

Modifying the Porcine Genome -  
Genome engineering towards placement and editing of  
coding and non-coding genes

David Pius Preisinger

Vollständiger Abdruck der von der TUM School of Life Sciences der Technischen Universität  
München zur Erlangung eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Harald Luksch

Prüfende der Dissertation:

1. Prof. Angelika Schnieke, Ph.D.
2. Prof. Dr. Benjamin Schusser

Die Dissertation wurde am 16.10.2023 bei der Technischen Universität München eingereicht  
und durch die TUM School of Life Sciences am 06.03.2024 angenommen.

# Table of contents

Table of contents .....	1
Abstract .....	4
Zusammenfassung .....	5
1 Introduction.....	6
1.1 Generation of genetically modified pigs .....	6
1.1.1 Microinjection.....	7
1.1.2 Generation of genetically modified pigs via SCNT .....	8
1.2 Tools to alter the genome.....	8
1.2.1 Knockouts.....	9
1.2.2 dCas9 variants for transcriptional control.....	11
1.2.3 Transgene addition .....	13
1.3 <i>In vivo</i> genome editing.....	16
1.3.1 <i>In vivo</i> genome editing by AAVs.....	16
1.3.2 <i>In vivo</i> genome editing using SpCas9-expressing pigs .....	17
1.4 Long non-coding RNAs .....	17
1.4.1 The regulation of maternally expressed gene 3 (MEG3) .....	18
1.4.2 MEG3 in cardiac fibrosis and hypertrophic cardiomyopathy.....	21
1.5 Objectives.....	23
2 Material and methods.....	24
2.1 Materials.....	24
2.1.1 Oligonucleotides .....	24
2.1.2 DNA vectors.....	27
2.1.3 Antibodies .....	29
2.1.4 Chemicals.....	29
2.1.5 Cells.....	31
2.1.6 Buffers.....	32
2.1.7 Tissue culture reagents .....	33
2.1.8 Tissue culture media .....	34

2.1.9	Kits.....	35
2.1.10	Consumables.....	35
2.1.11	Enzymes and enzyme buffers.....	36
2.1.12	Laboratory equipment.....	36
2.1.13	Software.....	37
2.2	Methods.....	39
2.2.1	Molecular biology.....	39
2.2.2	Microbiological methods for vector amplification.....	48
2.2.3	Cell culture.....	49
3	Results.....	54
3.1	Insertion of AsCas12a at the porcine <i>ROSA26</i> locus by CRISPlace and HDR.....	54
3.1.1	Homology-directed targeting for the insertion of AsCas12a at the <i>ROSA26</i> locus.....	55
3.1.2	CRISPlace for the insertion of AsCas12a at the <i>ROSA26</i> locus.....	56
3.2	Characterization of the AsCas12a-expressing pig line.....	58
3.2.1	Analysis of the transgene insertion site of founder animal #2026.....	58
3.2.2	AsCas12a F1 generation produces transgene mRNA and protein.....	59
3.2.3	AsCas12a shows low efficiency.....	62
3.2.4	Targeting $\beta$ 2M in porcine AsCas12a cells is inefficient compared to Cas9.....	66
3.2.5	Nuclear and cytoplasmic distribution of AsCas12a.....	66
3.3	Manipulation of the lncRNA <i>MEG3</i> via SpCas9 variants.....	67
3.3.1	Characterization of the porcine <i>MEG3</i> gene.....	68
3.3.2	Manipulation of <i>MEG3</i> via SpCas9 variants.....	73
3.3.3	Bulk RNA-Seq shows effect of <i>MEG3</i> KO D-1 and E-2 on fibrotic signaling.....	82
4	Discussion.....	84
4.1	CRISPlace is an efficient strategy for large transgene insertions.....	84
4.2	AsCas12a requires engineering to achieve Cas9-comparable performance.....	87
4.2.1	Neither pol III promoter nor backbone integration causes AsCas12a inefficiency.....	87
4.2.2	AsCas12a requires engineering to achieve high editing rates reliably.....	88
4.3	<i>MMP2</i> downregulation could be hindered by large T antigen.....	89

4.3.1	<i>MEG3</i> knockout reveals important regulatory regions .....	90
4.3.2	RNA-Seq does not show downregulation of <i>MMP2</i> via <i>MEG3</i> KO due to large T antigen immortalization of cardiac fibroblasts .....	92
5	Concluding remarks and outlook .....	94
6	Abbreviations .....	95
7	List of tables .....	97
8	List of figures .....	98
9	Supplementary .....	99
10	Acknowledgements.....	105
11	References .....	106

## Abstract

Transgenic pigs can provide valuable translational models for biomedicine. However, the generation of genetically modified pigs takes years and requires comprehensive resources. Here, I present CRISPlace, a non-homologous end joining (NHEJ) and CRISPR/Cas9-based transgene insertion strategy, which more efficiently inserts transgenes in porcine somatic cells than homology-mediated targeting. I successfully used CRISPlace to insert the sizable AsCas12a gene into the safe harbor locus *ROSA26* and prove its efficiency for other genes and another locus.

Cas-expressing transgenic pigs circumvent the production of numerous different knockout pig lines, as only sgRNAs need to be delivered to achieve knockouts *in vivo*. In contrast to Cas9, AsCas12a enables expression of sgRNA arrays from tissue specific pol II promoters and therefore knockouts in an organ or even cell-specific manner. Hence, an AsCas12a-expressing pig was generated for the first time. RT-PCR, western blot analysis, and immunohistochemistry proved expression of AsCas12a mRNA and protein in nearly all tissues and cells. Functionality was shown by transfecting primary cells, derived from the AsCas12a pig, with sgRNAs against the porcine *B2M* and *GGTA1* gene. The gene editing efficiency was generally lower compared to Cas9 and requires further optimization.

As about 90 % of the genome is transcribed as various types of non-coding RNAs, not all diseases or their treatment can be modelled through genetic modification of protein-coding genes. The long non-coding RNA (lncRNA) maternally expressed gene 3 (*Meg3*) is located at the imprinted locus *Dio3-Dlk1* and promotes cardiac fibrosis in mice. To manipulate porcine *MEG3*, its intron-exon structure was determined, and the first localization of the regulatory elements *MEG3* DMR and porcine *MEG3* promoter was carried out. Excision of parts of the promoter/DMR and the exon 2 using CRISPR/Cas9, led to loss of *MEG3* expression in immortalized porcine cardiac fibroblasts. Furthermore, using CRISPRa and CRISPRi, I successfully overexpressed or knocked down *MEG3*, which also affected the expression of the lncRNA *MEG8*, located at the same imprinted locus. It did not alter transcription of co-localized protein coding genes. Bulk RNA-Seq of *MEG3* edited cells showed downregulation of the main pro-fibrotic pathway *TGF- $\beta$* , the muscle contraction pathway, and the activated fibroblast marker *ACTA2*. Hence, knocking out *MEG3* could also prevent the progression of cardiac fibrosis in pigs.

## Zusammenfassung

Transgene Schweine sind ein wertvolles translationales Modell. Die Erzeugung von genetisch veränderten Schweinen dauert jedoch Jahre und erfordert umfangreiche Ressourcen. Hier stelle ich CRISPlace vor, eine non-homologous end joining (NHEJ) und CRISPR/Cas9-basierte Transgen-Insertionsstrategie, die Transgene effizienter in somatische Schweinezellen einfügt als homology-mediated Gene Targeting. Ich habe CRISPlace verwendet, um das AsCas12a-Gen in den Safe-Harbor-Lokus *ROSA26* einzufügen. Zudem habe ich die Effizienz von CRISPlace bei der Bearbeitung anderer Gene und an einem anderen genetischen Locus nachgewiesen.

Mit Cas9-exprimierenden transgenen Schweinen lässt sich die Herstellung zahlreicher einzelner Knockout-Schweinelinien vermeiden. Es genügt sgRNAs zu applizieren, um Knockouts *in vivo* zu erreichen. Im Gegensatz zu Cas9 ermöglicht Cas12a sogar gewebespezifische Knockouts und die Expression von sgRNA-Arrays von Pol-II-Promotoren. Daher wurde zum ersten Mal ein AsCas12a-exprimierendes Schwein erzeugt. Laut RT-PCR, Western Blot und Immunhistochemie exprimiert das Schwein AsCas12a mRNA und produziert Protein in vielen Geweben und Zellen. Als jedoch primäre Zellen des AsCas12a Schweins mit sgRNAs gegen das *B2M*- und *GGTA1*-Gen transfiziert wurden, war die Effizienz der Gen-Editierung im Vergleich zu Cas9 niedriger und erfordert Optimierung.

Das Einfügen von Transgenen oder das Ausschalten von proteinkodierenden Genen reicht nicht aus, um alle Krankheiten zu bekämpfen, wenn man bedenkt, dass 98 % des Genoms nicht-kodierend sind. Die lange nicht-kodierende RNA (lncRNA) maternally expressed gene 3 (*Meg3*) befindet sich im *Dio3-Dlk1* Locus und fördert die Herzfibrose bei Mäusen. Um *MEG3* im Schwein zu manipulieren, analysierte ich zuerst die Intron-Exon-Struktur des *MEG3*-Gens. Zum ersten Mal wurden auch die Lage der regulatorischen Elemente *MEG3* DMR und des *MEG3* Promotors im Schwein identifiziert. Als mithilfe von CRISPR/Cas9 Teile des Promotors/DMR und des Exons 2 entfernt wurden, erreichte ich einen Verlust der *MEG3* Expression in immortalisierten Herzfibroblasten des Schweins. Außerdem habe ich mithilfe von CRISPRa und CRISPRi *MEG3* erfolgreich überexprimiert und herunterreguliert. Mit Ausnahme von *MEG8* wurden dabei die Transkriptlevel der anderen Gene des *DLK1-DIO3*-Lokus nicht beeinflusst. RNA-Seq-Untersuchungen der Zellen, bei denen der Promotor entfernt wurde, zeigte eine Herunterregulierung des wichtigsten pro-fibrotischen Signalwegs *TGF- $\beta$* , des Muskelkontraktionswegs und des aktivierten Fibroblastenmarkers *ACTA2*. Das Ausschalten von *MEG3* könnte also das Fortschreiten der Herzfibrose auch im Schwein verhindern.

# 1 Introduction

## 1.1 Generation of genetically modified pigs

The importance of pigs for biomedical applications had become clear well before the first transplantation of a pig heart into a human in 2022<sup>1</sup> or the production of chimeric human-pig embryos in 2017<sup>2</sup>. In the early and middle 20<sup>th</sup> century, pigs were a source of hormones that could not be produced biotechnologically, like insulin or thyrotropin-releasing hormone (TRH). For studies concerning the isolation of TRH from pigs, the Nobel Prize was awarded in 1977<sup>3</sup>.

However, using pigs for biomedical applications often requires genetic modifications. For the above mentioned heart xenotransplantation even multiple genetic modifications including gene knockouts and gene insertions had to be carried out in the xenodonor pig<sup>1</sup>. Generation of genetically modified (GM) pigs with targeted alternations, however, is more challenging compared to mice, where murine embryonal stem (ES) cells enable a relatively high homologous recombination rate<sup>4</sup>.

Until recently, these cells were lacking in pigs<sup>5,6</sup>, and it still needs to be proven, if they are fully functional and can form germline chimeras<sup>7</sup>. Attempts to use the related induced pluripotent stem cells (iPSCs) for generating genetically modified pigs resulted in a very low yield of viable offspring<sup>8</sup>. Furthermore, genetic chimeras, which result from ES-based approaches, are not ideal for generating genetically modified porcine models due to the comparably long generation time of pigs<sup>9</sup>.

Due to the lack of porcine ES cells, the generation of genetically modified pigs still mainly relies on two techniques, somatic cell nuclear transfer (SCNT) and microinjection (MI) (Figure 1), which will be explained in detail in the following two chapters 1.1.1 and 1.1.2. Section 1.2.1 focuses on knocking out endogenous genes, while different ways to insert transgenes in somatic cells before SCNT are described in 1.2.3.

SCNT and MI both require expensive equipment and experienced, well-trained experimentalists. This bottleneck hampers advances in porcine research and makes genetic modifications protracted and inflexible. Consequently, several attempts have been made to circumvent this by applying genetic modifications *in vivo*, which chapter 1.3 focuses on.

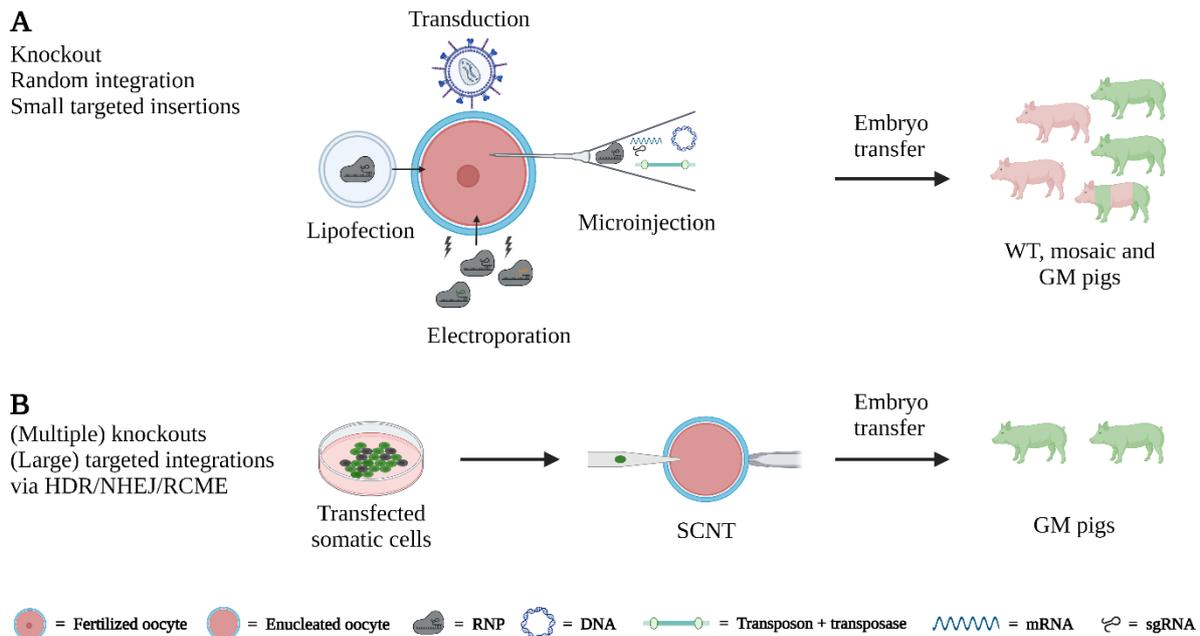


Figure 1 - Generation of genetically modified pigs. Introduction of genetic material into the fertilized oocyte. Different transfection methods can be used to achieve knockouts, random integrations, or targeted insertions of small transgenes. Since no thorough embryo analysis is possible, the resulting offspring is a mixture of wild-type, mosaic, and genetically modified (GM) pigs. (B) Somatic cells are first modified in cell culture for multiple knockouts or targeted large transgene insertions. Only correctly targeted clones are applied for somatic cell nuclear transfer (SCNT), resulting in 100 % GM offspring. Created with BioRender.com.

### 1.1.1 Microinjection

In 1985, the first transgenic pigs were generated via pronuclear microinjection (MI) shortly after the first transgenic mice<sup>10</sup>. Thereby, genetic material is inserted via an injection needle into the pronucleus of a fertilized oocyte. The embryo is then transferred to a surrogate sow (Figure 1).

However, pronuclear MI suffers from several disadvantages, like the production of mosaic animals, low efficiency of only 1 % transgenic animals<sup>11</sup>, and the challenging procedure itself. Furthermore, only random integrations of the donor constructs occur and it does not allow/enable targeted gene modifications<sup>9</sup>. Microinjection had lost relevance<sup>12</sup> with the development of somatic cell nuclear transfer (SCNT, see 1.1.2), resulting in the first cloned mammal from adult cells, “Dolly”<sup>13</sup>.

With the discovery of the CRISPR/Cas9 system<sup>14,15</sup> (see 1.2.1.2), the field was revolutionized by enabling target-specific gene editing via microinjection with unreached efficiency and feasibility<sup>16</sup>. Using CRISPR/Cas9, the material is injected into the cytoplasm instead of the pronucleus since the nuclear core complex takes up DNA/RNA<sup>17</sup> and Cas9 contains a nuclear localization signal<sup>18</sup>. This

facilitates easy injection into fertilized pig oocytes since their pronuclei are only visible after centrifugation due to the high lipid content of the ooplasm<sup>19</sup>.

So far, multiple homozygous KOs in a single step have been reported using not only cytoplasmic microinjection in pigs but also electroporation<sup>20</sup>. One drawback, using CRISPR/Cas9 in the embryo frequently produces mosaicism since genetic alterations can occur in several developmental stages, which increases the time and number of animals required<sup>21</sup>.

### **1.1.2 Generation of genetically modified pigs via SCNT**

In 2001, the first transgenic pig was generated by somatic cell nuclear transfer (SCNT)<sup>22</sup>. In this procedure, primary porcine cells are genetically modified, followed by transferring the transgenic cell into the perivitelline space of an enucleated recipient oocyte. Oocyte and cell are then fused, activated, and the resulting embryo is introduced into a surrogate sow via laparoscopic embryo transfer.

Since genetic modifications can already be introduced *in vitro*, more challenging genetic modifications like targeted insertions of large transgenes can be conducted. Furthermore, it is possible to thoroughly analyze the donor cells before using them for the generation of genetically modified animals, in contrast to fragile embryos. As the full genetic information is derived from the donor cell, 100 % genetically modified offspring are generated without the drawback of mosaicism. Consequently, SCNT is the primary method for generating gene targeted animals<sup>23</sup>.

## **1.2 Tools to alter the genome**

Translational research in pigs requires different types of genomic modifications. Chapter 1.2.1 focuses on how to perform knockouts of endogenous genes, which is nowadays mostly done using CRISPR/Cas. How to modulate gene expression instead of a knock-out is described in chapter 1.2.2.

Several ways to insert transgenes into the porcine genome have been developed. Section 1.2.3 describes these approaches which either use targeting vectors, employing cellular double-strand break (DSB) mechanisms, or site-specific recombinases.

## **1.2.1 Knockouts**

### **1.2.1.1 Conditional knockouts using the Cre/LoxP system**

Before the CRISPR/Cas system was developed, tissue-specific gene knockouts were often achieved by excising genomic sites via the Cre/LoxP system<sup>24</sup>. The target region is “floxed,” meaning that LoxP sites are added up- and downstream of the region to be excised. Upon addition of Cre, mainly by breeding with a Cre-expressing animal line, the LoxP sites are brought in spatial proximity, recombine and the fragment in between is excised. If a tissue-specific promoter drives Cre, conditional gene knockouts can be restricted to specific tissues. Also, inducible promoters can achieve temporal-specific recombination, especially for genes whose knockouts are associated with lethality during embryogenesis. This is particularly useful for cancer research, where organ-specific activation of oncogenes is necessary. This system has been used for decades in mice with great success<sup>25</sup>. For pigs, few Cre-expressing pig lines have been published, due to their laborious generation, but have proven valuable to biomedical research<sup>26–28</sup>.

### **1.2.1.2 CRISPR/Cas9**

The advent of CRISPR/Cas9 revolutionized the field of genome engineering due to its ease of use and adaptability. Originating from *Streptococcus pyogenes*' bacterial defense against viruses, CRISPR/Cas9 is a flexible tool for genome editing and gene regulatory applications. Cas9 is a multidomain Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) associated nuclease (Cas), which can be instructed to bind and cleave complementary DNA targets via a sgRNA. The Cas9 sgRNA consists of a 20 nt CRISPR-RNA complementary to the target site and a ~82 nt trans-activating crRNA sequence mediating the binding of the sgRNA to Cas9. Furthermore, binding of Cas9 to the target sequence requires a 3 bp NGG protospacer adjacent motif (PAM) directly downstream of the genomic target sequence.

The cleavage of genomic DNA and the resulting DNA double-strand breaks (DSB) can lead to indel mutations due to inaccurate DNA repair via non-homologous end joining (NHEJ). In coding regions, this often results in frameshifts leading to a premature translation stop or a non-sense protein, effectively a knockout (KO)<sup>14</sup>.

Before CRISPR/Cas9, TALEN and Zinc Finger Nucleases (ZNF) had been developed to induce site-specific DSBs, but they require laborious protein engineering to alter the target site whereas CRISPR/Cas only requires the adaption of the sgRNA sequence<sup>29</sup>. Because the guide RNA sequence can be so easily adapted, the function of the enzyme modified, and the relatively high specificity of the system, CRISPR/Cas9 is such a versatile and widely used tool. Besides Cas9, there is a broad

spectrum of Cas nucleases showing different properties such as PAM sequence recognition, sgRNA constitution, or sgRNA processing.

### 1.2.1.3 AsCas12a as alternative to SpCas9

Cas12a is part of the type V-A CRISPR family, and the three mainly used variants are derived either from *Acidaminococcus sp.* (AsCas12a), *Lachnospiraceae bacterium* (LbCas12a), or *Francisella novicida* (FnCas12a)<sup>30</sup>. Cas12a binds to a single mature 42–44 nt-long sgRNA, consisting of a 5' 19 nt direct repeat (DR) for the nuclease binding and a 3' terminal, 23 – 25 nt-long spacer responsible for target site binding (Figure 2A).

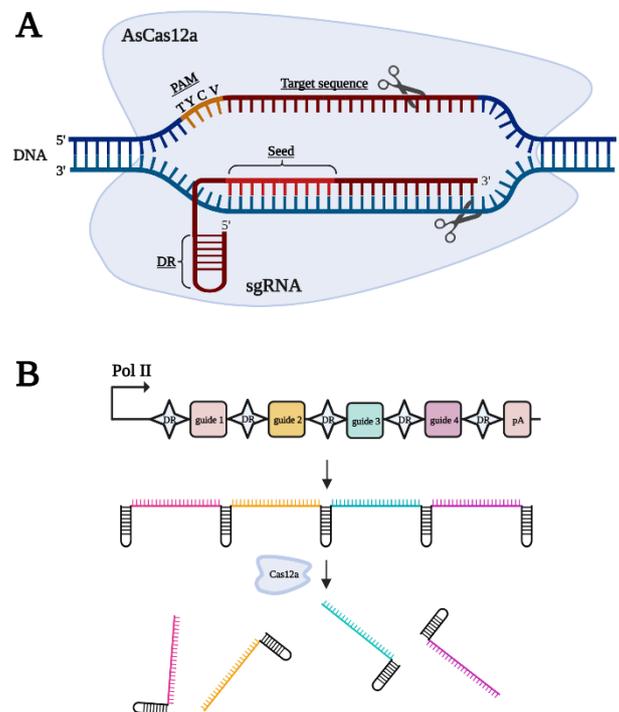


Figure 2 - AsCas12a (TYCV) structure and target recognition. (A) Engineered AsCas12a (S542R/K607R) recognizes a 5' TYCV PAM and produces staggered ends. (B) sgRNAs of Cas12a can be expressed as a single transcript by a polymerase II promoter. The guides are then processed to mature sgRNA by Cas12a's intrinsic ribonuclease domain. Created with BioRender.com.

The seed region is essential for the guide specificity and is defined by the 9-10 nt at the 5' end of the spacer-derived sequence<sup>31</sup>. Cas12a shows higher specificity and is less tolerant to mismatches in sgRNA-target hybridization than Cas9<sup>32</sup>.

In Cas9 and Cas12a, target recognition requires binding of the Cas nuclease to a short DNA motif, the PAM. While Cas9 requires a G-rich 3' NGG PAM sequence, Cas12a recognizes 5' T-rich motifs,

expanding possible genomic targets. Furthermore, the required TTTV (V= A, G, C) PAM for AsCas12a can be modified to recognizing TYCV and TATV PAMs by introducing the mutations S542R/K607R or S542R/K548V/N552R in the Cas12a coding sequence, respectively<sup>33</sup>. The AsCas12a (S542R/K607R) is depicted in Figure 2A.

Differences between Cas9 and Cas12a are also observed in the produced DSB ends. Cas12a cleaves the genomic target sequentially and produces staggered ends with a 5' overhang<sup>34</sup>, while Cas9 cleaves simultaneously and leaves blunt ends, 3 bp upstream of the PAM<sup>35</sup>.

Interestingly, Cas12a-induced DSBs mostly result in larger mutations between 7 and 11 bp within the target sequence since the DSB occurs distal to the PAM, which does not pair with the seed region<sup>36</sup>. Consequently, several rounds of binding and cleavage can occur until a deletion is introduced destroying the PAM proximal region. This is in sharp contrast to Cas9, which typically produces only short indels within the seed region, preventing further cutting<sup>35</sup>. Because of that, Cas9 is faster and more suitable for generating biallelic KO<sup>37</sup>. On the other hand, for HDR-mediated gene insertions in zebrafish and different plants, the staggered ends outclass the blunt ends of Cas9<sup>30,38,39</sup>.

Another considerable advantage of Cas12a is the ability to process CRISPR arrays<sup>40</sup> (Figure 2B). sgRNA maturation is enabled by the intrinsic ribonuclease activity, which allows Cas12a to process guides from a CRISPR array, while Cas9 lacks this and depends on other enzymes<sup>35</sup>. Hence, expression of sgRNAs can be driven by tissue-specific pol II promoters, making Cas12a a valuable tool if cell- or tissue-specific expression is desired. The efficiency of sgRNAs expressed by pol II and pol III promoters is comparable in mammalian cells<sup>41</sup>. In plants, pol II-expressed sgRNAs are even more efficient than sgRNAs expressed by pol III promoters<sup>42</sup>.

## 1.2.2 dCas9 variants for transcriptional control

Cas variants can also be used to modulate gene expression. This requires the usage of Cas9 or Cas12a proteins with inactivated nuclease domains, so called dCas9 or dCas12a proteins, which are fused to other effector domains and are not capable to promote a DSB. These fusion proteins enable precise modulation of gene expression by employing the target specificity of the CRISPR/Cas system and gene modulation activity of the effector domain (see Figure 3).

For transcription activation (CRISPRa), dCas9 or dCas12a is fused to transcriptional activators and guided to a promoter site<sup>43</sup>. An overview of the different CRISPRa fusion proteins is given in Figure 3B. With Synergistic Activation Mediator (SAM), SunTag, and VPR, the original VP64 system, has been further developed to increase efficiency. All systems are commonly used with

high efficiencies<sup>44-46</sup>. SAM seems best suited for activating single genes<sup>46</sup> and has been proven efficient in pigs<sup>47</sup>. However, due to its complicated aptamer structure it relies on several vectors when using viral vectors which impedes its usability. SunTag is an array of VP64 proteins, making it more efficient than the original VP64 system. However, the large size of the array reduces efficiency compared to SAM. The third system Vp64-p65-Rta (VPR), combines three different transcriptional activators working synergistically. It is not as efficient as SAM, but VPR exhibits a high level of activation<sup>46</sup>. It is the most common system due to its simple construction as a fusion and not as a multi-component system. A more current system is based on a different approach by fusing dCas9 to the histone acetyltransferase CREB binding protein (CBP), opening the chromatin structure and activating expression even stronger than SAM<sup>48</sup>.

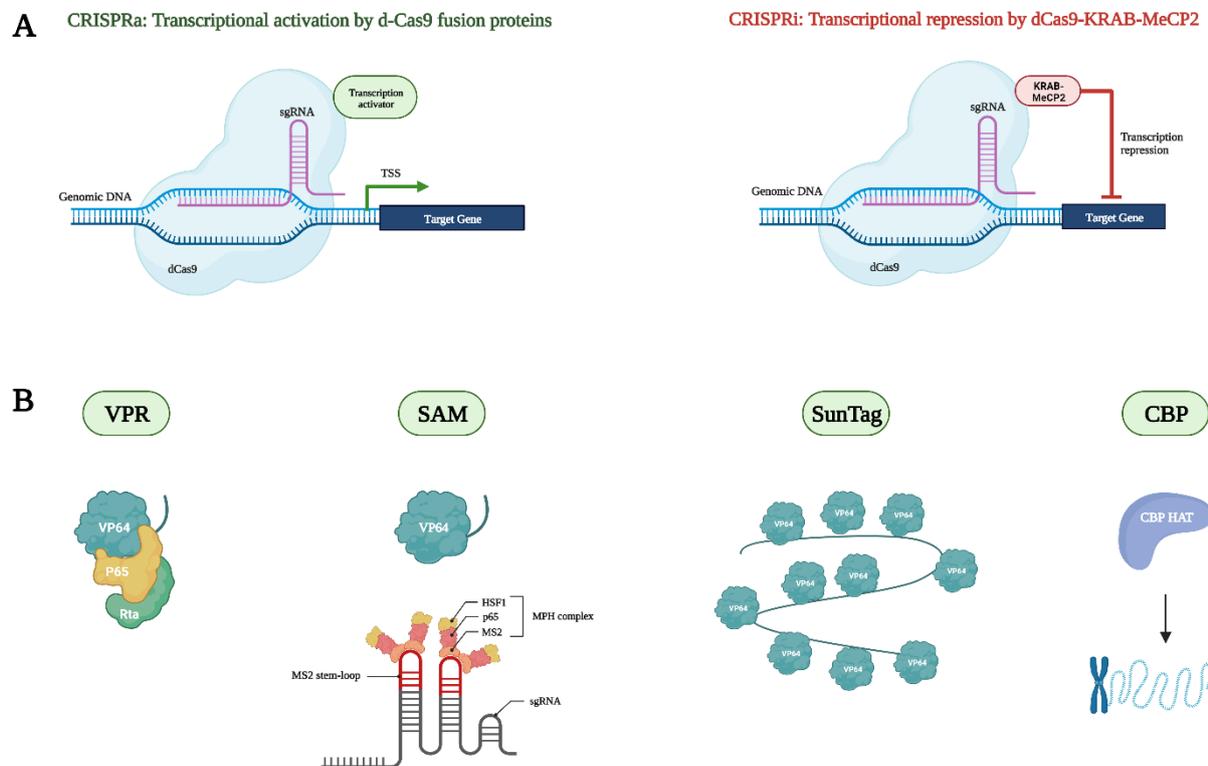


Figure 3 - Transcriptional modulation by dCas9 variants. (A) dCas9 variants can either be fused to transcriptional activators (CRISPRa) to activate expression or the transcriptional repressor KRAB-MeCp2 (CRISPRi) to inhibit expression. (B) Overview of different transcriptional activators fused to dCas9. VPR is a tripartite activator fusion protein of the activators VP64, P65, and Rta. Synergistic Activation Mediator (SAM) is a dual system based on engineered sgRNAs with aptamers that bind transcriptional activators and a dCas9 Vp64 fusion protein. SunTag is a 10-fold array of VP64 activator proteins. The histone acetyltransferase domain of the CBP protein leads to chromatin opening, followed by increased transcription. Adapted from BioRender and Addgene.org1.

On the other hand, CRISPR interference (CRISPRi) enables transcriptional repression of target genes. Binding of the dCas protein to the promoter site is sufficient to block RNA polymerase elongation, thereby inhibiting transcription<sup>49</sup>. The expression can be further reduced by fusing dCas to a transcriptional repressor called Krüppel-associated box (KRAB), that induces heterochromatin formation<sup>50</sup>. A bipartite repressor consisting of KRAB and methyl-CpG-binding protein 2 (KRAB-MeCP2) is even more effective<sup>51</sup>. MeCP2 binds to methylated CpG dinucleotides and recruits a set of histone modifiers, including DNA- and histone methyltransferases and deacetylases. Histone modification, in turn, leads to the formation of heterochromatin and gene repression<sup>52,53</sup>.

### **1.2.3 Transgene addition**

#### **1.2.3.1 Gene insertions via lentiviruses, transposons, and MI**

The production of genetically modified pig disease models often requires insertion of a single or even a series of transgenes<sup>54</sup>. Transgene vehicles like transposons, site-specific recombinases, and viruses can be employed via cytoplasmic microinjection to facilitate gene insertions.

Lentiviral transduction of pig embryos has been used for over two decades with impressively high transgenic offspring efficiencies of up to 70 %<sup>55,56</sup>. Recombinant lentiviruses carrying the transgene stably integrate their genetic information via reverse transcription of RNA into DNA into the host genome, often at several loci.

However, lentiviral transgenesis possesses many drawbacks and thus is not used regularly anymore. The most important drawback is low transgene expression, especially in the offspring, due to hypermethylation of the insert. Furthermore, transgenes might segregate in the offspring due to multiple integrations at different loci or due to their absence from the germline<sup>57</sup>. Even though the transgene packaging capacity of lentiviruses has increased over time, it is still restricted to a maximum of around 10 kb<sup>58</sup>. Hence, lentiviruses are not suitable for large transgenes. Finally, lentiviruses require safety level 2 handling, which is expensive and cannot be provided by every laboratory or large animal husbandry facility.

By contrast, transposon systems can be applied in biosafety 1 labs and work similarly to lentiviruses but lack some of the disadvantages. Transposons are non-viral, mobile DNA components with a relatively straightforward structure. They contain a transposase gene flanked by inverted terminal repeats (ITRs), which carry binding sites for the transposase. The *sleeping beauty* (SB) transposon is efficient and was used several times to generate transgenic pigs<sup>59-61</sup>. Upon injection of the transposase mRNA together with a DNA vector containing the transgene and flanking ITRs into the porcine embryo, the transposase excises the transgene from the vector and

inserts it into the porcine genome. Since most transposons do not depend on reverse transcription, they are not restricted in packaging capacity, even though larger transgenes are inserted less efficiently. Most importantly, transgene expression was persistent after several generations showing less mosaicism.

The SB transposon system shows lentivirus-comparable transgenic offspring efficiencies in pigs of up to 60 %<sup>62</sup>. However, lentiviruses and transposons lead to random and multiple integrations of the transgene, which in turn may result in position effects and transgene segregation during breeding.

### **1.2.3.2 Targeted transgene insertions in pigs using homologous recombination**

Site-specific integration of transgenes at a safe harbor locus enables the high and constitutive expression of the transgene without disturbing the function of essential endogenous genes and without variable expression levels due to transgene segregation during breeding. In mice<sup>63</sup> and pigs<sup>64</sup>, *ROSA26* is the most prominent of such a permissive site. It codes for a non-coding RNA and allows the insertion of transgenes by gene trapping. Even though other safe harbor loci like *H11*<sup>65</sup> have been reported, *ROSA26* seems to perform best<sup>66</sup>.

For a long time, gene targeting via homologous recombination (HR) has been the main method to generate targeted insertions at a permissive locus in porcine primary cells<sup>67</sup>. HR naturally occurs during meiosis but also during the repair of double-strand breaks, termed homology-directed repair (HDR). Thereby, recombination happens between two identical nucleotide sequences. If a DNA template is provided, which harbors a transgene being flanked by nucleotide sequences homologous to the genomic target locus, HR can precisely place the transgene into the genome<sup>68</sup>. However, the efficiency of gene targeting via HR in primary somatic cells is extremely low. Even with selection, the rate is only one in  $10^5$  -  $10^6$  cells<sup>69</sup>. This rate has been massively increased by provoking DSBs in the genomic target region using site-specific endonucleases like CRISPR/Cas9.

Moreover, the rate can be further optimized for example by chemical inhibition of the dominant but error-prone DSB repair mechanism, non-homologous end joining (NHEJ), or by synchronizing the cells to the S2 phase, to increase the HDR rate. Another way is recruiting cellular HDR factors or the DNA template near the DSB site<sup>70</sup>. Due to the precision of HDR, it is the main method for targeted transgene insertions.

Only two targeted transgene insertions via cytoplasmic microinjection in porcine embryos via HDR have been reported. Both were restricted to smaller inserts under 200 bp<sup>71</sup> or a maximum of 2400 bp<sup>72</sup>, indicating a lack of efficiency for targeted transgene insertion in porcine embryos. Consequently, targeted transgene insertion in somatic cells via HDR to generate donor cells for SCNT remains the method of choice.

### **1.2.3.3 Transgene insertion employing non-homologous end-joining**

Despite above mentioned improvements in HDR rate, efficiencies are still much lower compared to the prominent DSB repair mechanism non-homologous end joining (NHEJ). In NHEJ, both ends of the broken DNA are blunted and ligated. Thereby, deletion or insertion of nucleotides (indels) occur, which lead to a loss of the genomic integrity<sup>73</sup> or gene knockouts (KO)<sup>74</sup>. For targeted knock-ins, NHEJ was first neglected due to its inaccuracy, even though NHEJ is the dominant way of DSB repair.

When using NHEJ for transgene insertion, increasing insert size is not associated with decreasing efficiency, even though indels often occur at the insertion site<sup>75,76</sup>. If no in-frame insertions or base pair precise insertions are necessary, NHEJ-mediated insertion of transgenes is a suitable method. However, NHEJ can also be used for in-frame insertions called homology-independent integration (HITI). Using a suitable selection, this can be used for in-frame gene tagging called CRISPaint.

CRISPaint can be used for creating C-terminal tags of endogenous proteins via the canonical NHEJ pathway. Thereby, a donor vector is cleaved at three potential positions to enable variable reading-frame selection simultaneously to a genomic target site where the donor DNA is then incorporated<sup>77</sup>.

### **1.2.3.4 Genetic modifications using recombinase-mediated cassette exchange (RMCE)**

Targeted transgene insertions have also been achieved in pigs using site-specific recombinases, via recombinase-mediated cassette exchange (RMCE)<sup>78</sup>. Thereby, recombination between a donor vector and a specific target site in the genome is catalyzed by a site-specific recombinase. However, it has to be ensured that no recombination occurs at pseudo-target sites, which is a problem concerning the most commonly used recombinases, such as Cre or Flp. If very target-specific unidirectional integrases are used, like the Bxb1 integrase, the target site has to be integrated into the genome beforehand, mostly by drawing on CRISPR/Cas9<sup>79</sup> (Figure 4). Although large transgenes have been inserted in mice using the Bxb1 integrase<sup>80</sup>, there are no reports to date for transgenic pigs.

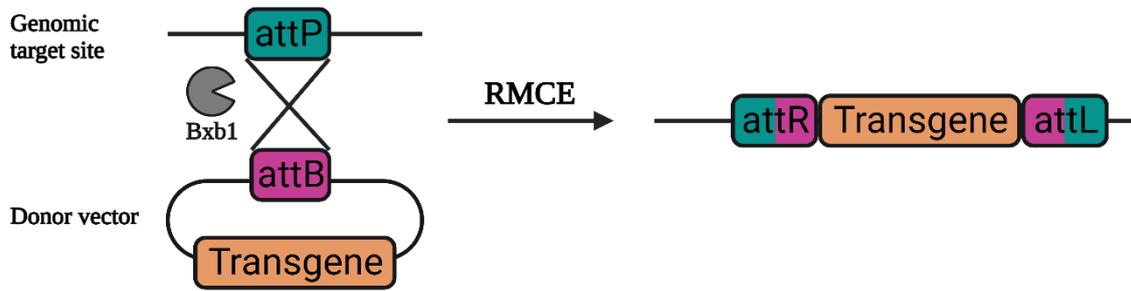


Figure 4 - Transgene insertion via recombination-mediated cassette exchange (RMCE). Adapted from Mulholland et al. (2015)<sup>79</sup>. Created with BioRender.com.

### 1.3 *In vivo* genome editing

Despite the fact that CRISPR/Cas9 has revolutionized the field of genome engineering, the generation of genetically modified pigs is hampered by the bottleneck of SCNT and MI. Hence, attempts have been made to perform modifications directly in pigs *in vivo* instead of in embryos or cell culture. There are two major approaches to achieve *in vivo* genome editing which use either viral vectors (1.3.1) or Cas9-expressing pigs (1.3.2).

#### 1.3.1 *In vivo* genome editing by AAVs

The first attempts for *in vivo* genome editing have been made using viral vectors. Cas9 and the sgRNA are delivered into wild-type animals or tissues via viral transduction. Even though lentiviruses and adenoviruses have been used for this purpose, Adeno-associated viruses (AAVs) are leading the field due to their serotype tissue specificity, low immunogenicity, and limited integration into the host genome<sup>81</sup>.

Currently the FDA has approved several different AAV gene therapy approaches in humans, and the number increases yearly<sup>82</sup>. However, the major drawback using this viral vector is the limited packaging capacity that is below 5 kb. Large transgenes or a promoter plus CRISPR/Cas9 coding sequences exceed this capacity, impeding efficient genome editing. One attempt to circumvent this is to split the Cas9 coding sequence using two separate AAVs (Split-Cas9). However, successful genome editing requires the presence and correct assembly of both Cas9 components plus the sgRNA within the cells and hence is associated with lower efficiencies as three components need to be delivered and assembled successfully<sup>83</sup>.

