop® Spectrophotometer ND-1000	Peqlab, Erlangen
Neubauer counting chamber	Marienfeld, Lauda-Königshofen
Nitrogen System	German-Cryo, Jüchen
PCR-Thermocycler	Landgraf, Langenhagen
Pipetboy acu	Integra Biosciences, Fernwald
Pipettes Research® (10µl; 20µl; 200µl; 1000µl)	Eppendorf, Hamburg
Power Pac™ HC Power Supply	BioRad, Munich
Precision Balance TE3102S	Sartorius, Göttingen
Rotator Reax2	Heidolph, Schwabach
Safety Cabinet Herasafe® KS Thermo Fisher Scientific	Thermo Fisher Scientific, Schwerte
Shaking incubator	Axon, Kaiserslautern
Submerged Gel Electrophoresis Apparatus SEA 2000®	Elchrom Scientific, Cham
TC10™ Automated Cell Counter	BioRad, Munich
Thermo Cycler TPersonal	Biometra, Göttingen
Thermo Mixer comfort	Eppendorf, Hamburg
Trans-Blot® semi-dry cell	BioRad, Munich
Transilluminator	Biotec-Fischer, Reiskirchen

Ultracentrifuge L8-M with Rotor SW27	Beckman Coulter, Krefeld
Vacuum pump	Merck Millipore, Darmstadt
Video documentation system	Peqlab, Erlangen
Vortexer MS1	IKA, Staufen
Water Bath Haake SWB25	Thermo Fisher Scientific, Schwerte
Water treatment plant	TKA, Niederelbert

3.1.2 Plastic and disposables

Cell Culture Flasks, EasyFlasks™ (25cm ² ; 75cm ² ; 175cm ²)	Nunc (Thermo Fisher Scientific), Schwerte
Cell Culture Plates (6-; 12-; 24-; 48-; 96-well)	Becton, Dickinson and Company, Heidelberg
nylon cell strainer (40-100 µm pore size)	Becton, Dickinson and Company, Heidelberg
cotton swabs, sterile	Noba Verbandmittel Danz GmbH, Wetter
Cover Glasses	Marienfeld, Lauda-Königshofen
Cryotubes	Corning, Kaiserslautern
FACS Tubes, BD™ Falcon™ Round-Bottom Tube (5ml)	Becton, Dickinson and Company, Heidelberg
Filter (0.22µm pore size)	Merck Millipore, Darmstadt
Hollow Needle	Becton, Dickinson and Company, Heidelberg
Hollow Needle, blunt end 18G	Becton, Dickinson and Company, Heidelberg
Lab Gloves (Nitril)	Microflex, Reno

Parafilm	Pechiney Plastic Packaging, Chicago
PCR reaction tube (V=0,2 ml)	Genaxxon, Ulm
Petri dish (d=10 cm, 3.5cm)	Genaxxon, Ulm
Petri dish Nunclon™ (d=15 cm)	Thermo Electron, Langenselbold
Pipette Tips (200µl, extended)	Thermo Fisher Scientific, Schwerte
Pipettes (V=2 - 50 ml)	Genaxxon, Ulm
Pipette tips (V=10/20/200/1000 µl)	Starlab, Hamburg
PVDF membrane	BioRad, Munich
scalpel	Feather Safety, Osaka
Stericup vacuum filtration system	Merck Millipore, Darmstadt
Syringe (2,5/5/20/50 ml) Omnifix® Solo	B. Braun, Melsungen
Twin.tec PCR plate 96 well	Eppendorf, Hamburg
Ultra Low Attachment Flasks	Corning, Kaiserslautern
Ultracentrifuge tubes	Beraneck Laborgeräte, Weinheim
Whatman filter paper Protean®	BioRad, Munich

3.1.3 Bacteria culture

1x LB Medium

<u>reagents</u>	<u>manufacturer</u>	<u>amount</u>
Luria Broth Base (LB)	Invitrogen, Darmstadt	25g
H ₂ O		Ad 1000ml

Autoclaved and stored at 4°C, antibiotic was added before use.

<u>reagents</u>	<u>manufacturer</u>
SOC-Medium (OneShot-Kit)	Invitrogen, Darmstadt

LB-Agar

<u>reagent</u>	<u>manufacturer</u>	<u>amount</u>
Luria Broth Base (LB)	Invitrogen, Darmstadt	25g
Agar	Sigma-Aldrich, Munich	12,5g
H ₂ O		Ad 1000ml

The solution was autoclaved, cooled down at 60 ° C in a water bath. Antibiotic was added before pouring to plates. After solidification the plates were kept at 4 ° C.

3.1.4 Cell cultureHEK-293T and HeLa cells

<u>reagent</u>	<u>manufacturer</u>	<u>amount</u>
Iscove's modified Dulbecco's Medium (IMDM)	Gibco [®] Invitrogen, Darmstadt	500ml
+ L-Glutamine (200mM)	Gibco [®] Invitrogen, Darmstadt	5ml
+ FBS, heat-inactivated, filter sterilized	Biosera, Sussex	50ml

32D cells

<u>reagent</u>	<u>manufacturer</u>	<u>amount</u>
Roswell Park Memorial Institute (RPMI)	Gibco [®] Invitrogen, Darmstadt	100ml
+ L-Glutamine (200mM)	Gibco [®] Invitrogen, Darmstadt	1ml
+ FBS, heat-inactivated, filter sterilized	Biosera, Sussex	10ml
+WEHI-3B supernatant	produced according to (3.2.2.3.3)	10ml

WEHI-3B cells

<u>reagent</u>	<u>manufacturer</u>	<u>amount</u>
Dulbecco's Modified Eagle Medium (DMEM)	Gibco® Invitrogen, Darmstadt	500ml
+ L-Glutamine (200mM)	Gibco® Invitrogen, Darmstadt	5ml
+ FBS, heat-inactivated, filter sterilized	Biosera, Sussex	50ml

FDCP-Mix cells

<u>reagent</u>	<u>manufacturer</u>	<u>amount</u>
IMDM	Gibco® Invitrogen, Darmstadt	500ml
+ L-Glutamine (200mM)	Gibco® Invitrogen, Darmstadt	5ml
+ FBS, heat-inactivated, filter sterilized	Biosera, Sussex	100ml
+ rmlL3	R&D Systems, Wiesbaden	10 ng/ml

3.1.5 Freezing and thawing medium**Freezing medium**

<u>reagent</u>	<u>amount</u>
respective cell medium	55%
FBS	30%
DMSO	15%

Thawing medium

<u>reagent</u>	<u>amount</u>
respective cell medium	80%
FBS	20%

3.1.6 Cytokines, medium supplements and solutions for cell culture**Cytokines**

<u>reagents</u>	<u>manufacturer</u>
rmFlt3L	R&D Systems, Wiesbaden
rmIL3	R&D Systems, Wiesbaden
rmSCF	R&D Systems, Wiesbaden
rhTPO	R&D Systems, Wiesbaden

Medium supplements and solutions for cell culture

<u>reagents</u>	<u>manufacturer</u>
Dulbecco's Modified Eagle Medium Pulver	Gibco [®] Invitrogen, Darmstadt
Dulbecco's phosphate buffered saline (DPBS)	Gibco [®] Invitrogen, Darmstadt
Fetal Bovine Serum (FBS)	PAN Biotech, Aidenbach; Biosera
Hank's balanced salt solution (H)	Sigma-Aldrich, Munich
MethoCult M3434	StemCell Technologies, Grenoble
MethoCult M3630	StemCell Technologies, Grenoble
MethoCult H4230	StemCell Technologies, Grenoble
Polybrene	Millipore, Schwalbach
Polyethylenimin (PEI)	Sigma-Aldrich, Munich
Protaminsulfate	Sigma-Aldrich, Munich
Trypsin-EDTA 0,05%	Invitrogen, Darmstadt

Antibiotics**reagents**Ampicillin (100mg/ml mit H₂O)**manufacturer**

Sigma-Aldrich, Munich

3.1.7 Antibodies**Antibodies for Western blot****reagent**mouse-anti-Human- α -Tubulin

EVL (C-17), sc-66527

PLCB4 (A-9), sc-166131

manufacturer

Invitrogen, Darmstadt

Santa Cruz, Heidelberg

Santa Cruz, Heidelberg

Antibodies for flow cytometry**reagent**

CD16/CD32

anti mouse CD45.1-PE

anti mouse CD45.2-PerCPCy5.5

anti mouse CD3-PerCPCy5.5

anti mouse Ter119-PECy7

anti mouse CD11b-PerCPCy5.5

anti mouse Ly6G-Alexa700

anti mouse CD117-PE

anti mouse Sca-1-Alexa Fluor700

anti mouse Lineage cocktail-APC

anti mouse CD150-PE/Cy7

manufacturer

BD Biosciences, Heidelberg

eBiosciences, Frankfurt

BD Biosciences, Heidelberg

Biolegend, Fell

anti mouse CD135-PE	BD Biosciences, Heidelberg
anti mouse CD48-PerCpCy5.5	BD Biosciences, Heidelberg
anti mouse CD34-FITC	BD Biosciences, Heidelberg
mouse IgG2a,FITC	BD Biosciences, Heidelberg
mouse IgG2a, κ -PE-Cy7	BD Biosciences, Heidelberg
mouse IgG2a, κ - Alexa Fluor700	BD Biosciences, Heidelberg
mouse IgG1, κ -PE	BD Biosciences, Heidelberg
mouse IgG2a, κ -PE	BD Biosciences, Heidelberg
mouse IgG2a, κ - PerCPCy5.5	BD Biosciences, Heidelberg
mouse IgG1, κ -APC	BD Biosciences, Heidelberg

3.1.8 PCR primers and reagents

RT-PCR primers

<u>designation</u>	<u>sequence(5'-3')</u>
PLCB4_Fwd	ATGCGGGTACCTTCTCAAGC
PLCB4_Rev	TTTCCGTATGGTGTTCGGTGG
ZNF217_Fwd	TGAGGATGGACTCCCTGACG
ZNF217_Rev	GCTGCGGCATACTCACAGAA
IRF2BPL_Fwd	AGATGCTAGCTGTCCCATGC
IRF2BPL_Rev	TGTTCCCTACCGAGCTTCAG
EVL1_Fwd	ATGAGTGAACAGAGTATCTGCC
EVL1_Rev	TCTTTGCCACAGACGGGGTT
ZNF217_opt Fwd	ACCCCGAAGTGCTGATGATG
ZNF217_opt Rev	ACTTGCTGTGAGGGCTGAAA
EVL1_opt Fwd	ATCTACCACAACACCGCCAG

EvL1_opt Rev	AGGTGGTGGCTTCCTCTTTG
IRF2BPL_opt Fwd	AAACAGAGCCGAGGAATGGG
IRf2BPL_opt Rev	GCCGGTGGGATACTCGATG
PLCB4_opt Fwd	GAAGTCCGAGGGCAAAGAGG
PLCB4_opt Rev	CACCATGTAGGTGAAGCCGA
LRR33_opt Fwd	CCGACAACAGACTGAGCGAG
LRR33_opt Rev	TCGAAGATGCTGTCGTCCAG
TBP-Mus_Fwd-new	CCT TCA CCA ATG ACT CCT ATG AC
TBP-Mus_Rev-new	CAA GTT TAC AGC CAA GAT TCA C

Genotype-PCR primers

designation

sequence(5'-3')

PLCb4-MT_Rev	GCT ACT TCC ATT TGT CAC GTC C
PLCb4-WT-MT_Fwd	TGT CGG TGT TGG GTG AGA C
PLCb4-WT_Rev	GGA TGT TCG TGG TAG CAC CT

Genotype-PCR reagents

reagents

manufacturer

dNTPs (mM)	Genaxxon, Ulm
Red Taq Master Mix	Genaxxon, Ulm

3.1.9 Plasmids

designation	function
p101 pMDL gag pol rre	CMV-driven HIV-1 construct expressing Gag and Pol with RRE (Rev responsive element)
p102 pRSV-rev	RSV-driven HIV-1 construct expressing Rev
p103 pMD2.VSVG	CMV-driven construct expressing VSV-G
pCCLsin.ppt.PGK.EVL.IRES.eGFP.Wpre	PGK-driven construct expressing EVL and GFP
pCCLsin.ppt.PGK.ZNF217.IRES.eGFP.Wpre	PGK-driven construct expressing ZNF217 and GFP
pCCLsin.ppt.PGK.IRF2BPL.IRES.eGFP.Wpre	PGK-driven construct expressing IRF2BPL and GFP
pCCLsin.ppt.PGK.PLCB4.IRES.eGFP.Wpre	PGK-driven construct expressing PLCB4 and GFP
pCCLsin.ppt.PGK.LRRC33.IRES.eGFP.Wpre	PGK-driven construct expressing LRRC33 and GFP
pCCLsin.ppt.PGK.IRES.eGFP.Wpre	PGK-driven construct expressing GFP

3.1.10 Chemicals and reagents

reagents

Agarose

Aqua ad injectabilia (aqua dest)

Baytril®

Bovine Serum Albumin

Dimethylsulfoxide (DMSO)

manufacturer

Serva, Heidelberg

B. Braun, Melsungen

Bayer, Leverkusen

Sigma-Aldrich, Munich

Sigma-Aldrich, Munich

DNA ladder 100 bp	Invitrogen, Darmstadt
DNA ladder 1kb	Invitrogen, Darmstadt
Ethanol	Sigma-Aldrich, Munich
Ethidiumbromide (0.07%) AppliChem	AppliChem, Darmstadt
Fluorogold	Invitrogen, Darmstadt
Isopropyl alcohol	Sigma-Aldrich, Munich
Loading Buffer (5x)	Elchrom Scientific, Cham
MgCl ₂ 100mM	Sigma-Aldrich, Munich
Propidiumiodid	Invitrogen, Darmstadt
RNase/DNase free water	Ambion, Darmstadt
Sodium azide pure (NaN ₃)	AppliChem, Darmstadt
Sodium Chloride (NaCl)	VWR International, Vienna
Tris HCl pH 7.5 (1M)	USBiological, Swampscott
Precision Plus Protein™ Standard	Bio Rad, Munich

3.1.11 Enzymes and reaction buffers for cloning

T4 DNA-Ligase + buffer	New England Biolabs, Frankfurt
Restriction endonuclease digestion buffers	New England Biolabs, Frankfurt
Restriction endonucleases	New England Biolabs, Frankfurt

3.1.12 Commercial kits

designation

Blood and Tissue Kit

EasySep

EndoFree® Plasmid Purification Kit

SYBR® Select Master Mix

Plasmid Miniprep DNA Purification Kit

QIAquick® Gel Extraction Kit

manufacturer

Qiagen, Hilden

StemCell Technologies, Grenoble

Qiagen, Hilden

Roche, Mannheim

GeneMATRIX

Qiagen, Hilden

QIAquick [®] PCR Purification Kit	Qiagen, Hilden
RNeasy Micro Kit	Qiagen, Hilden
RNeasy Mini Kit	Qiagen, Hilden
Super-Script III Synthesis-Kit	Invitrogen, Darmstadt
Topo-TA-Cloning [®] Kit	Invitrogen, Darmstadt
Western Lightning Plus-ECL	Perkin Elmer, Rodgau
ARCTURUS [®] PicoPure [®] RNA Isolation Kit	Life Technologies, Frankfurt
CellTiter-Glo [®] Luminescent	Promega, Mannheim
Dual-Light [®] Luciferase & β -Galactosidase Reporter Gene Assay System	Life Technologies, Frankfurt

3.1.13 Western blot buffers and reagents

RIPA-Buffer

<u>reagents</u>	<u>final concentration</u>
NaCl	150mM
Tris	50mM
Nonidet P-40	1%
Na-Deoxycholate	0,5%
Sodium monododecyl sulfate	1%
	pH7,5

1:25 Roche Complete Proteaseinhibitor was added to RIPA right before preparing the cell lysate.

<u>reagents</u>	<u>manufacturer</u>	<u>amount</u>
Complete [®] Proteaseinhibitor	Roche, Mannheim	1 tablet
H ₂ O		2 ml
Aliquote and stored at -20°C		

Lower Gel buffer (LGP)

<u>reagents</u>	<u>final concentration</u>
TRIS	1,5M
Natriumdodecylsulfat (20%)	0,4%
	pH8,8

Upper Gel Buffer (UPG)

<u>reagents</u>	<u>final concentration</u>
TRIS	0,5M
Sodium monododecyl sulfate (20%)	0,4%

Resolving gel (1x)

<u>reagents</u>	<u>5%</u>	<u>7,5%</u>	<u>10%</u>	<u>13%</u>	<u>15%</u>
LGB (ml)	2	2	2	2	2
H ₂ O (ml)	4,6	4	3,3	2,5	2
Acrylamide (30%; ml)	1,3	2	2,6	3,4	4
APS (10%; µl)	50	50	50	50	50
TEMED (µl)	5	5	5	5	5

Stacking gel

<u>reagents</u>	<u>5%</u>
UGB (ml)	1,25
H ₂ O (ml)	3
Acrylamide(30%; ml)	0,8
APS (10%; µl)	50
TEMED (µl)	5

SDS-running buffer (10x)

<u>reagents</u>	<u>final concentration</u>
TRIS	25mM
Glycine	190 mM
Sodium dodecyl sulfate (20%)	1%

Blotting buffer

<u>reagents</u>	<u>final concentration</u>	
TRIS	25mM	
Glycine	150mM	
Methanol	10%	optional
Sodium dodecyl sulfate (20%)	1%	optional
	pH8,5	

Washing buffer (PBS-T, 10x)

<u>reagents</u>	<u>final concentration</u>
NaCl	1,37M
Na ₂ HPO ₄	100mM
K ₂ HPO ₄	20mM
KCl	27mM
Tween-20	0,5%

Blocking buffer (TBS-T, 10x)

<u>reagents</u>	<u>final concentration</u>
NaCl	1,5M
TRIS-HCl pH7,5	100mM
Tween-20	0,5%

Solution for primary antibody

<u>reagents</u>	<u>final concentration</u>
BSA	1,25%
Washing buffer	1x PBS-T
Sodium azide 10%	1:1000

Solution for secondary antibody

<u>reagents</u>	<u>final concentration</u>
milk powder	5%
Washing buffer	1x PBS-T

3.1.14 Bacteria strains

<u>designation</u>	<u>genotype</u>
One Shot [®] TOP10 Chemically Competent <i>E. coli</i>	F ⁻ <i>mcrA</i> Δ (<i>mrr-hsdRMS-mcrBC</i>) Φ 80 <i>lacZ</i> Δ M15 Δ <i>lacX74</i> <i>recA1</i> <i>araD139</i> Δ (<i>ara leu</i>) 7697 <i>galU galK rpsL</i> (StrR) <i>endA1 nupG</i>

3.1.15 Cell lines

<u>designation</u>	<u>manufacturer</u>
293T	ATCC, Wesel/DSMZ, Braunschweig
HeLa	ATCC, Wesel/DSMZ, Braunschweig
FDCP-Mix	DSMZ, Braunschweig
32D	ATCC, Wesel/DSMZ, Braunschweig

C1498 ATCC, Wesel/DSMZ, Braunschweig

3.1.16 Mouse strains

designation

manufacturer

C57Bl/6J Jackson Laboratories, Bar Harbor, ME, USA

B6.SJL-Ptprca Pepcb/BoyJ Jackson Laboratories, Bar Harbor, ME, USA

B6.129S1-Plcb4tm1Dwu/J Jackson Laboratories, Bar Harbor, ME, USA

3.1.17 Surgical instruments

Cotton Swabs	Böttger, Bodenmais
Forceps, Moria Ultrafein	Fine Science Tools, Heidelberg
Forceps, standard anatomical	Fine Science Tools, Heidelberg
Forceps, standard surgical	Fine Science Tools, Heidelberg
Insulin Syringe 0.5ml, 27G	Becton, Dickinson and Company, Heidelberg
Scalpels	Feather Safety, Osaka
Scissors, standard surgical	Fine Science Tools, Heidelberg
Ear punch	Fine Science Tools, Heidelberg

3.2 Methods

3.2.1 Molecular biology methods

3.2.1.1 Transformation of competent *E.coli* bacteria

Bacteria exhibit distinct features which make them ideal organisms to amplify DNA. For this work, One Shot® TOP10 chemically competent *E. coli* was used. The competent state was achieved by Ca^{2+} -treatment leading to masking of the negatively charged outer bacterial membrane which consists of lipopolysaccharides and phospholipids. The negatively charged DNA molecules are attracted and can enter the cell upon heat shock which weakens the membrane structure and leads to pore formation.

Chemically competent *E.coli* were stored at -80°C and thawed on ice for 30min. 0.5-1 μg plasmid was added and incubated on ice for further 30min. Bacteria were heat shocked for 30s at 42°C . The mixture was chilled on ice for 1min. 250 μl super optimal broth (SOC)-medium was added and incubated in shaking block at 350 rpm, 37°C for 60min to help bacteria recover from the heat shock.

3.2.1.2 Cultivating transformed bacteria

If absorbed plasmids harbor a gene conferring antibiotic resistance, transformed bacteria can be enriched by antibiotic selection. The bacterial cell suspension is grown on agar plates containing the respective antibiotic and only plasmid containing bacteria will have the ability to metabolize the drug and form colonies. Single colonies can then be picked and amplified in liquid cultures. All plasmids used in this study harbored an ampicillin-resistance gene.

50-100 μl of the bacteria suspension was plated on ampicillin-containing agar plates (100 $\mu\text{g}/\text{ml}$). Plates were incubated upside down at 37°C for 12-16h. Colonies were picked by an autoclaved toothpick and transferred into 5ml LB medium containing 100 μg ampicillin/ml. Bacteria were grown for 8-14h at 37°C , 190rpm. Depending on the desired amount of DNA, liquid cultures were either used for DNA isolation (Mini Preparation) or used for inoculation of 250ml ampicillin-LB medium which was then kept on the shaking incubator for further 12-14h and finally used for DNA isolation (Maxi Preparation).

3.2.1.3 Plasmid isolation

Plasmids can be isolated by alkaline lysis. Addition of an alkaline solution containing a detergent like Sodium dodecyl sulfate (SDS) disrupts bacterial cell membranes. The chemical can enter the cell and nucleic acid is denatured. A neutralizing solution like sodium acetate allows the renaturing of plasmid DNA, but precipitates chromosomal DNA and the

detergent. The plasmid can then be purified using an ion-exchange polymer column. DNA is bound to the column, washed and finally eluted with a saline solution. Purified DNA is then pelleted by isopropanol-precipitation and centrifugation and dissolved in H₂O.

Bacterial cultures were centrifuges at 6000rpm, 10min, 4°C (2ml Mini culture) or at 4600rpm, 30min, 4°C (250ml Maxi culture) and either frozen and stored at -80°C or directly submitted to alkaline lysis according to the kit-manufacturers' protocol. The following kits were used:

- Mini Prep: GeneMATRIX Plasmid Miniprep DNA Purification Kit
- Maxi Prep: Qiagen EndoFree® Plasmid Purification Kit.

The concentrations of DNA were determined using the spectrophotometer NanoDrop® ND-1000. Purified DNA was stored at -20°C.

3.2.1.4 Restriction digestion

DNA can be digested enzymatically using restriction endonucleases. These enzymes recognize particular and unique sequences of DNA and act as molecular scissors producing DNA fragments of distinct sizes with 5'-, 3'- (sticky) or no overhang (blunt) end. Visualization of these DNA fragments using agarose gel electrophoresis and ethidium bromide as an intercalating agent determines fragment sizes which can then be compared to theoretically expected ones. Fragments can be further extracted from the gel and be used in further cloning experiments, which needs higher amount of DNA, or the gel is used to determine integrity of isolated DNA samples, lower amount of DNA.

<u>reagent</u>	<u>amount</u>
DNA	1µg
specific Enzyme buffer (10x)	2µl
BSA (10x), if necessary	2µl
Enzyme	5U
H2O	up to 20µl

The restriction digestion was Incubated for at least 1hr in the enzyme-specific reaction temperature in the heating block.

In order to have sufficient material for further cloning steps larger amount of DNA was used for restriction digestion.

<u>reagent</u>	<u>amount</u>
DNA	5 μ g
specific Enzyme buffer (10x)	5 μ l
BSA (10x), if necessary	5 μ l
Enzyme	20U
H2O	up to 50 μ l

Incubation of the mixture was performed for at least 3hrs at the optimum enzyme working temperature.

3.2.1.5 Agarose gel electrophoresis

DNA molecules are negatively charged due to their phosphate groups. They can thereby move in an electrical field. The velocity of the movement depends on DNA size (correlating with the molecules total charge and mass) and on the constitution of the medium. In agarose gel electrophoresis DNA is loaded onto an agarose gel, an electrical field is applied and the molecules start to move through the electrical field. Agarose is a natural polysaccharide found in red algae. As a gel, agarose medium is porous. Its pore size depends on the concentration of agarose (0.8 - 2%). Smaller molecules migrate faster than bigger molecules because they move much easier through the pore mesh. Thus, DNA fragments can be separated according to their size. A DNA intercalating dye visualizes the fragments and a marker of known sizes determines the size of these fragments. Fragments can be extracted from the gel and be used in further cloning experiments or the gel is simply used to determine integrity of isolated DNA samples like in this work.

Depending on the intended concentration, 0.8 – 2g of purified agarose were dissolved in 100ml 1x TBE buffer by heating in a microwave. After a short cool down, 1 drops of ethidiumbromide (1mg/ml) was added and the mixture was decanted carefully into a gel chamber. A comb was inserted for the formation of wells and the gel was solidified and cooled for 20min at RT. The chamber containing the gel was placed into the electrophoresis tank with 1x TBE buffer. 8 μ l of the digested DNA samples were mixed with 2 μ l loading buffer

5x Blue Run and loaded into the gel chambers. Gel electrophoresis was performed for 45min at 150V. Fragment sizes were estimated by visualization under UV light and comparing DNA bands to a DNA ladder loading control.

3.2.1.6 Ligation of linear DNA fragments

Inserting the desired DNA fragment into a vector was performed by using T4 DNA ligase. This enzyme catalyzes the formation of a phosphodiester bond between the terminal 5'-phosphate and 3'-hydroxyl groups of duplex DNA. T4 DNA ligase efficiently joins blunt and sticky ends and repairs single-stranded nicks in duplex DNA. The DNA fragment used for ligation should be at least 3-fold higher than the vector. After mixing the appropriate amounts of insert and vector, 2µl of reaction buffer (containing ATP), and 1µl of the ligase was added and the final volume was adjusted to 20µl. The reaction was started in a thermocycler. After incubation for 16 hrs at 16 ° C the enzyme was inactivated at 65°C for 10min. The ligation mixture was either stored at -20°C or directly used for bacteria transformation.

3.2.1.7 RNA isolation from cells

Depending on the number of cells, the isolation of mRNA was performed using, RNeasy Mini or Micro Kit (Qiagen) according to the manufacturer's instructions. Isolated RNA was then stored at -80 ° C.

RNA from murine hematopoietic stem cell populations ($<10^6$ cells) was isolated with PicoPure kit (Life Technologies). Cells were directly sorted into 50µl of extraction buffer provided by the kit. Isolation was done according to the manufacturer's protocol and RNA was eluted with 12-15µl of H₂O.

3.2.1.8 Photometric determination of DNA and RNA concentration

Determining the concentration of DNA was performed using the Nanodrop spectrophotometer. The ratio of absorbance at 260 nm and 280 nm was used to assess the purity of DNA and RNA. A ratio of about 1.8 is generally accepted for pure DNA and a ratio of 2.0 is accepted as pure for RNA. If the ratio is significantly lower in either case, it may indicate the presence of protein, phenol or other contaminants that absorb strongly at or near 280 nm.

3.2.1.9 Quantitative polymerase chain reaction

Quantitative polymerase chain reaction (qPCR) amplifies and simultaneously quantifies a targeted DNA molecule. Using specific primers and the LightCycler[®] 480, a sensitive detection and quantification of the desired DNA sequence can be performed. The PCR product size should not exceed 500 bp. Using SYBR[®] Green, a dye which intercalates between double-stranded DNA, the amount of product formed by a spectrometric detection can be accurately determined after each cycle. The quantity of target DNA or cDNA is either an absolute number of copies or a relative amount when normalized to DNA input or additional normalizing genes.

The qPCR was performed according to the following protocol:

<u>reagent</u>	<u>amount</u>
cDNA (1:10)	4 μ l
Primer (forward)	1 μ l
Primer (reverse)	1 μ l
SYBR [®] Green PCR Master Mix	10 μ l
H2O	4 μ l

3.2.1.10 Polymerase chain reaction (PCR)

PCR is a method based on using the ability of DNA polymerase to synthesize new strand of DNA complementary to the desired template strand. Because DNA polymerase can add a nucleotide only onto a preexisting 3'-OH group, it needs a primer to which it can add the first nucleotide. This requirement makes it possible to delineate a specific region of template sequence that should be amplified. At the end of the PCR reaction, the specific sequence will be accumulated in billions of copies (PCR amplicons).

To perform genotyping PCR on transgenic mice, 100-500ng DNA isolated from mouse tail (DNeasy Blood & Tissue Kit) was used as PCR template. PCR was performed according to the following protocol:

<u>reagents</u>	<u>amount</u>
DNA	100ng
Primer Mix (10 μ M each)	1.5 μ l

Red Taq master mix(2x)	12.5µl
H2O	up to 25µl

Before adding the Taq polymerase, the mixture was well mixed and then stored on ice until the start of the reaction in the thermocycler. The target sequence was amplified using the following program:

<u>reaction step</u>	<u>temperature</u>	<u>time</u>
denaturation	95°C	2 min
denaturation	95°C	45 s
hybridization	53° - 60°C	45 s
elongation	72°C	30 s
(30 cycles)		
final elongation	72°C	5 -7 min

3.2.1.11 Sequencing DNA fragment after cloning

Plasmids had to be tested for their integrity before amplification and usage in further experiments. DNA is replicated *in vitro* using a primer, a DNA polymerase, a mix of normal deoxynucleosidetriphosphates (dNTP) and only one type of radio-labeled dideoxynucleotides (ddNTP) lacking a 3'-OH group. If ddNTPs are incorporated into the newly synthesized DNA, replication is terminated since a new phosphodiester bond cannot be formed. These fragments are subsequently separated via electrophoresis and the sequence can directly be read from the gel picture.

