

**Hepatocellular carcinoma:
Establishing a new transplantation model for the study of benign and
malignant clonal regeneration of the murine liver *in vivo***

Dissertation
to obtain the academic degree of
Doctor of Natural Sciences (Dr. rer. nat.)

submitted to the Department of Biology
Universität Hamburg

by
Ludmilla Unrau

Hamburg

2018

This doctoral dissertation was supervised by Prof. Dr. Boris Fehse and performed in the Research Department Cell and Gene Therapy, Department of Stem Cell Transplantation, Center for Oncology, University Medical Center Hamburg-Eppendorf, in Hamburg, Germany from August 2014 until December 2017.

Reviewer of the dissertation:

Prof. Dr. Boris Fehse

Prof. Dr. Thomas Dobner

This thesis was successfully defended on October 19th, 2018 in the presence of the following committee members:

Prof. Dr. Julia Kehr

Prof. Dr. Boris Fehse

Prof. Dr. Maura Dandri

PD Dr. Hartwig Lüthen

PD Dr. Claudia Lange

Dr. Kerstin Cornils

Abstract

One major type of primary liver cancer is the hepatocellular carcinoma (HCC), which is the second most leading cause of cancer-related mortality worldwide. Though HCC development has been investigated extensively, the phenotypic and molecular heterogeneity driving HCC initiation still remain largely unknown.

The objective of this thesis was to induce HCC formation via lentiviral-mediated gene transduction and to take advantage of the red/green/blue (RGB) marking system to develop a small animal model to study both benign and malignant clonal regeneration of the liver *in vivo*.

One of three different transgenes, *CyclinA2*, *HRas-V12*, and *LargeT-antigen*, each co-expressed with a Green fluorescent gene, together with “empty” vectors expressing either Cherry (Red), Venus (yellow-Green) or Cerulean (Blue) fluorescent genes, were introduced into primary adult wild-type hepatocytes using the LeGO vector system. Intrasplenic transplantation into urokinase-type plasminogen activator/SCID/Beige (USB) transgenic mice was then performed. Tumor growth was monitored via repeated magnetic resonance imaging screenings. At the final stage, liver transaminases (ALT) and albumin (ALB) serum concentrations were determined and cryosections and molecular analyses of livers were performed.

Transplantation of primary murine hepatocytes resulted in efficient liver repopulation of all mice. A healthy liver phenotype was observed in the mock-treated group (n=6). RGB control mice (n=12, without oncogene expression) were repopulated by numerous clones with normal ALT and ALB levels. Explanted livers showed normal phenotype with single fibrotic areas. The three different transgenes mediated distinct kinetics of liver regeneration and HCC formation. *CyclinA2*-transduced hepatocytes were found in small repopulated areas (n=11). Mice had normal ALT and ALB values but minor liver damage was already observable macroscopically. Early signs of malignant transformation were only observed in mice transplanted with *HRas-V12*-transduced hepatocytes. MRI analyses showed formation of multiple tumor nodules within 3 weeks (n=12). In these mice, ALT was slightly elevated, whereas ALB was normal. *LargeT*-antigen-transduced hepatocytes caused severe liver damage in transplanted mice as indicated by 6.35-fold up-regulated ALT values but normal ALB and severe tumor formation (n=12). Molecular quantification of vector copy numbers by droplet-digital PCR in relation to a reference gene was in line with immunofluorescence data of GFP and oncoprotein levels: in *LargeT*-antigen- and RGB-transplanted livers maximum proviral vector copy numbers were found, followed by *HRas-V12*-modified hepatocytes and *CyclinA2*-transduced cells. The adapted Barcelona Clinic Liver Cancer (BCLC) staging classified *HRas-V12* and *LargeT* induced hepatocellular carcinomas to an intermediate stage in this experimental setting.

Based on these data, *HRas-V12* is a promising candidate to study early and intermediate phases of liver carcinogenesis by lentiviral-mediated oncogene transduction and marking.

The combination of sensitive cell-marking strategies and oncogene-mediated hepatocarcinogenesis in USB mice can provide new insights in the clonal evolution of HCC, and provides the necessary tools to establish a more realistic and patient-related *in-vivo* model for the investigation of liver cancer.

Zusammenfassung

Eine der häufigsten Formen von Leberkrebs ist das hepatozelluläre Karzinom (engl. *hepatocellular carcinoma*, kurz: HCC), welches außerdem weltweit als zweithäufigste Ursache krebsinduzierte Letalität bedingt. Auch wenn die Entstehung von HCCs bereits eingehend untersucht wurde, sind die phänotypische und molekulare Heterogenität, welche die Initiierung von HCC bedingen, weitgehend unbekannt.

Das Ziel dieser Arbeit war die Induktion der HCC-Entstehung mittels Lentiviral vermittelter Gentransduktion unter Verwendung des Rot/Grün/Blau (RGB) Markierungssystems, um ein Kleintiermodell für die *in vivo* Untersuchung der physiologischen aber auch der malignen klonalen Regeneration der Leber zu etablieren.

Eins von drei verschiedenen Transgenen, *CyclinA2*, *HRas-V12* und *LargeT-Antigen*, die jeweils ein grünes Fluoreszenzprotein ko-exprimierten, wurde zusammen mit „leeren“ Vektoren, welche das Cherry (Rot), Venus (Gelb-Grün) or Cerulean (Blau) Fluoreszenzgen exprimierten, mittels LeGO-Vektorsystem in primäre adulte Wildtyphepatozyten eingebracht. Die Hepatozyten wurden anschließend intrasplenal in vier Wochen alte transgene Urokinasetyp-Plasminogenaktivator/SCID/Beige (USB) Mäuse transplantiert. Das Tumorstadium wurde anhand wiederholter MRT-Messungen beobachtet. Lebertransaminasen (ALT) und Albumin (ALB) Konzentrationen wurden bei der finalen Blutabnahme im Serum bestimmt, sowie Kryoschnitte angefertigt und molekulare Analysen der entnommenen Lebern durchgeführt.

Die Transplantation primärer muriner Hepatozyten resultierte in effizienter Leberrepopulation in allen untersuchten Mäusen. Ein gesunder Leberphänotyp konnte in der Mock-behandelten Gruppe beobachtet werden (n=6). RGB-Kontrolltiere wiesen eine Vielzahl an transduzierten Klonen in der repopulierten Leber und außerdem normale ALT- und ALB-Werte auf (n=12, ohne Onkogenexpression). Die entnommenen Lebern waren phänotypisch gesund, jedoch mit einzelnen fibrotischen Arealen. Die drei verschiedenen Transgene zeigten jeweils reproduzierbare Kinetiken der Leberregeneration und HCC-Formation. *CyclinA2*-transduzierte Hepatozyten konnten in kleinen repopulierten Flächen wieder gefunden werden (n=11). Die Tiere zeigten normale ALT- und ALB-Werte und keine Tumorbildung trotz bereits makroskopisch sichtbaren geringen Leberschadens. Frühe Zeichen der malignen Transformation konnten ausschließlich in Tieren festgestellt werden, in die *HRas-V12*-transduzierte Leberzellen transplantiert wurden; MRT-Aufnahmen zeigten die Bildung multipler Tumore innerhalb von 3 Wochen nach Transplantation (n=12). In diesen Mäusen waren die ALT-Werte leicht erhöht, während Albumin auf normalem Level blieb. Mit *LargeT*-Antigen-transduzierte Hepatozyten verursachten massiven Leberschaden in transplantierten Mäusen, angezeigt durch exzessives Tumorstadium und einen 6,35-fach hoch regulierten ALT-Wert bei normalem ALB (n=12). Molekulare Quantifizierung von Vektorkopiezahlen durch digitale PCR im Vergleich zu einem Referenzgen entsprach den mittels Immunfluoreszenz nachgewiesenen Levels von GFP und den Onkoproteinen: die maximale provirale Vektorkopiezahl wurde in *LargeT*-Antigen- und RGB-transplantierten Lebern gefunden, gefolgt von *HRas-V12*-modifizierten Hepatozyten und *CyclinA2*-transduzierten Zellen. Anhand des angepassten Barcelona Krankenhaus Leberkrebsbewertungssystems (BCLC system) wurden *HRas-V12* und *LargeT* induzierte hepatozelluläre Karzinome als intermediäres Stadium eingestuft.

Auf diesen Daten basierend ist *HRas-V12* ein vielversprechender Kandidat für die Untersuchung früher und intermediärer Phasen der Leberkarzinogenese durch lentiviral-vermittelte Onkogen-transduktion und -markierung.

Die Kombination sensitiver Zellmarkierungsstrategien und Onkogen-induzierter Hepatokarzinogenese in USB-Mäusen kann zum Erkenntnisgewinn der klonalen Evolution von HCC führen und liefert die nötigen Hilfsmittel zur Etablierung eines realitätsnahen und Patientenbezogenen *in vivo* Modells zur Erforschung von Leberkrebs.

Contents

Abstract	i
Zusammenfassung	ii
Contents	iii
Abbreviations	vi
List of Figures	xi
List of Tables	xiii
1 Introduction	1
1.1 Hepatocellular carcinoma	1
1.1.1 Risk factors of liver disease	2
1.1.2 Treatment options for HCC	3
1.1.3 Molecular profile in liver malignancies.....	4
1.1.4 Role of viruses in carcinogenesis.....	7
1.2 Lentiviral vectors	8
1.2.1 <i>Lentivirus</i> virion structure and genome organization.....	8
1.2.2 Replication cycle of lentiviruses	10
1.2.3 Lentiviral integration preferences and insertional mutagenesis	11
1.2.4 Optimization of lentiviral vectors	12
1.2.5 RGB marking with LeGO vectors	14
1.3 Liver repopulation models	15
1.3.1 USB mouse model.....	16
1.4 Aim of this thesis	18
2 Materials and Methods	19
2.1 Materials	19
2.1.1 Disposables.....	19
2.1.2 Instruments and accessories	20
2.1.3 Software	21
2.1.4 Chemicals and reagents	22
2.1.5 Kits.....	23
2.1.6 Enzymes and oligos	23
2.1.7 Media and additives for cell culture	23
2.1.7 Buffer for viral vector production	24
2.1.8 Buffer and stock solutions for liver perfusion and hepatocyte isolation.....	24
2.1.9 Surgical kit	25

2.2 Methods.....	26
2.2.1 Cloning of LeGO-LargeT-iV2	26
2.2.1.1 PCR amplification	26
2.2.1.2 DNA fragment purification and digestion	26
2.2.1.3 Ligation and transformation.....	27
2.2.1.4 Sequencing and maxi preparation.....	27
2.2.2 Cell culture	28
2.2.2.1 Cultivation of adherent cells.....	28
2.2.2.2 Production of infectious vector particles	29
2.2.2.3 Concentration of infectious vector particles	30
2.2.2.4 Determination of infectious vector titers.....	30
2.2.2.5 Generation of control cells for molecular analysis.....	31
2.2.3 Mouse experiments	31
2.2.3.1 <i>In-situ</i> perfusion	32
2.2.3.2 Isolation of primary hepatocytes.....	33
2.2.3.3 Transduction of primary hepatocytes.....	34
2.2.3.4 Assessment of transduction rates of primary hepatocytes	34
2.2.3.5 Intrasplenic transplantation of primary hepatocytes	35
2.2.3.6 Magnetic resonance imaging	36
2.2.3.7 Blood/serum sampling	36
2.2.3.8 Euthanizing.....	36
2.2.3.9 Serum ALB and ALT measurements.....	36
2.2.3.10 RNA and DNA analyses of liver tissue.....	37
2.2.4 Droplet-digital PCR.....	37
2.2.5 Histology and microscopy	39
2.2.5.1 Preparation of liver tissue for microscopy	39
2.2.5.2 Immunohistochemical staining of liver tissue	40
2.2.5.3 Immunofluorescence staining of liver tissue	41
2.2.5.4 Calculation of positive staining in IF and IHC images	43
2.2.6 Statistics.....	43
2.2.7 Workflow	43
3 Results.....	46
3.1 Physiological weight gain in all experimental groups.....	46
3.2 HRas-V12- and LargeT-transplanted mice showed early symptoms of disease progression.....	47
3.3 Normal albumin values were observed in all groups	49
3.4 Highly elevated alanine aminotransferase levels in LargeT-transplanted mice	50
3.5 Magnetic resonance imaging revealed distinct group kinetics	51

3.7 Immunohistochemical staining confirmed engraftment of ex-vivo transduced transplanted hepatocytes.....	56
3.8 Confirmation of oncogene and fluorescence protein expressions by immunofluorescence analyses.....	59
3.9 Molecular quantification of integrated proviral DNA by droplet-digital PCR.....	61
3.9.1 Identification of specific primer/probe combinations on NIH control cells	61
3.9.2 Integration of lentiviral constructs was verified in murine liver samples.....	63
4 Discussion.....	65
4.1 Malignant clonal regeneration in HRas-V12 and LargeT recipient livers	66
4.2 Significantly increased ALT concentrations in LargeT-transplanted mice.....	68
4.3 Determination of transduction efficiency in transplanted hepatocytes	69
4.4 Benign and malignant clonal liver regeneration in a liver damage mouse model.....	70
4.5 Adaption of the Barcelona Clinic Liver Cancer classification system for the evaluation of hepatocarcinogenesis in a mouse model	72
4.6 Compilation of evaluation parameters in a radar chart	73
4.7 Conclusion.....	75
5 Bibliography.....	77
6 Online references.....	88
Supplementary.....	I
Danksagung	XXVI
Bestätigung der Korrektheit der Sprache.....	XXVIII
Eidesstattliche Versicherung	XXIX

Abbreviations

-/-	homozygous gene knock out
+/+	homozygous gene knock in
α	anti
Δ	delta, deleted
ψ	psi: packaging signal
\emptyset	diameter
% (v/v)	weight volume percent
% (w/v)	weight volume percent
AAV	adeno-associated vectors
AE buffer	elution buffer for DNA (Qiagen)
AIDS	acquired immune deficiency syndrome
AKT	protein kinase B
ALB/Alb	albumin
ALT	alanine transaminase
Amp	ampicillin
AmpR	ampicillin resistance
app.	approximately
BCLC	Barcelona Clinic Liver Cancer (classification)
BHQ	black hole quencher
bp	base pair(s)
BITREX	denatonium benzoate
C2	mCherry, red fluorescent protein
CCNA	CyclinA
CDK	Cyclin-dependent kinases
cDNA	complementary DNA
Cer2	Cerulean, blue fluorescent protein
CMV	cytomegalus virus promoter
cPPT	central polypurine tract
CT	computer tomography
DAB	3,3'-diaminobenzidine, catalyzed by peroxidase
ddH ₂ O	double distilled water
ddPCR	droplet-digital PCR
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
dNTP(s)	deoxyribonucleoside triphosphate(s)

DPBS	Dulbecco's Phosphate-Buffered Saline
dscDNA	double-stranded complementary DNA
dsDNA	double-stranded DNA
EB	elution buffer
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor
eGFP	enhanced green fluorescent protein
EGFR	epidermal growth factor receptor
EGTA	ethylene glycol-bis-(β -aminoethyl ether-)-N,N,N',N'-tetraacetic acid
EmGFP	emerald green fluorescent protein
Env	retroviral envelope protein, determines viral tropism, cleaved into transmembrane subunit (gp41) and surface subunit (gp120)
<i>Epo-R</i>	Erythropoietin receptor
ERK	extracellular signal-regulated kinase
eYFP	enhanced yellow fluorescent protein
FACS	fluorescence-activated cell sorting
Fah	fumarylacetoacetathydrolase
FAM	carboxyfluorescein; fluorophore, absorption max. 495 nm; emission max. 517 nm
FBS	fetal bovine serum
FH-hTERT	human telomerase reverse transcriptase-immortalized fetal hepatocytes
g	g-force
<i>gag/pol</i>	group antigens/polymerase, retroviral polyprotein
Gag	group-specific antigen: processed to matrix (p17) and capsid (p24) proteins, and nucleocapsid (p7)
gDNA	genomic DNA
GFP	(enhanced) Green fluorescent protein
HEX	carboxy-2,4,4,5,7,7-hexachlorofluorescein succinimidyl ester; fluorophore, absorption max. 535 nm, emission max. 556 nm
HBS	HEPES buffered saline
HBV	hepatitis B virus
HCC	hepatocellular carcinoma
HCV	hepatitis C virus
HEK293T	Human embryonic kidney cells, SV40 LargeT-antigen was stably introduced
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HIV-1	human immunodeficiency virus type 1
<i>HRas</i>	proto-oncogene, see HRas-V12
HRas-V12	GTPase, transforming protein p21, amino acid change from glycine to valine at residue 12
hSFM	HepatoZYME serum free medium
HTLV-I	human T cell leukemia virus type 1
i	IRES: internal ribosomal entry site
IF	immunofluorescence
IGF	insulin-like growth factor
IGFR	insulin-like growth factor receptor
IHC	immunohistochemistry
IL2	Interleukin-2
IL2rg	Interleukin-2 receptor common gamma chain
i.p.	intra peritoneal
IP	infectious particles
IRES	internal ribosomal entry site
kb	kilo base pair(s)
LAV	lymphadenopathy-associated virus
LB	lysogeny broth medium
LeGO	lentiviral gene ontology vector
LoxP	Cre-recombinase restriction site
LTR	long terminal repeat
mCherry	Cherry (red) fluorescent protein
MCS	multiple cloning site
MCT	monocrotaline
MDM2	E3 ubiquitin ligase
MEK	mitogen-activated protein kinase
MEK	methyl ethyl ketone/butanone
Mg	milligram
min	minute(s)
MOI	multiplicity of infection
MRI	Magnetic Resonance Imaging
mTOR	mammalian target of rapamycin
n	number, group size
n.d.	not determined
n.f.	not functional
NCBI	National Center for Biotechnology Information

Nef	negative regulatory factor
NIH3T3	primary mouse embryonic fibroblast cells; they were transferred (the "T") every 3 days (the first "3"), and inoculated at the rigid density of 3×10^5 cells per 20 cm ² dish (the second "3") continuously
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDGF	platelet-derived growth factor
PDGFR	platelet-derived growth factor receptor
PFA	paraformaldehyde
Pfp	perforin gene
PIC	pre-integration complex
PI3K	phosphoinositide-3-kinase
Pol	expressed as gag/poly polyprotein; codes for and protease (p10), reverse transcriptase (p51), ribonuclease (p15), and integrase (p32)
PPT	polypurine tract
psi	packaging signal
PTEN	phosphatase and tensin homolog
R	redundant sequence in lentiviruses
RAF	rapidly accelerated fibrosarcoma protein kinase
Rag2	recombination-activating gene 2
REF	reference
Rev	regulator of virion protein expression
RGB	red/green/blue
RNA	ribonucleic acid
RRE	Rev-response element
RSV	Rous sarcoma virus
RT	room temperature
s	second(s)
s.c.	subcutaneous
SCID	severe combined immunodeficiency
SCID-X1	severe combined immunodeficiency-X1
SFFV	spleen focus-forming virus promoter
SIN	self-inactivating
ssRNA	single-stranded RNA
SV40	simian vacuolating virus 40
Taq	<i>Thermus aquaticus</i> polymerase
T _m	melting temperature

TAE	Tris/acetate/EDTA buffer
Tat	<i>trans</i> -activator of transcription
TE	Tris-EDTA
Tris	Tris-(hydroxymethyl)-aminomethane
Tx	transplantation
U	Weiss Units
U3	unique 3' region
U5	unique 5' region
uPA	urokinase-type 1 plasminogen activator
USB	urokinase-type 1 plasminogen activator/SCID/beige
USG	uPA/SCID/Beige/Il2rg ^{-/-}
USRapG	uPA/SCID/Beige/Pfp ^{-/-} /Rag2 ^{-/-} /Il2rg ^{-/-}
V	Volt(s)
V2	Venus, green fluorescent protein
VCN	vector copy number
VEGF	vascular endothelial growth factor
VEGFR	vascular endothelial growth factor receptor
Vif	virion infectivity factor
Vpr	viral protein r
Vpu	viral protein u
vol	volume(s)
wPRE	woodchuck hepatitis virus post-transcriptional regulatory element
x	times

List of Figures

Figure 1.1 Age-standardized global mortality rates of male patients due to hepatocellular carcinoma in 2012.....	1
Figure 1.2 Progression of chronic liver disease.....	3
Figure 1.3 Cellular signaling pathways involved in the normal cell cycle and the pathogenesis of HCC	6
Figure 1.4 HIV-1 schematic exemplary of a <i>Lentivirus</i> virion.....	9
Figure 1.5 Replication cycle of HIV-1.....	11
Figure 1.6 Scheme of HIV-derived lentiviral vectors.....	13
Figure 1.7 Lentiviral gene ontology (LeGO) vector principle.....	14
Figure 1.8 RGB marking with LeGO vectors is efficient in different cell types.....	15
Figure 2.1 Expression plasmids used for viral vector production.....	31
Figure 2.2 Scheme of droplet-digital PCR.....	38
Figure 2.3 Workflow of mouse experiments.....	45
Figure 3.1 Relative weight changes [%] of transplanted mice during the course of time [days post Tx]	47
Figure 3.2 Kaplan-Meier survival curves [%] of transplanted animals [days post Tx].....	48
Figure 3.3 Final measurement of serum albumin levels (ALB) [g/L] confirmed normal values for all experimental groups	50
Figure 3.4 Liver alanine aminotransferase (ALT) levels [U/L] were determined from final serum samples	51
Figure 3.5 Representative magnetic resonance images of transplanted mice during the course of time	52
Figure 3.6 Livers of all animals of the single groups showed group-specific outcome at end of experiment.....	55
Figure 3.7 Representative peroxidase staining visualized multiple areas stained with different intensities in a USB mouse liver repopulated with RGB-transduced hepatocytes	57
Figure 3.8 CyclinA2-transduced hepatocytes engrafted in small repopulation areas of USB mouse livers 66 d post Tx	57
Figure 3.9 Repopulated USB mouse liver showed small areas of engrafted GFP-expressing HRas-V12-transduced hepatocytes 48 days post Tx	58
Figure 3.10 Large repopulation areas of LargeT-transduced hepatocytes transplanted into USB mouse livers were positive for Venus and LargeT	58
Figure 3.11 Fluorescence microscopy of liver cryosections revealed engraftment of transduced transplanted hepatocytes in all experimental groups	60
Figure 3.12 Transduced NIH3T3 control cells proved specificity of primer/probe combinations for detection of genes of interest	62
Figure 3.13 Integrated vector DNA quantified by droplet-digital PCR correlated well within groups	64
Figure 4.1 Radar chart with summarized study parameters	74

Figure S1 LeGO-LargeT-iV2 cloning strategy	IV
Figure S2 LeGO-C2 vector map.....	V
Figure S3 LeGO-V2 vector map	VIII
Figure S4 LeGO-Cer2 vector map	IX
Figure S5 LeGO-CyclinA2-iG2 vector map	XIV
Figure S6 LeGO-HRas-V12-iG2 vector map	XVIII
Figure S7 LeGO-LargeT-iV2 vector map.....	XXI
Figure S8 Titration examples for all vectors used in this thesis	XXV

List of Tables

Table 2.1 Peroxidase staining primary antibodies	40
Table 2.2 Overview of primary antibodies used for IF staining in this thesis.....	42
Table 2.3 Optimal dilution of secondary antibodies Alexa Fluor 488 and 633.....	42
Table 2.4 Minimum and maximum threshold adjustments of IHC and IF images.....	43
Table 3.1. Transduction efficiency of NIH3T3 cells as genetic control for molecular analysis measured 3 days post transduction	61
Table 4.1 Barcelona Clinic Liver Cancer staging system.....	72
Table 4.2 Customized Barcelona Clinic Liver Cancer staging system for evaluation of liver liver disease in USB mice	72
Table S1 Primer used for cloning and sequencing of LeGO-LargeT-iV2	I
Table S2 Primer/probe sets for amplification of gDNA by ddPCR for quantification of relative vector copy numbers (VCN) of integrated lentiviral vectors	I
Table S3 List of all constructs and titers [IP/mL] used for transduction experiments of primary murine hepatocytes or NIH3T3 cells	I
Table S4 Summary of experiments, recipient mouse strain, sex and MOI	II
Table S5 Liver weight of experimental groups Mock, RGB, CyclinA2, HRas-V12, and LargeT	III

1 Introduction

1.1 Hepatocellular carcinoma

Hepatocellular carcinoma (HCC) is the most common liver malignancy worldwide (Forner et al. 2012; Sia et al. 2017). With approximately 810,000 deaths in 2015, and with numbers increasing each year, it is the second leading cause of cancer-related deaths globally (Wang et al. 2016; online reference 1). HCC is 3 to 5 times more abundant in males compared to females. The disease risk is positively age-correlated (Forner et al. 2012) and has its highest prevalence in Africa and Asia with mortality rates of more than 12.7 cases per 100,000 inhabitants (Castelli et al. 2017; Fig. 1.1).

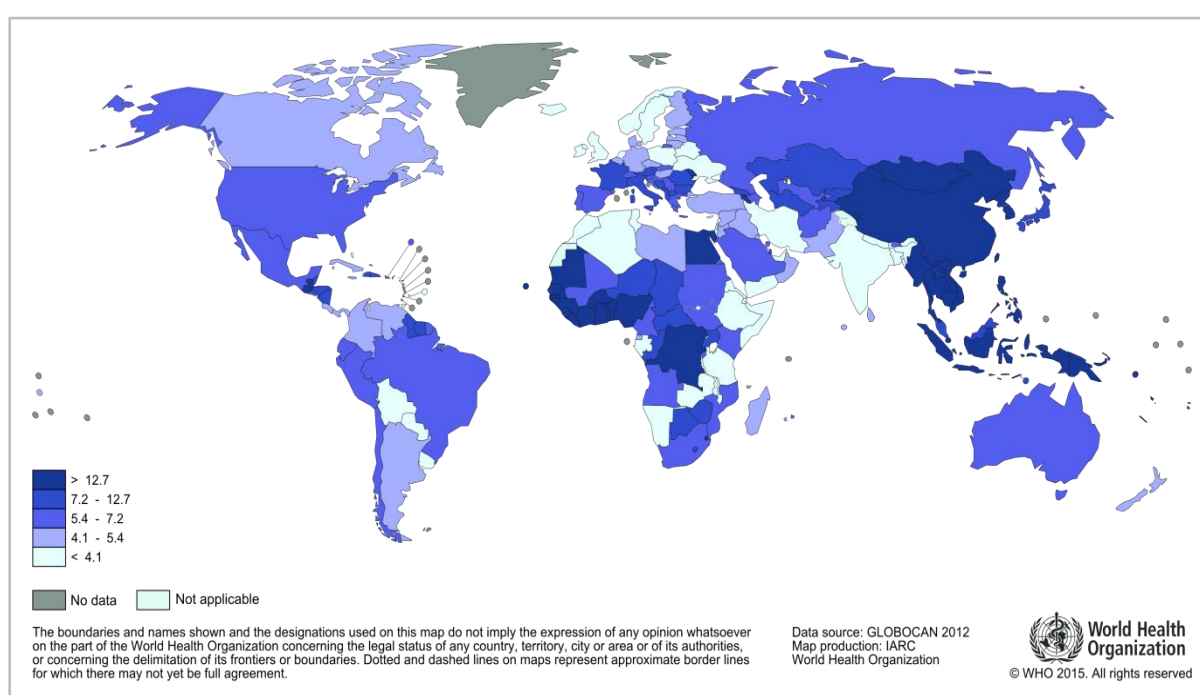


Figure 1.1 Age-standardized global mortality rates of male patients due to hepatocellular carcinoma in 2012. Incidences per 100,000 residents. Color modified (adapted from IARC 2017, online reference 1).

Region, incidence rate, sex and, most likely, ethnicity determine the global age distribution of HCC. In the United States or Canada, exemplary for low-risk populations, age-specific rates are highest among persons aged 75, and on the other hand, peak between ages 60 and 65 in high-risk African populations (El-Serag and Rudolph 2007). While HCC was less frequent in Europe and America during the past, incidence rates are rising and HCC has become of major interest in the whole world (Llovet and Bruix 2008). Although important risk factors for HCC development have been elucidated in recent years, the molecular mechanisms of HCC pathogenesis are still poorly understood (Ghouri et al. 2017).

1.1.1 Risk factors of liver disease

Already in 1953, Carl Nordling hypothesized that mutations in several genes are necessary for cancer development (Nordling 1953). Later in 1971, Alfred Knudson postulated the two-hit hypothesis concluding that cancer cells originate from cells with at least two pathological mutations (Knudson 1971). This applies also for liver malignancies (Marquardt et al. 2015). Hepatocarcinogenesis is a complex, long-lasting multistep process often associated with other liver diseases preceding the malignant transformation (Pellicoro et al. 2014; Fig. 1.2). HCC is the most prevalent inflammation-associated carcinoma, secondary to viral hepatitis infections, non-alcoholic steatohepatitis, autoimmune disorders, aflatoxin or liver cirrhosis (Ghouri et al. 2017; Pellicoro et al. 2014). Alcohol abuse is also promoting HCC development and has a synergistic effect in patients chronically infected with hepatitis B, hepatitis C, or both (de Lope et al. 2012; Pellicoro et al. 2014). Moreover, chronic long-term hepatitis B virus (HBV) infection is the most frequent risk factor for HCC, which pertains to more than 50% of all cases (Blumberg et al. 1975; Sherman 2010). Extended infection duration and increased viral load initiate latent HBV infection causing ongoing inflammation. Within this context, successively enhanced liver cell death, regeneration and repair result in premature liver senescence (Bonilla Guerrero and Roberts 2005). Associated with high cellular turnover, the chance for mutations is elevated promoting malignant transformation (Brecht et al. 2001; Chen et al. 2006; Forner et al. 2012).

Furthermore, the main characteristic for the pathomechanism of HCC is sustained hepatocyte damage by increased release of inflammatory mediators (Block et al. 2003). Hepatocytes usually persist in G₀-phase in adult livers, as proliferation occurs only at rare frequencies under physiological steady-state conditions, whereas enhanced proliferation is triggered during chronic liver inflammation (Ramboer et al. 2014). In addition, recurrent injury and perpetual wound healing caused by inflammatory damage can lead to hepatic stellate cell activation and matrix deposition initiating fibrosis (Block et al. 2003). In some cases though, removal of the underlying cause of fibrosis can resolve this state and patients' livers can return even to near-normal liver architecture. However, if fibrosis and apoptosis of parenchymal cells progress over many years, permanent liver damage will most likely result in cirrhosis (Block et al. 2003; Pellicoro et al. 2014; Fig. 1.2). Median time to develop cirrhosis has been estimated to be 30 years in chronic-hepatitis patients (Pellicoro et al. 2014). Obesity, epigenetic markers and genetic risk factors can promote disease progression. Patients often develop liver failure and portal hypertension. Furthermore, angiogenesis, and tumor macro- and microenvironment influence liver carcinogenesis (Castelli et al. 2017; Ghouri et al. 2017; Pellicoro et al. 2014). At advanced stages of disease, cancer cells can spread and induce metastases in the lung, portal vein, periportal nodes, bone, or brain.

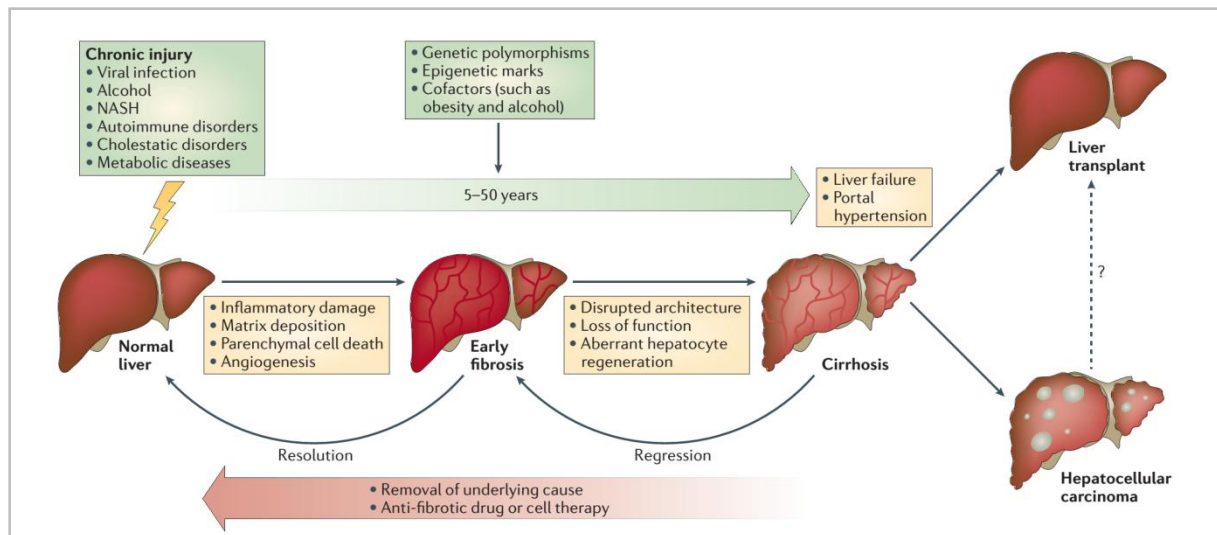


Figure 1.2 Progression of chronic liver disease. Viral infection, alcohol and non-alcoholic steatohepatitis are some of the causal factors inducing chronic injury in the liver. Recurrent inflammatory damage and matrix deposition, parenchymal cell death and angiogenesis induce progressive fibrosis. Fibrosis can be resolved if the cause of fibrosis is eliminated. Otherwise, cirrhosis can evolve due to ongoing inflammation and other cofactors minimizing the potential to reverse fibrosis and enhanced risk for developing hepatocellular carcinoma. So far, liver transplantation is the only existing treatment for liver failure. Adapted by permission from Springer Nature, Nature Reviews Immunology, *Liver fibrosis and repair: immune regulation of wound healing in a solid organ*, Pellicoro et al. © 2014 (license number: 4386521260649).

In summary, a combination of genetic, viral and environmental factors contributes to the complex and heterogeneous multistep process of hepatocarcinogenesis (Cha and Dematteo 2005). So far, there is no ultimate cure for HCC (Bruix et al. 2011; Forner and Bruix 2012).

1.1.2 Treatment options for HCC

To get optimal treatment for liver cancer patients, the severity of disease is categorized by a staging system based on patients' prognosis. There are several systems, but all aim at prolonging life expectancy based on optimal treatment customized for a given patient (Forner et al. 2010). Diagnostic confirmation of HCC can be performed either by biopsy, computer tomography (CT) or magnetic resonance imaging (MRI) (Bruix et al. 2011; de Lope et al. 2012). Patients diagnosed at an early stage can benefit from curative therapies as resection, transplantation and ablation, which today offer a potential cure (Forner et al. 2012; McKillop et al. 2006; Pellicoro et al. 2014). Surgical resection is exclusively applicable to non-cirrhotic patients as major resections are well tolerated (Kim et al. 2007). Nevertheless, about 50% to 80% of patients suffer a recurrence and die from it within 5 years after surgery (Blum 2005; Imamura et al. 2003; Kim et al. 2007). Feasibility of resection in cirrhosis is limited by liver function impairment if attempting to minimize morbidity and mortality. If cancer has already progressed too far, the remnant liver function could become insufficient and fail after surgery resulting in the death of the patient (de Lope et al. 2012). Intermediate-stage HCC can be treated with chemoembolization resulting in a delay of tumor progression (Bargellini et al. 2012; Llovet et al. 2002). Patients suffering from advanced HCC showing extrahepatic

dissemination, vascular invasion, or mild tumor-related symptoms with a preserved liver function have a median survival of about 6 to 8 months without treatment (de Lope et al. 2012; Llovet and Bruix 2008). These patients can benefit from sorafenib, which has anti-angiogenic and anti-proliferative effects (Cheng et al. 2009; Llovet et al. 2008). Sorafenib is a multikinase inhibitor affecting RAF (rapidly accelerated fibrosarcoma protein kinase), PDGF (platelet-derived growth factor), and VEGF (vascular endothelial growth factor) signaling, leading to an increased median life expectancy of 10 months (Bruix et al. 2011; Forner and Bruix 2012).

In most patients, HCC is detected at advanced stages leaving only palliative treatment as an option when cure is no longer considered possible (Llovet et al. 2003). Although early detection methods were introduced in recent years, treatment options have not been ameliorated to improve patient survival (Ghouri et al. 2017). Moreover, tumors are very heterogeneous and difficult to treat (Dhanasekaran et al. 2016). Thanks to analyses of HCC development in small-animal models and comparison of the results with human HCC data, insight was gained into tumor cell morphology and dysregulated signaling pathways involved in promotion of HCC (Ranzani et al. 2013; Rittelmeyer et al. 2013; Thorgeirsson et al. 2006).

1.1.3 Molecular profile in liver malignancies

The oncogenic potential of chronic inflammation ties back to alteration of many signaling cascades leading to a heterogeneous molecular profile in malignant degenerated cells (Forner et al. 2012; Liu et al. 2006; Meira et al. 2008; Sia et al. 2017). A gene involved in cell replication is termed proto-oncogene, when it has the potential to become an oncogene contributing to cancer, e.g., due to mutation or virus-mediated (over-)expression (Todd and Munger 2001). Continuously induced cell division through activation of various pro-proliferative signal cascades or inhibition of proliferation-regulating pathways can promote carcinogenesis (Meira et al. 2008). In 40-50% of liver cancer patients, the PI3K/AKT/mTOR (PI3K: phosphoinositide-3-kinase; AKT: protein kinase B; mTOR: mammalian target of rapamycin) pathway is dysregulated with regard to impaired upstream signaling e.g., by inactivation or mutation of the tumor suppressor PTEN (phosphatase and tensin homolog) (Carnero and Paramio 2014). Cell growth, survival, metabolism, and apoptosis are regulated by the PI3K/AKT/mTOR pathway (Forner et al. 2012). PI3K is triggered by EGFR (epidermal growth factor receptor) or IGFR (insulin-like growth factor receptor) binding (Singh et al. 2014; Fig. 1.3). Inactivation of PTEN results in constitutively activated AKT, which reduces cell death (Carnero and Paramio 2014; Toss and Cristofanilli 2015).

Additionally, in more than 50% of hepatocellular carcinomas, HRas (also known as transforming protein p21) signaling is activated (Liu et al. 2017; Mizushima et al. 2011). The GTPase HRas, encoded by the proto-oncogene *HRas*, is involved in cell proliferation

regulation in response to growth factor stimulation (Singh et al. 2014; Toss and Cristofanilli 2015; Fig. 1.3). Somatic mutations in the gene are known to direct unrestricted cell growth (Forner et al. 2012). One of the most common point mutations leads to an amino acid change from glycine to valine at residue 12 (HRas-V12) (Barbacid 1987; Bos 1989; Xin et al. 2017). This results in a permanently activated protein causing uncontrolled cell division and tumor development (Bos 1989). Already in 1989, mutated HRas was applied in mouse models and caused liver tumors (Sandgren et al. 1989).

Also, other cancer genes and cell-cycle-related genes are known to be dysregulated in liver tumors (Deshpande et al. 2005; Lee and Thorgeirsson 2006; Marquardt et al. 2015). Five percent of HCC cases show overexpression of *CCNA* (CyclinA) (Block et al. 2003; Sia et al. 2017). CyclinA1 is expressed during meiosis and embryogenesis, whereas CyclinA2 is present in dividing somatic cells and has functions in both S-phase and mitosis (Yam et al. 2002; Fig. 1.3). It regulates cell cycle progression by interaction with two different Cyclin-dependent kinases (CDK) (Chao et al. 1998; Yam et al. 2002). CDK1 is bound during the transition from G2 to M phase and CDK2 during S phase (Pagano et al. 1992). Overexpression of *CCNA2* contributes to high proliferative activity in affected cells (Wang et al. 2005). Phosphorylation of tumor suppressors like p53 or oncoproteins induced by CyclinA2-CDK can also lead to tumorigenesis (Wang and Prives 1995).

The tumor suppressor p53 is a down-stream target of PI3K/AKT/mTOR and Ras signaling, and is involved in cell cycle progression, proliferation and transformation (Ahuja et al. 2005; Toss and Cristofanilli 2015). p53 is often mutated in HCC tissue (Block et al. 2003; Hsu et al. 1993; Fig. 1.3). Furthermore, p53 was discovered as cellular protein bound to LargeT-antigen in SV40-transformed cells (SV40: simian vacuolating virus 40) (Ahuja et al. 2005; Lane and Crawford 1979; Linzer and Levine 1979). SV40 is a tumor-inducing DNA virus of the genus *Polyomavirus* that is dormant in rhesus monkeys but can also infect other species such as humans (Ahuja et al. 2005). The proto-oncogene LargeT-antigen is involved in its viral genome replication and regulation of host cell cycle and was previously connected with several malignancies (Ahuja et al. 2005; Javier and Butel 2008). The function of p53 in the regulation of gene expression or apoptosis can be blocked by multiple redundant mechanisms of SV40's LargeT-antigen (Bargonetti et al. 1992; Wang and Yang 2010; Fig. 1.3). Moreover, several groups have shown the cell transformation capabilities of SV40 LargeT-antigen (Giri and Bader 2014; Holczbauer et al. 2013; Pan et al. 2015). LargeT-antigen-treated cells become immortal, proliferate and survive contrary to untreated normal cells (Giri and Bader 2014; Meng et al. 2014). Treated cells are also able to induce tumor formation in transplanted mice (Ahuja et al. 2005; Lou et al. 2005; Sandgren et al. 1989).

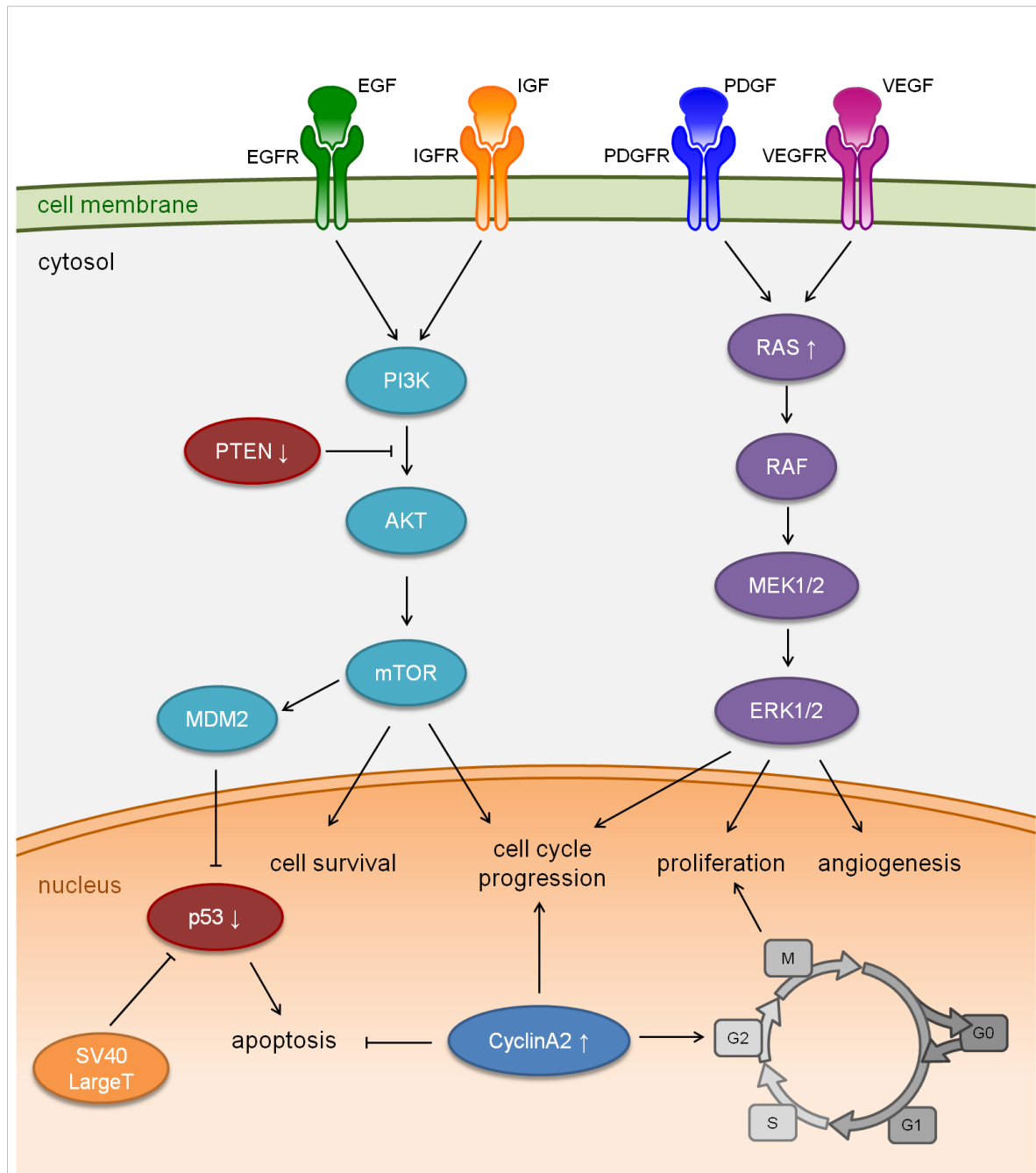


Figure 1.3 Cellular signaling pathways involved in the normal cell cycle and the pathogenesis of HCC. AKT, protein kinase B; EGF, epidermal growth factor; EGFR, EGF receptor; ERK 1/2, extracellular signal-regulated kinases 1/2; IGF, insulin-like growth factor; IGFR, IGF receptor; MDM2, E3 ubiquitin ligase; MEK, Mitogen-activated protein kinase; mTOR, mammalian target of rapamycin; p53, tumor suppressor; PDGF, platelet-derived growth factor; PDGFR, PDGF receptor; PI3K, phosphoinositide-3-kinase; PTEN, phosphatase and tensin homolog; RAF, rapidly accelerated fibrosarcoma protein kinase; RAS, GTPase, transforming protein 21; SV40 LargeT, simian virus 40 large T antigen; VEGF, vascular endothelial growth factor; VEGFR, VEGF receptor; white arrow down: down regulation; white arrow up: up regulation. Not to scale. Based on Ahuja et al. 2005; Carnero and Paramio 2014; Deshpande et al. 2005; Liu et al. 2017; Singh et al. 2014; Toss and Cristofanilli 2015.

These examples demonstrate the diverse molecular profile in hepatocellular carcinomas. Nevertheless there are also other sources of liver cancer induction like viruses. In fact, investigations on cancer-causing viruses even have started more than 100 years ago, and

many methods have been implemented how to modify and utilize viruses in science, especially in cancer research.

1.1.4 Role of viruses in carcinogenesis

Many viruses have been established as the cause or a contributing factor of various cancer types (Javier and Butel 2008). Research on tumor virology was initiated at the beginning of the 20th century. Peyton Rous' discovery of virus-mediated sarcoma in Plymouth Barred Rock fowls with potential intra- and interspecific transmission of the (today known) Rous sarcoma virus (RSV) was the first evidence for viruses with tumorigenic potential (Rous 1910; Rous 1911). During the next decades, tumor-inducing viruses were identified in rabbits, mice, cats and non-human primates (Javier and Butel 2008; Shope and Hurst 1933; Weiss and Vogt 2011). This resulted in extensive research regarding carcinogenesis-promoting viruses in humans. The Epstein-Barr virus was the first oncogenic human virus discovered by Tony Epstein, Yvonne Barr and colleagues (Epstein et al. 1964; Henle and Henle 1966). Later, studying retroviruses like RSV and associated animal tumors led to discovery of tumor-inducing genes, so-called oncogenes by David Baltimore, Renato Dulbecco and Howard M. Temin, who were rewarded with the Nobel Prize in 1975 (online reference 2). For their discovery of the cellular origin of retroviral oncogenes, the Nobel Prize in Physiology or Medicine 1989 was awarded jointly to J. Michael Bishop and Harold E. Varmus (online reference 3). Four years later, the human T cell leukemia-virus type 1 (HTLV-I) was described as first tumorigenic human retrovirus (Poiesz et al. 1980). Retroviruses were subjected to critical scrutiny after the AIDS-causing (acquired immune deficiency syndrome) human immunodeficiency virus type 1 (HIV-1) was identified the early 1980s (Barre-Sinoussi et al. 1983; Coffin et al. 1986). In 2008, Luc Montagnier and Françoise Barre-Sinoussi won the Nobel Prize for discovery of the virus (online reference 4). Although no known cancer is ultimately induced by HIV, other viruses were known to cause cancer in AIDS patients. Kaposi's sarcoma had been characterized as one of the AIDS-defining diseases and also has carcinogenic potential in skin, lymph nodes and other organs (Chang et al. 1994; Kaposi 1872). Later, co-infection of HIV-1 and HBV or hepatitis C virus (HCV) was suggested to facilitate malignant transformation of tumors e.g., in liver cancer patients (Cheruvu et al. 2007; Sherman et al. 2015).

As already mentioned, hepatitis B virus was identified as link to human hepatocellular carcinoma (Blumberg et al. 1975; Bonilla Guerrero and Roberts 2005). Interestingly, randomly distributed integrations of HBV sequences were found at high frequencies in HBV-related hepatocellular carcinomas (<90%) (Bonilla Guerrero and Roberts 2005). In these tumor cells enduring oncogenic damage promotes carcinogenesis according to specific overexpression of the HBx antigen and its impact on signaling cascades benefitting cell

survival and proliferation. Overall, the data indicate HCC development, induced in the context of a HBV infection, to be the result of multiple cooperative events (Bonilla Guerrero and Roberts 2005; Chen et al. 2010). Although the role of insertional mutagenesis by HBV remains elusive, it is a promising tool to investigate molecular mechanisms of liver carcinogenesis. Other independent observations with adeno-associated vectors (AAV) and more importantly with lentiviral vectors led to the suggestion of insertional mutagenesis to induce malignant transformation in hepatocytes (Donsante et al. 2007; Ranzani et al. 2013).

1.2 Lentiviral vectors

Retroviridae are a family of well-characterized viruses with a relative simple RNA genome and which have been usurped as vectors for gene transduction in widely different settings. For instance, retroviral vectors were used in several gene therapy studies (Naldini 2015) (e.g., in severe combined immunodeficiency-X1 (SCID-X1) (Cavazzana-Calvo et al. 2000; Wu and Dunbar 2011) or Wiskott-Aldrich syndrome (Boztug et al. 2010; Braun et al. 2014)). On the other hand, vectors have been used to transfer genetic information in order to induce a malignancy (Ranzani et al. 2013; Thomaschewski et al. 2017). Lentiviral vectors originate from lentiviruses, which are also part of the *Retroviridae* family and belong to Group VI viruses. These diploid single-stranded RNA (ssRNA) retroviruses encode a reverse transcriptase to produce a DNA intermediate in their replication cycle, and therefore were named after this characteristic enzyme. The principle of an RNA-dependent DNA synthesis and stable integration of viral cDNA was first described independently by two groups in 1970 (Baltimore 1970; Temin and Mizutani 1970). The *Retroviridae* family is divided in the two subfamilies Orthoretrovirinae and Spumaretrovirinae, including the seven genera *Alpha*-, *Beta*-, *Gamma*-, *Delta*-, *Epsilon*-, *Spumaretrovirus*, and *Lentivirus*. The viruses of all genera are characterized by differences in their pathogenicity, morphology and genetics. Depending on their genome, retroviruses are classified as simple or complex (Gabriel et al. 2012). Both classes carry the structural genes *gag*, *pol* and *env* (see below) and regulatory elements, whereas only complex viruses contain additional accessory genes for transcriptional activation of viral and host genes (Escors and Breckpot 2010; Freed 2001). These features facilitate the use of the different retrovirus genera as vehicle for gene delivery. For a better understanding how lentiviral vectors can be deployed in liver cancer research, some detailed information will be given in the following.

1.2.1 *Lentivirus* virion structure and genome organization

A lentivirus is a complex retrovirus with a diameter of 80-100 nm and contains two copies of a linear positive-sense ssRNA genome with a 5' cap and 3' polyadenylated tail densely packaged in the nucleocapsid, which protects RNA from degradation by cellular enzymes

(Escors and Breckpot 2010; Fig. 1.4). The capsid surrounds the RNA genome and essential enzymes for viral replication together with accessory enzymes. The capsid is, in turn, enclosed by a matrix protein layer, directly adjoining a cell-membrane derived lipid bilayer. The transmembrane and surface glycoproteins are anchored into the bilayer and form the lipid envelope (Fanales-Belasio et al. 2010).

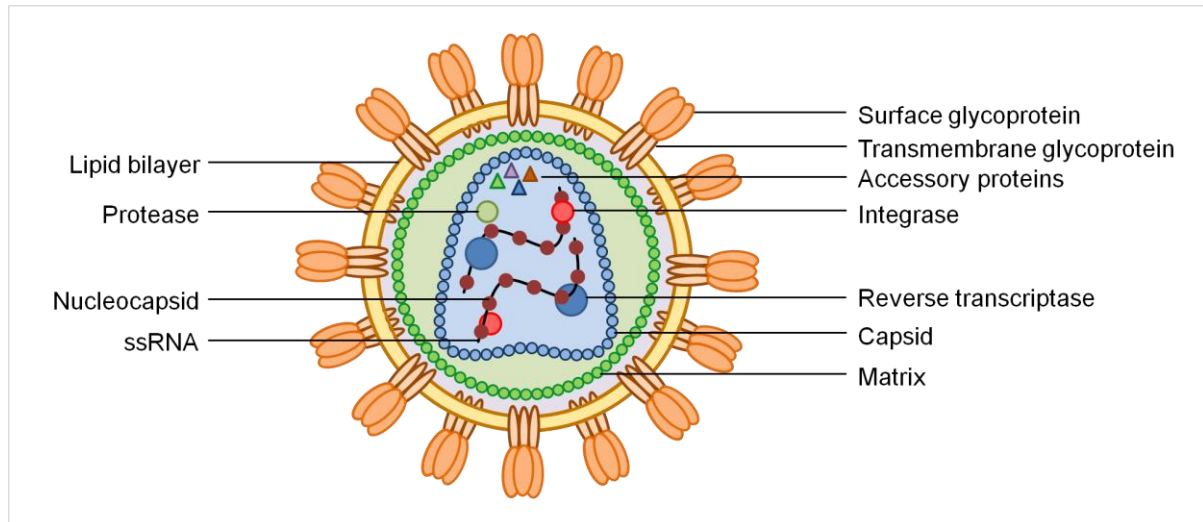


Figure 1.4 HIV-1 schematic exemplary of a *Lentivirus* virion. The surface and transmembrane proteins are anchored within the lipid bilayer. Matrix proteins surround the capsid, which encloses enzymes (accessory proteins, reverse transcriptase, protease, and integrase) required for integration and replication, as well as two copies of the ssRNA genome protected by nucleocapsid proteins. Not to scale. Based on Escors and Breckpot 2010.

The genome of complex lentiviruses (e.g., HIV-1) is organized in three different parts: 1) structural, 2) regulatory, and 3) accessory genes (Freed 2001). Transcription of the genome happens in different splice variants also using three reading frames (Petropoulos 1997). All retroviruses encode the structural genes *gag*, *pol* and *env* (O'Keefe 2013). The group-specific antigen (Gag) protein is a polyprotein, which is cleaved by the viral protease into the matrix (p17), capsid (p24), and RNA-binding nucleocapsid proteins (p7) (Kuzembayeva et al. 2014). The *pol* (polymerase) gene encodes the enzymatic proteins protease (p10), integrase (p32), reverse transcriptase (p51), and ribonuclease (p15), and is translated into a gag-pol polyprotein (Barmania and Pepper 2013). These proteins are required for DNA integration and viral replication in the host. Both polyproteins Gag and Gag-pol are translated from unspliced viral mRNA transcripts. The polyprotein precursor encoded by *env* (envelope) is cleaved into the envelope transmembrane subunit (gp41) facilitating membrane fusion during infection and the surface subunit (gp120), which is important for receptor binding and virus entry (Escors and Breckpot 2010). Regulatory elements include the following: *trans*-activator (Tat) and regulator of viral protein expression (Rev), both important for viral transcription and nuclear export (Escors and Breckpot 2010). The long terminal repeats (LTR), including the unique 3' region (U3), a redundant sequence (R) and unique 5' region (U5), as well as the polypurine tract (PPT) are required for viral

transcription, reverse transcription and integration (Escors and Breckpot 2010; Kuzembayeva et al. 2014). The U3, R, and U5 contain promoter regions and present binding sites for cellular transcription factors (Escors and Breckpot 2010). As its name says, the **packaging signal** (ψ) is required for packaging but also for genome dimerization (Sakuma et al. 2012). Accessory proteins (Vif, Vpr, Vpu, Nef) play a role in virulence (Fanales-Belasio et al. 2010).

To use lentiviral vectors in research, understanding and manipulating their replication cycle is essential.

1.2.2 Replication cycle of lentiviruses

The replication cycle of lentiviruses is similar to that of other retroviruses and has eight major steps (Barmania and Pepper 2013; O'Keefe 2013; Fig. 1.5). First, the virion's envelope gp120 binds the cellular receptor of the host cell, e.g., HIV-1 targets CD4⁺ T cells. The resulting conformational change in the envelope exposes the transmembrane subunit gp41 followed by fusion of the virion and host cell membrane. Second, the core is released into the cytoplasm, partially uncoats, and the *trans*-RNA genomes are reverse transcribed in a third step. The reverse transcriptase has a ribonuclease H active site and therefore is able to synthesize proviral complementary DNA (cDNA) with a (+)RNA/(-)DNA intermediate. Then, it breaks down the RNA strand from the RNA/DNA hybrid double helix, continuing the transcription into proviral double-stranded (ds) cDNA (dscDNA) (Fanales-Belasio et al. 2010). The infection goes on with the transport of the pre-integration complex (PIC) into the nucleus of the host cell. Importantly, complex retroviruses are able to infect both dividing and non-dividing cells whereas dscDNA of simple retroviruses can only enter the nucleus during mitosis, when the nuclear membrane is decomposed (Escors and Breckpot 2010). Once the viral dscDNA has reached the nucleus, it can stay in a circular form in a latency phase without integration or, in the fourth major step, it is stably integrated into a host cell chromosome by the integrase. This enzyme cleaves cellular dsDNA and mediates integration and ligation of viral dscDNA (Coiras et al. 2009). In the fifth step, the integrated form of the retrovirus, the so-called 'provirus' is transcribed and spliced by the host nuclear machinery into different mRNA splice variants including unspliced mRNA. Sixth, viral mRNAs are exported into the cytoplasm and translated into structural and regulatory viral proteins regulated by the viral protein Rev. Exported, unspliced mRNA serves as future ssRNA genome of new matured virions. Assembly of viral proteins and genomic RNA at the cellular membrane follows in the seventh step. In the final eighth step, new virions bud and finish the maturation process through viral protease activity (Coiras et al. 2009).