Moreover, *in vivo* genome editing via AAVs faces specific problems in pigs compared to mice since a huge viral load is needed. The size of the pig, the requirement for biosafety level 2 animal housing, and the labor-intensive production of AAVs contribute to increased costs for this technique. Hence, only a few studies used AAVs for somatic genome editing in pigs so far<sup>84-86</sup>.

### **1.3.2 *In vivo* genome editing using SpCas9-expressing pigs**

With the development of a constitutive Cas9-expressing pig line expressing SpCas9 at our chair<sup>87</sup> and a Cre-dependent Cas9-expressing pig line by others<sup>88</sup>, *in vivo* genome editing in pigs has reached the next level. These pigs can now more easily be used for *in vivo* genome editing, as only sgRNAs need to be delivered by AAVs, eliminating the issue with exceeding the packaging limit. Administration of only two sgRNA by AAVs in Cas9-expressing porcine hearts excised a 12.7 kb DNA fragment of the *MYBPC3* gene *in vivo*<sup>87</sup>.

Despite this success, AAVs cannot always ensure specific tissue penetration. Even though capsid engineering has improved viral tropism, AAV serotypes often transduce more than one tissue<sup>89</sup>. In particular, if genome editing in specific cell types is desired, Cas9 faces the problem that expression of sgRNAs depends on RNA pol III promoters like U6 or U3 snRNA promoter, which are ubiquitously expressed and lack spatiotemporal expression<sup>90</sup>.

Alternatively, Cas variants like AsCas12a, which do not depend on RNA pol III promoters, provide a suitable solution.

## **1.4 Long non-coding RNAs**

68 % of the human genes are classified as long non-coding RNA (lncRNA), RNAs which are longer than 200 nt, and are not coding for proteins<sup>91</sup>. Unlike mRNA, many lncRNAs are only poorly conserved and expressed. However, they share many features with mRNAs, like splicing, 5'-capping, or 3'-polyadenylation. These lncRNAs play various essential roles by interacting with mRNA, miRNA, DNA, and proteins through their complex 3D structure or nucleotide interactions. They act across chromatin remodeling, methylation, and transcriptional regulation<sup>92</sup>. However, they influence not only physiological but also pathological processes. One lncRNA known to be involved in numerous pathological processes is the imprinted lncRNA maternally expressed gene 3 (MEG3)<sup>93</sup>.

### 1.4.1 The regulation of maternally expressed gene 3 (MEG3)

Maternally expressed gene 3 (MEG3), as the name indicates, is an imprinted gene where only the maternal allele is expressed. In genomic imprinting, the origin of the allele – paternal or maternal – determines the regulation of gene expression. Imprinted genes like MEG3 are organized in clusters of up to 20 genes regulated by superior genetic elements, the imprinting control regions (ICR). All ICRs are differentially methylated regions (DMR) where the methylation depends on parental origin<sup>94</sup>. The ICRs act as epigenetic master switches to control the gene expression of multiple imprinted genes. Hence deletion of the imprinting mark is often lethal even though only 1 out of 100 mammalian genes is imprinted<sup>95</sup>.

#### 1.4.1.1 The human and murine MEG3 locus

MEG3 resides in the imprinted DLK1-DIO3 locus on chromosome 14 in humans, chromosome 12 in mice, and chromosome 7 in pigs<sup>96</sup>. Figure 5 shows a schematic presentation of the human and murine genomic regions.

The locus contains three protein-coding genes DLK1, DIO3, and RTL1, and, besides MEG3, the three other lncRNA genes MEG8, RTL1as, and MIRG. Furthermore, the locus harbors one of the largest known clusters of micro-RNAs (miRNA) and small nucleolar RNAs (snoRNAs)<sup>97,98</sup>. Except for RTL1, all genes are transcribed in 5' to 3' direction, and it is proposed that all non-coding RNAs within the locus are derived from a single polycistronic transcript starting with MEG3<sup>96,99</sup>.

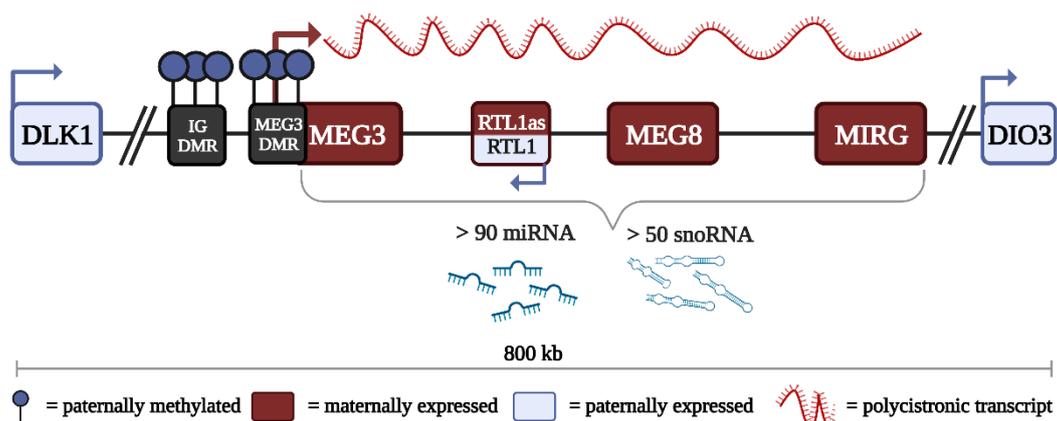


Figure 5 - Schematic presentation of the human and murine DLK1-DIO3 locus. All non-coding genes are maternally expressed, while all coding genes are paternally expressed. Locus imprinting and the hypermethylation of the paternal IG-DMR and MEG3-DMR controls expression. All non-coding RNAs are derived from a single polycistronic transcript driven by the MEG3 promoter. The DLK1-DIO3 locus contains an enormous cluster of miRNA and snoRNAs. DMR = differentially methylated region, IG-DMR = intergenic DMR, miRNA = microRNA, snoRNAs = small nucleolar RNA, DLK1

= delta-like non-canonical Notch ligand 1, MEG3 = maternally expressed gene 3, RTL1 = retrotransposon-like gene 1, RTL1as = RTL1 antisense, MEG8 = maternally expressed gene 8, MIRG = miRNA containing gene, DIO3 = iodothyronine deiodinase 3. Adapted from Sanli et al.<sup>96</sup>, Budkova et al.<sup>98</sup>, and MacDonald et al.<sup>94</sup>. Created with BioRender.com.

The regulation of the locus is highly complex. The region contains at least two ICRs, the intergenic DMR (IG-DMR) and the MEG3-DMR. The maternal expression of MEG3 is ensured by paternal hypermethylation and maternal hypomethylation of the primary germline-derived IC-DMR and the secondary post-fertilization-derived MEG3-DMR. The methylation of the MEG3-DMR is dependent on the methylation of the superior IG-DMR. Promoters of imprinted lncRNAs often overlap with the ICRs, which is true for the MEG3 promoter and the MEG3-DMR. Moreover, expression of the locus is regulated by 3D chromatin interactions between CCCTC-binding factor (CTCF) sites and the MEG3-DMR<sup>100,101</sup>. CTCFs are crucial for organizing the 3D structure of the genome to ensure intra-domain interactions and bring together genes and their regulatory elements<sup>100</sup>.

#### **1.4.1.2 Expression of MEG3 is mainly controlled by promoter methylation**

Besides the epigenetic control elements, a series of other factors influence MEG3 expression. Figure 6 gives an overview. Briefly, binding of transcription factors like STAT3<sup>102</sup> and NFκB<sup>103</sup> to the MEG3 promoter or cAMP<sup>104</sup> to cAMP response elements (CRE) increase MEG3 expression. Furthermore, MEG3 is upregulated upon miR-141-mediated inhibition of the cell cycle protein E2F3 expression which negatively regulates MEG3<sup>105</sup>.

Repression of MEG3 is mainly regulated by methylation of the CpG-rich promoter<sup>106</sup>. DNA methyltransferases (DNMTs) mediate the methylation of the MEG3 promoter<sup>107</sup>, which frequently happens in tumors to eliminate MEG3's tumor-suppressive function<sup>108</sup>. Promoter methylation can be inhibited by a series of miRNAs (miR-26a, miR-29a, miR-148a, miR-185) binding to the DNMT mRNA. Furthermore, protein-protein interaction of the retinoblastoma protein<sup>109</sup> (pRb) or ubiquitin-like with PHD and ring finger domains 1 (UHRF1) with DNMT can inhibit promoter methylation<sup>107,110</sup>.

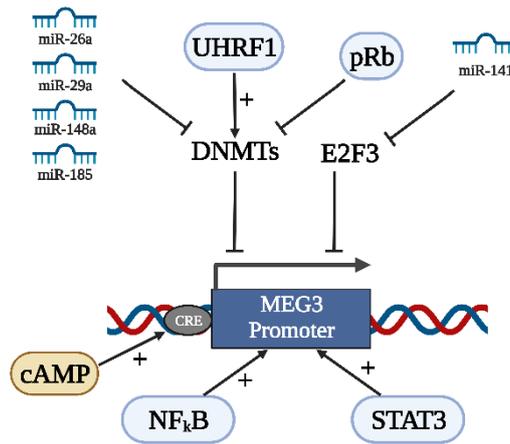


Figure 6 - Regulation of MEG3 expression. MEG3 expression is also regulated by cyclic adenosine monophosphate (cAMP) response elements (CRE) at the proximal promoter. Created with BioRender.com.

#### 1.4.1.3 MEG3 acts preferably in *trans*

Table 1 gives an overview of MEG3's action in *cis* and *trans*. MEG3 represses the upstream DLK1 expression in *cis* by recruiting CTCF to a maternal-specific CTCF binding site located in its own promoter DMR. This isolates DLK1 physically and represses its expression. Also, it ensures the expression of MEG3 itself since it isolates the active maternal MEG3 allele from the inactive maternal RTL1 and DIO3 allele<sup>94</sup>.

By-co-localization in the same 3D nuclear space, MEG3 influences other imprinted loci/lncRNA in *trans* as imprinting network<sup>111</sup>. Furthermore, the crucial signaling pathways TGF- $\beta$ , WNT/ $\beta$ -Catenin, and tumor suppressor TP53 are influenced by MEG3 in *trans*<sup>112-114</sup>. A possible way of action is by DNA-RNA triplex formation. MEG3 exhibits several triplex-binding sites across the genome, and it has been shown that MEG3 changes TGF- $\beta$  signaling by triplex formation with subsequent Polycomb repressive complex 2 (PRC2) recruitment<sup>112</sup>. PRC2 establishes the repressive chromatin modification H3K27me3 and silences target genes bound by the MEG3-RNA-DNA triplex<sup>115</sup>.

Another way of action for MEG3 is as competing endogenous RNA (ceRNA). Studies have demonstrated that MEG3 is a sponge for numerous miRNAs<sup>102,107,116</sup>. MEG3 binds these miRNAs, which hampers the miRNAs' binding to target mRNAs. As a result, the target mRNAs expression is increased<sup>117</sup>. Moreover, as indicated in Figure 5, the MEG3 transcript is also a precursor for smaller RNAs, like the miR-770-5p, which in turn influences pathological processes<sup>118</sup>.

Due to its ability to bind TP53 directly, MEG3 has been extensively studied in cancer<sup>119-121</sup>.

Table 1 - Selected *cis* and *trans* actions of Meg3.

<b><i>Cis/trans</i></b>	<b>Way of action</b>	<b>Effect</b>
<i>cis</i>	CTCF recruitment	Silencing of maternal DLK1
		Separation of silent RTL1/DIO3 allele from active MEG3 allele
<i>trans</i>	Co-localization with imprinted lncRNAs	Imprinted locus regulation
	Triplex formation	PRC2 recruitment
	Protein interaction	Protein stabilization
	miRNA sponging	Gene regulation
	Small RNA precursor	

#### 1.4.2 MEG3 in cardiac fibrosis and hypertrophic cardiomyopathy

Heart disease is one of the most difficult diseases to treat, mainly due to the incapacity of cardiomyocytes of the adult heart to proliferate<sup>122</sup>. Hence, new treatment options are urgently needed<sup>122,123</sup>. It is reported that MEG3 is an important player in cardiac diseases<sup>124</sup> and influences heart failure<sup>125</sup>, myocardial infarction<sup>126</sup>, and (diabetic) cardiomyopathy<sup>102,127</sup>. Moreover, cardiac fibrosis is linked to all these disorders<sup>124</sup>. In cardiac fibrosis, cardiac fibroblasts are activated after injury and transformed into myofibroblasts by an increase in crucial transforming growth factor  $\beta$  (TGF- $\beta$ ) signaling<sup>128</sup>. The activated myofibroblasts upregulate specific genes like alpha-smooth muscle actin ( $\alpha$ -SMA), a strongly contractile protein, or metalloproteases (TIMPs), and matrix metalloproteases (MMPs). Activation of these genes leads to massive deposition of the extracellular matrix. Ultimately, the heart becomes stiffer and progresses to heart failure<sup>129</sup>. Figure 7 gives an overview of the pro-fibrotic and pro-hypertrophic action of MEG3.

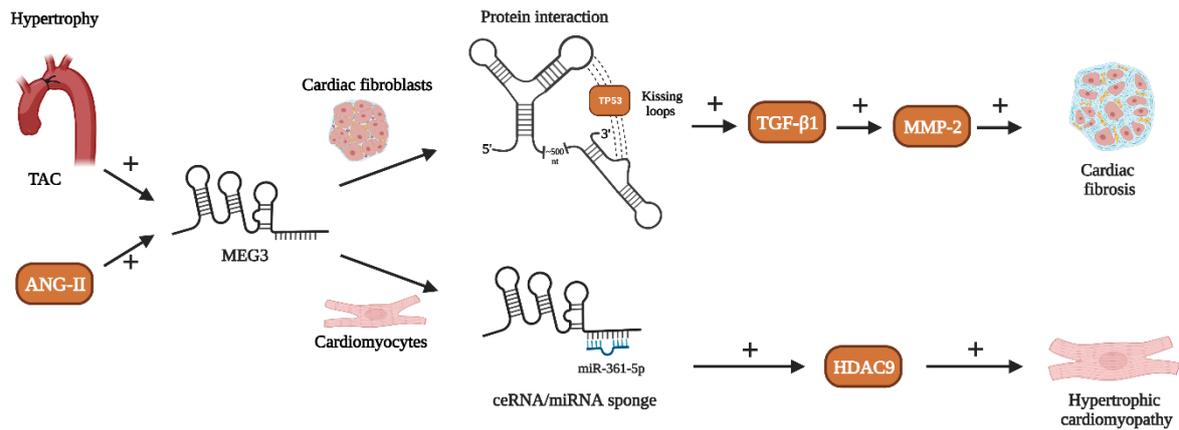


Figure 7 - Schematic illustration of the pro-fibrotic and pro-hypertrophic action of MEG3. After a hypertrophic stimulus either by angiotensin II or transverse aortic constriction (TAC), MEG3 expression is increased. Via direct TP53 protein interaction by binding in kissing loops in cardiac fibroblasts, expression of TGF-β1 genes and MMP2 is increased, ultimately resulting in cardiac fibrosis. In cardiomyocytes, MEG3 sponges miR-361-5p, which increases HDAC9 expression leading to hypertrophic cardiomyopathy. Created with BioRender.com.

Strikingly, it has not only been shown that MEG3 can bind regulatory elements of the TGF-β pathway by RNA-DNA triplex formation<sup>112</sup>. Moreover, MEG3 is directly involved in cardiac fibrosis. Six weeks after pressure overload induced by transverse aortic constriction (TAC), the TGF-β I induced upregulation of Mmp2 via TP53 protein-interaction is hampered when MEG3 is knocked down using an LNA-GapmeR<sup>130</sup>. Consequently, pathological remodeling of the heart and cardiac fibrosis were reduced. Another study indicates that selenium may reduce cardiac fibrosis by a decrease in MEG3 expression<sup>131</sup>.

Hypertrophic cardiomyopathy describes the pathological thickening of the heart<sup>132</sup>. MEG3 is upregulated in hypertrophic cardiomyocytes by the transcription factor STAT3. MEG3 acts as a ceRNA for miR-361-5P. When they bind each other, miR-361-5P cannot inhibit its target mRNA of the histone deacetylases 9 (HDAC9), leading to an upregulation of the hypertrophic HDAC9. When MEG3 is knocked down, sponging and HDAC9 upregulation are reduced, and hypertrophic cardiomyopathy is attenuated<sup>102</sup>.

Even though these studies uncover MEG3 as a promising therapeutic target, they were primarily conducted in mice. Due to the low sequence and function conservation of lncRNAs<sup>133</sup>, they need to be confirmed in other (large) species.

## 1.5 Objectives

The discovery and further development of the bacterial defense system CRISPR/Cas have revolutionized the field of gene editing. In this work, I aimed to adapt and develop CRISPR/Cas-based genome engineering tools for pigs to facilitate the generation of genetically modified pig models and to modulate coding and non-coding genes.

Generating transgenic pig models by inserting transgenes via homology-mediated gene targeting is time-consuming and expensive due to low efficiency and laborious screening. Alternatives are needed to insert transgenes more efficiently into the porcine genome. Therefore, the first aim of my thesis was to create a homology-independent CRISPR-based system for targeted gene addition into the porcine genome.

However, even though the generation of genetically modified cells can be improved, the generation of genetically modified pigs remains labor-intensive, inefficient and expensive. Therefore, a pig line for *in vivo* genome editing is desired enabling inactivation of various endogenous genes in specific organs or cell types in adult animals. The CRISPR/AsCas12a system provides an intrinsic ribonuclease activity enabling the expression of sgRNAs under the control of tissue-specific promoters. The addition of sgRNAs against various targets to an AsCas12a-expressing pig line could eliminate the need to produce a separate animal model for each disease reducing costs and the required animal number. Thus, my second objective was establishing and characterizing a transgenic AsCas12a-expressing pig line.

So far, knockouts of endogenous porcine genes were carried out in protein-coding genes. However, not only coding-, but also non-coding genes play essential roles in diseases. The lncRNA Meg3 is linked to the development of cardiac fibrosis<sup>130</sup> and hypertrophic cardiomyopathy<sup>134</sup> in mice. In my third objective, I aimed to translate these findings into the pig by applying a toolbox based on Cas9 to either overexpress, knock down, or knock out the porcine *MEG3*.

## 2 Material and methods

### 2.1 Materials

#### 2.1.1 Oligonucleotides

All oligonucleotides were purchased from MWG Eurofins, Ebersberg, GER. Synthetic guides were purchased from Integrated DNA Technologies, Inc., Coralville, USA and modified to prevent immune stimulation and reduce degradation. Probes for droplet digital PCR (ddPCR) were modified with a 5' fluorescence dye (HEX/FAM) and a 3' quencher (BHQ1). Reverse primers for bisulfite sequencing were 5' biotinylated.

##### 2.1.1.1 Primers and probes

Table 2 - Oligonucleotides for luciferase assay and methylation assays

<b>Luciferase reporter assay</b>		
<b>Amplicon</b>	<b>Oligonucleotide name</b>	<b>5' → 3' sequence</b>
DMR1	56B4-pMEG3-DMR-F1	tcgctcttggtctttagggg
	56B5-pMEG3-DMR-R1	caggcagcgacacatgttg
DMR2	51E4-pMEG3-prom-seq-F1	atctcatgtgccaagtccc
	50E7_pMEG3_E1_R1	aggttctcgcgtgggc
DMR3	51E7-pMEG3-E1-seq-F1	gcttttgagaaatgagcgc
	51F3-pMEG3-I1-seq-R2	tccttctgacatggcccaa
DMR4	56B3-pMEG3-DMR-F2	cctaggacgaggtgtgtgg
	51F4-pMEG3-I1-seq-R3	gaaaccttgctccaacc
DMR5	56B7-pMEG3-DMR-F3	ggtggtcaatggcagctctg
	56B8-pMEG3-DMR-R2	gccacgttaggtataaattgcc
DMR6	52H7-MEG3_T1_Scr_F1	ctctgaggtcctgagcgtc
	56B9-pMEG3-DMR-R3	aaatcagtcctgtgccctg

Bisulfite sequencing		
Methylation assay name	Oligonucleotide name	5' → 3' sequence
Meg3-F2-1	pMeg3_Pyro_CpG_F2	ggggtattttattttaggataataagt
Meg3-F2-2	pMeg3_Pyro_CpG_R2_Bio	BIO-taaaccccccaattctataacaattac
Meg3-F2-1	pMeg3_Pyro_CpG_seq_F2-1	agtagaaagtatta
Meg3-F2-2	pMeg3_Pyro_CpG_seq_F2-2	attaggtttaagtagggagaaa
Meg3-CTCF	pMeg3_CpG_CTCF_F	gtagttagtgttttggtagtt
	pMeg3_CpG_CTCF_R_5'_Bio	BIO-aacaaaacacaaatattctatatac
	pMeg3-CpG-CTCF-seq-F	ttgggtttgggtgt

Table 3 - q-RT-PCR and ddPCR oligonucleotides

q-RT-PCR		
Target gene	Oligonucleotide name	5' → 3' sequence
<i>GAPDH</i>	GAPDH_1F	ttcacgaccatggagaaggc
	GAPDH_1R	ggttcacgccccatcacaac
<i>MEG3</i>	qPCR-Meg3-F	cgagaacctccctacctgag
	qPCR-Meg3-R	ctggctggtcagtctctgt
<i>MEG3</i>	pMeg3-qPCR-F4	gatcccaccagcctacgaag
	pMeg3-qPCR-R4	agcatagcaaaggctcagggg
<i>MEG8</i>	pMeg8-qPCR-F4	ggagtgtggagttgcatgaa
	pMeg8-qPCR-R4	agtccttggctgtgtatcc
<i>DIO3</i>	Dio3-qPCR-F1	aatttcggaagctgcacctg
	Dio3-qPCR-R1	gggatgctgtagggagagtc
<i>RTL1as</i>	pRTL1as-qPCR-F2	aaggaggaagacagatgccg
	pRTL1as-qPCR-R2	ctcactcctctgggcag
<i>DLK1</i>	pDLK1-qPCR-F3	agtgcattgcaaggacggc
	pDLK1-qPCR-R3	caggttccattgttggcgca
<i>RTL1</i>	pRTL1-qPCR-F3	ggctggtcgggaagtctcctc
	pRTL1-qPCR-R3	cactcgcctactgcctgaa
<i>ACTG2</i>	ACTG2-qPCR-F2	agacacaccagcctcagtc
	ACTG2-qPCR-R2	ggagcgtcatctcctgcgaa
<i>MYL9</i>	MYL9-qPCR-F2	cgtgatccgcaacgcctt
	MYL9-qPCR-R2	gcgctcacggtacatctcg
<i>THBS1</i>	THBS1-qPCR-F1	acatggatggggttggcgat
	THBS1-qPCR-R1	ttgtcgtggtcagcctggtt
<i>ITGB5</i>	ITGB5-qPCR-F1	aagtggagctgtccgtctgg
	ITGB5-qPCR-R1	caggggtaacgtgtgctgtg
<i>SDC2</i>	SDC2-qPCR-F1	agagtggaaaccacgacgct
	SDC2-qPCR-R1	tcagctgggtccgttttctt
<i>CRABP1</i>	CRABP1-qPCR-F2	gccaggacgggatcaattct
	CRABP1-qPCR-R2	cccctcgagcagagtttgtg
<i>ID3</i>	ID3-qPCR-F2	ccaaacgaccttctgccact

	ID3-qPCR-R2	gacttcggccgtagagggt
<i>SMAD6</i>	SMAD6-qPCR-F1	ctctcggtgtctcctcgtg
	SMAD6-qPCR-R1	ttggtggcctccgtttcagt
<i>ACTA1</i>	ACTA1-qPCR-Fwd4	agcttcgtgtgccccagaa
	ACTA2-qPCR-Rev4	cagaggcgtagaggacagc
<i>MMP2</i>	pMMP2 qPCR F2	gctccagttaaaggcagcat
	pMMP2 qPCR R2	aggaggagaaggctgtgttc
<i>18srRNA</i>	61I3-18s-rRNA-qPCR-F	cccacggaatcgagaaagag
	61I4-18s-rRNA-qPCR-R	ttgacggaaggccacca
<b>Copy number determination</b>		
<i>Cas12a</i>	Cpf1_ddPCR_F1	gtctggcttctgttttacg
	Cpf1_ddPCR_R1	gtagtcagaaaagtcgaagc
	Cpf1_ddPCR_Probe	FAM-cgtggacccttcgtgtgga-BHQ1
<i>GAPDH</i>	ddpoGAPDH F1	ctcaacgaccactctgcaaa
	ddpoGAPDH R1	ccctgttctgtagccaaat
	ddGAPDH-HEX_probe	HEX-TGTGATCAAGTCTGGTGCCC-BHQ1
<i>Hygromycin resistance</i>	Hygro-F3	cagcttcgatgtagggggc
	Hygro-R3	tcttgcaacgtgacaccctg
	ddprobe-Hygro-1	FAM-GCGCCGATGGTTTCTACAAA-BHQ1

### 2.1.1.2 sgRNAs

Table 4 - CRISPR/Cas9 sgRNAs targeting *MEG3*

Guide name	Location within the <i>MEG3</i> gene	sgRNA 5' → 3' sequence
pMEG3-Prom T1	Promoter	ggggttcaaagtgtacgtgt
pMEG3-Prom-T2	Promoter	ggggtaaccctcgttaacg
pMEG3-Prom-T3	Promoter	tttgcgattgggagactcgg
pMEG3-Prom-T4	Promoter	ttgccgggggaactggacaa
pMEG3-Prom-T5	Promoter	gctcagggtgttggtcatgg
pMEG3-Prom-T6	Promoter	aatttgcataagaatctggg
pMEG3-DMR-2-T1	Promoter	ccgggcactgagtggaag
pMEG3-DMR-2-T2	Promoter	ccagcccctagcgcagacgg
pMEG3-E1-T1	Exon 1	caggtaggagggttctcgcg
pMEG3-I1-T0-2	Intron 1	gctaccttggaaacacggg
pMEG3-I1-T1	Intron 1	aatttcgtcgccgatgctcg
pMEG3-I1-T2	Intron 1	cttgattcgcaatccctag
pMEG3-I1-T3	Intron 1	tattagggccatcatgcag
Meg3_T1	Intron 1	gtggcaattatacctaagc
Meg3_T2	Exon 2	gaatgacgtcctgaacggaa
Meg3_T3	Exon 2	cgacgtcatgccatggaac
Meg3_T4	Exon 2	gaatgagatgaagtcgctc
Meg3_T5	Exon 2	ttactgcgtcaggcatacgt

Meg3_T6	Exon 2	gcgtttatctaccccaca
---------	--------	--------------------

Table 5 - CRISPR/Cas12a sgRNAs

<b>Cas12a pol III sgRNAs</b>	
<b>gRNA target site and name</b>	<b>sgRNA 5' → 3' sequence</b>
Cas12a-TYCV-CAG-pB2M-E1T1	accgccagcaccgctccagtagc
Cas12a-TYCV-CAG-pB2M-E1T2	ggccagacagtgagagcagcccg
Cas12a-TYCV-CAG-pGGTA1-E7T1	ataccactggagccttccatctg
Cas12a-TYCV-CAG-pGGTA1-E8T1	ttctcaacaaaccactaaaatc
<b>Synthetic sgRNAs</b>	
pGGTA1-Cas12a-E7T1	ataccactggagccttccatctg
pB2M-Cas12a-E1T2	ggccagacagtgagagcagcccg

## 2.1.2 DNA vectors

Table 6 - DNA vectors

<b>Name</b>	<b>Supplier</b>
AsCpf1(TYCV)(BB) (pY211)	
SP-dCas9-VPR (Plasmid #63798)	
SV40 1: pBSSVD2005 (Plasmid #21826)	
pB-CAGGS-dCas9-KRAB-MeCP2 (Plasmid #110824)	Addgene, Cambridge, USA
AsCas12a(TTTV)-Triplex-pCI108-SiT-Cas12a-[Cond]-(Plasmid #128407)	
pJET1.2/blunt	Thermo Fisher Scientific, Waltham, USA
OpCas12a-3xHA	
841-pX330-MCS-T2A-Puro	
px330-eGFP(w/o-stop)-Puro-CAG-sRgRsR	
px330-Cas12a-TYCV-CAG-sRgRsR	
pX330_T2A+_CIRSPR_3_ROSA26-(New-HA targeting)	
px330_ROSA26_old-HA_g1	Chair of Livestock Biotechnology, TUM, Freising, GER
ROSA26-SA-BS-LA	
Cas12a(TYCV)_ROSA26-SA-BS-LA_old_TV	
Cas12a(TYCV)_ROSA26-SA-BS-LA_new_TV	
CRISPlace-universal-donor-(Hygro)-2x-target-site-(PSL1180)	
CRISPlace-universal-donor-(Hygro)-1x-target-site-(PSL1180-Backbone)	

AsCas12a(TYCV)-CRISPlace-Targeting-Vector- (Hygro)-2x-target-site-(PSL1180)	
AsCas12a(TYCV)-CRISPlace-Targeting-Vector-(Hygro)-2x-target-site-(PSL1180)	
hUCP1-SV40-hygro-CRISPLace-Targeting-Vector	
CRISPlace-universal-guide+BbsI	Chair of Livestock Biotechnology, TUM, Freising, GER
CRISPlace-universal-guide+ROSA26-g3	
CRISPlace-universal-guide+UCP1_gRNA2	
CRISPRa-Sp-dCas9-VPR-Meg3-CRa	
CRISPRi-dCas9-KRAB-MeCP2-Meg3-CRi	
839_pcDNA3-eGFP	
709_pJET1.2-pCAGGS	
pcDNA3.1-hygro-cag-2xBG	
reverse-pidt-trna-grna-scaffold	
AsCas12a-6xNLS-E174RS542R-(pRG232)-P2A-Blast	
CRISPaint-mNeon-2APuro	Kindly provided by Jonathan L. Schmid-Burgk, Institute of Molecular Medicine, University of Bonn, Bonn, GER
NanoLuc-empty	Kindly provided by Sabrina Schleibinger, Chair of Reproductive Biotechnology, TUM, GER
PGK-Phirefly-Luciferase-empty	

### 2.1.3 Antibodies

Table 7 - Antibodies

Antibody	Dilution	Supplier
<b>Primary antibodies</b>		
Monoclonal anti-GAPDH antibody produced in mouse; clone GAPDH-71.1	1:3300 (WB)	Sigma-Aldrich, Steinheim, GER
Anti-HA tag antibody - ChIP Grade (ab9110)	1:1000 (WB)	Abcam, Cambridge, UK
Isolectin B4 (Bandeiraea simplicifolia, biotin conjugate, ALX-650-001B-MC05)	1:400 (FACS)	Enzo Life Sciences, Farmingdale, USA
Monoclonal Anti-B2M-Biotin antibody produced in mouse (SAB4700015)	1:50 (FACS)	Sigma-Aldrich, Steinheim, GER
<b>HRP-coupled secondary antibodies</b>		
m-IgGκ BP-HRP (sc-516102)	1:1000 (WB)	Santa Cruz Biotechnology, Dallas, USA
Goat Anti-Rabbit IgG-HRP (4030-05)	1:1000 (WB)	Southern Biotechnology Associates, Birmingham, USA
<b>Anti-biotin secondary antibody</b>		
PE Streptavidin (BD554061)	1:200 (FACS) 1:800 (FACS)	BD Bioscience, Frankling Lakes, USA

### 2.1.4 Chemicals

Table 8 - Chemicals

Name	Supplier
DAB, 3,3'-Diaminobenzidine enhanced liquid substrate system tetra hydrochloride	Sigma-Aldrich, Darmstadt, GER
Acidic acid	AppliChem, Darmstadt, GER
Agarose	Sigma-Aldrich, Darmstadt, GER
Boric acid	AppliChem, Darmstadt, GER
Bovine serum albumin, BSA (Fraction V)	Biomol, Hamburg, GER
Blocking solution immune cell dingens	
Advanced protein assay reagent	Cytoskeleton, Denver, USA
2-log DNA ladder	New England Biolabs, Frankfurt, GER
Deoxynucleotide (dNTP) solution mix	New England Biolabs, Frankfurt, GER
DTT, Dithiothreitol	Omnilab-Laborzentrum, Bremen, GER

EDTA	AppliChem, Darmstadt, GER
Ethanol absolute	Fisher Scientific, Loughborough, UK
Formaldehyde solution 37%	AppliChem, Darmstadt, GER
Gel loading dye purple (6x)	New England Biolabs, Frankfurt, GER
Glycine	Carl Roth, Karlsruhe, GER
Methanol	Sigma-Aldrich, Steinheim, GER
Milk powder, blocking grade	Carl Roth, Karlsruhe, GER
NE-PER™ nuclear and cytoplasmic extraction reagents	Thermo Fisher Scientific, Loughborough, UK
peqGREEN dye	VWR International, Ismaning, GER
Pierce ECL western blotting substrate	Thermo Fisher Scientific, Waltham, USA
Potassium chloride (KCl)	Carl Roth, Karlsruhe, GER
Propan-2-ol (C <sub>3</sub> H <sub>8</sub> O)	Thermo Fisher Scientific, Loughborough, UK
Proteinase inhibitor cocktail (PIC) tablets, Mini EASY pack	Roche Diagnostics, Mannheim, GER
SDS	Omnilab-Laborzentrum, Bremen, GER
Sodium chloride (NaCl)	AppliChem, Darmstadt, GER
TEMED (C <sub>6</sub> H <sub>16</sub> N <sub>2</sub> )	Carl Roth, Karlsruhe, GER
Tris, ultrapure (C <sub>4</sub> H <sub>11</sub> NO <sub>3</sub> )	AppliChem, Darmstadt, GER
Tris-HCl (C <sub>4</sub> H <sub>11</sub> NO <sub>3</sub> x HCl)	Sigma-Aldrich, Steinheim, GER
Triton-X-100	Omnilab-Laborzentrum, Bremen, GER
Trypan blue solution	Sigma-Aldrich, Steinheim, GER
Vectashield Antifade Mounting Medium with DAPI (H-1200-10)	Vector Laboratories, Burlingame, USA
X-ray developer T32	Calbe Chemie GmbH, Calbe, GER
X-ray fixing solution Superfix 25	Tetenal Europe, Norderstedt, GER
Igepal	Sigma-Aldrich, Steinheim, GER
Vanadyl Ribonucleoside Complex	New England Biolabs, Frankfurt, GER
Nonidet™ P 40 Substitute	Sigma-Aldrich, Steinheim, GER
Laemmli buffer, 4X	Bio-Rad Laboratories, Hercules, USA
MagicMark™ XP Western Protein Standard	Thermo Fisher Scientific, Loughborough, UK
Tri-sodium citrate dihydrate (C <sub>6</sub> H <sub>5</sub> NaO <sub>7</sub> *2 H <sub>2</sub> O)	AppliChem, Darmstadt, GER
Ribo Ruler high range RNA ladder	Thermo Scientific, Waltham, USA
ROTI®Histokitt	Carl Roth, Karlsruhe, GER

ROTI®Liquid barrier marker	Carl Roth, Karlsruhe, GER
ROTI®Histol xylene replacement	Carl Roth, Karlsruhe, GER
2x RNA loading dye	Thermo Scientific, Waltham, USA
Gel loading dye, purple 6X	New England Biolabs, Frankfurt, GER
Low molecular weight DNA ladder	New England Biolabs, Frankfurt, GER
Luria Broth, Base, Miller	Difco BD, Sparks, USA
LB agar, Miller (Luria-Bertani)	Difco BD, Sparks, USA
Ampicillin	Carl Roth, Karlsruhe, GER

## 2.1.5 Cells

Table 9 - Mammalian and bacterial cells

<b>Mammalian cells</b>		
<b>Cell type and animal number</b>	<b>Genotype</b>	<b>Supplier</b>
Porcine kidney fibroblasts (PKF, 912-Hyb)	WT	
Porcine adipose-derived mesenchymal stem cells (PADMSC, 1101)	WT hUCP1 <sup>+/-</sup>	Chair of Livestock Biotechnology, TUM, Germany
Porcine cardiac fibroblasts (PCF, 2185)	WT	
Porcine ear fibroblasts (PEF, 2417, 2026)	AsCas12a(S542R/K607R) <sup>+/-</sup> PAM: TYCV	
Porcine heart endothelial cells (PEC, 2417, 2026)	AsCas12a(S542R/K607R) <sup>+/-</sup> PAM: TYCV	
Porcine adipose-derived mesenchymal stem cells (PADMSC, 2417, 2026)	AsCas12a(S542R/K607R) <sup>+/-</sup> PAM: TYCV	Chair of Livestock Biotechnology, TUM, Germany
Porcine bladder epithelial cells (PBEC, 2417)	AsCas12a(S542R/K607R) <sup>+/-</sup> PAM: TYCV	Germany
Porcine immortalized cardiac fibroblasts (PCFi)	WT, SV40 large-antigen immortalized	
Porcine kidney 15 (PK-15)	ATCC CCL-33	Joachim Denner, Robert Koch Institut, Berlin, GER
<b>Bacterial cells</b>		
<b>Cell type</b>	<b>Genotype</b>	<b>Supplier</b>
E. coli ElectroMAX™ DH10B	F-mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 endA1	Thermo Fisher Scientific, Waltham, USA

araD139Δ(ara, leu) 7697  
galU galK λ-rpsL nupG

## 2.1.6 Buffers

Table 10 - Buffers

<b>Buffer</b>	<b>Component</b>
APS 10 %	0.1 g/ml (NH <sub>4</sub> ) <sub>2</sub> S <sub>2</sub> O <sub>8</sub> in ddH <sub>2</sub> O
FACS wash buffer	500 mg BSA 100 mg NaN <sub>3</sub> PBS, add to 100 ml
Cell lysis buffer	2 ml 1 M tris-HCL pH 8.5 0.2 ml 0.5 M EDTA 0.2 ml 20 % SDS 0.8 ml 5 M NaCl
NP-40 buffer	150 mM NaCl 1.0 % Nonidet™ P-40 50 mM tris-HCL, pH 8.0 in ddH <sub>2</sub> O
Nucleofection buffer pH 7.2	5 mM KCL 10 mM MgCl <sub>2</sub> 70 mM Na <sub>2</sub> HPO <sub>4</sub> 70 mM NaH <sub>2</sub> PO <sub>4</sub> in nuclease-free water sterile filtered 0.22 μm
RLN1	50 mM Tris-HCL pH 8.0 140 mM NaCl 1.5 mM MgCl <sub>2</sub> 0.5 % NP-40 0.5 % Igepal 2mM Vanadyl Ribonucleoside Complex in nuclease-free ddH <sub>2</sub> O
RLN2	50 mM Tris-HCL pH 8.0 500 mM NaCl 1.5 mM MgCl <sub>2</sub> 0.5 % NP-40 0.5 % Igepal 2mM Vanadyl Ribonucleoside Complex in nuclease-free ddH <sub>2</sub> O
Semi-dry blotting buffer	3 g tris 14.4 g glycine 200 ml methanol 1 g SDS ddH <sub>2</sub> O, add to 1 l
Sodium citrate buffer, 10 mM pH 6.0	2.9 g sodium citrate ddH <sub>2</sub> O, add to 1 l adjust pH to 6.0
TAE, 50X	242 g tris 100 ml 0.5M EDTA 57.1 ml acetic acid ddH <sub>2</sub> O, add to 1 l

TBE, 10X	545 g tris 275 g boric acid 39.2 g EDTA ddH <sub>2</sub> O, add to 5 l
TBS, 10X	24.2 g tris 80 g NaCl ddH <sub>2</sub> O, add to 1 l
TBST, 1X	100 ml TBS, 10X 1 ml Tween-20 ddH <sub>2</sub> O, add to 1 l
TE	10 mM Tris-CL, 1 M pH 8.0 1 mM EDTA, 0.5 M in ddH <sub>2</sub> O
Tris-HCL, 0.5 M pH 6.8	15.1 g tris ddH <sub>2</sub> O, add to 125 ml adjust pH to 6.8 using HCL ddH <sub>2</sub> O, add to 250 ml
Tris-HCL, 1 M pH 8.8	39.4 g tris ddH <sub>2</sub> O, add to 125 ml adjust pH to 8.8 using HCL ddH <sub>2</sub> O, add to 250 ml
TTE	242 mg tris 1 ml triton-X-100 584 mg EDTA ddH <sub>2</sub> O, add to 100 ml
Western blot running buffer, 10x, pH 8.3	30 g tris 144 g glycine 10 g SDS ddH <sub>2</sub> O, add to 1 l
Western blot sample loading buffer, 4X	Laemmli buffer, 4X 10 % 2M DTT

