

Aus der Klinik für Kardiologie und Pneumologie
(Prof. Dr. med. G. Hasenfuß)
der Medizinischen Fakultät der Universität Göttingen

**‘Knockout-first’ mouse model as a biological tool to study the
role of *KIAA0182* gene in hypoplastic left heart syndrome**

INAUGURAL-DISSERTATION
zur Erlangung des Doktorgrades
der Medizinischen Fakultät der
Georg-August-Universität zu Göttingen

vorgelegt von
Fouzi Alhour
aus
Damas Suburb, Syrien

Göttingen 2015

Dekan: Prof. Dr. rer. nat. H. K. Kroemer

I. Berichterstatter: Prof. Dr. E. Zeisberg

II. Berichterstatter: Prof. Dr. S. Johnsen

III. Berichterstatterin: Prof. Dr. M. Schön

Tag der mündlichen Prüfung: 16.03.2016

Table of contents

1. Introduction	1
1.1 Hypoplastic left heart syndrome	2
1.1.1 Definition.....	2
1.1.2 Incidence.....	3
1.1.3 Pathogenesis and etiology	3
1.1.4 Genetics	5
1.1.5 Clinical presentation.....	6
1.1.6 Diagnosis and management	6
1.2 Endocardial fibroelastosis (EFE)	7
1.2.1 Definition.....	7
1.2.2 Classification	7
1.2.3 Incidence.....	8
1.2.4 Etiology	8
1.2.5 Pathogenesis	9
1.2.6 Diagnosis and management	9
1.3 Endothelial to mesenchymal transition (EndMT)	10
1.3.1 Definition of EndMT	10
1.3.2 EndMT stimulants and mechanism	10
1.3.3 EndMT markers	12
1.3.4 EndMT and cardiac fibrosis	13
1.4 KIAA0182 gene	14
1.4.1 General information	14
1.4.2 Gse1 gene in mouse	15
1.4.3 <i>KIAA0182</i> and circular RNA.....	16
1.4.4 KIAA0182 and CoREST complex.....	17
1.4.5 <i>KIAA0182</i> and cardiovascular diseases	19
1.5 Gene trap mutagenesis	19
1.5.1 Mutagenesis strategies.....	19
1.5.2 Gene trapping.....	20
1.5.3 The ‘Knockout-first’ strategy.....	21
2. Materials and methods	23
2.1 Materials	23
2.1.1 Animals	23
2.1.2 Chemicals.....	24

2.1.3	Commercial kits	26
2.1.4	Cell culture mediums.....	26
2.1.5	Buffers	26
2.1.6	Instruments	27
2.1.7	Antibodies	28
2.1.8	Primers	28
2.1.9	Other materials.....	31
2.2	Methods.....	32
2.2.1	Genomic DNA extraction.....	32
2.2.2	RNA extraction	32
2.2.3	RNA reverse transcription.....	33
2.2.4	RNase R treatment	33
2.2.5	DNA extraction from agarose gel.....	34
2.2.6	Genotyping PCR	34
2.2.7	Short-range PCR.....	36
2.2.8	Long-range PCR.....	37
2.2.9	Reverse transcription PCR	37
2.2.10	Quantitive real-time PCR	38
2.2.11	Isolation of mouse fibroblasts	39
2.2.12	Cell culture	39
2.2.13	EndMT assay	40
2.2.14	Small interfering RNA transfection.....	40
2.2.15	Transduction of primary fibroblasts with Cre-recombinase adenovirus... 41	
2.2.16	Protein extraction.....	41
2.2.17	Western blotting.....	41
2.2.18	Ascending aortic constriction (AAC)	42
2.2.19	Masson's trichrome staining	43
2.2.20	Statistical analysis	43
3.	Results.....	45
3.1	Genotyping protocols for 'Knockout-first' mice	45
3.2	Quality control tests	46
3.2.1	Confirming the specificity of <i>Gse1</i> targeting.....	46
3.2.2	Confirming the structure of the trapping cassette	49
3.3	Genotyping results.....	49
3.4	Generating mice with <i>Gse1</i> ^{tm1c} allele.....	52

3.5	Generating <i>Gse1</i> ^{tm1b} allele in vitro	54
3.6	<i>Gse1</i> expression results	56
3.7	<i>Gse1</i> circular RNA results	59
3.8	The expression results of <i>Gse1</i> -neighboring genes	64
3.9	The role of <i>KIAA0182</i> in EndMT	65
3.10	Results of AAC operation	67
4.	Discussion	70
5.	Summary	88
6.	References	89

List of abbreviations

α-SMA	Alpha-smooth muscle actin
AAC	Ascending aortic constriction
ASD	Atrial septum defect
BHC	BRAF35-HDAC complex
BMP	Bone morphogenetic protein
bp	Basepairs
cDNA	Complementary DNA
CoREST	REST corepressor 1
DNA	Deoxyribonucleic acid
EFE	Endocardial fibroelastosis
EMT	Epithelial-mesenchymal transition
EN2SA	Engrailed 2 splice acceptor
EndMT	Endothelial to mesenchymal transition
ES	Embryonic stem cells
EUCOMM/KOMP-CSD	European conditional mouse mutagenesis program and knockout mouse program
FCS	Fetal calf serum
FRT	FLP-recombinase recognition target
FSP	Fibroblast specific protein
GFP	Green fluorescent protein
GRC	Genome reference consortium
HCAEC	Human coronary artery endothelial cells
HLHS	Hypoplastic left heart syndrome
HRP	Horseradish peroxidase
iPSCs	Induced pluripotent stem cells
IRES	Internal ribosome entry site
KOMP2	Knockout mouse project phase 2
LacZ	β -galactosidase
LoxP	Locus of crossover (x) in P1
MCEC	Mouse cardiac endothelial cells
miRNA	Micro RNA
MOI	Multiplicity of infection

mRNA	Messenger RNA
ncRNA	Noncoding RNA
PBS	Phosphate buffered saline
PBST	PBS with 1% Tween 20
PCR	Polymerase chain reaction
PDA	Patent ductus arteriosus
PFO	Patent foramen ovale
qRT-PCR	Quantitative real-time polymerase chain reaction
RE1	Repressor element 1
REST	RE1 silencing transcription factor
RNA	Ribonucleic acid
RT-PCR	Reverse transcription polymerase chain reaction
SA	Splice acceptor
siRNA	Small interfering RNA
TGF-β	Transforming growth factor-beta

1. Introduction

It was the German pathologist Bardeleben, who was the first to describe hypoplastic left heart syndrome (HLHS) more than 160 years ago. In his report about the autopsy findings in an infant after dying due to severe asphyxia, he was able to predict the reason of death accurately, showing clear understanding of the pathophysiology of this disease (Gehrmann et al. 2001). This pathophysiology is exemplified by the presence of a hypoplastic left ventricle, which implies the necessity of patent ductus arteriosus (PDA) for maintaining the systemic circulation, and consequently the survival of the affected patient. Since that time until now a lot of efforts were dedicated to understand this syndrome, especially in regard to its molecular mechanisms, and probably genetic etiology.

HLHS is furthermore characterized by the presence of so called endocardial fibroelastosis (EFE), which is a diffuse fibrotic thickening in the endocardium of left ventricle (Friehs et al. 2013). Endocardial fibroelastosis is a unique type of cardiac fibrosis, which was shown to be derived from aberrant endothelial to mesenchymal transition (EndMT) (Xu et al. 2015a). EndMT is known to be essential for the formation of cardiac cushions during embryonic life (Eisenberg and Markwald 1995; Armstrong and Bischoff 2004). Mutations in several genes related to EndMT, such as *NOTCH1* that represents an essential regulator for EndMT process during cardiac development, were suspected to be involved in the pathogenesis of HLHS (Garg et al. 2005; Iacone et al. 2012).

Another gene, *KIAA0182* (*Gse1* in mouse), is found to be mutated in HLHS (unpublished data from Iacone lab, Bergamo). However, the function of this gene with respect to its role in congenital heart diseases or any other biological process is still unknown. Interestingly, *KIAA0182* was reported before to be a potential component in CoREST complex that is involved together with the transcription factor *SNAIL* in a process similar to EndMT, which is epithelial-mesenchymal transition (EMT) (Hakimi et al. 2003; Lin et al. 2010; Yang et al. 2011; Yokoyama et al. 2008). This gene is also known to produce one abundant exonic circular RNA (Jeck et al. 2013; Memczak et al. 2013; Starke et al. 2015), but to the best of our knowledge no other study investigated before any possible specific function for this circularized exon or for *KIAA0182* gene in general.

Therefore, taking advantage of the 'Knockout-first' mouse model used in this project and additional *in vitro* experiments, this study was performed to answer some questions related to the function of *KIAA0182*, and how it may be involved in HLHS or in other diseases, which can be summarized in the following aims:

- To confirm the correct targeting of *Gse1* gene and the structure integrity of the trapping cassette in the 'Knockout-first' mice.
- To establish suitable genotyping protocols for the 'Knockout-first' mice with the mutated allele *Gse1*^{tm1a}, and for the mice with the mutated allele *Gse1*^{tm1c} after breeding with FLP-recombinase mice.
- To investigate the phenotype associated with *Gse1* trapping and the efficiency of the trapping.
- To elucidate the function of *Gse1* gene and its circular RNA with respect to its role in EndMT and in HLHS using *in vitro* and *in vivo* experiments.

1.1 Hypoplastic left heart syndrome

1.1.1 Definition

Hypoplastic left heart syndrome (HLHS) is a severe form of congenital heart diseases, and it is the most common form of single ventricle heart defects (Fixler et al. 2010). This syndrome is defined, according to the International Working Group for Mapping and Coding of Nomenclatures for Paediatric and Congenital Heart Disease, as "a spectrum of cardiac malformations with normally aligned great arteries without a common atrioventricular junction, characterized by underdevelopment of the left heart with significant hypoplasia of the left ventricle including atresia, stenosis, or hypoplasia of the aortic or mitral valve, or both valves, and hypoplasia of the ascending aorta and aortic arch". The term hypoplastic left heart complex is used to describe the milder forms of this syndrome without valvar stenosis or atresia (Tchervenkov et al. 2006). In HLHS patient, the oxygenated blood reaches the right atrium through patent foramen ovale (PFO) or nonrestrictive atrial septal defect (ASD), and after mixing with the deoxygenated blood, the right ventricle pumps the mixed blood to reach the systemic circulation through the patent ductus arteriosus (PDA). The endocardium of the left ventricle is commonly thickened, forming prominent fibro-elastic tissue called endocardial fibroelastosis (EFE), which represents a very special type of cardiac fibrosis and prevents the left ventricle from

growing (Friehs et al. 2013). Left untreated, this disease is associated with 100 % mortality. HLHS accounts for 20-25 % of deaths because of congenital heart defects, representing the first cause of mortality linked to these diseases in infants less than one year old (Boneva et al. 2001; Gordon et al. 2008)

1.1.2 Incidence

The incidence of hypoplastic left heart syndrome (HLHS) is estimated to be two to three cases per 10,000 live births in the United States, and it is responsible for 2-3 % of all congenital heart diseases (Barron et al. 2009; Hoffman and Kaplan 2002; Reller et al. 2008). Both males and females are affected with a male predominance (ratio about 3:2) (Karamlou et al. 2010; Tikkanen and Heinonen 1994). The disease is described in all ethnic groups with no difference reported in the incidence rate related to race in several studies (Botto et al. 2001; Shaw et al. 2002; Storch and Mannick 1992), but one study showed higher rate among Caucasians compared with other ethnic groups (Cronk et al. 2004). The reports about any seasonal variation are inconsistent, as different studies came to different results (Cronk et al. 2004; Eghtesady et al. 2011; Samánek et al. 1991; Tikkanen and Heinonen 1994). The reported incidence of this disease may represent underestimation of the true incidence, as the rate of spontaneous abortions is not determined and the option of elective termination of pregnancy is available in some countries if the disease is antenatally diagnosed (Boldt et al. 2002; Cox and Wilson 2007; Hoffman and Kaplan 2002).

1.1.3 Pathogenesis and etiology

The most prevalent theory to explain the development of hypoplastic left heart syndrome depends on the hypothesis that abnormal blood flow through the left side of the heart causes cardiac malformations, as it reduces the shear forces applied on the developing heart. The evidence for this theory is derived from animal studies carried out on chick and zebrafish embryos (Hove et al. 2003; Sedmera et al. 2002), and other studies on human fetuses with aortic valve stenosis, which has later developed into hypoplastic left heart syndrome (Mäkikallio et al. 2006; McElhinney et al. 2009). In embryo, the blood passes through the foramen ovale towards the left ventricle, and it was described that the foramen ovale in fetuses with HLHS is smaller

than that in normal fetuses; suggesting also that the resulting impaired flow negatively affects the growth of the left ventricle (Feinstein et al. 2012).

However, a primary problem in the myocardium of the left ventricle cannot be excluded, as it was also found that successful fetal aortic valvuloplasty for aortic stenosis was not associated with improvement in the growth rate of the left ventricle, and the biventricular circulation could not be maintained after birth in some cases (McElhinney et al. 2009). Another study has also demonstrated that cardiomyocytes derived from induced pluripotent stem cells (iPSCs), which were differentiated from the skin fibroblasts of a patient with HLHS, showed altered functional characteristics and expression levels of several markers. That was shown by comparing these cells with cardiomyocytes derived from human embryonic stem cells, and with cardiomyocytes derived from the iPSCs differentiated from the dermal fibroblasts of unaffected control, suggesting myocyte susceptibility as a contributing factor in the pathogenesis of HLHS (Jiang et al. 2014).

Another important process to be considered in this context is endothelial to mesenchymal transition (EndMT), which is necessary during embryonic heart development for the formation of valves and septum (Eisenberg and Markwald 1995). It was found that aberrant epigenetically-induced EndMT leads to the development of EFE tissue (Xu et al. 2015a), which plays an important role in the pathogenesis of HLHS. It was observed that the systolic and diastolic function of the left ventricular in borderline left heart disease (moderate severity of HLHS) improved when the EFE was resected, making it possible for those patients to avoid the conversion to univentricular circulation (Emani et al. 2009). Another study reported that the severity of EFE tissue, estimated by echocardiography for midgestation fetuses with evolving hypoplastic left heart syndrome, can be correlated with postnatal biventricular outcome (McElhinney et al. 2010).

The primary causes behind these processes are still not well understood, and several environmental and/or genetic risk factors can play a role. Maternal infections in the first trimester were reported to be associated with HLHS (Tikkanen and Heinonen 1994). Maternal exposure to solvents or degreasing agents, such as paint stripper; or the use of certain drugs during pregnancy, such as retinoids, were also described as potential causes for this disease (Cox and Wilson 2007; Wilson et al.

1998). No link was found between HLHS and maternal smoking, alcohol or coffee consumption (Kallen 1999; Tikkanen and Heinonen 1994).

1.1.4 Genetics

A large body of evidence supports the genetic etiology of HLHS. For example, several case reports about familial recurrence of HLHS described different modes of inheritance with different degrees of penetrance and phenotype severity (Grossfeld 1999). Higher risk for congenital heart diseases was also observed in the first degree relatives of a child diagnosed with HLHS, where most of the affected relatives had left-sided obstructive defects. This suggests the presence of common genetic etiology for this group of cardiac diseases (Boughman et al. 1987; Loffredo et al. 2004). More than 30 syndromes contain HLHS as a cardiac phenotype, such as Holt-Oram syndrome that is caused by mutations in *TBX5* gene (Bruneau et al. 1999); Rubinstein Taybi syndrome, which is caused by a mutation in *CBP* gene that is also named as CREB Binding Protein (CREBBP) (Hanauer et al. 2002); and Alagille syndrome as described in one case report (Robert et al. 2007). Chromosomal abnormalities were found in 10 % of children with HLHS, which include trisomies-21, -13, and -18; Turner syndrome; and deletion of distal 11q that is also called Jacobsen syndrome (Cox and Wilson 2007; Grossfeld et al. 2004; Natowicz et al. 1988).

Mutations in different genes were reported to associate with this syndrome. For example as described in one case of HLHS with ASD, heterozygous C-T transition at nucleotide 642 was found in *NKX2-5* gene on chromosome 5, whose homozygous disruption was reported to cause abnormal cardiac development and early embryonic lethality in mice. However, it is still not certain whether this mutation is pathologically significant or not, as it was found in another first degree relative without any cardiac defects; and no mutations was reported at all in this gene in a recent study (Elliott et al. 2003; lascone et al. 2012; Lyons et al. 1995). Among other suspected genes, a frameshift mutation in *NOTCH1* was also described in a family with different heart defects including hypoplastic left ventricle (Garg et al. 2005), and two *de novo* mutations in *NOTCH1* were additionally found among 53 patients with HLHS (lascone et al. 2012). *De novo* mutation in *KIAA0182* gene causing heterozygous T-C substitution was also detected in a child with HLHS (unpublished data from lascone lab, Bergamo). Mutations in *GJA1* gene (6q22), which is also called *connexin43*, were identified in one study including 14 children with HLHS, where 8 of

them had defects in this gene. This was found by analyzing genomic DNA extracted from the tissue of hearts explanted from HLHS patients upon heart transplantation (Dasgupta et al. 2001). A frameshift mutation in *HAND1* gene (5q33) was described in 24 out of 31 hypoplastic ventricles. However, in the previously mentioned work of lascone and her colleagues this gene was not found to be mutated using genomic DNA isolated from the blood of HLHS patients, suggesting that analyzing blood samples might not be sufficient to identify mutations that are responsible for this syndrome (lascone et al. 2012; Reamon-Buettner et al. 2008).

1.1.5 Clinical presentation

The presence and size of interatrial communication and the patency of the ductus arteriosus decide the timing of clinical presentation after birth. In most of the patients the size of the patent foramen ovale (PFO) or the atrial septum defect (ASD) is adequate to guarantee sufficient mixing of oxygenated and deoxygenated blood, and the infants are relatively free of symptoms at birth. However, the physiologic closure of the arterial duct and the reduction in vascular pulmonary resistance cause a dramatic decrease in systemic perfusion, which leads to hypoxia, acidosis and shock. HLHS patients with a restricted or no connection between left and right atrium manifest at birth with severe cyanosis and cardiogenic shock, and they will not be able to survive without an immediate intervention to create an adequate interatrial communication (Vlahos et al. 2004).

1.1.6 Diagnosis and management

Echocardiography is usually used to make the postnatal diagnosis. Detailed information about the anatomic features should be obtained to direct the future management, including the adequacy of the interatrial communication, the function of the valves and the size of the ascending aorta. Prenatal diagnosis can also be made by fetal echocardiography in the 2nd trimester of pregnancy (Blake et al. 1991), which gives the opportunity for parental education, for arranging the delivery in or near a specialized center for neonates with HLHS in order to improve their survival (Morris et al. 2014), or for the termination of pregnancy. The option of fetal intervention is also available in some centers for selected cases with critical aortic valve stenosis and left ventricular with accepted volume. Transcatheter aortic

valvuloplasty in some of those cases is described to prevent the development of HLHS (Tworetzky et al. 2004; McElhinney et al. 2009).

Apart from the option of no intervention, three major surgical management strategies are available postnatally after an initial stabilizing medical therapy, which aims to keep the ductus arteriosus patent by prostaglandin E1 (Alprostadil) infusion, and in some cases, to create an interatrial communication of sufficient size by transcatheter atrial septoplasty. The first surgical option is primary cardiac transplantation, which represents a curative option. However, it is restricted by many problems including the increased risk of death during waiting for the hearts of infant donor, which are very rare (Morrow et al. 1997). Therefore, the other option of palliative surgical therapy, divided in 3 stages, is currently more preferred (Prsa et al. 2010). This strategy is considered to be palliative as it restores the patient to univentricular system using the right ventricle to support in-series systemic and pulmonary circulations (Barron et al. 2009). The third option of biventricular repair may be considered for patients with mild severity of HLHS (Emani et al. 2009; Grossfeld et al. 2009; Tchervenkov et al. 2006)

1.2 Endocardial fibroelastosis (EFE)

1.2.1 Definition

The term endocardial fibroelastosis was presented for the first time almost 70 years ago by Weinberg and Himelfarb, who described two cases of infantile cardiomegaly in siblings (Weinberg and Himelfarb 1943), where the endocardium appeared milky white, glistening and opaque, not thin and transparent as in normal hearts. The characteristic pathology in this disorder is represented by different degrees of thickening in ventricular endocardium composed of fibroelastotic tissue and manifests with symptoms of heart failure in infants and children (Lurie 2010; Sellers et al. 1964; Steger et al. 2012).

1.2.2 Classification

Endocardial fibroelastosis can be classified into two types; the secondary form which is associated with congenital heart diseases, mostly hypoplastic left heart

syndrome and left ventricular outflow tract obstruction, and likely as a result of these defects. The other type is the primary form, which occurs in anatomically normal hearts without obvious reason and manifests usually as dilated cardiomyopathy, but it may rarely have restrictive or contracted form with left ventricular of small or normal size (Ni et al. 1997; Ursell et al. 1984). However, this classification may be misleading, as even primary forms are recognized nowadays to be secondary to another heart disorder (other than congenital heart diseases), which makes the term idiopathic more suitable for the cases without clear cause (Lurie 2010).

1.2.3 Incidence

In 1964 a study reported that the incidence of this disorder in the United States is almost one per 5000 live births (Moller et al. 1964). This high rate has dramatically decreased in recent decades without identified causes. However, that can be attributed probably to the declining of mumps, as it started particularly since the introduction of the mumps vaccine (Ni et al. 1997). This potential causal relation should be kept in mind, as the children of young adults affected by mumps during some epidemics in the last decade may be more susceptible to endocardial fibroelastosis (Lurie 2010). Actually, it was recently described that endocardial fibroelastosis is found in 25% of children undergoing heart transplantation with the diagnosis of dilated cardiomyopathy (Seki et al. 2013).

1.2.4 Etiology

Endocardial fibroelastosis can not be regarded as a distinct disease, but rather as a reaction in the endocardium in response to different heart stressors, which are more prominent in the left ventricle, where the stress is usually at the highest level (Lurie 2010; Ursell et al. 1984). Several factors can be involved in the etiology of endocardial fibroelastosis. Viral infections, such as infection with Coxsackie or mumps virus, are accused to be responsible for the development of this reaction, following a stage of myocarditis (Lurie 2010). Non-infectious etiologies can also play a role in the occurrence of endocardial fibroelastosis, as it was found to be the result of some metabolic abnormalities, such as the absence of lysosomal alpha-glucosidase or Carnitine deficiency (Dincsoy et al. 1965; Tripp et al. 1981). Placental immaturity and myocardial hypoxia in the fetus may also be responsible for this reaction (Perez et al. 2009). Immunological problems, such as maternal anti-Ro and

anti-La, were reported to be associated with endocardial fibroelastosis together with congenital heart block (Nield et al. 2002). Strong association was reported between reduced blood flow through the left ventricle during fetal state and the development of EFE (Friehs et al. 2013). Additionally, impaired cardiac lymph flow; myocardial infarction or physical injury, such as electrical shock, can cause this reaction in the heart (Hutchins and Bannayan 1971; Kline et al. 1964; Naguit and Dexheimer 1974). However, it should be kept in mind that these triggers may not be able to induce the development of EFE without the presence of genetic susceptibility. Mutations in several genes, such as *tafazzin* gene (*TAZ*) on chromosom X, β -cardiac myosin heavy chain gene, muscle *LIM* protein or *a-actinin-2* genes can be linked to this disorder (Brady et al. 2006; Kamisago 2006; Mohapatra et al. 2003).

1.2.5 Pathogenesis

The normal endocardium has five layers. Starting from the cavity of the heart towards outside, these layers are: the endothelium at first; a layer of loose connective tissue with a few cells; a layer of elastin and collagen fibers then another layer of smooth muscle cells. The last layer before the myocardium is loose connective tissue containing capillaries, unmyelinated nerves and a few cells. Several histologic changes can contribute to the pathogenesis of endocardial fibroelastosis, including inflammatory, infiltrative or scarring changes in the myocardium related to the primary disease. The thickening reaction in the endocardium manifests mostly in the layer of smooth muscle cells and in the sub-myocardial layer of loose connective tissue (Lurie 2010). In a recent study, the role of endothelial-mesenchymal transition was identified as a source of fibroblasts producing elastin and collagen fibers in endocardial fibroelastosis (Xu et al. 2015a).

1.2.6 Diagnosis and management

Previously, the diagnosis of endocardial fibroelastosis was only possible to be made postmortem by autopsy. However, the development in the technology of echocardiography made it possible to detect increased echodensity of the endocardium with reduced contractility of the ventricle as early as the age of 14 weeks in pregnancy (Rustico at al. 1995). It was emphasized on the importance of using the right terminology in reporting the diagnosis of endocardial fibroelastosis as a secondary reaction to another heart disease (if it is possible to be identified), not as

a distinct entity. It is almost impossible to separate the pathological effect of this reaction from the effect of the primary disease (Lurie et al. 2010). Similarly, the therapeutical interventions are usually directed against the primary etiology in combination with symptomatic treatment for heart failure. However, previous report described the possibility of completely reversing the development of endocardial fibroelastosis, which was caused by maternal anti-Ro and anti-La antibodies, by steroids therapy (Raboisson et al. 2005). Resecting the fibroelastic tissue itself was shown to have additional benefit regarding the function and the growth of the left ventricle after fetal valvotomy for congenital aortic stenosis in non-responding patients (Lurie et al. 2010; Tworetzky et al. 2005).

1.3 Endothelial to mesenchymal transition (EndMT)

1.3.1 Definition of EndMT

EndMT is defined as a differentiation process, in which endothelial cells drop off their endothelial nature, such as cell-cell junctions and cell polarity, and gain new mesenchymal characteristics. This includes the acquiring of invasive and migratory capacities to transform into mesenchymal or fibroblast-like cells with elongated spindle-shaped morphology, and represents a special form of epithelial-mesenchymal transition (EMT) (Medici and Kalluri 2012; Yu et al. 2014). This process was first described in a study about heart development 40 years ago (Markwald et al. 1975). It was thought before that the fibroblast phenotype represents the final fate of the endothelial cells undergoing EndMT; but it was recently shown that these endothelial cells are able to transform into stem-like cells. That indicates that EndMT is a dedifferentiation process, and the dedifferentiated endothelial cells can later differentiate again into other types of cells, such as osteoblasts, chondrocytes or adipocytes (Medici and Kalluri 2012).

1.3.2 EndMT stimulants and mechanism

Several stimulants can induce the endothelial cells to go into EndMT, such as hypoxia, which is shown to cause EndMT in pulmonary artery endothelial cells (Zhu et al. 2006). This is thought to be mediated by the downregulation of bone morphogenetic protein receptor type II, and it was suggested that EndMT can play a

role in the development of pulmonary arterial hypertension (Reynolds et al. 2012). Another important stimulant is inflammation, as it was found that interleukin-6 and tumor necrosis factor- α , which are inflammatory cytokines, can induce EndMT in porcine aortic valve endothelial cells, which indicates that this process may be involved as an early stage of the aortic valve disease (Mahler et al. 2013). High glucose can also cause the loss of endothelial cell markers and transform the human aortic endothelial cells into chondrocyte-like cells, and this may be important in the development of medial calcification in the vessels of diabetic patients (Tang et al. 2012; Yu et al. 2014).

The exact mechanism of EndMT is still not completely understood, but a strong evidence can be obtained from several studies *in vitro* and *in vivo* that transforming growth factor-beta (TGF- β), among other factors, is important in EndMT induction, which is to some degree mediated by the SNAIL family of transcription factors (Frid et al. 2002; Kokudo et al. 2008; Medici et al. 2011; Zeisberg et al. 2007b). The TGF- β superfamily comprises two groups; the bone morphogenetic proteins (BMPs) and TGF- β /activin A subfamilies. TGF- β is involved in the regulation of several biological events, such as cell proliferation, cell differentiation and cell apoptosis (Lin et al. 2012). It has 3 isoforms (TGF β 1, TGF β 2, TGF β 3), which were all shown to be involved in EMT, especially TGF β 1, but regarding EndMT, TGF β 2 is more important EndMT-inducer compared with TGF β 1, whereas TGF β 3 is found to be not essential for EndMT during embryonic development (Medici and Kalluri 2012). The signal of TGF- β superfamily members is mediated by different combinations of two types of transmembrane receptors, which are type I and type II serine/threonine kinase receptors, where the type II receptor phosphorylates the type I receptor upon binding to the ligand. There are seven type I receptors and five type II receptors that are coded by the human genome, and the specificity to certain ligand is determined by the combinations formed from them (van Meeteren and ten Dijke 2012).

TGF- β has two type I receptors in endothelial cells working with one TGF- β type II receptor. The first receptor is the activin receptor-like kinase 1 (ALK1) that activates SMAD 1/5 pathways and can induce endothelial cell proliferation and migration. The other type I receptor is ALK5 that activates SMAD 2/3 and plays an inhibitory role for the previous events, but it was shown that ALK5 is also necessary for TGF- β /ALK1-induced response (Goumans et al. 2003; Lin et al. 2012). Activation of these SMAD

proteins induces their interaction with SMAD 4, the common-mediator (Co) SMAD, forming a larger complex that translocates into the nucleus to activate the transcription factors that mediate the epigenetic response to TGF- β (van Meeteren and ten Dijke 2012). A long list of transcription factors can be involved in the cascade of EndMT and EMT in general, such as SNAIL, SLUG, TWIST, zinc finger E-box binding homeobox 1 and 2, and FOXC2. These have the common function of inhibiting the expression of epithelial- and endothelial-specific proteins; and upregulating the mesenchymal markers, and were used previously in several studies as markers for EndMT process (Cooley et al. 2014; Elliott et al. 2014; Kalluri and Weinberg 2009; Lee et al. 2015; Levet et al. 2015; Piera-Velazquez et al. 2011; Xu et al. 2015a; Xu et al. 2015b).

It is necessary to mention here that receptors other than previous receptors can bind to TGF- β , which are betaglycan and endoglin that are also expressed in endothelial cells and play a role in modulating the response to TGF- β mediated by type I and type II receptors (Wong et al. 2000). Other pathways are suggested to be involved in EndMT and EMT; such as *NOTCH* pathway, whose signal is important during heart development mediated also by *SLUG* activation (Nosedá et al. 2004; Timmerman et al. 2004; Yoshimatsu and Watabe 2011). The canonical *Wnt* signaling pathway is also described to induce EndMT in subepicardial endothelial cells after myocardial infarction, which may be important for the tissue repair in the infarcted area (Aisagbonhi et al. 2011). Additionally, *Wnt3a* is reported to activate EndMT in dermal microvascular endothelial cells, which may be involved in the formation of keloids and hypertrophic scars (Lee et al. 2015).

1.3.3 EndMT markers

In addition to using the above mentioned transcription factors as markers for the occurrence of EndMT, the expression of downstream genes is also important in this context. The suppression of VE-cadherin is a good marker for EndMT as it indicates the loss of adherens junctions' integrity in endothelial cells (Frid et al. 2002). The downregulation of CD31, which is also called platelet endothelial cell adhesion molecule-1, and Zonula occludens-1 can serve this mission. They are both expressed in the intercellular junctions, and their inhibition reflects the process of losing the endothelial phenotype. Similarly in case of EMT, the loss of the epithelial phenotype can be shown by the downregulation of Zonula occludens-1 or E-

cadherin, which are important elements in the intercellular junctions in epithelial cells (Newman 1994; Nosedá et al. 2004; Peinado et al. 2004; Roy et al. 2015; Tian and Phillips 2002; Yoshimatsu and Watabe 2011).

The gain of mesenchymal characteristics is also useful indicator for EndMT or EMT, and the upregulation of several genes represent a marker for these processes, such as fibroblast specific protein (FSP)-1, vimentin and alpha-smooth muscle actin (α -SMA). These markers were used in many studies but with a great debate about their specificity and sensitivity, as vimentin can also be positive in endothelial cells; and FSP1, which is a specific marker for fibroblasts in the heart, is not sensitive, as not all cardiac fibroblasts are FSP1-positive including EndMT-derived fibroblasts (Chen et al. 2015; Elliott et al. 2014; Krenning et al. 2010; Zeisberg and Neilson 2009; Zeisberg and Kalluri 2010; Zeisberg et al. 2007a; Zeisberg et al. 2007b).

1.3.4 EndMT and cardiac fibrosis

The role of EndMT in cardiovascular development is well known, as it was shown that the endocardial cells in the atrio-ventricular canal undergo EndMT generating thereby the mesenchymal cells of the heart cushion, which participate in the formation of the cardiac valves and septa (Eisenberg and Markwald 1995; Armstrong and Bischoff 2004). Additionally, its role in the development of embryonic pulmonary artery and aorta was reported (Arciniegas et al. 1989; Arciniegas et al. 2005). However, EndMT is also essential in several pathological processes, including cancer progression, as it is involved in producing carcinoma-associated fibroblasts (Zeisberg et al. 2007a); in addition to its important contribution to fibrogenesis in several organs, such as intestines (Rieder et al. 2011), lung (Hashimoto et al. 2010), kidney (Zeisberg et al. 2008) and heart (Zeisberg et al. 2007b).

Cardiac fibrosis is an essential component in most heart diseases caused by different factors such as ischemia, pressure overload, or cardiomyopathy (Beltrami et al. 1994; Ho et al. 2010; Vogt et al. 1993). Fibroblasts accumulation with their excess secretion of extracellular matrix is the most important factor in scar formation, which sequentially has deleterious effect on the fibrotic heart (Krenning et al. 2010). In the paper of Zeisberg and her colleagues it was reported that 27% to 35% of all cardiac fibroblasts were derived from endothelial cells. That was shown in an animal model for pressure overload using ascending aortic constriction (AAC) to induce heart

fibrosis. The induction of EndMT by TGF- β 1 was also possible in human coronary artery endothelial cells (HCAEC) and could be rescued by BMP-7 (Zeisberg et al. 2007b). BMP-7 was also suggested to be a useful treatment for endocardial fibroelastosis by the inhibition of EndMT, as it was shown as mentioned before that fibroblasts in this fibrotic tissue are derived from endocardial endothelial origin. That was investigated in a model of heterotopic transplantation using the hearts of newborn mice to imitate the reduced flow-state in HLHS (Xu et al. 2015a). Other models for cardiac fibrosis suggested a role for EndMT in hypertrophic cardiomyopathy (Teekakirikul et al. 2010), and diabetes-induced cardiac fibrosis (Widyantoro et al. 2010). This indicates the importance of investigating the potential advantages of EndMT inhibition on the prognosis of different cardiac diseases, as it was found that Irbesartan in a rat model for diabetes was able to reduce EndMT, representing a possible treatment for diabetic cardiomyopathy (Tang et al. 2013).

1.4 KIAA0182 gene

1.4.1 General information

KIAA0182 gene locates on chromosome 16 and consists of 23 exons, mapping to the band (16q24.1) in the position 85170003-85676206 on forward strand according to the second patch release for the human reference assembly from the genome reference consortium (GRCh38.p2), which was released In December 2014. It encodes for a protein called Gse1 Coiled-Coil Protein, which consists in its longest isoform of 1217 amino acids, and GSE is an acronym from the phrase: genetic suppressor element (<http://www.ncbi.nlm.nih.gov/gene/23199>). The sequences of nearly full length transcripts of this gene and other genes isolated from human immature myeloid cell line (KG-1) were first reported by Takahiro Nagase and his colleagues. Their project in Kazusa DNA Research Institute, from which the abbreviation KIAA is derived, aimed to identify the coding sequence of human genes that are still unknown. In their study an alternating Arg and Glu repeat was found in Gse1 Coiled-Coil Protein, which may indicate that GSE1 protein has RNA-binding activity (Nagase et al. 1996; Nagase et al. 2006).