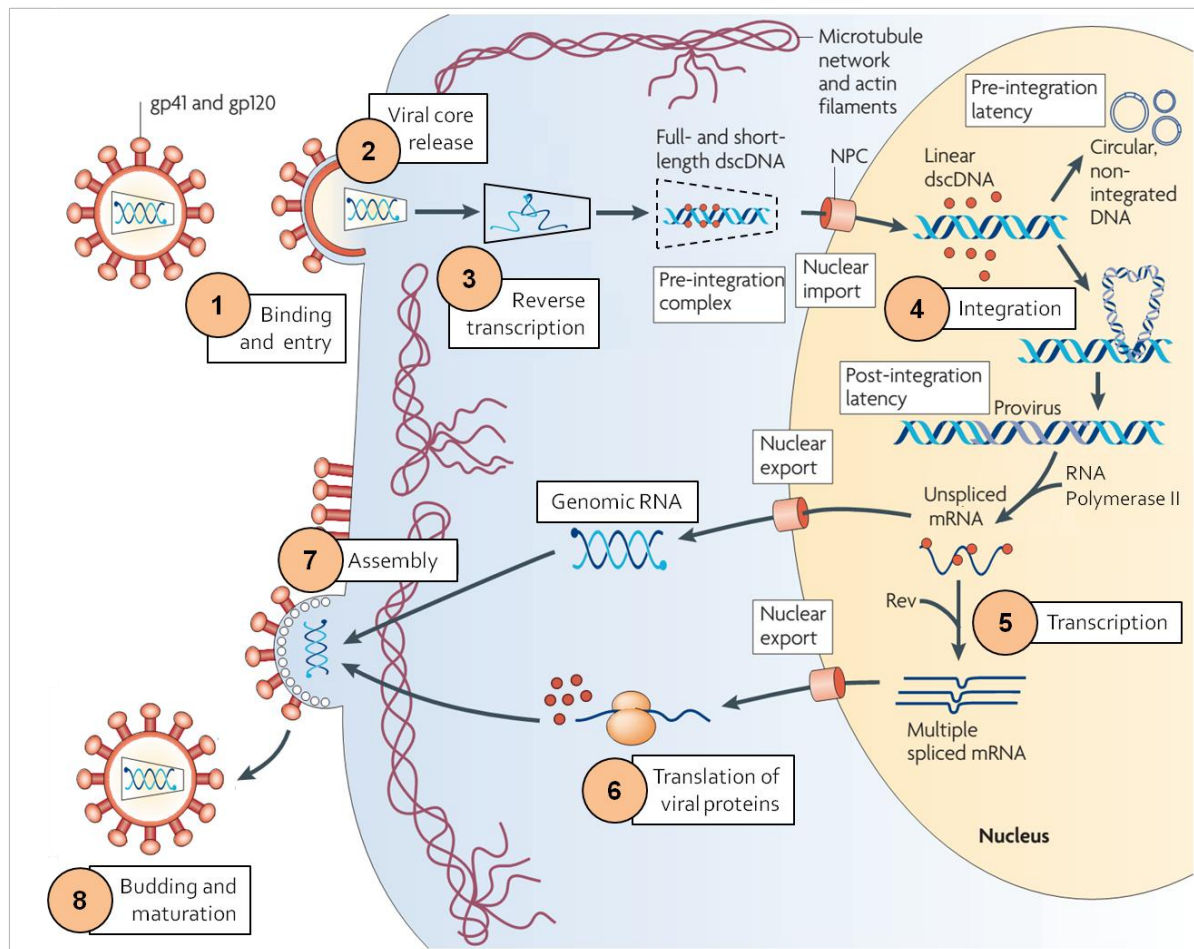


Figure 1.5 Replication cycle of HIV-1. The most important eight steps in the replication of the lentivirus HIV-1 are 1) binding and entry, 2) viral core release, 3) reverse transcription of ssRNA genome into dscDNA, 4) provirus integration, 5) transcription of spliced and unspliced viral mRNA, 6) export of mRNA and translation of viral proteins and 7) assembly of genomic RNA and viral proteins, and 8) budding and maturation. Not to scale. Adapted and modified by permission from Springer Nature, Nature Reviews Microbiology, *Understanding HIV-1 latency provides clues for the eradication of long-term reservoirs*, Coiras et al. © 2009 (license number: 4386590624714).

When working with lentiviral vectors, mostly viral genes are removed and LTRs are modified to prevent steps 6 to 8 from proceeding as maturation and replication of virions is not desired to protect the host cells from re-infection. In all applications lentiviral vectors are used, proviral DNA integration for stable transgene expression is aimed to be the final step (Sakuma et al. 2012).

1.2.3 Lentiviral integration preferences and insertional mutagenesis

As outlined above, integration of proviral DNA into the host genome is one essential step for replication of retroviruses (Coiras et al. 2009). Retroviral integration is a semi-random process (Cattoglio et al. 2010). Different retroviruses show diverse features regarding recognition of specific cellular receptors by the viral envelope (tropism) and integration patterns of proviral DNA (Felice et al. 2009; Gabriel et al. 2012; Suerth et al. 2012). The pre-integration complex (PIC) recognizes cellular components or host cell chromatin, which

varies between the different virus genera, and results in specific integration preferences (De Ravin et al. 2014; Felice et al. 2009; Florence and Faller 2001). Favored integration sites of lentiviral vectors are within active genes (Schröder et al. 2002; Suerth et al. 2012). Integrases of lentiviruses hijack host proteins to direct the PIC to specific chromatin contexts. Since integration can occur within any active gene, this process is unpredictable (Baum et al. 2003; Schröder et al. 2002). By definition, the insertion of additional foreign genetic information – in this case the provirus – into the host cell genome is mutagenic (Fehse and Roeder 2007). Particularly, insertional mutagenesis is defined as the introduction of any foreign DNA into preexisting host DNA. Multiple integrations favor insertional transformation (Modlich et al. 2005).

Advantages of lentiviral vectors include efficient gene delivery and constant transgene expression by stable integration into the host cell genome not only in dividing but also non-dividing cells (Baum et al. 2003). Nevertheless, various categories of risks related to vector-mediated genetic manipulation and transgene expression are known, e.g., genotoxicity, phenotoxicity, or dose-dependent side effects (Baum et al. 2003; Baum et al. 2006). Using replication-defective self-inactivating (SIN) lentiviral vectors in preclinical models, it was discovered that the insertion site preferences of SIN lentiviral vectors are most likely safer than of other virus-derived vectors (Baum et al. 2006; Modlich et al. 2009). Additionally, random generation of replication-competent retroviruses was considered a main safety concern (Temin 1990), this concern is alleviated by use of SIN lentiviral vectors. In general, some optimization steps were applied to avoid the risk of mutagenesis, and protocols are still modified especially since severe adverse events were observed in gene therapy trials with lentiviral vectors (Baum et al. 2006; Baum et al. 2004; Hacein-Bey-Abina et al. 2003; Uren et al. 2005; Vranckx et al. 2016).

1.2.4 Optimization of lentiviral vectors

Comparative analysis in mouse models revealed lentiviral vectors to be safer than *Gammaretrovirus*-derived vectors (Modlich et al. 2009; Wu and Dunbar 2011). To generate safe and effective lentiviral vectors for stable gene transfer, several optimizations were implemented to the wildtype form, pseudotyping included (Sakuma et al. 2012). Pseudotyping refers to the use of an envelope protein other than the original wildtype envelope (Coffin et al. 1997). By changing the viral envelope protein the cell tropism of the vector can be altered to target other cell types allowing a broader range of applications (O'Keefe 2013). Nowadays, most of the constructs used are second- or third-generation self-inactivating vectors lacking major viral genes and the capability to replicate (Sakuma et al. 2012; Fig. 1.6).

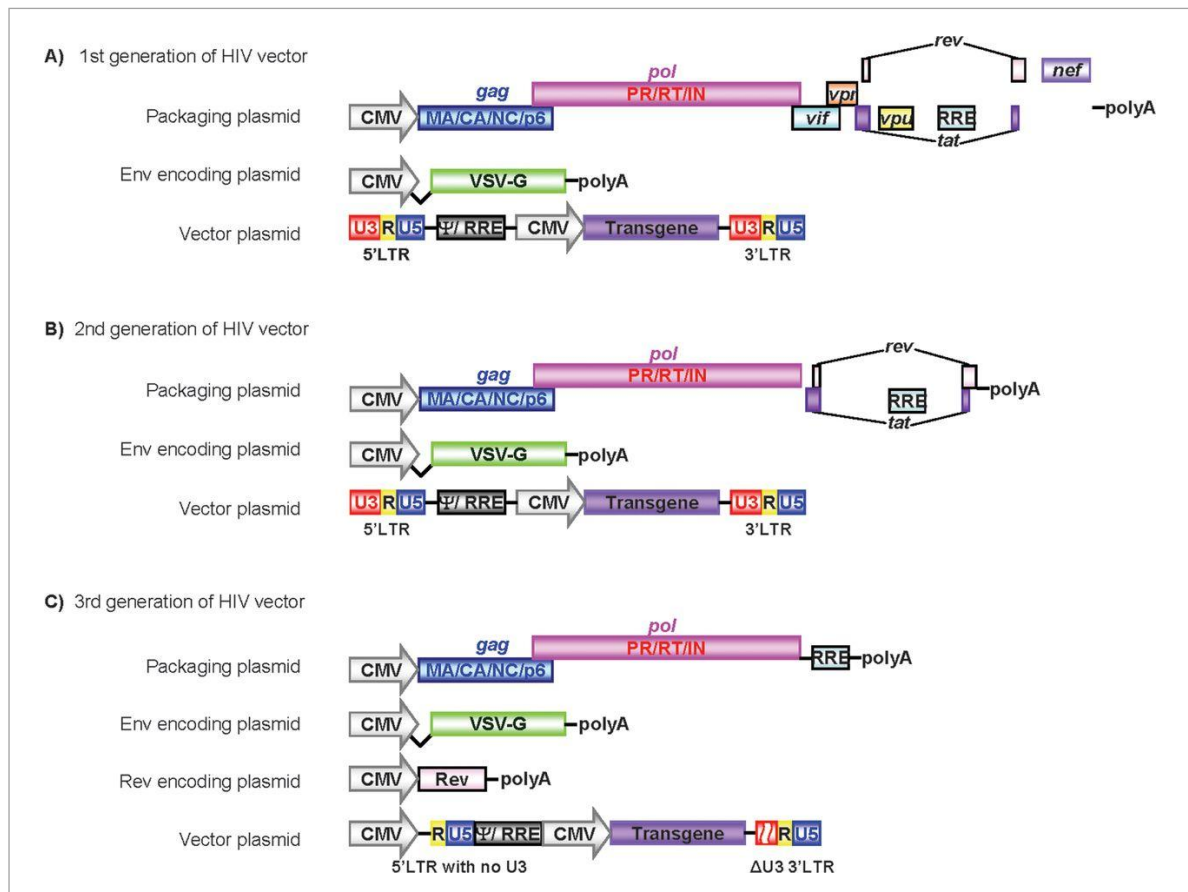


Figure 1.6 Scheme of HIV-derived lentiviral vectors. Three plasmids were used in first-generation HIV vectors. A packaging plasmid contained all viral proteins except Env protein. VSV-G, used to replace the HIV Env, is encoded by a second plasmid, together with a third, vector, plasmid containing the necessary *cis*-elements (LTR, RRE, ψ) and the transgene under a CMV promoter (A). For second-generation HIV vectors, all accessory proteins were removed from the packaging plasmid. Envelope and transgene were encoded by two different plasmids, similar to the first-generation vectors (B). Different from the first two generations, third generation of HIV vectors involved four different plasmids. An additional Rev protein was provided, encoded by a fourth plasmid. Rev is essential for regulation of viral protein expression and export of mRNA from the nucleus. In the vector plasmid, the U3 region was deleted from 5'-LTR together with partial deletion (Δ) of the 3'-LTR (so-called self-inactivating or SIN design) to provide replication-deficient viruses. A strong promoter such as CMV was inserted for expression of the vector (C). CMV, Cytomegalovirus promoter; Δ U3, deleted U3 region in the 3'-LTR; Env, envelope; LTR, long-terminal repeat; Ψ , psi, packaging signal; Rev, regulator of viral protein expression; RRE, Rev-responsive element. Figure was originally published in Biochemical Journal, *Lentiviral vectors: basic to translational*, Sakuma et al. © 2012.

Basic concept of these vectors is the local separation of the structural genes *gag* and *pol* from *env*, encoded by single vectors, and removal of viral accessory proteins as well as the deletion of the majority of their U3 region (Baum et al. 2006; Dull et al. 1998; Modlich et al. 2006). Probably, the deletions of viral enhancer and promoter elements were the most important modifications to generate safer vectors. To produce third-generation vectors four plasmids are used: 1) a *gag/pol* containing packaging plasmid; 2) an *env* plasmid; 3) a *rev* expression plasmid; and 4) a vector plasmid (Sakuma et al. 2012; Fig. 1.6).

Additional modifications, such as RRE (Rev-responsive element), cPPT (central polypurine tract), and wPRE (woodchuck posttranscriptional regulatory element), can be introduced to the vector plasmid for enhanced efficiency *in vitro* and *in vivo* (Sakuma et al. 2012;

Schambach et al. 2006a). Transgene expression of stably integrated lentiviral vectors can be influenced by the chosen promoter (Kostic et al. 2003; Kustikova et al. 2009; Modlich et al. 2009). Viral promoters often show stronger activities compared to physiological promoters and are favored to induce high expression levels correlating with a higher risk for insertional mutagenesis (Knight et al. 2010; Schambach et al. 2006b; Zychlinski et al. 2008).

Overall, modified third-generation self-inactivating vectors reduce the likelihood of LTR-driven transcriptional interference and suppression, insertional activation of adjacent cellular proto-oncogenes, and dissemination of spontaneously assembled replication-competent HIV-like viruses (Sakuma et al. 2012). On the other hand, high vector copy numbers resulting from multiple integrations increase the potential risk of insertional mutagenesis inducing malignant transformation (Fehse et al. 2004; Woods et al. 2003). Desired vector copy numbers per target cell can be achieved by altering the multiplicity of infection (MOI) considering the determined vector titers (Hamaguchi et al. 2000; Modlich et al. 2006).

This harnessing of eventual danger of undesired multiple integrations is the ultimate key to develop a reliable *in-vivo* model for hepatic tumorigenesis (Fehse et al. 2004).

1.2.5 RGB marking with LeGO vectors

Third-generation lentiviral SIN vectors were used to create the so-called LeGO (**l**entiviral **g**ene **o**ntology) system based on a building-block principle (Weber et al. 2008; Fig. 1.7).

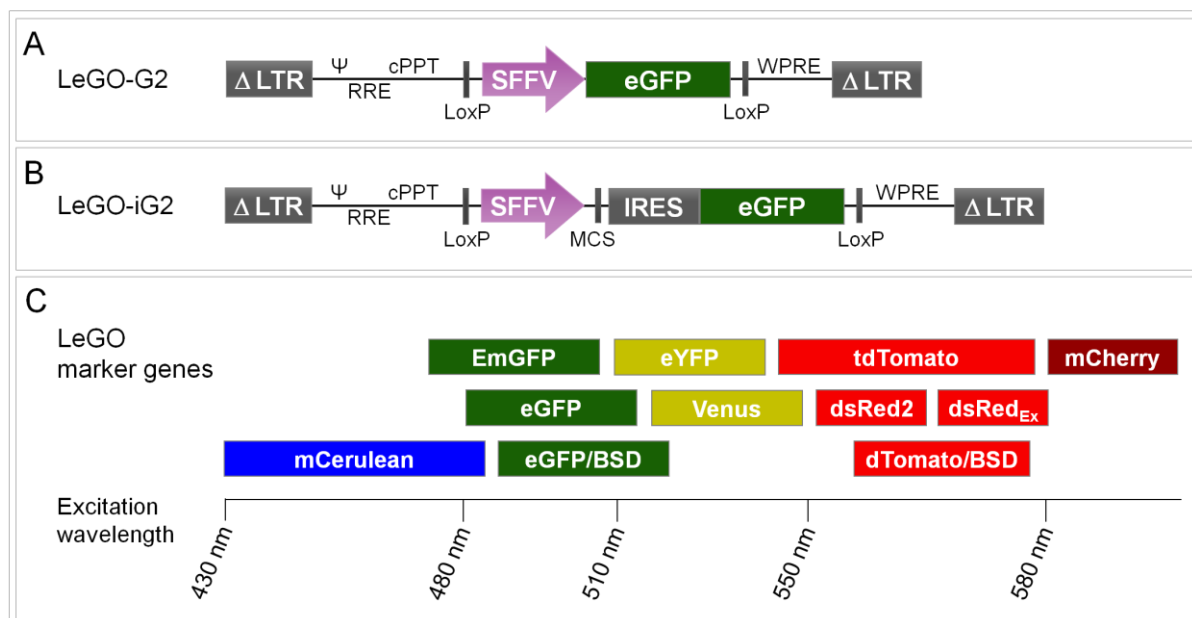


Figure 1.7 Lentiviral gene ontology (LeGO) vector principle. LeGO-G2 encodes an eGFP driven by an SFFV promoter and is used for cell marking (A). LeGO-iG2 encodes for the SFFV promoter, MCS, IRES, and eGFP, allowing for insertion of a gene of interest between MCS and IRES (B). Several marker genes can be chosen for LeGO vectors, all of them with different excitation wavelengths (C). cPPT, central polypurine tract; ΔLTR, self-inactivating-long-terminal repeat; eGFP, enhanced green fluorescent protein; EmGFP, emerald green fluorescent protein; eYFP, enhanced yellow fluorescent protein; IRES, internal ribosomal entry site; LoxP, Cre-recombinase restriction site; MCS, multiple cloning site; Ψ, psi, packaging signal; RRE, Rev-responsive element; SFFV, spleen focus-forming virus; WPRE, Woodchuck hepatitis virus post-transcriptional regulatory element. Not to scale. Based on Weber et al. 2008.

All elements of the vector, such as promoter, marker gene, or gene of interest can be exchanged using flanking unique restriction sites (Weber et al. 2008). In this system, a broad variety of fluorescent proteins can be used according to the desired color. LeGO vectors are a flexible system to introduce several genes within one vector. Simultaneous expression and/or suppression of genes can be achieved in one target cell taking advantage of the multiple cloning site (MCS) in combination with the internal ribosomal entry site (IRES) and a fluorescence protein (Weber et al. 2008; Weber et al. 2010). Considering color theories, simultaneous cell transduction with three LeGO vectors encoding red, green, or blue fluorescent proteins each results in a variety of possible combinations of mixed colors dependent on the expression strength of the individual colors (as a function of copy numbers and insertion sites). This principle is referred to as red-green-blue (RGB) marking (Weber et al. 2012; Weber et al. 2011). RGB marking allows for clonal cell tracking since color codes remain stable after cell division (Weber et al. 2011). Various cell types can be efficiently transduced with LeGO vectors leading to high expression levels of single and mixed fluorescence proteins as shown in Figure 1.8.

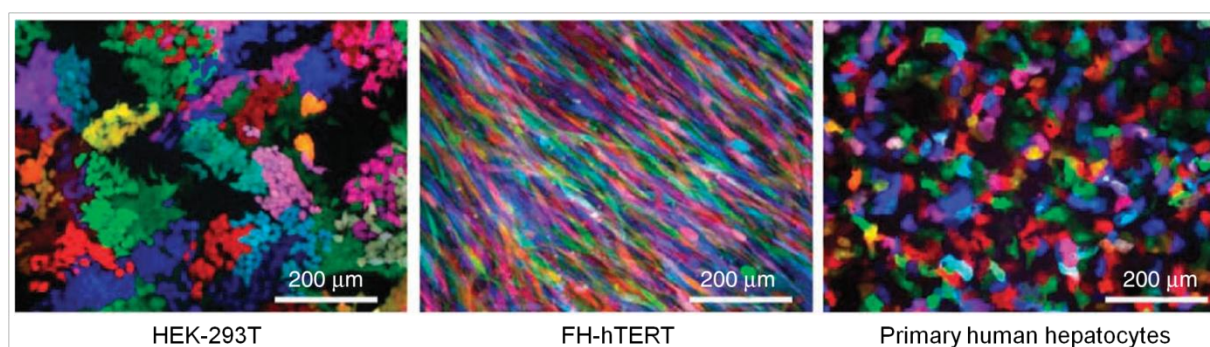


Figure 1.8 RGB marking with LeGO vectors is efficient in different cell types. HEK-293T, FH-hTERT and primary human hepatocytes were successfully transduced by LeGO-C2, LeGO-V2 and LeGO-Cer2 resulting in multiple differently colored cell clusters. C2: fluorescent protein mCherry; Cer2; fluorescent protein Cerulean; FH-hTERT: human telomerase reverse transcriptase-immortalized human fetal hepatocytes; HEK-293T: human embryonic kidney cells-derived cell line; LeGO: lentiviral gene ontology; V2: fluorescent protein Venus. Adapted and modified by permission from Springer Nature, Nature Medicine, *RGB marking facilitates multicolor clonal cell tracking*, Weber et al. © 2011 (license number: 4386610526793).

This principle is applicable to *in-vitro* and *in-vivo* experiments (Gomez-Nicola et al. 2014; Weber et al. 2012) and recently has been refined to optical barcoding (Mohme et al. 2017). Together with genetic barcoding, RGB marking was successfully used to analyze clonally regenerated humanized mouse livers and to characterize tumors (Cornils et al. 2014; Thomaschewski et al. 2017). To conclude, RGB marking with lentiviral vectors is a promising tool for gene marking and cell tracking in general, but also specifically in the liver.

1.3 Liver repopulation models

As there is no cure for advanced-stage HCC patients, strategies need to be developed and evaluated to improve patient survival. Transplantation of primary adult hepatocytes is thought

to be an alternative for orthotopic liver transplantation, since liver resection is not a feasible treatment in patients with advanced HCC and matching donors for liver transplantation are only available for a small proportion of patients (de Lope et al. 2012). In clinical trials, less than 5% of total liver mass were revealed as the maximum transplantable mass to reduce the risk of substantial intraportal cell infusion (Fox et al. 1998). To improve results after hepatocyte transplantation, competition between resident and transplanted hepatocytes had to be eliminated. In mouse models, the proportion of repopulated liver mass was increased by providing a selective advantage for transplanted hepatocytes over resident cells (Gilgenkrantz 2010). This finding emphasizes that a well working *in-vivo* model is required to assess the study of liver regeneration, tumor formation and to investigate treatment options.

During the past 20 years, different strategies have been examined to develop novel mouse models (Gilgenkrantz 2010). Application of chemicals prior to intrasplenic transplantation could be shown to enhance repopulation rates (Joseph et al. 2006; Kusano and Mito 1982). For example, the plant-derived pyrrolizidine alkaloid monocrotaline (MCT) in combination with carbon tetrachloride (CCl₄) was shown to induce endothelial toxicity in the lung, liver and kidney (Joseph et al. 2006; Scholten et al. 2015; Shah et al. 2005).

Another well-working approach is liver-specific transgene expression or targeted gene deletion in mouse models e.g., *Fah*^{-/-}/*Rag2*^{-/-}/*Il2*^{-/-}, and *Alb-uPA*^{+/+}/*Rag2*^{-/-} (Gilgenkrantz 2010; Overturf et al. 1996; Petersen et al. 1998). The role of insertional mutagenesis induced by lentiviral vectors in serial transplantation experiments into adult fumarylacetoacetate hydrolase-deficient immune-compromised (*Fah*^{-/-}/*Rag2*^{-/-}/*Il2*^{-/-}) mice was examined by Rittelmeyer and colleagues (Rittelmeyer et al. 2013). Polyclonality remained intact after four generations, but no tumor formation was observed in long-term studies. However, Ranzani and colleagues were able to identify oncogenes involved in murine and human HCC formation using lentiviral vectors in neonatal mice (Ranzani et al. 2013).

1.3.1 USB mouse model

Different mouse models were described to analyze repopulation and regeneration of the liver by transplantation of primary hepatocytes (Gilgenkrantz 2010). One of those models is the uPA mouse model, in which the *urokinase-type plasminogen activator* (uPA) transgene has been backcrossed with immunodeficient mice (Mercer et al. 2001; Petersen et al. 1998). In 1990, it was established to study neonatal bleeding disorders (Heckel et al. 1990). The characteristic of this model is the albumin enhancer/promoter-driven expression of the uPA transgene resulting in a high, liver-specific concentration of this protease causing neonatal hemorrhage and hypofibrinogenemia, and overall inducing massive hepatocyte damage and collapse with chronic hepatic insufficiency (Heckel et al. 1990). This functional liver damage

causes an advantage for transplanted hepatocytes leading to improved engraftment and proliferation rates, and allows for regeneration of the damaged recipient liver (Rhim et al. 1995). For transplantation of xenogenous donor cells, Alb-uPA^{+/+}/Rag2^{-/-}-transgenic mice were cross-bred with SCID/Beige mice lacking an adaptive immunity (Lütgehetmann et al. 2012; Lütgehetmann et al. 2010; Meuleman and Leroux-Roels 2008). In the following, the mouse line uPA/SCID/Beige is termed USB. Although both, hemizygous and homozygous, genotypes can be used for transplantation experiments, nearly complete liver repopulation (~90%) was achieved in homozygous mice, only (Dandri et al. 2001; Meuleman and Leroux-Roels 2008). Hemizygous USB mice showed liver reconstitution levels of only about 15% (Dandri et al. 2001) referable to an override of uPA expression due to somatic recombination in a small proportion of host hepatocytes (Sandgren et al. 1991). These cells preserve a growth advantage versus transgene-expressing cells and are able to regenerate liver tissue. In transplantation experiments, recombined hemizygous cells directly compete with transplanted cells and thereby impede repopulation by those. Already 2-4 weeks after birth, USB mice have to get hepatocyte transplantation, otherwise the mice would die from liver failure (Dandri et al. 2001; Meuleman et al. 2005). Not all transplanted cells survive, but already a small proportion of engrafted cells are sufficient to repopulate the host liver forming so-called regeneration nodes after proliferation (Meuleman et al. 2005). Active proliferation is observed only at the distal edge of the nodes adjacent to endogenous damaged liver tissue of the host. Transplanted primary hepatocytes form biliary tracts, which connect to the host's bile system. A functional liver metabolism resulting from integrated-donor hepatocytes is indicated by albumin increase to normal levels (Meuleman et al. 2005). Mainly, humanized USB mice were used for study of hepatitis infection (Dandri et al. 2001; Dandri and Petersen 2012). However, in the past seven years this model was also used for analysis of clonal evolution of transplanted murine liver cells (Cornils et al. 2014; Thomaschewski et al. 2017; Weber et al. 2012; Weber et al. 2011). Intrasplenically-transplanted primary murine hepatocytes repopulated successfully the recipient mouse liver after *ex-vivo* transduction with LeGO vectors (Weber et al. 2011). To avoid an immune response against the fluorescent proteins, the immune deficiency of the recipients was essential. Single cell analysis proved stably integrated proviral DNA to induce solid tumors in mice (Cornils et al. 2014). Moreover, isolated cells from lentivirally-transduced FH-hTERT (human telomerase reverse transcriptase-immortalized human fetal hepatocytes)-derived primary tumors were able to form tumors after re-transplantation into hemizygous mice retaining their color code (Thomaschewski et al. 2017). These studies have demonstrated the potential of modified primary murine hepatocytes to analyze liver repopulation after gene delivery via lentiviral vectors. RGB-transduced hepatocytes were able to form multicolor cell clones showing engraftment without tumor development implying that lentiviral insertion and

fluorescent protein overexpression together are not sufficient to induce tumorigenesis in the recipient liver.

To summarize, the combination of the USB mouse model with LeGO vectors is a well-working *in-vivo* approach to investigate physiological liver repopulation.

1.4 Aim of this thesis

Research on liver carcinogenesis and the molecular mechanisms behind HCC formation requires a suitable *in vivo* model. Although valuable insights have been gained into the mechanisms of liver repopulation and cancer development by using the mouse models described in the past (Gilgenkrantz 2010; Ranzani et al. 2013), those models still do not resemble disease progression in human patients closely. HCC occurs in patients at a median age of 65 years, but only neonatal mice were used in the existing studies with observable tumor formation (Ranzani et al. 2013). Hence, the need for a less artificial model is crucial for the study of hepatocellular carcinomas. In the long run, determining the underlying molecular mechanisms of HCC formation and developing putative therapies can only be accomplished with the most suitable model.

This thesis aims at the generation of a reliable liver tumor animal model. Fluorescence protein-expressing LeGO vectors that have already been used successfully in transplantation models to study the benign regeneration of the liver (Cornils et al. 2014) are to be deployed under the two-hit hypothesis using confirmed or potential oncogenes to induce HCC in USB mice. These transgenes – *CyclinA2*, *HRas-V12*, and *SV40 LargeT-antigen* – encode proteins that are found deregulated in signaling pathways of liver tumors, as previously mentioned (Sia et al. 2017), some of which have already been used in mouse studies (Dubois et al. 1991; Sandgren et al. 1989). Combining the already existing RGB marking system with oncogene-expressing lentiviral vectors, an *in-vivo* model can be developed for the investigation of hepatocarcinogenesis. From this approach two hypotheses emerge:

1. Introduction of lentivirally expressed oncogenes into adult primary hepatocytes induces hepatocarcinogenesis within the hepatotoxic environment of USB mice *in vivo*.
2. Tools for gene marking allow quantitative analyses of the physiological and malignant clonal regeneration of the liver.

These hypotheses will be tested in the above mouse model.

2 Materials and Methods

2.1 Materials

All materials used in this thesis are listed below. Waste was properly disposed according to safety guidelines.

2.1.1 Disposables

Disposable	Reference	Manufacturer
Bottle top filter, 0.22 µm	SCGPT05RE	Merck KGaA (Darmstadt, Germany)
Cell culture dish, 10 cm (adherent)	83.3902	Sarstedt AG & Co. (Nümbrecht, Germany)
Cell culture plate, 6-well (adherent)	83.3920.005	Greiner Bio-One GmbH (Frickenhäusen, Germany)
Cell culture plate, 24-well (adherent)	83.3922.005	Greiner Bio-One GmbH
Cell culture flask, T75 (adherent)	83.3911	Greiner Bio-One GmbH
Cell culture plate, 6-well (suspension)	83.3920.500	Greiner Bio-One GmbH
Cell culture flask, T75 (suspension)	83.3911.500	Greiner Bio-One GmbH
Cell scraper	83.1832	Sarstedt AG & Co.
Cell strainer, 40 µm	431750	Corning Inc. B.V. Life Sciences (Amsterdam, Netherlands)
Cell strainer, 100 µm	431752	Corning Inc. B.V. Life Sciences
Corex tube, 38 mL	3117-0380	Thermo Fisher Scientific (Karlsruhe, Germany)
Cover slip, 24x60 mm	LH26.1	Carl Roth GmbH & Co. KG (Karlsruhe, Germany)
Cryo tube vial, 1.0 mL	37722	Thermo Fisher Scientific
DG8 gaskets	1863009	Bio-Rad Laboratories Inc. (Hercules, USA)
DG8 cartridges	1864008	Bio-Rad Laboratories Inc.
Glass capillary	9201010	Hirschmann Laborgeräte (Eberstadt, Germany)
Lint-free tissue	07730012	GVS Großverbraucherspezialisten eG (Heidenheim, Germany)
Pasteur pipette, 5 mL	86.1253.001	Sarstedt AG & Co.
Pasteur pipette, 10 mL	86.1254.001	Sarstedt AG & Co.
Pasteur pipette, 25 mL	760180	Greiner Bio-One GmbH
PCR plate, 96-well	951020389	Eppendorf AG (Hamburg, Germany)
PCR Plate Heat Seal foil	1814040	Bio-Rad Laboratories Inc.
Reaction tube, 0.2 mL	0030124332	Eppendorf AG
Reaction tube, 1.5 mL	0030125150	Eppendorf AG
Reaction tube, 2.0 mL	0030120094	Eppendorf AG
Serum sample tubes	41.1501.005	Sarstedt AG & Co.
Superfrost/plus slide, 75x25 mm	42409110	Glaswarenfabrik Karl Hecht GmbH & Co. KG (Sondheim vor der Rhön, Germany)
Syringe filter, PVDF membrane, 0.22 µm	SLGV033RK	Merck KGaA
Syringe filter, PVDF membrane, 0.45 µm	SLHV033RB	Merck KGaA
Tube, 30 mL	201150	Greiner Bio-One GmbH
Tube, 50 mL	227289	Greiner Bio-One GmbH

2.1.2 Instruments and accessories

Device	Name	Manufacturer
7T MRI system	ClinScan	Bruker BioSpin GmbH (Ettlingen, Germany)
CCD Camera	Color View III	Soft Imaging Systems GmbH (Münster, Germany)
Centrifuge	Megaforce 1.0R Sorvall	Heraeus Instruments GmbH (Hanau, Germany)
Centrifuge	Rotanta 460R	Andreas Hettich GmbH & Co.KG (Tuttlingen, Germany)
Centrifuge	Sorvall RC-5C Plus (HB-6 swinging bucket rotor)	Heraeus Instruments GmbH
Clinical chemistry analyzer	COBAS Integra 400 plus	Roche Diagnostics Deutschland GmbH (Mannheim, Germany)
Cryo-microtome	HM550	Thermo Fisher Scientific Microm International GmbH (Walldorf, Germany)
Digital camera	FOcubusIEEE1394	PHASE GmbH (Lübeck, Germany)
Droplet generator	QX100 system	Bio-Rad Laboratories Inc.
Droplet reader	QX100 system	Bio-Rad Laboratories Inc.
DNA migration chamber	EASY Cast Electrophoresis system #B2	Owl Scientific Inc. (San Francisco, USA)
DNA migration chamber	PerfectBlue Gelsystem Mini S	Peqlab Biotechnologie GmbH (Erlangen, Germany)
FACS analyzer	BD FACS LSR Fortessa 405, 488, and 561 nm Laser	BD Biosciences (Heidelberg, Germany)
FACS sorter	BD FACS AriaIII 405, 488, and 561 nm Laser	BD Biosciences
Filter set DAPI	U-MNU2	Olympus Europa SE & Co. KG (Hamburg, Germany)
Filter set Cerulean	438/24 BrightLine HC 483/32 BrightLine HC Beamsplitter HC BS 458	AHF analysentechnik AG (Tübingen, Germany)
Filter set GFP/Venus	U-MNIBA2 Exciter: BP470-490 Dichroic: DM505 Emitter: BA510-550	Olympus Europa SE & Co. KG
Filter set mCherry	U-MWIG2 Exciter: BP520-550 Dichroic: DM565 Emitter: BA580IF	Olympus Europa SE & Co. KG
Filter set Tripleband mCherry/Venus/Cerulean	Exciter: CFP/YFP/mCherry ET Tripleband Emitter: ET Tripleband bs CFP/YFP/mCherry	AHF analysentechnik AG (Tübingen, Germany)

Fluorescence microscope	IX81	Olympus Europa SE & Co. KG
Fluorometer	Qubit 2.0	Thermo Fisher Scientific
Heating block	Thermomixer 5436	Eppendorf AG
Heating plate	PZ72	Harry Gestigkeit (Düsseldorf/Deutschland)
Incubator	CB150	BINDER GmbH (Tuttlingen, Germany)
Laminar flow hood	Herasafe	Thermo Fisher Scientific
Light Microscope	CKX41	Olympus Europa SE & Co. KG
Magnetic stirrer	IKAMAG REO	IKA-Werke GmbH & Co. KG (Staufen, Germany)
Spectrophotometer	Nanodrop ND-1000	Thermo Fisher Scientific
Spectrophotometer	Qubit 2.0	Thermo Fisher Scientific
Peristaltic pump	XX8200230	Merck Millipore (Billerica, USA)
pH meter	inoLab pH 720	WTW, Xylem Analytics Germany Sales GmbH & Co. KG (Weilheim, Germany)
Plate sealer	QX100 system	Bio-Rad Laboratories Inc.
Power supply	Biometra Standard Power Pack P25	Analytik Jena (Jena, Germany)
Silicone hose, platin-hardened, inner diameter 0.8 mm	OU-06371-00	Masterflex (Gelsenkirchen, Germany)
Table top centrifuge	1. 5418 2. 5417R	Eppendorf AG
Thermocycler	1. Biometra Professional 2. Biometra Professional gradient	Analytik Jena
UV transilluminator	312 nm UV lamp	PHASE GmbH
Vortex mixer	REAX control	Heidolph Instruments GmbH & Co. KG (Schwabach, Germany)
Water bath	Typ 1008	GFL Gesellschaft für Labordiagnostik GmbH (Burgwedel, Germany)
Water purifier	membraPure	membraPure GmbH (Henningsdorf, Germany)

2.1.3 Software

Software	Version	Manufacturer
BioEdit Sequence Alignment Editor	v7.2.5	Tom Hall Ibis Therapeutics (Carlsbad, USA)
FACSDiva	v4.2.2	BD Biosciences
FireCam Control software	v2.03	PHASE GmbH
Fire Wire		Soft Imaging Systems GmbH
GraphPad Prism	v6.0	STATCON (Witzenhausen, Germany)
ImageJ	v1.51j8	Wayne Rasband, National Institutes of Mental Health (Bethesda, USA)
K-PACS	v1.6.0	Image Information Systems Ltd. (London, United Kingdom)
Microsoft Excel	2007	Microsoft Corporation (Redmond, USA)
Microsoft Word	2007	Microsoft Corporation
Microsoft Powerpoint	2007	Microsoft Corporation
Microsoft Image Composite Editor	v2.0.3	Microsoft Corporation

Nanodrop ND-1000	v3.7.1	Thermo Fisher Scientific
pDRAW32 1.0 Revision	v1.1.133	AcaClone Software (Copenhagen, Denmark)
Primer Express	v3.0.1	Applied Biosystems by Thermo Fisher Scientific
QuantaSoft, Inc.	v1.7.4.0917 2014	Bio-Rad Laboratories Inc.
Qubit	v3.10	Thermo Fisher Scientific
SnapGene Viewer	v3.3.3	GSL Biotech LLC (Chicago, USA)

2.1.4 Chemicals and reagents

Chemical	Reference	Manufacturer
Acetone	9002-02	J.T. Baker Avantor Performance Materials B.V. (Deventer, Netherlands)
alpha-D-glucose	158968	Sigma-Aldrich (München, Germany)
Agarose Ultra Pure Agarose	16500500	Thermo Fisher Scientific
Ampicillin	HP62.1	Carl Roth GmbH & Co. KG
β -mercaptoethanol	M3148	Sigma-Aldrich
CaCl_2	1.02820.1000	Merck KGaA
ddPCR Droplet Reader Oil	1863004	Bio-Rad Laboratories Inc.
ddPCR Supermix for Probes	1863010	Bio-Rad Laboratories Inc.
DNA loading dye (6x)	R0611	Thermo Fisher Scientific
DNA/RNA-dye, peqGREEN	732-3196	Peqlab/VWR International GmbH (Hannover, Germany)
dNTPs 10 μM each: dATP; dCTP; dGTP; dTTP	R0142; R0151 R0161; R0171	Thermo Fisher Scientific
Droplet generation oil for probes	1863005	Bio-Rad Laboratories Inc.
EGTA	E4378	Sigma-Aldrich
GeneRuler 1 kb plus DNA Ladder (0.1 $\mu\text{g}/\mu\text{L}$)	SM1343	Thermo Fisher Scientific
KCl	6781	Carl Roth GmbH & Co. KG
LB_{Amp} agar (Fast Media Amp agar)	fas-am-s	InvivoGen (San Diego, USA)
LB medium	6669	Carl Roth GmbH & Co. KG
MgCl_2	AM9530G	Thermo Fisher Scientific
Methylbutane	721123	Sigma-Aldrich
NaCl (solid)	3624-01	J.T. Baker
Na_2HPO_4	T876	Carl Roth GmbH & Co. KG
$\text{Na}_2\text{HPO}_4 \cdot \text{H}_2\text{O}$	T877	Carl Roth GmbH & Co. KG
NaOH pellets	106482	Merck KGaA
Paraformaldehyde*	158127	Sigma-Aldrich
Sucrose**	US1SX1075/1	Merck KGaA
UltraPure DNA Typing Grade 50X TAE (Tris/Acetate/EDTA) Buffer, diluted in deionized H_2O to 1x TAE	24710030	Thermo Fisher Scientific
<i>Tissue-Tek Cryomold Intermediate embedding forms</i>	4566	Sakura (Torrance/USA)
<i>Tissue-Tek O.C.T. Compound</i>	4583	Sakura (Alphen aan den Rijn, Netherlands)

*For complete dissolution of the paraformaldehyde powder, the flask was incubated for 1.5 h at 60°C and shaken occasionally. As soon as the solution was cooled down to RT it was ready to use. For storage, the solution was aliquoted in 50-mL tubes and kept at -20°C.

**The 20% sucrose solution was prepared with double distilled water and sterile filtered through a 0.22- μm filter unit (Merck Millipore) under the hood. For storage, the solution was kept at 4°C.

2.1.5 Kits

All listed kits were used according to manufacturers' instructions.

Kit name	Reference	Manufacturer
DAKO EnVision ⁺ Dual Link System-HRP (DAB ⁺) Kit	K406511-2	Agilent Technologies Deutschland GmbH (Waldbronn, Germany)
PureLink HiPure Plasmid Maxiprep Kit	K210007	Thermo Fisher Scientific
QIAamp DNA Blood Mini Kit	51106	QIAGEN GmbH (Hilden, Germany)
QIAGEN AllPrep DNA/RNA Mini Kit	80204	QIAGEN GmbH
QIAprep Spin Miniprep Kit	27106	QIAGEN GmbH
QIAquick Gel Extraction Kit	28706	QIAGEN GmbH
QIAquick PCR Purification Kit	28106	QIAGEN GmbH
Qubit dsDNA BR Assay Kit	Q32850	Thermo Fisher Scientific
Qubit dsRNA HS Assay Kit	Q32852	Thermo Fisher Scientific

2.1.6 Enzymes and oligos

Enzyme	Reference	Manufacturer
<i>Eco</i> RI Fast Digest	FD0274	Thermo Fisher Scientific
FastAP Thermosensitive Alkaline Phosphatase (1 U/μL)	EF0651	Thermo Fisher Scientific
Liberase TM Research Grade	05401119001	Roche Diagnostics Deutschland GmbH
<i>Not</i> I Fast Digest	FD0593	Thermo Fisher Scientific
<i>Phusion</i> High-Fidelity DNA Polymerase (2,000 U/mL)	M0530	New England Biolabs (Frankfurt am Main, Germany)
T4 DNA Ligase (5 U/μL)	EL0011	Thermo Fisher Scientific

All oligonucleotides such as primers and probes were purchased from Eurofins mwg. Primers were ordered from Eurofins Genomics (Ebersberg, Germany) choosing "Salt Free" or HPLC purification method. Used oligonucleotides are listed in Supplementary Tables S1 and S2.

2.1.7 Media and additives for cell culture

Product	Reference	Manufacturer
Chloroquine	PHR1258	Sigma-Aldrich
DMEM, high glucose, GlutaMAX Supplement, pyruvate	31966021	Thermo Fisher Scientific
DPBS (1x) no calcium, no magnesium	14190094	Thermo Fisher Scientific
Fetal Bovine Serum	F7524	Sigma-Aldrich
L-Glutamine	25030149	Thermo Fisher Scientific
Hepatocyte Wash Medium	17704024	Thermo Fisher Scientific
HepatoZYME-SFM	17705021	Thermo Fisher Scientific
HEPES (1M)	15630056	Thermo Fisher Scientific
Penicillin-Streptomycin (100 IU/mL/100 μg/mL)	15140122	Thermo Fisher Scientific
Polybrene	TR-1003	Sigma-Aldrich
Trypan Blue Solution (0.4%)	15250061	Thermo Fisher Scientific
Trypsin-EDTA (0.05%), phenol red	25300054	Thermo Fisher Scientific

2.1.7 Buffer for viral vector production2x HBS (HEPES buffered saline)275.8 mM CaCl_2

10.2 mM KCl

1.41 μM NaH_2PO_4

42 mM HEPES

1.1 mM α -D-glucose

ad 1L with ddH₂O; pH 7.05 adjusted by 1 N NaOH and sterilized via 0.22 μm filtration, storage at -20°C

2.1.8 Buffer and stock solutions for liver perfusion and hepatocyte isolation

All buffers were adjusted to pH 7.4 with 1 N NaOH, sterile filtered with the *BottleTop*-filter system (0.22 μm), and stored at 4°C. Prior to use, the buffers were prewarmed to 39°C.

10x Leffert's stock solution

100 mM HEPES

30 mM KCl

1.3 M NaCl

10 mM $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ 100 mM α -D-Glucosead 1000 mL ddH₂OEGTA stock solution

100 mL 10x Leffert's stock solution

5 mM EGTA

ad 1000 mL ddH₂OCalcium chloride stock solution63 mM CaCl_2 ad 500 mL ddH₂OWashing buffer

45 mL 10x Leffert's stock solution

50 mL EGTA stock solution (final 0.5 mM)

ad 500 mL ddH₂O

Collagenase buffer

50 mL 10x Leffert's stock solution

5 mL CaCl₂ stock solution (final 0.63 mM)

ad 500 mL ddH₂O

2.1.9 Surgical kit

All material used in hepatocyte transplantation surgery is listed below.

Product	Reference	Manufacturer
26 G cannula; BD Microlance 3 1/2; 0.45x13 mm	305111	Becton Dickinson (Heidelberg, Germany)
27 G cannula; BD Microlance 3 3/4; 0.4x13 mm	302200	Becton Dickinson
Carprofen 50 mg/mL Rimadyl	400684.00.00	Pfizer Deutschland GmbH (Berlin, Germany)
Enrofloxacin (Baytril 2.5% oral solution)	3572731	Bayer Vital GmbH (Leverkusen, Germany)
Forene inhalation solution (active agent: isofluran)	10182054	Abbvie Deutschland GmbH & Co. KG (Wiesbaden, Germany)
Glass syringe 100 µL Model 1710 RN SYR, NDL, barrel inner diameter 1.46 mm, barrel outer diameter 7.75 mm; Needle G 28 /15mm/ pointstyle	7803-02	Hamilton by Sigma-Aldrich
Infusion set Vasofix Safety PUR, 22G, 0.9x25 mm	4269098S-01	B. Braun Melsungen AG (Melsungen, Germany)
Insulin syringe, 0.33 mm, 29G, U100	305932	Becton Dickinson
Ketamin 100 mg/mL	9089.01.00	Wirtsgenossenschaft deutscher Tierärzte eG (Garbsen, Germany)
Novaminsulfon (500 mg/mL)	08713863	Ratiopharm GmbH (Ulm, Germany)
Octenisept, alcohol-free	173702	Schülke & Mayr (Norderstedt, Germany)
Ophthalmic ointment (Bepanthen, Augen- und Nasensalbe)	01578681	Bayer Vital GmbH
PGA Resorba needle DS 18, 3/0 USP, 0.45 m	PA 11416	RESORBA Medical GmbH (Nürnberg, Germany)
PGA Resorba suture 3/0, 0.45 m0	PS 465	RESORBA Medical GmbH
Povidone-iodine unguent	04923204	Mundipharma GmbH (Limburg, Germany)
Sedaxylan 20 mg/mL	400650.00.00	Wirtsgenossenschaft deutscher Tierärzte eG (Garbsen, Germany)
Wound Clip (MikRon Autoclip, 9 mm)	427631	Becton Dickinson
Wound Clip Applier	427630	Becton Dickinson
Wound Clip Remover	427637	Becton Dickinson

2.2 Methods

If not mentioned otherwise all work was carried out in a biosafety level 2 laboratory following the applicable biosafety guidelines.

2.2.1 Cloning of LeGO-LargeT-iV2

For the mouse studies, the vector LeGO-LargeT-iV2 had to be generated. The 2154-bp insert with the complete coding sequence of LargeT-antigen from simian virus SV40 was obtained by amplification with specific primers (p248/p249) in a polymerase chain reaction (PCR) on the plasmid pZL52 LargeT-2A-GFP-i-CD34 kindly provided by Elke Grassmann (Assurex Health, Cincinnati, USA). The forward primer (p248 – LargeT-Fw) sequence included an overhang (TATA) followed by the *EcoRI* site (GAATTC) and added a Kozak sequence (GCCACC) as well as sequences for the two amino acids methionine (ATG) and aspartic acid (GAC) to generate a full wildtype sequence of LargeT-antigen. At the 3' end, a *NotI* site (GCGGCCGC) was inserted following an overhang (TATA) using the reverse primer (p249 – LargeT-Rv). The four nucleotide overhangs were added to facilitate the restriction by the enzymes and were cleaved during the cloning process. The cloning strategy is shown in Supplementary Figure S1, primers are listed in Supplementary Table S1.

2.2.1.1 PCR amplification

The PCR was carried out in a total volume of 20 µL with Phusion reaction buffer, 10 nM dNTPs (stock 10 mM each), 10 pM LargeT-Fw (stock 10 µM), 10 pM LargeT-Rv (stock 10 µM), 1 U *Phusion* High-Fidelity DNA polymerase (2,000 U/mL), and 20 ng plasmid template DNA. The PCR was performed under the following conditions: Initial denaturation for 5 min at 98°C, 25 cycles of denaturation for 30 sec at 95°C, following annealing for 30 sec at 55°C and elongation for 1.5 min at 72°C repeating, followed by a final elongation of 5 min at 72°C and hold at 22°C. The PCR product was analyzed on a 1% agarose gel. 1x Tris/Acetate/EDTA (TAE) buffer was used to prepare and run the gel (5-8.5 V/cm) in a PerfectBlue Gelsystem Mini S or EASY Cast electrophoresis system. Samples were mixed with 1/6 volume of 6x Loading Dye and applied onto the gel next to 1 µg of 1 kb DNA Ladder (0.1 µg/µL). For visualization of DNA fragments and molecular weight marker under UV light, the agarose gel was supplemented with peqGREEN in a dilution of 1:25,000 prior to gel polymerization. Resulting images were recorded using the FireCam Control software on a desktop computer.

2.2.1.2 DNA fragment purification and digestion

Desired DNA fragment bands were cut from the gel for purification limiting UV light exposure of the DNA to a minimum to avoid unwanted mutations. Using the DNA Gel extraction Kit

according to the manufacturer's protocol, the desired DNA fragments were purified and eluted in 30 μ L elution buffer (EB). DNA concentrations were measured spectrophotometrically (Nanodrop ND-1000). Restriction enzyme cleavage with *EcoRI* and *NotI* on 200 ng PCR product and on 3 μ g backbone plasmid DNA (LeGO-iV2, kindly provided by Kristoffer Riecken, Dept. of Cell and Gene Therapy, University Medical Center Hamburg-Eppendorf, Germany) was performed in 20-30 μ L of Fast Digest Green reaction buffer, already including loading dye, for 30 to 60 minutes at 37°C. The LeGO-iV2 contains an SFFV (spleen focus forming virus) promoter, IRES (i, internal ribosomal entry site) and a coding sequence for the fluorescence protein Venus (V2). Its restriction sites for *EcoRI* and *NotI* are located between the SFFV promoter and IRES. Digested DNA was analyzed by agarose gel electrophoresis and purified by gel extraction as described previously.

2.2.1.3 Ligation and transformation

Following purification, the plasmid backbone (200 ng) and insert were ligated in a molar ratio of 1:3 within 20 μ L ligase buffer using 5 U T4 DNA ligase (5 U/ μ L). The reaction was carried out at 22°C for at least 1 h. Afterwards, 10 μ L of the ligation mix were used for transformation of 100 μ L competent *Escherichia coli* (*E. coli*) bacteria Top10 F' (Thermo Fisher Scientific, Karlsruhe, Germany; 5.1×10^6 colony forming units/ μ g) thawed on ice. The transformation mix was incubated for 30 min on ice, carefully agitating the tube every few minutes. A heat shock for 30 sec at 42°C and an immediate cold shock on ice for 2 min followed. To enhance initial replication of the bacteria, 300 μ L unsupplemented LB medium (lysogeny broth) was added and the mix was incubated for 2 h at 37°C on a shaker at 300 rpm. The bacteria suspension was plated on an LB_{Amp} agar dish (Amp: ampicillin) and incubated over night at 37°C. The next day, 8 clones were picked and cultivated in 5 mL LB_{Amp} medium (100 μ g/mL Amp) over night at 37°C. To isolate plasmid DNA from bacteria, the QIAGEN Plasmid Mini Kit was used according to the manual. DNA was eluted in 50 μ L EB. Restriction enzyme digest with *EcoRI/NotI* and agarose gel electrophoresis were performed as described earlier, using 1 μ g plasmid DNA of each clone. The DNA fragment pattern exhibited the correct size for all clones.

2.2.1.4 Sequencing and maxi preparation

Two clones were randomly chosen for DNA sequencing. Both clones were sequenced with three different primers: p007, p250, and p251 (see Supplementary Table S1). For each reaction, 1 μ g plasmid DNA was mixed with 30 pM primer (10 μ M stock) and filled with ddH₂O to a total volume of 15 μ L. Sequencing was performed by SeqLab (Sequence Laboratories Göttingen GmbH, Göttingen, Germany). Results were analyzed and aligned to the reference sequence of the desired LeGO-LargeT-iV2 using the software BioEdit

Sequence Alignment Editor. Both clones did not contain unwanted mutations. A plasmid maxi culture of clone 2 was prepared, transforming 100 μ L of competent *E. coli*, thawed on ice, with 200 ng plasmid DNA. The suspension was incubated on ice for 10 min and transferred to 100 mL LB_{Amp} liquid culture, which incubated over night at 37°C shaking at 180 rpm. The next morning, plasmid DNA was purified according to the manufacturer's protocol for Plasmid Maxiprep Kit (Invitrogen), and in a final step, eluted in a volume of 500 μ L Tris/EDTA (TE) buffer. The final concentration of plasmid LeGO-LargeT-iV2 was 1.4 μ g/ μ L determined spectrophotometrically. A final control was performed, conducting restriction with *EcoRI*/*NotI* as described before. The DNA band pattern on the agarose gel showed the predicted fragment size. This plasmid was used for generation of infectious vector particles.

2.2.2 Cell culture

All handling steps of cell culturing took place in a HEPA filtered laminar air flow cabinet and used material and media were sterile.

"PBS" connotes 1x DPBS (Dulbecco's Phosphate-Buffered Saline). Cell culture medium of adherent cells consisted of 1x DMEM (Dulbecco's Modified Eagle Medium) with additional 10% (v/v) FBS (fetal bovine serum), 1% (v/v) penicillin/streptomycin and 25 μ M HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) and is referred to as "DMEM".

2.2.2.1 Cultivation of adherent cells

The used adherent cell lines, HEK293T (ACC 635, DSMZ GmbH, Braunschweig, Germany) and NIH3T3 (ACC 59, DSMZ GmbH, Braunschweig, Germany), were cultured at 37°C with an atmosphere of 5% CO₂ and a relative humidity of app. 90%. Cells were kept in culture plates or culture flasks for adherent cells and passaged according to cell density and proliferation rate. Each time liquid (supernatant) was aspirated from the culture vessels; special care was taken to avoid drying of the cell layer. To passage cells, the medium was removed with a sterile pipette and cell layers were rinsed with PBS. The PBS was discarded and the cell dissociation reagent trypsin/EDTA (0.05%) was added to detach cells from the plate. Cells were incubated for 1 min at 37°C, and detachment was confirmed using a light microscope. To dilute and inactivate trypsin, prewarmed FBS-supplemented DMEM was added in excess to the culture plate. Cells were resuspended by repeatedly and vigorously pipetting media onto them to achieve a single-cell suspension. For counting, the cells were centrifuged and washed with PBS. The supernatant was carefully removed to avoid disturbing of the cell pellet. All washing and centrifugation steps were performed for 5 min at RT (room temperature) and 335 g. To disperse the cells, DMEM was added and pipetted gently up and down.

For dead-cell discrimination, cells were counted and, based on their optical density, diluted 1:5 or 1:10 in trypan blue. About 10 μL were pipetted onto a Neubauer improved counting chamber (area 0.025 mm^2 , depth 0.1 mm; #06.400.60, Paul Marienfeld GmbH & Co. KG (Lauda-Königshofen, Germany)). Cell count was calculated using the following formula:

$$\text{concentration of cells in original mixture} = \frac{\text{number of cells counted}}{\text{counted large squares}} \times \text{dilution factor} \times 10^4 \text{ cells/mL.}$$

To seed the cells, an aliquot of the cells was diluted in an appropriate volume of fresh medium and transferred into a culture plate. Culture conditions were the same as mentioned before.

2.2.2.2 Production of infectious vector particles

For transduction of primary murine hepatocytes and NIH3T3 control cells vector particles were produced. HEK293T cells were transfected according to the calcium-phosphate-method (Pear et al. 1993) using four plasmids. The envelope plasmid pHCMV-VSV-G (Beyer et al. 2002) encoding the VSV-G (glycoprotein G of the Vesicular stomatitis virus) was co-transfected with the packaging plasmid pMDLg/pRRE (Dull et al. 1998) encoding the viral enzymes and core proteins Gag and Pol together with the plasmid pRSV-Rev (Dull et al. 1998) encoding the accessory HIV1 gene *rev* and the transfer plasmid delivering the transgene sequence. The Rev protein is important for late phase gene expression and also regulates the transport of mRNA encoding viral structural proteins from the nucleus to the cytoplasm. By using the two independent plasmids *gag/pol* and *rev*, safety was increased and third-generation lentiviral vector particles of the expression plasmid were produced. These particles are called pseudotyped, since the envelope proteins used by them do not originate from the wildtype virus itself. The following protocol applies to transfection of one 10-cm dish with 4×10^6 seeded cells resulting in app. 10 mL final viral supernatant. To prepare larger amounts, multiple dishes were prepared.

Before transfection, HEK293T cells were seeded in 10 mL DMEM, resulting in a confluency of around 70%. Sixteen hours after seeding, each dish was transfected with 2 μg VSV-G envelope plasmid DNA, 10 μg *gag/pol* packaging plasmid DNA, 5 μg *rev* plasmid DNA, and 15 μg of the respective transfer plasmid (Fig. 2.1; Supplementary Fig. S2-S7) solved in a total volume of 450 μL dH_2O , adding 250 mM CaCl_2 in a final step (see <http://www.lentigo-vectors.de/protocols.htm>; 16.07.2018, 5:00 p.m.).

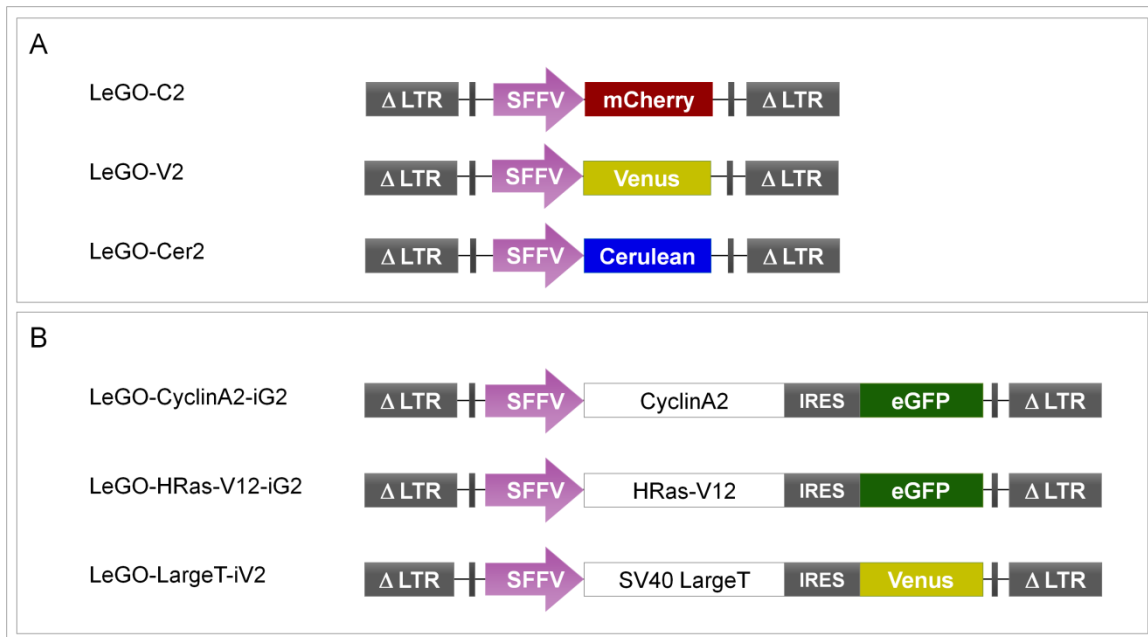


Figure 2.1 Expression plasmids used for viral vector production. RGB vectors (A) and oncogene vectors each coupled with a fluorescent protein (B) were used. LeGO, lentiviral gene ontology; IRES, internal ribosomal entry site; eGFP, enhanced green fluorescent protein. LeGO-LargeT-iV2 was cloned during this thesis all other vectors were kindly provided by Kristoffer Riecken (Dept. of Stem Cell Transplantation, University Medical Center Hamburg-Eppendorf, Germany); RGB vectors (A) have been published in Weber et al. 2008.