## 2.1.7 Tissue culture reagents

Table 11 - Tissue culture reagents and supplements

Name	Supplier
Accutase solution	Sigma-Aldrich, Steinheim, GER
Ala-Gln	Sigma-Aldrich, Steinheim, GER
Amphotericin B	Sigma-Aldrich, Steinheim, GER
Blasticidin S	InvivoGen, San Diego, USA
Cell culture water	Sigma-Aldrich, Steinheim, GER
Collagenase type IA (C2674)	Sigma-Aldrich, Steinheim, GER
Dimethylsulfoxid (DMSO)	Sigma-Aldrich, Steinheim, GER
DMEM/F12	Sigma-Aldrich, Steinheim, GER
Dulbecco's Modified Eagle's Medium high Glucose (DMEM)	Sigma-Aldrich, Steinheim, GER
Endothelial cell growth medium supplementMix	PromoCell GmbH, Heidelberg, Germany
Endothelin-1	Sigma-Aldrich, Steinheim, GER

FCS Superior	Biochrom GmbH, Berlin, GER
G418 sulphate	Genaxxon Bioscience, Ulm, GER
Gelatin 2 % solution from bovine skin cell	Sigma-Aldrich, Steinheim, GER
hEGF	Sigma-Aldrich, Steinheim, GER
Hydrocortisone	Lonza Group AG, Basel, CH
Hygromycin B solution	PanReac AppliChem, Darmstadt, GER
Lipofectamine 2000	Thermo Fisher Scientific, Waltham, USA
MEM non-essential amino acid Solution, x100	Sigma-Aldrich, Steinheim, GER
Opti-MEM reduced serum medium	Life technologies, Carlsbad, USA
Penicillin/Streptomycin	Sigma-Aldrich, Steinheim, GER
Phosphate-buffered saline (PBS)	Sigma-Aldrich, Steinheim, GER
Puromycin	InvivoGen, San Diego, USA
Sodium pyruvate solution	Sigma-Aldrich, Steinheim, GER
Trypan blue stain	Life Technologies, Paisley, UK

## 2.1.8 Tissue culture media

Table 12 - Tissue culture media

Medium	Components	Cells
Bladder cell medium, culturing	45 ml DMEM/F12 5 ml FCS 10 ng/ml hEGF 4 ng/ml hydrocortisone	PBC
Bladder cell medium, isolation	45 ml DMEM/F12 5 ml FCS 100 ng/ml hEGF 4 ng/ml hydrocortisone 500 µl pen-strep	PBC
Endothelial cell medium	100 ml DMEM 10 ml FCS 400 µl Endothelial cell growth medium supplementMix optional: 1 ml pen-strep	PEC
Fibroblast growth medium	500 ml DMEM 50 ml FCS 6 ml NEAA 6 ml sodium pyruvate 6 ml Ala-Gln optional: 6 ml pen-strep	PKF PEF PCF PADMSC
Freezing medium	70 ml FCS 20 ml DMEM 10 ml DMSO	All cells

## 2.1.9 Kits

Table 13 - Kits

<b>Name</b>	<b>Supplier</b>
GenElute mammalian genomic DNA miniprep Kit	Sigma-Aldrich, Steinheim, GER
Mix2Seq kit	Eurofins, Ebersberg, GER
NucleoBond® Xtra midi kit	Macherey-Nagel, Düren, GER
CloneJET PCR Cloning Kit	Thermo Fisher Scientific, Waltham, USA
FirstChoice® RLM-RACE kit	Invitrogen, Karlsruhe, GER
iBind™ Solution Kit	
Monarch DNA Gel Extraction Kit	New England Biolabs GmbH, Frankfurt am Main, GER
Monarch PCR & DNA Cleanup Kit	
Monarch Total RNA Miniprep Kit	
LunaScript® RT Master Mix Kit (Primer-free)	
NEBuilder® HiFi DNA assembly	Promega, Mannheim, GER
PureYield™ Plasmid Miniprep System	
Nano-Glo® dual luciferase assay	Zymo Research Europe GmbH, Freiburg, GER
EZ DNA Methylation-Direct KIT	
PyroMark PCR Kit	
VECTASTAIN elite ABC-HRP kit (peroxidase, standard)	Vector Laboratories, Burlingame, USA

## 2.1.10 Consumables

Table 14 - Consumables

<b>Name</b>	<b>Supplier</b>
Electroporation cuvettes, 2 mm	Peqlab Biotechnologie, Erlangen, GER
Electroporation cuvettes, 4 mm	
Western blot membrane	Biorad Laboratories, Munich, GER
Filter paper (extra thick blot paper)	
MicroAmp fast optical 96-Well reaction plates	Life Technologies, Carlsbad, USA
PVDF membrane 'Roti-PVDF' (0.45 µm)	Carl Roth, Karlsruhe, GER
X-Ray film 'Cronex 5'	Agfa Healthcare, Mortsels, BEL

### 2.1.11 Enzymes and enzyme buffers

Table 15 - Enzymes and enzyme buffers

<b>Name</b>	<b>Supplier</b>
Calf intestinal phosphatase (CIP)	New England Biolabs GmbH, Frankfurt am Main, GER
ddPCR supermix for probes (no dUTP) 2X	Bio-Rad Laboratories, Hercules, USA
DNA polymerase I large (Klenow) fragment	New England Biolabs GmbH, Frankfurt am Main, GER
GoTaq G2 DNA polymerase	Promega, Mannheim, GER
5x Green GoTaq reaction buffer	
Proteinase K (20 mg/ml)	Sigma-Aldrich, Steinheim, GER
Q5 high-fidelity DNA polymerase	New England Biolabs GmbH, Frankfurt am Main, GER
5x Q5 reaction buffer	
Restriction enzymes	
Restriction buffers 1.1, 2.1, 3.1, CutSmart	
RNase A	Sigma-Aldrich, Steinheim, GER
T4 DNA ligase	New England Biolabs GmbH, Frankfurt am Main, GER
T4 DNA ligase buffer 10X	New England Biolabs GmbH, Frankfurt am Main, GER
2x qPCRBio SyGreen Mix Lo-ROX	PCR Biosystems Ltd., London, UK

### 2.1.12 Laboratory equipment

Table 16 - Equipment

<b>Equipment</b>	<b>Supplier</b>
Attune NxT flow cytometer	Thermo Fisher Scientific, Waltham, USA
Automated cell counter 'Countess'	Invitrogen, Carlsbad, USA
Blue light table	Serva, Heidelberg, GER
BTX® ECM 630 electroporation system	BTX, Holliston, USA
Centrifuges 'Sigma 3-16', 'Sigma 1-15K', 'Sigma 1-15', 'Sigma 4K15'	Sigma, Osterode, GER
CO2 incubator 'Forma Steri-Cycle 371'	Thermo Fisher Scientific, Waltham, USA
Echo Revolve R4 microscope	Echo, San Diego, USA

Electrophoresis system (buffer chamber, gel trays, combs)	Peqlab Biotechnologie, Erlangen, GER
Eporator®	Eppendorf, Hamburg, Ger
Gel documentation imaging system 'Quantum ST5'	Vilber Lourmat, Eberhardzell, GER
iBind™ Western Device	Thermo Fisher Scientific, Waltham, USA
LUMIstar Omega Microplate Luminometer	BMG Labtech, Ortenberg, GER
Microscope 'Axiovert 40CLF', 'Axiovert 200M', 'Primo Star'	Carl Zeiss, Jena, GER
NanoDrop light spectrophotometer	Thermo Fisher Scientific, Waltham, USA
PCR cycler 'peqStar 2x'	Peqlab Biotechnologie, Erlangen, GER
PyroMark Q48 Autoprep	Qiagen, Hilden, GER
QuantStudio™ 5 Real-Time PCR Instrument	Thermo Fisher Scientific, Waltham, USA
QX200 Droplet Digital PCR System	Bio-Rad Laboratories, Hercules, USA
Rotary microtome 'Microm HM355'	Microm International, Walldorf, GER
Safety cabinet 'HERAsafe HS 12'	Kendro Laboratory Products, Hanau, GER
Western blot 'Mini PROTEAN tetra handcast system'	Biorad Laboratories, Munich, GER

### 2.1.13 Software

Table 17 - Software

<b>Name</b>	<b>Producer and/or link</b>
BLAST, Basic Local Search Tool	<a href="https://blast.ncbi.nlm.nih.gov">https://blast.ncbi.nlm.nih.gov</a>
Benchling, designing and analyzing nucleic acids	Benchling, San Francisco, USA <a href="https://www.benchling.com">https://www.benchling.com</a>
BioRender	Biorender.com
Citavi 6	Swiss Academic Software GmbH, Wädenswil, CH
CRISPOR, CRISPR design tool	<a href="http://crispor.tefor.net">http://crispor.tefor.net</a>
Ensembl genome browser	<a href="https://www.ensembl.org">https://www.ensembl.org</a>
FinchTV, Chromatogram viewer	Geospiza Inc., Seattle, USA <a href="https://digitalworldbiology.com/FinchTV">https://digitalworldbiology.com/FinchTV</a>
FlowJo™, flow cytometry software	BD, Ashland, USA
GraphPad, statistics tool	GraphPad Software, San Diego, USA
ICE, CRISPR analysis tool	<a href="https://www.synthego.com/products/bioinformatics/crispr-analysis">https://www.synthego.com/products/bioinformatics/crispr-analysis</a>
JASPAR, transcription factor binding site prediction	<a href="https://jaspar.genereg.net/">https://jaspar.genereg.net/</a>
Microsoft Office	Microsoft Corporation, Redmond, USA
Multi-user Reader Control and MARS Data Analysis Software	BMG Labtech, Ortenberg, GER
NCBI genome browser	<a href="https://www.ncbi.nlm.nih.gov/gene">https://www.ncbi.nlm.nih.gov/gene</a>

NCBI open reading frame finder	<a href="https://www.ncbi.nlm.nih.gov/orffinder/">https://www.ncbi.nlm.nih.gov/orffinder/</a>
Neural Network Promoter Prediction	<a href="https://www.fruitfly.org/seq_tools/promoter.html">https://www.fruitfly.org/seq_tools/promoter.html</a>
Primer3, primer design tool	<a href="https://primer3.ut.ee/">https://primer3.ut.ee/</a>
PyroMark Q48 Autoprep Instrument Software Version	Qiagen, Hilden, GER
QuantaSoft, ddPCR software	Bio-Rad Laboratories, Hercules, USA
QuantStudio 5 software,	Thermo Fisher Scientific, Waltham, USA
Quantum ST5, gel documentation software	Vilber Lourmat, Eberhardzell, GER
RNAhybrid	<a href="https://bibiserv.cebitec.uni-bielefeld.de/rnahybrid">https://bibiserv.cebitec.uni-bielefeld.de/rnahybrid</a> , Rehmsmeier, Marc and Steffen, Peter and Hoechsmann, Matthias and Giegerich, Robert Fast and effective prediction of microRNA/target duplexes RNA, RNA, 2004

## **2.2 Methods**

### **2.2.1 Molecular biology**

#### **2.2.1.1 Isolation of genomic DNA**

Genomic DNA from tissue was isolated using the GenElute™ mammalian genomic DNA miniprep kit according to the manufacturer's instructions.

For gDNA derived from cultured cells, half of the cells of a confluent 6-well were detached and pelleted at 300 rcf for 5 min. After aspiration of the supernatant, the pellet was resuspended in 30 µl QuickExtract® DNA extraction solution. Then, the solution was incubated at 68 °C for 15 min, followed by 98 °C for 8 min.

For highly pure gDNA, phenol/chloroform extraction was performed. Therefore, cells of a confluent T150 flask were pelleted as written above. Then, the pellet was washed with PBS once and resuspended in 500 µl lysis buffer containing 30 µl proteinase K (20 mg/ml). After ON incubation at 37 °C, 1 µl RNase A (20 mg/ml) was added and incubated at 25 °C for 5 min. To separate the DNA from other components, 500 µl of phenol-chloroform-isoamyl alcohol (25:24:1) was added, vortexed, and incubated for 10 min at RT followed by 15 min centrifugation at 10000 rcf. The DNA containing the upper water phase was then transferred in a new reaction tube, and again 500 µl phenol-chloroform-isoamyl alcohol was added and centrifuged for 10 min at 10000 rcf. Then, the upper 350 µl of the liquid phase were mixed with an equal volume of isopropanol to precipitate the DNA. Afterwards, the DNA was pelleted for 2 min at 10000 rcf, and the supernatant was discarded. Finally, the gDNA was dried for 10 min and resolved in 100 µl TE buffer.

#### **2.2.1.2 Plasmid preparation**

For smaller DNA amounts up to 15 µg, the PureYield™ miniprep system was used according to the manufacturer's instructions, using 4 ml of an overnight bacterial culture.

To isolate larger amounts of DNA of up to 300-400 µg from a 100 ml ON bacterial culture, the NucleoBond® Xtra Midi Kit was used according to the manufacturer. Both variants of isolated plasmid DNA provide a highly pure vector suitable for transfection.

#### **2.2.1.3 Purification of DNA**

DNA derived from digests or PCRs, PCR amplicons, ligation approaches, and other DNA, which required the removal of other components, were purified using the Monarch PCR & DNA cleanup kit. The Monarch DNA gel extraction kit was used if the DNA was derived from agarose gels. Either was performed following the manufacturer's recommendation.

#### **2.2.1.4 RNA isolation**

For RNA isolation, up to  $6 \times 10^6$  cultured cells were detached and pelleted for 5 min at 300 rcf. The pellet was either stored at  $-80\text{ }^{\circ}\text{C}$  until processing or used directly for isolation via the Monarch total RNA miniprep kit according to the manufacturer's protocol.

To isolate porcine tissue-derived RNA, a tissue homogenization step had to be done before the extraction. Therefore, 10 mg of the respective tissue was cut in liquid nitrogen and transferred into the recommended SpeedMill lysis tubes containing homogenization beads and 350  $\mu\text{l}$  lysis buffer of the Monarch total RNA miniprep kit. The tissue was then homogenized twice for 30 s followed by 30 s cooling time using the SpeedMill Plus homogenizer. Afterwards, the protocol was continued similarly to the isolation from cultured cells.

#### **2.2.1.5 RNA fractioning assay**

For RNA fractionation, cultured cells were detached, washed once with PBS, and pelleted at 300 rcf for 5 min. The cytoplasmic membrane was then lysed by resuspending the pellet in 175  $\mu\text{l}$  cold RLN1 per  $10^6$  cells, followed by incubation for 5 min on ice. After centrifuging again for 2 min at 300 rcf at  $4\text{ }^{\circ}\text{C}$ , the cytoplasmic RNA containing supernatant was moved to a fresh tube and stored on ice. To lyse the remaining pellet containing the nuclear RNA, the pellet was resuspended and incubated in 175  $\mu\text{l}$  cold RLN2 buffer per  $10^6$  cells for 5 min on ice. Another centrifugation step for 2 min at 16000 rcf and  $4\text{ }^{\circ}\text{C}$  was performed to separate the nuclear RNA from the nuclear debris. Finally, the nuclear RNA supernatant was transferred to a new tube and stored at  $-80\text{ }^{\circ}\text{C}$ .

#### **2.2.1.6 Determination of DNA and RNA concentration**

According to the manufacturer's instructions, DNA and RNA concentrations were determined via the NanoDrop light spectrophotometer.

#### **2.2.1.7 cDNA synthesis**

For cDNA synthesis, 400 ng of RNA were appointed using the LunaScript® RT master mix kit (primer-free) according to the manufacturer's instructions. The cDNA was diluted for downstream applications by adding 80  $\mu\text{l}$  of nuclease-free  $\text{H}_2\text{O}$ .

#### **2.2.1.8 Molecular cloning**

For restriction digests, 1-5  $\mu\text{g}$  of DNA and 3 U enzyme per  $\mu\text{g}$  DNA were appointed and incubated for 1.5 h. Temperature and buffer depended on the restriction enzyme used.

If restriction enzymes produced incompatible sticky ends for ligation, blunting the overhangs was performed using a DNA polymerase I large (Klenow) fragment. Therefore, the restriction digest was supplemented with 33  $\mu\text{M}$  of each dNTP and 1 U Klenow per  $\mu\text{g}$  DNA. After incubation for 15 min at 25  $^{\circ}\text{C}$ , the reaction was stopped by adding EDTA to a final concentration of 10 mM and heating to 75  $^{\circ}\text{C}$  for 20 min.

Ligation of classic sticky- and blunt-end cloning was performed using T4 DNA ligase according to the manufacturer's instructions. Therefore, 100 ng of vector DNA in a molar ratio of 3:1 to insert DNA were appointed. The ligation was incubated for ON at 16  $^{\circ}\text{C}$ .

Whenever PCR amplification of the insert was possible, NEBuilder® HiFi DNA assembly was used for error-free and facilitated cloning according to the manufacturer's protocol. Primers for amplification were designed with 15 bp homology to the backbone and insert, adding up to 30 bp primer length. 100 ng of vector DNA in a molar ratio of 2:1 to the insert DNA were appointed.

All reactions were purified as described in 2.2.1.3 before transformation into bacteria. All enzymes were purchased from New England Biolabs (Frankfurt am Main, Germany).

If PCR products were generated only in low abundance or for facilitated handling, PCR products were first subcloned into the pJet1.2 vector using the CloneJET™ PCR cloning kit according to manufacturer's instructions.

#### **2.2.1.9 Agarose gel electrophoresis**

In order to isolate desired fragments of a restriction digest or for analytical purposes, DNA fragments or PCR amplicons were separated via gel electrophoresis. For smaller DNA fragments up to 700 bp, 2 % agarose TBE gels were generated. For larger DNA fragments, 1 % agarose TBE gels were poured. If the sample was further processed, TAE was used instead of TBE. 4  $\mu\text{l}$ /100 ml PeqGreen dye was added to all gels to visualize nucleic acids under UV light. After loading the samples, 80-140 V were applied to the gel chambers according to their size for up to 2 h. Images were taken using the Quantum ST5 gel documentation imaging system.

To check the integrity of RNA before RNA-seq, RNA gels were used. Therefore, 0.8 % agarose TBE gels containing 800  $\mu\text{l}$  37 % formaldehyde were poured. Cold buffer was used to prevent RNA degradation during gel electrophoresis, and the chamber was put into ice. 500 g of RNA sample were mixed with 2X RNA loading dye and incubated at 70  $^{\circ}\text{C}$  for 10 min followed by a 5 min 4  $^{\circ}\text{C}$  incubation before loading directly on the gel. Then, the gel was applied to 80 V for 50 min. Two separated bands indicate 18S rRNA and 28S rRNA, and high-quality RNA.

### 2.2.1.10 Methylation assay

In order to identify differentially methylated regions in the porcine *MEG3* locus, a methylation assay, including a bisulfite DNA conversion followed by pyrosequencing, was performed. Therefore, 200 ng of sample DNA were first converted into bisulfite DNA using the EZ DNA Methylation-Direct KIT according to manufactures' instructions. Thereby, the unmethylated cytosines are converted into uracils. The region of interest was then amplified from the bisulfite DNA using PyroMark PCR (Table 20), with a biotinylated reverse primer to achieve a biotinylated product. Pyrosequencing, where the uracils were recognized as thymidines, was then performed using the PyroMark Q48 Autoprep system according to the manufacturer's instructions. Analysis was performed using the PyroMark Q48 Autoprep instrument software.

### 2.2.1.11 PCR

PCR was performed to amplify desired DNA from cDNA, gDNA, or plasmid templates. Depending on amplicon length, further downstream application, and template origin, either proofreading or non-proofreading polymerases were used. For short amplicons up to several hundred bp and screening purposes, GoTaq® DNA polymerase without proofreading was used. For longer amplicons, cloning, and further downstream processing like Sanger sequencing, the high-fidelity DNA polymerase Q5® with proofreading was used. If pyrosequencing was performed after amplification, PyroMark polymerase was used. 100 ng of template was used for amplification from gDNA, while for amplification from purified plasmid DNA or cDNA, 40 ng were appointed. For bisulfite-converted DNA in PyroMark PCR, 10-20 ng were appointed. In Table 18, Table 19, and Table 20, the according reaction conditions are depicted.

Table 18 - GoTaq® reaction components and conditions

Composition		GoTaq® DNA polymerase			
Component	Final concentration	Step	Temperature	Time	Cycles
5x Green GoTaq® reaction buffer	1X (1.5 mM MgCL <sub>2</sub> )	Initial denaturation	95 °C	2 min	1
dNTPs	0.2 mM each	Denaturation	95 °C	30 sec	35
Forward primer	0.2 µM	Annealing	60 °C	30 sec	35
Reverse primer	0.2 µM	Extension	72 °C	1 min/kb	35
GoTaq® DNA polymerase	0.75 U	Final extension	72 °C	5 min	1
Template DNA	40-100 ng	Storage	8 °C	∞	1
ddH <sub>2</sub> O to	Add to 25 µl				

Table 19 - Q5® reaction components and conditions

Q5® high-fidelity DNA polymerase					
Composition		Cycling conditions			
Component	Final concentration	Step	Temperature	Time	Cycles
5x Q5® reaction buffer	1X	Initial denaturation	98 °C	30 sec	1
dNTPs	0.2 mM each	denaturation	98 °C	10 sec	35
Forward primer	0.5 µM	Annealing	60 °C	30 sec	35
Reverse primer	0.5 µM	Extension	72 °C	30 sec/kb	35
Q5® high-fidelity DNA polymerase	0.01 U	Final extension	72 °C	2 min	1
Template DNA	40-100 ng	Storage	8 °C	∞	1
ddH <sub>2</sub> O to	Add to 25 µl				

Table 20 - PyroMark reaction components and conditions

PyroMark polymerase					
Composition		Cycling conditions			
Component	Final concentration	Step	Temperature	Time	Cycles
2X PyroMark PCR Master Mix	1X	Initial PCR activation	95 °C	15 min	1
10X CoraLoad concentrate	1X	Denaturation	94 °C	30 sec	45
Forward primer	0.2 µM	Annealing	56 °C	30 sec	45
Reverse primer	0.2 µM	Extension	72 °C	30 sec	45
Template DNA	10-20 ng bisulfite converted DNA	Final extension	72 °C	10 min	1
ddH <sub>2</sub> O to	Add to 25 µl	Storage	8 °C	∞	1

### 2.2.1.12 Droplet digital PCR

Droplet digital PCR (ddPCR) was performed to verify the correct transgene copy number. ddPCR employs fractioning of single DNA molecules in nanosized droplets. Thereby, PCR amplification happens in each droplet at the same time. Droplets positive for target DNA are normalized to the

housekeeping gene (GAPDH) of known copy number. TaqMan probes and primers were designed and labeled according to the manufacturer's protocol. To ensure that only one copy is amplified per droplet, sample DNA was first digested using HindIII (see 2.2.1.8) or any other restriction enzyme, which does not cut inside the amplicon and separates possible transgene arrays. After DNA digestion, ddPCR was setup according to Table 21. Subsequently, 70 µl of droplet generator oil and 20 µl of the TaqMan PCR setup were pipetted into the appropriate wells of the droplet generation cartridge. Droplet generation was performed using the QX200 droplet generator. The droplets were then transferred into a 96-well plate and sealed using the PX1™ PCR plate sealer and the aluminum foil provided by the manufacturer. Finally, after PCR was performed according to Table 21, the droplets were counted and analyzed via the QX200 droplet reader and the QuantaSoft software.

Table 21 - ddPCR components and conditions

Composition		Cycling conditions				
Component	Final concentration	Step	Temp.	Time	Ramp	Cycles
100 ng of digested DNA	1X	Initial enzyme activation	95 °C	10 min		1
2X ddPCR supermix for probes (no UTP)	1X	Denaturation	94 °C	30 sec	2 °C	40
20X FAM-labeled target probe + primers	250 nM probe 900 nM primers	Annealing/ extension	60 °C	1 min	/ sec	40
20X HEX-labelled reference probe +primers (GAPDH)	250 nM probe 900 nM primers	Enzyme inactivation	98 °C	10 min		1
H <sub>2</sub> O	Add to 23 µl	Storage	4°C	∞		1

### 2.2.1.13 3' Rapid amplification of cDNA ends (RACE)

3' RACE PCR was performed to identify 3' ends of the porcine *MEG3* transcripts. Therefore, the FirstChoice® RLM-RACE Kit was used according to the manufacturer's protocol. Successful products of 3' RACE PCR were sent for Sanger sequencing (see 2.2.1.14)

### 2.2.1.14 Sanger sequencing

Eurofins Genomics (Ebersberg, GER) performed the Sanger sequencing. All samples were prepared according to the Mix2Seq kit instructions.

### 2.2.1.15 Determination of INDEL efficiency

In order to determine the INDEL efficiency of the CRISPR/Cas9 or Cas12a system, a PCR was performed across the cut site, with at least 150 bp distance to it in both directions. Amplicons of both groups, targeted and wild-type control, were then sent for Sanger sequencing. The sequencing results were uploaded as .abi1 files to online tool interference of CRISPR edits (ICE, <https://ice.synthego.com>). The algorithm then shows the efficiency and the spectrum of INDELS detected. Furthermore, it shows the probability of causing a frameshift mutation and the R<sup>2</sup> Pearson correlation coefficient as a quality control metric. Only R<sup>2</sup> < 0.8 were considered reliable.

### 2.2.1.16 Quantitative real-time PCR (qPCR)

qPCR was performed using the qPCRBIO SyGreen Lo-ROX mix and the QuantStudio 5 system. The settings were set for 'SybrGreen' and 'Fast'. In Table 22, the qPCR components and cycling conditions are depicted. Each sample was measured in technical triplicates, and the average CT value was normalized to the housekeeping gene glyceraldehyde 3-phosphate dehydrogenase (GAPDH). For RNA derived from an RNA fractioning assay, the RNA was normalized to 18srRNA. The fold change was calculated using the  $\Delta\Delta CT$  method.

Table 22 - qPCR components and cycling conditions

Composition		Cycling conditions			
Component	Final concentration	Step	Temperature	Time	Cycles
2X qPCRBIO SyGreen mix	1X	Polymerase activation	95 °C	2 min	1
Template DNA	<100 ng	Denaturation	95 °C	1 sec	40
Forward primer	400 nM	Annealing/Extension	60 °C	20 sec	40
Reverse primer	400 nM				
ddH <sub>2</sub> O to	Add to 10 $\mu$ l				

### **2.2.1.17 Protein isolation**

To isolate western blot protein, a lysis buffer was prepared by freshly adding a proteinase inhibitor cocktail to cold NP-40 buffer to a final concentration of 1X.

For tissue protein, a 5 mg piece of tissue is homogenized in 300 µl of cold lysis buffer in a homogenization tube, and the tissue is disrupted as described in 2.2.1.4 and put on ice for 20 min. Subsequently, the sample is again homogenized and centrifuged for 10 min at 4 °C at 16000 rcf.

For protein from cultured cells, cells are detached, pelleted for 5 min at 300 rcf, and resuspended in 100 µl per 10<sup>6</sup> cells cold lysis buffer. Lysis is performed for 20 min at 4 °C followed by centrifugation for 10 min at 4 °C at 16000 rcf.

Finally, after lysis, the protein-containing supernatant is transferred into a fresh tube and stored at -80 °C.

For fractioned nuclear and cytoplasmic protein isolation, the NE-PER nuclear and cytoplasmic extraction reagents were used according to the manufacturer's instructions. Protein concentration was then determined using the advanced protein assay reagent. 2 µl of each protein extract was mixed with 198 µl 1X advanced protein assay reagent in a 96-well plate. After thorough mixing and 5 min incubation at RT, absorption was measured at 595 nm in triplicates using the FLUOstar Omega microplate reader. Protein concentration was then calculated by  $OD_{\text{Sample-Blank}} * 7.5 \mu\text{g}/\mu\text{l}$ .

### **2.2.1.18 Western blot**

Western blot was performed to verify and detect transgene protein production in different tissues and cells. In order to enable antibody-mediated protein detection, the proteins were first denatured. Therefore, 20-40 µg of each protein sample were mixed 3:1 with 4X Laemmli sample buffer containing 10 % SDS and incubated for 5 min at 95 °C. The samples were then loaded with the MagicMark™ XP western protein standard on a polyacrylamide gel, consisting of a 4 % collection gel and a 10 % separation gel. The gels were poured according to Table 23. Gel electrophoresis was then performed in 1X running buffer for 30 min at 80 V followed by 70 min at 140 V. According to the manufacturer's instructions, the separated proteins were transferred on a PVDF membrane in semi-dry blotting buffer using the mini Trans-Blot® cell. Blotting was run for 150 min at 80 V. Unspecific binding was then blocked for 1 h in 1X TBST containing 5 % milk, followed by 10 min washing in 1X TBST. Primary and secondary antibody binding was performed using the iBind™ western device according to the manufacturer's recommendation. Subsequently, the membrane was washed thrice in TBST for 10 min each. The chemiluminescence reaction, catalyzed by the horseradish peroxidase bound to the secondary antibody, was induced

by delivering the Pierce™ ECL western blotting substrate to the membrane. Finally, the membrane was exposed to an X-ray film for 5 sec to 15 min, developed, and fixed.

Table 23 - Composition of western blot separation- and collection gels

<b>Component</b>	<b>Volume for a 10 % separation gel</b>	<b>Volume for a 4 % collection gel</b>
H <sub>2</sub> O	2840 µl	2800 µl
40 % polyacrylamide (29:1)	2000 µl	400 µl
1 M Tris-Cl pH 8.8	3000 µl	/
0.5 M Tris-Cl pH 6.8	/	1000 µl
10 % SDS	80 µl	40 µl
10 % APS	80 µl	40 µl
TEMED	3.2 µl	4 µl

### 2.2.1.19 Immunohistochemistry

The protein production of Cas12a transgenic pigs was detected in various tissues by immunohistochemistry. Freshly isolated tissue was fixed with 4 % formaldehyde for at least 24 h and stored in 80 % ethanol until paraffin embedding. Embedded tissue was then sliced into 4 µm thick sections and let dry at RT for at least 24 h. Subsequently, the paraffin was melted for 10-15 min at 55-60 °C in a hybridization oven. To remove the paraffin and rehydrate the tissue, the sections are incubated twice for 10 min in ROTI®Histol, followed by incubation in 100 % ethanol twice for 5 min. The slides are then put for 5 min in 95 % ethanol, followed by a 5 min incubation in 70 % ethanol and two final incubations for 2 min in dH<sub>2</sub>O. For antigen unmasking, the slides were heated in 10 mM pH 6.0 sodium citrate buffer for 23 min at sub-boiling temperature and cooled down to RT for at least 30 min. The slides were washed twice for 5 min in dH<sub>2</sub>O and transferred for 15 min in 3 % H<sub>2</sub>O<sub>2</sub> without light exposure to saturating endogenous peroxidases. In the next step, the sections were washed for 5 min in dH<sub>2</sub>O and permeabilized for 10-15 min in PBS containing 0.4 % Triton X-100 and 0.1 % Tween-20, followed by washing for 5 min in PBS and for 5 min in dH<sub>2</sub>O. To prevent unspecific antibody-binding, the tissue was surrounded using the ROTI®Liquid barrier marker, and a blocking solution was added, consisting of 2 % goat serum and 1 % BSA in PBS. After incubation for 1 h at RT, the blocking solution was removed by washing for 5 min in PBS. The primary antibody diluted in blocking solution was then applied ON at 4 °C

and washed away with PBS twice for 5 min the next day. The HRP-coupled secondary antibody diluted in blocking solution was then incubated for 1 h at RT. Subsequently, the slides were washed twice for 5 min in PBS and then incubated for 25 min in the VECTASTAIN® ABC reagent, followed by washing in dH<sub>2</sub>O twice for 5 min. To visualize the location of the HRP, the DAB-enhanced liquid system was applied to the sections. After a few seconds of incubation, depending on antibody dilution and tissue, the DAB reaction was stopped by transferring the sections into dH<sub>2</sub>O. The slides were then counterstained for 30 sec in hematoxylin and washed for 5 min under tap water. Finally, the sections were dehydrated. Therefore, they were incubated in 70 % ethanol first, then in 90 % ethanol, followed by 100 % ethanol, each for 2 min. Finally, the sections were incubated twice for 5 min in ROTI®Histol and mounted with the Roti-Histokitt. Images were taken using the Echo Revolve R4 microscope.

#### **2.2.1.20 RNA sequencing**

Dr. Rupert Öllinger did the RNA sequencing at the TranslaTUM, TUM's sequencing core unit. Dr. Thomas Engleitner and Riccardo Trozzo performed the analysis. Poly(A)-RNA bulk sequencing library preparation was carried out as previously described.<sup>135</sup>

#### **2.2.1.21 Generation of CRISPR/Cas9/Cas12a components**

Promising genomic target sites were identified using CRISPOR or Benchling (Table 17). *Sus scrofa* 11.1 was selected as the reference genome. TYCV as PAM sequence for AsCas12a, NGG as PAM sequence for SpCas9. The sgRNAs with suitable overhangs were synthesized as oligonucleotides and cloned into appropriate Bbs-I digested vectors. Correct clones were identified as described in the following chapter, 2.2.2.

### **2.2.2 Microbiological methods for vector amplification**

All transformations were performed via electroporation of the electrocompetent *E.coli* strain DH10B. Furthermore, only vectors containing an ampicillin resistance gene were introduced. 50 µl of thawed DH10B were mixed with 10 ng of ligation approach or plasmid DNA and transferred to a cold electroporation cuvette with 2 mm electrode distance. Electroporation was performed at 2500 V for 5 ms using the Eporator®.

Immediately afterward, the bacterial suspension was incubated in a reaction tube containing 500 µl of LB medium at 37 °C for 30 min. Subsequently, the bacteria were plated on LB agar ampicillin-100 plates and incubated at 37 °C ON. Single colonies were then picked using a sterile toothpick, streaked, and cultured on a new agar plate.

Moreover, the toothpick was dipped into 30 µl of TTE buffer for DNA isolation. Subsequently, the bacteria were lysed at 95 °C for 5 min, and 2 µl of the mixture was used for colony PCR, with an amplicon spanning over the backbone and the insert, according to Table 18. Clones positive in colony PCR were inoculated in 5 ml 0.1 % ampicillin-LB medium and cultured ON at 37 °C and 220 rpm. Then, mini prep (2.2.1.2) and a subsequent Sanger sequencing (2.2.1.13) were performed for plasmid verification.

Finally, cell clones harboring the desired vector were inoculated in 100 ml 0.1 % ampicillin-LB medium and cultured ON at 37 °C and 220 rpm for midi prep (2.2.1.2). For future inoculations, 1 ml of the ON culture was mixed with 0.5 ml glycerol and stored at -80 °C.

## **2.2.3 Cell culture**

### **2.2.3.1 Isolation and cultivation of porcine cells**

All porcine cells were isolated from German Landrace pigs obtained from the TUM animal facility Thalhausen. To isolate fibroblasts derived from ear tissue (PEF) or kidney (PKF), a clean 1 cm<sup>3</sup> tissue piece was cleared from blood vessels, connective tissue, and fat with a scalpel. The tissue was then minced and digested for 30 min at 37 °C, stirring in 10 ml 1 mg/ml collagenase type 1A. Digest was then stopped with 10 ml fibroblast medium, and cells were pelleted for 5 min at 300 rcf. The pellet was then resuspended in 1 % pen-strep medium (Table 12) and distributed to two T25 flasks for PEF and two T150 flasks for PKF. The medium was exchanged daily during the first week and finally changed to an antibiotic-free medium seven days post-isolation.

For fibroblasts derived from cardiac tissue (PCF), 1 cm<sup>3</sup> of porcine heart tissue was washed with cold PBS and minced in 4 ml fibroblast medium to tissue pieces < 1 mm<sup>3</sup>. The suspension was then transferred into a 50 ml centrifuge tube, and the plate was rinsed twice with 5 ml of PBS, which was also transferred to the tube. The suspension was then centrifuged for 2 min at 200 rcf, and the supernatant was aspirated. The cells were then resuspended in 5 ml cold PBS and centrifuged with identical settings. Finally, the tissue pieces were resuspended in 24 ml warm 1 % pen-strep medium and distributed evenly to four 10 cm cell culture dishes coated with 0.1 % gelatin. After one week at 37 °C in an incubator without moving the plates, the medium was changed to an antibiotic-free fibroblast medium.

To isolate porcine heart endothelial cells (PEC), a freshly isolated heart was perfused with 500 ml warm PBS. Then, the upper third of the heart was cut off, and the heart chambers were spread open with clamps, with the heart body remaining in a beaker containing warm PBS. The heart chambers were then filled with 1 mg/ml collagenase type 1A and incubated for 20 min at 37 °C. Subsequently, the chambers were rinsed a few times with the used collagenase and one volume

endothelial cell medium (Table 12) and transferred into a 50 ml centrifuge tube. Once the cells were pelleted for 5 min at 300 rcf, they were resuspended in endothelial cell medium containing 1 % pen-strep and seeded on a T25 tissue flask.

Porcine adipose-derived mesenchymal stem cells (PADMSC) were isolated from neck fat tissue. Therefore, 6 g of fat, free of skin and blood vessels, was minced in 10 ml 1 mg/ml collagenase type 1A and stirred for 20 min at 37 °C. The cell suspension was then filtered through a 100 µm strainer to remove fat clumps and mixed with one volume of fibroblast medium. The cells were then centrifuged for 10 min at 1000 rcf, and the supernatant was disposed. Subsequently, the pellet was resuspended in 1 % pen-strep fibroblast medium and distributed evenly to two T75 tissue flasks. The next day, the flasks were washed twice with PBS, and the PADMSCs were cultured for 2 days using a medium containing antibiotics until it was changed to an antibiotic-free medium.

All primary cells were checked for mycoplasma contamination, two days after switching to antibiotic-free conditions. Furthermore, in a humidified incubator, all cells were cultured at 37 °C, 5 % CO<sub>2</sub>. Once reaching 90 % confluence, all cells were passaged by washing with PBS, followed by adding Accutase® cell detachment solution and incubating for 10 min at 37 °C. The reaction was stopped by adding medium, and cells were transferred to an appropriate culture vessel.

### **2.2.3.2 Cryopreservation**

For long-time storage, cells were detached and pelleted for 5 min at 300 rcf. The pellet was then resuspended in 1 ml freezing medium, transferred to a cryopreservation vial, and frozen at 1 °C per min with a freezing container to -80 °C. The vials were either stored at -80 °C or in liquid nitrogen. Thawing of the cells was performed quickly at 37 °C. Once all ice was melted, the cells were immediately diluted in medium and seeded on an appropriate tissue flask.

### **2.2.3.3 Transfection**

Transfection with DNA was either performed by lipofection or nucleofection.

For nucleofection, 10<sup>6</sup> cells were detached and pelleted for 5 min at 300 rcf. The pellet was then resuspended in 100 µl nucleofection buffer and mixed with 3-5 µg of DNA for transfection. The mixture was then transferred to a 4 mm electrode distance cuvette and inserted into the BTX® ECM 630 electroporation system. The electroporation was performed at 300 V, at a pulse length of 1 ms, and at a pulse interval of 100 ms for three pulses. Cells were then plated on a T25, and the medium was changed the day after.

For lipofection, cells were seeded to reach a 60-80 % density the next day. Then, the medium was changed to Opti-MEM™ reduced-serum media. The lipofection was then performed using

Lipofectamine™ 2000 according to the manufacturer's instructions. Thereby, 1 µg of DNA and 2 µl of Lipofectamine™ 2000 were appointed for a 6-well and 3-5 µg DNA and 6 µl Lipofectamine™ 2000 for a 10 cm dish, respectively. 4 h after adding the mixture dropwise to the cells, the culture medium was added to a regular volume. The next day, the medium was exchanged.

Transfection of synthetic sgRNAs was performed using the Stemfect™ RNA transfection kit. The cells were seeded on a 12-well plate and cultured ON. 1 h prior to transfection the next day, the medium was exchanged to fresh medium. Therefore, 25 µl of transfection buffer was mixed with 1 µl transfection reagent in one reaction tube and 25 µl transfection buffer with 25 pmol sgRNA in another. Both mixtures were combined, resuspended, and incubated for 15 min at RT. Finally, the transfection mix was added to the 12-well and incubated ON at 37 °C. The medium was changed the day after.

#### **2.2.3.4 Transient selection for knockouts and sgRNA testing**

Whenever cells were transfected with vectors containing a puromycin resistance cassette for knockouts or CRISPR sgRNA testing, it aimed only for transient expression of the CRISPR/Cas system. Therefore, 24 h post-transfection, the selection was started with a medium containing 4.5 µg/ml puromycin. The selection was preserved for 48 h to select only cells which had taken up the construct. Subsequently, the cells were cultured until a sufficient number for DNA isolation had grown.

#### **2.2.3.5 Selection for stable integrations**

When transfections aimed for stable integrations in the porcine genome and harbored an antibiotic resistance gene, the selection was started 24 h post-transfection. Therefore, the appropriate antibiotic (Blasticidin S, Neomycin/G418, or Hygromycin B) was added to the medium in a concentration first determined in a killing curve experiment, depending on cell type and isolate. The medium was changed every two days, and selection was preserved for at least one week until all non-transfected control cells were eradicated.