KIAA0182 has at least 12 splice variants; 6 of them are protein coding without overlapping between some isoforms. The transcript Gse1-002 is the longest,

comprising 16 exons with the length of 7495 basepairs (bp). It is almost identical to the transcripts Gse1-001 and Gse1-003, but the most important difference is the absence of the second exon in the last 2 transcripts, which is circularized to form the circular RNA of *KIAA0182*. Both Gse1-001 and Gse1-003 variants contain 15 exons, coding for a smaller proteins with 1144 and 1113 amino acids respectively. The other isoforms represent truncated proteins at the 5' end (Gse1-005 and Gse1-011) or the 3' end (Gse1-006), consisting of 12, 7 or 5 coding exons respectively. The relative expression of these isoforms is still not well known. Several other transcripts are described which contain retained introns only and do not code for any proteins (http://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core:g=ENSG00000131149;r=16:85611409-85676204, Memczak et al. 2013). The RNA of this gene is expressed in all major tissues, with relatively high expression in placenta and thymus, and no expression in breast. Several orthologs for *KIAA0182* can be identified in other species. The similarity ratio of these orthologs to the human gene ranges between 60% in zebrafish and 97% in chimpanzee (<http://www.genecards.org/cgi-bin/carddisp.pl?gene=GSE1>).

The reports about the role for this gene in any human disease are very few. It was shown in one study that the homozygosity of a single nucleotide polymorphism (SNP) within the 3'-UTR of *KIAA0182* (rs709805) was associated with increased risk for colorectal cancer. The studied SNPs in this study were chosen according to several criteria including which mutated genes could have a role in carcinogenesis in the colorectum; which SNPs are predicted to locate in micro RNA (miRNA) binding sites and which miRNAs are expressed in the colorectum (Landi et al. 2012). Another paper described a duplication in chromosome 16, which has the size of 250 kb comprising seven genes (*KIAA0182*, *GINS2*, *c16orf74*, *COX4NB*, *COX4I1*, *MIR1910* and *IRF8*), in a woman with mental retardation, spastic paraplegia, severe epilepsy, a narrow and arched palate, malar hypoplasia, little subcutaneous fat and arachnodactyly (Quéméner-Redon et al. 2013).

1.4.2 Gse1 gene in mouse

The ortholog of *KIAA0182* in mouse is called *Gse1*. It has great similarity to the human gene *KIAA0182* (84% of nucleotides and 88% of amino acids). (<http://www.genecards.org/cgi-bin/carddisp.pl?gene=GSE1>). Mouse *Gse1* locates on chromosome 8 (8 E1) on the forward strand in the position 120230536-120581390

according to the genome reference consortium for mouse 38 patch release 3 (GRCm38.p3) and consists of 20 exons (<http://www.ncbi.nlm.nih.gov/gene/382034>).

Gse1 is described to have at least 6 splicing variants, 4 of them are protein coding. The longest transcript is *Gse1-002*, comprising 7127 bp in 16 exons, coding for 1223 amino acids. The transcripts *Gse1-001* and *Gse1-003* are almost identical to *Gse1-002* transcript, except for small differences in the first coding exon, consisting of 1213 and 1210 amino acids respectively. The fourth coding transcript is very short, comprising 3 exons that code for 16 amino acids only (http://www.ensembl.org/Mus_musculus/Gene/Summary?db=core;g=ENSMUSG00000031822;r=8:120230536-120581390). This gene is also not broadly studied as its human ortholog, and only one study reported that its expression level changes temporarily upon induction by retinoic acid in neurites, suggesting an important role for this gene in neurite outgrowth (Imai et al. 2005).

1.4.3 *KIAA0182* and circular RNA

More than 95% of total RNA is noncoding RNA (ncRNA) (Warner 1999). Circular RNAs that are comprised of circularized exonic sequence (also called scrambled exons) form an important part of this ncRNA, which were described in a lot of previous studies since more than thirty years ago but are still not well understood. These circular RNAs are formed by a process called backsplicing, in which a downstream 5' splice site (splice donor) and an upstream 3' splice site (splice acceptor) are joined together (Hsu and Coca-Prados 1979; Lasda and Parker 2014; Nigro et al. 1991). It was reported using Genome-wide RNA-sequencing analysis that more than 10 % of the transcribed genes in human fibroblasts have backspliced exons, producing together more than 25000 different circular RNAs (Jeck et al. 2013).

It was thought previously that the expression level of all circular RNA isoforms in human is very low compared with the canonical linear RNAs, but recent studies have shown that the circular RNA represents the dominant form in many genes with a different circular/total RNA ratio. This ratio ranged in the case of *KIAA0182* between more than 50% in 3 leukocyte cell types (naive B cells (CD19+), hematopoietic stem cells (CD34+) and neutrophils) and almost 100% in other 3 human cell lines (BJ-T, HEK293, and HeLa). The circular RNA of *KIAA0182* consists of one exon of 219 bp length (the second exon in the isoform *Gse1-002*), generated by splicing the donor

site to the acceptor site of this exon (Jeck et al. 2013; Salzman et al. 2012; Starke et al. 2015). It was found that this circular RNA is more stable against actinomycin D treatment compared with the linear isoforms and not translated (Jeck et al. 2013). It is noteworthy to mention here that *KIAA0182* circular RNA is conserved in mouse, where it is formed from one exon corresponding to the same circularized exon in human. This conservation was also described for 69 different circular RNAs, and can raise a question about a conserved characteristic in the pre-mRNA that enhances backsplicing, suggesting the presence of an important function for this circular RNA (Jeck et al. 2013; Lasda and Parker 2014; Memczak et al. 2013).

Another important question to be asked in this context is whether all these circular RNAs detected by RNA-sequencing are really circles. Several criteria are suggested to prove that, which includes reverse transcription polymerase chain reaction (RT-PCR) using outward-facing primers or Northern blot using a probe spanning the scrambled exonic junction. The second criterion for circularity can be the resistance of circular RNAs to exonuclease activity, represented in several studies by RNase R treatment. Another criterion that would also be useful is that the detected scrambled exons do not have the characteristics of linear messenger RNAs (mRNAs), such as polyadenylation (Lasda and Parker 2014). It is nowadays not well known, what exact roles the circular RNAs can play, which were previously regarded as transcriptional noise or RT-PCR artifacts only. According to the available literature, the first well proved function is their role as micro RNA (miRNA) sponge, taking advantage of their high stability to compete with the mRNAs targeted by this miRNA (Hansen et al. 2013; Memczak et al. 2013).

1.4.4 KIAA0182 and CoREST complex

The exact function of Gse1 Coiled-Coil Protein is not known until now, but according to few studies it is described to have a leucine-zipper domain and it may participate in a large multi-proteins complex named as CoREST complex, or BRAF35-HDAC complex (BHC) as it is called in some publications (Hakimi et al. 2003; Yang et al. 2011; Yokoyama et al. 2008). CoREST complex is one of several histone modifying complexes, and it regulates the expression of several genes through modulating two distinct processes, which are histone acetylation and methylation. Histone acetylation on lysine residues, mostly within the N-terminal tail of the histones, causes gene activation by reducing the positive charge of the histone, making thereby the

chromatin more open. This is regulated by two groups of enzymes: histone acetyltransferases and histone deacetylases (HDACs). Histone methylation on lysine or arginine residues is associated with both gene activation or inactivation, depending on which lysine residue is methylated, and this process is controlled by histone methyltransferases and histone demethylases enzymes (Delcuve et al. 2012; Hayakawa and Nakayama 2011; Kouzarides 2007; Lakowski et al. 2006).

The essential components of CoREST/BHC complex as often described in most reviews are the class I HDACs, HDAC1 and/or HDAC2, that have 82% similarity and usually exist in homo- or heterodimer; LSD1 that is also named as BHC110 and has the function of demethylating dimethylated lysine 4 in histone H3 (H3K4me₂); and CoREST protein that is also named as the REST corepressor1 and has ELM2 domain and two SANT domains, which are essential for the interaction with HDAC1 and hypoacetylated histone tails to stimulate thereby the demethylation activity of LSD1 (Brunmeir et al. 2009; Delcuve et al. 2012; Lakowski et al. 2006; Shi et al. 2005; You et al. 2001; Yu et al. 2003). Additional components can include among others: BHC80, which is also called PHF21A and may negatively regulate the activity of other components (Shi et al. 2005); BRAF35 (HMG20B), which is mentioned to have HMG domain with a binding ability to DNA that is necessary for the repressive role of the complex (Hakimi et al. 2002).

CoREST complex was first mentioned to be recruited to a DNA site consisting of 21–23 nucleotides called repressor element 1 (RE1) by RE-1 silencing transcription factor (REST). It should play an important role in suppressing the expression of neuronal genes in non-neuronal terminally-differentiated cells. It can also play an activating role in embryonic stem cells and neural stem cells using help from other proteins, such as histone methyltransferases and iBRAF (HMG20A). This protein is found to be expressed in the brain and can compete with BRAF35 for binding with RE1-containing genes (Delcuve et al. 2012; Lakowski et al. 2006; Wynder et al. 2005). Another important transcription factor, SNAIL, was found recently to be able to recruit CoREST complex through its SNAG domain to suppress the target gene, E-cadherin, inducing thereby epithelial-mesenchymal transition (EMT) (Lin et al. 2010). Similarly, KIAA0182 and other components of CoREST complex were reported to be recruited by the SNAG domain of the transcriptional factor Insm1 in AtT-20 cells, and that is important for the differentiation of endocrine cells. It was suggested in this

study that KIAA0182 plays a regulatory role for this complex (Welcker et al. 2013). All this can argue for the presence of several other functions for this complex according to its partners, the transcriptional factors involved and the targeted genes.

1.4.5 KIAA0182 and cardiovascular diseases

KIAA0182 could be involved in the pathogenesis of HLHS, as it was found to be heterozygously *de novo* mutated in a child diagnosed with this disease. The missense T-C mutation was located in exon 9 from the isoform Gse1-002 causing the amino acid leucine to be replaced by proline. This mutation is not one of the mentioned SNPs for *KIAA0182*, which count more than 2200 SNPs (<http://www.genecards.org/cgi-bin/carddisp.pl?gene=GSE1>). In silico prediction for the mutated protein could not decide whether it is a loss-of-function- or gain-of-function-mutation (unpublished data from lascone lab, Bergamo). The importance of *KIAA0182* in embryonic cardiac development was further confirmed by an experiment performed on zebrafish, in which it was shown that *KIAA0182* knockout was associated with clear cardiac phenotype, represented by reduced size of the ventricle and enlarged atrium (unpublished data from Marc-Phillip Hitz, Sanger institute).

Additional evidence can be extrapolated from the data about heterozygous mice produced for the Europhenome Mouse Project, which has one allele mutated due to the insertion of a trapping cassette in *Gse1* gene, and this is supposed to create a reporter knockout allele. Those mice were found to have hearts with greater weight compared with wild type animals in both genders. Other abnormalities were detected also in these mice, such as higher systolic blood pressure; in addition to higher pulse rate in males, and lower body weight and lower bone density in females (<http://www.europhenome.org/databrowser/viewer.jsp?set=true&m=true&x=Both-Split&ln=Gse1&project=All&zygosity=All&m=true&l=10946>).

1.5 Gene trap mutagenesis

1.5.1 Mutagenesis strategies

To study the role of one mutation in the pathogenesis of any human disease, or to determine the function of one unknown gene several ways can be followed. Mutagenesis represents a frequently used method to achieve this purpose. This

process can be defined as the modification of the genomic DNA of the organism in a stable way to get a mutation, which was described for the first time more than 80 years ago (Muller 1927).

Old methods to induce mutagenesis included using X-ray radiation or applying chemical mutagenesis by chlorambucil, but these strategies have the problem that several genes can be affected, which makes studying one single gene or mutation so difficult (Russell et al. 1989; Stanford et al. 2001). In contrast with previous methods, ethylnitrosourea (ENU) can cause point mutations or small deletions (20–50 bp) in spermatogonial stem cells, but it has the disadvantage that it gives no landmark for identifying the mutated genes (Russell et al. 1979; Stanford et al. 2001). The first report about using retrovirus to introduce exogenous DNA into the mouse germ cells was almost 40 years ago, and it was possible using this method to recover the affected genes, opening the road for the wide use of insertional mutagenesis (Jaenisch 1976; Spence et al. 1989). The first transgenic mouse was produced in 1981 by the microinjection of DNA into fertilized oocytes (Gordon and Ruddle 1981; Wagner et al. 1981), but identifying the affected gene after that was still not easy, which made the strategy of homologous recombination in embryonic stem cells (ES) preferred in order to mutate a specific gene (Stanford et al. 2001).

1.5.2 Gene trapping

This strategy can help to produce embryonic stem cells (ES) with random mutations by the inserting of a trapping vector into the genomic DNA using electroporation or retroviral infection. The reporter after the splicing acceptor inside the trapping vector gives a signal indicating its presence in a transcriptionally active gene, and mutating thereby this gene by disrupting the splicing process, taking advantage of the polyA tail after the reporter. The trapped gene can be identified by sequencing the mRNA product using a technique named as rapid amplification of cDNA ends and primers located in the trapping cassette. The trapping vector may not be completely successful to inactivate the affected gene and hypomorphic allele, rather than a null allele, can be generated due to the occurrence of alternative splicing, especially when the trapping vector is inserted into an intron. Different vectors can be used, with variable characteristics and efficacy, which can be classified basically into 3 types (Stanford et al. 2001; <http://www.genetrap.org/tutorials/overview.html>).

The enhancer-trap vector includes inducible minimum promoter that needs to be inserted in an intronic region near to a cis-acting enhancer element (the enhancer of affected gene), which derives the expression of the reporter gene β -galactosidase (LacZ). It makes the affected gene usually hypomorphic, which made it not frequently used. The gene-trap vector produces a fusion transcript between the upstream exon of the mutated gene and the promoterless reporter LacZ by the insertion of a trapping cassette that contains a splice acceptor (SA) upstream of LacZ in an intronic region also. In contrast to that, the promoter-trap vector is inserted into an exon of the affected gene producing thereby a fusion protein with the LacZ reporter. It should be always kept in mind that the mutated protein generated by these different vectors may still be functional depending on the location of its domains (Stanford et al. 2001).

1.5.3 The 'Knockout-first' strategy

Several modifications can be applied on the previous trapping vectors, such as the integration of homologous recombination sites, facilitating thereby specific genes to be targeted. The mutated allele can also be further altered so that it can be reverted back to wild type phenotype then reverted again to the null allele-state using different systems, such as Cre- and FLP-Recombinase systems, which recognize system-specific sites inside the trapping cassette. The mechanisms of these two systems are similar, including DNA recombination in an irreversible way by strand cleavage, exchange and ligation. The targeted sites, LoxP (locus of crossover (x) in P1) and FRT (FLP-recombinase recognition target), share common structure, as they consist of two inverted repeats of 13 bp size, flanking an asymmetrical core of 8 bp (Branda and Dymecki 2004; Skarnes et al. 2011; Stanford et al. 2001; Testa et al. 2004).

The 'Knockout-first' allele (tm1a as named mostly), whose design is based upon the structure of gene-trap vector, takes advantage of these previous advancements and it is proposed, as its name suggests, to behave as a null allele from the beginning without any further modifications. As it is possible to convert it into a conditional allele (tm1c) by FLP-recombinase in ES cells or by breeding with transgenic FLP mice, restoring thereby its normal function as wild type allele, this system can possess the characteristics of knockout and conditional alleles in one mouse. The conditional allele after that, taking advantage of Cre-recombinase, can be reverted into true null allele (tm1d) by deleting the critical exon. This is achieved through the generation of frame-shift mutation and nonsense-mediated decay of the mutated protein, which can

be controlled temporarily and spatially according to the Cre mice used (Figure 1.1). LacZ-tagged null allele (tm1b) can also be obtained directly from the ‘Knockout-first’ allele by Cre-recombinase. The trapping cassette is inserted in one of the introns of the gene of interest, which should avoid causing deletion of regulatory elements in that region, and this needs a lot of efforts to understand the structural details of the targeted gene. Computer programs are usually used applying algorithms that predict the most suitable site for the insertion of the trapping cassette and the homologous recombination strategy around the critical exon. This exon is usually chosen to be the 5'-most exon that is common to all mRNA isoforms and its deletion can cause disruption of at least 50% of the protein structure. These designing criteria are only applicable in 60 % of protein-coding genes (Skarnes et al. 2011; Testa et al. 2004).

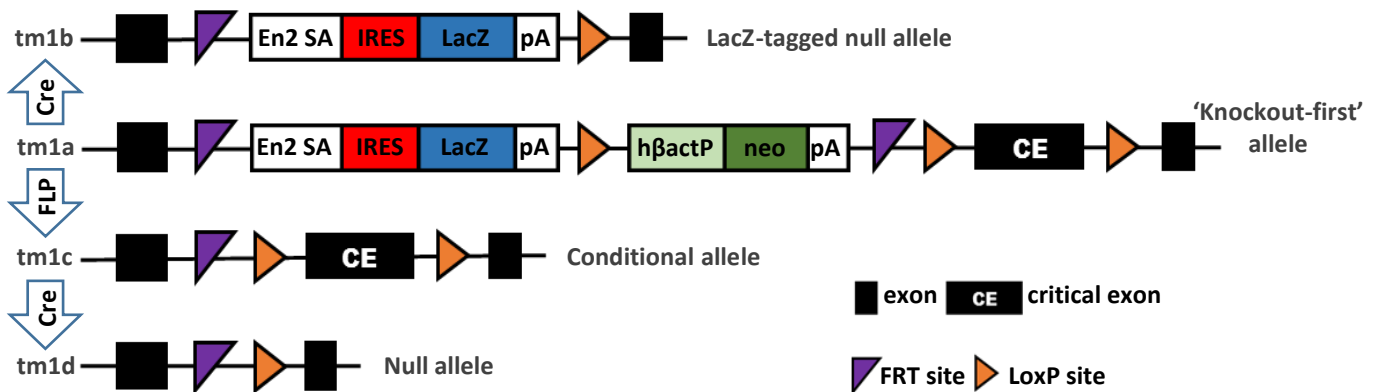


Figure 1.1: The ‘Knockout-first’ allele structure and its possible allelic series (Adapted and modified from Ryder et al. 2013).

This allele is the most common type of mutated alleles in the collection of embryonic stem cells (ES) of the European conditional mouse mutagenesis program and the knockout mouse program (EUCOMM/KOMP-CSD), which represents the main source for generating targeted mutations in mice by the International Mouse Phenotyping Consortium (IMPC). The aims of this program are to produce knockout models for all protein-coding genes in the mouse (Brown and Moore 2012; Ryder et al. 2013; Skarnes et al. 2011), and the ‘Knockout-first’ mice were recently used in several studies as a knockout model without further changes by FLP or Cre mice (Maguire et al. 2014; Nijnik et al. 2012; Rainger et al. 2011; Wheway et al. 2013).

2. Materials and methods

2.1 Materials

2.1.1 Animals

In this study, we used mice generated from the embryonic stem cell (ES) resources of EUCOMM/KOMP-CSD (cell clone: EPD0557_2_C07, project number: 71610) with a C57BL/6NTac background. The mutated allele in these mice ($Gse1^{tm1a(EUCOMM)Wtsj}$) has the design of 'Knockout-first' allele, and it is produced by the insertion of a trapping cassette (L1L2_Bact_P) in *Gse1* gene using homologous recombination strategy. This allele in its basic state is coded $Gse1^{tm1a}$ and it includes beta-galactosidase (LacZ) cassette with a polyA tail downstream from Engrailed 2 splice acceptor (EN2SA) and an internal ribosome entry site (IRES) mediating the initiation of LacZ protein translation with controlled expression by *Gse1* promoter. After that, there is the coding sequence for neomycin resistance driven by the human beta-actin autologous promoter ($h\beta actP$) with another polyA tail, which are separated from the whole EN2SA-IRES-LacZ cassette by a LoxP site. All this sequence is flanked by two FRT sites, and followed by two other LoxP sites flanking the targeted exon (the third exon in the isoform *Gse1*-002) (Figure 2.1). This design is supposed to knockout *Gse1* gene on transcription level due to the presence of the previously described trapping cassette, and the generated protein should be truncated including only the first 85 amino acids of the wild type protein translated from the longest isoform *Gse1*-002. Several quality control tests had been performed on the ES cells by the providing company including 5' end long-range PCR, 3' end long-range PCR, LoxP sites confirmation, karyotyping and Southern blot. All information about this allele can be found in this website: <http://www.knockoutmouse.org/martsearch/project/71610>.

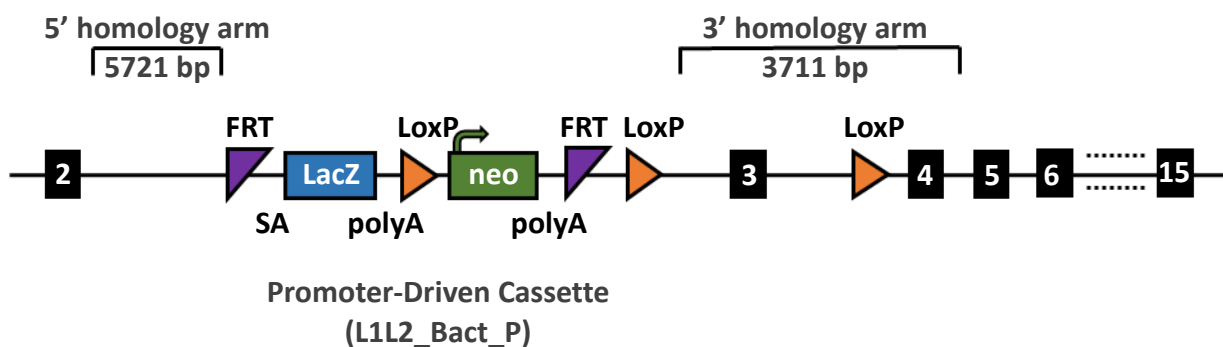


Figure 2.1: The design of the mutated allele $Gse1^{tm1a(EUCOMM)Wtsi}$ (Adapted and modified from: <http://www.knockoutmouse.org/martsearch/project/71610>).

This design, as described before, enables also the pre-conditional allele $Gse1^{tm1a}$ to be converted to the conditional allele $Gse1^{tm1c}$ by FLP-recombinase. That will remove the whole cassette except for the last two LoxP sites, which can be used later to convert it by Cre-recombinase to the null allele without the targeted exon $Gse1^{tm1d}$. Direct breeding with Cre mice is also an option as mentioned before, producing a null allele tagged with LacZ protein by removing the cassette of neomycin resistance and the flanked exon in the optimal case, which is named as $Gse1^{tm1b}$ allele. However, other possibilities can happen, as only the targeted exon alone or the cassette of neomycin resistance alone may be removed, generating in this case the undesired alleles $Gse1^{tm1b.1}$ or $Gse1^{tm1b.2}$ respectively, which are not consistent with the goals of the knockout mouse project phase 2 (KOMP2) (Ryder et al. 2014). Rosa-FLP-e mice expressing FLP-recombinase that were used in this study were kindly given by Prof. Dr. André Fischer (UMG). Mice breeding and all other investigation were performed according to the standards of the Institutional Animal Care and Use Committee of the University of Göttingen.

2.1.2 Chemicals

Table 2.1: List of the chemicals used in this study in alphabetical order with the providing supplier and the ordering number.

Name	Provider	Ordering number
20X LumiGLO and 20X Peroxide	Cell Signaling	7003
2-Mercaptoethanol	ROTH	4227.3
Acetic acid solution	MERCK	1.37035.2500
Agarose gel	Biozym	840004
Aniline blue solution	Sigma-Aldrich	HT154-250ML
Bouin's solution	Sigma-Aldrich	HT10132-1L
Chloroform	Sigma-Aldrich	C2432-500ML
Collagenase II	Worthington	LS004176
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich	D8418-50ML
Dithiothreitol (DTT) 0.1 M	Invitrogen	y00147

Name	Provider	Ordering number
dNTP Mix 10 Mm	Invitrogen	18427-013
EDTA 25 Mm	Invitrogen	y02353
Ethanol	ROTH	9065.2
Fast SYBR green Master Mix	Applied Biosystems	4385612
Fetal calf serum	Sigma-Aldrich	F4135-500ML
Formaldehyde	ROTH	P087.3
Fuchsin solution	Sigma-Aldrich	HT151-250ML
Glycin	MERCK	1.04201.0100
Isoflurane	Abbvie	B506
Isopropanol	ROTH	6752.2
Lipofectamine 2000	Invitrogen	11668-019
Magnesium Chloride Solution 25 mM	Sigma-Aldrich	M8787-1.5ML
Nuclease-Free Water	QIAGEN	129114
Oligo (dT) 12-18 Primers	Invitrogen	18418-012
Penicillin/Streptomycin solution	Sigma-Aldrich	P4333-100ML
Phosphomolybdic acid solution	Sigma-Aldrich	HT153-250ML
Phosphotungstic acid solution	Sigma-Aldrich	HT152-250ML
Powdered milk	ROTH	T145.2
Random Hexamers 50 µM	Invitrogen	N8080127
TaqMan GTXpress Master Mix	Applied Biosystems	4403311
TGF-β1	R&D Systems	240-B-010/CF
Trizol	Ambion	15596-026
Trypsin-EDTA 0.25%	Sigma-Aldrich	T4049-100ML
Tween 20	ROTH	9127.2
Weigert's iron hematoxylin solution	Sigma-Aldrich	HT107-500ML
Xylol	ROTH	9713.3

2.1.3 Commercial kits

Table 2.2: List of the commercial kits used in this in alphabetical order study with the providing supplier and the ordering number.

Kit	Provider	Ordering number
DNeasy Blood & Tissue Kit	QIAGEN	69506
Phusion High-Fidelity PCR Kit	NEB	E0553L
Pierce™ BCA Protein Assay Kit	Thermo SCIENTIFIC	23225
PureLink™ RNA Mini Kit	Ambion	12183025
QIAquick Gel Extraction Kit	QIAGEN	28706

2.1.4 Cell culture mediums

Table 2.3: List of the cell culture mediums used in this study in alphabetical order with the providing supplier and the ordering number.

Medium	Provider	Ordering number
DMEM	Gibco	41965-039
Endothelial Cell Basic Medium	Genlantis	PM210500
HCAEC Growth Medium	Genlantis	PM212500
OPTI-MEM I (1X)	Gibco	31985-062

2.1.5 Buffers

Table 2.4: List of the buffers used in this study in alphabetical order with the providing supplier and the ordering number.

Buffer	Provider	Ordering number
NP-40 lysis buffer	Invitrogen	FNN0021
NuPAGE LDS Sample Buffer (4X)	Novex	NP0007
PCR buffer without MgCl ₂ 10X	Sigma-Aldrich	P2317-1VL

Buffer	Provider	Ordering number
Phosphate buffered saline (PBS) for cell culture	Gibco	14190-094
Pierce Fast Semi-Dry Transfer Buffer	Thermo SCIENTIFIC	35035
TAE buffer 10X	ROTH	T845.2

2.1.6 Instruments

Table 2.5: List of the instruments used in this study with the providing supplier in alphabetical order according to the function.

Function	Instrument	Provider
Cell culture hood	BIOWIZARD	KOJAIR
Centrifugation	Centrifuge 5424R	Eppendorf
Centrifugation for cell culture	ROTOFIX 32 A	Hettich
DNA and RNA Concentration Measurement	NanoDrop 2000	Thermo Scientific
DNA electrophoresis	Owl Easycast B2	Thermo Scientific
Gel imaging	BioDocAnalyze	Biometra
Heating block	Thermomixer comfort	Eppendorf
Incubator for cell culture	InCu safe	EWALD
Microscope	BX43	OLYMPUS
PCR cyclers	Veriti 96 Well Thermo Cyclers Mastercycler eppgradient S	Applied Biosystems Eppendorf
Protein electrophoresis	XCell SureLock™ Electrophoresis Cell	Invitrogen
Protein transferring	Fastblot B44	Biometra
Real-time PCR cycler	StepOnePlus Real-Time PCR System	Applied Biosystems
Tissue disruption	TissueLyser LT	QIAGEN
Vortexing	Vortex Genie 2™	Bender & Hobein AG

2.1.7 Antibodies

Table 2.6: List of the antibodies used in this study with the providing supplier and the working dilution.

Antibody	Provider	Ordering number	Dilution
Anti-Gse1 (rabbit)	Sigma-Aldrich	SAB2101244-50UG	1:2000
Anti-Gse1 (rabbit)	LifeSpan Biosciences	LS-C178849	1:2000
Anti-β-Actin (mouse)	Sigma-Aldrich	A2228-200UL	1:5000
Anti-α-Tubulin (mouse)	Sigma-Aldrich	T5168-.2ML	1:5000
Anti-Mouse (rabbit)	Dako	P0161	1:5000
Anti-Rabbit (goat)	Dako	P0448	1:2500

2.1.8 Primers

Table 2.7: List of the primers used in this study for genotyping PCR and reverse transcription polymerase chain reaction (RT-PCR) with the sequence and annealing temperature.

Primer	Sequence	Annealing Temperature
Gse1-5arm-WTF	CCGCAGGTAGCATGATTGT	55°C
Gse1-Crit-WTR	GTGGATCGGCCATATTGAGT	
Tm1a-5mut-R	GAACTTCGGAATAGGAACTTCG	
Tm1c-EN2-R	TGTTAGTCCCAACCCCTTCC	55°C
Gse1-F	GGCCCTGGTAGTCATGTGAA	60°C
Gse1-ttR	CTCCAGACTCCATTGACGGT	60°C
Gse1-R	GGAAGGGAAGCTCTATGGGG	60°C
CSD-neoF	GGGATCTCATGCTGGAGTTCTTCG	60°C
CSD-lacF	GCTACCATTACCAGTTGGTCTGGTGTC	60°C
CSD-loxF	GAGATGGCGCAACGCAATTAATG	60°C
Gse1-GR	TGCAGATGCTGAGAAGTGGGAACC	68°C
5FRT_F	AGGCGCATAACGATACCACGAT	55°C
5FRT_R	CCACAACGGGTTCTTCTGTT	

Primer	Sequence	Annealing Temperature
LacZ_2_small_F	ATCACGACGCGCTGTATC	55°C
LacZ_2_small_R	ACATCGGGCAAATAATATCG	
Floxed_PNF	ATCCGGGGGTACCGCGTCGAG	55°C
Floxed_LR	ACTGATGGCGAGCTCAGACC	
FLP-Geno-F	GTGACAGAGACAAAGACAAGCGTTAG	55°C
FLP-Geno-R	AATTGCCGGTCCTATTTACTCGTT	
m.linearRNA-R	CCATTGACGGTTTTAGTGGGG	55°C

Table 2.8: List of the primers used in this study for quantitative real-time polymerase chain reaction (qRT-PCR) with the sequence and the product size. The annealing temperature was 60°C for all primers.

Primer	Sequence	Product Size
h.KIAA0182-F	ACCGTGACCGAGGACTACC	79 bp
h.KIAA0182-R	GGGGAGGCAGTGAGGACAT	
h.GAPDH	Undisclosed (ordered from Primer Design)	
h.Snail-F	GGCAATTTAACAATGTCTGAAAAGG	105 bp
h.Snail-R	GAATAGTTCTGGGAGACACATCG	
h. α -SMA-F	AAGCACAGAGCAAAGAGGAAT	76 bp
h. α -SMA-R	ATGTCGTCCCAGTTGGTGAT	
m.Gse1-F	GCTCTCCCACTCGTCCCT	111 bp
m.Gse1-R	TCCCGCTCTCGCTCTCT	
m.GAPDH-F	TGTAGACCATGTAGTTGAGGTCA	123 bp
m.GAPDH-R	AGGTCGGTGTGAACGGATTTG	
LacZ-F	AAATATGATGAAAACGGCAACC	203 bp
LacZ-R	AACAGGTATTCGCTGGTCACTT	
m.Foxp4-F	ATCGGCAGCTGACGCTAAATGAGA	121 bp
m.Foxp4-R	AAACACTTGTGCAGGCTGAGGTTG	

Primer	Sequence	Product Size
m.Thy1-F	GCCGCCATGAGAATAACA	166 bp
m.Thy1-R	GCTAGGGTAAGGACCTTGAT	
m.Kmt2d-F	CCACACCAGATGGCTGACTT	105 bp
m.Kmt2d-R	GCAGCAGCGACGATTGTTAG	
m.Tfap2b-F	CCCACACGCCATCATCGG	123 bp
m.Tfap2b-R	GCAGAGGGTTCAGGGAGTAG	
m.Actn4-F	GACTACTATGACTCCCACAACG	111 bp
m.Actn4-R	CTCTGTTTTCTCCAAGGCTTCC	
m.Wnt1-F	GCCTCGGAGCCATTGAACA	95 bp
m.Wnt1-R	GCAAATTATTTACACAGTGATGAGGA	
m.circularRNA-F	CGCTCCATCCTCCAGCTTT	117 bp
m.circularRNA-R	GTCGCCGTGGAAAGCATC	
Probe (FAM labeled)	CAGAGGCATGAGCCATGAG	
Gse1-LOA-F	ATACGCACCCGGTACTGGAA	88 bp
Gse1-LOA-R	GCAAGCAACGATGAATTTGTGG	
Probe (FAM labeled)	AGGATTAAGGCAACCCTTACATCAGAC	
Dot1l-ctl-F	GCCCCAGCACGACCATT	68 bp
Dot1l-ctl-R	TAGTTGGCATCCTTATGCTTCATC	
Probe (FAM labeled)	CCAGCTCTCAAGTCG	
m.Usp10-F	GTCATCGAACCTAGTGAGGGG	89 bp
m.Usp10-R	CCAAGGATAAATTCGGGTGCC	
m.Zdhhc7-F	TCGTCTATGCAGACTTCGTGG	142 bp
m.Zdhhc7-R	GGGTCAGTGAGCATGGTTCT	
m.Crispld2-F	GGCCATCCCCATGTCAGAC	104 bp
m.Crispld2-R	CATCCAAGTCATGTGTTCCA	
m.Irf8-F	TGACACCAACCAGTTCATCCGAGA	136 bp
m.Irf8-R	CACCAGAATGAGTTTGGAGCGCAA	

Primer	Sequence	Product Size
m.Cox4i1-F	AGAAGGCGCTGAAGGAGAA	80 bp
m.Cox4i1-R	CTGGATGCGGTACAACACTGAA	
m.Gins2-F	CATGCTTCTGACAACATCCC	146 bp
m.Gins2-R	AAGGTCAGGTTGTCCAGCTT	

2.1.9 Other materials

Table 2.9: List of the materials used in this study and not included in previous tables in alphabetical order with the providing supplier and the ordering number.