2.5µg of the DNA sample were diluted in 30µl H₂O in a 1.5ml reaction tube. Sequencing was performed at GATC Biotech in Konstanz. Primers were chosen based on the following criteria: 17-28 bases in length, G/C-content of about 50-60%, location about 50bp up- or downstream of region to be sequenced, melting temperature preferentially 52 – 58°C and a G- or C-3'end was favored. If not available, primers were also synthesized by GATC Biotech.

3.2.1.12 cDNA synthesis

A rapid method to analyze gene expression in cells is the amplification of RNA by PCR. However, since RNA cannot serve as a template for Taq polymerase, it must first be transcribed to DNA by reverse transcriptases (RT).

Up to 5µg of RNA was used as template for cDNA synthesis. Oligo(dT) nucleotides which consist 12-18 dTs and bind specifically to the poly-A tail of eukaryotic RNA were used as primers. cDNA was synthesized using SuperScript® III kit according to the manufacturer's protocol.

3.2.1.13 Global gene expression profiling

RNA was harvested according to section 3.2.1.7. RNA integrity was determined by the Aligent 2100 Bioanalyzer at the DKFZ Genomics and Proteomics Core Facility. RIN values ranging from 9.5 to 10 were used for gene expression profile.

500 ng of RNA was submitted to the DKFZ Genomics and Proteomics core facility for the expression profile on MouseWG-6 v1.1 Expression BeadChip (Illumina™).

3.2.1.14 Western blot

Cells were lysed and resuspended in RIPA buffer at 10^6 cells/100µl. Cell lysate was centrifuged at 13,000 rpm, 4°C, 10 min. The supernatant containing protein was used for Western blot. 15-20 µg of protein lysate was mixed with 4x loading buffer and heated at 95°C for 6 min. The samples were centrifuged at 12000 rpm for 5 min and placed on ice. The first step was polyacrylamide gel electrophoresis. The stacking and resolving gels were made according to the protocol. APS and TEMED were added right before casting the gel. First the resolving gel was poured and ethanol was added gently on top of the gel. The gel was let to solidify and then ethanol was removed. The stacking gel was poured and the comb was immediately inserted. The gel was let to solidify for 20-30 min. The comb was removed the wells were washed with running buffer to remove any piece of remaining gel. Electrophoresis tank was filled with running buffer. Protein lysate samples and 10µl of protein ladder were loaded in wells. Gel electrophoresis was run with 120V for 90min. The gel was blotted with PVDF membrane with 25V for 1hr. the membrane was incubated in blocking buffer for 1hr at RT and then washed with washing buffer 3x, 5min each. The membrane was incubated with primary antibody overnight at 4°C on a roller followed by 3x

washing, 5min each. Secondary antibody was diluted in antibody buffer and incubated with the membrane for 1hr at RT and then washed 3X, 5min each. Post washing, the protein bands were detected by developing the membrane using 1:1 volume from A and B ECL solution and ChemiDoc XRS with Image Lab software.

3.2.2 Cell biology methods

3.2.2.1 Freezing and thawing cells

In order to have back up stocks of primary cells and cell lines, they can be viably frozen and kept at very low temperature. Cells should be frozen in complete growth medium in the presence of dimethylsulfoxide (DMSO). DMSO is used as a cryoprotective agent; it allows slower cooling rates by reducing ice crystal formation which can damage the cells.

To freeze the cells respective cell suspensions were centrifuged or adherent cells were harvested (3.2.2.3.1) and resolved in 750 μ l growth-medium. The cell suspension was transferred into a cryotube and mixed with 750 μ l of the respective freezing-medium diluting DMSO to 7.5%. Cells were placed in a freezing box filled with isopropanol providing a cooling rate of 1 $^{\circ}$ C per min which is required for successful cryopreservation of cells. The freezing box was then kept at -80 $^{\circ}$ C for 24-48hrs. Frozen cells were transferred to the vapor phase of liquid nitrogen for long-term storage.

For thawing the cells, the respective cryotubes were placed in a 37 $^{\circ}$ C water bath until only a small ice piece was left. Cells were transferred in a 50ml falcon tube and 1ml growth-medium or in case of adherent cells conditioned thawing medium was added gently to thaw the remaining cells completely. In order to further dilute the DMSO from the freezing medium, another 5ml of thawing medium was gradually added during 1 minute. Finally, 20ml medium was added to the cell suspension which was then centrifuged. Cells were then cultured in humidified incubator at 37 $^{\circ}$ C, 5% CO₂.

3.2.2.2 Determination the number of cells

The dye exclusion test is used to determine the number of viable cells present in a cell suspension. It is based on the principle that living cells possess intact cell membranes, therefore certain dyes, such as trypan blue cannot enter the cell.

Suspended cells were mixed with trypan blue in specific dilution (1:2 – 1:100). A cover slip was placed on a Neubauer chamber and 10 μ l of the cell mixture was loaded into the

counting chamber. Light, viable cells were counted in 4 quadrants under a light microscope. The number of cells/ml in cells suspension was calculated according to this formula:

$$\frac{\text{number of cells}}{\text{ml}} = \frac{\text{number of cells}}{4 \text{ quadrants}} \times \text{dilution factor} \times 10^4$$

1 quadrant holds 0.1 μ l (= 10⁻⁴ml). The average number of cells per quadrant multiplied with the dilution factor gives the average number of cells per 0.1 μ l. In order to calculate the average number of cells per ml, the average number of cells per 0.1 μ l has to be multiplied by 10⁴.

3.2.2.3 Cell line and hematopoietic stem cell culture

3.2.2.3.1 Adherent cell culture

In this work adherent HEK-293T (293T) and HeLa cells were used. 293T cells are human embryonic kidney cells which stably express the SV40 T-antigen. HeLa cells are adherent cells derived from human cervical carcinoma.

Culturing of adherent cell lines was performed in cell culture flasks in an incubator at humidified atmosphere (37°C, 5% CO₂). At 80% of confluency, medium was aspirated and the cell layer was washed with 10ml PBS. 5ml 0.025% Trypsin was added and cells were incubated at 37°C for 5min. The reaction was stopped by the addition of 20ml stopping medium. The cell suspension was centrifuged (1200rpm, 5min, 4°C) and resuspended in 5ml growth medium. 1/100 – 1/10 of the cell suspension was transferred into a new flask containing 15ml growth medium and culturing was continued at 37°C.

3.2.2.3.2 Suspension cell culture

The suspension cells used in this work comprised C1498, FDCP-Mix, WEHI-3B and 32D. C1498 are murine lymphoblasts suspension cells derived from C57BL/6J mouse with acute myeloid leukemia. FDCP-Mix cells are murine, multipotent, interleukin-3 (IL3)-dependent progenitor cells established from long-term bone marrow cultures. 32D cells are immortalized myeloblast-like cells originally derived from long-term cultures of murine bone marrow. WEHI-3B is a macrophage-like myelomonocytic leukemia cell line established from

inbred BALB/c mice. Due to the integration of a retroviral genome in the vicinity of the IL3 gene this cell line constitutively produces IL3.

Suspension cell cultures can be maintained by adding fresh medium or replacement of medium. Alternatively, cell cultures can be established by centrifugation with subsequent resuspension in fresh medium at 1 to 3 x 10⁵ viable cells/mL. A saturation density of 1 to 2 X 10⁶ was obtainable. Depending on cell density fresh medium can be added to cell suspension every 2 to 3 days.

3.2.2.3.3 Generating IL3 containing supernatant from WEHI cells

WEHI cells were counted and resuspended to a final concentration of 1x10⁶ cells/ml. Cells were incubated at 37°C, 5% CO₂ for 24hrs. The cells were spun down at 1200 rpm, 4°C, 5min and the supernatant was collected and filter sterilized (0.22µm). Aliquotes of WEHI supernatant were stored and stored at -20°C. It was later used as IL3 source in 32D cells conditioned medium.

3.2.2.3.4 Hematopoietic stem and progenitor cell culture

Hematopoietic stem and progenitor cells were sorted from the mouse BM (3.2.2.8, single or bulk cells (up to 10⁴/ well) were cultured in 96-well plate in 100 µl StemSpan™ medium containing 1% pen-strep, 100ng/ml rmSCF, 100ng/ml Flt3ligand, 100ng/ml rhTPO, and 20ng/ml IL3. Cell medium was changed every third day or when the cell number reached 10⁵.

3.2.2.3.4.1 LSK cell transduction

10⁴ LSK cells were suspended in 100 µl StemSpan™ medium containing 1% pen-strep, 100ng/ml rmSCF, 100ng/ml Flt3ligand, 100ng/ml rhTPO, 20ng/ml IL3 and Protamine sulfate 8µg/ml in one well of a 96-well plate. Viral particles were added to cells with MOI 60 and the plate was spun at 1000 rpm, 1hr, 4°C. The cells were incubated at 37°C, 5% CO₂.

3.2.2.4 Colony forming assay (CFC)

Hematopoietic stem and progenitor cells isolated from murine BM cells can be differentiated *in vitro* using semi-solid MethoCult™ medium containing cytokines. Depending on the cytokines, cells can be differentiated to various blood lineages and form colonies. The type and number of colonies can be determined by light microscope.

Three types of MethoCult™ semi-solid medium were used in this work. MethoCult™ M3630, which is optimized for the detection and quantification of mouse pre-B progenitor cells in

BM. MethoCult™ GF M3434 which has been formulated to support optimal growth of erythroid progenitors (BFU-E), granulocyte-macrophage progenitors (CFU-GM, CFU-M, CFU-G) and multi-potential granulocyte, erythroid, macrophage, megakaryocyte progenitors (CFU-GEMM). MethoCult™ H4230 does not contain any cytokine and provides the basic medium for CFC assay.

The MethoCult™ aliquots were thawed overnight at 4 °C or a few hours before use at RT. After a brief mixing of the medium, the aliquots were let stand to lose the air bubbles. To perform pre-B CFC, 2×10^5 murine total BM cells were used per dish. Cells were suspended at 3x concentration in 300µl of IMDM+2% FBS and then added to 3ml MethoCult™ 3630. The mixture was briefly mixed and let stand for 2-3 min till the air bubbles disappeared. 2x 1.1ml of the cells and semi-solid medium mixture were distributed to 2x 3.5cm dishes in a 5ml syringe with a blunt end needle. To prevent the semi-solid medium drying out, 3ml H₂O was added to another 3.5cm dish without a lid. The dishes containing cell suspension and H₂O were placed in a 10cm dish and incubated at 37°C and 5% CO₂. Colonies were observed and counted under light microscope after 10-12 days.

Myeloid CFC was performed with 2.5×10^4 total BM cells and MethoCult™ GF M3434 following the same procedure as pre-B CFC.

Similar protocol was followed to perform IL3 independent CFC assay using MethoCult™ H4230 and 10^5 32D cells/dish. The colonies were counted after 10-12 days. The semi-solid medium was once washed with 25ml cold PBS, and washed with 10ml PBS centrifuged at 1200rpm, 4°C, 5min

3.2.2.5 Lentiviral vector production

In order to stably express transgenes in primary hematopoietic stem cells and cell lines, lentiviral vectors (LV) encoding the respective expression cassette were produced. These vectors were used as a gene shuttle delivering the genetic material of interest into the target cells.

6×10^6 293T-cells were seeded into 15cm-dishes. Cells were grown in 15ml IMDM growth medium for 24hrs at 37°C and 5% CO₂. Medium was replaced with 13ml fresh growth medium. Polyethyleneimine (PEI, 179.25µg/500µl blank IMDM) was used as a transfection agent. PEI packs the negatively charged DNA into positively charged particles facilitating cellular uptake by endocytosis. 500µl plasmid-mix was prepared with IMDM. The solution was filter sterilized (0.22µm pore diameter) and mixed with an equal volume of PEI/IMDM. After 20min incubation at RT, the mixture was added into the cells and the dish was gently shaken in order to distribute the DNA among the cells.

<u>plasmid designation</u>	<u>function</u>	<u>amount per dish (μg)</u>
101	gag-pol	12.5
102	reverse transcriptase	6.25
103	VSV-G	9
transfer plasmid	encoding the transgenes	32

Medium was refreshed 16h after incubation (15ml). After further 48hours of incubation, medium containing lentiviral particles was harvested and filter sterilized (stericup vacuum filtration system). It was divided equally to ultracentrifugation tubes and lentiviral particles were pelleted at 20.000rpm, 2h, 20°C. The supernatant was discarded and tubes were put upside down on a piece of tissue. After 10min incubation at RT, residual medium was removed by a sterile cotton-tip. 50 μl PBS was added to each tube, tubes were incubated at RT for 30min. The pellet was pipetted up and down 25 times with the PBS. The supernatants from all centrifuge tubes was collected in a 1.5ml-reaction tube and rotated at RT for further 20min. The concentrated viral supernatant was aliquotted and stored at -80°C.

3.2.2.6 Determination of Functional Lentiviral Titers

The efficiency of lentiviral infection depends on the amount of functional and infectious viral particles present at the time of infection. This amount might differ depending on production and harvest conditions. Therefore, the amount of infectious particles in a given solution (functional titer) should be determined. The titer can be quantified using a serial limiting dilution transduction assay under standardized conditions.

Adherent HeLa cells were grown in liquid medium according to manufacturers' protocol in an incubator at humidified atmosphere (37°C, 5% CO₂). At about 70-80% of confluency, medium was aspirated and the cell layer was washed with 10ml PBS. 5ml 0.05% Trypsin was added and cells were incubated at 37°C for 5min. The reaction was stopped by the addition of 20ml stopping medium (PBS+10% FBS). The cell suspension was centrifuged with 1200rpm, 5min, at 4°C and either 5x10⁴ (24hrs prior to transduction) or 10⁵ cells (4hrs prior to transduction) were seeded in a 6-well plate.

Growth medium was replaced by 500 μl conditioned growth medium containing 16 μg Polybrene/ ml. Concentrated LV was diluted in a 24-well-plate, 2 μl of virus stock was added into 1ml growth medium and mixed properly. Five serial 1:10-dilutions were prepared by transferring 100 μl of the previous well into 900 μl fresh growth medium (dilution factor 10⁻³

to 10^{-7}). Subsequently, 500 μ l of each viral dilution was added into the wells containing the HeLa cells in 500 μ l conditioned medium containing Polybrene. Transduction was allowed for 72hrs at 37°C, 5% CO₂. Then, cells were trypsinized and prepared for FACS analysis (3.2.2.7.1).

The functional titer of the lentiviral stock solution was calculated with the following formula from the sample when 1% to 25% (y %) of all living cells were transduced:

$$\text{LV concentration (TU/ml)} \\ = \text{number of cells (time of infection)} \times \text{dilution factor} \times \frac{\text{y \% GFP}^+ \text{ cells}}{100}$$

3.2.2.7 Flow Cytometry

Flow Cytometry (fluorescence activated cell sorting, FACS) is a technology that measures cells or particles in liquid suspension. A flow cytometer, like the BD™ LSR II, is composed of three subsystems: fluidics, optics and electronics. Samples are kept in suspension and loaded into the cytometer's fluidic system which brings the sample to a point of interrogation where it is detected by a laser beam of a certain wavelength. Because of the samples fluorescent properties or previous treatment of the sample with fluorescent dyes, the fluorescent components are excited by the laser light and emit fluorescent light which is subsequently modified by mirrors, a photo multiplier tube system and filters. The modified light is then detected by collection components of the optics subsystem. The electronic subsystem converts the light signal into an electronic one and digitizes the data which is then displayed by a computer using BD FACSDiva™ software. The experimenter can measure fluorescence intensity, counts, the relative complexity and the relative size of particles or cells.

In addition to the BD™ LSR II, the BD FACS Aria™ II Cell Sorter was also used in this work. A cell sorter has a similar function to a common cytometer. The optics system contains additional features for sorting the samples according to different characteristics of choice. The sample drop is charged depending on the emitted fluorescence signal after excitation and deflected in a subsequent electrical field. Tubes at respective positions collect the sorted samples.

For flow cytometric measurement or sorting of viable cells, specific fluorescent dyes can be used to exclude dead cells. Membrane integrity of viable cells guarantees that these dyes

cannot invade a cell. Cell death results in loss of membrane integrity. Thus, the stain enters the dead cells which can be identified by the FACS machine.

To exclude dead cells either Fluorogold or Propidium iodide (PI) was used in this work. PI is excited at 488 nm and emits at a maximum wavelength of 617 nm. Fluorogold was only used in combination with dyes occupying the red laser (635nm).

3.2.2.7.1 Preparation of cells for flow cytometry

Cells were centrifuged and the pellet was washed once with 1ml PBS. Subsequently, cells were resuspended in 500µl PI and spun down. Samples were resuspended in 300µl PBS and stored on ice until FACS analysis.

3.2.2.8 Purification of hematopoietic stem cell populations from bone marrow

In order to isolate various hematopoietic stem and progenitor cell populations, mouse tibia, femur and ilium bones were harvested and flushed with 5ml syringe (0.6mm/23G needle) in 3ml HF solution in a round bottom tube. BM cells were counted at 1:100 ratio with Türk solution according to previously described protocol (3.2.2.2). The blood lineages were depleted from the whole BM cells using EasySep™ kit. Lineage depleted cells were counted with Trypan blue with 1:10 ratio and incubated with blocking antibody solution at concentration of 10^7 cells/100µl for 30min at 4°C. The mixture of cells and blocking solution was then incubated with 10x pre-diluted respective antibody cocktail for 30min at 4°C. Post antibody staining, the cell suspension was washed with HF containing Fluorogold, filtered with 40µm cell strainer and resuspended in HF. Cells were then sorted using FACS Aria II™ cell sorter.

3.2.2.9 Cell Titer Glo assay

C1498 cells were seeded at 5×10^4 /ml concentration. On days 1 to 4 the ATP activity level and cell viability was measured by CellTiter-Glo™ assay. 15µl of cell suspension was pipetted in a 384-well plate for 30min at RT and mixed with 15µl CellTiter-Glo™ reagent which was previously thawed and reached RT. The cell mixture was shaken for 2min and incubated at RT for 10min to stabilize luminescence signal. Post incubation, luminescence level was measured at luminometer.

3.2.2.10 Staining peripheral blood cells

To stain the blood cell with lymphoid and myeloid antibody cocktails, heparin- blood was transferred from micro cuvette to FACS tubes. Red blood cells were lysed with lysis buffer 2 x 5 min and washed with HF. Post lysis, the cells were incubated in C16/32 blocking antibody

solution at 4°C for 30min, and stained with lymphoid, myeloid and isotype control cocktails at 4°C, for 30min. Cells were then washed with HF and stained with Fluorogold. Stained blood cells were resuspended in 200µl HF and measured with BD™ LSR II FACS.

3.2.2.11 NF-κB activity assay

5x10⁵ 293T cells were seeded in a 6-well plate and incubated for 4-5 hrs at 37°C, 5% CO₂ till the cells attach. Cells were transfected with combinations of LRRC33, TLR4, and GFP (5µg each) and NF-κB-Firefly (5µg) and LacZ (2µg) plasmids according to the same protocol described in section 3.2.2.5. 2 days post transfection cells were stimulated with Lipopolysaccharides (LPS) for 24hrs and NF-κB activity level was measured using Dual-Light™ kit according to the manufacturer's protocol.

3.2.3 Mouse experiments

The mouse strains used for the animal experiments were B6.129S1-*Plcb4*^{tm1Dwu}/J, B6.SJL-*Ptprc*^a *Pepc*^b/BoyJ, and C57Bl/6J mice (from the Jackson Laboratory). All animal experiments were approved by the "Regierungspräsidium Karlsruhe" (Application number: G-160/08 and G-175/13).

3.2.3.1 Blood collection from mice

To collect the mouse blood for further analysis, the mouse was fixed in a 50ml tube with an open end. Its leg was shaved and the peripheral blood was collected by puncturing the saphenous vein with a fine needle (0.45 mm/26G) in an EDTA or heparin-coated micro cuvette.

3.2.3.2 Determination of the blood cells count

The blood cell count was performed by Advia Analyzer at Analysis Center of the Heidelberg University Hospital. To prepare the blood to be measured, 60µl of EDTA-blood was mixed with 240µl of 0.9% NaCl solution (1:5).

3.2.3.3 Intravenous bone marrow transplantation

Four to six hours before transplantation the mice were lethally irradiated with 950 cGy. To perform intravenous injection the mouse tail was heated with a red lamp. The mouse was fixed in a restrainer and injection of murine bone marrow cells into the tail vein (in maximum 200µl PBS) was performed using a needle (0.45 mm/26G). All transplanted mice were fed with 0.5 ml of Baytril per liter in drinking water.

4 Results

4.1 Establishment of a selection strategy for identification of novel candidate stem cell regulatory genes

Major progress in the efficiency of gene transfer using integrating vectors has allowed achieving a number of genetically corrected hematopoietic cell clones *in vivo* [105, 106, 109, 122, 123].

The integration of gammaretroviral vectors has an overall preference for the transcription start site (TSS) of genes (\pm 5-10 kb), and a preference for the transcribed region of genes at the time of infection [99]. The flanking genomic genes can therefore get activated by the strong enhancer and promoter sequences within the long terminal repeats (LTRs) in gammaretroviral vectors. Depending on the activated gene this can cause clonal selection and dominance up to malignant transformation. Analyzing the integration sites (IS) repertoire, we detected genetic loci which are highly overrepresented, indicating for a selective advantage of clones carrying these common vector integration sites (CIS) *in vivo*. We therefore hypothesized that significantly enriched CIS in genetically modified blood cells point to a regulatory function of nearby genes in post-embryonic hematopoiesis.

To select hematopoietic candidate regulatory genes we systematically analyzed the IS repertoire of gene corrected peripheral blood (PB) and bone marrow (BM) cells within a cohort of 10 gene therapy patients with Wiskott-Aldrich-Syndrome (WAS) using linear amplification–mediated PCR (LAM-PCR) technique. After exclusion of genes based on stringent criteria, which will be described in details, we selected 12 novel candidate genes with hematopoietic regulatory potential. To investigate the long term clonal activity in clinical gene therapy patients every individual clone harboring IS close to the selected candidate genes was followed up in a 4-year time period. The expression of each gene in hematopoietic stem (HSC) and progenitor cell (HPC) populations was assessed using whole transcriptome datasets of well-defined hematopoietic populations. Finally, after intensive literature research the most novel regulatory candidate genes in hematopoiesis were selected (Figure 9).

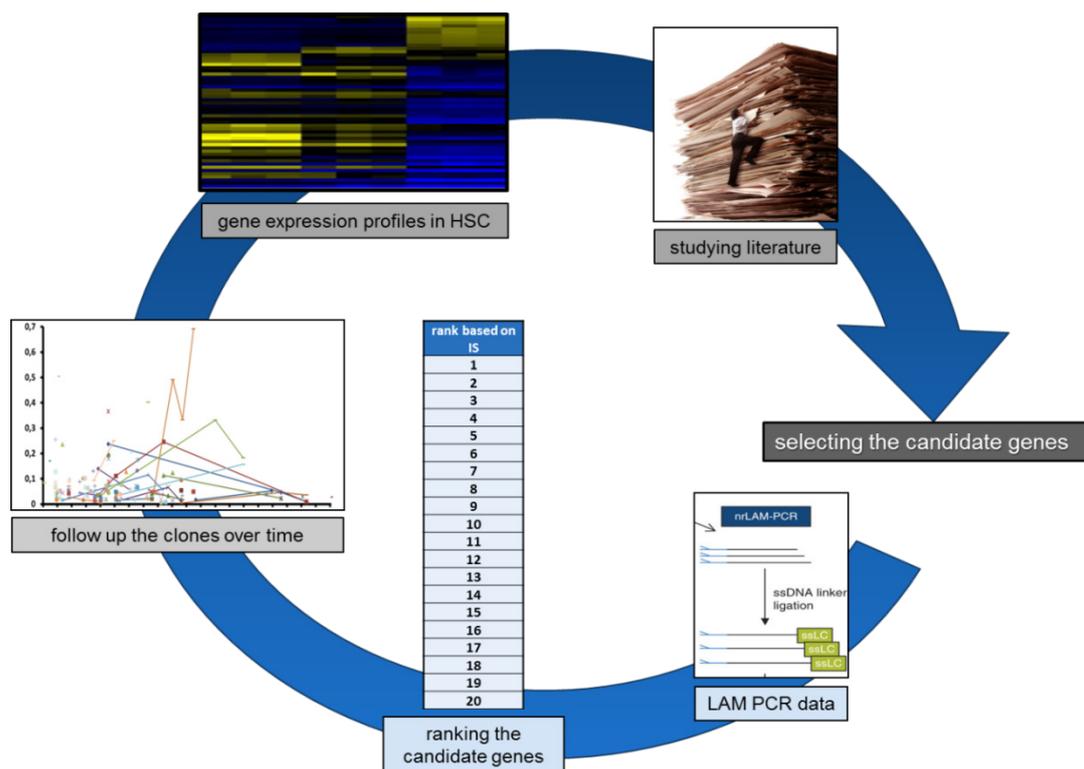


Figure 9: An overview of sequential steps of selecting candidate stem cell regulatory genes. The CIS repertoire of gene therapy patients was studied to select the top candidates of hematopoietic regulatory genes. The hematopoiesis contribution of HSC clones which carry integrations near candidate genes was investigated up to 4 years. The expression of candidate genes in different HSC populations was assessed and after studying the literature novel genes with unknown functions in hematopoiesis were selected for further analysis.

Using highly sensitive LAM-PCR combined with high-throughput sequencing, a total of 12.887 unique IS in vicinity of 3.268 genes were identified. Since in this study we used the combined dataset from 10 WAS gene therapy patients we set the 10 IS cutoff in vicinity of each gene to increase the probability of selecting candidate genes from all 10 WAS patients. Therefore, only genes with at least 10 different IS within a 200 kb window from their TSS were considered (n=588).

It has been demonstrated that the enhancer or promoter elements in LTR region of integrated vectors may trans-activate host genes even if they are quite distant (~ 300kb) [124]. In our selection procedure we narrowed this window to 50 kb to only select genes which were more likely to be affected by the RV integration. Thus, we chose genes with at least 10 IS within a 50 kb window around their TSS (n=424).

In order to establish the selection strategy with the highest specificity, gene clusters harboring the same IS in their proximity were not included in the analyses. To detect these gene-rich areas, UCSC genome browser (<http://genome.ucsc.edu>) was used. This browser visualized the precise location of every IS in human genome and demonstrated the genome-

wide distribution of IS clusters and their neighboring genes along each respective chromosome (Figure 10). After analyzing all IS clusters and their nearby genes, only single genes were selected (n=32).

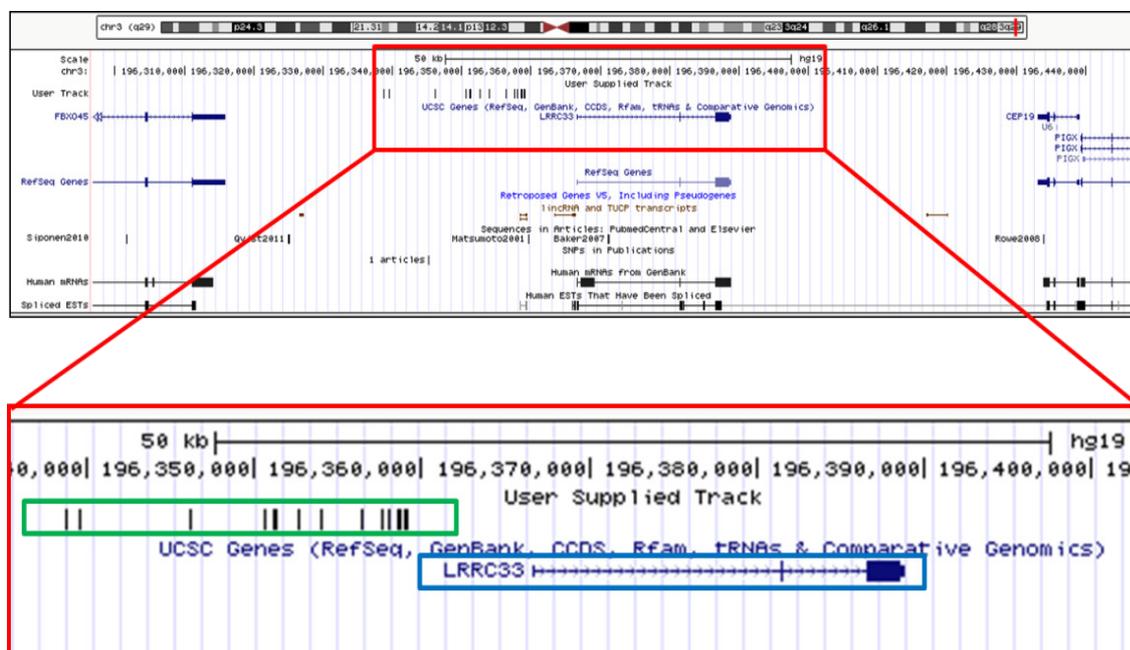


Figure 10: The position of the IS and their flanking genes were observed along the chromosome. Using UCSC genome browser the IS cluster (green box) and the genes in their vicinity (blue box) were localized.

Taken all together, genes excluded in this selection procedure were either genes with less than 10 IS in vicinity, or genes harboring at least 10 IS but in a larger 50 kb-window from their TSS. To increase the specificity of our selection strategy, clusters of two or more genes flanking the same IS were excluded. Ultimately, only single genes neighboring 10 or more IS in a 50 kb window from their TSS were selected for further functional analysis (Figure 11).

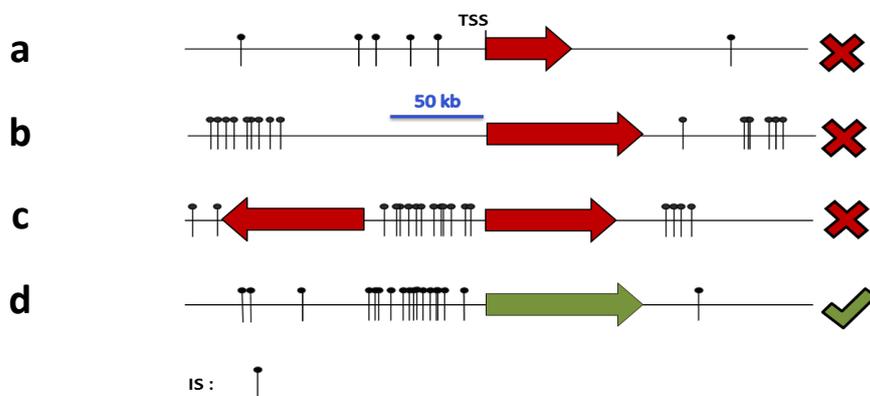


Figure 11: Schematic view of genes in vicinity of IS excluded or included as candidate hematopoietic regulatory genes. Genes as in red were not selected as candidate genes because of a low number of IS in their vicinity (a), harboring IS in a window larger than 50 kb from their TSS (b) or being located in a gene-rich area (c).