#### **2.2.3.6 Immortalization**

Since primary porcine cardiac fibroblasts only grow for up to two passages properly, an immortalized cardiac fibroblast cell line was generated. Therefore, primary porcine cardiac fibroblasts were transfected with a vector containing the SV40 large T antigen using lipofection (see 2.2.3.3). Then single-cell clones were isolated (see 2.2.3.8) and cultured for five passages. A well-growing and morphologically fibroblast-resembling cell clone was used for further experiments as cardiac fibroblast cell line (CFi).

### **2.2.3.7 Activation of immortalized cardiac fibroblasts**

CFis were activated by culturing them in fibroblast medium containing 100 nM of endothelin-1 for 24 h before extracting RNA.

### **2.2.3.8 Isolation of single cell clones**

In order to get genetically identical cells, knockouts, or immortalization, single-cell clones were isolated using cloning rings after stable transfections. Therefore, well-separated cells were first marked on the bottom of the dish. Then the plate was washed with PBS, and the cloning rings were dipped into sterile silicone grease to place them onto the colonies. Subsequently, the cloning rings were filled with Accutase® solution, and the plate was incubated at 37 °C until the cells had detached. The reaction was then stopped by adding one volume of medium, and each colony was transferred to a medium containing 12-well. Once confluent, 50 % of the cells were used for DNA isolation via QuickExtract® (see 2.2.1.1), and 50 % were transferred to a 6-well for further cultivation.

### **2.2.3.9 Nano-Glo® dual luciferase assay**

Nano-Glo® dual luciferase reporting assay system was used to identify active promoter sites in the porcine *MEG3* gene. Therefore, DNA sequences were amplified from porcine gDNA via PCR, where active promoter fragments were assumed. Then, the fragments were cloned into a vector in front of the highly sensitive, promoter-lacking NanoLuc® luciferase (Nluc). A PGK promoter-driven firefly luciferase LUC2 (Fluc) containing vector was used as an internal control.

In the next step, 50000 cells/well of HEK293 and porcine primary cardiac fibroblasts were seeded into a 24-well plate. The next day, 500 ng of overall plasmid DNA of the respective Nluc and Fluc were transfected via lipofection in each well, with a molar ratio of 200:1, each sample in triplicates. 24 h post-transfection, 80 µl or 20 % of the cells were detached and transferred into a 96-well plate enabling luminescence measurement. To each well, the same volume or 80 µl of ONE-Glo™ EX reagent was added to start the Fluc-driven luminescence emission. After incubation for 5 min at RT at 300 rpm on an orbital shaker, the Fluc activity was measured using the Omega FLUOstar and the settings for endpoint luminescence measurement. To stop the Fluc reaction and initiate Nluc-driven light emission, same volume or 80 µl of NanDLR™ Stop & Glo reagent® were added to each well. After 5 minutes incubation at RT and 300 rpm on an orbital shaker, luminescence of Nluc was measured with the same settings as before. The samples with the highest ratio of Nluc:Fluc signal indicate an active promoter site.

### **2.2.3.10 Preparation of porcine cells for somatic cell nuclear transfer (SCNT)**

Correctly targeted cell clones suitable for SCNT were seeded to be confluent two days before the SCNT. Then cell cycle synchronization to G<sub>0</sub> phase was induced by serum reduction to 0.5 % FCS in the standard medium for two days. SCNT was performed by the group of Prof. Eckhard Wolf (Chair for Molecular Animal Breeding and Biotechnology, LMU, Munich, GER), and reconstructed embryos were transferred into hormone-synchronized donor sows as described previously<sup>23</sup>.

### **2.2.3.11 Flow cytometry**

Flow cytometry was performed to validate the determined indel efficiency of Cas12a-mediated knockout experiments on the protein level. Therefore, PECs derived from a Cas12a<sup>+/-</sup> pig were transfected with vectors containing a guide against *GGTA1* or *β2M*. After transient selection (see 2.2.3.4), 0.25\*10<sup>5</sup> cells per well were transferred to a 96-well plate, pelleted for 1 min at 700 rcf, and the supernatant was disposed. Blocking was not performed to prevent the transmission of antigens. Then, 20 µl of primary antibody were added per well, the cells were shaken up and incubated for 20 min at 4 °C, avoiding light exposure. Subsequently, washing was performed using 100 µl of PBS per well. Cells were pelleted again for 1 min at 700 rcf, and the supernatant was disposed. Then, 20 µl secondary antibody were added per well and incubated for 20 min at 4 °C, avoiding light exposure. Subsequently, the cells were pelleted again for 1 min at 700 rcf and washed thrice using 100 µl PBS per well. Finally, the cells were resuspended in 500 µl PBS and transferred into reaction tubes for flow cytometry measurement. The measurement was performed using the Attune™ NxT flow cytometer and analyzed using FlowJo™.

### 3 Results

Genetic modification of pigs is generally time-consuming and inefficient due to its dependence on SCNT for targeted transgenes insertions and the lack of porcine stem cell techniques<sup>136</sup>. In order to enable tissue-specific or even cell-type specific knockouts in adult pigs *in vivo* and therefore providing a shortcut, avoiding the need to establish a knockout animal line for every desired modification, an AsCas12a expressing transgenic pig was generated.

To improve precise transgene placement efficiencies, I developed the CRISPlace system, described in section 3.1.2. Both, homology-directed targeting and CRISPlace was used to insert AsCas12a in the porcine *ROSA26* locus, see section 3.1. Once the AsCas12a-expressing pig was successfully generated, transgenic offspring were characterized in section 3.2.2.

The pig has long been physiological relevant model for cardiovascular disease and treatment thereof and helps to verify studies performed in mice<sup>137</sup>. Many of these investigations in the past decade focused on long non-coding RNAs (lncRNAs) due to their crucial role in regulating gene expression<sup>138</sup>. The lncRNA *Meg3* is expressed in cardiac fibroblasts and involved in cardiac fibrosis in mice<sup>130</sup>. To modify the lncRNA in a porcine model, I first characterized the complex porcine *MEG3* locus in 3.3.1. Then, different Cas9 variants were used to either knockout, knockdown, or overexpress *MEG3* in porcine cardiac fibroblasts in section 3.3.

#### 3.1 Insertion of AsCas12a at the porcine *ROSA26* locus by CRISPlace and HDR

Pigs with constitutive expression of *Acidaminococcus sp.* derived Cas12a (AsCas12a) allow performing knockouts (KOs) in adult animals by delivering the target-specific guide RNA only. Consequently, producing an individual pig line for every desired organ specific KO should become unnecessary. Dr. Beate Rieblinger (Chair of Livestock Biotechnology, TUM, Freising, GER) has successfully produced a SpCas9-expressing transgenic pig line and shown functionality<sup>87</sup>. An AsCas12a expressing pig line would be superior if cell type-specific KOs are desired since it provides an intrinsic ribonuclease activity enabling the processing of sgRNAs expressed from tissue-specific pol II promoters<sup>139</sup>. Furthermore, AsCas12a recognizes a T-rich PAM sequence, contrary to SpCas9 and its G-rich PAM sequence. To increase the genomic targeting range, I decided to use the TYCV PAM-recognizing engineered AsCas12a variant of Gao et al. <sup>33</sup>, which is called AsCas12a in this thesis, unless otherwise stated.

In order to generate pigs expressing transgenes, correctly targeted porcine primary cells must be created first in cell culture. Placing the transgenes at a single, specific locus minimizes the risk of

segregation. The transgenes should further be placed at a safe harbor locus to enable stable expression in the following generations. The porcine *ROSA26* locus is such a permissive locus<sup>64</sup>. Furthermore, the endogenous *ROSA26* promoter can drive antibiotic resistance genes for gene trap vectors, and disrupting the endogenous *ROSA26* gene is not detrimental for the animal<sup>140</sup>. Due to these reasons, the locus was chosen for AsCas12a insertion.

### 3.1.1 Homology-directed targeting for the insertion of AsCas12a at the *ROSA26* locus

The HDR gene trap vector used for the insertion of AsCas12a is shown in Figure 8. The transgene was inserted with the help of a vector containing Cas9 and a sgRNA targeting *ROSA26*. Briefly, porcine kidney fibroblasts (PKFs) from #912 hybrid pigs were co-transfected with both vectors and selected for stable AsCas12 transgene integration via Blasticidin S. Single-cell clones were isolated and analyzed for correct insertion via PCR across the 5' and 3' junctions (see Figure S. 1). Overall, 10 % of all analyzed clones showed correct 5' insertion, while only 1.7 % showed correct 3' insertion (Table 24).

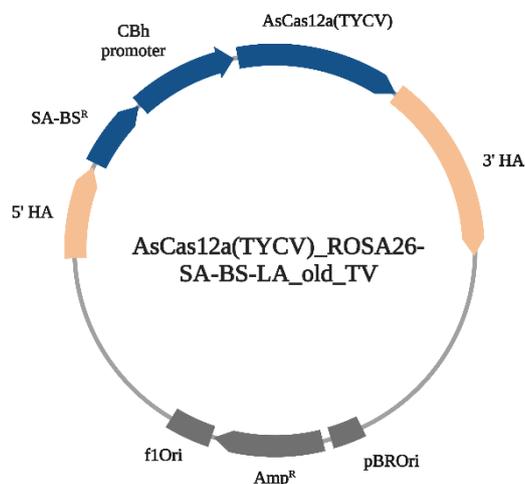


Figure 8 - Structure of the *ROSA26*-AsCas12a HDR gene trap vector. SA = Splice acceptor, HA = Homology arm, BSR = Blasticidin S resistance, Amp<sup>R</sup>= ampicillin resistance. Created with BioRender.com.

The efficiency of a CRISPR/Cas9 mediated insertion of transgenes via homology-directed repair (HDR) generally drops more drastically with the size of the insert<sup>75</sup> than with homology-independent integration. Furthermore, it is challenging to screen for correct insertion due to large homology arms. Given that the AsCas12a gene alone already comprises 3990 bp, and the 5'- and

3'-homology arms are 2.2 kb and 4.7 kb in length, respectively, I concurrently developed two strategies: one based on homology-independent integration and another utilizing homology-directed targeting. It would have the advantage that homology arms became obsolete, and possibly improve targeting efficiency. The system was called CRISPlace.

### 3.1.2 CRISPlace for the insertion of *AsCas12a* at the *ROSA26* locus

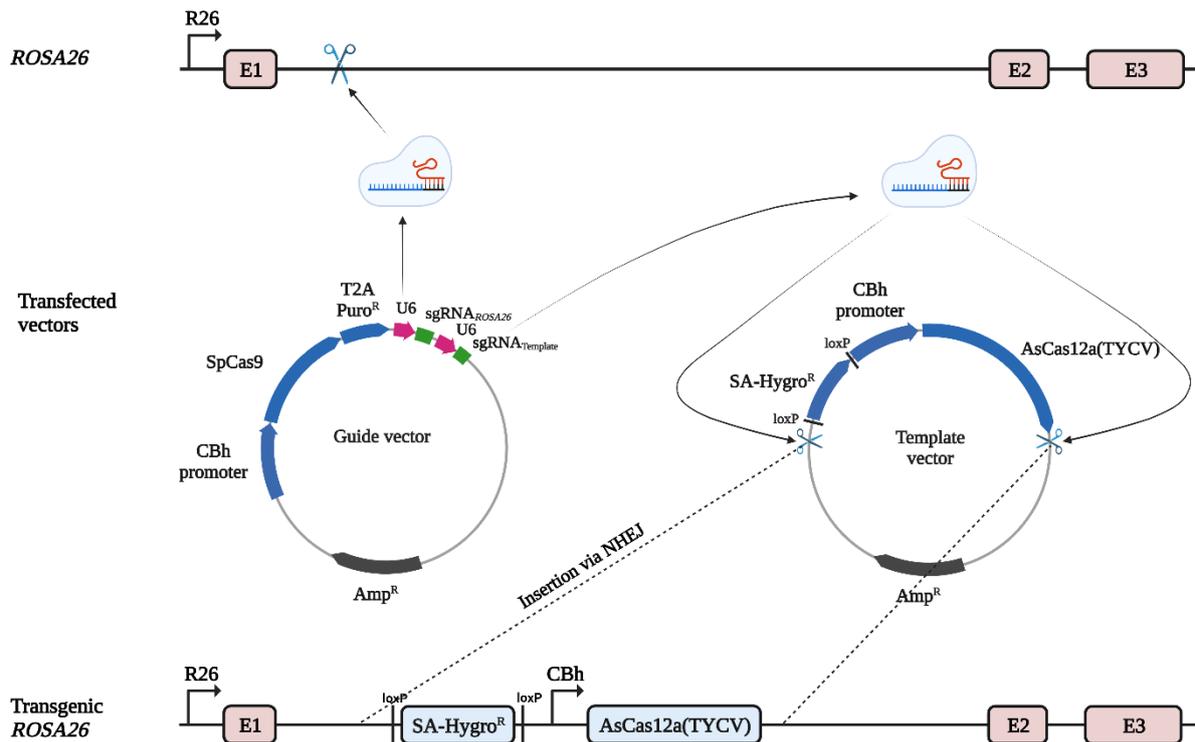


Figure 9 - The CRISPlace principle for homology-independent based transgene insertion. Two vectors, the guide and the template vector are co-transfected into porcine primary cells. The guide vector encodes the Cas9 protein and two effective sgRNAs. Cas9 is directed to the endogenous target locus *ROSA26* by the first guide. At the same time, the transgenic cassette is excised from the template vector by the second guide recognizing two target sites there. The cassette is inserted via homology-independent integration (HITI) at the genomic DSB site. Selection via Blasticidin S ensures that only clones survive carrying the transgene at the correct locus since the endogenous *ROSA26* promoter drives the antibiotic resistance. SA = Splice acceptor, R26 = *ROSA26*, E = Exon, U6 = U6 promoter. Created with BioRender.com.

The CRISPlace system is based on the CRISPaint system for gene tagging of Schmid-Burgk et al., who also kindly provided the original vectors<sup>77</sup>. The CRISPlace principle is depicted in Figure 9. Thereby, porcine primary cells are co-transfected with two plasmids, the guide, and the template plasmid. The guide plasmid encodes for two efficient gRNAs and the Cas9 protein. The first guide (sgRNA<sub>ROSA26</sub>) directs SpCas9 to the endogenous target locus *ROSA26*. The second guide

(sgRNA<sub>Template</sub>) recognizes two target sites on the template plasmid leading to the excision of the AsCas12a-containing transgene cassette from the plasmid, which is then inserted via NHEJ at the genomic cleavage site. The construct contains a splice acceptor (SA) and a promoter-less hygromycin-resistance cassette. Upon correct transgene insertion, the expression of the selectable marker gene will be controlled by the endogenous *ROSA26* promoter. Contrary to the CRISPaint system, both sgRNAs were united on one vector to simplify transfection and handling. Furthermore, to simplify the procedure and avoid the need for minicircle DNA production required in the original system, two identical target sites were introduced—both before and after the desired insert—that were targeted by the same sgRNA. Similar to 3.1.1, PKFs were co-transfected with both CRISPlace vectors and selected for stable transgene integration, this time with hygromycin. Single-cell clones were analyzed via 5'- and 3' junction PCR, showing a higher correct insertion rate (Table 24 and Figure S. 2 ). In 5' PCR, 62 % and 3' 31 % of the clones showed insertion at the *ROSA26* locus. When the template vector was linearized before transfection with a restriction enzyme, the 5' positive rate increased to 78 % and the 3' positive rate to 38 %. Sanger sequencing of six clones indicated that the transgenes were inserted at the cleavage site with indels of up to 90 bp in the majority of clones. However, one out of the six clones exhibited seamless insertion (Figure S. 3).

Table 24 - Transgene insertion efficiency of the gene targeting and CRISPlace approach. CRISPlace is more efficient than gene targeting, particularly in combination with a linearized donor template.

<b>Insertion</b>	<b>Positive in 5' junction PCR</b>	<b>Positive in 3' junction PCR</b>
Homology-mediated targeting	10 %	1.7 %
CRISPlace	62 %	31 %
CRISPlace linearized template	78 %	38 %

It has been shown that double-strand breaks (DSBs) and subsequent NHEJ may increase the insertion of concatemers or multiple copies<sup>141</sup>. Hence, the AsCas12a copy number was checked via droplet digital PCR (ddPCR) in the clones intended for SCNT. Indeed, the NHEJ-based approach CRISPlace showed a significantly higher copy number with a mean copy number of 4 copies compared to 1 copy in the HDR approach (Figure 10). However, only clones containing a single transgene insertion were chosen for generation of GM pigs. Clones from both strategies, CRISPlace, and gene targeting, were used for somatic cell nuclear transfer (SCNT). One healthy



of the AsCas12a protein (Figure 11F). All these findings qualified piglet #2026 as a founder of an AsCas12a pig line.

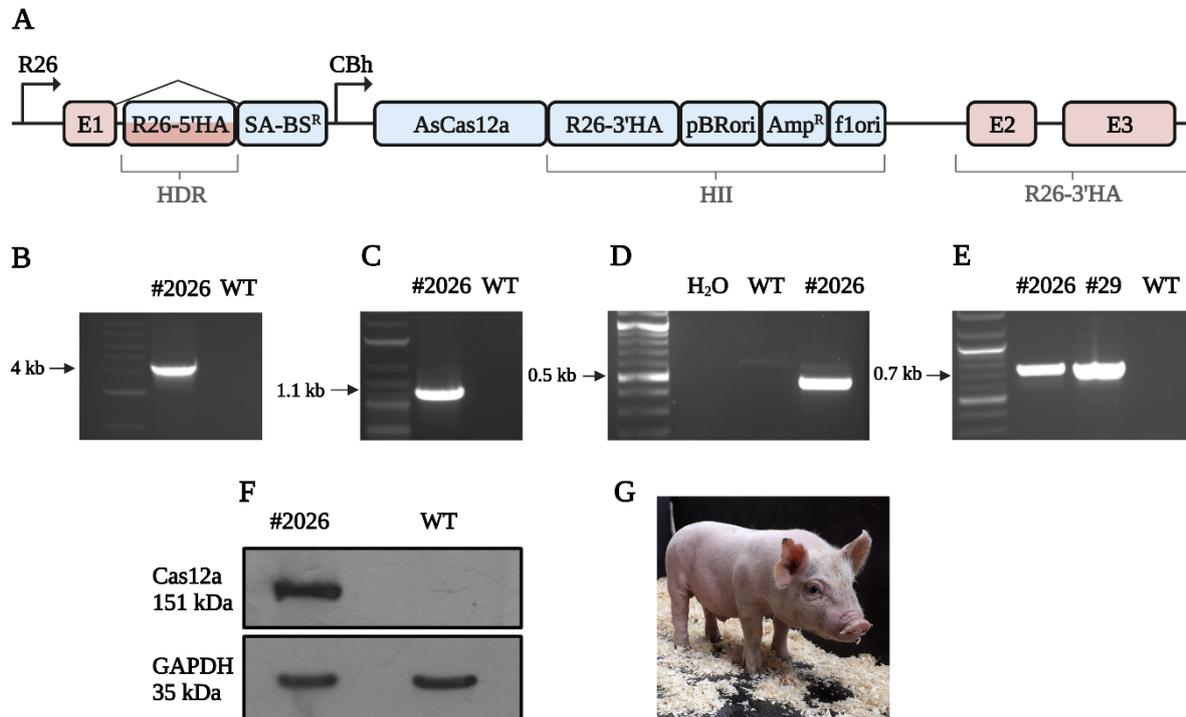


Figure 11 - The founder pig #2026 is suitable for breeding. (A) Structure of the transgenic *ROSA26* locus of the founder pig. The transgene was inserted via HDR at the 5' end while it was integrated homology-independently at the 3' end containing the template vector backbone. (B) 5' junction PCR. (C) 3' PCR spanning from the vector backbone into intron 1 of *ROSA26*. (D) Correct splicing of *ROSA26* exon 1 to the SA-BS<sup>R</sup> cassette on cDNA. (E) mRNA expression of AsCas12a in nuclear transfer clone #29, of which founder animal #2026 was derived. (F) Western blot of porcine ear fibroblast derived protein. (G) Founder animal #2026. E = Exon, R26 = *ROSA26*, HA = Homology arm, HDR = Homology directed repair, HII = Homology independent integration, SA = Splice acceptor. Created with BioRender.com.

### 3.2.2 AsCas12a F1 generation produces transgene mRNA and protein

Once pig #2026 generated offspring, these were genotyped and one of the four transgenic piglets (#2417) was sacrificed. RT-PCR (Figure 12A) and RT-qPCR (Figure 12B) analysis were used to compare the expression of AsCas12a mRNA in different tissues to the average expression across all tissues. The pancreas and the esophagus showed the most robust expression with a fold change of 23.3 and 18.8, respectively. In contrast, the spleen showed practically no expression, and the brain had a low expression level, with a fold change of 0.1. AsCas12a expression in other tissues lay between these values. Analyzed #2417 cells, e.g. epithelial cells (PBEC), endothelial cells

(PAEC, PEC), and fibroblasts (PKF) showed low to intermediate expression between 0.1 and 0.6. Taken together, variable AsCas12a expression was observed in various tissues.

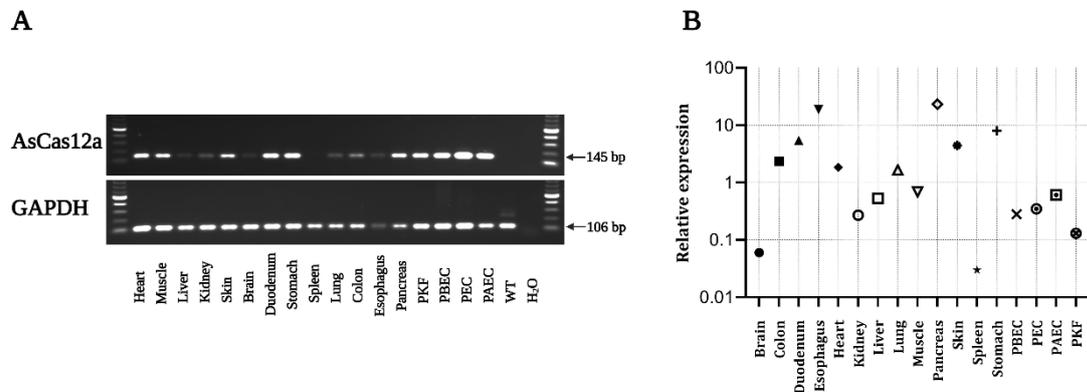


Figure 12 - Expression of AsCas12a in various tissues of the F1 generation (#2417). (A) RT-PCR of #2417 tissue and cells. (B) RT-qPCR of the same samples. PKF = Porcine kidney fibroblasts, PBEC = Porcine bladder epithelial cells, PEC = Porcine heart endothelial cells, PAEC = Porcine aortic endothelial cells, WT = PKFs from wild type pig.

Next AsCas12a protein expression was analyzed by western blot and immunohistochemistry. Since the protein was fused to a 3' 3xHA-tag, I used an anti-HA-tag antibody (ab91110) for all antibody-dependent detection methods. Immunohistochemical evaluation showed AsCas12a protein in all tissues analyzed from pig #2417 (Figure 13A). Protein expression levels varied between tissues, see western blot (Figure 12B). While both mRNA and protein analyses confirmed high expression in the pancreas and esophagus, the remaining protein expression data did not fully align with the mRNA expression. For example, muscle tissue showed moderate mRNA expression but high protein expression. Conversely, the duodenum exhibited high mRNA expression but only a weak signal in the western blot.

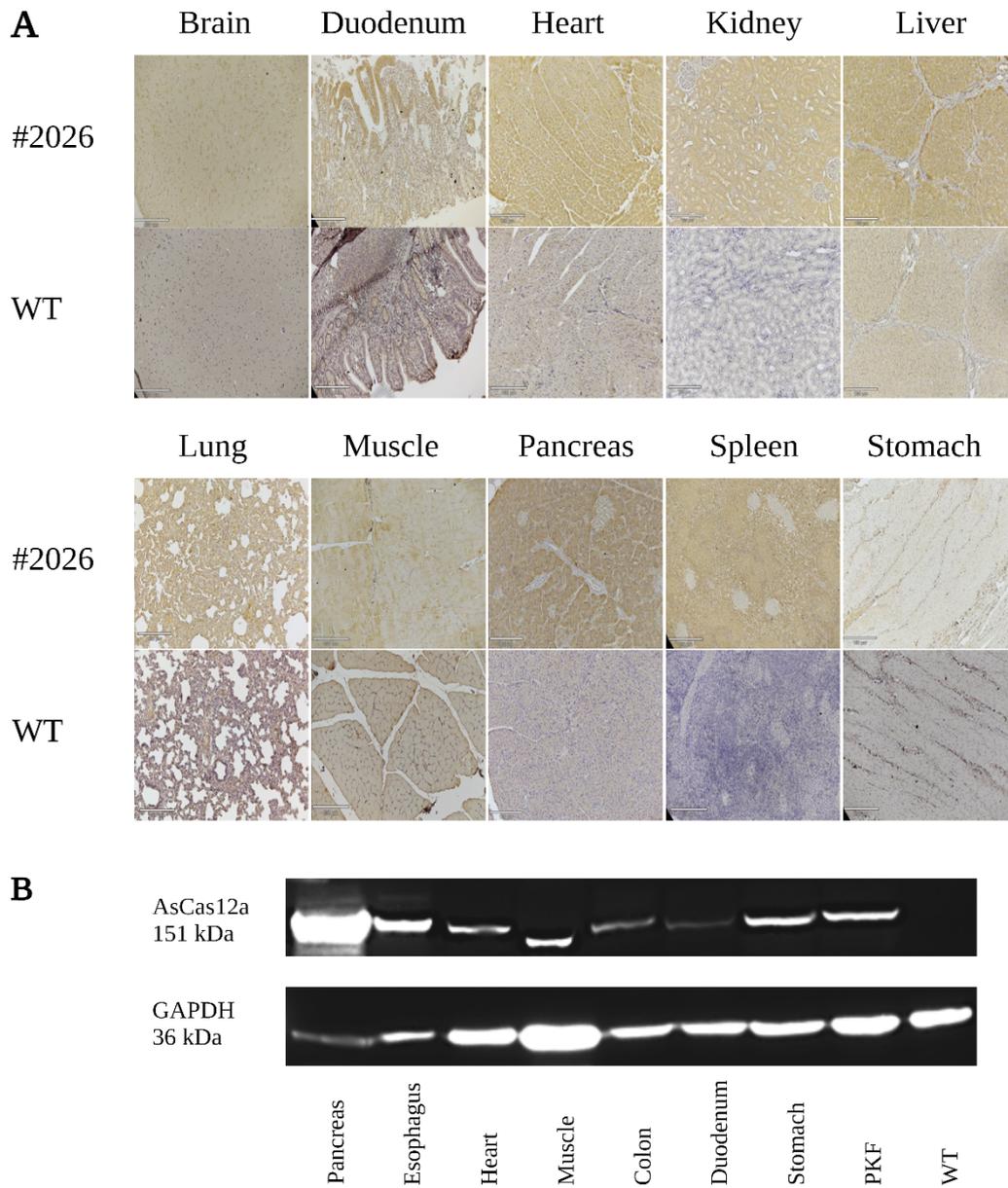


Figure 13 - AsCas12a protein is present in various tissues of F1 generation (#2417). (A) Immunohistochemical detection of AsCas12a using an anti-HA tag antibody. Browning indicates the presence of protein by horseradish peroxidase (HRP) mediated reaction. (B) AsCas12a and GAPDH western blot of selected tissues.

### 3.2.3 AsCas12a shows low efficiency

To validate the functionality of the AsCas12a protein, efficient sgRNAs must be designed. One advantage of AsCas12a over SpCas9 is the option to express sgRNAs from tissue-specific pol II promoters. This is enabled by AsCas12a's intrinsic ribonuclease recognizing 19 nt direct repeats (DR), placed 5' and 3' of the spacer sequence<sup>142</sup>. All sgRNAs for AsCas12a which were used, were driven by the strong pol II CMV enhancer/chicken  $\beta$ -actin (CAG) promoter<sup>143</sup>.

The AsCas12a sgRNAs used in this study were aimed at porcine  $\beta 2M$  or *GGTA1* genes. These genes were previously targeted by Dr. Beate Rieblinger (Chair of Livestock Biotechnology, TUM, Freising, GER) in combination with SpCas9 (<sup>87</sup> and unpublished data). This would enable a direct comparison of the two Cas pig lines. Furthermore, both genes provide a surface marker that can be detected and quantified via FACS analysis. *B2M* is part of the porcine major histocompatibility complex class I (MHC I) gene complex. *GGTA1* has been frequently targeted in pigs for xenotransplantation. It encodes an enzyme producing the xenoreactive  $\alpha$ -Gal epitope<sup>144</sup>.

Six different AsCas12a sgRNAs targeting each individual gene were tested in WT PKF cells to identify the most efficient one for subsequent experiments. These cells were transfected with a single vector that encodes for both AsCas12a and the corresponding sgRNA. After transient selection for vector uptake, the DNA of the cells was sequenced, and editing efficiency was determined using the inference of CRISPR edits (ICE) algorithm. With 58 % and 54 % of edited alleles for  $\beta 2M$ -E1T2 and *GGTA1*-E7T1, respectively, the most efficient sgRNAs for each gene were selected to proof the function of AsCas12a in cells derived from the transgenic pig line. To further improve the efficiency, adding a U<sub>4</sub>AU<sub>4</sub> 3'-overhang for each guide was tested, according to Su Bin Moon et al<sup>145</sup>. Contrary to their findings, the addition of overhangs reduced the editing efficiency by 20 to 30 % and thus was not further used.

The sgRNAs  $\beta 2M$ -E1T2 and *GGTA1*-E7T1 were then used to prove the function of AsCas12a in different cell types isolated from the transgenic pig #2417 and founder animal #2026. These cells were cultured and transfected with a vector encoding for a polycistronic construct of a puromycin resistance and eGFP driven by the hybrid CBA promoter (CBh) and the respective guide expressed the CAG promoter (see Figure 14).

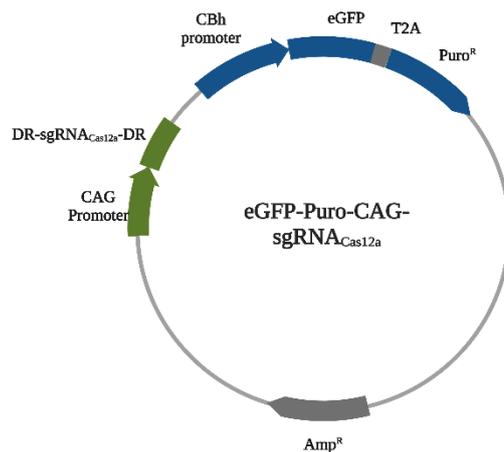


Figure 14 - Map of the vector used to evaluate the functionality of cells derived from the Cas12a-transgenic pig.  
DR = direct repeat.

The cells were either selected by fluorescence via eGFP expression in cell sorting or puromycin. Furthermore, AsCas12a PKFs and PEFs were also transfected with modified synthetic guides from Synthego's Alt-R® to increase RNA stability. Since these experiments did not show any editing, this approach was not pursued further. In total, 67 transfections were performed, of which 16 ( $\cong$  24 %) showed indels. These 16 experiment are summarized in Table 25.

Table 25 - sgRNA efficiencies in AsCas12a derived cells for B2M and GGTA1.

Animal	Cell type	Target site	Selection	Indel efficiency
2417	PEC	GGTA1 (E7T1)	Puro	40 %
2417	PEC	B2M (E1T1)	Puro	37 %
2417	PEF	B2M (E1T2)	eGFP	20 %
2026	PADMSC	B2M (E1T2)	Puro	14 %
2026	PADMSC	GGTA1 (E7T1)	Puro	12 %
2417	PBEC	B2M (E1T2)	eGFP	10 %
2417	PEC	B2M (E1T2)	Puro	9 %
2026	PEF	B2M (E1T2)	Puro	6 %
2026	PEF	GGTA1 (E7T1)	Puro	6 %
2417	PADMSC	B2M (E1T2)	eGFP	5 %
2026	PADMSC	GGTA1 (E7T1)	Puro	5 %
2026	PADMSC	B2M (E1T2)	Puro	5 %
2026	PADMSC	B2M (E1T2)	Puro	5 %
2417	PEC	B2M (E1T2)	eGFP	3 %

2417	PKF	GGTA1 (E7T1)	eGFP	1 %
2026	PADMSC	GGTA1 (E7T1)	Puro	1 %

For the 67 transfections indel efficiency ranged between 0 and 40 %, with the majority lying between 1 and 10 %. Of the analyzed cells, the porcine adipose tissue-derived mesenchymal stem cells (PADMSCs) showed the highest frequency of editing. In PKF cells, the lowest editing rate was observed with 0-1%. In contrast, in one experiment, PECs showed the highest efficiency of 40 % indels. Flow cytometry of these cells was performed to verify the ICE-determined indel efficiency at the protein level. Indeed, 37 % of PECs showed loss of the  $\alpha$ -Gal epitope, indicating that in most of the edited cells, a homozygous KO occurred (Figure 15). Only deletions with more than 3 bp up to 8 bp occurred, showing a typical deletion pattern which others have also observed<sup>36</sup>. Overall, KO experiments in AsCas12a-derived primary cells generally show low efficiency, with only 24 % editing amongst all experiments.

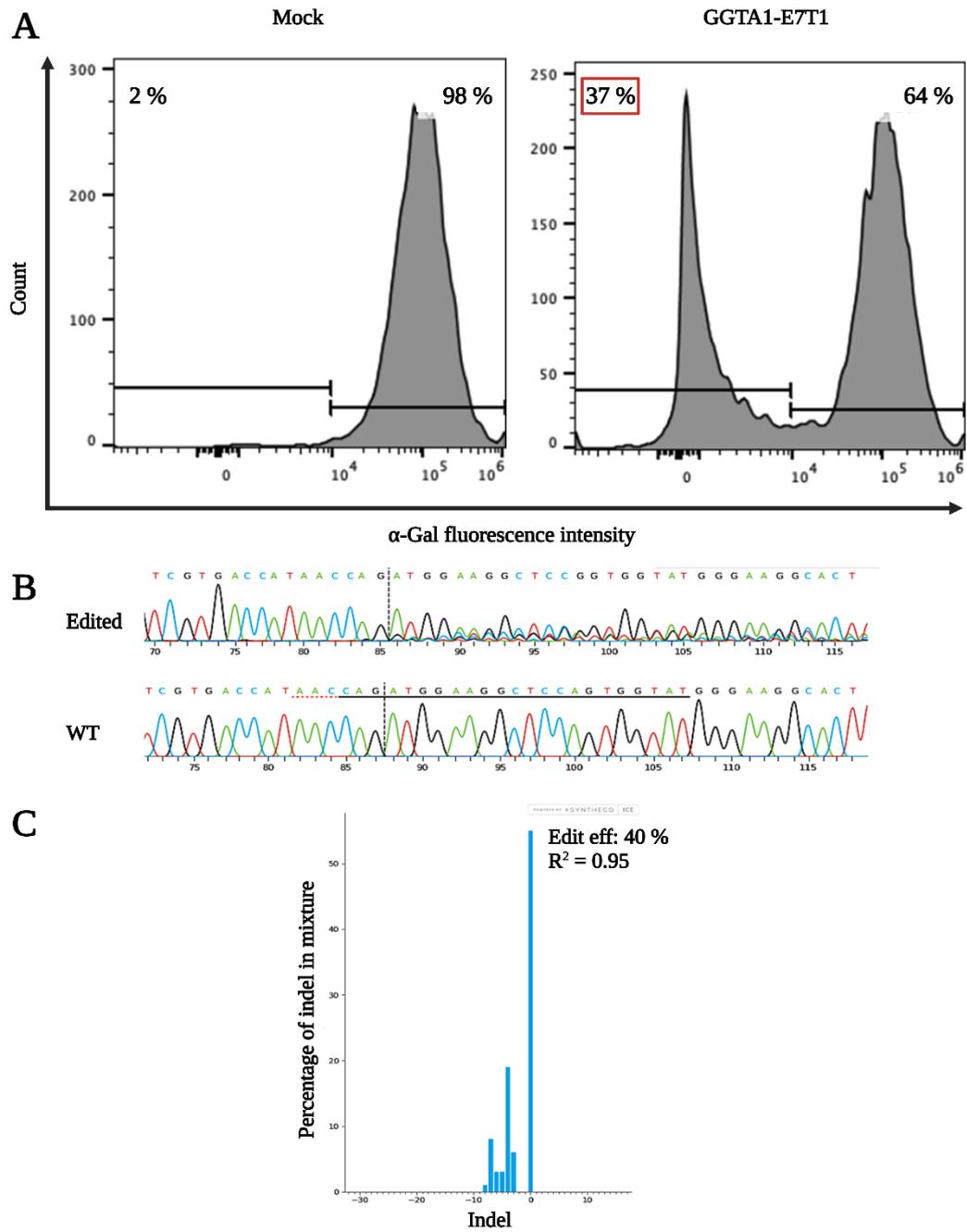


Figure 15 - AsCas12a functionality in primary porcine heart epithelial cells (PEC) derived from AsCas12a transgenic pig #2417. (A) PECs derived from AsCas12a transgenic #2417 animal transfected with a vector comprising a sgRNA against *GGTA1* or *B2M* (Mock). (B) Sanger sequencing and indel distribution (C) of the cut-side PCR.

### 3.2.4 Targeting $\beta 2M$ in porcine AsCas12a cells is inefficient compared to Cas9

Targeting the porcine  $\beta 2M$ , Dr. Beate Rieblinger achieved 63 % homozygous inactivation after puromycin selection in PADMSCs, 36 % in PAECs, and 16 % in PKFs derived from Cas9-expressing pigs<sup>87</sup>. Targeting the same exon to exclude position effects in AsCas12a transgenic cells, I achieved 37 % in PECs, 14 % in PADMSCs, and 6 % in PEF (Table 26). Even though Cas9 porcine aortic endothelial cells (PAECs) and AsCas12a porcine heart endothelial cells (PEC) achieved comparable efficiencies, this was successful in one out of 19 attempts. 15 of the 19 transfections did not show any editing after selection. In AsCas12a pig-derived PKFs no editing was detected, while 16 % of Cas9 cells had a homozygous knockout.

Table 26 - Efficiency comparison in porcine Cas9 and Cas12a cells. The most efficient AsCas12a experiments were selected and compared to the values of the literature.

Cell type	Cas9	AsCas12a
PEC/PAEC	36 %	37 %
PADMSC	63 %	14 %
PKF	16 %	0 %
PEF	59 % (synthetic sgRNA)	6 %

Furthermore, Dr. Beate Rieblinger successfully used synthetic guides. Cas9 porcine ear fibroblasts (PEF) showed 59 % homozygous knockout of  $B2M$  without previous selection. Using synthetic guides no editing was detected in porcine AsCas12a cells.

Taken together, genome editing in AsCas12a cells appears only infrequently (24 % of transfections) and has lower efficiency than in Cas9-expressing cells.

### 3.2.5 Nuclear and cytoplasmic distribution of AsCas12a

Gier et al. suggested that the lower efficiency of AsCas12a is caused by an adverse distribution of the protein due to nuclear export signals (NES) encoded in the amino acid sequence and a weak nuclear localization signal (NLS)<sup>146</sup>. As a consequence, they replaced the original SV40 NLS with six copies of the stronger c-Myc NLS. To investigate the suboptimal nuclear protein distribution, I tested their modified AsCas12a (OpCas12a). Since N-terminal tags on Cas12a have been shown in the literature to not decrease editing efficiency<sup>147</sup>, I added a 3xHA tag at the N-terminus, identical to that of AsCas12a, to detect both proteins with the same anti-HA-tag antibody. An immortal

epithelial porcine kidney cell line, PK15, was transfected with both Cas12a variants, and nuclear and cytoplasmic fractions of the protein were isolated.

The protein distribution was compared to cells derived from the transgenic #2026 pig (Figure 16). Remarkably, for both AsCas12a and OpCas12a, Cas12a seems more prominent in the cytoplasmic than nuclear fraction. Contrary, for protein derived from #2026 cells, the signal is more robust in the nuclear fraction. These findings did not indicate that the low AsCas12a efficiency was due to a suboptimal nuclear protein localization. As a consequence, OpCas12a was not used for further efficiency studies and based on the low and variable editing efficiency, no *in vivo* genome editing in AsCas12a pigs was carried out.

