

Gene Editing in pig models of inherited retinal diseases

von Hannah Franziska Auch

Inaugural-Dissertation zur Erlangung der Doktorwürde
der Tierärztlichen Fakultät
der Ludwig-Maximilians-Universität München

Gene Editing in pig models of inherited retinal diseases

von Hannah Franziska Auch

aus Schwäbisch Gmünd

München 2023

Aus dem Veterinärwissenschaftlichen Department der Tierärztlichen
Fakultät der Ludwig-Maximilians-Universität München

Lehrstuhl für Molekulare Tierzucht und Biotechnologie

Arbeit angefertigt unter der Leitung von: Univ.-Prof. Dr. Eckhard Wolf

Mitbetreuung durch: Univ.-Prof. Dr. Nikolai Klymiuk

Gedruckt mit Genehmigung der Tierärztlichen Fakultät
der Ludwig-Maximilians-Universität München

Dekan: Univ.-Prof. Dr. Reinhard K. Straubinger, Ph.D.

Berichterstatter: Univ.-Prof. Dr. Eckhard Wolf

Korreferenten: Univ.-Prof. Dr. Cornelia A. Deeg
Univ.-Prof. Dr. Laurent Franz
Univ.-Prof. Dr. Mathias Ritzmann
Univ.-Prof. Dr. Bernd Kaspers

Tag der Promotion: 11. Februar 2023

Für Grovat

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	REVIEW OF THE LITERATURE	3
1.	Diseases in the focus	3
1.1.	Usher syndrome.....	3
1.1.1.	Clinical phenotype.....	3
1.1.2.	Genetics and Pathophysiology	4
1.1.3.	The USH1C pig model.....	5
1.2.	IRDs based on <i>GUCY2D</i>	7
1.2.1.	Genetics and Pathophysiology	7
1.2.2.	Modelling <i>GUCY2D</i> defects	9
2.	Genome editing techniques	12
2.1.	CRISPR/Cas – the basic principle.....	12
2.2.	Combination of CRISPR/Cas with ssODN or BAC	15
2.3.	Therapeutic applications of CRISPR/Cas	16
2.4.	CRISPR without DSB: Base Editing and Prime Editing	17
III.	ANIMALS, MATERIALS AND METHODS	21
1.	Animals	21
2.	Cells	21
3.	Bacteria	21
4.	Materials	21
4.1.	Devices	21
4.2.	Consumables	22
4.3.	Buffers, Chemicals, Media and Solutions.....	23
4.4.	Enzymes	25
4.5.	Oligonucleotides.....	25
4.6.	Kits	28
4.7.	Software	28
5.	Methods	29
5.1.	DNA Isolation	29
5.2.	Verification techniques	29

5.2.1.	qPCR	29
5.2.2.	End-point PCR	31
5.2.3.	Sequencing	33
5.3.	Nucleofection and generation of single cell clones (SCCs).....	34
5.4.	Lipofection	34
IV.	RESULTS.....	35
1.	Generation of the <i>GUCY2D</i>^{E837D/R838S} pig model	35
1.1.	From design to final vector	37
1.2.	Single cell clones (SCCs).....	41
1.2.1.	Validation of SCCs.....	42
1.3.	Somatic Cell Nuclear Transfer (SCNT).....	47
1.4.	Validation of the genotype in founder piglets.....	48
2.	Therapy approach in the <i>GUCY2D</i>^{E837D/R838S} pig model.....	51
2.1.	Design of disruptive CRISPR/Cas9 therapy	51
2.2.	Efficacy of sgRNA constellations	52
3.	Therapy approach in the USH1C pig model	60
3.1.	Production and use of the USH1C pig model in research.....	60
3.2.	Design of reconstituting CRISPR/Cas9 therapy	64
3.3.	Efficacy of ssODNs.....	65
3.4.	Generation and analysis of SCCs.....	68
3.5.	Additional therapy approach: Prime Editing.....	70
V.	DISCUSSION	73
VI.	SUMMARY.....	83
VII.	ZUSAMMENFASSUNG	85
VIII.	INDEX OF FIGURES.....	87
IX.	INDEX OF TABLES.....	89
X.	REFERENCES	91
XI.	APPENDIX	107
XII.	ACKNOWLEDGEMENTS.....	133

INDEX OF ABBREVIATIONS

AAV	adeno-associated virus
ABE	adenine base editors
ABR	auditory brainstem response
adCD	autosomal dominant cone dystrophy
adCRD	autosomal dominant cone rod dystrophy
aEJ	alternative end-joining
asgRNA	assistant guide ribonucleic acid
ASSR	auditory steady state response
BAC	bacterial artificial chromosome
BE	base editor
bp	base pair
Cas	CRISPR-associated proteins
CBE	cytidine base editors
cDNA	complementary deoxyribonucleic acid
cGMP	cyclic guanosin monophosphate
CHD	choroideal dystrophy
cm	centimeters
COD	cone dystrophy
CORD6	dominant cone-rod dystrophy based on <i>GUCY2D</i>
CORRECT	consecutive re-guide/re-Cas steps to erase CRISPR/Cas-blocked targets
CP	calyceal process
CRISPR	clustered regularly interspaced short palindromic repeats
dB	decibel
Del	deletion
DFG	Deutsche Forschungsgemeinschaft
DNA	deoxyribonucleic acid
DSB	double strand break
Epo	electroporation
ERG	electroretinography
ET	embryo transfer
FDA	United States Food and Drug Administration
FLPe	flippase
FRT	flippase recognition target

fw	forward
GCAP	guanylate-cyclase activating protein
gDNA	genomic deoxyribonucleic acid
GE	Gene Editing
GFP	green fluorescent protein
HA	homology arm
HDR	homology-directed repair
I-%	intensity percentage
I-Max	intensity maximum
indel	insertion/deletion
iPSC	induced pluripotent stem cell
IRD	inherited retinal disease
I-Vol	intensity volume
kb	kilo bases
LCA	Leber congenital amaurosis
LOWA	loss of wild-type allele
M	molar
MD	macular degeneration
mg	milligram
min	minutes
mL	milliliters
mM	millimolar
mRNA	messenger ribonucleic acid
mut	mutant
n.d.	not determined
neoR	neomycin resistance
NHEJ	non-homologous end-joining
nt	nucleotide
NTC	no template control
OAE	oto acoustic emission
OCT	optical coherence tomography
PAM	protospacer adjacent motif
PBS	primer binding site
PCR	polymerase chain reaction

PE	prime editor
pegRNA	prime editing guide ribonucleic acid
PR	photoreceptor
qPCR	quantitative polymerase chain reaction
RETGC1	retinal-specific guanylate-cyclase 1
rev	reverse
RNA	ribonucleic acid
RP	retinitis pigmentosa
RT	reverse transcriptase
RTT	reverse transcriptase template
s	seconds
SCC	single cell clone
SCNT	somatic cell nuclear transfer
SD	Stargardt disease
sgRNA	single guide ribonucleic acid
SNP	single nucleotide polymorphism
SPP2127	Schwerpunktprogramm 2127
SSA	single-strand annealing
ssODN	single-stranded oligodeoxynucleotide
USH	Usher syndrome
WT	wild type
μ -CT	micro computed tomography
μ g	microgram
μ M	micromolar
μ L	microliter

I. INTRODUCTION

Visual perception is among the most important senses that we possess, and humans as well as many other species rely heavily on its function. While damage from environmental influences can often be prevented by suitable protection or reverted by surgical treatment, we are rendered relatively helpless when it comes to genetic causes of ocular diseases. Inherited retinal diseases (IRDs) cover a broad spectrum of symptoms, ranging from severe forms resulting in legal blindness to milder forms including night blindness or colour vision abnormalities. IRDs affect approximately one in 2000 people worldwide (BERGER et al., 2010). Various genetic reasons and the wide spectrum of clinical features pose huge challenges for researchers and clinicians. The development of suitable animal models has led to a deeper understanding of physiological processes, as well as insight into the genetic and molecular basis of disease mechanisms allowing for progress of many therapeutic strategies (SLIJKERMAN et al., 2015).

Novel treatments like Gene Augmentation and Gene Editing (GE) are on the rise, with therapies for several IRDs already being tested in clinical trials, as reviewed in HU et al. (2021). One of the first trials involving human subjects was treatment of Leber Congenital Amaurosis (LCA1) by supplementary gene therapy (BAINBRIDGE et al., 2008), which also became the first clinically approved gene therapy less than ten years later (FDA, 2017).

The eye is an ideal target for therapeutic GE, as it is easily accessible for surgery and its function can be monitored *in vivo* by non-invasive standard techniques like optical coherence tomography (OCT) and electroretinography (ERG), according to international guidelines for examination and data processing (FERCHER et al., 2003; HOOD et al., 2012). Distinct viral vectors can be placed very close to the targeted cells either by subretinal or intravitreal delivery, with each route and multiple vehicles posing their own advantages and disadvantages (GUPTA & HUCKFELDT, 2017). Although the eye is being claimed to be immune-privileged, any inflammatory process arising from gene therapy might result in disastrous loss of post-mitotic and therefore irreplaceable cells like photoreceptors (PRs). Therefore, the need for nonimmunogenic vectors and a close monitoring of immunological responses in the eye after gene therapy is of utmost importance. Both adeno-associated virus (AAV) and lentivirus, the most frequently used

vectors, have shown to cause rather mild inflammatory symptoms in the eye and are therefore promising for further studies (BENNETT, 2003).

Importantly, the status of the PR itself as a terminally differentiated cell does not only pose challenges, but also opportunities. A change made in the genome of such a cell will remain for the lifetime of the patient. Specifically, GE revolutionized by the discovery of CRISPR/Cas (JINEK et al., 2012) and recently broadened by Base and Prime Editing (KOMOR et al., 2016; ANZALONE et al., 2019), offers very precise targeting and treating of pathological variants and would cause long lasting benefits. In principle, GE offers two main possibilities: either disrupting a dominant negative mutation or, more challenging, repairing a nonsense or missense mutation to bring back physiological function of the concerned pathway and improve quality of life for affected patients (GUPTA & HUCKFELDT, 2017; MCCULLOUGH, 2019).

The way to clinical trials on GE in human patients is certainly long and winding, as it starts with proof of concept studies in cell culture and moves on to pre-clinical trials in suitable animal models. The aim of this thesis is to evaluate the contribution of GE towards therapy of IRDs. I therefore worked with an existing pig model for Usher syndrome that was generated at our institute (GROTZ et al., 2022), and contributed in the generation of a new model for cone-rod dystrophy based on a dominant negative mutation in the *GUCY2D* gene and tested initial GE approaches there.

II. REVIEW OF THE LITERATURE

1. Diseases in the focus

Numerous IRDs are known in human beings, with diverse disease mechanisms and progression rates. Although recent achievements in understanding and treating certain IRDs have been made, many substantial and disease specific difficulties remain in the field. For this reason, our group has decided to promote research by working on pig models with two complementary IRDs: Usher Syndrome (USH) and Cone-Rod Dystrophy (CORD6).

1.1. Usher syndrome

1.1.1. Clinical phenotype

Quite a while ago, the Scottish ophthalmologist Charles Usher described a combined impairment of the auditory, vestibular and visual systems and proposed a hereditary origin (USHER, 1914). Nowadays, Usher syndrome (USH) has been identified as the major cause of combined congenital deafness with developing additional blindness (REINERS et al., 2006), with a worldwide prevalence estimated at 4-17 in 100,000 (TOMS et al., 2020a). Blindness is based on retinitis pigmentosa (RP), a disorder with a highly diverse background, the common feature being a degeneration of rod and cone PRs with USH accounting for about 20-40% of cases (HARTONG et al., 2006). Classification of the three USH subtypes is based on clinical features, assessing hearing, vision and vestibular function. RP onset and visual prognosis differ between subtypes, but advancing course of disease is a common hallmark. Initial night blindness and decline of rods in the periphery is later followed by destruction of cones in the center, adding colour vision loss to the symptoms. For the first time, USH was categorized in the early 1980's (FISHMAN et al., 1983), and a study on a larger cohort of patients aiming at quantifying the extend of vision loss was performed by EDWARDS et al. (1998), with both authors concentrating on visual acuity and visual field impairment. More recently, ERG measurements became useful especially in the beginning of the disease, as they can show reduced and delayed amplitudes even before clinical symptoms are recognized (TOMS et al., 2020a).

Table 1: Usher subtypes based on clinical features

	<u>Hearing</u>	<u>Vision</u>	<u>Vestibular function</u>
USH1	Severe, congenital deafness	Early onset of RP, legal blindness in fourth decade	Vestibular hypofunction, delayed motor development
USH2	Moderate, congenital deafness	Onset of RP in second decade, legal blindness in sixth decade	normal
USH3	Later-onset, progressive, variable hearing loss	Variable onset of RP	Variable, usually mild

Adapted from TOMS et al. (2020a)

1.1.2. Genetics and Pathophysiology

Although genetically very heterogeneous, with at least ten genes being involved, all USH types are inherited in a recessive manner (TOMS et al., 2015). Depending on the respective gene, distinct protein families are affected. In general, USH proteins tend to be located where plasma membranes are being tightly stacked or folded. They work closely together in a network, the so-called Usher complex, and fulfill different functions, such as being scaffold, matrix or cell-to-cell adhesion proteins. Most of them are spliced in several variations and expressed in multiple tissues, but they have all been proven to be of great importance in sensory neurons bearing cilia, therefore affecting both inner ear and retina (REINERS et al., 2006). In the inner ear, stimuli are processed by hair cells possessing highly specialized hair bundles. Each consisting of several stereocilia of different lengths and a kinocilium, the hair bundles mechanically react to sound or movement. By the opening of cation channels, they transfer the mechanical stimulus into an electrical shift of polarization. USH proteins have been shown to be important for the links in between the stereocilia and with the kinocilium during early development as well as in mature cells (MATHUR & YANG, 2015). In the retina, the outer segments of PRs constitutes a modified sensory cilium with membrane discs where the proteins for phototransduction are present. It is connected to the inner segment with a connecting cilium, which plays a role in transportation of newly synthesized components from the inner to the outer segment. In contrast to the inner ear, reports of the exact localization of the USH proteins have been inconsistent, but with a

strong tendency to localize USH proteins to the calyceal process (CP) and sensory cilium, which resemble stereocilia at a structural level. Importantly, CPs are absent in rodents, which may be one of the reasons why existing mouse models of USH poorly represent the retinal phenotype observed in human patients (WILLIAMS, 2008), making these models essential for studying the inner ear defect, but contributing little to understanding the function of USH protein complexes in the retina (SAHLY et al., 2012; EL-AMRAOUI & PETIT, 2014). In addition to the abundance of USH proteins near the sensory compartment of hair cells and PRs, they are also present in the synaptic ribbon, a structure located in the presynaptic terminals of both cells, where synaptic vesicles are being accumulated to the active zone of the synapse (ZANAZZI & MATTHEWS, 2009). It is suggested that USH proteins might be involved in the synaptic vesicle trafficking or in regulating calcium transport. Various studies especially in mice have been shining light onto the complicated interactions in the USH protein network, but the highly variable expression poses a challenge to tracking their presence with established antibodies (COSGROVE & ZALLOCCHI, 2014). There is definitely need of further research to understand the function of the acoustic, visual and vestibular system at the molecular as well as at the clinical level.

1.1.3. The USH1C pig model

To overcome the limitations of the retinal phenotype in USH mouse models, a pig model for *USH1C* deficiency causing the lack of harmonin has been generated (GROTZ et al., 2022). Harmonin has been shown to interact with all USH1 and USH2 proteins, linking the two types on a molecular basis and underlining its fundamental role in the development of the characteristic phenotype (REINERS & WOLFRUM, 2006). For this reason, a knock-out model of its main splice variants USH1C_a, _b and _c was designed on the base of a patient-specific mutation. Importantly, different harmonin isoforms take part in stabilizing the actin filaments of stereocilia already in the growing state, fulfilling a crucial role in physiological hair bundle architecture in hair cells (EL-AMRAOUI & PETIT, 2005), but little was known about distinct harmonin isoforms in the eye. Genetic modification of the porcine genome took place via CRISPR/Cas induced homologous recombination of a bacterial artificial chromosome (BAC) vector in porcine kidney cells (VOCHOZKOVA et al., 2019). The established modification carries the humanized exon 2 and its adjoining intronic regions instead of the orthologous

porcine sequence. Having implemented the c.C91T/p.R31X nonsense mutation and three intronic SNPs identified in a human patient, it copies a clinically relevant mutation.

During the breeding process, genotyping of the first heterozygous litters from founder animals with wild type pigs revealed a knock-in of the R31X mutation on only one allele, whereas the other one suffered a large deletion in the *USH1C* gene (GROTZ, 2021). As this deletion also results in a functional knock-out, absence of harmonin is achieved nevertheless, and even reflects the situation in R31X patients better, as all of them are compound heterozygous. With further breeding the following generations, animals carrying a homozygous R31X mutation are present by now, as well as animals with deletion on one allele, which allows for further investigation of a potentially different phenotype. All *USH1C* offspring, whether carrying homozygous R31X mutation or compound heterozygosity with deletion of exon 2, show pronounced circling right after birth. This indicates an even more prominent vestibular phenotype than seen in human patients, but ceases during the first few days of life compensated by vision and statokinetic sensing (GROTZ et al., 2022). During complementary analysis of auditory brainstem response (ABR) for proof of hearing deficits, *USH1C* pigs showed no response to a click stimulus of up to 120 dB sound pressure level from a very early age on. This is being supported by daily observations in the stable, where *USH1C* piglets do not react to their mothers' nurturing calls but have to be woken up by nudging by the caretaker or their littermates.

Mainly, Grotz' thesis aimed at characterizing the vision of the *USH1C* pig model and compare it to clinical characteristics in human patients. Considering the limitation of established tools such as ERG and OCT to estimate immediate vision, and the difficulties of transferring common vision tests in humans to pigs, animals were trained to run in distinct obstacle courses under different light conditions. Several parameters such as time, obstacle contacts and general movement characteristics like straightness and anticipation were analyzed. In short, *USH1C* pigs needed more time than wild type (WT) pigs and they touched the obstacles more often. Especially under dark conditions, frontal contacts to barriers were occurring more often. A behaviour that was almost exclusive in *USH1C* pigs was turning around and walking in circles in the obstacle course, sometimes resulting in prolonged loss of orientation indicating the re-appearance of balance problems when compensatory mechanisms were challenged under a new environment,

especially in the dark. Trajectory analysis of the routes the pigs took in the course showed the USH1C pigs to walk straight towards an obstacle and abruptly change direction shortly before them, whereas WT pigs anticipated them from a distance and chose a more bending and “elegant” path around. Impaired vision was also supported by ERG at 12 months of age revealing both rod and cone photoreceptor cell responses to be declined compared to age-matched WT controls, this being in line with an early stage of the disease in human patients. OCT imaging showed only minor differences in retinal thickness with the structural architecture remaining intact. At a histological level, however, the absence of harmonin seems to be causing changes in the conformation of the synapse and the connecting cilium in cones, but most obvious in the horizontal stacking of the membrane discs located in the outer segment in rods. Together, the clinical and the structural alterations in the eye indicate a progressing rod-cone dystrophy bearing the characteristics of a true retinal ciliopathy as described in BUJAKOWSKA et al. (2017).

1.2. IRDs based on *GUCY2D*

1.2.1. Genetics and Pathophysiology

In contrast to the well and longly characterized USH, mutations in the *GUCY2D* gene causing IRDs have been described only several years ago. The complicated process of phototransduction in the PR includes tightly regulated homeostasis of calcium by a cascade of regulatory enzymes, messenger molecules and membrane channels that heavily rely on each other, reviewed in ARSHAVSKY and BURNS (2012). Coding for the retinal-specific guanylate-cyclase 1 (RETGC1), *GUCY2D* plays a major role in the calcium metabolism and is responsible for several characteristic symptoms (KELSELL et al., 1998). RETGC1 produces cyclic guanosine monophosphate (cGMP), a second messenger molecule responsible for opening calcium channels and allowing the influx of calcium ions into the cell. Low calcium levels are recognized by several GC-activating proteins (GCAPs), which stimulate RETGC1 activity, therefore restoring cGMP levels quickly after light adaption and allowing for a new depolarization to take place (DIZHOOR & PESHENKO, 2021). The process of phototransduction is mainly located in the outer segments of PRs (DIZHOOR et al., 1994), but RETGC1 has also been reported to be present in the synaptic region connecting the PR to the bipolar cell (DUDA et al., 2002).

Of note, mutations in the *GUCY2D* gene can affect function in two different ways. First, in the case of deleterious mutations, *GUCY2D* is associated with LCA1 (PERRAULT et al., 1996; PERRAULT et al., 2000). The authors describe missense or frameshift mutations that cause lack of RETGC1 when showing up on both alleles, which will eventually lead to permanently closed cGMP-dependent cation channels in the retina. In addition to the severe phenotype in homozygous patients, reduced ERG amplitudes have been reported in heterozygous carriers at varying levels (GALVIN et al., 2005). An interesting hypothesis has been postulated to explain variation on the basis of the causative mutation (SHARON et al., 2018). On the one hand, gene products might be removed by nonsense-mediated messenger RNA (mRNA) decay or lead to an unstable protein that will be degraded. In those cases, RETGC1 with a normal structure, but just half of the amount is present, still keeping up normal photoreceptor function (haplosufficiency). On the other hand, heterozygous carriers of missense mutations might produce mutated proteins that actually interfere with the physiological process. Since RETGC1 is forming dimers, WT and mutant (mut) forms will comprise normally active WT/WT combinations, WT/mut versions with slightly reduced activity and probably inactive mut/mut dimers. According that scenario, the remaining WT/WT dimers, along with the partially active WT/mut dimers will keep up a substantial, but still decreased level of function, resulting in a mild but notable phenotype. Overall, there is common sense that the degree of impairment is rather low and indeed, most heterozygous carriers deny significant clinical symptoms.

In contrast, mutations in the dimerization domain mapped to exons 12 to 14 have a consistent dominant effect causing autosomal dominant cone and cone-rod dystrophies (adCDs, adCRDs) (SALEHI CHALESHTORI et al., 2019). Characterized by the initial degeneration of cone PRs, causing early loss of visual acuity and colour vision (no S cone), later stages also involve degeneration of rod PRs leading to myopia, nyctalopia and peripheral visual field loss (GREGORY-EVANS et al., 2000). Affected individuals report nystagmus from childhood on and early onset of cone dysfunction before six years of age. This is supported by ERG responses that indicate a primary loss of cones, followed by rods afterwards, which might be due to the fact that RETGC1 is primarily expressed in cones (PERRAULT et al., 1996). Other reports include a morphologically relatively intact retina, but with impaired function and in contrast to previous statements a preserved outer visual field border (KITIRATSCHKY et al., 2008).