The genes as shown in green were selected as candidate genes since they were single genes with at least 10 IS in a 50 kb window from their TSS (d). IS: integration site; TSS: transcription start site.

Interestingly, we observed that 20 out of 32 final candidate genes were already known to be involved in hematopoiesis regulation which strongly validates our selection strategy. As an example, the 4 highest-ranked candidate genes EVI1, CCND2, LMO2 and MDS1 are well known as hematopoietic regulators and key factors in malignant transformation upon activation.

The hematopoietic candidate regulatory gene selection procedure is summarized in Figure 12.

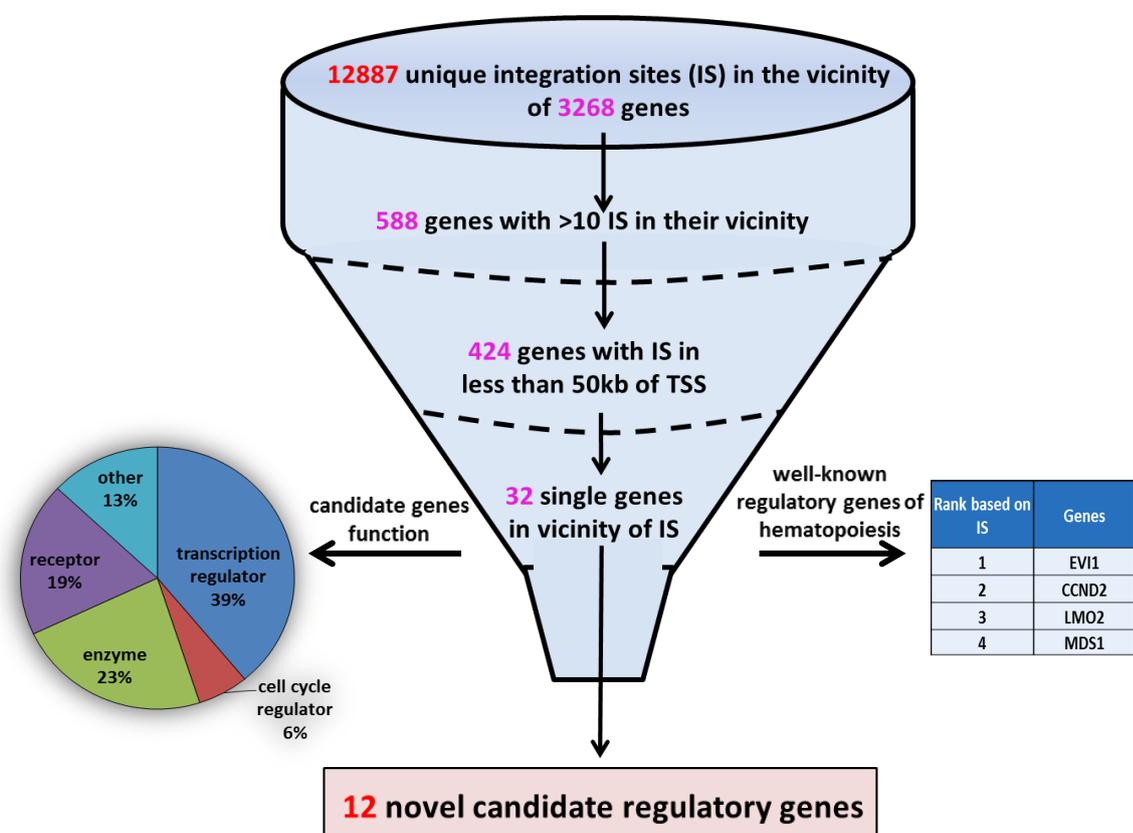


Figure 12: Selecting candidate hematopoietic regulatory genes by analyzing the IS repertoire. The primary list of all the genes neighboring IS was first refined for the genes harboring the highest number of IS (10 or more) in vicinity and then the genes with more than 10 IS within a 50 kb window. In the next step, only single genes which were not close to other genes were considered as potential candidates. The 32 selected genes had various functions as transcription regulators, receptors and enzymes. 68% of these genes are well-known hematopoietic regulators, such as EVI1, CCND2, LMO2 and MDS1. Finally, 12 novel regulatory genes were selected for further functional analyses.

Detecting long term durable candidate-gene-associated clones might point to a regulatory role for the selected candidate genes in hematopoiesis. Therefore, the IS repertoire of PB

genomic DNA from 10 WAS gene therapy patients was examined 102 times within 4 years post bone marrow transplantation (BMT). Analyzing the over-time clonal frequency of each individual 32 candidate genes revealed that these clones were detectable multiple times, 15 (Notch2) to 93 (MDS1), in a total of 102 time points (Figure 13). This long term durability can strengthen the possibility that these candidate genes play a role in hematopoiesis regulation.

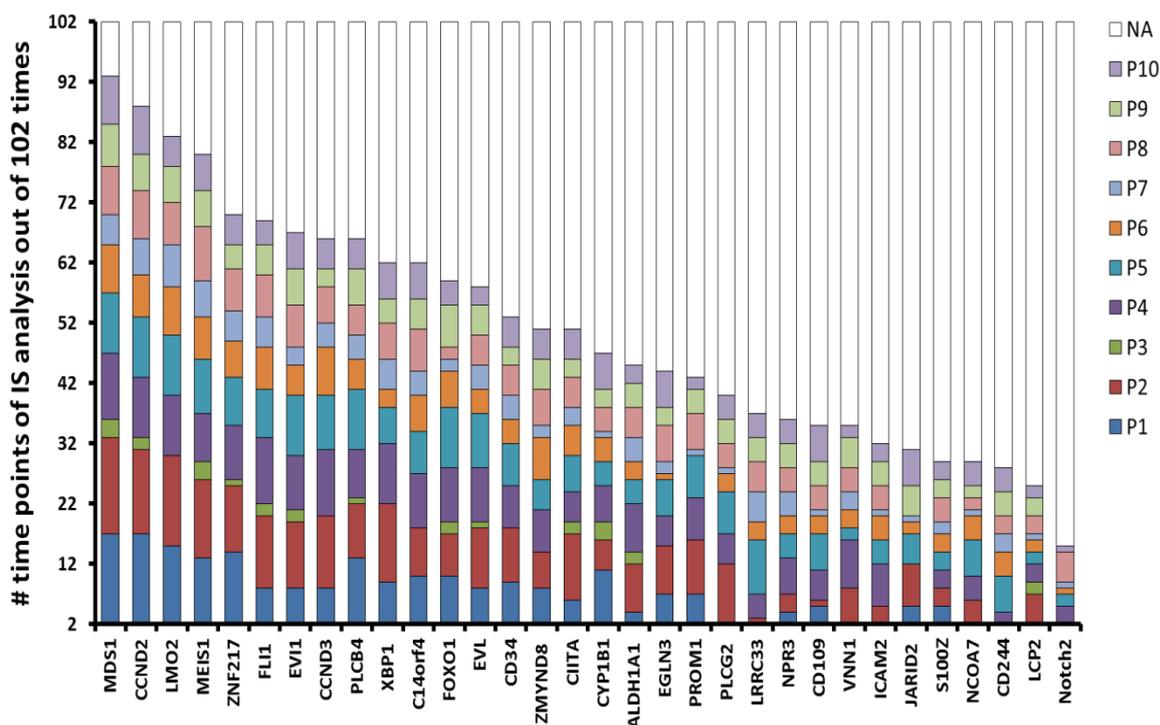


Figure 13: The distribution of 32 highly-ranked candidate genes in individual WAS patients in four years. Candidate genes were detectable at different time points that each patient was analyzed, demonstrating clonal activity of HSC clones carrying IS near candidate genes. P: gene therapy patient.

In order to investigate the biological role of selected candidate genes in hematopoiesis, five out of 12 novel candidate genes were selected for further functional experiments. According to our selection criteria, these genes were among the highest-ranked genes. They belonged to various protein families with different functions. The genes selected for further validation and functional analyses were ZNF217, LRRC33, PLCB4, EVL, and IRF2BPL (Table 2).

Table 2: The 5 candidate hematopoietic regulatory genes selected for further *in vitro* and *in vivo* analyses

candidate genes	full name
ZNF217	Zinc Finger Protein
LRR33	Leucine Rich Repeat33
PLCB4	Phospho Lipase β 4
EVL	Ena-Vasp Like
IRF2BPL	Interferon regulatory factor 2 binding protein-like

4.1.1 Assessing the contribution of candidate-gene-associated clones in hematopoiesis

To evaluate the contribution of selected candidate genes in post-transplant hematopoiesis, each individual candidate gene-associated clone marked with a specific IS was monitored in 4 years after gene therapy. Specific clones associated to ZNF217, EVL, and PLCB4 were detected at multiple time points between days 90 to 1655 post-transplant. IRF2BPL and LRR33-associated clones were detectable up to day 1289 and 950 respectively (Figure 14).

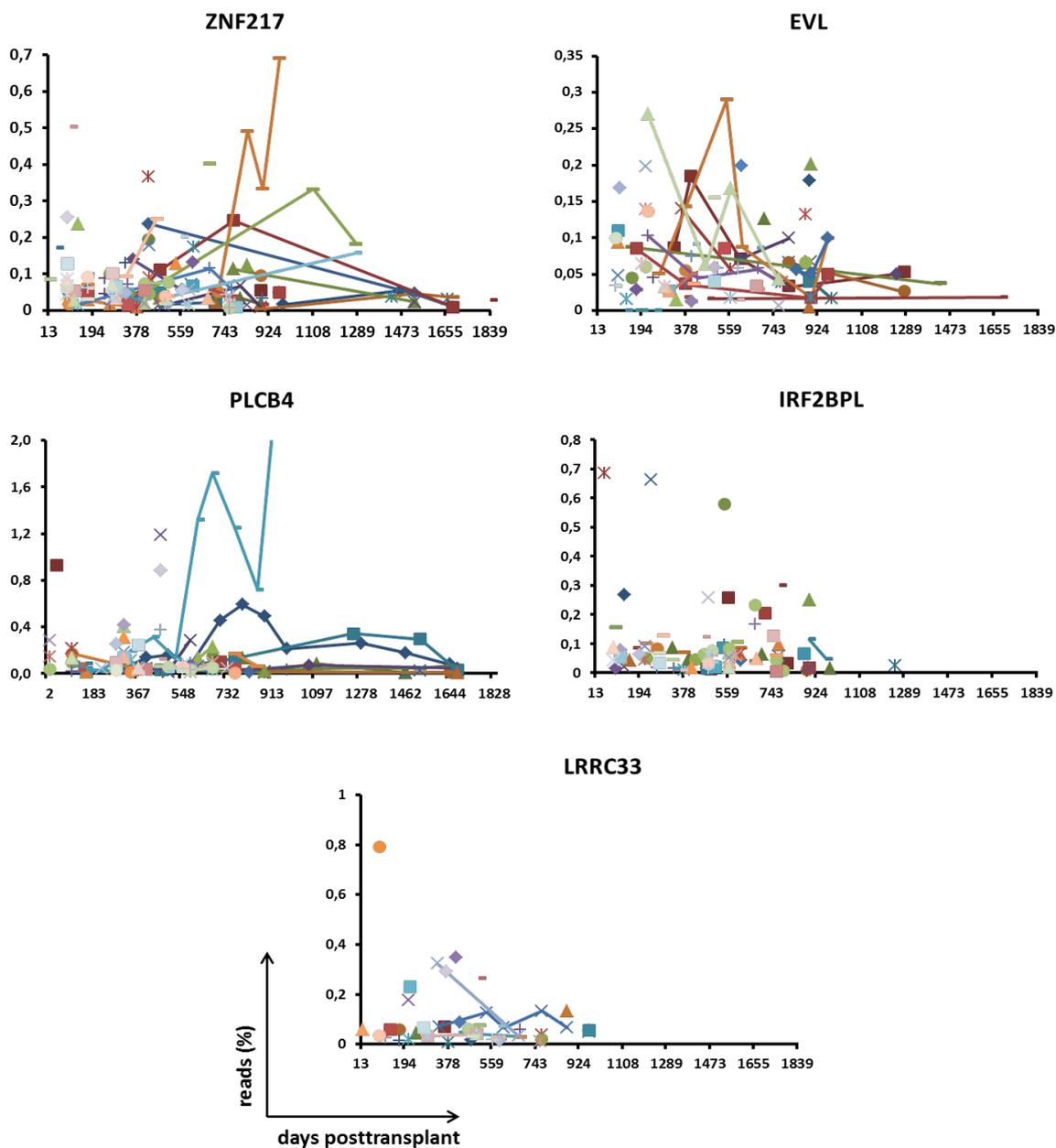


Figure 14: Long term clonal activity of the candidate genes. To evaluate the hematopoietic activity of clones carrying IS near identified candidate genes in post-transplant hematopoiesis, their contribution to blood formation were monitored at different time points after transplantation. Each dot indicates a specific clone. Lines represent the same clones detected at different time points.

Clones recurrently detected at several time points over 4 years might indicate for a long term clonal activity and their contribution in blood formation. Furthermore, at each time point new clones harboring IS near identified candidate genes were also detected.

4.1.2 Analyzing long-term multilineage contribution of hematopoietic stem and progenitor cells *in vivo*

The selected candidate genes were hypothesized to have regulatory functions in hematopoiesis; therefore their high expression in any of HSC or HPC populations might suggest a role for these genes in hematopoiesis. So far, various marker schemes have been defined for sorting HSC and HPC cells. To validate the long-term engraftment, differentiation and self-renewal properties of isolated HSC and HPC using defined cell surface markers, hematopoietic populations expressing various combinations of cell surface markers were purified from murine BM and transplanted into lethally irradiated mice at different numbers (Table 3). The engraftment of transplanted cells and their contribution to hematopoiesis in primary and secondary recipients were assessed over time. By using the mice with the same genetic background but different expression of CD45 isoforms (CD45.1 or CD45.2) on their nucleated hematopoietic cells, the engraftment of donor-derived cells (CD45.1) can be distinguished from host cells (CD45.2).

Table 3: HSC and HPC marker combinations used for isolating various hematopoietic populations

designation	population type	cell surface markers	number of transplanted cells
HSC_1	hematopoietic stem cells	CD34 ⁻ CD48 ⁻ CD135 ⁻ CD150 ⁺ LSK	560
HSC_2	hematopoietic stem cells	Lin ⁻ CD150 ⁺ CD48 ⁻	600
HSC_3	hematopoietic stem cells	Lin ⁻ CD150 ⁺ CD48 ⁻ CD244 ⁻	500
HPC_1	hematopoietic progenitor cells	CD34 ⁺ CD48 ⁺ CD135 ⁺ CD150 ⁻ LSK	7500
HPC_2	hematopoietic progenitor cells	Lin ⁻ CD150 ⁻ CD48 ⁺ CD244 ⁺	5000

Although all HSC and HPC populations were engrafted in transplanted mice, their efficiency and kinetics which was assessed by the chimerism of donor-derived cells, varied between different populations. Approximately 5 months after primary bone marrow transplantation, the PB of mice transplanted with HSC_1, HSC_2 and HSC_3 showed 62.7%, 43.4% and 38% donor-derived hematopoietic cells respectively. Transplanting $1.5-2 \times 10^7$ whole BM cells from these mice into secondary recipient mice exhibited 70% engraftment in HSC_1 and around 11% engraftment in HSC_2 and HSC_3, which confirmed self-renewal HSC properties (Figure 15a).

Within one month after BMT, the PB of mice transplanted with HPC_1 and HPC_2 populations comprised 19.75% and 16.25% of donor-derived cells respectively. 3 months after BMT the activity of transplanted cells decreased to 6.6% and 1.6% and no secondary engraftment was detected 1 month after BMT into recipient mice (Figure 15b).

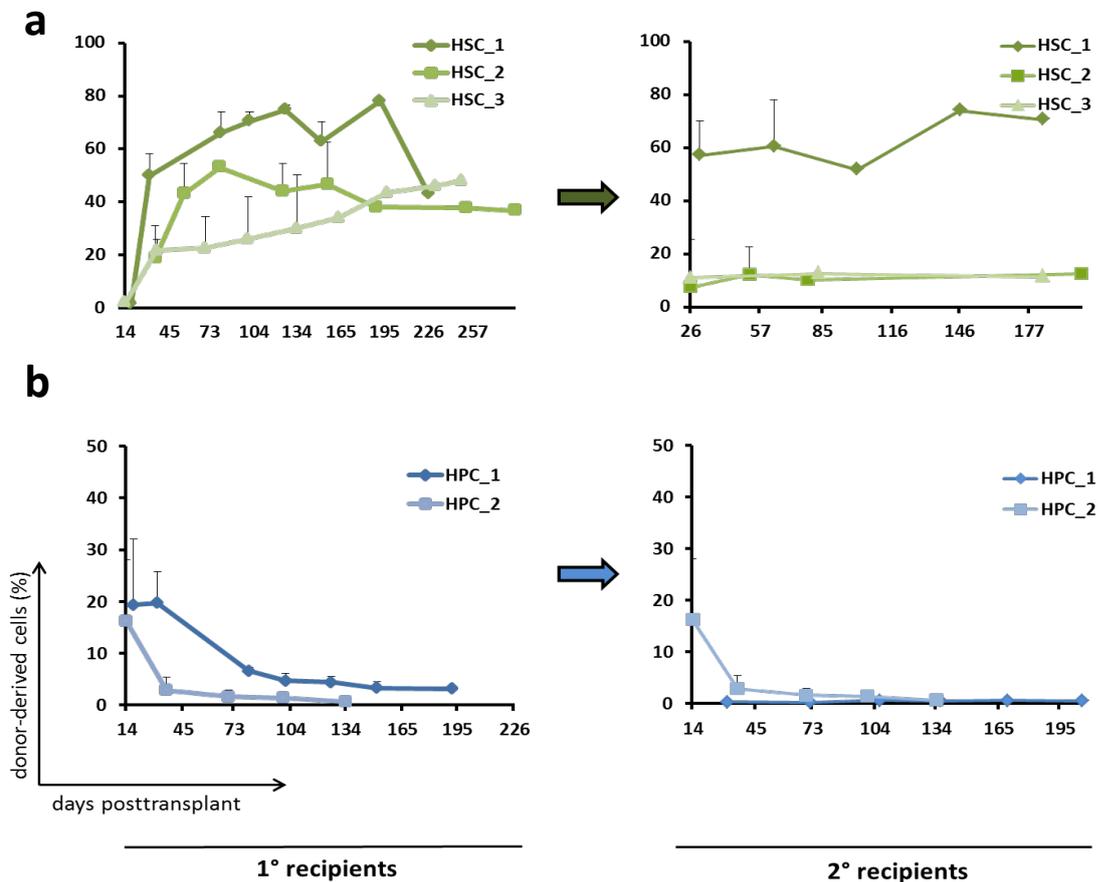


Figure 15: The engraftment of transplanted HSC and HPC with various combinations of stem cell markers in recipient mice. a) Approximately 4 months after bone marrow transplantation (BMT), HSC_1, HSC_2 and HSC_3 sorted populations showed 74.8, 44.1 and 15.8 percent engraftment of donor-derived hematopoietic cells respectively, whole BM cells of primary transplanted mice were transplanted into secondary irradiated mice and the self-renewal potential of the initially transplanted cells was evaluated over time (left). BM cells originated from HSC_1 population showed about 70% of engraftment in secondary transplanted mice whereas the engraftment decreased to around 11% for HSC_2 and HSC_3 in secondary transplantation (right). b) The activity of HPC_1 and HPC_2 sorted fractions was restricted to the first 12 weeks after transplantation (left). No secondary reconstitution was observed in both HPC populations 1 month after secondary transplantation (right).

To show the contribution of transplanted HSC and HPC to hematopoiesis, their differentiation potential to lymphoid and myeloid lineages was investigated by analyzing PB of recipient mice. We detected multilineage reconstitution in all transplanted HSC_1, HSC_2 and HSC_3 populations (Figure 16). Eight weeks after BMT, transplanted HSC_1 cells gave

rise to 17.5% of T-cells, 46% of B-cells, and 15% of macrophages respectively. These lineages retained a stable kinetics within 41 weeks after BMT, whereas the contribution of donor-derived granulocytes decreased from 73% (week 8) to 16% (week 41).

In 1 month after primary BMT, HPC_1 cells gave rise to 29.7% of T-cells, 56% of B-cells, 54% of granulocytes and 15.9% of macrophages. Since HPC_1 did not exhibit a long term activity, no donor-derived lymphoid and myeloid differentiation was detected 41 weeks post-transplant (Figure 16).

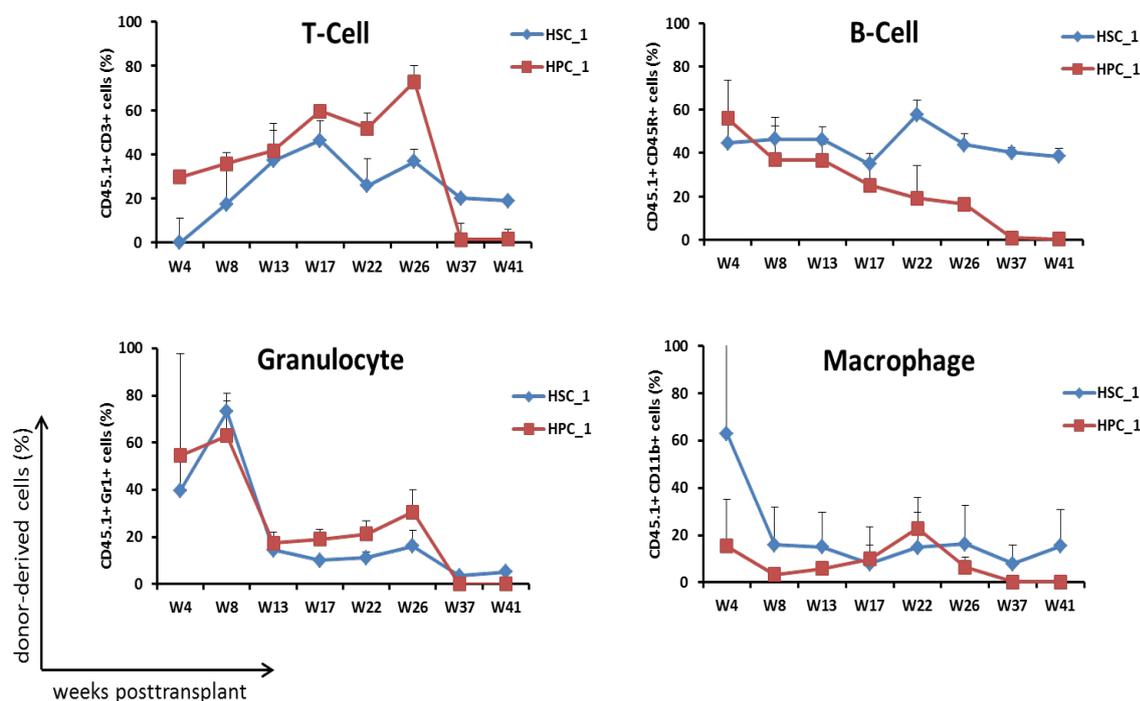


Figure 16: Differentiation capability of HSC_1 and HPC_1 cells into recipient mice hematopoiesis. Kinetics of donor-derived lymphoid cells (T cells and B-cells) and myeloid cells (granulocytes and macrophages) in the PB of the recipient mice over nine months. The mice were transplanted with either 1600 HSC expressing CD34⁻CD48⁻CD135⁻CD150⁺LSK, or 6000 HPC expressing CD34⁺CD48⁺CD135⁺CD150⁻LSK on their surface.

Taken all together, we observed that within all the transplanted HSC with various phenotypes, CD34⁻CD48⁻CD135⁻CD150⁺LSK purified HSC showed the highest multilineage engraftment in primary and secondary recipient mice. Moreover, as previously described, isolated HPC_1 population (CD34⁺CD48⁺CD135⁺CD150⁻LSK) was only active in the first 12 weeks after BMT in the primary mice. No engraftment was observed in secondary recipients.

4.1.3 Evaluating the expression of candidate genes in hematopoietic stem and progenitor cells

According to our initial hypothesis, selected candidate genes are involved in HSC and HPC regulation, thus we tested their expression in these populations. To elucidate the expression of candidate genes in murine HSC and HPC, global transcriptome dataset from highly enriched HSC, MPP (multipotent progenitor cell) 1-4 populations were analyzed (Cabezas-Wallscheid et al. submitted).

Interestingly, the overall RNA expression of LRRC33, in examined HSC and MPP1-4 populations was between 7 to 79 times higher than other candidate genes.

ZNF217 and LRRC33 showed an ascending expression pattern from HSC towards MPP4 population. We observed a nearly 2-fold increase at their RNA expression level in MPP4 compared to HSC.

Unlike ZNF217 and LRRC33, PLCB4 had a descending expression pattern from HSC and MPP1 towards MPP4. There was a 2-fold decline between the average of PLCB4 expression in HSC and MPP1 compared to its expression level in MPP4.

Although EVL almost showed no expression in HSC, MPP1, MPP2 and MPP3 populations, its expression had a dramatic 11-fold increase in MPP4.

Among all the candidate genes IRF2BPL was stably expressing in all of the investigated HSC and HPC populations (Figure 17).

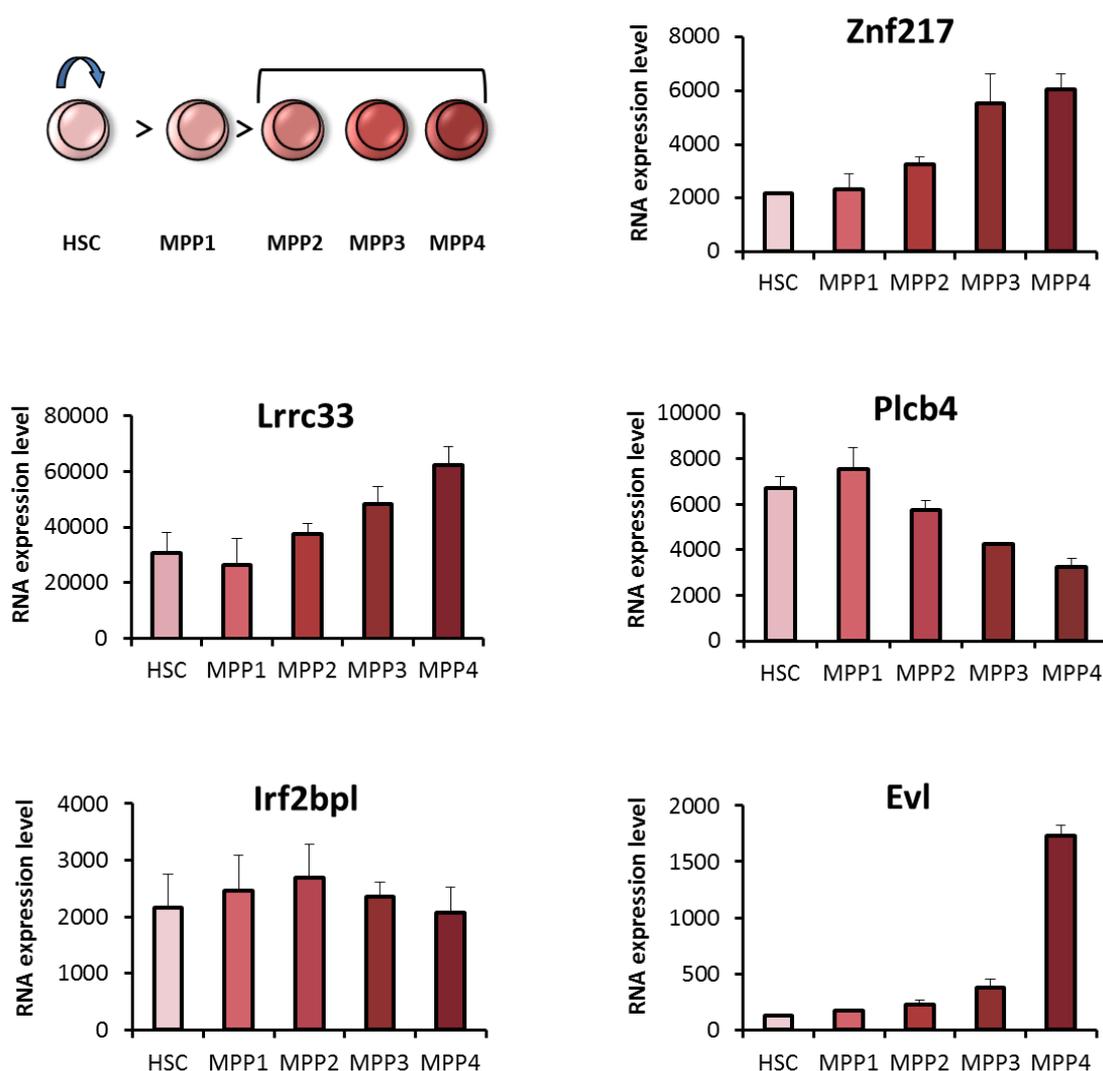


Figure 17: RNA expression of the selected candidate genes in HSC and HPC populations. The expression of each candidate gene in HSC and HPC was evaluated by whole transcriptome analysis. HSC: hematopoietic stem cell, MPP: multipotent progenitor (Cabezas-Wallscheid et al. submitted).

Apparently, all candidate genes were expressing at least in one of these stem and progenitor cell populations.

4.2 Investigating the function of candidate hematopoietic regulatory genes

To validate the involvement of selected candidate genes in hematopoiesis, various *in vitro* and *in vivo* assays were performed. In addition to hematopoietic cell lines, murine hematopoietic stem and progenitor cells were transduced with the lentiviral vectors (LV) encoding for the candidate genes to achieve a stable long-term expression. Following LV production, the function of candidate genes in cell proliferation, differentiation, cytokine

dependency, clonogenicity, long term multilineage reconstitution and self-renewal was assessed *in vitro* and *in vivo* (Figure 18).