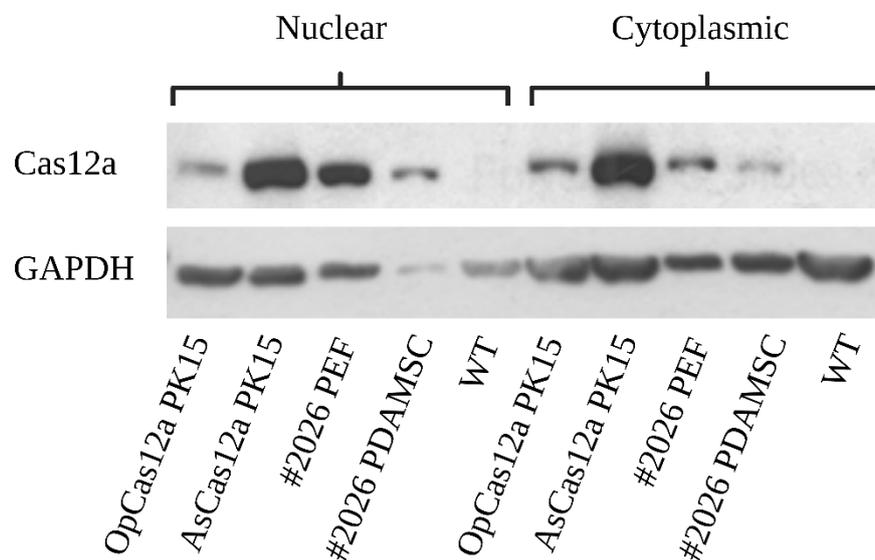


Figure 16 - Nuclear and cytoplasmic Cas12a distribution in OpCas12a vs. AsCas12a. PEF = porcine ear fibroblast, PADMSC = porcine adipose-derived mesenchymal stem cells, PK15 = immortal epithelial porcine kidney cell line.

### 3.3 Manipulation of the lncRNA *MEG3* via SpCas9 variants

Given the advantages of CRISPR technologies for targeted gene manipulation, both Cas9 and the newly derived AsCas12a pig lines were to be used for *in vivo* experiments, such as the editing of non-coding RNAs implicated in cardiac diseases. Specifically, *MEG3* is was chosen, as previous research has established its association with hypertrophic cardiomyopathy<sup>102</sup> and heart fibrosis<sup>130</sup>.

In parallel to developing an AsCas12a transgenic pig line, I characterized (3.3.1) and edited (3.3.2) the porcine lncRNA *MEG3* in cell culture as a possible future target for *in vivo* experiments. The

porcine *MEG3* gene is located on chromosome 7 within the *DLK1-DIO3* locus. The equivalent human locus is found on chromosome 14q32. It is imprinted, regulated by several differentially methylated regions (DMR), and harbors the largest cluster of miRNAs in the human genome (see 1.4.1). Because of its complexity, it was important to assess if a knockout of *MEG3* effects the expression of the other major lncRNA *MEG8* and the coding genes at this locus.

### 3.3.1 Characterization of the porcine *MEG3* gene

When I started this study in Nov-2019, porcine *MEG3* was only partially annotated in NCBI and Ensembl genome browsers. A genomic sequence for one *MEG3* transcript (NR\_021488.1, released Dec-2017) with a length of 1380 bp and comprising two exons was available. In humans, 15 *MEG3* transcript variants with up to 13 exons were identified, and in mice, three transcripts with up to 11 exons. Therefore, mice and human exons were aligned to the porcine reference genome Sscrofa11.1 (release 02-Jul-2017, Swine Genome Sequencing Consortium) via BLAST and Benchling's MAFFT and Clustal Omega algorithm. For genomic regions including introns showing at least 70 % of sequence homology, PCRs were designed spanning a maximum of 1000 bp amplicons overlapping each other. Then, RT-PCR assays were performed using cDNA from heart of WT German Landrace pigs, due to the association of *MEG3* with cardiac disease, and were verified in brain tissue due to the robust expression of *MEG3* there (Figure S. 6). The amplicons were then sent for Sanger sequencing, and the results aligned to the reference genome (Figure S. 4). The uncovered intron-exon structure is depicted to scale in Figure 17.

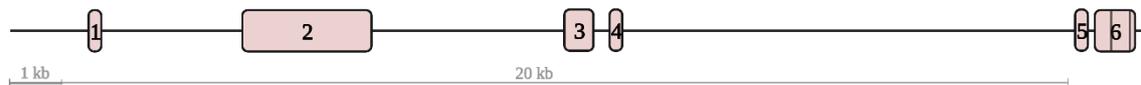


Figure 17 - Porcine *MEG3* intron-exon structure. Lines in exon 6 indicate alternative transcript ends. Created with BioRender.com.

Six different exons were identified, including a large second exon comprising 2390 bp. Using 3'RACE PCR, three alternative 3'ends of the transcripts were detected, defining a transcript length for *MEG3* of either 3526, 3938, or 4009 bp. In humans, the length of *MEG3* transcript varies between 1519 bp and 9701 bp and comprises up to 13 exons.

Recently, an updated porcine *MEG3* transcript (ENSSSCG00000051274, Ensembl 109, release Feb-2023) has been published. The newest version shows a similar intron-exon structure with five to seven exons and alternative transcript ends, as shown in Figure 17.

Even though termed non-coding, some lncRNAs provide short open reading frames (ORFs), thereby coding for short peptides<sup>148</sup>. To exclude that the porcine *MEG3* is also coding for a peptide, NCBI's ORF finder was used to identify possible open reading frames in the *MEG3* exon 1 and 2. The algorithm identified 22 possible ORFs, the longest with 282 amino acids (Figure S. 8). When aligned to known peptides of the SwissProt database, no hits were found via NCBI's blastp. To further exclude possible peptide coding sequences empirically, the subcellular localization of the *MEG3* transcript was checked. lncRNAs located almost exclusively in the nucleus are unlikely to code for peptides, as these are produced in the cytoplasm<sup>149</sup>. Therefore, the nuclear and cytoplasmic subcellular RNA fractions from porcine immortalized cardiac fibroblasts (CFi) (see 2.2.3.6) were isolated and an RT-qPCR was performed (Figure 18). Indeed, 97 % of the *MEG3* transcript was localized in the nucleus, similar to the known nuclear lncRNA nuclear enriched abundant transcript 1 (NEAT1)<sup>150</sup>. Consequently, it is unlikely that lncRNA *MEG3* is coding for peptides.

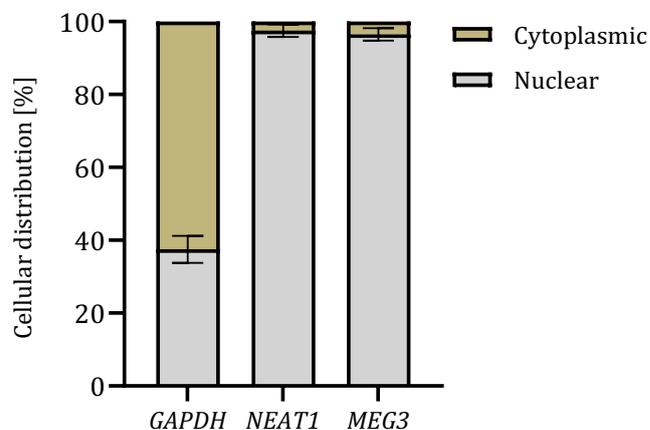


Figure 18 - Cellular distribution of selected RNA. *MEG3* is almost exclusively located in the nucleus. N = 4 biological replicates. Created with BioRender.com.

Next, the regulatory regions of pig *MEG3* were investigated. Gene promoters are suitable targets to knock out lncRNAs since simple indel mutations caused by CRISPR/Cas9 are primarily insufficient to lead to functional impairment of lncRNAs due to their lack of coding<sup>151</sup>. The Neural Network Promoter Prediction was used for *in silico* prediction of promoter regions.

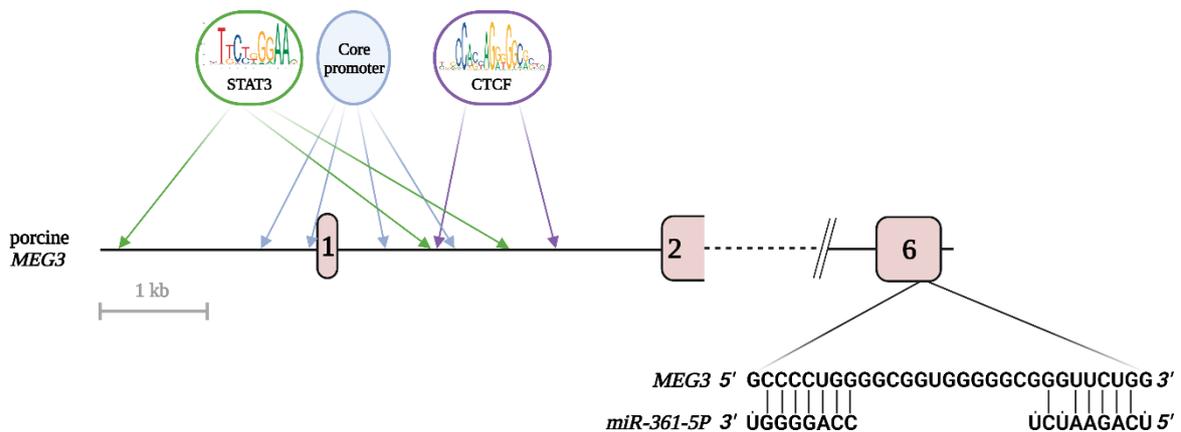


Figure 19 - *In silico* predictions for *MEG3*. Gene depicted to scale. Possible core promoters predicted with the Neural Network Promoter Prediction and possible STAT3 and CTCF predicted with JASPAR (search area +/- 3 kb from exon 1). The possible miR-361-5P binding sites were predicted with RNAhybrid (search area entire gene). Created with BioRender.com.

All four predicted core promoters with a score greater than 0.9 are depicted in Figure 19. Interestingly, they all lie within 1.2 kb around exon 1. The putative core promoter with the highest score of 1.0 was predicted directly upstream of exon 1.

In the next step, a luciferase assay was used to determine the activity of the promoters empirically (Figure 20A). Therefore, the region of interest around exon 1 was divided into six 1 kb fragments overlapping each other. Then, the fragments were cloned upstream of a NanoLuc luciferase reporter gene. Each vector was co-transfected with an internal control vector containing a PGK promoter-driven firefly luciferase. The luminescence ratio of NanoLuc:firefly in HEK293T and porcine primary cardiac fibroblasts (PEC) is shown in Figure 20C. Fragments five and six could harbor a weak promoter due to their low but measurable signal. However, the strongest luciferase activity was derived from fragment 2, which encompasses the region including exon 1 and 1 kb upstream. Furthermore, two of the four core promoter predictions with the highest score are located also in this area. This indicates the localization of the main *MEG3* promoter in this area.

After locating the promoter, other elements that regulate *MEG3* expression had to be determined.

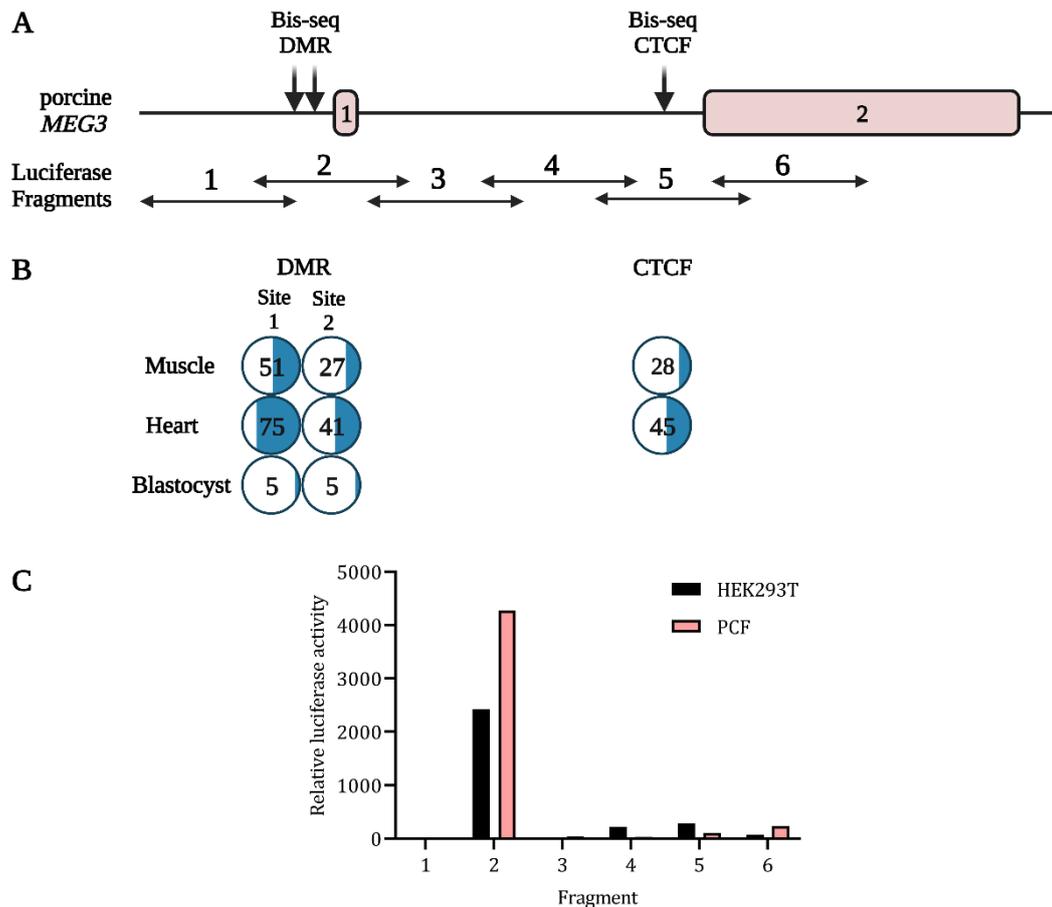


Figure 20 - Luciferase- and methylation assay of the regulatory region of *MEG3*. (A) Areas of bisulfite sequencing (Bis-seq) and luciferase fragments which were analyzed. (B) Percentage of methylated CpGs in muscle, heart, and blastocyst embryonic day 7. (C) The relative luciferase activity of NanoLuc:firefly in porcine primary cardiac fibroblasts (PCF) and HEK293T.

A differentially methylated region (DMR) is defined by showing different methylation levels across different samples, such as distinctive development stages or tissues<sup>152</sup>. Human *Meg3* expression is mainly controlled by two DMRs, the close *Meg3*-DMR, and the more distant but dominant intergenic DMR (IG-DMR)<sup>101</sup>. Recently, the porcine IG-DMR, but not the *MEG3*-DMR, has been described<sup>153</sup>.

I focused primarily on identifying the *MEG3*-DMR to keep the effects of the intended genome engineering of the locus restricted to *MEG3*. To identify the *MEG3*-DMR, I looked for differential methylation in the early *MEG3* gene region, similar to the human *MEG3*-DMR. DNA was isolated from various tissue samples including muscle and heart were used from female and male adult pigs, six-week-old female and male embryos, and blastocysts of embryonal day 7. Bisulfite sequencing was performed to identify the methylation pattern in a region close to the *MEG3* gene. The results of the bisulfite sequencing are shown in Table S. 1, the most important findings in

Figure 20. Overall, sites 1 and 2 showed differential DNA methylation at CpGs in the supposed *MEG3*-DMR, with higher methylation rates at site 1. Up to 24 % higher methylation levels were observed in the heart compared to muscle tissue at site 1. Compared to blastocyst embryonic day (ed) 7, where both sites showed only 5 % CpG methylation, the difference was even more significant. Site 2 showed 14% higher methylation at the CpGs in the heart compared to muscle tissue and 36% higher methylation compared to day 7 blastocysts. The data suggest the presence of DMRs at sites 1 and 2. Interestingly, this region overlaps with the promoter identified by the in luciferase assay.

Oshima et al. postulate that CCCTC-binding factor (CTCF) binding sites within the *Meg3*-DMR are crucial in controlling the non-coding RNAs of the *Dlk1-Dio3* locus in humans<sup>154</sup>. Therefore, the CTCF binding sites are putative targets for genome editing. The JASPAR algorithm was used to predict putative CTCF binding sites in the porcine *MEG3*-locus (Figure 19). Two positions were identified in intron 1. Since the function of CTCF is dependent on accessibility factors such as the methylation status of its binding sites<sup>155</sup>, DNA samples isolated from both female and male porcine tissues were used to analyze the methylation levels at the position with the higher JASPAR score. Looking at the methylation of this putative CTCF site located in intron 1 close to exon 2, variable methylation could also be observed, with the heart showing 17 % higher CpG methylation than muscle tissue.

To sum up, the luciferase assay locates the *MEG3* promoter in the region upstream of exon 1. Furthermore, *MEG3*-DMR was identified in the same region. Moreover, the DMR extended to intron 1 where differential methylation occurs at the putative CTCF binding site.

Zhang et al.<sup>134</sup> discovered the interaction of murine *Meg3* with miR-361-5P and the transcription factor Stat3 plays a role in hypertrophic cardiomyopathy. The miR-361-5P sequence is 100 % evolutionary conserved among human, mice and pigs (miRBase). The likeliest interaction site of porcine *MEG3* with miR-361-5P was predicted using RNAhybrid<sup>156</sup> and was located in exon 6, see Figure 19. Interestingly, only two of the three transcripts identified by 3' RACE PCR harbor this putative interaction site. JASPAR was used to predict possible STAT3 transcription factor binding sites with a score > 0.9. Three possible binding sites were found located either 1.7 kb upstream of exon 1, or in the middle of intron 1, close to the putative CTCF binding sites (Figure 19). In mice, the Stat3 binding site is located 1 kb upstream of *Meg3* exon 1. The miR-361-5P binding site in mice was found within the largest exon, which is exon 3; this corresponds to exon 2 in pigs<sup>102</sup>.

In summary, as binding sites for STAT3, CTCF, and miR-361-5P were identified, their interaction with *MEG3* could be presumed. Furthermore, I located the regulatory elements of *MEG3*, the *MEG3*-DMR, and the *MEG3* promoter. The excision of these elements by CRISPR/Cas9 could lead

to a loss of *MEG3* expression, therefore I focused on the upstream *MEG3* region for designing KO strategies.

### **3.3.2 Manipulation of *MEG3* via SpCas9 variants**

In the following chapter, 3.3.2.1, I describe different excision strategies for the regulatory elements of *MEG3*. In 3.3.2.2, I use CRISPRi to knock down *MEG3*, and CRISPRa to overexpress *MEG3*. All three types of gene expression alterations gave functional insights into the role of *MEG3* in cardiac fibrosis 3.3.3.

#### **3.3.2.1 Efficient knockout of *MEG3* by several different strategies**

The findings of 3.3.1 provided several options for a CRISPR/Cas9-mediated *MEG3* knockout. An overview of the different knockout strategies is given in Figure 21A. The *MEG3* promoter/DMR region, located upstream of exon one and including exon 1 itself to remove the transcriptional start site<sup>151</sup>, is thereby the most obvious target, to knock out all *MEG3* transcripts. Targeting this region will likely lead to the most robust effects and effective KO.

However, promoter knockouts are not always suitable for lncRNAs, as some possess bidirectional promoters that are also required for the expression of other genes<sup>157</sup>. Moreover, targeting this region could also cause severe alterations in the expression of the non-coding RNAs of the *DLK1-DIO3* locus. In humans, BruUV-seq data of the chromosome 14 cluster shows only one transcriptional start site located in the first exon of *MEG3*<sup>99</sup> for all non-coding RNAs of this locus. Furthermore, the *MEG3*-DMR controls the imprinting of the entire locus in mice<sup>158</sup>. Hence, excision of that region would most likely inactivate all non-coding RNAs of the region.

Therefore, strategies to target alternative regions were also designed. This could be the excision of the two CTCF binding sites located in intron 1, which control *MEG3* expression<sup>154</sup>, or parts of the lncRNA sequence itself. More than 50 % of the *MEG3* transcript is encoded by the large exon 2. Important features of the *MEG3* function are likely located within this exon, and its excision could lead to structural impairment. Furthermore, fewer side effects are expected from targeting only exon 2, since it does not harbor as many genetic control elements as the region upstream exon 1. However, the remaining exons can still retain some function, or the deletion of one exon could result in a new transcripts<sup>157</sup>.

In total, nine different knockout (KO) strategies were designed, each consisting of two sgRNAs combined in a single px330 vector, which also contains Cas9 and a puromycin resistance gene. Three strategies aimed to excise exon 1 or 2 (E1-1, E1-2, E-2), five to excise the promoter and DMR

region (E1-1, E1-2, D-1, D-2, D-3) and five the two CTCF sites (E1-1, E1-2, I-1, I-2, I-3) (see Figure 21A). Each vector was transfected into a porcine immortalized cardiac fibroblast (CFi) cell line (see 2.2.3.6), and the cell pool was selected via puromycin. Then, DNA was isolated from the cell pools and PCRs spanning the excision sites were performed. The results are depicted in Figure 21B. For each strategy, a band lower than the WT band of the expected size is visible, indicating a successful excision of the target region.

Furthermore, RNA was isolated from the cell pools and an RT-qPCR was performed to identify the strategy for the most efficient knockout of *MEG3* compared to a mock control (Figure 21C). As expected, all strategies targeting an exon (E1-1, E1-2, E-2) show significantly less *MEG3* transcript than the mock control, with E-2 showing the most substantial reduction by 63 %, followed by E1-2 by 50 % and E1-1 by 47 %. Since *MEG3* expression was measured using primers that span exons 1 and 2, in order to capture all transcript variants, including those with early termination, excision of either of the primer binding sites on the genomic level disables the amplification. Targeting the DMR/promoter region, strategy D-3, even though not significant, shows downregulation of 53 %, while D-1 and D-2 do not show a reduction. Interestingly, the intron strategies displayed an inverse picture. Strategy I-2, targeting the second putative CTCF site, showed a significant upregulation of *MEG3* of around 60 %. However, excising the first (I-1) or both (I-3) putative CTCF binding sites did not change *MEG3* expression significantly.

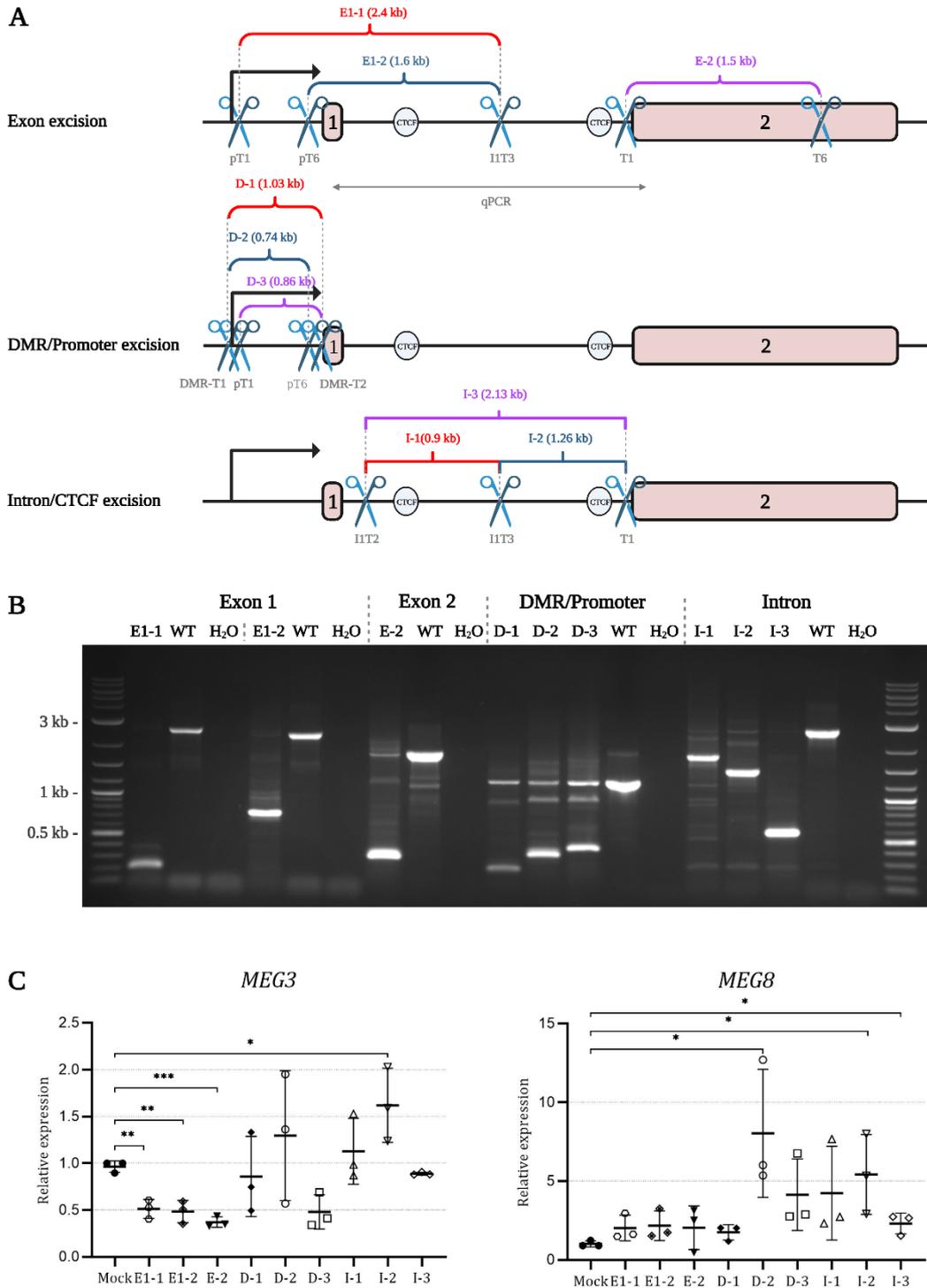


Figure 21 - Different KO strategies for *MEG3*. (A) All strategies for a *MEG3* knockout target the excision of either an exon, the DMR/promoter region, or the intron/CTCF region. (B) PCR spanning the cuts sites of CFi cell pools transfected with the *MEG3* knockout vectors. All strategies lead to an efficient excision of the target region. (C) RT-qPCR of *MEG3* and *MEG8* of the same cell pools. DMR = differentially methylated region. N = 3 biological replicates. One-sided Student T-test was used to perform statistical comparison. \*  $p < 0.05$ , \*\*  $p < 0.005$ , \*\*\*  $p < 0.0005$ . Created with BioRender.com.

Given the intricate regulatory mechanisms governing non-coding RNAs in the *DLK1-DIO3* locus, I also examined potential consequences of modified *MEG3* levels on *MEG8* lncRNA. Notably, every strategy aimed at knocking out *MEG3* led to elevated *MEG8* levels relative to mock controls, with D-2, I-2, and I-3 showing statistical significance. Compared to the mock control, the most robust increase can be observed for D-2 with an 8-fold increase in expression.

As the RT-qPCR data were obtained from mixed cell pools, results might have been obscured due to differences in efficiency of the sgRNAs or reduced efficiency, when larger fragments were excised. Consequently, single cell clones were analyzed for every KO approach. RT-qPCR was performed on RNA isolated from three single-cell colonies, shown to have a homozygous excision of the respective DNA fragment (Figure 22 and Figure S. 7). The expression data showed a complete loss of *MEG3* expression for strategy E1-1, E1-2, E-2, D-1, and D-3. This aligns with the data from the cell pools, where these five sgRNA combinations resulted in the highest reduction. To exclude that the loss of expression was not due to the excision of the primer binding sites in E1-1, E1-2, and E-2, *MEG3* RT-qPCR was performed with an alternative primer pair, binding at the 3' end of exon 2. Indeed, loss of expression could again be observed there in E1-1 and E-2, but not in E1-2. For this reason, E1-2 was excluded as a suitable *MEG3* KO candidate (Figure S. 8).

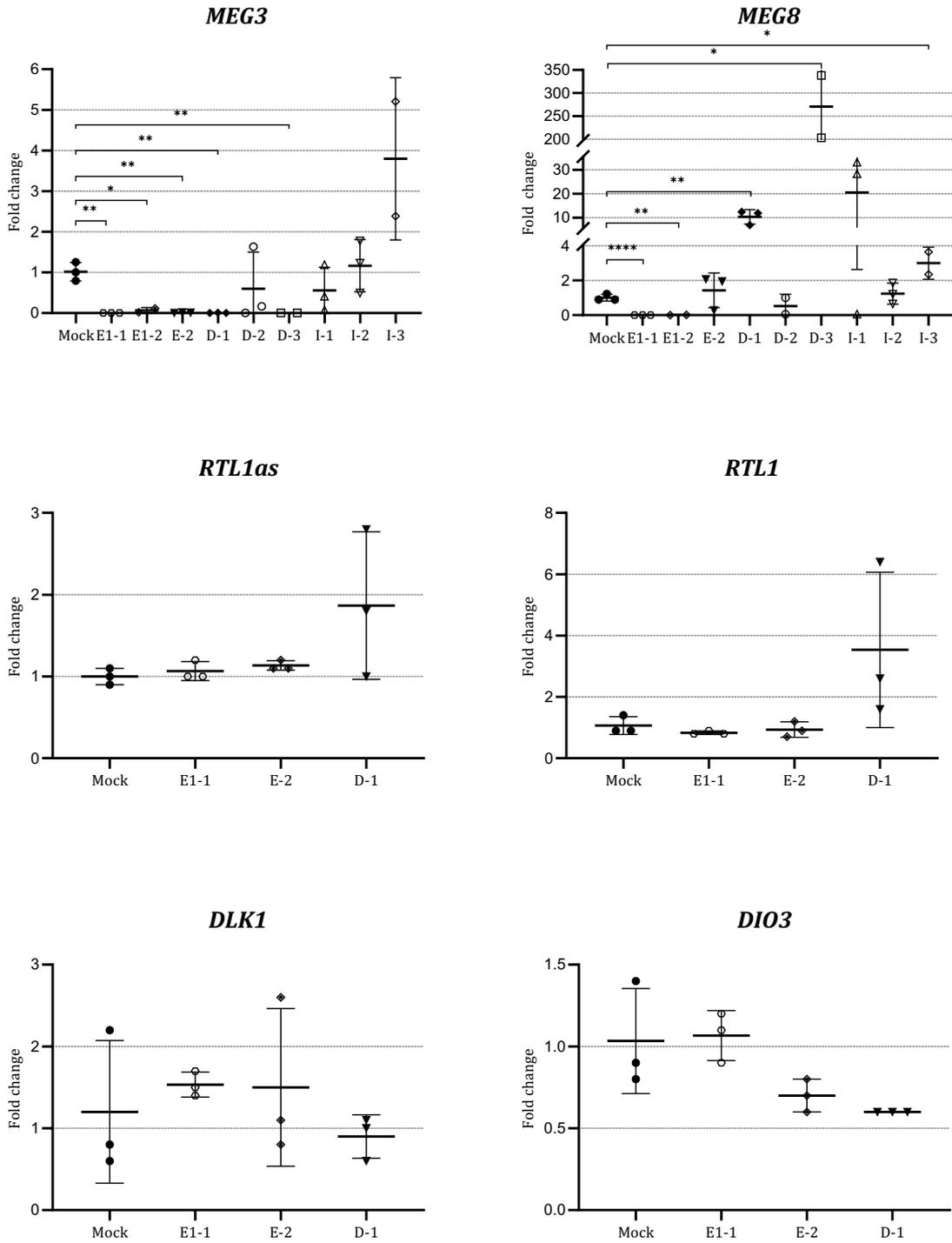


Figure 22 - Expression of relevant genes of the *DLK1-DIO3* locus in *MEG3* KO single-cell clones. Five different strategies lead to a loss of *MEG3*. The KO strategies also alter *MEG8* expression. Besides *MEG3* and *MEG8*, the KO strategies did not significantly affect any other gene located in the *DLK1-DIO3*. N=3 single-cell clones. One-sided Student T-test was used to perform statistical comparison. \* p < 0.05, \*\* p < 0.005.

Interestingly, when targeting the promoter region, *MEG3* expression was lost in strategy D-1 and D-3 but not in D-2. This indicates that a crucial genetic element for *MEG3* expression is located directly 100 bp upstream of exon 1, the region that is not excised in D-2 but in D-1 and D-3. The core promoter with the highest score was also predicted there (3.3.1). As a consequence, D-2 was also excluded as a suitable KO strategy.

Not all RT-PCRs showed a loss of *MEG3* expression. Moreover, in I-3, even an upregulation of *MEG3* could be observed, so the intron strategies were excluded for further analysis.

To examine which combinations have the lowest impact on the expression of the other genes of the *DIO3-DLK1* locus, the *MEG8* expression was also investigated in the single cell clones. The expression of *MEG8* was generally influenced by KO of *MEG3*. While strategy E-2, D-2, I-1, and I-2 did not change *MEG8* expression significantly, E1-1 and E1-2 resulted in downregulation of *MEG8* expression, and D-1, D-3 and I-3 even in a significant upregulation of *MEG8* by up to 250-fold for D-3 and tenfold for D-1. Because of the considerable deregulation of *MEG8* in D-3, I decided to exclude this strategy for *MEG3* KO.

As a result, E1-1, E-2, and D-1 were considered as suitable strategies for a *MEG3* KO. While E1-1 also lead to a loss of *MEG8* expression, E-2 did not affect *MEG8* expression, and D-1 even resulted in a significant upregulation. Subsequently, I investigated whether these strategies also influence the expression of the other genes of the locus besides *MEG8*. There was no significant change in gene expression of the coding genes *DLK1*, *DIO3*, and *RTL1* within the *DLK1-DIO3* locus. Interestingly, the other non-coding RNA in the locus, *RTL1as*, was also not affected by the *MEG3* KO (see Figure 22). To assess the influence of a *MEG3* KO (+/- altered *MEG8* expression) on cardiac fibrosis RNA sequencing was carried out on RNA samples from the single cell clones (CFis).

Genomic alterations in the *MEG3* gene can influence the whole locus and thereby impede unraveling the exact role of the *MEG3* lncRNA. As a consequence, CRISPRi was used to knock down *MEG3* alone. Furthermore, CRISPRa was used to overexpress *MEG3* from its endogenous locus and to assess how this affects overall gene expression in CFis in comparison to either a *MEG3* knockout or knockdown.

### 3.3.2.2 Efficient knockdown and overexpression of *MEG3* by CRISPRi and CRISPRa

Section 1.2.2 gives an overview over CRISPR interference (CRISPRi) and CRISPR activation (CRISPRa). A deactivated Cas9 (dCas9) is fused to a transcriptional regulator in both strategies. For gene activation, dCas9 is fused to the tripartite activator VPR consisting of VP64-p65-Rta<sup>46</sup>. For gene downregulation, dCas9 is C-terminally fused to a bipartite repressor domain, KRAB-MeCP2<sup>51</sup>.

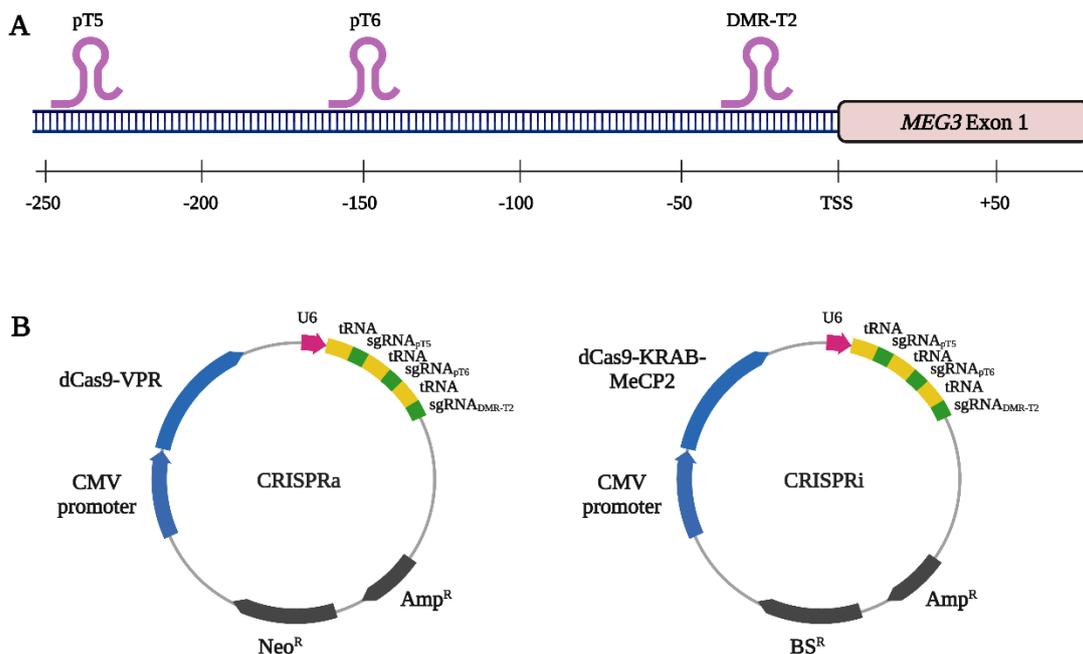


Figure 23 - *MEG3* target area for sgRNAs (A) using CRISPRa and CRISPRi and corresponding vectors (B). All three sgRNAs were used as a set. Neo<sup>R</sup> = neomycin resistance, VPR = Vp64-p65-Rta, BS<sup>R</sup> = Blasticidin S resistance. Created with BioRender.com.

Figure 23A depicts the binding sites for the used sgRNAs. Since both dCas9 variants interact within the promoter region, the same set of guides was used for activation and knockdown. Off-targets are generally of lesser concern than in SpCas9 due to the unlikelihood that the off-target is located in a promoter region<sup>159</sup>.

Furthermore, gene expression regulation might be controlled by the most potent guide in the set<sup>51</sup>. Hence, three guides were used instead of one, binding up to 250 bp upstream of exon 1, the generally recommended target region for the best outcome<sup>46,51</sup>. All guides were multiplexed under the control of a single U6 promoter using the tRNA system<sup>160</sup>. Then, these guides were combined with the respective dCas9 variant on a single vector (see Figure 23B).

Once the vectors were generated, they were transfected into porcine immortalized cardiac fibroblasts (CFi). Then, cells were selected for stable integration via Blasticidin S for CRISPRi and neomycin for CRISPRa. Six days after transfection, RNA was isolated and checked for efficient knockdown or overexpression of *MEG3* compared to a guideless mock control via RT-qPCR (Figure 24). Indeed, the CRISPRa approach led to a significant increase of *MEG3* expression of up to 40-fold compared to the guideless mock control. Interestingly, as for the KO approaches, *MEG8* expression is also affected. Moreover, *MEG8* is even upregulated up to 100-fold compared to the mock control. Again, similar to the KO, the other genes of the *DLK1-DIO3* locus were not affected significantly by overexpression of *MEG3* and *MEG8* (Figure S. 9).

For CRISPRi, *MEG3* expression showed a 49% reduction compared to the mock control, albeit not statistically significant. For *MEG8*, no reduction was observed. Also, the expression of other genes of the locus *DLK1*, *RTL1*, *RTL1as*, and *DIO3* did not change significantly (Figure S. 9). Summing up, an effective strategy was generated to overexpress *MEG3* as well as *MEG8* up to 40 and 100 fold, respectively, from their natural locus using CRISPR activation. Furthermore, via CRISPRi, *MEG3* expression was reduced by 49 % without producing DSBs in the genome. Finally, with three suitable *MEG3* knockout strategies all tools are available to get deeper insights into the role of *MEG3* in cardiac fibrosis.

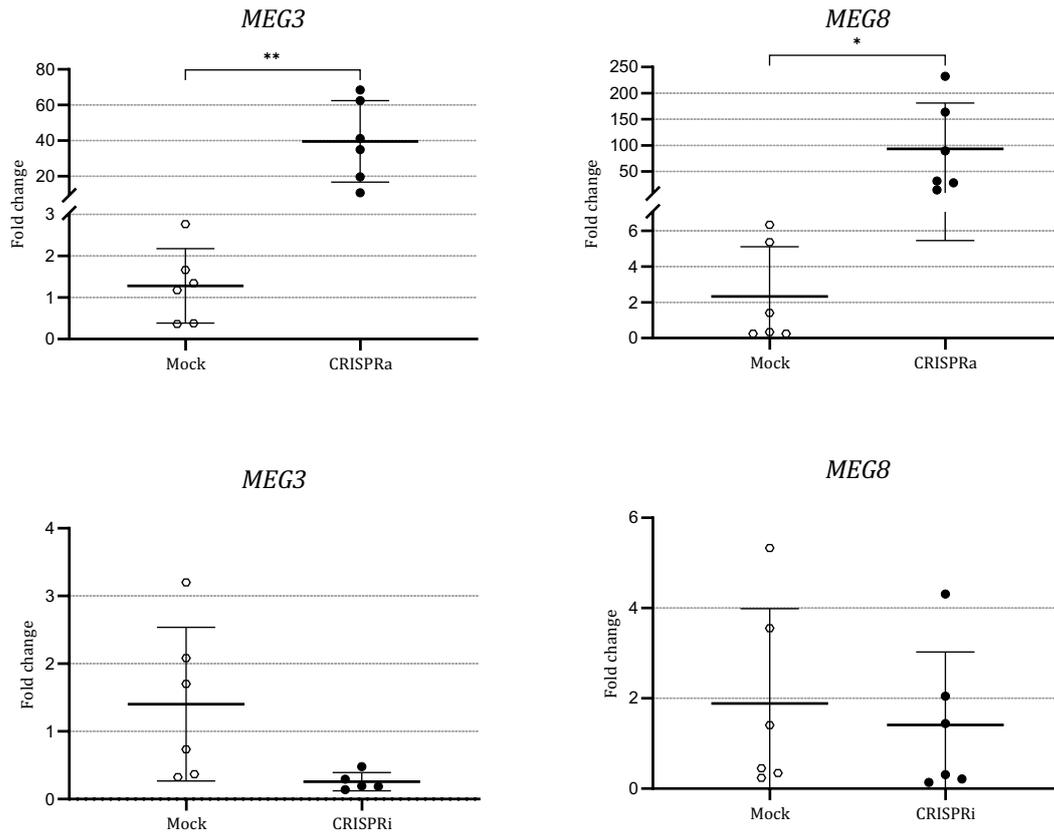


Figure 24 - Overexpression and knockdown of lncRNA *MEG3*. N = 6 individual experiments. One-sided Student T-test was used to perform statistical comparison. \*  $p < 0.05$ , \*\*  $p < 0.005$ .