While the recessive mutations in *GUCY2D* plainly represent a version of the multiple loss of function IRD pathologies, the dominant negative variants gained specific interest for their mechanistic properties. In its dominant negative form, RETGC1 is already active at higher calcium levels, and the constant production of cGMP leads to constantly open calcium channels, resulting in extremely high intracellular calcium levels (SHARON et al., 2018). The degeneration and ultimately destruction of PRs is supposedly due to either the high calcium levels or the cytotoxic potential of cGMP itself (XU et al., 2013; GIARMARCO et al., 2017). When looking at the broad spectrum of diseases caused by the manifold changes in *GUCY2D*, it becomes clear that this gene surely is a hotspot for mutations. Throughout the dimerization domain, codon 838 seems to be especially sensitive for modifications and variants of the original R838 appear in independent IRD families, either alone or in combination with altered codons 837 and 839. The codon 838 is proposed to suffer frequent erroneous deamination of methylated cytosine at DNA level, eventually leading to the respective mutations at protein level and accounting for about one third of adCDs and adCRDs (KITIRATSCHKY et al., 2008). One of the most common mutations here is the heterozygous amino acid substitution from glutamine to aspartic acid at codon 837 (E837D) and from arginine to serine at codon 838 (R838S) causing CORD6 (GREGORY-EVANS et al., 2000), evoking interest in research for this particular disease.

1.2.2. Modelling *GUCY2D* defects

Several animal models have been generated for both the dominant and recessive forms of mutated *GUCY2D* in distinct species (SHARON et al., 2018). Given the variety of RETGC-encoding proteins and the inconsistent nomenclature of *GUCY2*-genes, care must be taken of the correct assignment. Apparently, an X-linked gene is consistently designated as *GUCY2F/Gucy2f* across species, lacking any association to human disease. What has been designated as human *GUCY2D* finds its counterpart in the conserved *KCN3B-CENTROB-GUCY2D-ALOX18B-ALOX12B* locus in most mammals. The corresponding gene, however, is annotated as *Gucy2e* in rodents and *gucy2f* in zebrafish. In rodents, an intact gene *Gucy2d* encodes for an alternative RETGC, but this gene has been found inactivated in primates, ungulates and carnivores. For this reason, examining the RETGC function might be compromised in rodents due to the existence of a potentially redundant RETGC.

Table 2: Animal models for retinal guanylate cyclase defects

<u>animal</u>	<u>mutation</u>	<u>method</u>	<u>phenotype</u>
mouse	<i>Gucy2e</i> ^{-/-} knock-out	site-directed	cone loss, rods viable but reduced ERG
mouse	<i>Gucy2f</i> ^{-/-} knock-out	site-directed	normal retinal morphology and ERG
mouse	<i>Gucy2e</i> ^{-/-} / <i>Gucy2f</i> ^{-/-} double knock-out	site-directed	degeneration of outer segments, no ERG response
mouse	<i>Gucy2e</i> ^{R838S}	additive transgene	advancing retinal degeneration with loss of vision
pig	<i>GUCY2D</i> ^{E837D/R838S}	additive transgene	very variable phenotype, reduced cone function and disturbed organization
chicken	<i>GUCY1B1</i> knock-out	spontaneous	normal retinal morphology but no ERG response
zebrafish	Blocking of translation and splicing (knock-down)	induced	vision loss, degeneration of outer segments
zebrafish	<i>gucy2f</i> ^{E837D/R838S}	additive transgene	reduced number and irregular cone morphology, no vision loss detected

Adapted from SHARON et al. (2018)

In a first *in vitro* examination on the effect of dominant negative *GUCY2D* mutations, RETGC1 with a R838C mutation showed a dramatically increased affinity to CGAP-1, resulting in abnormal high levels of cGMP and, finally cell death (TUCKER et al., 1999). This was confirmed in a transgenic mouse model, where cDNA coding for the mutant human RETGC1 (R838S) was inserted under the control of a rod specific promoter (DIZHOOR et al., 2016). In this publication, the authors claimed potential influence of two intact *Gucy2e* alleles and therefore crossed the *Gucy2e*^{R838S} transgene into a *Gucy2e*^{-/-} background to eliminate mouse RETGC1. Here, the transgenic RETGC1 is localized in the rod outer segments, resembling the situation in WT mice. Increased RETGC1 activity results in higher

calcium levels than in the control group and reduced rod ERG responses were already detected at one month of age. The total number of PR nuclei is dramatically reduced, leaving only one third remaining by the end of six months. To further investigate the role of CGAPs, the R838S-transgenic mouse lines were cross-bred to mice having reduced or lacking expression of one or both CGAPs (SATO et al., 2018). Here, a complete absence of CGAPs prevents calcium feedback on RETGC1 and retinal morphology stays intact even after five months. Importantly, the classical additive transgenic approach resulted in lines with different expression levels, resulting in distinct speed of retinal deterioration and severities of phenotype. A further drawback is the usage of a rod specific promoter, preventing insight into potentially essential pathogenic processes in cones. A complementary cone-specific approach in an E837D/R838S-transgenic zebrafish model resulted in reduced numbers and irregular morphology of cone PRs (COLLERY et al., 2013). However, no decline of visual function was reported in 5-day-old larvae.

Notably, the dominant negative form of *GUCY2D* was also among the first transgenic approaches in a large animal species (KOSTIC et al., 2013). The generation of transgenic pigs was accomplished via lentiviral transgenesis of the E837D/R838S cDNA under a cone specific promoter. An early onset reduction of cone ERG amplitude was observed in the transgenic group, although with high variability among individual founder animals. Behavioural tests were conducted to assess the useful vision at different time points. While no significant difference showed at an age of 11 weeks, the transgenic pigs later seem to develop a visual impairment to some extent, as their performance becomes more unsteady and especially a kind of behaviour called “alternative prospection”, meaning evaluation of obstacles by other senses than vision, rises in frequency along with time needed to complete the obstacle course. Still, results are difficult to justify and vary widely between individuals. While OCT measurements did not reveal any significant difference, a change in the cellular structure of the outer and inner nuclear layer of PRs was observed in some of the animals, as well as Müller glia cell activation, a sign of immunoreactions in a diseased retina. Overall, the phenotypic heterogeneity reported is in line with random and variable insertion and expression of the transgene in models generated by additive gene transfer.

2. Genome editing techniques

The aforementioned drawbacks for transgenic animal models were mainly caused by the previously limited genetic toolbox for the creation of models and the consequent phenotypical variation, depending on transgene integration sites, copy numbers and expression levels. In the past decade, however, innovative tools have been developed that did not only facilitate the tailoring of novel animal models, but also propose usage as therapeutics. Among the variety of designer nucleases, CRISPR/Cas undeniably represents the tool with the utmost potential and perspectives.

2.1. CRISPR/Cas – the basic principle

Clustered Regularly Interspaced Short Palindrome Repeats (CRISPR)/Cas is a biological defense mechanism in prokaryotes that has gained outstanding relevance in biotechnology when it was adapted as a genetic engineering tool. It was not so long ago that CRISPR/Cas has been identified as a kind of “adaptive immune system” in prokaryotes (MAKAROVA et al., 2006). Basically, CRISPR/Cas integrates fragments of invading phages and plasmids into the host genome in a clustered manner. These fragments are then expressed in conjunction with adapter elements that bind to Cas proteins, comprising the potential to create double-strand breaks (DSBs) in target DNA. As a consequence, the endogenized fragments guide the disruptive Cas activity only to exogenous phage and plasmid DNA, thereby preventing re-infection (BARRANGOU et al., 2007; BROUNS et al., 2008). The breakthrough for using CRISPR/Cas in biotechnology occurred, when JINEK et al. (2012) demonstrated CRISPR/Cas9 to be a programmable DNA endonuclease, guiding Cas to its cutting site by a customized site-specific RNA molecule. Importantly, this site-specific RNA molecule is relatively flexible in its design as long as it is adjacent to the so-called protospacer adjacent motif (PAM), a short domain of a few nucleotides length that is indispensable for interaction of Cas with DNA. In the case of the Cas9 used by JINEK et al. (2012), PAM is a “GG” dinucleotide. Shortly afterwards, CONG et al. (2013) and MALI et al. (2013) proved the prokaryotic system to be working also in higher, eukaryotic cells and refined the characteristics and parameters making it such an effective, precise and versatile tool. Since then, CRISPR/Cas has revolutionized the field of genetics in a

way that has been hardly if ever seen for any other tools. Besides scientific merits, this outstanding potential gained also economic consequences. Both groups from the University of Berkeley, California as well as the one from the Broad Institute of MIT and Harvard in Cambridge, Massachusetts filed patent applications, sparking a tremendous battle over who owns the rights to this technology that is still ongoing by now (LEDFORD, 2016, 2022).

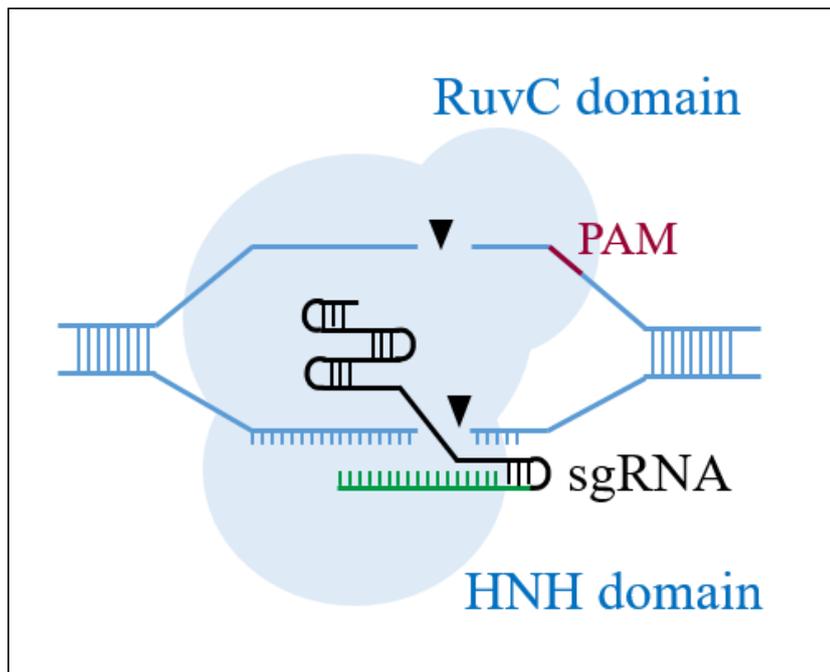


Figure 1: Basic principle of CRISPR/Cas

Adapted from HSU et al. (2014)

RuvC domain: named after an *E. coli* protein involved in DNA repair, cuts the nontarget strand;
HNH domain: named after its composition of histidine and asparagine, cuts the target strand;
sgRNA: single guide ribonucleic acid

Fueled by rapid progress in design and new variants, applicability of different CRISPR/Cas systems for genome engineering ranges from examining the physiological function of genes to enhancing production capacity or resistance against infections in agriculture, as well as the studying of pathologic consequences of disease-causing variants in model organisms or therapeutic approaches in medicine, reviewed in HSU et al. (2014). In its fundamental principle, CRISPR/Cas makes use of the distinct mechanisms that cells possess to repair DNA double strand breaks (DSBs). The distinct DSBs can occur due to endogenous or exogenous reasons, and the cell's viability depends on fast and tightly regulated repair pathways (SHRIVASTAV et al., 2008). Two main pathways have been described

and their variations are used for specific GE tasks. One major pathway is non-homologous end-joining (NHEJ), where the two ends of the DNA strand are fused back together, often altering the sequence by adding or deleting a few nucleotides, resulting in frameshift or nonsense mutations (BIBIKOVA et al., 2002). The alternative repair pathway is homology-directed repair (HDR), where a homologous sequence is used as a repair template (SHRIVASTAV et al., 2008). Along with several minor repair pathways like single-strand annealing (SSA) or alternative end-joining (aEJ), NHEJ and HDR compete for the repair of a DSB, and the decision which path is being chosen is influenced by multiple factors like cell type, stage of cell cycle or DNA structure, reviewed in SCULLY et al. (2019).

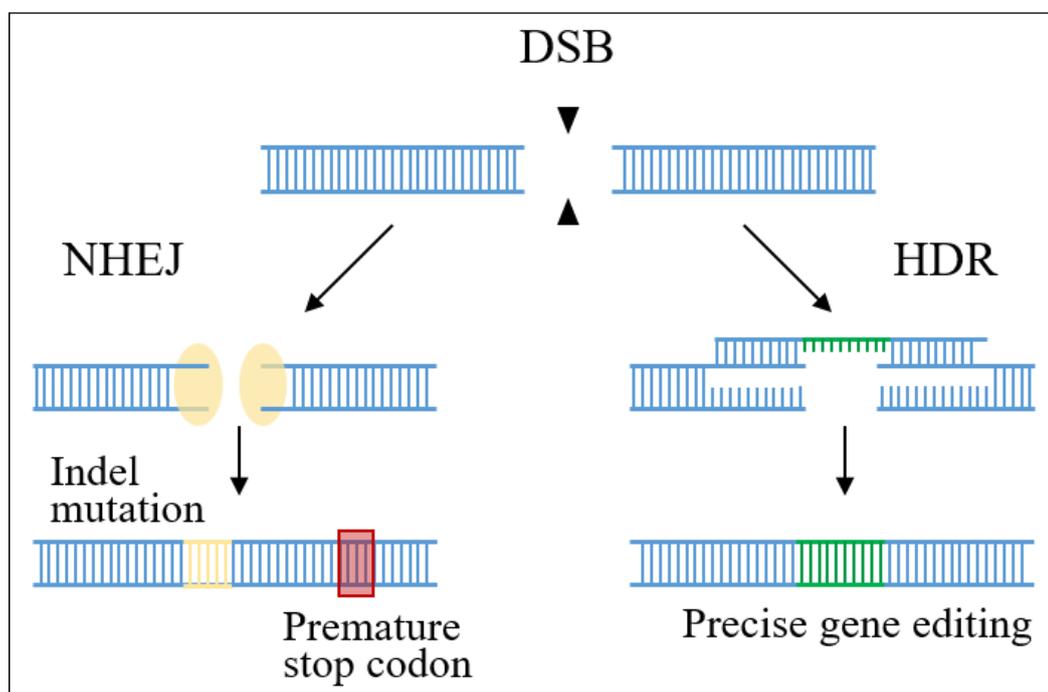


Figure 2: Repair mechanism of the cell

Adapted from HSU et al. (2014)

Initially, the use of CRISPR/Cas in biotechnology has been made along the dominant DSB repair mechanisms NHEJ and HDR. In the case loss of function of a target gene is sufficient to achieve the desired effect, GE by NHEJ is favoured. For this reason, the efficiency of the disruptive effect mainly depends on the power to introduce DSBs. Several animal models have been generated that way, starting from simple knock-out of one gene to multiple loss of function in different genes, using either one or several sgRNAs working alone or in combination with each other in different species like mouse, zebrafish and cynomolgus monkey (JAO et

al., 2013; LI et al., 2013; WANG et al., 2013; NIU et al., 2014). In the case a specific modification is desired, a repair template is needed to introduce the intended variant via HDR. This approach mostly fits better for the introduction of a defined mutation, but in addition to the efficacy to introduce DSBs, the success rate of homologous recombination depends on the activity of HDR mediated repair mechanisms in the target cell type and the design of the repair template.

2.2. Combination of CRISPR/Cas with ssODN or BAC

The creation of knock-in models for certain diseases is of special interest for mimicking the phenotype seen in human patients. The exact change of specific nucleotides or the replacement of a large part of the genome, however, pose a greater challenge than simple disruption of a specific locus. A relatively simple strategy for creation of small changes in the genome is the use of single-stranded oligodeoxynucleotides (ssODNs) as a donor template (WANG et al., 2013). To improve the comparably low efficacy of HDR, many attempts were made to improve the design of the donor DNA. Importantly, RICHARDSON et al. (2016) showed that Cas9, after cutting the target DNA, will asymmetrically release the 3' end of the strand that is not complementary to the sgRNA. Therefore, best results are obtained when the ssODN is complementary to the non-target strand and located about one third on the PAM-distal and two thirds on the PAM-proximal side. To overcome the problem of potential re-cutting of a successful edit due to the CRISPR/Cas system staying active, it is advisable to introduce a blocking mutation located at the PAM or the gRNA binding sequence, preventing the Cas9 from targeting the locus a second time (PAQUET et al., 2016). The blocking mutation needs to be a silent mutation in the case the PAM is within a coding region to avoid unwanted side effects by a change in the amino acid sequence. An alternative to silent blocking mutations is their subsequent removal by CORRECT (consecutive re-guide or re-Cas steps to erase CRISPR/Cas-blocked targets), enabling scarless GE.

An alternative for achieving more complex modifications in the genome is a combination of CRISPR/Cas with classical targeting vectors carrying large modification segments, extended homologous arms and a cassette for positive selection. For historical reasons, bacterial artificial chromosomes (BACs) are an appropriate tool for generating targeting vectors for primary cells. BACs are

circular DNA molecules with a length of 150-300 kb (SHIZUYA et al., 1992), and have previously been used to decipher the genome of humans, pigs and other species with large genomes. Although current sequencing techniques work without them, BAC libraries are still a valuable asset for gene modification (LANDER et al., 2001; ARCHIBALD et al., 2010). The considerable size of BACs allows for the introduction of large modifications and facilitates the incorporation of complex modifications into the genome by providing long homologous arms on both sides. For genetic engineering, BACs are modified by bacterial recombination ("recombineering") (ZHANG et al., 1998; WARMING et al., 2005) to carry the intended mutation. Several models for important human diseases have been created by using modified BACs without any stimulus (SONG et al., 2010; KLYMIUK et al., 2012). The combination with CRISPR/Cas empowered the capacity of BAC vectors substantially (GROTZ et al., 2022).

2.3. Therapeutic applications of CRISPR/Cas

When moving from the generation of models towards therapeutic applications of CRISPR/Cas, the consequences of DSBs are of significant concern, in particular if those appear not only at the desired site, but somewhere else in the genome (off-target effects). Mismatches between the gRNA and the complementary target sequence are tolerated by the CRISPR/Cas system according to location and pattern relative to the cutting site (HSU et al., 2013). Variants of CRISPR/Cas being more stringent for gRNA specificity have been developed to overcome this problem (TSAI & JOUNG, 2016). Several web-based tools, such as CRISPOR or the Cas-OFF finder are available today that check for genome-wide off-target effects in many species (BAE et al., 2014; CONCORDET & HAEUSSLER, 2018).

In therapy, similar as in the generation of models, two different strategies are possible that go along the main pathways for DSB repair. Making use of HDR, ssODN-mediated repair of a point mutation was successfully reported in the rd1 mouse model of RP (WU et al., 2016), while BAKONDI et al. (2016) utilized NHEJ for the knock-out of the dominant negative mutation in the rhodopsin gene of a rat model for autosomal dominant RP. Following the disruptive approach, the first clinical trial of *in vivo* GE in humans has started enrollment in 2019 and will treat patients suffering from LCA10 by removing the disease-causing aberrant splice site in the *CEP290* gene (ClinicalTrials.gov Identifier: NCT03872479). A variant for

applying CRISPR in a therapeutic approach is the *ex vivo* treatment of induced pluripotent stem cells (iPSCs) gained from patients that are being transplanted back into the host after successful editing (BASSUK et al., 2016; BURNIGHT et al., 2017), with treatment for Sickle Cell Disease being in clinical trials (ClinicalTrials.gov Identifier: NCT03745287). Still, a number of severe diseases affect non-dividing cells, where efficacy of HDR is low (CONG et al., 2013; RAN et al., 2013). Along with concerns about the long-term effects of applying DNA-cleaving proteins, recent focus is being put on the idea of GE without DSBs.

2.4. CRISPR without DSB: Base Editing and Prime Editing

As Cas9 contains two different nuclease domains, HNH and RuvC (JINEK et al., 2012), each of them cleaving one DNA strand, it is possible to tailor its cutting abilities to cut either the gRNA-binding strand or the opposite strand. This sparked the further development of Cas9 as a biotechnological tool by creating catalytically impaired nickases that only nick one strand of DNA instead of creating a full DSB. As an off-target nick is much less critical than the off-target effects of a DSB, adopting Cas-proteins to nickase function is a tremendous progress. The first approach were cytidine base editors (CBEs), facilitating the change of one specific nucleotide to another without the need for DNA cleavage and donor DNA. CBEs combine dCas9 covalently linked to a cytidine deaminase enzyme converting cytidine to uridine, thereby creating a C to T change in the DNA (KOMOR et al., 2016). Advantages of this technique are the very low occurrence of NHEJ, as indel formation is dependent on DSBs, no danger of re-cut, since the product is no longer able to bind, as well as its function in non-dividing cells (YEH et al., 2018). Disadvantages are the relatively broad activity window of five nucleotides, where multiple Cs will be converted if present, the need for a compatible PAM and, of course, the defined base conversion as the most obvious limitation (KOMOR et al., 2016). Subsequently, newer generations of base editors (BE1-4) were developed with improved editing efficacy, for example by additionally nicking the non-edited strand (KOMOR et al., 2018). For spatiotemporal control of the activity level, a ligand responsive, self-cleaving BE minimizes the risk of unwanted by-products and off-target effects (TANG et al., 2017). Substantial increase in the versatility of base editing came with the development of adenine base editors (ABEs) converting A to G. By now, the 7th generation of ABEs achieve up to 68% efficacy with fewer

than 1% indel frequency, compared to classical CRISPR-mediated HDR at the same five loci with efficacy up to only 4% with indel frequency up to 10% (GAUDELLI et al., 2017). Together CBEs and ABEs can mediate four possible transition mutations: C to T, A to G, T to C and G to A. First *in vivo* applications in mouse models for Hutchinson Gilford syndrome and Sickle Cell Disease show promising results, but translation to the clinic has not yet been achieved (KOBLAN et al., 2021; NEWBY et al., 2021).