Material	Provider	Ordering number
100bp DNA ladder	PROMEGA	G210A
1kb DNA ladder	PROMEGA	G571A
Ad-Cre-IRES-GFP adenovirus	Vector Biolabs	1710
DNA loading dye 6X	PROMEGA	G190A
DNase I	Invitrogen	18068-015
Eukaryotic 18S rRNA Endogenous Control	Applied Biosystems	4310893E
GSE1 Trilencer-27 Human siRNA	Origene	SR308085
Human Coronary Artery Endothelial Cells (HCAEC)	Genlantis	PH30005A
Nitrocellulose blotting membrane	Amersham	10600003
NuPAGE 4-12% Bis-Tris Protein Gel	Novex	NP0335BOX
Protein standards	Invitrogen	LC5800
RNase R	Epicentre	RNR07250
RNaseOUT™	Invitrogen	10777-019
SuperScript™ II Reverse Transcriptase	Invitrogen	18064-014

Material	Provider	Ordering number
SuperScript™ III Reverse Transcriptase	Invitrogen	18080-044
Taq DNA Polymerase	Sigma-Aldrich	D4545

2.2 Methods

2.2.1 Genomic DNA extraction

Genomic DNA was extracted from the tail of newborn mice obtained 4 weeks after birth, or from cultured cells using DNeasy Blood & Tissue Kit according to the manufacturer instructions. In details, the tissue sample was incubated with tissue lysis buffer and proteinase K at 56°C with shaking overnight or until it was completely lysed, or in the case of cultured cells proteinase K alone with PBS buffer were used without incubation, then lysis buffer was added and mixed thoroughly by vortexing. After that, 100 % ethanol was added and mixed, then the whole mixture was pipetted into the spin column, which was centrifuged for one minute then the flow-through was discarded. After two washing steps the precipitated DNA was eluted using AE elution buffer into new micro-centrifuge tube. The DNA concentration and quality were verified after that using NanoDrop 2000 machine, and the samples were stored for further experiments in 4°C fridge.

2.2.2 RNA extraction

RNA was extracted from tissue samples or cultured cells using PureLink™ RNA Mini Kit according to the manufacturer instructions. In details, after an initial homogenizing step in the case of tissue samples using the TissueLyser LT machine in 30Hz for 5 minutes, tissue samples or cultured cells were incubated with one ml of Trizol for 5 minutes at room temperature, then chloroform was added and the tube was shaken vigorously for 15 seconds. After three minutes of incubation at room temperature, the mixture was centrifuged for 15 minutes at 4°C and the aqueous phase that contains the RNA was transferred into another tube. After that, ethanol was added until it reaches the concentration of 35% in the sample and vortexed well.

The sample was then transferred to the spin cartridge and centrifuged for 15 seconds and the flow-through was discarded. After three times of washing and one centrifugation step for one minute to dry the membrane, the spin cartridge was transferred into a recovery tube and the RNA was eluted by adding 30–50 μ l of nuclease-free water and incubating the tube in room temperature for one minute followed by centrifuging for two minutes. After measuring the concentration and verifying the RNA quality by NanoDrop 2000 machine, the sample was kept on ice to be used for further experiments or stored in -80°C fridge. All steps were performed using disposable RNase-free pipette tips and micro-centrifuge tubes. The embryonic RNA samples, which were extracted from the left lower limb of E15.5 embryos, were kindly given by Eva Baier, who is currently working on Gse1 project to investigate the timepoint and the reason of the observed lethality during embryonic development.

2.2.3 RNA reverse transcription

Before cDNA synthesis, RNA samples were treated by 1 μ l of DNase I with 1 μ l of 10X reaction buffer in 10 μ l reaction volume at room temperature for 15 minutes to remove any contamination with genomic DNA, then the reaction was stopped by adding 1 μ l of EDTA 25 mM and heating the sample to 65°C for 10 minutes. After that, 200-1000 ng RNA according to the sample concentration were reverse-transcribed by adding 1 μ l of oligo (dT) primers and 1 μ l of dNTP Mix and incubating the mixtures for 5 minutes at 65°C . Then after chilling the sample on ice for at least one minute, 4 μ l of 5X First-Strand Buffer, 2 μ l of DTT and 1 μ l of RNaseOUT™ were added and mixed gently and incubated for two minutes at 42°C . After that 1 μ l of SuperScript™ II Reverse Transcriptase was added and the sample was incubated at 42°C for 50 minutes, then for 15 minutes at 70°C to stop the reaction.

For circular RNA experiments SuperScript™ III Reverse Transcriptase was used according to the manufacturer instructions, with random hexamers, oligo (dT) primers or both of them as described in the results section (Jeck et al. 2013; Salzman et al. 2012). All cDNA samples were diluted in 1:10 ratio before using them in further experiments, and they were stored in -20°C fridge.

2.2.4 RNase R treatment

Regarding circular RNA, the RNA sample was treated in some experiments with RNase R before the reverse transcription. That was performed using 5-10 μ g of RNA,

which were incubated with 3 Unit/ μg of RNase R and 5 μl of 10X reaction buffer in total reaction volume of 50 μl at 37°C for three hours, then the RNA was recovered using Trizol and PureLink™ RNA Mini Kit as described above. Another sample was prepared with the same steps but without RNase R (0 Unit/ μg of RNase R) as a control for the whole procedure, then 200 ng of RNA for each sample were used for cDNA synthesis and further experiments (Starke et al. 2015).

2.2.5 DNA extraction from agarose gel

DNA was extracted from agarose gel using QIAquick Gel Extraction Kit according to the manufacturer instructions. In details, where indicated in the results sections, the gel slice containing the PCR product of interest was cut and incubated with suitable volume of buffer QG at 50°C until the gel was dissolved, then isopropanol was added to the sample and mixed. The mixture was loaded onto a spin column and centrifuged for one minute with discarding the flow-through. After that, washing with buffer QG then with buffer PE followed by one centrifugation step to get rid of the residual ethanol were performed, then the DNA was eluted by nuclease-free water into a new collection tube. Suitable amount of the extracted DNA depending on the size of the PCR product (18 ng per 100 bases) was used for sequencing after mixing it with 3 μl of sequencing primer diluted in 10 pmol/ μl aliquots in total reaction volume of 15 μl according to the suggested protocol by the sequencing laboratory (Seqlab-Göttingen, <http://www.seqlab.de/index.php?id=93>). The sequencing results were analyzed using the program ApE- A plasmid Editor V2.0.47.

2.2.6 Genotyping PCR

To genotype our ‘Knockout-first’ mice a set of three primers was used. The first primer was Gse1-5arm-WTF which is located in the 5’ homology arm and can give a PCR product of 241 bp from the wild type allele when used with the second primer, Gse1-Crit-WTR (WT PCR1), which is located in the region before the critical exon. The mutated allele Gse1^{tm1a} cannot give a product using those two primers due to the insertion of the trapping cassette, which makes the product too big to be amplified in the conditions used in the genotyping PCR. The third primer, Tm1a-5mut-R, is universal mutant-specific reverse primer that is located in the sequence just after the 5’ homology arm before the first FRT site, and can give with the primer Gse1-5arm-WTF a PCR product of 127 bp from the mutant allele Gse1^{tm1a} and the mutant allele

$Gse1^{tm1c}$ (Mutant PCR1) (Figure 2.2). The PCR was performed in 25 μ l reaction volume using 50-100 ng of genomic DNA, 1 μ l of each one of the three primers after diluting them in 10 pmol/ μ l aliquots, 2.5 μ l of Magnesium Chloride Solution, 2.5 μ l of 10X PCR buffer without $MgCl_2$, 0.5 μ l of DMSO, 1 μ l of dNTP Mix and 0.3 μ l of Taq DNA Polymerase. The PCR conditions were as follows: 1. 95°C 10 min, 2. 95°C 15 sec, 3. 55°C 30 sec, 4. 72°C 30 sec, 5. Steps 2,3, and 4 repeated for 35 cycles, 6. 72°C 7 min, 7. Hold at 4 °C. After that, samples were mixed with 5 μ l of DNA loading dye 6X, then 15 μ l of the mixture were loaded with 100bp DNA ladder on 2.5% agarose gel prepared in TAE buffer 1X, and left to run for 15 minutes at 80 volt then one hour at 120 volt. After that the gel was imaged to determine the genotype and the images were documented.

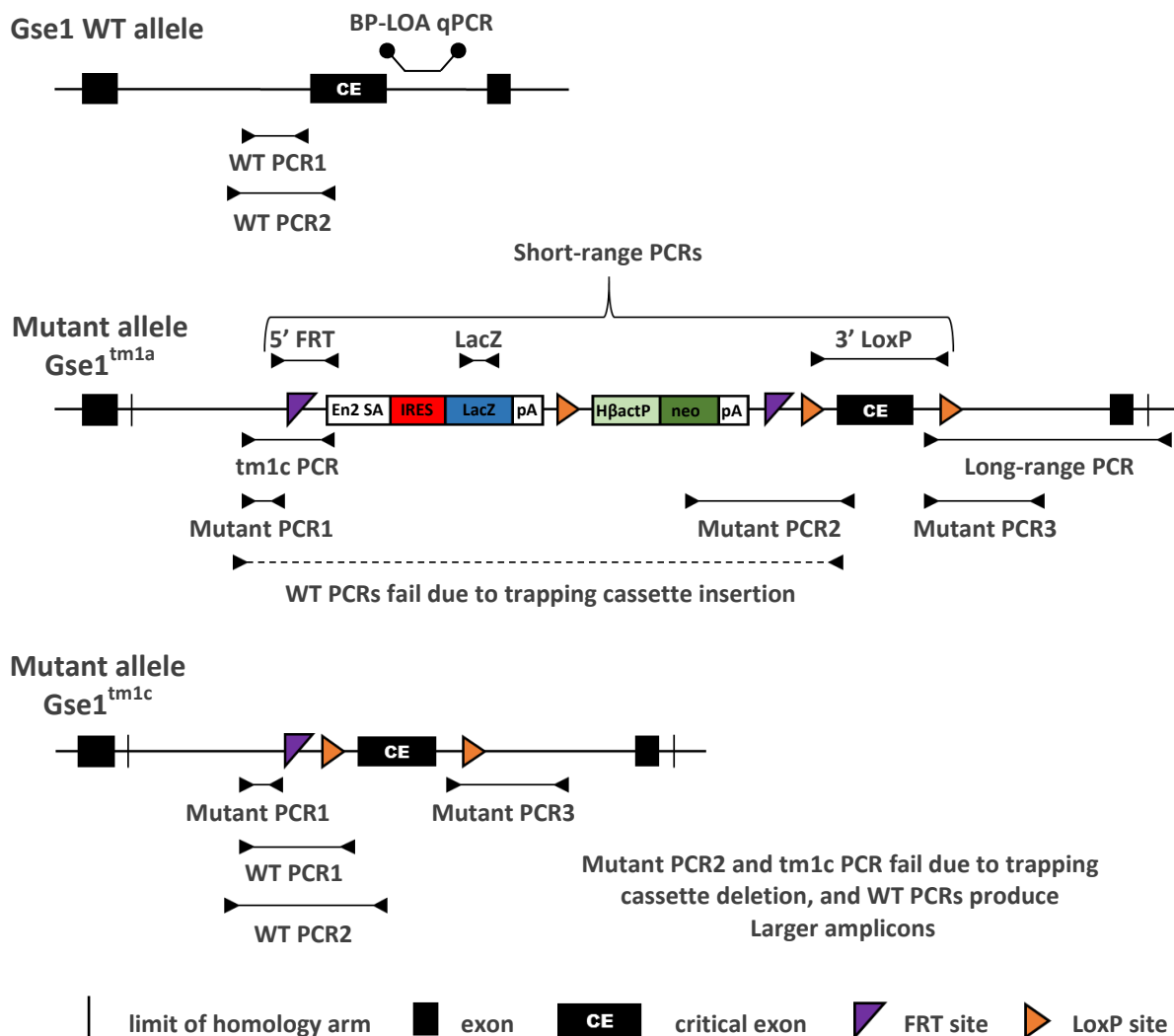


Figure 2.2: The genotyping protocols for $Gse1$ gene and the quality control tests used in this study (Adapted and modified from Ryder et al. 2013).

To genotype the newborn mice after breeding between the 'Knockout-first' mice and mice expressing FLP-recombinase, the same three previous primers were used. In this case, the primers Gse1-5arm-WTF and Gse1-Crit-WTR will give an amplicon from the allele Gse1^{tm1c} but with the size 427 bp, which is larger than the product from the wild type allele due to the presence of the remaining FRT and LoxP sites. Additionally, another primer, the Tm1c-EN2-R, which is located in En2SA sequence, was used in separated PCR reaction with the primer Gse1-5arm-WTF applying the same conditions (tm1c PCR). This reaction will not give a product from the allele Gse1^{tm1c} because the trapping cassette is deleted, but it will give a 231 bp product from the mutant allele Gse1^{tm1a} (Figure 2.2). Another pair of primers was also used to detect the cassette coding for FLP-recombinase, which were FLP-Geno-F and FLP-Geno-R primers. These primers are supposed to give an amplicon of 139 bp if the cassette is inside the genomic DNA of the mouse. This PCR reaction was also performed separately using the same PCR conditions described above.

The genotyping was also confirmed using second protocol with 5 primers: the primers Gse1-F and Gse1-ttR that are located in the 5' homology arm and in the targeted exon respectively, and can give a PCR product of 796 bp from the wild type allele or a larger product of 982 bp from the allele Gse1^{tm1c} (WT PCR2); the universal mutant-specific primer CSD-neoF, which is located in the cassette of neomycin resistance and can give with the primer Gse1-ttR a product of 1074 bp from the mutant allele Gse1^{tm1a} only, but not from the wild type allele or the mutant allele Gse1^{tm1c} (Mutant PCR2); the universal mutant-specific primer CSD-loxF and the primer Gse1-R, which are located in the region of the third LoxP site and in the 3' homology arm respectively and can give a product of 1003 bp from both mutant alleles Gse1^{tm1a} and Gse1^{tm1c}, but not from the wild type allele (Mutant PCR3). The PCR steps were as described before for the first protocol except that three PCR reactions were performed separately, and the annealing temperature was 60°C with one minute as extension time, which is also not enough to get an amplicon from the mutant allele Gse1^{tm1a} using the primers Gse1-F and Gse1-ttR, and the samples were loaded after that on 1.5% agarose gel with a suitable DNA ladder.

2.2.7 Short-range PCR

To confirm the structure of our trapping cassette, several short-range PCR reactions were performed detecting different parts of this cassette. The following primers were

used: the primers 5FRT_F and 5FRT_R for the 5' end of the cassette that should give an amplicon of 204 bp; the primers LacZ_2_small_F and LacZ_2_small_R for the LacZ cassette that should give an amplicon of 108 bp; and the primers Floxed_PNF and Floxed_LR for the region of 3' LoxP that should give an amplicon of 1078 bp (Figure 2.2) (Ryder et al. 2013). These 3 PCR reactions were performed separately applying the PCR conditions and the following steps as in the first genotyping protocol for the first two pairs of primers, and as in the second genotyping protocol for the third pair. The annealing temperatures for these reactions are described in the primers list.

2.2.8 Long-range PCR

This PCR was performed to prove that our gene is correctly targeted (Figure 2.2) (Ryder et al. 2013). In order to do that a primer specific for *Gse1* located after the 3' homologous region, the primer Gse1-GR, was used with universal mutant-specific primer, CSD-loxF, and these two primers are supposed to give a product of 4363 bp when the trapping cassette is in the right position. The reaction was performed using Phusion High-Fidelity PCR Kit with applying the following protocol. 200 ng of genomic DNA, 2.5 μ l of each one of the two primers after diluting them in 10 pmol/ μ l aliquots, 10 μ l of 5X Phusion GC Buffer, 2.5 μ l of DMSO, 1 μ l of dNTP Mix and 0.5 μ l of Phusion DNA Polymerase were mixed in 50 μ l reaction volume. Phusion Polymerase was added at last to prevent primer degradation caused by its 3' \rightarrow 5' exonuclease activity, and then the mixture was rapidly incubated in a pre-heated thermocycler to the denaturation temperature. The PCR conditions were as follows: 1. 98°C 3 min, 2. 98°C 15 sec, 3. 68°C 30 sec, 4. 72°C 4 min, 5. Steps 2,3, and 4 repeated for 35 cycles, 6. 72°C 10 min, 7. Hold at 4 °C. After that, 10 μ l of DNA loading dye 6X were added to each sample and mixed, then 30 μ l of the mixture were loaded on 0.8% agarose gel with 1kb DNA ladder.

2.2.9 Reverse transcription PCR

Reverse transcription PCR (RT-PCR) was performed in some experiments to prove the circularity of the RNA detected by the primers m.circularRNA-F and m.circularRNA-R, which are supposed to give an amplicon with the size of 117 bp. After cDNA synthesis by SuperScript™ III Reverse Transcriptase, 2 μ l of the diluted cDNA were used in 25 μ l reaction with 1 μ l of each primer after diluting them in 10

pmol/μl aliquots, 2.5 μl of Magnesium Chloride Solution, 2.5 μl of PCR buffer without MgCl₂ 10X, 1 μl of dNTP Mix and 0.3 μl of Taq DNA Polymerase. The PCR conditions were as follows: 1. 95°C 3 min, 2. 95°C 15 sec, 3. 55°C 30 sec, 4. 72°C 20 sec, 5. Steps 2,3, and 4 repeated for 45 cycles, 6. 72°C 7 min, 7. Hold at 4 °C, and the PCR products were loaded with 100bp DNA ladder on 2.5% agarose gel.

2.2.10 Quantitative real-time PCR

After diluting the cDNA in 1:10 ratio, 5 μl were used for each reaction, mixed with 10 μl of SYBR Green Master Mix, 1 μl of the forward primer and similar amount of the reverse primer after diluting them in 4 pmol/μl aliquots in total volume of 20 μl per reaction. The fast protocol with the following PCR conditions was used: 1. 95°C 20 sec, 2. 95°C 3 sec, 3. 60°C 30 sec, 4. Steps 2 and 3 repeated for 40 cycles. Melting curve was added at the end of the PCR reaction to ensure the specificity of the primers. The results were analyzed using $\Delta\Delta C_T$ method, where the expression of the gene of interest in the tested samples was normalized to another housekeeping gene (*GAPDH*) and then compared with one control sample.

For the wild type break point loss of allele (BP- LOA) assay, which was performed in order to confirm the correct position of the trapping cassette in our gene, qRT-PCR with TaqMan probe was used (Figure 2.2) (Ryder et al. 2013). This experiment depends on the fact that some loss in the genomic DNA of the targeted gene will happen due to the insertion of the trapping cassette, and the sequence of this deleted region will not exist in the mutated allele, provided that the desired gene is correctly targeted, and the probe is designed to span this lost sequence. In each reaction for this assay 100 ng of genomic DNA were used with 5 μl of TaqMan master mix, 0.4 μl of the FAM-labeled probe diluted in 5 μM concentration, 0.9 μl of the forward primer and 0.9 μl of the reverse primer after diluting them into 10 pmol/μl aliquots in total reaction volume of 10 μl. *Dot1L* gene was used as internal control. The results were analyzed using $\Delta\Delta C_T$ method as described before, and the PCR conditions were as follows using the standard protocol: 1. 50°C 2 min, 2. 95°C 10 min, 3. 95°C 15 sec, 4. 60°C 60 sec, 5. Steps 3 and 4 repeated for 45 cycles.

For analyzing the expression of *Gse1* circular RNA, TaqMan probe covering the junction between the beginning and the end of the circularized exon was used (Jeck et al. 2013). In each reaction 2 μl of cDNA were used with 5 μl of TaqMan master

mix, 0.5 μ l of the FAM-labeled probe for the circular RNA diluted in 5 μ M concentration, 0.5 μ l of VIC-labeled probe together with primers for 18S ribosomal RNA as housekeeping gene, 1 μ l of the forward primer (m.circularRNA-F) and 1 μ l of the reverse primer (m.circularRNA-R) after diluting them into 10 pmol/ μ l aliquots in total reaction volume of 10 μ l. The PCR conditions were the same as for the (BP-LOA) assay, and the results were analyzed using $\Delta\Delta C_T$ method.

2.2.11 Isolation of mouse fibroblasts

Mouse fibroblasts were isolated from adult mouse kidneys using the following protocol: isoflurane anesthesia then cervical dislocation were performed to the adult mice then kidneys were isolated (among other organs such as heart, lung and liver, which were used for other experiments) and washed well with PBS. After that they were cut into very small pieces that can pass easily through 10-ml cell culture pipet tip then were incubated with 3-5 ml of Collagenase II (diluted to the concentration of 5 mg/ml in PBS) for 15 minutes at 37°C with repeated mixing every three minutes. After that they were centrifuged for 5 minutes using 1000 rpm velocity and the supernatant was discarded. Then after one step of washing with PBS and discarding the supernatant by centrifuging, they were incubated with three ml of Trypsin for 5 minutes at 37°C. After that the trypsin was neutralized by similar amount of complete growth medium supplemented with 20% fetal calf serum (FCS). After centrifugation, the supernatant was discarded and the pellet was suspended in 5 ml of complete growth medium and cultured in 10-cm dish in the cell culture incubator, then the medium was changed in the next day and every two days thereafter.

2.2.12 Cell culture

HCAEC cells were cultured in T75 flasks or 10-cm dishes using 10 or 6 ml of HCAEC Growth Medium respectively, which was supplemented with 1% Penicillin/Streptomycin and changed every two days, and the cells were incubated at 37°C with 5% CO₂ in a humidified incubator. The cells were passaged every 5-7 days in 1:2 ratio. That was done by washing with phosphate buffered saline (PBS) for cell culture then incubating with 4 ml of trypsin at 37°C for 2-3 minutes, then the trypsin was neutralized by equivalent amount of growth medium and collected in 15-ml falcon tube and centrifuged for 5 minutes at 1300 rpm. After that, the supernatant was discarded and the cell pellet was suspended in 1 ml of growth medium and

divided into the two new flasks. All the experiments in this study were performed using HCAEC cells of early passages (passage 3-6). Primary mouse fibroblasts were cultured in DMEM medium supplemented with 20% fetal calf serum and 4% Penicillin/Streptomycin and passaged every 4-5 days in 1:3 ratio.

2.2.13 EndMT assay

To induce EndMT in HCAEC, an established protocol in our lab was used (Xu et al. 2015a). In this protocol TGF- β 1 was used, which is described in previous studies to be able to activate EndMT in endothelial cells (Moonen et al. 2010; Zeisberg et al. 2007b). 1.2×10^5 cells were cultured in one well of 6-well plate using HCAEC Growth Medium, and in the next day the medium was changed into Endothelial Cell Basic Medium to starve the cells. After 24 hours, the medium was changed again into HCAEC Growth Medium diluted in Endothelial Cell Basic Medium in 1:10 ratio, and 10 ng/mL of TGF- β 1 was added. The treatment was repeated every two days. Cells incubated in the same medium without TGF- β 1 were used as a control. Samples were collected for RNA isolation after 2 and 6 days. RNA samples from mouse cardiac endothelial cells (MCEC) were kindly given by Elisa Sanchez Sendin, who established the protocol of EndMT assay for this cell line in our lab.

2.2.14 Small interfering RNA transfection

HCAEC were transfected with small interfering RNA against *KIAA0182* using Lipofectamine 2000 according to the manufacturer protocol. 1.5×10^5 cells were cultured in one well of 6-well plate 24 hours before the transfection using HCAEC growth medium without antibiotics. In the next day 50 pmol of each of the three provided small interfering RNAs (siRNAs) were diluted together in 250 μ l of OPTI-MEM medium. Three μ l of Lipofectamine 2000 were also diluted in 250 μ l of OPTI-MEM medium and mixed well, then after 5 minutes of incubation in room temperature both previous dilutions were combined together. During the following 20 minutes of incubation in room temperature the growth medium for the cells was changed into endothelial cell basic medium, then the Lipofectamine-siRNA-OPTI-MEM mixture was added to the cells and mixed gently. After 4 hours of incubation the medium was changed back into HCAEC growth medium then the cells were incubated again. Samples were collected for RNA isolation after 24, 48 or 72 hours. All previous steps were optimized using different amounts of OPTI-MEM medium, Lipofectamine and

KIAA0182-siRNA; different confluency of the cells and different incubation period with the Lipofectamine-siRNA-OPTI-MEM mixture to get the highest possible efficiency for the transfection with the lowest cell toxicity. Cells transfected with equivalent amount of scrambled siRNA were used as a control.

2.2.15 Transduction of primary fibroblasts with Cre-recombinase adenovirus

In order to prove the conditionality of our 'Knockout-first' mice, 1×10^5 fibroblasts isolated from the kidneys of adult mice from different genotypes (wild type or heterozygous) were seeded in one well of 6-well plate using complete growth medium. In the next day after washing with DMEM medium without any serum, the cells were incubated in one ml of basic medium after adding of 5 μ l of Ad-Cre-IRES-GFP adenovirus achieving thereby a multiplicity of infection (MOI) of 500 according to the manufacturer instructions. After 24 hours the cells were washed with PBS and two ml of complete growth medium were added. The transduction efficiency was estimated in the next day depending on the signal of green fluorescent protein (GFP) in the transduced cells, and then genomic DNA was extracted from those cells as described above to be used in further experiments.

2.2.16 Protein extraction

NP-40 lysis buffer was used to extract protein from tissue samples. TissueLyser LT machine was used for homogenizing then the tissue samples were incubated in 200-400 μ l of lysis buffer for 40 minutes on ice with vortexing every 10 minutes for 10 seconds. After sonication for 1.5 minutes (15 seconds on alternatively with 15 seconds off) the lysate was centrifuged for 10 minutes at 4°C, then the supernatant was collected and used for further experiments or stored in -80°C fridge. Pierce™ BCA Protein Assay Kit was used to determine the concentration of the protein samples according to the manufacturer instruction.

2.2.17 Western blotting

10-20 μ g of each protein sample with equal volume of 2X sample buffer, which was prepared by mixing 2.5 ml of NuPAGE LDS Sample Buffer (4X) with 1.5 ml of water and one ml of 2-Mercaptoethanol, were loaded into the wells of NuPAGE 4-12% Bis-Tris Protein Gel with protein standards after denaturation step at 95°C for 5 minutes followed by brief centrifugation. The gel was left to run for 15 minutes at 80 volt then

for 75 minutes at 120 volt, and after that the proteins were transferred to nitrocellulose blotting membrane using Pierce Fast Semi-Dry Transfer Buffer diluted 10 times in water. That was done by running in the transfer chamber for 15 minutes at 150 mA then for 45 minutes at 250 mA after confirming that no air bubbles were trapped. Afterwards, the membrane was blocked by incubating with 5% powdered milk diluted in PBST (PBS with 1% Tween 20) for one hour at room temperature then incubated with the primary antibody diluted in the blocking buffer at 4°C overnight. After three times of washing for 10 minutes each with 2% powdered milk in PBST, the membrane was incubated with HRP-linked secondary antibody diluted in the blocking buffer at room temperature for one hour then washed three times again with 2% powdered milk in PBST and one time with PBS for 5 minutes. After that the membrane was developed by applying the chemiluminescent substrate (20X LumiGLO and 20X Peroxide in 1:1 ratio diluted to 1X in water) and imaged. When necessary, the membrane was stripped by incubating with Glycin (0.1 mM solution adjusted to PH=3) for 30 minutes at room temperature, then after washing for 5 minutes with PBS it was blocked again and incubated with another primary antibody.

2.2.18 Ascending aortic constriction (AAC)

This operation was used in several previous studies as a mouse model for pressure overload on left ventricle mimicking aortic stenosis (Jiang et al. 2007; Takahashi et al. 2007; Zeisberg et al. 2007b). It is possible to make the constriction on different levels, such as the ascending, transverse or descending aorta, but the constriction of the ascending aorta is reported to cause significant and fast pressure overload on the left ventricle producing thereby a high degree of hypertrophy after 48 hours (Tarnavski et al. 2004). All the operations in our study were kindly performed by Julia Steinbrecher, and age-matched male animals were always operated. In short details, after the mice were anesthetized and ventilated, a left thoracotomy was performed then a ligation around the ascending aorta was applied using a silk suture and a 23-gauge blunted needle, then the needle was removed. In the case of sham operation, all steps were the same, except that no ligation was made. The gradient across the constriction site was measured two days after the operation, and detailed echocardiography for the operated mice was kindly performed by Marcel Zoremba, Roland Blume and Beate Knocke after 7 days. The animals were sacrificed two weeks after the operation by cervical dislocation after isoflurane anesthesia. Samples from the heart were taken to

be used later for RNA and protein extraction, and after washing with PBS the hearts were fixed in 10% formaldehyde followed by paraffin embedding.

2.2.19 Masson's trichrome staining

This staining method has been used in a lot of studies to quantify the fibrosis induced by different mechanisms in different organs, such as liver, kidney and heart, where the collagen fibers will be stained in blue color and the nuclei in black with red background (Amin and Mahmoud-Ghoneim 2011; Tampe et al. 2014; Xu et al. 2015b). The established protocol for this staining in our lab was as follows: transverse sections with 3 μm thickness were prepared from the embedded hearts, then the sections were deparaffinized and rehydrated by immersing them two times in xylol for 10 minutes each, two times in 100% ethanol for 5 minutes each, one time for 5 minutes in 95% ethanol, one time for 5 minutes in 80% ethanol, one time for 5 minutes in 50% ethanol and finally one time for 5 minutes in 30% ethanol. After washing in distilled water, the samples were immersed in Bouin's solution, which was prewarmed to 56°C, for 15 minutes. After that, they were left to cool down in normal water for 5 minutes, and then washed using running tap water for 5-10 minutes until the yellow color was removed. Immersion in Weigert's iron hematoxylin solution for 5 minutes followed, then washing again by running tap water for 5 minutes. After that, three times of washing in distilled water were done, followed by immersion in acid Fuchsin solution for 5 minutes then three times of washing with distilled water were repeated. The sections after that were immersed in phosphomolybdic-phosphotungstic solution (prepared by mixing one volume of both acids with two volumes of water) for 5 minutes, followed by 10 minutes of immersion in Aniline blue solution. Then the sections were rinsed for very short time once in acetic acid solution, once in distilled water, twice in 96% ethanol, twice in 100% ethanol and at the end twice in xylol for 5 minutes each. After that, the sections were covered by cover slips and left to dry overnight. The fibrosis was quantified microscopically using CellSens Dimension 1.6 computer program, where the ratio of fibrosis was defined as the fibrotic area / the area of the myocardium in the left ventricle.

2.2.20 Statistical analysis

Statistical analysis and figures production were performed using GraphPad Prism 6 software. Chi-square test was applied to assess the statistical significance of the

deviation in the genotyping results from the expected Mendelian ratio. For the results of qRT-PCR and the characteristics and results of the mice objected to AAC and sham operation one-way analysis of variance (one-way ANOVA) followed by post hoc analysis with Bonferroni correction was used when more than 2 different groups of animals were included (Takahashi et al. 2007; Xu et al. 2015b). Otherwise, two-tailed one-sample t-test, was used to compare the results of the tested samples with a hypothetical mean of 1 or 100% in the wild type animals or the control group of cells as shown in each figure (Ewald et al. 2015; Ge et al. 2011). The cutpoint of p-value less than 0.05 was used to determine whether the difference was statistically significant. In all comparisons at least 3 age- and sex-matched animals from each genotype, which were not siblings, were tested.

3. Results

3.1 Genotyping protocols for 'Knockout-first' mice

As described previously, 2 independent protocols were used to genotype our 'Knockout-first' mice. In the first one we used 3 primers: Gse1-5arm-WTF, Gse1-Crit-WTR and Tm1a-5mut-R, which give one product of 241 bp in the mice that have 2 Gse1 wild type alleles ($Gse1^{WT/WT}$), and these mice are named as wild type mice thereafter in our study. In the mice that have 2 $Gse1^{tm1a(EUCOMM)Wtsi}$ mutated alleles ($Gse1^{tm1a/tm1a}$), one amplicon of 127 bp is amplified, and these mice are named in our study as homozygous mice. The third genotype is found in the mice that have one mutated and one wild type allele ($Gse1^{tm1a/WT}$), and these mice will give both products using this genotyping PCR and are named in our study as heterozygous mice (Figure 3.1). In the genotyping PCR one additional sample was included serving as a negative control, in which nuclease-free water was used instead of genomic DNA and should give none of the products. In the second protocol we used 5 primers in 3 separated reactions: Gse1-F and Gse1-ttR together; CSD-neoF and Gse1-ttR together; and CSD-loxF and Gse1-R together. Three products are to be obtained in the heterozygous mice with the sizes 796, 1074 and 1003 bp respectively. Only one amplicon from the first pair in the wild type animals, or 2 products from the last 2 pairs in the homozygous mice should be detected (Figure 3.1).

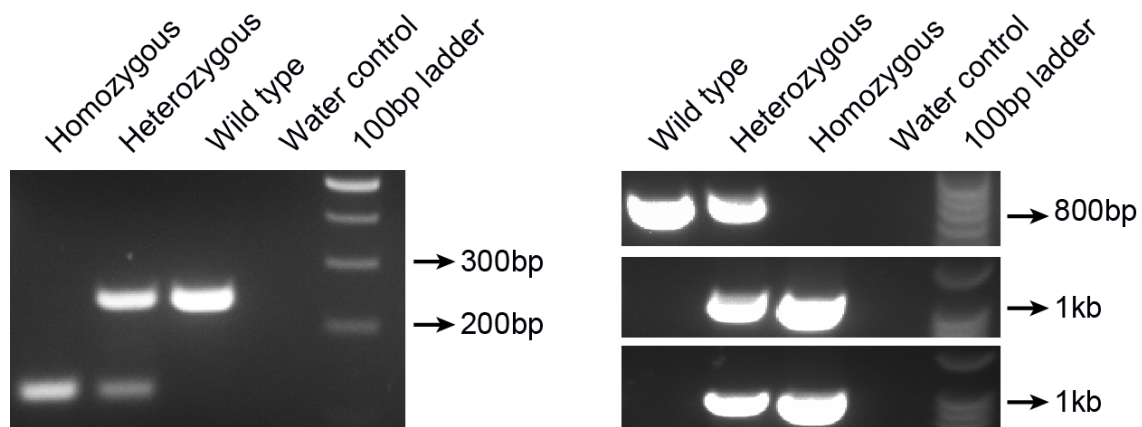


Figure 3.1: The genotyping protocols used in our study. Left: In the heterozygous mice 2 bands were detected using the primers Gse1-5arm-WTF, Gse1-Crit-WTR and Tm1a-5mut-R, but only one band of the large or small size is detected in the wild type or the homozygous mice respectively. Right: In the second protocol 3 PCR reactions were performed. One band from the primers Gse1-F and Gse1-ttR (upper lane); or two bands from the primers CSD-neoF and Gse1-ttR (middle lane) and the primers

CSD-loxF and Gse1-R (lower lane) were obtained in the wild type or the homozygous mice respectively. All three bands were detected in the reactions of heterozygous mice.

3.2 Quality control tests

The ES cells used to generate our mice were tested in several ways by the supplier, but it was found by Ryder and his colleagues that more than 10% of mice generated from ES cells with similar design to ours had problems related to incorrect targeting of the gene of interest or deletion of the 5' end of the trapping cassette (Ryder et al. 2013). Therefore, we performed several quality control tests on the heterozygous mice from the first generation (G1) contributing to expand our colony to exclude these problems.

3.2.1 Confirming the specificity of Gse1 targeting

As described in the methods section, long-range PCR was performed to prove that the trapping cassette was correctly inserted in Gse1 gene. Positive band with the right size was detected after PCR performed with samples from G1 heterozygous animals, but not with samples from a wild type mouse or with the negative control (Figure 3.2), and this band was sequenced to confirm its specificity (Figure 3.3).