The DNA- CaCl_2 mix was given dropwise to 500 μL 2x HBS (HEPES buffered saline) buffer while simultaneously blowing air bubbles into the solution with a Pasteur pipette attached to a motorized pipette controller increasing the surface of the solution. The transfection mix was incubated at RT for 15 to 20 min after vigorously vortexing. Directly before transfection, medium was changed and 10 mL of new media were given to the cells with additional 25 μM chloroquine. The complete transfection solution (1 mL) was added dropwise to the HEK293T cells and mixed carefully. After 6 – 8 h, the medium was replaced by new DMEM. 24 hours later, the viral-vector containing supernatant was harvested and filtrated with a 0.45- μm sterile filter and frozen in aliquots at -80°C until use.

2.2.2.3 Concentration of infectious vector particles

To enhance viral-vector titers, the supernatant harvested 24 h after transfection was concentrated, if necessary. Supernatant was collected and applied to 30-mL corex glass tubes for centrifugation over night at 8,000 g and 4°C . 14 h later the supernatant was carefully removed, leaving 500 μL liquid behind. In this volume, the viral particle pellet was resuspended. All supernatant was pooled and frozen in 100- μL aliquots at -80°C .

2.2.2.4 Determination of infectious vector titers

The titer of vector supernatant was measured in HEK293T cells via flow-cytometric analysis. 5×10^4 cells per well were seeded in a 24-well cell culture dish suspended in 500 μL DMEM containing 8 μM polybrene. The cells were allowed to sediment and attach for 2 h before

applying increasing volumes of viral supernatant (0.1, 0.3, and 1.0 μL) in triplicates. Then, the plate was centrifuged for 1 h at 931 g at RT. Subsequently, the cells were incubated at 37°C over night. The next day (after 16 h), the medium was replaced by fresh DMEM without polybrene and cells were incubated for another 48 h. The supernatant was removed and cells were washed with PBS. To detach the cells from the culture dish surface, 200 μL 0.05% trypsin/EDTA were added, and the cells were incubated for 2 min at 37°C. The cells were resuspended, transferred to 1.5-mL tubes, and centrifuged at 335 g at RT for 5 min. The supernatant was discarded, the cell pellet resuspended in 500 μL PBS, and vortexed immediately. Next, transduction rates were measured via flow cytometry using the LSR Fortessa analyzer. The amount of IP/mL (infectious particles per mL) was calculated using the following formula:

$$\text{Titer [IP/mL]} = \frac{\text{seeded cells per well} \times \frac{\text{positive cells [\%]}}{100}}{\text{viral supernatant [mL]}}.$$

The viral supernatant was used for transduction of primary murine hepatocytes for transplantation or NIH3T3 cells to generate a genetic control for molecular analyses. Supplementary Figure S8 displays flow-cytometric data of titer determination for the vectors LeGO-C2, LeGO-V2, LeGO-Cer2, LeGO-HRas-V12 and LeGO-SV40-LargeT-iV2 3 d post transduction. Applying 0.1 μL of supernatant, 9.8% of the transduced cells were positive for mCherry, 9.1% were Venus-positive and 4.6% were positive for Cerulean (Supplementary Fig. S8, A-C; Supplementary Table S3). Similar numbers were achieved for GFP-expressing LeGO-HRas-V12-iG2 (8.3%) and LeGO-SV40-LargeT-iG2 (8.2%) transduced cells using 0.3 μL viral supernatant each (Supplementary Fig. S8, D, E; Supplementary Table S3).

2.2.2.5 Generation of control cells for molecular analysis

Per construct, 1×10^6 NIH3T3 cells were seeded in a 10-cm dish. MOIs 1 and 20 were tested. Transduction was performed according to the protocol described above. Viral supernatant was applied to the cells, followed by 1 h centrifugation at 335 g at RT. 24 h after transduction, medium was exchanged. On day 3, flow-cytometric analysis was conducted, and transduction efficiency was calculated. Genomic DNA (gDNA) was extracted using 5×10^6 cells per sample with the DNA Blood Mini Kit. All steps were performed according to the manufacturer's protocol. Elution was completed with 200 μL of AE buffer.

2.2.3 Mouse experiments

All mouse experiments were approved by the authorities of the Hamburg State (Behörde für Gesundheit und Verbraucherschutz) and permitted by the Tierversuchsgenehmigung G057/11 and G014/16. All animals were bred at the animal facility of the UKE and maintained in accordance with animal welfare guidelines. Transplantation of primary murine

hepatocytes was performed in a biosafety level 2 animal laboratory. 10 days after transplantation, recipients were transferred to a biosafety level 1 animal facility. General health conditions of animals were screened daily; weight was determined two to three times a week. For primary hepatocyte isolation C57Bl6/J_2014 mice were used; recipient animals were derived from the USB (uPA/SCID/Beige; urokinase-type plasminogen activator/severe combined immunodeficiency/Beige) mutant mouse strain. Overexpression of the murine uPA gene under control of the albumin enhancer/promoter results in chronic hepatic insufficiency (Sandgren et al., 1991). Transplantation has to be completed soon after birth; otherwise homozygous animals succumb due to their phenotype. In this study, the SCID/Beige mutation is necessary to prevent an immunological reaction against the transgenes, such as GFP, expressed by the lentiviral constructs. Mice harboring this mutation show a severe deficiency of natural killer (NK) cells (Roder, 1979), as well as defective cytotoxic T cell and antibody responses to allogeneic tumor cells (Carlson et al., 1984; Saxena et al., 1982), but also defective and reduced bactericidal activity of granulocytes (Gallin et al., 1974). For the first experiment, two further recipient mouse strains were used to evaluate transplantation efficiency. The USG (uPA/SCID/Beige/Il2rg^{-/-}; Il2rg, IL2 receptor common gamma chain) and USRapG (uPA/SCID/Beige/Pfp^{-/-}/Rag2^{-/-}/Il2rg^{-/-}; Pfp, Perforin gene; Rag2, recombination activating gene 2) immunodeficient mouse strains lack mature B, T and NK immunity achieved through several additional immunodeficiency-inducing knock-outs. However, no difference in outcomes was observed comparing the three mouse strains. In total, 6 mock mice, 11 RGB-, 11 CyclinA2-, 12 HRas-V12-, and 12 LargeT-transplanted mice were analyzed. An overview of recipient numbers per groups and mouse strains is given in Supplementary Table S4.

2.2.3.1 *In-situ* perfusion

Primary hepatocytes were isolated by *in-situ* perfusion according to a modified two-step collagenase-perfusion protocol (Seglen 1976; Aiken et al. 1990; Li et al. 2010). Eight-to-twelve weeks old C57Bl/6J_2014 female wildtype mice were used as donors. From one donor mouse enough cells were obtained to transplant up to 24 recipient animals. Prior to the cell isolation process, the hose system (0.8 mm inner diameter) was rinsed with 80% ethanol, followed by PBS and sterile ddH₂O. The flow rate of the peristaltic pump was adjusted to 5 to 7 mL per minute. Prior to cell isolation, all perfusion buffers were kept at 39°C in a water bath connected to the peristaltic pump and hose system. The intake hose was placed into the wash buffer flask channeled through the peristaltic pump and ended in an air trapping. The second hose was led through the peristaltic pump, starting at the bottom of the air trapping and ending in the wash buffer flask to generate a stable temperature within the circulating system. During the perfusion, the second hose was connected to the infusion

set (22G cannula). Collagenase I and II Liberase enzyme mix was resuspended in pre-warmed collagenase buffer directly before use. A lethal dose of anesthetics (10 mL/kg body weight; 120 mg Ketamin/16 mg Sedaxylan dissolved in 0.9% NaCl solution (B. Braun)) was applied i.p. (intraperitoneally). After reaching surgical anesthesia indicated by loss of the pedal withdrawal reflex of the hind limb, the procedure was undertaken. The mouse was set dorsal on a 39°C heating plate using adhesive tape ventral facing the experimenter. After disinfection with Octenisept, the abdominal cavity was opened with sterile scissors, and the intestinal loop was luxated with a blunt forceps to uncover the portal vein (*vena portae*) and inferior caval vein (*vena cava inferior*) without damaging them. The thorax was opened along the midaxillary line (*linea axillaris media*) with the same equipment to reach the still beating heart. A ligature was placed around the superior caval vein (*vena cava superior*) with hydrated 3-0 silk suture needed for latter fixation of the infusion set. The venous catheter was introduced into the right atrium. The cannula was removed and the infusion system was connected to the *vena cava superior* fastened with the 3-0 suture prepared earlier. The end of the second hose was fixed carefully to the infusion system, avoiding any trapping of air bubbles. The *vena cava inferior* was disconnected using a bulldog clamp (Fine Science Tools, Heidelberg, Germany) and the *vena portae* was cut open immediately after swelling of the liver to drain blood and buffer. Afterwards, the gall bladder was disrupted and the surface of the liver was rinsed with washing buffer. About 70 mL washing buffer were used to perfuse the liver for 10 min. Successful perfusion was marked by a color change of the liver from dark to light brown. During the last minute of perfusion with washing buffer, 24 mL prewarmed collagenase buffer were mixed with 1 mL of collagenase solution (1 mg/mL; final concentration 40 µg/mL) and the intake hose was changed quickly to the collagenase-buffer-mix. When the first mix was nearly empty, a second batch was prepared the same way. Perfusion with collagenase did not exceed 10 minutes. The perfusion was stopped, when the extracellular matrix was digested and hepatocytes began to detach from parenchyma.

2.2.3.2 Isolation of primary hepatocytes

After digestion with collagenase, the liver was removed from the body, carefully transferred to a Petri dish filled with Hepatocyte Wash medium, and placed on ice. All following steps were performed on ice under a laminar flow hood or at 4°C. Hepatocytes were handled with special care as excessive pipetting was avoided. The liver was retained with a forceps, and by cautiously shaking the organ the hepatocytes were released into the buffer. Additionally, a cell scraper was used to detach the hepatocytes. Hepatocytes were filtered through a 100-µm cell strainer first, followed by a 40-µm cell strainer and collected in a 50-mL tube. The tube was filled with 30 mL Hepatocyte wash medium centrifuged for 5 min at 50 g and 4°C. The supernatant was collected in a new tube. By tapping, the cell pellet was

resuspended. For a second time, the tube was filled with 30 mL Wash medium and both cell fraction and supernatant were centrifuged as described above. The supernatants were then discarded; both pellets were resuspended in the remaining buffer, pooled in one tube and mixed with 30 mL wash medium followed by another straining step (100 μ m). Cells were counted by trypan blue staining in a Neubauer improved counting chamber. Another centrifugation followed and supernatant was discarded with the hepatocytes now being ready for *ex-vivo* transduction.

2.2.3.3 Transduction of primary hepatocytes

After isolation, 4×10^6 hepatocytes were resuspended in supplemented HepatoZYME culture medium (hSFM, + 200 mM glutamine + 1% P/S + 4 μ g/mL polybrene) and seeded in a T75 flask for non-adherent cells. The volume of viral supernatant was considered for the final volume of 25 mL. The respective volumes of viral supernatant were applied to ensure MOIs of 40 or 60, respectively (see Supplementary Table S3). In every experiment untransduced hepatocytes were seeded additionally as control for microscopy and flow cytometry analyses. The untransduced control hepatocytes were resuspended in supplemented hSFM without application of viral supernatant. Cells were cultured for 1 h at 37°C with 5% CO₂ atmosphere. During the incubation time of experiment III, the cells were transduced for 1.5 h instead of 1 h due to a fire alert. After incubation, three aliquots of 5×10^4 cells from all groups each were seeded as control in a 24-well adherent cell culture plate cells and resuspended in 500 μ L hSFM to check for the transduction efficiency 3d post transplantation (Tx) by flow cytometry and fluorescence microscopy. The medium was replaced after 6 h incubation with prewarmed supplemented William's E (+GlutaMAX, +1% P/S, +10% FBS) and 2d post Tx for a second time. The remaining transduced cells were resuspended and transferred to a new 50-mL tube filled with 30 mL hSFM. Cells were centrifuged as described before and supernatants were discarded. Cell pellets were resuspended in 1 mL hSFM each and slowly pipetted into a 1.5-mL reaction tube. Subsequent to another centrifugation step in a table top centrifuge for 5 min at 50 g at RT, the cells were resuspended in a final volume of 20 μ L per 5×10^5 hepatocytes and pipetted into a conical cryo tube with internal thread. Final volume was dependent on group size. Cells were kept on ice until transplantation.

2.2.3.4 Assessment of transduction rates of primary hepatocytes

As mentioned before, of each group aliquots of 5×10^4 hepatocytes were kept in triplicates to measure the transduction efficiency 3 days after transduction by flow cytometry. First, the medium was removed and hepatocytes were washed with PBS. A thin liquid layer was kept on top of the cell surface for microscopy of the cells. To detach the hepatocytes from the culture plates afterwards, PBS was removed and 200 μ L 0.05% Trypsin-EDTA were added.

After incubation for 5 min at 37°C, hepatocytes were resuspended and 500 µL supplied William's E medium was added. Hepatocytes were centrifuged as described before, and the pellet was resuspended in 300 µL PBS. Before analyzing the hepatocytes on the FACS Aria III, the tubes were vortexed.

2.2.3.5 Intrasplenic transplantation of primary hepatocytes

Four-week old homozygous USB mice were used as recipient animals for transplantation experiments. In general, hepatocyte transplantation into the spleen leads to cell migration via the splenic and portal veins into the liver with hepatocytes disseminating dispersively into the hepatic sinusoids. As only a small number of transplanted cells successfully translocates to the liver plate, no more than few hepatocytes can eventually integrate in the liver parenchyma and repopulate the damaged organ (Gupta et al., 1999). Nevertheless, the proportion of engrafted cells has been shown to be sufficient for liver regeneration (Meuleman and Leroux-Roels 2008). Per mouse, 20 µL with 5×10^5 transduced cells were transplanted carefully using a glass syringe with a 28G/15mm/point style needle. As antiseptic prophylaxis, 5 mg/kg Carprofen (Pfizer) were applied s.c. (subcutaneously) with a 27G cannula prior to surgery. 24 h post transplantation, 200 mg/kg Metamizol could be applied s.c. to aid in pain reduction. To prevent bacterial infections, additionally Enrofloxacin was dissolved in the drinking water (final concentration 0.2 mg/mL).

During surgery, the animals were anesthetized by inhalation anesthesia (2% isoflurane/98% oxygen at a flow rate of 500 mL/min). Protection against dehydration of the eyes was achieved by using ophthalmic ointment. Once the animals were anesthetized, a small area of 1.5 cm² of the left flank was shaved and disinfected with povidone-iodine unguent. Animals were kept on a heating plate at 37°C during transplantation until they awoke afterwards. Transduced cells were resuspended and 20 µL containing 5×10^5 hepatocytes were taken up with the syringe. A small 0.5-cm cut was introduced beneath the thorax on the upper abdomen. First, the epidermis was opened. Once the spleen was visible, a small cut was introduced directly above it to open the peritoneum. The spleen was taken out carefully with a blunt forceps. A ligature was formed around the distal pole of the spleen with absorbable 3/0 suture, and the cells were slowly injected. After removing the cannula, the ligature was closed with two surgical knots and the spleen was placed back into the abdominal cavity. The peritoneum was sewed with 4 stitches, and the epidermis was closed with a 9-mm wound closure clip. The wound-closure clip was removed latest 2 weeks after transplantation.

2.2.3.6 Magnetic resonance imaging

Tumor growth was monitored via MRI (magnetic resonance imaging) 3, 5, 9, and 15 weeks after transplantation. All MR images were acquired with a small-animal 7T MRI system using a mouse body coil. Analyses were conducted under the supervision of Dr. Jan Dieckhoff (Department of Radiology, University Medical Center Hamburg-Eppendorf, Germany). Anesthesia of mice was performed by inhalation of 2% isoflurane and 98% oxygen at a flow rate of 500 mL/min. Respiration was monitored using a pressure pad (SA Instrument Inc., Stony Brook, USA) and controlled to a rate of 30-40 breath cycles per minute by isoflurane concentration changes. Constant animal warming was ensured through a warmed waterbed. The MRI protocol consisted of a survey scan and a 2D T2-weighted scan in transverse orientation. The T2-weighted scan was based on a turbo spin echo sequence with following parameters: 30 mm FOV (field of view). 256 x 256 matrix, 28 slices, 0.8 mm slice thickness (no gap), 5000 ms TR (repetition time), 77 ms TE (echo time), turbo factor of 11 and fat saturation. Eyes of animals were protected against dehydration using ophthalmic ointment.

2.2.3.7 Blood/serum sampling

Retrobulbar punctation was performed on anesthetized animals to collect blood or serum samples. To reach the retrobulbar vein plexus, a short glass capillary inserted retrobulbary with gentle rotation. To collect final serum samples, 1.3-mL sample tubes were used containing a clotting activator. Tubes were inverted after blood collection and incubated for 10 min at RT. Afterwards, tubes were centrifuged for 5 min at 1,000 g. After centrifugation, clear serum was transferred to a second sample tube containing clotting activator and the procedure was repeated. For storage, serum samples were pipetted into a 2.0-mL reaction tube and kept at -80°C.

2.2.3.8 Euthanizing

Animals were anesthetized with a mix of 80% CO₂ and 20% O₂ and sacrificed by cervical dislocation 6 – 28 weeks after transplantation. Final serum samples were taken as described before. Liver, spleen and lung were removed, weighed and photographed from top and bottom view. The *lobus medialis* was separated, weighed, immediately frozen in liquid nitrogen and stored at -80°C for later RNA and DNA analyses. A small piece of the spleen and one lung were flash frozen as well. The remaining organs were fixated in 4% PFA (paraformaldehyde) for later histology and microscopy analyses.

2.2.3.9 Serum ALB and ALT measurements

Determination of liver ALB (albumin) and ALT (alanine transaminase) concentration in serum was performed by Laura Berkhout and Carsten Rothkegel (both: Institute for Experimental

Immunology and Hepatology, University Medical Center Hamburg-Eppendorf, Germany) on a COBAS Integra 400 plus. A total volume of 120 μ L was required per read. Serum samples were diluted 1:10 in ddH₂O and measured in duplicates. Statistical analyses were performed by one-way ANOVA, Bartlett's test for equal variances, and Dunnett's Multiple Comparison Test using GraphPad Prism v5.02.

2.2.3.10 RNA and DNA analyses of liver tissue

To extract RNA and DNA from flash-frozen liver tissue, 30 mg tissue was used. According to the manufacturer's instructions (AllPrep DNA/RNA Mini Kit) 600 μ L RLT buffer supplemented with β -mercaptoethanol were added and tissue was grinded on dry ice with a metal pestle (Th. Geyer GmbH & Co. KG) fitting to a 1.5-mL Eppendorf tube. Extraction was performed following the manufacturer's protocol. DNA samples were eluted in 50-80 μ L EB, whereas RNA was eluted in 50 μ L RNase-free water. The DNA was used for droplet-digital PCR (ddPCR). Concentrations of extracted DNA and RNA from liver tissue samples were measured using the Qubit 2.0 fluorospectrometer according to manufacturer's instructions.

2.2.4 Droplet-digital PCR

DNA of liver tissue was analyzed by ddPCR. This method allows for sensitive quantification of target gene concentration in comparison to a reference gene using duplex assays (Fig. 2.2). The QX100 system and corresponding equipment/consumables from the company Bio-Rad were used. To validate primers and probes, gDNA of control cells was used. NIH3T3 were transfected with the corresponding viral vectors used in the mouse experiments. Transduction was performed as described before. 100 ng of control DNA from NIH3T3 cells were used per reaction. For PCR, 50 ng DNA of murine liver samples were used respectively. PCR was performed within a total volume of 25 μ L consisting of 12.5 μ L ddPCR Supermix for Probes, 900 nM forward primer and 900 nM reverse primer, 0.5 μ L *Eco*RI fast digest enzyme, 250 nM of each probe and 3 mM MgCl₂. Sequences of primers and probes are listed in Supplementary Table S2. FAM- and HEX-labelled probes were used for detection of either the gene of interest or the reference (REF) gene. In all samples, duplex measurements were conducted to compare concentrations between the REF gene (murine *Epo-R* (*erythropoietin receptor*)), and the gene of interest. Prior to droplet generation, gDNA restriction was performed at 37°C for 15 min. Subsequently, droplets were generated according to the manufacturer's instructions. PCR was performed in a 96-well plate. Plates were sealed and the PCR started with initial denaturation at 95°C for 10 min, 40 cycles of 94°C for 30 s and 60°C for 120 s followed by a final step of 98°C for 10 min and hold on 4°C. Readout was carried out on a QX100 Reader. Relative VCN (vector copy number) was

calculated by the following formula:

$$VCN = \frac{\text{Copies}_{\text{gene of interest}}}{\text{Copies}_{\text{murine reference gene}} / \text{karyotype}}$$

Since NIH3T3 cells have a hypertriploid karyotype, VCN were computed based on 3N. Most hepatocytes have a polyploid karyotype as well. After weaning, about 85% of hepatocytes have a 4N karyotype and at least two nuclei per cell (Epstein 1967; Gentric and Desdouets 2014). To properly compute VCN for hepatocytes 6.8N was defined as karyotype in this study based on this calculation: 4N x 2 (nuclei) x 85%.

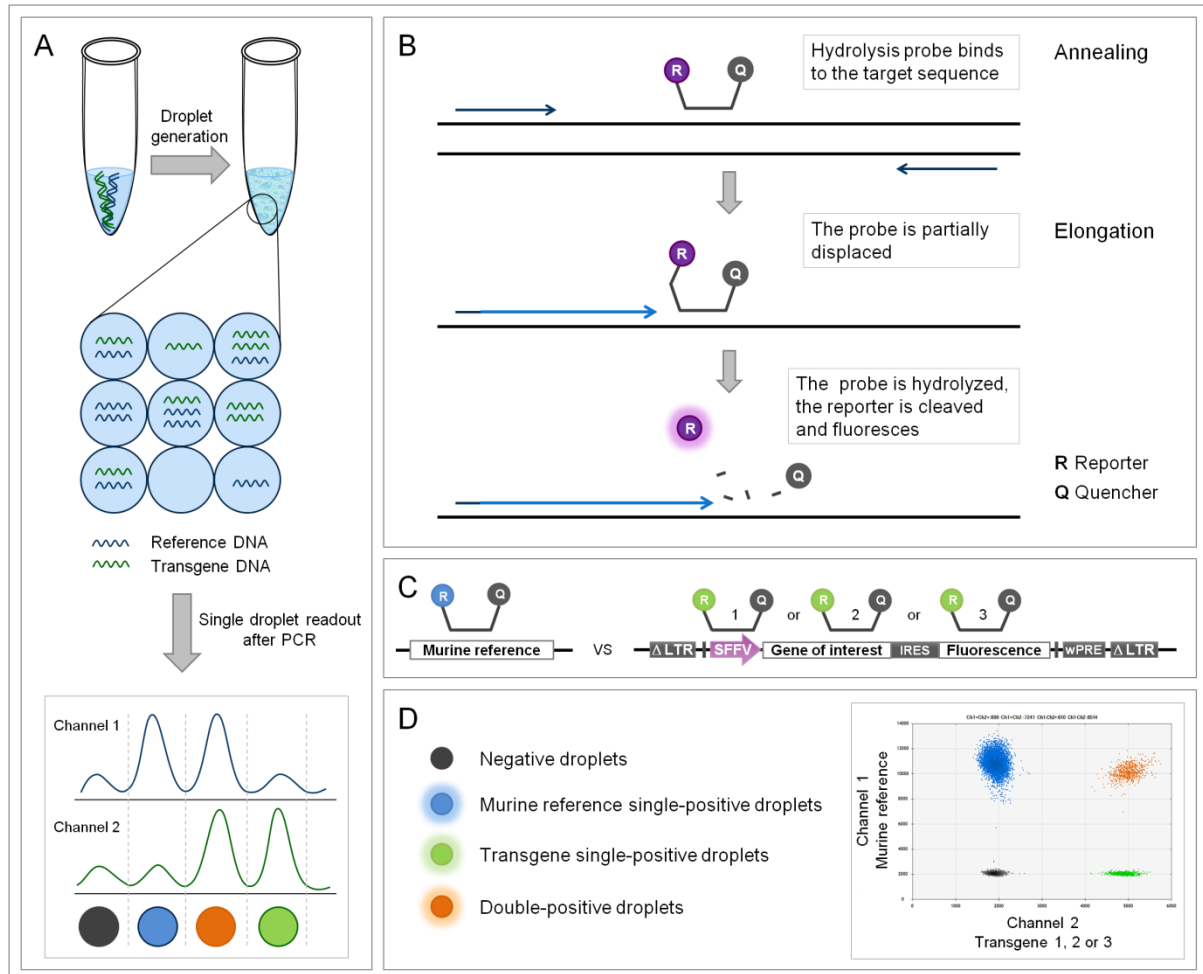


Figure 2.2 Scheme of droplet-digital PCR. Prior to PCR, sample is partitioned into app. 20,000 droplets with target and reference DNA randomly distributed among the droplets. Following amplification, each droplet presents a fluorescent positive or negative signal indicating the presence of reference or transgene target DNA. Each droplet provides an independent digital measurement (A). Sequence-specific probes are used containing a fluorescent dye (reporter) at the 5' end and a black hole quencher (quencher) at the 3' end. During annealing the probe binds to DNA, is displaced partially during elongation resulting finally in cleavage through the exonuclease activity of *Taq* polymerase and release of the reporter, which then gives a fluorescent signal (B). In this study, duplex analyses were performed using a specific probe targeting a murine reference together with one specific probe targeting either 1) the SFFV promoter, 2) CyclinA2, HRas-V12, or LargeT (gene of interest), or 3) Venus or GFP (fluorescence cassette) (C). Example for plate readout and detection of fluorescent signals after PCR shows four different populations either negative, single-positive or double-positive (D). *Taq*, polymerase isolated from *Thermus aquaticus*. Based on Droplet Digital PCR Applications Guide by Bio-Rad (Bulletin 6407 Ver B US/EG).

2.2.5 Histology and microscopy

All microscopic images were taken with a CCD Camera attached to an Olympus IX81 fluorescence microscope and processed with Fire Wire software or ImageJ 1.51j8.

2.2.5.1 Preparation of liver tissue for microscopy

1. Fixation

Remaining spleen, lung and other liver lobes, *lobus lateralis sinister*, *lobus lateralis dexter* and *lobus caudatae*, were separated and placed in one tube with 5 mL 4% PFA in PBS covering the whole tissue. Incubation with PFA lasted for 4 h at 4°C for fixation. The 4% (w/v) PFA solution was prepared by diluting PFA powder in PBS. Following the fixation, the PFA solution was discarded. To prevent freezing damage of the tissue, the liver lobes were dehydrated in 20% sucrose solution for 48 h at 4°C. The 20% sucrose solution was prepared with ddH₂O and sterile filtered through a 0.22-µm filter unit under the fume hood. For storage, the solution was kept at 4°C. Right before freezing, the lobes were rinsed once in PBS.

2. Freezing

To freeze the liver lobes, single lobes were transferred into *Tissue-TekCryomold Intermediate* embedding forms and covered with *Tissue-Tek O.C.T. Compound*. The embedding forms were frozen in methylbutane, pre-cooled with dry ice. The embedding forms were removed from the methylbutane and transferred into a pre-labeled plastic bag. Liver lobes were stored in zip lock plastic bags (Th. Geyer GmbH & Co. KG) at -80°C.

3. Tissue sectioning

Frozen liver sections were used for sectioning with a cryo-microtome. Before sectioning, the cryo-chamber, specimen clamping and other working accessories were preconditioned to a stable temperature around -20°C. The specimen acclimated to the temperature inside the chamber before any cutting processes; otherwise the specimen could break due to the sudden difference in temperature or stick to the metallic working accessories. For sectioning, the frozen tissue was prepared according to the manufacturer's instructions. Trimming was performed at a thickness of 20 µm. Of each lobe, several sections were prepared cutting the tissue into 7-µm serial slices with a microtome blade (C35, FEATHER Safety Razor Co. Ltd., Osaka, Japan).

Three to four consecutive sections of the same liver tissue were picked up per superfrost/plus slide. The slide was kept at RT until it was full. Afterwards, the slide was stored in the cooling chamber of the microtome at -20°C in a slide box. 12 to 20 sections per individual liver tissue were prepared. The slide boxes were temporarily stored in a container with dry ice and afterwards transferred to -80°C for long-term storage.

4. Mounting

The frozen slides with liver tissue were removed from -80°C storage and thawed at room temperature. Condensed water was dried with lint-free tissue, carefully wiping around the specimen. For fluorescence microscopy, the tissue sections were mounted with a bubble-free drop of Fluoromount-G (SBA-0100-01 Southern Biotechnology Associates Inc. by Biozol Diagnostica Vertrieb GmbH, Eching, Germany) and covered by fitting a cover slip in a 45° angle lowering it carefully with a blunt end forceps. Excess mounting medium was removed. After initial drying for at least 30 min in the dark at RT slides could be used for fluorescence microscopy.

2.2.5.2 Immunohistochemical staining of liver tissue

IHC (immunohistochemical) peroxidase staining was performed according to the manufacturer's protocol (DAKO EnVision⁺ Dual Link System-HRP (DAB⁺) Kit). All reagents had to adjust to RT before starting the procedure.

1. Endogenous enzyme block

Frozen liver tissue sections were removed from -80°C storage and immediately transferred into a cuvette filled with PBS for thawing.

The slides were removed from the PBS cuvette and excess buffer was tapped off. Using a lint-free tissue, any remaining liquid was removed by carefully wiping around the specimen. Drops of Dual Endogenous Enzyme Block were applied to cover the specimen and incubated for 5 to 10 min. The slides were then gently rinsed with distilled water from a wash bottle avoiding direct targeting of the tissue sections and then placed in a fresh PBS bath.

2. Primary-antibody staining

Slides were wiped dry as described before. Diluted primary antibody (in PBS final volume 200 µL) was applied to each specimen making sure to cover it completely (Tab. 2.1). Primary-antibody staining lasted 30 min at RT. Next, the slides were gently rinsed with PBS from a wash bottle and placed in a fresh PBS bath. Flow was not focused directly on tissue.

Table 2.1 Peroxidase staining primary antibodies. Respective dilution, reference, clonality, and concentration are listed. All antibodies were purchased from Thermo Fisher Scientific.

Antibody	Concentration	Reference / lot	Clonality
Anti-CyclinA2 rabbit IgG	2 µg	PA1-30395 / #RK22968753	polyclonal
Anti-LargeT mouse IgG	1 µg	MA1-90661 / #RK2294473	monoclonal
Anti-GFP rabbit IgG	20 µg	A-11122 / 1661233	polyclonal

3. Labeled polymer-HRP

Slides were wiped dry as described before. Labeled polymer was applied to cover the specimen; incubation followed for 30 min. Slides were rinsed as in step 2 and placed in PBS bath for 5 min.

4. Substrate-chromogen

Slides were wiped as described before. The substrate-chromogen solution was applied to cover the specimen. Slides were incubated for 5 to 10 min and then rinsed gently with distilled water from a wash bottle. Flow was not focused directly on tissue.

5. Counterstain

Slides were washed three times with PBS in a cuvette in the dark for 5 min. Following the final washing step, a 1:20,000 dilution (40 µg) of Hoechst33258 (Pentahydrate (bis-Benzimidazole) [10 mg/mL in water], 352/461nm; #H3569, Thermo Fisher Scientific, Karlsruhe, Germany) in PBS was prepared in a cuvette to stain the nuclei. The slides were incubated for 2 min in the dark at RT. To stop the staining, the slides were placed in a cuvette with fresh PBS and kept in the dark.

6. Mounting

Slides were mounted as previously described with Fluoromont-G. The slides were kept in the folder horizontally for one day at RT until they were dried and then stored in the dark at 4°C.

2.2.5.3 Immunofluorescence staining of liver tissue

IF (immunofluorescence) staining was performed for liver slices of all groups.

1. Fixation

Frozen liver tissue sections were removed from -80°C storage and immediately transferred into a cuvette filled with ice-cold acetone for fixation. They were incubated for 10 minutes at RT under the chemical hood.

2. Washing

Tissue slides were removed from the acetone and put into a new cuvette filled with PBS. After 5 min of incubation at RT, the PBS was discarded. Washing was repeated three times, using fresh PBS each time.

3. Primary-antibody staining

Respective the desired dilution (Tab. 2.2), the primary antibody was diluted in a final volume of 200 μ L per slide using PBS.

Table 2.2 Overview of primary antibodies used for IF staining in this thesis. Dilution, reference, clonality and concentration are mentioned. IF, immunofluorescence. All antibodies were purchased from Thermo Fisher Scientific.

Antibody	Concentration	Reference / lot	Clonality
Anti-CyclinA2 rabbit IgG	2 μ g	PA1-30395 / #RK22968753	polyclonal
Anti-LargeT mouse IgG	2 μ g	MA1-90661 / #RK2294473	monoclonal
Anti-GFP rabbit IgG	20 μ g	A-11122 / 1661233	polyclonal

After the final washing step, a single slide was taken out of the cuvette filled with PBS. Any remaining liquid was removed by carefully wiping around the specimen with a lint-free tissue. The slide was placed in a light-protected staining chamber and the diluted primary antibody (200 μ L) was pipetted dropwise onto the sections, taking care that each section was covered completely with the solution. The slides were prepared one by one to prevent dehydration. Slides were incubated over night in the dark at 4°C at high humidity.

4. Secondary-antibody staining

After 15-18 hours, the slides were washed three times for 5 min in a cuvette filled with PBS as described above. Per slide, a final volume of 200 μ L of the desired antibody dilution (Tab. 2.3) was required.

Table 2.3 Optimal dilution of secondary antibodies Alexa Fluor 488 and 633. All antibodies were purchased from Thermo Fisher Scientific.

Antibody	Concentration	Reference / lot	Clonality
Alexa Fluor 488 chicken anti-rabbit IgG (H+L)	5 μ g	A-21441 / 1697089	polyclonal
Alexa Fluor 488 goat anti-mouse IgG, IgA, IgM (H+L)	5 μ g	A-10667 / 1723707	polyclonal

Post final washing, the slides were removed from the PBS cuvette and dried carefully with lint-free tissue without disturbing the sections. The antibody solution was pipetted dropwise onto the sections covering them completely. As described above, one slide after the other was prepared. Again, the slides were placed in the light-protected staining chamber. The slides were incubated for 45 to 60 min in the dark at 4°C.

5. Nucleus staining

Sections were counterstained with Hoechst33258 as described for immunohistochemical staining.

6. Mounting

Mounting was performed as described for immunohistochemical staining.

2.2.5.4 Calculation of positive staining in IF and IHC images

Positively stained areas in immunohistochemical and fluorescence microscopy analyses were calculated by ImageJ 1.51j8 (Wayne Rasband, National Institutes of Health). IHC images were converted into 8-bit images and the stained areas were calculated in square millimeters and percent in relation to the displayed area of liver slices by adjustment of the threshold. Stained areas of IF images were calculated by adjustment of the color threshold (Tab. 2.4).

Table 2.4 Minimum and maximum threshold adjustments of IHC and IF images. IF, immunofluorescence; IHC, immunohistochemistry.

Group	Staining	Minimum Threshold	Maximum Threshold
RGB	IHC αGFP	0	210
	IF	43	255
CyclinA2	IHC αGFP	0	148
	IHC αCyclinA2	0	148
	IF αCyclinA2	55	255
HRas-V12	IHC αGFP	0	129
	IF αGFP	72	255
LargeT	IHC αGFP	0	202
	IHC αLargeT	0	202
	IF αLargeT	60	255

2.2.6 Statistics

Data were analyzed with the GraphPad Prism v5.02 software. Statistical analyses were performed with one-way ANOVA for non-parametric data set, Bartlett's test for equal variances and post-test Dunnett's multiple comparison test for group-wise comparison. P values <0.05 were considered statistically significant.

2.2.7 Workflow

Freshly isolated primary murine hepatocytes derived from wildtype mice were transduced with different lentiviral vectors. A scheme of the experimental workflow is depicted in Figure 2.3.

As a control, one group of USB (urokinase-type plasminogen activator/severe combined immunodeficiency/beige) mice received non-transduced hepatocytes (Mock); 6 out of 8 animals could be analyzed. For the RGB-control group, hepatocytes transduced with vectors LeGO-C2, LeGO-V2, and LeGO-Cer2 encoding red, green, and blue (RGB) fluorescent proteins were transplanted into 15 recipient mice; 11 transplanted animals were analyzed. In the following, this control group will be termed "RGB". The putative oncogenes *CyclinA2*, mutated *HRas-V12*, or *SV40-LargeT-antigen* all in combination with a green fluorescent

protein, either enhanced GFP (eGFP) or Venus, were chosen for transduction of primary murine hepatocytes followed by intrasplenic transplantation into 4-week old homozygous USB mice. Since the fluorescent protein expression cassette was linked via an IRES (internal ribosomal entry site) to the oncogene driven by an SFFV promoter, the oncogene and fluorescent proteins were co-expressed in all cells, which were stably transduced with these vectors. In one group, primary hepatocytes were transduced with LeGO-CyclinA2-iG2 vector and transplanted into 22 recipient mice. Of these 22 animals, 11 animals survived until final analysis. This group was named “CyclinA2” after the oncogene *CyclinA2*. For stable integration of mutated HRas-V12, LeGO-HRas-V12-iG2 vector particles were used for transduction of primary murine hepatocytes. In total, 16 USB mice were transplanted. Four animals died during the first week after transplantation, hence 12 animals were analyzed at the final time point. “HRas-V12” will be used as term for this group in the following. 12 out of 15 USB mice transplanted with LeGO-LargeT-iV2 transduced hepatocytes were finally analyzed, as 3 animals died shortly after transplantation. This group will be referred to as “LargeT”. Unexpected deaths in all groups could be explained due to the damaged liver phenotype of transgenic mice, transplantation failure, or by individual susceptibility to analgetic medication.

General health conditions of mice were checked daily, whereas weight was determined one to three times a week. Livers and tumor growth were monitored via repeated magnetic resonance imaging (MRI) examinations. Mice were euthanized when they displayed symptoms of stress and worsened general health status. Severity of the symptoms was evaluated according to a rating system approved by the authorities of the Hamburg State (Behörde für Gesundheit und Verbraucherschutz): 5 points each: weight loss >10%, bristled coat, or abnormal behaviour; 10 points each: weight loss >20%, limited locomotion, isolation, decreased reaction, or dehydration; 20 points: tumor size >2.5 mm. If mice were rated a 20 or higher, they had to be euthanized. At final analysis, liver albumin (ALB) and alanine aminotransferase (ALT) concentrations were determined in the serum. Livers from all experimental groups were investigated by microscopy of immunohistochemically and immunofluorescently stained cryosections, as well as quantification of integrated vector copy numbers by droplet-digital PCR.

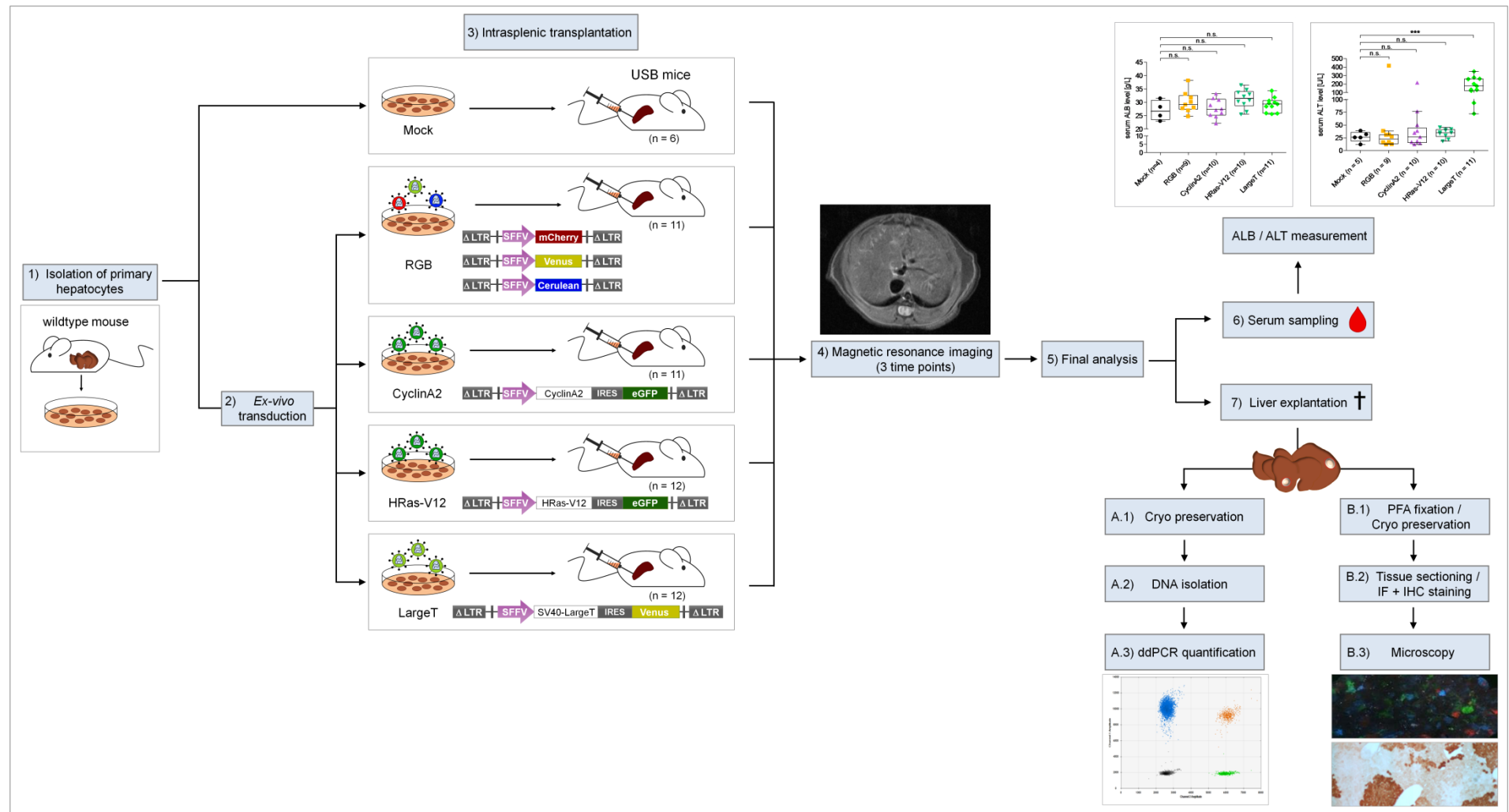


Figure 2.3 Workflow of mouse experiments. 1) Primary murine hepatocytes were isolated from C57Bl6/J_2014 wildtype mice by *in-situ* perfusion. 2) Subsequently, hepatocytes were transduced *ex vivo* with red-green-blue (RGB) vectors or the oncogene vectors CyclinA2, HRas-V12 or SV40-LargeT (LargeT), and 3) transplanted into the spleen of USB mice. Proviral DNA scheme is displayed below the respective groups. Untransduced hepatocytes (mock) were transplanted in a separate control group. Group size of analyzed animals is indicated in brackets. 4) Starting three weeks after transplantation, MRI screenings were performed at 3 time points. 5) Depending on health status and/or tumor size, animals were finally analyzed. 6) Blood was taken for 6.1) serum ALB/ALT measurements. 7) Explanted liver lobes were separated. Lobes were used either for A.1) cryo preservation for A.2) DNA isolation and A.3) ddPCR quantification, or for B.1) PFA fixation and cryo preservation to continue with B.2) tissue sectioning and B.3) microscopy of IHC and IF stained samples. ALB: albumin; ALT: alanine aminotransferase; ddPCR: digital droplet PCR; IHC: immunohistochemical; IF: immunofluorescent; MRI: magnetic resonance imaging; PFA: paraformaldehyde.

3 Results

Incidence and mortality rates of HCC are steadily rising around the whole world (Castelli et al. 2017). Although the field of hepatocarcinogenesis is well studied, no curative therapy is available for the complex and multi-step process of HCC formation yet (Forner and Bruix 2012). This demonstrates the importance of finding a suitable *in-vivo* model to gain a better understanding of underlying mechanisms and ultimately to improve treatment options for HCC. Using neonatal mice, Ranzani and colleagues were able to identify oncogenes involved in murine HCC formation that match human HCC by using lentiviral vectors (Ranzani et al. 2013). However, their artificial model does not resemble the situation of human patients closely. *In-vivo* studies with RGB marking in connection with lentiviral vectors have demonstrated that this is a useful approach to investigate liver regeneration, but also tumor formation (Cornils et al. 2014; Thomaschewski et al. 2017; Weber et al. 2011). The aim of this thesis was to take advantage of RGB marking in combination with oncogene-expressing lentiviral vectors to induce liver tumor formation in USB mice and to study the benign and malignant clonal regeneration of the liver *in vivo* and thus establishing a new tumor model. The workflow of experiments is summarized in Figure 2.3, p. 45.

3.1 Physiological weight gain in all experimental groups

Health status of transplanted USB mice was monitored daily, and weight was determined at least weekly (Fig. 3.1). During the first two weeks, mice stayed at the same weight level, whereas four weeks after transplantation (Tx) mice gained >200% of their initial weight. HRas-V12-transplanted mice lost up to 30% body weight during the first 9 days but doubled their weight 7 weeks post Tx. This progress was expected considering the age of the newly transplanted mice, which was 4 weeks, and normal physiological development. However, if the transplantation had failed and no viable hepatocytes would have reached the liver, the mice would have died due to their enduring liver damage caused by uPA transgene expression. Hence, weight gain indicated successful transplantation. 8 weeks after transplantation weight gain decreased to 1-20% per week since the mice reached adulthood. In all groups, constant weight gain could be observed with standard deviations of $\pm 20\%$. Comparison of animals from all groups showed weight changes to a similar extent within the groups. Mean comparison of all groups revealed HRas-V12-transplanted mice as lightest during the whole course of experiments. However, it remained open whether this was related to the transgene or represented normal biological variations.

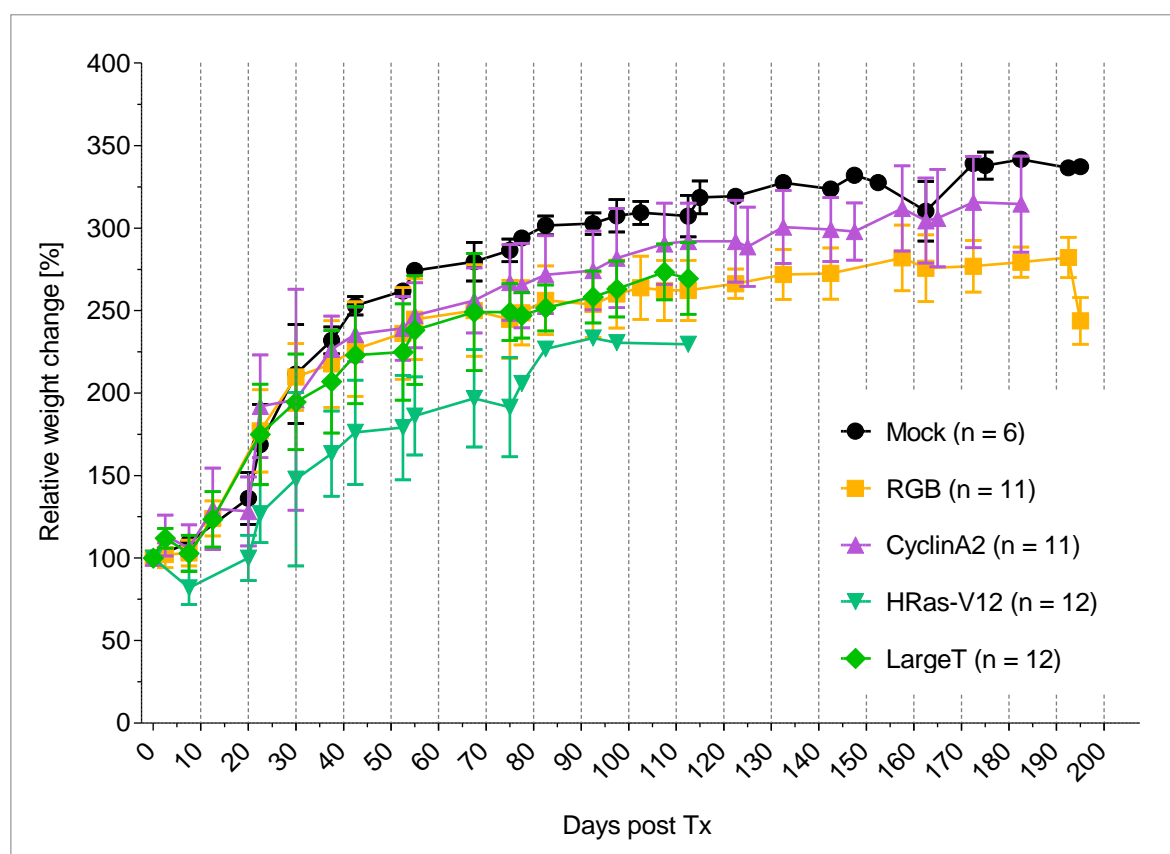


Figure 3.1 Relative weight changes [%] of transplanted mice during the course of time [days post Tx]. 4-week old USB recipient mice were transplanted with either untransduced (mock, black circles), RGB- (orange squares), CyclinA2- (violet triangles), HRas-V12- (turquoise triangles), or LargeT-transduced (green rhombus) primary murine hepatocytes. Numbers in brackets indicate analyzed animals per group. Mean \pm SD for each group (GraphPad Prism v5.02).

Summarizing, mock control mice transplanted with non-transduced hepatocytes were the heaviest of all animals with constant weight gain. Second highest weight increase was observed for CyclinA2-transplanted mice. Trends of relative weight gain were similar for RGB- and LargeT-mice. At final time point, the weight of RGB-mice decreased by 14% to the previous time point. All in all, relative weight gain was within normal physiological limits for all groups indicating successful transplantation and engraftment of transduced hepatocytes.

3.2 HRas-V12- and LargeT-transplanted mice showed early symptoms of disease progression

End analysis was executed according to health status and/or tumor size measured via magnetic resonance imaging and decided individually for each animal. Signs of a worsened general health status included: reduced weight, a bristled coat, abnormal behavior (isolation, limited locomotion, or slow reaction), dehydration, or tumors >2.5 mm. The animals were sacrificed if at least three of these criteria occurred and were rated a 20 or higher according to the applied rating system (see chapter 2.2.7 Workflow, p.44). Mice transplanted with HRas-V12- or LargeT-transduced hepatocytes displayed symptoms of worsening general

health condition earlier than the other groups. As reference, mock and RGB control-mice or even CyclinA2- or HRas-V12-transplanted mice were sacrificed at similar time points indicated by “#” in Figure 3.2. Remaining animals of all groups were maintained until they showed a worsened condition, as well.

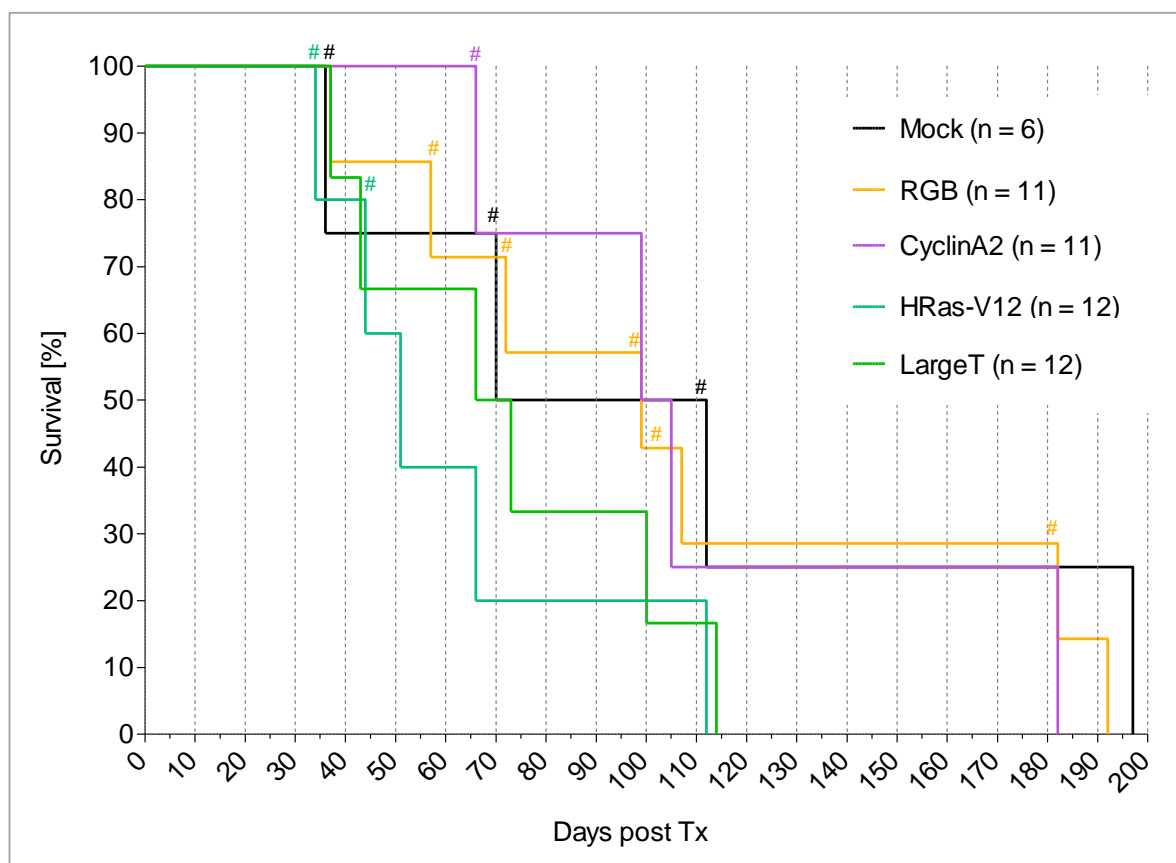


Figure 3.2 Kaplan-Meier survival curves [%] of transplanted animals [days post Tx]. HRas-V12- and LargeT-transplanted mice showed signs of stress and worsening health condition first compared to all groups. # indicates sacrifice of animals as reference. Remaining animals of all groups were maintained until they showed signs of worsening general health condition (GraphPad Prism v5.02).

34 days after transplantation, two HRas-V12-transplanted mice were sacrificed, and one animal of the mock group was sacrificed 36 d post Tx as reference. One RGB- and LargeT-transplanted animal were sacrificed 37 d post Tx each. 43 d post Tx, another LargeT-mouse was euthanized. Two HRas-V12-transplanted animals were sacrificed 44 d post Tx as a reference. Another mouse of the same group had to be euthanized 51 days after transplantation. As reference, two RGB-transplanted animals were sacrificed 57 d post Tx. Three LargeT-transplanted mice and six mice transplanted with HRas-V12-transduced hepatocytes were euthanized 66 d after Tx. Three CyclinA2-mice were sacrificed as reference. On day 70 after transplantation one mock-transplanted mouse was sacrificed, followed by sacrifice of one mouse transplanted with RGB-transduced cells on day 72, and two LargeT-transplanted animals 73 d post Tx. One CyclinA2-transplanted animal and two RGB-transplanted animals were sacrificed 99 d post Tx. 100 days after Tx, two mice

belonging to the LargeT group had to be sacrificed. Next, three mice of the CyclinA2-group were sacrificed on day 105 post Tx. One RGB-mouse was sacrificed 107 d after transplantation as reference. The final remaining HRas-V12-transplanted mouse was sacrificed together with two mock-mice 112 d post Tx. Last three surviving LargeT-animals were euthanized 114 d post Tx. Final analysis of the remaining four CyclinA2-transplanted mice was performed on day 182 post Tx. Longest survival was observed in RGB-mice as the remaining four animals had to be euthanized 192 d post Tx, and mock-mice, of which the last remaining two animals were sacrificed 197 d post Tx.

Survival of all analyzed mice demonstrated groups LargeT and HRas-V12 to exhibit early symptoms of worsening health condition, e.g., decreased weight, bristled coat as first mice had to be sacrificed 37 and 51 days after transplanation already. Mice of these groups survived until 112 and 114 days maximum. CyclinA2-transplanted mice showed first signs of stress 99 d post Tx and final sacrifice had to be executed 182 d post Tx. Mice of control groups mock (197 d) and RGB (192 d) were sacrificed latest due to worsening general health condition. These observations lead to the conclusion that early liver disease progression would be expected in HRas-V12 and LargeT groups, which could be confirmed by the following analyses.

3.3 Normal albumin values were observed in all groups

Albumin (ALB) is an important parameter for general health status. Strongly decreased ALB levels indicate ongoing inflammation and predict poor disease outcome (Levitt and Levitt 2016). At final bleeding, serum samples of transplanted mice were taken and analyzed to determine ALB concentrations. Some samples could not be investigated as the serum volume was not sufficient for analysis. Figure 3.3 presents ALB levels [g/L] for all analyzed animals.

Mock-transplanted animals showed 27.0 ± 3.7 g/L ALB in final serum. Similar ALB values were found for RGB-control mice (30.0 ± 4.0 g/L) and CyclinA2-mice (27.8 ± 3.4 g/L). Albumin levels of 31.3 ± 3.5 g/L were measured for HRas-V12-transplanted mice. Besides, ALB levels of mice transplanted with LargeT-transduced hepatocytes were normal as well as 29.1 ± 2.7 g/L ALB verified (Fig. 3.3).

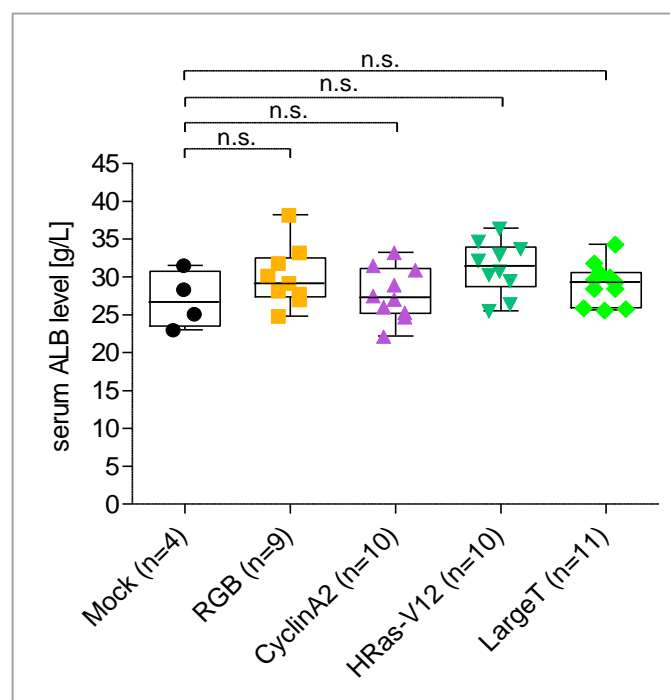


Figure 3.3 Final measurement of serum albumin levels (ALB) [g/L] confirmed normal values for all experimental groups. Samples were analyzed on a Cobas Integra 400 plus (Roche Diagnostics Deutschland GmbH). Median, quartiles, minimum and maximum values are displayed. One-way ANOVA for non-parametric data set; $\alpha = 0.05$, Bartlett's test for equal variances, Dunnett's multiple comparison test (GraphPad Prism v5.02); n.s., not significant.

Overall, normal albumin levels were determined for all mice in final serum samples indicating a healthy phenotype.

3.4 Highly elevated alanine aminotransferase levels in LargeT-transplanted mice

Elevated liver alanine aminotransferase levels (ALT [U/L]) in the serum, which were measured after final bleeding of the mice, are an indicator of ongoing liver inflammation and/or damage (Fig. 3.4). Groups of RGB-control and CyclinA2-transplanted mice showed one outlier each, which were excluded from statistical analysis to avoid bias.