To overcome the limited spectrum of edits possible with BEs, additional combinations of Cas9 and other enzymes led to the development of an alternative modification of the CRISPR/Cas system. Prime Editing can introduce a wide variety of changes without DSBs or donor templates. The technique was first described by ANZALONE et al. (2019), based on a fusion protein consisting of a dCas9 and a reverse transcriptase (RT), combined with a prime editing guide RNA (pegRNA) that specifies the target site and encodes the desired editing template. The pegRNA consists of a spacer sequence that binds the target, a primer binding site (PBS) and a template sequence for the reverse transcriptase (RTT) with the encoded mutations. Similar to BEs, prime editors (PEs) of several generations (PE1-3) have been developed, with PE3 increasing efficacy by additionally nicking the non-edited strand. Efficacy in general is claimed to be lower than current generation CBEs or ABEs, but Prime Editing is much more flexible since introduction of insertions up to 44 base pairs, deletions of up to 80 base pairs and any of the twelve possible nucleotide transitions or transversions and combinations thereof are possible. The targeted locus is more freely to be chosen, as locations from 3 upstream to 29 base pairs downstream of the PAM appear working, and modification is also more specific than BEs which is limited when multiple identical nucleotides are present inside the activity window of 5 nt. Compared to HDR via CRISPR/Cas plus ssODN, prime editing seems to range in similar efficacy levels, but produces far less indel by-products, and its application in post-mitotic neurons in mice by the same group holds promise for therapeutic applications in the future.

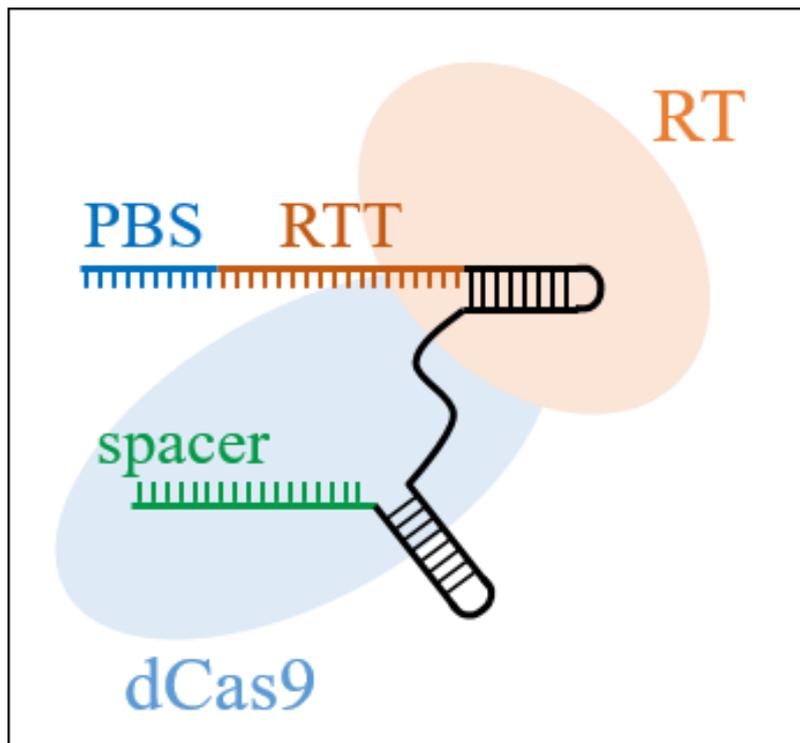


Figure 3: Components of Prime Editing

Adapted from ANZALONE et al. (2019)

RT: reverse transcriptase; dCas9: catalytically impaired Cas9; PBS: primer binding site; RTT: reverse transcriptase template

III. ANIMALS, MATERIALS AND METHODS

1. Animals

All animal experiments were carried out in compliance with the German Animal Protection Law, approved by the Regierung von Oberbayern, the responsible animal welfare authority (AZ 55.2-1-54-2532-70-12 and AZ 02-17-136).

Heterozygous female *GUCY2D*^{E837D/R838S} piglets were produced by somatic cell nuclear transfer (SCNT) followed by embryo transfer (ET) according to standard procedures (KUROME et al., 2015). No apparent phenotype was shown in the first week after birth.

2. Cells

Primary fibroblasts of porcine kidneys derived from the cell lines 0407 (WT), 539 (WT), 5615 (*USH1C*^{R31X/USH1C^{Del}) and 12431 and 12433 (*GUCY2D*^{E837D/R838S}).}

3. Bacteria

E. coli strain DH10B original, heat-shock competent (own production)

E. coli strain SW106 original, electro-competent, recombineering competent (own production)

4. Materials

4.1. Devices

Accu-jet® pro	Brand GmbH, Wertheim
Analytik Jena US UVP GelStudio Plus	Thermo Fisher Scientific, USA
Barnstead™ Easypure™ II	Thermo Fisher Scientific, USA
Benchtop 96 tube working rack	Stratagene, USA
Bunsenbrenner Gasprofi 1SCS	WLD TEC, Arenshausen
Cellavista® automated cell culture microscope	Synentec, Elmshorn
Corning® CoolCell™	Sigma Aldrich, USA

Daewoo KOC-154K microwave	Daewoo, South Korea
Eppendorf Centrifuge 5417 R	Eppendorf, Hamburg
Eppendorf Centrifuge 5424	Eppendorf, Hamburg
Eppendorf Centrifuge 5804	Eppendorf, Hamburg
Eppendorf Centrifuge 5910 R	Eppendorf, Hamburg
Eppendorf Eporator®	Eppendorf, Hamburg
Erlenmeyer flask 500 mL, 1 L	Schott AG, Mainz
Gel chambers EasyCast™ B2	Thermo Fisher Scientific, USA
Grant JB Nova 5 water bath	Grant Instruments Ltd, UK
Heraeus Biofuge pico	Heraeus, Hanau
HeraSafe workbench	Heraeus, Hanau
Incubator with CO ₂ and humidity regulation	Binder, Tuttlingen
Labcyler thermocycler	SensoQuest GmbH, Göttingen
Light Cycler 96® qPCR	Roche, Switzerland
Nucleofector® 2b Device	Lonza, Switzerland
Photometer GeneQuant Pro	Amersham BioSciences, UK
Pipettes	Gilson Inc., USA
Plastic cuvettes for photometer	Carl Roth, Karlsruhe
Select vortexer	Select BioProducts, USA
Shaking Incubator GFL 3031 with orbital motion	Lauda-GFL, Burgwedel
SimpliNano™ spectrophotometer	Biochrom GmbH, Berlin
Stationary incubator	Memmert, Schwabach
Thermoblock HTM	HTA-BioTec, Bovenden
VHC Pro vacuum pump	vacuubrand GmbH, Wertheim

4.2. Consumables

Cell culture plates	
(10cm, 6 well, 96well half area, 96well full area)	Sigma-Aldrich, USA
Centrifuge tubes with round bottom 50mL	TPP, Switzerland
Cryoconservation vials 1.5 mL	Thermo Fisher Scientific, USA
Culture tubes 12 mL	Carl Roth, Karlsruhe
Electroporation cuvettes	Carl Roth, Karlsruhe
Glass pipets	Thermo Fisher Scientific, USA
Greiner Bio-One Cellstar™ tubes 15 mL, 50 mL	Thermo Fisher Scientific, USA
Inoculation loops	Carl Roth, Karlsruhe

Latex gloves	Brightway, Malaysia
PCR reaction tubes (0.2 mL)	Brand GmbH, Wertheim
Petri-dish 10 cm for bacteria	Sigma-Aldrich, USA
Pipet tips	Eppendorf, Hamburg
Pipet tips with filter	Eppendorf, Hamburg
qPCR plates 96 wells	Eppendorf, Hamburg
Safe-Lock reaction tubes 1.5 mL, 5 mL	Eppendorf, Hamburg

4.3. Buffers, Chemicals, Media and Solutions

2-Mercaptoethanol	Sigma Aldrich, USA
Ampicillin	Carl Roth, Karlsruhe
Bromphenolblue	Carl Roth, Karlsruhe
Cell culture grade water	Biowest, France
Chloramphenicol	Carl Roth, Karlsruhe
Collagen	Thermo Fisher Scientific, USA
DMSO (Dimethylsulfoxide)	Sigma Aldrich, USA
dNTP mix (100mM)	Agilent Technologies, USA
dNTPs (dATP, dCTP, dGTP, dTTP)	Thermo Fisher Scientific, USA
double-distilled water	by Barnstead™ Easypure™ II
DTT (Dithiothreitol)	Thermo Fisher Scientific, USA
EDTA (Ethylenediaminetetraacetic acid)	Carl Roth, Karlsruhe
EtOH (Ethanol 99.8%)	Carl Roth, Karlsruhe
Fetal Calf Serum	Thermo Fisher Scientific, USA
GelRed® Nucleic Acid Gel Stain	Biotium, USA
Gene Ruler™ 1 kb DNA ladder	Thermo Fisher Scientific, USA
Geneticin™ Selective Antibiotic	Thermo Fisher Scientific, USA
Gibco™ DMEM GlutaMAX™	Thermo Fisher Scientific, USA
Gibco™ Hepes Buffer Solution	Thermo Fisher Scientific, USA
Gibco™ MEM NEAA	Thermo Fisher Scientific, USA
Glucose	Carl Roth, Karlsruhe
Glycerol	Carl Roth, Karlsruhe
HCl (Hydrochloric acid)	Bernd Kraft GmbH, Duisburg
Herculase II Reaction Buffer	Agilent Technologies, USA
HoAc (Acetic acid)	Carl Roth, Karlsruhe
Isoamyl alcohol	Carl Roth, Karlsruhe

iPrOH (Isopropanol)	Carl Roth, Karlsruhe
Kanamycin	Carl Roth, Karlsruhe
KOAc (Potassium acetate)	Carl Roth, Karlsruhe
LB (lysogeny broth) Agar	Carl Roth, Karlsruhe
LB (lysogeny broth) Medium	Carl Roth, Karlsruhe
MgCl ₂ (Magnesium chloride)	Qiagen, Hilden
NaCl (Sodium chloride)	Carl Roth, Karlsruhe
NaOAc (Sodium acetat)	Carl Roth, Karlsruhe
NaOH (Sodium hydroxide)	Carl Roth, Karlsruhe
PEG 8000 (Polyethylene glycol)	Sigma Aldrich, USA
Phenol	Carl Roth, Karlsruhe
PBS (Phosphate-buffered saline)	Sigma Aldrich, USA
Q-Solution	Agilent Technologies, USA
Restriction enzyme reaction buffers	Thermo Fisher Scientific, USA
SDS (Sodium dodecyl sulfate)	Carl Roth, Karlsruhe
Sequencing buffer	Applied Biosystems, USA
SOB (super optimal broth) Medium	Carl Roth, Karlsruhe
T4 Ligation buffer	Thermo Fisher Scientific, USA
Trichlormethan/Chloroform	Carl Roth, Karlsruhe
Tris (Tris-hydroxymethyl-aminomethane)	Carl Roth, Karlsruhe
Trypsin	Thermo Fisher Scientific, USA
Universal Agarose	Bio&SELL, Nuremberg

In-house preparations:

PCiA: phenol, chloroform, isoamyl alcohol (25:24:1)

PK Buffer: 200mM Tris, 1M NaCl, 40mM EDTA

Plasmid A: 50 mM Glucose, 25 mM Tris/HCl pH 8.0, 10 mM EDTA/NaOH pH 8.0

Plasmid B (freshly prepared): 0.1 M NaOH, 0.5 % SDS

Plasmid C: 3 M KOAc pH 4.8 with 9M HOAc

STE: 10 mM Tris/HCl pH 8.0, 100 mM NaCl, 1 mM EDTA/NaOH pH 8.0

T Buffer: 10mM Tris/HCl

TAE Buffer: 2 M Tris, 50 mM EDTA, 1 M HoAc

4.4. Enzymes

BigDye® Terminator v3.1	Applied Biosystems, USA
Fast Start SYBR® Green Master	Roche, Switzerland
Herculase II Fusion DNA Polymerase	Agilent Technologies, USA
Proteinase K	Agilent Technologies, USA
Restriction enzymes (<i>BpiI</i> , <i>EcoRI</i> , <i>FspI</i> , <i>KpnI</i> , <i>NheI</i> , <i>NotI</i>)	Thermo Fisher Scientific, USA
RNAse A	Thermo Fisher Scientific, USA
T4 Ligase	Thermo Fisher Scientific, USA
Uracil-DNA Glycosylase	Thermo Fisher Scientific, USA

4.5. Oligonucleotides

gRNAs were designed with the web-based tools CHOPCHOP, CRISPOR, e-crisp and RGENOME and then purchased from BioCat GmbH, Heidelberg. Sequence is shown 5' to 3' and cut site is indicated in bold letters.

asgRNA1	tctccctgaggtctg CT at
asgRNA2	ctttgtctcagggag CC ct
asgRNA3	gatgggtgttctgag AC ag
asgRNA4	ctgaggtctgctatgg GT gg
coGUCY12	gtgggaaggggtgggc TG gg
coGUCY17	tcacccactctacct GG ct
coGUCY26	cccatgaggggggcat AA ag
coGUCY28	actattcatcaagcac CC cg
coGUCY35	ggatctgatccgggac AG ca
coGUCY36	ccagctcctccgtgct GT cc
coGUCY38	tctgatccgggacagc AC gg
gGUCY6	gtcaccacccggct TG gc
gGUCY7	cctccgtgggtgccga TG gc
urg1	ggaccagcacactta CT gg
urg2	gtaccaccagtaagt TG ct

pegRNAs were ordered from BioCat GmbH, Heidelberg. Sequence is shown from 5' to 3' and cut site and mutations in the RTT are indicated in bold letters.

ushpegRNA1	ggaccagcacacttaCTgggtttaagagctatgctggaaacagcatagcaagttaaa taaggctagtccgttatcaacttgaaaaagtgaccgagtcggtgcttgCgaatgtaTc accagtaagtgtgc
ushpegRNA2	ggaccagcacacttaCTgggtttaagagctatgctggaaacagcatagcaagttaaa taaggctagtccgttatcaacttgaaaaagtgaccgagtcggtgcttgCgaatgtaTc accagtaagtgtgctg
ushpegRNA3	ggaccagcacacttaCTgggtttaagagctatgctggaaacagcatagcaagttaaa taaggctagtccgttatcaacttgaaaaagtgaccgagtcggtgcttgCgaatgtaTc accagtaagtgtgctgggt
ushpegRNA4	ggaccagcacacttaCTgggtttaagagctatgctggaaacagcatagcaagttaaa taaggctagtccgttatcaacttgaaaaagtgaccgagtcggtgctatgtgctgCgaa tgtaTcaccagtaagtgtgc
ushpegRNA5	ggaccagcacacttaCTgggtttaagagctatgctggaaacagcatagcaagttaaa taaggctagtccgttatcaacttgaaaaagtgaccgagtcggtgctatgtgctgCgaa tgtaTcaccagtaagtgtgctg
ushpegRNA6	ggaccagcacacttaCTgggtttaagagctatgctggaaacagcatagcaagttaaa taaggctagtccgttatcaacttgaaaaagtgaccgagtcggtgctatgtgctgCgaa tgtaTcaccagtaagtgtgctgggt

ssODNs were ordered from biomers.net GmbH, Ulm. Sequence is shown from 5' to 3' and the mutations are indicated in bold letters.

urt1.1	tctcctccctgaggtctgctatgggtgggggtgctgagcctggagctgtgattctgctattggatttc caggtggattttctgattgaaatgatgcagagaaggactatctctatgatgtgctgCgaatgtaTca ccagtaagtgtgctgggtccagctcttgtggccacttgggtt
urt1.2	atgggtgggggtgctgagcctggagctgtgattctgctattggatttccaggtggattttctgattga aaatgatgcagagaaggactatctctatgatgtgctgCgaatgtaTcaccagtaagtgtgctgggt ccagctcttgtggccacttgggttcttctcaggagccc
urt1.3	tctgattgaaatgatgcagagaaggactatctctatgatgtgctgCgaatgtaTcaccagtaagt

	tgctgggtccagctcttggtggccactgggttcctttgtcttcagggagcctgggatgggtgttct gagacagaggagctcagagggtggatgctcacggctcctggaaa
urt1.4	aaccaagtggcccacaagagetggaccagcacacttactggtgAtacattcGcagcacatcat agagatagtccttctctgcatcatttcaatcagaaaatccacctggaaaatccaatagcagaatcac agctccaggctcagcaccaccatagcagacctcagggaggaga
urt1.5	tttcaggagccgtgagcatccacctctgagctcctctgtctcagaacaacctcccagggtcc ctgaagacaaaggaaccaagtggcccacaagagetggaccagcacacttactggtgAtacat tcGcagcacatcatagatagtccttctctgcatcatttcaatcaga

Primers were designed with the web-based tools Primer-BLAST by NCBI and Primer3Plus and then purchased from biomers.net GmbH, Ulm. Sequence is shown 5' to 3'.

crp2f	TGCATATACGATACAAGGCTG
huGUCY1f	TGACAGAAAGACCCTTGGCC
huUSH2f	CCTTGCTCTGTTACCCGTTCC
huUSH2r	GTTCTGTCCCAACAATCATGC
huUSHrev	CACTGTCTTACCTGATGG
neokanF	GACAATAGCAGGCATGCTG
neokanR	GTGGATGTGGAATGTGTGC
NGqf6	GAGGCTTCACTTGTTAAGGG
NGqr4	CTGAGATCAGGGTAGACATAC
poGUCY1f	AAGCCACTTGTCACGCTAAC
poGUCY2f	CTAACTTCCCCACCTGCTGA
poGUCY1r	TGTCAGGAGGCCTAAACTGG
poGUCY1s	CTGAGCATGAGATGGAAGG
poGUCY2r	GAGGAGTTACCCGAGTGCAG

qGUCY5f	GCTCCGGGTGGGAGAAATA
qGUCY5r	CGCTGTGTTTCCCATGCTATTA
USH1Cq1f	CTTCTTGGAGCAGATGGGATAAA
USH1Cq1r	GAATGCAGATTTCTGGTTTCCAC
ush1s	GTGCCTGGCCACATCTGGA
ushHS1f	CCATGAAACTGACTAGTGG

4.6. Kits

Amaxa™ Basic Nucleofector™ Kit	Lonza, Switzerland
BioSell Double Pure Kombi Kit	Bio&Sell GmbH, Nürnberg
DNeasy® Blood & Tissue Kit	Quiagen, Hilden
Endofree Plasmid Maxi Kit	Quiagen, Hilden
Large Construct Kit	Quiagen, Hilden
Nexttec™ Isolation Kit Tissue & Cells	nexttec Biotechnologie GmbH, Hilgertshausen
Plasmid Mini, Midi and Maxi Kit	Quiagen, Hilden
TransMessenger Transfection Reagent Kit	Quiagen, Hilden

4.7. Software

BioEdit Sequence Alignment Editor 7.0.5.3	Informer Technologies Inc., USA
Endnote20	Clarivate Analytics, UK
FinchTV 1.4.0	Geospiza Inc., USA
Light Cycler 96® Software 1.1.0.1320	Roche Diagnostics Internat., Switzerland
Macromedia Freehand MX 11.0.2	Adobe Inc., USA
Microsoft Office 2016	Microsoft Corporation, USA
SnapGene Viewer 6.0	Graph Pad Software, USA
Vision Works® Acquisition and Analysis 9.0	Analytic Jena GmbH, Jena

5. Methods

Most of the used methods are described in detail in the manuscript AUCH et al. (2022) in the Methods in Molecular Biology book series (Volume 2495: Applications of Genome Modulation and Editing) attached in the appendix. Additional methods as well as detailed PCR protocols are specified below:

5.1. DNA Isolation

DNA Isolation of piglets was performed from tissue samples of tails with the Nexttec™ Isolation Kit Tissue & Cells according to manufacturer's instructions.

DNA Isolation of mixed cell populations was performed with DNeasy® Blood & Tissue Kit according to manufacturer's instructions.

DNA isolation of single cell clones was performed by high salt precipitation according to the following protocol: Resuspend the cell pellet in 100 µL PK-Buffer, add 10 µL of 10% SDS and 4.4 µL of 1M DTT, mix thoroughly and incubate at 60°C for 1 hour. Add 2 µL Proteinase K (20mg/mL), mix and incubate for 1 hour. Add 30 µL 4.5M NaCl, shake, put on ice for 10 min, then centrifuge for 20 min at 16,000 x g. Transfer the supernatant into a new tube, add 0.7 times the sample volume of iPrOH, mix thoroughly and centrifuge for 20 min at 16,000 x g. Remove the supernatant, add 500 µL of EtOH and incubate at 4°C overnight. Centrifuge for 2.5 min at 16,000 x g, remove the supernatant and air-dry for 6 min. Resolve the pellet in 45 µL of T-Buffer and incubate at 60°C for 1 hour.