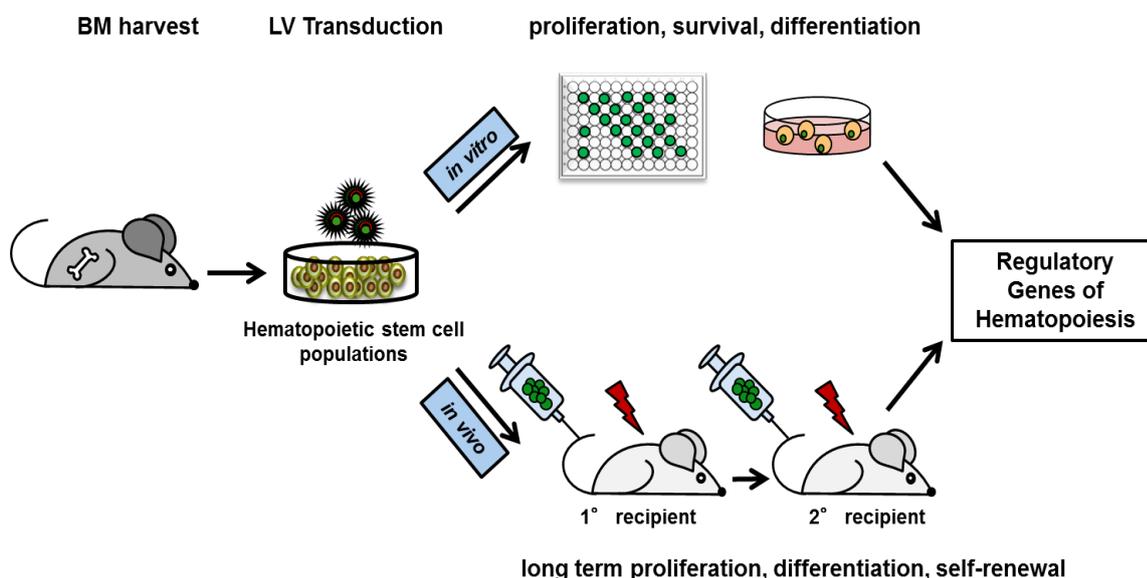


Figure 18: Experimental layout to investigate the role of selected candidate genes *in vitro* and *in vivo*. HSC and HPC populations were harvested, transduced with the LV of the candidate genes and their effect in hematopoiesis, such as proliferation, differentiation and self-renewal, was assessed *in vitro* and *in vivo*.

4.2.1 Examining proliferative activity and survival of single LSK cells *in vitro*

To investigate the role of selected candidate genes on proliferation and growth kinetics of HSC and HPC populations *in vitro*, we initially optimized un-manipulated LSK cell culture under stem cell conditions. We further analyzed and clustered the growth kinetics of individual LSK clones. Following optimization and analyzing the growth kinetics of single-LSK cell-derived clones, we performed a similar assay with mock-transduced and cyclin D2 (CCND2) overexpressing LSK cells (as positive control). This cyclin has a critical role in expansion of hematopoietic cells by functioning as a regulatory subunit of cyclin dependent kinase (CDK) 4 or CDK6 [125].

4.2.1.1 Evaluating growth kinetics of LSK cells cultured *in vitro*

LSK ($\text{Lin}^- \text{Sca-1}^+ \text{cKit}^+$) is a heterogeneous hematopoietic population comprising HSC and HPC. LSK cells transduced with the candidate genes and GFP control LV can be used to study the effect of the candidate genes on the proliferation of hematopoietic cells *in vitro*. First, we cultured un-manipulated single LSK cells under stem conditions and examined their proliferation kinetics *in vitro*. 120 individual single-cell-derived clones were purified from

mouse BM, cultured with stem cell cytokines and monitored for up to 90 days. The cells in each clone were counted every day in the first week and 2-3 times in the following weeks. Individual single-cell derived clones were diluted with 1:10 ratio when they grew to 10^5 cells. Growth kinetics of clones survived up to 21, 41, 63, and 90 days were analyzed and clustered based on the cells life span (Figure 19).

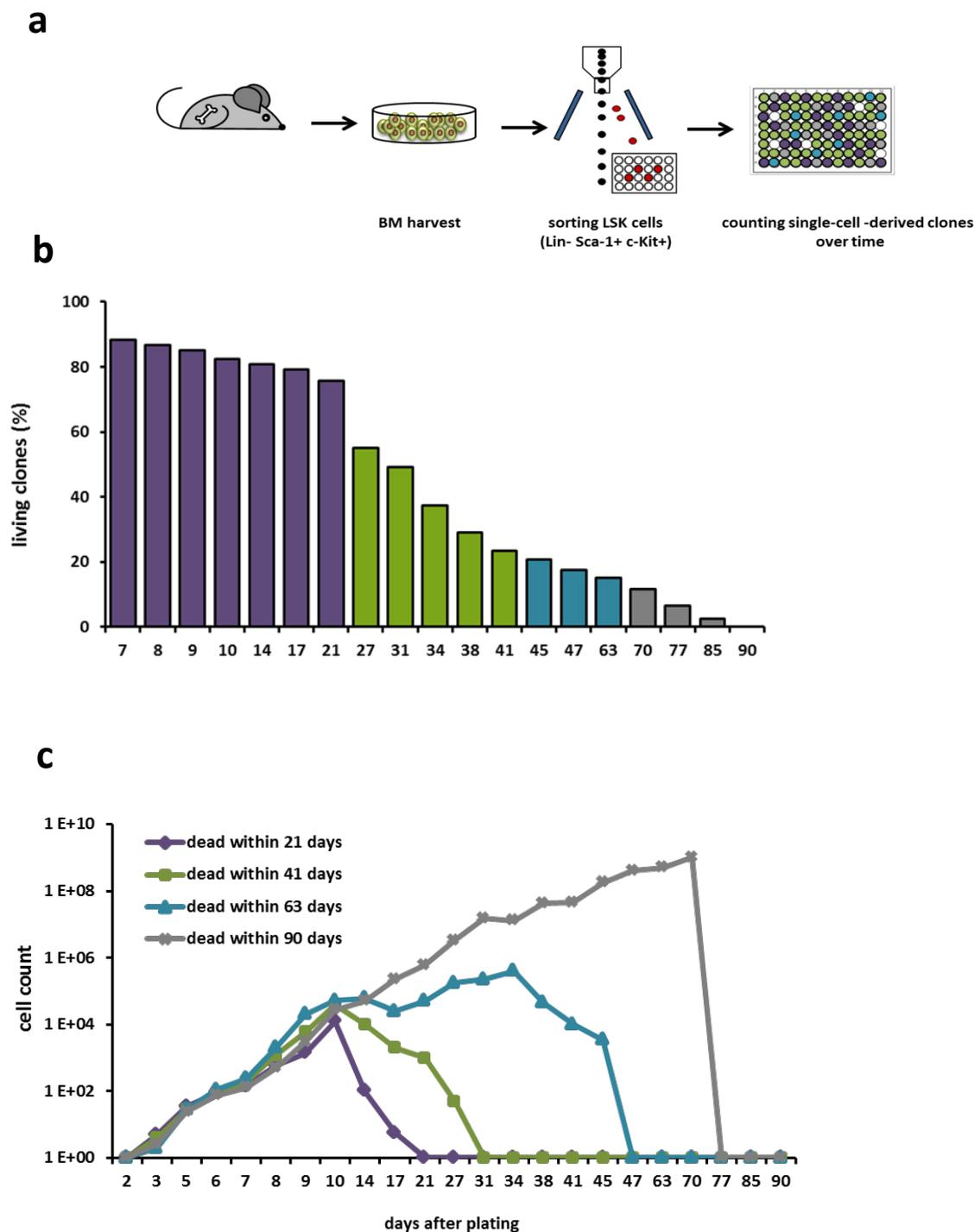


Figure 19: Proliferation kinetics of single-cell-derived LSK clones. a) BM was harvested and purified for LSK cells. Sorted single cells were cultured in stem cell medium and counted for up to 3 months to analyze their

growth kinetics *in vitro*. b) LSK single cell-derived clones were clustered into 4 groups based on their life span. c) Proliferation kinetics of the cultured clones in 90 days. All the clones showed the same kinetics in the first 10 days of *in vitro* culturing; however they showed different growth patterns after 10 days.

We observed that after initially identical growth kinetics for up to day 10 (d10), 24.2% of the sorted LSK clones died within 21 days, 52,5% of clones died between d21 to d41 of culture, 8.3% died between d41 and d63, and 15% of clones survived for up to d90.

4.2.1.2 Analyzing survival and growth kinetics of CCND2-transduced LSK cells

Following successful optimization of LSK cell culture *in vitro* and analyzing the growth kinetics of individual clones, we performed a similar assay using CCND2-transduced LSK cells as positive control to test whether any alterations in cell survival and growth kinetics could be detected. CCND2 has shown to have a role in expansion of hematopoietic cells by functioning as a regulatory subunit of cyclin dependent kinase (CDK) 4 or CDK6, whose activity is required for cell cycle G1/S transition [125]. However, in this assay we did not detect any significant difference between survival rates of CCND2 overexpressing cells compared to GFP control cells (Figure 20).

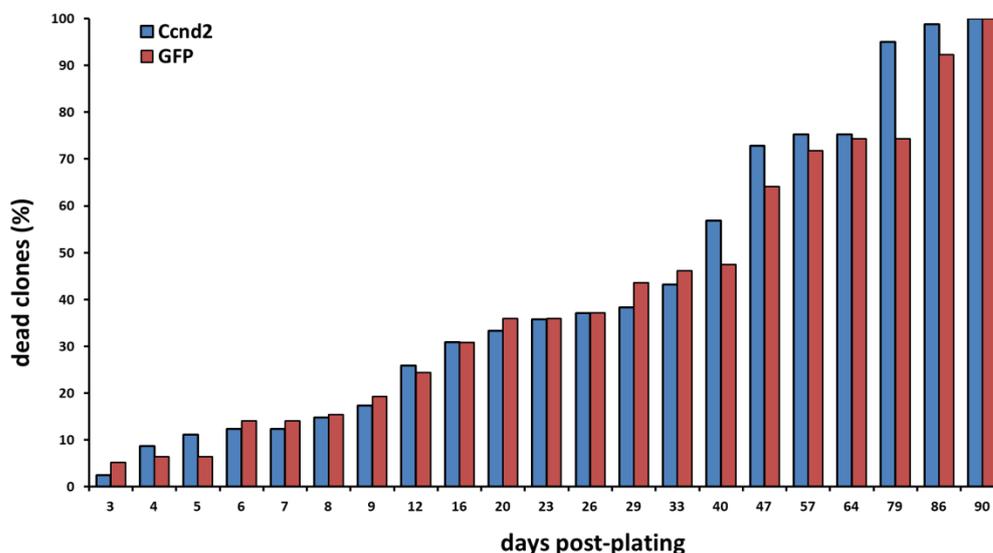


Figure 20: Survival of CCND2-transduced-LSK cells compared to mock transduced control. The frequency of dead single-cell derived LSK cells transduced with the LV encoding for CCND2 or GFP control over 3 months.

Furthermore, CCND2 and GFP transduced LSK cells were counted over time and the growth kinetics were generated and clustered into different groups based on their life span. We did not detect any dramatic growth changes in CCND2 transduced cells compared to GFP control cells.

4.2.2 Producing functional lentiviral overexpression vectors encoding for the candidate genes

To study the effect of selected candidate regulatory genes, coding sequences of EVL, LRRC33, PLCB4, ZNF217, and IRF2BPL were successfully cloned into a PGK-driven LV backbone by excision at BamH1 and SbfI restriction sites (Figure 21). Third generation LV particles of all vectors were produced with titers ranging from 2×10^7 to 8×10^8 transduction unit (TU)/ml.

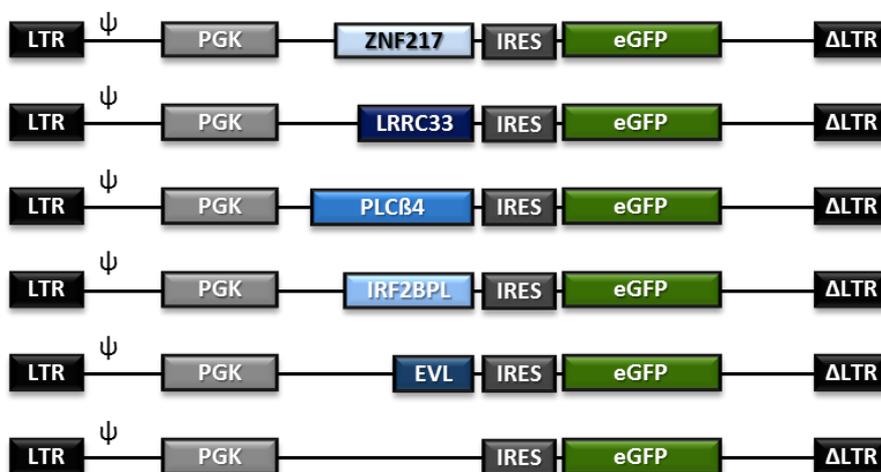


Figure 21: LV constructs encoding for individual candidate transgenes: pCCL-PGK-Transgene-IRES-eGFP. The coding sequences of five selected candidate genes were cloned into a LV backbone followed by an IRES sequence and eGFP as a marker gene. LTR: long terminal repeat; ψ : packaging signal; PGK: phosphoglycerate kinase; IRES: internal ribosomal entry site; eGFP: enhanced green fluorescent protein.

The transcription and translation of the transgenes were verified by examining transduced cell lines. A 1.000 to 100.000-fold increase of relative mRNA expression was observed by performing qPCR on cDNA obtained from 293T cells transduced with multiplicity of infection (MOI) 10 and sorted for GFP⁺ cells one week after transduction.

In addition, protein overexpression of PLCB4 and EVL, which had available antibodies, was detected in transduced 293T cells or hematopoietic C1498 cell (Figure 22).

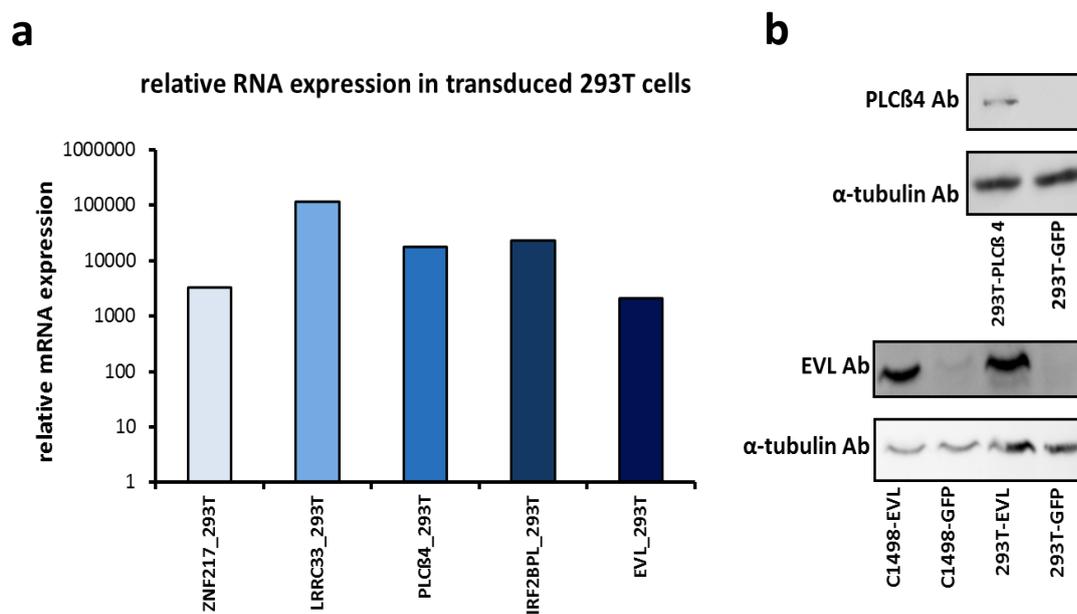


Figure 22: Overexpression of the candidate genes in transduced cells at RNA and protein level. a) The RNA expression of each candidate gene in 293T transduced cells was quantified using qPCR. The expression levels are normalized to mock control. b) Overexpression of PLCB4 and EVL proteins in 293T and C1498 transduced cells.

4.2.3 Studying the impact of candidate genes on hematopoiesis *in vitro*

In order to address whether candidate genes function in hematopoiesis regulation, *in vitro* assays were performed to investigate proliferation, colony forming capacity and cytokine independency of hematopoietic cells overexpressing candidate genes.

4.2.3.1 Studying the viability of hematopoietic cells overexpressing candidate genes

To evaluate the effect of candidate genes' overexpression on hematopoietic cell viability, we performed Cell Titer-Glo assay which is used to measure ATP activity level of cells. ATP is an indicator of metabolically active and viable cells [126]. ZNF217, LRCC33 and EVL transduced C1498 hematopoietic cells showed significant 1.9, 1.5, and 1.3-fold increase in their ATP level compared to GFP transduced control cells (Figure 23).

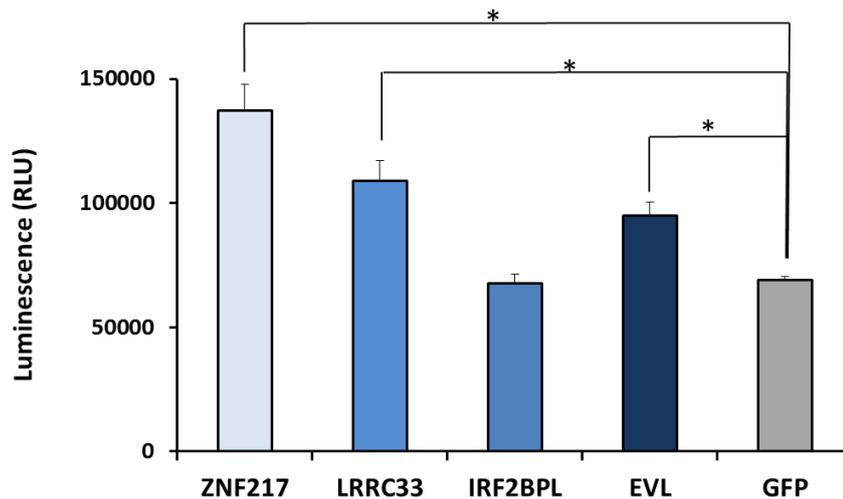


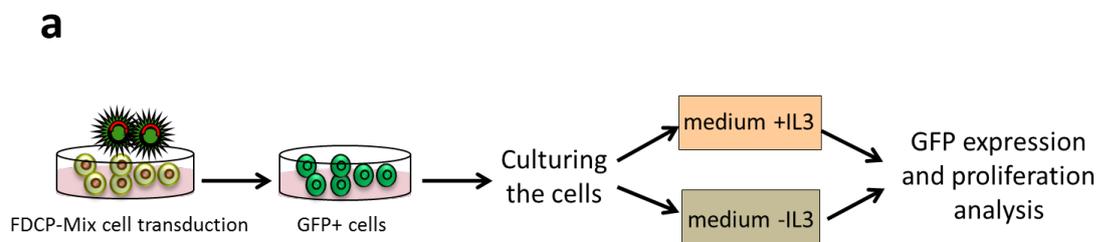
Figure 23: Metabolic activity of transduced hematopoietic C1498 cells. ZNF217, LRRC33 and EVL transduced C1498 cells show significantly higher ATP activity compared to mock control cells using Cell Titer-Glo assay.

Therefore, C1498 cells overexpressing ZNF217, LRRC33 and EVL demonstrated increased viability compared to GFP control cells.

4.2.3.2 Analyzing cytokine dependency in hematopoietic cells overexpressing candidate genes

To further explore the effect of candidate genes on the proliferation of hematopoietic cells, interleukin 3 (IL3)-dependent mouse FDCP-Mix progenitor cells were transduced with the LVs encoding for the candidate genes and cultured in the medium containing IL3. Moreover, to test whether the candidate genes overexpression can transform hematopoietic cells, transduced FDCP-Mix cells were cultured in the absence of IL3. The growth kinetics of transduced cells were evaluated over time.

We observed that cells transduced with ZNF217, IRF2BPL and c-Myc showed increased proliferation when cultured with IL3. Remarkably, ZNF217, IRF2BPL and c-Myc transduced FDCP-Mix cells survived and proliferated in IL3 free medium for 2 up to weeks (Figure 24).



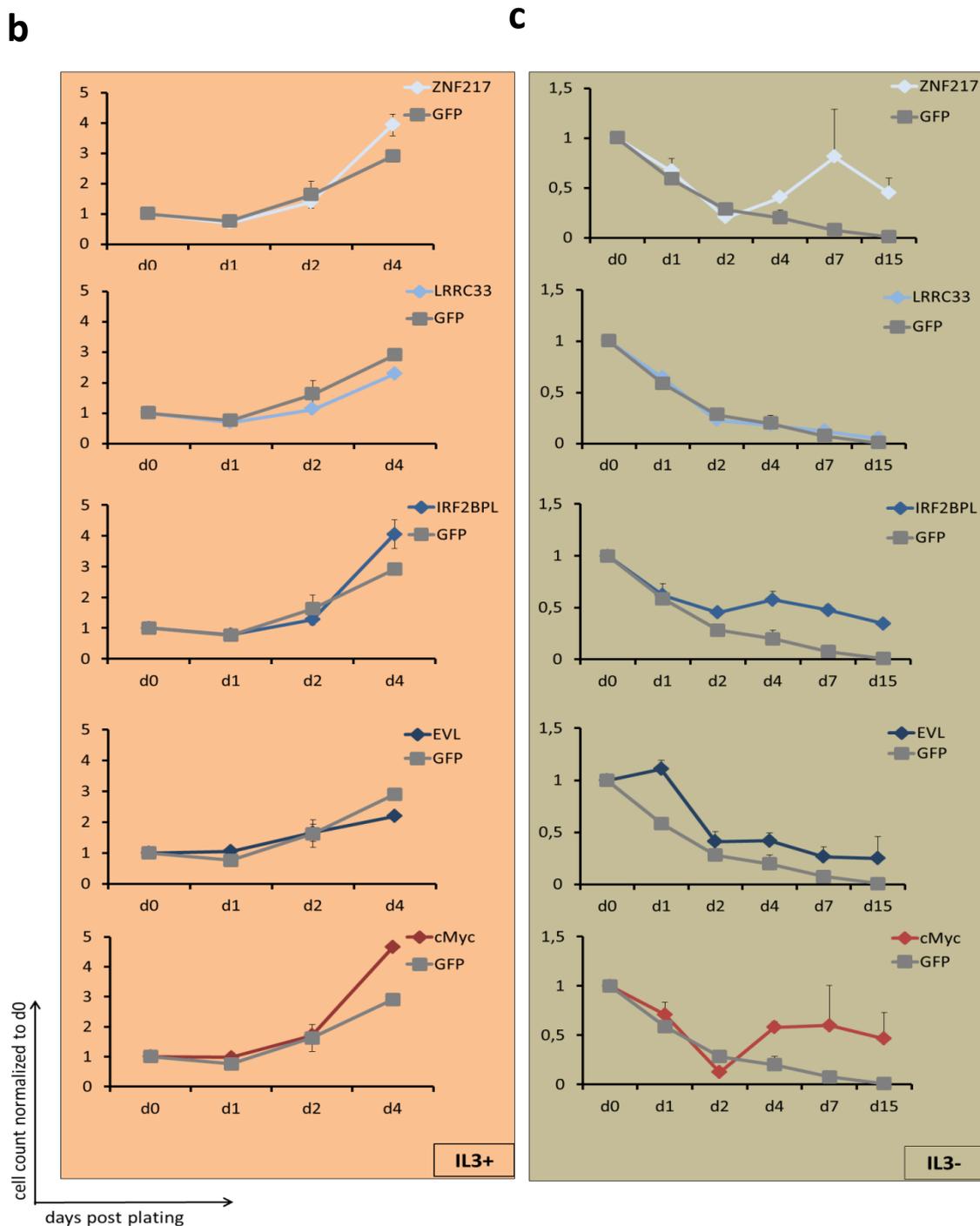


Figure 24: Growth kinetics of FDCP-Mix cells transduced with the LV of candidate genes in the presence (left) or absence (right) of IL3. In addition to c-Myc transduced cells, cells transduced with ZNF217 show higher growth rate in presence of IL3. ZNF217 transduced cells show higher survival rate compared to mock transduced cells when cultured in IL3- medium.

To examine whether overexpression of the selected candidate genes induces cytokine independency in another cytokine dependent cell line, IL3 dependent hematopoietic 32D

cells were transduced with the LV of the candidate genes and colony forming assay (CFC) was performed in semisolid medium lacking IL3.

It has been shown that 32D cells undergo cell cycle arrest and apoptosis when cultured in the absence of IL3 [127]. Moreover, IL3 dependent cells overexpressing c-Myc, a known proto-oncogene- acquire the ability to survive and form colonies in IL3-free medium[128]. Therefore, in this assay c-Myc overexpressing cells were used as positive control. We observed that in addition to c-Myc, ZNF217 transduced cells formed a significant higher number of IL3 independent colonies compared to GFP transduced cells. However, cells overexpressing all other candidate genes did not survive in the absence of IL3 (Figure 25).

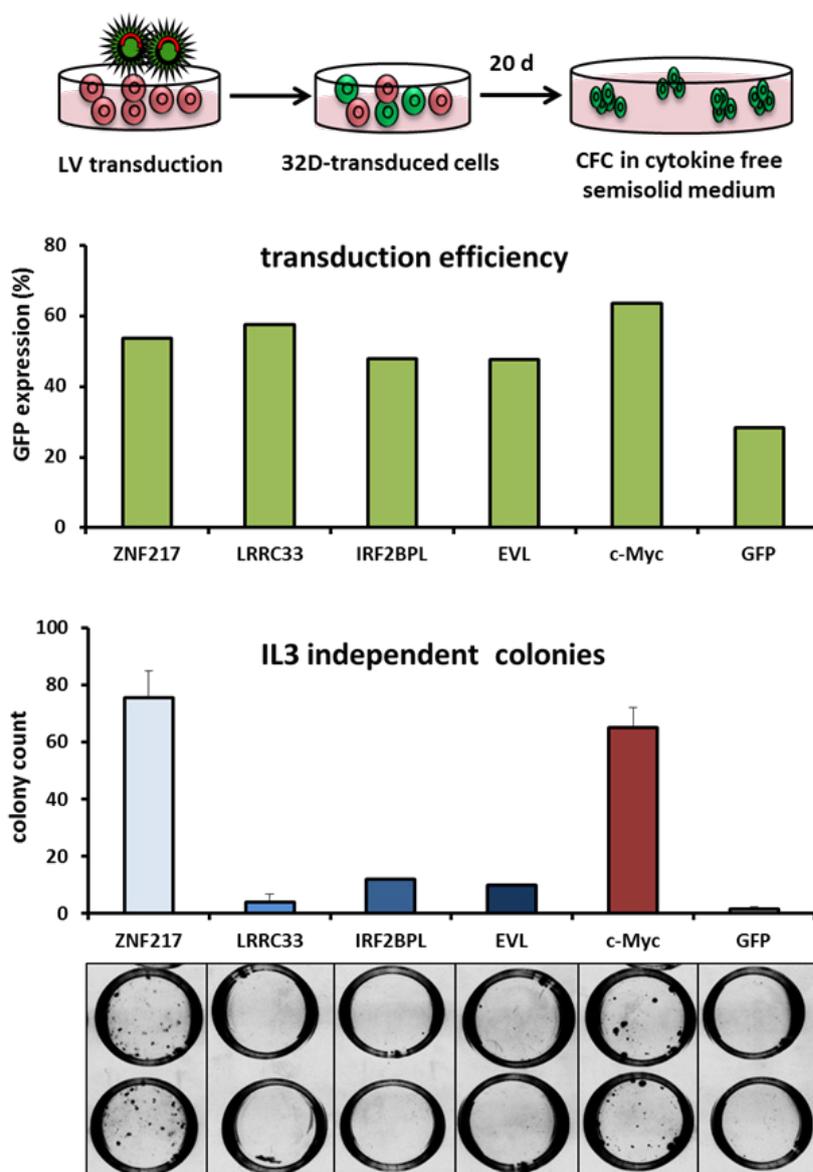


Figure 25: CFC assay on transduced 32D cells transduced with the LV of four candidate genes in absence of IL3 cytokine. Hematopoietic 32D cells were transduced with LV of ZNF217, LRRC33, IRF2BPL, and EVL candidate genes and cultured in semisolid medium in the absence of IL3 cytokine. The number of colonies was counted after 20 days.

4.2.3.3 Assessing gene expression profiling of ZNF217-transduced cells

ZNF217 is known to be a transcription regulator. To gain a better understanding of up or down regulated cellular pathways and ZNF217 target genes in hematopoietic cells, we performed a global expression profiling on hematopoietic 32D cells transduced with ZNF217 or GFP control vector. We observed that the number of downregulated genes ($n=337$) was 4 times more than upregulated genes ($n=85$) in ZNF217 overexpressing cells compared to control cells (Figure 27).

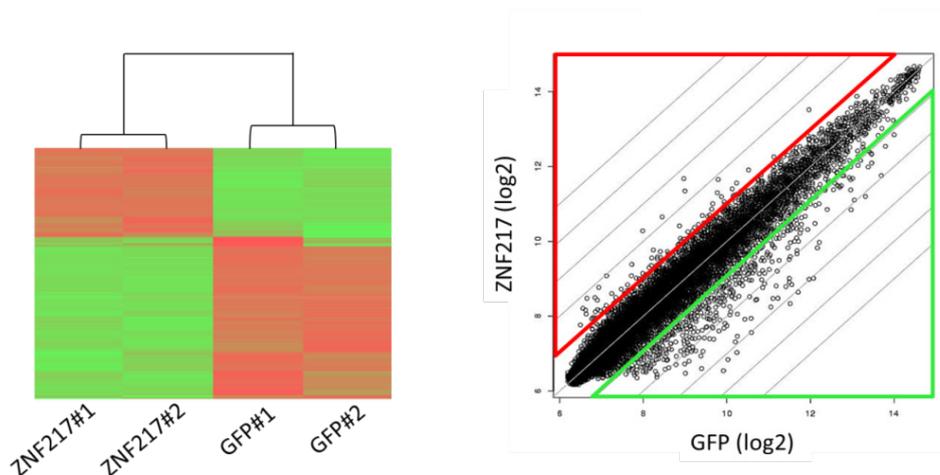


Figure 27: Global gene expression profiling of hematopoietic 32D cells overexpressing ZNF217. Up (red) and down (green) regulated genes in ZNF217 and GFP transduced cells.