Mice in the mock control group demonstrated ALT levels of 26.9 ± 9.9 U/L. ALT values of RGB-control mice were determined to be within physiological limits (23.3 ± 10.0 U/L). In CyclinA2- (32.1 ± 21.1 U/L) and HRas-V12-mice (34.1 ± 8.7 U/L), ALT was slightly elevated compared to control groups mock and RGB. No significant difference was found in ALT concentration of groups RGB, CyclinA2 and HRas-V12 compared to the mock control group. In contrast, mice transplanted with LargeT-transduced hepatocytes displayed severe liver damage and significantly increased ALT levels (190.0 ± 88.1 U/L) by factor 6.35 compared to mock-mice (Fig. 3.4).

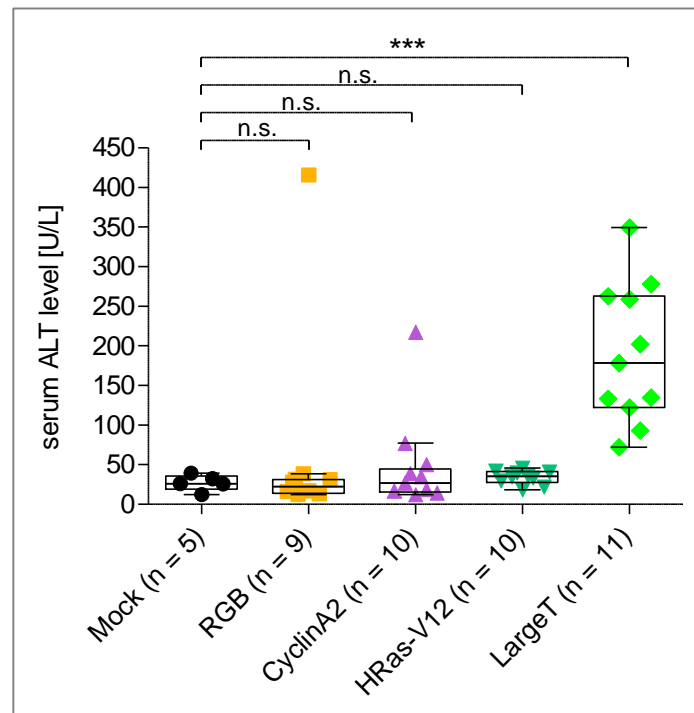


Figure 3.4 Liver alanine aminotransferase (ALT) levels [U/L] were determined from final serum samples. All groups showed normal ALT levels except for LargeT-transplanted mice displaying significantly elevated ALT levels compared to mock control group. Samples were analyzed on a Cobas Integra 400 plus (Roche Diagnostics Deutschland GmbH). Median, quartiles, minimum and maximum values are displayed. Statistical analysis of RGB and CyclinA2 groups excluded one outlier each. One-way ANOVA for non-parametric data set, $\alpha = 0.05$, Bartlett's test for equal variances, Dunnett's multiple comparison test (GraphPad Prism v5.02); n.s., not significant; ***, $p < 0.0001$.

In conclusion, liver alanine aminotransferase levels of groups mock, RGB, CyclinA2, and HRas-V12 were within normal range. Only LargeT-transplanted mice exhibited strongly increased ALT values pointing towards liver damage and inflammation.

3.5 Magnetic resonance imaging revealed distinct group kinetics

To monitor cell engraftment and tumor growth, transplanted mice were examined by magnetic resonance imaging (MRI) regularly (Fig. 3.5). Due to a different cell architecture, tumors are expected to show a different contrast in MRI compared to healthy tissue, and thus appear bright in T2-weighted images. If the transplantation failed, severe liver damage and apoptosis would be observable in these animals. Transplanted animals were scanned at three time points starting 3 weeks after transplantation.

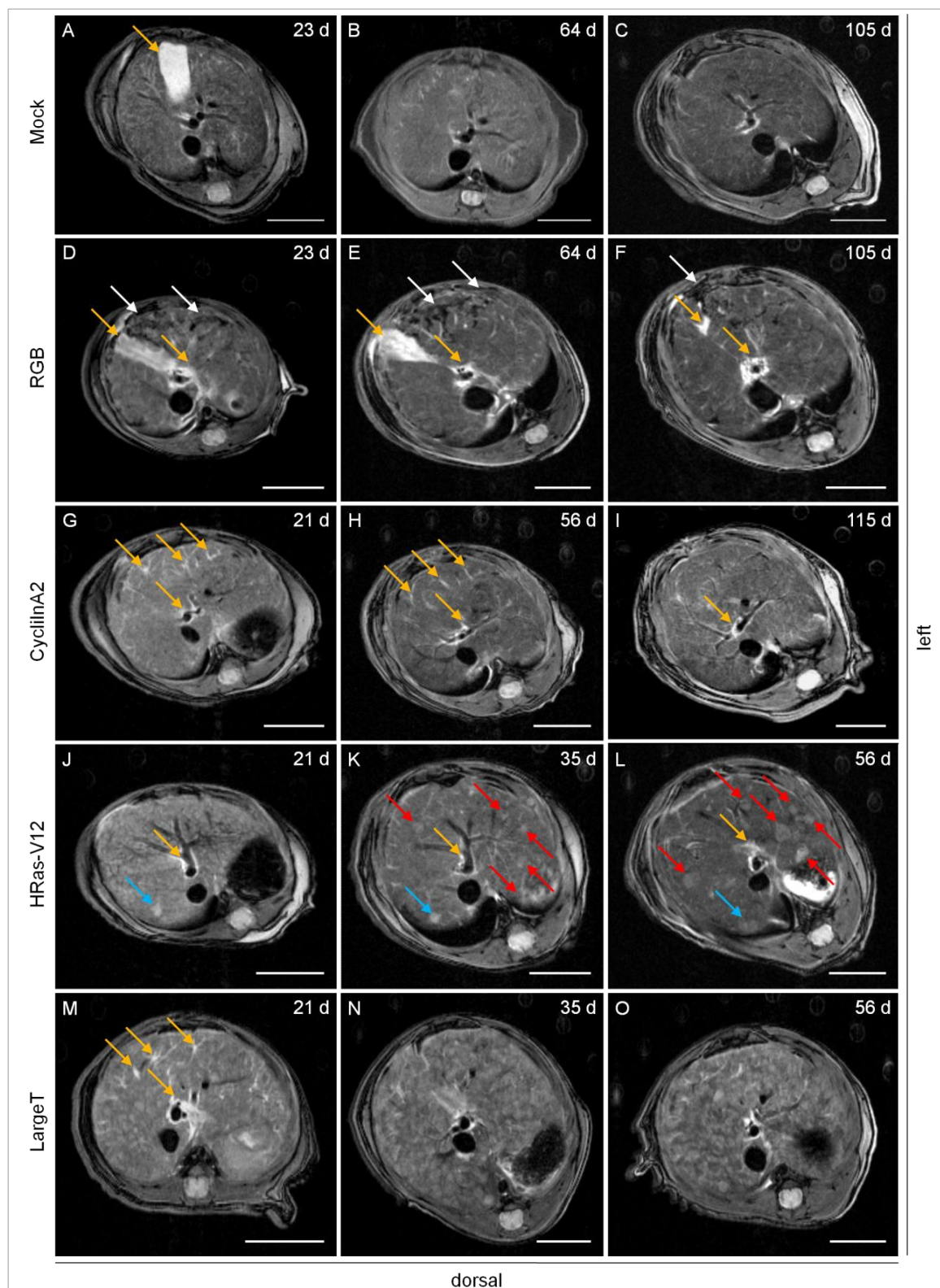


Figure 3.5 Representative magnetic resonance images of transplanted mice during the course of time. Mock control mice showed healthy liver regeneration 23 (A), 64 (B) and 105 (C) days post Tx with highlighted bile ducts (A, orange arrow). RGB-transduced hepatocytes resulted in regeneration of liver tissue with small injured areas (white arrows), which were reduced over time, 23 (D), 64 (E) to 105 (F) days post Tx. Healthy liver regeneration with minor inflammation signs (orange arrows) was observed in mice transplanted with CyclinA2-transduced hepatocytes 21 (G), 56 (H) and 115 (I) days post Tx. Liver cancer progression in HRas-V12-transplanted mice was apparent with single tumors (cyan, red arrows) 21 (J), 35 (K) and 56 (L) days post Tx. LargeT-mice showed severe liver damage and rapid tumor development 21 (M, orange arrows, highlighted bile ducts), 35 (N) and 56 (O) days after transplantation Tx. All T2-weighted images were taken in transversal orientation by a 7T MRI system (ClinScan, Bruker BioSpin GmbH). Scale bars: 0.5 cm.

Over time, non-transduced hepatocytes uniformly repopulated the mock recipient livers and resulted in a healthy phenotype (Fig. 3.5 A-C). 23 days after transplantation, bile ducts were highlighted in the MRI (Fig. 3.5 A, orange arrow) indicating high bile acids production, which are a sign of liver damage or injury. This observation was not unusual since recipient mice expressed the uPA transgene, which induced hepatocyte death. After hepatocyte transplantation biliary tracts also regenerate induced by donor cells and connect to endogenous bile ducts (Meuleman et al. 2005). All follow-up MRI data indicated that generation of bile canaliculi reduces over time, when full liver repopulation is achieved and obviously cell proliferation stops.

Transplantation of RGB-transduced primary murine hepatocytes resulted in liver repopulation with small areas of intense liver damage 23 days post Tx (Fig. 3.5 D, white arrows). In course of time, the injured tissue recovered and mice developed a healthy phenotype with minor bile excess indicated by orange arrows (Fig. 3.5 D-F). Liver damage could be caused by inflammation due to uPA transgene expression and hepatocyte death. Another explanation could be blocked sinusoids of distal liver lobules. This could have happened after transplantation if hepatocytes agglutinated and inhibited blood flow resulting in lack of nourishment causing hepatocyte death. Over time, the liver tissue could recover and damaged areas were repopulated (Fig. 3.5 D-F, white arrows). CyclinA2-transduced hepatocytes engrafted and did not show any tumor characteristics on MRI and declining inflammation over time (Fig. 3.5 G-I, orange arrows). In contrast, HRas-V12-transplanted mice showed tumorigenesis, even though at a relatively low level. First tumors were found 21 days post Tx (Fig. 3.5 J, cyan arrow) with a size of 0.5 mm. Size of this tumor rose via 160% (0.8 mm) at day 35 (Fig. 3.5 K, cyan arrow) to 350% (1.7 mm) 56 days after transplantation (Fig. 3.5 L, cyan arrow). Over time, several tumors emerged (Fig. 3.5 K, L, red arrows) with indication of constant inflammation (Fig. 3.5 J-L, orange arrows), however surrounding liver tissue remained healthy. Mostly impaired liver architecture was observed in recipients of LargeT-transduced hepatocytes, which exhibited accelerated tumor growth displaying innumerable tumor load (Fig. 3.5 M-O). High bile activity was observed 21 d post Tx (Fig. 3.5 M, orange arrows).

To conclude, mock-transplanted hepatocytes resulted in a healthy liver phenotype of recipient mice. Although liver damage occurred in RGB-transplanted mice early after transplantation, it was reduced over time and livers remained healthy. Despite some signs of inflammation, livers of mice transplanted with CyclinA2-transduced hepatocytes had a conserved healthy phenotype. Multiple slowly growing tumors were identified in HRas-V12-recipients. In contrast, a disrupted liver architecture was observed in livers of LargeT-transplanted mice. In summary, magnetic resonance imaging facilitated visualization of distinct kinetics of liver regeneration in the different experimental groups.

3.6 Liver phenotype characteristics were group-specific

At the final time point of analysis, examination of livers revealed a healthy phenotype in the mock group with minor fibrosis 112 d and 197 d after transplantation (Fig. 3.6 A, B, black circles). RGB-transplanted mice displayed a healthy phenotype with single fibrotic areas 112 d post Tx (Fig. 3.6 C, black circle) and small fibrotic areas at later time points (Fig. 3.6 D, black circle). 115 d after transplantation, CyclinA2-mice showed few fibrotic areas (Fig. 3.6 E, black circles), but apart from that a healthy phenotype. However, the whole liver surface of CyclinA2-transplanted mice was covered with fibrotic lesions 183 d post Tx (Fig. 3.6 F). In line with the magnetic resonance images, at least 12 single tumors surrounded by healthy liver tissue were detected 66 d after transplantation in HRas-V12-mice ranging from 0.6 mm to 1.8 mm in diameter (\varnothing) (Fig. 3.6 G, red circles). During progression, liver phenotype impaired and individual tumors as large as \varnothing 3.6 mm at 112 d post Tx (Fig. 3.6 H, red arrow). Livers of mice transplanted with LargeT-transduced hepatocytes demonstrated a completely disrupted architecture already 66 d post Tx shown in the 1.5x zoom in (Fig. 3.6 I). In line with the previous MRI findings, red circles and arrows indicate numerous tumors from \varnothing 0.4 mm to \varnothing 2.2 mm at or beneath the surface of the liver. Moreover, at day 101 after transplantation LargeT-mice exhibited countless tumors of different sizes; one even up to 6.5 mm in diameter (Fig. 3.6 J, red arrow).

Phenotypical characteristics were detected for livers in all groups. Untransduced (mock) and RGB-transduced hepatocytes resulted in healthy recipient livers with small fibrotic areas. Countless small fibrotic lesions were found in CyclinA2-transplanted mice. Although livers of HRas-V12-transplanted mice showed a healthy phenotype, several tumors emerged. Worst outcome was detected in mice transplanted with LargeT-transduced hepatocytes, which gave rise to innumeral fast growing tumors.

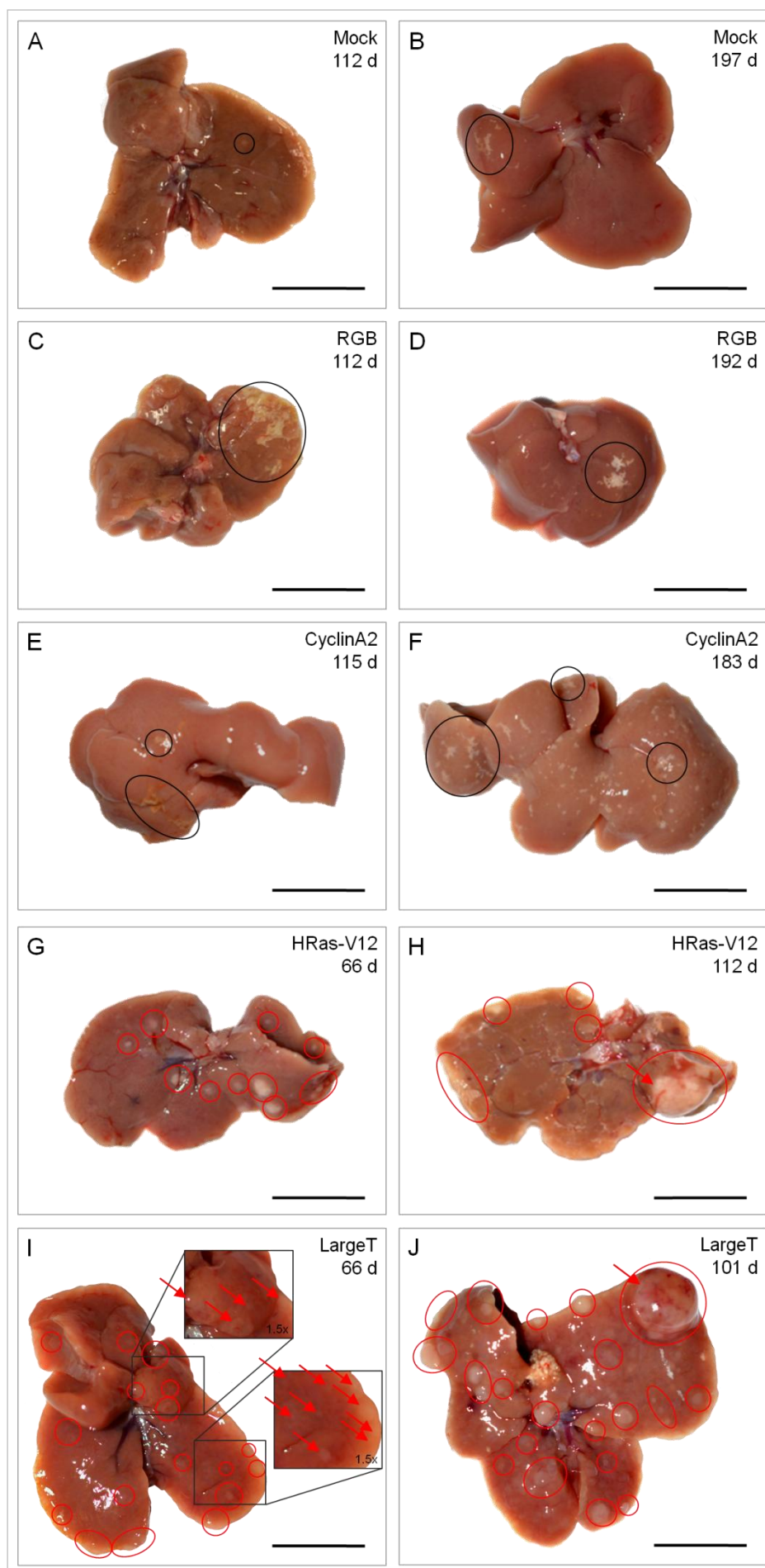


Figure 3.6 Livers of all animals of the single groups showed group-specific outcome at end of experiment.

Representative samples are shown for each group. Control mice transplanted with untransduced hepatocytes (mock) exhibited healthy liver regeneration with minor scar formation 112 (A) and 197 (B) days after transplantation. RGB-transduced hepatocytes led to regeneration of liver tissue with some fibrotic areas 112 d (C) and 192 d (D) post Tx, but in general mediated a healthy liver phenotype. Despite minor fibrotic areas in livers transplanted with

CyclinA2-transduced hepatocytes, these mice showed healthy liver regeneration 115 d post Tx (E). 183 d post Tx fibrotic areas were found all over the liver surface of CyclinA2-mice (F). Single tumors emerged from HRas-V12-transduced hepatocytes 66 (G) and 112 (H) days after transplantation. Livers of LargeT-transplanted mice exhibited malformed liver reconstitution 66 d post Tx (I) and innumerable fast growing tumors 101 d post Tx (J). Livers from A, C, E, G, and I originated from the same animals as shown in Fig. 3.5. Black circles, fibrotic liver damage; red circles and arrows, tumors. Scale bars: 1.0 cm.

3.7 Immunohistochemical staining confirmed engraftment of *ex-vivo* transduced transplanted hepatocytes

Several immunohistochemical (IHC) staining methods could be used for localization of expressed proteins of interest in sectioned tissue samples. In this work, these were the fluorescent proteins and the oncogenes introduced to the experimental groups by transduction of donor hepatocytes with the respective lentiviral vector constructs. One advantage of immunoperoxidase staining is based on the catalyzed reaction from colorless DAB (3,3'-diaminobenzidine) to a brown precipitate, which allows sensitive detection and can be enhanced by extended incubation. IHC staining was performed with a specific antibody to detect expressed fluorescence proteins in all groups, except for the mock group as these mice got untransduced hepatocytes and thus, transplanted hepatocytes would not express any fluorescence protein. GFP, Venus and Cerulean were detected with an α GFP antibody, which recognizes all green fluorescent proteins and its derivatives. Hence, the same antibody could be used to stain on the one hand Venus and Cerulean in RGB-transplanted mice, and on the other hand co-expressed GFP in CyclinA2- and HRas-V12 groups, as well as Venus in LargeT-mice. Moreover, detection of the oncogenes CyclinA2 and LargeT was accomplished by specific antibodies. No suitable antibody was available for α HRas-V12 staining, therefore mouse livers of the HRas-V12 group were stained for expression of the green fluorescent protein only. Percentages of stained areas were determined using the ImageJ software. Each image represented an area of 2.39 mm². As expected, multiple stained areas were found in RGB-transplanted mice 101 d post Tx (Fig. 3.7). Different staining grades could be caused either by 1) multiple vector integrations per cell, which resulted in varying expression levels of detectable fluorescence proteins, 2) enhanced expression due to transactivation depending on the integration site, and/or 3) differing antibody binding affinity for Venus or Cerulean. Evaluation of stained areas revealed 54.0% of the total representative area shown in Figure 3.7 to be positive for Venus, or Cerulean, or both in one cell.

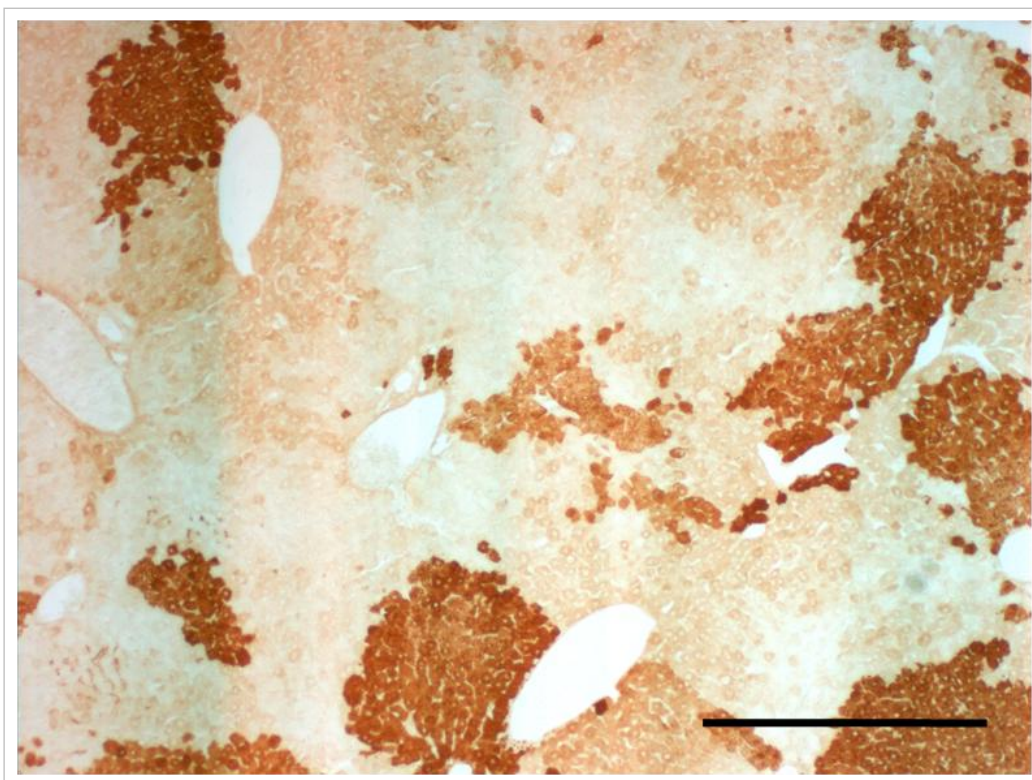


Figure 3.7 Representative peroxidase staining visualized multiple areas stained with different intensities in a USB mouse liver repopulated with RGB-transduced hepatocytes. This mouse was sacrificed 101 days post Tx. α GFP-staining of 7- μ m cryo-slice. Scale bar: 500 μ m.

Figure 3.8 displays consecutive slices of the same animal transplanted with CyclinA2-transduced hepatocytes stained for GFP (A) and CyclinA2 (B) 66 d post Tx. An excellent co-localization was demonstrated for both antibodies. Calculation of GFP-positive stained area resulted in 18.4% and matched computed 19.1% positive CyclinA2-staining of the displayed area.

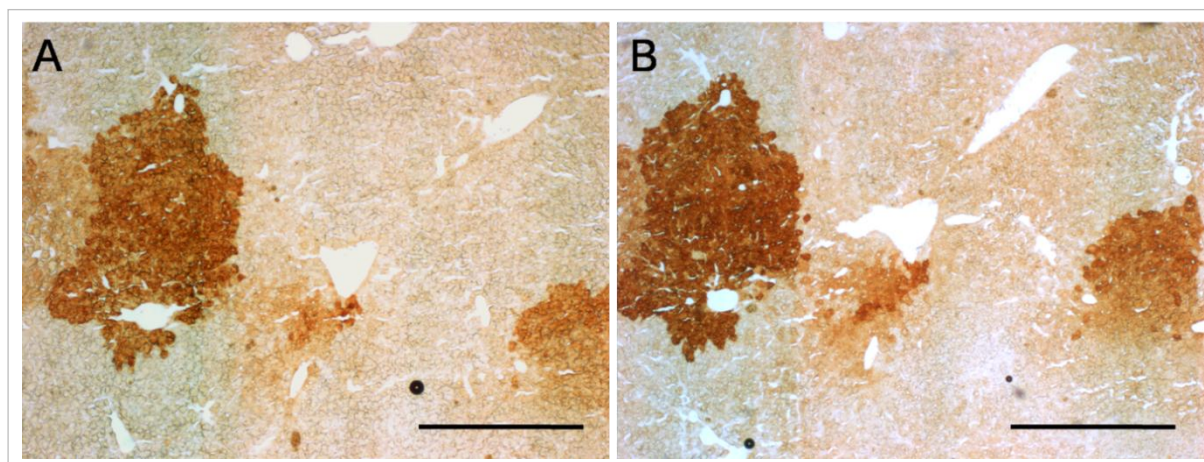


Figure 3.8 CyclinA2-transduced hepatocytes engrafted in small repopulation areas of USB mouse livers 66 d post Tx. Representative images are shown. α GFP-staining (A) and α CyclinA2-staining (B) of consecutive 7- μ m cryo-slices. Scale bars: 500 μ m.

Single repopulated areas, 13.8%, were stained positive for GFP in HRas-V12-transplanted mice 48 d post Tx (Fig. 3.9).

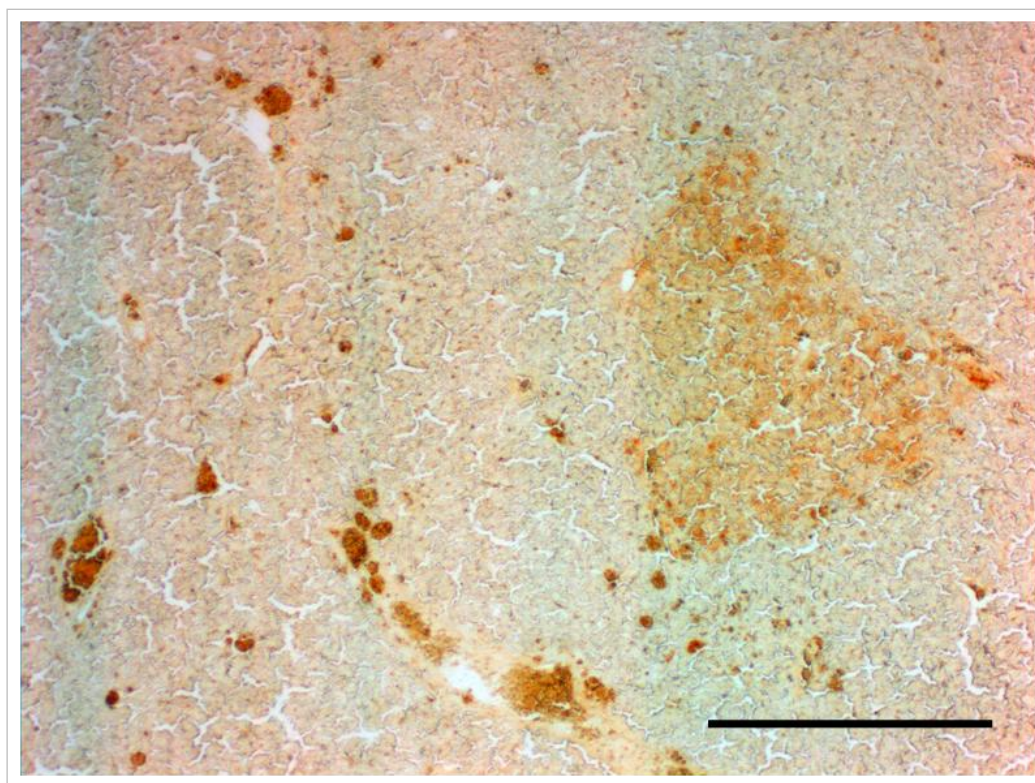


Figure 3.9 Repopulated USB mouse liver showed small areas of engrafted GFP-expressing HRas-V12-transduced hepatocytes 48 days post Tx. A representative image is shown. α GFP-staining of 7- μ m cryo-slice. Scale bar: 500 μ m.

The highest values for positively stained areas were found in LargeT-transplanted mice. 72.0% of the displayed representative area were found to be Venus-positive (Fig. 3.10 A, 101 d post Tx). Furthermore, 66 d after transplantation 63.9% of the area of another mouse liver were stained positive for LargeT (Fig. 3.10 B).

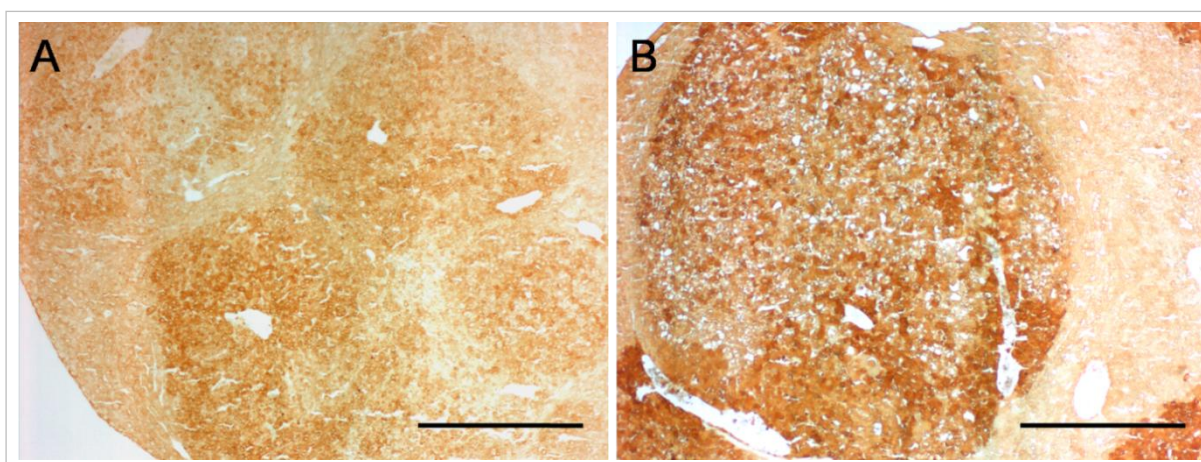


Figure 3.10 Large repopulation areas of LargeT-transduced hepatocytes transplanted into USB mouse livers were positive for Venus (A, α GFP, 101 d post Tx) and LargeT (B, α LargeT 66 d post Tx). Representative images are shown. 7- μ m cryo-slices were stained. Scale bars: 500 μ m.

In summary, immunohistochemical staining with specific antibodies as α GFP, α CyclinA2 or α LargeT allowed visualization of repopulated areas expressing the respective transgenes in all experimental groups. RGB-transplanted mice showed high engraftment rates, whereas CyclinA2 and HRas-V12 groups displayed single repopulated areas. As expected based on MRI and phenotypic examinations, analyses of LargeT-transplanted livers yielded in positively stained areas considerably higher than in the other oncogene groups.

3.8 Confirmation of oncogene and fluorescence protein expressions by immunofluorescence analyses

In addition to immunohistochemical analyses, fluorescence microscopy was performed. Since the results of immunoperoxidase staining confirmed oncoprotein-, and fluorescence protein-positive repopulation areas, these should be visualized by fluorescence microscopy as well. Indeed, microscopy of RGB-transplanted livers revealed fluorescent signals in repopulated areas without any antibody staining. The same primary antibodies (α GFP, α CyclinA2, α LargeT) as used for immunoperoxidase staining were applied for immunofluorescence (IF) staining of CyclinA2-, HRas-V12- and LargeT-transplanted mouse livers in combination with an Alexa-Fluor 488 secondary antibody resulting in positively stained areas. ImageJ software was used to determine percentages of stained areas

Liver cryosections of RGB-transplanted mice revealed manifold multicolor repopulated areas in the recipient livers 101 d post Tx (Fig. 3.11 A). Color spectrum of cell clones ranged from yellow, orange, pink and red to green, violet and blue. The mixed colors and their diverse intensities originated from various integrations of different vector combinations into the same donor cell (Weber et al. 2011). After transplantation, the hepatocytes engrafted, started proliferation and thus, clonal regeneration of the recipient livers occurred. Besides, the intensity of the expressed fluorescent proteins is also influenced in a great part by the integration sites of the proviral DNA (Weber et al. 2011), which were not investigated in this study. Evaluation by ImageJ confirmed 61.2% repopulated liver tissue with fluorescence protein-expressing cells in view of the total slice area of 1430.92 mm². 66 d post Tx, liver cryosections of CyclinA2-transplanted mice were to 16.5% CyclinA2-positive considering the depicted area of 2.39 mm² (Fig. 3.11 B). In contrast, 22.6% of 2.39 mm² liver sections of mice transplanted with HRas-V12-transduced hepatocytes were GFP-positive at day 66 after transplantation (Fig. 3.11 C). The displayed area of 1.91 mm² was considered for calculation of LargeT-positive hepatocytes in mice of the respective group; 29.8% were found to be positive 72 d post Tx (Fig. 3.11 D).

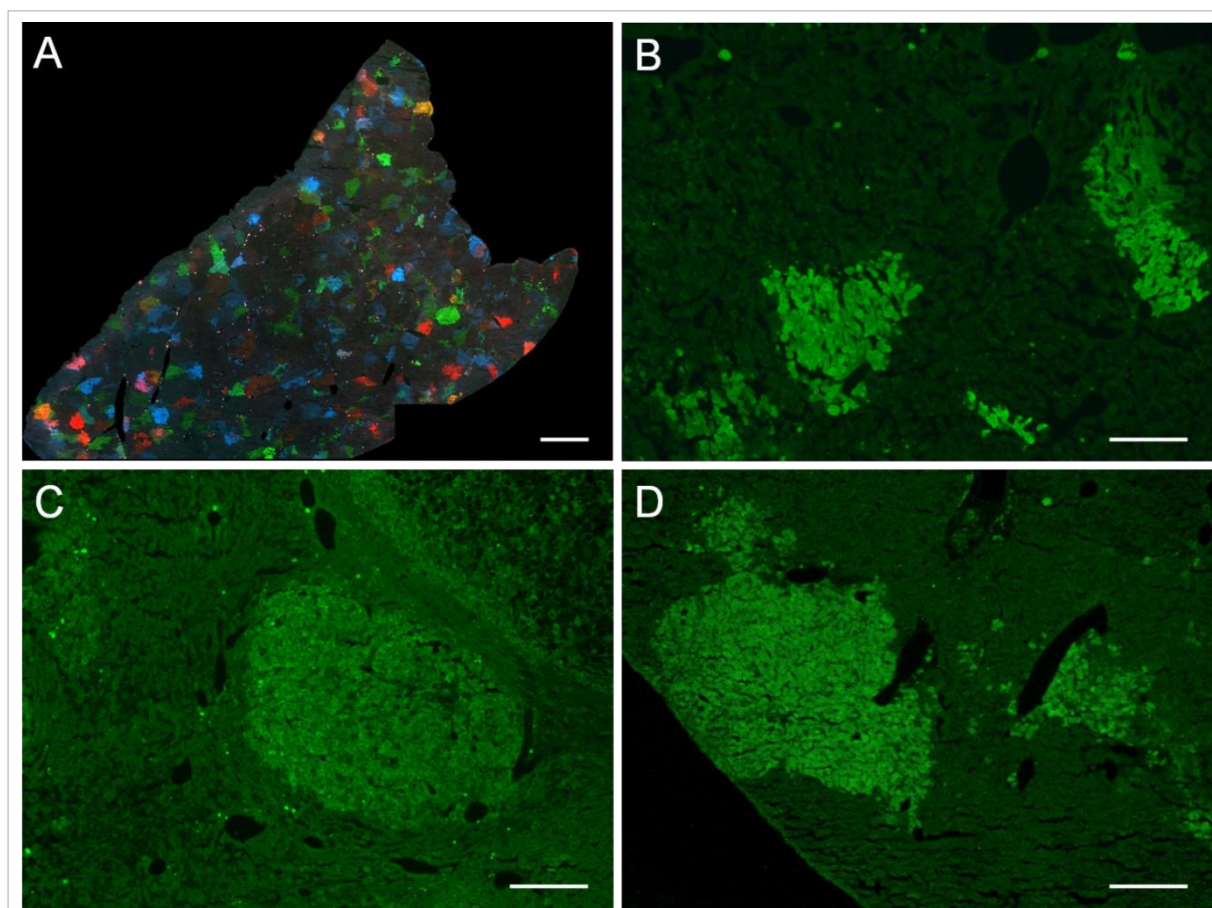


Figure 3.11 Fluorescence microscopy of liver cryosections revealed engraftment of transduced transplanted hepatocytes in all experimental groups. Representative images are shown. RGB-transduced hepatocytes repopulated livers of USB mice with numerous cell clones of different colors (A, no staining; 101 d post Tx, stitched overlay images by Microsoft Image Composite Editor v2.0.3). CyclinA2-transduced hepatocytes gave rise to only relatively small repopulation areas (B, α CyclinA2-staining; 66 d post Tx), whereas in livers of mice that got HRas-V12 (C, α GFP-staining; 66 d post Tx.) or LargeT (D, α LargeT-staining; 72 d post Tx) tumors originating from the oncogene-transduced hepatocytes were found. Scale bars: 4 mm (A); 200 μ m (B-D).

Comparison of immunofluorescence data demonstrated highest repopulation levels in livers of RGB-transplanted mice. The fluorescent proteins were expressed at high rates due to the strong SFFV (spleen focus forming virus) promoter activity of the LeGO vectors. Moreover, the applied high multiplicity of infection allowed for multiple integrations of all vectors. This indeed resulted in more combination possibilities and a highly differentiated color pattern in the transplanted and repopulated recipient livers. CyclinA2 and HRas-V12 groups showed repopulation with modified hepatocytes to a lower extent, whereas examination of livers of LargeT-transplanted mice confirmed one third to be repopulated with transgene-expressing cells. These numbers differed from results of immunoperoxidase staining, which could be most likely explained by biological divergence as livers of different animals were stained in both procedures. Nonetheless, these analyses contributed to the identification of single cell clones in transplanted mice.

3.9 Molecular quantification of integrated proviral DNA by droplet-digital PCR

3.9.1 Identification of specific primer/probe combinations on NIH control cells

To quantify integrated vector copy numbers (VCN) on a molecular level, droplet digital PCR (ddPCR) was performed. To establish the method, first NIH3T3 cells were transduced with viral vectors at MOIs of 1 and 20 for each construct. Percentages of positively transduced cells were measured by flow-cytometric analyses (LSR Fortessa, Becton Dickinson; Tab. 3.1). Specific primer/probe combinations targeting either i) the fluorescent proteins *GFP* or *Venus*, ii) the oncogenes *CyclinA2*, *HRas-V12*, or *LargeT*, or iii) the *SFFV* promoter were mixed with a primer/probe set specific for the murine reference gene *Erythropoietin receptor* (*Epo-R*) for duplex analyses. All primer/probe combinations were identified using the Primer Express software and later on were established on gDNA of transduced NIH3T3 control cells. To do so, gDNA of NIH3T3 cells transduced with an MOI of 1 was used, since these cells showed fluorescent-protein expressions of 15.5 - 30.3% (Tab. 3.1), and ddPCR analyses require a certain amount of negative droplets to be reliable.

Table 3.1. Transduction efficiency of NIH3T3 cells as genetic control for molecular analysis measured 3 days post transduction. C2, mCherry, red fluorescent protein; Cer2, Cerulean, blue fluorescent protein; G2, eGFP; i, internal ribosomal entry site; IP, infectious particles; LeGO, lentiviral gene ontology; MOI, multiplicity of infection; V2, Venus, green fluorescent protein. *Viral supernatant of LeGO-CyclinA2-iG2 was kindly provided by Kristoffer Riecken (Dept. of Cell and Gene Therapy, University Medical Center Hamburg-Eppendorf, Germany).

Construct	Titer [IP/mL]	MOI 1 positive cells [%]	MOI 20 positive cells [%]
LeGO-C2	2.52E+07	16.5	89.9
LeGO-V2	2.38E+07	16.3	92.4
LeGO-Cer2	2.97E+07	15.5	92.6
LeGO-CyclinA2-iG2*	2.00E+07	30.3	97.2
LeGO-HRas-V12-iG2	1.44E+07	27.0	89.8
LeGO-SV40-LargeT-iV2	1.26E+07	17.6	89.8

Relative vector copy number was calculated by division of measured copies of the gene of interest by copies found for *Epo-R*. It is known that NIH3T3 cells have a hypertriploid karyotype, thus this result provided the number of integrated vector copies per triploid genome. Results were background corrected against ddPCR results performed on gDNA extracted from untransduced NIH3T3 cells.

RGB-transduced NIH3T3 cells showed a relative vector copy number of 0.085 for *GFP*. As expected, VCN for *Venus* (0.215) were about one third of *SFFV* (0.636) (Fig. 3.12). As these cells were transduced with three vectors expressing either mCherry, Venus or Cerulean fluorescence proteins, the false-positive result for *GFP* could be explained by 18 bp sequence homology of the *GFP* probe to the *Cerulean* cDNA sequence and/or 12 bp

sequence homology to *Venus* leading to detection of those integrated cDNAs instead. ddPCR with primer/probe combinations for *CyclinA2*, *HRas-V12* and *LargeT* were negative in the RGB sample.

High numbers of *GFP*-integrations (VCN = 0.636) were found for *CyclinA2*-transduced NIH3T3 equivalent to VCN for *CyclinA2* (VCN = 0.629), and VCN for *SFFV* (VCN = 0.602). On very low level, samples were positive for *Venus* (VCN = 0.001). Again, sequence homology of *GFP* and *Venus* probes could result in false-positive VCN for *Venus*, which was probably due to the aforementioned sequence homology. *HRas-V12* and *LargeT* primer/probe sets were negative for gDNA of *CyclinA2*-transduced NIH3T3. VCN comparison of *GFP* (VCN = 0.405), *HRas-V12* (VCN = 0.445), and *SFFV* (VCN = 0.469) showed matching results of control cells transduced with *HRas-V12*. ddPCR with *CyclinA2* and *LargeT* specific primer/probe sets was negative for the *HRas-V12*-control cells. Finally, vector copy numbers of 0.289 and 0.317 were determined for *Venus* and *SFFV* by ddPCR on gDNA from *LargeT*-transduced NIH3T3 cells; all other primer/probe combinations remained negative.

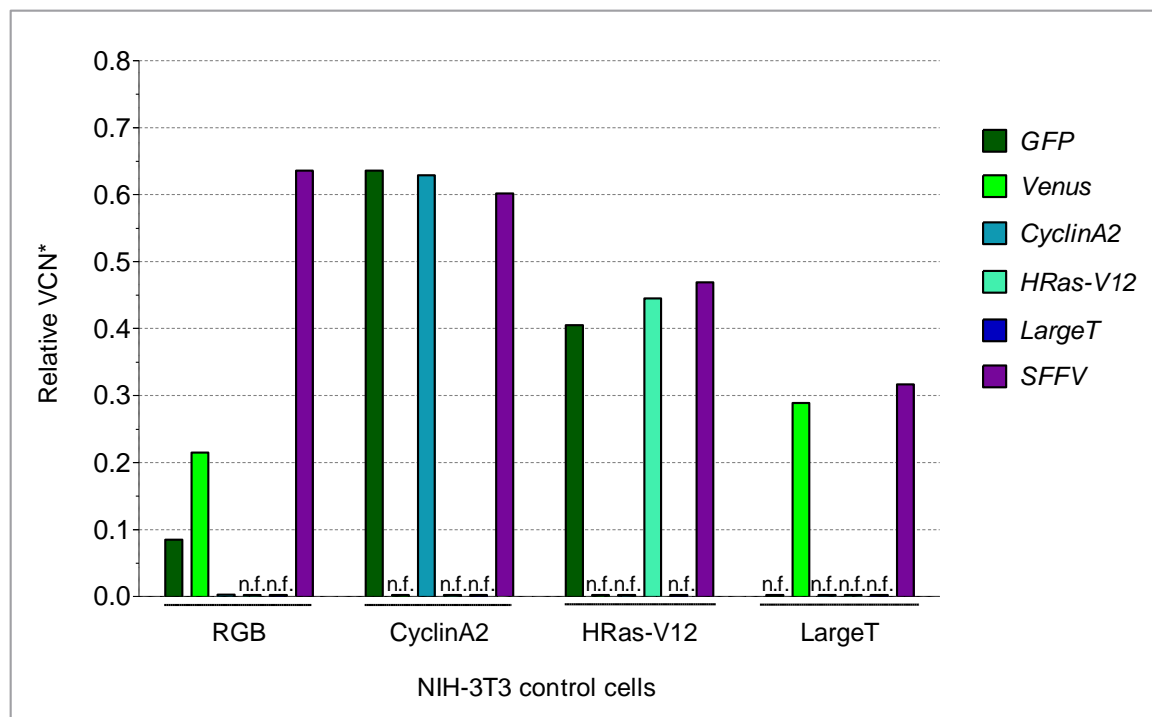


Figure 3.12 Transduced NIH3T3 control cells proved specificity of primer/probe combinations for detection of genes of interest. *GFP*, *Venus*, *CyclinA2*, *HRas-V12*, *LargeT*, and *SFFV* were analyzed. Background of non-transduced NIH3T3 gDNA was subtracted for calculation of relative vector copy number (VCN) by dividing copy numbers for genes of interest by reference gene *Epo-R* copy numbers and karyotype of 3N. *Epo-R*, Erythropoietin receptor; *GFP*, green fluorescent protein; *SFFV*, spleen-focus forming virus promoter; VCN, vector copy number, *VCN = 1 means 1 copy per genome; n.f., not found.

Altogether, analyses of NIH3T3 control cell ddPCR results proved that most of the chosen primer/probe sets were excellently suited to quantify integrated copy numbers of the introduced genes. Results were in line with percentages of positively transduced cells measured by flow cytometry (Tab. 3.1). False-positive levels of *GFP* found in RGB-transduced cells as well as false-positive *Venus* in CyclinA2-transduced NIH3T3 control cells were due to sequence homology of these two probes. Even so, the cross-reaction of these two primer/probe sets was not relevant to the experiment itself. Only in the RGB group more than one vector construct was introduced; however, it was not necessary to distinguish precisely between the integrated genes encoding for the three different fluorescent proteins as this experiment aimed at quantification of vector copy numbers in general. The groups CyclinA2, HRas-V12, and LargeT each had received a vector with only a single fluorescent-protein cDNA, so detection of this specific gene would be limited to either the *Venus* or *GFP* primer/probe set. Since mice and transplanted cells were not exposed to any other fluorescent-protein cDNA, any cross-reaction would have been improbable. *LargeT* could not be detected by ddPCR at all; therefore this primer/probe combination was excluded and all future analyses were performed with the remaining sets.

3.9.2 Integration of lentiviral constructs was verified in murine liver samples

Genomic DNA of mouse livers was examined with the relevant primer/probe combinations for each group, respectively. Murine samples of RGB-transplanted livers were analyzed with primer/probe sets for *Venus* and *SFFV*. CyclinA2- and HRas-V12-transduced liver samples were tested for expected 1) *GFP*, 2) *SFFV* and 3) *CyclinA2* or *HRas-V12* integrations. Primer/probe sets for *Venus* and *SFFV* were used for examination of LargeT-transduced livers. To calculate vector copy numbers, the polyploidy has to be considered. By 85% chance, hepatocytes have a 4N karyotype and at least two nuclei per cell (Epstein 1967; Gentric and Desdouets 2014). Hence, VCN of genes of interest were calculated in relation to reference gene *Epo-R* copy numbers. Results of relative vector copy numbers for all groups were background corrected against VCN of mock control group and are summarized in Figure 3.13.

In RGB-transplanted livers, vector copy numbers of 0.229 ± 0.140 of integrated *Venus* and 0.798 ± 0.456 of *SFFV* promoter were found. Analyses of CyclinA2-transplanted mice presented similar results with low standard deviation for all primer/probe sets. Vector copy number of *GFP* was measured (0.264 ± 0.091) and coincided with VCN of *CyclinA2* (0.322 ± 0.114) and *SFFV* (0.314 ± 0.100). Using the three primer/probe sets *GFP*, *HRas-V12* and *SFFV*, comparable VCN were determined in livers of HRas-V12-transplanted mice ($GFP = 0.255 \pm 0.183$; $HRas-V12 = 0.365 \pm 0.195$; $SFFV = 0.50 \pm 0.307$). In contrast, a

vector copy number 0.867 ± 0.555 of *Venus* was identified in LargeT-transplanted livers. VCN of *SFFV* was determined as 0.997 ± 0.699 in these mice.

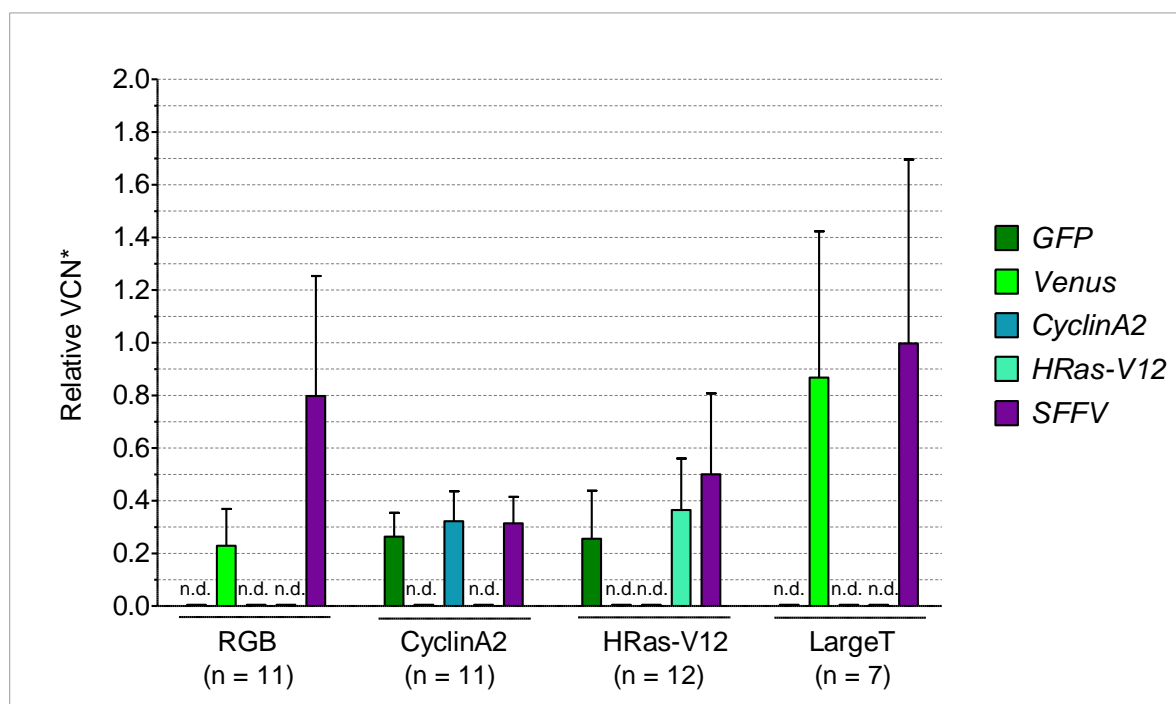


Figure 3.13 Integrated vector DNA correlated well within groups quantified by droplet-digital PCR. Primer/probe combinations specific for the fluorescent protein expression cassette (GFP or Venus), the oncogenes *CyclinA2* or *HRas-V12*, or the *SFFV* promoter were tested for all groups. With an 85% chance karyotype of binucleate hepatocytes is 4N, thus VCN were calculated by division of gene of interest VCN by reference gene Epo-R VCN and 6.8. Calculated relative VCN were background corrected against ddPCR results on gDNA of mock-transduced mouse liver tissue. Mean + SD are displayed. GFP, green fluorescent protein; SFFV, spleen-focus forming virus; VCN, vector copy number, *VCN =1 means 1 copy per genome in reference to murine Erythropoietin receptor copy number; n.d. not determined.

In conclusion, ddPCR allowed molecular quantification of integrated vector copy numbers for all experimental groups. In RGB-transplanted mice, *SFFV* was found at about 3-fold higher numbers compared to *Venus*. This was expected as all three vectors (LeGO-C2, LeGO-V2, and LeGO-Cer2) used for transduction of transplanted cells contained the *SFFV* promoter. Conversely, primer/probe combinations specific for the *SFFV* promoter, the fluorescent protein GFP and the genes of interest of *CyclinA2*- and *HRas-V12* resulted in comparable vector copy numbers in transplanted mice, respectively. *Venus* and *SFFV* were present in corresponding levels in mice transplanted with LargeT. VCN of LargeT-transplanted mice compared to groups *CyclinA2* and *HRas-V12* were at least doubled, which correlated well with the proportion of positively stained areas in immunoperoxidase and immunofluorescence analyses of all groups.

4 Discussion

Understanding the underlying mechanisms of disease development is crucial to prevent progression and optimize treatment options for any type of cancer. This holds also true for hepatocellular carcinomas. Although at early stages of the disease, liver resection, transplantation, and ablation can be performed to offer a potential cure for HCC patients (Forner et al. 2012; Pellicoro et al. 2014), about 50 to 80% incur disease recurrence and die from it within 5 years after surgery (Blum 2005; Kim et al. 2007). Patients suffering from advanced HCC have a median survival of about 6 to 8 months without treatment (de Lope et al. 2012; Llovet and Bruix 2008). Whereas survival rates can be increased by drug treatment, there is no ultimate cure for HCC (Bruix et al. 2011). In this study, a novel liver tumor mouse model was established, forming the basis to gain knowledge about molecular mechanisms involved in HCC formation and to eventually design new strategies to fight liver cancer.

Cellular and molecular heterogeneity is one major pitfall in liver tumors. Several risk factors contribute to hepatocarcinogenesis and result in complex and diverse tumors difficult to treat (Ghouri et al. 2017); these include deregulation of many cellular signaling cascades leading to a heterogeneous molecular profile in malignant degenerated cells (Forner et al. 2012; Liu et al. 2006; Meira et al. 2008; Sia et al. 2017). To perform *in-vivo* investigations, three different transgenes, *CyclinA2*, *HRas-V12*, and *SV40 LargeT-antigen*, of which the latter two are described oncogenes, were chosen for the present study. In more than 50% of hepatocellular carcinomas HRas signaling is activated (Liu et al. 2017; Mizushima et al. 2011). The mutated permanently-activated HRas-V12 oncogene is involved in cell-proliferation regulation causing uncontrolled cell division and tumor development (Bos 1989; Singh et al. 2014; Toss and Cristofanilli 2015). A small percentage of HCC cases show overexpression of *CyclinA2*, which contributes to high proliferative activity in affected cells (Sia et al. 2017; Wang et al. 2005). The proto-oncogene *SV40 LargeT-antigen* was previously connected with several malignancies (Ahuja et al. 2005; Javier and Butel 2008), and a number of groups have shown the cell-transformation capabilities of LargeT-antigen (Giri and Bader 2014; Holczbauer et al. 2013; Pan et al. 2015). LargeT-antigen-treated cells become immortal, proliferate, and survive, contrary to untreated normal cells (Giri and Bader 2014; Meng et al. 2014). Treated cells are also able to cause tumor formation in transplanted mice (Ahuja et al. 2005; Lou et al. 2005). In addition, other groups have investigated lentiviral-vector mediated oncogenesis in the liver using different mouse models. Ranzani and colleagues were able to identify four cancer-driving genes in HCCs by administration of lentiviral vectors to three different mouse models (Ranzani et al. 2013). However, only neonatal mice were used in their study, which is not close to the situation of human patients where the median age of HCC formation is about 65 (El-Serag and Rudolph 2007). Serial

transplantation of primary hepatocytes treated with lentiviral vectors expressing the *FAH* gene into adult *Fah*^{-/-} mice revealed maintenance of vector insertions' polyclonality, but did not result in tumor development (Rittelmeyer et al. 2013). Furthermore, lentiviral vectors have already been used successfully in several transplantation experiments exploring the clonal regeneration of the murine liver via RGB marking, thus highlighting the advantages of lentiviral vectors and their broad application spectrum (Cornils et al. 2014; Thomaschewski et al. 2017; Weber et al. 2011).

Based on this knowledge, a new reliable transplantation model for HCC studies could be generated in this thesis. Considering the two-hit hypothesis more than one event is required to induce tumorigenesis (Knudson 1971). In this study, *CyclinA2*, *HRas-V12*, and *SV40 LargeT-antigen* (short: LargeT) were used in the background of a liver damage mouse model. These proteins have already been described in HCC cases, and also were applied already in mouse studies in the past (Ahuja et al. 2005; Bakiri and Wagner 2013; Sandgren et al. 1989). Immunocompromized 4 week-old USB mice were transplanted intrasplenically either with untransduced (mock) wildtype hepatocytes, RGB-marked hepatocytes, or hepatocytes transduced with LeGO vectors containing one transgene combined with a green fluorescent protein. In total, five groups – mock, RGB, *CyclinA2*, *HRas-V12*, and LargeT – were analyzed each consisting of 6 to 12 animals. The examination of this mouse model confirmed fast hepatocarcinogenesis within the inflammatory milieu of USB transgenic mice when applying one of the two oncogenes *HRas-V12* and LargeT-antigen. Furthermore, the physiological and malignant clonal regeneration of the liver could be easily analyzed at the histological and molecular level, using RGB marking. The combination of sensitive cell-marking strategies using lentiviral vectors, and oncogene-mediated hepatocarcinogenesis in USB mice can provide new insights in the clonal evolution of the HCC, and potentially presents the necessary tools to establish a more realistic and patient-related *in-vivo* model for the investigation of liver cancer progression.

4.1 Malignant clonal regeneration in *HRas-V12* and LargeT recipient livers

The most critical step in establishing a new transplantation model is successful engraftment of transplanted cells. Hence, surgery was performed with special diligence, and animals were examined closely, particularly during the first two weeks after transplantation. Transplantation failure would have resulted in death of the animals due to their hepatotoxic phenotype caused by sustained uPA-transgene expression. Successful transplantation was evident already not only by the mice surviving but also by examining weight data of all transplanted mice. Up to 10 days after transplantation, mice in all groups showed only slight weight gain. This was most likely induced by the stress associated with the surgery. However, this lag period might also reflect the time required for the rescue of the ongoing liver failure by the

transplanted hepatocytes. In fact, not only engraftment but also proliferation of transplanted hepatocytes is required to substitute a sufficient proportion of the impaired endogenous liver tissue and to initiate physiologic organ function, which can take up to 7 days (Weber et al. 2009). However, after four to seven weeks mice in all groups had gained >200% weight, which was similar to wildtype C57Bl6/J mice under normal physiological conditions (online reference 5) indicating hepatocyte engraftment and proliferation. After reaching adulthood, the body-weight gain decreased to 1-20% weekly until the end of experiment, which was also expected (online reference 5).

Intra-group-consistent kinetics were ascertained with the observation of general health status, MRI screenings, and histological analyses. Mice receiving either mock-, RGB-, or CyclinA2-treated donor cells showed good recovery and survival rates. Mock mice were expected to survive the longest and to demonstrate efficient hepatocyte engraftment, since transplanted wildtype hepatocytes were not treated with any vector. Transduced cells are influenced in several ways by vector application, e.g., reduced cell survival, decreased engraftment potential and/or functionality due to genotoxicity or phenotoxicity (Baum et al. 2003). Furthermore, VSV-G-pseudotyped vectors can lead to cytotoxicity at high concentrations (Sakuma et al. 2012). In addition to hepatocyte isolation, treatment with any vector and culture medium mixed with viral supernatant causes stress to the hepatocytes (Weber et al. 2009). Hence, mock mice received less stressed hepatocytes and had higher chances for high engraftment rates and fast liver repair.