5.2. Verification techniques

5.2.1. qPCR

Identification of modified *GUCY2D*^{E837D/R838S} single cell clones and estimation of FLPe efficacy:

Table 3: GUCY2D qPCR

H ₂ O	3.075 µL
SYBR green	6.250 µL
UNG	0.075 µL
qGUCY5f (5µM)	0.3 µL
qGUCY5r (5µM)	0.3 µL
DNA sample	2.5 µL

Table 4: Neomycin-resistance (neoR) qPCR

H ₂ O	3.075 µL
SYBR green	6.250 µL
UNG	0.075 µL
qGUCY5f (5µM)	0.3 µL
neoKanR (5µM)	0.3 µL
DNA sample	2.5 µL

Table 5: NANOG qPCR

H ₂ O	2.475 µL
SYBR green	6.250 µL
UNG	0.075 µL
NGqf6 (5µM)	0.6 µL
NGqr4 (5µM)	0.6 µL
DNA sample	2.5 µL

Table 6: Cyclor program for GUCY2D, neoR and NANOG qPCR

Preincubation	50°C	120 s	
Preincubation	95°C	600 s	
Two Step Amplification	95°C	10 s	45 x
	60°C	90 s	
Melting	97°C	1 s	
Cooling	37°C	30 s	

Table 7: USH1C qPCR

H ₂ O	2.275 µL
SYBR green	6.250 µL
UNG	0.075 µL
USH1Cq1f (5µM)	0.7 µL
USH1Cq1r (5µM)	0.7 µL
DNA sample	2.5 µL

Table 8: Cyclor program for USH1C qPCR

Preincubation	50°C	120 s	
Preincubation	95°C	600 s	
Two Step Amplification	95°C	10 s	45 x
	63°C	90 s	

Melting	97°C	1 s	
Cooling	37°C	30 s	

Genotyping of USH1C pigs with Deletion of exon 2 on one allele:

Table 9: Humanized USH1C qPCR

H ₂ O	2.675 µL
SYBR green	6.250 µL
UNG	0.075 µL
ushHS1f	0.5µL
huUSHrev	0.5µL
DNA sample	2.5 µL

Table 10: Cyclor program for humanized USH1C qPCR

Preincubation	50°C	120 s	
Preincubation	95°C	600 s	
Two Step Amplification	95°C	10 s	45 x
	60°C	90 s	
Melting	97°C	1 s	
Cooling	37°C	30 s	

5.2.2. End-point PCR

Verification of correctly modified huGUCYfrtNeo-BAC, genotyping of *GUCY2D*^{E837D/R838S} single cell clones and piglets and analysis of mixed cell population after electroporation:

Table 11: GUCY2D PCR 1

H ₂ O	17.7 µL
dNTP (100mM)	0.25 µL
Q-Solution	0 µL
5x Herculase II Reaction Buffer	5 µL
poGUCY2f (10µM)	0.4 µL
neokanR (10µM)	0.4 µL
Herculase II	0.25 µL
DNA sample	1 µL

Table 12: GUCY2D PCR 2

H ₂ O	16.3 μ L
dNTP (100mM)	0.25 μ L
Q-Solution	1 μ L
5x Herculase II Reaction Buffer	5 μ L
neokanF (10μM)	0.6 μ L
poGUCY1r (10μM)	0.6 μ L
Herculase II	0.25 μ L
DNA sample	1 μ L

Table 13: GUCY2D PCR 3

H ₂ O	17.7 μ L
dNTP (100mM)	0.25 μ L
Q-Solution	0 μ L
5x Herculase II Reaction Buffer	5 μ L
huGUCYf (10μM)	0.4 μ L
poGUCY1r (10μM)	0.4 μ L
Herculase II	0.25 μ L
DNA sample	1 μ L

Table 14: Cycler program for GUCY2D PCRs 1, 2 and 3

Denaturation	95°C	5 min	
Denaturation	95°C	30 s	35 x
Annealing	61°C	30 s	
Elongation	72°C	55 s	
Final elongation	72°C	10 min	
Termination	4°C	5 min	

Table 15: GUCY2D PCR 4

H ₂ O	17.2 μ L
dNTP (100mM)	0.25 μ L
Q-Solution	0.5 μ L
5x Herculase II Reaction Buffer	5 μ L
poGUCY1f (10μM)	0.4 μ L
poGUCY2r (10μM)	0.4 μ L
Herculase II	0.25 μ L
DNA sample	1 μ L

Table 16: Cyclor program for GUCY2D PCR 4

Denaturation	95°C	5 min	
Denaturation	95°C	30 s	35 x
Annealing	63°C	20 s	
Elongation	72°C	1 min 15 s	
Final elongation	72°C	10 min	
Termination	4°C	5 min	

Analysis of *USH1C*^{R31X} mixed cell population after electroporation:

Table 17: USHR31X PCR

H ₂ O	17.7 µL
dNTP (100mM)	0.25 µL
Q-Solution	0 µL
5x Herculanse II Reaction Buffer	5 µL
huUSH2f (10µM)	0.4 µL
huUSH2r (10µM)	0.4 µL
Herculanse II	0.25 µL
DNA sample	1 µL

Table 18: Cyclor program for USHR31X PCR

Denaturation	95°C	5 min	
Denaturation	95°C	30 s	35 x
Annealing	59°C	30 s	
Elongation	72°C	1 min	
Final elongation	72°C	10 min	
Termination	4°C	5 min	

5.2.3. Sequencing

Table 19: Sequencing reaction

H ₂ O	2 µL
5x Sequencing Buffer	4 µL
Primer (10µM)	1 µL
Big Dye	1 µL
DNA sample	2 µL

Table 20: Cyclor program for sequencing reaction

Denaturation	95°C	1 min	
Denaturation	95°C	5 s	40 x
Annealing	54°C	10 s	
Elongation	60°C	4 min	
Termination	4°C	5 min	

DNA was purified afterwards with EtOH precipitation and sequenced at the Sequencing Service of the LMU Biozentrum, Martinsried.

5.3. Nucleofection and generation of single cell clones (SCCs)

Nucleofection into porcine kidney cells was performed according to (VOCHOZKOVA et al., 2019). In short, seed 1×10^6 cells on a 10-cm dish and split after 24 hours. After another 24 hours, harvest the cells and perform nucleofection with program U-12 according to the instructions from the Amaxa™ Basic Nucleofector™ Kit. Use a maximum of 8 µg of DNA and seed 0.5×10^5 cells per reaction into a 6-well dish. After 24 hours, either exchange medium or harvest the cells according to growth speed and viability. For generation of SCCs with antibiotic selection, seed 50-100 cells per well on a 96-well half area plate, without antibiotic selection seed 0.5-1 cell per well. After identification of SCCs by regular scanning with the Cellavista cell culture microscope, clones are split at approximately 75% confluence on two wells of a 96-well full area plate. When reaching 90% confluence, harvest one well as cells for storage in liquid nitrogen, the other one for DNA analysis.

5.4. Lipofection

Lipofection of cells before SCNT was performed with the TransMessenger Transfection Reagent Kit according to manufacturer's instructions.

IV. RESULTS

Testing the therapeutic perspectives of GE in retinal diseases requires complementary models for reconstituting and disruptive GE approaches. Therefore, a recessive as well as a dominant model are desired, ideally with humanized regions around the mutation site in both cases to allow the testing of human specific tools. With the starting of my PhD thesis, a recessive *USH1C* model was already existing. The generation of a complementary dominant model for progressive retinal degeneration was funded by the “SPP2127: Gene and cell based therapies to counteract neuroretinal degeneration” (<https://www.spp2127.de/>), supported by the “Deutsche Forschungsgemeinschaft” (DFG) and became a significant component of my thesis.

1. Generation of the *GUCY2D*^{E837D/R838S} pig model

When deciding on the exact design of a model for a progressive retinal degeneration, multiple candidate genes were taken into consideration (Table 21). The gene and its physiological function should show a high degree of conservation between species to facilitate translational interpretation of the findings in the model. Specifically, the causative mutation and its intimate surroundings should be located in a region that is very conserved among species. Second, the course of the disease should show rapid progression with full or high penetrance and onset in early childhood, giving reason to expect significant signs of degeneration in pigs within the first months or even weeks of age. If possible, the mutation should affect retinal function alone, to avoid complications in reproduction as it is the case for *USH1C*, where piglets need intensive care after birth for their pronounced vestibular dysfunction in the neonatal age. Finally, access to patient data and material within members of the SPP2127 was seen as highly desirable. The selection procedure resulted in preference for *CORD6*, i.e. dominant negative mutations in the dimerization domain of *GUCY2D*. I favored a pig model carrying the *GUCY2D*^{E837D/R838S} mutation as the best possible option, going along with KOSTIC et al. (2013), who worked on the same mutation. The much less sophisticated tools that were available when this group generated their pig model resulted in a varying and inconsistent phenotype. Therefore, a pig model created by specifically replacing the porcine locus with the humanized mutated sequence appeared a much

more valuable asset to research of CORD6.

Table 21: Candidate genes for a complementary model

candidate gene	inheritance	gene function	disease type*	disease onset**	human	conservation***	pig	access to patients
<i>EYS</i>	recessive	PR-specific, unknown function	RP	slow (>30 years)	chr6:66Mb, intact	chr1:3.1Mb, disrupted	chr1:46Mb, intact (?)§	no
<i>RAX2</i>	dominant	eye development	CORD, MD	slow (>60 years)	chr19:3.8Mb, intact	chr10:81.1Mb, disrupted	chr2:74.9Mb, intact	no
<i>CRX</i>	dominant	PR-specific development	CORD, LCA	slow (>35 years)	chr19:47.8Mb, intact	chr7:15.6Mb, intact	chr6:53.5Mb, intact	no
<i>PITPNM3</i>	dominant	Ca ²⁺ -binding	CORD	very early (childhood)	chr17:6.5Mb, intact	chr11:72.0Mb, intact	chr12:50.9Mb, intact	no
<i>PRPF31</i>	dominant	mRNA splicing	RP	early (teens)	chr19:54.1Mb, intact	chr7:3.6Mb, intact	chr6:56.0Mb, intact	no
<i>PROM1</i>	dominant	cell differentiation	CORD, MD, SD	diverse (9-58 years)	chr4:16.0Mb, intact	chr5:44.3Mb, intact	chr8:11.2Mb, intact	no
<i>RIMS1</i>	dominant	exocytosis, Ca ²⁺ -gating	CORD	diverse (20-40 years)	chr6:72.2Mb, intact	chr1:22.5Mb, intact	chr1:52.1Mb, intact	no
<i>PRPH2</i>	dominant	cell development, growth, activation, motility	CHD, LCA, MD, RP	diverse, according phenotype	chr6:42.7Mb, intact	chr17:47.2Mb, intact	chr7:37.8Mb, intact	no
<i>UNC119</i>	dominant	PR-specific neurotransmitter release	CORD	slow (>40 years)	chr17:28.5Mb, intact	chr11:78.2Mb, intact	chr12:44.8Mb, intact	no
<i>AIPL1</i>	dominant	PR-specific chaperone and nuclear transfer	CORD, LCA, RP	early (teens)	chr17:6.4Mb, intact	chr11:71.9Mb, intact	chr12:50.9Mb, intact	no
<i>GUCA1A</i>	dominant	guanylate cyclase	COD, CORD	very early (2-17 years)	chr6:42.1Mb, intact	chr17:47.7Mb, intact	chr6:169.1Mb, intact	yes
<i>GUCY2D</i>	dominant	PR-specific guanylate cyclase	CHD, CORD, LCA	very early (< 7 years)	chr6:18.0Mb, intact	chr11:69.1Mb, intact	chr12:53.2Mb, intact	yes

* RP - retinitis pigmentosa, CORD - cone-rod dystrophy, MD - macular degeneration, LCA - Leber congenital amaurosis, SD - Stargardt disease, CHD - choroideal dystrophy, COD - cone dystrophy

** disease progression according to OMIM database and references therein (OMIM)

*** conservation determined by BLAT search on human, mouse, pig in ensembl.org and gene annotation in GenBank at ncbi.nlm.nih.gov (ENSEMBL; NCBI)

§ gene annotated, but large and complex gene structure and disruption in rodent, ungulate and primate species
green: desired property, red: exclusion criteria

1.1. From design to final vector

A manuscript on protocols for generating large targeting vectors, involving the selection of the BAC, the generation of the modification plasmid and the modification of the BAC has been published in the Methods in Molecular Biology book series (Volume 2495: Applications of Genome Modulation and Editing, AUCH et al. (2022), attached in the appendix). Therefore, in the following only brief descriptions with specific details on the *GUCY2D* approach are given.

Selection of BAC: Three BACs potentially covering the porcine *GUCY2D* locus were selected (Figure 4A) and purchased for further characterization. They were prepared for long-term storage and examined by sequencing of both ends to determine the exact location in the pig genome (Table 22). BAC end sequencing indeed confirmed exact genomic breakpoints at *Mbo*I sites as described in the generation protocol of the BAC library (BACPAC-RESOURCES). The BACs did not fully match the proposed localization in the map, with only CH242-60A12 and CH242-124B22 covering the porcine *GUCY2D* locus (Figure 4B, Table 22). CH242-60A12 was finally chosen as basis for the modification vector.

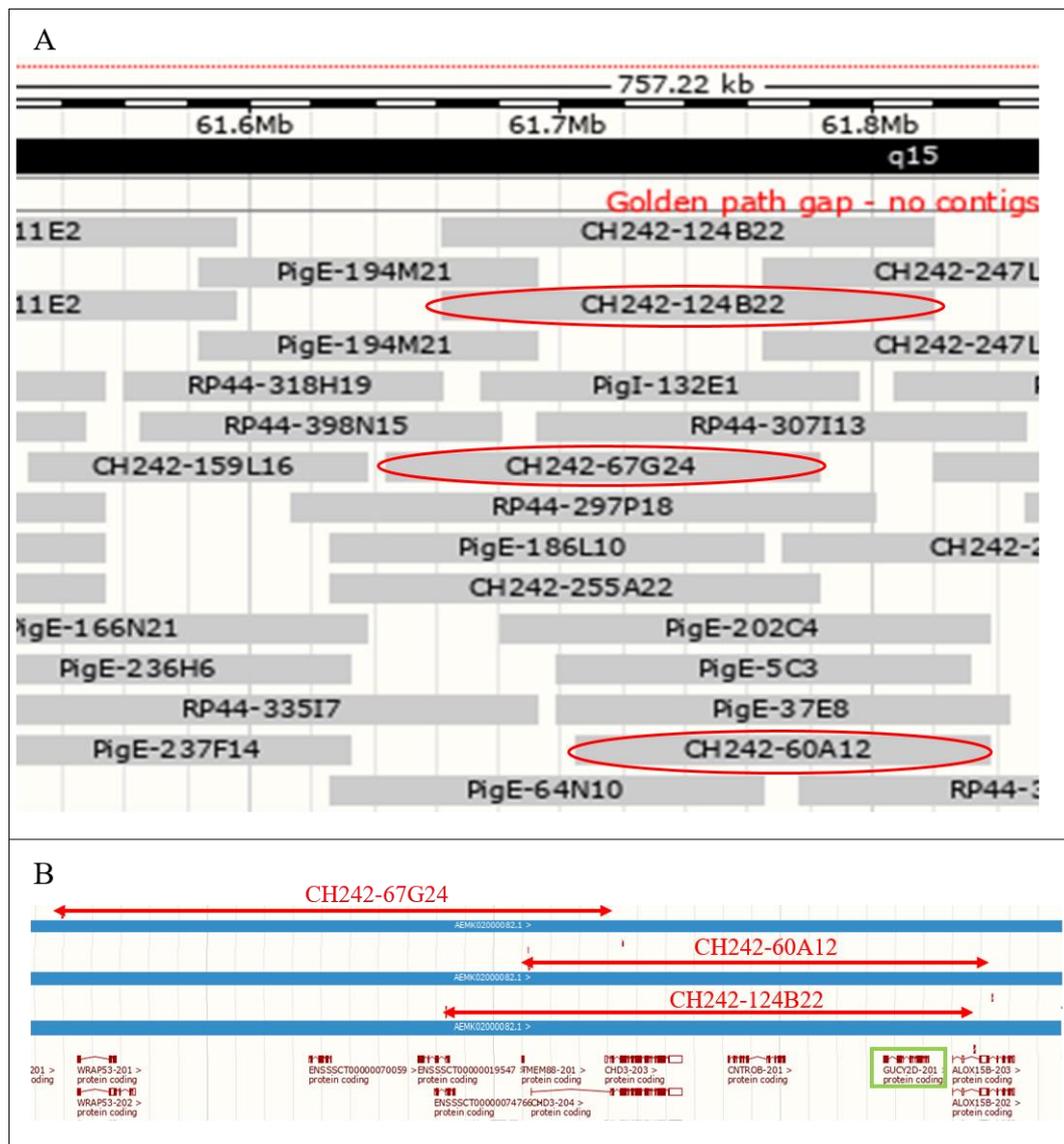


Figure 4: Location of the BACs

(A) Suggested positions of BACs from different libraries according to the PigBAC Pre map. Three of them supposedly covering the *GUCY2D* locus were ordered (red circles). (B) Precise positioning of the BACs on the porcine genome after BAC end sequencing. Red arrows show the exact location of the BACs determined by end sequencing. While CH242-60A12 and CH242-124B22 cover the *GUCY2D* locus (green box), CH242-67G24 is not placed as suggested in the BAC library and does therefore not fit for the construction of the vector.

Table 22: Location of the BACs

BAC	Location	HA left side	HA right side
CH242-67G24	12:52949118-53145181	none	none
CH242-60A12	12:53112209-53274169	135 kb	27 kb
CH242-124B22	12:53267865-53083626	164 kb	21 kb

Modification plasmid: A plasmid carrying essential components of the desired modification was commercially synthesized (Figure 5). The key part of this construct was the human *GUCY2D* exon 13 with partial introns 12 and 13 (blue), including the causative mutation (red). Further it carried homologous arms for recombineering with the BAC and restriction enzyme sites for introducing a neomycin selection cassette and for final excision of the completed construct from the plasmid backbone. Two FRT sites were placed in a way that they finally flank the neomycin cassette to enable its removal after stable integration of the modification into the genome.

```
GCGGCCGCaaaagtgtctgggagagcagccggaactcggccctccatggatcgacacctcgacctggtcaggggctgggatcgggcaagg
gctgggcccacttgcccaaaaagccagcaccctgggtggggatgggggctgcagcaaggcaggctccagattcttgacctccaggct
gcttcctaaggggtggcctgagcaccaaaggttgggtggggcagggtagcatggggcatggagacctctggaatgagagatggagctccgg
gtgggagaaatatttaaccaatttaaatGAAGTTCCTATTctagaaaGTATAGGAACTTCaggtctgaagaggagtttacgtcca
gccaaGCTAGCttggGAAGAcaatagcaggcatgctgggatcggtggctctatggcttctgaggcggaagaaccagctgggctcg
actagagcttgccaacccctcGAAGTTCCTATTctagaaaGtATAGGAACTTCttgagaaccaagtatgtctggcctgctgtga
cagaaagacctggcctgggagcccaacgattgggctggctccagtgcctgtcaattactagctgagatcaactgacctgggaacctcatt
tcccacgtgctcctaactgtgtctgaaaacacagtgcccagcaccggggctgctgatgaatagtagatgaatgtggcagcgggggtgggt
tcagagtgaacagccccatgagaggcccatgagggggcataaagaggcatggcaaccaggctctcagcagcttaccagctccttctact
gctagTTCAAGAACATCAACAAGGGCCGGAAGACGAACATCATTGACTCGATGCTTCGGATGCTGGAGCAGTAC
TCTAGTAACCTGGAGGATCTGATCCGGGcaGCACGGAGGAGCTGGAGCTGGAAAAGCAGAAGACAGACCCG
GCTGCTTACACAGATGCTGCCTCCgtgggtccagtgggaaggggtgggctgggagggcagctggagcccagccaggtAGAGtgg
ggtgacctccgggaacctcaccctgccatccccaggtctgtggcagaggctctgaagatggggaccctgtggagcctgagtatttcagga
ggtgacctgtactcagtgacattgtgggtttcaccacatctcagccatgagtgagcccatcgaagtctggacctactcaacacctctacac
gcttttcgatgccatcatcggtcccatgatgtctacaaggtggggcgtgtagggagcaagcccttctcagcgtccccctcaccagcccgacc
cctcttgccGCGGCCGC
```

Figure 5: Sequence as ordered for commercial synthesis for the modification plasmid

green: homologous arms for bacterial recombineering; blue: human *GUCY2D* fragment with capital letters exon 13; red: causative mutation sites; pink: *NheI* and *BpiI* sites for integration of neomycin; brown: FRT sites; purple: *NotI* sites for the final excision of the entire modification. The fragment was synthesized into a plasmid backbone carrying ampicillin resistance for selection.

Neomycin resistance comprising a bacterial EM7 for expression in *E.coli* and a murine PGK promoter for activity in mammalian cells was excised from the plasmid pL451 (pFNF #22687, Addgene) with *NheI* and *BpiI* and ligated into the modification plasmid that was linearized with the same restriction enzymes. After heat-shock transformation into *E.coli*, ampicillin resistant clones propagated and were checked for correct integration by digestion with *FspI* (Figure 6) and sequencing.

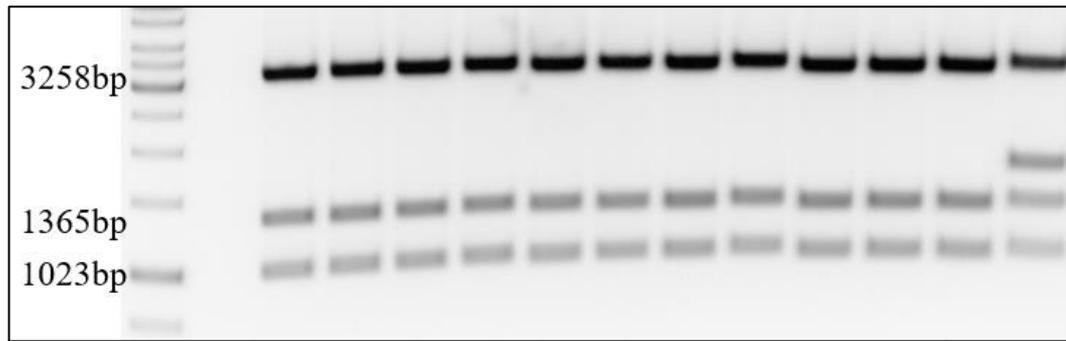


Figure 6: huGUCYfrtNeo plasmid digestion

Exemplary digestion of 11 correctly and one incorrectly modified plasmid with *FspI*

Modified BAC: The BAC CH242-60A12 was prepared from the original DH10B *E.coli* strain and transferred into the recombineering competent SW106. Integrity of the BAC clones in SW106 was confirmed by the same restriction digestion pattern as the original clone. CH242-60A12 in SW106 was then prepared for recombineering and electroporation. The verified huGUCYfrtNeo fragment was excised from the plasmid backbone with *NotI*, purified and electroporated into the recombineering competent SW106 containing CH242-60A12. Bacteria were selected on kanamycin agar plates and several clones were checked by end-point PCR for correct integration with primers spanning across the homologous arms. Integrity of the modified BAC was determined by restriction enzyme digestion which revealed the disappearance or appearance of specific bands as a result of the correct integration of the modification fragment (Figure 7).