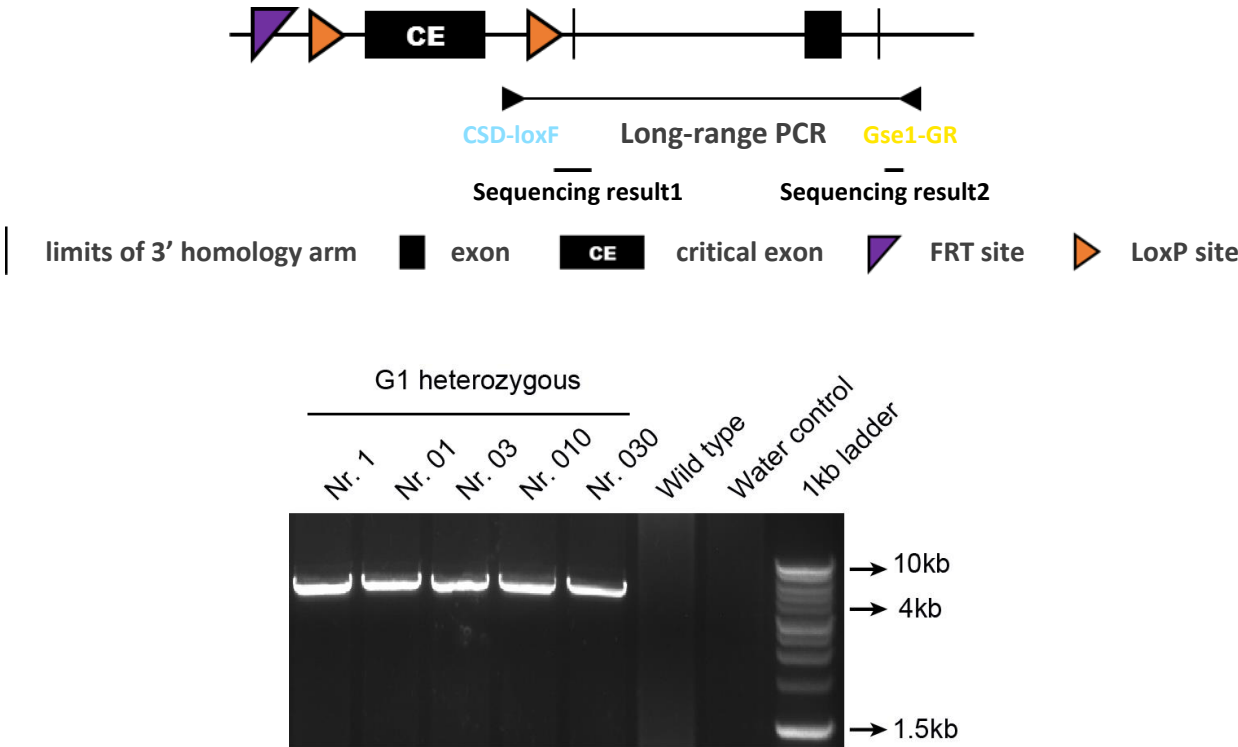


Figure 3.2: Upper: Schematic representation of the long-range PCR showing the locations of the primers Gse1-GR and CSD-loxF, which were also used to sequence the PCR product. Lower: The results of long-range PCR using these primers, showing that a positive band was detected with the size of more than 4 kb from the tested G1 heterozygous animals, but not in the reactions of the wild type animal or the negative control using nuclease-free water.

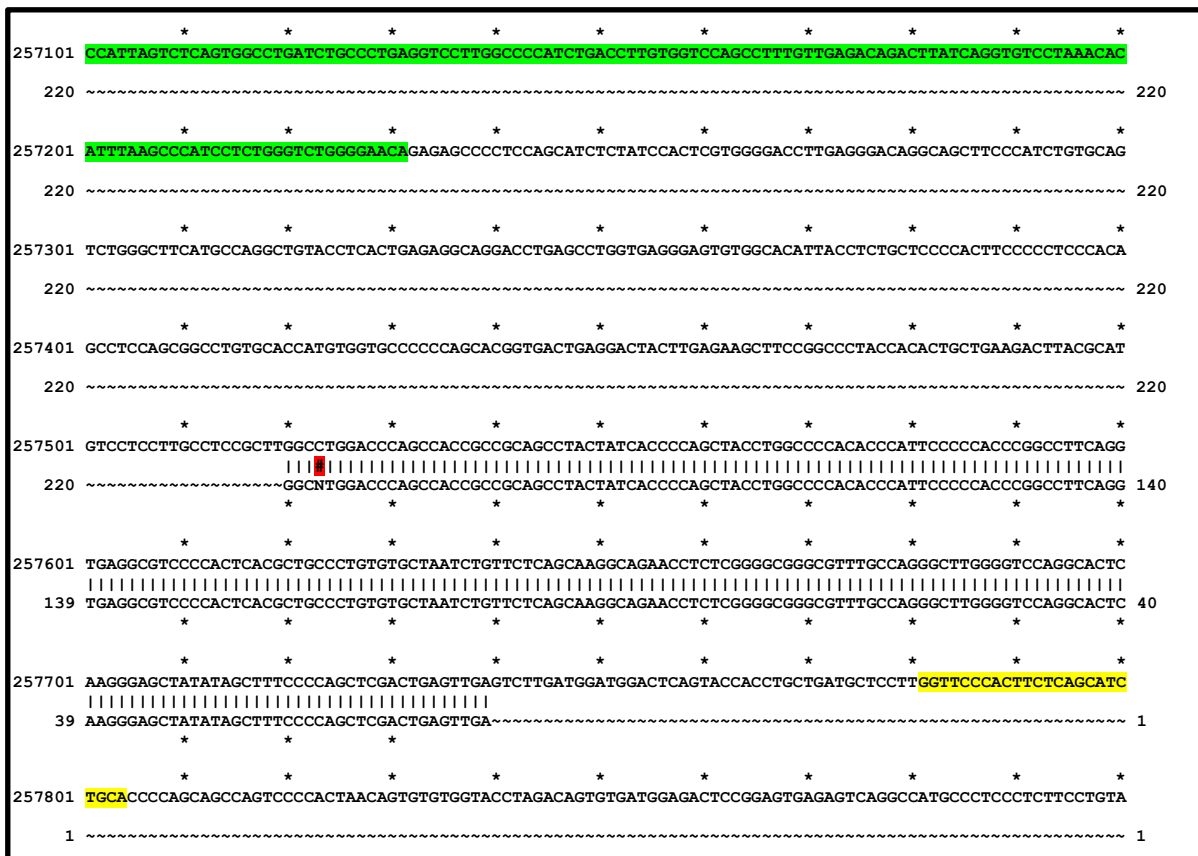
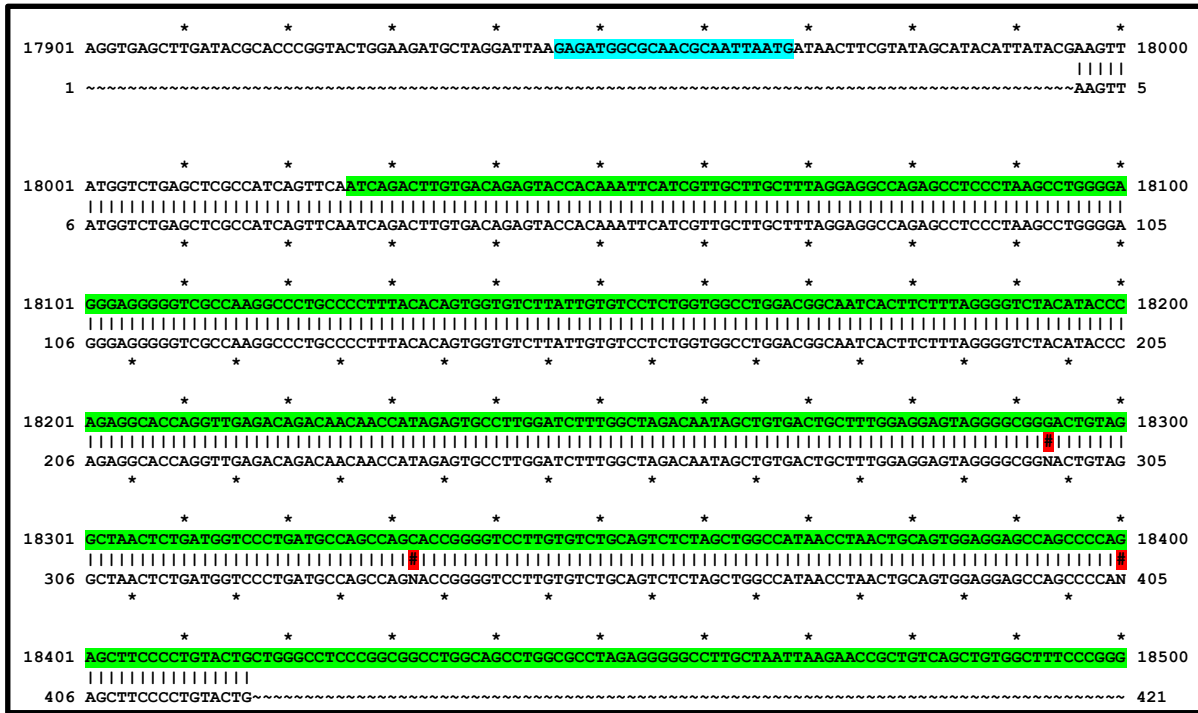


Figure 3.3: Upper: Alignment between the first sequencing result of the amplicon from long-range PCR (lower sequence) obtained by CSD-loxF primer (sequence in blue), and the vector PG00193_Z_8_B01 with the trapping cassette L1L2_Bact_P that was used to generate the mutated allele $Gse1^{tm1a(EUCOMM)Wtsi}$ (upper sequence) (https://www.i-dcc.org/imits/targ_rep/alleles/2232/targeting-vector-genbank-file). Lower: Alignment between the reverse complement of the second sequencing result of the same amplicon (lower sequence), obtained by Gse1-GR primer (sequence in yellow) and Gse1 genomic DNA (upper sequence) (http://www.ensembl.org/Mus_musculus/Gene/Sequence?db=core;g=ENSMUSG0000031822;r=8:120230536-120581390). Sequence in green indicates the 3' homologous arm.

Another experiment was also performed to confirm *Gse1* targeting, which was qRT-PCR with TaqMan probe detecting the break sequence in *Gse1* wild type allele (the deleted region from the genomic DNA of *Gse1* due to the insertion of the trapping cassette), which is named as break point loss of allele (BP- LOA) assay. It was proven that this break region cannot be detected in the samples from homozygous animals, and it was also shown that this sequence was found almost two times more in the wild type animals compared with the heterozygous, as they have two wild type alleles compared with one in the heterozygous mice (Figure 3.4).

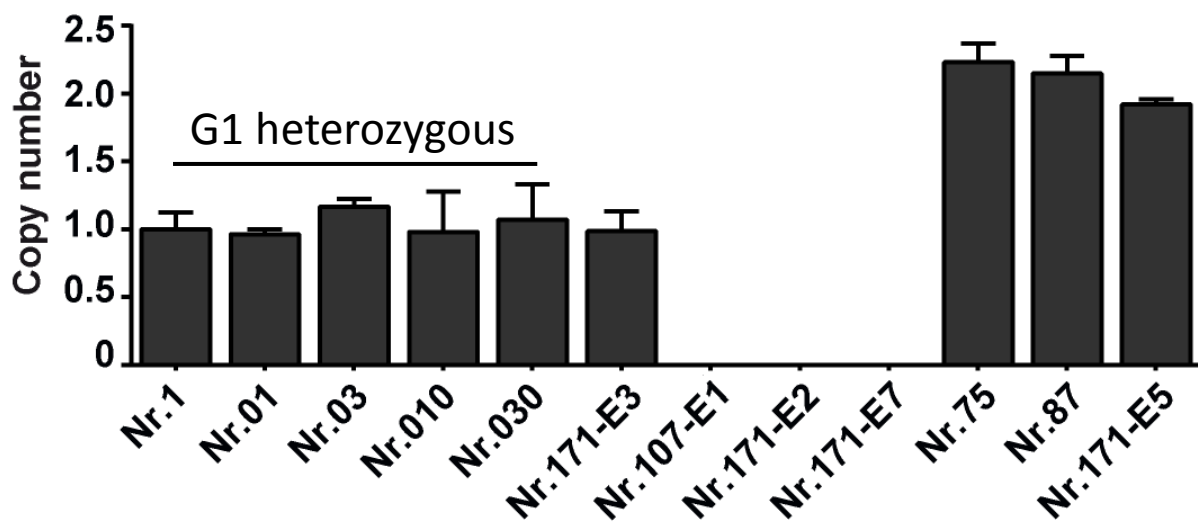


Figure 3.4: The results of (BP- LOA) assay. Mean value and standard error for each animal were normalized to the first heterozygous animal and represented. The same amount of genomic DNA that includes the break sequence of the wild type allele (regarded as one copy of this allele) was detected in the tested G1 heterozygous animals numbers 1, 01, 03, 010 and 030, which were involved in expanding our mice colony. Similar amount of genomic DNA including the break sequence was found in the heterozygous embryo number 171-E3. The double amount was found in the wild type animals (the adult mice number 75 and 87 and the embryo number 171-E5), and this sequence was not detected at all in homozygous animals (the embryos number 171-E2, 171-E7 and 107-E1).

3.2.2 Confirming the structure of the trapping cassette

Several short-range PCRs were performed to confirm that no part of our trapping cassette is missing using different pairs of primers detecting the 5' end, the LacZ cassette and the region of 3' LoxP, which were 5FRT_F and 5FRT_R; LacZ_2_small_F and LacZ_2_small_R; Floxed_PNF and Floxed_LR respectively. All tested samples from G1 heterozygous mice showed the expected amplicon from all these PCRs indicating that the trapping cassette was complete (Figure 3.5). In addition to that, the PCR using the primers CSD-neoF and Gse1-ttR was performed during genotyping for all G1 heterozygous mice and other animals proving that the cassette of neomycin resistance was also not deleted.

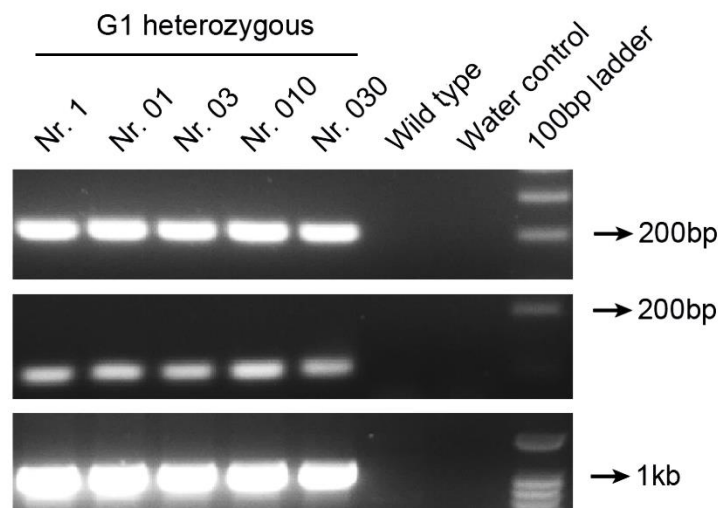


Figure 3.5: The results of short-range PCRs using the primers 5FRT_F and 5FRT_R (upper lane); the primers LacZ_2_small_F and LacZ_2_small_R (middle lane) and the primers Floxed_PNF and Floxed_LR (lower lane), indicating that none of the following regions in the trapping cassette; the 5'end, the LacZ cassette and the region of 3' LoxP was missing respectively, confirming thereby the structure integrity of this cassette in all tested heterozygous mice from the first generation.

3.3 Genotyping results

Genotyping was performed for the newborn mice from breeding between heterozygous male and heterozygous female using the two protocols mentioned above. Interestingly, out of more than 130 newborn mice from more than 20 litters and different males and females in breeding, no homozygous newborn mouse was observed (Table 3.1). The heterozygous animals accounted for 62% of all the genotyped newborn mice from heterozygous with heterozygous matings and the rest

were wild type mice representing significant deviation from the Mendelian ratio 1:2:1 with p-value less than 0.0001, and indicating clearly that the homozygosity of *Gse1* trapping is lethal. 52% of the animals were males with no significant deviation from the ratio 1:1 (p-value = 0.6042). No lethality in the heterozygous mice could be proven as there were no significant deviation from the expected ratio 2:1 of the surviving heterozygous and wild type newborn mice (p-value = 0.2458). The average size of the litter was 5.82 ± 2.05 .

Male & Female	Litters	Wild type		Heterozygous		Homozygous		Total
		M	F	M	F	M	F	
03 & 01	3	3	3	8	3	0	0	17
35 & 42	4	8	5	11	6	0	0	30
35 & 49	2	2	4	3	5	0	0	14
030 & 64	1	4	0	0	0	0	0	4
030 & 55	1	0	1	2	2	0	0	5
85 & 81	1	1	1	2	1	0	0	5
85 & 82	1	2	1	1	2	0	0	6
80 & 74	4	3	6	6	7	0	0	22
80 & 71	2	3	0	2	5	0	0	10
176 & 150	1	1	0	1	3	0	0	5
147 & 182	1	0	2	1	2	0	0	5
164 & 219	1	1	0	1	2	0	0	4
215 & 192	1	0	0	4	3	0	0	7
Total	23	28	23	42	41	0	0	134

Table 3.1: The genotyping results of the newborn mice from 23 different litters of 13 matings between heterozygous male and heterozygous female, showing that no homozygous newborn mouse was found.

Representative results of genotyping PCR for 20 mice chosen from 5 different litters are shown in figure 3.6, where we see that the DNA samples of all genotyped animals produced a positive band from the primers *Gse1*-5arm-WTF and *Gse1*-Crit-WTR, meaning that all these animals had at least one wild type allele. This band was also sequenced to prove that it was specific for *Gse1*. The two genotyping protocols

provided the same results (except for very few cases related mostly to bad quality of the DNA sample, which required repeating the PCR using more DNA or applying higher numbers of cycles or with a new sample of genomic DNA). The sequencing results of the band from the primers Gse1-F and Gse1-ttR, which was also positive in all tested animals, were also specific for *Gse1* gene.

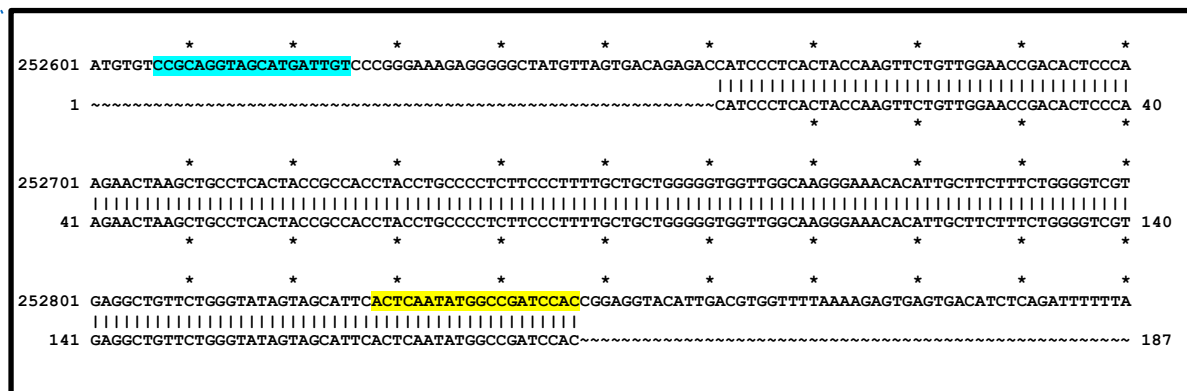
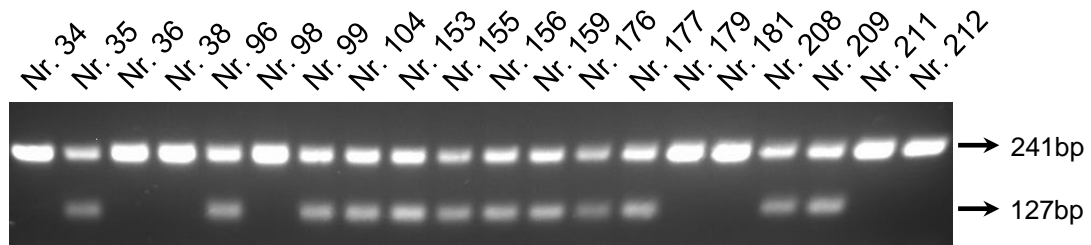


Figure 3.6: Upper: The genotyping results of 20 newborn mice from breeding between heterozygous male and heterozygous female, using the primers Gse1-5arm-WTF, Gse1-Crit-WTR and Tm1a-5mut-R. None of them was homozygous, as all of them gave positive band from the primers Gse1-5arm-WTF and Gse1-Crit-WTR (the upper band). Lower: Alignment between *Gse1* genomic DNA (upper sequence) (http://www.ensembl.org/Mus_musculus/Gene/Sequence?db=core;g=ENSMUSG0000031822;r=8:120230536-120581390) and the sequencing result of the positive band from the primers Gse1-5arm-WTF (sequence in blue) and Gse1-Crit-WTR (sequence in yellow), which was obtained using the primer Gse1-5arm-WTF for sequencing (lower sequence).

The genotyping results of the newborn mice when one of the parents was wild type and the other one was heterozygous showed that 57% of them were heterozygous with no significant deviation from the Mendelian ratio 1:1 (p-value = 0.1253), indicating also that the lethality does not affect the heterozygous mice (Table 3.2). 46% of the newborn animals were males with no significant deviation from the ratio 1:1 (p-value = 0.4171). The average size of the litter was 6.15 ± 2.52 .

Male & Female	Litters	Wild type		Heterozygous		Total
		M	F	M	F	
030 & A4	2	5	4	5	3	17
030 & A3	1	2	0	2	2	6
40 & A11	1	1	2	6	0	9
43 & A13	2	3	3	3	6	15
39 & A12	3	4	4	4	11	23
197 & 231	1	1	3	1	3	8
A8 & 64	6	7	7	6	9	29
A7 & 1	1	0	1	1	1	3
A9 & 010	1	2	0	0	1	3
233 & 229	2	2	2	2	4	10
Total	20	27	26	30	40	123

Table 3.2: The genotyping results of the newborn mice from 20 different litters of 10 matings between heterozygous male and wild type female or the opposite, showing that no lethality can be proved in the heterozygous animals.

3.4 Generating mice with $Gse1^{tm1c}$ allele

In accordance with the aims of the knockout mouse project phase 2 (KOMP2), several matings between heterozygous mice and mice expressing FLP-recombinase were arranged. However, when the newborn mice were positive for FLP-recombinase and heterozygous for $Gse1$, the mutated allele $Gse1^{tm1a}$ was not converted completely to $Gse1^{tm1c}$. It was shown according to our genotyping results that both types of mutated allele were present in the same mouse, which means that we had mosaic mice, where the mutated allele $Gse1^{tm1a}$ were converted to the mutated allele $Gse1^{tm1c}$ in some cells but not in the others. Those results were confirmed several times by both genotyping protocols in different litters from the first generation after breeding with FLP-recombinase mice. The genotyping result of the first litter in this mouse line is presented in the figure 3.7.

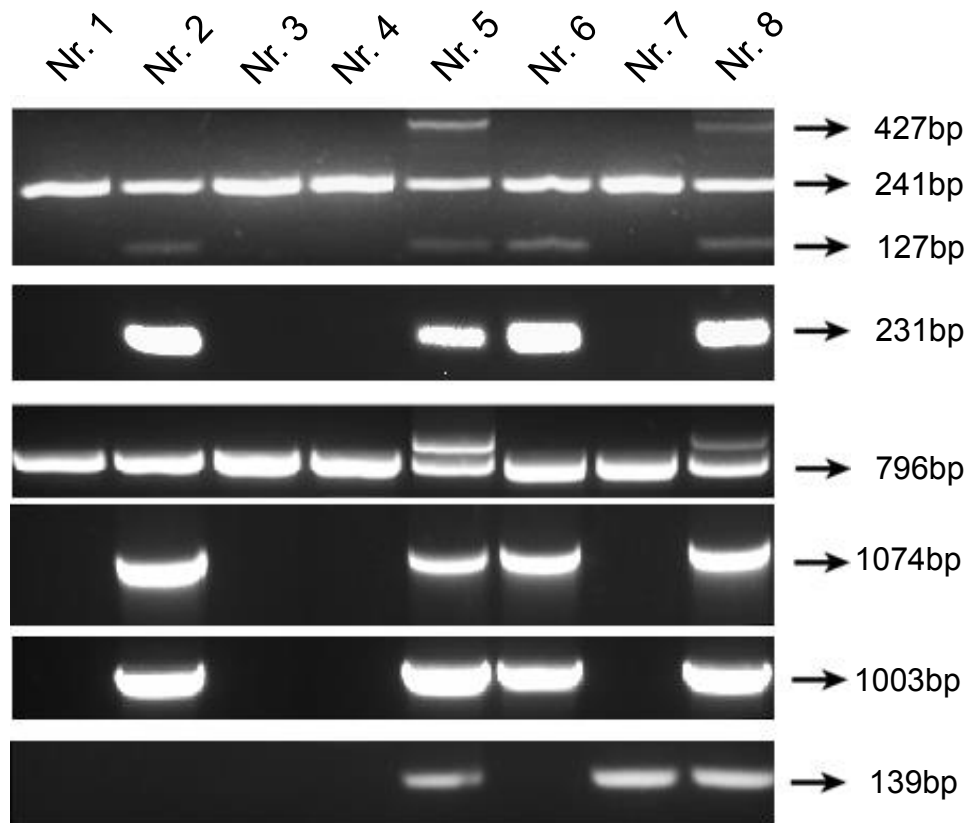


Figure 3.7: The genotyping results of the newborn mice from the mating between a heterozygous male ($Gse1^{tm1a/WT}$) and a FLP-positive female, showing that we had mosaic mice (number 5 and 8). Three bands were detected using the primers Gse1-5arm-WTF, Gse1-Crit-WTR and Tm1a-5mut-R (first lane); and two bands using the primers Gse1-F with Gse1-ttR (third lane). The bands with the larger size from both reactions (427 bp and 982 bp respectively) indicated that we had the mutated allele $Gse1^{tm1c}$, as this allele will produce larger amplicon compared with the wild type allele due to the remaining FRT and LoxP sites. However, an amplicon was still detected using the primers Gse1-5arm-WTF with Tm1c-EN2-R (second lane), and the primers CSD-neoF with Gse1-ttR (4th lane), which should be negative if the trapping cassette is completely removed by FLP-recombinase. The 5th lane is the PCR results using the primers CSD-loxP and Gse1-R, which give positive band from both mutant alleles ($Gse1^{tm1a}$ and $Gse1^{tm1c}$). The 6th lane is the PCR results using the primers FLP-Geno-F and FLP-Geno-R, which give positive band when FLP-recombinase cassette is inside the genomic DNA of the mouse. Our genotyping protocols are described in details in the methods section (Paragraph 2.2.6, Figure 2.2)

After that these mosaic mice were bred with wild type mice, and we were able to identify using our genotyping protocols several heterozygous animals with pure mutant $Gse1^{tm1c}$ allele among the newborn mice, such as number 116 and 120 in figure 3.8. These results were also confirmed by other primers, such as LacZ_2_small_F and LacZ_2_small_R primers. Some of these mice were negative for FLP-recombinase, such as number 120, and will be used in further matings to

generate the mice with the mutant allele $Gse1^{tm1d}$ in order to confirm the causative relationship between the observed lethality and the trapping of $Gse1$.

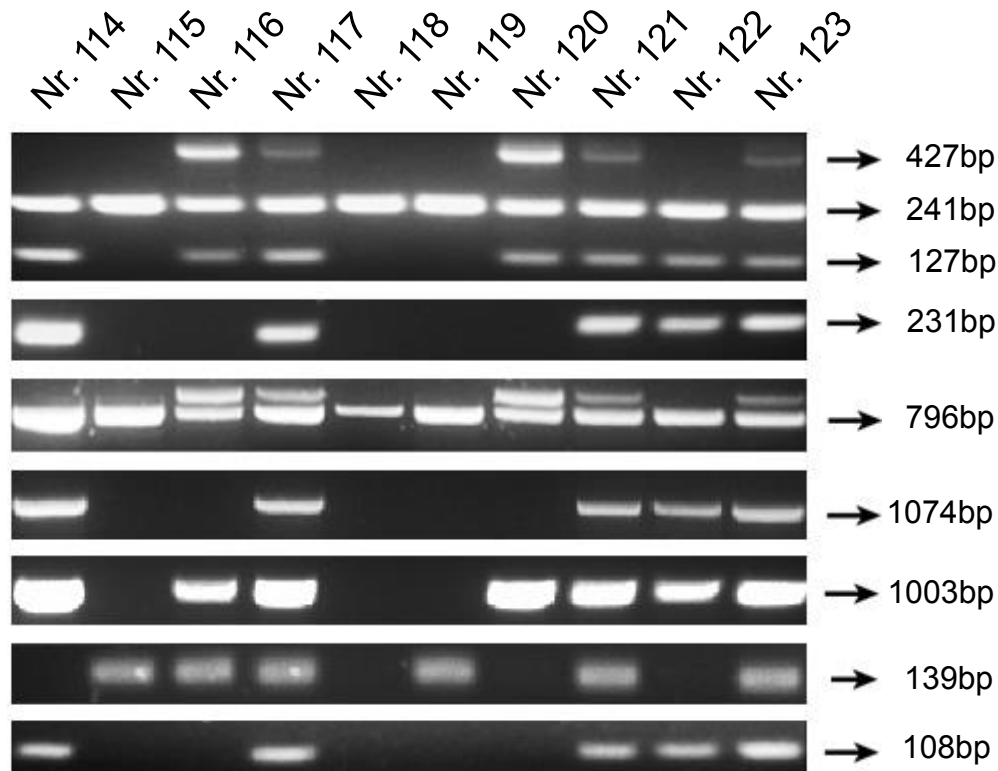


Figure 3.8: The genotyping results of newborn animals from breeding between one of the previously described mosaic mice and wild type mouse, showing that we had heterozygous mice with pure $Gse1^{tm1c}$ allele (number 116 and 120). No amplicon was detected using the primers $Gse1$ -5arm-WTF and Tm1c-EN2-R (second lane), the primers CSD-neoF and $Gse1$ -ttR (4th lane) or the primers LacZ_2_small_F and LacZ_2_small_R (7th lane) from these two animals. The PCR results using the primers $Gse1$ -5arm-WTF, $Gse1$ -Crit-WTR and Tm1a-5mut-R; the primers $Gse1$ -F and $Gse1$ -ttR; the primers CSD-loxP and $Gse1$ -R and the primers FLP-Geno-F and FLP-Geno-R are shown in the first, third, 5th and 6th lane respectively.

3.5 Generating $Gse1^{tm1b}$ allele *in vitro*

Our mice were already tested to confirm the presence of the LoxP sites, which allow them to work as a conditional mouse model, but we wanted to confirm that *in vitro* before breeding with Cre-recombinase mice. Therefore, fibroblasts from the kidneys of wild type and heterozygous mice were isolated and transduced by Cre-recombinase virus, which is supposed to convert the mutant allele $Gse1^{tm1a}$ to $Gse1^{tm1b}$ allele. We were able to achieve a transduction efficiency of 70% as estimated by GFP signal from the transduced fibroblasts (Figure 3.9).

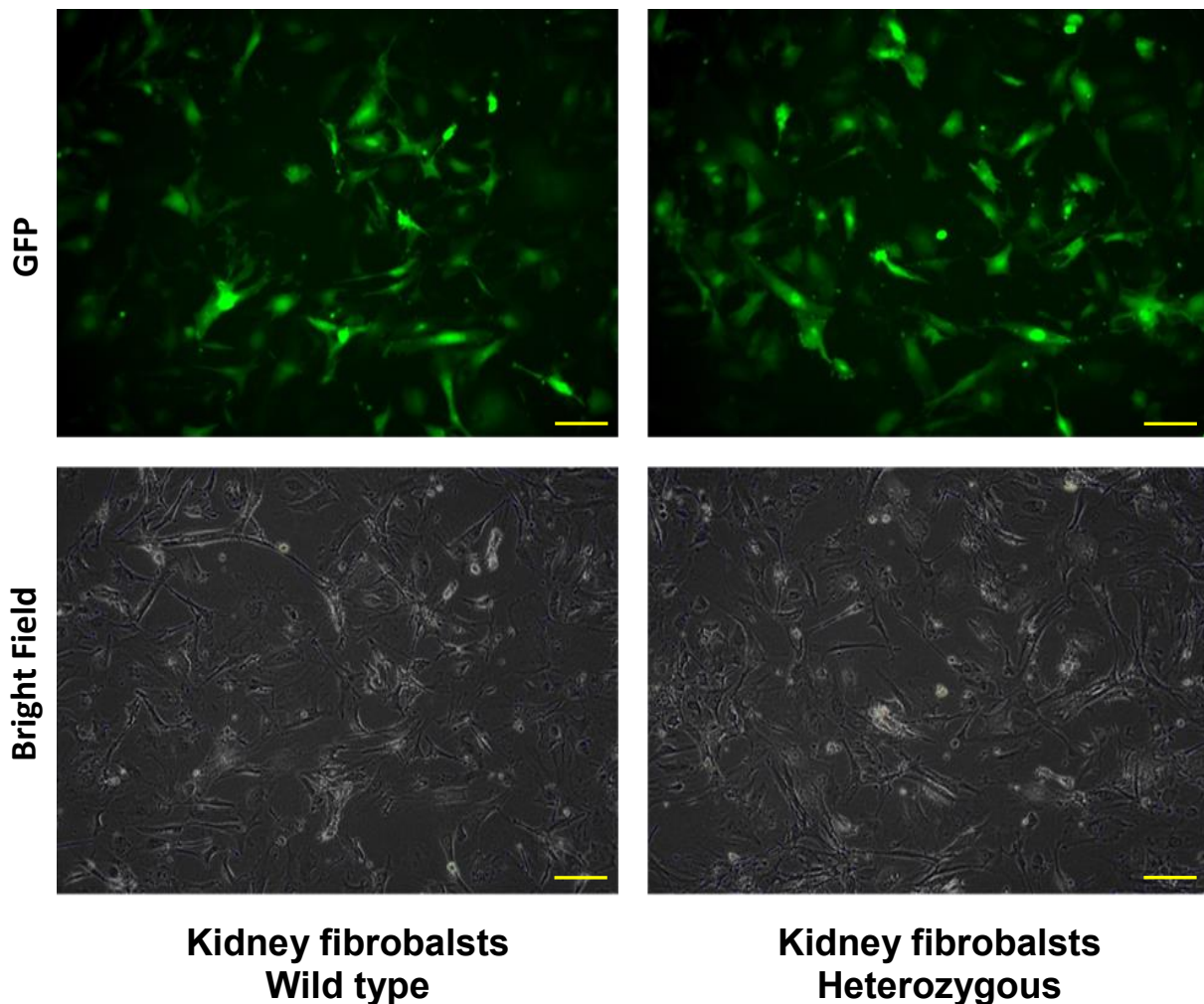


Figure 3.9: Kidney fibroblasts from heterozygous and wild type animals transduced with adenovirus expressing Cre-recombinase, showing a transduction efficiency of 70% percent approximately in both groups of cells according to how many cells are GFP-positive.