Although mice were healthy and showed a physiological liver phenotype at final analysis, RGB- and CyclinA2-transplanted animals experienced liver damage, which was visible partly in MRI screenings but also macroscopically. Recurrent injury and perpetual wound healing caused by organ damage, including ongoing inflammation due to apoptosis, can lead to hepatic stellate cell activation and matrix deposition initiating fibrosis (Block et al. 2003). Hence, scarred tissue was to be expected to some extent. In some mice, increased bile production as well as apoptotic liver tissue was found during the first weeks after transplantation, which was reduced over time. This observation correlated well with the literature: transplanted primary hepatocytes form biliary tracts, which connect to the host's bile system (Meuleman et al. 2005), and transplantation of hepatocytes can cause also ischemia-reperfusion injury by obstructing hepatic sinusoids, which are 6 to 9 μm in diameter in contrast to hepatocytes with a size of 20 to 40 μm (Weber et al. 2009). Over time, the damaged tissue is replaced. Consecutive MRI confirmed reduction of apoptotic tissue most likely related to liver regeneration by engrafted healthy transplanted wildtype hepatocytes over time.

LargeT-transplanted mice, followed by HRas-V12 recipients, first had signs of a bristled coat, limited locomotion, and isolated themselves. These were all indicators for an ongoing

disease progression. This observation was reinforced by magnetic resonance imaging showing tumor development already 21 days post Tx, which was confirmed after final liver examinations, and histology, and molecular analyses as well.

4.2 Significantly increased ALT concentrations in LargeT-transplanted mice

Indicators for ongoing liver inflammation also can be found in serum parameters. One of these is albumin (ALB), which comprises the majority of vertebrate blood plasma proteins and is produced in the liver (Levitt and Levitt 2016). Albumin offers various key metabolic functions, including general binding and transport of a broad variety of compounds, maintenance of colloidal osmotic pressure, and antioxidant properties (Taverna et al. 2013). Worsening hypoalbuminemia is associated with poorer outcome in a wide variety of diseases such as liver disease, rheumatoid arthritis, Wiskott-Aldrich syndrome or inflammatory diseases (Levitt and Levitt 2016; Nicholson et al. 2000). A study demonstrated albumin to predict mortality in non-surgical as well as post-surgical patients as survival decreased when ALB concentrations declined below 3.5 mg/dL, which is considered to be the lower limit of normal 4.3 mg/dL albumin levels in humans (Fulks et al. 2010). Therefore, measurement of ALB concentrations in the serum assists in observation of general health conditions. Notably, all groups in this study presented ALB levels in the physiological range. Mock-transplanted mice presented ALB levels similar to the other groups RGB, CyclinA2, HRas-V12, and LargeT closely resembling earlier reports of 35.5 ± 4.45 g/L albumin measured in the serum of untreated adult C57BL/6J wildtype mice (retro-orbital bleeding; male: 32.0 ± 4.1 g/L; female: 39.0 ± 4.8 g/L; Schnell et al. 2002).

Another sensitive and important parameter of liver damage is the alanine aminotransferase (ALT) level (Chen et al. 2006). In the liver, ALT plays a major role in the tricarboxylic acid cycle by catalyzing the enzymatic reaction from L-alanine and α -Ketoglutarat to pyruvate and L-glutamate (Liu et al. 2014). It has its highest concentration in the cytosol of hepatocytes and only is released to the serum when the liver is damaged. Various factors like medication, alcohol consumption or viral hepatitis influence ALT activity; hence, ALT is used as general marker to evaluate liver function and screen for liver disease. 27.1 ± 7.0 U/L are considered as normal serum ALT levels in healthy patients (Liu et al. 2014). As increased ALT concentrations in the serum reflect liver damage and maybe also ongoing inflammation, this parameter has become a relevant indicator of disease in diagnostics. The same applies for mouse models. Healthy mice present ALT concentrations of 69.0 ± 35.0 U/L (retro-orbital bleeding; male: 77.0 ± 36.0 U/L; female: 59.0 ± 34.0 U/L; Schnell et al. 2002). Transgene uPA/SCID mice have a serum ALT peak at 4 weeks of age (Weglarz et al. 2000), which declines reciprocally with repopulation levels of newly engrafted healthy hepatocytes. Hence, normal ALT levels indicate a healthy liver phenotype in USB mice. This observation was true

for the groups mock, RGB, CyclinA2, and HRas-V12 in this study. All analyzed samples for these groups were within the normal range. On the other hand, ALT levels in LargeT-mice were upregulated highly significant ($p < 0.0001$) compared to mock-mice indicating persisting liver damage. This was in contrast to the normal albumin levels in LargeT-transplanted mice (see above). High increase of ALT with normal albumin implies current liver damage but no liver toxic phenotype in LargeT-mice.

Despite tumorigenesis in HRas-V12 mice, serum parameters were stable and corresponded to physiological measures. These findings make HRas-V12 a promising candidate to investigate early and intermediate stages of liver disease.

4.3 Determination of transduction efficiency in transplanted hepatocytes

Previous studies demonstrated only about 30% of transplanted cells enter the hepatic chords, engraft and proliferate (Meuleman and Leroux-Roels 2008; Weglarz et al. 2000). About 70% of transplanted cells remain trapped in liver sinusoids if transplanted via the spleen or portal vein (Gupta et al. 2013; Weber et al. 2009). Transplanted wildtype hepatocytes have a selective growth advantage against endogenous transgene-carrying cells in the USB mouse model due to the pathophysiology of the endogenous plasminogen hyperactivation during early liver development resulting in their cell's death and 90% liver repopulation by allogeneic transplantation in homozygous mice (Heckel et al. 1990; Sandgren et al. 1991). The hepatic remodeling is complete with transplanted hepatocytes becoming indistinguishable from host endogenous cells on histological level within 3 to 7 days (Weber et al. 2009). On the other hand, hemizygous USB mice showed liver reconstitution levels of only about 15% (Dandri et al. 2001) referable to an override of uPA expression due to somatic recombination in a small proportion of host hepatocytes (Sandgren et al. 1991). Strikingly, the deletion of the uPA transgene can occur also in homozygous mice at any time, followed by clonal replication of the transgene-deficient hepatocytes (Sandgren et al. 1991). Normal repopulation in mice that lose transgene expression is reached after 2-3 months (Sandgren et al. 1991).

In this thesis, successful repopulation of transduced hepatocytes was proven on molecular and histological level as well. However, if only 30% of transplanted hepatocytes survive and engrafted cells are not longer distinguishable on histological level, the question needs to be answered, how many of these potential 30% of transplanted cells were successfully transduced. In this study, aliquots of transduced hepatocytes were kept in culture to analyze transduction rates microscopically and by flow cytometry after three days. It is not possible to calculate the transduction efficiency of newly transduced cells within one hour after vector application, since vector integration and stable gene expression take two to three days (Follenzi et al. 2000; Nguyen et al. 2002). Thus, fluorescence proteins are not detectable

directly after transduction. The same applies for flow cytometry with the additional disadvantage of losing a high proportion of cells by this analysis leaving an even smaller population for transplantation. Quantification of transduction efficiency is not even possible on molecular level at such early time points; vectors, which entered the target cell but are not integrated and remain in the cytosol or nucleus, and residual vector DNA persisting in the supernatant would interfere with ddPCR on genomic DNA as well and lead to an overestimation of the actual transduction efficiency. Hence, aliquots of transduced hepatocytes of all groups were seeded in triplicates prior to transplantation, and cultured for three days allowing the hepatocytes to express the integrated transgenes to a detectable niveau. However, on day 3 microscopic analyses revealed >95% dead cells with barely visible fluorescent protein expression. Flow cytometric analyses confirmed <5% viable cells with max. 3% of fluorescence protein expression in all groups. Thus this data confirms the high sensitivity of hepatocytes to environmental changes (Rothe et al. 2012; Severgnini et al. 2012). In addition, hepatocyte morphology changes over time in culture including de-differentiation and cell polarization (Bhandari et al. 2001; Li et al. 2010; Severgnini et al. 2012; Talamini et al. 1997; Wallace et al. 2010). Thus, a true transduction efficiency of transplanted hepatocytes could not be determined by this method. Another option to determine the proportion of transplanted cells, that are also positively transduced, could be additional IHC staining of explanted liver tissue in future experiments. A specific antibody staining against uPA should be close-to negative in wildtype cells but also in endogenous cells, which lost the transgene, but still hepatocytes highly positive for the transgene would be identified. Simultaneously stained liver slices with an antibody against Rag2 would not show any expression in endogenous cells as USB mice are null mutants contrary to wildtype hepatocytes, which express Rag2. An overlay of the uPA and Rag2 stainings, and subsequent liver slices positively stained with antibodies against introduced transgenes would reveal the proportion of efficiently transduced hepatocytes, which repopulated the recipient liver. In addition, chimerism of recipient and donor cells could be distinguished on molecular level with quantitative RT-PCR (Fehse et al. 2001) or digital PCR (Stahl et al. 2015).

The proportion of transgenic protein-expressing cells could also be measured by FACS analysis after preparing a single cell suspension of transplanted livers allowing for the calculation of engraftment and transduction rates.

4.4 Benign and malignant clonal liver regeneration in a liver damage mouse model

A new transplantation model was established in this thesis to address the need of a reliable *in-vivo* model for the investigation of liver cancer. The expression of lentiviral-integrated oncogenes was a key factor in this model. Multiple vector integrations enhance expression

levels of introduced transgenes and result in high vector copy numbers. Vector integrations directly relate to a high multiplicity of infection (Fehse et al. 2004), and as this study partly aimed at the induction of liver carcinogenesis, multiplicities of infection of 40 and 60 were applied in the performed experiments. However, high MOIs were not only chosen to induce enhanced integration levels, but also to affirm successful transduction.

Droplet-digital PCR validated relative VCN of integrated *Venus* around one third of *SFFV* promoter VCN in RGB-transplanted mice. CyclinA2- and HRas-V12-transplanted mice had comparable VCN for *GFP*, *SFFV*, and the respective oncogenes *CyclinA2* and *HRas-V12*. RGB-control mice presented highest VCN of *SFFV* and *Venus* together with mice transplanted with LargeT-transduced hepatocytes.

Despite a strong promoter and an MOI of 40 for each vector construct resulting in a total MOI of 120 and multiple integrations indicated by colorful FM data, no RGB recipients developed any HCC, but exhibited physiological liver repopulation. Obviously, multiple vector integrations alone are not sufficient to induce malignant transformation in the inflammatory background of the USB mice. In contrast, severe tumor formation was observed macroscopically in LargeT recipients contributing to the conclusion of LargeT-antigen's high transformation capability (Wang and Yang 2010). IHC and IF analyses confirmed ddPCR results: In RGB and LargeT recipients the largest areas were found repopulated by transgene-expressing hepatocytes.

Although VCN were quite similar in CyclinA2 and HRas-V12 groups corresponding well with IHC and IF analyses, no tumors developed in CyclinA2-transplanted mice. Only recipients of HRas-V12-transplanted hepatocytes favored tumorigenesis and displayed single nodules in the liver implying the relevance of the chosen oncogene and its function are crucial, and that not any introduced foreign gene can promote tumor formation. Again, this confirms the hypothesis that multiple hits – lentiviral integration and transient oncogene expression in a liver-damage background – are required to cause hepatocarcinogenesis.

Possibly, also CyclinA2 could induce tumors in mice if higher MOIs would be used. However, at some point the transduction efficiency would reach a plateau regardless of the vector particle excess as rodent adult primary hepatocytes are less susceptible to transduction by lentiviral vectors compared to human hepatocytes (Haridass et al. 2009; Nguyen et al. 2002). Transplantation success would also be decreased, if the DMEM:SFM ratio was too high during transduction with >60 MOI. Hepatocytes have to be cultivated in supplemented serum-free medium, but lentiviral vector particles are harvested in DMEM (Weber et al. 2011). DMEM has not optimal conditions to culture hepatocytes (Weber et al. 2009), and thus would negatively influence hepatocyte vitality. To preserve the hepatic phenotype and viability, primary hepatocytes were *ex-vivo* transduced polybrene for one

hour, even though incubation periods of at least 24h could improve transduction rates (Rothe et al. 2012).

4.5 Adaption of the Barcelona Clinic Liver Cancer classification system for the evaluation of hepatocarcinogenesis in a mouse model

In human patients, liver cancer is categorized into different stages. Both, the American and European Association for the Study of Liver Disease recommend the Barcelona Clinic Liver Cancer (BCLC) classification although there are several other staging systems (Kinoshita et al. 2015). The severity of liver disease, extent of tumor spread, and patient's general health status are evaluated according to BCLC (Llovet et al. 1999; Tab. 4.1).

Table 4.1 Barcelona Clinic Liver Cancer staging system. Based on Llovet et al. 1999.

Tumor stage	General health status	Tumor characteristics	Child stage
0 Very early	Good	Single nodule ≤ 2 cm	A - B
A Early	Good	Single nodule ≤ 5 cm, or 3 nodules ≤ 3 cm	A - B
B Intermediate	Good	Large, multiple nodules	A - B
C Advanced	Reduced	Vascular invasion, extrahepatic secondaries	A - B
D Terminal	Severely reduced	Any form	C

Unfortunately, no such system is established in mouse models. However, on the basis of the BCLC system a classification for liver tumor staging in the USB mouse model can be calculated. Tumor characteristics were adjusted to mean mouse liver weight. In this study, explanted mouse livers had an average weight of 1.49 g (Supplementary Table S5). A human liver with an average weight of 1.4 kg (Molina and DiMaio 2012) is considered as reference. Tumor characteristics and general health status according to the rating system used in this thesis were considered to score the USB mouse livers, since no data were available for α -fetoprotein, creatinine, bilirubin and alkaline phosphatase, which are evaluated in the Child scoring system (Selcuk 2017). Instead, ALB and ALT levels were considered. The adapted BCLC staging system is presented in Table 4.2.

Table 4.2 Adapted BCLC staging system for evaluation of liver disease in USB mice.

Tumor stage	General health status	Tumor characteristics	ALB level [g/L] *	ALT level [U/L] *
0 Very early	Good (0 points)	Single nodule ≤ 21 μ m	35.5 ± 4.45	69.0 ± 35.0
A Early	Good (0-5 points)	Single nodule ≤ 53 μ m, or 3 nodules ≤ 32 μ m	35.5 ± 4.45	69.0 ± 35.0
B Intermediate	Good (5-10 points)	Large, multiple nodules	35.5 ± 4.45	69.0 ± 35.0
C Advanced	Reduced (10-20 points)	Vascular invasion, extrahepatic secondaries	$<35.5 \pm 4.45$	$>69.0 \pm 35.0$
D Terminal	Severely reduced (>20 points)	Any form	$<35.5 \pm 4.45$	$>69.0 \pm 35.0$

*Schnell et al. 2002

At final analysis, all mock-, RGB-, and CyclinA2-transplanted mice showed a good general health status indicated by appearance, behavior, and also ALT and ALB concentrations; but they did not fulfill any other criteria as no tumors developed. General health status of HRas-V12-transplanted mice was good as well with ALB and ALT levels within physiological measures. However, as already mentioned MRI showed one tumor in HRas-V12 mice 21 days after transplantation. However, 35 days after transplantation, more than three nodules, all $\geq 32 \mu\text{m}$ in diameter, became apparent contributing to intermediate stage classification of HRas-V12-induced tumors. The conditions stayed the same until the final analysis. Recipient mice transplanted with LargeT-transduced hepatocytes showed signs of a decreased general health status already three weeks after transplantation with a score of 5 points. In addition, these recipients developed multiple tumors and displayed highly impaired liver architecture measured by MRI 21 days post Tx. Therefore, LargeT recipients were rated as intermediate stage liver cancer already three weeks after transplantation. Over time, their condition worsened in contrast to the other groups: five to nine weeks after transplantation, mice had a reduced general health status rated 10 to 20 points with significantly elevated ALT levels ($p < 0.001$), but no vascular invasion or extrahepatic metastases at the final analysis. Finally, two out of four evaluated factors (tumor characteristics and ALB level) pointed to an intermediate stage rating, and the other two (general health status and ALT level) to an advanced stage of liver disease in the LargeT group.

Comparing HRas-V12- and LargeT-induced outcomes, the evaluated parameters characterize LargeT as a fast and more aggressive oncogene and point towards HRas-V12 as eligible candidate for future investigations of early to intermediate stage liver cancer in the USB mouse model.

4.6 Compilation of evaluation parameters in a radar chart

The results of the experiments conducted in this thesis are summarized in Figure 4.1 highlighting the relevance of the investigated parameters to characterize the transplanted groups. Liver cancer progression and severity of disease were rated using the adapted BCLC score. To reveal the tumor onset weeks after transplantation, MRI analyses were evaluated. Final ALB and ALT levels are displayed next to each other and were measured to investigate general health conditions and ongoing inflammation. Immunofluorescence and immunohistochemical staining data suggest the quantity of engrafted modified hepatocytes, whereas vector copy numbers represent the product of transduction rates and cell expansion of lentivirally-modified hepatocytes.

RGB-mice received hepatocytes transduced with three vectors encoding for a red, green, and blue fluorescent protein using MOIs of 40, respectively. This was necessary to obtain a broad color spectrum and high fluorescent protein expression rates of transduced cells.

Therefore, VCN and engraftment rates measured by IHC and IF staining are higher compared to the other groups. In contrast, a high VCN measured in the CyclinA2, HRas-V12, or LargeT group has a different relevance and higher impact, since the likelihood of a tumor is increased with elevated VCN in these groups. Consequently, mice of the RGB group can not be compared directly to the other groups, and thus are displayed using a dashed line in Figure 4.1.

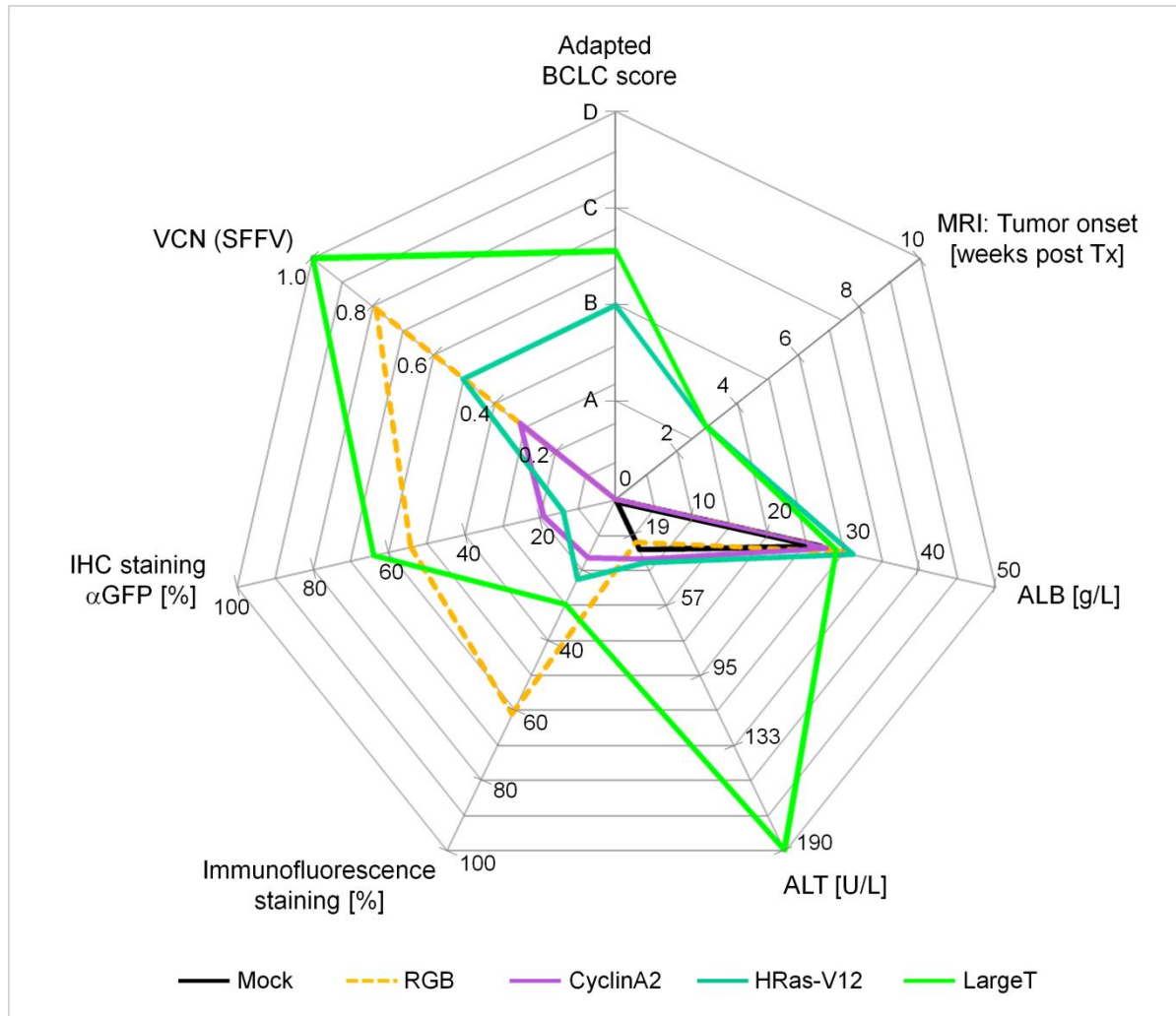


Figure 4.1 Radar chart with summarized study parameters. The adapted BCLC score (0-D), tumor onset found by MRI [weeks post Tx], ALB [g/L], and ALT levels [U/L], immunofluorescence [%] and αGFP-IHC staining [%], and finally relative VCN of SFFV were displayed on a separate axis each. ALB, albumin; ALT, alanine aminotransaminase; BCLC, Barcelona Clinic Liver Cancer staging system; IHC, immunohistochemical staining; SFFV, spleen-focus forming virus promoter; VCN, vector copy number.

Looking at the radar chart, the group-specific kinetics are evident. All transplanted mice had physiological albumin levels, and also normal serum ALT levels, except for elevated ALT in LargeT-mice (Fig. 4.1), which indicates a functional liver metabolism resulting from integrated-donor hepatocytes (Meuleman et al. 2005). Histological and molecular analyses demonstrated uniform group-specific liver repopulation levels and vector copy numbers of the transduced hepatocytes. *SFFV* vector copy numbers of LargeT recipients were more than twice as high as VCN of HRas-V12 or CyclinA2 recipients. MRI and the adapted BCLC

scoring confirmed physiological liver repopulation of mock, RGB, and CyclinA2 recipients. In contrast, HRas-V12-transplanted mice were rated with intermediate stage liver cancer. Mice transplanted with LargeT-modified hepatocytes were scored between stage B and stage C. In HRas-V12 and LargeT recipients, tumor development was identified via MRI within 3 weeks after transplantation (Fig. 4.1). Again, these results highlight the severity of LargeT-induced hepatocarcinogenesis and emphasize the relevance of HRas-V12 as a promising candidate for liver cancer studies.

4.7 Conclusion

The aim of this study was the establishment of a new, reliable HCC mouse model to investigate the benign and malignant clonal regeneration of the liver *in vivo*.

I could verify, that transplantation of primary adult wildtype hepatocytes expressing lentiviral-transduced oncogenes, such as *HRas-V12* and *SV40 Large T-antigen*, induces hepatocarcinogenesis *in vivo* within the hepatotoxic environment of USB mice. Even though both groups, HRas-V12 and LargeT, were classified with intermediate stage liver cancer, using HRas-V12 for induction of cancer would be more meaningful. Some adjustments (e.g., transduction with decreased MOIs and reduced incubation period) could ameliorate survival and decelerate tumor onset contributing to stage A tumor development in homozygous mice. To date, there have been no studies providing a reliable animal model for early stage hepatocarcinogenesis with the characteristics displayed by the HRas-V12-USB model of a healthy phenotype, normal serum parameters, and early tumor onset within 3 weeks after orthotopic transplantation of modified primary hepatocytes expressing only a single oncogene (Brown et al. 2018; Ju et al. 2016). Although breeding of USB mice can be challenging, the model described in this thesis is a realistic setting for tumor-development investigations, despite the immunodeficient background and young age of mice, which have not reached adulthood at time of transplantation. The main characteristic for the pathomechanism of HCC is sustained hepatocyte damage by increased release of inflammatory mediators (Block et al. 2003). In adult livers, hepatocyte proliferation occurs only at rare frequencies under physiological steady-state conditions, whereas enhanced proliferation is triggered during chronic liver inflammation (Ramboer et al. 2014) supporting engraftment of transplanted hepatocytes. Consequently, the properties of the USB mouse model ideally fulfill these conditions (Lütgehetmann et al. 2010; Meuleman et al. 2005).

Furthermore, RGB marking was successfully used to analyze clonal regeneration in homozygous USB mice, broadening its application spectrum as previous analyses were performed with hemizygous mice only (Cornils et al. 2014; Weber et al. 2011). With this tool several possibilities emerge to study clonal liver regeneration by benign but also malignant cell clones. The combination with genetic barcoding (Cornils et al. 2014) would facilitate

clonal cell tracking and integration site analysis. Also transcriptome sequencing would benefit by barcoded tumors to analyze signaling pathways deregulated in liver tumors induced by oncogene expression.

To summarize, this thesis comprises two main findings: 1) lentiviral integration of HRas-V12 and LargeT oncogenes result in HCC formation in a liver damage mouse model with persistent uPA-transgene expression, and 2) transplantation of RGB-marked hepatocytes led to physiological clonal regeneration of USB mouse livers.

In conclusion, the approach presented here is promising to develop a clinically relevant transplantation model inducing oncogene-mediated carcinogenesis after orthotopic transplantation of primary hepatocytes. A reliable HCC model would also enable research on treatment options effecting tumor onset, growth or even clearance.

5 Bibliography

- Ahuja, D., Saenz-Robles, M. T. and Pipas, J. M. 2005.** SV40 large T antigen targets multiple cellular pathways to elicit cellular transformation. *Oncogene* 24(52):7729-7745.
- Bakiri, L. and Wagner, E. F. 2013.** Mouse models for liver cancer. *Mol Oncol* 7(2):206-23.
- Baltimore, D. 1970.** Viral RNA-dependent DNA Polymerase: RNA-dependent DNA Polymerase in Virions of RNA Tumour Viruses. *Nature* 226(5252):1209-1211.
- Barbacid, M. 1987.** ras Genes. *Annual Review of Biochemistry* 56(1):779-827.
- Bargellini, I., Sacco, R., Bozzi, E., Bertini, M., Ginanni, B., Romano, A., Cicorelli, A., Tumino, E., Federici, G., Cioni, R. and others. 2012.** Transarterial chemoembolization in very early and early-stage hepatocellular carcinoma patients excluded from curative treatment: a prospective cohort study. *Eur J Radiol* 81(6):1173-8.
- Bargonetti, J., Reynisdóttir, I., Friedman, P. N. and Prives, C. 1992.** Site-specific binding of wild-type p53 to cellular DNA is inhibited by SV40 T antigen and mutant p53. *Genes and Development* 6:1886-1898.
- Barmania, F. and Pepper, M. S. 2013.** C-C chemokine receptor type five (CCR5): An emerging target for the control of HIV infection. *Appl Transl Genom* 2:3-16.
- Barre-Sinoussi, F., Chermann, J., Rey, F., Nugeyre, M., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C. and others. 1983.** Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 220(4599):868-871.
- Baum, C., Dullmann, J., Li, Z., Fehse, B., Meyer, J., Williams, D. A. and von Kalle, C. 2003.** Side effects of retroviral gene transfer into hematopoietic stem cells. *Blood* 101(6):2099-114.
- Baum, C., Kustikova, O., Modlich, U., Li, Z. and Fehse, B. 2006.** Mutagenesis and oncogenesis by chromosomal insertion of gene transfer vectors. *Hum Gene Ther* 17(3):253-63.
- Baum, C., von Kalle, C., Staal, F. J., Li, Z., Fehse, B., Schmidt, M., Weerkamp, F., Karlsson, S., Wagemaker, G. and Williams, D. A. 2004.** Chance or necessity? Insertional mutagenesis in gene therapy and its consequences. *Mol Ther* 9(1):5-13.
- Beyer, W. R., Westphal, M., Ostertag, W. and von Laer, D. 2002.** Oncoretrovirus and lentivirus vectors pseudotyped with lymphocytic choriomeningitis virus glycoprotein: generation, concentration, and broad host range. *J Virol* 76(3):1488-95.
- Bhandari, R. N., Riccalton, L. A., Lewis, A. L., Fry, J. R., Hammond, A. H., Tendler, S. J. and Shakesheff, K. M. 2001.** Liver tissue engineering: a role for co-culture systems in modifying hepatocyte function and viability. *Tissue Eng* 7(3):345-57.
- Block, T. M., Mehta, A. S., Fimmel, C. J. and Jordan, R. 2003.** Molecular viral oncology of hepatocellular carcinoma. *Oncogene* 22(33):5093-107.
- Blum, H. E. 2005.** Treatment of hepatocellular carcinoma. *Best Pract Res Clin Gastroenterol* 19(1):129-45.
- Blumberg, B. S., Larouze, B., London, W. T., Werner, B., Hesser, J. E., Millman, I., Saimot, G. and Payet, M. 1975.** The relation of infection with the hepatitis B agent to primary hepatic carcinoma. *Am J Pathol* 81(3):669-82.
- Bonilla Guerrero, R. and Roberts, L. R. 2005.** The role of hepatitis B virus integrations in the pathogenesis of human hepatocellular carcinoma. *J Hepatol* 42(5):760-77.

- Bos, J. L. 1989.** ras Oncogenes in Human Cancer: A Review. *Cancer Research* 49(17):4682-4689.
- Boztug, K., Schmidt, M., Schwarzer, A., Banerjee, P. P., Diez, I. A., Dewey, R. A., Bohm, M., Nowrouzi, A., Ball, C. R., Glimm, H. and others. 2010.** Stem-cell gene therapy for the Wiskott-Aldrich syndrome. *N Engl J Med* 363(20):1918-27.
- Braun, C. J., Boztug, K., Paruzynski, A., Witzel, M., Schwarzer, A., Rothe, M., Modlich, U., Beier, R., Gohring, G., Steinemann, D. and others. 2014.** Gene therapy for Wiskott-Aldrich syndrome--long-term efficacy and genotoxicity. *Sci Transl Med* 6(227):227ra33.
- Brechot, C., Thiers, V., Kremsdorf, D., Nalpas, B., Pol, S. and Paterlini-Brechot, P. 2001.** Persistent hepatitis B virus infection in subjects without hepatitis B surface antigen: clinically significant or purely "occult"? *Hepatology* 34(1):194-203.
- Brown, Z. J., Heinrich, B. and Greten, T. F. 2018.** Mouse models of hepatocellular carcinoma: an overview and highlights for immunotherapy research. *Nat Rev Gastroenterol Hepatol*.
- Bruix, J., Sherman, M. and American Association for the Study of Liver, D. 2011.** Management of hepatocellular carcinoma: an update. *Hepatology* 53(3):1020-2.
- Carnero, A. and Paramio, J. M. 2014.** The PTEN/PI3K/AKT Pathway in vivo, *Cancer Mouse Models*. *Front Oncol* 4:252.
- Castelli, G., Pelosi, E. and Testa, U. 2017.** Liver Cancer: Molecular Characterization, Clonal Evolution and Cancer Stem Cells. *Cancers (Basel)* 9(9).
- Cattoglio, C., Pellin, D., Rizzi, E., Maruggi, G., Corti, G., Miselli, F., Sartori, D., Guffanti, A., Di Serio, C., Ambrosi, A. and others. 2010.** High-definition mapping of retroviral integration sites identifies active regulatory elements in human multipotent hematopoietic progenitors. *Blood* 116(25):5507-17.
- Cavazzana-Calvo, M., Hacein-Bey, S., de Saint Basile, G., Gross, F., Yvon, E., Nusbaum, P., Selz, F., Hue, C., Certain, S., Casanova, J. L. and others. 2000.** Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* 288(5466):669-72.
- Cha, C. and Dematteo, R. P. 2005.** Molecular mechanisms in hepatocellular carcinoma development. *Best Pract Res Clin Gastroenterol* 19(1):25-37.
- Chang, Y., Cesarman, E., Pessin, M. S., Lee, F., Culpepper, J., Knowles, D. M. and Moore, P. S. 1994.** Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science* 266(5192):1865-9.
- Chao, Y., Shih, Y.-L., Chiu, J.-H., Chau, G.-Y., Lui, W.-Y., Yang, W. K., Lee, S.-D. and Huang, T.-S. 1998.** Overexpression of Cyclin A but not Skp 2 Correlates with the Tumor Relapse of Human Hepatocellular Carcinoma. *Cancer Research* 58(5):985-990.
- Chen, C. J., Yang, H. I., Su, J., Jen, C. L., You, S. L., Lu, S. N., Huang, G. T., Iloeje, U. H. and Group, R.-H. S. 2006.** Risk of hepatocellular carcinoma across a biological gradient of serum hepatitis B virus DNA level. *JAMA* 295(1):65-73.
- Chen, J. D., Yang, H. I., Iloeje, U. H., You, S. L., Lu, S. N., Wang, L. Y., Su, J., Sun, C. A., Liaw, Y. F., Chen, C. J. and others. 2010.** Carriers of inactive hepatitis B virus are still at risk for hepatocellular carcinoma and liver-related death. *Gastroenterology* 138(5):1747-54.
- Cheng, A. L., Kang, Y. K., Chen, Z., Tsao, C. J., Qin, S., Kim, J. S., Luo, R., Feng, J., Ye, S., Yang, T. S. and others. 2009.** Efficacy and safety of sorafenib in patients in the Asia-Pacific region with advanced hepatocellular carcinoma: a phase III randomised, double-blind, placebo-controlled trial. *Lancet Oncol* 10(1):25-34.

- Cheruvu, S., Marks, K. and Talal, A. H. 2007.** Understanding the pathogenesis and management of hepatitis B/HIV and hepatitis B/hepatitis C virus coinfection. *Clin Liver Dis* 11(4):917-43, ix-x.
- Coffin, J., Haase, A., Levy, J. A., Montagnier, L., Oroszlan, S., Teich, N., Temin, H., Toyoshima, K., Varmus, H., Vogt, P. and others. 1986.** Human immunodeficiency viruses. *Science* 232(4751):697.
- Coffin, J. M., Hughes, S. H. and Varmus, H. E. 1997.** Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY).
- Coiras, M., Lopez-Huertas, M. R., Perez-Olmeda, M. and Alcamí, J. 2009.** Understanding HIV-1 latency provides clues for the eradication of long-term reservoirs. *Nat Rev Microbiol* 7(11):798-812.
- Cornils, K., Thielecke, L., Huser, S., Forgber, M., Thomaschewski, M., Kleist, N., Hussein, K., Riecken, K., Volz, T., Gerdes, S. and others. 2014.** Multiplexing clonality: combining RGB marking and genetic barcoding. *Nucleic Acids Res* 42(7):e56.
- Dandri, M., Burda, M. R., Torok, E., Pollok, J. M., Iwanska, A., Sommer, G., Rogiers, X., Rogler, C. E., Gupta, S., Will, H. and others. 2001.** Repopulation of mouse liver with human hepatocytes and in vivo infection with hepatitis B virus. *Hepatology* 33(4):981-8.
- Dandri, M. and Petersen, J. 2012.** Chimeric mouse model of hepatitis B virus infection. *Journal of Hepatology* 56(2):493-495.
- de Lope, C. R., Tremosini, S., Forner, A., Reig, M. and Bruix, J. 2012.** Management of HCC. *J Hepatol* 56 Suppl 1:S75-87.
- De Ravin, S. S., Su, L., Theobald, N., Choi, U., Macpherson, J. L., Poidinger, M., Symonds, G., Pond, S. M., Ferris, A. L., Hughes, S. H. and others. 2014.** Enhancers are major targets for murine leukemia virus vector integration. *J Virol* 88(8):4504-13.
- Deshpande, A., Sicinski, P. and Hinds, P. W. 2005.** Cyclins and cdks in development and cancer: a perspective. *Oncogene* 24(17):2909-15.
- Dhanasekaran, R., Bando, S. and Roberts, L. R. 2016.** Molecular pathogenesis of hepatocellular carcinoma and impact of therapeutic advances. *F1000Res* 5.
- Donsante, A., Miller, D. G., Li, Y., Vogler, C., Brunt, E. M., Russell, D. W. and Sands, M. S. 2007.** AAV vector integration sites in mouse hepatocellular carcinoma. *Science* 317(5837):477.
- Dubois, N., Bennoun, M., Allemand, I., Molina, T., Grimber, G., Daudet-Monsac, M., Abelanet, R. and Briand, P. 1991.** Time-course development of differentiated hepatocarcinoma and lung metastasis in transgenic mice. *J Hepatol* 13(2):227-39.
- Dull, T., Zufferey, R., Kelly, M., Mandel, R. J., Nguyen, M., Trono, D. and Naldini, L. 1998.** A third-generation lentivirus vector with a conditional packaging system. *J Virol* 72(11):8463-71.
- El-Serag, H. B. and Rudolph, K. L. 2007.** Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. *Gastroenterology* 132(7):2557-76.
- Epstein, C. J. 1967.** Cell size, nuclear content, and the development of polyploidy in the mammalian liver. *Proceedings of the National Academy of Sciences of the United States of America* 57(2):327-334.
- Epstein, M. A., Achong, B. G. and Barr, Y. M. 1964.** Virus Particles in Cultured Lymphoblasts from Burkitt's Lymphoma. *The Lancet* 283(7335):702-703.
- Escors, D. and Breckpot, K. 2010.** Lentiviral vectors in gene therapy: their current status and future potential. *Arch Immunol Ther Exp (Warsz)* 58(2):107-19.

- Fanales-Belasio, E., Raimondo, M., Suligoi, B. and Butto, S. 2010.** HIV virology and pathogenetic mechanisms of infection: a brief overview. *Ann Ist Super Sanita* 46(1):5-14.
- Fehse, B., Chukhlovina, A., Kuhlcke, K., Marinetz, O., Vorwig, O., Renges, H., Kruger, W., Zabelina, T., Dudina, O., Finckenstein, F. G. and others. 2001.** Real-time quantitative Y chromosome-specific PCR (QYCS-PCR) for monitoring hematopoietic chimerism after sex-mismatched allogeneic stem cell transplantation. *J Hematother Stem Cell Res* 10(3):419-25.
- Fehse, B., Kustikova, O. S., Bubenheim, M. and Baum, C. 2004.** Poisson - it's a question of dose. *Gene Ther* 11(11):879-81.
- Fehse, B. and Roeder, I. 2007.** Insertional mutagenesis and clonal dominance: biological and statistical considerations. *Gene Ther* 15(2):143-53.
- Felice, B., Cattoglio, C., Cittaro, D., Testa, A., Miccio, A., Ferrari, G., Luzi, L., Recchia, A. and Mavilio, F. 2009.** Transcription factor binding sites are genetic determinants of retroviral integration in the human genome. *PLoS One* 4(2):e4571.
- Florence, B. and Faller, D. V. 2001.** You bet-cha: a novel family of transcriptional regulators. *Front Biosci* 6:D1008-18.
- Follenzi, A., Ailles, L. E., Bakovic, S., Geuna, M. and Naldini, L. 2000.** Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. *Nat Genet* 25(2):217-22.
- Forner, A. and Bruix, J. 2012.** Biomarkers for early diagnosis of hepatocellular carcinoma. *Lancet Oncol* 13(8):750-1.
- Forner, A., Llovet, J. M. and Bruix, J. 2012.** Hepatocellular carcinoma. *The Lancet* 379(9822):1245-1255.
- Forner, A., Reig, M. E., de Lope, C. R. and Bruix, J. 2010.** Current strategy for staging and treatment: the BCLC update and future prospects. *Semin Liver Dis* 30(1):61-74.
- Fox, I. J., Chowdhury, J. R., Kaufman, S. S., Goertzen, T. C., Chowdhury, N. R., Warkentin, P. I., Dorko, K., Sauter, B. V. and Strom, S. C. 1998.** Treatment of the Crigler-Najjar syndrome type I with hepatocyte transplantation. *N Engl J Med* 338(20):1422-6.
- Freed, E. O. 2001.** HIV-1 replication. *Somat Cell Mol Genet* 26(1-6):13-33.
- Fulks, M., Stout, R. L. and Dolan, V. F. 2010.** Albumin and all-cause mortality risk in insurance applicants. *J Insur Med* 42(1):11-7.
- Gabriel, R., Schmidt, M. and von Kalle, C. 2012.** Integration of retroviral vectors. *Curr Opin Immunol* 24(5):592-7.
- Gentric, G. and Desdouets, C. 2014.** Polyploidization in Liver Tissue. *The American Journal of Pathology* 184(2):322-331.
- Ghouri, Y. A., Mian, I. and Rowe, J. H. 2017.** Review of hepatocellular carcinoma: Epidemiology, etiology, and carcinogenesis. *J Carcinog* 16:1.
- Gilgenkrantz, H. 2010.** Rodent models of liver repopulation. *Methods Mol Biol* 640:475-90.
- Giri, S. and Bader, A. 2014.** Immortalization of Human Fetal Hepatocyte by Ectopic Expression of Human Telomerase Reverse Transcriptase, Human Papilloma Virus (E7) and Simian Virus 40 Large T (SV40 T) Antigen Towards Bioartificial Liver Support. *J Clin Exp Hepatol* 4(3):191-201.
- Gomez-Nicola, D., Riecken, K., Fehse, B. and Perry, V. H. 2014.** In-vivo RGB marking and multicolour single-cell tracking in the adult brain. *Sci Rep* 4:7520.
- Gupta, S. S., Maetzig, T., Maertens, G. N., Sharif, A., Rothe, M., Weidner-Glunde, M., Galla, M., Schambach, A., Cherepanov, P. and Schulz, T. F. 2013.** Bromo- and

- extraterminal domain chromatin regulators serve as cofactors for murine leukemia virus integration. *J Virol* 87(23):12721-36.
- Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M. P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C. S., Pawliuk, R., Morillon, E. and others. 2003.** LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* 302(5644):415-9.
- Hamaguchi, I., Woods, N. B., Panagopoulos, I., Andersson, E., Mikkola, H., Fahlman, C., Zufferey, R., Carlsson, L., Trono, D. and Karlsson, S. 2000.** Lentivirus vector gene expression during ES cell-derived hematopoietic development in vitro. *J Virol* 74(22):10778-84.
- Haridass, D., Yuan, Q., Becker, P. D., Cantz, T., Iken, M., Rothe, M., Narain, N., Bock, M., Norder, M., Legrand, N. and others. 2009.** Repopulation efficiencies of adult hepatocytes, fetal liver progenitor cells, and embryonic stem cell-derived hepatic cells in albumin-promoter-enhancer urokinase-type plasminogen activator mice. *Am J Pathol* 175(4):1483-92.
- Heckel, J. L., Sandgren, E. P., Degen, J. L., Palmiter, R. D. and Brinster, R. L. 1990.** Neonatal bleeding in transgenic mice expressing urokinase-type plasminogen activator. *Cell* 62(3):447-56.
- Henle, G. and Henle, W. 1966.** Immunofluorescence in cells derived from Burkitt's lymphoma. *J Bacteriol* 91(3):1248-1256.
- Holczbauer, A., Factor, V. M., Andersen, J. B., Marquardt, J. U., Kleiner, D. E., Raggi, C., Kitade, M., Seo, D., Akita, H., Durkin, M. E. and others. 2013.** Modeling pathogenesis of primary liver cancer in lineage-specific mouse cell types. *Gastroenterology* 145(1):221-31.
- Hsu, H. C., Tseng, H. J., Lai, P. L., Lee, P. H. and Peng, S. Y. 1993.** Expression of p53 gene in 184 unifocal hepatocellular carcinomas: association with tumor growth and invasiveness. *Cancer Res* 53(19):4691-4.
- Imamura, H., Matsuyama, Y., Tanaka, E., Ohkubo, T., Hasegawa, K., Miyagawa, S., Sugawara, Y., Minagawa, M., Takayama, T., Kawasaki, S. and others. 2003.** Risk factors contributing to early and late phase intrahepatic recurrence of hepatocellular carcinoma after hepatectomy. *J Hepatol* 38(2):200-7.
- Javier, R. T. and Butel, J. S. 2008.** The history of tumor virology. *Cancer Res* 68(19):7693-706.
- Joseph, B., Kumaran, V., Berishvili, E., Bhargava, K. K., Palestro, C. J. and Gupta, S. 2006.** Monocrotaline promotes transplanted cell engraftment and advances liver repopulation in rats via liver conditioning. *Hepatology* 44(6):1411-20.
- Ju, H. L., Han, K. H., Lee, J. D. and Ro, S. W. 2016.** Transgenic mouse models generated by hydrodynamic transfection for genetic studies of liver cancer and preclinical testing of anti-cancer therapy. *Int J Cancer* 138(7):1601-8.
- Kaposi, M. 1872.** Idiopathisches multiples Pigmentsarkom der Haut. *Archiv für Dermatologie und Syphilis* 4(2):265-273.
- Kim, R. D., Reed, A. I., Fujita, S., Foley, D. P., Mekeel, K. L. and Hemming, A. W. 2007.** Consensus and controversy in the management of hepatocellular carcinoma. *J Am Coll Surg* 205(1):108-23.
- Kinoshita, A., Onoda, H., Fushiya, N., Koike, K., Nishino, H. and Tajiri, H. 2015.** Staging systems for hepatocellular carcinoma: Current status and future perspectives. *World J Hepatol* 7(3):406-24.

- Knight, S., Bokhoven, M., Collins, M. and Takeuchi, Y. 2010.** Effect of the internal promoter on insertional gene activation by lentiviral vectors with an intact HIV long terminal repeat. *J Virol* 84(9):4856-9.
- Knudson, A. G., Jr. 1971.** Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A* 68(4):820-3.
- Kostic, C., Chiodini, F., Salmon, P., Wiznerowicz, M., Deglon, N., Hornfeld, D., Trono, D., Aebischer, P., Schorderet, D. F., Munier, F. L. and others. 2003.** Activity analysis of housekeeping promoters using self-inactivating lentiviral vector delivery into the mouse retina. *Gene Ther* 10(9):818-21.
- Kusano, M. and Mito, M. 1982.** Observations on the fine structure of long-survived isolated hepatocytes inoculated into rat spleen. *Gastroenterology* 82(4):616-28.
- Kustikova, O. S., Schiedlmeier, B., Brugman, M. H., Stahlhut, M., Bartels, S., Li, Z. and Baum, C. 2009.** Cell-intrinsic and vector-related properties cooperate to determine the incidence and consequences of insertional mutagenesis. *Mol Ther* 17(9):1537-47.
- Kuzembayeva, M., Dilley, K., Sardo, L. and Hu, W. S. 2014.** Life of psi: how full-length HIV-1 RNAs become packaged genomes in the viral particles. *Virology* 454-455:362-70.
- Lane, D. P. and Crawford, L. V. 1979.** T antigen is bound to a host protein in SV40-transformed cells. *Nature* 278(5701):261-263.
- Lee, J. S. and Thorgeirsson, S. S. 2006.** Comparative and integrative functional genomics of HCC. *Oncogene* 25(27):3801-9.
- Levitt, D. G. and Levitt, M. D. 2016.** Human serum albumin homeostasis: a new look at the roles of synthesis, catabolism, renal and gastrointestinal excretion, and the clinical value of serum albumin measurements. *Int J Gen Med* 9:229-55.
- Li, W. C., Ralphs, K. L. and Tosh, D. 2010.** Isolation and culture of adult mouse hepatocytes. *Methods Mol Biol* 633:185-96.
- Linzer, D. I. and Levine, A. J. 1979.** Characterization of a 54K dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. *Cell* 17(1):43-52.
- Liu, L., Cao, Y., Chen, C., Zhang, X., McNabola, A., Wilkie, D., Wilhelm, S., Lynch, M. and Carter, C. 2006.** Sorafenib blocks the RAF/MEK/ERK pathway, inhibits tumor angiogenesis, and induces tumor cell apoptosis in hepatocellular carcinoma model PLC/PRF/5. *Cancer Res* 66(24):11851-8.
- Liu, L., Liao, J. Z., He, X. X. and Li, P. Y. 2017.** The role of autophagy in hepatocellular carcinoma: friend or foe. *Oncotarget* 8(34):57707-57722.
- Liu, Z., Que, S., Xu, J. and Peng, T. 2014.** Alanine aminotransferase-old biomarker and new concept: a review. *Int J Med Sci* 11(9):925-35.
- Llovet, J. M., Bru, C. and Bruix, J. 1999.** Prognosis of hepatocellular carcinoma: the BCLC staging classification. *Semin Liver Dis* 19(3):329-38.
- Llovet, J. M. and Bruix, J. 2008.** Novel advancements in the management of hepatocellular carcinoma in 2008. *J Hepatol* 48 Suppl 1:S20-37.
- Llovet, J. M., Burroughs, A. and Bruix, J. 2003.** Hepatocellular carcinoma. *The Lancet* 362(9399):1907-1917.
- Llovet, J. M., Real, M. I., Montana, X., Planas, R., Coll, S., Aponte, J., Ayuso, C., Sala, M., Muchart, J., Sola, R. and others. 2002.** Arterial embolisation or chemoembolisation versus symptomatic treatment in patients with unresectable hepatocellular carcinoma: a randomised controlled trial. *Lancet* 359(9319):1734-9.
- Llovet, J. M., Ricci, S., Mazzaferro, V., Hilgard, P., Gane, E., Blanc, J. F., de Oliveira, A. C., Santoro, A., Raoul, J. L., Forner, A. and others. 2008.** Sorafenib in advanced hepatocellular carcinoma. *N Engl J Med* 359(4):378-90.

- Lou, D. Q., Molina, T., Bennoun, M., Porteu, A., Briand, P., Joulin, V., Vasseur-Cognet, M. and Cavard, C. 2005. Conditional hepatocarcinogenesis in mice expressing SV 40 early sequences. *Cancer Lett* 229(1):107-14.
- Lütgehetmann, M., Mancke, L. V., Volz, T., Helbig, M., Allweiss, L., Bornscheuer, T., Pollok, J. M., Lohse, A. W., Petersen, J., Urban, S. and others. 2012. Humanized chimeric uPA mouse model for the study of hepatitis B and D virus interactions and preclinical drug evaluation. *Hepatology* 55(3):685-94.
- Lütgehetmann, M., Volz, T., Kopke, A., Broja, T., Tigges, E., Lohse, A. W., Fuchs, E., Murray, J. M., Petersen, J. and Dandri, M. 2010. In vivo proliferation of hepadnavirus-infected hepatocytes induces loss of covalently closed circular DNA in mice. *Hepatology* 52(1):16-24.
- Marquardt, J. U., Andersen, J. B. and Thorgeirsson, S. S. 2015. Functional and genetic deconstruction of the cellular origin in liver cancer. *Nat Rev Cancer* 15(11):653-67.
- McKillop, I. H., Moran, D. M., Jin, X. and Koniaris, L. G. 2006. Molecular pathogenesis of hepatocellular carcinoma. *J Surg Res* 136(1):125-35.
- Meira, L. B., Bugni, J. M., Green, S. L., Lee, C. W., Pang, B., Borenshtein, D., Rickman, B. H., Rogers, A. B., Moroski-Erkul, C. A., McFaline, J. L. and others. 2008. DNA damage induced by chronic inflammation contributes to colon carcinogenesis in mice. *J Clin Invest* 118(7):2516-25.
- Meng, F. Y., Liu, L., Yang, F. H., Li, C. Y., Liu, J. and Zhou, P. 2014. Reversible immortalization of human hepatocytes mediated by retroviral transfer and site-specific recombination. *World J Gastroenterol* 20(36):13119-26.
- Mercer, D. F., Schiller, D. E., Elliott, J. F., Douglas, D. N., Hao, C., Rinfret, A., Addison, W. R., Fischer, K. P., Churchill, T. A., Lakey, J. R. and others. 2001. Hepatitis C virus replication in mice with chimeric human livers. *Nat Med* 7(8):927-33.
- Meuleman, P. and Leroux-Roels, G. 2008. The human liver-uPA-SCID mouse: a model for the evaluation of antiviral compounds against HBV and HCV. *Antiviral Res* 80(3):231-8.
- Meuleman, P., Libbrecht, L., De Vos, R., de Hemptinne, B., Gevaert, K., Vandekerckhove, J., Roskams, T. and Leroux-Roels, G. 2005. Morphological and biochemical characterization of a human liver in a uPA-SCID mouse chimera. *Hepatology* 41(4):847-56.
- Mizushima, N., Yoshimori, T. and Ohsumi, Y. 2011. The role of Atg proteins in autophagosome formation. *Annu Rev Cell Dev Biol* 27:107-32.
- Modlich, U., Bohne, J., Schmidt, M., von Kalle, C., Knoss, S., Schambach, A. and Baum, C. 2006. Cell-culture assays reveal the importance of retroviral vector design for insertional genotoxicity. *Blood* 108(8):2545-53.
- Modlich, U., Kustikova, O. S., Schmidt, M., Rudolph, C., Meyer, J., Li, Z., Kamino, K., von Neuhoff, N., Schlegelberger, B., Kuehlcke, K. and others. 2005. Leukemias following retroviral transfer of multidrug resistance 1 (MDR1) are driven by combinatorial insertional mutagenesis. *Blood* 105(11):4235-46.
- Modlich, U., Navarro, S., Zychlinski, D., Maetzig, T., Knoess, S., Brugman, M. H., Schambach, A., Charrier, S., Galy, A., Thrasher, A. J. and others. 2009. Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors. *Mol Ther* 17(11):1919-28.
- Mohme, M., Maire, C. L., Riecken, K., Zapf, S., Aranyosy, T., Westphal, M., Lamszus, K. and Fehse, B. 2017. Optical Barcoding for Single-Clone Tracking to Study Tumor Heterogeneity. *Mol Ther* 25(3):621-633.

- Molina, D. K. and DiMaio, V. J. 2012.** Normal organ weights in men: part II-the brain, lungs, liver, spleen, and kidneys. *Am J Forensic Med Pathol* 33(4):368-72.
- Naldini, L. 2015.** Gene therapy returns to centre stage. *Nature* 526(7573):351-60.
- Nguyen, T. H., Oberholzer, J., Birraux, J., Majno, P., Morel, P. and Trono, D. 2002.** Highly efficient lentiviral vector-mediated transduction of nondividing, fully reimplantable primary hepatocytes. *Mol Ther* 6(2):199-209.
- Nicholson, J. P., Wolmarans, M. R. and Park, G. R. 2000.** The role of albumin in critical illness. *Br J Anaesth* 85(4):599-610.
- Nordling, C. O. 1953.** A New Theory on the Cancer-inducing Mechanism. *British Journal of Cancer* 7(1):68-72.
- O'Keefe, E. P. 2013.** Nucleic Acid Delivery: Lentiviral and Retroviral Vectors. *Materials and Methods* 3:174.
- Overturf, K., Al-Dhalimy, M., Tanguay, R., Brantly, M., Ou, C. N., Finegold, M. and Grompe, M. 1996.** Hepatocytes corrected by gene therapy are selected in vivo in a murine model of hereditary tyrosinaemia type I. *Nat Genet* 12(3):266-73.
- Pagano, M., Pepperkok, R., Verde, F., Ansorge, W. and Draetta, G. 1992.** Cyclin A is required at two points in the human cell cycle. *The EMBO Journal* 11(3):961-971.
- Pan, X., Wang, Y., Yu, X., Li, J., Zhou, N., Du, W., Zhang, Y., Cao, H., Zhu, D., Chen, Y. and others. 2015.** Establishment and characterization of an immortalized human hepatic stellate cell line for applications in co-culturing with immortalized human hepatocytes. *Int J Med Sci* 12(3):248-55.
- Pellicoro, A., Ramachandran, P., Iredale, J. P. and Fallowfield, J. A. 2014.** Liver fibrosis and repair: immune regulation of wound healing in a solid organ. *Nat Rev Immunol* 14(3):181-94.
- Petersen, J., Dandri, M., Gupta, S. and Rogler, C. E. 1998.** Liver repopulation with xenogenic hepatocytes in B and T cell-deficient mice leads to chronic hepatitis B virus infection and clonal growth of hepatocellular carcinoma. *Proc Natl Acad Sci U S A* 95(1):310-5.
- Petropoulos, C. 1997.** Retroviral Taxonomy, Protein Structures, Sequences, and Genetic Maps. *in* J. M. Coffin, S. H. Hughes, H. E. Varmus, eds. *Retroviruses*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY).
- Poiesz, B. J., Ruscetti, F. W., Gazdar, A. F., Bunn, P. A., Minna, J. D. and Gallo, R. C. 1980.** Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc Natl Acad Sci U S A* 77(12):7415-9.
- Ramboer, E., De Craene, B., De Kock, J., Vanhaecke, T., Berx, G., Rogiers, V. and Vinken, M. 2014.** Strategies for immortalization of primary hepatocytes. *J Hepatol* 61(4):925-43.
- Ranzani, M., Cesana, D., Bartholomae, C. C., Sanvito, F., Pala, M., Benedicenti, F., Gallina, P., Sergi, L. S., Merella, S., Bulfone, A. and others. 2013.** Lentiviral vector-based insertional mutagenesis identifies genes associated with liver cancer. *Nat Methods* 10(2):155-61.
- Rhim, J. A., Sandgren, E. P., Palmiter, R. D. and Brinster, R. L. 1995.** Complete reconstitution of mouse liver with xenogeneic hepatocytes. *Proceedings of the National Academy of Sciences* 92(11):4942-4946.
- Rittelmeyer, I., Rothe, M., Brugman, M. H., Iken, M., Schambach, A., Manns, M. P., Baum, C., Modlich, U. and Ott, M. 2013.** Hepatic lentiviral gene transfer is associated with clonal selection, but not with tumor formation in serially transplanted rodents. *Hepatology* 58(1):397-408.

- Rothe, M., Rittelmeyer, I., Iken, M., Rudrich, U., Schambach, A., Glage, S., Manns, M. P., Baum, C., Bock, M., Ott, M. and others. 2012.** Epidermal growth factor improves lentivirus vector gene transfer into primary mouse hepatocytes. *Gene Therapy* 19(4):425-434.
- Rous, P. 1910.** A Transmissible Avian Neoplasm. (Sarcoma of the Common Fowl.). *J Exp Med* 12(5):696-705.
- Rous, P. 1911.** A Sarcoma of the Fowl Transmissible by an Agent Separable from the Tumor Cells. *J Exp Med* 13(4):397-411.
- Sakuma, T., Barry, M. A. and Ikeda, Y. 2012.** Lentiviral vectors: basic to translational. *Biochem J* 443(3):603-18.
- Sandgren, E. P., Palmiter, R. D., Heckel, J. L., Daugherty, C. C., Brinster, R. L. and Degen, J. L. 1991.** Complete hepatic regeneration after somatic deletion of an albumin-plasminogen activator transgene. *Cell* 66(2):245-56.
- Sandgren, E. P., Quaife, C. J., Pinkert, C. A., Palmiter, R. D. and Brinster, R. L. 1989.** Oncogene-induced liver neoplasia in transgenic mice. *Oncogene* 4(6):715-24.
- Schambach, A., Bohne, J., Baum, C., Hermann, F. G., Egerer, L., von Laer, D. and Giroglou, T. 2006a.** Woodchuck hepatitis virus post-transcriptional regulatory element deleted from X protein and promoter sequences enhances retroviral vector titer and expression. *Gene Ther* 13(7):641-5.
- Schambach, A., Bohne, J., Chandra, S., Will, E., Margison, G. P., Williams, D. A. and Baum, C. 2006b.** Equal potency of gammaretroviral and lentiviral SIN vectors for expression of O6-methylguanine-DNA methyltransferase in hematopoietic cells. *Mol Ther* 13(2):391-400.
- Schnell, M. A., Hardy, C., Hawley, M., Propert, K. J. and Wilson, J. M. 2002.** Effect of blood collection technique in mice on clinical pathology parameters. *Hum Gene Ther* 13(1):155-61.
- Scholten, D., Trebicka, J., Liedtke, C. and Weiskirchen, R. 2015.** The carbon tetrachloride model in mice. *Lab Anim* 49(1 Suppl):4-11.
- Schröder, A. R., Shinn, P., Chen, H., Berry, C., Ecker, J. R. and Bushman, F. 2002.** HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* 110(4):521-9.
- Selcuk, H. 2017.** Prognostic Factors and Staging Systems in Hepatocellular Carcinoma. *Exp Clin Transplant* 15(Suppl 2):45-49.
- Severgnini, M., Sherman, J., Sehgal, A., Jayaprakash, N. K., Aubin, J., Wang, G., Zhang, L., Peng, C. G., Yucius, K., Butler, J. and others. 2012.** A rapid two-step method for isolation of functional primary mouse hepatocytes: cell characterization and asialoglycoprotein receptor based assay development. *Cytotechnology* 64(2):187-95.
- Shah, M., Patel, K. and Sehgal, P. B. 2005.** Monocrotaline pyrrole-induced endothelial cell megalocytosis involves a Golgi blockade mechanism. *Am J Physiol Cell Physiol* 288(4):C850-62.
- Sherman, K. E., Rockstroh, J. and Thomas, D. 2015.** Human immunodeficiency virus and liver disease: An update. *Hepatology* 62(6):1871-82.
- Sherman, M. 2010.** Hepatocellular carcinoma: epidemiology, surveillance, and diagnosis. *Semin Liver Dis* 30(1):3-16.
- Shope, R. E. and Hurst, E. W. 1933.** Infectious Papillomatosis of Rabbits : With a Note on the Histopathology. *J Exp Med* 58(5):607-24.
- Sia, D., Villanueva, A., Friedman, S. L. and Llovet, J. M. 2017.** Liver Cancer Cell of Origin, Molecular Class, and Effects on Patient Prognosis. *Gastroenterology* 152(4):745-761.