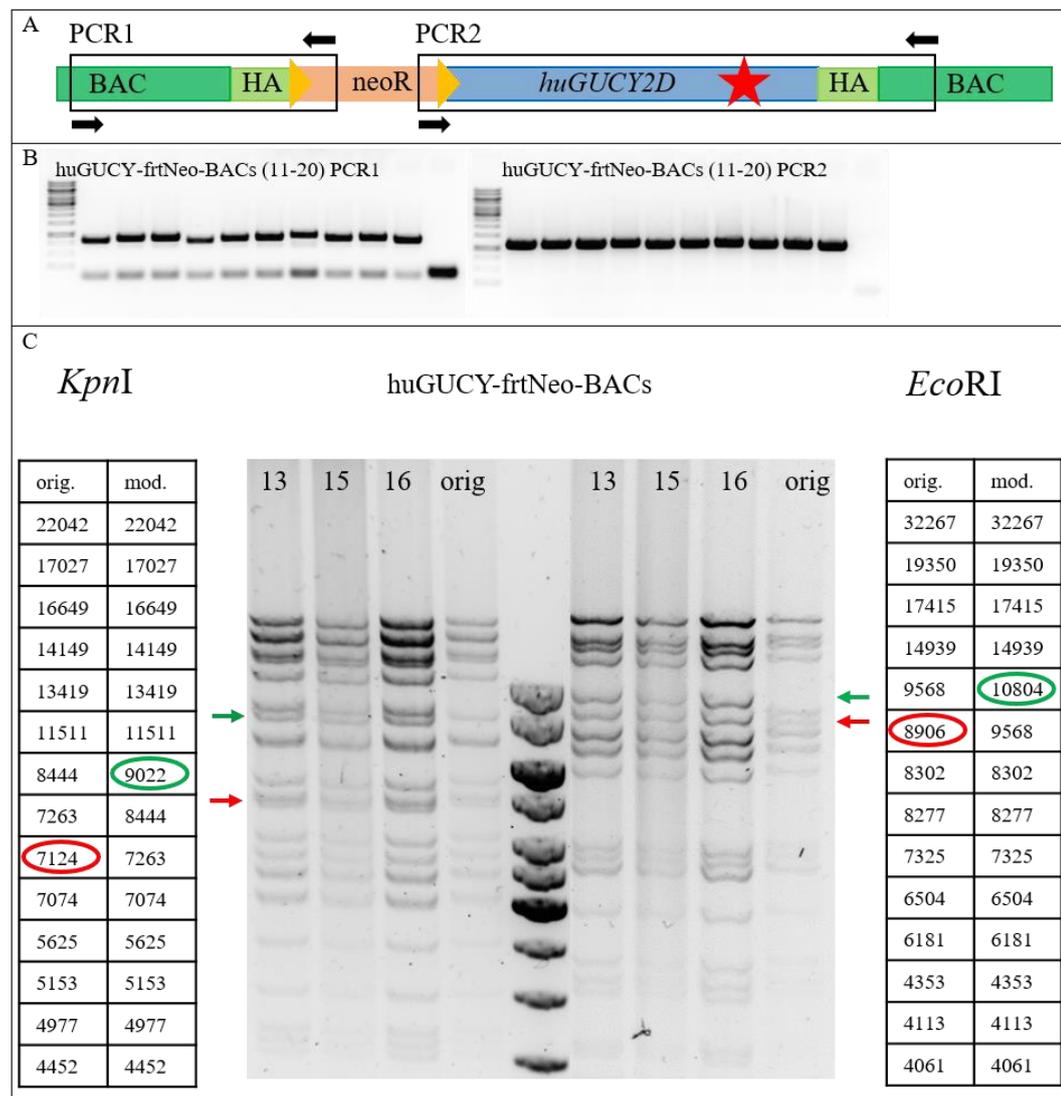


Figure 7: Verification of modified BAC

(A) Structure of *GUCY2D* exon in BAC CH242-60A12 and modification plasmid, boxes indicating PCRs across 5' end (PCR1) and 3' end (PCR2) for verification of correct recombineering. (B) Exemplary screening data of representative clones in end-point PCR indicating correct recombineering. (C) BAC clones were digested with *Kpn*I and *Eco*RI for verification of integrity of the BAC. Disappearance of the 7124bp or 8906bp (red circles and arrow) and appearance of the 9022bp or 10804bp band (green circles and arrow) respectively indicate changes in the digestion pattern as a result of correct recombineering. Band sizes are given according to *in silico* digest pattern; fragment sizes smaller than 4 kb are not indicated as they were difficult to identify for their low band intensity.

1.2. Single cell clones (SCCs)

For introducing the desired modification into the pig genome, the modified BAC vector huGUCYfrtNeo_13 was prepared at large scale and endotoxin-free status,

linearized and co-nucleofected with a plasmid expressing Cas9 and a plasmid expressing either one of the sgRNAs gGUCY6 or gGUCY7 into pig primary cells.

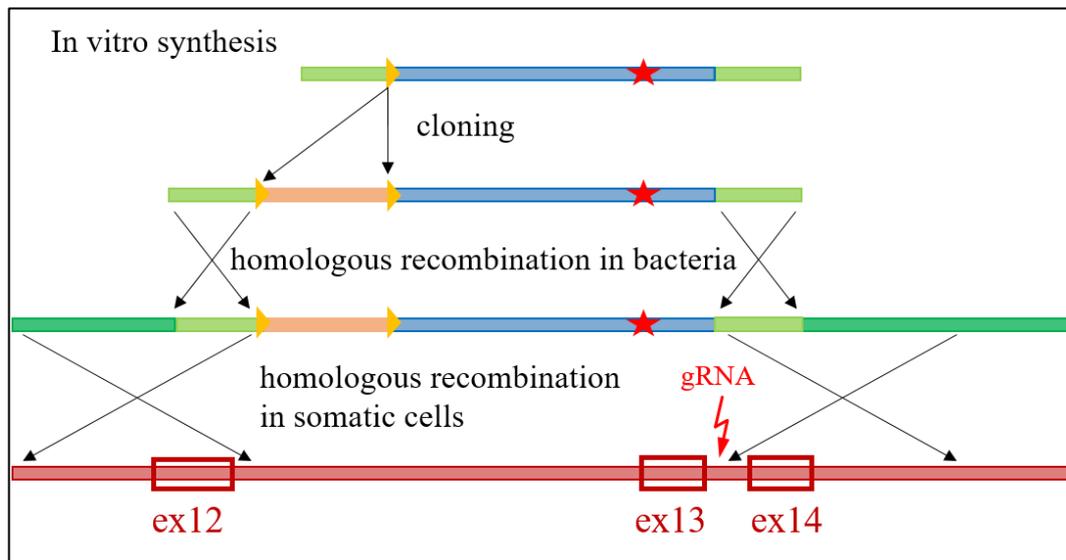


Figure 8: Strategy of generating the *GUCY2D* pig model

The general workflow from the commercially ordered sequence to the neomycin resistant modification vector and the modified BAC vector that was used to modify the porcine *GUCY2D* gene. Blue: human *GUCY2D* fragment with causative mutation (red star); orange: neomycin selection cassette; yellow: FRT sites; light green: porcine homologous arms; dark green: BAC; red: porcine genome with exons indicated by boxes

Cells were seeded with a concentration of 50 cells per well onto 20 96-well plates and kept under Geneticin selection according to VOCHOZKOVA et al. (2019). In total, 216 SCCs grew to confluence and were passaged. 202 of them propagated further and were harvested, conserved for potential SCNT and screened for the correct modification.

1.2.1. Validation of SCCs

After isolation of gDNA from the SCCs' analysis batches, correctly modified SCCs are identified in a consecutive manner: first all clones are screened by qPCR, determining the number of modified alleles, followed by end-point PCRs for identifying the constellation of porcine alleles and finally verifying the abundance of the desired modification by sequencing of the PCR products.

For the qPCR-based loss-of-wild-type-allele (LOWA) approach, the copy number

of the target site is compared to two reference loci, with the ideal ratio of reference locus to targeted locus being 2:1 (Figure 9). qPCR is a fast method for screening a lot of samples, excluding most of the SCCs from further analysis and indicating candidate clones that need further verification. Here I chose the *USH1C* and the *NANOG* loci as reference sites, for which robust qPCRs have been established before.

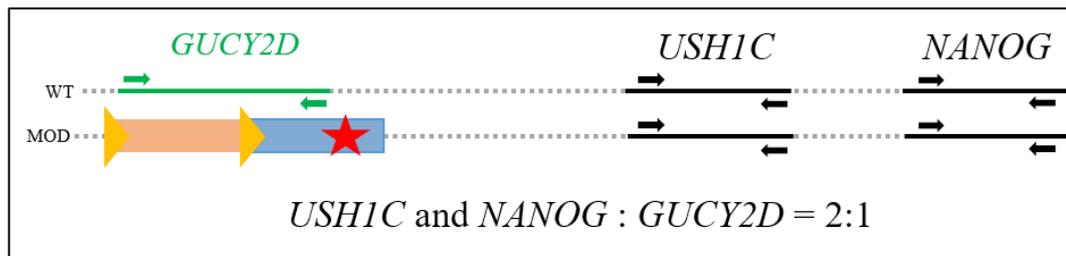


Figure 9: Reference loci

qPCR-based LOWA-approach comparing the porcine *GUCY2D* locus (green) to two reference loci, *USH1C* and *NANOG* (black). On the modified allele, FRT sites (yellow), the neomycin selection cassette (orange) and the *huGUCY2D* fragment (blue) with causative mutation (red star) prevent the porcine specific primers from binding the *GUCY2D* locus.

This way, a total number of 202 SCCs were analyzed by LOWA, with 21 of them indicating a potential modification of the target site (Figure 10).

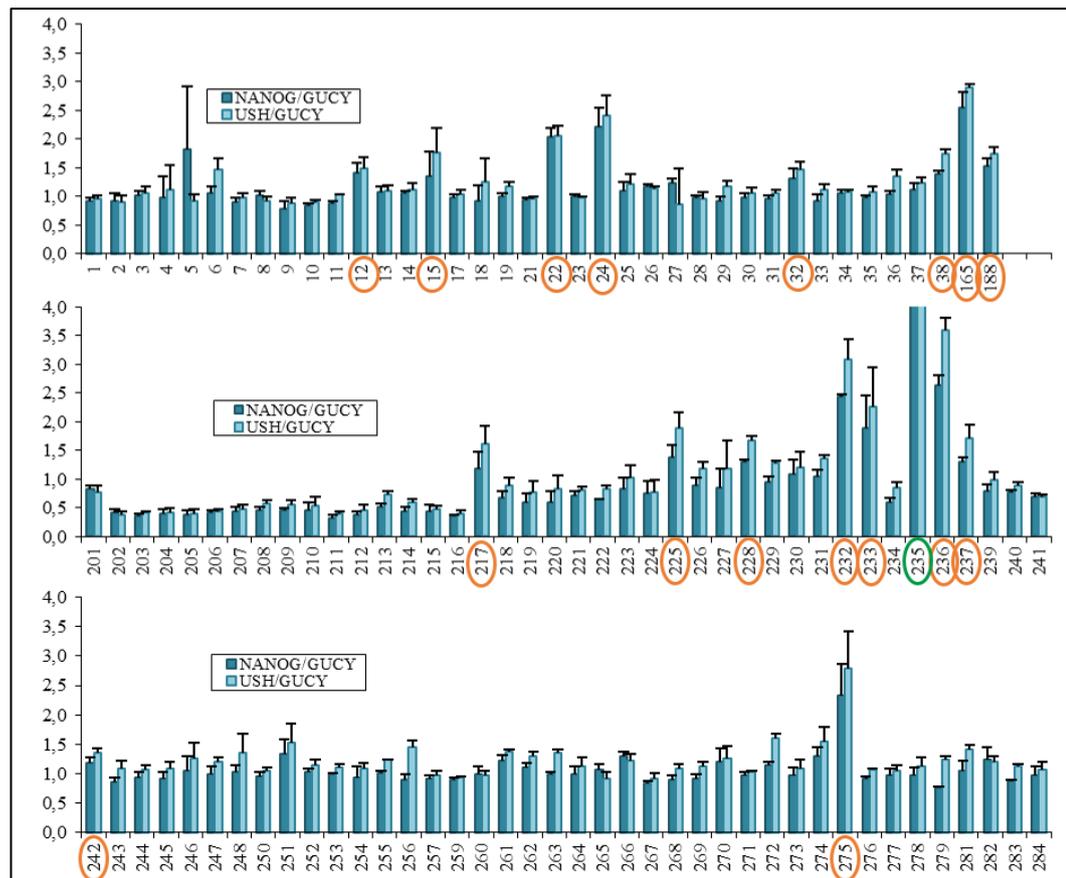


Figure 10: Analysis of SCCs with LOWA-approach

Exemplary screening data of SCCs with qPCR clearly illustrates the advantage of the LOWA method, which lies in rapidly excluding most of the samples for further analysis. A number of clones are indicated to have a mono-allelic modification (orange), while one clone is a candidate for bi-allelic modification (green).

As a next step, the abundance of the porcine allele(s) in the 21 candidate clones was verified by Sanger sequencing of a pig-specific PCR product (GUCY2D PCR 4). The presence of several single nucleotide polymorphisms (SNPs) on the porcine alleles around the target site enabled confirmation of the number of porcine alleles and, moreover, the haplotype of the remaining allele(s). Of the 21 candidate SCCs suggested by qPCR, only seven proved to have one single porcine allele remaining. Three of them showed the GGACC pattern on the remaining porcine allele (Figure 11) and four of them the alternative CAGAA pattern, suggesting that there was no clear preference for recombination with the modified BAC vector. At the cutting site of the gRNA in intron 13, three of these seven clones did not show any NHEJ-induced modification, while the other showed either an insertion of one nucleotide or a deletion of 25 nucleotides. Being located in the intronic region, no influence on the expression of the allele is assumed, nevertheless I prefer to choose a non-

affected clone for the generation of the animal model. The seven heterozygous and the one supposedly homozygous clone were further examined for the existence of the intended humanization.

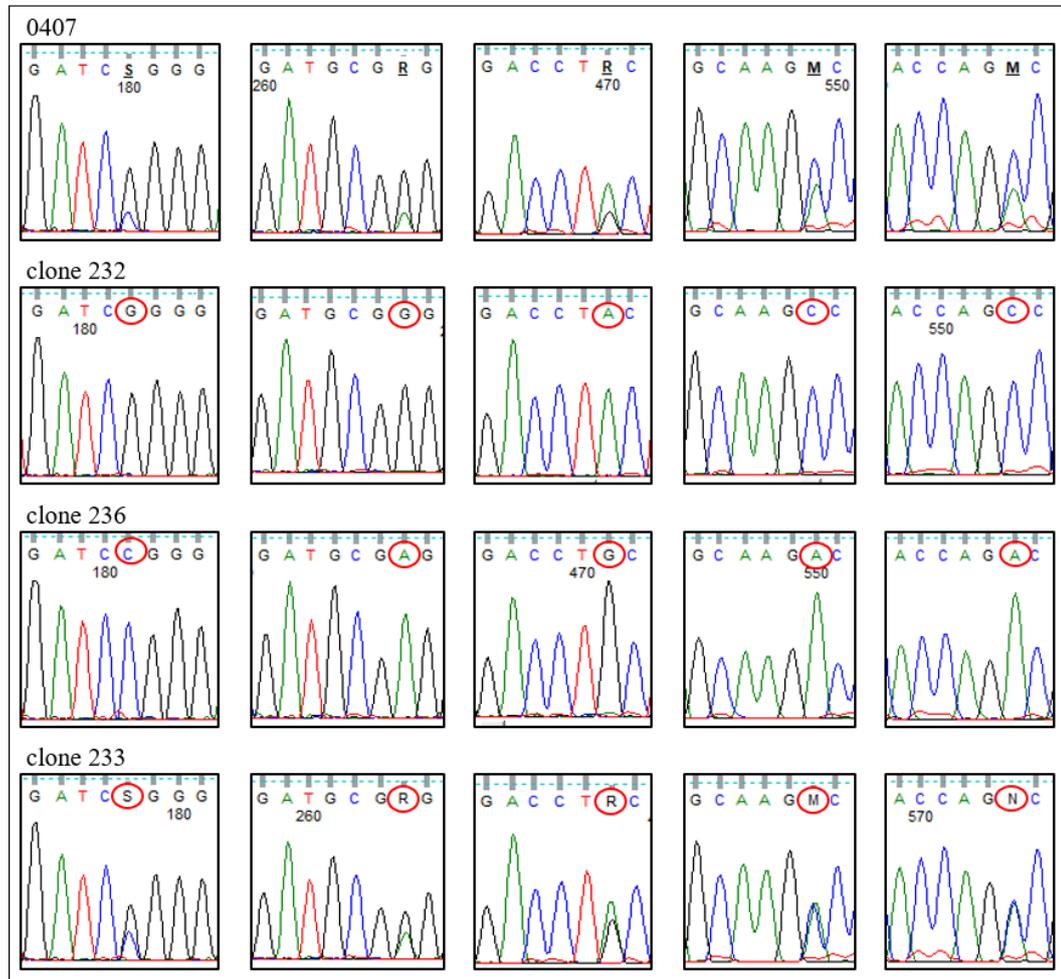


Figure 11: Allelic constellation of *poGUCY2D* in candidate SCCs

Exemplary constellations showing clone 232 to have a “GGACC” constellation on the remaining porcine allele, whereas clone 236 has a “CAGAA” constellation. Clone 233 shows two porcine *GUCY2D* alleles, indicating classification as false positive in qPCR.

As a final step, the same end-point PCRs that were used for the analysis of the modified BAC vector (Figure 7A) were used to verify the abundance of the desired modification in the genome of the SCC. Sequencing of both PCR products indicated incorporation of the entire construct in six SCCs that were used for further preparation for somatic cell nuclear transfer (SCNT).

clone	qPCR ratios		single poGUCY2D allele	Verification of humanization
	NANOG/GUCY2D	USHIC/GUCY2D		
12	1,41	1,49		
15	1,35	1,77		
22	2,03	2,05		
24	2,21	2,40		
32	1,32	1,47		
38	1,39	1,74		
165	2,55	2,90		
188	1,53	1,74		
217	1,18	1,62		
225	1,39	1,89		
228	1,31	1,68		
232	2,45	3,09		
233	1,88	2,26		
235	4,16	5,58		
236	2,64	3,59		
237	1,31	1,72		
242	1,19	1,36		
270	1,19	1,26		
272	1,15	1,61		
274	1,29	1,55		
275	2,34	2,79		

Out of the 21 candidate clones indicated by qPCR (light green), eight proved to have only one single porcine *GUCY2D* allele remaining in Sanger sequencing of a pig-specific PCR product (*GUCY2D* PCR 4). Two of them were excluded due to negative PCRs targeting the humanized region (*GUCY2D* PCR 1 and 2).

Figure 13: Summary of sequential SCC characterization

1.3. Somatic Cell Nuclear Transfer (SCNT)

The six correctly modified SCCs were thawed on wells of a 96 well plate, propagated and split twice to produce aliquots of cells for repeated SCNT. One well of each clone was used for nucleofection with a plasmid expressing flippase (FLPe) (#13787, pCAG-Flpe, Addgene) to remove the neomycin selection cassette. Alternatively, SCC batches were treated with *FLPe*-mRNA via lipofection with the TransMessenger Transfection Reagent in 96-well format. Based on previous experiments we estimated to have approximately 30% of founder animals lacking the selection cassette (GROTZ et al. (2022), RUNA-VOCHOZKOVA et al., unpublished data).

For producing *GUCY2D*^{E837D/R838S} founder pigs (called *GUCY2D* pigs in the

following), four SCNT experiments were performed at the Institute of Molecular Animal Breeding and Biotechnology (head: Prof. Eckhard Wolf, cloning team: Tuna Güngör, Arne Hinrichs, Barbara Keßler, Mayuko Kurome and Valeri Zakhartchenko). One pregnancy was established and resulted in a litter of three live and one stillborn piglet on November 2nd, 2021. Although properly viable at birth, the founder animals were bitten by the mother sow or died of an infection within 14 days after birth. While disappointing, such a scenario is not uncommon for SCNT offspring, but still the birth of the founder animals left the opportunity for sampling tissue and primary cells for further investigations.

1.4. Validation of the genotype in founder piglets

Despite repeated efforts I was not able to gain DNA or primary cells from the stillborn animal. The genotype of the other three piglets was confirmed by the end-point PCRs used for BAC and SCC characterization (Figure 7A). The humanized fragment proved to be consistently abundant in all three tested animals. Sequencing of the porcine allele confirmed the CAGAA haplotype pattern and the lack of any NHEJ-mediated modification at the gRNA cutting site, showing that all three animals derive from clone 236 (Figure 14).

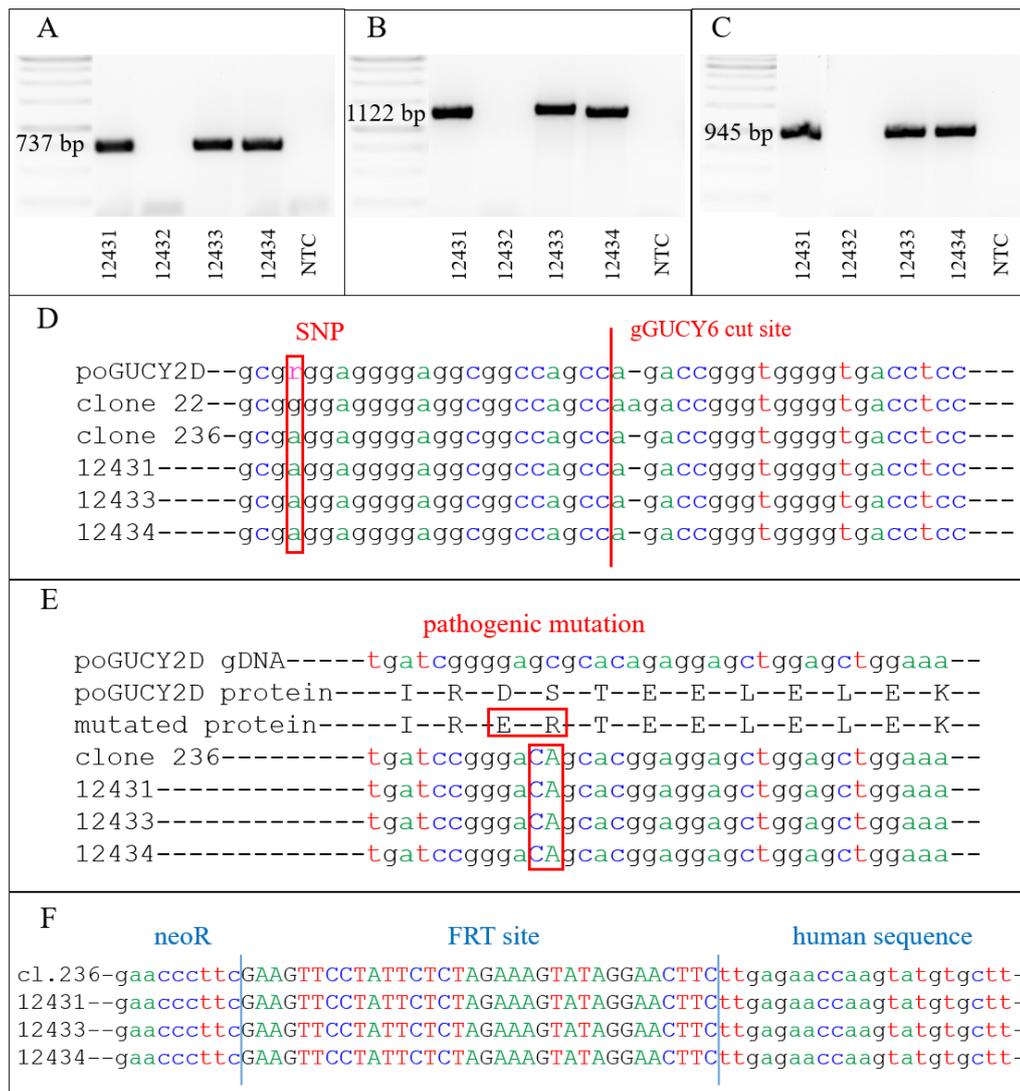


Figure 14: Validation of the genotype in founder piglets

(A) GUCY2D PCR1 between the 5' end of the neomycin selection cassette and the upstream porcine sequence; (B) GUCY2D PCR2 between the 3' end of the neomycin selection cassette and the downstream porcine sequence, spanning the humanized part; (C) GUCY2D PCR3 between the humanized fragment and the downstream porcine sequence. This PCR is necessary for the verification of the humanization in case the neomycin selection cassette was removed before SCNT; (D) sequencing of a pig-specific PCR product (GUCY2D PCR4, see Figure 11) shows all animals derived from clone 236, the sequences being identical considering the aforementioned SNPs and no unwanted modification at the cut site. (E) and (F) show the correct sequence of the pathogenic mutation and the FRT site at the border of the neomycin selection cassette and the humanized sequence.