Afterwards, genomic DNA was isolated from those cells and the triple-PCR-reactions of the second protocol were performed as described before. Additionally, one PCR was performed using Gse1-R primer with the primer CSD-lacF, which is located in the LacZ cassette. This reaction is supposed to give an amplicon of 1301 bp if we have the mutated allele $Gse1^{tm1b}$, in which the cassette of neomycin resistance and the targeted exon should be removed by Cre-recombinase. However, in case of the allele $Gse1^{tm1a}$ the PCR product will be more than 4 kb, which is not possible to be amplified because of our PCR conditions. The results are shown in figure 3.10 suggesting that we obtained the desired $Gse1^{tm1b}$ allele but not in all cells, as we were still able to detect a positive band, but weaker one, from the reactions using the primers CSD-neoF with Gse1-ttR; and the primers CSD-loxF with Gse1-R.

That reflected likely the genomic DNA extracted from the cells, where the transduction with Cre-recombinase virus was not successful.

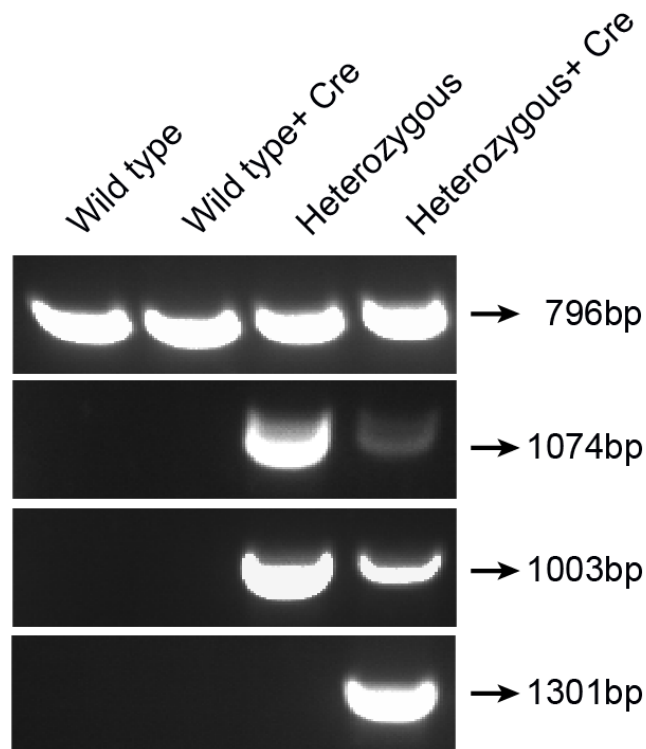


Figure 3.10: The genotyping results of the transduced fibroblasts, showing that we had cells with the allele $Gse1^{tm1b}$, indicated by the positive band from the PCR using the primers Gse1-R and CSD-lacF (4th lane). The PCR results using the primers Gse1-F and Gse1-ttR; the primers CSD-neoF and Gse1-ttR; and the primers CSD-loxP and Gse1-R are shown in the first, second and third lane respectively.

3.6 *Gse1* expression results

To prove the connection between the lethality in the homozygous mice and the trapping of *Gse1*, RNA samples from embryonic tissue were analyzed by qRT-PCR after reverse transcription. The primers, m.*Gse1*-F and m.*Gse1*-R, are located after the trapping cassette in exon 6 and 7 from the isoform *Gse1*-002 respectively. That should enable us to detect the difference in the expression level among mice of different genotypes. Western blotting was also performed using embryonic protein samples. It was found that *Gse1* expression was significantly lower in the homozygous embryos compared with the wild type, but the mRNA and the protein of *Gse1* were still detected in these mice (with expression level of more than 50 % compared with that in wild type embryos) (Figure 3.11). Additionally, the expression

results of LacZ reporter, which showed clear upregulation in the heterozygous and homozygous, suggested that the trapping cassette is working (Figure 3.12), and it can be concluded that the trapping system generated a hypomorphic allele of *Gse1* instead of null allele.

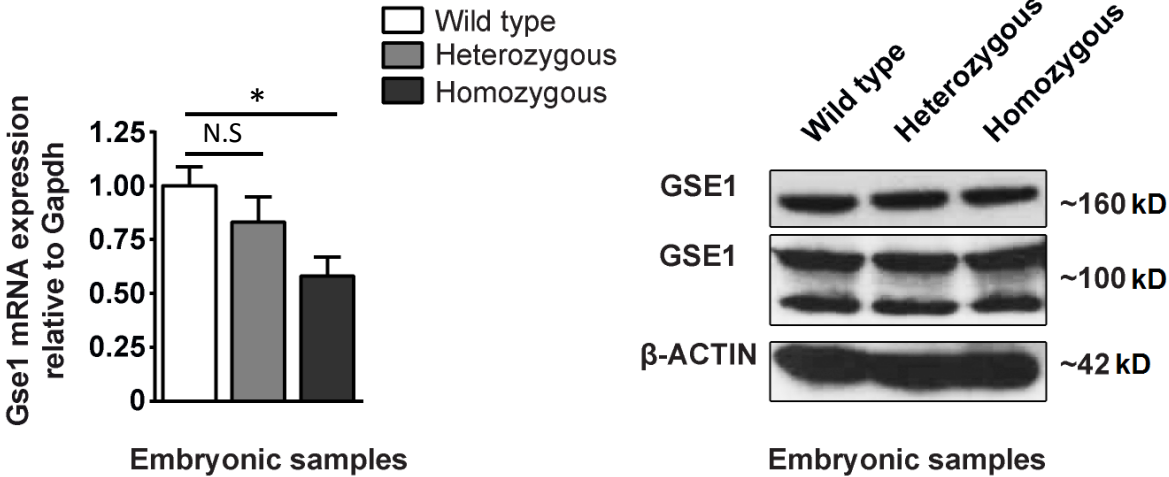


Figure 3.11: Left: The expression level of *Gse1* in samples from wild type, heterozygous and homozygous embryos, represented by means and standard errors, where (*) indicates p-value less than 0.05. Right: Western blotting showing minimal difference among the three genotypes using two different antibodies for GSE1 protein (Sigma-Aldrich in the upper lane and LifeSpan Biosciences in the middle) against two different epitopes, and antibody against β -ACTIN as loading control in the lower lane.

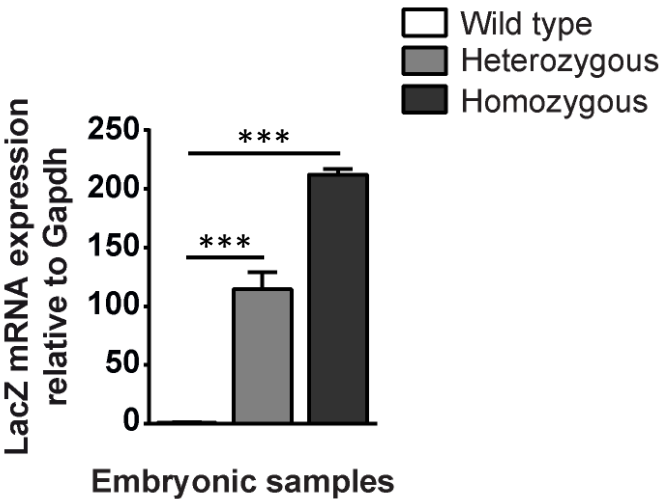


Figure 3.12: The expression level of LacZ in samples from wild type, heterozygous and homozygous embryos, represented by means and standard errors, where (***) indicates p-value less than 0.001.

The expression level of *Gse1* was also tested in adult mice by qRT-PCR and Western blot, which showed that the highest expression is in the lung and liver among the tested organs, and suggested that different protein isoforms are expressed in these organs (Figure 3.13). It was also found, that the heterozygous adult mice have significantly lower *Gse1* expression in heart and kidney, but not in lung and liver, compared with the wild type (Figure 3.14).

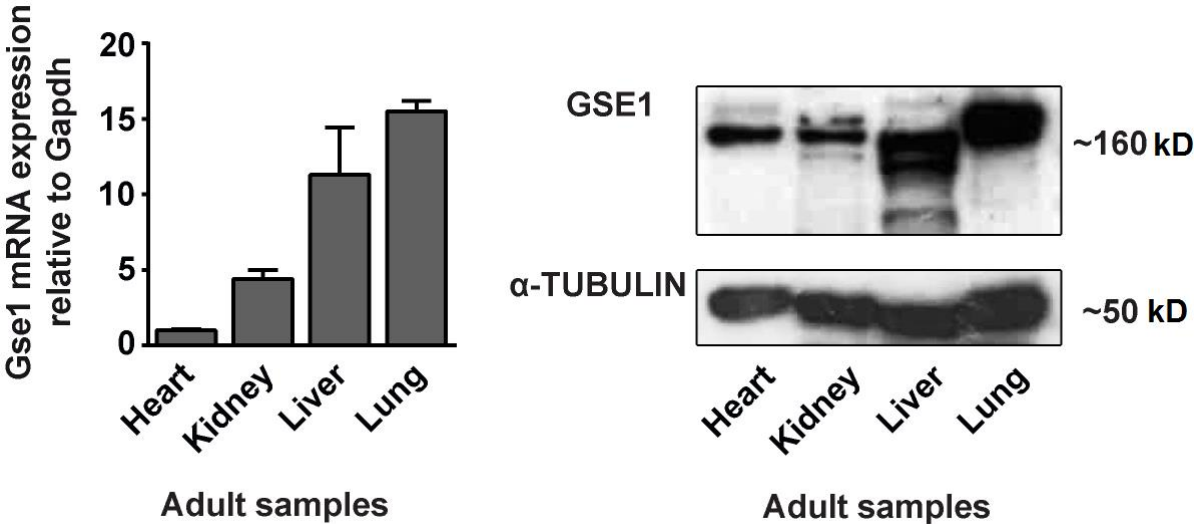


Figure 3.13: Left: The expression level of *Gse1* in samples from different organs of adult mice estimated by qRT-PCR and represented by means and standard errors. Right: Western blotting of protein samples from different organs of adult mouse using antibody for GSE1 protein (Sigma-Aldrich) in the upper lane and antibody against α -TUBULIN as loading control in the lower lane.

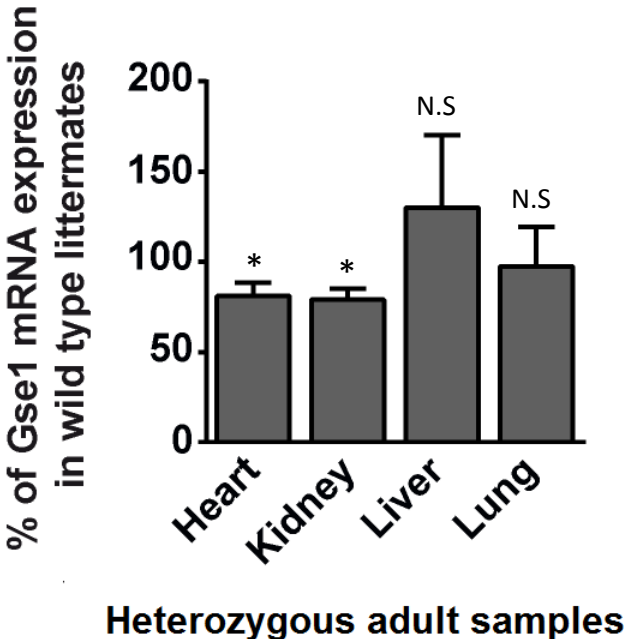


Figure 3.14: The expression level of *Gse1* in heart, kidney, liver and lung using samples from heterozygous adult mice compared with wild type littermates and represented by means and standard errors.

3.7 *Gse1* circular RNA results

Our gene is described to produce circular RNA from the second exon of the isoform *Gse1*-002 (Memczak et al. 2013), which is located on genomic level just upstream from the intron, in which the trapping cassette was inserted. Therefore, it could be possible that the splicing process that controls the expression of this circular RNA is affected, and in this case that may be involved in the observed lethality of the homozygous mice. To test our hypothesis we started by proving that *Gse1* circular RNA is expressed during embryonic development. We designed 2 outward-facing primers, m.circularRNA-F and m.circularRNA-R, and performed RT-PCR with cDNA synthesized from embryonic RNA samples using either random hexamers, oligo (dT), or both of them for reverse transcription (Jeck et al. 2013; Starke et al. 2015). As it is shown in figure 3.15, we were able to obtain an amplicon, and only from the samples prepared using random hexamers, which indicates that we were detecting non-polyadenylated RNA. The band was extracted from the gel and proved by sequencing that it belongs to the exon we are looking for, but in scrambled form (Figure 3.16). Another PCR using primers for mouse *Gapdh* was performed for the same samples as a control.

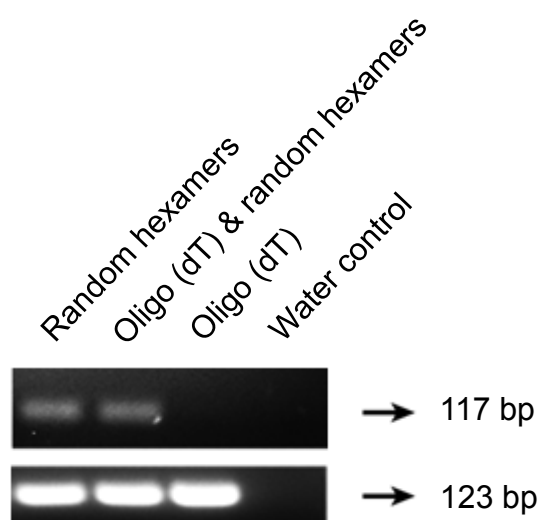


Figure 3.15: The RT-PCR results using embryonic RNA after reverse transcription with random hexamers, oligo (dT), or both of them, showing that we had the amplicon of the scrambled exon from the outward-facing primers, m.circularRNA-F and

m.circularRNA-R, only in samples reverse transcribed by random hexamers (upper lane). That reflects the absence of polyA tail in the detected RNA, which is a criterion for its circularity. The RT-PCR using mouse Gapdh primers (lower lane) was performed as a control, which is supposed to give an amplicon of 123 bp.

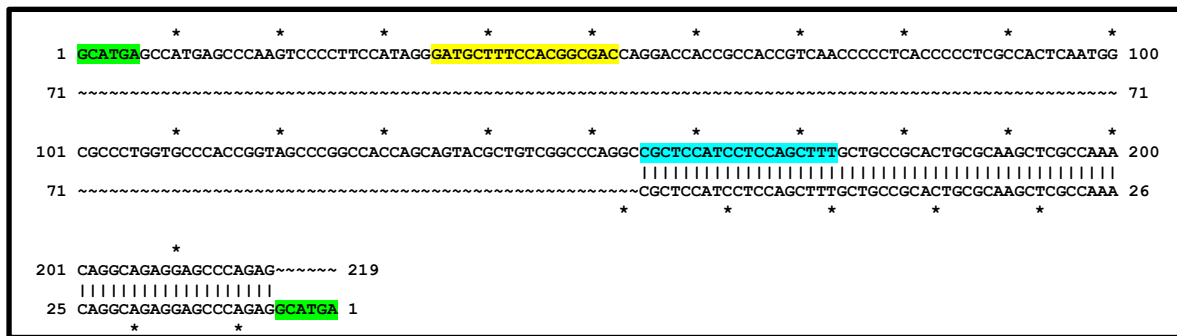
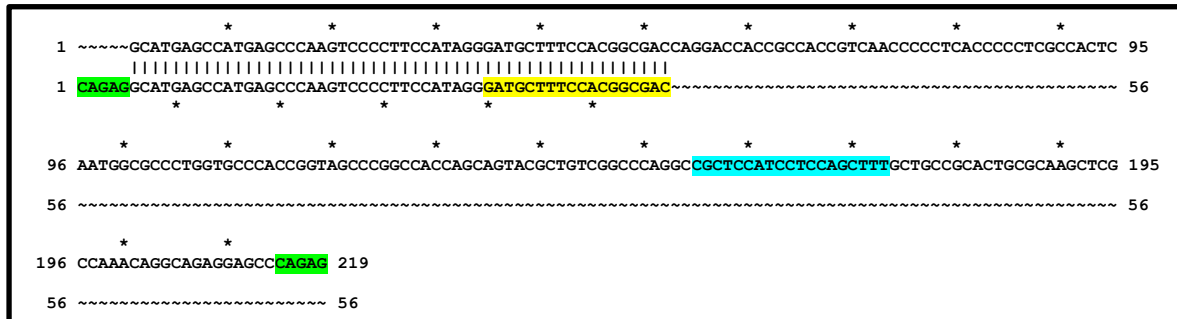


Figure 3.16: Upper: Alignment between the sequencing result of the product from the primers m.circularRNA-F (sequence in blue) and m.circularRNA-R (sequence in yellow), which was obtained using the primer m.circularRNA-F (lower sequence), and the sequence of the circularized exon of *Gse1* (upper sequence) (http://www.ensembl.org/Mus_musculus/Gene/Sequence?db=core;g=ENSMUSG0000031822;r=8:120230536-120581390). The alignment between the sequences in green at the beginning of the sequencing result and at the end of the circularized exon proves the scrambled junction between the beginning and the end of the exon. Lower: Alignment between the reverse complement of the sequencing result for the same product obtained by the primer m.circularRNA-R (lower sequence) and the sequence of the circularized exon of *Gse1* (upper sequence). The alignment between the sequences in green confirms further the scrambled junction of this exon.

Afterwards, another experiment was performed to confirm further the circularity of the scrambled exon. RNA sample from wild type embryo was treated by RNase R before reverse transcription, which is supposed to digest the linear RNAs only. Another sample was prepared using the same conditions but without adding the RNase R to be used as a control. To prove that the digestion actually worked another PCR was performed for both samples as a control using the primer m.circularRNA-F, which is located near the end of the circularized exon, and another primer located in

the next exon, which was m.linearRNA-R. These primers are supposed to give an amplicon of 240 bp, detecting thereby the linear RNA isoforms of *Gse1* that contain the same circularized exon (Starke et al. 2015). According to our results the linear RNA isoforms were almost completely depleted in the sample treated with RNase R, but the amplicon from the primers that detect the scrambled exon was not affected, which confirms further that we detect the circular RNA of *Gse1* (Figure 3.17).

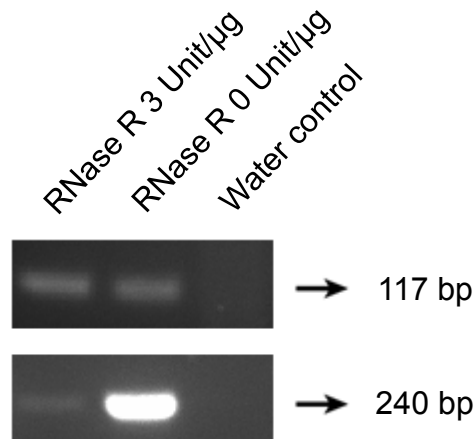


Figure 3.17: The RT-PCR results using embryonic RNA after treating with RNase R and reverse transcription using random hexamers and oligo (dT), showing that we still can get the amplicon of the scrambled exon using the outward-facing primers, m.circularRNA-F and m.circularRNA-R, in the sample treated with RNase R (upper lane). That indicates the resistance of the detected RNA to RNase R, which is another criterion for its circularity. The amplicon generated by the primers m.circularRNA-F and m.linearRNA-R that detect the linear RNAs of *Gse1* was almost completely absent in the RNase R-treated sample, indicating the efficiency of the treatment (lower lane). Sample prepared using the same steps but without RNase R was used as a control.

After that, we used the same outward-facing primers mentioned above to quantify the expression level of the circular RNA of *Gse1*. QRT-PCR with TaqMan probe detecting the junction between the end and the beginning of the scrambled exon was performed to investigate the expression in embryos from different genotypes. According to our results, the expression of the circular RNA was significantly upregulated in the homozygous animals compared with the wild type (Figure 3.18), indicating that the trapping cassette, which is located after the circularized exon, caused more circular RNA and less linear RNAs of *Gse1* to be generated. We are aware of the possibility that the primers we used to detect the linear isoforms may detect other circular RNA isoforms for *Gse1* gene, but these isoforms are not described previously in the literature.

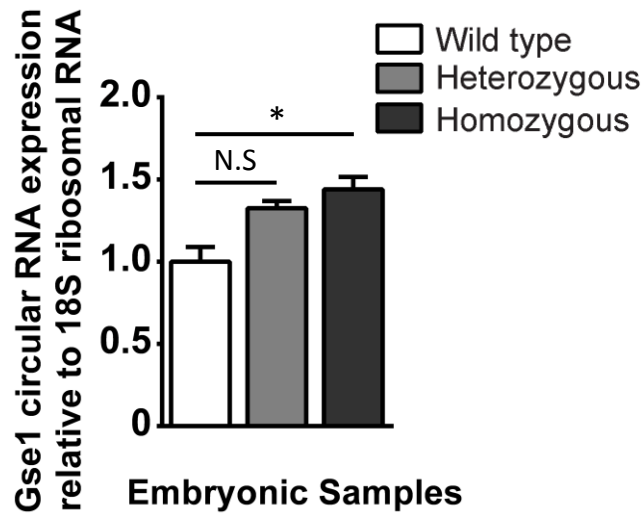


Figure 3.18: The expression level of the circular RNA of *Gse1* in samples from wild type, heterozygous and homozygous embryos, represented by means and standard errors, showing significant upregulation in the homozygous embryos compared with the wild type.

The expression of the circular RNA of *Gse1* was also analyzed in different tissues from adult mice, and the highest expression was also found in the lung, but the lowest expression among the tested organs was in the liver (Figure 3.19).

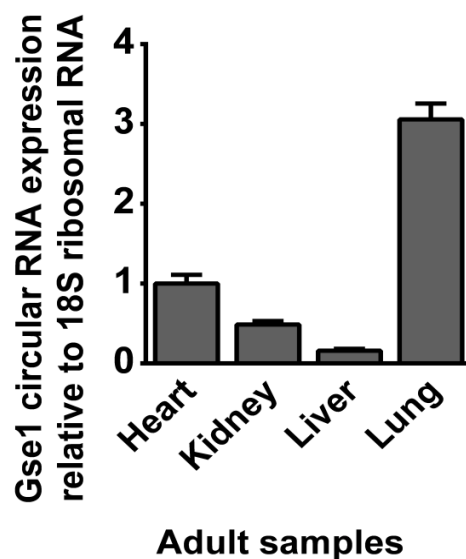


Figure 3.19: The expression level of *Gse1* circular RNA in samples from different organs of adult mice estimated by qRT-PCR and represented by means and standard errors.

As the most well-described function of circular RNA is to work as miRNA sponge (Lasda and Parker 2014), we searched for the miRNAs, which can be sequestered by the circularized exon of *Gse1* using its sequence as a query in the database for miRNAs (miRBase) (<http://www.mirbase.org/search.shtml>). Several candidate miRNAs were suggested, which were: mmu-miR-1934-3p, mmu-miR-6917-5p, mmu-

miR-3113-3p, mmu-miR-224-3p, mmu-miR-5110, mmu-miR-1947-3p, mmu-miR-6930-3p, mmu-miR-7235-5p, mmu-miR-3081-5p, mmu-miR-6392-5p, mmu-miR-6968-5p and mmu-miR-7215-3p. The miRNA mmu-miR-6968-5p was chosen from the predicted list, as it showed the best alignment in its seed region, and then we tried to search for the genes that could be targeted by this miRNA. More than 100 genes were found to be potential targets according to the online database, miRDB (<http://mirdb.org/miRDB/>). The expression level of the 6 genes at the top of the list according to their targeting score was investigated in embryonic RNA samples from homozygous embryos, where the circular RNA was proved to be upregulated, in comparison with wild type littermates. These genes were *Foxp4*, *Thy1*, *Actn4*, *Wnt1*, *Kmt2d* and *Tfap2b*. Only one of these genes, *Tfap2b*, showed upregulation trend, with almost significantly higher expression in the homozygous embryos compared with the wild type (p-value = 0.059). This can support the hypothesis about possible role for the circular RNA of *Gse1* as miRNA sponge, as its expression was increased in the homozygous embryos and should cause the suspected genes as miRNA targets to be upregulated by protecting them from the supposed inhibiting effect of this sequestered miRNA. Another gene, *Thy1*, was found to be significantly downregulated in the homozygous embryos, and the gene *Actn4* showed downregulation trend (p-value = 0.072) (Figure 3.20).

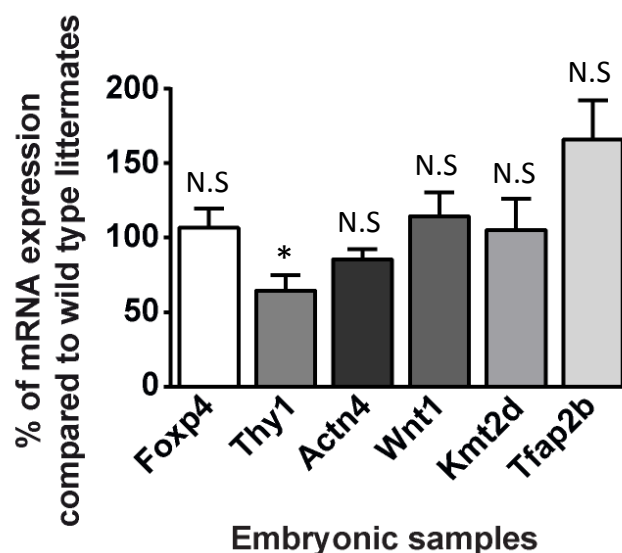


Figure 3.20: The expression level of *Foxp4*, *Thy1*, *Actn4*, *Wnt1*, *Kmt2d* and *Tfap2b* in samples from homozygous embryos compared with wild type littermates and represented by means and standard errors, showing significant downregulation in the expression of the gene *Thy1*, and upregulation and downregulation pattern in the homozygous embryos for the genes *Tfap2b* and *Actn4* respectively.

3.8 The expression results of *Gse1*-neighboring genes

The mutated allele *Gse1*^{tm1a} still includes the selection cassette of neomycin resistance with its autonomous promoter, making it possible that the observed lethality in the homozygous embryos was secondary to off-target effect of this promoter. To exclude this possibility, the expression level of several genes upstream and downstream from *Gse1* on chromosome 8 was analyzed by qRT-PCR. Those genes were *Usp10*, *Crispld2*, *Zdhhc7*, *Gins2*, *Cox4i1* and *Irf8*. None of the tested genes showed significant change in homozygous embryos compared with wild type (Figure 3.21).

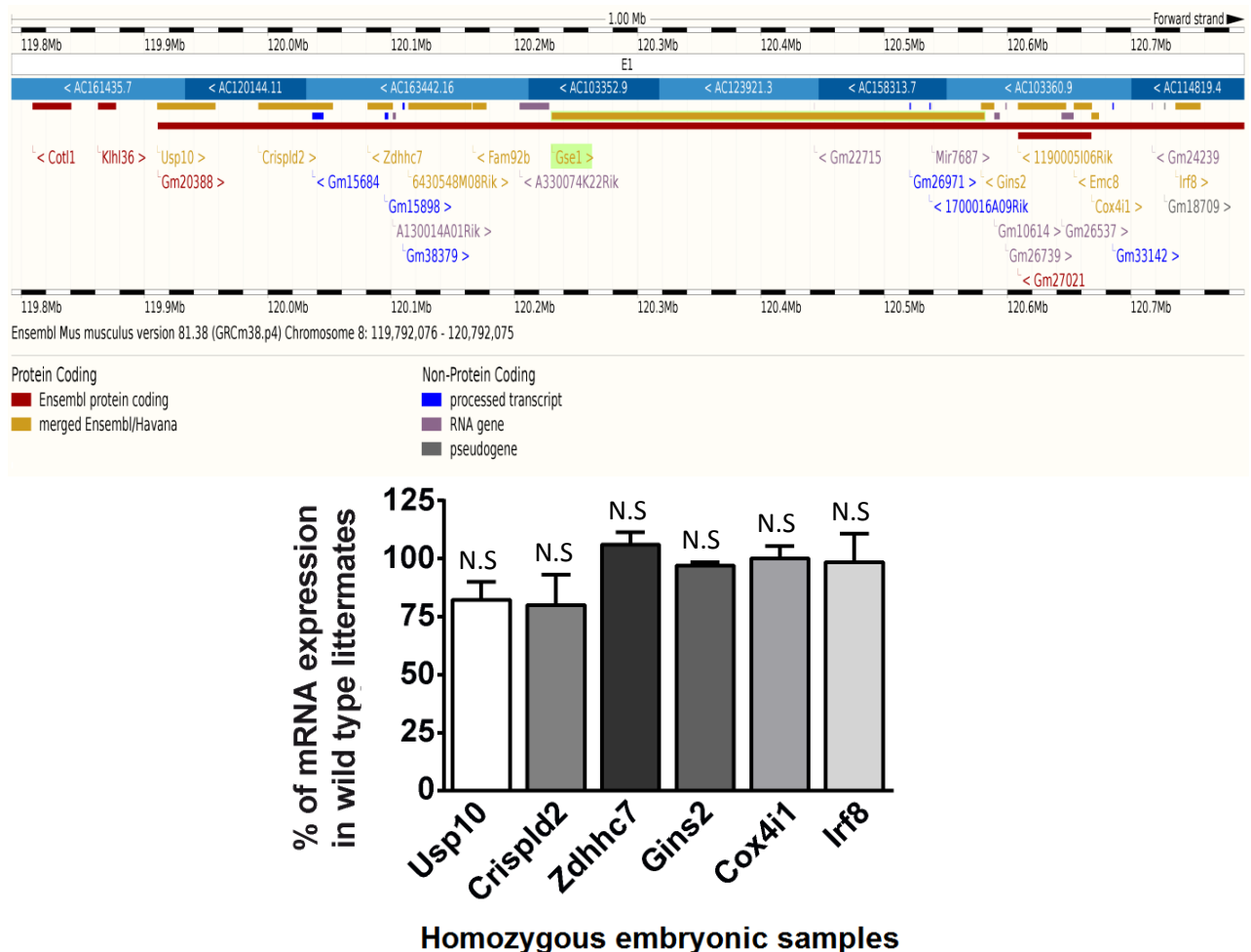


Figure 3.21: Upper: 1 Mb interval of chromosome 8 around *Gse1* gene showing the candidate genes to be affected by the trapping cassette (Adapted from http://www.ensembl.org/Mus_musculus/Component/Location/Web/ViewTop/main?db=core:g=ENSMUSG00000042269;r=8:120281341-120292410;update_panel=1;export=png-5). Lower: The expression level of *Usp10*, *Crispld2*, *Zdhhc7*, *Gins2*, *Cox4i1* and *Irf8* in samples from homozygous embryos compared with wild type littermates and represented by means and standard errors.

3.9 The role of *KIAA0182* in EndMT

EndMT was found to be responsible for the formation of endocardial fibroelastosis, which is involved in the pathogenesis of HLHS. Therefore, we wanted to explore whether *KIAA0182* has a role in EndMT process. First, we wanted to investigate how the expression of *KIAA0182* will be changed by treating HCAEC cells with TGF- β 1, which is already an established protocol to induce EndMT in our lab. Interestingly, it was found that *KIAA0182* was progressively upregulated upon TGF- β 1 treatment, which caused significant upregulation of EndMT markers *SNAIL* and *SLUG* (Figure 3.22), indicating that this gene may be involved in mediating and/or regulating the response to TGF- β 1 treatment in HCAEC cells.

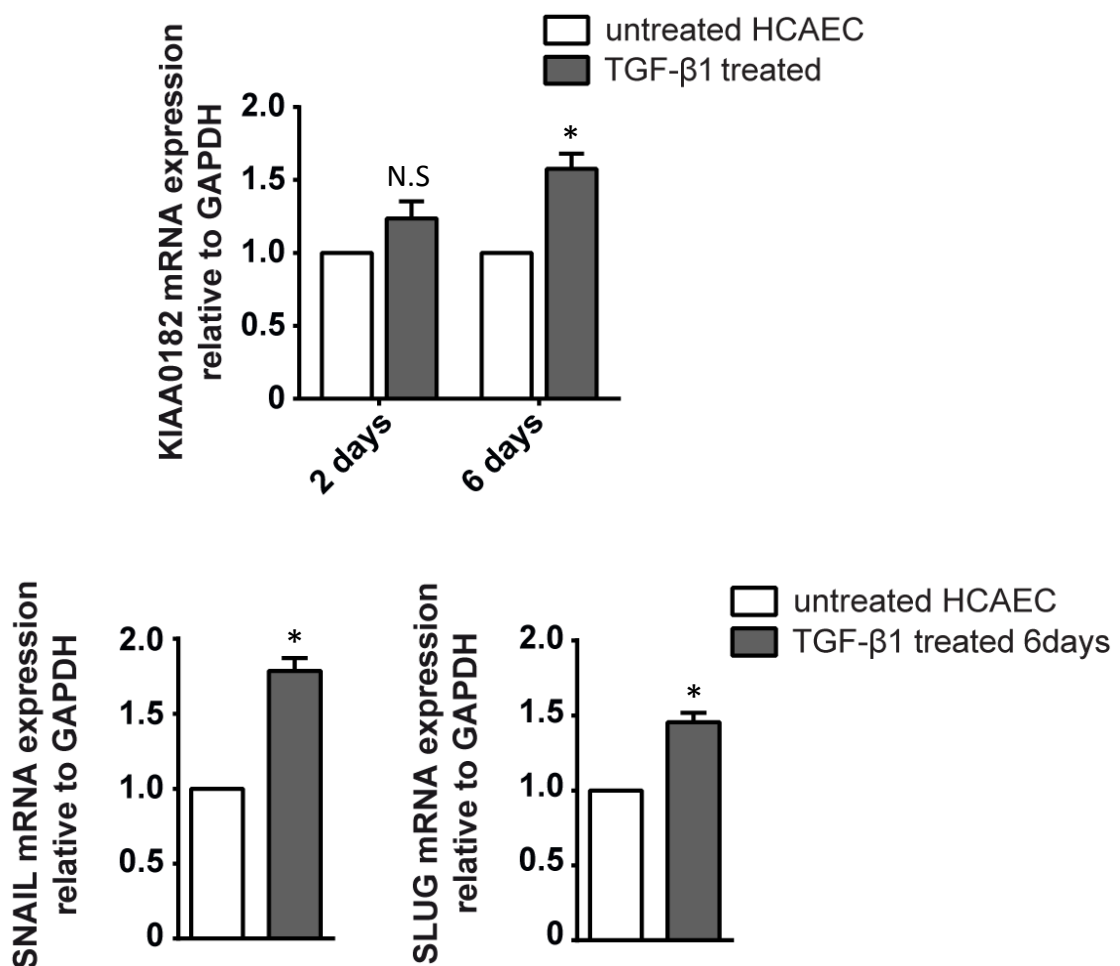


Figure 3.22: Upper: The expression level of *KIAA0182* in human coronary arterial endothelial cells after treating them by TGF- β 1 represented by means and standard errors, showing significant upregulation of *KIAA0182* after 6 days of treatment compared with untreated cells. Lower: The expression level of EndMT markers *SNAIL* and *SLUG* in human coronary arterial endothelial cells after treating them by TGF- β 1 represented by means and standard errors.