- Singh, S., Singh, P. P., Roberts, L. R. and Sanchez, W. 2014.** Chemopreventive strategies in hepatocellular carcinoma. *Nat Rev Gastroenterol Hepatol* 11(1):45-54.
- Stahl, T., Bohme, M. U., Kroger, N. and Fehse, B. 2015.** Digital PCR to assess hematopoietic chimerism after allogeneic stem cell transplantation. *Exp Hematol* 43(6):462-8 e1.
- Suerth, J. D., Maetzig, T., Brugman, M. H., Heinz, N., Appelt, J. U., Kaufmann, K. B., Schmidt, M., Grez, M., Modlich, U., Baum, C. and others. 2012.** Alpharetroviral self-inactivating vectors: long-term transgene expression in murine hematopoietic cells and low genotoxicity. *Mol Ther* 20(5):1022-32.
- Talamini, M. A., Kappus, B. and Hubbard, A. 1997.** Repolarization of hepatocytes in culture. *Hepatology* 25(1):167-172.
- Taverna, M., Marie, A. L., Mira, J. P. and Guidet, B. 2013.** Specific antioxidant properties of human serum albumin. *Ann Intensive Care* 3(1):4.
- Temin, H. M. 1990.** Safety considerations in somatic gene therapy of human disease with retrovirus vectors. *Hum Gene Ther* 1(2):111-23.
- Temin, H. M. and Mizutani, S. 1970.** Viral RNA-dependent DNA Polymerase: RNA-dependent DNA Polymerase in Virions of Rous Sarcoma Virus. *Nature* 226(5252):1211-1213.
- Thomaschewski, M., Riecken, K., Unrau, L., Volz, T., Cornils, K., Ittrich, H., Heim, D., Wege, H., Akgün, E., Lütgehetmann, M. and others. 2017.** Multi-color RGB marking enables clonality assessment of liver tumors in a murine xenograft model. *Oncotarget* 8 [2017/12/].
- Thorgeirsson, S. S., Lee, J. S. and Grisham, J. W. 2006.** Functional genomics of hepatocellular carcinoma. *Hepatology* 43(2 Suppl 1):S145-50.
- Todd, R. and Munger, K. 2001.** *Oncogenes*. eLS. John Wiley & Sons, Ltd.
- Toss, A. and Cristofanilli, M. 2015.** Molecular characterization and targeted therapeutic approaches in breast cancer. 60 pp.
- Uren, A. G., Kool, J., Berns, A. and van Lohuizen, M. 2005.** Retroviral insertional mutagenesis: past, present and future. *Oncogene* 24(52):7656-72.
- Vranckx, L. S., Demeulemeester, J., Debyser, Z. and Gijssbers, R. 2016.** Towards a Safer, More Randomized Lentiviral Vector Integration Profile Exploring Artificial LEDGF Chimeras. *PLoS One* 11(10):e0164167.
- Wallace, K., Fairhall, E. A., Charlton, K. A. and Wright, M. C. 2010.** AR42J-B-13 cell: an expandable progenitor to generate an unlimited supply of functional hepatocytes. *Toxicology* 278(3):277-87.
- Wang, H., Naghavi, M., Allen, C., Barber, R. M., Bhutta, Z. A., Carter, A., Casey, D. C., Charlson, F. J., Chen, A. Z., Coates, M. M. and others. 2016.** Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet* 388(10053):1459-1544.
- Wang, H. C., Chang, W. T., Chang, W. W., Wu, H. C., Huang, W., Lei, H. Y., Lai, M. D., Fausto, N. and Su, I. J. 2005.** Hepatitis B virus pre-S2 mutant upregulates cyclin A expression and induces nodular proliferation of hepatocytes. *Hepatology* 41(4):761-70.
- Wang, J. and Yang, J. 2010.** Interaction of Tumor Suppressor p53 with DNA and Proteins. *Current Pharmaceutical Biotechnology* 11(1):122-127.
- Wang, Y. and Prives, C. 1995.** Increased and altered DNA binding of human p53 by S and G2/M but not G1 cyclin-dependent kinases. *Nature* 376(6535):88-91.

- Weber, A., Groyer-Picard, M. T., Franco, D. and Dagher, I. 2009.** Hepatocyte transplantation in animal models. *Liver Transpl* 15(1):7-14.
- Weber, K., Bartsch, U., Stocking, C. and Fehse, B. 2008.** A multicolor panel of novel lentiviral "gene ontology" (LeGO) vectors for functional gene analysis. *Mol Ther* 16(4):698-706.
- Weber, K., Mock, U., Petrowitz, B., Bartsch, U. and Fehse, B. 2010.** Lentiviral gene ontology (LeGO) vectors equipped with novel drug-selectable fluorescent proteins: new building blocks for cell marking and multi-gene analysis. *Gene Ther* 17(4):511-20.
- Weber, K., Thomaschewski, M., Benten, D. and Fehse, B. 2012.** RGB marking with lentiviral vectors for multicolor clonal cell tracking. *Nat Protoc* 7(5):839-49.
- Weber, K., Thomaschewski, M., Warlich, M., Volz, T., Cornils, K., Niebuhr, B., Tager, M., Lutgehetmann, M., Pollok, J. M., Stocking, C. and others. 2011.** RGB marking facilitates multicolor clonal cell tracking. *Nat Med* 17(4):504-9.
- Weglarz, T. C., Degen, J. L. and Sandgren, E. P. 2000.** Hepatocyte Transplantation into Diseased Mouse Liver : Kinetics of Parenchymal Repopulation and Identification of the Proliferative Capacity of Tetraploid and Octaploid Hepatocytes. *The American Journal of Pathology* 157(6):1963-1974.
- Weiss, R. A. and Vogt, P. K. 2011.** 100 years of Rous sarcoma virus. *J Exp Med* 208(12):2351-5.
- Woods, N. B., Muessig, A., Schmidt, M., Flygare, J., Olsson, K., Salmon, P., Trono, D., von Kalle, C. and Karlsson, S. 2003.** Lentiviral vector transduction of NOD/SCID repopulating cells results in multiple vector integrations per transduced cell: risk of insertional mutagenesis. *Blood* 101(4):1284-9.
- Wu, C. and Dunbar, C. E. 2011.** Stem cell gene therapy: the risks of insertional mutagenesis and approaches to minimize genotoxicity. *Front Med* 5(4):356-71.
- Xin, B., Yamamoto, M., Fujii, K., Ooshio, T., Chen, X., Okada, Y., Watanabe, K., Miyokawa, N., Furukawa, H. and Nishikawa, Y. 2017.** Critical role of Myc activation in mouse hepatocarcinogenesis induced by the activation of AKT and RAS pathways. *Oncogene* 36(36):5087-5097.
- Yam, C. H., Fung, T. K. and Poon, R. Y. 2002.** Cyclin A in cell cycle control and cancer. *Cell Mol Life Sci* 59(8):1317-26.
- Zychlinski, D., Schambach, A., Modlich, U., Maetzig, T., Meyer, J., Grassman, E., Mishra, A. and Baum, C. 2008.** Physiological promoters reduce the genotoxic risk of integrating gene vectors. *Mol Ther* 16(4):718-25.

6 Online references

- 1 World Health Organization, International Agency for Research on Cancer (IARC). 2017. http://gco.iarc.fr/today/online-analysis-map?mode=population&mode_population=continents&population=900&sex=1&cancer=7&type=1&statistic=0&prevalence=0&color_palette=default&projection=natural-earth (29.09.2017, 10:01 p.m.)
- 2 The Nobel Prize in Physiology or Medicine 1975. *Nobelprize.org*. Nobel Media AB 2014. http://www.nobelprize.org/nobel_prizes/medicine/laureates/1975/ (19.07.2018, 1:34 p.m.)
- 3 The Nobel Prize in Physiology or Medicine 1989. *Nobelprize.org*. Nobel Media AB 2014. http://www.nobelprize.org/nobel_prizes/medicine/laureates/1989/ (19.07.2018, 1:36 p.m.)
- 4 The Nobel Prize in Physiology or Medicine 2008. *Nobelprize.org*. Nobel Media AB 2014. http://www.nobelprize.org/nobel_prizes/medicine/laureates/2008/ (19.07.2018, 1:38 p.m.)
- 5 The Jackson Laboratory. <https://www.jax.org/jax-mice-and-services/strain-data-sheet-pages/body-weight-chart-000664#> (25.06.2018, 00:39 a.m.)

Supplementary

Table S1 Primer used for cloning and sequencing of LeGO-LargeT-iV2. LeGO, lentiviral gene ontology; i, IRES, internal ribosomal entry site; V2, fluorescent protein Venus; T_m, melting temperature.

Target	Oligo	Oligoname	Sequence 5' → 3'	T _m [°C]
<i>SFFV</i>	Forward primer	p007	GAGCTCACAACCCCTCACTC	61.8
<i>LargeT (+ EcoRI)</i>	Forward primer	p248	TATAGAATTCGCCACCATGGACAAAGTTTAAACAGAGAGGAATCTTGCAGCT	71.5
<i>LargeT (+ NotI)</i>	Reverse primer	p249	TATAGCGGCCGCTTATGTTTCAGGTCAGGGGGAGG	73.9
<i>LargeT</i>	Forward primer	p250	AGTGCCTTGACTAGAGATCCATT	59.9
<i>LargeT</i>	Forward primer	p251	TGAGGATGTAAAGGGCACTGGA	62.1

Table S2 Primer/probe sets for amplification of gDNA by ddPCR for quantification of relative vector copy numbers (VCN) of integrated lentiviral vectors.

Target	Oligo	Oligoname	Sequence 5' → 3'	Modification
<i>Epo-R</i>	Forward primer	mEpo-f	GCAGGCGGGGTCGCTACTC	-
	Reverse primer	mEpo-r	CGCCTGTGCAGATCCGATAA	-
	Probe	mEpo-p FAM	TTCTGAGGCGCCACTTTTGCAAGACC	5' FAM 3' BHQ1
	Probe	mEpo-p HEX	TTCTGAGGCGCCACTTTTGCAAGACC	5' HEX 3' BHQ1
<i>GFP/Venus</i>	Forward primer	Lu52-ddGFP-F	CTGCTGCCCCGACAACCA	-
<i>GFP/Venus</i>	Reverse primer	Lu53-ddGFP-R	TGTGATCGCGCTTCTCGTT	-
<i>GFP</i>	Probe	GFP-p HEX	TACCTGAGCACCAGTCCGCCCT	5' HEX 3' BHQ1
<i>Venus</i>	Probe	Venus-p HEX	TACCTGAGCTACCAGTCCGCCCT	5' HEX 3' BHQ1
<i>CyclinA2</i>	Forward primer	Lu58-hCyclinA2-FW	GGTGCACCCCTTAAGGATCTT	-
	Reverse primer	Lu48-CycA2-R	TGAACGCAGGCTGTTTACTGT	-
	Probe	hCyclinA2-p HEX	TGAGCATGTACCGTTCCTCCTTGG	5' HEX 3' BHQ1
<i>HRas-V12</i>	Forward primer	Lu59-HRas-FW	AGGGCTTCCTGTGTGTGTTG	-
	Reverse primer	Lu49-HRas-R	CCGTTTGATCTGCTCCCTGTA	-
	Probe	HRas-p HEX	CAACACCAAGTCTTTTGAGGACATCCACCA	5' HEX 3' BHQ1
<i>LargeT</i>	Forward primer	LargeT-FW	ATGGAAGACTCAGGGCATGAA	-
	Reverse primer	Lu50-LT-SV40-R	TGGTATGGCTGATTATGATCATGAA	-
	Probe	LargeT-p HEX	ACAGTCCCAAGGCTCATTTCAGGCC	5' HEX 3' BHQ1
<i>SFFV</i>	Forward primer	SFFV-FW	ACCCTGCGCCTTATTTGAATT	-
	Reverse primer	SFFV-R	CTTCTGCTTCCCGAGCTCTATAA	-
	Probe	SFFV-p FAM	CCAATCAGCCTGCTTCTCGCTTCTGTT	5' FAM 3' BHQ1

Table S3 List of all constructs and titers [IP/mL] used for transduction experiments of primary murine hepatocytes or NIH3T3 cells. * Kindly provided by Kristoffer Riecken (Dept. of Cell and Gene Therapy, University Medical Center Hamburg-Eppendorf, Germany); # Titer was determined after concentration by centrifugation.

Construct	1. Production Titer [IP/mL]	2. Production Titer [IP/mL]
LeGO-C2	5.89E+06	2.52E+07
LeGO-V2	7.45E+06	2.38E+07
LeGO-Cer2	3.93E+06	2.97E+07
LeGO-CyclinA2-iG2*	2.00E+07	-
LeGO-HRasV12-iG2	1.44E+07	1.26E+07
LeGO-SV40-LargeT-iV2	1.26E+07	1.59E+08 [#]

Table S4 Summary of experiments, recipient mouse strain, sex and MOI. Numbers in brackets indicate the number of transplanted mice, which died early after transplantation due to an unexpected death. Analysis showed no different outcome in the groups between USB, USRapG, and USG mice. USB, uPA/SCID/Beige; USRapG, uPA/SCID/Beige/Pfp^{-/-}/Rag2^{-/-}/Il2rg^{-/-}; USG, uPA/SCID/Beige/Il2rg^{-/-}; uPA, urokinase-type plasminogen activator; SCID, severe combined immunodeficiency; Pfp, Perforin gene; Rag2, recombination-activating gene 2; Il2rg, Interleukin-2 receptor common gamma chain.

Experiment	Group	USB		USRapG		USG		MOI	Analyzed animals
		male	female	male	female	male	female		
II	Mock	4 (1†)	0	0	0	0	0		
IV	Mock	2 (1†)	0	0	0	0	0		
	total	6	0	0	0	0	0		6
I	RGB	2	0	1	1	0	2	40	
II	RGB	0	4	0	0	0	0	40	
IV	RGB	0 (2†)	1 (2†)	0	0	0	0	40	
	total	2	5	1	1	0	2		11
I	CyclinA2	3	0	0	2	3	0	40	
III	CyclinA2	1 (3†)	1 (3†)	0	0	0	0	60	
IV	CyclinA2	0 (1†)	1 (4†)	0	0	0	0	40	
	total	4	2	0	2	3	0		11
II	HRas-V12	2	6	0	0	0	0	40	
III	HRas-V12	2	1 (1†)	0	0	0	0	60	
IV	HRas-V12	1 (2†)	0	0	0	0	0	40	
	total	5	7	0	0	0	0		12
I	LargeT	0	3	0	2	0	2	40	
II	LargeT	2	2	0	0	0	0	40	
III	LargeT	1 (1†)	0 (2†)	0	0	0	0	60	
	total	3	5	0	2	0	2		12

Table S5 Liver weight of experimental groups Mock, RGB, CyclinA2, HRas-V12, and LargeT.

Mouse #	Group	Sacrifice [days post Tx]	Liver weight [g]
7363	Mock	43	1.39
7373	Mock	72	1.37
7400	Mock	112	1.38
7406	Mock	112	1.20
7371	Mock	197	1.47
7372	Mock	197	1.35
7366	RGB	43	1.15
292	RGB	67	1.60
7338	RGB	67	1.40
7365	RGB	72	0.97
2515	RGB	101	1.00
7336	RGB	101	1.10
7411	RGB	112	1.16
2513	RGB	183	1.29
289	RGB	183	1.43
7367	RGB	193	0.85
7369	RGB	193	0.83
559	CyclinA2	66	1.36
288	CyclinA2	67	1.20
7333	CyclinA2	67	1.30
2518	CyclinA2	101	1.40
2514	CyclinA2	101	1.40
7403	CyclinA2	112	1.22
567	CyclinA2	115	1.20
2517	CyclinA2	183	1.77
7334	CyclinA2	183	1.74
7335	CyclinA2	183	1.46
291	CyclinA2	183	1.44
7361	HRas-V12	43	1.27
7362	HRas-V12	43	0.93
569	HRas-V12	43	1.02
561	HRas-V12	66	2.02
562	HRas-V12	66	1.32
7353	HRas-V12	71	2.47
7354	HRas-V12	71	1.37
7355	HRas-V12	71	1.41
7359	HRas-V12	71	1.67
7360	HRas-V12	71	1.20
7364	HRas-V12	71	1.53
7417	HRas-V12	112	1.38
7368	LargeT	43	0.93
7357	LargeT	50	1.69
572	LargeT	66	1.95
2516	LargeT	67	1.20
7339	LargeT	67	1.30
7358	LargeT	71	1.52
7370	LargeT	72	0.90
7337	LargeT	101	1.90
290	LargeT	101	2.70
2512	LargeT	115	1.30
287	LargeT	115	4.90
7340	LargeT	115	3.20
Mean			1.49
SD			0.65

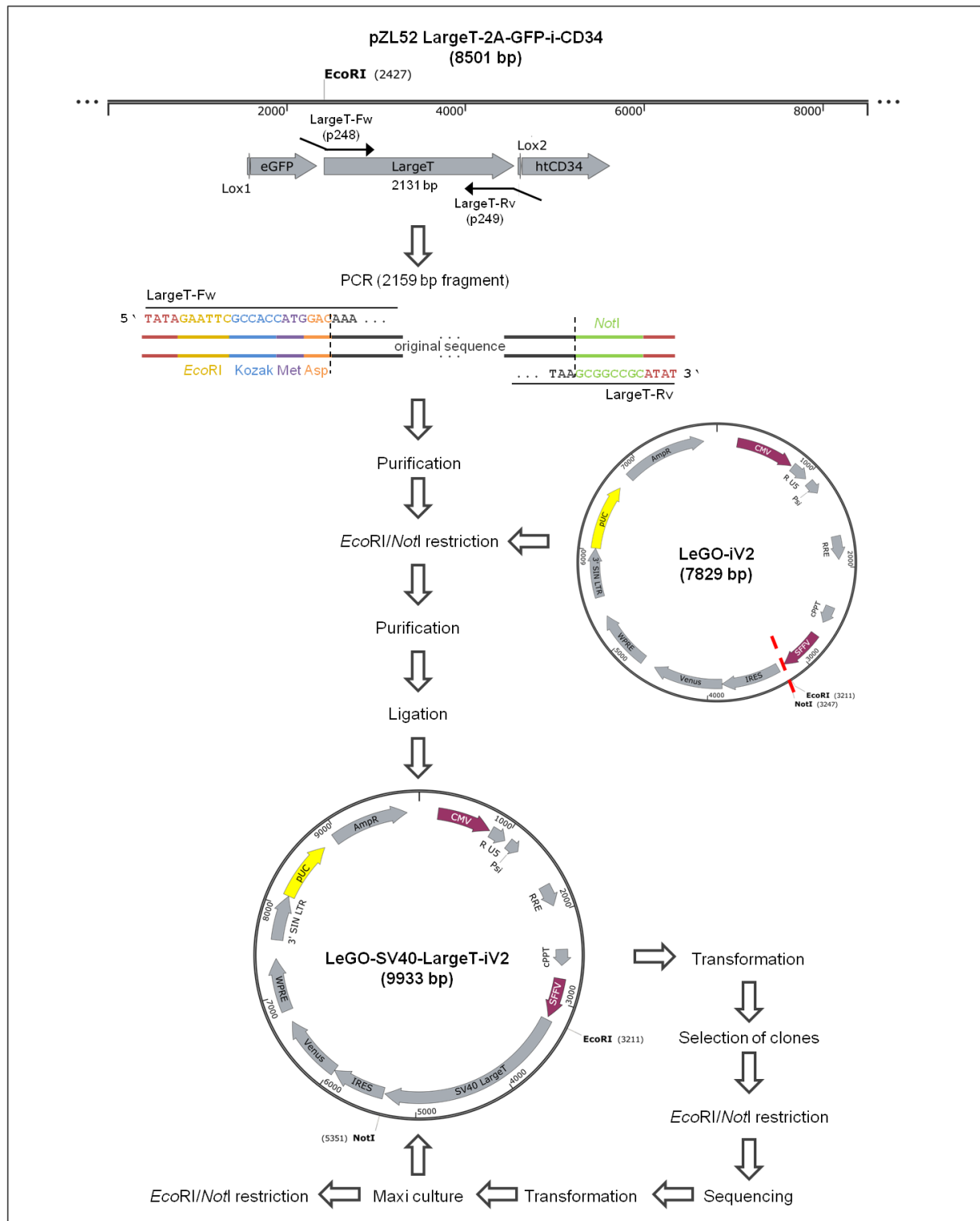


Figure S1 LeGO-LargeT-iV2 cloning strategy. First, a PCR was performed with primers LargeT-Fw/LargeT-Rv (p248/p249) on pZL52 LargeT-2A-GFP-i-CD34. The PCR product was purified and sequentially digested with *EcoRI*/*NotI* (red line). The backbone LeGO-iV2 was also digested with *EcoRI*/*NotI*. After purification, the fragments were ligated to transform competent bacteria afterwards. Single colonies were picked and cultured for plasmid mini preparation. All clones were screened via *EcoRI*/*NotI* digestion and two clones with correct fragment patterns were sequenced. One clone was selected for transformation and plasmid maxi preparation and finally analyzed by *EcoRI*/*NotI* digestion. Vector maps were generated with SnapGene Viewer v3.3.3. LeGO, lentiviral gene ontology; i / IRES, internal ribosomal entry site; V2, fluorescent protein Venus.

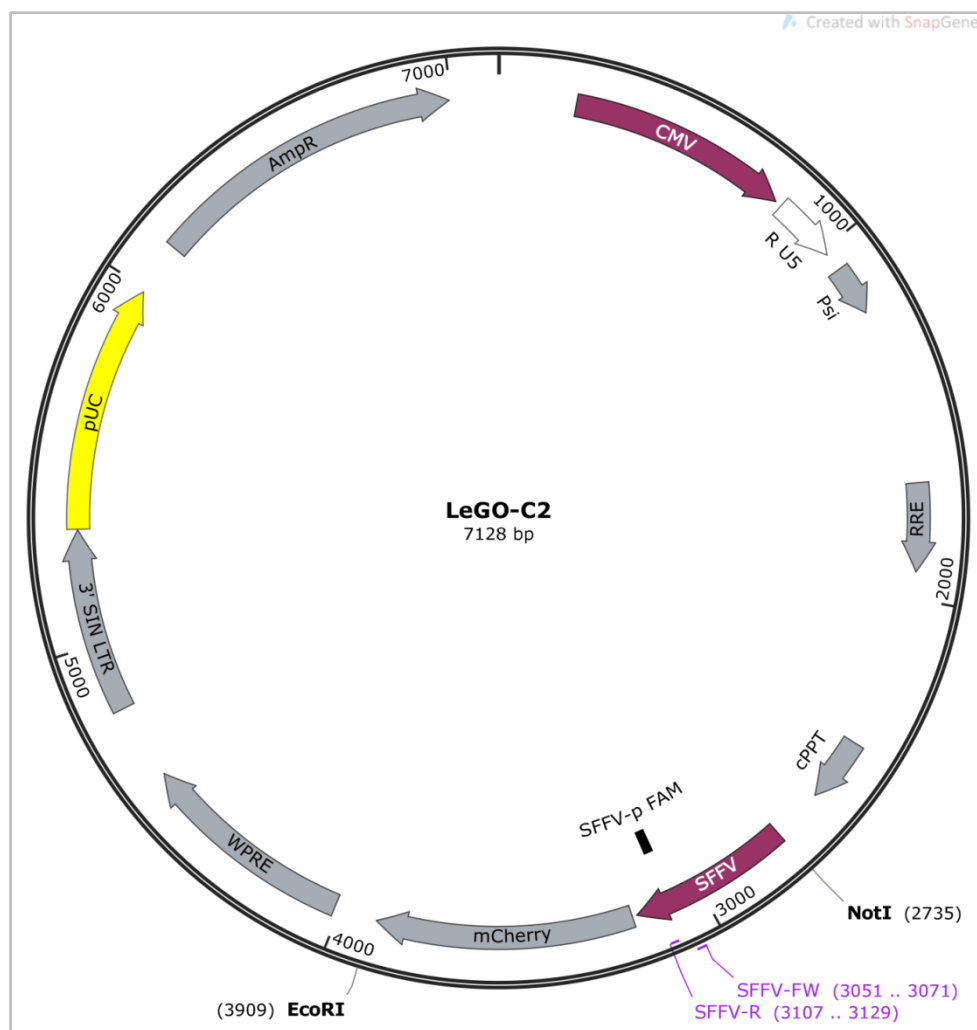


Figure S2 LeGO-C2 vector map. 3' SIN LTR, self-inactivating long-terminal repeat; AmpR, ampicillin resistance; CMV, cytomegalus virus promoter; cPPT, central polypurine tract; LeGO, lentiviral gene ontology; psi, packaging signal; pUC, origin of replication; R, redundant sequence; RRE, Rev-responsive element; SFFV, spleen focus-forming virus; SFFV-FW, forward primer for ddPCR; SFFV-p FAM, FAM-labeled ddPCR probe for SFFV; SFFV-R, reverse primer for ddPCR; U5, unique region; WPRE, Woodchuck hepatitis virus post-transcriptional regulatory element.

LeGO-C2 vector sequence

```

1  GTCGACGGAT  CGGGAGATCT  CCCGATCCCC  TATGGTGCAC  TCTCAGTACA
51  ATCTGCTCTG  ATGCCGCATA  GTTAAGCCAG  TATCTGCTCC  CTGCTTGTTG
101  GTTGGAGGTC  GCTGAGTAGT  GCGCGAGCAA  AATTTAAGCT  ACAACAAGGC
151  AAGGCTTGAC  CGACAATTGC  ATGAAGAATC  TGCTTAGGGT  TAGGCGTTTT
201  GCGCTGCTTC  GCGATGTACG  GGCCAGATAT  ACGCGTTGAC  ATTGATTATT
251  GACTAGTTAT  TAATAGTAAT  CAATTACGGG  GTCATTAGTT  CATAGCCCAT
301  ATATGGAGTT  CCGCGTTACA  TAACTTACGG  TAAATGGCCC  GCCTGGCTGA
351  CCGCCCAACG  ACCCCCGCCC  ATTGACGTCA  ATAATGACGT  ATGTTCCCAT
401  AGTAACGCCA  ATAGGGACTT  TCCATTGACG  TCAATGGGTG  GAGTATTTAC
451  GGTAAGTGC  CCACTTGGCA  GTACATCAAG  TGTATCATAT  GCCAAGTACG
501  CCCCTATTG  ACGTCAATGA  CGGTAAATGG  CCCGCTGGC  ATTATGCCCC
551  GTACATGACC  TTATGGGACT  TTCCTACTTG  GCAGTACATC  TACGTATTAG
601  TCATCGCTAT  TACCATGGTG  ATGCGGTTTT  GGCAGTACAT  CAATGGGCGT
651  GGATAGCGGT  TTGACTCACG  GGGATTTCCA  AGTCTCCACC  CCATTGACGT
701  CAATGGGAGT  TTGTTTTTGG  ACCAAAATCA  ACGGGACTTT  CCAAAAATGTC
751  GTAACAACTC  CGCCCCATTG  ACGCAAATGG  GCGGTAGGCG  TGTACGGTGG
801  GAGGTCTATA  TAAGCAGCGC  GTTTTGCCTG  TACTGGGTCT  CTCTGGTTAG

```

```

851 ACCAGATCTG AGCCTGGGAG CTCTCTGGCT AACTAGGGAA CCCACTGCTT
901 AAGCCTCAAT AAAGCTTGCC TTGAGTGCTT CAAGTAGTGT GTGCCCGTCT
951 GTTGTGTGAC TCTGGTAACT AGAGATCCCT CAGACCCCTT TAGTCAGTGT
1001 GGAAAATCTC TAGCAGTGGC GCCCGAACAG GGACTTGAAA GCGAAAAGGGA
1051 AACCAGAGGA GCTCTCTCGA CGCAGGACTC GGCTTGCTGA AGCGCGCACG
1101 GCAAGAGGCG AGGGGCGGCG ACTGGTGAGT ACGCCAAAAA TTTTGACTAG
1151 CGGAGGCTAG AAGGAGAGAG ATGGGTGCGA GAGCGTCAGT ATTAAGCGGG
1201 GGAGAATTAG ATCGCGATGG GAAAAAATTC GGTTAAGGCC AGGGGGAAAG
1251 AAAAAATATA AATTAAAACA TATAGTATGG GCAAGCAGGG AGCTAGAACG
1301 ATTCGCAGTT AATCCTGGCC TGTTAGAAAC ATCAGAAGGC TGTAGACAAA
1351 TACTGGGACA GCTACAACCA TCCCTTCAGA CAGGATCAGA AGAACTTAGA
1401 TCATTATATA ATACAGTAGC AACCCTCTAT TGTGTGCATC AAAGGATAGA
1451 GATAAAAGAC ACCAAGGAAG CTTTAGACAA GATAGAGGAA GAGCAAAACA
1501 AAAGTAAGAC CACCGCACAG CAAGCGGCCG GCCGCGCTGA TCTTCAGACC
1551 TGGAGGAGGA GATATGAGGG ACAATTGGAG AAGTGAATTA TATAAATATA
1601 AAGTAGTAAA AATTGAACCA TTAGGAGTAG CACCACCAA GGCAAAGAGA
1651 AGAGTGGTGC AGAGAGAAAA AAGAGCAGTG GGAATAGGAG CTTTGTTCCT
1701 TGGGTTCTTG GGAGCAGCAG GAAGCACTAT GGGCGCAGCG TCAATGACGC
1751 TGACGGTACA GGCCAGACAA TTATTGTCTG GTATAGTGCA GCAGCAGAAC
1801 AATTTGCTGA GGGCTATTGA GGCGCAACAG CATCTGTTGC AACTCACAGT
1851 CTGGGGCATC AAGCAGCTCC AGGCAAGAAT CCTGGCTGTG GAAAGATACC
1901 TAAAGGATCA ACAGCTCCTG GGGATTTGGG GTTGCTCTGG AAAACTCATT
1951 TGCACCACTG CTGTGCCTTG GAATGCTAGT TGGAGTAATA AATCTCTGGA
2001 ACAGATTTGG AATCACACGA CCTGGATGGA GTGGGACAGA GAAATTAACA
2051 ATTACACAAG CTTAATACAC TCCTTAATTG AAGAATCGCA AAACCAGCAA
2101 GAAAAGATG AACAAGAATT ATTGGAATTA GATAAATGGG CAAGTTTGTG
2151 GAATTGGTTT AACATAACAA ATTGGCTGTG GTATATAAAA TTATTCATAA
2201 TGATAGTAGG AGGCTTG GTA GGTTTAAGAA TAGTTTTTGC TGACTTTTCT
2251 ATAGTGAATA GAGTTAGGCA GGGATATTCA CCATTATCGT TTCAGACCCA
2301 CCTCCCAACC CCGAGGGGAC CCGACAGGCC CGAAGGAATA GAAGAAGAAG
2351 GTGGAGAGAG AGACAGAGAC AGATCCATTC GATTAGTGAA CGGATCGGCA
2401 CTGCGTGCGC CAATTCTGCA GACAAATGGC AGTATTCATC CACAATTTTA
2451 AAAGAAAAGG GGGGATTGGG GGGTACAGTG CAGGGGAAAG AATAGTAGAC
2501 ATAATAGCAA CAGACATACA AACTAAAGAA TTACAAAAAC AAATTACAAA
2551 AATTCAAAAT TTTCGGGTTT ATTACAGGGA CAGCAGAGAT CCAGTTTGGT
2601 TAGTACCGGG CCCGCTCTAG TCGAGGTCGA CGGTATCGAT AAGCTCGCTT
2651 CACGAGATTC CAGCAGGTCG AGGGACCTAA TAACTTCGTA TAGCATACAT
2701 TATACGAAGT TATATTAAGG GTTCCAAGCT TAAGCGGCCG CTGAAAGACC
2751 CCACCTGTAG GTTTGGCAAG CTAGCTGCAG TAACGCCATT TTGCAAGGCA
2801 TGGAAAAATA CCAAACCAAG AATAGAGAAG TTCAGATCAA GGGCGGGTAC
2851 ATGAAAATAG CTAACGTTGG GCCAAACAGG ATATCTGCGG TGAGCAGTTT
2901 CGGCCCCGGC CCGGGGCCAA GAACAGATGG TCACCGCAGT TTCGGCCCCG
2951 GCCCCAGGCC AAGAACAGAT GGTCCCCAGA TATGGCCCCA CCCTCAGCAG
3001 TTTCTTAAGA CCCATCAGAT GTTTCCAGGC TCCCCAAGG ACCTGAAATG
3051 ACCCTGCGCC TTATTTGAAT TAACCAATCA GCCTGCTTCT GCCTTCTGTT
3101 CGCGCGCTTC TGCTTCCCGA GCTCTATAAA AGAGCTCACA ACCCCTCACT
3151 CGGCGCGCCA GTCCTCCGAT TGACTGAGTC GCCCGGATCC CGCCACCATG
3201 GTGAGCAAGG GCGAGGAGGA TAACATGGCC ATCATCAAGG AGTTCATGCG
3251 CTTCAAGGTG CACATGGAGG GCTCCGTGAA CGGCCACGAG TTCGAGATCG
3301 AGGGCGAGGG CGAGGGCCGC CCCTACGAGG GCACCCAGAC CGCCAAGCTG
3351 AAGGTGACCA AGGGTGGCCC CCTGCCCTTC GCCTGGGACA TCCTGTCCCC
3401 TCAGTTCATG TACGGCTCCA AGGCCTACGT GAAGCACCCC GCCGACATCC
3451 CCGACTACTT GAAGCTGTCC TTCCCCGAGG GCTTCAAGTG GGAGCGCGTG
3501 ATGAACTTCG AGGACGGCGG CGTGGTGACC GTGACCCAGG ACTCCTCCCT
3551 GCAGGACGGC GAGTTCATCT ACAAGGTGAA GCTGCGCGGC ACCAACTTCC
3601 CCTCCGACGG CCCCCTAATG CAGAAGAAGA CCATGGGCTG GGAGGCCTCC
3651 TCCGAGCGGA TGTACCCCGA GGACGGCGCC CTGAAGGGCG AGATCAAGCA
3701 GAGGCTGAAG CTGAAGGACG GCGGCCACTA CGACGCTGAG GTCAAGACCA
3751 CCTACAAGGC CAAGAAGCCC GTGCAGCTGC CCGGCGCCTA CAACGTCAAC
3801 ATCAAGTTGG ACATCACCTC CCACAACGAG GACTACACCA TCGTGGAAAC
3851 GTCACGAACGC GCCGAGGGCC GCCACTCCAC CGGCGGCATG GACGAGCTGT
3901 ACAAGTAAGA ATTCGTCGAG GGACCTAATA ACTTCGTATA GCATACATTA
3951 TACGAAGTTA TACATGTTTA AGGGTTCCGG TTCCACTAGG TACAATTCGA

```

```

4001 TATCAAGCTT ATCGATAATC AACCTCTGGA TTACAAAATT TGTGAAAGAT
4051 TGACTGGTAT TCTTAACATAT GTTGCTCCTT TTACGCTATG TGGATACGCT
4101 GCTTTAATGC CTTTGTATCA TGCTATTGCT TCCCGTATGG CTTTCATTTT
4151 CTCCTCCTTG TATAAATCCT GGTTGCTGTC TCTTTATGAG GAGTTGTGGC
4201 CCGTTGTCAG GCAACGTGGC GTGGTGTGCA CTGTGTTTGC TGACGCAACC
4251 CCCACTGGTT GGGGCATTGC CACCACCTGT CAGCTCCTTT CCGGGACTTT
4301 CGCTTTCCCC CTCCTATTG CCACGGCGGA ACTCATCGCC GCCTGCCTTG
4351 CCCGCTGCTG GACAGGGGCT CGGCTGTTGG GCACTGACAA TTCCGTGGTG
4401 TTGTGCGGGA AATCATCGTC CTTTCCTTGG CTGCTCGCCT GTGTTGCCAC
4451 CTGGATTCTG CGCGGGACGT CCTTCTGCTA CGTCCCTTCG GCCCTCAATC
4501 CAGCGGACCT TCCTTCCCCG GGCCTGCTGC CGGCTCTGCG GCCTCTTCCG
4551 CGTCTTCGCC TTCGCCCTCA GACGAGTCGG ATCTCCCTTT GGGCCGCCTC
4601 CCCGCATCGA TACCGTCGAC CTCGATCGAG ACCTAGAAAA ACATGGAGCA
4651 ATCACAAGTA GCAATACAGC AGCTACCAAT GCTGATTGTG CCTGGCTAGA
4701 AGCACAAGAG GAGGAGGAGG TGGGTTTTCC AGTCACACCT CAGGTACCTT
4751 TAAGACCAAT GACTTACAAG GCAGCTGTAG ATCTTAGCCA CTTTTTAAAA
4801 GAAAAGGGGG GACTGGAAGG GCTAATTCAC TCCAACGAA GACAAGATAT
4851 CTTTGATCTG TGGATCTACC ACACACAAGG CTACTTCCCT GATTGGCAGA
4901 ACTACACACC AGGGCCAGGG ATCAGATATC CACTGACCTT TGGATGGTGC
4951 TACAAGCTAG TACCAGTTGA GCAAGAGAAG GTAGAAGAAG CCAATGAAGG
5001 AGAGAACACC CGCTTGTTAC ACCCTGTGAG CCTGCGTGGG ATGGATGACC
5051 CGGAGAGAGA AGTATTAGAG TGGAGGTTTG ACAGCCGCCT AGCATTTTCAT
5101 CACATGGCCC GAGAGCTGCA TCCGGACTGT ACTGGGTCTC TCTGGTTAGA
5151 CCAGATCTGA GCCTGGGAGC TCTCTGGCTA ACTAGGGAAC CCACTGCTTA
5201 AGCCTCAATA AAGCTTGCCCT TGAGTGCTTC AAGTAGTGTG TGCCCGTCTG
5251 TTGTGTGACT CTGGTAACTA GAGATCCCTC AGACCTTTT AGTCAGTGTG
5301 GAAAATCTCT AGCAGCATGT GAGCAAAAGG CCAGCAAAAG GCCAGGAACC
5351 GTAAAAAGGC CGCGTTGCTG GCGTTTTTCC ATAGGCTCCG CCCCCGTGAC
5401 GAGCATCACA AAAATCGACG CTCAAGTCAG AGGTGGCGAA ACCCGACAGG
5451 ACTATAAAGA TACCAGGCGT TTCCCCCTGG AAGCTCCCTC GTGCGCTCTC
5501 CTGTTCCGAC CCTGCCGCTT ACCGGATACC TGTCCGCTT TCTCCCTTCG
5551 GGAAGCGTGG CGCTTTCTCA TAGCTCACGC TGTAGGTATC TCAGTTCCGGT
5601 GTAGGTGCTT CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCCGTTTCAGC
5651 CCGACCGCTG CGCCTTATCC GGTAACATAT GTCTTGAGTC CAACCCGGTA
5701 AGACACGACT TATCGCCACT GGCAGCAGCC ACTGGTAACA GGATTAGCAG
5751 AGCGAGGTAT GTAGGCGGTG CTACAGAGTT CTTGAAGTGG TGGCTAACT
5801 ACGGCTACAC TAGAAGAACA GTATTTGGTA TCTGCGCTCT GCTGAAGCCA
5851 GTTACCTTCG GAAAAAGAGT TGGTAGCTCT TGATCCGGCA AACAAACCAC
5901 CGCTGGTAGC GGTGGTTTTT TTGTTTGCAA GCAGCAGATT ACGCGCAGAA
5951 AAAAAGGATC TCAAGAAGAT CCTTTGATCT TTTCTACGGG GTCTGACGCT
6001 CAGTGGAACG AAAACTCACG TTAAGGGATT TTGGTCATGA GATTATCAAA
6051 AAGGATCTTC ACCTAGATCC TTTTAAATTA AAAATGAAGT TTTAAATCAA
6101 TCTAAAGTAT ATATGAGTAA ACTTGGTCTG ACAGTTACCA ATGCTTAATC
6151 AGTAGGCAC CTATCTCAGC GATCTGTCTA TTTGCTTCAT CCATAGTTGC
6201 CTGACTCCCC GTCGTGTAGA TAACTACGAT ACGGGAGGGC TTACCATCTG
6251 GCCCCAGTGC TGCAATGATA CCGCGAGACC CACGCTCACC GGCTCCAGAT
6301 TTATCAGCAA TAAACCAGCC AGCCGGAAGG GCCGAGCGCA GAAAGTGGTCC
6351 TGCAACTTTA TCCGCTCCA TCCAGTCTAT TAATTGTTGC CGGGAAGCTA
6401 GAGTAAGTAG TTCGCCAGTT AATAGTTTGC GCAACGTTGT TGCCATTGCT
6451 ACAGGCATCG TGGTGTACG CTCGTCGTTT GGTATGGCTT CATTCAGCTC
6501 CGGTTCCCAA CGATCAAGGC GAGTTACATG ATCCCCATG TTGTGCAAAA
6551 AAGCGGTTAG CTCCTTCGGT CCTCCGATCG TTGTCAGAAG TAAGTTGGCC
6601 GCAGTGTTAT CACTCATGGT TATGGCAGCA CTGCATAATT CTCTTACTGT
6651 CATGCCATCC GTAAGATGCT TTTCTGTGAC TGGTGAGTAC TCAACCAAGT
6701 CATTCTGAGA ATAGTGATAG CGGCGACCGA GTTGCTCTTG CCCGCGCTCA
6751 ATACGGGATA ATACCGCGCC ACATAGCAGA ACTTTAAAAG TGCTCATCAT
6801 TGGAAAACGT TCTTCGGGGC GAAAACCTCT AAGGATCTTA CCGCTGTTGA
6851 GATCCAGTTC GATGTAACCC ACTCGTGAC CCAACTGATC TTCAGCATCT
6901 TTTACTTTCA CCAGCGTTTC TGGGTGAGCA AAAACAGGAA GGCAAAATGC
6951 CGCAAAAAAG GGAATAAGGG CGACACGGAA ATGTTGAATA CTCATACTCT
7001 TCCTTTTTTCA ATATTATTGA AGCATTTATC AGGGTTATTG TCTCATGAGC
7051 GGATACATAT TTGAATGTAT TTAGAAAAAT AAACAAATAG GGGTTCGCGC
7101 CACATTTCCC CGAAAAGTGC CACCTGAC

```

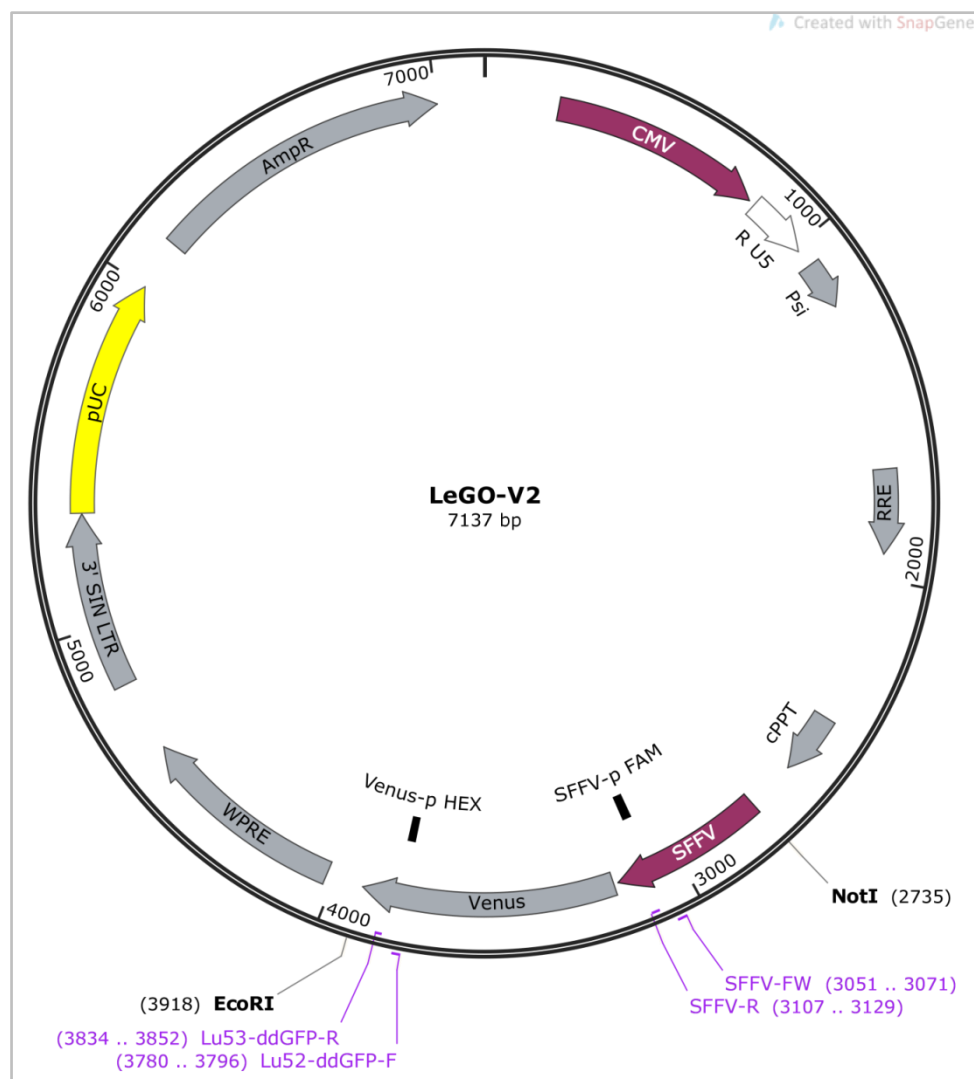


Figure S3 LeGO-V2 vector map. 3' SIN LTR, self-inactivating long-terminal repeat; AmpR, ampicillin resistance; CMV, cytomegalus virus promoter; cPPT, central polypurine tract; LeGO, lentiviral gene ontology; Lu52-ddGFP-F, forward primer for ddPCR; Lu53-ddGFP-R, reverse primer for ddPCR; psi, packaging signal; pUC, origin of replication; R, redundant sequence; RRE, Rev-responsive element; SFFV, spleen focus-forming virus; SFFV-FW, forward primer for ddPCR; SFFV-p FAM, FAM-labeled ddPCR probe for SFFV; SFFV-R, reverse primer for ddPCR; U5, unique region; V2, Venus, green-fluorescent protein; Venus-p HEX, HEX-labeled ddPCR probe for Venus; WPRE, Woodchuck hepatitis virus post-transcriptional regulatory element.

LeGO-V2 vector sequence

```

1  GTCGACGGAT  CGGGAGATCT  CCCGATCCCC  TATGGTGCAC  TCTCAGTACA
51  ATCTGCTCTG  ATGCCGCATA  GTTAAGCCAG  TATCTGCTCC  CTGCTTGTGT
101 GTTGGAGGTC  GCTGAGTAGT  GCGCGAGCAA  AATTTAAGCT  ACAACAAGGC
151 AAGGCTTGAC  CGACAATTGC  ATGAAGAATC  TGCTTAGGGT  TAGGCGTTTT
201 GCGCTGCTTC  GCGATGTACG  GGCCAGATAT  ACGCGTTGAC  ATTGATTATT
251 GACTAGTTAT  TAATAGTAAT  CAATTACGGG  GTCATTAGTT  CATAGCCCAT
301 ATATGGAGTT  CCGCGTTACA  TAACTTACGG  TAAATGGCCC  GCCTGGCTGA
351 CCGCCCAACG  ACCCCCACGC  ATTGACGTCA  ATAATGACGT  ATGTTCCCAT
401 AGTAACGCCA  ATAGGGACTT  TCCATTGACG  TCAATGGGTG  GAGTATTTAC
451 GGTAAGTGC  CCACTTGGCA  GTACATCAAG  TGTATCATAT  GCCAAGTACG
501 CCCCTATTG  ACGTCAATGA  CGGTAAATGG  CCCGCTGGC  ATTATGCCCC
551 GTACATGACC  TTATGGGACT  TTCCTACTTG  GCAGTACATC  TACGTATTAG
601 TCATCGCTAT  TACCATGGTG  ATGCGGTTTT  GGCAGTACAT  CAATGGGCGT
651 GGATAGCGGT  TTGACTCACG  GGGATTTCAC  AGTCTCCACC  CCATTGACGT

```

```

701 CAATGGGAGT TTGTTTTGGC ACCAAAATCA ACGGGACTTT CCAAAATGTC
751 GTAACAACTC CGCCCCATTG ACGCAAATGG GCGGTAGGCG TGTACGGTGG
801 GAGGTCTATA TAAGCAGCGC GTTTTGCCTG TACTGGGTCT CTCTGGTTAG
851 ACCAGATCTG AGCCTGGGAG CTCTCTGGCT AACTAGGGAA CCCACTGCTT
901 AAGCCTCAAT AAAGCTTGCC TTGAGTGCTT CAAGTAGTGT GTGCCCCGCT
951 GTTGTGTGAC TCTGGTAACT AGAGATCCCT CAGACCCTTT TAGTCAGTGT
1001 GGAAAATCTC TAGCAGTGGC GCCCGAACAG GGACTTGAAA GCGAAAGGGA
1051 AACCAGAGGA GCTCTCTCGA CGCAGGACTC GGCTTGCTGA AGCGCGCACG
1101 GCAAGAGGCG AGGGGCGGCG ACTGGTGAGT ACGCCAAAAA TTTTGACTAG
1151 CGGAGGCTAG AAGGAGAGAG ATGGGTGCGA GAGCGTCAGT ATTAAGCGGG
1201 GGAGAATTAG ATCGCGATGG GAAAAAATTC GGTTAAGGCC AGGGGGAAAG
1251 AAAAAATATA AATTA AAAACA TATAGTATGG GCAAGCAGGG AGCTAGAACG
1301 ATTCGCAGTT AATCCTGGCC TGTTAGAAAC ATCAGAAAGG TGTTAGACAAA
1351 TACTGGGACA GCTACAACCA TCCCTTCAGA CAGGATCAGA AGAACTTAGA
1401 TCATTATATA ATACAGTAGC AACCTCTAT TGTGTGCATC AAAGGATAGA
1451 GATAAAAGAC ACCAAGGAAG CTTTAGACAA GATAGAGGAA GAGCAAAACA
1501 AAAGTAAGAC CACCGCACAG CAAGCGGCCG GCCGCGCTGA TCTTCAGACC
1551 TGGAGGAGGA GATATGAGGG ACAATTGGAG AAGTGAATTA TATAAATATA
1601 AAGTAGTAAA AATTGAACCA TTAGGAGTAG CACCCACCAA GGCAAAGAGA
1651 AGAGTGGTGC AGAGAGAAAA AAGAGCAGTG GGAATAGGAG CTTTGTTCCT
1701 TGGGTTCTTG GGAGCAGCAG GAAGCACTAT GGGCGCAGCG TCAATGACGC
1751 TGACGGTACA GGCCAGACAA TTATTGTCTG GTATAGTGCA GCAGCAGAAC
1801 AATTTGCTGA GGGCTATTGA GGCGCAACAG CATCTGTTGC AACTCACAGT
1851 CTGGGGCATC AAGCAGCTCC AGGCAAGAAT CCTGGCTGTG GAAAGATACC
1901 TAAAGGATCA ACAGCTCCTG GGGATTGGG GTTGCTCTGG AAAACTCATT
1951 TGCACCACTG CTGTGCCTTG GAATGCTAGT TGGAGTAATA AATCTCTGGA
2001 ACAGATTTGG AATCACACGA CCTGGATGGA GTGGGACAGA GAAATTAACA
2051 ATTACACAAG CTTAATACAC TCCTTAATTG AAGAATCGCA AAACCAGCAA
2101 GAAAAGAATG AACAAGAATT ATTGGAATTA GATAAATGGG CAAGTTTGTG
2151 GAATTGGTTT AACATAACAA ATTGGCTGTG GTATATAAAA TTATTCATAA
2201 TGATAGTAGG AGGCTTGGA GGTTTAAGAA TAGTTTTTGC TGTACTTTCT
2251 ATAGTGAATA GAGTTAGGCA GGGATATTCA CCATTATCGT TTCAGACCCA
2301 CCTCCAACC CCGAGGGGAC CCGACAGGCC CGAAGGAATA GAAGAAGAAG
2351 GTGGAGAGAG AGACAGAGAC AGATCCATTC GATTAGTGAA CGGATCGGCA
2401 CTGCGTGCGC CAATTCTGCA GACAAATGGC AGTATTCATC CACAATTTTA
2451 AAAGAAAAGG GGGGATTGGG GGGTACAGTG CAGGGGAAAG AATAGTAGAC
2501 ATAATAGCAA CAGACATACA AACTAAAGAA TTACAAAAAC AAATTACAAA
2551 AATTCAAAAT TTTCGGGTTT ATTACAGGGA CAGCAGAGAT CCAGTTTGGT
2601 TAGTACCGGG CCCGCTCTAG TCGAGGTCGA CGGTATCGAT AAGCTCGCTT
2651 CACGAGATTC CAGCAGGTCG AGGGACCTAA TAACTTCGTA TAGCATACAT
2701 TATACGAAGT TATATTAAGG GTTCCAAGCT TAAGCGGCCG CTGAAAGACC
2751 CCACCTGTAG GTTTGGCAAG CTAGCTGCAG TAACGCCATT TTGCAAGGCA
2801 TGGAAAAATA CCAAACCAAG AATAGAGAAG TTCAGATCAA GGGCGGGTAC
2851 ATGAAAAATAG CTAAGCTTGG GCCAAACAGG ATATCTGCGG TGAGCAGTTT
2901 CGGCCCCGCG CCGGGGCCAA GAACAGATGG TCACCGCAGT TTCGGCCCCG
2951 GCCCGAGGCC AAGAACAGAT GGTCCCCAGA TATGGCCCCA CCTCAGCAG
3001 TTTCTTAAGA CCCATCAGAT GTTTCCAGGC TCCCCAAGG ACCTGAAATG
3051 ACCCTGCGCC TTATTTGAAT TAACCAATCA GCCTGCTTCT CGCTTCTGTT
3101 CGCGCGCTTC TGCTTCCCGA GCTCTATAAA AGAGCTCACA ACCCTCACT
3151 CGGCGCGCCA GTCCTCCGAT TGACTGAGTC GCCCGGATCC CGCCACCATG
3201 GTGAGCAAGG GCGAGGAGCT GTTCACGGG GTGGTGCCCA TCCTGGTCGA
3251 GCTGGACGGC GACGTAAACG GCCACAAGTT CAGCGTGTCC GCGAGGGCG
3301 AGGGCGATGC CACCTACGGC AAGCTGACCC TGAAGCTGAT CTGCACCACC
3351 GGCAAGCTGC CCGTGCCCTG GCCCACCCTC GTGACCACCC TGGGCTACGG
3401 CCTGCAGTGC TTCGCCCCTG ACCCCGACCA CATGAAGCAG CACGACTTCT
3451 TCAAGTCCGC CATGCCCCGA GGCTACGTCC AGGAGCGCAC CATCTTCTTC
3501 AAGGACGACG GCAACTACAA GACCCGCGCC GAGGTGAAGT TCGAGGGCGA
3551 CACCCTGGTG AACCGCATCG AGCTGAAGGG CATCGACTTC AAGGAGGACG
3601 GCAACATCCT GGGGCACAAG CTGGAGTACA ACTACAACAG CCACAACGTC
3651 TATATCACCG CCGACAAGCA GAAGAACGGC ATCAAGGCCA ACTTCAAGAT
3701 CTGCCACAAC ATCGAGGACG GCGGCGTGCA GCTCGCCGAC CACTACCAGC
3751 AGAACACCCC CATCGGCGAC GGCCCCGTGC TGCTGCCCGA CAACCACCTAC
3801 CTGAGCTACC AGTCCGCCCT GAGCAAAGAC CCCAACGAGA AGCGCGATCA

```



```

3851 CATGGTCCTG CTGGAGTTTC TGACCGCCGC CGGGATCACT CTCGGCATGG
3901 ACGAGCTGTA CAAGTAAGAA TTCGTCGAGG GACCTAATAA CTTCTATATAG
3951 CATAATTAT ACGAAGTTAT ACATGTTTAA GGGTTCCGGT TCCACTAGGT
4001 ACAATTTCGAT ATCAAGCTTA TCGATAATCA ACCTCTGGAT TACAAAAATTT
4051 GTGAAAGATT GACTGGTATT CTTAACTATG TTGCTCCTTT TACGCTATGT
4101 GGATACGCTG CTTTAATGCC TTTGTATCAT GCTATTGCTT CCCGTATGGC
4151 TTTTCATTTTC TCCTCCTTGT ATAAATCCTG GTTGCTGTCT CTTTATGAGG
4201 AGTTGTGGCC CGTTGTCAGG CAACGTGGCG TGGTGTGCAC TGTGTTTGCT
4251 GACGCAACCC CCACTGGTTG GGGCATTGCC ACCACCTGTC AGCTCCTTTC
4301 CGGGACTTTC GCTTTCCCCC TCCCTATTGC CACGGCGGAA CTCATCGCCG
4351 CCTGCCTTGC CCGCTGCTGG ACAGGGGCTC GGCTGTTGGG CACTGACAAT
4401 TCCGTGGTGT TGTGCGGGAA ATCATCGTCC TTTCTTGGC TGCTCGCCTG
4451 TGTTGCCACC TGGATTCTGC GCGGGACGTC CTTCTGCTAC GTCCCTTCGG
4501 CCCTCAATCC AGCGGACCTT CTTCCCGCG GCCTGCTGCC GGCTCTGCGG
4551 CCTCTTCCGC GTCTTCGCCT TCGCCCTCAG ACGAGTCGGA TCTCCCTTTG
4601 GGCCGCTCC CCGCATCGAT ACCGTCGACC TCGATCGAGA CCTAGAAAAA
4651 CATGGAGCAA TCACAAGTAG CAATACAGCA GCTACCAATG CTGATTGTGC
4701 CTGGCTAGAA GCACAAGAGG AGGAGGAGGT GGGTTTTCCTA GTCACACCTC
4751 AGGTACCTTT AAGACCAATG ACTTACAAGG CAGCTGTAGA TCTTAGCCAC
4801 TTTTTAAAG AAAAGGGGGG ACTGGAAGGG CTAATTCACCT CCCAACGAAG
4851 ACAAGATATC CTTGATCTGT GGATCTACCA CACACAAGGC TACTTCCCTG
4901 ATTGGCAGAA CTACACACCA GGGCCAGGGA TCAGATATCC ACTGACCTTT
4951 GGATGGTGCT ACAAGCTAGT ACCAGTTGAG CAAGAGAAGG TAGAAGAAGC
5001 CAATGAAGGA GAGAACACCC GCTTGTTACA CCCTGTGAGC CTGCATGGGA
5051 TGGATGACCC GGAGAGAGAA GTATTAGAGT GGAGGTTTGA CAGCCGCCTA
5101 GCATTTTCATC ACATGGCCCG AGAGCTGCAT CCGGACTGTA CTGGGTCTCT
5151 CTGGTTAGAC CAGATCTGAG CCGGGAGCT CTCTGGCTAA CTAGGGAACC
5201 CACTGCTTAA GCCTCAATAA AGCTTGCTT GAGTGCTTCA AGTAGTGTGT
5251 GCGCGTCTGT TGTGTGACTC TGGTAACTAG AGATCCCTCA GACCCCTTTA
5301 GTCAGTGTGG AAAATCTCTA GCAGCATGTG AGCAAAAAGG CAGCAAAAAGG
5351 CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC
5401 CCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA
5451 CCCGACAGGA CTATAAAGAT ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG
5501 TGCCTCTCC TGTTCCGACC CTGCCGCTTA CCGGATACCT GTCCGCCTTT
5551 CTCCCTTCGG GAAGCGTGGC GCTTTCTCAT AGCTCACGCT GTAGGTATCT
5601 CAGTTCGGTG TAGGTCGTTT GCTCCAAGCT GGGCTGTGTG CACGAACCCC
5651 CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACCTATCG TCTTGAGTCC
5701 AACCCGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG
5751 GATTAGCAGA GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT
5801 GGCCTAACTA CGGCTACACT AGAAGAACAG TATTTGGTAT CTGCGCTCTG
5851 CTGAAGCCAG TTACCTTCGG AAAAAGAGTT GGTAGCTCTT GATCCGGCAA
5901 ACAAACCACC GCTGGTAGCG GTGGTTTTTT TGTGTGCAAG CAGCAGATTA
5951 CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG
6001 TCTGACGCTC AGTGAACGA AAACCTACGT TAAGGGATT TGGTCATGAG
6051 ATTATCAAAA AGGATCTTCA CCTAGATCCT TTTAAATTAA AAATGAAGTT
6101 TTAAATCAAT CTAAAGTATA TATGAGTAAA CTTGGTCTGA CAGTTACCAA
6151 TGCTTAATCA GTGAGGCACC TATCTCAGCG ATCTGTCTAT TTCGTTTCATC
6201 CATAGTTGCC TGAATCCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT
6251 TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG
6301 GCTCCAGATT TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG
6351 AAGTGGTCTT GCAACTTTAT CCGCTCCAT CCAGTCTATT AATTGTTGCC
6401 GGAAGCTAG AGTAAGTAGT TCGCCAGTTA ATAGTTTGGC CAACGTTGTT
6451 GCCATTGCTA CAGGCATCGT GGTGTCACGC TCGTCGTTTG GTATGGCTTC
6501 ATTCAGCTCC GGTTCCTAAC GATCAAGGCG AGTTACATGA TCCCCCATGT
6551 TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC CTCCGATCGT TGTCAGAAGT
6601 AAGTTGGCCG CAGTGTTATC ACTCATGGTT ATGGCAGCAC TGCATAATTC
6651 TCTTACTGTC ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT
6701 CAACCAAGTC ATTCTGAGAA TAGTGATATG GCGGACCGAG TTGCTCTTGC
6751 CCGGCGTCAA TACGGGATAA TACCGCGCCA CATAGCAGAA CTTTAAAAGT
6801 CGTCATCATT GGAAAACGTT CTTGCGGGCG AAAACTCTCA AGGATCTTAC
6851 GCTGTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT
6901 TCAGCATCTT TTACTTTTAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG
6951 GCAAAATGCC GCAAAAAAGG GAATAAGGGC GACACGGAAA TGTGAATAC

```

```

7001 TCATACTCTT CCTTTTTCAT TATTATTGAA GCATTATCA GGGTTATTGT
7051 CTCATGAGCG GATACATATT TGAATGTATT TAGAAAAATA AACAAATAGG
7101 GGTTCCGCGC ACATTTCCCC GAAAAGTGCC ACCTGAC

```

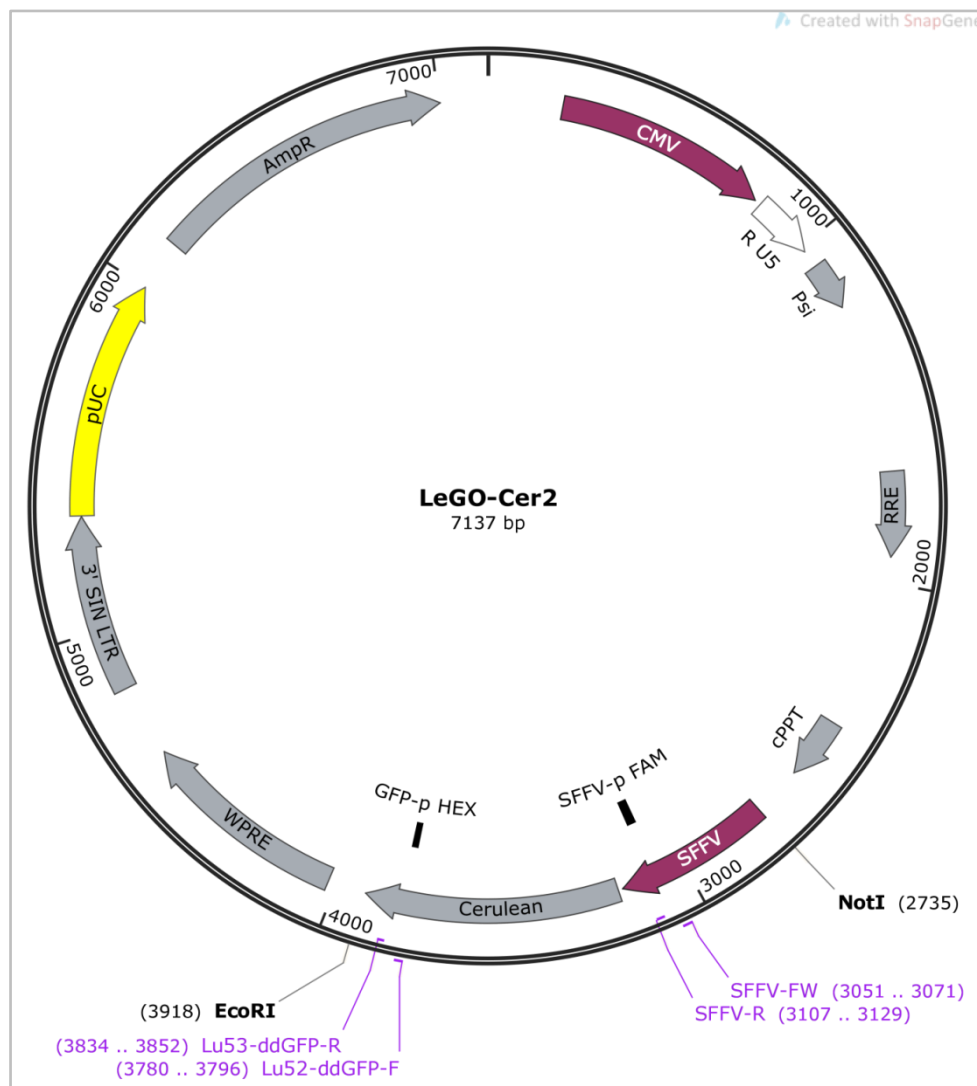


Figure S4 LeGO-Cer2 vector map. 3' SIN LTR, self-inactivating long-terminal repeat; AmpR, ampicillin resistance; Cer2, Cerulean, blue-fluorescent protein; CMV, cytomegalus virus promoter; cPPT, central polypurine tract; GFP-p HEX, HEX-labeled ddPCR probe for GFP (also binds to Cerulean due to sequence homology); LeGO, lentiviral gene ontology; Lu52-ddGFP-F, forward primer for ddPCR; Lu53-ddGFP-R, reverse primer for ddPCR; psi, packaging signal; pUC, origin of replication; R, redundant sequence; RRE, Rev-responsive element; SFFV, spleen focus-forming virus; SFFV-FW, forward primer for ddPCR; SFFV-p FAM, FAM-labeled ddPCR SFFV probe for SFFV; SFFV-R, reverse primer for ddPCR; U5, unique region; WPRE, Woodchuck hepatitis virus post-transcriptional regulatory element.