Using Ingenuity Pathway Analysis (IPA) software the most prominent deregulated pathways in ZNF217 overexpressing cells were detected by categorizing deregulated genes based on their biological functions. The top deregulated networks in ZNF217 overexpressing cells were identified by associating up and downregulated genes to biological functions. The networks with the highest scores were involved in cancer, cellular movement, inflammatory response and immune cell trafficking (Figure 28).

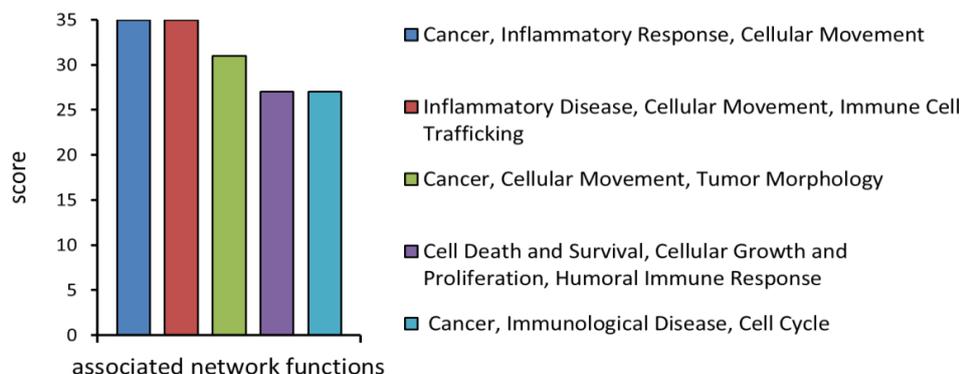


Figure 28: Top deregulated biological networks identified in ZNF217 overexpressing cells.

Moreover, based on IPA classification, deregulated genes were evaluated in 3 different categories: molecular and cellular functions, physiological system development, and diseases and disorders. The contribution of each deregulated pathway in its respective category was determined according to the total number of detected deregulated genes involved in that pathway. In the category of molecular and cellular functions, the highest number of deregulated genes played roles in cell death and survival (24.3%). In physiological system development category most of the deregulated genes had functions in hematological system development (29.8%). In diseases and disorders category, genes involved in immunological diseases had the highest contribution among all other deregulated genes (Figure 29).

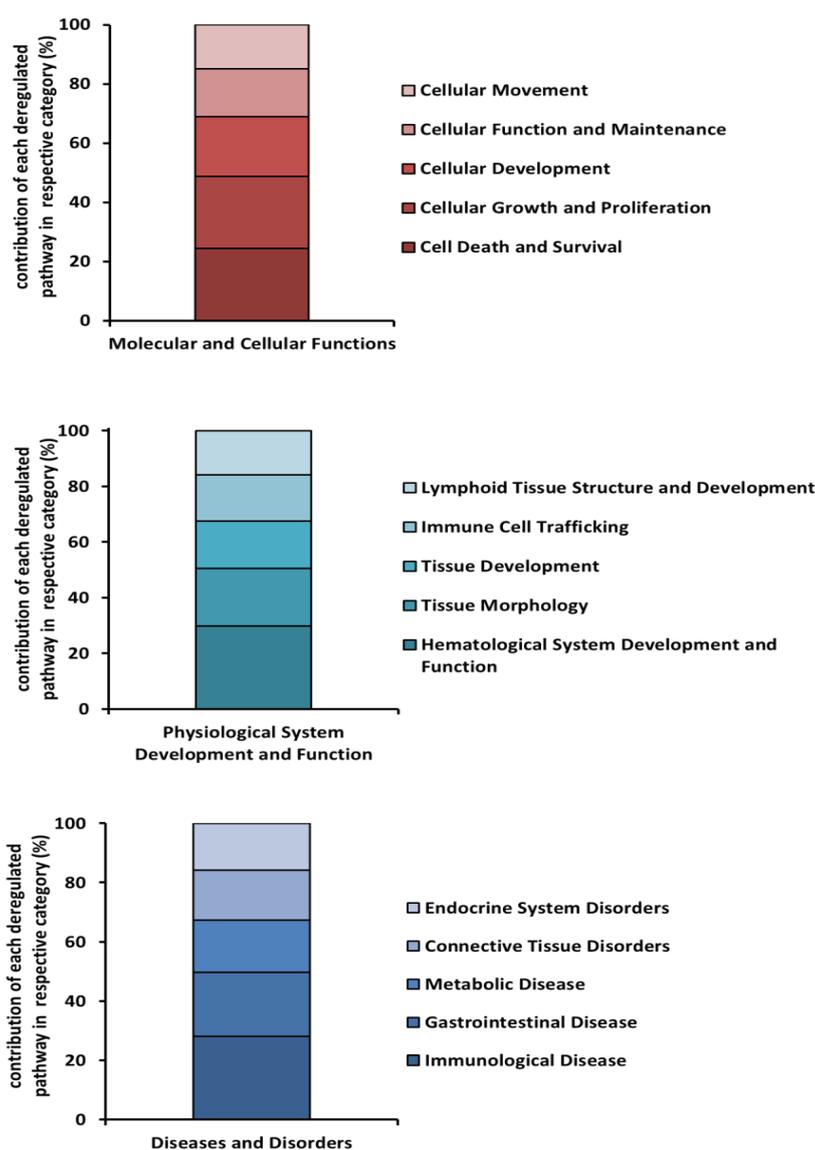


Figure 29: Categories and subcategories of ZNF217 deregulated genes. Using IPA software, up or down regulated genes in ZNF217 overexpressing cells were detected and classified into several groups based on their biological functions. IPA: Ingenuity Pathway Analysis.

4.2.3.4 Evaluating the role of EVL in colony forming activity

We previously demonstrated that EVL overexpressing C1498 hematopoietic cells have higher ATP metabolic activity (section 4.2.3.1). To examine the colony forming ability of these cells we performed a soft agar CFC assay. C1498 hematopoietic cells were transduced with LV encoding for EVL. GFP⁺ cells were sorted and cells were plated at 50, 100, and 250 cells per dish. The number of colonies was counted 20 days after plating. EVL transduced cells formed 2-fold higher number of colonies compared to GFP transduced cells when 250 cells/dish were seeded (Figure 30).

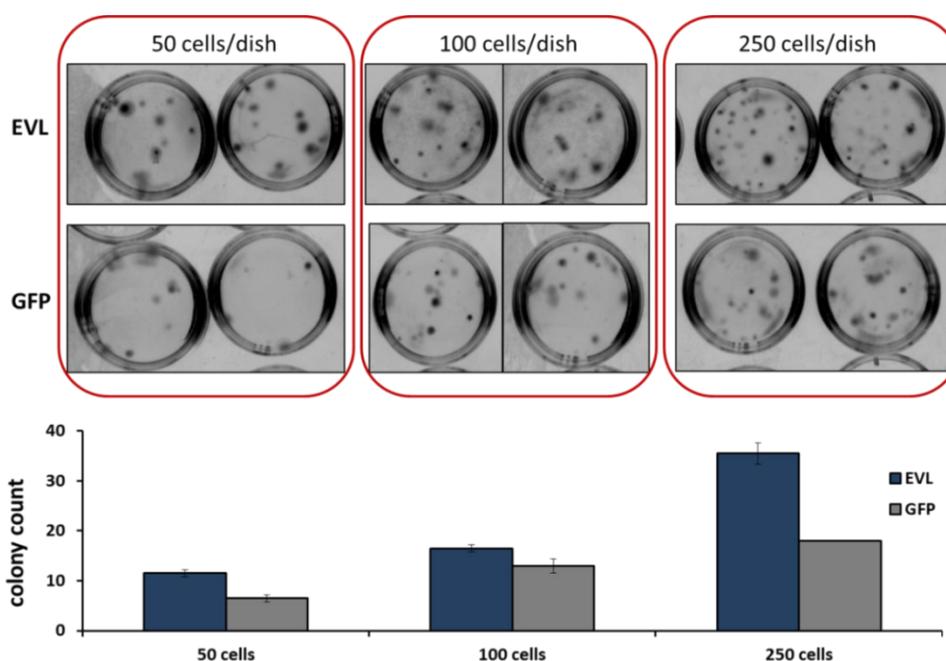


Figure 30: CFC assay with transduced C1498 hematopoietic cells overexpressing EVL. After transducing and purifying GFP⁺, C1498 cells, CFC assay was performed plating different number of cells in soft agar.

Therefore by performing this CFC assay we observed that C1498 cells overexpressing EVL showed higher clonogenicity ability in comparison with GFP control cells.

4.2.3.5 Analyzing the impact of LRRC33 on NF- κ B activity

LRRC33 protein structure resembles toll-like receptor (TLR) protein family members [129, 130]. TLR4 is one of the toll-like receptors which plays a critical role in the activation of innate immunity and NF- κ B pathway [131]. Therefore, we tested whether LRRC33 can function as a TLR4 co-receptor and affect NF- κ B activity in cells stimulated with lipopolysaccharides (LPS).

To validate the TLR-mediated inhibitory effect of murine LRRC33 on NF- κ B activity, 293T cells were co-transfected with various combinations of LRRC33, GFP and TLR4 constructs together with both LacZ and NF- κ B firefly luciferase reporter plasmids. The plasmids used to transfect 293T cells were either one of LRRC33, GFP or TLR4 or the combination of LRRC33 and TLR4 or GFP and TLR4. NF- κ B activity level of cells was determined 16 hours after LPS stimulation.

We observed that NF- κ B activity decreases in LPS-stimulated cells expressing both LRRC33 and TLR4 in comparison with the same transfected cells which are not treated with LPS and also mock control cells transfected with both GFP and TLR4 constructs in either presence or absence of LPS (Figure 31).

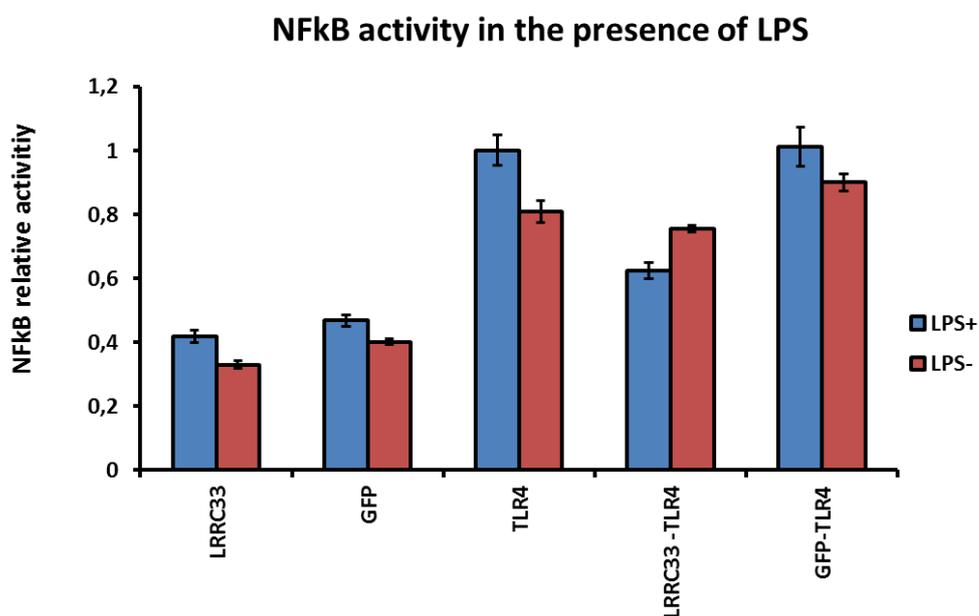


Figure 31: NF- κ B activity in transfected 293T cells. 293T cells were transfected with LRRC33, GFP, TLR4 or co-transfected with LRRC33 and TLR4 or GFP and TLR4. Cells were stimulated with LPS 2 days after transfection and NF- κ B activity level was measured.

Therefore we detected an inhibition of NF- κ B activity in cells transfected with both LRRC33 and TLR4 vectors after LPS stimulation.

4.2.4. Investigating the function of candidate regulatory genes *in vivo*

To investigate the role of the candidate genes on hematopoiesis *in vivo*, LSK cells were transduced with LV encoding for the candidate genes and 1000 cells were transplanted into lethally irradiated recipient mice the day after transduction. PB from transplanted mice were collected and examined monthly for evaluating the donor-derived multilineage engraftment (Figure 32a).

LSK cells were transduced with the LVs of EVL, LRRC33, IRF2BPL, ZNF217, c-Myc and GFP with the transduction efficiency of 1.9%, 36.7%, 1.5%, 0.1%, 5.5%, and 19.6% respectively (Figure 32b). Donor-derived hematopoietic cells comprised about 40% of PB cells 31 weeks post-transplant (Figure 32c).

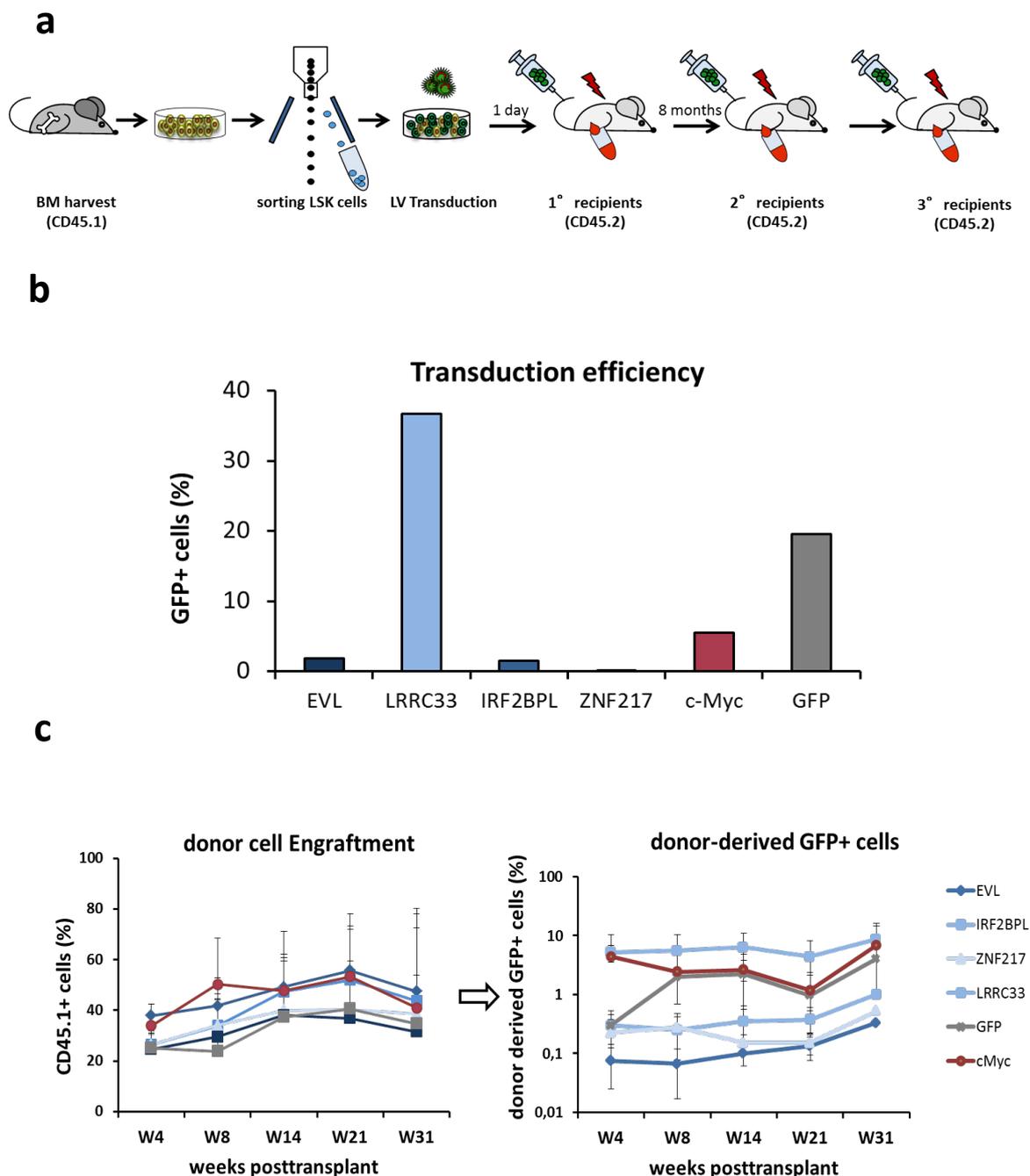


Figure 32: Investigating the function of the candidate genes *in vivo*. a) LSK cells were sorted from BL6 mice BM (CD45.1), transduced with the LV of the candidate genes and transplanted into lethally irradiated recipient mice. The effect of the candidate genes in hematopoiesis of transplanted mice was studied by analyzing the mice PB over time. b) Transduction efficiency of transduced LSK cells with various LV of candidate genes. c)

Engraftment of the transplanted cells into primary recipient mice over 8 months (left), donor-derived GFP⁺ cells in PB of transplanted mice (right).

In total EVL, ZNF217, IRF2BPL and GFP-transduced LSK cells were transplanted into 4 mice each and LRRC33 and c-Myc transduced LSK cells were transplanted into 5 mice each.

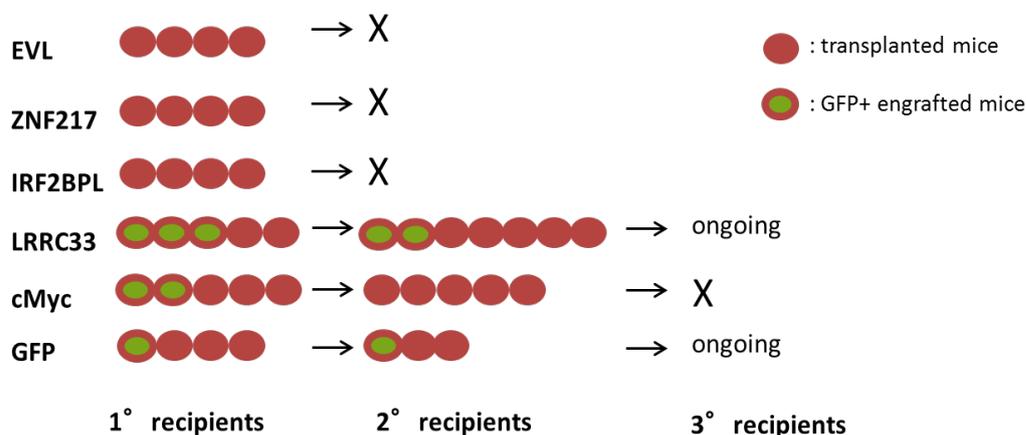


Figure 33 : Number of mice transplanted and engrafted with transduced LSK cells. Of all the primary recipient mice 3 out of 5 LRRC33, 2 out of 5 c-Myc and 1 out of 4 GFP mice showed GFP⁺ cells in their PB. After secondary transplantation, 2 out of 7 LRRC33 and 1 out of 3 GFP mice showed > 1% GFP⁺ engrafted cells in PB samples.

After examining the mice PB 3/5, 2/5, and 1/4 of primary recipient mice transplanted with LRRC33, c-Myc and GFP-transduced LSK cells respectively exhibited engraftment of GFP⁺ donor-derived cells (Figure 33). We observed that mice transplanted with EVL, ZNF217, and IRF2BPL-transduced LSK cells did not show an efficient GFP⁺ donor cell engraftment which was due to the initial low transduction efficiency. The hematopoietic multilineage engraftment of mice with efficient GFP⁺-donor-derived cell engraftment (LRRC33 and GFP control) was evaluated in primary and secondary mice.

4.2.4.1 Analyzing hematopoiesis in mice transplanted with LRRC33-transduced LSK cells

According to preliminary serial transplantation results, primary recipient mice transplanted with LRRC33 and GFP-transduced LSK cells respectively showed 5% and 12% of donor-derived GFP expressing cells in their PB at week 36 post BMT. At weeks 31 and 36, the ratio of GFP⁺ T-cells was about 2-fold reduced in LRRC33 transplanted mice compared to GFP control mice. In contrast, the ratio of GFP⁺ macrophages was 2 and 5 fold increased in LRRC33 transplanted mice compared to GFP control vector transplants at weeks 31 and 36, respectively (Figure 34).

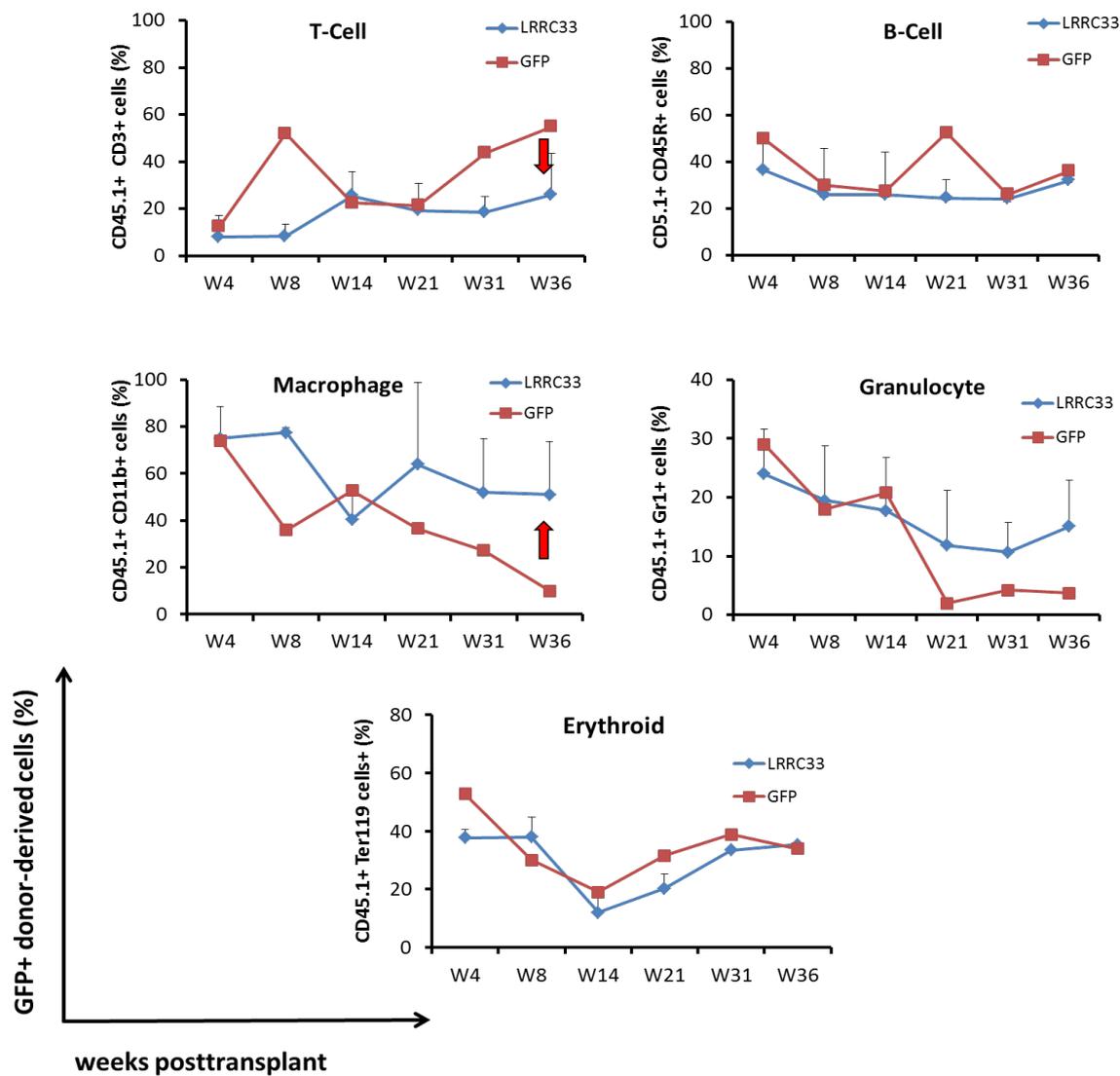


Figure 34: Multilineage engraftment in primary LRRC33 and control GFP transplants. Ratio of GFP⁺ donor-derived peripheral T- cells (CD45.1⁺ CD3⁺), B-cells (CD45.1⁺ CD3⁺), macrophages (CD45.1⁺ CD11b⁺), granulocytes (CD45.1⁺ Gr1⁺) and Erythroides (CD45.1⁺ Ter119⁺) 4 – 36 weeks after BM transplants into primary recipients.

At week 36 whole BM from LRRC33 and GFP primary mice was harvested and 8×10^6 cells were transplanted into each secondary recipient mouse, 16 weeks after BMT 4% to 36% of GFP⁺ cells were detected in mouse PB. Comparable to primary recipients, T-cells were 2 to 3 fold reduced in LRRC33 transplanted secondary recipients compared to GFP control mice, whereas the ratio of GFP⁺ macrophages was 6 to 12-fold increased in LRRC33 transplanted mice at weeks 6 to 16 after BMT (Figure 35).

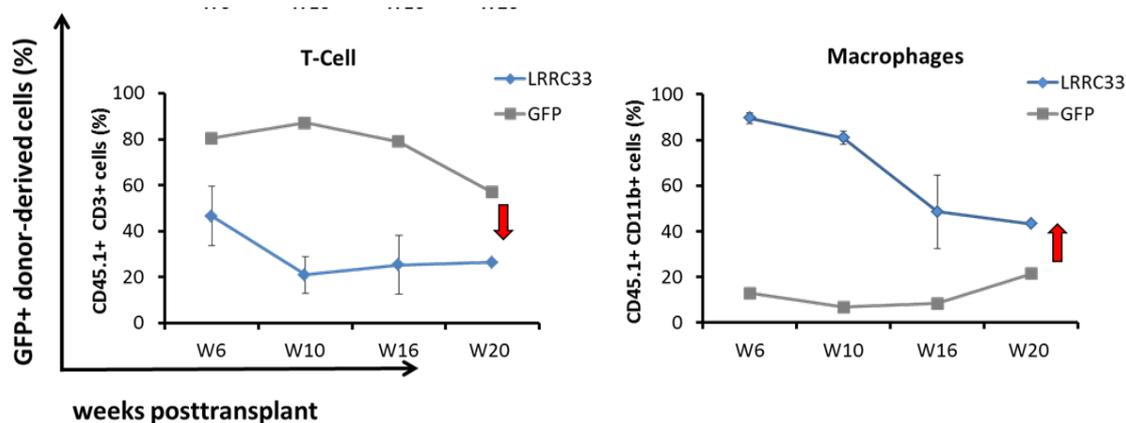


Figure 35: Multilineage engraftment in secondary LRRC33 and control GFP transplants. Ratio of GFP⁺ donor-derived peripheral T- cells (CD45.1⁺ CD3⁺), and macrophages (CD45.1⁺ CD11b⁺) 6 - 20 weeks post transplants into secondary recipients.

Taken together, we observed that LRRC33-overexpressing LSK cells show higher differentiation towards macrophages and lower differentiation towards T-cells compared to control-transduced cells in primary and secondary transplants.

4.2.4.2 Characterizing BM cells from PLCB4 knockout mice *in vitro* and *in vivo*

Comparing the expression of the candidate gene PLCB4 in five HSC and HPC populations, we observed that PLCB4 had the highest expression level in HSC and MPP1 populations which may indicate for its critical role in these populations (section 4.1.2). Therefore, knocking down the expression of PLCB4 in hematopoietic stem cell populations might have effects on hematopoiesis process.

PLCB4 knockout (KO) mice were already generated [132], but never reported for any changes in their hematopoietic system. Due to the disruption of the catalytic domain in an exon of PLCB4 gene, there is no PLCB4 protein expression in KO mice (Figure 36a). To understand the role of PLCB4 in hematopoiesis, we analyzed the BM from PLCB4 KO mice in comparison with wild type and heterozygote mice *in vitro* and *in vivo*.

Compared to wild type (+/+) and heterozygote (+/-) mice, PLCB4 knockout (-/-) mice show lower body weight +/- (Figure 36b). Furthermore, we observed that breeding +/- mice leads to about 11% of 21-day old -/- progenies which is 14% less than expected according to the mendelian ratio of about 25% (Figure 36c).