These results were also reproduced in mouse endothelial cells, as it was found that TGF-β1 treatment in mouse cardiac endothelial cells (MCEC) to induce EndMT was associated with *Gse1* overexpression (Figure 3.23)

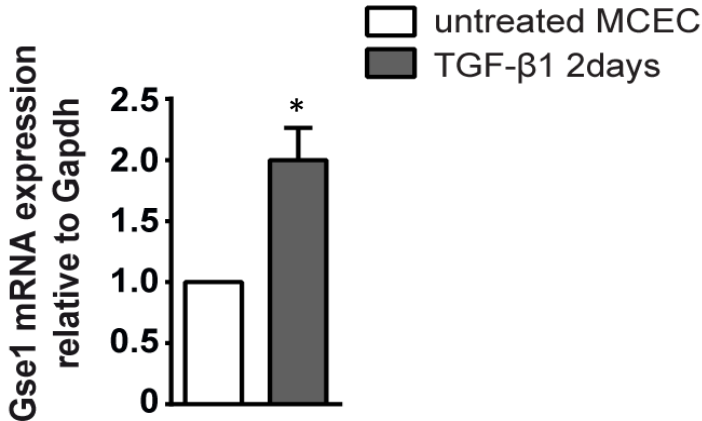


Figure 3.23: The expression level of *Gse1* in mouse cardiac endothelial cells after treating them by TGF-β1 represented by means and standard errors, showing significant upregulation after 2 days of treatment compared with untreated cells.

After that, we tried to find out what the consequences of *KIAA0182* downregulation will be, using HCAEC cells as *in vitro* model. Numerous trials were performed until we were successful in transfecting these cells by siRNA targeting *KIAA0182* due to the difficulty of obtaining good transfection efficiency in endothelial cells. According to our results, we obtained the maximum downregulation 48 hours after the transfection. Surprisingly, that was followed directly in the next day by an increase in the expression of *KIAA0182* to mildly higher levels compared with the cells transfected by scrambled siRNA (Figure 3.24), and this changing pattern was repeated in several trials.

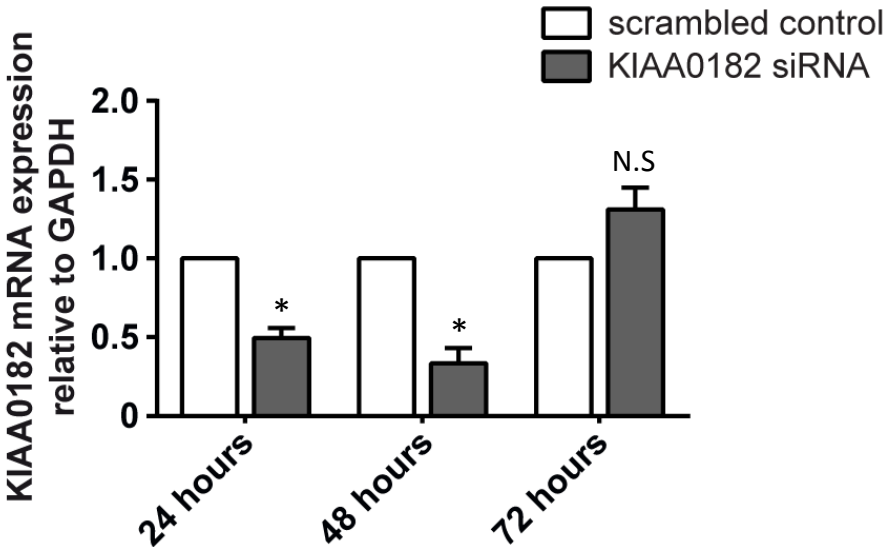


Figure 3.24: The expression level of *KIAA0182* in human coronary arterial endothelial cells after transfecting them by *KIAA0182*-siRNA represented by means and standard errors, showing significant downregulation after one and two days.

We analyzed after that whether the downregulation of *KIAA0182* was associated with any changes in the expression of EndMT markers, and it was found, that the transcription factor *SNAIL* only was slightly but significantly upregulated in the cells transfected by *KIAA0182*-siRNA. That was accompanied by upregulation of the mesenchymal marker α -SMA (Figure 3.25), which may represent spontaneous EndMT in HCAEC cells induced by the downregulation of *KIAA0182*, suggesting a potential function for this gene as a negative regulator of the EndMT process.

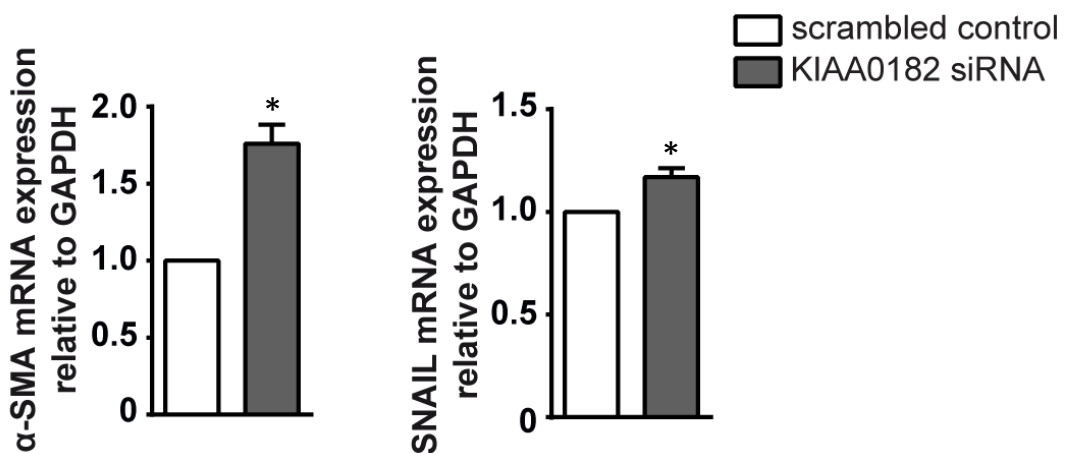


Figure 3.25: The expression level of *SNAIL* and α -SMA in human coronary arterial endothelial cells 2 days after transfecting them by *KIAA0182*-siRNA represented by means and standard errors, showing significant upregulation compared with cells transfected by scrambled siRNA.

3.10 Results of AAC operation

The homozygosity of the mutated allele *Gse1*^{tm1a} was associated with mouse lethality during embryonic development, in spite of the fact that the mutated allele was only hypomorphic. The hearts of the heterozygous animals in our study showed small but significant downregulation in the expression of *Gse1*. Those mice in the Europhenome Mouse Project had a cardiac phenotype represented by heavier hearts, and in our mice colony it was found that the heart weight of the mice was mildly more in heterozygous males and females, but without significant difference compared with wild type animals (Figure 3.26). Therefore, we wanted to confirm our *in vitro* EndMT results by performing AAC operation on these mice as *in vivo* model

for cardiac fibrosis. Heterozygous and wild type age-matched male mice were subjected to AAC or sham surgery, using several echocardiographic parameters, representing the systolic function and cardiac hypertrophy of the left ventricle, and the degree of cardiac fibrosis after two weeks as endpoints for this experiment.

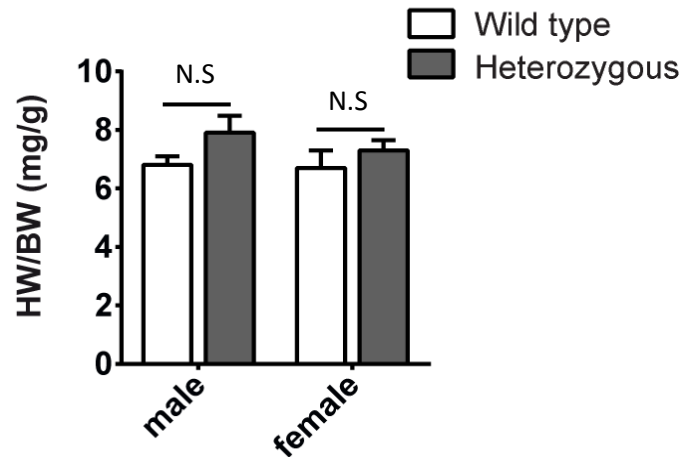


Figure 3.26: The heart weight in heterozygous and wild type mice of both genders normalized to body weight and represented by means and standard errors.

The AAC operation was performed on 20 mice with 35% mortality during or directly after the operation, mostly because of bleeding, and the sham operation was performed on 11 animals with no operative mortality leaving 24 animals alive in total, creating 4 groups with 5-7 animals in each group. There was no statistically significant difference regarding the age of the mice among the 4 groups (Figure 3.27). The pressure gradient across the constriction site was significantly higher in the groups of AAC operation compared with the groups of sham operation, but without any significant difference between the heterozygous group and the wild type group of mice undergoing AAC operation or sham operation (Figure 3.27).

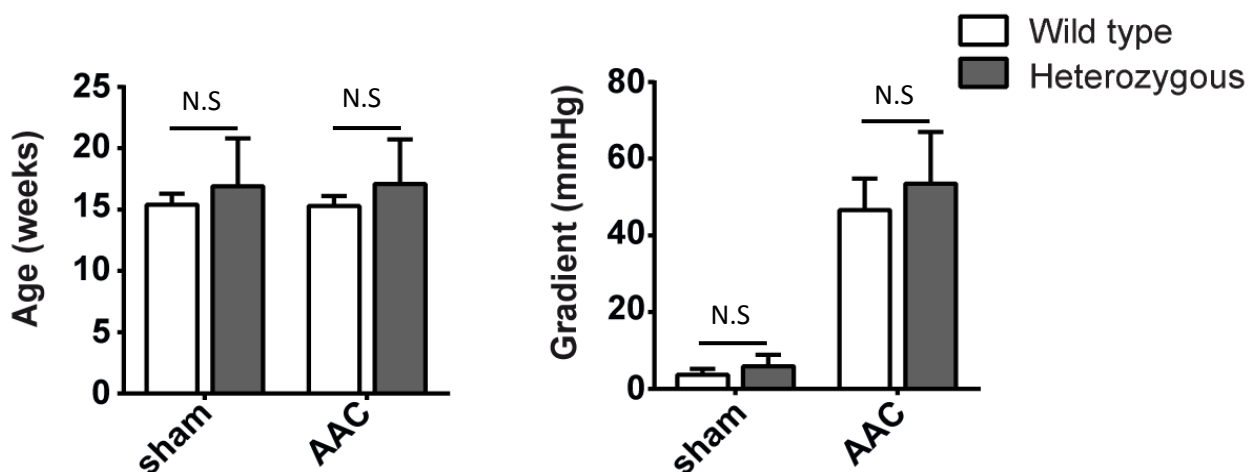


Figure 3.27: The characteristics of the operated mice in this study represented by means and standard deviations, showing that the 4 groups were not significantly different regarding the age of the animals. The pressure gradient was as supposed significantly higher in both AAC groups compared with sham groups, but not significantly different between the heterozygous and wild type groups.

The echocardiography study for the mice one week after the operation showed that the heterozygous mice were not significantly different regarding the systolic function of the left ventricle after AAC operation and left ventricle hypertrophy, compared with the wild type mice. That was estimated depending on the ejection fraction and fractional shortening of the left ventricle; and the thickness of the anterior (AWT) and posterior wall (PWT) respectively (Figure 3.28). No difference was also detected regarding end-diastolic (LVEDD) and end-systolic dimensions (LVESD) of the left ventricle or the mass of the left ventricle normalized to the body weight of the animal. Similarly, the quantification of the fibrosis in the left ventricle did not show any statistically significant difference between the heterozygous and the wild type animals.

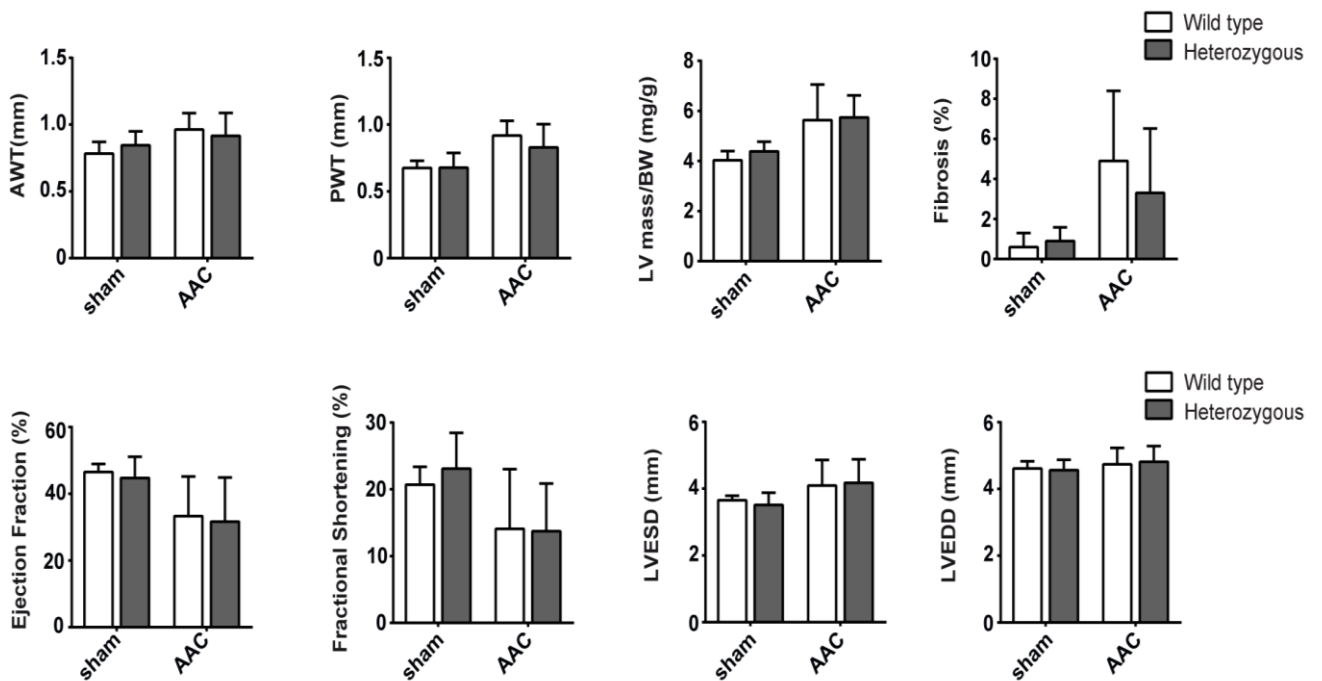


Figure 3.28: The results of the AAC operation in this study, represented by means and standard deviations, showing that there was no significant difference between the heterozygous and wild type animals regarding all analyzed echocardiographic parameters of the left ventricle, and regarding the fibrosis ratio two weeks after the operation.

4. Discussion

The scientific basis of our study depends on the unpublished data that a *de novo* mutation in *KIAA0182* gene was identified in HLHS, and it was not possible to be predicted whether it is a loss-of-function- or gain-of-function-mutation. However, the cardiac phenotype of small ventricle with large atrium in zebrafish model of *KIAA0182* knockout suggested a relevance of *KIAA0182* knockout or downregulation in congenital heart diseases.

In this study we aimed to find a link between this gene and HLHS by utilizing the 'Knockout-first' mice with the mutant allele $Gse1^{tm1a(EUCOMM)Wtsi}$ as *in vivo* model. These mice represent a useful biological tool to explore the phenotypes associated with inactivation of unknown genes, such as *KIAA0182* or *Gse1* in mouse, and to the best of our knowledge, this is the first study trying to answer this question using this model. Additionally, we tried to investigate the importance of this gene in EndMT process, which is described to be involved in the pathogenesis of HLHS. Our results indicated that the homozygosity of the allele $Gse1^{tm1a(EUCOMM)Wtsi}$ was clearly lethal. This mutated allele behaved as a hypomorphic allele, not as a null allele as it is supposed to do. It was associated with upregulation of the circular RNA of *Gse1* gene, which may be functional as miRNA sponge, and might play a role in the observed lethality.

Before making any investigations using these mice, and in order to ensure the reliability of our results, it was important and time-worthy to exclude problems reported before in animals generated from ES cells with mutated allele of the same design as the allele $Gse1^{tm1a}$. According to the study of Ryder and his colleagues, the targeting of wrong gene or the absence of the 5'end of the trapping cassette represented the most frequent reasons for the failure in the quality control tests applied to the mice (Ryder et al. 2013). It should be taken into consideration that the ES cells in that study were only tested by long-range PCR and sequencing before using them to produce the mice. However, the ES cell clone (EPD0557_2_C07), from which the mice in our study were generated, passed more detailed tests including Southern blot and karyotyping, which were performed by the providing company (http://www.mousephenotype.org/data/alleles/qc_data/es_cell/EPD0557_2_C07/).

The quality control tests performed by us in this study included several short-range PCRs, detecting different parts of the trapping cassette. These parts were the 5' end, the LacZ cassette and the region of 3' LoxP, in addition to the cassette of neomycin resistance which was confirmed through our genotyping protocols. According to our results, none of the previously mentioned parts was missing in the mutated allele, confirming the integrity of the trapping cassette. To exclude the possibility of wrong targeting, which was reported to be relatively common, we performed two experiments, which were break point loss of allele (BP-LOA) assay and long-range PCR. In the first experiment it was shown that all the tested heterozygous animals had the same number of copies of the wild type allele (approximately one copy) compared with two copies in the wild type animals and no copies in the homozygous embryos. That was determined according to qRT-PCR results using TaqMan probe, which detects the deleted region from *Gse1* gene due to the insertion of the trapping cassette.

The second experiment was 3' long-range PCR, where a product with the size of more than 4 kb was amplified from the heterozygous but not from the wild type animals, indicating the specificity of this PCR, which was also confirmed by the sequencing results of the amplicon. Another evidence for the correct targeting of *Gse1* can also be extrapolated from the genotyping results, as we were able by two genotyping protocols for *Gse1* gene to identify homozygous animals (during embryonic development only). This would not be possible if another gene was targeted and a product will always be amplified in that case by PCR detecting the wild type allele of the gene of interest (*Gse1* gene in our study), as the trapping cassette is not inserted in this gene to interrupt the amplifying process (Ryder et al. 2013).

After proving that *Gse1* was for sure correctly targeted, our genotyping results showed that none of the genotyped newborn mice was homozygous, and these results were produced using two protocols without any contradictive events. It should be noticed that there are some redundancy in the results of the genotyping PCRs detecting the wild type allele in these two protocols. Both pairs of primers, *Gse1*-5arm-WTF with *Gse1*-Crit-WTR in the first protocol and *Gse1*-F with *Gse1*-ttR in the second, are actually amplifying the same region of the wild type allele. Therefore, both of them will automatically give the same positive or negative result, provided that

the required PCR conditions are fulfilled. These redundant results could have been avoided by using the (BP-LOA) assay as a genotyping modality, not only as a quality control test. This PCR can be used together with another qRT-PCR assay with TaqMan probe that detects the LacZ or the neomycin resistance cassettes to decide the copies number of the mutated allele. However, we cannot neglect in this case the economic burden of using such protocol to genotype each animal in the colony. Regarding the mutant allele, the problem of redundancy was avoided. The primers used to detect the mutant allele in the first protocol, Gse1-5arm-WTF and Tm1a-5mut-R, are amplifying the 5' end of the trapping cassette. The two pairs of primers used in the second protocol, Gse1-ttR with CSD-neoF and Gse1-R with CSD-loxF, are detecting different regions, which are the cassette of neomycin resistance and the 3' LoxP site respectively. This indicates that all heterozygous genotyped mice, not only G1 heterozygous mice tested by the quality control PCRs, were proved to have these regions inside their genome.

In total, 134 animals born from several heterozygous with heterozygous matings were genotyped, and no homozygous newborn mice were found. According to that, it can be safely concluded that we observe embryonic lethality in the homozygous animals. It might also be suspected that unspecific product, which has the same size predicted to be obtained from the primers pairs in wild type reactions, was amplified, and this might be the reason why all the genotyped animals were found to have wild type allele. However, this is highly unlikely to happen for two pairs of primers, and the sequencing results of the product from both of them proved the specificity of our genotyping protocols, which represents one of the most important strengths in this study.

It can also be concluded that heterozygous animals are not affected by this lethality, as the numbers of the surviving mice are not significantly deviated from the expected ratio of 1:2 in previous matings, and no deviation from the ratio of 1:1 was found when matings were between heterozygous and wild type parents. Our results indicated also that both genders are involved in this phenotype, as the numbers of the surviving males and females are not significantly different, meaning that the lethality happened in similar numbers of animals from both genders. It should be kept in mind in this context that the genotyping was performed for the newborn mice 4 weeks after birth. This keeps the door open for the possibility that the homozygous

animals died after birth during the time period before genotyping. However, it was clearly clarified with the responsible caretakers of the animals that each newborn mouse should be reported and genotyped. Actually, in three cases we were informed about weak animals, which did not seem to be able to survive after birth, and those animals were also genotyped and all of them were heterozygous.

Detailed investigations are currently undergoing to answer the questions regarding when, why, and how the homozygous animals died during embryonic development. However, one of the most essential questions also is whether the observed lethality can be caused by the incomplete knockout of *Gse1* gene, which was detected according to our results. At the beginning we should notice that the expression of different splicing variants in different organs can be predicted depending on our Western blot results. The antibody used in that experiment is against an epitope located in the second and third exon of the isoform Gse1-002 in mouse and human. It can be predicted that this polyclonal antibody detects also the isoforms Gse1-001, Gse1-003 and Gse1-006 in human; and the isoforms Gse1-001 and Gse1-003 in mouse, as these isoforms contain the previously mentioned exons, but not the other isoforms. Similarly, the primers that we have used for qRT-PCR in human and mouse show this limitation. Therefore, it should be always kept in mind by the interpretation of our present results that we may not be detecting all expressed isoforms of *Gse1* gene. Further efforts are certainly needed to establish primers that can determine the expression level of each isoform in specific way, because different isoforms can have different functions as described in previous studies for other genes (Castelli et al. 2014; Heemskerk et al. 2011).

However, the primers m.Gse1-F and m.Gse1-R that we used to analyze the expression level of *Gse1* in different genotypes are located in exon 6 and 7 from the isoform Gse1-002 respectively. The trapping cassette in the mutated allele $Gse1^{tm1a(EUCOMM)Wtsi}$, which is based on the design of gene-trap vector, is inserted in the intron upstream from the third exon. That is supposed to produce truncated mRNA consisting of the first two exons and the LacZ reporter (Figure 4.1) (Skarnes et al. 2011; Stanford et al. 2001). Therefore, these primers should be able to show the difference in the expression level due to the insertion of this trapping cassette, and we found that the expression of *Gse1* in the RNA samples from the homozygous embryos was less than 50% downregulated compared with the wild type. That was

shown using samples extracted from the left limb of E15.5 embryos, and it was also confirmed on protein level by detecting GSE1 protein in the homozygous samples, meaning that the mutated allele behaved as a hypomorphic allele. The results in the heterozygous embryos showed that we have less than 20% downregulation effect. We tested also different organs from adult wild type mice, and they showed different expression level for *Gse1*. Additionally, the tested organs showed variable level of *Gse1* trapping, as the difference in *Gse1* expression between wild type and heterozygous animals was significant in heart and kidney, but not in lung and liver. This hypomorphic state was previously reported in different previous papers about gene trapping. When the trapping cassette is inserted in an intron, which is the case in our mutated allele, different splicing pattern can happen, making the trapping not completely efficient (Hanstein et al. 2013; Maguire et al. 2014; Stanford et al. 2001). However, as we found that *Gse1* had different trapping efficiency in adult mice, this can also be the case in embryonic tissues. This difference may be important in determining the reason of the lethality, which can be related to the organ with the most needed expression of *Gse1* during certain stage of embryonic development.

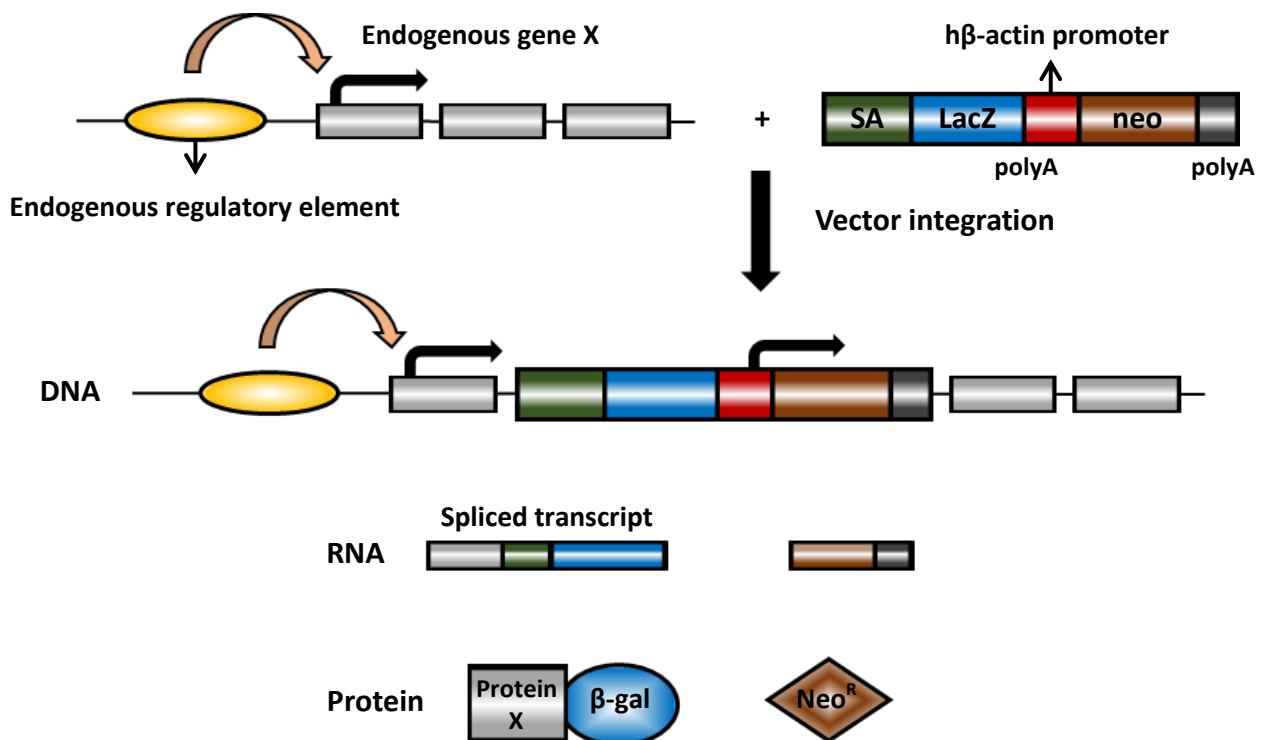


Figure 4.1: The design of gene-trap vector, consisting of a splice acceptor (SA) site and promoterless LacZ cassette followed by polyA tail, which is supposed to disrupt the expression of the targeted gene on mRNA and protein level (Adapted and modified from Stanford et al. 2001).

In a recent study using transgenic mice with similar design to our mutated allele (the 'Knockout-first' allele) for different gene, it was reported that the downregulation effect was relatively similar in different organs (around 65 % decrease in the expression level of the targeted gene in homozygous animals compared with wild type). However, it was more than the decrease measured using samples from the tail of the animals (around 50%). Careful inspection of the data presented in that study shows also that the trapping effect in the bladder was less than 55 %, which was clearly lower than the reported downregulation in the liver (more than 70 %) (Hanstein et al. 2013). Similar differences were also described in another study using the same trapping mechanism but without an independent promoter upstream from the sequence for neomycin resistance, but this difference was more clearly manifested in the heterozygous animals, as the homozygous mice showed more than 95 % trapping effect in all tested tissues in that study (Chen et al. 2013).

All this argues for the possibility that trapping efficiency differs from tissue to tissue. Therefore, we cannot assume that the observed decrease in *Gse1* expression, as it was estimated in the tested embryonic RNA from the animals limbs, necessarily reflects the trapping state in each tissue or cell type and at each timepoint of the embryonic life. The apparently inadequate trapping of *Gse1* in the homozygous embryos does not absolutely mean that the lethality is not caused by *Gse1* downregulation, but may on the opposite suggest that the expression of *Gse1* is so important temporarily or spatially for the embryo, that only less than 50 % decrease in this expression is enough to cause embryonic lethality.

On the other hand, we should not neglect the fact that the selection cassette with its promoter is still inside our mutated allele. Consequently, it is not usually recommended in this stage to consider that the generated phenotype is surely caused by the observed downregulation of *Gse1* gene. In principle, it is possible that the observed lethality is due to off-target effect of this promoter, which may have bidirectional activity and can cause changes in the expression profile of one or more of the neighboring genes located more than 100 kb apart from the targeted gene (Johnson and Friedmann 1990, Pham et al. 1996; Ren et al. 2002; Ryder et al. 2014; Scacheri et al. 2001).

This undesired scenario was described in details in the paper of Maguire and his colleagues, and a lot of conclusions can be drawn by interpreting the following results

of that study. The researchers showed in their study that the homozygous animals with two mutated alleles of *Slc25a21* gene (*Slc25a21*^{tm1a}) had a phenotype comprising orofacial abnormalities, hearing defect and inflammation in the middle ear. However, in spite of the fact that the expression of *Slc25a21* was more than 70 % downregulated, the expected symptoms of 2-oxoadipate acidaemia were not found in these mice. In order to further investigate the mechanism behind these results, the mice were bred with transgenic mice expressing FLP- and Cre-recombinase. Surprisingly, the observed phenotype was not reproduced in the homozygous mice for the mutated allele *Slc25a21*^{tm1b} or *Slc25a21*^{tm1d}, although they had lower expression of the studied gene compared with the homozygous mice for the allele *Slc25a21*^{tm1a} (almost 0 % and 13 % of the expression level in the wild type mice respectively). These findings suggested that the phenotype was caused by off-target effect of the selection cassette. In order to prove that, the expression level of several genes located within 1 Mb interval around *Slc25a21* was investigated. Only one of these genes, *Pax9*, was found to be affected exclusively in the *Slc25a21*^{tm1a}-homozygous animals. *Pax9* downregulation was known before to cause similar phenotype in mice (Maguire et al. 2014).

The first conclusion from this study is that the magnitude of the trapping effect in the targeted gene, neither in the case when it seems adequately high nor when it is relatively low as we observed for *Gse1*, is enough as evidence for or against the causal relationship between the gene of interest and the observed phenotype. Second, it is not possible to predict exactly how the insertion of the trapping cassette will affect the splicing pattern of the targeted gene, either before or after breeding with Cre mice, and wide spectrum of trapping efficiency can be expected using the same design of trapping cassette. This may range from less than 50 % downregulation according to our results up to more than 95 % as shown in another publication (Nijnik et al. 2012).

Third, it may be necessary to analyze the expression level of the genes flanking *Gse1* on chromosome 8 (Figure 3.21), and according to the results in our study, none of the tested neighboring genes was affected in the homozygous embryos. However, the detection of any mis-regulated gene will not be adequate to prove that it is caused by the presence of the selection cassette, as it can be simply a secondary consequence of the decreased expression of *Gse1*. If it turns out that one of the

neighboring genes is associated with a phenotype similar to the one reported in our study, this gene should be kept in mind as a possible target for this potential off-target effect. To the best of our knowledge, none of the genes shown in figure 3.21 is described before to cause embryonic lethality in mouse model (<http://www.informatics.jax.org/>). All this together argues more in favour of the conclusion that a direct causal relation between *Gse1* knockdown and the lethality phenotype actually exists.

The last and the most important conclusion is that the phenotyping of the mice should be continued by investigating the presence or the absence of any similar or non-similar phenotypes after breeding with Cre mice. This represents the logical method to confirm our results in the 'Knockout-first' mice. Additionally, finding out the exact function of the studied gene (*Gse1* in our case) and in which biological processes it is involved represents very essential step. Only that can prove with high reliability that the observed lethality was actually caused by the change in the expression of *Gse1* gene, taking into consideration that several possible genomic mechanisms may be involved in this phenotype. Interestingly, it was recently reported that the short sequence of LoxP or FRT sites, not only the neomycin cassette with its promoter, can cause unpredicted off-target effects far away from the targeted gene, indicating that we should be very careful by interpreting any phenotype in these transgenic mice (Meier et al. 2010).

Therefore, we started breeding the heterozygous mice with FLP mice to get rid of the trapping cassette. This will leave the animals with the mutant allele *Gse1*^{tm1c}, which only differs from the wild type allele by the presence of one FRT and two LoxP sites, restoring thereby the normal expression of the gene. In order to do that ROSA-FLPe mice were used, which contain the ROSA26 promoter that is supposed to guarantee moderate and ubiquitous expression of FLP-recombinase in all tissues of embryos and adult mice (Casola 2010; Giel-Moloney et al. 2007; Nyabi et al. 2009). According to our genotyping results, we obtained mosaic mice in the first generation after breeding with FLPe mice, where the FLP-recombinase was not efficient in all cells. Similar findings were described before for Cre and FLP mice, and especially for FLPe-recombinase (enhanced FLP), which has less recombining efficiency compared with FLPo-recombinase (optimized FLP) (Farley et al. 2000; Ryder et al. 2014; Schaft et al. 2001; Wu et al. 2009).

However, in the following generation after breeding the mosaic mice with wild type mice, we were able to identify pure heterozygous mice for the mutant allele $Gse1^{tm1c}$ without FLP-recombinase, indicating that the recombination of the FRT sites occurred in the germ line of the parent. Only these mice are suitable to be used for generating and phenotyping the mice with $Gse1^{tm1c}$ allele without any possible off-target effects due to the presence of the FLP-recombinase cassette inside their genomic DNA. Phenotyping these mice should not show any difference compared with wild type animals. What should also be noticed in this context is that our genotyping protocols were suitable to distinguish between the mosaic and non-mosaic mice after breeding with FLP mice without redundancy or contradiction in the results. That was also confirmed by performing additional PCR using the primers LacZ_2_small_F and LacZ_2_small_R to further ensure the complete deletion of the trapping cassette from the genome of the mice from the second generation. The success after all these steps in establishing this mice strain represents therefore another valuable result from this study.

In addition to the previously shown efficient *in vivo* generating of the mice with $Gse1^{tm1c}$ allele, which will be used later to produce the null allele $Gse1^{tm1d}$, we also performed an experiment to get the LacZ-tagged null allele $Gse1^{tm1b}$ *in vitro*. This was performed using Cre-recombinase-adenovirus and fibroblasts isolated from the kidneys of heterozygous and wild type adult mice (fibroblasts isolated from other organs such as heart or from the skin of the embryos were also used but their survival before or after the transduction were not sufficient for establishing this protocol). Our results, along with the quality control tests performed by us and by the providing company of the mice, represent adequate evidence for the conditionality of the 'Knockout-first' allele.

As previously mentioned, *Gse1* gene has a circular RNA isoform in mouse generated from the same exon as in human, which therefore may be functional (Jeck et al. 2013; Lasda and Parker 2014; Memczak et al. 2013). In our study, a protocol to detect the circular RNA of *Gse1* was established, and a difference in the expression level of this circular RNA was observed among mice from different genotypes. That might contribute to the observed lethality in the homozygous mice, as circular RNA can play the role of miRNA sponge. We were able to show that the RNA of the scrambled exon detected by our outward-facing primers is resistant to the activity of

RNase R and not reverse-transcribed by oligo (dT). This is considered as an adequate proof for the circularity of this RNA (Lasda and Parker 2014). These experiments were important to be performed because the detection of a scrambled exon by outward-facing primers does not necessarily mean that we detect circular RNA. The proof about the presence of a scrambled exon, which only means that the sequence of the detected RNA is not in the same order compared with the corresponding genome, is not sufficient evidence for its circularity, as this could also exist in RNA with linear form. That can happen because of DNA rearrangements, tandem duplications in the DNA or trans-splicing (Figure 4.2) (Guo et al. 2014; Lasda and Parker 2014; Salzman et al. 2012).