LeGO-Cer2 vector sequence

```

1  GTCGACGGAT CGGGAGATCT CCCGATCCCC TATGGTGAC TCTCAGTACA
51 ATCTGCTCTG ATGCCGCATA GTTAAGCCAG TATCTGCTCC CTGCTTGTGT
101 GTTGGAGGTC GCTGAGTAGT GCGCGAGCAA AATTTAAGCT ACAACAAGGC
151 AAGGCTTGAC CGACAATTGC ATGAAGAATC TGCTTAGGGT TAGGCGTTTT
201 GCGCTGCTTC GCGATGTACG GGCCAGATAT ACGCGTTGAC ATTGATTATT
251 GACTAGTTAT TAATAGTAAT CAATTACGGG GTCATTAGTT CATAGCCCAT
301 ATATGGAGTT CCGCGTTACA TAACTTACGG TAAATGGCCC GCCTGGCTGA
351 CCGCCAACG ACCCCGCCCC ATTGACGTCA ATAATGACGT ATGTTCCCAT

```

```

401 AGTAACGCCA ATAGGGACTT TCCATTGACG TCAATGGGTG GAGTATTTAC
451 GGTAAACTGC CCACTTGGCA GTACATCAAG TGTATCATAT GCCAAGTACG
501 CCCCTATTG ACGTCAATGA CGGTAAATGG CCCGCTGGC ATTATGCCCCA
551 GTACATGACC TTATGGGACT TTCTACTTGG GCAGTACATC TACGTATTAG
601 TCATCGCTAT TACCATGGTG ATGCGGTTTT GGCAGTACAT CAATGGGCGT
651 GGATAGCGGT TTGACTCACG GGGATTTCCA AGTCTCCACC CCATTGACGT
701 CAATGGGAGT TTGTTTTGGC ACCAAAATCA ACGGGACTTT CCAAAAATGTC
751 GTAACAACTC CGCCCCATTG ACGCAAATGG GCGGTAGGCG TGTACGGTGG
801 GAGGTCTATA TAAGACGGC GTTTTGCCTG TACTGGGTCT CTCTGGTTAG
851 ACCAGATCTG AGCCTGGGAG CTCTCTGGCT AACTAGGGAA CCCACTGCTT
901 AAGCCTCAAT AAAGCTTGCC TTGAGTGCTT CAAGTAGTGT GTGCCCGTCT
951 GTTGTGTGAC TCTGGTAACT AGAGATCCCT CAGACCCCTT TAGTCAGTGT
1001 GGAAAATCTC TAGCAGTGGC GCCCGAACAG GGACTTGAAA GCGAAAAGGGA
1051 AACCAGAGGA GCTCTCTCGA CGCAGGACTC GGCTTGCTGA AGCGCGCACG
1101 GCAAGAGGCG AGGGGCGGCG ACTGGTGAGT ACGCCAAAAA TTTTGACTAG
1151 CGGAGGCTAG AAGGAGAGAG ATGGGTGCGA GAGCGTCAGT ATTAAGCGGG
1201 GGAGAATTAG ATCGCGATGG GAAAAAATTC GGTAAAGGCC AGGGGGAAAG
1251 AAAAAATATA AATTAAACA TATAGTATGG GCAAGCAGGG AGCTAGAACG
1301 ATTCGCAGTT AATCCTGGCC TGTTAGAAAC ATCAGAAGGC TGTAGACAAA
1351 TACTGGGACA GCTACAACCA TCCCTTCAGA CAGGATCAGA AGAACTTAGA
1401 TCATTATATA ATACAGTAGC AACCTCTAT TGTGTGCATC AAAGGATAGA
1451 GATAAAAGAC ACCAAGGAAG CTTTAGACAA GATAGAGGAA GAGCAAAACA
1501 AAAGTAAGAC CACCGCACAG CAAGCGGCCG GCCGCGCTGA TCTTCAGACC
1551 TGGAGGAGGA GATATGAGGG ACAATTGGAG AAGTGAATTA TATAAATATA
1601 AAGTAGTAAA AATTGAACCA TTAGGAGTAG CACCCACCAA GGCAAAGAGA
1651 AGAGTGGTGC AGAGAGAAAA AAGAGCAGTG GGAATAGGAG CTTTGTTCCT
1701 TGGGTTCTTG GGAGCAGCAG GAAGCACTAT GGGCGCAGCG TCAATGACGC
1751 TGACGGTACA GGCCAGACAA TTATTGTCTG GTATAGTGCA GCAGCAGAAC
1801 AATTTGCTGA GGGCTATTGA GGCGCAACAG CATCTGTTGC AACTCACAGT
1851 CTGGGGCATC AAGCAGCTCC AGGCAAGAAT CCTGGCTGTG GAAAGATACC
1901 TAAAGGATCA ACAGCTCCTG GGGATTTGGG GTTGCTCTGG AAAACTCATT
1951 TGCACCACTG CTGTGCCTTG GAATGCTAGT TGGAGTAATA AATCTCTGGA
2001 ACAGATTTGG AATCACACGA CCTGGATGGA GTGGGACAGA GAAATTAACA
2051 ATTACACAAG CTTAATACAC TCCTTAATTG AAGAATCGCA AAACCAGCAA
2101 GAAAAGAATG AACAAGAATT ATTGGAATTA GATAAATGGG CAAGTTTGTG
2151 GAATTGGTTT AACATAACAA ATTGCTGTG GTATATAAAA TTATTCATAA
2201 TGATAGTAGG AGGCTTGGTA GGTTTAAGAA TAGTTTTTGC TGTACTTTCT
2251 ATAGTGAATA GAGTTAGGCA GGGATATTCA CCATTATCGT TTCAGACCCA
2301 CCTCCCAACC CCGAGGGGAC CCGACAGGCC CGAAGGAATA GAAGAAGAAG
2351 GTGGAGAGAG AGACAGAGAC AGATCCATTC GATTAGTGAA CGGATCGGCA
2401 CTGCGTGCGC CAATTCTGCA GACAAATGGC AGTATTCATC CACAAATTTA
2451 AAAGAAAAGG GGGGATTGGG GGGTACAGTG CAGGGGAAAG AATAGTAGAC
2501 ATAATAGCAA CAGACATACA AACTAAAGAA TTACAAAAAC AAATTACAAA
2551 AATTCAAAAT TTTCGGTTTT ATTACAGGGA CAGCAGAGAT CCAGTTTGGT
2601 TAGTACCGGG CCCGCTCTAG TCGAGGTGCA CGGTATCGAT AAGCTCGCTT
2651 CACGAGATTC CAGCAGGTG AGGGACCTAA TAACTTCGTA TAGCATACAT
2701 TATACGAAGT TATATTAAGG GTTCCAAGCT TAAGCGGCCG CTGAAAGACC
2751 CCACCTGTAG GTTTGGCAAG CTAGCTGCAG TAACGCCATT TTGCAAGGCA
2801 TGGAAAAATA CCAAACCAAG AATAGAGAAG TTCAGATCAA GGGCGGGTAC
2851 ATGAAAATAG CTAACGTTGG GCCAAACAGG ATATCTGCGG TGAGCAGTTT
2901 CGGCCCCGGC CCGGGGCCAA GAACAGATGG TCACCGCAGT TTCGGCCCCG
2951 GCCCGAGGCC AAGAACAGAT GGTCCCCAGA TATGGCCCAA CCCTCAGCAG
3001 TTTCTTAAGA CCCATCAGAT GTTTCCAGGC TCCCCAAGG ACCTGAAATG
3051 ACCCTGCGCC TTATTTGAAT TAACCAATCA GCCTGCTTCT CGCTTCTGTT
3101 CGCGCGCTTC TGCTTCCCGA GCTCTATAAA AGAGCTCACA ACCCTCACT
3151 CGGCGCGCCA GTCCTCCGAT TGACTGAGTC GCCCGGATCC CGCCACCATG
3201 GTGAGCAAGG GCGAGGAGCT GTTCACCGGG GTGGTGCCCA TCCTGGTCGA
3251 GCTGGACGGC GACGTAAACG GCCACAAGTT CAGCGTGTCC GGCGAGGGCG
3301 AGGGCGATGC CACCTACGGC AAGCTGACCC TGAAGTTCAT CTGCACCACC
3351 GGCAAGCTGC CCGTGCCCTG GCCCACCTC GTGACCACC TGACCTGGGG
3401 CGTGAGTGC TTCCGCCGCT ACCCCGACCA CATGAAGCAG CACGACTTCT
3451 TCAAGTCCGC CATGCCCCGA GGCTACGTCC AGGAGCGCAC CATCTTCTTC
3501 AAGGACGACG GCAACTACAA GACCCGCGCC GAGGTGAAGT TCGAGGGCGA

```

```

3551 CACCCTGGTG AACCGCATCG AGCTGAAGGG CATCGACTTC AAGGAGGACG
3601 GCAACATCCT GGGGCACAAG CTGGAGTACA ACGCCATCAG CGACAACGTC
3651 TATATCACCG CCGACAAGCA GAAGAACGGC ATCAAGGCCA ACTTCAAGAT
3701 CCGCCACAAC ATCGAGGACG GCAGCGTGCA GCTCGCCGAC CACTACCAGC
3751 AGAACACCCC CATCGGCGAC GGCCCCGTGC TGCTGCCCCG CAACCACTAC
3801 CTGAGCACCC AGTCCAAGCT GAGCAAAGAC CCCAACGAGA AGCGCGATCA
3851 CATGGTCCTG CTGGAGTTCG TGACCGCCGC CGGGATCACT CTCGGCATGG
3901 ACGAGCTGTA CAAGTAAGAA TTCGTCGAGG GACCTAATAA CTTCTGATAG
3951 CATACTTAT ACGAAGTTAT ACATGTTTAA GGGTTCCGGT TCCACTAGGT
4001 ACAATTCGAT ATCAAGCTTA TCGATAATCA ACCCTGCGAT TACAAAATTT
4051 GTGAAAGATT GACTGGTATT CTTAACTATG TTGCTCCTTT TACGCTATGT
4101 GGATACGCTG CTTTAATGCC TTTGTATCAT GCTATTGCTT CCCGTATGGC
4151 TTTTATTTTC TCCTCCTTGT ATAAATCCTG GTTGCTGTCT CTTTATGAGG
4201 AGTTGTGGCC CGTTGTCAGG CAACGTGGCG TGGTGTGCAC TGTGTTTGCT
4251 GACGCAACCC CCACTGGTTG GGGCATTGCC ACCACCTGTC AGCTCCTTTC
4301 CGGGACTTTC GCTTTCCCCC TCCCTATTGC CACGGCGGAA CTCATCGCCG
4351 CCTGCCTTGC CCGCTGCTGG ACAGGGGCTC GGCTGTTGGG CACTGACAAT
4401 TCCGTGGTGT TGTCGGGGAA ATCATCGTCC TTTCCCTGGC TGCTCGCCTG
4451 TGTTGCCACC TGGATTCTGC GCGGGACGTC CTTCTGCTAC GTCCCTTCGG
4501 CCCTCAATCC AGCGGACCTT CTTTCCCGCG GCCTGCTGCC GGCTCTGCGG
4551 CCTCTTCCGC GTCTTCGCCT TCGCCCTCAG ACGAGTCGGA TCTCCCTTTG
4601 GGCCGCCTCC CCGCATCGAT ACCGTCGACC TCGATCGAGA CCTAGAAAAA
4651 CATGGAGCAA TCACAAGTAG CAATACAGCA GCTACCAATG CTGATTGTGC
4701 CTGGCTAGAA GCACAAGAGG AGGAGGAGGT GGGTTTTCCA GTCACACCTC
4751 AGGTACCTTT AAGACCAATG ACTTACAAGG CAGCTGTAGA TCTTAGCCAC
4801 TTTTTAAAAG AAAAGGGGGG ACTGGAAGGG CTAATCACT CCCAACGAAG
4851 ACAAGATATC CTTGATCTGT GGATCTACCA CACACAAGGC TACTCCCTG
4901 ATTGGCAGAA CTACACACCA GGGCCAGGGA TCAGATATCC ACTGACCTTT
4951 GGATGGTGCT ACAAGCTAGT ACCAGTTGAG CAAGAGAAGG TAGAAGAAGC
5001 CAATGAAGGA GAGAACACCC GCTTGTTACA CCCTGTGAGC CTGCATGGGA
5051 TGGATGACCC GGAGAGAGAA GTATTAGAGT GGAGGTTTGA CAGCCGCCTA
5101 GCATTTTCATC ACATGGCCCG AGAGCTGCAT CCGGACTGTA CTGGGTCTCT
5151 CTGGTTAGAC CAGATCTGAG CCTGGGAGCT CTCTGGCTAA CTAGGGAACC
5201 CACTGCTTAA GCCTCAATAA AGCTTGCCCT GAGTGCTTCA AGTAGTGTGT
5251 GCCCGTCTGT TGTGTGACTC TGGTAACTAG AGATCCCTCA GACCTTTTTA
5301 GTCAGTGTGG AAAATCTCTA GCAGCATGTG AGCAAAAGGC CAGCAAAAGG
5351 CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC
5401 CCCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA
5451 CCCGACAGGA CTATAAAGAT ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG
5501 TGCGCTCTCC TGTTCCGACC CTGCCGCTTA CCGGATACCT GTCCGCTTTT
5551 CTCCCTTCGG GAAGCGTGGC GCTTTCTCAT AGCTCACGCT GTAGGTATCT
5601 CAGTTCGGTG TAGGTCGTTT GCTCCAAGCT GGGCTGTGTG CACGAACCCC
5651 CCCTTCAGCC CGACCGCTGT GCCTTATCCG GTAACATATG TCTTGAGTCC
5701 AACCCGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTACAG
5751 GATTAGCAGA GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT
5801 GGCCTAACTA CGGCTACACT AGAAGAACAG TATTTGGTAT CTGCGCTCTG
5851 CTGAAGCCAG TTACCTTCGG AAAAAGAGTT GGTAAGTCTT GATCCGGCAA
5901 ACAAACCACC GCTGGTAGCG GTGGTTTTTT TGTGCAAG CAGCAGATTA
5951 CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG
6001 TCTGACGCTC AGTGAACGA AAACACAGT TAAGGGATTT TGGTCATGAG
6051 ATTATCAAAA AGGATCTTCA CCTAGATCCT TTTAAATTAA AAATGAAGTT
6101 TTAAATCAAT CTAAAGTATA TATGAGTAAA CTTGGTCTGA CAGTTACCAA
6151 TGCTTAATCA GTGAGGCACC TATCTCAGCG ATCTGTCTAT TTCGTTTCATC
6201 CATAGTTGCC TGACTCCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT
6251 TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACC
6301 GCTCCAGATT TATCAGCAAT AAACAGCCA GCCGGAAGGG CCGAGCGCAG
6351 AAGTGGTCCT GCAACTTTAT CCGCTCCAT CCAGTCTATT AATTGTTGCC
6401 GGGAAGCTAG AGTAAGTAGT TCGCCAGTTA ATAGTTTGCG CAACGTTGTT
6451 GCCATTGCTA CAGGCATCGT GGTGTCACGC TCGTCGTTTG GTATGGCTTC
6501 ATTCAGCTCC GGTTCCTAAC GATCAAGGCG AGTTACATGA TCCCCATGT
6551 TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC CTCCGATCGT TGTGAGAAGT
6601 AAGTTGGCCG CAGTGTTATC ACTCATGGTT ATGGCAGCAC TGCATAATTC
6651 TCTTACTGTC ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT

```

```

6701 CAACCAAGTC ATTCTGAGAA TAGTGTATGC GGCGACCGAG TTGCTCTTGC
6751 CCGGCGTCAA TACGGGATAA TACCGCGCCA CATAGCAGAA CTTTAAAAAGT
6801 GCTCATCATT GGAAAACGTT CTTTCGGGGCG AAAACTCTCA AGGATCTTAC
6851 CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT
6901 TCAGCATCTT TTACTTTCAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG
6951 GCAAAATGCC GCAAAAAGG GAATAAGGC GACACGGAAA TGTTGAATAC
7001 TCATACTCTT CCTTTTCAA TATTATTGAA GCATTATCA GGGTTATTGT
7051 CTCATGAGCG GATACATATT TGAATGTATT TAGAAAAATA AACAAATAGG
7101 GGTTCGCGC ACATTTCCCC GAAAAGTGCC ACCTGAC

```

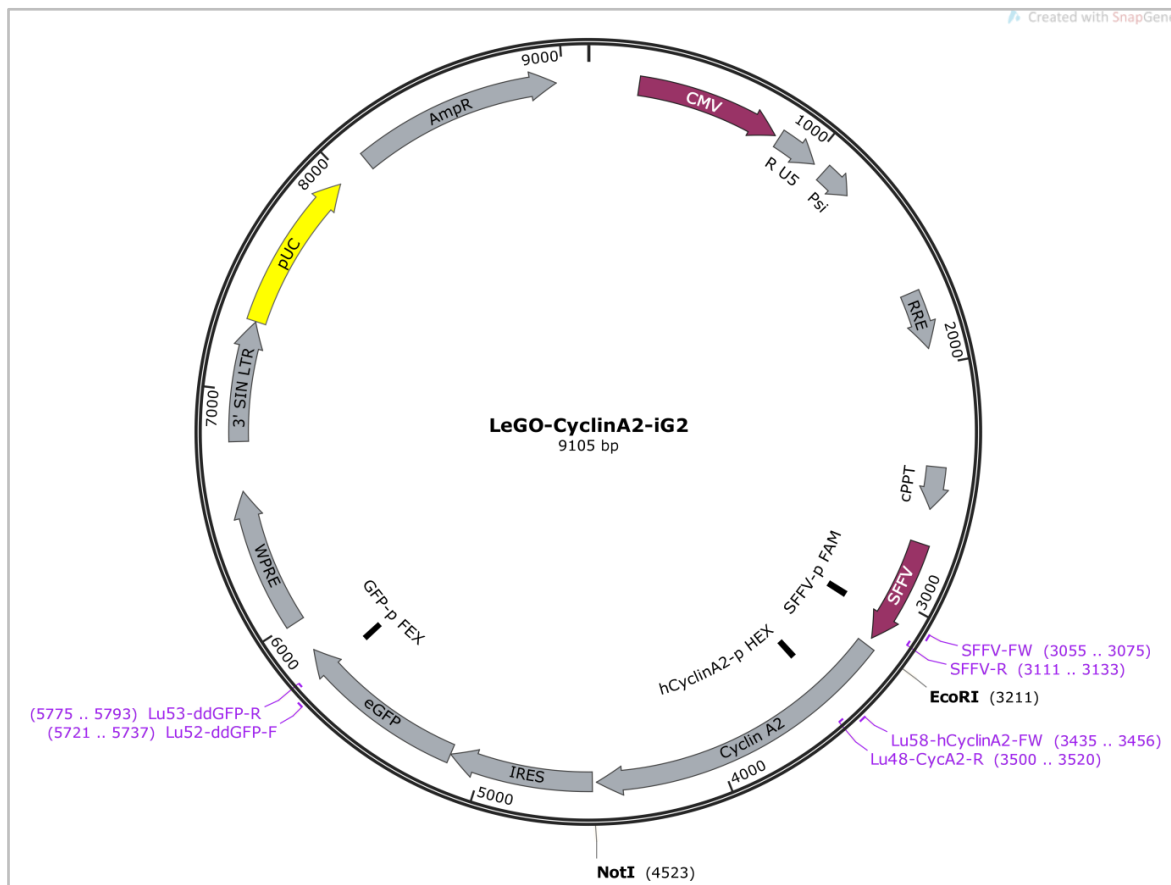


Figure S5 LeGO-CyclinA2-iG2 vector map. 3' SIN LTR, self-inactivating long-terminal repeat; AmpR, ampicillin resistance; CMV, cytomegalus virus promoter; cPPT, central polypurine tract; G2, GFP, green-fluorescent protein; GFP-p HEX, HEX-labeled ddPCR probe for GFP; hCyclinA2-p HEX, HEX-labeled ddPCR probe for CyclinA2; i / IRES, internal-ribosomal entry site; LeGO, lentiviral gene ontology; Lu52-ddGFP-F, forward primer for ddPCR; Lu53-ddGFP-R, reverse primer for ddPCR; Lu48-CycA2-R, reverse primer for ddPCR; Lu58-hCyclinA2-FW, forward primer for ddPCR; psi, packaging signal; pUC, origin of replication; R, redundant sequence; RRE, Rev-responsive element; SFFV, spleen focus-forming virus; SFFV-FW, forward primer for ddPCR; SFFV-p FAM, FAM-labeled ddPCR probe for SFFV; SFFV-R, reverse primer for ddPCR; U5, unique region; WPRE, Woodchuck hepatitis virus post-transcriptional regulatory element.

LeGO-CyclinA2-iG2 vector sequence

```

1  GTCGACGGAT CGGGAGATCT CCCGATCCCC TATGGTGAC TCTCAGTACA
51  ATCTGCTCTG ATGCCGCATA GTTAAGCCAG TATCTGCTCC CTGCTTGTGT
101 GTTGGAGGTC GCTGAGTAGT GCGCGAGCAA AATTTAAGCT ACAACAAGGC
151 AAGGCTTGAC CGACAATTGC ATGAAGAATC TGCTTAGGGT TAGGCGTTTT
201 GCGCTGCTTC GCGATGTACG GGCCAGATAT ACGCGTTGAC ATTGATTATT
251 GACTAGTTAT TAATAGTAAT CAATTACGGG GTCATTAGTT CATAGCCCAT
301 ATATGGAGTT CCGCGTTACA TAACTTACGG TAAATGGCCC GCCTGGCTGA
351 CCGCCCAACG ACCCCCGCCC ATTGACGTCA ATAATGACGT ATGTTCCCAT

```

```

401 AGTAACGCCA ATAGGGACTT TCCATTGACG TCAATGGGTG GAGTATTTAC
451 GGTAAACTGC CCACTTGGCA GTACATCAAG TGTATCATAT GCCAAGTACG
501 CCCCTATTG ACGTCAATGA CGGTAAATGG CCCGCTGGC ATTATGCCCCA
551 GTACATGACC TTATGGGACT TTCTACTTGG GCAGTACATC TACGTATTAG
601 TCATCGCTAT TACCATGGTG ATGCGGTTTT GGCAGTACAT CAATGGGCGT
651 GGATAGCGGT TTGACTCACG GGGATTTCCA AGTCTCCACC CCATTGACGT
701 CAATGGGAGT TTGTTTTGGC ACCAAAATCA ACGGGACTTT CCAAAAATGTC
751 GTAACAACTC CGCCCCATTG ACGCAAATGG GCGGTAGGCG TGTACGGTGG
801 GAGGTCTATA TAAGCAGCGC GTTTTGCCCTG TACTGGGTCT CTCTGGTTAG
851 ACCAGATCTG AGCCTGGGAG CTCTCTGGCT AACTAGGGAA CCCACTGCCTT
901 AAGCCTCAAT AAAGCTTGCC TTGAGTGCTT CAAGTAGTGT GTGCCCGTCT
951 GTTGTGTGAC TCTGGTAACT AGAGATCCCT CAGACCCCTT TAGTCAGTGT
1001 GGAAAATCTC TAGCAGTGGC GCCCGAACAG GGACTTGAAA GCGAAAGGGA
1051 AACCAGAGGA GCTCTCTCGA CGCAGGACTC GGCTTGCTGA AGCGCGCACG
1101 GCAAGAGGCG AGGGGCGGCG ACTGGTGAGT ACGCCAAAAA TTTTGACTAG
1151 CGGAGGCTAG AAGGAGAGAG ATGGGTGCGA GAGCGTCAGT ATTAAGCGGG
1201 GGAGAATTAG ATCGCGATGG GAAAAAATTC GGTAAAGGCC AGGGGGAAAG
1251 AAAAAATATA AATTAACAACA TATAGTATGG GCAAGCAGGG AGCTAGAACG
1301 ATTCGCAGTT AATCCTGGCC TGTAGAAAC ATCAGAAGGC TGTAGACAAA
1351 TACTGGGACA GCTACAACCA TCCCTTCAGA CAGGATCAGA AGAACTTAGA
1401 TCATTATATA ATACAGTAGC AACCTCTAT TGTGTGCATC AAAGGATAGA
1451 GATAAAGAC ACCAAGGAAG CTTTAGACAA GATAGAGGAA GAGCAAAACA
1501 AAAGTAAGAC CACCGCACAG CAAGCGGCCG GCCGCGCTGA TCTTCAGACC
1551 TGGAGGAGGA GATATGAGGG ACAATTGGAG AAGTGAATTA TATAAATATA
1601 AAGTAGTAAA AATTGAACCA TTAGGAGTAG CACCCACCAA GGCAAAGAGA
1651 AGAGTGGTGC AGAGAGAAAA AAGAGCAGTG GGAATAGGAG CTTTGTTCCT
1701 TGGGTTCTTG GGAGCAGCAG GAAGCACTAT GGGCGCAGCG TCAATGACGC
1751 TGACGGTACA GGCCAGACAA TTATTGTCTG GTATAGTGCA GCAGCAGAAC
1801 AATTTGCTGA GGGCTATTGA GGCGCAACAG CATCTGTTGC AACTCACAGT
1851 CTGGGGCATC AAGCAGCTCC AGGCAAGAAT CCTGGCTGTG GAAAGATACC
1901 TAAAGGATCA ACAGCTCCTG GGGATTTGGG GTTGCTCTGG AAAACTCATT
1951 TGCACCACTG CTGTGCCTTG GAATGCTAGT TGGAGTAATA AATCTCTGGA
2001 ACAGATTTGG AATCACACGA CCTGGATGGA GTGGGACAGA GAAATTAACA
2051 ATTACACAAG CTTAATACAC TCCTTAATTG AAGAATCGCA AAACCAGCAA
2101 GAAAAGAATG AACAAGAATT ATTGGAATTA GATAAATGGG CAAGTTTGTG
2151 GAATTGGTTT AACATAACAA ATTGCTGTG GTATATAAAA TTATTCATAA
2201 TGATAGTAGG AGGCTTGGTA GGTTTAAGAA TAGTTTTTGC TGTACTTTCT
2251 ATAGTGAATA GAGTTAGGCA GGGATATTCA CCATTATCGT TTCAGACCCA
2301 CCTCCCAACC CCGAGGGGAC CCGACAGGCC CGAAGGAATA GAAGAAGAAG
2351 GTGGAGAGAG AGACAGAGAC AGATCCATTC GATTAGTGAA CGGATCGGCA
2401 CTGCGTGCGC CAATTCTGCA GACAAATGGC AGTATTCATC CACAATTTTA
2451 AAAGAAAAGG GGGGATTGGG GGGTACAGTG CAGGGGAAAG AATAGTAGAC
2501 ATAATAGCAA CAGACATACA AACTAAAGAA TTACAAAAAC AAATTACAAA
2551 AATTCAAAAT TTTCGGTTTT ATTACAGGGA CAGCAGAGAT CCAGTTTGGT
2601 TAGTACCGGG CCCGCTCTAG TCGAGGTGCA CGGTATCGAT AAGCTCGCTT
2651 CACGAGATTC CAGCAGGTCG AGGGACCTAA TAACTTCGTA TAGCATACAT
2701 TATACGAAGT TATATTAAGG GTTCCAAGCT TAAGCGGCCG GCCGCTGAAA
2751 GACCCACCT GTAGGTTTGG CAAGCTAGCT GCAGTAACGC CATTTTGCAA
2801 GGCATGGAAA AATACCAAAC CAAGAATAGA GAAGTTCAGA TCAAGGGCGG
2851 GTACATGAAA ATAGCTAACG TTGGGCCAAA CAGGATATCT GCGGTGAGCA
2901 GTTTCGGCCC CGGCCCGGG CCAAGAACAG ATGGTCACCG CAGTTTCGGC
2951 CCCGGCCCGA GGCCAAGAAC AGATGGTCCC CAGATATGGC CCAACCTCA
3001 GCAGTTTCTT AAGACCCATC AGATGTTTCC AGGCTCCCCC AAGGACCTGA
3051 AATGACCCTG CGCCTTATTT GAATTAACCA ATCAGCCTGC TTCTCGCTTC
3101 TGTTTCGCGC CTTCTGCTTC CCGAGCTCTA TAAAAGAGCT CACAACCCCT
3151 CACTCGGCGC GCCAGTCTC CGATTGACTG AGTCGCCCCG ATCCCAGTGT
3201 GGTGGTACGG GAATTGCGCA CCATGTTGGG CAACTCTGCG CCGGGGCCTG
3251 CGACCCGCGA GCGGGGCTCG GCGCTGCTAG CATTCAGCA GACGGCGCTC
3301 CAAGAGGACC AGGAGAATAT CAACCCGGA AAGGCAGCGC CCGTCCAACA
3351 ACCCGGACC CGGCCGCGC TGGCGGTACT GAAGTCCGGG AACC CGCGG
3401 GTCTAGCGCA GCAGCAGAGG CCGAAGACGA GACGGGTTGC ACCCTTAAG
3451 GATCTTCCTG TAAATGATGA CGATGTCACC GTTCCCTCCTT GGAAAGCAA
3501 CAGTAAACAG CCTGCGTTCA CCATTCATGT GGATGAAGCA GAAAAGAAG

```

```

3551 CTCAGAAGAA GCCAGCTGAA TCTCAAAAAA TAGAGCGTGA AGATGCCCTG
3601 GCTTTTAATT CAGCCATTAG TTTACCTGGA CCCAGAAAAC CATTTGGTCCC
3651 TCTTGATTAT CCAATGGATG GTAGTTTTGA GTCACCACAT ACTATGGACA
3701 TGTCAATTAT ATTAGAAGAT GAAAAGCCAG TGAGTGTTAA TGAAGTACCA
3751 GACTACCATG AGGATATTCA CACATACCTT AGGGAAATGG AGGTAAATG
3801 TAAACCTAAA GTGGGTTACA TGAAGAAACA GCCAGACATC ACTAACAGTA
3851 TGAGAGCTAT CCTCGTGGAC TGGTTAGTTG AAGTAGGAGA AGAATATAAA
3901 CTACAGAATG AGACCCTGCA TTTGGCTGTG AACTACATTG ATAGGTTCCCT
3951 GTCTTCCATG TCAGTGCTGA GAGGAAAAC TACAGTTGTG GGCAGTGCTG
4001 CTATGCTGTT AGCCTCAAAG TTTGAAGAAA TATACCCCCC AGAAGTAGCA
4051 GAGTTTGTGT ACATTACAGA TGATACCTAC ACCAAGAAAC AAGTTCTGAG
4101 AATGGAGCAT CTAGTTTTGA AAGTCCTTAC TTTTGACTTA GCTGCTCCAA
4151 CAGTAAATCA GTTTCTTACC CAATACTTTC TGCATCAGCA GCCTGCAAAC
4201 TGCAAAGTTG AAAGTTTAGC AATGTTTTTG GGAGAATTAA GTTTGATAGA
4251 TGCTGACCCA TACCTCAAGT ATTTGCCATC AGTTATTGCT GGAGCTGCCT
4301 TTCATTTAGC ACTCTACACA GTCACGGGAC AAAGCTGGCC TGAATCATTA
4351 ATACGAAAGA CTGGATATAC CCTGGAAGT CTTAAGCCTT GTCTCATGGA
4401 CCTTCACCAG ACCTACCTCA AAGCACCACA GCATGCACAA CAGTCAATAA
4451 GAGAAAAGTA CAAAAATTCA AAGTATCATG GTGTTTCTCT CCTCAACCCA
4501 CCAGAGACAC TAAATCTGTA AGCGGCCGCT ACGTAAATTC CGCCCCCCCC
4551 CCCCCTCTCC CTCCCCCCCC CCTAACGTTA CTGGCCGAAG CCGCTTGGA
4601 TAAGGCCGGT GTGCGTTTGT CTATATGTTA TTTTCCACCA TATTGCCGTC
4651 TTTTGGCAAT GTGAGGGCCC GGAAACCTGG CCCTGTCTTC TTGACGAGCA
4701 TTCCTAGGGG TCTTTCCCCT CTCGCCAAAG GAATGCAAGG TCTGTTGAAT
4751 GTCGTGAAGG AAGCAGTTCC TCTGGAAGCT TCTTGAAGAC AAACAACGTC
4801 TGTGCGAGC CTTTGCAGGC AGCGGAACCC CCCACCTGGC GGCAGTGCC
4851 TCTGCGGCCA AAAGCCAGCT GTATAAGATA CACCTGCAAA GCGGCGACAA
4901 CCCAGTGCC ACGTTGTGAG TTGGATAGTT GTGGAAGAG TCAATGGCT
4951 CTCCTCAAGC GTATTCAACA AGGGGCTGAA GGATGCCCAG AAGGTACCCC
5001 ATTGTATGGG ATCTGATCTG GGGCCTCGGT GCACATGCTT TACATGTGTT
5051 TAGTCGAGGT TAAAAAACG TCTAGGCCCC CCGAACCACG GGGACGTGGT
5101 TTTCTTTGA AAAACACGAT GATAATATGG CCACAACCAT GGTGAGCAAG
5151 GGCGAGGAGC TGTTACCCGG GGTGGTGCCC ATCTGGTCG AGCTGGACGG
5201 CGACGTAAAC GGCCACAAGT TCAGCGTGTC CGGCGAGGGC GAGGGCGATG
5251 CCACCTACGG CAAGCTGACC CTGAAGTTCA TCTGCACCAC CGGCAAGCTG
5301 CCCGTGCCCT GGCCACCCCT CGTGACCACC CTGACCTACG GCGTGCAGTG
5351 CTTAGCCGC TACCCCGACC ACATGAAGCA GCACGACTTC TTCAAGTCCG
5401 CCATGCCCCG AGGCTACGTC CAGGAGCGCA CCATCTTCTT CAAGGACGAC
5451 GGCAACTACA AGACCCGCGC CGAGGTGAAG TTCGAGGGCG ACACCCTGGT
5501 GAACCGCATC GAGCTGAAGG GCATCGACTT CAAGGAGGAC GGCAACATCC
5551 TGGGGCACAA GCTGGAGTAC AACTACAACA GCCACAACGT CTATATCATG
5601 GCCGACAAGC AGAAGAACGG CATCAAGGTG AACTTCAAGA TCCGCCACAA
5651 CATCGAGGAC GGCAGCGTGC AGCTCGCCGA CCACTACCAG CAGAACACCC
5701 CCATCGGCGA CGGCCCGTGC CTGCTGCCCC ACAACCACTA CCTGAGACCC
5751 CAGTCCGCCC TGAGCAAAGA CCCCACGAG AAGCGCGATC ACATGGTCCT
5801 GCTGGAGTTC GTGACCGCCG CCGGGATCAC TCTCGGCATG GACGAGCTGT
5851 ACAAGTAAAG CGGCCGCGC CCAGCACAGT GGTCGAAATT CGTCGAGGGA
5901 CCTAATAACT TCGTATAGCA TACATTATAC GAAGTTATAC ATGTTTAAGG
5951 GTTCCGGTTC CACTAGGTAC AATTCGATAT CAAGCTTATC GATAATCAAC
6001 CTCTGGATTA CAAAATTTGT GAAAGATTGA CTGGTATTCT TAACATGTT
6051 GCTCCTTTTA CGCTATGTGG ATACGCTGCT TTAATGCCTT TGTATCATGC
6101 TATTGCTTCC CGTATGGCTT TCATTTTCTC CTCCTTGAT AAATCCTGGT
6151 TGCTGTCTCT TTATGAGGAG TTGTGGCCCC TTGTCAGGCA ACGTGGCGTG
6201 GTGTGCACTG TGTTTGCTGA CGCAACCCCC ACTGGTTGGG GCATTGCCAC
6251 CACCTGTCAG CTCCTTTCCG GGACTTTCGC TTTCCCCCTC CCTATTGCCA
6301 CGGCGGAAC CATCGCCGCC TGCCTTGCCC GCTGCTGGAC AGGGGCTCGG
6351 CTGTTGGGCA CTGACAATTC CGTGGTGTTG TCGGGGAAAT CATCGTCCTT
6401 TCCTTGGCTG CTCGCTGTG TTGCCACCTG GATTCTGCGC GGGACGTCCCT
6451 TCTGCTACGT CCCTTCGGCC CTCAATCCAG CGGACCTTCC TTCCCGCGGC
6501 CTGCTGCCGG CTCTGCGGCC TCTTCCGCGT CTTGCGCTTC GCCCTCAGAC
6551 GAGTCGGATC TCCCTTTGGG CCGCCTCCCC GCATCGATAC CGTCGACCTC
6601 GATCGAGACC TAGAAAAACA TGGAGCAATC ACAAGTAGCA ATACAGCAGC
6651 TACCAATGCT GATTGTGCCT GGCTAGAAGC ACAAGAGGAG GAGGAGGTGG

```

```

6701 GTTTTCCAGT CACACCTCAG GTACCTTTAA GACCAATGAC TTACAAGGCA
6751 GCTGTAGATC TTAGCCACTT TTTAAAAGAA AAGGGGGGAC TGGAAGGGCT
6801 AATTCACCTC CAACGAAGAC AAGATATCCT TGATCTGTGG ATCTACCACA
6851 CACAAGGCTA CTTCCCTGAT TGGCAGAACT ACACACCAGG GCCAGGGATC
6901 AGATATCCAC TGACCTTTGG ATGGTGCTAC AAGCTAGTAC CAGTTGAGCA
6951 AGAGAAGGTA GAAGAAGCCA ATGAAGGAGA GAACACCCGC TTGTTACACC
7001 CTGTGAGCCT GCATGGGATG GATGACCCGG AGAGAGAAGT ATTAGAGTGG
7051 AGGTTTGACA GCCGCCTAGC ATTTTCATCAC ATGGCCCGAG AGCTGCATCC
7101 GGACTGTACT GGGTCTCTCT GGTAGACCA GATCTGAGCC TGGGAGCTCT
7151 CTGGCTAACT AGGGAACCCA CTGCTTAAGC CTCAATAAAG CTTGCCCTGA
7201 GTGCTTCAAG TAGTGTGTGC CCGTCTGTTG TGTGACTCTG GTAAC TAGAG
7251 ATCCCTCAGA CCCTTTTAGT CAGTGTGGAA AATCTCTAGC AGCATGTGAG
7301 CAAAAGGCCA GCAAAAGGCC AGGAACCGTA AAAAGGCCGC GTTGCTGGCG
7351 TTTTTCATA GGCTCCGCCC CCCTGACGAG CATCACAAAA ATCGACGCTC
7401 AAGTCAGAGG TGGCGAAACC CGACAGGACT ATAAAGATAC CAGGCGTTTC
7451 CCCCTGGAAG CTCCCTCGTG CGCTCTCCTG TTCCGACCCT GCCGCTTACC
7501 GGATACCTGT CCGCCTTTCT CCCTTCGGGA AGCGTGGCGC TTTCTCATAG
7551 CTCACGCTGT AGGTATCTCA GTTCGGTGTA GGTCTTCGC TCCAAGCTGG
7601 GCTGTGTGCA CGAACCCCCC GTTCAGCCCG ACCGCTGCGC CTTATCCGGT
7651 AACTATCGTC TTGAGTCCAA CCCGGTAAGA CACGACTTAT CGCCACTGGC
7701 AGCAGCCACT GGTAACAGGA TTAGCAGAGC GAGGTATGTA GGCGGTGCTA
7751 CAGAGTTCTT GAAGTGGTGG CCTAACTACG GCTACACTAG AAGAACAGTA
7801 TTTGGTATCT GCGCTCTGCT GAAGCCAGTT ACCTTCGGAA AAAGAGTTGG
7851 TAGCTCTTGA TCCGGCAAAC AAACCACCGC TGGTAGCGGT GGTTTTTTTG
7901 TTTGCAAGCA GCAGATTACG CGCAGAAAAA AAGGATCTCA AGAAGATCCT
7951 TTTGCTTTT CTACGGGGTC TGACGCTCAG TGGAACGAAA ACTCAGTTA
8001 AGGGATTTTG GTCATGAGAT TATCAAAAAG GATCTTCACC TAGATCCTTT
8051 TAAATTAAAA ATGAAGTTTT AAATCAATCT AAAGTATATA TGAGTAAACT
8101 TGGTCTGACA GTTACCAATG CTTAATCAGT GAGGCACCTA TCTCAGCGAT
8151 CTGTCTATTT CGTTCATCCA TAGTTGCCTG ACTCCCCGTC GTGTAGATAA
8201 CTACGATACG GGAGGGCTTA CCATCTGGCC CCAGTGCTGC AATGATACCG
8251 CGAGACCCAC GCTCACCGGC TCCAGATTTA TCAGCAATAA ACCAGCCAGC
8301 CGGAAGGGCC GAGCGCAGAA GTGGTCCTGC AACTTTATCC GCCTCCATCC
8351 AGTCTATTAA TTGTTGCCGG GAAGCTAGAG TAAGTAGTTC GCCAGTTAAT
8401 AGTTTGCGCA ACGTTGTTGC CATTGCTACA GGCATCGTGG TGTACGCTC
8451 GTCGTTTGGT ATGGCTTCAT TCAGCTCCGG TTCCCAACGA TCAAGGCGAG
8501 TTACATGATC CCCCATGTTG TGCAAAAAAG CGGTTAGCTC CTTCGGTCCT
8551 CCGATCGTTG TCAGAAGTAA GTTGGCCGCA GTGTTATCAC TCATGGTTAT
8601 GGCAGCACTG CATAATTCTC TTAGTGTGAT GCCATCCGTA AGATGCTTTT
8651 CTGTGACTGG TGAGTACTCA ACCAAGTCAT TCTGAGAATA GTGTATGCGG
8701 CGACCGAGTT GCTCTTGCCC GCGTCAATA CGGGATAATA CCGCGCCACA
8751 TAGCAGAACT TTAAAAGTGC TCATCATTGG AAAACGTTCT TCGGGGCGAA
8801 AACTCTCAAG GATCTTACCG CTGTTGAGAT CCAGTTCGAT GTAACCCACT
8851 CGTGCACCCA ACTGATCTTC AGCATCTTTT ACTTTCACCA GCGTTTCTGG
8901 GTGAGCAAAA ACAGGAAGGC AAAATGCCGC AAAAAAGGGA ATAAGGGCGA
8951 CACGGAATG TTGAATACTC ATACTCTTCC TTTTTCATAA TTATTGAAGC
9001 ATTTATCAGG GTTATTGTCT CATGAGCGGA TACATATTTG AATGTATTTA
9051 GAAAAATAAA CAAATAGGGG TTCCGCGCAC ATTTCCCCGA AAAGTGCCAC
9101 CTGAC

```

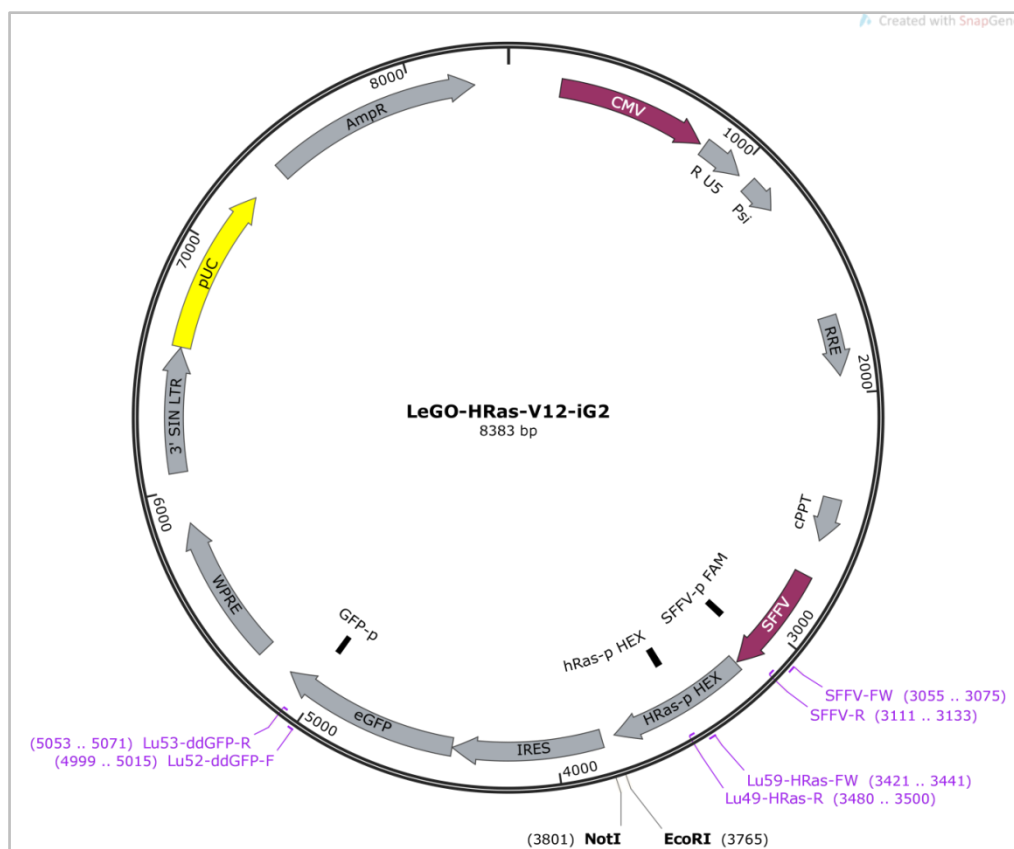



Figure S6 LeGO-HRas-V12-iG2 vector map. 3' SIN LTR, self-inactivating long-terminal repeat; AmpR, ampicillin resistance; CMV, cytomegalus virus promoter; cPPT, central polypurine tract; G2, GFP, green-fluorescent protein; GFP-p HEX, HEX-labeled ddPCR probe for GFP; hRAS-p HEX, HEX-labeled ddPCR probe for HRas-V12; i / IRES, internal-ribosomal entry site; LeGO, lentiviral gene ontology; Lu52-ddGFP-F, forward primer for ddPCR; Lu53-ddGFP-R, reverse primer for ddPCR; Lu49-HRas-R, reverse primer for ddPCR; Lu59-HRas-FW, forward primer for ddPCR; psi, packaging signal; pUC, origin of replication; R, redundant sequence; RRE: Rev-responsive element; SFFV, spleen focus-forming virus; SFFV-FW, forward primer for ddPCR; SFFV-p FAM, FAM-labeled ddPCR probe for SFFV; SFFV-R, reverse primer for ddPCR; U5, unique region; WPRE, Woodchuck hepatitis virus post-transcriptional regulatory element.