### 3.3.3 Bulk RNA-Seq shows effect of *MEG3* KO D-1 and E-2 on fibrotic signaling

To investigate *MEG3*'s role in cardiac fibrosis, RNA sequencing was performed on two cell clones derived from each of the three successful knockouts (E1-2, E-2, D-1) in cardiac fibroblasts. Figure 25 depicts the results. Principal component analysis (PCA) analysis (Figure 25A) and significance analysis (Figure 25B) showed that there was no significant change in gene expression in KO strategy E1-2, even though this is the only guide combination which lead to a double KO of *MEG3* and *MEG8*. As a consequence, this strategy was excluded from the pathway analysis.

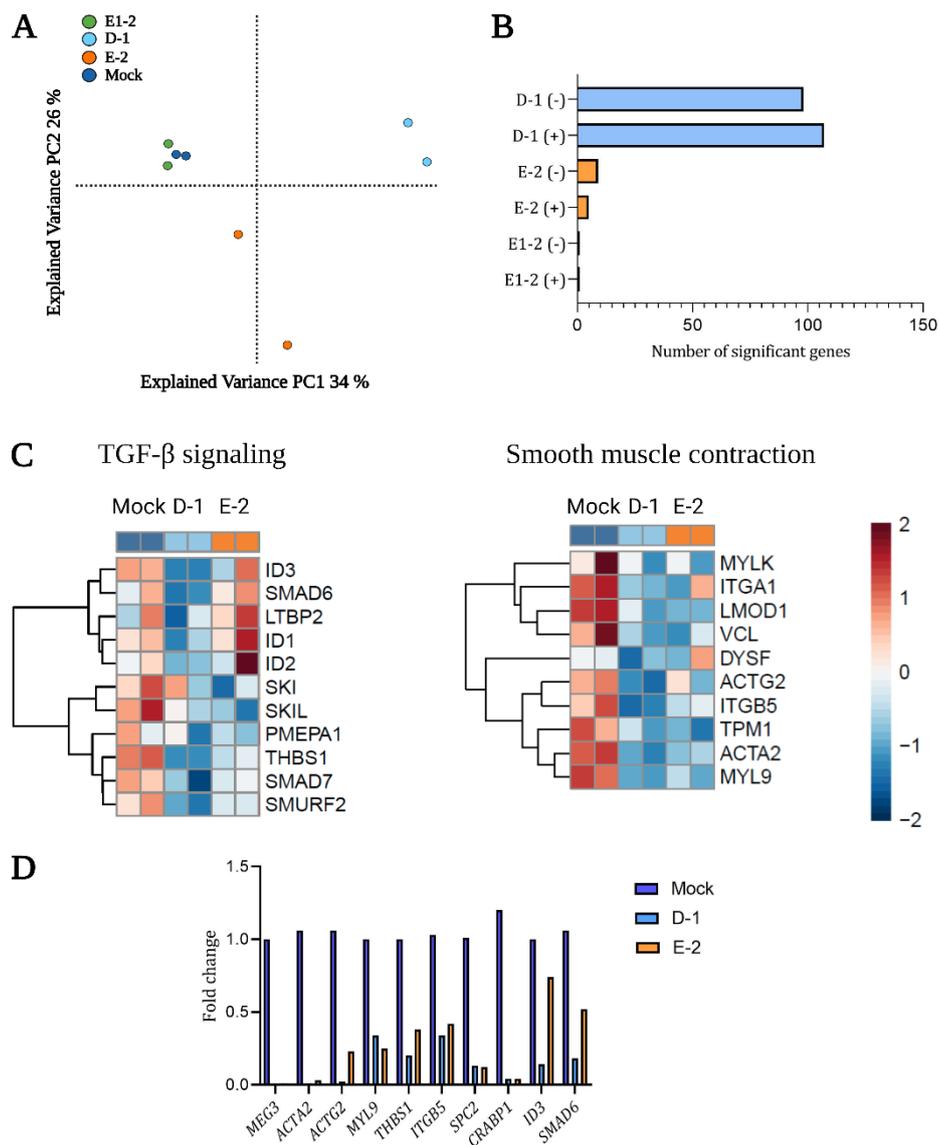


Figure 25 - *MEG3* knockout in cardiac fibroblast attenuates fibrotic signaling in bulk RNA-seq. (A) PCA analysis and the number of significant genes (B) reveal the most differential expression in *MEG3* promoter KO D-1, while there is no significant change in KO E-1-2. Genes with a  $p_{adj} < 0.05$  were considered significant. (C) Genes of the pro-fibrotic *TGF-β* pathway are downregulated in KO D-1 more substantially than in KO E-2. The smooth muscle contraction pathway is reduced in D-1 and E-2 comparably. (D) Verification of selected genes via RT-qPCR. (+) = upregulation, (-) = downregulation.

The most robust changes were observed in the PCA of the promoter KO using the D-1 strategy, which lead to a *MEG3* KO and a *MEG8* upregulation. Furthermore, more than 107 genes were significantly upregulated, and more than 98 genes were significantly downregulated.

In PCA, E-2 shows substantial differences compared to the mock control. However, there are high inter-individual variances between the E-2 duplicates. For E-2, the exon two excision, only five genes are significantly upregulated, and nine genes are significantly downregulated.

Pathway analysis shows that both KOs D-1 and E-2 affected regulatory pathways relevant to cardiac fibrosis. The *TGF- $\beta$*  pathway is pivotal for cardiac fibrosis, while the smooth muscle contraction indicates activated myofibroblasts<sup>161</sup>. Genes of both pathways were downregulated in both KOs, while D-1 shows a more consistent downregulation than E-2 for *TGF- $\beta$* . Furthermore, Hedgehog signaling, which is connected with promoting cardiac fibrosis, is also reduced in D-1 (Figure S. 10)<sup>162</sup>. Other pathways, like the anti-inflammatory *IL-10* pathway, are upregulated in D-1 and E-2. To confirm the results from the RNA-Seq, qPCR analysis of selected genes was performed (Figure 25D). All genes analyzed showed similar changes in expression as determined by RNA-seq. With 96 % downregulation in D-1 and 93 % in E-2, respectively, the most significant downregulation among all genes can be observed for the *ACTA2* gene, the most prominent activated fibroblast marker. This indicates an impact of a *MEG3* KO on the activation of cardiac fibroblasts. On the other hand, Piccoli et al. showed that cardiac fibrosis in mice is promoted by activation of Tgf- $\beta$  via Mmp2 (see 1.4.2)<sup>130</sup>. However, no significant change in *MMP2* expression was observed for D-1 and E-2.

Overall, two of the three *MEG3* KOs downregulated pathways relevant for cardiac fibrosis, D-1 more than E-2. However, this is not linked with the downregulation of *MMP2*.

In order to analyze the consequences of an overexpression (OE) or a knockdown (KD) in comparison to a *MEG3* knockout in cardiac fibroblasts, RNA-Seq for samples from the OE and KD experiments was performed. Unexpectedly, similar to the KO E1-2, the OE, and KD did not show any significant changes in gene expression. Consequently, the OE and KD RNA-Seq data could not be compared to the KO data.

## 4 Discussion

The first aim of my project was to generate an AsCas12a-expressing pig line to enable tissue-specific *in vivo* knockouts (KO) in pigs. Therefore, the CRISPlace system was developed for transgene insertions to accelerate the generation of donor cells for somatic cell nuclear transfer (SCNT). The efficiency of the CRISPlace system is discussed in the following chapter, 4.1. Once the AsCas12a pig was successfully generated, I characterized and tested the efficiency of the AsCas12a-derived cells. Section 4.2 discusses whether this pig line is a valuable alternative to the Cas9-expressing pig.

One immediate goal for the generation of the AsCas12a transgenic pig line was an *in vivo*, cardiac cell specific knockout of non-coding RNAs. Therefore, in the second part of my project, the porcine cardiac disease-relevant long non-coding RNA (lncRNA) *MEG3* was characterized. Based on this, different KO strategies for *MEG3* were generated and assessed in a porcine cardiac fibroblast cell line. In addition, the effects of CRISPRa- and CRISPRi-based overexpression and knockdown of *MEG3* was analyzed. Chapter 4.3.2 discusses the characterization of the locus, the most vital strategies for modulating *MEG3*, and their influences on cardiac fibrosis pathways.

### 4.1 CRISPlace is an efficient strategy for large transgene insertions

The precise insertion of entire genes or DNA fragments into the genomes of livestock animals has been challenging due to several factors. These include the absence of pluripotent stem cells, the requirement for gene targeting in somatic cells—which have a short lifespan and low HR efficiencies—and the for necessity SCNT. The CRISPR/Cas system can streamline cellular manipulation, thereby accelerating gene insertion in somatic cells and reducing the consumption of resources.

Insertion of transgenes is based on providing suitable repair templates after a CRISPR/Cas9-induced DSB at a target locus. Most transgenes are inserted via HDR since NHEJ is error-prone. However, HDR is generally less efficient than NHEJ<sup>163</sup>. Targeting the safe harbor locus *ROSA26* enables stable transgene expression<sup>66</sup>, without the need for an error-free transgene insertion, when a gene trap vector is used.

As described in 3.1, I developed the NHEJ and gene trap-based insertion strategy CRISPlace for inserting the AsCas12a-TYCV variant to increase somatic cell targeting efficiency. Strikingly, considering 5' and 3' junction PCR, CRISPlace was up to 22-fold more efficient in generating AsCas12a transgenic donor cells for SCNT than a similar HDR approach. This significant boost in efficiency could be attributed to several factors, the most prominent might be the size of the

transgene. The large AsCas12a transgene cassette, including resistance gene and CBh promoter-driven AsCas12a, comprises 6951 bp. It has been reported that the efficiency for large transgene insertions is lower for HDR than for NHEJ<sup>75,164</sup>. In a study in plant cells, the NHEJ approach compared to HDR achieved double insertion efficiency for fragments up to 25 kb<sup>174</sup>. The insertion efficiency of the NHEJ approach could be even higher, since in this study, a method was used restricted to the detection of single copy integrations, and NHEJ often results in insertion of concatemers<sup>165</sup> (See 3.1.2).

Second, the homology-independent integration of the 3' homology arms in the homology-mediated approach could lead to false negatives during clone screening, complicating the identification of positive clones. For example, cell clone #29, which eventually gave rise to the transgenic AsCas12a pig, displayed this issue. This clone was not initially identified in 3' junction PCR screening due to the large size of the 4.7 kb 3' homology arm and additional backbone integration. While this issue partially accounts for the efficiency difference in 3' junction PCR, CRISPlace still outperforms the HDR method. As both vectors depend on a gene trap construct, the cells were selected for correct 5' insertion via antibiotics. When considering only the 5' positive rate, CRISPlace is 7.8-fold more efficient.

In order to exclude transgene size effects, the CRISPlace system was also tested for another gene. Agnieszka Bak (Chair of Livestock Biotechnology, TUM, Munich) replaced the 3.9 kb AsCas12a with the 317 bp xenoprotective transgene LEA-29Y and used CRISPlace to insert it at the *ROSA26* locus (unpublished data). Here, all 17 analyzed clones were positive in 5' junction PCR, and 35 % were positive for 3' junction PCR, thereby showing comparable efficiency to the insertion of AsCas12a. Therefore, the efficiency at the *ROSA26* locus does not drop substantially with the gene size.

To exclude genomic position effects, a modified CRISPlace vector was used in another approach. Instead of using a gene trap vector, the resistance cassette included the SV40 promoter. This vector was then employed to insert the human UCP1-derived minigene into the porcine *UCP1* pseudogene locus (unpublished data). Even without the selection advantage provided by a promoter trap vector, a targeting efficiency of 26 % (5' junction PCR) was achieved. Contrary, only 10 % positive cells in 5' junction PCR for AsCas12a were achieved with the homology-mediated targeting vector.

To sum up, regardless of gene size and transgene location, the CRISPlace system achieves consistently higher insertion rates than the HDR-based gene targeting approach.

One reason for the higher rate is, as mentioned before in 3.1.2, that NHEJ is the predominant repair pathway in the cell. The sister chromatid as a natural HDR template is only available in the late S and G2 phase of the cell cycle<sup>166</sup>.

The second reason for the higher efficiency is the dual-cut strategy. The CRISPlace vector contains the transgene flanked by two sgRNA recognition sites, leading to the transgene excision from the vector upon co-transfection with the guide vector. The simultaneous cut of the genome and the vector increases the insertion rates. For HDR templates, this strategy increased in one study the efficiency two- to five-fold<sup>167</sup>. In another study, it increased the efficiency for targeted integration via NHEJ more than for HDR. They also tested for off-target integration of the transgene, which is possible, but was low compared to on-target integration<sup>75</sup>. Since CRISPlace can employ a promoter trap vector, selection prevents off-target integration.

A drawback of the NHEJ-based integration is the risk of multi-copy insertion compared to a single copy with the HDR approach (Figure 10). In transgenic fish lines, long tail-to-head tandem integrations tended to decrease expression, probably by methylation-dependent silencing. On the other hand, in some clones tandem arrays also increased expression<sup>168</sup>.

However, the five-fold transgenic pigs expressing human complement inhibitors and anti-inflammatory genes generated by Dr. Konrad Fischer (Chair of Livestock Biotechnology, TUM, Munich) stably and highly express the transgenes even with copy numbers as high as 26 (unpublished data) over several generations<sup>169</sup>. Here concatemers or array integrations do not hamper transgene expression. Consequently, the transgene copy number alone is insufficient to forecast transgene expression. Hence, it is advisable to check single cell clones for transgene expression prior to SCNT.

Concluding, the CRISPlace system is a valuable tool for the insertion of large transgenes with high efficiency. It can be used for sizable transgenes and at different genomic locations. It does neither require linearization of donor vectors prior to transfection, nor any homology arms. The CRISPlace system reduces laborious cloning and targeting of porcine somatic cells, therefore accelerates the generation of donor cells for SCNT.

## 4.2 AsCas12a requires engineering to achieve Cas9-comparable performance

The Cas9-expressing pigs generated by Dr. Beate Rieblinger (Chair of Livestock Biotechnology, TUM, Freising) are a valuable tool for *in vivo* genome editing of pigs. Using these pigs circumvents the production of separate knockout lines for each desired genotype and reduces the number of required animals. However, the expression of sgRNAs is restricted to ubiquitous pol III promoters, preventing the expression from tissue-specific promoters<sup>87</sup>.

In contrast, AsCas12a can process whole CRISPR arrays driven by tissue-specific pol II promoters and enables editing of alternative target sites to Cas9 since it recognizes a T-rich PAM. An AsCas12a-expressing pig complements Cas9 as an *in vivo* genome editing tool.

The AsCas12a-expressing pig (see 3.2) derived from the gene targeting cell clone #29. The transgene was inserted at the safe harbor locus *ROSA26* and was correctly spliced. The offspring showed broad mRNA expression in tissues and cells. AsCas12a protein was present in all tissues analyzed via western blot and immunohistochemistry.

However, although multiple transfections of AsCas12a-expressing cells with gRNAs targeting different genes were performed, overall efficiency was low.

### 4.2.1 Neither pol III promoter nor backbone integration causes AsCas12a inefficiency

Section 3.2.4 demonstrated that AsCas12a is less efficient than Cas9 in knocking out the *B2M* gene. Notably, both Cas9 and AsCas12a are driven by the same CBh promoter and are inserted at the same locus, *ROSA26*. However, the AsCas12a transgene locus also contains parts of the bacterial vector backbone. While some studies suggest that transgenes without backbone integration exhibit higher expression<sup>170</sup>, others have found no such effect<sup>171</sup>. Expression analysis using RT-qPCR, western blot, and immunohistochemistry confirmed the presence of AsCas12 mRNA and protein in porcine tissues. However, the potential influence of the backbone sequence on expression cannot be ruled out. For instance, AsCas12a showed no mRNA expression in spleen tissue as detected by RT-PCR, unlike Cas9, which was clearly expressed in the spleen of Cas9 pigs<sup>87</sup>. Nevertheless, immunohistochemistry detected AsCas12a protein in the spleen.

Also, varying transfection efficiency is unlikely since Dr. Beate Rieblinger and I used the same protocol, equipment, and transfection reagent. Hence, reduced expression levels are not the reason for the differing efficiency.

Another reason could be the type of promoter. While the pol III promoter U6 expresses Cas9 guide RNAs, the strong pol II promoter CAG expressed the AsCas12a guides used in this study. In yeasts, the editing efficiency of AsCas12a is up 30 % higher when using a pol II promoter for guide expression than a pol III promoter<sup>142</sup> In mammalian HEK 293T cells, editing efficiencies for DNMT and EGFP were similar, regardless of the type of promoter used, e.g. U6 (pol III) or EF1a (pol II)<sup>41</sup>. Consequently, the expression of gRNAs via a pol II promoter probably did not reduce the efficiency in porcine cells, but remains to be confirmed.

AsCas12a has been used to generate transgenic pigs. This was achieved by transient transfection of cells used for SCNT, the pigs themselves did not express Cas12a<sup>172</sup>. Even for the transient expression of Cas12a, low editing efficiencies of the porcine genome were reported<sup>40</sup>. Moreover, no mammalian Cas12a-expressing model has been published so far. The first Cas9-expressing transgenic mouse was already published in 2014, shortly after the development of CRISPR/Cas9 for genome editing in mice<sup>173</sup>, and other species followed, such as pigs<sup>87,88</sup>, chickens<sup>87</sup>, or *Drosophila*<sup>174</sup>. Only for *Drosophila*, a transgenic Cas12a-expressing line exists. It is possible that others have tried and failed to generate functional Cas12a transgenic lines. However, failed experiments are rarely reported. Recent experiments to improve gene editing outcomes suggest, that the reasons for the low efficiency could be an innate property of Cas12a itself.

#### **4.2.2 AsCas12a requires engineering to achieve high editing rates reliably**

Several groups recently reported low editing efficiency using Lb/AsCas12a<sup>145,146,175-179</sup>, particularly in primary cells<sup>175</sup>. Consequently, various engineered variants of Cas12a were produced, which show higher activity<sup>175-179</sup>.

Luk et al. and Gier et al. claim a suboptimal nuclear localization of Cas12a as the reason for the mitigated efficiency of Cas12a<sup>146,175</sup>. The AsCas12a amino acid sequence contains two problematic nuclear export signals (NES), one at the conserved catalytic RuvC-II domain<sup>146</sup>. To prevent the nuclear export, they modified the original single nuclear localization signals (NLS) of AsCas12a. While Luk et al. fused 3 NLS signals to the C-terminus of the AsCas12a protein<sup>175</sup>, Gier et al. even used 6 NLS signals. Both groups replaced the original SV40 NLS with the more potent c-Myc NLS. Using the 'opAsCas12a' containing six C-terminal c-Myc NLS, Gier et al. achieved up to 32-fold higher knockout efficiency<sup>146</sup>.

DeWeirdt et al. also modified the NLS signals. For their 'enCas12a', they used two NLS, one N-terminal and one C-terminal. Furthermore, they introduced the mutations E174R, S542R, and

K548R. Gene editing efficiency of enCas12a was comparable to Cas9, when using a genome-wide library<sup>176</sup>.

I tested the nuclear localization hypothesis in 3.2.5. A 3xHA-tag was fused to the opAsCas12a to detect the opAsCas12a protein and its localization was compared to the protein expressed in cells from the AsCas12a transgenic pig. I could not detect a substantial increase in nuclear localization suggesting that the nuclear export is responsible for the low efficiency AsCas12a in porcine cells. Even though the presence of an N-terminal 3xHA-tag could potentially influence nuclear localization, this seems unlikely given that others have successfully used enCas12a with a 3xHA-tag<sup>180</sup>. Nonetheless, this possibility cannot be ruled out without further experimental validation.

Zhang et al. engineered AsCas12a with the mutations M537R and F870L, which increased efficacy without disturbing the high specificity of 'AsCas12a ultra'. They achieved editing efficiencies of up to 100 % and high knock-in rates in primary human cells, probably by selecting a variant that efficiently recognizes a TTTT PAM<sup>178</sup>.

Huang et al. used structure-guided protein engineering to generate 'AsCas12a-Plus'<sup>179</sup>, compared to directed evolution-derived AsCas12a ultra<sup>178</sup>. AsCas12a-Plus possesses the mutations E174R, R951K, and R955A, which increase specificity and editing efficiency 2-fold compared to WT AsCas12a<sup>179</sup>. Note that different mutations were used to increase efficiency.

These reports show that genetic engineering of AsCas12a has been required to generate an efficient genome editing tool. An optimal AsCas12a variant contains, in addition to an improved nuclear localization signal, mutations of amino acids to raise its efficacy comparable to Cas9. Furthermore, since AsCas12a provides an unspecific ssDNA *trans* cleavage activity, this must also be disrupted to achieve high knock-in rates using ssDNA repair templates<sup>177</sup>. The PAM site should be restricted to a distinct motif, unlike in the TYCV variant used in this study. Otherwise, the specificity could drop<sup>178</sup>. All this should result in an optimal AsCas12a version, which can be used for effective knockouts and knock-ins, comparable to Cas9, but with different attractive features. Unfortunately, due to this project's time scope, it was not possible to create a new pig line expressing an optimized version of Cas12a.

### **4.3 MMP2 downregulation could be hindered by large T antigen**

lncRNAs are a promising therapeutic target for hardly treatable cardiac disease. Pigs are an excellent translational model for this disease due to human-similar size and anatomy of the heart. The imprinted lncRNA maternally expressed gene 3 (Meg3) boosts cardiac fibrosis and

hypertrophic cardiomyopathy in mice<sup>102,130</sup>. Translating these observations into pigs could promote therapeutic approaches.

### 4.3.1 MEG3 knockout reveals important regulatory regions

Using luciferase assay indicated that the active *MEG3*-promoter resided in a 1 kb DNA fragment, including sequences 5' of exon 1 plus exon 1. Different knockout approaches (D-1, D-2, D-3) (see Figure 26 and Table 27) then narrowed down the region to 100 bp 5' of exon1. While D-1 and D-3 resulted in a loss of *MEG3* expression, *MEG3* expression was unaltered if 100 bp prior to exon 1 remained (D-2). Hence, the important proximal promoter element is located there. This aligned with the two other knockout strategies, E1-1 and E1-2, which also excised this region, in addition to exon 1, and also lead to a complete loss of *MEG3* and *MEG8* expression.

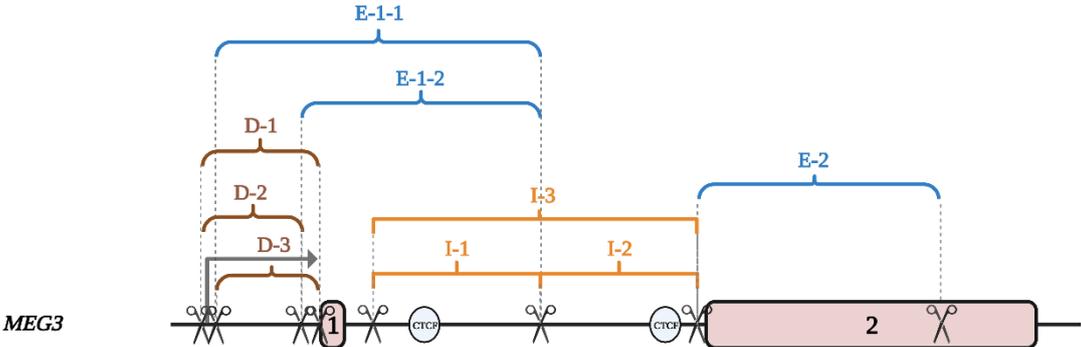


Figure 26 - Simplified scheme of *MEG3* KO strategies.

Table 27 - Impact of different *MEG3* KO approaches on *MEG3* and *MEG8* gene expression.

KO strategy	<i>MEG3</i> expression	<i>MEG8</i> expression
E-1-1	KO	KO
E-1-2	KO	KO
E-2	KO	→
D-1	KO	↑
D-2	→	→
D-3	KO	↑↑
I-1	→	→
I-2	→	→
I-3	→	↑

The change in *MEG3* expression levels in the D-1 and D-3 knockouts is unlikely to be the sole cause of altered *MEG8* expression. This is because different *MEG3* knockouts have resulted in varied *MEG8* expression patterns: no change in the strategy E-2, complete KO for E-1-2, and upregulation for D-1. Therefore, it is more likely that disruption of the promoter/DMR region, rather than changes at the transcript level, affects *MEG8* expression. BruUV-seq data shows that human *MEG3*, *MEG8*, and *MIRG* are expressed as one operon from a common transcriptional start site (TSS)<sup>99</sup>. Indeed, only the E-1-2 approach, which excises *MEG3* exon 1, including the putative TSS, leads to a combined loss of *MEG3* and *MEG8* expression, indicating a common TSS in porcine exon 1. Consequently, *MEG8* and *MEG3* are not driven by the same promoter but share the same TSS.

At least in mice, the DMR does not extend into the exon 2<sup>158</sup>. However, bisulfite sequencing (Figure 20) showed that one putative CTCF binding site shortly upstream of exon 2 is differentially methylated. Hence, the porcine *MEG3*-DMR includes most of intron 1.

In mice, two earlier approaches from Zhou et al.<sup>181</sup> and Takahashi et al.<sup>182</sup> use a PGK-promoter-driven neomycin HDR cassette to delete *Meg3* in embryonic stem cells (ESCs). However, the PGK-promoter probably enables the expression of the local genes even when the endogenous *Meg3*-promoter is excised<sup>96</sup>. The newer CRISPR/Cas9 strategy of Sanli et al. removes either the intron 1, the promoter, or almost the entire gene of *Meg3* in mice, similar to my technique<sup>96</sup>. All came to different conclusions regarding the regulation of *Dlk1* by *Meg3*. Takahashi and I did not observe differences in *Dlk1* expression in the *Meg3* knockouts cells<sup>182</sup> (Figure 22). However, *DLK1* expression was generally low in porcine cardiac fibroblasts. Zhou et al. did not observe a change in *Dlk1* expression for paternal deletion of *Meg3*, while maternal *Meg3* deletion resulted in an upregulation of *Dlk1*<sup>181</sup>. Sanli et al. used various CRISPR/Cas9 excisions to knock out *Meg3* in mice. While they observed a reduction in *Dlk1* expression, they did not see a complete loss of *Dlk1* in the *Meg3* KO cells<sup>96</sup>.

Since the approach from Sanli is most comparable to my approach using CRISPR/Cas9, I will mainly compare my results to theirs. They achieved a loss of *Meg3*, either by a homozygous 198 bp promoter deletion, a homozygous or maternal 2.2 kb intron 1 deletion, or a homozygous 10 kb deletion from intron 1 to 9. These excisions also led to a loss of *Meg8* and *Mirg* expression.

Contrary to Sanli et al., homozygous knockout of pig *MEG3* intron 1 strategies did not result in the loss of expression of any genes. Since the two putative CTCF sites are located in these excisions, there are probably other reasons for the differences. Either there are still other crucial CTCF sites downstream, which I did not target in my approaches<sup>183</sup>. Or, the intron 1, which is retained in mouse ESC<sup>96</sup> and essential to organize the function of *Meg3* in foci<sup>138</sup>, is not retained in pigs. Hence the excision of the intron could promote *Meg3* degradation in mice<sup>184</sup>.

In none of the porcine KO strategies, I saw a significant change in the expression of the protein-coding genes of the locus, *DLK1*, *DIO3*, and *RTL1*. However, these genes are highly tissue-specific and generally show a low expression in pig hearts, according to the pig RNA atlas<sup>185</sup>. This indicates a minor role of the protein-coding genes for porcine cardiac tissue.

#### **4.3.2 RNA-Seq does not show downregulation of *MMP2* via *MEG3* KO due to large T antigen immortalization of cardiac fibroblasts**

Although all three KO approaches, E1-2, D-1, and E-2, lead to a loss of *MEG3* expression (Figure 22), there was no significant change in gene expression in RNA-Seq (Figure 25). One possibility could be an interaction between the *MEG3* and the *MEG8* transcript to change the expression of other genes. However, such an interaction has yet to be reported.

These inconsistent results are in line with the literature<sup>96</sup>. Sanli et al., who performed similar knockout studies in mouse ESCs, also did not observe consistent changes in gene expression by RNA-Seq<sup>96</sup>, except for the genes within the gene edited locus.

In contrast, for strategies D-1 and E-2, I observed consistent changes in gene expression. On single gene level, *ACTA2* was the most downregulated gene for both approaches. *ACTA2* is not crucial for differentiating cardiac fibroblasts into myofibroblast but it is the most widely accepted marker of myofibroblast differentiation<sup>186</sup>. Consequently, cardiac fibroblasts harboring the modification D-1 and E-2 might differentiate less into myofibroblasts, the cell type playing the major role in pathological cardiac remodeling<sup>187</sup>.

Cardiac fibrosis is also defined by a massive deposition of ECM, such as collagen. *COL5A1* and *COL8A1* are significantly reduced in D-1, and a lack of *COL8* reduces cardiac fibrosis<sup>188</sup>.

In pathway analysis, Piccoli used LNA-GapmeRs to downregulate *Meg3* in murine cardiac fibroblasts *in vitro*. When they knocked down *Meg3* *in vivo*, they could reduce pathological cardiac fibrosis via the reduction of *Mmp2*.

Even though *TGF-β* pathway genes are downregulated in porcine *MEG3* KO D-1, the KO of *MEG3* did not lead to a decrease in *MMP2* expression. This discrepancy between mice and pigs could be caused by the different cells used. Piccoli et al. used primary cardiac fibroblasts to knockdown *Meg3*, while I used SV40 large T-antigen immortalized porcine cardiac fibroblasts for the knockout. Interestingly, the large T-antigen reduces the binding of P53 to DNA, thereby preventing target gene regulation<sup>189</sup>. Indeed, Piccoli et al. showed that binding of P53 to the *Mmp2* promoter induces its expression when stimulated with *TGF-β*. Consequently, the immortalization

could prevent the *TGF-β*-mediated binding of P53 to the *MMP2* promoter in my knockout approaches.

Furthermore, this could explain why there is no significant change in gene expression in the overexpression and knockdown experiments (CRISPRa / CRISPRi). Even though there was a direct impact on *MEG8* due to local activation/repression of *MEG3* gene expression, there was no significant gene expression change elsewhere in the genome. Since *MEG3* is known to interact with P53<sup>114,120,190</sup>, the immortalization approach could have prevented the *MEG3*'s effect via P53 on the regulation of other genes.

Concluding, I found effective knockout, overexpression, and knockdown strategies for porcine *MEG3*. One strategy, D-1, even reduced the main pathway in cardiac fibrosis, *TGF-β*, similar to experiments in mice. To further investigate the impact on cardiac fibrosis and rule out the possible impacts of the large T antigen-mediated immortalization, the best option is a hypertrophic porcine *in vivo* model, which has already been published<sup>137</sup>. Alternatively, the immortalization approach using telomerase reverse transcriptase (TERT), which does not impact *P53*, could be applied to porcine cardiac fibroblasts to generate a new *in vitro* model<sup>191</sup>.

## 5 Concluding remarks and outlook

In this study, I demonstrated that the CRISPlace approach efficiently boosts transgene insertion in porcine somatic cells to generate donor cells for SCNT. Using microhomologous sequences could allow a further development of this method, e.g. combining the efficiency of NHEJ with the accuracy of HDR<sup>192</sup>.

Studies on pigs must meet high ethical standards. Reducing the number of animals needed for translational studies is desirable. This can be achieved by *in vivo* genome editing of using pig lines expressing Cas proteins. Even though I successfully generated an AsCas12a-expressing pig line, editing cells derived from these pigs showed low efficiency. A new pig line expressing an engineered Cas12a variant like Cas12a-Plus would be potentially beneficial. However, the generation of new transgenic pigs takes years, and the development of new CRISPR/Cas-based methods advances quickly. To circumvent the generation of new pigs, the expression of sgRNAs for Cas9 can also be achieved by tissue-specific pol II promoters<sup>193</sup>, which could be delivered into the existing Cas9-expressing pigs<sup>87</sup>. Other approaches could be the delivery of RNP complexes via engineered nanoparticles<sup>194</sup> or the delivery of hypercompact Cas12a with guides using AAVs in wild-type pigs<sup>195</sup>.

Targeting lncRNAs in difficult to treat cardiac fibrosis could be game-changing. Therefore, the successfully generated CRISPR/Cas-based strategies modifying the expression of *MEG3* are promising therapeutic targets for cardiac fibrosis. I provide an efficient strategy for knocking out *MEG3*, decreasing the main fibrotic pathway *TGF-β*.

Currently, the European Union even funded a 2.5 million project for the preclinical development of an antisense oligonucleotide targeting *MEG3* in cardiac fibrosis<sup>196</sup>, showing the therapeutic potential of *MEG3*.

Furthermore, *Meg3* is also linked to other diseases. For example, *Meg3* is dysregulated in pancreatic ductal adenocarcinoma (PDAC)<sup>190,197</sup>. The provided modulation strategies could be used to investigate its role in a porcine PDAC model, which has been developed at our chair. Preliminary tests have already shown aberrant methylation and expression patterns of *MEG3* in PDAC tissue.

## 6 Abbreviations

AAVs	.....	<i>adeno-associated viruses</i>
CAG	.....	<i>CMV enhancer/chicken beta-actin</i>
Cas	.....	<i>CRISPR-associated nucleases</i>
ceRNA	.....	<i>competing endogenous RNA</i>
CFi	.....	<i>immortalized porcine cardiac fibroblasts</i>
CRE	.....	<i>cAMP response elements</i>
CRISPR	.....	<i>Clustered Regularly Interspaced Short Palindromic Repeats</i>
CRISPRa	.....	<i>CRISPR activation</i>
CRISPRi	.....	<i>CRISPR interference</i>
CTCF	.....	<i>CCCTC-binding factor</i>
ddPCR	.....	<i>droplet digital PCR</i>
Dio3	.....	<i>iodothyronine deiodinase 3</i>
Dlk1	.....	<i>delta like non-canonical Notch ligand 1</i>
DMR	.....	<i>differentially methylated region</i>
DNMT	.....	<i>DNA methyltransferases</i>
DR	.....	<i>direct repeat</i>
DSB	.....	<i>double-strand break</i>
Fluc	.....	<i>firefly luciferase LUC2</i>
GAPDH	.....	<i>Glycerinaldehyd-3-phosphat-Dehydrogenase</i>
GGTA	.....	<i>Alpha-1,3-galactosyltransferase 1</i>
GM	.....	<i>genetically modified</i>
HA	.....	<i>hemagglutinin</i>
HDAC	.....	<i>histone deacetylases</i>
HDR	.....	<i>homology-directed repair</i>
HITI	.....	<i>homology-independent integration</i>
ICE	.....	<i>interference of CRISPR edits</i>
ICR	.....	<i>imprinting control regions</i>
indels	.....	<i>insertion-deletion</i>
iPSC	.....	<i>induced pluripotent stem cells</i>
iPSCs	.....	<i>Induced pluripotent stem cell</i>
ITR	.....	<i>inverted terminal repeats</i>
KD	.....	<i>knockdown</i>
KRAB	.....	<i>Krüppel-associated box</i>
lncRNA	.....	<i>long non-coding RNA</i>
MeCP2	.....	<i>methyl-CpG-binding protein 2</i>
Meg8	.....	<i>maternally expressed gene 8</i>
MI	.....	<i>pronuclear microinjection</i>
MMP	.....	<i>matrix metalloproteases</i>
NEAT1	.....	<i>nuclear enriched abundant transcript 1</i>
NES	.....	<i>nuclear export signals</i>
NFκB	.....	<i>nuclear factor 'kappa-light-chain-enhancer' of activated B-cells</i>
NHEJ	.....	<i>non-homologous end joining</i>
NLS	.....	<i>nuclear localization signal</i>
Nluc	.....	<i>NanoLuc® luciferase</i>
OE	.....	<i>overexpression</i>
PADMSC	.....	<i>Porcine adipose-derived mesenchymal stem cells</i>
PAM	.....	<i>protospacer adjacent motif</i>
PBEC	.....	<i>porcine bladder epithelial cells</i>
PCA	.....	<i>principal component analysis</i>
PCF	.....	<i>porcine cardiac fibroblast</i>

PDAC .....	<i>pancreatic ductal adenocarcinoma</i>
PEC.....	<i>porcine heart endothelial cells</i>
PEF.....	<i>porcine ear fibroblast</i>
PKF.....	<i>porcine kidney fibroblast</i>
pRb.....	<i>retinoblastoma protein</i>
PRC2.....	<i>Polycomb repressive complex 2</i>
RACE.....	<i>Rapid amplification of cDNA ends</i>
Rian.....	<i>RNA imprinted and accumulated in nucleus</i>
RMCE.....	<i>recombinase-mediated cassette exchange</i>
ROSA26.....	<i>Reverse Oriented Splice Acceptor, Clone 26</i>
RT .....	<i>room temperature</i>
Rtl1 .....	<i>retrotransposon Gag like 1</i>
Rtl1as.....	<i>Rtl1 antisense</i>
SA.....	<i>splice acceptor</i>
SAM.....	<i>Synergistic Activation Mediator</i>
SB.....	<i>sleeping beauty</i>
snoRNAs .....	<i>small nucleolar RNAs</i>
STAT3 .....	<i>signal transducer and activator of transcription 3</i>
TAC .....	<i>transverse aortic constriction</i>
TERT.....	<i>telomerase reverse transcriptase</i>
TGF- $\beta$ .....	<i>Transforming Growth Factor beta</i>
TIMP.....	<i>tissue inhibitors of metalloproteinases</i>
TRH.....	<i>thyrotropin-releasing hormone</i>
TSS .....	<i>transcriptional start site</i>
UCP1.....	<i>Uncoupling Protein 1</i>
UHRF1 .....	<i>ubiquitin-like with PHD and ring finger domains 1</i>
VPR .....	<i>Vp64-p65-Rta</i>
$\alpha$ -SMA .....	<i>alpha-smooth muscle actin</i>
$\beta$ 2M .....	<i>Beta-2-Microglobulin</i>

## 7 List of tables

Table 1 - Selected <i>cis</i> and <i>trans</i> actions of Meg3.....	21
Table 2 - Oligonucleotides for luciferase assay and methylation assays .....	24
Table 3 - q-RT-PCR and ddPCR oligonucleotides .....	25
Table 4 - CRISPR/Cas9 sgRNAs targeting <i>MEG3</i> .....	26
Table 5 - CRISPR/Cas12a sgRNAs .....	27
Table 6 - DNA vectors .....	27
Table 7 - Antibodies .....	29
Table 8 - Chemicals.....	29
Table 9 - Mammalian and bacterial cells.....	31
Table 10 - Buffers .....	32
Table 11 - Tissue culture reagents and supplements.....	33
Table 12 - Tissue culture media.....	34
Table 13 - Kits .....	35
Table 14 - Consumables .....	35
Table 15 - Enzymes and enzyme buffers .....	36
Table 16 - Equipment.....	36
Table 17 - Software.....	37
Table 18 - GoTaq® reaction components and conditions.....	42
Table 19 - Q5® reaction components and conditions .....	43
Table 20 - PyroMark reaction components and conditions.....	43
Table 21 - ddPCR components and conditions.....	44
Table 22 - qPCR components and cycling conditions.....	45
Table 23 - Composition of western blot separation- and collection gels.....	47
Table 24 - Transgene insertion efficiency of the gene targeting and CRISPlace approach.....	57
Table 25 - sgRNA efficiencies in AsCas12a derived cells for B2M and GGTA1.....	63
Table 26 - Efficiency comparison in porcine Cas9 and Cas12a cells.....	66
Table 27 - Impact of different <i>MEG3</i> KO approaches on <i>MEG3</i> and <i>MEG8</i> gene expression. ....	90

## 8 List of figures

Figure 1 - Generation of genetically modified pigs.....	7
Figure 2 - AsCas12a (TYCV) structure and target recognition.....	10
Figure 3 - Transcriptional modulation by dCas9 variants.....	12
Figure 4 - Transgene insertion via recombination-mediated cassette exchange (RMCE).....	16
Figure 5 - Schematic presentation of the human and murine DLK1-DIO3 locus.....	18
Figure 6 - Regulation of MEG3 expression.....	20
Figure 7 - Schematic illustration of the pro-fibrotic and pro-hypertrophic action of MEG3.....	22
Figure 8 - Structure of the ROSA26-AsCas12a HDR gene trap vector.....	55
Figure 9 - The CRISPlace principle for homology-independent based transgene insertion. ....	56
Figure 10 - AsCas12a transgene copy number in CRISPlace compared to gene targeting. ....	58
Figure 11 - The founder pig #2026 is suitable for breeding. ....	59
Figure 12 - Expression of AsCas12a in various tissues of the F1 generation (#2417).....	60
Figure 13 - AsCas12a protein is present in various tissues of F1 generation (#2417).....	61
Figure 14 - Map of the vector used to evaluate the functionality of cells derived from the Cas12a-transgenic pig.....	63
Figure 15 - AsCas12a functionality in primary porcine heart epithelial cells (PEC) derived from AsCas12a transgenic pig #2417.....	65
Figure 16 - Nuclear and cytoplasmic Cas12a distribution in OpCas12a vs. AsCas12a. ....	67
Figure 17 - Porcine MEG3 intron-exon structure.....	68
Figure 18 - Cellular distribution of selected RNA.....	69
Figure 19 - <i>In silico</i> predictions for MEG3.....	70
Figure 20 - Luciferase- and methylation assay of the regulatory region of MEG3. ....	71
Figure 21 - Different KO strategies for MEG3. ....	75
Figure 22 - Expression of relevant genes of the DLK1-DIO3 locus in MEG3 KO single-cell clones.....	77
Figure 23 - MEG3 target area for sgRNAs (A) using CRISPRa and CRISPRi and corresponding vectors (B).....	79
Figure 24 - Overexpression and knockdown of lncRNA MEG3. ....	81
Figure 25 - MEG3 knockout in cardiac fibroblast attenuates fibrotic signaling in bulk RNA-seq. ....	82
Figure 26 - Simplified scheme of MEG3 KO strategies. ....	90

## 9 Supplementary

Table S. 1 - Overview of the MEG3 methylation assays.