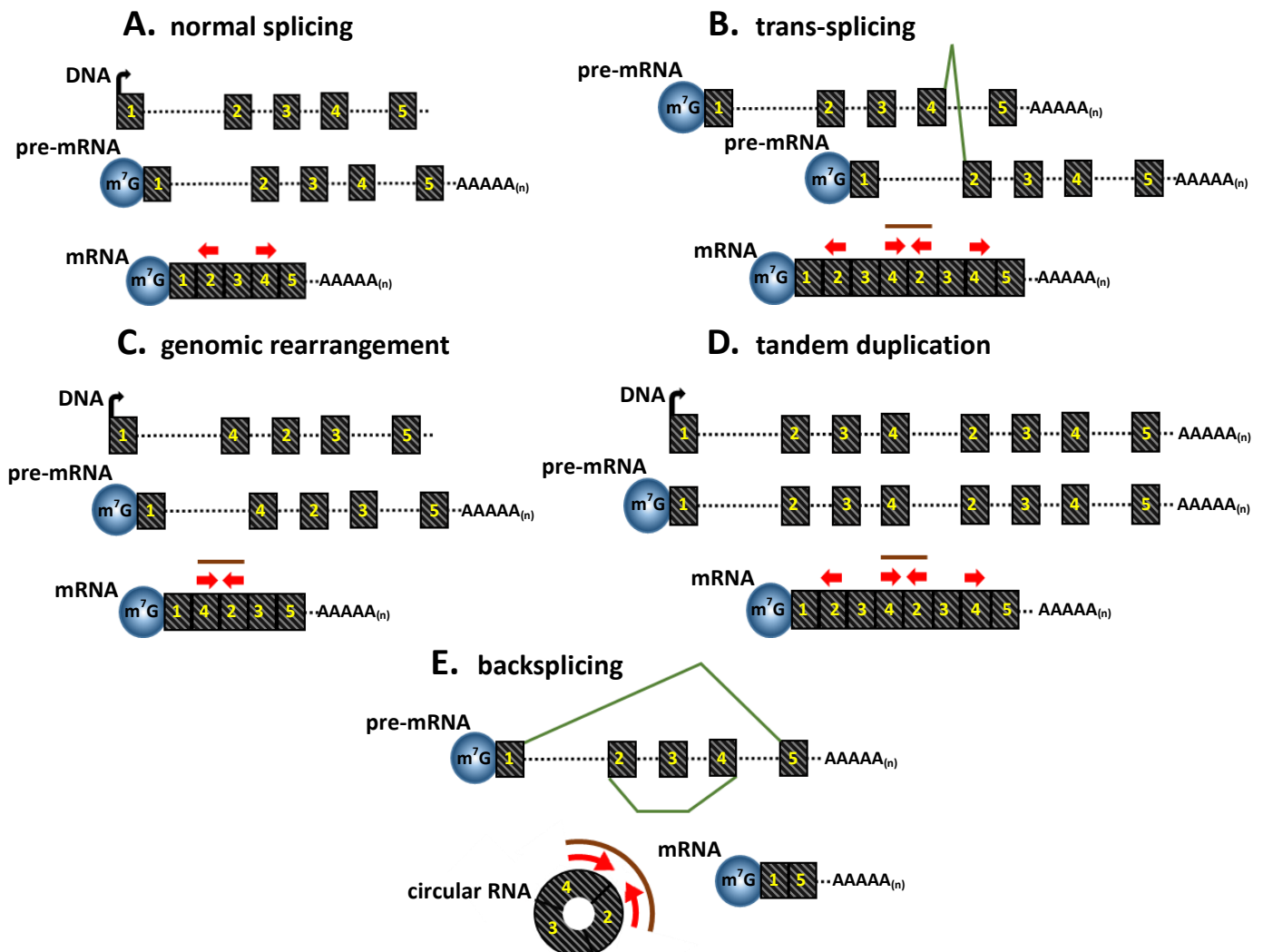


Figure 4.2: Possible mechanisms for generating scrambled exons detected by outward-facing primers (Adapted and modified from Lasda and Parker 2014).

The resistance of circular RNAs to the exonuclease RNase R can further confirm the circularity, because these RNAs have no 5' and 3' ends. This causes relative enrichment of the circular RNAs in the treated samples compared with untreated samples, as only linear RNAs are degraded by RNase R treatment (Jeck et al. 2013; Salzman et al. 2012; Suzuki and Tsukahara 2014). The second proof for the circularity, which was provided in our study, was the depletion of the scrambled exon in oligo (dT)-reverse-transcription-primed samples compared with samples prepared using random hexamers. Such results could not be obtained if we were detecting linear RNAs containing this scrambled exon, including that generated by DNA rearrangements, tandem duplications or trans-splicing (Guo et al. 2014; Jeck et al. 2013; Lasda and Parker 2014). Other methods are described to provide additional evidences for the presense of circular RNAs, such as Northern blot with probe complementary to the circularized exon, or virtual Northern as described in some studies. That depends on the fact that circular RNAs will migrate faster in the gel due to their smaller size compared with their normal corresponding linear RNAs, or compared with the trans-spliced transcripts which may be even larger than normal transcripts (Jeck et al. 2013; Memczak et al. 2013). Another suggested method is TRAP electrophoresis using solidifying agarose mixed with RNA, which will make the circular RNAs trapped, and only the linear RNAs will be able to migrate into the gel (Hansen et al. 2011).

After confirming the circularity of the detected scrambled exon, we found in our study that *Gse1* circular RNA is expressed more in the embryonic RNA samples from the homozygous compared with the wild type animals. That was proved using specific TaqMan probe for the junction between the end and the beginning of the scrambled exon. These findings need to be further investigated by testing more samples from other embryonic and adults mice tissues, as the circular RNA is also described to have variable expression among different tissues (Bachmayr-Heyda et al. 2015; Lasda and Parker 2014). Another method to confirm the results regarding the upregulation of the circular RNA in homozygous samples is needed. Northern blot as a semi-quantitative analysis is helpful to achieve this purpose (Hansen et al. 2013; Memczak et al. 2013, Salzman et al. 2012).

In the future, we should certainly also investigate the mechanism behind the upregulation of the circular RNA. The process of backsplicing is suggested to be

controlled on the first hand by the presence of complementary sequences, which cause pairing between the introns flanking the circularized exon or exons, inducing thereby the circularization. On the other hand, other pairing events between complementary sequences within these flanking introns, which are mediating the normal splicing, are in competition with the pairing between introns (Jeck et al. 2013; Zhang et al. 2014). Therefore, one explanation for the observed upregulation of the circular RNA can be that pairing within the intron downstream of the circularized exon is disrupted due to the insertion of the trapping cassette in this particular location, or pairing between the two introns around this exon is augmented. Another mechanism can be the lack of self-regulatory effect exerted by the linear isoforms of *Gse1* that can control the splicing process of the pre-mRNA, and this cannot be confirmed or excluded based on our results. However, we showed that the organ with the lowest expression of the circular RNA among the tested organs in the mouse was the liver, which also had high expression of linear RNA isoforms (depending on the results obtained using the primers m.Gse1-F and m.Gse1-R, which are supposed to detect linear isoforms only). The heart showed on the opposite relatively high expression of the circular RNA, and the lowest expression of the linear isoforms. This may suggest the presence of a tissue-specific mechanism controlling the ratio of the circular/linear isoforms of *Gse1* gene.

We tried also in this study to investigate whether the circular RNA of *Gse1* is functional or not. The most studied function for circular RNAs in general is their role as miRNA sponge. It was shown by Hansen and his colleagues that one circular RNA, called ciRS-7 or CDR1as, which arises from the antisense transcript of *CDR1* gene, can modulate the activity of one miRNA called miR-7. This process is mediated by sequestering miR-7, as ciRS-7 has more than 70 conserved seed matches for this miRNA. Similar role was also described for another testis-specific circular RNA, which comes from the murine *Sry* (Sex-determining region Y) gene (Hansen et al. 2013). Other suggested functions for circular RNA are the regulation of alternative splicing by competing with other splicing options and altering the structure of the remaining pre-mRNA splice, which can simply have different function or no function at all (Ashwal-Fluss et al. 2014; Lasda and Parker 2014). They might also be able to bind to different RNA-binding proteins, as delivery vehicles or as regulators for their activities (Hentze and Preiss 2013).

The expression results of the group of genes tested in our study; *Foxp4*, *Thy1*, *Actn4*, *Wnt1*, *Kmt2d* and *Tfap2b* are considered as a first and essential step to prove that the circular RNA of *Gse1* is functional as a miRNA sponge, representing another novelty for our research. In order to do that, we followed the strategy of identifying the miRNAs that may be sequestered by the circularized exon of *Gse1* according to the database miRBase (<http://www.mirbase.org/search.shtml>). Thereafter, the miRNA, mmu-miR-6968-5p, which showed the best match for its seed region with our circular RNA, was chosen. Previous genes were analyzed as they are the top 6 genes on the list of possibly targeted genes by this miRNA depending on the database miRDB (<http://mirdb.org/miRDB/>).

According to our results, most of these genes showed no difference in the expression level between homozygous and wild type embryos, and only one gene, *Thy1*, showed a significant decrease in the homozygous samples. However, this change is in the opposite direction to the expected change according to the hypothesis that the upregulated circular RNA of *Gse1* works as a sponge for the miRNA mmu-miR-6968-5p, inhibiting thereby its suppressive activity. The last gene, *Tfap2b*, was almost significantly upregulated in the RNA samples from the homozygous embryos, which may provide the first evidence for the rightness of this hypothesis. Other experiments are certainly needed to prove the causal relation between the increased expression of the circular RNA, and the upregulation of *Tfap2b*. These can include overexpressing the circular RNA of *Gse1* or knocking it down using siRNA designed to target the sequence formed by the backsplicing of the circularized exon (Hansen et al. 2013; Jeck et al. 2013). After that, we can investigate whether a corresponding change will happen in the expression level of *Tfap2b* or another possible target.

It is important to mention here that the negative effect of miRNAs on the target genes can be mediated through two distinct mechanisms, which are translational repression and/or mRNA degradation (McDanel 2009; van Rooij 2011). This means that it may be insufficient to use qRT-PCR for investigating the expression of the genes that can potentially be targeted by the miRNA mmu-miR-6968-5p. It is possible that this miRNA does not induce the mRNAs of those genes to be degraded, but it may only suppress the translation of their proteins. Therefore, we need also to analyze the expression of suspected genes on protein level, otherwise we may miss

some genes, by which only the protein level is altered due to the upregulation of the circular RNA, and consequently its augmented role as a sponge for miRNAs. Interestingly, it was also found that the effect of miRNAs on gene expression is not always negative. According to the cell conditions and the recruited micro-ribonucleoproteins, some genes were found to be upregulated by one miRNA and downregulated by another; the same miRNA can also cause increase or decrease in the expression of its target genes (McDanel 2009; Valinezhad Orang et al. 2014). This makes the interpretation of our results even more complicated, as we should at first define the basic response of the suspected genes to the studied miRNA in each cell condition. That can be achieved by transfecting the cells with miRNA mimics, and identifying thereafter the change pattern in the expression of targeted genes as it was done in the previously mentioned work of Hansen and his colleagues (Hansen et al, 2013). The importance of the circular RNA as miRNA sponge can then be determined by modulating the expression of this circular RNA and investigating any potential alteration in the basic change pattern induced by this miRNA.

Reviewing the literature about the previously mentioned 6 genes showed interesting results. *Tfap2b* gene, which was found in our study to be almost significantly upregulated in the samples from homozygous embryos, is a transcriptional factor that is important for the development of the aortic arch and ductus arteriosus. It was shown to be mutated in Char syndrome, which includes PDA, facial dysmorphism and hand anomalies (Zhao et al. 2011). Additionally, in another study about knockout mouse model for the gene *CITED2* (CBP/p300-interacting transactivator with ED-rich tail 2) cardiac defects were reported. *CITED2* gene was found to be essential for the normal activation of the three isoforms of *Tfap2* (*Tfap2a*, *Tfap2b* and *Tfap2c*) by *EP300* and *CREBBP*. Mutations in *CREBBP* gene are responsible for Rubinstein-Taybi syndrome, where HLHS is described as one of its cardiac abnormalities, suggesting that a part of the observed defects in this syndrome can be related to *Tfap2b* gene (Bamforth et al. 2001).

Interestingly, all the previous results led us back to the point where we started, which is HLHS, where *KIAA0182* was found to be mutated. This could be the first clue regarding the role of *KIAA0182* gene in the development of HLHS, which supports further the hypothesis about the genetic etiology of this disease. As it was previously mentioned, several genes were accused to be responsible for HLHS, but

the evidences regarding the involvement of any of these genes are still not sufficiently conclusive. It was suggested before that HLHS may have a heterogeneous genetic etiology with a complex pattern of inheritance and penetrance, another hypothesis was that the accumulation of several mutations or rare variants is necessary to cause the disease. The exact determination of the involved genes can be helpful to understand the molecular mechanism behind HLHS, and possibly to make the right decision regarding the management plan (Grossfeld et al. 2009; Iacone et al. 2012).

We should keep in mind in this context that in spite of the improvement in surgical and medical care, the five-year survival rate for HLHS remains only 50-70 % after surgical repair. The outcome in these patients is strongly related to their characteristics, such as gestational age and genetic susceptibility (Feinstein et al. 2012). One possible surgical approach for HLHS is the strategy of two-ventricle repair for the patients with patent mitral and aortic valves and mild hypoplasia in the left ventricle, as it was shown that left ventricle can resume growing in some cases after birth (Minich et al. 1997). The left ventricle in this approach may not be able later to support the systemic circulation, making the other choice of single-ventricle repair safer for the surgeon. However, the information about long-term prognosis after single-ventricle palliative repair including Fontan operation are still not completely promising, as the complications of “failing Fontan” are challenging, and may be increased in HLHS patients who underwent this operation (Feinstein et al. 2012). That makes it clearly essential to define the exact etiology, and genetic cause if possible, for this disease, which can help to decide the most suitable treatment plan for each patient, taking into consideration the great heterogeneity of this syndrome. The therapeutic decisions can be better guided by developing some criteria regarding the genetic etiology of the disease and its pathogenesis, not only according to the anatomic phenotype (Grossfeld et al. 2009).

EndMT was also described before to be involved in the formation of EBF tissue in HLHS (Xu et al. 2015a), and understanding the role of *KIAA0182* in this process may also have important therapeutic implications in HLHS or other diseases in general. Therefore, we started to study the importance of *KIAA0182* knockdown using *in vitro* model for EndMT. According to the findings in our study, *KIAA0182* gene (or *Gse1* in mouse cells) was activated by the induction of EndMT using TGF- β 1 in both HCAEC

and MCEC cells, suggesting a role for this gene in regulating this process. The upregulation of *SNAIL* and α -SMA upon *KIAA0182* inhibiting could mean that it is involved in suppressing EndMT in endothelial cells. These results can be further confirmed by treating the endothelial cells with TGF- β after downregulating our gene of interest *KIAA0182* and then comparing the response of these cells with that of control cells. This can be optimally performed using endothelial cells isolated from the mice with *Gse1* null allele, but establishing an endothelial cell line with stable *KIAA082*-shRNA expression by virus transduction represents another useful option. We tried to establish this method before but it is not completely successful until now due to the short survival of the transduced endothelial cells and their low response to TGF- β after virus infection, especially in the cells of late passages. One interesting point to be remembered in this context is that GSE1 protein may be a component in CoREST complex (Hakimi et al. 2003; Yang et al. 2011; Yokoyama et al. 2008). This complex is reported to be recruited by *SNAIL* to inhibit the expression of *E-cadherin* during EMT (Lin et al. 2010), and it may be also similarly involved in EndMT. Therefore, another important experiment to be performed is Co-immunoprecipitation (Co-IP), which can help to prove the involvement of GSE1 protein in this complex, possibly as a regulator or as another transcriptional factor that recruits CoREST complex to the targeted genes.

The results of the AAC operation in the heterozygous and wild type animals should have helped us to get better understanding for the role of *Gse1* in EndMT process and in cardiac fibrogenesis in general, but they were inconclusive. We could not detect any difference between the two genotypes regarding the heart weight, the systolic function, the systolic and diastolic dimensions, the hypertrophy or the fibrosis in the left ventricle. Several negative factors can be involved in that; first of all the decrease in the expression of *Gse1* in the hearts of the heterozygous animals may not be adequate to cause any difference regarding the previously mentioned parameters, especially using such small sample size. Second, the results of the AAC and sham operations were analyzed only two weeks after the surgery. It may be possible that waiting for 4 weeks could have enabled us to detect more fibrosis in the AAC groups, as it was done in previous studies (Jiang et al. 2007; Takahashi et al. 2007; Xu et al. 2015b; Zeisberg et al. 2007b), which could have made a potential difference between those two groups more prominent.

At last, the results of the operation were not consistent, as it was shown by the high variance in the fibrosis ratio, the ejection fraction and other echocardiographic parameters (Figure 3.28). This may be the consequence of the heterogeneity among the operated mice regarding the age and the pressure gradient across the constriction, especially in the heterozygous groups (Figure 3.27). Unfortunately, the high mortality of the operation in this study (35% compared with 16-20% in previous studies, which could be explained by the higher age of our mice) (Jiang et al. 2007; Takahashi et al. 2007) resulted in such small and heterogeneous cohort of animals. Therefore, that can be insufficient for investigating the effect of *Gse1* downregulation on cardiac fibrogenesis *in vivo*, which represents, in addition to the other previously mentioned problems, a clear limitation in this study that needs to be managed in the future.

Another important limitation is that the expression results of GSE1 protein as estimated by Western blot are not completely trustable. Both *Gse1* antibodies shown in previous section, among others tested, did not give a band with the expected size for GSE1 protein (around 140 kDa), which could be also explained by protein post-translational modifications (PTMs) or by the presence of different isoforms for GSE1 protein with unexpected sizes. In fact, all commercially purchased antibodies for *Gse1* were not used in any previous study, and all of them were tested for Western blot and for immunofluorescence staining using different protocols, but it was not possible to find a specific and reliable one, especially for staining. The low trapping efficiency in the mice did not enable us to use them as a validation tool for the specificity of the antibody, which is necessary before interpreting the expression results of GSE1 protein, and *Gse1* circular and linear RNAs to decide the major player in the observed phenotype.

Another critical task to be done in the future, in addition to the generation of the *Gse1* null allele on genomic level, is to determine the expression profile of *Gse1* in different tissues during several stages of embryonic development. This can be performed taking advantage of the presence of LacZ reporter in the trapping cassette, which will be expressed under the control of *Gse1* promoter in homozygous and heterozygous mice. That will definitely help us to know when and where this gene is active in the embryo, and what the consequences of trapping are, which potentially caused the embryonic lethality.

The last thing to be remembered is that we started this study because of a heterozygous mutation in *KIAA1082* gene found in a child with HLHS. Therefore, even if we are hopefully able to answer all the questions asked before regarding the function of this gene and the effect of its downregulation *in vivo* and *in vitro*, further experiments should be performed to find out the answers regarding any possible relevance of the mutation itself. This can be accomplished using endothelial cells and cardiac myocytes derived from iPSCs, which can be differentiated from fibroblasts obtained from that child. All these answers represent important pieces, and not necessarily interchangeable, in the puzzle to be solved in order to understand the role of this mysterious gene in HLHS, and in cardiac development and pathogenesis.

5. Summary

Objective: Hypoplastic left heart syndrome (HLHS) is one of the most severe congenital heart diseases. The genetic etiology for this syndrome was strongly suggested by several previous studies, and a lot of genes were suspected to be involved in its pathogenesis. *De novo* mutation in *KIAA0182* gene (*Gse1* in mouse) was found in this disease, but its exact role in HLHS or any other biological process is still to be determined. The aim of this research is to investigate the consequences of *KIAA0182* downregulation *in vivo* and *in vitro*. **Methods:** The 'Knockout-first' design was used in this study to produce supposedly null allele of *Gse1* ($Gse1^{tm1a(EUCOMM)Wtsi}$) in mouse model by disrupting its function on expression level. Human coronary arterial endothelial cells (HCAEC) were used to explore the role of *KIAA0182* in endothelial to mesenchymal transition (EndMT), which was found previously to play a role in the development of endocardial fibroelastosis in HLHS. Ascending aorta constriction (AAC) was performed in mice as a model for cardiac fibrosis *in vivo* to confirm the results observed *in vitro*. **Results:** After several quality control tests to prove the correct targeting of *Gse1* by the trapping cassette and to confirm its integrity, the genotyping results of 134 newborn mice from several matings between heterozygous ($Gse1^{tm1a/WT}$) male and female using two different protocols indicated the presence of embryonic lethality in the homozygous mice ($Gse1^{tm1a/tm1a}$). The expression level of *Gse1* in the homozygous embryos was significantly downregulated compared with the wild type, but less than 50 % downregulation effect was observed, suggesting that the mutated allele $Gse1^{tm1a}$ behaved as a hypomorphic allele. That was associated with upregulation of *Gse1* circular RNA, which might have the function of micro RNA (miRNA) sponge. *KIAA0182* was activated in HCAEC upon treatment with TGF- β 1 to induce EndMT, and *KIAA0182* knockdown was associated with upregulation of some EndMT markers. However, no difference was found between heterozygous and wild type animals regarding fibrosis, systolic function or hypertrophy after AAC operation. **Conclusions:** The homozygosity of *Gse1* trapping is associated with embryonic lethality, but further experiments are needed to prove the causal relation between *Gse1* gene and this phenotype, where the altered expression of *Gse1* circular RNA could also be involved. *KIAA0182* gene may play a critical role in EndMT as a negative regulator, but confirming the importance of these findings *in vivo* requires further investigations.

6. References

- Aisagbonhi O, Rai M, Ryzhov S, Atria N, Feoktistov I, Hatzopoulos AK (2011): Experimental myocardial infarction triggers canonical Wnt signaling and endothelial-to-mesenchymal transition. *Dis Model Mech* 4, 469-483.
- Amin A, Mahmoud-Ghoneim D (2011): Texture analysis of liver fibrosis microscopic images: a study on the effect of biomarkers. *Acta Biochim Biophys Sin (Shanghai)* 43, 193-203.
- Arciniegas E, Servin M, Argüello C, Mota M (1989): Development of the aorta in the chick embryo: structural and ultrastructural study. *Atherosclerosis* 76, 219-235.
- Arciniegas E, Neves CY, Carrillo LM, Zambrano EA, Ramírez R (2005): Endothelial-mesenchymal transition occurs during embryonic pulmonary artery development. *Endothelium* 12, 193-200.
- Armstrong EJ, Bischoff J (2004): Heart valve development: endothelial cell signaling and differentiation. *Circ Res* 95, 459-470.
- Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M, Evantal N, Memczak S, Rajewsky N, Kadener S (2014): circRNA biogenesis competes with pre-mRNA splicing. *Mol Cell* 56, 55-66.
- Bachmayr-Heyda A, Reiner AT, Auer K, Sukhbaatar N, Aust S, Bachleitner-Hofmann T, Mesteri I, Grunt TW, Zeillinger R, Pils D (2015): Correlation of circular RNA abundance with proliferation--exemplified with colorectal and ovarian cancer, idiopathic lung fibrosis, and normal human tissues. *Sci Rep* 5, 8057.
- Bamforth SD, Bragança J, Eloranta JJ, Murdoch JN, Marques FI, Kranc KR, Farza H, Henderson DJ, Hurst HC, Bhattacharya S (2001): Cardiac malformations, adrenal agenesis, neural crest defects and exencephaly in mice lacking *Cited2*, a new *Tfap2* co-activator. *Nat Genet* 29, 469-474.
- Barron DJ, Kilby MD, Davies B, Wright JG, Jones TJ, Brawn WJ (2009): Hypoplastic left heart syndrome. *Lancet* 374, 551-564.
- Beltrami CA, Finato N, Rocco M, Feruglio GA, Puricelli C, Cigola E, Quaini F, Sonnenblick EH, Olivetti G, Anversa P (1994): Structural basis of end-stage failure in ischemic cardiomyopathy in humans. *Circulation* 89, 151-163.

Blake DM, Copel JA, Kleinman CS (1991): Hypoplastic left heart syndrome: prenatal diagnosis, clinical profile, and management. *Am J Obstet Gynecol* 165, 529-534.

Boldt T, Andersson S, Eronen M (2002): Outcome of structural heart disease diagnosed in utero. *Scand Cardiovasc J* 36, 73-79.

Boneva RS, Botto LD, Moore CA, Yang Q, Correa A, Erickson JD (2001): Mortality associated with congenital heart defects in the United States: trends and racial disparities, 1979-1997. *Circulation* 103, 2376-2381.

Botto LD, Correa A, Erickson JD (2001): Racial and temporal variations in the prevalence of heart defects. *Pediatrics* 107, e32.

Boughman JA, Berg KA, Astemborski JA, Clark EB, McCarter RJ, Rubin JD, Ferencz C (1987): Familial risks of congenital heart defect assessed in a population-based epidemiologic study. *Am J Med Genet* 26, 839-849.

Brady AN, Shebata BM, Fernhoff PM (2006): X-linked foetal cardiomyopathy caused by a novel mutation in the TAZ gene. *Prenat Diagn* 26, 462–465.

Branda CS, Dymecki SM (2004): Talking about a revolution: The impact of site-specific recombinases on genetic analyses in mice. *Dev Cell* 6, 7-28.

Brown SD, Moore MW (2012): The International Mouse Phenotyping Consortium: past and future perspectives on mouse phenotyping. *Mamm Genome* 23, 632-640.

Bruneau BG, Logan M, Davis N, Levi T, Tabin CJ, Seidman JG, Seidman CE (1999): Chamber-specific cardiac expression of Tbx5 and heart defects in Holt-Oram syndrome. *Dev Biol* 211, 100-108.

Brunmeir R, Lagger S, Seiser C (2009): Histone deacetylase HDAC1/HDAC2-controlled embryonic development and cell differentiation. *Int J Dev Biol* 53, 275-289.

Casola S (2010): Mouse models for miRNA expression: the ROSA26 locus. *Methods Mol Biol* 667, 145-163.

Castelli M, Piobbico D, Bartoli D, Pieroni S, Brunacci C, Bellet MM, Chiacchiaretta M, Della Fazia MA, Servillo G (2014): Different functions of HOPS isoforms in the cell: HOPS shuttling isoform is determined by RIP cleavage system. *Cell Cycle* 13, 293-302.

Chen J, Ingham N, Clare S, Raisen C, Vancollie VE, Ismail O, McIntyre RE, Tsang SH, Mahajan VB, Dougan G et al. (2013): Mcph1-deficient mice reveal a role for MCPH1 in otitis media. *PLoS One* 8, e58156.

Chen X, Cai J, Zhou X, Chen L, Gong Y, Gao Z, Zhang H, Huang W, Zhou H (2015): Protective Effect of Spironolactone on Endothelial-to-Mesenchymal Transition in HUVECs via Notch Pathway. *Cell Physiol Biochem* 36, 191-200.

Cooley BC, Nevado J, Mellad J, Yang D, St Hilaire C, Negro A, Fang F, Chen G, San H, Walts AD et al. (2014): TGF- β signaling mediates endothelial-to-mesenchymal transition (EndMT) during vein graft remodeling. *Sci Transl Med* 6, 227ra34.

Cox H, Wilson DI (2007): Genetics of hypoplastic left heart syndrome. *Fetal and Maternal Medicine Review* 18, 103-119.

Cronk CE, Pelech AN, Malloy ME, McCarver DG (2004): Excess birth prevalence of hypoplastic left heart syndrome in eastern Wisconsin for birth cohorts 1997-1999. *Birth Defects Res A Clin Mol Teratol* 70, 114-120.

Dasgupta C, Martinez AM, Zuppan CW, Shah MM, Bailey LL, Fletcher WH (2001): Identification of connexin43 (alpha1) gap junction gene mutations in patients with hypoplastic left heart syndrome by denaturing gradient gel electrophoresis (DGGE). *Mutat Res* 479, 173-186.

Delcuve GP, Khan DH, Davie JR (2012): Roles of histone deacetylases in epigenetic regulation: emerging paradigms from studies with inhibitors. *Clin Epigenetics* 4, 5.

Dincsoy MY, Dincsoy HP, Kessler AD, Jackson MA, Sidbury JB Jr (1965): Generalized glycogenosis and associated endocardial fibroelastosis. Report of 3 cases with biochemical studies. *J Pediatr* 67, 728-740.

Eghtesady P, Brar A, Hall M (2011): Seasonality of hypoplastic left heart syndrome in the United States: a 10-year time-series analysis. *J Thorac Cardiovasc Surg* 141, 432-438.

Eisenberg LM, Markwald RR (1995): Molecular regulation of atrioventricular valvuloseptal morphogenesis. *Circ Res* 77, 1-6.

Elliott DA, Kirk EP, Yeoh T, Chandar S, McKenzie F, Taylor P, Grossfeld P, Fatkin D, Jones O, Hayes P et al. (2003): Cardiac homeobox gene NKX2-5 mutations and

congenital heart disease: associations with atrial septal defect and hypoplastic left heart syndrome. *J Am Coll Cardiol* 41, 2072-2076.

Elliott GC, Gurtu R, McCollum C, Newman WG, Wang T (2014): Foramen ovale closure is a process of endothelial-to-mesenchymal transition leading to fibrosis. *PLoS One* 9, e107175.

Emani SM, Bacha EA, McElhinney DB, Marx GR, Tworetzky W, Pigula FA, del Nido PJ (2009): Primary left ventricular rehabilitation is effective in maintaining two-ventricle physiology in the borderline left heart. *J Thorac Cardiovasc Surg* 138, 1276-1282.

Ewald CY, Landis JN, Porter Abate J, Murphy CT, Blackwell TK (2015): Dauer-independent insulin/IGF-1-signalling implicates collagen remodelling in longevity. *Nature* 519, 97-101.

Farley FW, Soriano P, Steffen LS, Dymecki SM (2000): Widespread recombinase expression using FLPeR (flipper) mice. *Genesis* 28, 106-110.

Feinstein JA, Benson DW, Dubin AM, Cohen MS, Maxey DM, Mahle WT, Pahl E, Villafañe J, Bhatt AB, Peng LF et al. (2012): Hypoplastic left heart syndrome: current considerations and expectations. *J Am Coll Cardiol* 59, S1-42.

Fixler DE, Nembhard WN, Salemi JL, Ethen MK, Canfield MA (2010): Mortality in first 5 years in infants with functional single ventricle born in Texas, 1996 to 2003. *Circulation* 121, 644-650.

Frid MG, Kale VA, Stenmark KR (2002): Mature vascular endothelium can give rise to smooth muscle cells via endothelial-mesenchymal transdifferentiation: in vitro analysis. *Circ Res* 90, 1189-1196.

Friehs I, Illigens B, Melnychenko I, Zhong-Hu T, Zeisberg E, Del Nido PJ (2013): An animal model of endocardial fibroelastosis. *J Surg Res* 182, 94-100.

Garg V, Muth AN, Ransom JF, Schluterman MK, Barnes R, King IN, Grossfeld PD, Srivastava D (2005): Mutations in NOTCH1 cause aortic valve disease. *Nature* 437, 270-274.

Ge Y, Sun Y, Chen J (2011): IGF-II is regulated by microRNA-125b in skeletal myogenesis. *J Cell Biol* 192, 69-81.

Gehrmann J, Krasemann T, Kehl HG, Vogt J (2001): Hypoplastic left-heart syndrome: the first description of the pathophysiology in 1851; translation of a publication by Dr. Bardeleben from Giessen, Germany. *Chest* 120, 1368-1371.

Giel-Moloney M, Krause DS, Chen G, Van Etten RA, Leiter AB (2007): Ubiquitous and uniform in vivo fluorescence in ROSA26-EGFP BAC transgenic mice. *Genesis* 45, 83-89.

Gordon BM, Rodriguez S, Lee M, Chang RK (2008): Decreasing number of deaths of infants with hypoplastic left heart syndrome. *J Pediatr* 153, 354-358.

Gordon JW, Ruddle FH (1981): Integration and stable germ line transmission of genes injected into mouse pronuclei. *Science* 214, 1244-1246.

Goumans MJ, Valdimarsdottir G, Itoh S, Lebrin F, Larsson J, Mummery C, Karlsson S, ten Dijke P (2003): Activin receptor-like kinase (ALK)1 is an antagonistic mediator of lateral TGFbeta/ALK5 signaling. *Mol Cell* 12, 817-828.

Grossfeld PD (1999): The genetics of hypoplastic left heart syndrome. *Cardiol Young* 9, 627-632.

Grossfeld PD, Mattina T, Lai Z, Favier R, Jones KL, Cotter F, Jones CA (2004): The 11q terminal deletion disorder: a prospective study of 110 cases. *Am J Med Genet A* 129A, 51-61.

Grossfeld PD, Ye M, Harvey R (2009): Hypoplastic left heart syndrome: new genetic insights. *J Am Coll Cardiol* 53, 1072-1074.

Guo JU, Agarwal V, Guo H, Bartel DP (2014): Expanded identification and characterization of mammalian circular RNAs. *Genome Biol* 15, 409.

Hakimi MA, Bochar DA, Chenoweth J, Lane WS, Mandel G, Shiekhattar R (2002): A core-BRAF35 complex containing histone deacetylase mediates repression of neuronal-specific genes. *Proc Natl Acad Sci U S A* 99, 7420-7425.

Hakimi MA, Dong Y, Lane WS, Speicher DW, Shiekhattar R (2003): A candidate X-linked mental retardation gene is a component of a new family of histone deacetylase-containing complexes. *J Biol Chem* 278, 7234-7239.

Hanauer D, Argilla M, Wallerstein (2002): Rubinstein-Taybi syndrome and hypoplastic left heart. *R Am J Med Genet* 112, 109-111.

Hansen TB, Wiklund ED, Bramsen JB, Villadsen SB, Statham AL, Clark SJ, Kjems J (2011): miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *EMBO J* 30, 4414–4422.

Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J. (2013): Natural RNA circles function as efficient microRNA sponges. *Nature* 495, 384–388.

Hanstein R, Negoro H, Patel NK, Charollais A, Meda P, Spray DC, Suadicani SO, Scemes E (2013): Promises and pitfalls of a Pannexin1 transgenic mouse line. *Front Pharmacol* 4, 61.

Hashimoto N, Phan SH, Imaizumi K, Matsuo M, Nakashima H, Kawabe T, Shimokata K, Hasegawa Y (2010): Endothelial-mesenchymal transition in bleomycin-induced pulmonary fibrosis. *Am J Respir Cell Mol Biol* 43, 161–172.

Hayakawa T, Nakayama J (2011): Physiological roles of class I HDAC complex and histone demethylase. *J Biomed Biotechnol* 2011, 129383.

Heemskerk JW, Harper MT, Cosemans JM, Poole AW (2011): Unravelling the different functions of protein kinase C isoforms in platelets. *FEBS Lett* 585, 1711-1716.

Hentze MW, Preiss T (2013): Circular RNAs: splicing's enigma variations. *EMBO J* 32, 923–925.