Taking into account the abundance of the neomycin selection cassette in all founder animals, additional SCNT experiments for generating GUCY2D pigs lacking the neomycin cassette was desired anyway. For this, primary cells were isolated from

the kidneys and nucleofected with the plasmid pCAG-Flpe. Analysis with qPCR indicated that 80-90% of the cells do not retain the neomycin cassette, making these cells a perfect source for a second round of SCNT experiments to finally deliver neo-free GUCY2D pigs (Figure 15).

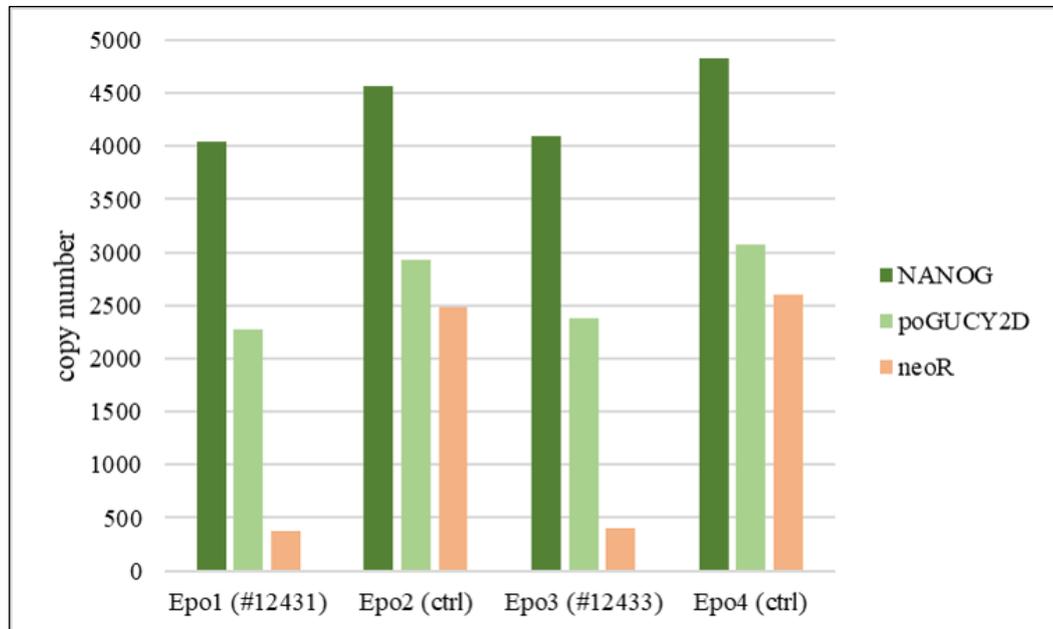


Figure 15: qPCR results of FLPe-treated pig primary cells

Both available cell lines from animal 12431 and 12433 were nucleofected and a non-treated aliquot of the respective cell line served as control. According to viability and growth speed, cells were harvested 24-48 h after nucleofection. Each cell population was partly frozen for long term storage, propagated further on a 48-well plate or DNA was extracted for analysis. Removal of the neomycin selection cassette appeared to be working at a similar level of efficacy in both cell lines, indicated by a significant drop in copy numbers compared to two reference loci. gDNA from animal 12431 was used for the standard curve.

2. Therapy approach in the *GUCY2D*^{E837D/R838S} pig model

The primary cells isolated from the founder animals have also provided a perfect source for testing therapeutic Gene Editing (GE) efficacy. The dominant negative fashion of the E837D/R838S mutation is an ideal test case for disruptive GE. The goal is to destroy the mutated allele without influencing the WT allele, thus leaving only one healthy copy of the gene. The specific abundance of a recessive *GUCY2D* variant without physiological hallmarks proves that such a disruptive approach on a dominant negative variant of the gene is feasible, since the presence of one unmodified allele is enough to keep up physiological function in heterozygous carriers of a knock-out mutation.

2.1. Design of disruptive CRISPR/Cas9 therapy

I decided on two different approaches: one was to design one sgRNA placed directly at the site of the causative mutation that will result in a non-functional transcript as a consequence of erroneous NHEJ within exon 13. The other was to place two sgRNAs inside the intronic regions up and downstream of exon 13 and to cut out the whole mutated exon by direct joining of free genomic arms. This strategy is based on the presence of discriminating SNPs between the two alleles, therefore specific sgRNAs need to be designed for each patient. In this case, the humanized and the porcine allele differ substantially in large intronic sequences, making the design of the sgRNAs relatively flexible. Although the approach is artificial to a certain extent, it still tests the hypothesis of inactivating the mutated allele not by the causative mutation alone, but also on the basis of correlating SNPs.

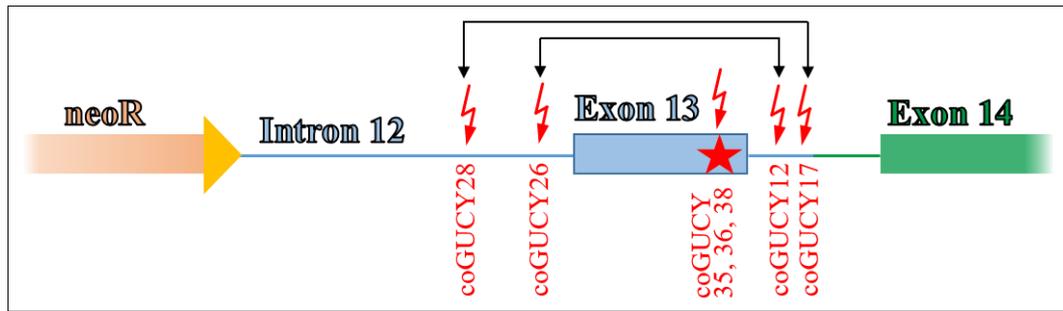


Figure 16: *GUCY2D* disruptive gene editing strategy

The three sgRNAs coGUCY35, coGUCY36 and coGUCY38 are targeted directly at exon 13 (blue), where the only discrimination between the two alleles is the causative mutation itself (red star). The other sgRNAs are used in pairs and due to the differences in the intronic sequences they are specific only for the humanized allele.

2.2. Efficacy of sgRNA constellations

Experiments were performed on the cell line 12431 (*GUCY2D*^{E837D/R838S}). A plasmid expressing Cas9 and one or two plasmids expressing the sgRNAs were co-electroporated into the cells in distinct constellations (Table 23) and cells were harvested as mixed cell populations after 24 hours. Out of the mixed cell population, gDNA was isolated and used for amplification of *GUCY2D* PCR2 (Figure 7A).

Table 23: Composition of *GUCY2D* disruptive GE experiments

<u>experiment</u>	<u>components</u>
Epo1	Cas9 + coGUCY12 + coGUCY26
Epo2	Cas9 + coGUCY17 + coGUCY28
Epo3	Cas9 + coGUCY35
Epo4	Cas9 + coGUCY36
Epo5	Cas9 + coGUCY38
Epo6	Cas9 + H ₂ O

Even from separation of the PCR products on agarose gel, it became obvious that in Epo 1 and 2, comprising the double-gRNAs for excision of exon 13, a substantial number of cells underwent the desired deletion of exon 13.

The Epos 3-5, comprising the single gRNAs for inactivating the reading frame within exon13, showed a single band of the expected size of 1122bp (Figure 17).

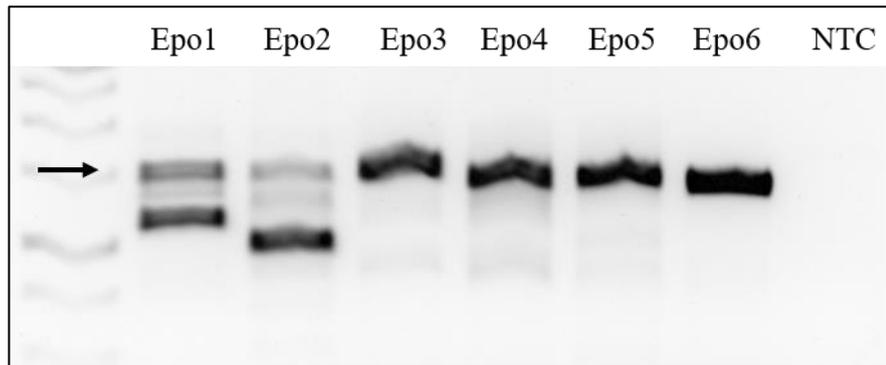


Figure 17: GUCY2D disruptive GE

A band of 1122bp (indicated by the arrow) was expected for GUCY PCR2. For Epos 1 and 2, only a minority of PCR products constitute this size. For Epos 3-5 only small indels are expected, resulting in a band size similar to the control experiment Epo 6.

Amplicons of the mixed cell population were then sequenced and the electropherograms were analyzed with the web-based tool ICE by Synthego (CONANT et al., 2022) and by measuring peak heights with the drawing software FreeHand MXa. For Epos 3-5, cutting efficacy was estimated to be between 23-67% as calculated with the Synthego tool, while FreeHandMXa suggested a mutation rate between 50-70%.



Figure 18: Analysis of Epo 3-5 with Synthego

The Synthego tool uses the electropherograms from Sanger sequencing to calculate indel contributions.

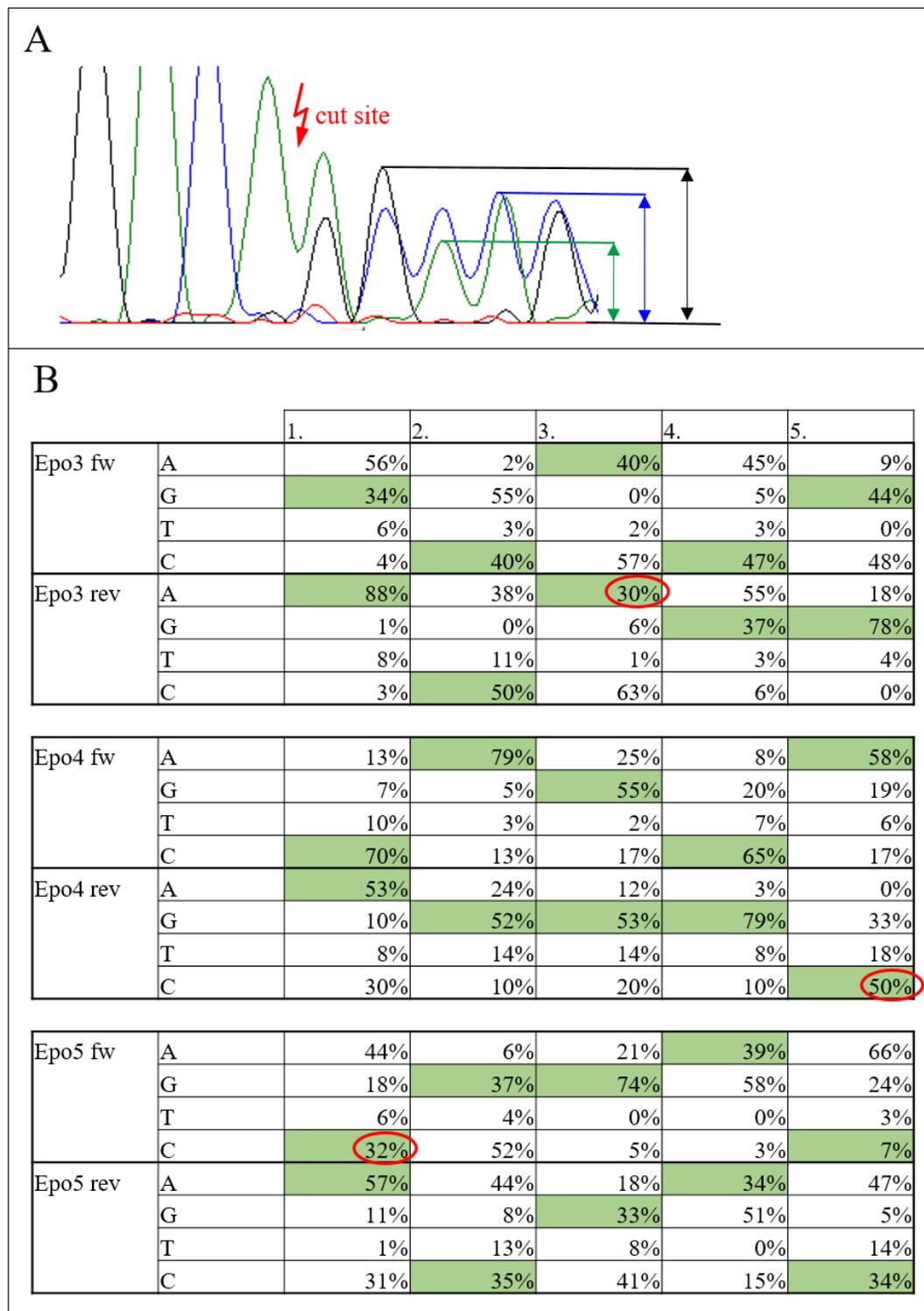


Figure 19: Analysis of Epo 3-5 with FreeHandMXa

(A) shows a representative electropherogram of Epo3. For the alternative analysis with FreeHandMXa, the five nucleotides after the cut site were measured. This was done with electropherograms from sequencing with the forward and the reverse primer for each experiment and results are summarized in (B). The percentage of the original nucleotide is marked in green. The lowest percentage was taken as the true value of the unedited sequence (red circle).

In contrast to Epos 3-5 that constitute modifications within exon 13 and make any mutation an inactivating event, Epos 1 and 2 are more difficult to analyze, as both sgRNAs in each approach cut in an intronic sequence. The band intensity was analyzed on the Analytik Jena US UVP GelStudio Plus with the Vision Works® Acquisition and Analysis Software that is pre-installed. It is calculated that for both Epo 1 and Epo 2, the shorter band makes up for more than 40% of the DNA in the respective lane, which is emphasized by the peak heights graphically (Figure 20).

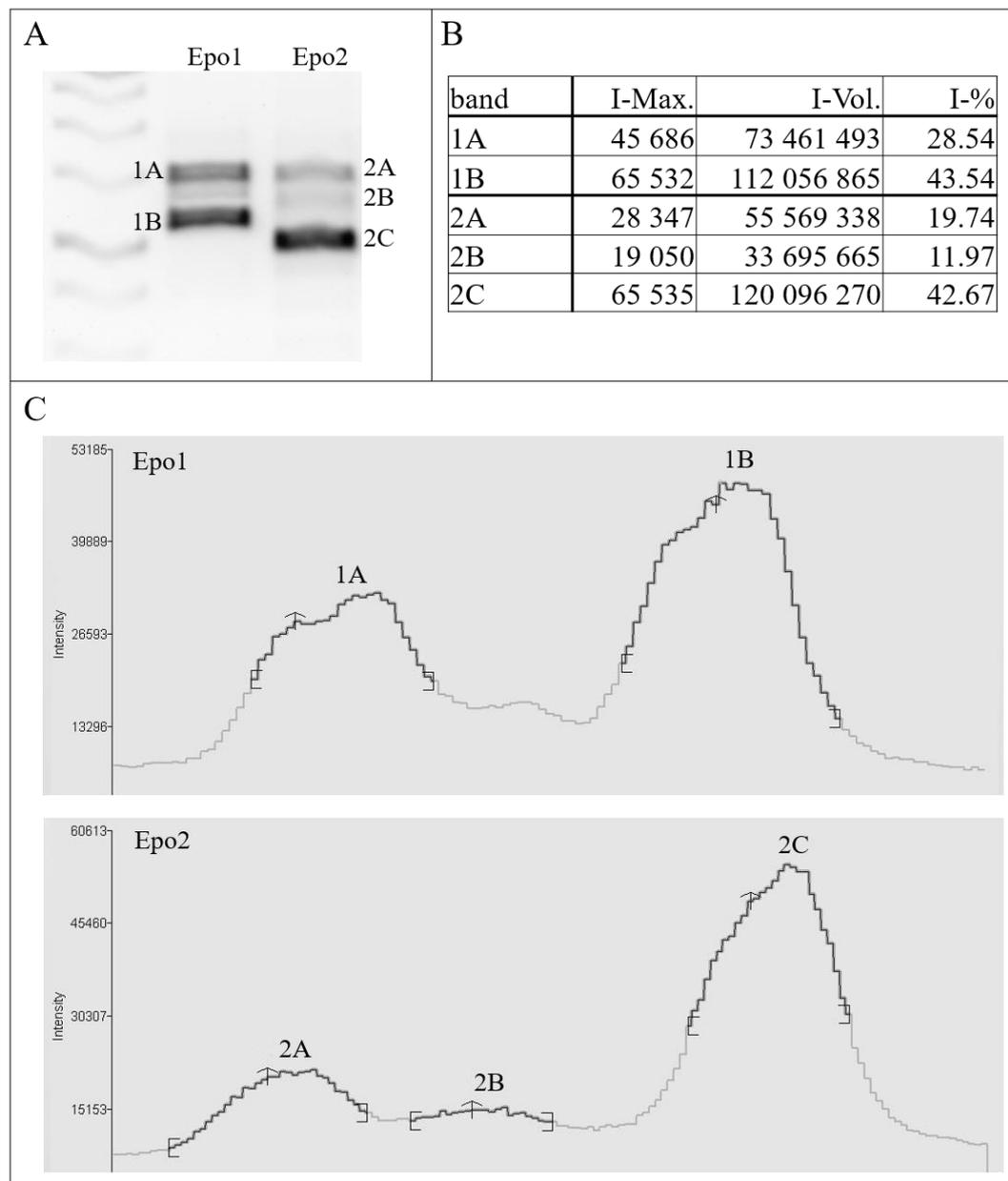


Figure 20: Analysis of band intensity for Epo 1 and Epo 2

(A) The major bands in each lane were identified and intensity was measured. On first estimation, most of the product seem to lack the entire or at least a partial region between the cutting sites (248bp between the cutting sites of coGUCY12 and coGUCY26, and 358bp between the cutting sites of coGUCY17 and coGUCY28). (B) I-Max (Intensity maximum): maximum pixel intensity in the band; I-Vol (Intensity volume): sum of all pixel intensities in the lane or band; I-% (Intensity percentage): $(\text{Band I-Vol} / \text{Lane I-Vol}) * 100$

(C) Graphic display of the band intensities

Amplicons of the mixed cell population were sequenced and analyzed with Synthego as well. Theoretically, small indels might appear at both sites, while the exon 13 is retained and remains functional. Only events that completely excise exon 13, leaving a gap between the cutting sites reliably inactivate the dominant

Overall, the efficacy in all five experiments was outstandingly high, indicating mutation rates between 23 and 77% (Table 24). Sequencing of the porcine WT allele in the mixed cell population confirmed the strict specificity of all sgRNAs for the humanized allele.

Table 24: Efficiency of *GUCY2D* knock-out experiments

	<u>Synthego</u>	<u>FreeHandMXa</u>	<u>Band intensity</u>
Epo1	66%	n.d.	44%
Epo2	77%	n.d.	43%
Epo3	52%	70%	n.d.
Epo4	23%	50%	n.d.
Epo5	67%	68%	n.d.

n.d.: not determined

The high editing rates for both the single site and excision approach illustrated the potential of disruptive gene editing. A translation *in vivo*, according (MORETTI et al., 2020) on treating a pig model of Duchenne muscular dystrophy, is the obvious next step once *GUCY2D* pigs have been reconstituted by a second round of SCNT. Both described approaches, however, are limited to dominant mutations. For this reason, I also aimed at testing reconstituting GE, which would fit both dominant and recessive mutations.

3. Therapy approach in the USH1C pig model

Exploratory experiments on therapeutic GE in *USH1C* have been performed in a similar setting as I have done for *GUCY2D*, using primary cells of an existing pig model for a recessive *USH1C* loss of function mutation and have been published in GROTZ et al. (2022). In addition, I was responsible for propagating the USH1C pig breeding herd and the supply of experimental animals for several pre-clinical studies conducted within national and international collaborations.

3.1. Production and use of the USH1C pig model in research

After the initial production of USH1C founder pigs via embryo transfer, breeding started with the first litter of heterozygous animals born in 2020. From breeding the heterozygous animals with each other and also with the homozygous founder animals, first homozygous piglets were born in 2021. Since the founder animals carry the R31X mutation on one allele, and a large deletion on the other, this will be passed on according to Mendelian rules to the offspring. Due to the different allelic constellations in the breeding herd, it became necessary to perform adapted genotyping with qPCR for the F2 generation in case end-point PCRs were inconclusive. To discriminate WT/WT animals from WT/Del and R31X/R31X from R31X/Del, copy numbers of either the unmodified or the humanized *USH1C* locus respectively were compared to the *NANOG* locus, reliably identifying the genotype.