		MEG3 methylation assay FR2-1								MEG3 methylation assay FR2-2				
	% CpG methylated	Pos 1	Pos 2	Pos 3	Pos 4	Pos 5	Pos 6	Pos 7	Pos 8	Pos 1	Pos 2	Pos 3	Pos 4	Pos 5
Female adult	Brain	33	27	53	15	18	28	18	16	30	28	28	25	29
	Kidney	42	29	72	16	23	31	20	18	34	27	26	26	31
	Spleen	37	30	68	15	23	31	19	18	31	27	25	24	31
	Lung	42	30	70	16	24	32	2	19	33	30	29	26	34
	Liver	34	26	60	15	21	27	19	17	31	26	28	23	32
	Pancreas	39	26	62	15	19	28	20	18	31	28	29	25	33
	Muscle	28	20	47	12	15	23	16	15	28	24	24	22	27
	Heart	41	28	68	16	23	30	22	19	34	32	32	31	36
	Lymph node	44	33	75	18	26	32	22	20	34	32	31	28	37
	Colon	38	26	58	15	20	27	17	17	26	25	26	21	28
Stomach	33	30	58	19	21	40	28	19	33	28	28	28	33	
Male adult	Brain	34	32	50	20	11	39	29	19	34	30	31	28	31
	Kidney	35	35	57	21	19	37	31	17	31	27	26	26	30
	Spleen	34	39	70	24	21	43	32	21	31	28	29	25	32
	Lung	37	34	56	19	22	36	27	18	33	29	27	27	32
	Liver	41	29	64	17	24	33	22	17	31	27	26	25	32
	Pancreas	42	30	63	18	20	33	23	18	32	30	30	27	31
	Muscle	24	22	47	15	18	25	18	14	27	18	23	21	25
	Heart	37	30	60	19	23	25	23	18	35	31	31	29	35
	Lymph node	36	29	65	15	21	31	20	16	34	30	29	27	33
	Colon	40	30	64	17	22	34	21	19	38	33	34	30	36
Stomach	34	27	60	16	23	30	21	17	35	33	32	30	34	
Female embryo	Liver	36	27	57	19	24	34	25	19	34	30	32	26	33
	Muscle	38	32	50	20	23	33	26	18	35	35	32	28	33
	Kidney	44	37	62	20	25	38	26	19	37	33	34	30	35
	Heart	44	33	65	19	25	34	24	18	36	33	35	31	37
	Bone	44	33	64	18	26	36	24	20	36	31	30	27	32
Lung	45	35	68	20	28	39	26	22	39	34	35	33	35	
Male embryo	Liver	39	27	64	18	25	34	23	19	35	32	33	28	34
	Muscle	44	35	69	20	27	40	26	21	37	35	36	34	36
	Kidney	46	38	68	21	27	40	25	20	40	37	40	34	39
	Bone	37	31	60	18	20	36	23	19	34	31	30	27	32
	Lung	32	29	52	17	23	34	22	18	32	30	31	29	31
Female embryo	Liver	35	31	63	20	25	33	21	20	34	32	32	28	36
	Muscle	35	32	51	20	24	41	28	19	35	31	34	30	35
	Kidney	50	43	71	24	27	44	31	25	45	40	40	34	42

	Heart	43	39	69	23	27	43	26	21	39	36	35	32	39
	Bone	41	36	59	22	24	41	28	20	37	33	37	31	32
	Lung	45	34	67	19	25	38	24	19	39	34	32	31	37
Male embryo	Liver	42	27	68	17	23	31	21	19	34	32	32	28	36
	Muscle	33	30	52	19	22	34	26	19	36	29	32	28	33
	Kidney	41	33	67	22	18	41	28	22	39	35	36	33	36
	Heart	41	33	64	20	25	39	26	31	35	29	31	28	33
	Bone	42	31	60	18	34	24	23	21	32	27	29	24	29
	Lung	45	37	79	21	19	38	26	26	39	34	32	31	37
Blastocysts (7 days old)	Bl. Nr.1	6	5	9	4	3	6	4	3	4	4	6	3	3
	Bl. Nr.2	6	6	7	2	4	5	4	3	3	6	5	3	4
	Bl. Nr.3	13	6	10	4	5	9	6	4	5	4	5	4	6
	Bl. Nr.4	6	5	9	4	3	8	5	3	7	5	7	5	5
	Bl. Nr.5	4	7	6	3	3	6	4	4	6	5	5	2	6
	Bl. Nr.6	5	6	7	3	16	6	12	4	5	4	4	4	4
	Bl. Nr.7	24	22	39	11	11	21	16	12	22	18	20	15	21
	Bl. Nr.8	15	9	8	5	3	0	9	6	7	5	6	5	5
MEG3 methylation assay CTCF														
Female adult	% CpG methylated	Pos 1	Pos 2	Pos 3	Pos 4	Pos 5	Pos 6	Pos 7	Pos 8	Pos 9				
	Brain	14	14	10	14	22	26	28	13	24				
	Heart	13	21	25	11	21	15	44	23	19				
	Lymph node	15	14	22	9	27	11	40	13	15				
	Muscle	13	11	23	6	22	12	44	9	14				
	Spleen	15	17	24	8	26	12	44	14	16				
Male adult	Brain	8	24	30	16	24	22	37	21	24				
	Heart	14	18	23	11	23	17	38	18	19				
	Lymph node	15	17	28	9	24	14	43	15	19				
	Muscle	12	13	21	6	19	10	36	9	13				
	Spleen	14	18	31	10	27	15	46	18	22				

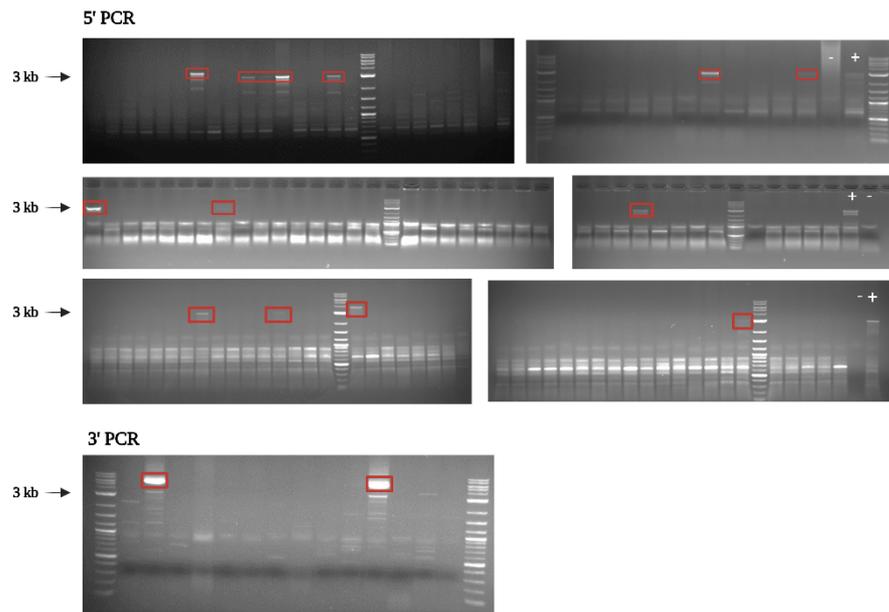


Figure S. 1 - 5'- and 3'-junction PCR to identify correctly targeted AsCas12a clones by homology-mediated targeting. Red boxes indicate clones considered as positive.

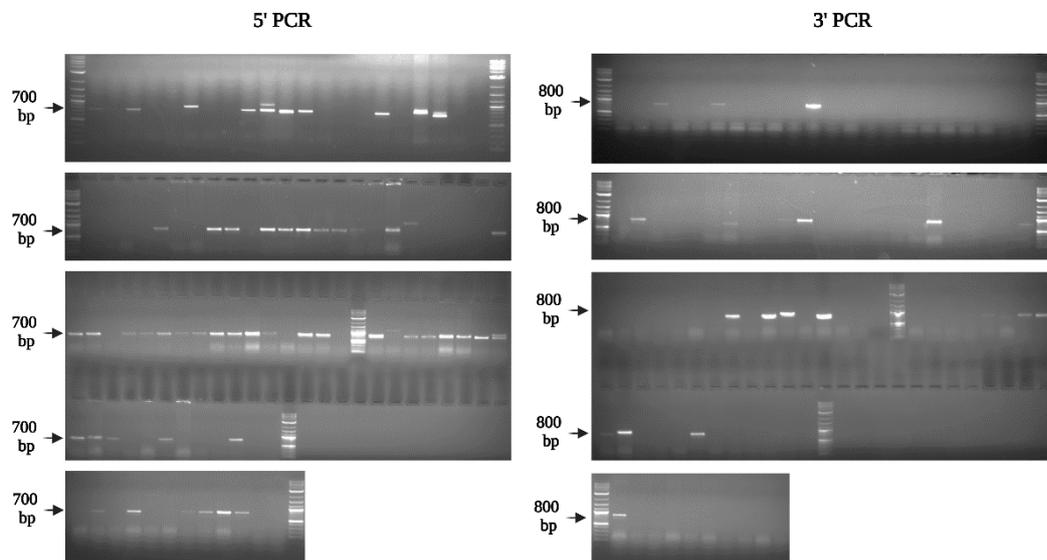


Figure S. 2 - 5'- and 3'-junction PCR to identify correctly targeted AsCas12a clones by CRISPlace.

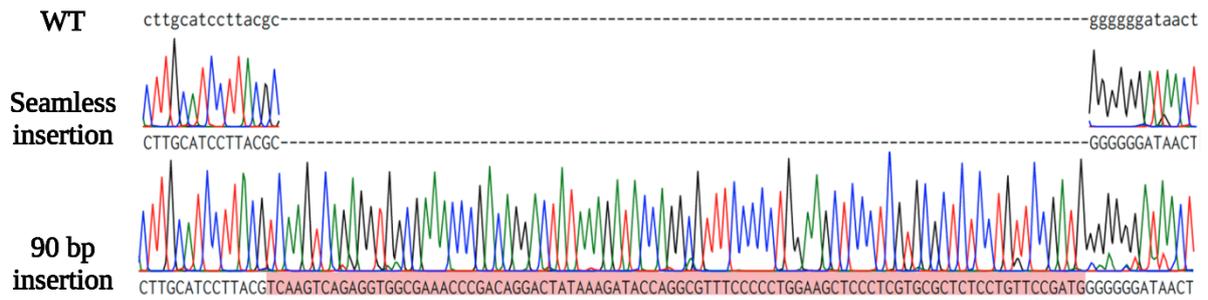


Figure S. 3 - Sequencing of CRISPlace insertion sites reveals insertions up to 90 bp but also seamless insertions.

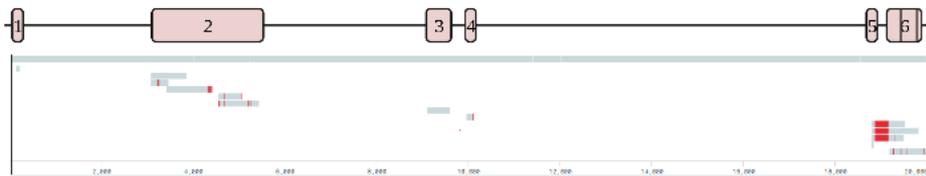
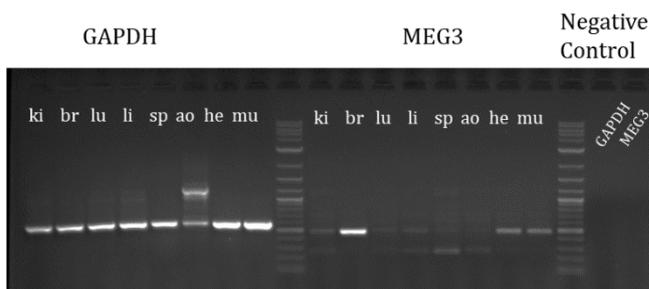


Figure S. 4 - Sequence alignments of cDNA amplicons for porcine MEG3 intron exon structure.



Figure S. 5 - Open reading frames in porcine MEG3 exon 1 and 2.



Ki = kidney, br = brain, lu = lung, li = liver, sp = spleen, ao = aorta, he = heart, mu = muscle

Figure S. 6 - RT-PCR of GAPDH and MEG3 in different WT pig tissues.

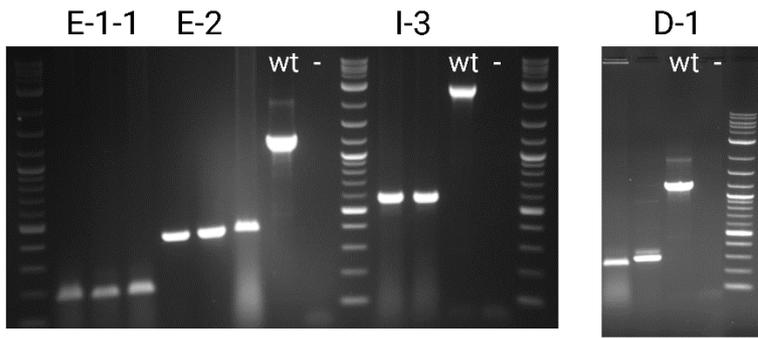


Figure S. 7 - PCR to verify homozygous excision of different *MEG3* KO single cell clones.

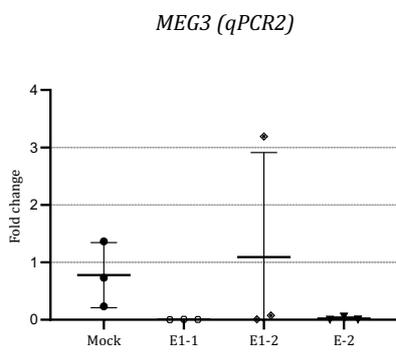


Figure S. 8 - *MEG3* alternative qPCR of single cell clones

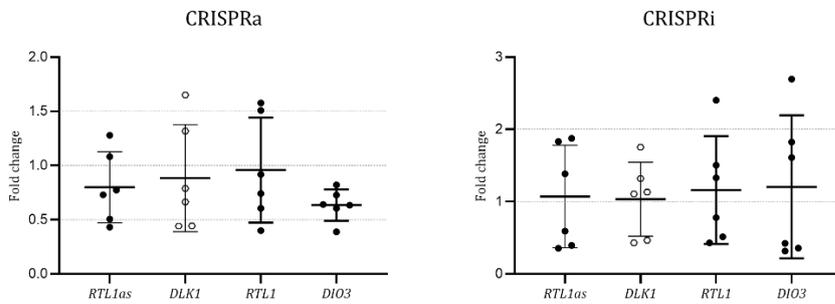


Figure S. 9 - Relative expression of genes in the DLK1-DIO3 locus in knockdown and activation of *MEG3*.



## 10 Acknowledgements

Zuerst möchte ich mich ganz herzlich bei Prof. Angelika Schnieke bedanken, die mir ermöglicht hat eine Promotion, in einem der spannendsten Felder die es gibt, durchzuführen. Ihre klugen Einfälle, aber auch ihre Bodenständigkeit trotz zahlloser Erfolge sind wirklich vorbildhaft. Danke auch an Prof. Benjamin Schusser, für das kommissarische Management unseres Lehrstuhls, das viel extra Arbeit für ihn war. Weiterhin danke ich PD Dr. habil. Krzysztof Flisikowski für die Korrekturen und die anregenden Gespräche über Forschung im Allgemeinen.

Ganz besonderer Dank gebührt natürlich meiner Betreuerin Dr. Beate Rieblinger. Du warst immer freundlich, zuverlässig und hilfsbereit. Keine Mühe war dir zu groß, um mich zu unterstützen. Ich weiß, dass es nur wenige Betreuerinnen gibt, die sich so einsetzen für ihre Studenten, wie Du. Ich wünsche dir nur das Beste für deine weitere Karriere, das junge Familienglück und das Leben im Spargelparadies.

Danke auch an meine Freunde im Embryoteam, Thomas, Alex, Wei und Bernhard. Ohne euren spätabendlichen Einsatz wäre die Arbeit an unserem Lehrstuhl nicht möglich gewesen. Was unseren Lehrstuhl besonders gemacht hat, war der Zusammenhalt und das tolle Kollegium, das wir hatten. Die Arbeitsatmosphäre war hilfsbereit, freundlich und frei. Dazu haben wirklich alle beigetragen, unsere Postdocs, unsere TAs, unsere Doktoranden und Doktorandinnen. Danke dafür. Hier will ich mich ganz besonders bei Laura bedanken, die mir eine treue Freundin geworden ist. Mit dir waren die Ergebnisse, die oft zum Verzweifeln waren, nur halb so schlimm. Deine fröhliche Art, der gemeinsame Kaffee und das gemeinsame Büro, werden mir am meisten fehlen.

Weiterhin danke ich meinem Studenten Paul Wätzig und meinen Studentinnen Inna Nikitina und Anika Eckstein. Die Arbeit mit euch gehörte zum besten Teil meiner Promotion, hat mir super viel Freude bereitet und mich extrem weitergebracht, sowohl persönlich als auch durch eure Beiträge für mein Projekt. Ich hoffe, dass ihr auch durch mich das ein oder andere gelernt habt, das ihr in eurem Lebensweg verwenden könnt.

Über alles Berufliche hinaus möchte ich mich ganz besonders bei meiner besten Freundin, Partnerin und Frau Konstanze bedanken. Die unzähligen Male, die du mich wiederaufgebaut und mit mir gelacht hast, sind von unschätzbarem Wert. Kein Acknowledgement könnte ausreichend sein, um zu beschreiben wie wichtig du mir bist. Es gibt keine bessere Partnerin als dich.

Zuletzt aber am allerwichtigsten, Danke an meine Eltern. Mein Papa Ludwig und meine Mama Vefi haben unermüdlich auf dem Milchviehbetrieb gearbeitet, um mir diese Ausbildung zu ermöglichen. Trotz eines Jobs, der oft undankbar ist, habt ihr mich ohne zu zögern bedingungslos unterstützt. Das werde ich euch nie vergessen.

## 11 References

1. Reardon S. First pig-to-human heart transplant: what can scientists learn? *Nature*. 2022;601(7893):305–306. doi:10.1038/d41586-022-00111-9.
2. Wu J, Platero-Luengo A, Sakurai M, Sugawara A, Gil MA, Yamauchi T, Suzuki K, Bogliotti YS, Cuello C, Morales Valencia M, et al. Interspecies Chimerism with Mammalian Pluripotent Stem Cells. *Cell*. 2017;168(3):473-486.e15. doi:10.1016/j.cell.2016.12.036.
3. KAROLINSKA INSTITUTET, editor. The Nobel Prize in Physiology or Medicine 1977. [place unknown]: [publisher unknown]. <https://www.nobelprize.org/prizes/medicine/1977/press-release/>.
4. Skarnes WC. Is mouse embryonic stem cell technology obsolete? *Genome Biol*. 2015;16(1):109. <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-015-0673-6>. doi:10.1186/s13059-015-0673-6.
5. Choi K-H, Lee D-K, Oh J-N, Kim S-H, Lee M, Woo S-H, Kim D-Y, Lee C-K. Pluripotent pig embryonic stem cell lines originating from in vitro-fertilized and parthenogenetic embryos. *Stem Cell Research*. 2020;49:102093. <https://www.sciencedirect.com/science/article/pii/S1873506120303949>. doi:10.1016/j.scr.2020.102093.
6. Gao X, Nowak-Imialek M, Chen X, Chen D, Herrmann D, Ruan D, Chen ACH, Eckersley-Maslin MA, Ahmad S, Lee YL, et al. Establishment of porcine and human expanded potential stem cells. *Nat Cell Biol*. 2019;21(6):687–699. doi:10.1038/s41556-019-0333-2.
7. Zhi M, Zhang J, Tang Q, Yu D, Gao S, Gao D, Liu P, Guo J, Hai T, Gao J, et al. Generation and characterization of stable pig pregastrulation epiblast stem cell lines. *Cell Res*. 2022;32(4):383–400. <https://www.nature.com/articles/s41422-021-00592-9>. doi:10.1038/s41422-021-00592-9.
8. Fan N, Chen J, Shang Z, Dou H, Ji G, Zou Q, Wu L, He L, Wang F, Liu K, et al. Piglets cloned from induced pluripotent stem cells. *Cell Res*. 2013;23(1):162–166. <https://www-nature-com.eaccess.tum.edu/articles/cr2012176>. doi:10.1038/cr.2012.176.
9. Perleberg C, Kind A, Schnieke A. Genetically engineered pigs as models for human disease. *Dis Model Mech*. 2018;11(1). <https://journals-biologists-com.eaccess.tum.edu/dmm/article/11/1/dmm030783/2587/Genetically-engineered-pigs-as-models-for-human>. doi:10.1242/dmm.030783.
10. Hammer RE, Pursel VG, Rexroad CE, Wall RJ, Bolt DJ, Ebert KM, Palmiter RD, Brinster RL. Production of transgenic rabbits, sheep and pigs by microinjection. *Nature*. 1985;315(6021):680–683. <https://www-nature-com.eaccess.tum.edu/articles/315680a0>. doi:10.1038/315680a0.
11. Nagashima H, Fujimura T, Takahagi Y, Kurome M, Wako N, Ochiai T, Esaki R, Kano K, Saito S, Okabe M, et al. Development of efficient strategies for the production of genetically modified

- pigs. *Theriogenology*. 2003;59(1):95–106. <https://www.sciencedirect.com/science/article/pii/S0093691X0201261X>. doi:10.1016/S0093-691X(02)01261-X.
12. Kurome M, Geistlinger L, Kessler B, Zakhartchenko V, Klymiuk N, Wuensch A, Richter A, Baehr A, Kraehe K, Burkhardt K, et al. Factors influencing the efficiency of generating genetically engineered pigs by nuclear transfer: multi-factorial analysis of a large data set. *BMC Biotechnol*. 2013;13(1):43. <https://bmcbiotechnol.biomedcentral.com/articles/10.1186/1472-6750-13-43>. doi:10.1186/1472-6750-13-43.
  13. Wilmut I, Schnieke AE, McWhir J, Kind AJ, Campbell KH. Viable offspring derived from fetal and adult mammalian cells. *Nature*. 1997;385(6619):810–813. doi:10.1038/385810a0.
  14. Le Cong, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013;339(6121):819–823. doi:10.1126/science.1231143.
  15. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012;337(6096):816–821. doi:10.1126/science.1225829.
  16. Jon Cohen. 'Any idiot can do it.' Genome editor CRISPR could put mutant mice in everyone's reach. [place unknown]: [publisher unknown]; 2016.
  17. Wei X, Henke VG, Strübing C, Brown EB, Clapham DE. Real-time imaging of nuclear permeation by EGFP in single intact cells. *Biophys J*. 2003;84(2 Pt 1):1317–1327. <https://www.sciencedirect.com/science/article/pii/S0006349503749479>. doi:10.1016/S0006-3495(03)74947-9.
  18. Maggio I, Zittersteijn HA, Wang Q, Liu J, Janssen JM, Ojeda IT, van der Maarel SM, Lankester AC, Hoeben RC, Gonçalves, Manuel A. F. V. Integrating gene delivery and gene-editing technologies by adenoviral vector transfer of optimized CRISPR-Cas9 components. *Gene Ther*. 2020;27(5):209–225. <https://www.nature.com/articles/s41434-019-0119-y>. doi:10.1038/s41434-019-0119-y.
  19. Kikuchi K, Ekwall H, Tienthai P, Kawai Y, Noguchi J, Kaneko H, Rodriguez-Martinez H. Morphological features of lipid droplet transition during porcine oocyte fertilisation and early embryonic development to blastocyst in vivo and in vitro. *Zygote*. 2002;10(4):355–366. doi:10.1017/S0967199402004100.
  20. Hirata M, Wittayarat M, Tanihara F, Sato Y, Namula Z, Le QA, Lin Q, Takebayashi K, Otoi T. One-step genome editing of porcine zygotes through the electroporation of a CRISPR/Cas9 system with two guide RNAs. *In Vitro Cell Dev Biol Anim*. 2020;56(8):614–621. <https://link.springer.com/article/10.1007/s11626-020-00507-9>. doi:10.1007/s11626-020-00507-9.

21. Mehravar M, Shirazi A, Nazari M, Banan M. Mosaicism in CRISPR/Cas9-mediated genome editing. *Developmental biology*. 2019;445(2):156–162. <https://www.sciencedirect.com/science/article/pii/S0012160618302513>. doi:10.1016/j.ydbio.2018.10.008.
22. Park KW, Cheong HT, Lai L, Im GS, Kühholzer B, Bonk A, Samuel M, Rieke A, Day BN, Murphy CN, et al. Production of nuclear transfer-derived swine that express the enhanced green fluorescent protein. *Anim Biotechnol*. 2001;12(2):173–181. doi:10.1081/abio-100108344.
23. Kurome M, Kessler B, Wuensch A, Nagashima H, Wolf E. Nuclear Transfer and Transgenesis in the Pig. In: *Nuclear Reprogramming*. [place unknown]: Humana Press, New York, NY; 2015. p. 37–59.
24. Yarmolinsky M, Hoess R. The Legacy of Nat Sternberg: The Genesis of Cre-lox Technology. *Annu Rev Virol*. 2015;2(1):25–40. doi:10.1146/annurev-virology-100114-054930.
25. Kim H, Kim M, Im S-K, Fang S. Mouse Cre-LoxP system: general principles to determine tissue-specific roles of target genes. *Lab Anim Res*. 2018;34(4):147–159. doi:10.5625/lar.2018.34.4.147.
26. Luo W, Li Z, Li P, Huang Y, Han Y, Yao C, Zhang Z, Yan H, Pang D, Ouyang H, et al. Expression of Cre recombinase in alveolar epithelial cells of the AQP2-Cre transgenic mini-pigs. *Cell Physiol Biochem*. 2014;34(5):1597–1613. doi:10.1159/000366363.
27. Kalla D, Flisikowski K, Yang K, Sangüesa LB, Kurome M, Kessler B, Zakhartchenko V, Wolf E, Lickert H, Saur D, et al. The Missing Link: Cre Pigs for Cancer Research. *Front Oncol*. 2021;11:755746. doi:10.3389/fonc.2021.755746.
28. Song Y, Lai L, Li L, Huang Y, Wang A, Tang X, Pang D, Li Z, Ouyang H. Germ cell-specific expression of Cre recombinase using the VASA promoter in the pig. *FEBS Open Bio*. 2016;6(1):50–55. doi:10.1002/2211-5463.12005.
29. Gaj T, Gersbach CA, Barbas CF. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol*. 2013;31(7):397–405. doi:10.1016/j.tibtech.2013.04.004.
30. Khan S, Sallard E. Current and Prospective Applications of CRISPR-Cas12a in Pluricellular Organisms. *Mol Biotechnol*. 2023;65(2):196–205. <https://link.springer.com/article/10.1007/s12033-022-00538-5>. doi:10.1007/s12033-022-00538-5.
31. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A, et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*. 2015;163(3):759–771. doi:10.1016/j.cell.2015.09.038.
32. Kim D, Kim J, Hur JK, Been KW, Yoon S-H, Kim J-S. Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat Biotechnol*. 2016;34(8):863–868. doi:10.1038/nbt.3609.
33. Gao L, Cox DB, Yan WX, Manteiga JC, Schneider MW, Yamano T, Nishimasu H, Nureki O, Crosetto N, Zhang F. Engineered Cpf1 variants with altered PAM specificities increase genome targeting range. *Nat Biotechnol*. 2017;35(8):789–792. doi:10.1038/nbt.3900.

34. Wörle E, Jakob L, Schmidbauer A, Zinner G, Grohmann D. Decoupling the bridge helix of Cas12a results in a reduced trimming activity, increased mismatch sensitivity and impaired conformational transitions. *Nucleic Acids Res.* 2021;49(9):5278–5293. doi:10.1093/nar/gkab286.
35. Swarts DC, Jinek M. Cas9 versus Cas12a/Cpf1: Structure-function comparisons and implications for genome editing. *Wiley Interdiscip Rev RNA.* 2018:e1481. doi:10.1002/wrna.1481.
36. Bernabé-Orts JM, Casas-Rodrigo I, Minguet EG, Landolfi V, Garcia-Carpintero V, Gianoglio S, Vázquez-Vilar M, Granell A, Orzaez D. Assessment of Cas12a-mediated gene editing efficiency in plants. *Plant Biotechnol J.* 2019;17(10):1971–1984. doi:10.1111/pbi.13113.
37. Endo A, Masafumi M, Kaya H, Toki S. Efficient targeted mutagenesis of rice and tobacco genomes using Cpf1 from *Francisella novicida*. *Sci Rep.* 2016;6:38169. doi:10.1038/srep38169.
38. Begemann MB, Gray BN, January E, Gordon GC, He Y, Liu H, Wu X, Brutnell TP, Mockler TC, Oufattole M. Precise insertion and guided editing of higher plant genomes using Cpf1 CRISPR nucleases. *Sci Rep.* 2017;7(1):11606. doi:10.1038/s41598-017-11760-6.
39. Huang T-K, Armstrong B, Schindele P, Puchta H. Efficient gene targeting in *Nicotiana tabacum* using CRISPR/SaCas9 and temperature tolerant LbCas12a. *Plant Biotechnol J.* 2021;19(7):1314–1324. doi:10.1111/pbi.13546.
40. Zetsche B, Heidenreich M, Mohanraju P, Fedorova I, Kneppers J, DeGennaro EM, Winblad N, Choudhury SR, Abudayyeh OO, Gootenberg JS, et al. Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat Biotechnol.* 2017;35(1):31–34. doi:10.1038/nbt.3737.
41. Campa CC, Weisbach NR, Santinha AJ, Incarnato D, Platt RJ. Multiplexed genome engineering by Cas12a and CRISPR arrays encoded on single transcripts. *Nat Methods.* 2019;16(9):887–893. doi:10.1038/s41592-019-0508-6.
42. Tang X, Lowder LG, Zhang T, Malzahn AA, Zheng X, Voytas DF, Zhong Z, Chen Y, Ren Q, Li Q, et al. A CRISPR-Cpf1 system for efficient genome editing and transcriptional repression in plants. *Nature Plants.* 2017;3(3):17018. <https://www-nature-com.eaccess.tum.edu/articles/nplants201718>. doi:10.1038/nplants.2017.18.
43. Kanafi MM, Tavallaei M. Overview of advances in CRISPR/deadCas9 technology and its applications in human diseases. *Gene.* 2022;830:146518. doi:10.1016/j.gene.2022.146518.
44. Chavez A, Tuttle M, Pruitt BW, Ewen-Campen B, Chari R, Ter-Ovanesyan D, Haque SJ, Cecchi RJ, Kowal EJK, Buchthal J, et al. Comparison of Cas9 activators in multiple species. *Nat Methods.* 2016;13(7):563–567. doi:10.1038/nmeth.3871.
45. Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C, Panning B, Ploegh HL, Bassik MC, et al. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell.* 2014;159(3):647–661. doi:10.1016/j.cell.2014.09.029.

46. Chavez A, Scheiman J, Vora S, Pruitt BW, Tuttle M, P R Iyer E, Lin S, Kiani S, Guzman CD, Wiegand DJ, et al. Highly efficient Cas9-mediated transcriptional programming. *Nat Methods*. 2015;12(4):326–328. doi:10.1038/nmeth.3312.
47. Jiang J, Sun Y, Xiao R, Wai K, Ahmad MJ, Khan FA, Zhou H, Li Z, Zhang Y, Zhou A, et al. Porcine antiviral activity is increased by CRISPRa-SAM system. *Biosci Rep*. 2019;39(8). doi:10.1042/BSR20191496.
48. Sajwan S, Mannervik M. Gene activation by dCas9-CBP and the SAM system differ in target preference. *Sci Rep*. 2019;9(1):18104. doi:10.1038/s41598-019-54179-x.
49. Pickar-Oliver A, Gersbach CA. The next generation of CRISPR-Cas technologies and applications. *Nat Rev Mol Cell Biol*. 2019;20(8):490–507. doi:10.1038/s41580-019-0131-5.
50. Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*. 2013;154(2):442–451. doi:10.1016/j.cell.2013.06.044.
51. Yeo NC, Chavez A, Lance-Byrne A, Chan Y, Menn D, Milanova D, Kuo C-C, Guo X, Sharma S, Tung A, et al. An enhanced CRISPR repressor for targeted mammalian gene regulation. *Nat Methods*. 2018;15(8):611–616. doi:10.1038/s41592-018-0048-5.
52. Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*. 1998;393(6683):386–389. doi:10.1038/30764;https://doi.org/10.26508/lsa.202101321.
53. Fuks F, Hurd PJ, Wolf D, Nan X, Bird AP, Kouzarides T. The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *J Biol Chem*. 2003;278(6):4035–4040. doi:10.1074/jbc.M210256200.
54. Niemann H, Petersen B. The production of multi-transgenic pigs: update and perspectives for xenotransplantation. *Transgenic Res*. 2016;25(3):361–374. doi:10.1007/s11248-016-9934-8.
55. Hofmann A, Kessler B, Ewerling S, Weppert M, Vogg B, Ludwig H, Stojkovic M, Boelhauve M, Brem G, Wolf E, et al. Efficient transgenesis in farm animals by lentiviral vectors. *EMBO Rep*. 2003;4(11):1054–1058. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1326377/>. doi:10.1038/sj.embor.7400007.
56. Cabot RA, Kühholzer B, Chan AW, Lai L, Park KW, Chong KY, Schatten G, Murphy CN, Abeydeera LR, Day BN, et al. Transgenic pigs produced using in vitro matured oocytes infected with a retroviral vector. *Anim Biotechnol*. 2001;12(2):205–214. doi:10.1081/ABIO-100108347.
57. Hofmann A, Kessler B, Ewerling S, Kabermann A, Brem G, Wolf E, Pfeifer A. Epigenetic regulation of lentiviral transgene vectors in a large animal model. *Molecular Therapy*.

- 2006;13(1):59–66. <https://www.sciencedirect.com/science/article/pii/S1525001605014097>. doi:10.1016/j.ymthe.2005.07.685.
58. Kalidasan V, Ng WH, Ishola OA, Ravichantar N, Tan JJ, Das KT. A guide in lentiviral vector production for hard-to-transfect cells, using cardiac-derived c-kit expressing cells as a model system. *Sci Rep*. 2021;11(1):19265. <https://www.nature.com/articles/s41598-021-98657-7>. doi:10.1038/s41598-021-98657-7.
59. Amberger M, Ivics Z. Latest Advances for the Sleeping Beauty Transposon System: 23 Years of Insomnia but Prettier than Ever: Refinement and Recent Innovations of the Sleeping Beauty Transposon System Enabling Novel, Nonviral Genetic Engineering Applications. *BioEssays*. 2020;42(11):e2000136. <https://onlinelibrary.wiley.com/doi/full/10.1002/bies.202000136>. doi:10.1002/bies.202000136.
60. Carlson DF, Garbe JR, Tan W, Martin MJ, Dobrinsky JR, Hackett PB, Clark KJ, Fahrenkrug SC. Strategies for selection marker-free swine transgenesis using the Sleeping Beauty transposon system. *Transgenic Res*. 2011;20(5):1125–1137. <https://experts.umn.edu/en/publications/strategies-for-selection-marker-free-swine-transgenesis-using-the>. doi:10.1007/s11248-010-9481-7.
61. Garrels W, Mátés L, Holler S, Dalda A, Taylor U, Petersen B, Niemann H, Izsvák Z, Ivics Z, Kues WA. Germline transgenic pigs by Sleeping Beauty transposition in porcine zygotes and targeted integration in the pig genome. *PLOS ONE*. 2011;6(8):e23573. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0023573>. doi:10.1371/journal.pone.0023573.
62. Ivics Z, Garrels W, Mátés L, Yau TY, Bashir S, Zidek V, Landa V, Geurts A, Pravenec M, Rüllicke T, et al. Germline transgenesis in pigs by cytoplasmic microinjection of Sleeping Beauty transposons. *Nat Protoc*. 2014;9(4):810–827. <https://www-nature-com.eaccess.tum.edu/articles/nprot.2014.010>. doi:10.1038/nprot.2014.010.
63. van Chu T, Weber T, Graf R, Sommermann T, Petsch K, Sack U, Volchkov P, Rajewsky K, Kühn R. Efficient generation of Rosa26 knock-in mice using CRISPR/Cas9 in C57BL/6 zygotes. *BMC Biotechnol*. 2016;16:4. doi:10.1186/s12896-016-0234-4.
64. Li S, Flisikowska T, Kurome M, Zakhartchenko V, Kessler B, Saur D, Kind A, Wolf E, Flisikowski K, Schnieke A. Dual fluorescent reporter pig for Cre recombination: transgene placement at the ROSA26 locus. *PLOS ONE*. 2014;9(7):e102455. <https://journals.plos.org/plosone/article/figures?id=10.1371/journal.pone.0102455>. doi:10.1371/journal.pone.0102455.
65. Ruan J, Li H, Xu K, Wu T, Wei J, Zhou R, Liu Z, Mu Y, Yang S, Ouyang H, et al. Highly efficient CRISPR/Cas9-mediated transgene knockin at the H11 locus in pigs. *Sci Rep*. 2015;5:14253. doi:10.1038/srep14253.