Ho CY, López B, Coelho-Filho OR, Lakdawala NK, Cirino AL, Jarolim P, Kwong R, González A, Colan SD, Seidman JG et al. (2010): Myocardial fibrosis as an early manifestation of hypertrophic cardiomyopathy. *N Engl J Med* 363, 552-563.

Hoffman JI, Kaplan S (2002): The incidence of congenital heart disease. *J Am Coll Cardiol* 39, 1890-1900.

Hove JR, Köster RW, Forouhar AS, Acevedo-Bolton G, Fraser SE, Gharib M (2003): Intracardiac fluid forces are an essential epigenetic factor for embryonic cardiogenesis. *Nature* 421, 172-177.

Hsu MT, Coca-Prados M (1979): Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells. *Nature* 280, 339-340.

Hutchins GM, Bannayan G (1971): Development of endocardial fibroelastosis following myocardial infarction. *Arch Pathol* 91, 113–118.

Iascone M, Ciccone R, Galletti L, Marchetti D, Seddio F, Lincasso AR, Pezzoli L, Vetro A, Barachetti D, Boni L et al. (2012): Identification of de novo mutations and rare variants in hypoplastic left heart syndrome. *Clin Genet* 81, 542-554.

Imai K, Kawai M, Tada M, Nagase T, Ohara O, Koga H (2005): Temporal change in mKIAA gene expression during the early stage of retinoic acid-induced neurite outgrowth. *Gene* 364, 114-122.

Jaenisch R (1976): Germ line integration and Mendelian transmission of the exogenous Moloney leukemia virus. *Proc Natl Acad Sci U S A* 73, 1260-1264.

Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, Marzluff WF, Sharpless NE (2013): Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* 19, 141-157.

Jiang Y, Reynolds C, Xiao C, Feng W, Zhou Z, Rodriguez W, Tyagi SC, Eaton JW, Saari JT, Kang YJJ (2007): Dietary copper supplementation reverses hypertrophic cardiomyopathy induced by chronic pressure overload in mice. *J Exp Med* 204, 657-666.

Jiang Y, Habibollah S, Tilgner K, Collin J, Barta T, Al-Aama JY, Tesarov L, Hussain R, Trafford AW, Kirkwood G et al. (2014): An induced pluripotent stem cell model of hypoplastic left heart syndrome (HLHS) reveals multiple expression and functional differences in HLHS-derived cardiac myocytes. *Stem Cells Transl Med* 3, 416-423.

Johnson P, Friedmann T (1990): Limited bidirectional activity of two housekeeping gene promoters: human HPRT and PGK. *Gene* 88, 207-213.

Kallen K (1999): Maternal smoking and congenital heart defects. *Eur J Epidemiol* 15, 731-737.

Kalluri R, Weinberg RA (2009): The basics of epithelial-mesenchymal transition. *J Clin Invest* 119, 1420-1428.

Kamisago M, Schmitt JP, McNamara D, Seidman C, Seidman JG (2006): Sarcomere protein gene mutations and inherited heart disease: a beta-cardiac myosin heavy

chain mutation causing endocardial fibroelastosis and heart failure. Novartis Found Symp 274, 176-189.

Karamlou T, Diggs BS, Ungerleider RM, Welke KF (2010): Evolution of treatment options and outcomes for hypoplastic left heart syndrome over an 18-year period. J Thorac Cardiovasc Surg 139, 119-125.

Kline IK, Miller AJ, Pick R, Katz LK (1964): The relationship between human endocardial fibroelastosis and obstruction of the cardiac lymphatics. Circulation 30, 728-735.

Kokudo T, Suzuki Y, Yoshimatsu Y, Yamazaki T, Watabe T, Miyazono K (2008): Snail is required for TGFbeta-induced endothelial-mesenchymal transition of embryonic stem cell-derived endothelial cells. J Cell Sci 121, 3317-3324.

Kouzarides T (2007): Chromatin modifications and their function. Cell 128, 693-705.

Krenning G, Zeisberg EM, Kalluri R (2010): The origin of fibroblasts and mechanism of cardiac fibrosis. J Cell Physiol 225, 631-637.

Lakowski B, Roelens I, Jacob S (2006): CoREST-like complexes regulate chromatin modification and neuronal gene expression. J Mol Neurosci 29, 227-239.

Landi D, Gemignani F, Pardini B, Naccarati A, Garritano S, Vodicka P, Vodickova L, Canzian F, Novotny J, Barale R et al. (2012): Identification of candidate genes carrying polymorphisms associated with the risk of colorectal cancer by analyzing the colorectal mutome and microRNAome. Cancer 118, 4670-4680.

Lasda E, Parker R (2014): Circular RNAs: diversity of form and function. RNA 20, 1829-1842.

Lee WJ, Park JH, Shin JU, Noh H, Lew DH, Yang WI, Yun CO, Lee KH, Lee JH (2015): Endothelial-to-Mesenchymal Transition induced by Wnt 3a in Keloid pathogenesis. Wound Repair Regen 23, 435-342.

Levet S, Ouarné M, Ciais D, Coutton C, Subileau M, Mallet C, Ricard N, Bidart M, Debillon T, Faravelli F et al. (2015): BMP9 and BMP10 are necessary for proper closure of the ductus arteriosus. Proc Natl Acad Sci U S A 112, E3207-3215.

Lin F, Wang N, Zhang TC (2012): The role of endothelial-mesenchymal transition in development and pathological process. Life 64, 717-723.

Lin Y, Wu Y, Li J, Dong C, Ye X, Chi YI, Evers BM, Zhou BP (2010): The SNAG domain of Snail1 functions as a molecular hook for recruiting lysine-specific demethylase 1. *EMBO J* 29, 1803-1816.

Loffredo CA, Chokkalingam A, Sill AM, Boughman JA, Clark EB, Scheel J, Brenner JI (2004): Prevalence of congenital cardiovascular malformations among relatives of infants with hypoplastic left heart, coarctation of the aorta, and d-transposition of the great arteries. *Am J Med Genet* 124A, 225-230.

Lurie PR (2010): Changing concepts of endocardial fibroelastosis. *Cardiol Young* 20, 115-123.

Lyons I, Parsons LM, Hartley L, Li R, Andrews JE, Robb L, Harvey RP (1995): Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeo box gene *Nkx2-5*. *Genes Dev* 9, 1654-1666.

Maguire S, Estabel J, Ingham N, Pearson S, Ryder E, Carragher DM, Walker N; Sanger MGP *Slc25a21* Project Team, Bussell J, Chan WI et al. (2014): Targeting of *Slc25a21* is associated with orofacial defects and otitis media due to disrupted expression of a neighbouring gene. *PLoS One* 9, e91807.

Mahler GJ, Farrar EJ, Butcher JT (2013): Inflammatory cytokines promote mesenchymal transformation in embryonic and adult valve endothelial cells. *Arterioscler Thromb Vasc Biol* 33, 121-130.

Mäkikallio K, McElhinney DB, Levine JC, Marx GR, Colan SD, Marshall AC, Lock JE, Marcus EN, Tworetzky W (2006): Fetal aortic valve stenosis and the evolution of hypoplastic left heart syndrome: patient selection for fetal intervention. *Circulation* 113, 1401-1405.

Markwald RR, Fitzharris TP, Smith WN (1975): Structural analysis of endocardial cytodifferentiation. *Dev Biol* 42, 160-180.

McDaneld TG (2009): MicroRNA: mechanism of gene regulation and application to livestock. *J Anim Sci* 87, E21-28.

McElhinney DB, Marshall AC, Wilkins-Haug LE, Brown DW, Benson CB, Silva V, Marx GR, Mizrahi-Arnaud A, Lock JE, Tworetzky W (2009): Predictors of technical success and postnatal biventricular outcome after in utero aortic valvuloplasty for

aortic stenosis with evolving hypoplastic left heart syndrome. *Circulation* 120, 1482-1490.

McElhinney DB, Vogel M, Benson CB, Marshall AC, Wilkins-Haug LE, Silva V, Tworetzky W (2010): Assessment of left ventricular endocardial fibroelastosis in fetuses with aortic stenosis and evolving hypoplastic left heart syndrome. *Am J Cardiol* 106, 1792-1797.

Medici D, Kalluri R (2012): Endothelial-mesenchymal transition and its contribution to the emergence of stem cell phenotype. *Semin Cancer Biol* 22, 379-384.

Medici D, Potenta S, Kalluri R (2011): Transforming growth factor- β 2 promotes Snail-mediated endothelial-mesenchymal transition through convergence of Smad-dependent and Smad-independent signaling. *Biochem J* 437, 515-520.

Meier ID, Bernreuther C, Tilling T, Neidhardt J, Wong YW, Schulze C, Streichert T, Schachner M (2010): Short DNA sequences inserted for gene targeting can accidentally interfere with off-target gene expression. *FASEB J* 24, 1714-1724.

Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M et al. (2013): Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 495, 333-338.

Minich LL, Tani LY, Hawkins JA, Shaddy RE (1997): Possibility of postnatal left ventricular growth in selected infants with non-apex-forming left ventricles. *Am Heart J* 133, 570-574.

Mohapatra B, Jimenez S, Lin JH, Bowles KR, Coveler KJ, Marx JG, Chrisco MA, Murphy RT, Lurie PR, Schwartz RJ et al. (2003): Mutations in the muscle LIM protein and α -actinin-2 genes in dilated cardiomyopathy and endocardial fibroelastosis. *Molec Genet Metab* 80, 207–215.

Moller JH, Lucas RV, Adams P, Anderson RC, Jorgens J, Edwards JE (1964): Endocardial fibroelastosis: a clinical and anatomic study of 47 patients with emphasis on its relationship to mitral insufficiency. *Circulation* 30, 759-82.

Moonen JR, Krenning G, Brinker MG, Koerts JA, van Luyn MJ, Harmsen MC (2010): Endothelial progenitor cells give rise to pro-angiogenic smooth muscle-like progeny. *Cardiovasc Res* 86, 506-515.

Morris SA, Ethen MK, Penny DJ, Canfield MA, Minard CG, Fixler DE, Nembhard WN (2014): Prenatal diagnosis, birth location, surgical center, and neonatal mortality in infants with hypoplastic left heart syndrome. *Circulation* 129, 285-292.

Morrow WR, Naftel D, Chinnock R, Canter C, Boucek M, Zales V, McGiffin DC, Kirklin JK (1997): Outcome of listing for heart transplantation in infants younger than six months: predictors of death and interval to transplantation. The Pediatric Heart Transplantation Study Group. *J Heart Lung Transplant* 16, 1255-1266.

Muller HJ (1927): Artificial transmutation of the gene. *Science* 66, 84-87.

Nagase T, Seki N, Ishikawa K, Tanaka A, Nomura N (1996): Prediction of the coding sequences of unidentified human genes. V. The coding sequences of 40 new genes (KIAA0161-KIAA0200) deduced by analysis of cDNA clones from human cell line KG-1. *DNA Res* 3, 17-24.

Nagase T, Koga H, Ohara O (2006): Kazusa mammalian cDNA resources: towards functional characterization of KIAA gene products. *Brief Funct Genomic Proteomic* 5, 4-7.

Naguit EC, Dexheimer H (1974): Adult endocardial fibroelastosis following electric shock. *IMJ Ill Med J* 145, 115–117.

Natowicz M, Chatten J, Clancy R, Conard K, Glauser T, Huff D, Lin A, Norwood W, Rorke LB, Uri A et al. (1988): Genetic disorders and major extracardiac anomalies associated with the hypoplastic left heart syndrome. *Pediatrics* 82, 698-706.

Newman PJ (1994): The role of PECAM-1 in vascular cell biology. *Ann N Y Acad Sci* 714, 165-174.

Ni J, Bowles NE, Kim YH, Demmler G, Kearney D, Bricker JT, Towbin JA (1997): Viral infection of the myocardium in endocardial fibroelastosis. Molecular evidence for the role of mumps virus as an etiologic agent. *Circulation* 95, 133-139.

Nield LE, Silverman ED, Taylor GP, Smallhorn JF, Mullen JB, Silverman NH, Finley JP, Law YM, Human DG, Seaward PG et al. (2002): Maternal anti-Ro and anti-La antibody-associated endocardial fibroelastosis. *Circulation* 105, 843-848.

Nigro JM, Cho KR, Fearon ER, Kern SE, Ruppert JM, Oliner JD, Kinzler KW, Vogelstein B (1991): Scrambled exons. *Cell* 64, 607-613.

Nijnik A, Clare S, Hale C, Raisen C, McIntyre RE, Yusa K, Everitt AR, Mottram L, Podrini C, Lucas M et al. (2012): The critical role of histone H2A-deubiquitinase Mym1 in hematopoiesis and lymphocyte differentiation. *Blood* 119, 1370-1379.

Nosedá M, McLean G, Niessen K, Chang L, Pollet I, Montpetit R, Shahidi R, Dorovini-Zis K, Li L, Beckstead B et al. (2004): Notch activation results in phenotypic and functional changes consistent with endothelial-to-mesenchymal transformation. *Circ Res* 94, 910-917.

Nyabi O, Naessens M, Haigh K, Gembarska A, Goossens S, Maetens M, De Clercq S, Drogat B, Haenebalcke L, Bartunkova S et al. (2009): Efficient mouse transgenesis using Gateway-compatible ROSA26 locus targeting vectors and F1 hybrid ES cells. *Nucleic Acids Res* 37, e55.

Peinado H, Portillo F, Cano A (2004): Transcriptional regulation of cadherins during development and carcinogenesis. *Int J Dev Biol* 48, 365-375.

Perez MH, Boulos T, Stucki P, Cotting J, Osterheld MC, Di Bernardo S (2009): Placental immaturity, endocardial fibroelastosis and fetal hypoxia. *Fetal Diagn Ther* 26, 107-110.

Pham CT, MacIvor DM, Hug BA, Heusel JW, Ley TJ (1996): Long-range disruption of gene expression by a selectable marker cassette. *Proc Natl Acad Sci U S A* 93, 13090-13095.

Piera-Velazquez S, Li Z, Jimenez SA (2011): Role of endothelial-mesenchymal transition (EndoMT) in the pathogenesis of fibrotic disorders. *Am J Pathol* 179, 1074-1080.

Prsa M, Holly CD, Carnevale FA, Justino H, Rohlicek CV (2010): Attitudes and practices of cardiologists and surgeons who manage HLHS. *Pediatrics* 125, e625-630.

Quéméner-Redon S, Bénech C, Audebert-Bellanger S, Friocourt G, Planes M, Parent P, Férec C (2013): A small de novo 16q24.1 duplication in a woman with severe clinical features. *Eur J Med Genet* 56, 211-215.

Raboisson MJ, Fouron JC, Sonesson SE, Nyman M, Proulx F, Gamache S (2005): Foetal Doppler echocardiographic diagnosis and successful steroid therapy of

Luciani-Wenckebach phenomenon and endocardial fibroelastosis related to maternal anti-Ro and anti-La antibodies. *J Am Soc Echocardiogr* 18, 375–380.

Rainger J, van Beusekom E, Ramsay JK, McKie L, Al-Gazali L, Pallotta R, Saponari A, Branney P, Fisher M, Morrison H et al. (2011): Loss of the BMP antagonist, SMOC-1, causes Ophthalmo-acromelic (Waardenburg Anophthalmia) syndrome in humans and mice. *PLoS Genet* 7, e1002114.

Reamon-Buettner SM, Ciribilli Y, Inga A, Borlak J (2008): A loss-of-function mutation in the binding domain of HAND1 predicts hypoplasia of the human hearts. *Hum Mol Genet* 17, 1397-1405.

Reller MD, Strickland MJ, Riehle-Colarusso T, Mahle WT, Correa A (2008): Prevalence of congenital heart defects in metropolitan Atlanta, 1998-2005. *J Pediatr* 153, 807-813.

Ren SY, Angrand PO, Rijli FM (2002): Targeted insertion results in a rhombomere 2-specific Hoxa2 knockdown and ectopic activation of Hoxa1 expression. *Dev Dyn* 225, 305-315.

Reynolds AM, Holmes MD, Danilov SM, Reynolds PN (2012): Targeted gene delivery of BMPR2 attenuates pulmonary hypertension. *Eur Respir J* 39, 329-343.

Rieder F, Kessler SP, West GA, Bhilocha S, de la Motte C, Sadler TM, Gopalan B, Stylianou E, Fiocchi C (2011): Inflammation-induced endothelial-to-mesenchymal transition: a novel mechanism of intestinal fibrosis. *Am J Pathol* 179, 2660-2673.

Robert ML, Lopez T, Crolla J, Huang S, Owen C, Burvill-Holmes L, Stumper O, Turnpenny PD (2007): Alagille syndrome with deletion 20p12.2-p12.3 and hypoplastic left heart. *Clin Dysmorphol* 16, 241-246.

Roy O, Leclerc VB, Bourget JM, Thériault M, Proulx S (2015): Understanding the process of corneal endothelial morphological change in vitro. *Invest Ophthalmol Vis Sci* 56, 1228-1237.

Russell LB, Hunsicker PR, Cacheiro NL, Bangham JW, Russell WL, Shelby MD (1989): Chlorambucil effectively induces deletion mutations in mouse germ cells. *Proc Natl Acad Sci U S A* 86, 3704-3708.

Russell WL, Kelly EM, Hunsicker PR, Bangham JW, Maddux SC, Phipps EL (1979): Specific-locus test shows ethylnitrosourea to be the most potent mutagen in the mouse. *Proc Natl Acad Sci U S A* 76, 5818-5819.

Rustico MA, Benettoni A, Bussani R, Maieron A, Mandruzzato G (1995): Early fetal endocardial fibroelastosis and critical aortic stenosis: a case report. *Ultrasound Obstet Gynecol* 5, 202-205.

Ryder E, Gleeson D, Sethi D, Vyas S, Miklejewska E, Dalvi P, Habib B, Cook R, Hardy M, Jhaveri K et al. (2013): Molecular characterization of mutant mouse strains generated from the EUCOMM/KOMP-CSD ES cell resource. *Mamm Genome* 24, 286-294.

Ryder E, Doe B, Gleeson D, Houghton R, Dalvi P, Grau E, Habib B, Miklejewska E, Newman S, Sethi D et al. (2014): Rapid conversion of EUCOMM/KOMP-CSD alleles in mouse embryos using a cell-permeable Cre recombinase. *Transgenic Res* 23, 177-185.

Salzman J, Gawad C, Wang PL, Lacayo N, Brown PO (2012): Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS One* 7, e30733.

Samánek M, Slavík Z, Krejčíř M (1991): Seasonal differences in the incidence of congenital heart defects. *Czech Med* 14, 146-155.

Scacheri PC, Crabtree JS, Novotny EA, Garrett-Beal L, Chen A, Edgemon KA, Marx SJ, Spiegel AM, Chandrasekharappa SC, Collins FS (2001): Bidirectional transcriptional activity of PGK-neomycin and unexpected embryonic lethality in heterozygote chimeric knockout mice. *Genesis* 30, 259-263.

Schaft J, Ashery-Padan R, van der Hoeven F, Gruss P, Stewart AF (2001): Efficient FLP recombination in mouse ES cells and oocytes. *Genesis* 31, 6-10.

Sedmera D, Hu N, Weiss KM, Keller BB, Denslow S, Thompson RP (2002): Cellular changes in experimental left heart hypoplasia. *Anat Rec* 267, 137-145.

Seki A, Patel S, Ashraf S, Perens G, Fishbein MC (2013): Primary endocardial fibroelastosis: an underappreciated cause of cardiomyopathy in children. *Cardiovasc Pathol* 22, 345-350.

Sellers FJ, Keith JD, Manning JA (1964): The diagnosis of primary endocardial fibroelastosis. *Circulation* 29, 49-59.

Shaw GM, Carmichael SL, Nelson V (2002): Congenital malformations in offspring of Vietnamese women in California, 1985-97. *Teratology* 65, 121-124.

Shi YJ, Matson C, Lan F, Iwase S, Baba T, Shi Y (2005): Regulation of LSD1 histone demethylase activity by its associated factors. *Mol Cell* 19, 857-864.

Skarnes WC, Rosen B, West AP, Koutsourakis M, Bushell W, Iyer V, Mujica AO, Thomas M, Harrow J, Cox T et al. (2011): A conditional knockout resource for the genome-wide study of mouse gene function. *Nature* 474, 337-342.

Spence SE, Gilbert DJ, Swing DA, Copeland NG, Jenkins NA (1989): Spontaneous germ line virus infection and retroviral insertional mutagenesis in eighteen transgenic Srev lines of mice. *Mol Cell Biol* 9, 177-184.

Stanford WL, Cohn JB, Cordes SP (2001): Gene-trap mutagenesis: past, present and beyond. *Nat Rev Genet* 2, 756-768.

Starke S, Jost I, Rossbach O, Schneider T, Schreiner S, Hung LH, Bindereif A (2015): Exon circularization requires canonical splice signals. *Cell Rep* 10, 103-111.

Steger CM, Antretter H, Moser PL (2012): Endocardial fibroelastosis of the heart. *Lancet* 379, 932.

Storch TG, Mannick EE (1992): Epidemiology of congenital heart disease in Louisiana: an association between race and sex and the prevalence of specific cardiac malformations. *Teratology* 46, 271-276.

Suzuki H, Tsukahara T (2014): A view of pre-mRNA splicing from RNase R resistant RNAs. *Int J Mol Sci* 15, 9331-9342.

Takahashi R, Asai T, Murakami H, Murakami R, Tsuzuki M, Numaguchi Y, Matsui H, Murohara T, Okumura K (2007): Pressure overload-induced cardiomyopathy in heterozygous carrier mice of carnitine transporter gene mutation. *Hypertension* 50, 497-502.

Tampe B, Tampe D, Müller CA, Sugimoto H, LeBleu V, Xu X, Müller GA, Zeisberg EM, Kalluri R, Zeisberg M (2014): Tet3-mediated hydroxymethylation of

epigenetically silenced genes contributes to bone morphogenic protein 7-induced reversal of kidney fibrosis. *J Am Soc Nephrol* 25, 905-912.

Tang R, Gao M, Wu M, Liu H, Zhang X, Liu B (2012): High glucose mediates endothelial-to-chondrocyte transition in human aortic endothelial cells. *Cardiovasc Diabetol* 11, 113.

Tang RN, Lv LL, Zhang JD, Dai HY, Li Q, Zheng M, Ni J, Ma KL, Liu BC (2013): Effects of angiotensin II receptor blocker on myocardial endothelial-to-mesenchymal transition in diabetic rats. *Int J Cardiol* 162, 92-99.

Tarnavski O, McMullen JR, Schinke M, Nie Q, Kong S, Izumo S (2004): Mouse cardiac surgery: comprehensive techniques for the generation of mouse models of human diseases and their application for genomic studies. *Physiol Genomics* 16, 349-360.

Tchervenkov CI, Jacobs JP, Weinberg PM, Aiello VD, Béland MJ, Colan SD, Elliott MJ, Franklin RC, Gaynor JW, Krogmann ON et al. (2006): The nomenclature, definition and classification of hypoplastic left heart syndrome. *Cardiol Young* 16, 339-368.

Teekakirikul P, Eminaga S, Toka O, Alcalai R, Wang L, Wakimoto H, Naylor M, Konno T, Gorham JM, Wolf CM et al. (2010): Cardiac fibrosis in mice with hypertrophic cardiomyopathy is mediated by non-myocyte proliferation and requires Tgf- β . *J Clin Invest* 120, 3520-3529.

Testa G, Schaft J, van der Hoeven F, Glaser S, Anastassiadis K, Zhang Y, Hermann T, Stremmel W, Stewart AF (2004): A reliable lacZ expression reporter cassette for multipurpose, knockout-first alleles. *Genesis* 38, 151-158.

Tian YC, Phillips AO (2002): Interaction between the transforming growth factor-beta type II receptor/Smad pathway and beta-catenin during transforming growth factor-beta1-mediated adherens junction disassembly. *Am J Pathol* 160, 1619-1628.

Tikkanen J, Heinonen OP (1994): Risk factors for hypoplastic left heart syndrome. *Teratology* 50, 112-117.

Timmerman LA, Grego-Bessa J, Raya A, Bertrán E, Pérez-Pomares JM, Díez J, Aranda S, Palomo S, McCormick F, Izpisúa-Belmonte JC et al. (2004): Notch

promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation. *Genes Dev* 18, 99-115.

Tripp ME, Katcher ML, Peters HA, Gilbert EF, Arya S, Hodach RJ, Shug AL (1981): Systemic carnitine deficiency presenting as familial endocardial fibroelastosis: a treatable cardiomyopathy. *N Engl J Med* 305, 385-390.

Tworetzky W, Wilkins-Haug L, Jennings RW, van der Velde ME, Marshall AC, Marx GR, Colan SD, Benson CB, Lock JE, Perry SB (2004): Balloon dilation of severe aortic stenosis in the fetus: potential for prevention of hypoplastic left heart syndrome: candidate selection, technique, and results of successful intervention. *Circulation* 110, 2125-2131.

Tworetzky W, del Nido PJ, Powell AJ, Marshall AC, Lock JE, Geva T (2005): Usefulness of magnetic resonance imaging of left ventricular endocardial fibroelastosis in infants after foetal intervention for aortic valve stenosis. *Am J Cardiol* 96, 1568-1570.

Ursell PC, Neill CA, Anderson RH, Ho SY, Becker AE, Gerlis LM (1984): Endocardial fibroelastosis and hypoplasia of the left ventricle in neonates without significant aortic stenosis. *Br Heart J* 51, 492-497.

Valinezhad Orang A, Safaralizadeh R, Kazemzadeh-Bavili M (2014): Mechanisms of miRNA-Mediated Gene Regulation from Common Downregulation to mRNA-Specific Upregulation. *Int J Genomics* 2014, 970607.

Van Meeteren LA, ten Dijke P (2012): Regulation of endothelial cell plasticity by TGF- β . *Cell Tissue Res* 347, 177-186.

Van Rooij E (2011): The art of microRNA research. *Circ Res* 108, 219-234.

Vlahos AP, Lock JE, McElhinney DB, van der Velde ME (2004): Hypoplastic left heart syndrome with intact or highly restrictive atrial septum: outcome after neonatal transcatheter atrial septostomy. *Circulation* 109, 2326-2330.

Vogt M, Motz WH, Schwartzkopf B, Strauer BE (1993): Pathophysiology and clinical aspects of hypertensive hypertrophy. *Eur Heart J* 14 Suppl D, 2-7.

Wagner EF, Stewart TA, Mintz B (1981): The human beta-globin gene and a functional viral thymidine kinase gene in developing mice. *Proc Natl Acad Sci U S A* 78, 5016-5120.

Warner JR (1999): The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci* 24, 437-440.

Weinberg T, Himelfarb AJ (1943): Endocardial fibroelastosis (so-called fetal endocarditis): a report of two cases occurring in siblings. *Bull Johns Hopkins Hosp* 72, 299-306.

Welcker JE, Hernandez-Miranda LR, Paul FE, Jia S, Ivanov A, Selbach M, Birchmeier C (2013): *Insm1* controls development of pituitary endocrine cells and requires a SNAG domain for function and for recruitment of histone-modifying factors. *Development* 140, 4947-4958.

Wheway G, Abdelhamed Z, Natarajan S, Toomes C, Inglehearn C, Johnson CA (2013): Aberrant Wnt signalling and cellular over-proliferation in a novel mouse model of Meckel-Gruber syndrome. *Dev Biol* 377, 55-66.

Widyantoro B, Emoto N, Nakayama K, Anggrahini DW, Adiarto S, Iwasa N, Yagi K, Miyagawa K, Rikitake Y, Suzuki T et al. (2010): Endothelial cell-derived endothelin-1 promotes cardiac fibrosis in diabetic hearts through stimulation of endothelial-to-mesenchymal transition. *Circulation* 121, 2407-2418.

Wilson PD, Loffredo CA, Correa-Villaseñor A, Ferencz C (1998): Attributable fraction for cardiac malformations. *Am J Epidemiol* 148, 414-423.

Wong SH, Hamel L, Chevalier S, Philip A (2000): Endoglin expression on human microvascular endothelial cells association with betaglycan and formation of higher order complexes with TGF-beta signalling receptors. *Eur J Biochem* 267, 5550-5560.

Wu Y, Wang C, Sun H, LeRoith D, Yakar S (2009): High-efficient FLPo deleter mice in C57BL/6J background. *PLoS One* 4, e8054.

Wynder C, Hakimi MA, Epstein JA, Shilatifard A, Shiekhattar R (2005): Recruitment of MLL by HMG-domain protein iBRAF promotes neural differentiation. *Nat Cell Biol* 7, 1113-1117.

Xu X, Friehs I, Zhong Hu T, Melnychenko I, Tampe B, Alnour F, Iacone M, Kalluri R, Zeisberg M, Del Nido PJ et al. (2015a): Endocardial fibroelastosis is caused by aberrant endothelial to mesenchymal transition. *Circ Res* 116, 857-866.

Xu X, Tan X, Tampe B, Nyamsuren G, Liu X, Maier LS, Sossalla S, Kalluri R, Zeisberg M, Hasenfuss G et al. (2015b): Epigenetic balance of aberrant Rasal1 promoter methylation and hydroxymethylation regulates cardiac fibrosis. *Cardiovasc Res* 105, 279-291.

Yang P, Wang Y, Chen J, Li H, Kang L, Zhang Y, Chen S, Zhu B, Gao S (2011): RCOR2 is a subunit of the LSD1 complex that regulates ESC property and substitutes for SOX2 in reprogramming somatic cells to pluripotency. *Stem Cells* 29, 791-801.

Yokoyama A, Takezawa S, Schüle R, Kitagawa H, Kato S (2008): Transrepressive function of TLX requires the histone demethylase LSD1. *Mol Cell Biol* 28, 3995-4003.

Yoshimatsu Y, Watabe T (2011): Roles of TGF- β signals in endothelial-mesenchymal transition during cardiac fibrosis. *Int J Inflam* 2011, 724080.

You A, Tong JK, Grozinger CM, Schreiber SL (2001): CoREST is an integral component of the CoREST- human histone deacetylase complex. *Proc Natl Acad Sci U S A* 98, 1454-1458.

Yu J, Li Y, Ishizuka T, Guenther MG, Lazar MA (2003): A SANT motif in the SMRT corepressor interprets the histone code and promotes histone deacetylation. *EMBO J* 22, 3403-3410.

Yu W, Liu Z, An S, Zhao J, Xiao L, Gou Y, Lin Y, Wang J (2014): The endothelial-mesenchymal transition (EndMT) and tissue regeneration. *Curr Stem Cell Res Ther* 9, 196-204.

Zeisberg EM, Kalluri R (2010): Origins of cardiac fibroblasts. *Circ Res* 107, 1304-1312.

Zeisberg EM, Potenta S, Xie L, Zeisberg M, Kalluri R (2007a): Discovery of endothelial to mesenchymal transition as a source for carcinoma-associated fibroblasts. *Cancer Res* 67, 10123-10128.

Zeisberg EM, Tarnavski O, Zeisberg M, Dorfman AL, McMullen JR, Gustafsson E, Chandraker A, Yuan X, Pu WT, Roberts AB et al. (2007b): Endothelial-to-mesenchymal transition contributes to cardiac fibrosis. *Nat Med* 13, 952-961.

Zeisberg EM, Potenta SE, Sugimoto H, Zeisberg M, Kalluri R (2008): Fibroblasts in kidney fibrosis emerge via endothelial-to-mesenchymal transition. *J Am Soc Nephrol* 19, 2282-2287.

Zeisberg M, Neilson EG (2009): Biomarkers for epithelial-mesenchymal transitions. *J Clin Invest* 119, 1429-1437.

Zhang XO, Wang HB, Zhang Y, Lu X, Chen LL, Yang L (2014): Complementary sequence-mediated exon circularization. *Cell* 159,134-147.

Zhao F, Bosserhoff AK, Buettner R, Moser M (2011): A heart-hand syndrome gene: *Tfap2b* plays a critical role in the development and remodeling of mouse ductus arteriosus and limb patterning. *PLoS One* 6, e22908.

Zhu P, Huang L, Ge X, Yan F, Wu R, Ao Q (2006): Transdifferentiation of pulmonary arteriolar endothelial cells into smooth muscle-like cells regulated by myocardin involved in hypoxia-induced pulmonary vascular remodelling. *Int J Exp Pathol* 87, 463-474.

Acknowledgment

At the end of my journey I would like to express, my deepest appreciation, gratitude and my sincere thanks to my supervisor Prof. Dr. Elisabeth Zeisberg. Working with Prof. Zeisberg was an exceptional opportunity for me to be engaged in such an advanced research field. I would like also to sincerely thank Prof. Dr. Gerd Hasenfuß for giving me this great chance to do this study in the department of cardiology in the medical school of Goettingen university.

I have learned a lot from Prof. Zeisberg, to have genuine perspectives, to be inquisitive, never to take my results for granted, but to think in a critical way to understand the limitations of my work and to plan logically to get the right answers after asking the right questions. I have learned from her to work in an organized way, to be precise and to interpret my results from an objective unbiased point of view to get reliable data, even when the experiments came to unexpected and sometimes seemingly contradictory results. Professor Zeisberg was my mentor not only in my project but also on personal level, she was a friend and working under her supervision was a great help for me to adjust in a new challenging field, after years of clinical practice.

I also would like to thank Dr. Xingbo Xu, who has led me through my work, step by step to get familiar with the new techniques, and showed me how to deal with possible obstacles. He taught me how to reevaluate my assumptions, to be updated and to have a comprehensive theoretical background about my project. His great experience and his wonderful abilities as a teacher helped me so much to adapt into the magnificent world of molecular biology.

I would like to express my thanks to the great team in Zeisberg's lab and all my colleagues in the cardiology department, who were very supportive during the whole project. The encouraging environment, which you created, made me feel that we became as one family after more than two years of working together.

Curriculum vitae

I, Fouzi Alnour, was born in Damas-Suburb/ Syria on 24.05.1983 as the only child of my parents, Tawfik Alnour and Sahar Barakeh.

During the period between 1997 and 2000, I studied in Subhi Oadeh high school, Alnabek, and in the year 2000 I got my high school certificate with the grade 254/260. Afterwards, I was accepted in the faculty of medicine of Damascus University, Damascus / Syria, and I received the license of Doctor (Bachelor degree) in human medicine in 2006 with the grade "Excellent" (85.41%) after 12 semesters including a practical phase of 3 years.

In December 2006 I started my residency in cardiology at the hospitals of Damascus University. The residency lasted 5 years and consisted of two phases; two years in the department of internal medicine and three years in the cardiology department. In the year 2012, I got the certificate of post-graduate studies (Master degree) in cardiology with the grade "Excellent" (86.67%). The title of my master thesis was as follows: "Comparison of the predictive value of TIMI and PREDICT risk scores for early mortality of non-high risk patients with ST-elevation acute myocardial infarction".

Between December 2011 and July 2013 I worked as a cardiologist in Alkalmoon University hospital, where I also participated in the training of medical students in addition to my responsibilities in the clinic, the department of cardiology and the intensive care unit.

Since August 2013 I have started working on my medical thesis under the supervision of Prof. Dr. Elisabeth Zeisberg in the department of cardiology at the University of Goettingen, Germany, where I got a scholarship from the German center for cardiovascular research.

Publications

Xu X, Friehs I, Zhong Hu T, Melnychenko I, Tampe B, Alnour F, Iascone M, Kalluri R, Zeisberg M, Del Nido PJ et al. (2015a): Endocardial fibroelastosis is caused by aberrant endothelial to mesenchymal transition. *Circ Res* 116, 857-866.