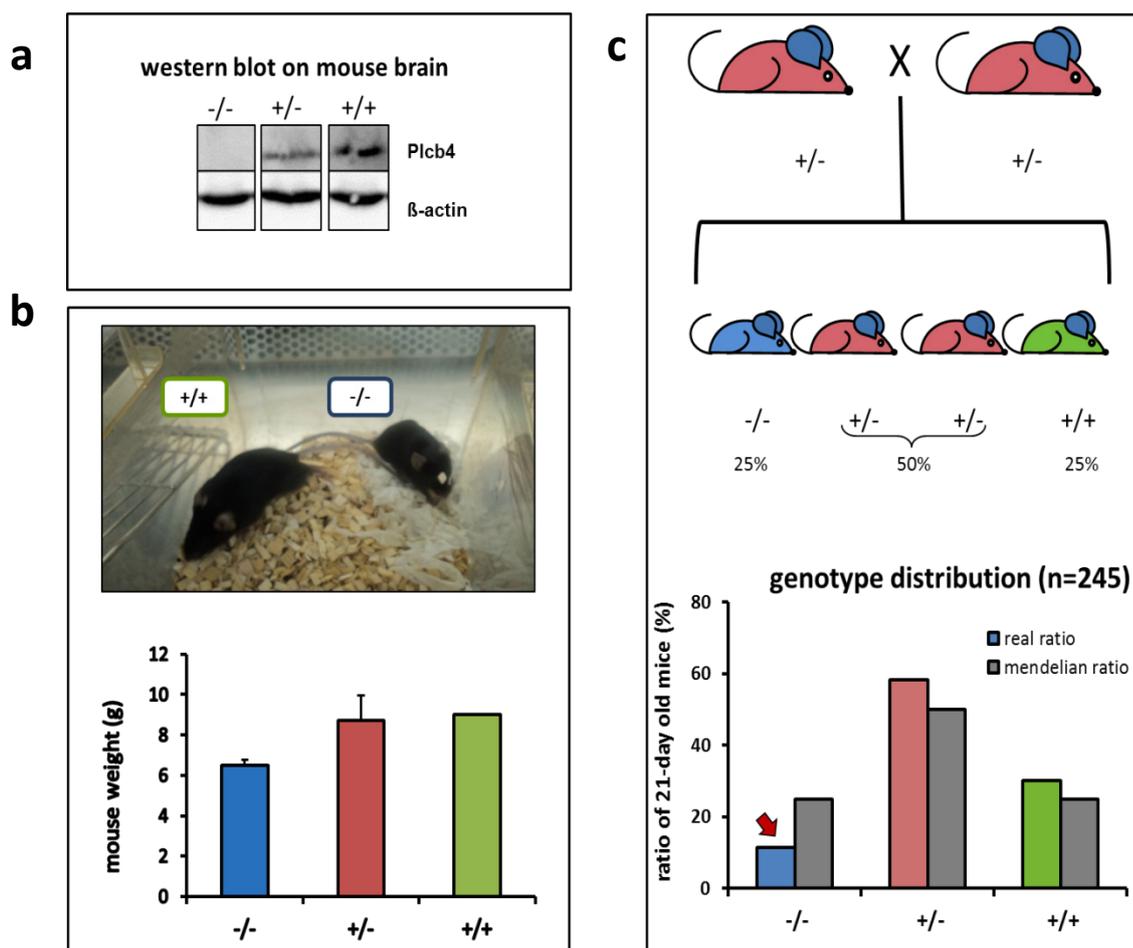


Figure 36: PLCB4 knockout mice. a) No PLCB4 protein expression is detected in KO mice. b) -/- mice show a reduced body size compared to +/- and +/+ mice. c) Breeding +/- mice leads to lower rate of mutant 21-day old mice.

4.2.4.2.1 Examining BM cells from PLCB4 knockout mice in pre-B cell CFC assay

Analyzing whole BM of -/-, +/- and +/+ at different ages revealed a 2-fold decrease in the LSK population from 0.2% in +/+ and +/- to 0.08% in -/- 18 and 25-day old mice. However, no reduction was detected in 40-day old mice (Figure 37).

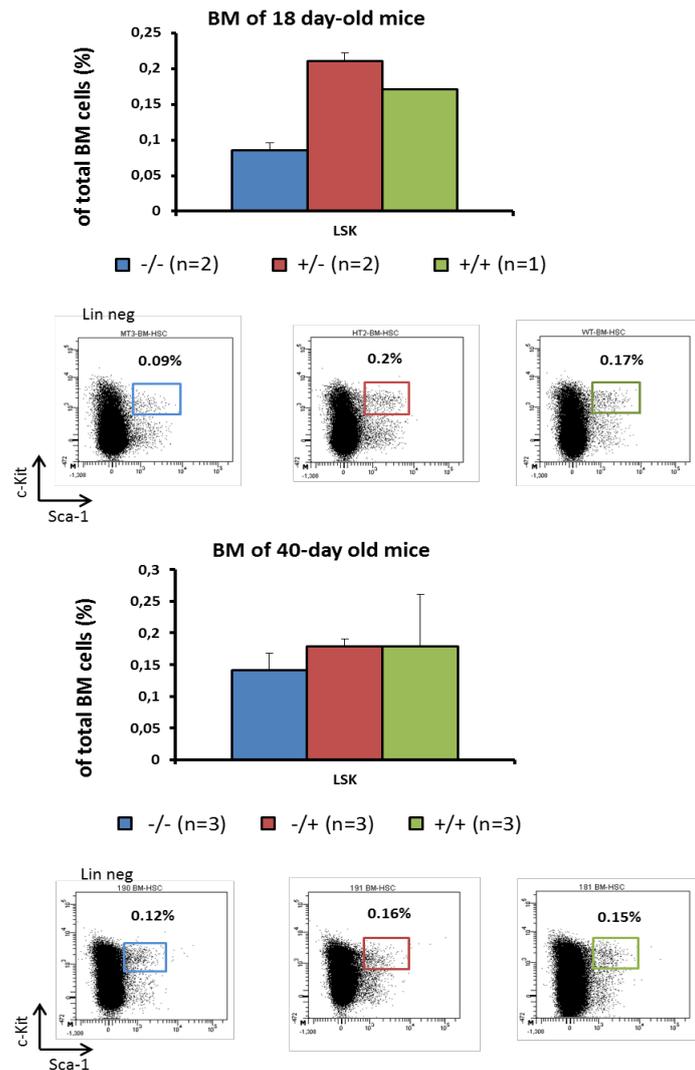


Figure 37: Analyzing LSK population in BM from PLCB4 mice. a) 18-day old -/- mice showed a 2 fold reduction in the LSK cells in their BM compared to +/- and +/+ PLCB4 mice. B) No significant difference was observed between -/-, +/- and +/+ 40-day old mice. -/- : knockout; +/-: heterozygote; +/+ : wild type PLCB4 mice.

To test the potential of hematopoietic cells from PLCB4 mutant mice in forming preB colonies, preB-cell CFC assay was performed with BM cells from 18-day old mice. It was shown that there is a 2 and 3-fold reduction in the number of B-cell colonies in -/- vs. +/- , and -/- vs. +/+ mice respectively. Similarly B-cell colonies from 25-day old -/- mice were about 2 folds less than colonies from +/- mice. However, no significant difference was observed in B-cell colony count in 40-day old mice (Figure 38).

Unlike pre-B CFC assay, myeloid colonies detected in myeloid CFC assay did not exhibit any significant differences in the number and type of various myeloid colonies in -/-, +/- and +/+ PLCB4 mice (Figure 38).

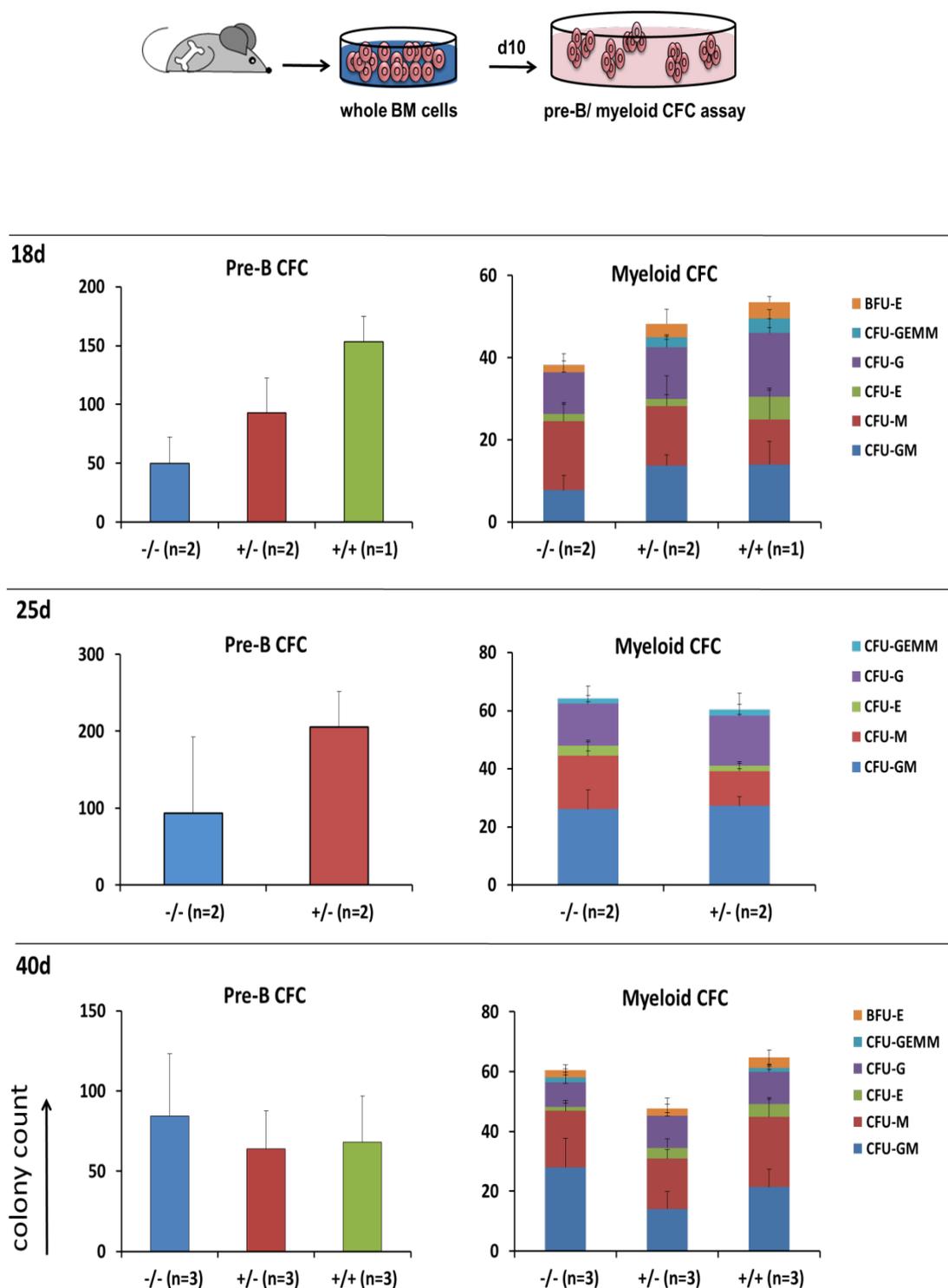


Figure 38: Myeloid and B-cell CFC assay with the whole BM cells from -/-, +/- and +/+ PLCB4 mice at different ages. Whole BM cells derived from -/-, +/- and +/+ PLCB4 mice at ages of 18 days, 25 days and 40 days were seeded in pre-B cell or myeloid semisolid medium and the number of colonies was counted 10 days later. In CFC myeloid assay, the type of colonies was also determined. -/- : knockout; +/-: heterozygote; +/+ : wild type PLCB4 mice.

4.2.4.2.2 Investigating multilineage engraftment of BM cells from PLCB4 knockout mice

To investigate how PLCB4 BM cells contribute to hematopoietic lineages differentiation *in vivo*, 10^6 BM cells from each of $-/-$, $+/-$ and $+/+$ mice were mixed with 10^5 BM carrier cells and injected into the tail vein of lethally irradiated mice expressing CD45.1 on their hematopoietic cells. Transplanted mice were bled every month and the chimerism of donor-derived cells and rate of hematopoietic lineages differentiated from donor-derived cells were examined for up to 6 months (Figure 39).

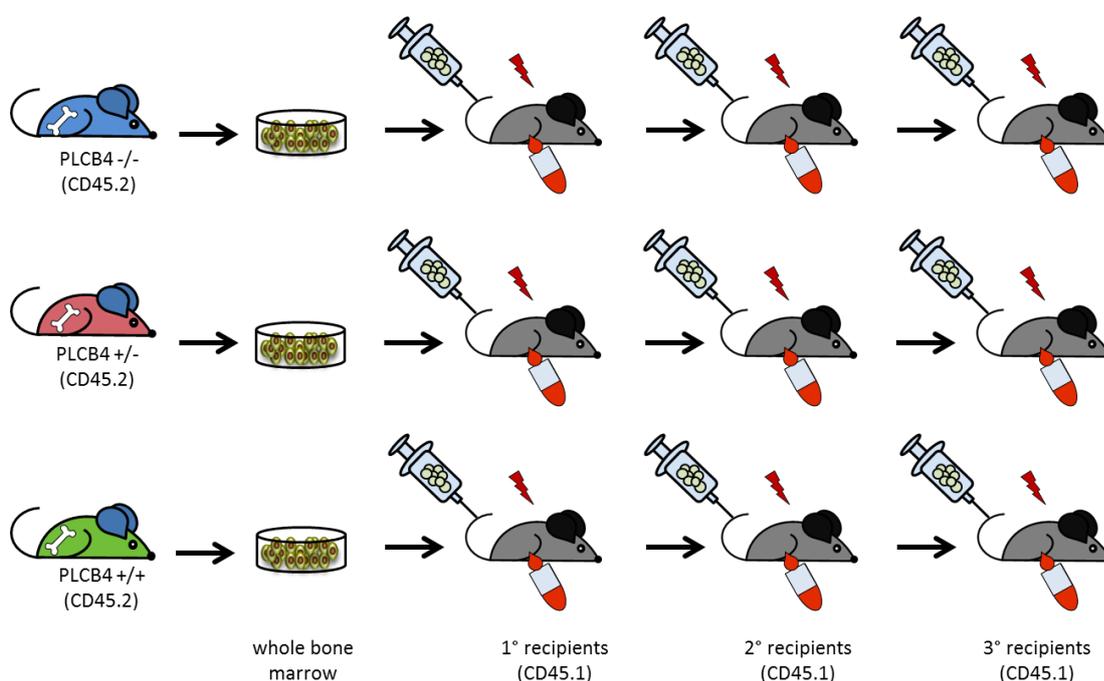


Figure 39: Transplantation of BM cells obtained from $-/-$, $+/-$ and $+/+$ 40-day old PLCB4 mice into lethally irradiated recipients. Whole BM cells were harvested from $-/-$, $+/-$ and $+/+$ PLCB4 mice, 10^6 BM CD45.2 cells from PLCB4 mice and 10^5 CD45.1 carrier wBM cells were transplanted into lethally irradiated CD45.1 recipient mice. The chimerism of donor-derived cells in the transplanted mice PB and their contribution in hematopoiesis was followed up every 4 weeks. $-/-$: knockout; $+/-$: heterozygote; $+/+$: wild type PLCB4 mice.

10^7 BM cells from 40 day-old $-/-$, $+/-$ and $+/+$ PLCB4 mice were transplanted into irradiated recipients to evaluate multilineage engraftment and self-renewal capacity of these cells during serial transplantation. We detected no significant difference in long term multilineage reconstitution in the primary recipients 6 months after BMT (Figure 40).

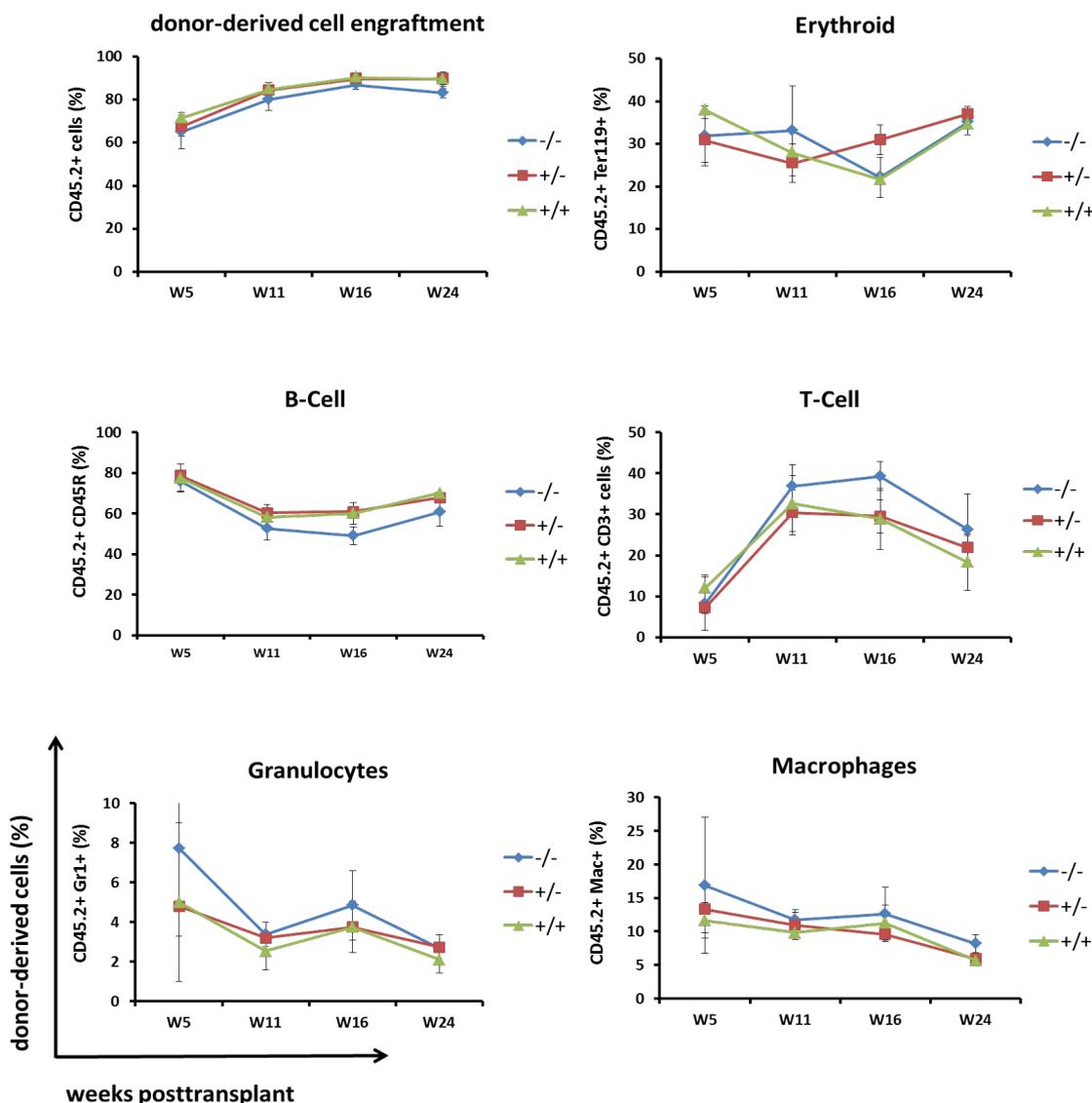


Figure 40: Long-term multilineage reconstitution of transplanted PLCB4-knockout BM cells in primary recipient mice. Long term engraftment and hematopoietic differentiation of BM cells from PLCB4 $-/-$ mice compared to BM cells from $+/-$ and $+/+$ was evaluated in primary transplanted mice in 6 months. $-/-$: knockout; $+/-$: heterozygote; $+/+$: wild type PLCB4 mice.

In summary, analyzing BM cells from PLCB4 KO mice at different ages *in vitro* revealed distinct differences in the frequency of LSK and pre-B colony forming cells in 18 day old mice which were not observed in 40-day old mice.

Moreover, we performed *in vivo* analysis by transplanting BM cells from 40-day old PLCB4 $-/-$, $+/-$ and $+/+$ mice into lethally irradiated recipients. We did not detect a significant alteration in the engraftment rate and differentiation to multiple hematopoietic lymphoid and myeloid lineages in the primary transplants.

5. Discussion

5.1. The integration site repertoire from gene therapy patients can be used for selecting candidate hematopoietic regulatory genes

In this study we established a strategy for selecting novel hematopoietic regulatory candidate genes based on systematically investigating the genome-wide retroviral vector (RV) integration site (IS) repertoire obtained from 10 gene therapy patients with Wiskott Aldrich Syndrome (WAS). The IS repertoire in gene therapy patients is usually analyzed to monitor the clonal dynamics after gene therapy and to detect clonal selection in individual patients. We hypothesized that re-analysis of this unique dataset can reveal candidate genes involved in the regulation of hematopoiesis.

The IS repertoires of WAS and other diseases such as X-linked severe combined immunodeficiency (X-SCID) and X-linked chronic granulomatous disease (X-CGD) treated with retroviral vectors (RV) have so far been studied. These studies assessed the vector-associated clonal genotoxicity, common integration sites and *in vivo* selection [108, 133-136]. All these investigations showed that a clustering of IS in certain genomic loci was detectable, and possibly provided a selective advantage for the affected cell clones. Moreover, it was indicated that RV integrate into genomic DNA in a non-random pattern and with a preference for active genes and regulatory regions [100, 137].

To identify the novel hematopoietic regulators among 3.268 genes which were annotated in vicinity of IS, we analyzed the IS from all 10 WAS gene therapy patients in detail. We strictly selected single genes harboring 10 or more than 10 IS in a 50 kb window from their transcription start site (TSS). Moreover, we showed that clones associated to 32 selected genes were durable in a 4-year time period after gene therapy. By tracing each individual clone marked with a unique IS, we detected some clones at multiple time points within 4 years indicating their long term clonal activity and highlighting their role as potential hematopoietic stem cell regulators.

Our group previously analyzed IS repertoire from WAS patients to evaluate the long-term efficacy and genotoxicity of RV used in WAS gene therapy trial [97]. Studying the RV integration pattern from each individual patient, they identified the genes with the highest number of IS in a 200 kb window from their TSS [97]. However, by analyzing the combined dataset from all patients, we set the 10 IS cutoff in vicinity of each gene to increase the

probability of selecting candidate genes from all 10 WAS patients. Moreover, in other studies all the IS located in a 200 kb window from the TSS of neighboring genes were included in the analyses [136, 138]. It is known that the long tandem repeats (LTR) region of RV can impact the expression of neighboring genes which can be quite distant (e.g. 300 kb) due to its strong internal enhancer and promoter [124]. However, TSS of genes such as CCND2, LMO2 and EVI1 which were shown to become biologically activated after bone marrow transplantation (BMT) were located within a 50 kb window from IS [97], therefore in our analysis we restricted the distance between IS and TSS to 50 kb to increase the probability of selecting genes which were more likely to get transcriptionally activated.

Furthermore, we observed that some of the IS clusters showed frequent and close occurrence to multiple genes. Therefore, to increase specificity on the expense of sensitivity, we only selected single genes which were not located in gene-rich areas. In previous studies the IS repertoire of RV was screened to investigate the clonal dynamics and not to select the hematopoietic candidate genes with possible regulatory effects on hematopoiesis.

Constitutively activated genes in vicinity of RV integrations can cause adverse effects after BMT [124]. It was observed that years after BMT in WAS patients, clones carrying the IS in vicinity of specific genes such as LMO2 and MDS1/EVI1 became dominant over other clones due to insertional mutagenesis. Activation of LMO2, which is one of the most frequently detected loci, leads to T-cell acute lymphoblastic leukemia (T-ALL). MDS1/EVI1 activation is involved in the development of acute myeloid leukemia (AML) [107, 115, 136, 139, 140]. However, besides clones with activated LMO2 or MDS1/EVI1 and more recently MN1 [97], we do not have evidence for malignant transformation of other clones.

A strong validation for our established selection strategy was detecting 20 out of 32 genes, such as MEIS1, CCND3, FLI1, XBP1, and FOXO1, which were previously known to have function in hematopoiesis regulation [141-149]. To have a wide range of genes covering various cellular functions, 5 out of 12 novel highly-ranked candidate genes (ZNF217, LRRC33, EVL, IRF2BPL, and PLCB4) from different protein families were selected for further functional analyses. Interestingly, ZNF217, PLCB4 and IRF2BPL were also reported in the IS analysis of X-SCID gene therapy trial [136].

Since the WAS gene therapy patients and patients from other gene therapy trials are still being followed up by analyzing their RV integration pattern over time, the systematic selection strategy that we successfully established in this thesis can be used to identify the novel hematopoietic regulatory genes in meta-datasets from a larger IS repertoire containing the most recent detected IS.

5.2. Evaluation of the candidate genes' expression in defined hematopoietic stem and progenitor cell populations

5.2.1. Efficient purification of hematopoietic stem cell population using CD34⁻CD48⁻CD135⁻CD150⁺LSK surface markers

The expression of pre-selected candidate genes in murine hematopoietic stem and progenitor cells (HSC and HPC) populations might point to their regulatory role in hematopoiesis. In order to evaluate the expression of selected candidate genes in these populations, obtaining highly purified HSC and HPC is necessary. So far, various combinations of cell surface markers for isolating these cells have been described [3, 31, 38, 39, 150, 151]. To obtain highly enriched populations, we examined different combinations of cell surface markers and functionally determined their purity in BMT experiments.

All examined phenotypes of HSC showed long term hematopoietic activity in primary and secondary recipients. However, by transplanting the same number of cells, CD34⁻CD48⁻CD135⁻CD150⁺LSK bone marrow (BM) cells exhibited higher hematopoietic multilineage engraftment in primary recipient mice and an efficient activity of donor-derived cells during serial transplantation. Our results showed the same multilineage donor-derived reconstitution for CD34⁻CD48⁻CD135⁻CD150⁺LSK population which was observed in other studies [6, 39]. Therefore this phenotype of HSC can be further studied for the expression of the selected candidate genes.

5.2.1.1. All five candidate genes are expressed at least in one of hematopoietic stem an progenitor cell populations

We asked whether the selected candidate genes are involved in HSC and HPC regulation, thus we tested their expression in HSC and HPC populations kindly provided by Dr. Trumpp (Cabezas-Wallscheid et al. submitted). Since different MPP populations have been shown to be either lymphoid or myeloid biased [6, 38, 152], high expression level or a significant change in any of these populations might indicate an important role for the candidate genes in the respective HSC or HPC populations.

ZNF217, one of the candidate genes, showed lower expression in HSC. Its expression level increased towards more mature MPP which might suggest for its more significant roles in these hematopoietic progenitor cells (MPP3 and MPP4).

ZNF217 was demonstrated that ZNF217 functions as a transcriptional repressor that associates with a core group of proteins including the co-repressors C-terminal binding protein (CtBP1) and CoREST (corepressor with REST, RE1 silencing transcription factor)[153].

Banck et al. suggested that ZNF217 assembles a distinct set of histone modifying proteins at target DNA sites that act synergistically in transcriptional repression [154-157]. Vendrell et al. reported a prognostic value of ZNF217 mRNA expression levels in numerous independent breast cancer cohorts [158].

DNA amplifications regions are often sites of oncogenes that give proliferative or survival advantages in cells harboring these amplifications in their genome. Amplifications involving 20q13 region encoding for ZNF217 in human are detected in a number of different cancer types, such as breast, ovarian, and squamous cell cancers [159, 160] [161-163]. The amplification of ZNF217 has been shown to correlate with tumor size in gastric cancer [164, 165]. However, ZNF217 was never studied for its impact on hematopoiesis regulation.

Interestingly, the overall RNA expression of LRRC33, the second candidate gene, in examined HSC and MPP1-4 populations was between 7 to 79 times higher than other candidate genes and in the same range of *Oaz1* and *Sdha* housekeeping genes. This could indicate a substantial role for LRRC33 in HSC and HPC. In addition, LRRC33 expression increased towards MPP3 and MPP4 which might suggest its prominent function in more differentiated hematopoietic progenitor cells. Liu et al. examined LRRC33 RNA expression in different mouse tissues and organs. They observed that LRRC33 was highly expressed in bone marrow, thymus, liver, lung, intestine, and spleen. These results could suggest that LRRC33 might play a role in the development and maintenance of homeostasis of these organs [130].

EVL expression in transcriptome analyses of different HSC and MPP populations showed almost no RNA expression in HSC, MPP1, MPP2 and MPP3 while it exhibited a drastic 11-fold increase in MPP4. MPP4 was already shown to be functionally lymphoid-biased [6, 166]. In addition, it was indicated that the average expression of EVL in mouse lymphoid lineages is 14.5 fold higher than myeloid lineages [167]. Therefore, EVL might have an effect on MPP4 differentiation towards B- and T-cells.

Human EVL is a member of ENA/VASP protein family and is involved in the process of actin remodeling [168]. Actin regulatory proteins have a major function in neuronal migration. Ena/VASP proteins bind actin and control the assembly of F-actin networks by disturbing the function of capping proteins and bundling actin filaments. They are necessary for formation of filopodia in various cell types such as fibroblasts and neuronal growth cones [169-172]. It is reported that EVL also plays a role in homologous recombination. Therefore, EVL can have a dual function in actin remodeling in cytoplasm and homologous recombination in nucleus [173]. Despite all these findings on EVL function, its role in hematopoiesis is yet to be studied.

We observed that PLCB4 was the only candidate gene with a descending expression pattern from long term HSC towards more mature HPC populations. Interestingly, PLCB4 RNA expression in human BM and cord blood from healthy donors exhibited the same pattern (Simon Raffel, personal communication). This might imply that PLCB4 plays a regulatory role in long term HSC. Beside its abundant expression in retina, PLCB4 has been shown to be mainly expressed in several limited regions in central nervous system including Purkinje cells [174]. Although the expression of PLCB4 in neural stem cells was not specifically studied, it was shown that in the developing mouse brain at embryonic stages PLCB4 express in areas which produce the most neurons and microglia [175]. Moreover, PLCB4 expression level decreased when myoblasts were differentiated towards myotubes *in vitro* [176]. This indicates that PLCB4 might play a role in other types of stem cells as well as HSC. Besides normal cells, PLCB4 expression has been shown to have a 4-fold increase in leukemic stem cell (LSC) compared to non-LSC fraction (Simon Raffel, personal communication).

PLCB4 protein is one of the phospholipase C (PLC) enzymes which hydrolyzes the inner membrane component of phosphatidylinositol-4, 5-bisphosphate (PIP2) and generate the second messengers inositol-1,4,5-triphosphate (IP3) and diacylglycerol (DAG). IP3 binds to IP3-specific receptors and causes the release of intracellular Ca^{2+} . DAG remains associated to the membrane and activates protein kinase C with increasing Ca^{2+} level [177]. This procedure leads to the regulation of numerous physiological processes such as muscle contraction [178], chemotaxis [179], opioid sensitivity [180] and cell proliferation and survival [181, 182]. However, the role of PLCB4 in hematopoiesis was never reported.

The fifth candidate gene, IRF2BPL, reveals a steady level of RNA expression in all HSC and HPC indicating that its expression is not specific to any of these populations. However, since the expression amount is considerable the role of IRF2BPL in hematopoiesis regulation should be further studied.

Less than 5% of human genes are intronless, IRF2BPL is also an intronless gene and does not need posttranscriptional splicing; therefore it can be efficiently transcribed with higher level of protein expression. Furthermore, intronless genes are not likely to undergo aberrant splicing, thus leading to greater transcriptional fidelity [183]. It is demonstrated that this gene is a transcriptional factor which functions within neuronal networks of the neuroendocrine brain and is required for the appropriate initiation of female puberty and maintenance of female reproductive cyclicity. Based on these findings, this gene was also designated as enhanced at puberty 1 (EAP1) [184]. However, the role of IRF2BPL in hematopoiesis and blood formation process was never addressed before.