66. Ma X, Zeng W, Wang L, Cheng R, Zhao Z, Huang C, Sun Z, Tao P, Wang T, Zhang J, et al. Validation of reliable safe harbor locus for efficient porcine transgenesis. *Funct Integr Genomics*. 2022;22(4):553–563. doi:10.1007/s10142-022-00859-3.
67. Richter A, Kurome M, Kessler B, Zakhartchenko V, Klymiuk N, Nagashima H, Wolf E, Wuensch A. Potential of primary kidney cells for somatic cell nuclear transfer mediated transgenesis in pig. *BMC Biotechnol*. 2012;12:84. doi:10.1186/1472-6750-12-84.
68. Li X, Heyer W-D. Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res*. 2008;18(1):99–113. doi:10.1038/cr.2008.1.
69. Denning C, Burl S, Ainslie A, Bracken J, Dinnyes A, Fletcher J, King T, Ritchie M, Ritchie WA, Rollo M, et al. Deletion of the alpha(1,3)galactosyl transferase (GGTA1) gene and the prion protein (PrP) gene in sheep. *Nat Biotechnol*. 2001;19(6):559–562. doi:10.1038/89313.
70. Reuven N, Shaul Y. Selecting for CRISPR-Edited Knock-In Cells. *Int J Mol Sci*. 2022;23(19). doi:10.3390/ijms231911919.
71. Park K-E, Powell A, Sandmaier SES, Kim C-M, Mileham A, Donovan DM, Telugu BP. Targeted gene knock-in by CRISPR/Cas ribonucleoproteins in porcine zygotes. *Sci Rep*. 2017;7:42458. doi:10.1038/srep42458.
72. Peng J, Wang Y, Jiang J, Zhou X, Song L, Wang L, Ding C, Qin J, Liu L, Wang W, et al. Production of Human Albumin in Pigs Through CRISPR/Cas9-Mediated Knockin of Human cDNA into Swine Albumin Locus in the Zygotes. *Sci Rep*. 2015;5:16705. doi:10.1038/srep16705.
73. Branzei D, Foiani M. Regulation of DNA repair throughout the cell cycle. *Nat Rev Mol Cell Biol*. 2008;9(4):297–308. <https://pubmed.ncbi.nlm.nih.gov/18285803/>. doi:10.1038/nrm2351.
74. Ishibashi A, Saga K, Hisatomi Y, Li Y, Kaneda Y, Nimura K. A simple method using CRISPR-Cas9 to knock-out genes in murine cancerous cell lines. *Sci Rep*. 2020;10(1):22345. doi:10.1038/s41598-020-79303-0.
75. He X, Tan C, Wang F, Wang Y, Zhou R, Cui D, You W, Zhao H, Ren J, Feng B. Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair. *Nucleic Acids Res*. 2016;44(9):e85. doi:10.1093/nar/gkw064.
76. Carroll D. Genome engineering with targetable nucleases. *Annu Rev Biochem*. 2014;83:409–439. doi:10.1146/annurev-biochem-060713-035418.
77. Jonathan L. Schmid-Burgk, Klara Höning, Thomas S. Ebert, Veit Hornung. CRISPaint allows modular base-specific gene tagging using a ligase-4-dependent mechanism. *Nat Commun*. 2016;7(1):1–12. En;en. <https://www.nature.com/articles/ncomms12338.pdf>. doi:10.1038/ncomms12338.
78. Jakobsen JE, Johansen MG, Schmidt M, Dagnæs-Hansen F, Dam K, Gunnarsson A, Liu Y, Kragh PM, Li R, Holm IE, et al. Generation of minipigs with targeted transgene insertion by recombinase-mediated cassette exchange (RMCE) and somatic cell nuclear transfer (SCNT).

- Transgenic Res. 2013;22(4):709–723. <https://link.springer.com/article/10.1007/s11248-012-9671-6>. doi:10.1007/s11248-012-9671-6.
79. Mulholland CB, Smets M, Schmidtman E, Leidescher S, Markaki Y, Hofweber M, Qin W, Manzo M, Kremmer E, Thanisch K, et al. A modular open platform for systematic functional studies under physiological conditions. *Nucleic Acids Res.* 2015;43(17):e112. doi:10.1093/nar/gkv550.
  80. Low BE, Hosur V, Lesbirel S, Wiles MV. Efficient targeted transgenesis of large donor DNA into multiple mouse genetic backgrounds using bacteriophage Bxb1 integrase. *Sci Rep.* 2022;12(1):5424. doi:10.1038/s41598-022-09445-w.
  81. Zhao Z, Anselmo AC, Mitragotri S. Viral vector-based gene therapies in the clinic. *Bioeng Transl Med.* 2022;7(1):e10258. doi:10.1002/btm2.10258.
  82. Mendell JR, Al-Zaidy SA, Rodino-Klapac LR, Goodspeed K, Gray SJ, Kay CN, Boye SL, Boye SE, George LA, Salabarria S, et al. Current Clinical Applications of In Vivo Gene Therapy with AAVs. *Mol Ther.* 2021;29(2):464–488. doi:10.1016/j.ymthe.2020.12.007.
  83. Zetsche B, Volz SE, Zhang F. A split-Cas9 architecture for inducible genome editing and transcription modulation. *Nat Biotechnol.* 2015;33(2):139–142. doi:10.1038/nbt.3149.
  84. Moretti A, Fonteyne L, Giesert F, Hoppmann P, Meier AB, Bozoglu T, Baehr A, Schneider CM, Sinnecker D, Klett K, et al. Somatic gene editing ameliorates skeletal and cardiac muscle failure in pig and human models of Duchenne muscular dystrophy. *Nat Med.* 2020;26(2):207–214. doi:10.1038/s41591-019-0738-2.
  85. Yan S, Zheng X, Lin Y, Li C, Liu Z, Li J, Tu Z, Zhao Y, Huang C, Chen Y, et al. Cas9-mediated replacement of expanded CAG repeats in a pig model of Huntington's disease. *Nat Biomed Eng.* 2023. doi:10.1038/s41551-023-01007-3.
  86. Watano R, Ohmori T, Hishikawa S, Sakata A, Mizukami H. Utility of microminipigs for evaluating liver-mediated gene expression in the presence of neutralizing antibody against vector capsid. *Gene Ther.* 2020;27(9):427–434. doi:10.1038/s41434-020-0125-0.
  87. Rieblinger B, Sid H, Duda D, Bozoglu T, Klinger R, Schlickerrieder A, Lengyel K, Flisikowski K, Flisikowska T, Simm N, et al. Cas9-expressing chickens and pigs as resources for genome editing in livestock. *PNAS.* 2021;118(10):e2022562118. <https://www.pnas.org/doi/10.1073/pnas.2022562118>. doi:10.1073/pnas.2022562118.
  88. Wang K, Jin Q, Ruan D, Yang Y, Liu Q, Wu H, Zhou Z, Ouyang Z, Liu Z, Zhao Y, et al. Cre-dependent Cas9-expressing pigs enable efficient in vivo genome editing. *Genome Res.* 2017;27(12):2061–2071. doi:10.1101/gr.222521.117.
  89. Li C, Samulski RJ. Engineering adeno-associated virus vectors for gene therapy. *Nat Rev Genet.* 2020;21(4):255–272. <https://www.nature.com/articles/s41576-019-0205-4>. doi:10.1038/s41576-019-0205-4.

90. Gao Y, Zhao Y. Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. *J. Integr. Plant Biol.* 2014;56(4):343–349. <https://onlinelibrary.wiley.com/doi/10.1111/jipb.12152>. doi:10.1111/jipb.12152.
91. Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y, Barrette TR, Prensner JR, Evans JR, Zhao S, et al. The landscape of long noncoding RNAs in the human transcriptome. *Nat Genet.* 2015;47(3):199–208. doi:10.1038/ng.3192.
92. Fang Y, Fullwood MJ. Roles, Functions, and Mechanisms of Long Non-coding RNAs in Cancer. *Genomics Proteomics Bioinformatics.* 2016;14(1):42–54. doi:10.1016/j.gpb.2015.09.006.
93. Li Z, Gao J, Di Sun, Jiao Q, Ma J, Cui W, Lou Y, Xu F, Li S, Li H. LncRNA MEG3: Potential stock for precision treatment of cardiovascular diseases. *Front Pharmacol.* 2022;13:1045501. doi:10.3389/fphar.2022.1045501.
94. MacDonald WA, Mann MRW. Long noncoding RNA functionality in imprinted domain regulation. *PLoS Genet.* 2020;16(8):e1008930. doi:10.1371/journal.pgen.1008930.
95. Monk D, Mackay DJG, Eggermann T, Maher ER, Riccio A. Genomic imprinting disorders: lessons on how genome, epigenome and environment interact. *Nat Rev Genet.* 2019;20(4):235–248. doi:10.1038/s41576-018-0092-0.
96. Sanli I, Lalevée S, Cammisa M, Perrin A, Rage F, Llères D, Riccio A, Bertrand E, Feil R. Meg3 Non-coding RNA Expression Controls Imprinting by Preventing Transcriptional Upregulation in cis. *Cell Rep.* 2018;23(2):337–348. doi:10.1016/j.celrep.2018.03.044.
97. Honda S, Chatterjee A, Leichter AL, Miyagi H, Minato M, Fujiyoshi S, Ara M, Kitagawa N, Tanaka M, Tanaka Y, et al. A MicroRNA Cluster in the DLK1-DIO3 Imprinted Region on Chromosome 14q32.2 Is Dysregulated in Metastatic Hepatoblastomas. *Front. Oncol.* 2020;10:513601. <https://www.frontiersin.org/articles/10.3389/fonc.2020.513601/full>. doi:10.3389/fonc.2020.513601.
98. Budkova Z, Sigurdardottir AK, Briem E, Bergthorsson JT, Sigurdsson S, Magnusson MK, Traustadottir GA, Gudjonsson T, Hilmarsdottir B. Expression of ncRNAs on the DLK1-DIO3 Locus Is Associated With Basal and Mesenchymal Phenotype in Breast Epithelial Progenitor Cells. *Front Cell Dev Biol.* 2020;8:461. doi:10.3389/fcell.2020.00461.
99. Magnuson B, Veloso A, Kirkconnell KS, Andrade Lima LC de, Paulsen MT, Ljungman EA, Bedi K, Prasad J, Wilson TE, Ljungman M. Identifying transcription start sites and active enhancer elements using BruUV-seq. *Sci Rep.* 2015;5(1):17978. <https://www.nature.com/articles/srep17978>. doi:10.1038/srep17978.
100. Llères D, Moindrot B, Pathak R, Piras V, Matelot M, Pignard B, Marchand A, Poncelet M, Perrin A, Tellier V, et al. CTCF modulates allele-specific sub-TAD organization and imprinted gene activity at the mouse Dlk1-Dio3 and Igf2-H19 domains. *Genome Biol.* 2019;20(1):272. doi:10.1186/s13059-019-1896-8.

101. Kagami M, O'Sullivan MJ, Green AJ, Watabe Y, Arisaka O, Masawa N, Matsuoka K, Fukami M, Matsubara K, Kato F, et al. The IG-DMR and the MEG3-DMR at human chromosome 14q32.2: hierarchical interaction and distinct functional properties as imprinting control centers. *PLoS Genet.* 2010;6(6):e1000992. doi:10.1371/journal.pgen.1000992.
102. Jingchang Zhang, Yi Liang, Xuecheng Huang, Xiaoyan Guo, Yang Liu, Jiming Zhong, Jieli Yuan. STAT3-induced upregulation of lncRNA MEG3 regulates the growth of cardiac hypertrophy through miR-361-5p/HDAC9 axis. *Sci Rep.* 2019;9(1):1–11. En;en. <https://www.nature.com/articles/s41598-018-36369-1.pdf>. doi:10.1038/s41598-018-36369-1.
103. He C, Yang W, Yang J, Ding J, Li S, Wu H, Zhou F, Jiang Y, Teng L, Yang J. Long Noncoding RNA MEG3 Negatively Regulates Proliferation and Angiogenesis in Vascular Endothelial Cells. *DNA Cell Biol.* 2017;36(6):475–481. doi:10.1089/dna.2017.3682.
104. Zhao J, Zhang X, Zhou Y, Ansell PJ, Klibanski A. Cyclic AMP stimulates MEG3 gene expression in cells through a cAMP-response element (CRE) in the MEG3 proximal promoter region. *Int J Biochem Cell Biol.* 2006;38(10):1808–1820. doi:10.1016/j.biocel.2006.05.004.
105. Zhou X, Ji G, Ke X, Gu H, Jin W, Zhang G. MiR-141 Inhibits Gastric Cancer Proliferation by Interacting with Long Noncoding RNA MEG3 and Down-Regulating E2F3 Expression. *Dig Dis Sci.* 2015;60(11):3271–3282. doi:10.1007/s10620-015-3782-x.
106. Zhao J, Dahle D, Zhou Y, Zhang X, Klibanski A. Hypermethylation of the promoter region is associated with the loss of MEG3 gene expression in human pituitary tumors. *J Clin Endocrinol Metab.* 2005;90(4):2179–2186. doi:10.1210/jc.2004-1848.
107. Moradi M-T, Fallahi H, Rahimi Z. Interaction of long noncoding RNA MEG3 with miRNAs: A reciprocal regulation. *J Cell Biochem.* 2019;120(3):3339–3352. doi:10.1002/jcb.27604#jcb27604-bib-0008.
108. Guo W, Dong Z, Liu S, Qiao Y, Kuang G, Guo Y, Shen S, Liang J. Promoter hypermethylation-mediated downregulation of miR-770 and its host gene MEG3, a long non-coding RNA, in the development of gastric cardia adenocarcinoma. *Mol Carcinog.* 2017;56(8):1924–1934. doi:10.1002/mc.22650.
109. Kruer TL, Dougherty SM, Reynolds L, Long E, Silva T de, Lockwood WW, Clem BF. Expression of the lncRNA Maternally Expressed Gene 3 (MEG3) Contributes to the Control of Lung Cancer Cell Proliferation by the Rb Pathway. *PLOS ONE.* 2016;11(11):e0166363. doi:10.1371/journal.pone.0166363.
110. Zhuo H, Tang J, Lin Z, Jiang R, Zhang X, Ji J, Wang P, Sun B. The aberrant expression of MEG3 regulated by UHRF1 predicts the prognosis of hepatocellular carcinoma. *Mol Carcinog.* 2016;55(2):209–219. doi:10.1002/mc.22270.
111. Yvette Lahbib-Mansais, Harmonie Barasc, Maria Marti-Marimon, Florence Mompert, Eddie Iannuccelli, David Robelin, Juliette Riquet, Martine Yerle-Bouissou. Expressed alleles of imprinted IGF2, DLK1 and MEG3 colocalize in 3D-preserved nuclei of porcine fetal cells. *BMC*

- Cell Biol. 2016;17(1):1–15. En;en. <https://bmcmolcellbiol.biomedcentral.com/track/pdf/10.1186/s12860-016-0113-9>. doi:10.1186/s12860-016-0113-9.
112. Tanmoy Mondal, Santhilal Subhash, Roshan Vaid, Stefan Enroth, Sireesha Uday, Björn Reinius, Sanhita Mitra, Arif Mohammed, Alva Rani James, Emily Hoberg, et al. MEG3 long noncoding RNA regulates the TGF- $\beta$  pathway genes through formation of RNA–DNA triplex structures. *Nat Commun.* 2015;6(1):1–17. En;en. <https://www.nature.com/articles/ncomms8743.pdf>. doi:10.1038/ncomms8743.
113. Gao Y, Lu X. Decreased expression of MEG3 contributes to retinoblastoma progression and affects retinoblastoma cell growth by regulating the activity of Wnt/ $\beta$ -catenin pathway. *Tumour Biol.* 2016;37(2):1461–1469. doi:10.1007/s13277-015-4564-y.
114. Zhou Y, Zhong Y, Wang Y, Zhang X, Batista DL, Gejman R, Ansell PJ, Zhao J, Weng C, Klibanski A. Activation of p53 by MEG3 non-coding RNA. *J Biol Chem.* 2007;282(34):24731–24742. doi:10.1074/jbc.M702029200.
115. Margueron R, Reinberg D. The Polycomb complex PRC2 and its mark in life. *Nature.* 2011;469(7330):343–349. doi:10.1038/nature09784.
116. Huang C, Liao X, Jin H, Xie F, Zheng F, Li J, Zhou C, Jiang G, Wu X-R, Huang C. MEG3, as a Competing Endogenous RNA, Binds with miR-27a to Promote PHLPP2 Protein Translation and Impairs Bladder Cancer Invasion. *Mol Ther Nucleic Acids.* 2019;16:51–62. doi:10.1016/j.omtn.2019.01.014.
117. Wang Q-C, Wang Z-Y, Xu Q, Chen X-L, Shi R-Z. lncRNA expression profiles and associated ceRNA network analyses in epicardial adipose tissue of patients with coronary artery disease. *Sci Rep.* 2021;11(1):1567. doi:10.1038/s41598-021-81038-5.
118. Li H, Li B, Zhu D, Xie H, Du C, Xia Y, Tang W. Downregulation of lncRNA MEG3 and miR-770-5p inhibit cell migration and proliferation in Hirschsprung's disease. *Oncotarget.* 2017;8(41):69722–69730. doi:10.18632/oncotarget.19207.
119. Xu J, Wang X, Zhu C, Wang K. A review of current evidence about lncRNA MEG3: A tumor suppressor in multiple cancers. *Front Cell Dev Biol.* 2022;10:997633. doi:10.3389/fcell.2022.997633.
120. Uroda T, Anastasakou E, Rossi A, Teulon J-M, Pellequer J-L, Annibale P, Pessey O, Inga A, Chillón I, Marcia M. Conserved Pseudoknots in lncRNA MEG3 Are Essential for Stimulation of the p53 Pathway. *Molecular Cell.* 2019;75(5):982–995.e9. <https://www.sciencedirect.com/science/article/pii/S1097276519305635>. doi:10.1016/j.molcel.2019.07.025.
121. Sherpa C, Rausch JW, Le Grice SFJ. Structural characterization of maternally expressed gene 3 RNA reveals conserved motifs and potential sites of interaction with polycomb repressive complex 2. *Nucleic Acids Res.* 2018;46(19):10432–10447. doi:10.1093/nar/gky722.

122. Raziyeva K, Kim Y, Zharkinbekov Z, Temirkhanova K, Saparov A. Novel Therapies for the Treatment of Cardiac Fibrosis Following Myocardial Infarction. *Biomedicines*. 2022;10(9). doi:10.3390/biomedicines10092178.
123. Dobaczewski M, Gonzalez-Quesada C, Frangogiannis NG. The extracellular matrix as a modulator of the inflammatory and reparative response following myocardial infarction. *J Mol Cell Cardiol*. 2010;48(3):504–511. doi:10.1016/j.yjmcc.2009.07.015.
124. Li J, Liu W, Peng F, Cao X, Xie X, Peng C. The multifaceted biology of lncR-Meg3 in cardio-cerebrovascular diseases. *Front. Genet*. 2023;14:400. <https://www.frontiersin.org/articles/10.3389/fgene.2023.1132884/full>. doi:10.3389/fgene.2023.1132884.
125. Zou L, Ma X, Lin S, Wu B, Chen Y, Peng C. Long noncoding RNA-MEG3 contributes to myocardial ischemia-reperfusion injury through suppression of miR-7-5p expression. *Biosci Rep*. 2019;39(8). doi:10.1042/BSR20190210.
126. Wu H, Zhao Z-A, Liu J, Hao K, Yu Y, Han X, Li J, Wang Y, Lei W, Dong N, et al. Long noncoding RNA Meg3 regulates cardiomyocyte apoptosis in myocardial infarction. *Gene Ther*. 2018;25(8):511–523. doi:10.1038/s41434-018-0045-4.
127. Chen Y, Zhang Z, Zhu D, Zhao W, Li F. Long non-coding RNA MEG3 serves as a ceRNA for microRNA-145 to induce apoptosis of AC16 cardiomyocytes under high glucose condition. *Biosci Rep*. 2019;39(6). doi:10.1042/BSR20190444.
128. Yousefi F, Shabaninejad Z, Vakili S, Derakhshan M, Movahedpour A, Dabiri H, Ghasemi Y, Mahjoubin-Tehran M, Nikoozadeh A, Savardashtaki A, et al. TGF- $\beta$  and WNT signaling pathways in cardiac fibrosis: non-coding RNAs come into focus. *Cell Commun Signal*. 2020;18(1):87. <https://biosignaling.biomedcentral.com/articles/10.1186/s12964-020-00555-4>. doi:10.1186/s12964-020-00555-4.
129. Travers JG, Kamal FA, Robbins J, Yutzey KE, Blaxall BC. Cardiac Fibrosis: The Fibroblast Awakens. *Circ Res*. 2016;118(6):1021–1040. doi:10.1161/CIRCRESAHA.115.306565.
130. Piccoli M-T, Gupta SK, Viereck J, Foinquinos A, Samolovac S, Kramer FL, Garg A, Remke J, Zimmer K, Batkai S, et al. Inhibition of the Cardiac Fibroblast-Enriched lncRNA Meg3 Prevents Cardiac Fibrosis and Diastolic Dysfunction. *Circ Res*. 2017;121(5):575–583. doi:10.1161/CIRCRESAHA.117.310624.
131. Li W, Li Y, Cui S, Liu J, Tan L, Xia H, Zhang C. Se alleviates homocysteine-induced fibrosis in cardiac fibroblasts via downregulation of lncRNA MEG3. *Experimental and Therapeutic Medicine*. 2021;22(5):1269. <https://www.spandidos-publications.com/10.3892/etm.2021.10704>. doi:10.3892/etm.2021.10704.
132. Marian AJ, Braunwald E. Hypertrophic Cardiomyopathy: Genetics, Pathogenesis, Clinical Manifestations, Diagnosis, and Therapy. *Circ Res*. 2017;121(7):749–770. doi:10.1161/CIRCRESAHA.117.311059.

133. Johnsson P, Lipovich L, Grandér D, Morris KV. Evolutionary conservation of long non-coding RNAs; sequence, structure, function. *Biochim Biophys Acta*. 2014;1840(3):1063–1071. doi:10.1016/j.bbagen.2013.10.035.
134. Zhang J, Liang Y, Huang X, Guo X, Liu Y, Zhong J, Yuan J. STAT3-induced upregulation of lncRNA MEG3 regulates the growth of cardiac hypertrophy through miR-361-5p/HDAC9 axis. *Sci Rep*. 2019;9(1):460. doi:10.1038/s41598-018-36369-1.
135. Parekh S, Ziegenhain C, Vieth B, Enard W, Hellmann I. The impact of amplification on differential expression analyses by RNA-seq. *Sci Rep*. 2016;6:25533. doi:10.1038/srep25533.
136. Xu J, Yu L, Guo J, Xiang J, Zheng Z, Gao D, Shi B, Hao H, Jiao D, Zhong L, et al. Generation of pig induced pluripotent stem cells using an extended pluripotent stem cell culture system. *Stem Cell Res Ther*. 2019;10(1):193. doi:10.1186/s13287-019-1303-0.
137. Gyöngyösi M, Pavo N, Lukovic D, Zlabinger K, Spannbaauer A, Traxler D, Goliash G, Mandic L, Bergler-Klein J, Gugerell A, et al. Porcine model of progressive cardiac hypertrophy and fibrosis with secondary postcapillary pulmonary hypertension. *J Transl Med*. 2017;15(1):202. doi:10.1186/s12967-017-1299-0.
138. Statello L, Guo C-J, Chen L-L, Huarte M. Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol*. 2021;22(2):96–118. doi:10.1038/s41580-020-00315-9.
139. Fatemeh Safari, Khadijeh Zare, Manica Negahdaripour, Mazyar Barekati-Mowahed, Younes Ghasemi. CRISPR Cpf1 proteins: structure, function and implications for genome editing. *Cell Biosci*. 2019;9(1):1–21. En;en. <https://cellandbioscience.biomedcentral.com/track/pdf/10.1186/s13578-019-0298-7>. doi:10.1186/s13578-019-0298-7.
140. Li X, Yang Y, Bu L, Guo X, Tang C, Song J, Fan N, Zhao B, Ouyang Z, Liu Z, et al. Rosa26-targeted swine models for stable gene over-expression and Cre-mediated lineage tracing. *Cell Res*. 2014;24(4):501–504. <https://www.nature.com/articles/cr201415>. doi:10.1038/cr.2014.15.
141. Medert R, Thumberger T, Tavhelidse-Suck T, Hub T, Kellner T, Oguchi Y, Dlugosz S, Zimmermann F, Wittbrodt J, Freichel M. Efficient single copy integration via homology-directed repair (scHDR) by 5' modification of large DNA donor fragments in mice. *Nucleic Acids Res*. 2022. <https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkac1150/6931855>. doi:10.1093/nar/gkac1150.
142. Zhao Y, Boeke JD. CRISPR-Cas12a system in fission yeast for multiplex genomic editing and CRISPR interference. *Nucleic Acids Res*. 2020;48(10):5788–5798. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7261154/>. doi:10.1093/nar/gkaa329.
143. Klymiuk N, Böcker W, Schönitzer V, Bähr A, Radic T, Fröhlich T, Wünsch A, Keßler B, Kurome M, Schilling E, et al. First inducible transgene expression in porcine large animal models. *FASEB journal: official publication of the Federation of American Societies for*

- Experimental Biology. 2012;26(3):1086–1099. <https://pubmed.ncbi.nlm.nih.gov/22138035/>. doi:10.1096/fj.11-185041.
144. Fischer K, Rieblinger B, Hein R, Sfriso R, Zuber J, Fischer A, Klinger B, Liang W, Flisikowski K, Kurome M, et al. Viable pigs after simultaneous inactivation of porcine MHC class I and three xenoreactive antigen genes GGTA1, CMAH and B4GALNT2. *Xenotransplantation*. 2020;27(1):e12560. <https://pubmed.ncbi.nlm.nih.gov/31591751/>. doi:10.1111/xen.12560.
145. Su Bin Moon, Jeong Mi Lee, Jeong Gu Kang, Nan-Ee Lee, Dae-In Ha, Do Yon Kim, Sun Hee Kim, Kwangsun Yoo, Daesik Kim, Jeong-Heon Ko, et al. Highly efficient genome editing by CRISPR-Cpf1 using CRISPR RNA with a uridylate-rich 3'-overhang. *Nat Commun*. 2018;9(1):1–11. En;en. <https://www.nature.com/articles/s41467-018-06129-w.pdf>. doi:10.1038/s41467-018-06129-w.
146. Gier RA, Budinich KA, Evitt NH, Cao Z, Freilich ES, Chen Q, Qi J, Lan Y, Kohli RM, Shi J. High-performance CRISPR-Cas12a genome editing for combinatorial genetic screening. *Nat Commun*. 2020;11(1):3455. doi:10.1038/s41467-020-17209-1.
147. Li J, Liang Q, Zhou H, Zhou M, Huang H. Profiling the impact of the promoters on CRISPR-Cas12a system in human cells. *Cell Mol Biol Lett*. 2023;28(1):41. doi:10.1186/s11658-023-00454-9.
148. Hartford CCR, Lal A. When Long Noncoding Becomes Protein Coding. *Mol Cell Biol*. 2020;40(6). doi:10.1128/MCB.00528-19.
149. Cabili MN, Dunagin MC, McClanahan PD, Biaisch A, Padovan-Merhar O, Regev A, Rinn JL, Raj A. Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. *Genome Biol*. 2015;16(1):20. <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-015-0586-4>. doi:10.1186/s13059-015-0586-4.
150. Clemson CM, Hutchinson JN, Sara SA, Ensminger AW, Fox AH, Chess A, Lawrence JB. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Molecular Cell*. 2009;33(6):717–726. doi:10.1016/j.molcel.2009.01.026.
151. Rosenlund IA, Calin GA, Dragomir MP, Knutsen E. CRISPR/Cas9 to Silence Long Non-Coding RNAs. *Methods Mol Biol*. 2021;2348:175–187. doi:10.1007/978-1-0716-1581-2\_12.
152. Rakyán VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. *Nat Rev Genet*. 2011;12(8):529–541. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3508712/>. doi:10.1038/nrg3000.
153. Li J, Yu D, Wang J, Li C, Wang Q, Du W, Zhao S, Pang Y, Hao H, Zhao X, et al. Identification of the porcine IG-DMR and abnormal imprinting of DLK1-DIO3 in cloned pigs. *Front Cell Dev Biol*. 2022;10:964045. doi:10.3389/fcell.2022.964045.

154. Oshima G, Poli EC, Bolt MJ, Chlenski A, Forde M, Jutzy JMS, Biyani N, Posner MC, Pitroda SP, Weichselbaum RR, et al. DNA Methylation Controls Metastasis-Suppressive 14q32-Encoded miRNAs. *Cancer Res.* 2019;79(3):650–662. doi:10.1158/0008-5472.CAN-18-0692.
155. Holwerda SJB, Laat W de. CTCF: the protein, the binding partners, the binding sites and their chromatin loops. *Philos Trans R Soc Lond B Biol Sci.* 2013;368(1620):20120369. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3682731/>. doi:10.1098/rstb.2012.0369.
156. Rehmsmeier M, Steffen P, Hochsmann M, Giegerich R. Fast and effective prediction of microRNA/target duplexes. *RNA.* 2004;10(10):1507–1517. doi:10.1261/rna.5248604.
157. Awwad DA. Beyond classic editing: innovative CRISPR approaches for functional studies of long non-coding RNA. *Biol Methods Protoc.* 2019;4(1):bpz017. <https://academic.oup.com/biomethods/article/4/1/bpz017/5651012>. doi:10.1093/biomethods/bpz017.
158. Zhu W, Botticelli EM, Kery RE, Mao Y, Wang X, Yang A, Wang X, Zhou J, Zhang X, Soberman RJ, et al. Meg3-DMR, not the Meg3 gene, regulates imprinting of the Dlk1-Dio3 locus. *Developmental biology.* 2019;455(1):10–18. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6754764/>. doi:10.1016/j.ydbio.2019.07.005.
159. Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, Hsu PD, Habib N, Gootenberg JS, Nishimasu H, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature.* 2015;517(7536):583–588. <https://www-nature-com.eaccess.tum.edu/articles/nature14136>. doi:10.1038/nature14136.
160. Xie K, Minkenberg B, Yang Y. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *PNAS.* 2015;112(11):3570–3575. <https://www.pnas.org/doi/10.1073/pnas.1420294112>. doi:10.1073/pnas.1420294112.
161. Bachmann JC, Baumgart SJ, Uryga AK, Bosteen MH, Borghetti G, Nyberg M, Herum KM. Fibrotic Signaling in Cardiac Fibroblasts and Vascular Smooth Muscle Cells: The Dual Roles of Fibrosis in HFpEF and CAD. *Cells.* 2022;11(10). doi:10.3390/cells11101657.
162. Li L, Gan H. Intact Fibroblast Growth Factor 23 Regulates Chronic Kidney Disease-Induced Myocardial Fibrosis by Activating the Sonic Hedgehog Signaling Pathway. *J Am Heart Assoc.* 2022;11(18):e026365. doi:10.1161/JAHA.122.026365.
163. Liu M, Rehman S, Tang X, Gu K, Fan Q, Chen D, Ma W. Methodologies for Improving HDR Efficiency. *Front. Genet.* 2018;9:691. doi:10.3389/fgene.2018.00691.
164. Schiermeyer A, Schneider K, Kirchhoff J, Schmelter T, Koch N, Jiang K, Herwartz D, Blue R, Marri P, Samuel P, et al. Targeted insertion of large DNA sequences by homology-directed repair or non-homologous end joining in engineered tobacco BY-2 cells using designed zinc finger nucleases. *Plant Direct.* 2019;3(7):e00153. doi:10.1002/pld3.153.
165. Dai J, Cui X, Zhu Z, Hu W. Non-homologous end joining plays a key role in transgene concatemer formation in transgenic zebrafish embryos. *Int J Biol Sci.* 2010;6(7):756–768. doi:10.7150/ijbs.6.756.

166. Heyer W-D, Ehmsen KT, Liu J. Regulation of homologous recombination in eukaryotes. *Annu Rev Genet.* 2010;44:113–139. doi:10.1146/annurev-genet-051710-150955.
167. Zhang J-P, Li X-L, Li G-H, Chen W, Arakaki C, Botimer GD, Baylink D, Zhang L, Wen W, Fu Y-W, et al. Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol.* 2017;18(1):35. doi:10.1186/s13059-017-1164-8.
168. Cho YS, Kim DS, Nam YK. Concatemer-Associated Transgene Expression Patterns in Transgenic Marine Medaka *Oryzias dancena* Strains. *Fisheries and aquatic sciences.* 2015;18(1):73–80. doi:10.5657/FAS.2015.0073.
169. Fischer K, Kraner-Scheiber S, Petersen B, Rieblinger B, Buermann A, Flisikowska T, Flisikowski K, Christan S, Edlinger M, Baars W, et al. Efficient production of multi-modified pigs for xenotransplantation by ‘combineering’, gene stacking and gene editing. *Sci Rep.* 2016;6. doi:10.1038/srep29081.
170. Jakobsen J, Mikkelsen J, Nielsen A. Elimination of the plasmid bacterial backbone in site-directed transgenesis. *Biotechniques.* 2010;48(4):313–316. doi:10.2144/000113386.
171. Tasic B, Hippenmeyer S, Wang C, Gamboa M, Zong H, Chen-Tsai Y, Luo L. Site-specific integrase-mediated transgenesis in mice via pronuclear injection. *PNAS.* 2011;108(19):7902–7907. doi:10.1073/pnas.1019507108.
172. Li W, Shi L, Zhuang Z, Wu H, Lian M, Chen Y, Li L, Ge W, Jin Q, Zhang Q, et al. Engineered Pigs Carrying a Gain-of-Function NLRP3 Homozygous Mutation Can Survive to Adulthood and Accurately Recapitulate Human Systemic Spontaneous Inflammatory Responses. *J Immunol.* 2020;205(9):2532–2544. doi:10.4049/jimmunol.1901468.
173. Platt RJ, Chen S, Zhou Y, Yim MJ, Swiech L, Kempton HR, Dahlman JE, Parnas O, Eisenhaure TM, Jovanovic M, et al. CRISPR-Cas9 Knockin Mice for Genome Editing and Cancer Modeling. *Cell.* 2014;159(2):440–455. doi:10.1016/j.cell.2014.09.014.
174. Kondo S, Ueda R. Highly improved gene targeting by germline-specific Cas9 expression in *Drosophila*. *Genetics.* 2013;195(3):715–721. doi:10.1534/genetics.113.156737.
175. Luk K, Liu P, Zeng J, Wang Y, Maitland SA, Idrizi F, Ponninselvan K, Zhu LJ, Luban J, Bauer DE, et al. Optimization of Nuclear Localization Signal Composition Improves CRISPR-Cas12a Editing Rates in Human Primary Cells. *GEN Biotechnology.* 2022;1(3):271–284. doi:10.1089/genbio.2022.0003.
176. DeWeirdt PC, Sanson KR, Sangree AK, Hegde M, Hanna RE, Feeley MN, Griffith AL, Teng T, Borys SM, Strand C, et al. Optimization of AsCas12a for combinatorial genetic screens in human cells. *Nat Biotechnol.* 2021;39(1):94–104. doi:10.1038/s41587-020-0600-6.
177. Ma E, Chen K, Shi H, Stahl EC, Adler B, Trinidad M, Liu J, Zhou K, Ye J, Doudna JA. Improved genome editing by an engineered CRISPR-Cas12a. *Nucleic Acids Res.* 2022;50(22):12689–12701. doi:10.1093/nar/gkac1192.

178. Zhang L, Zuris JA, Viswanathan R, Edelstein JN, Turk R, Thommandru B, Rube HT, Glenn SE, Collingwood MA, Bode NM, et al. AsCas12a ultra nuclease facilitates the rapid generation of therapeutic cell medicines. *Nat Commun.* 2021;12(1):3908. doi:10.1038/s41467-021-24017-8.
179. Huang H, Huang G, Tan Z, Hu Y, Shan L, Zhou J, Zhang X, Ma S, Lv W, Huang T, et al. Engineered Cas12a-Plus nuclease enables gene editing with enhanced activity and specificity. *BMC Biol.* 2022;20(1):91. doi:10.1186/s12915-022-01296-1.
180. Kleinstiver BP, Sousa AA, Walton RT, Tak YE, Hsu JY, Clement K, Welch MM, Horng JE, Malagon-Lopez J, Scarfò I, et al. Engineered CRISPR-Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. *Nat Biotechnol.* 2019;37(3):276–282. doi:10.1038/s41587-018-0011-0.
181. Zhou Y, Cheunsuchon P, Nakayama Y, Lawlor MW, Zhong Y, Rice KA, Zhang L, Zhang X, Gordon FE, Lidov HGW, et al. Activation of paternally expressed genes and perinatal death caused by deletion of the Gtl2 gene. *Development.* 2010;137(16):2643–2652. <https://journals.biologists.com/dev/article/137/16/2643/43908/Activation-of-paternally-expressed-genes-and>. doi:10.1242/dev.045724.
182. Takahashi N, Okamoto A, Kobayashi R, Shirai M, Obata Y, Ogawa H, Sotomaru Y, Kono T. Deletion of Gtl2, imprinted non-coding RNA, with its differentially methylated region induces lethal parent-origin-dependent defects in mice. *Hum Mol Genet.* 2009;18(10):1879–1888. doi:10.1093/hmg/ddp108.
183. Noordermeer D, Feil R. Differential 3D chromatin organization and gene activity in genomic imprinting. *Curr Opin Genet Dev.* 2020;61:17–24. doi:10.1016/j.gde.2020.03.004.
184. Jacob AG, Smith CWJ. Intron retention as a component of regulated gene expression programs. *Hum Genet.* 2017;136(9):1043–1057. doi:10.1007/s00439-017-1791-x.
185. Karlsson M, Sjöstedt E, Oksvold P, Sivertsson Å, Huang J, Álvarez MB, Arif M, Li X, Lin L, Yu J, et al. Genome-wide annotation of protein-coding genes in pig. *BMC Biol.* 2022;20. doi:10.1186/s12915-022-01229-y.
186. Li Y, Li C, Liu Q, Wang L, Bao AX, Jung JP, Dodlapati S, Sun J, Gao P, Zhang X, et al. Loss of Acta2 in cardiac fibroblasts does not prevent the myofibroblast differentiation or affect the cardiac repair after myocardial infarction. *J Mol Cell Cardiol.* 2022;171:117–132. doi:10.1016/j.yjmcc.2022.08.003.
187. Liu M, López de Juan Abad B, Cheng K. Cardiac fibrosis: Myofibroblast-mediated pathological regulation and drug delivery strategies. *Adv Drug Deliv Rev.* 2021;173:504–519. doi:10.1016/j.addr.2021.03.021.
188. Skrbic B, Engebretsen KVT, Strand ME, Lunde IG, Herum KM, Marstein HS, Sjaastad I, Lunde PK, Carlson CR, Christensen G, et al. Lack of collagen VIII reduces fibrosis and promotes

- early mortality and cardiac dilatation in pressure overload in mice. *Cardiovasc Res.* 2015;106(1):32–42. doi:10.1093/cvr/cvv041.
189. Sheppard HM, Corneillie SI, Espiritu C, Gatti A, Liu X. New insights into the mechanism of inhibition of p53 by simian virus 40 large T antigen. *Mol Cell Biol.* 1999;19(4):2746–2753. doi:10.1128/mcb.19.4.2746.
190. Hu D, Su C, Jiang M, Shen Y, Shi A, Zhao F, Chen R, Shen Z, Bao J, Tang W. Fenofibrate inhibited pancreatic cancer cells proliferation via activation of p53 mediated by upregulation of LncRNA MEG3. *Biochem Biophys Res Commun.* 2016;471(2):290–295. doi:10.1016/j.bbrc.2016.01.169.
191. He S, Li Y, Chen Y, Zhu Y, Zhang X, Xia X, Sun H. immortalization of pig fibroblast cells by transposon-mediated ectopic expression of porcine telomerase reverse transcriptase. *Cytotechnology.* 2016;68(4):1435–1445. doi:10.1007/s10616-015-9903-8.
192. Yao X, Wang X, Hu X, Liu Z, Liu J, Zhou H, Shen X, Wei Y, Huang Z, Ying W, et al. Homology-mediated end joining-based targeted integration using CRISPR/Cas9. *Cell Res.* 2017;27(6):801–814. doi:10.1038/cr.2017.76.
193. Hsieh-Feng V, Yang Y. Efficient expression of multiple guide RNAs for CRISPR/Cas genome editing. *aBIOTECH.* 2020;1(2):123–134. doi:10.1007/s42994-019-00014-w.
194. Wei T, Cheng Q, Min Y-L, Olson EN, Siegwart DJ. Systemic nanoparticle delivery of CRISPR-Cas9 ribonucleoproteins for effective tissue specific genome editing. *Nat Commun.* 2020;11(1):3232. doi:10.1038/s41467-020-17029-3.
195. Kim DY, Lee JM, Moon SB, Chin HJ, Park S, Lim Y, Kim D, Koo T, Ko J-H, Kim Y-S. Efficient CRISPR editing with a hypercompact Cas12f1 and engineered guide RNAs delivered by adeno-associated virus. *Nat Biotechnol.* 2022;40(1):94–102. doi:10.1038/s41587-021-01009-z.
196. Thomas Thum. Targeting cardiac fibrosis with next generation RNA therapeutics. [place unknown]: [publisher unknown]; [date unknown]. <https://cordis.europa.eu/project/id/101058103>.
197. Pan H, Diao H, Zhong W, Wang T, Wen P, Wu C. A Cancer Cell Cluster Marked by LincRNA MEG3 Leads Pancreatic Ductal Adenocarcinoma Metastasis. *Front Oncol.* 2021;11:656564. doi:10.3389/fonc.2021.656564.