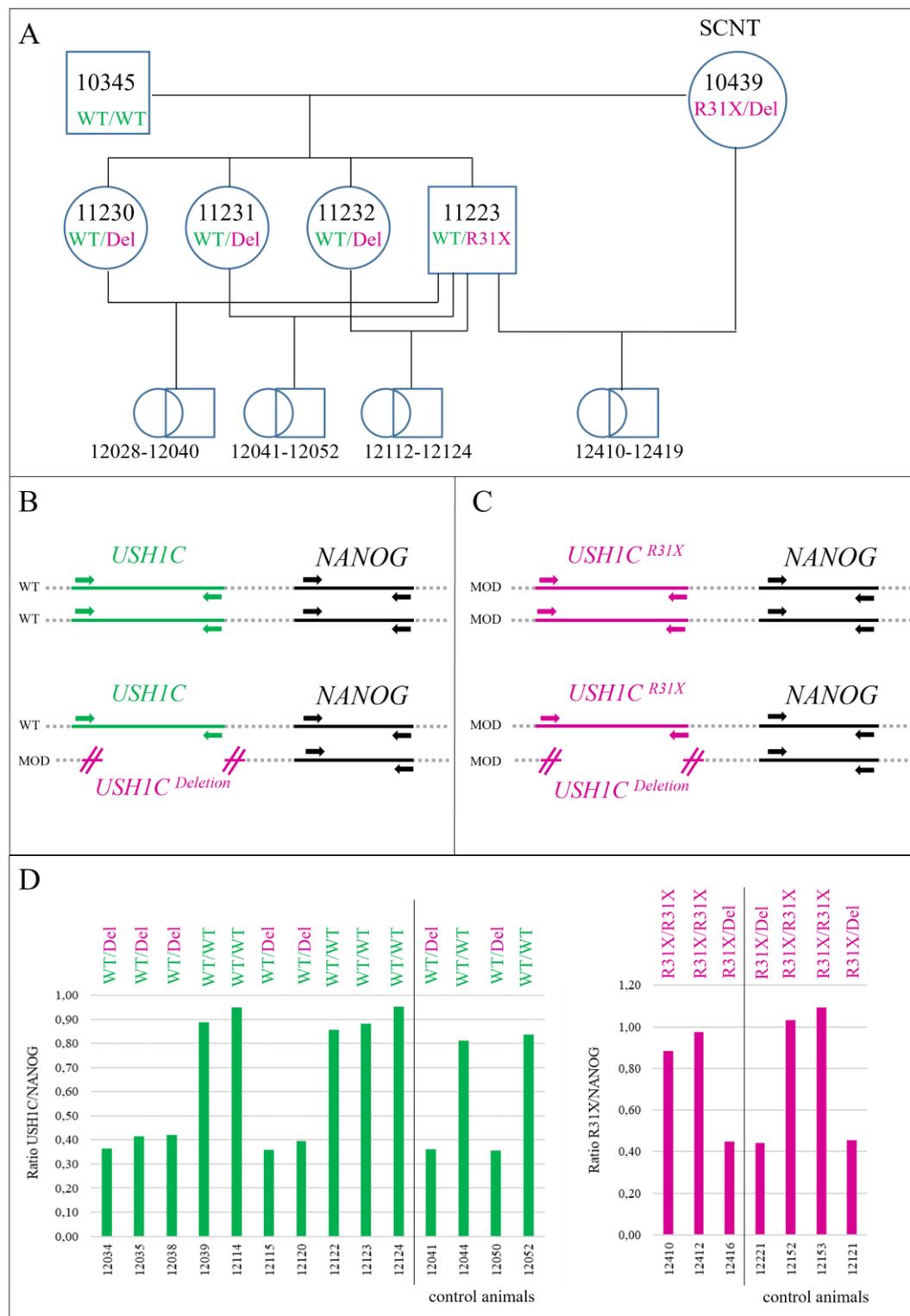


Figure 22: Genotyping of F2 generation

Adapted from Grotz et al., 2022 Appendix Fig. S2

(A) Exemplary pedigree showing the founder animal 10439 bred with a wild-type boar to produce heterozygous F1 offspring. Female animals are indicated with circles, male animals with boxes. These animals inherit either the R31X mutation or the deletion from their mother. When mated with each other, piglets of the F2 generation without apparent phenotype can be WT/WT or WT/Del (B). When mated with their homozygous ancestor, piglets of the F2 generation showing apparent

phenotype can be either R31X/R31X or R31X/Del (C). Exemplary qPCR results for the unmodified *USH1C* locus (green) or the humanized version (magenta) clearly identify the allelic constellation (D). gDNA of an animal with R31X/R31X genotype that has been confirmed by end-point PCR was used for the standard curve for the humanized *USH1C* qPCR. Reliability of both qPCR protocols is verified by adding control animals, whose genotype has been confirmed likewise.

Since then, a total of 11 litters has been produced, providing animals for research in different projects (Table 25).

Table 25: Use of animals in research

<u>animals</u>	<u>purpose</u>	<u>procedures</u>	<u>collaboration</u>
9 USH1C + 5 WT	Gene therapy study (Munich)	OCT/ERG, AAV injection, tissue sampling	Prof. Dr. Dr. M. Dominik Fischer (Oxford), Dr. Tobias Peters (Tübingen), Prof. Dr. Uwe Wolfrum (Mainz)
9 USH1C + 4 WT	Natural history study (Munich)	OCT/ERG, tissue sampling	Dr. Tobias Peters (Tübingen), Prof. Dr. Uwe Wolfrum (Mainz), Prof. Dr. Andreas Parzefall (Munich)
4 USH1C	PR transplantation (Dresden)	OCT, transplantation, tissue sampling	Prof. Dr. Marius Ader (Dresden)
4 USH1C + 6 WT	Inner ear study (Munich)	ABR, ASSR, OAE, perfusion fixation, tissue sampling, μ -CT of inner ear	Prof. Dr. Andrea Fischer (Munich), Prof. Dr. Andrea Meyer-Lindenberg (Munich), Prof. Dr. Andreas Parzefall (Munich), Dr. Aziz El Amraoui (Paris)
10 USH1C + 14 WT*	Gene therapy study (Libechov, Czech Rep.)	OCT/ERG, AAV injection, tissue sampling	Prof. Dr. Jan Motlik (Libechov), Prof. Dr. Dr. M. Dominik Fischer (Oxford), Prof. Dr. Uwe Wolfrum (Mainz)

* animals produced in a sister herd at PIGMOD Centre, Libechov

For the gene therapy study, recombinant AAV vectors with *USH1C* mRNA were injected into the subretinal space and the pigs are being examined over the course of several months with OCT/ERG. The goal is to induce the expression of harmonin and therefore slow down the disease progression. This study is done in collaboration with Prof. Dr. M. Dominik Fischer (Oxford Eye Hospital, Oxford University, Oxford, UK) and Dr. Tobias Peters (University Eye Hospital, Tübingen) who have excellent expertise in ophthalmic research as well as Prof. Dr. Uwe Wolfrum (Institute of Molecular Physiology, Johannes-Gutenberg University, Mainz) providing the knowledge for morphological and molecular analysis. Additionally, *USH1C* pigs are evaluated with OCT/ERG and sampled at defined time points in life to get a longitudinal overview over the natural progression of the disease. A better understanding of the correlation between phenotype and morphological and molecular changes is necessary to define suitable end-points for future pre-clinical studies.

Another collaboration is done with the group of Prof. Dr. Marius Ader (Center for Regenerative Therapies, Technical University, Dresden) who transplanted photoreceptor cells derived from human retinal organoids into the subretinal space of wild-type and *USH1C* pigs. They investigate the structural and functional integration of donor cells into degenerated and non-degenerated retinæ, as well as inflammatory reactions to the transplantation.

Besides the visual impairment, another interesting aspect of USH is the congenital deafness. Therefore Prof. Dr. Andrea Fischer (Clinic of Small Animal Medicine, Ludwig-Maximilian University, Munich) investigates the inner ear phenotype, using methods such as Auditory Brainstem Response (ABR), Auditory Steady State Response (ASSR) and Oto Acoustic Emission (OAE) that are regularly used in human patients. Expertise in Pathology comes from Prof. Dr. Andreas Parzefall (Institute of Veterinary Pathology, Ludwig-Maximilian University, Munich), who takes a closer look on the inner ear's ultrastructure together with Prof. Dr. Andrea Meyer-Lindenberg (Clinic for Small Animal Surgery and Reproduction, Ludwig-Maximilian University, Munich) in micro computer tomography (μ -CT). A close collaboration is maintained with Prof. Dr. Jan Motlik (Institute of Animal Physiology and Genetics, PIGMOD Centre, Czech Academy of Science, Libechev, Czech Republic), who funded his *USH1C* breeding herd on three founder sows generated by SCNT in Munich.

3.2. Design of reconstituting CRISPR/Cas9 therapy

Loss of function of the *USH1C* gene becomes clinically relevant when appearing in a homozygous manner. For a therapeutic effect, the causative c.C91T/p.R31X nonsense mutation needs to be repaired on at least one allele, but not necessarily on both. Two sgRNAs were designed and tested for their ability to introduce NHEJ mutations, indicating similar efficacy for both (20-30%). Thus, the gRNA located closer to the c.C91T site was combined with five different ssODNs. These were designed according to the guidelines described in RICHARDSON et al. (2016), suggesting higher efficacy for a ssODN that is asymmetric to the target site, as well as complementary to the sgRNA non-target strand. As recommended in PAQUET et al. (2016), I always implemented a silent blocking mutation located at the PAM to prevent re-cutting after a successful edit.

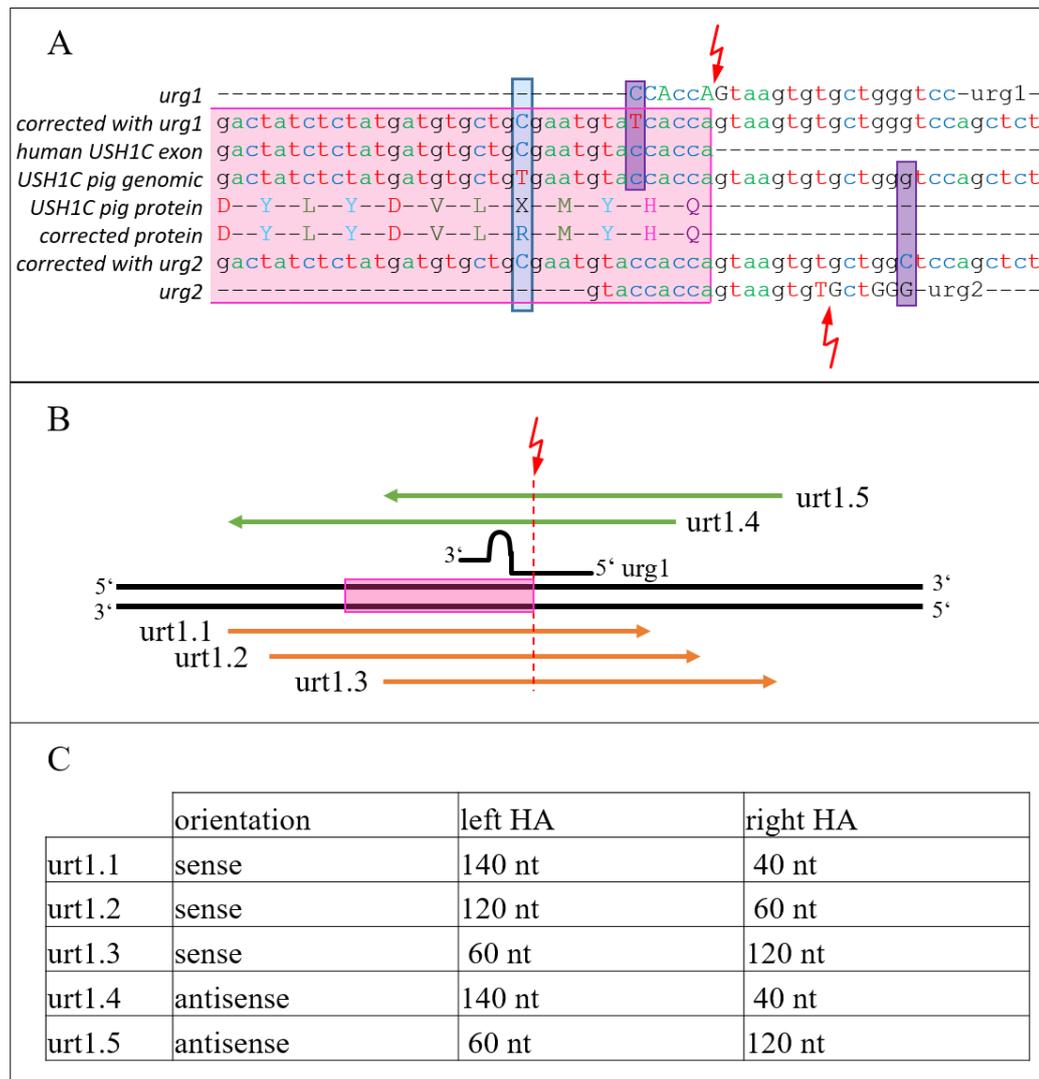


Figure 23: *USH1C* reconstituting GE strategy

As in GROTZ et al. (2022), Figure EV4 and Figure 7

(A) 2 oppositely oriented sgRNAs were tested for their efficacy to introduce NHEJ-mediated indel formation. Cut sites are located directly at the end of exon 2 (pink box) or a few nucleotides after the exon-intron border. The distinct positions at which the respective ssODNs for *urg1* or *urg2* should introduce blocking mutations are indicated by magenta boxes. The causative mutation is indicated by a blue box. (B) After deciding on *urg1*, fitting ssODNs of varying length and position binding to the target (green) or non-target strand (orange) were compared. (C) Exact positions are given in the table.

3.3. Efficacy of ssODNs

Experiments were performed on the cell line 5615 (*USH1C*^{R31X}/*USH1C*^{Del}) and the results are published in GROTZ et al. (2022). A plasmid expressing Cas9, a plasmid expressing the sgRNA *urg1* and one of the five different ssODNs were co-

electroporated into the cells in distinct constellations (Table 26). Cells were harvested as mixed cell populations after 24 hours. Out of the mixed cell population, gDNA was isolated and PCR products of USHR31X PCR were used for Sanger sequencing and analysis of modification rate. Experiments were repeated twice in the same constellations, but at different cell passages.

Table 26: Composition of *USH1C* reconstituting GE experiments

<u>experiment</u>	<u>components</u>
Epo1	Cas9 + urg1 + urt1.1
Epo2	Cas9 + urg1 + urt1.2
Epo3	Cas9 + urg1 + urt1.3
Epo4	Cas9 + urg1 + urt1.4
Epo5	Cas9 + urg1 + urt1.5
Epo6	Cas9 + H ₂ O

Analysis of the electropherograms from Sanger sequencing with the ICE tool by Synthego showed HDR rates of up to 30%, which was confirmed by analysis with FreeHand MXa. Rates of HDR and NHEJ declined with higher passage, but the different ssODNs kept similar efficacy compared to each other. In line with the expectations postulated by RICHARDSON et al. (2016), the asymmetrically and PAM opposite ssODN urt1.3 performed best. Estimation of successful editing is relatively easy in this case, as a clear change in the sequence at the causative mutation as well as at the blocking mutation proves the reconstitution of the gene by orienting on the repair template. The substantial amount of NHEJ occurring instead of or in parallel to the desired HDR has to be considered.

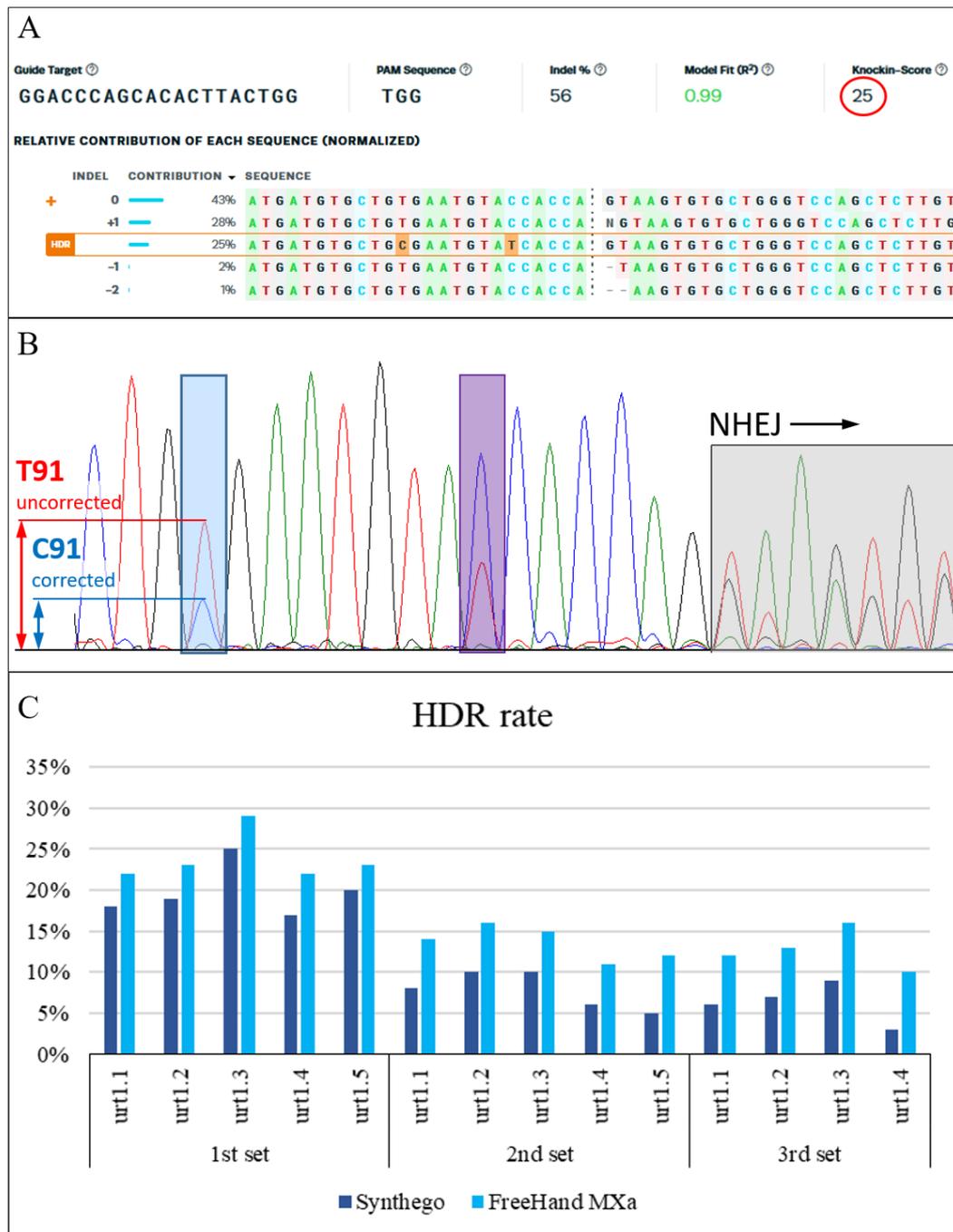


Figure 24: Analysis of HDR rates

Adapted from GROTZ et al. (2022), Figure 7

(A) Exemplary analysis of an electropherogram from Sanger sequencing with ICE by Synthego. The sequence of the used ssODN was put in as “donor template”. (B) shows exemplary analysis of the same electropherogram with FreeHand MXa. The causative mutation is marked with a blue box, the blocking mutation in magenta. Peak heights of the corrected and the uncorrected nucleotides were measured. From the cut site onwards, the sequence becomes unclear due to the manifold variants of NHEJ (grey box). (C) shows a comparison of the two methods for analyzing HDR rates in all three sets of experiments.

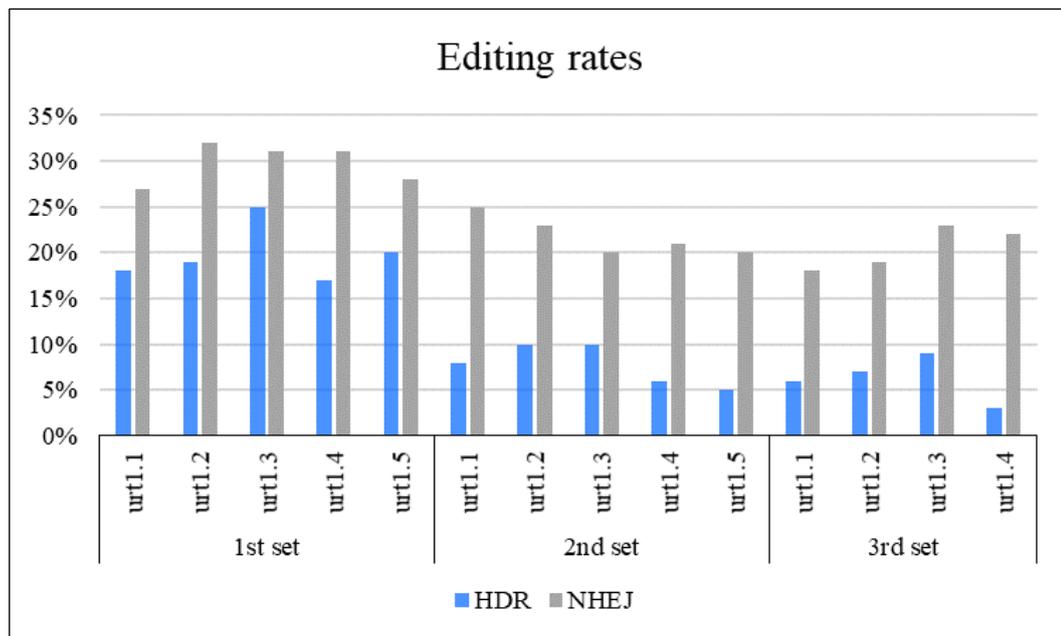


Figure 25: Editing rates in *USH1C* reconstituting GE experiments

Adapted from GROTZ et al. (2022), Figure EV4

Here the rates of HDR and NHEJ analyzed with Synthego are compared over the course of all three experiments. Rates of both HDR and NHEJ decline with higher cell passage, still, urt1.2 and urt1.3 seem to be favorable.

3.4. Generation and analysis of SCCs

For examining the exact changes on single alleles, I generated SCCs from the mixed cell population that was treated with urg1 and urt1.3. 72 SCCs were analyzed with Sanger sequencing. 29 SCCs (40%) underwent HDR, in three cases accompanied by additional changes leading to disruption of the correct amino acid sequence. In another two cases, small additional changes occurred in the intronic sequence, with presumably no effect on the protein. In three cases only the blocking mutation was incorporated, probably due to its localization closer to the cutting site. 23 SCCs (32%) underwent NHEJ, while another 20 SCCs (28%) remained unaffected (Figure 26).

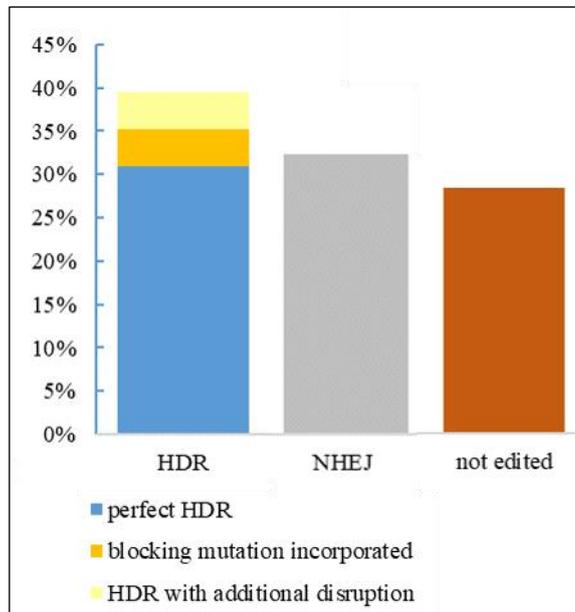


Figure 26: Efficacy of *USH1C* reconstituting GE in SCCs

Adapted from GROTZ et al. (2022), Fig. 7

Comparison of SCCs that underwent HDR (including only partly or comprising additional changes), those that underwent NHEJ and not edited SCCs.