In brief, all candidate genes are expressed at least in one of the examined HSC and HPC populations which might point to their importance in hematopoiesis. Performing *in vitro* and

in vivo analyses will allow gaining a better understanding to the biological function of the selected candidate genes.

5.3. Investigating the role of candidate genes *in vitro* and *in vivo*

5.3.1. Single LSK ($\text{Lin}^- \text{Sca-1}^+ \text{c-Kit}^+$) cells show different growth kinetics under stem cell condition *in vitro*

To investigate the role of selected candidate genes on proliferation and growth kinetics of HSC and HPC populations *in vitro*, we first cultured un-manipulated single LSK-cell-derived clones under stem cell conditions *in vitro*. We observed that all living single-cell derived clones showed the same growth kinetics in the first 10 days. These cells started exhibiting different proliferation kinetics in a period between day 10 and day 90. We categorized the clones into 4 groups based on their life span and proliferation kinetics. This is due to the heterogeneity of LSK population comprising long and short term repopulating hematopoietic stem cells [39]. Although all mouse HSC and HPC appear to have the $\text{Lin}^- \text{Sca1}^+ \text{c-Kit}^+$ phenotype, all cells with this phenotype are not HSC or HPC [185, 186]. Some cells only provide short-term hematopoietic repopulation and some have no demonstrable stem cell activity [187].

After optimizing single LSK cell culture *in vitro*, we aimed to know if single-LSK cell-derived colonies overexpressing selected candidate genes could be used as an appropriate model to detect candidate genes induced changes in survival and proliferation. We first performed the experiment using cyclin D2 (CCND2), as a positive control. This cyclin has a critical role in expansion of hematopoietic cells by functioning as a regulatory subunit of cyclin dependent kinase (CDK) 4 or CDK6, whose activity is required for cell cycle G1/S transition [125]. We observed that the proliferation kinetics of CCND2-transduced LSK cells did not show a significant difference compared to their mock counterpart. Although transduced-single LSK cells were successfully cultured *in vitro*, this model does not seem to be appropriate to assess the impact of transgenes on proliferation and growth kinetics. This might be due to the unstable expression of CCND2 in the initially transduced cells.

5.3.2. ZNF217 overexpression increased survival and proliferation of hematopoietic cells *in vitro*

5.3.2.1. Hematopoietic cells overexpressing ZNF217 show increased growth activity

Compared to GFP-transduced control cells we detected a significant 1.9-fold increase in ATP activity of ZNF217-transduced hematopoietic C1498 cells. ZNF217 has been shown to function as a proto-oncogene in several cancers such as breast and ovarian carcinoma [156, 188-190]. However, its role in normal or malignant hematopoiesis has never been addressed. To test the effect of ZNF217 overexpression on the viability of hematopoietic cells, we analyzed hematopoietic C1498 cells overexpressing ZNF217 by measuring their ATP activity level. Since ATP is an indicator of metabolically active and viable cells [126], this increase can be due to either more viable or more active cells.

In addition to higher ATP activity level, our results indicated increased proliferation and growth potential in hematopoietic cells overexpressing ZNF217. To further test if ZNF217 leads to a higher proliferation rate in hematopoietic cells, we examined the proliferation kinetics of ZNF217-transduced FDCP-Mix cells compared to GFP-transduced control cells. ZNF217 overexpressing cells showed a 1.4-fold increased proliferation 4 days after seeding in comparison to GFP-transduced control cells. This is in accordance with previous findings in non-hematopoietic cells showing that the proliferation ability of ovarian carcinoma ZNF217-transfected cells was significantly higher than control cells [189, 191]. However, the detailed mechanism of ZNF217 on the proliferation of hematopoietic cells remains to be determined.

5.3.2.2. Stable ZNF217 expression induces IL3 independency of FDCP-Mix and 32D cells

We observed that 32D and FDCP-Mix hematopoietic cells which are highly interleukin 3 (IL3) dependent showed signs of transformation by acquiring the ability to become independent of IL3 when transduced with the LV encoding for ZNF217.

The IL3 independency was examined and proved by performing two separate experiments on 32D and FDCP-Mix cells. Two weeks after plating 32D cells overexpressing ZNF217 in IL3-free semi-solid medium, IL3-independent colonies were formed, whereas GFP-transduced control cells formed no colonies. The enrichment for ZNF217 overexpressing cells from 33%-93% in the absence of IL3 indicated that only these cells could survive and proliferate without this cytokine. Hematopoietic 32D cells used in this experiment have previously served as a model cell system for studying the transforming properties of various leukemia-

associated oncogenes such as Bcr/Abl [192, 193]. When 32D cells are cultured in the absence of IL3, they undergo cell cycle arrest and apoptosis [127].

Our findings were in accordance with Littlepage et al. results demonstrating ZNF217 overexpression contributes to immortalization in both breast cancer cells and *in vivo* mouse transplant models [190]. We observed the same frequency of cytokine-independent cell colonies in ZNF217 compared to the positive control c-Myc, which is a known proto-oncogene [128]. We conclude that ZNF217 overexpression might lead to transformation in hematopoietic cells.

5.3.2.3. Expression profiling of ZNF217 overexpressing hematopoietic cells indicated deregulation of pathways involved in cellular movement

Since ZNF217 has been described as a transcriptional regulator [157], we performed global gene expression profiling on ZNF217-transduced 32D hematopoietic cells to detect possible target genes and subsequently deregulated pathways. We observed that the number of downregulated genes was 4 times higher than upregulated genes in hematopoietic 32D cells overexpressing ZNF217 which confirms its previously known role as a transcription repressor [155, 157].

We further observed that the networks deregulated in these cells are mainly involved in cellular movement. The role of ZNF217 in promoting epithelial-mesenchymal transition (EMT) in human mammary epithelial cells was previously shown [158, 190]. It was demonstrated that the frequency of increased ZNF217 copy number is higher in liver metastasis originated from colorectal cancer than in the primary colorectal tumors. None of the colorectal cancers which did not give rise to liver metastasis showed ZNF217 copy number amplifications [194]. However, in our gene expression profile cell movement networks affected by ZNF217 overexpression are downregulated. This might introduce the opposite effect of ZNF217 in hematopoietic cells.

5.3.3. EVL overexpression increases cell viability and colony forming ability

To assess the effect of EVL on hematopoietic cell viability and clonogenicity, hematopoietic C1498 cells transduced with the LV encoding for EVL were analyzed. We observed that in addition to higher ATP activity, EVL transduced C1498 cells formed more colonies compared to GFP transduced control cells in colony forming assay. So far, EVL has been reported to play roles in homologous recombination, DNA repair, actin remodeling and cell motility [168-173].

To investigate whether the overexpression of EVL in hematopoietic cells led to deregulation of other genes and if these genes were involved in specific cellular pathways, global gene expression profiling of EVL overexpressing 32D cells was performed. We observed that EVL overexpression did not lead to any significant deregulation of the transcription of other genes. This can be due to the fact that EVL is not a transcription regulator molecule; therefore its overexpression does not affect up or downregulation of numerous genes at RNA level and might have post-transcriptional effects.

5.3.4. Analyzing the effect of LRRC33 *in vivo* and *in vitro*

5.3.4.1. LRRC33 inhibits NF- κ B activity through toll-like receptor 4

LRRC33 is a member of LRR (Leucine-rich repeat) protein family and subgrouped in LRR_Tollkin (clustering with the Toll proteins) [195]. Its protein structure is similar to toll-like receptor (TLR) proteins [129, 130]. Therefore, we tested whether LRRC33 can function as a TLR co-receptor and affect NF- κ B activity when the cells were stimulated with lipopolysaccharides (LPS). We observed that LRRC33 decreased NF- κ B activity via TLR4. 293T cells transfected with TLR4 and murine LRRC33 showed reduced level of NF- κ B activity compared to cells transfected with TLR4 and GFP after LPS stimulation. Our data on murine LRRC33 suggested the same reported inhibitory effect of human LRRC33 indicating that this protein might have an effect on reducing pro-inflammatory signals [130], which can be one of the proposed roles for LRRC33 in innate immunity according to its protein structure.

NF- κ B has been shown to be active in various immune cells as well as non-hematopoietic cells and to function as a critical player in coordinating inflammation and immune cell functions [196]. However, the role of NF- κ B in HSC has not been profoundly explored. It was shown that the function of microRNA 146a, which downregulates NF- κ B, is necessary for maintaining HSC homeostasis and longevity under steady state conditions in murine hematopoietic system. This implies that NF- κ B inhibition might be important in maintenance of HSC pool and homeostasis of blood cells [197, 198]. Therefore further analyses of LRRC33 role under stressed and steady state conditions can clarify how LRRC33 functions in hematopoiesis regulation.

5.3.4.2. LRRC33 influences macrophage and T-cell differentiation *in vivo*

In order to study the effect of the selected genes on long term multilineage reconstitution and self-renewal of LSK cells *in vivo*, we transduced LSK cells with the LV of LRRC33 and transplanted the cells into recipient mice. Analyzing the role of LRRC33 *in vivo*, we observed that LRRC33-overexpressing LSK cells show higher differentiation towards macrophages and

lower differentiation towards T-cells compared to control-transduced cells in a small cohort of 10 recipient mice. Since it was shown that NF- κ B pathway plays roles in myeloid or lymphoid lineage development [199, 200] and considering our previous finding of LRRC33 inhibitory role on NF- κ B activity *in vitro*, we conclude that this difference in T-cell and macrophage differentiation might be due to the changes in NF- κ B activity.

The role of NF- κ B family members in different stages of T-cell development was subsequently evaluated in detail using several transgenic mouse models for different members of NF- κ B protein family. For instance, knocking out one of the NF- κ B protein family members, RelB, inhibits T-cells development to CD4⁻CD8⁺ cells in peripheral blood [199]. Therefore, performing further *in vitro* analyses to study the effect of LRRC33 on hematopoietic cell differentiation towards myeloid and lymphoid lineages is suggested. In addition, LRRC33- transduced LSK cells should be transplanted into larger number of recipient mice to further validate the detected lineage bias.

5.3.5. Analyzing PLCB4 function *in vitro* and *in vivo* using PLCB4 knockout mice

The role of PLCB4 in hematopoietic system has been never profoundly investigated. Analyzing RNA expression level of PLCB4 revealed its higher expression in HSC and MPP1 populations. In this study we used PLCB4 knockout mice to determine the role of PLCB4 in hematopoiesis. Moreover, since the coding sequence of PLCB4 gene is large (3.5 kb), producing LV of PLCB4 at high titer was challenging and studying its overexpression effect on hematopoietic cells could not be achieved efficiently.

5.3.5.1. Eighteen-day old PLCB4 knockout mice show decreased level of LSK cell count

Analyzing PLCB4 knockout mice [201], we observed that breeding +/- mice led to 11% of 21-day old -/- progenies which was 14% less than expected mendelian ratio. Moreover, the -/- mice are clearly smaller than +/- or +/+ mice and show retarded body growth. Since we could not detect a significant difference in the death rate of newborn and younger than 21-day old pups, and the newborn -/- pups have lower body weight, we conclude that there was a significantly higher death rate of fetuses before birth. Therefore, genotyping fetuses will clarify the death of -/- PLCB4 mice at embryonic stages. Since we previously observed signs of hematopoietic abnormalities in mice at young ages, we can further study if hematopoietic deficiencies are involved in the death of mice before birth.

The function of Phospholipase C beta4 has been well studied in central nervous system [202, 203]. Nuclear PLCB1, another member of phospholipase proteins, has been shown to play an

important role in controlling the balance between cell cycle progression and apoptosis in hematopoietic myelodysplastic syndrome cells [204]. Okada et al. performed a genome-wide single nucleotide polymorphism (SNP) analysis on patients with various diseases such as different types of cancers and myocardial infarction. In their study the subjects who were homozygous for neutrophil-increasing alleles in both of the SNPs had 1.17-fold higher neutrophil counts when compared with the subjects homozygous for neutrophil-decreasing alleles. They demonstrated that PLCB4 locus is significantly associated with the regulation of neutrophil count [205].

We observed that BM cells from 18-day old mice contained 2-fold less LSK cells compared to 40-day old mice. Moreover, a 2-fold decrease in the number of preB-cell colonies was detected in the BM cells from -/- vs. +/- , and -/- vs. +/+ 18 and 25-day old mice compared to preB-cell colony count in 40-day old mice. Since preB-cell colony forming cell (CFC) assay was performed with the whole BM cells, which we showed contain less LSK cells, the reduction observed in pre-B colonies could be explained by lower number of LSK cells which are responsible for giving rise to pre-B cells. Unlike preB-cell CFC assay, myeloid colonies detected in myeloid CFC did not exhibit any significant differences in the number and type of various myeloid colonies in -/-, +/- and +/+ PLCB4 mice, however myeloid colony forming potential in younger mice and fetuses are still to be investigated.

These results strongly suggest examining the hematopoiesis in the BM of newborn and younger than 18 day-old mice. Since the process of HSC development during embryogenesis engages various anatomical sites such as yolk sac, placenta, and fetal liver [206-208], analyzing hematopoietic system of PLCB4 knockout mice at early embryonic stages can elucidate the role of PLCB4 protein in various hematopoietic developmental stages.

In addition to *ex vivo* assays, BM cells from 40 day-old -/-, +/- and +/+ PLCB4 mice were transplanted into irradiated recipients to evaluate multilineage engraftment and self-renewal during serial transplantation. No significant difference in long term multilineage reconstitution was observed in the primary recipients 6 months after BMT, however secondary and tertiary recipient mice are still to be evaluated to study the self-renewal ability of BM cells derived from PLCB4 KO mice. It was demonstrated that BM from young mice has different potential in multilineage reconstitution compared to older mice [209, 210]. For instance, older mice BM have been shown to contain more HSC by measuring their competitive repopulating ability *in vivo* [211]. However, according to our *in vitro* analyses, 18-day old -/- mice contain less HSC compared to 40-day old mice which can be due to the effect of PLCB4 candidate gene. Therefore, performing limiting dilution in competitive repopulating assays in young and old mice can clarify the role of PLCB4 in long-term repopulation and multilineage reconstitution in hematopoiesis.

In summary, the BM from PLCB4 KO mice was shown to contain less LSK cells compared to wild type mice. Since this difference could be only detected in 18-day old mice and not older mice, analyzing the hematopoietic system in adult mice at young ages and murine embryos will unravel the role of PLCB4 in hematopoiesis.

Conclusion

In this study we demonstrated that clinical integration site datasets from gene therapy patients can be used to identify regulatory genes of hematopoiesis. We developed a strategy to select hematopoietic candidate regulatory genes by analyzing the genome wide IS repertoire from 10 WAS gene therapy patients. Following the selection of various candidate genes, we examined whether these genes play a role in the regulation of hematopoiesis. Performing various *in vitro* and *in vivo* analyses to address proliferation, differentiation and self-renewal, we demonstrated that four out of five candidate genes have an influence on hematopoietic primary cells. This clearly validates the established selection strategy and proves its significance. The identified genes will be further evaluated for their hematopoietic stem cell regulatory potential on molecular level. We could clearly show that this strategy can be applied for protein-coding genes and we aim to extend the identification of novel non-coding regulators in future. The systematic identification of novel hematopoietic regulatory genes in meta-datasets derived from a large number of gene therapy studies and subsequent *in vitro* and *in vivo* validation will allow gaining new insights into the biology of post-embryonic hematopoiesis.

References

1. Weissman, I.L., D.J. Anderson, and F. Gage, *Stem and progenitor cells: origins, phenotypes, lineage commitments, and transdifferentiations*. *Annu Rev Cell Dev Biol*, 2001. **17**: p. 387-403.
2. Benz, C., et al., *Hematopoietic stem cell subtypes expand differentially during development and display distinct lymphopoietic programs*. *Cell Stem Cell*, 2012. **10**(3): p. 273-83.
3. Dykstra, B., et al., *Long-term propagation of distinct hematopoietic differentiation programs in vivo*. *Cell Stem Cell*, 2007. **1**(2): p. 218-29.
4. Muller-Sieburg, C.E., et al., *Stem cell heterogeneity: implications for aging and regenerative medicine*. *Blood*, 2012. **119**(17): p. 3900-7.
5. Becker, A.J., C.E. Mc, and J.E. Till, *Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells*. *Nature*, 1963. **197**: p. 452-4.
6. Adolfsson, J., et al., *Identification of Flt3+ Lympho-Myeloid Stem Cells Lacking Erythro-Megakaryocytic Potential: A Revised Road Map for Adult Blood Lineage Commitment*. *Cell*, 2005. **121**(2): p. 295-306.
7. Akashi, K., et al., *A clonogenic common myeloid progenitor that gives rise to all myeloid lineages*. *Nature*, 2000. **404**(6774): p. 193-7.
8. Reya, T., et al., *Stem cells, cancer, and cancer stem cells*. *Nature*, 2001. **414**(6859): p. 105-11.
9. Katsura, Y. and H. Kawamoto, *Stepwise lineage restriction of progenitors in lympho-myelopoiesis*. *Int Rev Immunol*, 2001. **20**(1): p. 1-20.
10. Yamamoto, R., et al., *Clonal Analysis Unveils Self-Renewing Lineage-Restricted Progenitors Generated Directly from Hematopoietic Stem Cells*. *Cell*, 2013. **154**(5): p. 1112-1126.
11. Dykstra, B., et al., *Clonal analysis reveals multiple functional defects of aged murine hematopoietic stem cells*. *J Exp Med*, 2011. **208**(13): p. 2691-703.
12. Lai, A.Y. and M. Kondo, *Asymmetrical lymphoid and myeloid lineage commitment in multipotent hematopoietic progenitors*. *J Exp Med*, 2006. **203**(8): p. 1867-73.
13. Till, J.E. and C.E. Mc, *A direct measurement of the radiation sensitivity of normal mouse bone marrow cells*. *Radiat Res*, 1961. **14**: p. 213-22.
14. Jones, R.J., et al., *Characterization of mouse lymphohematopoietic stem cells lacking spleen colony-forming activity*. *Blood*, 1996. **88**(2): p. 487-91.
15. Ploemacher, R.E. and R.H. Brons, *Separation of CFU-S from primitive cells responsible for reconstitution of the bone marrow hemopoietic stem cell compartment following irradiation: evidence for a pre-CFU-S cell*. *Exp Hematol*, 1989. **17**(3): p. 263-6.
16. Copley, M.R., P.A. Beer, and C.J. Eaves, *Hematopoietic stem cell heterogeneity takes center stage*. *Cell Stem Cell*, 2012. **10**(6): p. 690-7.
17. Siminovitch, L., E.A. McCulloch, and J.E. Till, *THE DISTRIBUTION OF COLONY-FORMING CELLS AMONG SPLEEN COLONIES*. *J Cell Physiol*, 1963. **62**: p. 327-36.
18. Till, J.E., E.A. McCulloch, and L. Siminovitch, *A STOCHASTIC MODEL OF STEM CELL PROLIFERATION, BASED ON THE GROWTH OF SPLEEN COLONY-FORMING CELLS*. *Proc Natl Acad Sci U S A*, 1964. **51**: p. 29-36.

19. Humphries, R.K., et al., *CFU-S in individual erythroid colonies derived in vitro from adult mouse marrow*. *Nature*, 1979. **279**(5715): p. 718-20.
20. Ogawa, M., P.N. Porter, and T. Nakahata, *Renewal and commitment to differentiation of hemopoietic stem cells (an interpretive review)*. *Blood*, 1983. **61**(5): p. 823-9.
21. Dick, J.E., et al., *Introduction of a selectable gene into primitive stem cells capable of long-term reconstitution of the hemopoietic system of W/W^v mice*. *Cell*, 1985. **42**(1): p. 71-9.
22. Lemischka, I.R., D.H. Raulet, and R.C. Mulligan, *Developmental potential and dynamic behavior of hematopoietic stem cells*. *Cell*, 1986. **45**(6): p. 917-27.
23. Osawa, M., et al., *Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell*. *Science*, 1996. **273**(5272): p. 242-5.
24. Yang, L., et al., *Identification of Lin(-)Sca1(+)kit(+)CD34(+)Flt3- short-term hematopoietic stem cells capable of rapidly reconstituting and rescuing myeloablated transplant recipients*. *Blood*, 2005. **105**(7): p. 2717-23.
25. Jordan, C.T. and I.R. Lemischka, *Clonal and systemic analysis of long-term hematopoiesis in the mouse*. *Genes Dev*, 1990. **4**(2): p. 220-32.
26. Muller-Sieburg, C.E., et al., *Deterministic regulation of hematopoietic stem cell self-renewal and differentiation*. *Blood*, 2002. **100**(4): p. 1302-9.
27. Okada, S., et al., *In vivo and in vitro stem cell function of c-kit- and Sca-1-positive murine hematopoietic cells*. *Blood*, 1992. **80**(12): p. 3044-50.
28. Bryder, D., D.J. Rossi, and I.L. Weissman, *Hematopoietic stem cells: the paradigmatic tissue-specific stem cell*. *Am J Pathol*, 2006. **169**(2): p. 338-46.
29. Arai, F., et al., *Tie2/angiopoietin-1 signaling regulates hematopoietic stem cell quiescence in the bone marrow niche*. *Cell*, 2004. **118**(2): p. 149-61.
30. Balazs, A.B., et al., *Endothelial protein C receptor (CD201) explicitly identifies hematopoietic stem cells in murine bone marrow*. *Blood*, 2006. **107**(6): p. 2317-21.
31. Kiel, M.J., et al., *SLAM Family Receptors Distinguish Hematopoietic Stem and Progenitor Cells and Reveal Endothelial Niches for Stem Cells*. *Cell*, 2005. **121**(7): p. 1109-1121.
32. Goodell, M.A., et al., *Isolation and functional properties of murine hematopoietic stem cells that are replicating in vivo*. *J Exp Med*, 1996. **183**(4): p. 1797-806.
33. Weksberg, D.C., et al., *CD150- side population cells represent a functionally distinct population of long-term hematopoietic stem cells*. *Blood*, 2008. **111**(4): p. 2444-51.
34. Lin, K.K. and M.A. Goodell, *Detection of hematopoietic stem cells by flow cytometry*. *Methods Cell Biol*, 2011. **103**: p. 21-30.
35. Camargo, F.D., et al., *Hematopoietic stem cells do not engraft with absolute efficiencies*. *Blood*, 2006. **107**(2): p. 501-7.
36. Uchida, N., et al., *Different in vivo repopulating activities of purified hematopoietic stem cells before and after being stimulated to divide in vitro with the same kinetics*. *Exp Hematol*, 2003. **31**(12): p. 1338-47.
37. Wagers, A.J., R.C. Allsopp, and I.L. Weissman, *Changes in integrin expression are associated with altered homing properties of Lin(-/lo)Thy1.1(lo)Sca-1(+)c-kit(+) hematopoietic stem cells following mobilization by cyclophosphamide/granulocyte colony-stimulating factor*. *Exp Hematol*, 2002. **30**(2): p. 176-85.
38. Wilson, A., et al., *Dormant and self-renewing hematopoietic stem cells and their niches*. *Ann N Y Acad Sci*, 2007. **1106**: p. 64-75.
39. Wilson, A., et al., *Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair*. *Cell*, 2008. **135**(6): p. 1118-29.

40. Rieger, M.A. and T. Schroeder, *Exploring hematopoiesis at single cell resolution*. Cells Tissues Organs, 2008. **188**(1-2): p. 139-49.
41. Yin, T. and L. Li, *The stem cell niches in bone*. J Clin Invest, 2006. **116**(5): p. 1195-201.
42. Zhu, J. and S.G. Emerson, *Hematopoietic cytokines, transcription factors and lineage commitment*. Oncogene, 2002. **21**(21): p. 3295-313.
43. Palani, S. and C.A. Sarkar, *Integrating extrinsic and intrinsic cues into a minimal model of lineage commitment for hematopoietic progenitors*. PLoS Comput Biol, 2009. **5**(9): p. e1000518.
44. Rieger M.A. and Schroeder, T., *Hämatopoetische Stammzellen*. Biospektrum, 2007. **3**: p. 254-256.
45. Enver, T., C.M. Heyworth, and T.M. Dexter, *Do stem cells play dice?* Blood, 1998. **92**(2): p. 348-51; discussion 352.
46. Fukuchi, Y., et al., *Comprehensive analysis of myeloid lineage conversion using mice expressing an inducible form of C/EBP alpha*. EMBO J, 2006. **25**(14): p. 3398-410.
47. Rosenbauer, F. and D.G. Tenen, *Transcription factors in myeloid development: balancing differentiation with transformation*. Nat Rev Immunol, 2007. **7**(2): p. 105-17.
48. Medvedovic, J., et al., *Pax5: a master regulator of B cell development and leukemogenesis*. Adv Immunol, 2011. **111**: p. 179-206.
49. Miyamoto, T., et al., *Myeloid or lymphoid promiscuity as a critical step in hematopoietic lineage commitment*. Dev Cell, 2002. **3**(1): p. 137-47.
50. Metcalf, D., *Hematopoietic cytokines*. Blood, 2008. **111**(2): p. 485-91.
51. Robb, L., *Cytokine receptors and hematopoietic differentiation*. Oncogene, 2007. **26**(47): p. 6715-23.
52. Batard, P., et al., *TGF-(beta)1 maintains hematopoietic immaturity by a reversible negative control of cell cycle and induces CD34 antigen up-modulation*. J Cell Sci, 2000. **113** (Pt 3): p. 383-90.
53. Foulds, L., *Neoplastic development*. New York: Academic Press. Vol. 1. 1969.
54. Foulds, L., *The experimental study of tumor progression: a review*. Cancer Res, 1954. **14**(5): p. 327-39.
55. Irons, R.D. and W.S. Stillman, *The process of leukemogenesis*. Environ Health Perspect, 1996. **104 Suppl 6**: p. 1239-46.
56. Harrison, C.J., *Cytogenetics of paediatric and adolescent acute lymphoblastic leukaemia*. Br J Haematol, 2009. **144**(2): p. 147-56.
57. Graux, C., et al., *Cytogenetics and molecular genetics of T-cell acute lymphoblastic leukemia: from thymocyte to lymphoblast*. Leukemia, 2006. **20**(9): p. 1496-510.
58. Harrison, C.J. and L. Foroni, *Cytogenetics and molecular genetics of acute lymphoblastic leukemia*. Rev Clin Exp Hematol, 2002. **6**(2): p. 91-113; discussion 200-2.
59. Mullighan, C.G., *Genomic analysis of acute leukemia*. Int J Lab Hematol, 2009. **31**(4): p. 384-97.
60. Greif, P.A., et al., *Identification of recurring tumor-specific somatic mutations in acute myeloid leukemia by transcriptome sequencing*. Leukemia, 2011. **25**(5): p. 821-7.
61. Gross, S., et al., *Cancer-associated metabolite 2-hydroxyglutarate accumulates in acute myelogenous leukemia with isocitrate dehydrogenase 1 and 2 mutations*. J Exp Med, 2010. **207**(2): p. 339-44.
62. Ley, T.J., et al., *DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome*. Nature, 2008. **456**(7218): p. 66-72.

63. Anderson, K., et al., *Genetic variegation of clonal architecture and propagating cells in leukaemia*. *Nature*, 2011. **469**(7330): p. 356-61.
64. Notta, F., et al., *Evolution of human BCR-ABL1 lymphoblastic leukaemia-initiating cells*. *Nature*, 2011. **469**(7330): p. 362-7.
65. Massaad, M.J., N. Ramesh, and R.S. Geha, *Wiskott-Aldrich syndrome: a comprehensive review*. *Ann N Y Acad Sci*, 2013. **1285**: p. 26-43.
66. Sullivan, K.E., *Recent advances in our understanding of Wiskott-Aldrich syndrome*. *Curr Opin Hematol*, 1999. **6**(1): p. 8-14.
67. Modell, V., et al., *Global study of primary immunodeficiency diseases (PI)—diagnosis, treatment, and economic impact: an updated report from the Jeffrey Modell Foundation*. *Immunologic Research*, 2011. **51**(1): p. 61-70.
68. Parolini, O., et al., *Expression of Wiskott-Aldrich syndrome protein (WASP) gene during hematopoietic differentiation*. *Blood*, 1997. **90**(1): p. 70-5.
69. Stewart, D.M., et al., *Studies of the expression of the Wiskott-Aldrich syndrome protein*. *J Clin Invest*, 1996. **97**(11): p. 2627-34.
70. Ochs, H.D. and A.J. Thrasher, *The Wiskott-Aldrich syndrome*. *J Allergy Clin Immunol*, 2006. **117**(4): p. 725-38; quiz 739.
71. Shcherbina, A., F.S. Rosen, and E. Remold-O'Donnell, *WASP levels in platelets and lymphocytes of wiskott-aldrich syndrome patients correlate with cell dysfunction*. *J Immunol*, 1999. **163**(11): p. 6314-20.
72. Burns, S., et al., *Maturation of DC is associated with changes in motile characteristics and adherence*. *Cell Motility and the Cytoskeleton*, 2004. **57**(2): p. 118-132.
73. Notarangelo, L.D., C.H. Miao, and H.D. Ochs, *Wiskott-Aldrich syndrome*. *Curr Opin Hematol*, 2008. **15**(1): p. 30-6.
74. Bosticardo, M., et al., *Recent advances in understanding the pathophysiology of Wiskott-Aldrich syndrome*. *Blood*, 2009. **113**(25): p. 6288-95.
75. Albert, M.H., L.D. Notarangelo, and H.D. Ochs, *Clinical spectrum, pathophysiology and treatment of the Wiskott-Aldrich syndrome*. *Curr Opin Hematol*, 2011. **18**(1): p. 42-8.
76. Lutskiy, M.I., F.S. Rosen, and E. Remold-O'Donnell, *Genotype-phenotype linkage in the Wiskott-Aldrich syndrome*. *J Immunol*, 2005. **175**(2): p. 1329-36.
77. Zhu, Q., et al., *Wiskott-Aldrich syndrome/X-linked thrombocytopenia: WASP gene mutations, protein expression, and phenotype*. *Blood*, 1997. **90**(7): p. 2680-9.
78. Wengler, G.S., et al., *High prevalence of nonsense, frame shift, and splice-site mutations in 16 patients with full-blown Wiskott-Aldrich syndrome*. *Blood*, 1995. **86**(10): p. 3648-54.
79. Lemahieu, V., J.M. Gastier, and U. Francke, *Novel mutations in the Wiskott-Aldrich syndrome protein gene and their effects on transcriptional, translational, and clinical phenotypes*. *Hum Mutat*, 1999. **14**(1): p. 54-66.
80. Albert, M.H., et al., *X-linked thrombocytopenia (XLT) due to WAS mutations: clinical characteristics, long-term outcome, and treatment options*. *Blood*, 2010. **115**(16): p. 3231-8.
81. Antoine, C., et al., *Long-term survival and transplantation of haemopoietic stem cells for immunodeficiencies: report of the European experience 1968-99*. *Lancet*, 2003. **361**(9357): p. 553-60.
82. Filipovich, A.H., et al., *Impact of donor type on outcome of bone marrow transplantation for Wiskott-Aldrich syndrome: collaborative study of the International Bone Marrow Transplant Registry and the National Marrow Donor Program*. *Blood*, 2001. **97**(6): p. 1598-603.