LeGO-HRas-V12-iG2 vector sequence

```

1  GTGCACGGAT  CGGGAGATCT  CCCGATCCCC  TATGGTGCAC  TCTCAGTACA
51  ATCTGCTCTG  ATGCCGCATA  GTTAAGCCAG  TATCTGCTCC  CTGCTTGTGT
101 GTTGGAGGTC  GCTGAGTAGT  GCGCGAGCAA  AATTTAAGCT  ACAACAAGGC
151 AAGGCTTGAC  CGACAATTGC  ATGAAGAATC  TGCTTAGGGT  TAGGCGTTTT
201 GCGCTGCTTC  GCGATGTACG  GGCCAGATAT  ACGCGTTGAC  ATTGATTATT
251 GACTAGTTAT  TAATAGTAAT  CAATTACGGG  GTCATTAGTT  CATAGCCCAT
301 ATATGGAGTT  CCGCGTTACA  TAACCTACGG  TAAATGGCCC  GCCTGGCTGA
351 CCGCCCAACG  ACCCCCACCC  ATTGACGTCA  ATAATGACGT  ATGTTCCCAT
401 AGTAACGCCA  ATAGGGACTT  TCCATTGACG  TCAATGGGTG  GAGTATTTAC
451 GGTAAGCTGC  CCACTTGGCA  GTACATCAAG  TGTATCATAT  GCCAAGTACG
501 CCCCTATTG  ACGTCAATGA  CGGTAAATGG  CCCGCTGGC  ATTATGCCCC
551 GTACATGACC  TTATGGGACT  TTCCTACTTG  GCAGTACATC  TACGTATTAG
601 TCATCGCTAT  TACCATGGTG  ATGCGGTTTT  GGCAGTACAT  CAATGGGCGT
651 GGATACGGT  TTGACTCACG  GGGATTTCCT  AGTCTCCACC  CCATTGACGT
701 CAATGGGAGT  TTGTTTTGGC  ACCAAAATCA  ACGGGACTTT  CCAAAATGTC
751 GTAACAACCT  CGCCCCATTG  ACGCAAATGG  GCGGTAGGCG  TGTACGGTGG
801 GAGGTCTATA  TAAGCAGCGC  GTTTTGCCTG  TACTGGGTCT  CTCTGGTTAG
851 ACCAGATCTG  AGCCTGGGAG  CTCTCTGGCT  AACTAGGGAA  CCCACTGCTT
901 AAGCCTCAAT  AAAGCTTGCC  TTGAGTGCTT  CAAGTAGTGT  GTGCCCGTCT
951 GTTGTGTGAC  TCTGGTAACT  AGAGATCCCT  CAGACCCTTT  TAGTCAGTGT

```

```

1001 GGAAAATCTC TAGCAGTGGC GCCCGAACAG GGACTTGAAA GCGAAAGGGA
1051 AACCAGAGGA GCTCTCTCGA CGCAGGACTC GGCTTGCTGA AGCGCGCACG
1101 GCAAGAGGCG AGGGGCGGCG ACTGGTGAGT ACGCCAAAAA TTTTGACTAG
1151 CGGAGGCTAG AAGGAGAGAG ATGGGTGCGA GAGCGTCAGT ATTAAGCGGG
1201 GGAGAATTAG ATCGCGATGG GAAAAAATTC GGTTAAGGCC AGGGGGAAAAG
1251 AAAAAATATA AATTA AAAACA TATAGTATGG GCAAGCAGGG AGCTAGAACG
1301 ATTCGCAGTT AATCCTGGCC TGTTAGAAAC ATCAGAAGGC TGTAGACAAA
1351 TACTGGGACA GCTACAACCA TCCCTTCAGA CAGGATCAGA AGAACTTAGA
1401 TCATTATATA ATACAGTAGC AACCTCTAT TGTGTGCATC AAAGGATAGA
1451 GATAAAAGAC ACCAAGGAAG CTTTAGACAA GATAGAGGAA GAGCAAAACA
1501 AAAGTAAGAC CACCGCACAG CAAGCGGCCG GCCGCGCTGA TCTTCAGACC
1551 TGGAGGAGGA GATATGAGGG ACAATTGGAG AAGTGAATTA TATAAATATA
1601 AAGTAGTAAA AATTGAACCA TTAGGAGTAG CACCCACCAA GGCAAAGAGA
1651 AGAGTGGTGC AGAGAGAAAA AAGAGCAGTG GGAATAGGAG CTTTGTTCCT
1701 TGGGTTCTTG GGAGCAGCAG GAAGCACTAT GGGCGCAGCG TCAATGACGC
1751 TGACGGTACA GGCCAGACAA TTATTGTCTG GTATAGTGCA GCAGCAGAAC
1801 AATTTGCTGA GGGCTATTGA GCGCAACAG CATCTGTTGC AACTCACAGT
1851 CTGGGGCATC AAGCAGCTCC AGGCAAGAAT CCTGGCTGTG GAAAGATACC
1901 TAAAGGATCA ACAGCTCCTG GGGATTTGGG GTTGCTCTGG AAAACTCATT
1951 TGCACCACTG CTGTGCCTTG GAATGCTAGT TGGAGTAATA AATCTCTGGA
2001 ACAGATTTGG AATCACACGA CCTGGATGGA GTGGGACAGA GAAATTAACA
2051 ATTACACAAG CTTAATACAC TCCTTAATTG AAGAATCGCA AAACCAGCAA
2101 GAAAAGAATG AACAAGAATT ATTGGAATTA GATAAATGGG CAAGTTTGTG
2151 GAATTGGTTT AACATAACAA ATTTGGCTGTG GTATATAAAA TTATTCATAA
2201 TGATAGTAGG AGGCTTGGTA GGTTTAAGAA TAGTTTTTGC TGTACTTTCT
2251 ATAGTGAATA GAGTTAGGCA GGGATATTCA CCATTATCGT TTCAGACCCA
2301 CCTCCAACC CCGAGGGGAC CCGACAGGCC CGAAGGAATA CGGATCGGCA
2351 GTGGAGAGAG AGACAGAGAC AGATCCATTC GATTAGTGAA CGGATCGGCA
2401 CTGCGTGCGC CAATTCTGCA GACAAATGGC AGTATTCATC CACAATTTTA
2451 AAAGAAAAGG GGGGATTGGG GGGTACAGTG CAGGGGAAAAG AATAGTAGAC
2501 ATAATAGCAA CAGACATACA AACTAAAGAA TTACAAAAAC AAATTACAAA
2551 AATTCAAAAT TTTCGGGTTT ATTACAGGGA CAGCAGAGAT CCAGTTTGGT
2601 TAGTACCGGG CCCGCTCTAG TCGAGGTCGA CGGTATCGAT AAGCTCGCTT
2651 CACGAGATTC CAGCAGGTCG AGGGACCTAA TAACTTCGTA TAGCATACAT
2701 TATACGAAGT TATATTAAGG GTTCCAAGCT TAAGCGGCCG GCCGCTGAAA
2751 GACCCACCT GTAGGTTTGG CAAGCTAGCT GCAGTAACGC CATTTTGCAA
2801 GGCATGGAAA AATACCAAAC CAAGAATAGA GAAAGTCAGA TCAAGGGCGG
2851 GTACATGAAA ATAGCTAACG TTGGGCCAAA CAGGATATCT GCGGTGAGCA
2901 GTTTCGGCCC CGGCCCGGGG CCAAGAACAG ATGGTCACCG CAGTTTCGGC
2951 CCCGGCCCGA GGCCAAGAAC AGATGGTCCC CAGATATGGC CCAACCTCA
3001 GCAGTTTCTT AAGACCCATC AGATGTTTCC AGGCTCCCCC AAGGACCTGA
3051 AATGACCCTG CGCCTTATTT GAATTAACCA ATCAGCCTGC TTCTCGCTTC
3101 TGTTTCGCGC CTTCTGCTTC CCGAGCTCTA TAAAAGAGCT CACAACCTCT
3151 CACTCGGCGC GCCAGTCCTC CGATTGACTG AGTCGCCCCG ATCCATGACG
3201 GAATATAAGC TGGTGGTGGT GGGCGCCGTC GGTGTGGGCA AGAGTGCGCT
3251 GACCATCCAG CTGATCCAGA ACCATTTTGT GGACGAATAC GACCCCACTA
3301 TAGAGGATTC CTACCGGAAG CAGGTGGTCA TTGATGGGGA GACGTGCCTG
3351 TTGGACATCC TGGATACCGC CGGCCAGGAG GAGTACAGCG CCATGCGGGA
3401 CCAGTACATG CGCACCGGGG AGGGCTTCCT GTGTGTGTTT GCCATCAACA
3451 ACACCAAGTC TTTTGAGGAC ATCCACCAGT ACAGGGAGCA GATCAAACGG
3501 GTGAAGGACT CGGATGACGT GCCCATGGTG CTGGTGGGGA ACAAGTGTGA
3551 CCTGGCTGCA CGCACTGTGG AATCTCGGCA GGCTCAGGAC CTCGCCCCGAA
3601 GCTACGGCAT CCCCTACATC GAGACCTCGG CCAAGACCCG GCAGGGAGTG
3651 GAGGATGCCT TCTACACGTT GGTGCGTGAG ATCCGGCAGC ACAAGCTGCG
3701 GAAGCTGAAC CCTCCTGATG AGAGTGGCCC CGGCTGCATG AGCTGCAAGT
3751 GTGTGCTCTC CTGAGAATTC CTGCAGGCCT CGACGAGGGC CGGCGCGCCG
3801 CGGCCGCTAC GTAAATTCCG CCCCCCCCCC CCCTCTCCCT CCCCCCCCCC
3851 TAACGTTACT GGCCGAAGCC GCTTGAATA AGGCCGGTGT GCGTTTGTCT
3901 ATATGTTATT TTCCACCATA TTGCCGTCTT TTGGCAATGT GAGGGCCCGG
3951 AAACCTGGCC CTGTCTTCTT GACGAGCATT CCTAGGGGTC TTTCCCTCT
4001 CGCCAAAGGA ATGCAAGGTC TGTTGAATGT CGTGAAGGAA CAGTTTCCTC
4051 TGGAAGCTTC TTGAAGACAA ACAACGCTG TAGCGACCTT TTGACGGCAG
4101 CGGAACCCCC CACCTGGCGA CAGGTGCCTC TCGGCCAAA AGCCACGTGT

```

```

4151 ATAAGATACA CCTGCAAAGG CGGCACAACC CCAGTGCCAC GTTGTGAGTT
4201 GGATAGTTGT GGAAAGAGTC AAATGGCTCT CCTCAAGCGT ATTC AACAAAG
4251 GGGCTGAAGG ATGCCCAGAA GGTACCCCAT TGTATGGGAT CTGATCTGGG
4301 GCCTCGGTGC ACATGCTTTA CATGTGTTTA GTCGAGGTTA AAAAAACGTC
4351 TAGGCCCCC GAACCACGGG GACGTGGTTT TCCTTTGAAA AACACGATGA
4401 TAATATGGCC ACAACCATGG TGAGCAAGGG CGAGGAGCTG TTCACCGGGG
4451 TGGTGCCCAT CCTGGTCGAG CTGGACGGCG ACGTAAACGG CCACAAGTTC
4501 AGCGTGTCCG GCGAGGGCGA GGGCGATGCC ACCTACGGCA AGCTGACCCT
4551 GAAGTTCATC TGCACCACCG GCAAGCTGCC CGTGCCCTGG CCCACCTCG
4601 TGACCACCTT GACCTACGGC GTGCAGTGCT TCAGCCGCTA CCCCACCAC
4651 ATGAAGCAGC ACGACTTCTT CAAGTCCGCC ATGCCC GAAG GCTACGTCCA
4701 GGAGCGCACC ATCTTCTTCA AGGACGACGG CAACTACAAG ACCCGCGCCG
4751 AGGTGAAGTT CGAGGGCGAC ACCCTGGTGA ACCGCATCGA GCTGAAGGGC
4801 ATCGACTTCA AGGAGGACGG CAACATCCTG GGGCACAAGC TGGAGTACAA
4851 CTACAACAGC CACAACGTCT ATATCATGGC CGACAAGCAG AAGAACGGCA
4901 TCAAGGTGAA CTTCAAGATC CGCCACAACA TCGAGGACGG CAGCGTGCAG
4951 CTCGCCGACC ACTACCAGCA GAACACCCCC ATCGGCGACG GCCCGTGCT
5001 GCTGCCCCGAC AACCCTACC TGAGCACCCA GTCCGCCCTG AGCAAAGACC
5051 CCAACGAGAA GCGCGATCAC ATGGTCCTGC TGGAGTTCTG GACCGCCGCC
5101 GGGATCACTC TCGGCATGGA CGAGCTGTAC AAGTAAAGCG GCCGCGCCG
5151 AGCACAGTGG TCGAAATTCG TCGAGGGACC TAATAACTTC GTATAGCATA
5201 CATTATACGA AGTTATACAT GTTTAAGGGT TCCGGTTCCA CTAGGTACAA
5251 TTCGATATCA AGCTTATCGA TAATCAACCT CTGGATTACA AAATTTGTGA
5301 AAGATTGACT GGTATTCTTA ACTATGTTGC TCCTTTTACG CTATGTGGAT
5351 ACGTGCTTTT AATGCCTTTT TATCATGCTA TTGCTTCCCG TATGGCTTTC
5401 ATTTTCTCCT CTTGTATATA ATCCTGGTTG CTGTCTCTTT ATGAGGAGTT
5451 GTGGCCCGTT GTCAGGCAAC GTGGCGTGGT GTGCACTGTG TTTGTGACG
5501 CAACCCCCAC TGTTTGGGGC ATTGCCACCA CCTGTCAGCT CTTTCCGGG
5551 ACTTTTCGCTT TCCCCCTCCC TATTGCCACG GCGGAACTCA TCGCCGCCCTG
5601 CTTTGCCCGC TGCTGGACAG GGGCTCGGCT GTTGGGCACT GACAATTCCG
5651 TGGTGTTGTC GGGGAAATCA TCGTCCTTTC CTTGGCTGCT CGCCTGTGTT
5701 GCCACCTGGA TTCTGCGCGG GACGTCCTTC TGCTACGTCC CTTGCGCCCT
5751 CAATCCAGCG GACCTTCCTT CCCGCGGCTT GCTGCCGGCT CTGCGGCCCTC
5801 TTCCGCGTCT TCGCCTTCGC CCTCAGACGA GTCGGATCTC CTTTGGGCC
5851 GCCTCCCCGC ATCGATACCG TCGACCTCGA TCGAGACCTA GAAAACATG
5901 GAGCAATCAC AAGTAGCAAT ACAGCAGCTA CCAATGCTGA TTGTGCCCTG
5951 CTAGAAGCAC AAGAGGAGGA GGAGGTGGGT TTTCCAGTCA CACCTCAGGT
6001 ACCTTTAAGA CCAATGACTT ACAAGGCAGC TGTAATCTT AGCCACTTTT
6051 TAAAAGAAAA GGGGGGACTG GAAGGGCTAA TTCACTCCCA ACGAAGACAA
6101 GATATCCTTG ATCTGTGGAT CTACCACACA CAAGGCTACT TCCCTGATTG
6151 GCAGAACTAC ACACCAGGGC CAGGGATCAG ATATCCACTG ACCTTTGGAT
6201 GGTGCTACAA GCTAGTACCA GTTGAGCAAG AGAAGGTAGA AGAAGCCAAT
6251 GAAGGAGAGA ACACCGCTT GTTACACCTT GTGAGCCTGC ATGGGATGGA
6301 TAGACCGGAG AGAGAAGTAT TAGAGTGGAG GTTTGACAGC CGCCTAGCAT
6351 TTCATCACAT GGCCCGAGAG CTGCATCCGG ACTGTACTGG GTCTCTCTGG
6401 TTAGACCAGA TCTGAGCCTG GGAGCTCTCT GGCTAACTAG GGAACCCACT
6451 GCTTAAGCCT CAATAAAGCT TGCCTTGAGT GCTTCAAGTA GTGTGTGCCC
6501 GTCTGTTGTG TGACTCTGGT AACTAGAGAT CCCTCAGACC CTTTGTAGTCA
6551 GTGTGAAAAA TCTCTAGCAG CATGTGAGCA AAAGGCCAGC AAAAGGCCAG
6601 GAACCGTAAA AAGGCCGCGT TGCTGGCGTT TTTCCATAGG CTCCGCCCCC
6651 CTGACGAGCA TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG
6701 ACAGGACTAT AAAGATACCA GCGCTTTCCC CCTGGAAGCT CCCTCGTGCG
6751 CTCTCCTGTT CCGACCCTGC CGCTTACCGG ATACCTGTCC GCCTTCTCTC
6801 CTTGCGGAAG CGTGGCGCTT TCTCATAGCT CACGCTGTAG GTATCTCAGT
6851 TCGGTGTAGG TCGTTGCTC CAAGCTGGGC TGTGTGCACG AACCCCCCGT
6901 TCAGCCCGAC CGCTGCGCCT TATCCGGTAA CTATCGTCTT GAGTCCAACC
6951 CGGTAAGACA CGACTTATCG CCACTGGCAG CAGCCACTGG TAACAGGATT
7001 AGCAGAGCGA GGTATGTAGG CCGTGCTACA GAGTTCTTGA AGTGGTGGCC
7051 TAACTACGGC TACACTAGAA GAACAGTATT TGGTATCTGC GCTCTGCTGA
7101 AGCCAGTTAC CTTGCGAAAA AGAGTTGGTA GCTCTTGATC CGGCAACAA
7151 ACCACGCTG GTAGCGGTGG TTTTTTTGTT TGCAAGCAGC AGATTACGCG
7201 CAGAAAAAAA GGATCTCAAG AAGATCCTTT GATCTTTTCT ACGGGGTCTG
7251 ACGCTCAGTG GAACGAAAAC TCACGTAAAG GGATTTTGGT CATGAGATTA

```

```

7301 TCAAAAAGGA TCTTCACCTA GATCCTTTTA AATTAAAAAT GAAGTTTAA
7351 ATCAATCTAA AGTATATATG AGTAAACTTG GTCTGACAGT TACCAATGCT
7401 TAATCAGTGA GGCACCTATC TCAGCGATCT GTCTATTTTCG TTCATCCATA
7451 GTTGCCCTGAC TCCCCGTCGT GTAGATAACT ACGATACGGG AGGGCTTACC
7501 ATCTGGCCCC AGTGCTGCAA TGATACCGCG AGACCCACGC TCACCGGCTC
7551 CAGATTTATC AGCAATAAAC CAGCCAGCCG GAAGGGCCGA GCGCAGAAAT
7601 GGTCTGCAA CTTTATCCGC CTCCATCCAG TCTATTAATT GTTGCCGGGA
7651 AGCTAGAGTA AGTAGTTCGC CAGTTAATAG TTTGCGCAAC GTTGTGCGCA
7701 TTGCTACAGG CATCGTGGTG TCACGCTCGT CGTTTGATAT GGCTTCATTC
7751 AGCTCCGGTT CCAACGATC AAGGCGAGTT ACATGATCCC CCATGTTGTG
7801 CAAAAAAGCG GTTAGCTCCT TCGGTCTCC GATCGTTGTC AGAAGTAAGT
7851 TGGCCGAGT GTTATCACTC ATGGTTATGG CAGCACTGCA TAATTCTCTT
7901 ACTGTCATGC CATCCGTAAG ATGCTTTTCT GTGACTGGTG AGTACTCAAC
7951 CAAGTCATTC TGAGAATAGT GTATGCGGCG ACCGAGTTGC TCTTGCCCGG
8001 CGTCAATACG GGATAATACC GCGCCACATA GCAGAACTTT AAAAGTGCTC
8051 ATCATTGGAA AACGTTCTTC GGGGCGAAAA CTCTCAAGGA TCTTACCGCT
8101 GTTGAGATCC AGTTCGATGT AACCCTCTCG TGCACCAAC TGATCTTCAG
8151 CATCTTTTAC TTTCACCAGC GTTCTGCGGT GAGCAAAAAC AGGAAGGCAA
8201 AATGCCGCAA AAAAGGGAAT AAGGGCGACA CGGAAATGTT GAATACTCAT
8251 ACTCTTCCTT TTTCAATATT ATTGAAGCAT TTATCAGGGT TATTGTCTCA
8301 TGAGCGGATA CATATTTGAA TGTATTTAGA AAAATAAACA AATAGGGGTT
8351 CCGCGCACAT TTCCCCGAAA AGTGCCACCT GAC

```

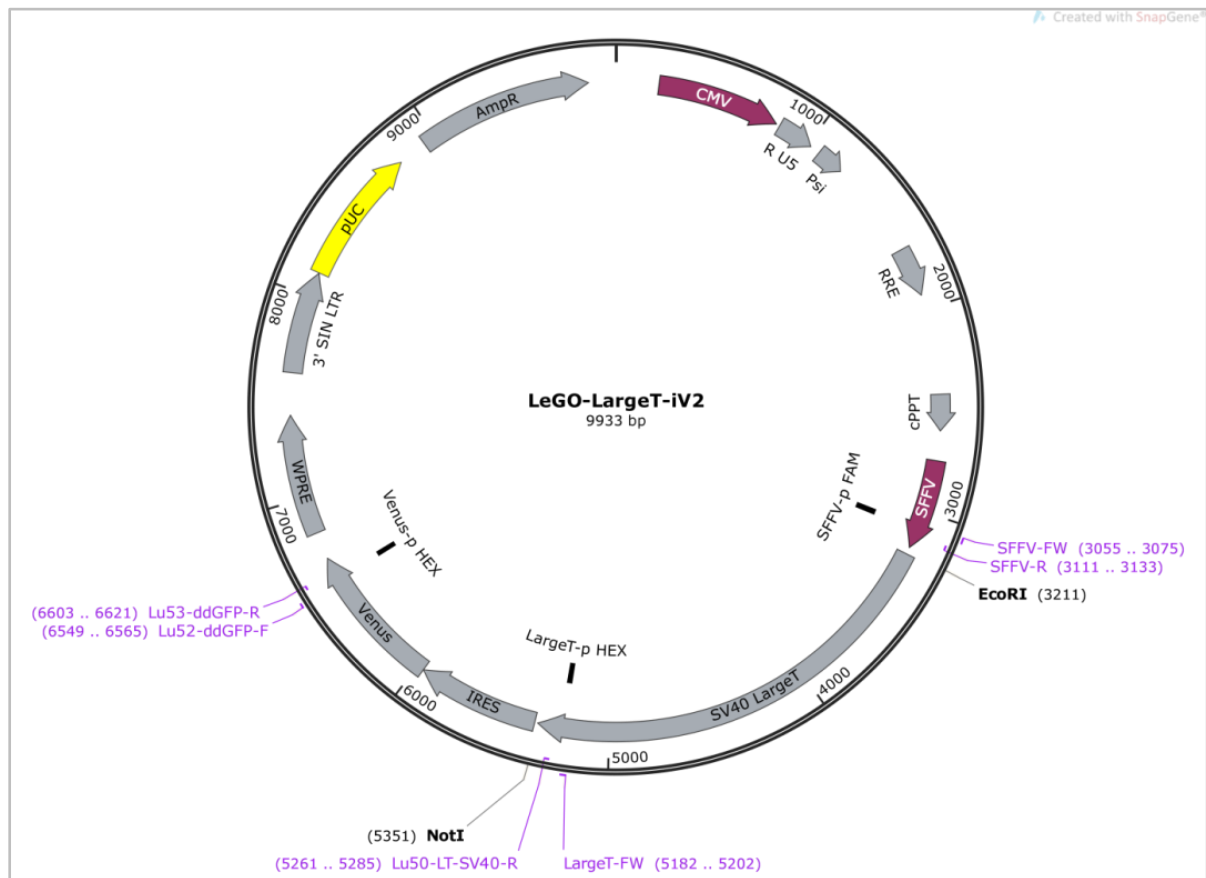


Figure S7 LeGO-LargeT-iv2 vector map. 3' SIN LTR, self-inactivating long-terminal repeat; AmpR, ampicillin resistance; CMV, cytomegalus virus promoter; cPPT, central polypurine tract; i / IRES, internal-ribosomal entry site; LargeT-FW, forward primer for ddPCR; LargeT-p HEX, HEX-labeled ddPCR probe for LargeT; LeGO, lentiviral gene ontology; Lu52-ddGFP-F, forward primer for ddPCR; Lu53-ddGFP-R, reverse primer for ddPCR; Lu50-LT-SV40-R, reverse primer for ddPCR; psi, packaging signal; pUC, origin of replication; R, redundant sequence; RRE, Rev-responsive element; SFFV, spleen focus-forming virus; SFFV-FW, forward primer for ddPCR; SFFV-p FAM, FAM-labeled ddPCR probe for SFFV; SFFV-R, reverse primer for ddPCR; U5, unique region; WPRE, Woodchuck hepatitis virus post-transcriptional regulatory element.

LeGO-LargeT-iV2 vector sequence

```

1  GTCGACGGAT CGGGAGATCT CCCGATCCCC TATGGTGCAC TCTCAGTACA
51 ATCTGCTCTG ATGCCGCATA GTTAAGCCAG TATCTGCTCC CTGCTTGTGT
101 GTTGGAGGTC GCTGAGTAGT GCGCGAGCAA AATTTAAGCT ACAACAAGGC
151 AAGGCTTGAC CGACAATTGC ATGAAGAATC TGCTTAGGGT TAGGCGTTTT
201 GCGCTGCTTC GCGATGTACG GGCCAGATAT ACGCGTTGAC ATTGATTATT
251 GACTAGTTAT TAATAGTAAT CAATTACGGG GTCATTAGTT CATAGCCCAT
301 ATATGGAGTT CCGCGTTACA TAACTTACGG TAAATGGCCC GCCTGGCTGA
351 CCGCCCAACG ACCCCCGCC ATTGACGTCA ATAATGACGT ATGTTCCCAT
401 AGTAACGCCA ATAGGGACTT TCCATTGACG TCAATGGGTG GAGTATTTAC
451 GGTAAACTGC CCACTTGGCA GTACATCAAG TGTATCATAT GCCAAGTACG
501 CCCCTATTG ACGTCAATGA CGGTAAATGG CCCGCTGGC ATTATGCCCA
551 GTACATGACC TTATGGGACT TTCTTACTTG GCAGTACATC TACGTATTAG
601 TCATCGCTAT TACCATGGTG ATGCGGTTTT GGCAGTACAT CAATGGGCGT
651 GGATAGCGGT TTGACTCACG GGGATTTCCA AGTCTCCACC CCATTGACGT
701 CAATGGGAGT TTGTTTTGGC ACCAAAATCA ACGGGACTTT CCAAAATGTC
751 GTAACAACCT CGCCCCATTG ACGCAAATGG GCGGTAGGCG TGTACGGTGG
801 GAGGTCTATA TAAGCAGCGC GTTTTGCCTG TACTGGGTCT CTCTGGTTAG
851 ACCAGATCTG AGCCTGGGAG CTCTCTGGCT AACTAGGGAA CCCACTGCTT
901 AAGCCTCAAT AAAGCTTGCC TTGAGTGCTT CAAGTAGTGT GTGCCCGTCT
951 GTTGTGTGAC TCTGGTAACT AGAGATCCCT CAGACCTTT TAGTCAGTGT
1001 GGAAAATCTC TAGCAGTGGC GCCCGAACAG GGACTTGAAA GCGAAAAGGA
1051 AACCAGAGGA GCTCTCTCGA CGCAGGACTC GGCTTGCTGA AGCGCGCACG
1101 GCAAGAGGCG AGGGGCGGCG ACTGGTGAGT ACGCCAAAAA TTTTGACTAG
1151 CGGAGGCTAG AAGGAGAGAG ATGGGTGCGA GAGCGTCAGT ATTAAGCGGG
1201 GGAGAATTAG ATCGCGATGG GAAAAAATTC GGTAAAGGCC AGGGGGAAAG
1251 AAAAAATATA AATTAACA TATAGTATGG GCAAGCAGGG AGCTAGAACG
1301 ATTCGCAGTT AATCCTGGCC TGTTAGAAAC ATCAGAAGGC TGTAGACAAA
1351 TACTGGGACA GCTACAACCA TCCCTTCAGA CAGGATCAGA AGAAGTTAGA
1401 TCATTATATA ATACAGTAGC AACCCTCTAT TGTGTGCATC AAAGGATAGA
1451 GATAAAAGAC ACCAAGGAAG CTTTAGACAA GATAGAGGAA GAGCAAAACA
1501 AAAGTAAGAC CACCGCACAG CAAGCGGCCG GCCGCGCTGA TCTTCAGACC
1551 TGGAGGAGGA GATATGAGGG ACAATTGGAG AAGTGAATTA TATAAATATA
1601 AAGTAGTAAA AATTGAACCA TTAGGAGTAG CACCCACCAA GGCAAAGAGA
1651 AGAGTGGTGC AGAGAGAAAA AAGAGCAGTG GGAATAGGAG CTTTGTTCCT
1701 TGGGTCTTTC GGAGCAGCAG GAAGCACTAT GGGCGCAGCG TCAATGACGC
1751 TGACGGTACA GGCCAGACAA TTATTGTCTG GTATAGTGCA GCAGCAGAAC
1801 AATTTGCTGA GGGCTATTGA GCGCAACAG CATCTGTTGC AACTCACAGT
1851 CTGGGGCATC AAGCAGCTCC AGGCAAGAAT CCTGGCTGTG GAAAGATACC
1901 TAAAGGATCA ACAGCTCCTG GGGATTTGGG GTTGCTCTGG AAAACTCATT
1951 TGCACCACTG CTGTGCCTTG GAATGCTAGT TGGAGTAATA AATCTCTGGA
2001 ACAGATTTGG AATCACACGA CCTGGATGGA GTGGGACAGA GAAATTAACA
2051 ATTACACAAG CTTAATACAC TCCTTAATTG AAGAATCGCA AAACCAGCAA
2101 GAAAAGAATG AACAAGAATT ATTGGAATTA GATAAATGGG CAAGTTTGTG
2151 GAATTGGTTT AACATAACAA ATTTGGCTGTG GTATATAAAA TTATTCATAA
2201 TGATAGTAGG AGGCTTGCTA GGTTTAAGAA TAGTTTTTGC TGTACTTTCT
2251 ATAGTGAATA GAGTTAGGCA GGGATATTCA CCATTATCGT TTCAGACCCA
2301 CCTCCCAACC CCGAGGGGAC CCGACAGGCC CGAAGGAATA GAAGAAGAAG
2351 GTGGAGAGAG AGACAGAGAC AGATCCATTC GATTAGTGAA CGGATCGGCA
2401 CTGCGTGCGC CAATTCTGCA GACAAATGGC AGTATTCATC CACAATTTTA
2451 AAAGAAAAGG GGGGATTGGG GGGTACAGTG CAGGGGAAAG AATAGTAGAC
2501 ATAATAGCAA CAGACATACA AACTAAAGAA TTACAAAAAC AAATTACAAA
2551 AATTCAAAAT TTTCCGGTTT ATTACAGGGA CAGCAGAGAT CCAGTTTGGT
2601 TAGTACCGGG CCCGCTCTAG TCGAGGTCGA CGGTATCGAT AAGCTCGCTT
2651 CACGAGATTC CAGCAGGTCG AGGGACCTAA TAACTTCGTA TAGCATACAT
2701 TATACGAAGT TATATTAAGG GTTCCAAGCT TAAGCGGCCG GCCGCTGAAA
2751 GACCCACCT GTAGGTTTGG CAAGCTAGCT GCAGTAACGC CATTTTGCAA
2801 GGCATGGAAA AATACCAAAC CAAGAATAGA GAAGTTCAGA TCAAGGGCGG
2851 GTACATGAAA ATAGCTAACG TTGGGCCAAA CAGGATATCT GCGGTGAGCA
2901 GTTTCGGCCC CGGCCCGGG CCAAGAACAG ATGGTCACCG CAGTTTCGGC
2951 CCCGGCCCGA GGCCAAGAAC AGATGGTCCC CAGATATGGC CCAACCTCA
3001 GCAGTTTCTT AAGACCCATC AGATGTTTCC AGGCTCCCCC AAGGACCTGA

```

```

3051 AATGACCTTG CGCCTTATTT GAATTAACCA ATCAGCCTGC TTCTCGCTTC
3101 TGTTTCGCGC CTTCTGCTTC CCGAGCTCTA TAAAAGAGCT CACAACCCCT
3151 CACTCGGCGC GCCAGTCCTC CGATTGACTG AGTCGCCCCG ATCCCAGTGT
3201 GGTGGTACGG GAATTCGCCA CCATGGACAA AGTTTTAAAC AGAGAGGAAT
3251 CTTTGCAGCT AATGGACCTT CTAGGTCTTG AAAGGAGTGC CTGGGGGAAT
3301 ATTCTCTGA TGAGAAAGGC ATATTTAAAA AAATGCAAGG AGTTTCATCC
3351 TGATAAAGGA GGAGATGAAG AAAAAATGAA GAAAATGAAT ACTCTGTACA
3401 AGAAAATGGA AGATGGAGTA AAATATGCTC ATCAACCTGA CTTTGGAGGC
3451 TTCTGGGATG CAACTGAGAT TCCAACCTAT GGAACGTATG AATGGGAGCA
3501 GTGGTGGAAAT GCCTTTAATG AGGAAAACCT GTTTTGCTCA GAAGAAATGC
3551 CATCTAGTGA TGATGAGGCT ACTGCTGACT CTCAACATTC TACTCCTCCA
3601 AAAAAGAAGA GAAAGGTAGA AGACCCCAAG GACTTTCCTT CAGAATTGCT
3651 AAGTTTTTTG AGTCATGCTG TGTTTAGTAA TAGAACTCTT GCTTGCTTTG
3701 CTATTTACAC CACAAAGGAA AAAGCTGCAC TGCTATACAA GAAAATTATG
3751 GAAAAATATT CTGTAACCTT TATAAGTAGG CATAACAGTT ATAATCATAA
3801 CATACTGTTT TTTCTTACTC CACACAGGCA TAGAGTGTCT GCTATTAATA
3851 ACTATGCTCA AAAATTGTGT ACCTTTAGCT TTTTAATTTG TAAAGGGGTT
3901 AATAAGGAAT ATTTGATGTA TAGTGCCCTG ACTAGAGATC CATTTTCTGT
3951 TATTGAGGAA AGTTTGCCAG GTGGGTAAA GGAGCATGAT TTTAATCCAG
4001 AAGAAGCAGA GGAAACTAAA CAAGTGCTCT GGAAGCTTGT AACAGAGTAT
4051 GCAATGGAAA CAAAATGTGA TGATGTGTTG TTATTGCTTG GGATGTACTT
4101 GGAATTTTCA TACAGTTTTG AAATGTGTTT AAAATGTATT AAAAAAGAAC
4151 AGCCCAGCCA CTATAAGTAC CATGAAAAGC ATTATGCAAA TGCTGCTATA
4201 TTTGCTGACA GCAAAAACCA AAAAACATA TGCCAACAGG CTGTTGATAC
4251 TGTTTTAGCT AAAAAGCGGG TTGATAGCCT ACAATTAAC AGAGAACAAA
4301 TGTTAACAAA CAGATTTAAT GATCTTTTGG ATAGGATGGA TATAATGTTT
4351 GTTTCTACAG GCTCTGCTGA CATAGAAGAA TGGATGGCTG GAGTTTCTTG
4401 GCTACACTGT TTGTTGCCCA AAATGGATTC AGTGGTGTAT GACTTTTTAA
4451 AATGCATGGT GTACAACATT CCTAAAAAAA GATACTGGCT GTTTAAAGGA
4501 CCAATTGATA GTGGTAAAC TACATTAGCA GCTGCTTTGC TTGAATTATG
4551 TGGGGGAAA GCTTTAAATG TTAATTTGCC CTTGGACAGG CTGAACTTTG
4601 AGCTAGGAGT AGCTATTGAC CAGTTTTTAG TAGTTTTTGA GGATGTAAAG
4651 GGCAGTGGAG GGGAGTCCAG AGATTTGCC TCCAGTCAGG GAATTAATAA
4701 CCTGGACAAT TTAAGGGATT ATTTGGATGG CAGTGTTAAG GTAACTTAG
4751 AAAAGAAACA CCTAAATAAA AGAACTCAA TATTTCCCC TGGAATAGTC
4801 ACCATGAATG AGTACAGTGT GCCTAAAACA CTGCAGCCA GATTTGTAAA
4851 ACAAATAGAT TTTAGGCCCA AAGATTATTT AAAGCATTGC CTGGAACGCA
4901 GTGAGTTTTT GTTAGAAAAG AGAATAATTC AAAGTGGCAT TGCTTTGCTT
4951 CTTATGTTAA TTTGGTACAG ACCTGTGGCT GAGTTTGCTC AAAGTATTCA
5001 GAGCAGAATT GTGGAGTGA AAGAGAGATT GGACAAAGAG TTTAGTTTGT
5051 CAGTGTATCA AAAAATGAAG TTTAATGTGG CTATGGGAAT TGGAGTTTAA
5101 GATTGGCTAA GAAACAGTGA TGATGATGAT GAAGACAGCC AGGAAAATGC
5151 TGATAAAAAT GAAGATGGTG GGGAGAAGAA CATGGAAGAC TCAGGGCATG
5201 AAACAGGCAT TGATTCACAG TCCCAAGGCT CATTCAGGC CCCTCAGTCC
5251 TCACAGTCTG TTCATGATCA TAATCAGCCA TACCACATTT GTAGAGGTTT
5301 TACTTGCTTT AAAAAACCTC CCACACCTCC CCCTGAACCT GAAACATAAG
5351 CGGCCGCTAC GTAAATTCCG CCCCCCCCCC CCCTCTCCCT CCCCCCCCCC
5401 TAACGTTACT GGCCGAAGCC GCTTGAATA AGGCCGGTGT GCGTTTGTCT
5451 ATATGTTATT TTCCACCATA TTGCCGTCTT TTGGCAATGT GAGGGCCCGG
5501 AAACCTGGCC CTGTCTTCTT GACGAGCATT CCTAGGGGTC TTTCCCTCT
5551 CGCCAAAGGA ATGCAAGGTC TGTTGAATGT CGTGAAGGAA GCAGTTCCTC
5601 TGGAAGCTTC TTGAAGACAA ACAACGTCTG TAGCGACCCT TTGCAGGCAG
5651 CGGAACCCCC CACCTGGCGA CAGGTGCCTC TGCGGCCAAA AGCCACGTGT
5701 ATAAGATACA CCTGCAAAGG CGGCACAACC CCAGTGCCAC GTTGTGAGTT
5751 GGATAGTTGT GGAAAGAGTC AAATGGCTCT CCTCAAGCGT ATTCAACAAG
5801 GGGCTGAAGG ATGCCAGAA GGTACCCCAT TGTATGGGAT CTGATCTGGG
5851 GCCTCGGTGC ACATGCTTTA CATGTGTTTA GTCGAGGTTA AAAAAACGTC
5901 TAGGCCCCC GAACCACGGG GACGTGGTTT TCCTTTGAAA AACACGATGA
5951 TAATATGGCC ACAACCATGG TGAGCAAGGG CGAGGAGCTG TTCACCGGGG
6001 TGGTGCCCAT CCTGGTCGAG CTGGACGGCG ACGTAAACGG CCACAAGTTC
6051 AGCGTGTCCG GCGAGGGCGA GGGCGATGCC ACCTACGGCA AGCTGACCTT
6101 GAAGCTGATC TGCACCACCG GCAAGCTGCC CGTGCCCTGG CCCACCTCG
6151 TGACCACCTT GGGCTACGGC CTGCAGTGCT TCGCCCGCTA CCCCAGCCAC

```

```

6201 ATGAAGCAGC ACGACTTCTT CAAGTCCGCC ATGCCCCAAG GCTACGTCCA
6251 GGAGCGCACC ATCTTCTTCA AGGACGACGG CAACTACAAG ACCCGCGCCG
6301 AGGTGAAGTT CGAGGGCGAC ACCCTGGTGA ACCGCATCGA GCTGAAGGGC
6351 ATCGACTTCA AGGAGGACGG CAACATCCTG GGGCACAAAG TGGAGTACAA
6401 CTACAACAGC CACAACGTCT ATATCACCGC CGACAAGCAG AAGAACGGCA
6451 TCAAGGCCAA CTTCAAGATC CGCCACAACA TCGAGGACGG CGGCGTGCAG
6501 CTCGCCGACC ACTACCAGCA GAACACCCCC ATCGGCGACG GCCCCGTGCT
6551 GCTGCCCCGAC AACCCTACC TGAGCTACCA GTCCGCCCTG AGCAAAGACC
6601 CCAACGAGAA GCGCGATCAC ATGGTCCTGC TGGAGTTCGT GACCGCCGCC
6651 GGGATCACTC TCGGCATGGA CGAGCTGTAC AAGTAAAGCG GCCGCGCCGC
6701 AGCACAGTGG TCGAAATTCG TCGAGGGACC TAATAACTTC GTATAGCATA
6751 CATTATACGA AGTTATACAT GTTTAAGGGT TCCGGTTCCA CTAGGTACAA
6801 TTCGATATCA AGCTTATCGA TAATCAACCT CTGGATTACA AAATTTGTGA
6851 AAGATTGACT GGTATTCTTA ACTATGTTGC TCCTTTTACG CTATGTGGAT
6901 ACGCTGCTTT AATGCCTTTG TATCATGCTA TTGCTTCCCG TATGGCTTTC
6951 ATTTTCTCCT CTTGTATAA ATCCTGGTTG CTGTCTCTTT ATGAGGAGTT
7001 GTGGCCCGTT GTCAGGCAAC GTGGCGTGGT GTGACTGTG TTTGCTGACG
7051 CAACCCCCAC TGGTTGGGGC ATTGCCACCA CCTGTCAGCT CCTTCCGGG
7101 ACTTTCGCTT TCCCCCTCCC TATTGCCACG GCGGAAGTCA TCGCCGCTG
7151 CTTTGCCCGC TGCTGGACAG GGGCTCGGCT GTTGGGCACT GACAATTCCG
7201 TGGTGTGTG GGGGAAATCA TCGTCCTTTC CTTGGCTGCT CGCCTGTGTT
7251 GCCACCTGGA TTCTGCGCGG GACGTCTTTC TGCTACGTCC CTTGCGCCCT
7301 CAATCCAGCG GACCTTCTTT CCCGCGGCTT GCTGCCGGCT CTGCGGCCCT
7351 TTCCGCGTCT TCGCCTTCGC CCTCAGACGA GTCGGATCTC CCTTTGGGCC
7401 GCCTCCCCGC ATCGATAACG TCGACCTCGA TCGAGACCTA GAAAAACATG
7451 GAGTATCAC AAGTAGCAAT ACAGCAGCTA CCAATGCTGA TTGTGCCTGG
7501 CTAGAAGCAC AAGAGGAGGA GGAGGTGGGT TTTCCAGTCA CACCTCAGGT
7551 ACCTTTAAGA CCAATGACTT ACAAGGCAGC TGTAGATCTT AGCCACTTTT
7601 TAAAAGAAAA GGGGGGACTG GAAGGGCTAA TTCACTCCCA ACGAAGACAA
7651 GATATCCTTG ATCTGTGGAT CTACCACACA CAAGGCTACT TCCCTGATTG
7701 GCAGAACTAC ACACCAGGGC CAGGGATCAG ATATCCACTG ACCTTTGGAT
7751 GGTGCTACAA GCTAGTACCA GTTGAGCAAG AGAAGGTAGA AGAAGCCAAT
7801 GAAGGAGAGA ACACCCGCTT GTTACACCTT GTGAGCCTGC ATGGGATGGA
7851 TGACCCGGAG AGAGAAGTAT TAGAGTGGAG GTTTGACAGC CGCCTAGCAT
7901 TTCATCACAT GGCCCGAGAG CTGCATCCGG ACTGTACTGG GTCTCTCTGG
7951 TTAGACCAGA TCTGAGCCTG GGAGCTCTCT GGCTAACTAG GGAACCCACT
8001 GCTTAAGCCT CAATAAAGCT TGCCTTGAGT GCTTCAAGTA GTGTGTGCCC
8051 GTCTGTTGTG TGACTCTGGT AACTAGAGAT CCCTCAGACC CTTTGTAGTCA
8101 GTGTGGAAAA TCTCTAGCAG CATGTGAGCA AAAGGCCAGC AAAAGGCCAG
8151 GAACCGTAAA AAGGCCGCGT TGCTGGCGTT TTTCCATAGG CTCCGCCCCC
8201 CTGACGAGCA TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG
8251 ACAGGACTAT AAAGATACCA GCGCTTCCC CCTGGAAGCT CCTCGTGCG
8301 CTCTCCTGTT CCGACCTGCG CGCTTACCGG ATACCTGTCC GCCTTCTCC
8351 CTTGCGGAAG CGTGGCGCTT TCTCATAGCT CACGCTGTAG GTATCTCAGT
8401 TCGGTGTAGG TCGTTCGCTC CAAGCTGGGC TGTGTGCACG AACCCCGCT
8451 TCAGCCCGAC CGCTGCGCCT TATCCGGTAA CTATCGTCTT GAGTCCAACC
8501 CGGTAAGACA CGACTTATCG CCACTGGCAG CAGCCACTGG TAACAGGATT
8551 AGCAGAGCGA GGTATGTAGG CGGTGCTACA GAGTTCTTGA AGTGGTGGCC
8601 TAACTACGGC TACACTAGAA GAACAGTATT TGGTATCTGC GCTCTGTGTA
8651 AGCCAGTTAC CTTGCGAAAA AGAGTTGGTA GCTCTTGATC CGGCAACAA
8701 ACCACCGCTG GTAGCGGTGG TTTTTTTGTT TGCAAGCAGC AGATTACGCG
8751 CAGAAAAAAA GGATCTCAAG AAGATCCTTT GATCTTTTCT ACGGGGTCTG
8801 ACGCTCAGTG GAACGAAAAC TCACGTAAAG GGATTTTGGT CATGAGATTA
8851 TCAAAAAGGA TCTTCACCTA GATCCTTTTA AATTAAAAAT GAAGTTTAA
8901 ATCAATCTAA AGTATATATG AGTAAACTTG GTCTGACAGT TACCAATGCT
8951 TAATCAGTGA GGCACCTATC TCAGCGATCT GTCTATTTTC TTCTATCCATA
9001 GTTGCCCTGAC TCCCCGTCGT GTAGATAACT ACGATACGGG AGGGCTTACC
9051 ATCTGGCCCC AGTGCTGCAA TGATACCGCG AGACCCACGC TCACCGGCTC
9101 CAGATTTATC AGCAATAAAC CAGCCAGCCG GAAGGGCCGA GCGCAGAAGT
9151 GGTCTCTGCA CTTTATCCGC CTCCATCCAG TCTATTAATT GTTGCCGGGA
9201 AGCTAGAGTA AGTAGTTCGC CAGTTAATAG TTTGCGCAAC GTTGTGTTGA
9251 TTGTACAGG CATCGTGGTG TCACGCTCGT CGTTTGGTAT CGTTTCATT
9301 AGCTCCGGTT CCAACGATC AAGGCGAGTT ACATGATCCC CCATGTTGTG

```

```

9351 CAAAAAGCG GTTAGCTCCT TCGGTCCTCC GATCGTTGTC AGAAGTAAGT
9401 TGGCCGCAGT GTTATCACTC ATGGTTATGG CAGCACTGCA TAATTCTCTT
9451 ACTGTCATGC CATCCGTAAG ATGCTTTTCT GTGACTGGTG AGTACTCAAC
9501 CAAGTCATTC TGAGAATAGT GTATGCGGCG ACCGAGTTGC TCTTGCCCGG
9551 CGTCAATACG GGATAATACC GCGCCACATA GCAGAACTTT AAAAGTGCTC
9601 ATCATTGGAA AACGTTCTTC GGGGCGAAAA CTCTCAAGGA TCTTACCGCT
9651 GTTGAGATCC AGTTCGATGT AACCCTCTCG TGCACCCAAC TGATCTTCAG
9701 CATCTTTTAC TTTCACCAGC GTTCTGCGGT GAGCAAAAAC AGGAAGGCAA
9751 AATGCCGCAA AAAAGGGAAT AAGGGCGACA CGGAAATGTT GAATACTCAT
9801 ACTCTTCCTT TTTCAATATT ATTGAAGCAT TTATCAGGGT TATTGTCTCA
9851 TGAGCGGATA CATATTTGAA TGTATTTAGA AAAATAAACA AATAGGGGTT
9901 CCGCGCACAT TTCCCCGAAA AGTGCCACCT GAC

```

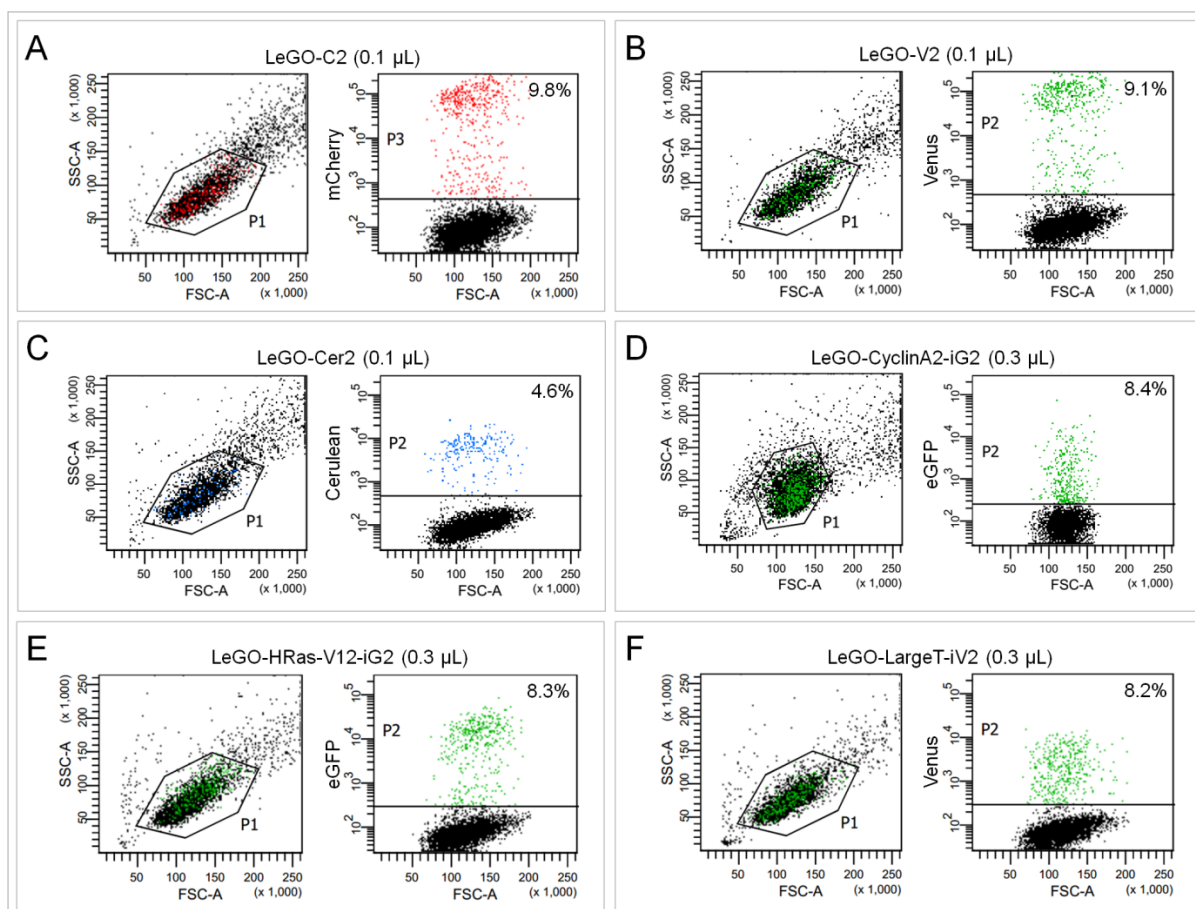


Figure S8 Titration examples for all vectors used in this thesis. LeGO-C2 (A), LeGO-V2 (B) and LeGO-Cer2 (C) were used in combination for RGB transduction of cells. LeGO-CyclinA2-iG2 (D, kindly provided by K. Riecken, Dept. of Stem Cell Transplantation, University Medical Center Hamburg-Eppendorf, Germany), LeGO-HRas-V12-iG2 (E) and LeGO-LargeT-iV2 (F) were used for solo transduction of cells. For each construct: SSC-A vs FSC-A on the left, fluorescence signal vs FSC-A on the right. Percentages of positive cell proportions were listed for each sample on the right top corner. SSC, side scatter; FSC, forward scatter; A, Area. Cells were analyzed on a LSR Fortessa and FACS Canto II (BD Biosciences).

Danksagung

Mein besonderer Dank geht an Prof. Dr. Boris Fehse für die Betreuung und die Möglichkeit in den letzten vier Jahren meine Promotion in der AG Zell- und Gentherapie anfertigen zu dürfen. Bei der Bearbeitung dieses herausfordernden Themas habe ich sehr viel gelernt und einige Hürden überwunden, durfte vieles ausprobieren und von dem mir gewährten Freiraum profitieren. Für die Ermöglichung der Teilnahme an Kongressen und Seminaren und das Korrekturlesen dieser Arbeit möchte ich mich ebenfalls bei Boris bedanken.

Prof. Dr. Thomas Dobner danke ich für die Tätigkeit als Zweitgutachter dieser Arbeit. Für die Mitwirkung in meiner Prüfungskommission danke ich Prof. Dr. Maura Dandri, PD Dr. Hartwig Lüthen und Prof. Dr. Julia Kehr als Vorsitzende. Den Fragestellern Dr. Kerstin Cornils und Dr. Claudia Lange danke ich ebenfalls.

Ein großes Dankeschön richte ich an Dr. Carol Stocking-Harbers für das Korrekturlesen dieser Arbeit und die Bescheinigung des angemessenen Sprachniveaus. Außerdem möchte ich Carol danken, dass sie Anfang 2014 den Kontakt zu Boris hergestellt hat, wodurch mir die Promotion erst ermöglicht wurde.

Dr. Tim Aranyossy, Lorenz Bausch und vor allem David Wozniak danke ich ebenso für das Korrekturlesen und Feedback zu dieser Arbeit.

Meinen ehemaligen Kollegen und meiner Chefin Dr. Jutta Höffler aus der TecMic möchte ich an dieser Stelle ebenfalls danken, da sie mich nicht nur im Vorfeld bedingungslos unterstützt und ermutigt haben, sondern mir auch in den Monaten nach meiner Zeit am UKE neue Motivationshilfe gaben.

Der Deutschen Forschungsgemeinschaft danke ich für die Finanzierung dieser Arbeit im Rahmen des Sonderforschungsbereichs 841 und dem dazu gehörigen Graduiertenkolleg unter Leitung von Prof. Dr. Gisa Tiegs mit wertvoller Unterstützung von Andrea Pfannkuch, Julia Hegemeister und Jutta Hüller-Dietrich.

Aus der AG von Prof. Dr. Tiegs möchte ich mich außerdem bei Elena Tasika bedanken, deren Geheimtipp dazu beitrug mein eigenes Protokoll zur Leberzellisolation zu verbessern. Außerdem danke ich Carsten Rothkegel und Laura Berkhout, die mir uneingeschränkten Zugang zum Cobas gewährt und die Serumpuben für mich gemessen haben. Laura Berkhout und Birgit Schiller möchte ich auch für den regelmäßigen Austausch inner- und außerhalb des GRKs herzlich danken.

Prof. Dr. Maura Dandri möchte ich auch für das Bereitstellen der USB-Mauslinie danken. Außerdem danke ich Dr. Tassilo Volz und Dr. Lena Allweiss für hilfreiche Tipps im Umgang mit den Mäusen und Einblick in die Leberzellisolation. Vielen, vielen Dank an Dr. Katja Giersch und Claudia Dettmer, die meine ersten beiden Mausexperimente gemeinsam mit mir durchgeführt haben. Claudia möchte ich an dieser Stelle doppelt danken, da sie mir außerdem die Kniffe beim Kryoschneiden beibrachte.

Dr. Jan Dieckhoff aus der Radiologie unter Leitung von Prof. Dr. Gerhard Adam und PD Dr. Harald Ittrich hat die zahlreichen Stunden am Kleintier-MRT nur so verfliegen lassen und einen wertvollen Beitrag zu meiner Arbeit geleistet: Vielen Dank!

Ohne die herausragende Arbeit in der Tierhaltung wäre mein Projekt nicht realisierbar gewesen. Dr. Mareike Budack und Christiane Willmann haben es ermöglicht, dass ich „mein eigenes“ S2-Mauslabor bekam und dafür bin ich ihnen zu großem Dank verpflichtet! Es war ein langer Weg und hat viele UKE-Mitarbeiter beschäftigt, um dies zu erreichen. Zu diesen gehörten auch Dr. Bastian Tiemann, der den Großteil einer Isoflurananlage zur Verfügung stellte, und Jens Mangels aus der Klimatechnik, der sich stets freundlich und hilfsbereit allen Fragen und Organisatorischem rund um die S2-Bench angenommen hat. Insbesondere Carol danke ich, dass ich die Bench überhaupt übernehmen durfte, und Ulla Müller vom HPI, die den Transport selbiger betreut hat. Für die Hilfe bei TBase-Problemen oder die Erfüllung von diesbezüglichen Sonderwünschen möchte ich Dr. Christian Peters danken. Bei Dr. Andreas Haemisch bedanke ich mich für die Unterstützung bei meinem Tierversuchsantrag. PD Dr. Claudia Lange hat mit viel Ausdauer und vielen Nerven alle Behördenfragen geklärt und sich um das Einhalten von Sonderauflagen gekümmert, vielen Dank dafür! Darüber hinaus gilt mein Dank besonders Nicole Lüder, die in der S1-Barriere für einen reibungslosen Ablauf sorgte, Nils Jäger, der den temporären Aufenthalt der Mäuse in der S2-Barriere ganz unkompliziert gestaltete, sowie Roswitha Reusch im Zenti, die sich gewissenhaft um die USB-Mauszucht kümmerte. Danke auch an Johannes Polke für all die Wochenendkontrollen im Mausstall!

Hinter den Kulissen haben die Spülfrauen, vor allem Manuela Sola Hermosa, unermüdlichen Einsatz gezeigt und für alle Selbstverständlichkeiten und Annehmlichkeiten in unserem Laboralltag gesorgt, wofür ich mich aufrichtig bedanken möchte.

Darüber hinaus möchte ich mich bei allen Mitgliedern der Arbeitsgruppe, aber auch den ehemaligen Kollegen und Studenten, für die gemeinsame Zeit bedanken.

„Hepatozyten facsen ist doof.“ Und trotzdem haben sie mir mit Rat und Tat zur Seite gestanden; lieben Dank an Melanie Lachmann, Susanne Roscher, Julia Spötter und Regine Thiele aus der FACS Core. In den ersten Jahren hatten wir so manche Besprechung und haben auch einige Stunden zusammen Mäuse transplantiert; für den Input und die Zusammenarbeit danke ich PD Dr. Daniel Bente. Für die Bereitstellung der Plasmide und Virusüberstände, sowie technischen Support bedanke ich mich bei Dr. Kristoffer Riecken.

Dr. Ercan Akgün danke ich, dass er mir wenige Tage vor seinem Urlaub und Ausscheiden aus der AG - noch vor meinem eigentlichen Promotionsstart - die Abläufe eines Mausexperimentes gezeigt hat. Bei Anita Badbaran, die ich als gute Seele des Labors kennengelernt habe, bedanke ich mich herzlich für das gemeinsame Designen der Sonden und Primer für die digitale PCR, für die Hilfe bei allen Fragen rund um RNA-Isolation, aber auch die Unterstützung generell bei Schwierigkeiten jedweder Art.

Tanja Sonntag half jederzeit auch kurzfristig beim Blutabnehmen bei meinen Mäusen, war bei Fragen immer ansprechbar und hat jeden Bestellmarathon mitgemacht – tausend Dank dafür, für ihre ansteckende Fröhlichkeit, Spontanität und die gemeinsamen Unternehmungen.

Dr. Stefan Horn danke ich für die zahlreichen kleinen Gespräche zwischen Tür und Angel über Mausexperimente, Wissenschaft im Allgemeinen und seinen aufmunternden Zuspruch, der manchmal unerwartet, aber genau im richtigen Moment kam.

Allen Doktoranden aus der Arbeitsgruppe danke ich für den Zusammenhalt, die wertvolle Unterstützung und den wissenschaftlichen Austausch der letzten Jahre!

Allzeit hilfsbereit und ein toller Büroplatznachbar war Dr. Tim Dominic Aranyossy, mit dem ich viele Fragestellungen erörtern und von dessen etablierten Protokollen profitieren konnte. Danke für mein „Nein!“-Schild (ich werde immer besser darin ;-)), die zahlreichen Gespräche, unsere gemeinsamen U-Bahn-Fahrten und spontan umgesetzte „Schnaps“-Ideen.

Wir hatten viele Diskussionen über den Ablauf unserer Experimente, haben Kriegsräte gehalten, uns gegenseitig unter die Arme gegriffen und regelmäßig unsere Mittagspause für den anderen verschoben – dafür und für ihre Ehrlichkeit, Offenheit und enge Freundschaft bedanke ich mich ganz herzlich bei Dr. Lucía Gallo Llorente de Wojtczak.

Unser Quartett im Büro hat Lea-Isabell Schwarze komplettiert. Dabei durfte ich ihre besonnene, herrlich unkomplizierte und freundliche Art kennen und schätzen lernen. Danke für unsere Freundschaft, jede Unterhaltung, ob beruflicher oder privater Natur, alle netten Worte, den realistischen Optimismus und die Geduld, meinem (teilweise) nicht enden wollendem Redeschwall zuzuhören.

Jill-Sandra Schoefinius ist mir besonders in den letzten Jahren sehr ans Herz gewachsen. Danke für unsere stundenlangen Telefonate, Lob und Tadel gleichermaßen, die Bereitschaft jedes Mal „meine“ Leberlappen umzutopfen, die (manchmal übertriebene, aber lebenswerte) Fürsorge und alle uns noch bevorstehenden Albernheiten.

Dr. Kerstin Cornils kann ich nur sagen: Danke für Alles. Ihr offenes Ohr, konstruktive Diskussionen, neue Ideen, tröstende Aufmunterung, rosa Flamingos und ihre Ausdauer, ellenlange Nachrichten zu lesen und auch noch zu beantworten.

Mein aufrichtiger Dank gilt auch Dr. Julia Preukschas, ohne die ich wahrscheinlich eine Doktorarbeit gar nicht erst in Erwägung gezogen hätte. Danke für die Ermutigung und tolle Zusammenarbeit im Vorfeld, die mich für meine wissenschaftliche Arbeit geprägt hat. Für unsere regelmäßigen Donnerstagstreffen in den letzten Jahren danke ich auch Marion Ziegler und Dr. Nilgün Tekin-Bubenheim, die jede Woche zu etwas Besonderem gemacht haben. Das wird mir fehlen!

Meiner Familie, allen Freunden und besonders David danke ich für die unschätzbare Unterstützung, die Liebe und das Verständnis, auf das ich jederzeit bauen konnte.

An alle, die mich in den letzten Jahren begleitet, mir Mut gemacht oder geduldig zugehört, all die Beschwerden ertragen und Freudentänze mitgefeiert haben: Danke.

Bestätigung der Korrektheit der Sprache

Dr. Carol Stocking-Harbers

Woldsenweg 13 • 20249 Hamburg • Telefon: 040 474896
E-Mail: csharbers@googlemail.com



Studienbüro Biologie
z.H. Frau Sült-Wüpping
MIN Fakultät
Universität Hamburg
Biozentrum Klein Flottbek
Ohnhorststraße 18
22609 Hamburg

31. Juli 2018

Sehr geehrte Damen und Herren,

hiermit bestätige ich, dass die von Frau Ludmilla Unrau mit dem Titel "Hepatocellular carcinoma: Establishing a new transplantation model for the study of benign and malignant clonal regeneration of the murine liver *in vivo*" vorgelegte Doktorarbeit in korrektem Englisch geschrieben ist.

Mit freundlichen Grüßen,

Dr. Carol Stocking
Klinik für Stammzelltransplantation
Universitätsklinikum Hamburg-Eppendorf
(Amerikanerin)

Eidesstattliche Versicherung

Hiermit bestätige ich, dass die vorliegende Arbeit von mir selbständig verfasst wurde und ich keine anderen als die angegebenen Hilfsmittel – insbesondere keine im Quellenverzeichnis nicht benannten Internet-Quellen – benutzt habe und die Arbeit von mir vorher nicht einem anderen Prüfungsverfahren eingereicht wurde. Die eingereichte schriftliche Fassung entspricht der auf dem elektronischen Speichermedium.

Hamburg, den 31. Juli 2018

Ludmilla Unrau