83. Ozsahin, H., et al., *Long-term outcome following hematopoietic stem-cell transplantation in Wiskott-Aldrich syndrome: collaborative study of the European Society for Immunodeficiencies and European Group for Blood and Marrow Transplantation*. *Blood*, 2008. **111**(1): p. 439-45.
84. Moratto, D., et al., *Long-term outcome and lineage-specific chimerism in 194 patients with Wiskott-Aldrich syndrome treated by hematopoietic cell transplantation in the period 1980-2009: an international collaborative study*. *Blood*, 2011. **118**(6): p. 1675-84.
85. Fischer, A., S. Hacein-Bey-Abina, and M. Cavazzana-Calvo, *Gene therapy of primary T cell immunodeficiencies*. *Gene*, 2013. **525**(2): p. 170-3.
86. Aiuti, A., et al., *Lentiviral Hematopoietic Stem Cell Gene Therapy in Patients with Wiskott-Aldrich Syndrome*. *Science*, 2013. **341**(6148).
87. Boztug, K., et al., *Stem-cell gene therapy for the Wiskott-Aldrich syndrome*. *N Engl J Med*, 2010. **363**(20): p. 1918-27.
88. Buchbinder, D., D.J. Nugent, and A.H. Phillipovich, *Wiskott-Aldrich syndrome: diagnosis, current management, and emerging treatments*. *Appl Clin Genet*, 2014. **7**: p. 55-66.
89. Mukherjee, S. and A.J. Thrasher, *Gene therapy for PIDs: progress, pitfalls and prospects*. *Gene*, 2013. **525**(2): p. 174-81.
90. Pfeifer, G.P., et al., *Genomic sequencing and methylation analysis by ligation mediated PCR*. *Science*, 1989. **246**(4931): p. 810-3.
91. Silver, J. and V. Keerikatte, *Novel use of polymerase chain reaction to amplify cellular DNA adjacent to an integrated provirus*. *J Virol*, 1989. **63**(5): p. 1924-8.
92. Mueller, P.R. and B. Wold, *In vivo footprinting of a muscle specific enhancer by ligation mediated PCR*. *Science*, 1989. **246**(4931): p. 780-6.
93. Schmidt, M., et al., *Polyclonal long-term repopulating stem cell clones in a primate model*. *Blood*, 2002. **100**(8): p. 2737-43.
94. Bushman, F.D., *Retroviral integration and human gene therapy*. *J Clin Invest*, 2007. **117**(8): p. 2083-6.
95. Cattoglio, C., et al., *Hot spots of retroviral integration in human CD34+ hematopoietic cells*. *Blood*, 2007. **110**(6): p. 1770-8.
96. Modlich, U., et al., *Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors*. *Mol Ther*, 2009. **17**(11): p. 1919-28.
97. Braun, C.J., et al., *Gene therapy for Wiskott-Aldrich syndrome--long-term efficacy and genotoxicity*. *Sci Transl Med*, 2014. **6**(227): p. 227ra33.
98. Ariga, T., *A possible turning point in the hematopoietic stem cell gene therapy for primary immunodeficiency diseases? Lentiviral vectors could take the place of retroviral vectors*. *Expert Rev Clin Immunol*, 2013. **9**(11): p. 1015-8.
99. Wu, X., et al., *Transcription start regions in the human genome are favored targets for MLV integration*. *Science*, 2003. **300**(5626): p. 1749-51.
100. Abel, U., et al., *Real-time definition of non-randomness in the distribution of genomic events*. *PLoS One*, 2007. **2**(6): p. e570.
101. Schmidt, M., et al., *Clonality analysis after retroviral-mediated gene transfer to CD34+ cells from the cord blood of ADA-deficient SCID neonates*. *Nat Med*, 2003. **9**(4): p. 463-8.

102. Schmidt, M., et al., *Detection and direct genomic sequencing of multiple rare unknown flanking DNA in highly complex samples*. Hum Gene Ther, 2001. **12**(7): p. 743-9.
103. Schmidt, M., et al., *High-resolution insertion-site analysis by linear amplification-mediated PCR (LAM-PCR)*. Nat Methods, 2007. **4**(12): p. 1051-7.
104. John M Coffin, S.H.H.a.H.E.V., *Retroviruses*. Cold Spring Harbor Laboratory Press, 1997.
105. Cavazzana-Calvo, M., et al., *Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease*. Science, 2000. **288**(5466): p. 669-72.
106. Gaspar, H.B., et al., *Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector*. Lancet, 2004. **364**(9452): p. 2181-7.
107. Hacein-Bey-Abina, S., et al., *LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1*. Science, 2003. **302**(5644): p. 415-9.
108. Howe, S.J., et al., *Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients*. J Clin Invest, 2008. **118**(9): p. 3143-50.
109. Ott, M.G., et al., *Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1*. Nat Med, 2006. **12**(4): p. 401-9.
110. Cleveland, S.M., et al., *Lmo2 induces hematopoietic stem cell-like features in T-cell progenitor cells prior to leukemia*. Stem Cells, 2013. **31**(5): p. 882-94.
111. Nam, C.H. and T.H. Rabbitts, *The role of LMO2 in development and in T cell leukemia after chromosomal translocation or retroviral insertion*. Mol Ther, 2006. **13**(1): p. 15-25.
112. Zychlinski, D., et al., *Physiological promoters reduce the genotoxic risk of integrating gene vectors*. Mol Ther, 2008. **16**(4): p. 718-25.
113. Montini, E., et al., *Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration*. Nat Biotechnol, 2006. **24**(6): p. 687-96.
114. Paruzynski, A., et al., *Genome-wide high-throughput integrome analyses by nrLAM-PCR and next-generation sequencing*. Nat Protoc, 2010. **5**(8): p. 1379-95.
115. Li, Z., et al., *Murine leukemia induced by retroviral gene marking*. Science, 2002. **296**(5567): p. 497.
116. Cartier, N., et al., *Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy*. Science, 2009. **326**(5954): p. 818-23.
117. Deichmann, A., et al., *Vector integration is nonrandom and clustered and influences the fate of lymphopoiesis in SCID-X1 gene therapy*. J Clin Invest, 2007. **117**(8): p. 2225-32.
118. Cavazzana-Calvo, M., et al., *Transfusion independence and HMGA2 activation after gene therapy of human beta-thalassaemia*. Nature, 2010. **467**(7313): p. 318-22.
119. Stein, S., et al., *Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease*. Nat Med, 2010. **16**(2): p. 198-204.
120. Wang, G.P., et al., *Dynamics of gene-modified progenitor cells analyzed by tracking retroviral integration sites in a human SCID-X1 gene therapy trial*. Blood, 2010. **115**(22): p. 4356-66.

121. Bartholomae, C.C., et al., *Insertion site pattern: global approach by linear amplification-mediated PCR and mass sequencing*. *Methods Mol Biol*, 2012. **859**: p. 255-65.
122. Hacein-Bey-Abina, S., A. Fischer, and M. Cavazzana-Calvo, *Gene therapy of X-linked severe combined immunodeficiency*. *Int J Hematol*, 2002. **76**(4): p. 295-8.
123. Aiuti, A., et al., *Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning*. *Science*, 2002. **296**(5577): p. 2410-3.
124. Nienhuis, A.W., C.E. Dunbar, and B.P. Sorrentino, *Genotoxicity of retroviral integration in hematopoietic cells*. *Mol Ther*, 2006. **13**(6): p. 1031-49.
125. Kozar, K., et al., *Mouse development and cell proliferation in the absence of D-cyclins*. *Cell*, 2004. **118**(4): p. 477-91.
126. Kabakov, A.E. and V.L. Gabai, *Heat-shock proteins maintain the viability of ATP-deprived cells: what is the mechanism?* *Trends Cell Biol*, 1994. **4**(6): p. 193-6.
127. Askew, D.S., et al., *Constitutive c-myc expression in an IL-3-dependent myeloid cell line suppresses cell cycle arrest and accelerates apoptosis*. *Oncogene*, 1991. **6**(10): p. 1915-22.
128. Chang, D.W., et al., *The c-Myc transactivation domain is a direct modulator of apoptotic versus proliferative signals*. *Mol Cell Biol*, 2000. **20**(12): p. 4309-19.
129. Su, X., et al., *Epigenetically modulated LRRC33 acts as a negative physiological regulator for multiple Toll-like receptors*. *J Leukoc Biol*, 2014.
130. Liu, J., et al., *Identification and characterization of a unique leucine-rich repeat protein (LRRC33) that inhibits Toll-like receptor-mediated NF-kappaB activation*. *Biochem Biophys Res Commun*, 2013. **434**(1): p. 28-34.
131. Alexander, C. and E.T. Rietschel, *Bacterial lipopolysaccharides and innate immunity*. *J Endotoxin Res*, 2001. **7**(3): p. 167-202.
132. Jiang, H., et al., *Phospholipase C beta 4 is involved in modulating the visual response in mice*. *Proc Natl Acad Sci U S A*, 1996. **93**(25): p. 14598-601.
133. Hacein-Bey-Abina, S., et al., *Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1*. *J Clin Invest*, 2008. **118**(9): p. 3132-42.
134. Hacein-Bey-Abina, S., et al., *A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency*. *N Engl J Med*, 2003. **348**(3): p. 255-6.
135. Baum, C., et al., *[Gene therapy of SCID-X1]*. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*, 2007. **50**(12): p. 1507-17.
136. Deichmann, A., et al., *Insertion sites in engrafted cells cluster within a limited repertoire of genomic areas after gammaretroviral vector gene therapy*. *Mol Ther*, 2011. **19**(11): p. 2031-9.
137. Ambrosi, A., C. Cattoglio, and C. Di Serio, *Retroviral integration process in the human genome: is it really non-random? A new statistical approach*. *PLoS Comput Biol*, 2008. **4**(8): p. e1000144.
138. Hacein-Bey-Abina, S., et al., *Efficacy of gene therapy for X-linked severe combined immunodeficiency*. *N Engl J Med*, 2010. **363**(4): p. 355-64.
139. Schwarzwaelder, K., et al., *Gammaretrovirus-mediated correction of SCID-X1 is associated with skewed vector integration site distribution in vivo*. *J Clin Invest*, 2007. **117**(8): p. 2241-9.
140. Kang, H.J., et al., *Retroviral gene therapy for X-linked chronic granulomatous disease: results from phase I/II trial*. *Mol Ther*, 2011. **19**(11): p. 2092-101.
141. Brunsing, R., et al., *B- and T-cell development both involve activity of the unfolded protein response pathway*. *J Biol Chem*, 2008. **283**(26): p. 17954-61.

142. Gonzalez-Farre, B., et al., *In vivo intratumoral Epstein-Barr virus replication is associated with XBP1 activation and early-onset post-transplant lymphoproliferative disorders with prognostic implications*. 2014.
143. Lopez-Guillermo, A., et al., *A multiepitope of XBP1, CD138 and CS1 peptides induces myeloma-specific cytotoxic T lymphocytes in T cells of smoldering myeloma patients*. *Mod Pathol*, 2014.
144. Ariki, R., et al., *Homeodomain transcription factor Meis1 is a critical regulator of adult bone marrow hematopoiesis*. *PLoS One*, 2014. **9**(2): p. e87646.
145. Wang, Z., et al., *Cyclin D3 is essential for megakaryocytopoiesis*. *Blood*, 1995. **86**(10): p. 3783-8.
146. Pereira, R., et al., *FLI-1 inhibits differentiation and induces proliferation of primary erythroblasts*. *Oncogene*, 1999. **18**(8): p. 1597-608.
147. Peterson, L.F., et al., *The hematopoietic transcription factor AML1 (RUNX1) is negatively regulated by the cell cycle protein cyclin D3*. *Mol Cell Biol*, 2005. **25**(23): p. 10205-19.
148. Sankaran, V.G., et al., *Cyclin D3 coordinates the cell cycle during differentiation to regulate erythrocyte size and number*. *Genes Dev*, 2012. **26**(18): p. 2075-87.
149. Welinder, E., et al., *The transcription factors E2A and HEB act in concert to induce the expression of FOXO1 in the common lymphoid progenitor*. *Proc Natl Acad Sci U S A*, 2011. **108**(42): p. 17402-7.
150. Adolfsson, J., et al., *Upregulation of Flt3 expression within the bone marrow Lin(-)Sca1(+)-kit(+) stem cell compartment is accompanied by loss of self-renewal capacity*. *Immunity*, 2001. **15**(4): p. 659-69.
151. Papathanasiou, P., et al., *Evaluation of the long-term reconstituting subset of hematopoietic stem cells with CD150*. *Stem Cells*, 2009. **27**(10): p. 2498-508.
152. Morita, Y., H. Ema, and H. Nakauchi, *Heterogeneity and hierarchy within the most primitive hematopoietic stem cell compartment*. *J Exp Med*, 2010. **207**(6): p. 1173-82.
153. Thillainadesan, G., et al., *TGF-beta-dependent active demethylation and expression of the p15ink4b tumor suppressor are impaired by the ZNF217/CoREST complex*. *Mol Cell*, 2012. **46**(5): p. 636-49.
154. Collins, C., et al., *Comprehensive genome sequence analysis of a breast cancer amplicon*. *Genome Res*, 2001. **11**(6): p. 1034-42.
155. Quinlan, K.G., et al., *Specific recognition of ZNF217 and other zinc finger proteins at a surface groove of C-terminal binding proteins*. *Mol Cell Biol*, 2006. **26**(21): p. 8159-72.
156. Cowger, J.J., et al., *Biochemical characterization of the zinc-finger protein 217 transcriptional repressor complex: identification of a ZNF217 consensus recognition sequence*. *Oncogene*, 2007. **26**(23): p. 3378-86.
157. Banck, M.S., et al., *The ZNF217 oncogene is a candidate organizer of repressive histone modifiers*. *Epigenetics*, 2009. **4**(2): p. 100-6.
158. Vendrell, J.A., et al., *ZNF217 is a marker of poor prognosis in breast cancer that drives epithelial-mesenchymal transition and invasion*. *Cancer Res*, 2012. **72**(14): p. 3593-606.
159. Matsumura, K., *[Detection of DNA amplifications and deletions in oral squamous cell carcinoma cell lines by comparative genomic hybridization (CGH)]*. *Kokubyo Gakkai Zasshi*, 1995. **62**(4): p. 513-31.
160. Shinomiya, T., et al., *Comparative genomic hybridization of squamous cell carcinoma of the esophagus: the possible involvement of the DPI gene in the 13q34 amplicon*. *Genes Chromosomes Cancer*, 1999. **24**(4): p. 337-44.

161. Courjal, F., et al., *DNA amplifications at 20q13 and MDM2 define distinct subsets of evolved breast and ovarian tumours*. Br J Cancer, 1996. **74**(12): p. 1984-9.
162. Courjal, F., et al., *Mapping of DNA amplifications at 15 chromosomal localizations in 1875 breast tumors: definition of phenotypic groups*. Cancer Res, 1997. **57**(19): p. 4360-7.
163. Kallioniemi, A., et al., *Detection and mapping of amplified DNA sequences in breast cancer by comparative genomic hybridization*. Proc Natl Acad Sci U S A, 1994. **91**(6): p. 2156-60.
164. Zhu, Y.Q., et al., *[Study on amplification of ZNF217 in primary gastric carcinoma]*. Zhonghua Wei Chang Wai Ke Za Zhi, 2005. **8**(1): p. 60-2.
165. Quinlan, K.G.R., et al., *Amplification of zinc finger gene 217 (ZNF217) and cancer: When good fingers go bad*. Biochimica et Biophysica Acta (BBA) - Reviews on Cancer, 2007. **1775**(2): p. 333-340.
166. Georgiou, G., et al., *Serial determination of FLT3 mutations in myelodysplastic syndrome patients at diagnosis, follow up or acute myeloid leukaemia transformation: incidence and their prognostic significance*. Br J Haematol, 2006. **134**(3): p. 302-6.
167. Bagger, F.O., et al., *HemaExplorer: a database of mRNA expression profiles in normal and malignant haematopoiesis*. Nucleic Acids Res, 2013. **41**(Database issue): p. D1034-9.
168. Kwiatkowski, A.V., F.B. Gertler, and J.J. Loureiro, *Function and regulation of Ena/VASP proteins*. Trends Cell Biol, 2003. **13**(7): p. 386-92.
169. Krause, M., et al., *Ena/VASP proteins: regulators of the actin cytoskeleton and cell migration*. Annu Rev Cell Dev Biol, 2003. **19**: p. 541-64.
170. Mejillano, M.R., et al., *Lamellipodial versus filopodial mode of the actin nanomachinery: pivotal role of the filament barbed end*. Cell, 2004. **118**(3): p. 363-73.
171. Applewhite, D.A., et al., *Ena/VASP proteins have an anti-capping independent function in filopodia formation*. Mol Biol Cell, 2007. **18**(7): p. 2579-91.
172. Chang, C., et al., *MIG-10/lamellipodin and AGE-1/PI3K promote axon guidance and outgrowth in response to slit and netrin*. Curr Biol, 2006. **16**(9): p. 854-62.
173. Takaku, M., et al., *Recombination activator function of the novel RAD51- and RAD51B-binding protein, human EVL*. J Biol Chem, 2009. **284**(21): p. 14326-36.
174. Hashimoto, K., et al., *Roles of phospholipase Cbeta4 in synapse elimination and plasticity in developing and mature cerebellum*. Mol Neurobiol, 2001. **23**(1): p. 69-82.
175. Watanabe, M., et al., *Patterns of expression for the mRNA corresponding to the four isoforms of phospholipase Cbeta in mouse brain*. Eur J Neurosci, 1998. **10**(6): p. 2016-25.
176. Faenza, I., et al., *Expression of phospholipase C beta family isoenzymes in C2C12 myoblasts during terminal differentiation*. J Cell Physiol, 2004. **200**(2): p. 291-6.
177. Lyon, A.M. and J.J. Tesmer, *Structural insights into phospholipase C-beta function*. Mol Pharmacol, 2013. **84**(4): p. 488-500.
178. Berridge, M.J., M.D. Bootman, and H.L. Roderick, *Calcium signalling: dynamics, homeostasis and remodelling*. Nat Rev Mol Cell Biol, 2003. **4**(7): p. 517-29.
179. Jiang, H., et al., *Roles of phospholipase C beta2 in chemoattractant-elicited responses*. Proc Natl Acad Sci U S A, 1997. **94**(15): p. 7971-5.
180. Murthy, K.S. and G.M. Makhlof, *Opioid mu, delta, and kappa receptor-induced activation of phospholipase C-beta 3 and inhibition of adenylyl cyclase is mediated by Gi2 and G(o) in smooth muscle*. Mol Pharmacol, 1996. **50**(4): p. 870-7.

181. Palaniyandi, S.S., et al., *Protein kinase C in heart failure: a therapeutic target?* Cardiovasc Res, 2009. **82**(2): p. 229-39.
182. Braz, J.C., et al., *PKC-alpha regulates cardiac contractility and propensity toward heart failure.* Nat Med, 2004. **10**(3): p. 248-54.
183. Rampazzo, A., et al., *Characterization of C14orf4, a novel intronless human gene containing a polyglutamine repeat, mapped to the ARVD1 critical region.* Biochem Biophys Res Commun, 2000. **278**(3): p. 766-74.
184. Heger, S., et al., *Enhanced at puberty 1 (EAP1) is a new transcriptional regulator of the female neuroendocrine reproductive axis.* J Clin Invest, 2007. **117**(8): p. 2145-54.
185. Krause, D.S., et al., *Multi-organ, multi-lineage engraftment by a single bone marrow-derived stem cell.* Cell, 2001. **105**(3): p. 369-77.
186. Li, C.L. and G.R. Johnson, *Rhodamine123 reveals heterogeneity within murine Lin-, Sca-1+ hemopoietic stem cells.* J Exp Med, 1992. **175**(6): p. 1443-7.
187. Randall, T.D. and I.L. Weissman, *Characterization of a population of cells in the bone marrow that phenotypically mimics hematopoietic stem cells: resting stem cells or mystery population?* Stem Cells, 1998. **16**(1): p. 38-48.
188. Collins, C., et al., *Positional cloning of ZNF217 and NABC1: genes amplified at 20q13.2 and overexpressed in breast carcinoma.* Proc Natl Acad Sci U S A, 1998. **95**(15): p. 8703-8.
189. Li, J., L.L. Song, and M. Zhong, *[ZNF217 expression correlates with the biological behavior of human ovarian cancer cells].* Zhonghua Zhong Liu Za Zhi, 2013. **35**(3): p. 170-4.
190. Littlepage, L.E., et al., *The transcription factor ZNF217 is a prognostic biomarker and therapeutic target during breast cancer progression.* Cancer Discov, 2012. **2**(7): p. 638-51.
191. Li, P., et al., *Multiple roles of the candidate oncogene ZNF217 in ovarian epithelial neoplastic progression.* Int J Cancer, 2007. **120**(9): p. 1863-73.
192. Laneuville, P., N. Heisterkamp, and J. Groffen, *Expression of the chronic myelogenous leukemia-associated p210bcr/abl oncoprotein in a murine IL-3 dependent myeloid cell line.* Oncogene, 1991. **6**(2): p. 275-82.
193. Yamamoto, Y., et al., *Activating mutation of D835 within the activation loop of FLT3 in human hematologic malignancies.* Blood, 2001. **97**(8): p. 2434-9.
194. Hidaka, S., et al., *Differences in 20q13.2 copy number between colorectal cancers with and without liver metastasis.* Clin Cancer Res, 2000. **6**(7): p. 2712-7.
195. Dolan, J., et al., *The extracellular leucine-rich repeat superfamily; a comparative survey and analysis of evolutionary relationships and expression patterns.* BMC Genomics, 2007. **8**: p. 320.
196. Baltimore, D., *NF-kappaB is 25.* Nat Immunol, 2011. **12**(8): p. 683-5.
197. Ruland, J., *Return to homeostasis: downregulation of NF-kappaB responses.* Nat Immunol, 2011. **12**(8): p. 709-14.
198. Zhao, J.L., et al., *MicroRNA-146a acts as a guardian of the quality and longevity of hematopoietic stem cells in mice.* Elife, 2013. **2**: p. e00537.
199. Bottero, V., S. Withoff, and I.M. Verma, *NF-kappaB and the regulation of hematopoiesis.* Cell Death Differ, 2006. **13**(5): p. 785-97.
200. Gerondakis, S., R.J. Grumont, and A. Banerjee, *Regulating B-cell activation and survival in response to TLR signals.* Immunol Cell Biol, 2007. **85**(6): p. 471-5.
201. Kim, D., et al., *Phospholipase C isozymes selectively couple to specific neurotransmitter receptors.* Nature, 1997. **389**(6648): p. 290-3.

202. Miyata, M., et al., *Role of thalamic phospholipase C[beta]4 mediated by metabotropic glutamate receptor type 1 in inflammatory pain*. J Neurosci, 2003. **23**(22): p. 8098-108.
203. Han, S.K., V. Mancino, and M.I. Simon, *Phospholipase Cbeta 3 mediates the scratching response activated by the histamine H1 receptor on C-fiber nociceptive neurons*. Neuron, 2006. **52**(4): p. 691-703.
204. Follo, M.Y., et al., *Nuclear PI-PLCbeta1 and myelodysplastic syndromes: genetics and epigenetics*. Curr Pharm Des, 2012. **18**(13): p. 1751-4.
205. Okada, Y., et al., *Common variations in PSMD3-CSF3 and PLCB4 are associated with neutrophil count*. Hum Mol Genet, 2010. **19**(10): p. 2079-85.
206. Mikkola, H.K. and S.H. Orkin, *The journey of developing hematopoietic stem cells*. Development, 2006. **133**(19): p. 3733-44.
207. Alvarez-Silva, M., et al., *Mouse placenta is a major hematopoietic organ*. Development, 2003. **130**(22): p. 5437-44.
208. Kikuchi, K. and M. Kondo, *Developmental switch of mouse hematopoietic stem cells from fetal to adult type occurs in bone marrow after birth*. Proc Natl Acad Sci U S A, 2006. **103**(47): p. 17852-7.
209. Liang, Y., G. Van Zant, and S.J. Szilvassy, *Effects of aging on the homing and engraftment of murine hematopoietic stem and progenitor cells*. Blood, 2005. **106**(4): p. 1479-87.
210. Morrison, S.J., et al., *The aging of hematopoietic stem cells*. Nat Med, 1996. **2**(9): p. 1011-6.
211. de Haan, G. and G. Van Zant, *Dynamic changes in mouse hematopoietic stem cell numbers during aging*. Blood, 1999. **93**(10): p. 3294-301.

Publications

“EVI1 inhibits apoptosis induced by antileukemic drugs via upregulation of CDKN1A/p21/WAF in human myeloid cells.”

Rommer A, Steinmetz B, Herbst F, Hackl H, Heffeter P, Heilos D, Filipits M, Steinleitner K, Hemmati S, Herbacek I, Schwarzingler I, Hartl K, Rondou P, Glimm H, Karakaya K, Krämer A, Berger W, Wieser R. *PLoS One*. 2013 Feb;8(2):e56308.

“Optimization of gene transfection in murine myeloma cell lines using different transfection reagents.”

Shabani M, Hemmati S, Hadavi R, Amirghofran Z, Jeddi-Tehrani M, Rabbani H, Shokri F. *Avicenna J Med Biotechnol*. 2010 Jul; 2(3): 123-30.

“Indoleamine 2,3-dioxygenase (IDO) is expressed at feto-placental unit throughout mouse gestation: An immunohistochemical study.”

Hemmati S, Mahmood JT, Ebrahim T, Jamileh G, Golnaz Ensieh KS, Parivash D, Leila BY, Mohammad Mehdi A, Amir Hassan Z. *J Reprod Infertil*. 2009 Oct; 10(3):177-83.

“Ectopic Expression of Sortilin 1 (NTR-3) in Patients with Ovarian Carcinoma.” Hemmati S, Zarnani AH, Mahmoudi AR, Sadeghi MR, Soltanghoreae H, Akhondi MM, Tarahomi M, Jeddi-Tehrani M, Rabbani H. *Avicenna J Med Biotechnol*. 2009 Jul;1(2):125-31.

Conferences

“Investigating the role of candidate regulators of post-embryonic hematopoiesis”

Shayda Hemmati, Friederike Herbst, Claudia R Ball, Manfred Schmidt, Christof von Kalle and Hanno Glimm. *DKFZ PhD Retreat, July 2011, Weil der Stadt, Germany*

“Control of stem cell fate by key regulators of post-embryonic hematopoiesis “

Shayda Hemmati, Friederike Herbst, Claudia R Ball, Manfred Schmidt, Christof von Kalle and Hanno Glimm. *DKFZ PhD Student Poster Presentation, November 2011, Heidelberg, Germany*

“Identification of novel regulatory genes of human post-embryonic hematopoiesis.”

Shayda Hemmati, Friederike Herbst, Claudia R. Ball¹, Anna Paruzynski, Kaan Boztug, Christoph Klein, Manfred Schmidt, Christof von Kalle, Hanno Glimm. *42nd Annual Scientific Meeting of the Society for Hematology and Stem Cells (ISEH) Poster Presentation. August 2013, Vienna, Austria.*

Declaration

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig angefertigt habe. Es wurden nur die in der Arbeit ausdrücklich benannten Quellen und Hilfsmittel benutzt. Wörtlich oder sinngemäß übernommenes Gedankengut habe ich als solches kenntlich gemacht.

Ich erkläre außerdem, dass diese Arbeit weder in dieser noch in einer anderen Form anderweitig als Dissertation oder Prüfungsarbeit verwendet oder einer anderen Fakultät als Dissertation vorgelegt wurde.

Heidelberg, _____

Ort, Datum

Unterschrift

Acknowledgments

I would like to express my special appreciation and thanks to my advisor Dr. Hanno Glimm. I thank you for encouraging my research and for allowing me to grow as an independent research scientist. This work would not have been possible without your great support and wise discussions.

I would like to thank my TAC committee members, Dr. Andreas Trumpp, Dr. Christof von Kalle, and Dr. Hartmut Goldschmidt for their insightful comments during my PhD.

I thank Dr. Suat Özbek and Dr. Martin Müller for accepting to be my thesis examiners.

I thank Dr. Claudia Ball for her endless support and supervision. I admire her passion and dedication in organizing our research laboratory in the best way.

Dr. Friederike Herbst who gave me an enormous positive energy and motivation during my PhD journey. Thanks for being by my side during all these years. Your intelligent comments and supervision was invaluable to me. Vielen Dank Fee 😊!

My special appreciation goes to our wonderful and supportive technicians: Tim, Sylvia, Sabrina, Annika, Galina, Stefanie, Rebecca, and Silja. I really appreciate your help and understanding.

I would like to thank our animal care takers especially Susanne, Sandra, Sabine and Nina for their wonderful help.

My wonderful friends and colleagues Oksana, Taronish, Klara, Felix and Chris: I always enjoyed your company. Thanks a million!

I would like to thank Sarah, Caro, Jenny, Sebastian, Roland, Alex, Peer, Elias, Tonio, and Lino for making my time in the lab so pleasurable.

I thank Miko for his great help under any conditions. You are awesome Miko!

Anna, Eva, Rebecca, Paula and Svantje thank you very much for all your support and encouragement during cheerful and stressful moments of the thesis and the rest of the time. I enjoyed every moment of my time that I spent with you!

And last but not least: words are not enough to thank my parents, brother and sister for always believing in me and being on my side during this challenging period and far beyond that! Thank you very much!