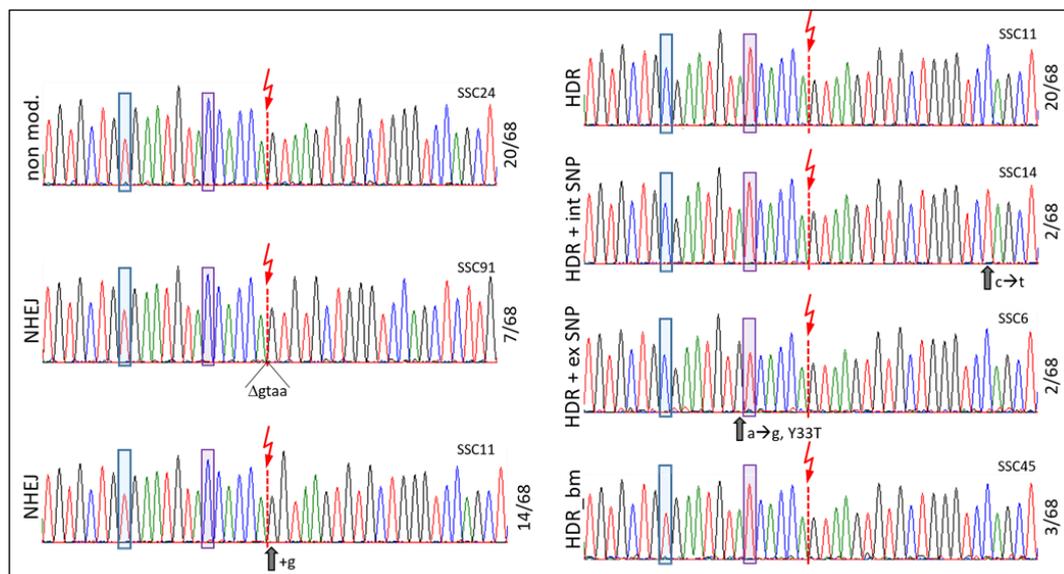


Figure 27: Analysis of individual SCC sequences

As in GROTZ et al. (2022), Fig. EV4D

Exemplary electropherograms from SCCs show a variety of modifications. The cut site of the sgRNA is marked with an arrow, the causative mutation and the blocking mutation with boxes as in Figure 24B. The total number of clones which underwent the shown modification is indicated on the right side of each electropherogram.

The proof of concept for reconstituting GE at the *USH1C*^{R31X} locus *in vitro* was quite successful, proposing to move therapy testing towards pre-clinical *in vivo* studies and clinical trials. The low rate of HDR in photoreceptors and the safety concerns raised from DSB caused by classical CRISPR/Cas9 suggests the adaptation of this strategy. For this reason, I further explored GE in *USH1C* primary cells with an adapted variant of CRISPR/Cas avoiding DSBs.

3.5. Additional therapy approach: Prime Editing

In addition to the classical CRISPR/Cas plus ssODN approach, I tested the possibilities of prime editing for the correction of the *USH1C* mutation. Based on the previously used sgRNA *urg1* I designed six different prime editing guide RNAs (pegRNAs), using the successfully tested *urg1* as the spacer sequence and varying the lengths of the reverse transcriptase template (RTT) and primer binding sequence (PBS) (see Figure 3). RTT were either 16 or 22 nucleotides long, the PBS ranged from 10 to 15 nucleotides. Experiments were performed on the same cell line 5615 (*USH1C*^{R31X}/*USH1C*^{Del}) used for HDR experiments previously. A plasmid expressing the modified Cas9 with the reverse transcriptase (kindly provided by Dr. F. Giesert, Institute of Developmental Genetics, Helmholtz Center Munich) and a plasmid expressing the respective pegRNAs were co-electroporated into the cells.

In the first set of experiments I did not detect any desired modification with the ICE tool by Synthego. In a second set of experiments, I therefore combined one of the pegRNAs with two different assistant sgRNAs to additionally nick the non-edited strand in either intron 1 or 2. The plasmids were co-electroporated into the cells in the following combinations:

Table 27: Composition of prime editing experiments

<u>experiment</u>	<u>components</u>
Epo1	PE2 + ushpegRNA2
Epo2	PE2 + ushpegRNA2 + asgRNA2
Epo3	PE2 + ushpegRNA2 + asgRNA4

While co-transfection of a plasmid expressing *asgRNA4* did not cause improvement, usage of *asgRNA2* resulted in a gene correction rate of approximately 3% (Figure 28).

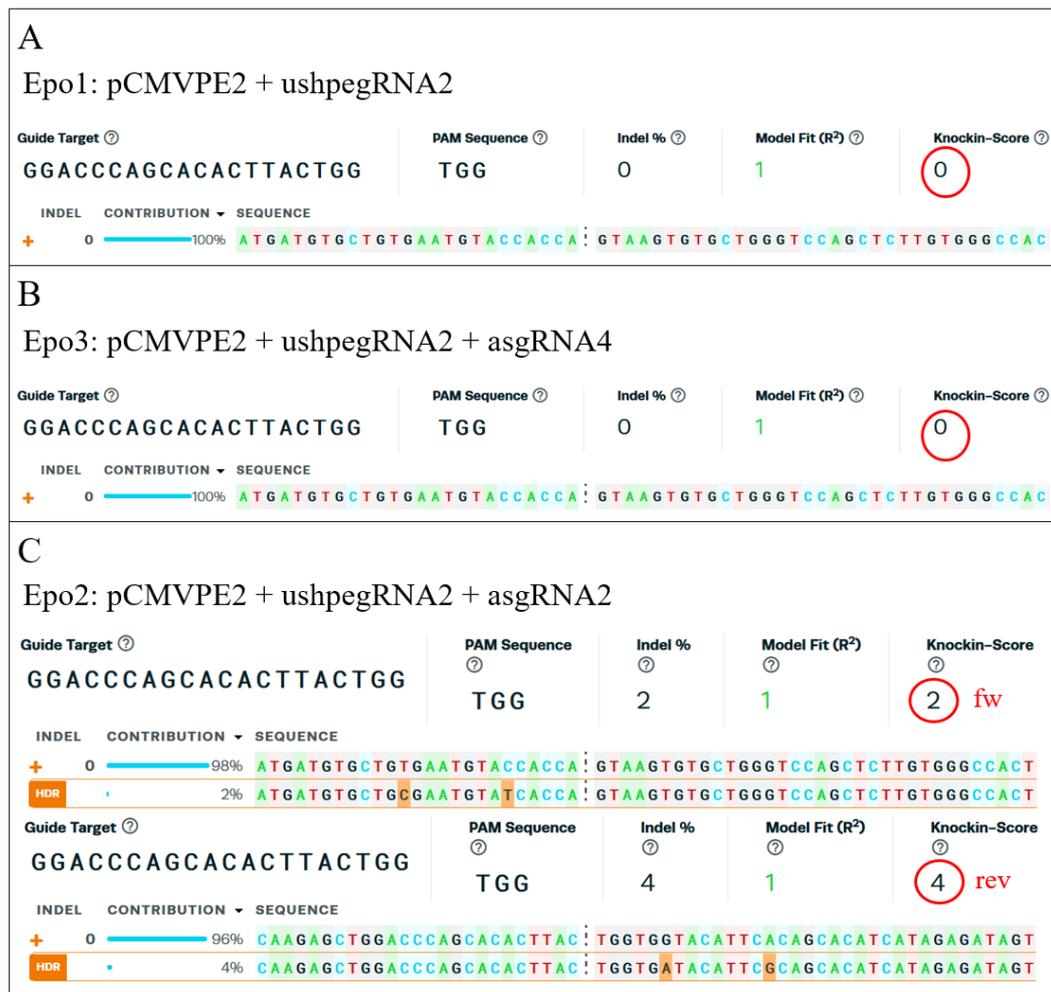


Figure 28: Analysis of Epo 1-3 with Synthego

Analysis of Epo 1 using only ushpegRNA2 and Epo 3 using the combination of ushpegRNA2 and asgRNA4 (placed in intron 1) did not show any modification. Analysis of Epo 2 combining ushpegRNA2 and asgRNA2 (placed in intron 2) indicated a correction rate of 2% when sequenced with the forward primer and 4% when sequenced with the reverse primer.

Due to the low efficacy, no detailed examination was carried out in SCCs. Obviously, the prime editing approach needs considerable optimization for reaching an efficacy rate similar as the classical CRISPR/Cas regimen as shown above.

V. DISCUSSION

Application of GE in pig models of IRDs covers a wide spectrum of research opportunities. By now, GE has not only revolutionized the prospects of therapies, but also the generation of animal models. Numerous variations of the CRISPR/Cas technology have been presented, including plain NHEJ-mediated disruption of target genes as well as the introduction of defined modifications by HDR using ssODNs, classical targeting vectors or BACs as recombination templates. My work aims at the generation of a tailored pig model to gain deeper insight into pathological mechanisms, as well as exploring the prospects of therapeutical approaches *in vitro*. Special focus is on the translational aspect, aiming at the adjustment of the *in vitro* tools for *in vivo* application. The high value of genetically modified pigs as models for translational biomedical research is appreciated in various human disease patterns, such as neurodegenerative and cardiovascular diseases, cystic fibrosis and diabetes mellitus (AIGNER et al., 2010; HRYHOROWICZ et al., 2020). With the tools for the generation and the models themselves becoming more and more sophisticated, their use for studying the molecular mechanisms and the clinical phenotype in detail is growing. Particular focus can be placed on testing safety and efficacy of novel therapies, including gene replacement, gene repair, exon skipping, translational read-through and many more (KLYMIUK et al., 2015).

Generation of a tailored pig model - technical aspects

BACs as targeting vector proved a valuable tool even without CRISPR/Cas, used by SONG et al. (2010) in human embryonic stem cells and in the generation of several pig models in our institute (CFTR: KLYMIUK et al. (2012), DMD: KLYMIUK et al. (2013)). Recent approaches combining BACs with CRISPR/Cas increased efficiency substantially (USH1C: GROTZ et al. (2022), Dr. Petra Runa-Vochozková, TU Munich, personal communication). The large size of BACs creates several advantages and disadvantages, compared to classical targeting vectors that comprise only several kb of homologous arms on either side. In the case of the latter, screening for HDR can be done with end-point PCRs spanning across the homologous arms, but the establishing of such screening tools often requires the separate generation of an additional mock-construct as positive control.

Further, the efficacy of recombination might be impaired in the case of unknown sequence variation in the targeting region in the genome. This is of particular relevance in the pig, where significant genetic variation between individuals requires exploration of that region specifically in the primary cell population of interest. All these restrictions are prevented by the use of modified BACs as targeting vectors. The handling of such large vectors (> 180 kb) is definitely challenging, but if procedures are carried out appropriately, disruption of the BAC structure remains a rare event, as shown exemplarily in my thesis (Figure 7). The fact that after final modification of the BAC, most examined clones and all SCNT derived founder pigs were correctly modified, supports the prediction that BACs are reliable in handling and retain a stable composition when modified (GIRALDO & MONTOLIU, 2001). When choosing a BAC for recombineering, it is necessary to determine the true location in the pig genome. Relying on the proposed localization by the PigBAC Pre map may be misleading, as I experienced with BAC CH242-67G24 that apparently did not cover the porcine *GUCY2D* locus. Even the two other purchased BACs covered the target site in a rather asymmetric way (Table 22), but previous experience with a similar situation when generating the DMD pig model (KLYMIUK et al., 2013) suggested to proceed with the available clones rather than to restart searching for a more balanced distribution of HA with uncertain perspectives. The relatively low number of 6 correctly modified clones out of 202 (Figure 13) may be a consequence of that constellation, but was still high enough to provide suitable SCCs for successful SCNT and the generation of founder animals.

A further prerequisite for modifying BACs is the usage of positive selection in prokaryotic cells for bacterial recombineering of the BAC and in eukaryotic cells for homologous recombination of the BAC with the target genome. This can be conveniently achieved by a selection cassette that is functional in both types of cells and driven by distinct promoters in prokaryotic as well as eukaryotic cells. Addition of a neomycin selection cassette with a prokaryotic and a mammalian promoter permits selection of both bacterial colonies when assembling the correct modification vector, as well as the cells in culture after nucleofection. Importantly, pre-testing of cell types for sensitivity to antibiotic concentration is necessary. In contrast to LANZA et al. (2013), who recommend zeocin when comparing hygromycin B, neomycin, puromycin and zeocin in cell line development in human cells, I experienced some difficulties with zeocin, especially in bacterial selection.

Puromycin was excluded for the same reason. Based on the successful experience in our lab, I decided to stick to neomycin, which was the basis for numerous models. Good results were also obtained with blasticidin S (KLYMIUK et al., 2012), but simply due to its rare application, it remains the second best choice.

An alternative approach to identify transgenic cells is fluorescent-activated cell sorting (FACS). Selection is then based on expression of a reporter gene either in the cell itself or tagged to antibodies binding on cell surface proteins (COLOSIMO et al., 2000). This requires several preconditions, i.e. either the desired transgene with the reporter or the detected surface epitope needs to actually be expressed in the sorted cell type. Both of these preconditions are not fulfilled in our setting, since RETGC1 expression is almost exclusively limited to the retina (Lucie Casalta, TU Munich, personal communication), and the mutated dimerization domain is only located in the intracellular compartment (LAURA et al., 1996). Further, placement of a reporter might impair subsequent applications of the models. In pre-clinical trials for AAV mediated gene therapy, a testing of transfection efficacy is common practice, often with AAVs tagged by fluorescent reporters like GFP. A model already expressing a fluorescent reporter is much less suited for such an application. The aspiration to create a genetic constellation that is as close as possible to the situation in patients also indicates to reduce the collateral side effects of the genetic modification process. This is definitely also true for the neomycin selection cassette, that is indispensable for the modification process itself, but has no further use after stable integration of the modification into the genome. Indeed, remaining neomycin selection cassettes have been eyed with suspicion for a long period of time (VALERA et al., 1994), but we presume less effect than a residual reporter gene might have. The intended removal of the neomycin selection cassette via FLPe/FRT-mediated excision lessens the concerns about these unwanted side-effects. Both the Cre/loxP and the FLPe/FRT system are long known and well established, especially in transgenic mice (BIRLING et al., 2009). The preference for the FLPe/FRT system in my strategy was based mainly on the aspect that the Cre/loxP system has been successfully implemented at our institute before, and we wanted to see if working with FLPe/FRT was similarly efficient, in case of need for both systems in parallel in a future attempt.

At first sight, the need for stepwise screening procedures during the BAC manipulation and the evaluation of a large number of SCCs after nucleofection may seem very laborious, but the application of the qPCR-based LOWA-approach

allows for a fast scanning and exclusion of about 90% of SCCs from detailed analysis (Figure 10). The more time-consuming and more informative analysis by Sanger sequencing of PCR products can then be limited to a small set of candidate clones. In fact, the efficient workflow from design to verified SCCs, even topped by successful SCNT and birth of founder animals within the first funding period allowed for the application and approval for an independent grant for our *GUCY2D* pig model.

Overall, the chosen approach of modifying the porcine *GUCY2D* gene with a combination of modified BAC vector plus CRISPR/Cas may be seen as technically challenging, but provides a safe and fast-forward strategy to tailor a pig model in a very sophisticated way and allows the implementation of any desired features.

The *GUCY2D* pig model - prospects and possibilities

The *GUCY2D* pig model has been designed to complement the existing USH1C pig in terms of disease mechanism and progress. I am confident that it will contribute to the refinement of clinical readout parameters such as OCT/ERG and behavioural tests to evaluate the natural progression of retinal diseases. Being suited for cell replacement therapies as well as direct treatment of the causative mutation at the genomic level, both reconstituting and disruptive, it will help in comparing and quantifying the impact of different therapies. The *GUCY2D* pig model will benefit from the already established interdisciplinary expertise around the USH1C pig and the ongoing support from the SPP2127. Specialists in molecular and cellular analyses as well as in physiological and anatomical characterization are part of our network and our growing experience in conducting pre-clinical studies will add substantially to research in the field of IRDs (Table 25).

When evaluating the possibilities of disruptive GE as a therapy approach in this model, I compared the different strategies for reaching a maximum of successful edits. Successful targeting of the mutation site itself with one sgRNA proves accessibility of the locus and offers the possibility of developing reconstituting GE strategies. Considering the variability of dominant negative *GUCY2D* mutations (see OMIM database), the direct targeting of the causative mutation will differ between mutation variants. The approach of excising essential genomic regions on the basis of allele-discriminating SNPs, exemplified in my thesis by deletion of exon 13, is a valuable alternative providing more flexibility. The challenge of nucleofecting a total of three plasmids in parallel did not cause substantial

restrictions due to the high nucleofection efficacy *in vitro*. The necessity to design sgRNAs based on specific SNPs reliably differentiating between the healthy and the diseased allele pays off with the good performance of 2 sgRNAs working in combination, almost reaching 80% (Table 24).

The mechanistic properties of RETGC1 as a dimer-forming protein are of special interest considering phenotype variations. As mutated RETGC1 supposedly interacts with the wild-type form, dimers of mutant and WT proteins might hinder the physiological process. In the case of a destruction and complete degradation of the mutated RETGC1, existence of mut-WT dimers would not occur, possibly being beneficial for the outcome (SHARON et al., 2018). For this reason, future experiments will evaluate the possibility of destroying the *GUCY2D* locus further upstream, which will also reflect the constellation in recessive *GUCY2D* patients more closely.

Plans for characterizing the *GUCY2D* pig will include analysis of retinal samples obtained from the founder piglets for structural deficits and the progression of cone degradation, but for more systematic analysis, examination of defined cohorts of *GUCY2D* pigs is necessary. I expect that this will not only give insight into the immediate degeneration process in *CORD6*, but also reveal to which extent function of the cone rich visual streak is indicative for diseases of the fovea centralis in human beings (SANCHEZ et al., 2011). This is one of the major unresolved questions in cone affecting diseases, as a structure resembling the human macula is only present in primates. Alternative animal models comprising comparable characteristics would be highly appreciated. Moreover, focus will be on characterizing the functional deficits with already established ophthalmological and behavioural analyses as well as the localization of a potential therapeutic window. Expression analyses and phenotype characterization will be performed in comparison to neo-free founder pigs as soon as they are available. qPCR results (Figure 15) strongly indicate the probability of this in the next round of SCNT, however, still intact humanization after FLPe treatment has of course to be verified in the new founder pigs.

Although phenotype characterization and the establishment of a breeding herd are goals still to be fulfilled, generation of first founder animals with verified genotype and storage of cells and tissue provides a robust basis for further research. With proof of concept for GE therapy shown in the pig cell line, this calls for individual design of GE treatment for each single patient in future clinical applications.

Breeding, raising and handling the USH1C pig

A substantial part of my scientific and veterinary work in the past three years was promoting the USH1C pig model in our facility for its usage in various pre-clinical studies (Table 25). What makes the USH1C pig such a valuable model for studying the disease is also its most challenging characteristic: the presence of both inner ear and eye impairment as seen in human patients. On the one hand, it offers the possibility to study the inner ear's structure, its functionality and the consequences of impairment as done in collaborative studies with Prof. Dr. Andrea Fischer and Prof. Dr. Andreas Parzefall. Experiments have stimulated further collaborations, e.g. with Dr. Aziz El Amraoui (Department of Hearing Institute, Institut Pasteur, Paris, France) and Dr. Oliver Profant (Department of Auditory Neuroscience, Charles University, Prague, Czech Republic). On the other hand, piglets need intensive care during the first few days in their life. For the congenital abundance of vestibular areflexia and the precocial behaviour of newborn piglets, they are not able to feed on their own due to pronounced circling. Together with my former colleague Dr. Sophia Grotz, we tried two different approaches of raising USH1C piglets: either separate them from their mother right after birth and raise them with commercially available milk replacers in a motherless rearing system ("rescue deck") as described in EGERER et al. (2018). The other possibility is to let them with the mother sow, supporting them in finding and holding on to the teat for the process of suckling, which is even more time consuming and labour intensive. Although general dangers of injury by the sow or the environment are more pronounced in this scenario, we came to the conclusion that the overall viability and growth of the piglets is much better when being nurtured by the sow. Even with highest hygienic standards, we were never able to overcome problems with diarrhea in the rescue decks, which always led to growth retardation and ongoing digestive problems for many months later. Interestingly, similar if not equally poor results are obtained when the piglets are being reared with milk replacer while being at the sow, which was explored in the case the sow had limited milking capacity after birth. After all, the nutritional benefit of sow milk seems to be so essential that the laborious first days paid off well. Importantly, raising with the sow as well as the general challenge of overcoming the first few critical days limits the number of animals that can be maintained at once. This implicates the need to find a balance between the number of sows giving birth simultaneously and the demand for larger cohorts for studies.

When aiming to produce larger numbers of animals for different studies, it is necessary to breed. ET is an excellent way of producing founder animals but too inefficient for production at large scale. Further, cloned offspring often suffer from clonal artefacts, litters are small and suffer poor viability, and nuclear reprogramming in the SCNT process might influence all kinds of expression patterns (KUROME et al., 2013; HÖRMANSEDER et al., 2017). The major reason for this is seen in impaired epigenetic reprogramming. Mating or artificial insemination is evidently the most natural and therefore most reliable form of breeding. Still, propagating USH1C pigs this way needs some considerations. When mating heterozygous carriers, the allelic constellations being inherited by Mendelian manners create USH1C pigs and control littermates with the same genetic background in one litter. Given the gene variability in pig populations and the necessity to maintain this variability to avoid detrimental effects on fertility and health as a consequence of inbreeding, the comparative examination of littermates is a well-established way to reduce the number of experimental animals in a study. Regularly, more USH1C pigs than control animals are needed in certain studies, e.g. dose finding in AAV therapy. Here, mating of homozygous USH1C founder animals with heterozygous offspring provides 50% of USH1C pigs in each litter. Breeding with homozygous USH1C females proved difficult at first, because the sows did not show all typical heat symptoms and especially refused the natural mating. This was overcome by careful observation of the sows and synchronizing estrous with commercially available hormones, followed by artificial insemination. Due to USH1C pigs being more susceptible to stress and new surroundings, accompanying the homozygous sows through the time of pregnancy and especially preparing them for a calm birth is of high importance. A close relationship between the sow and the individual caretaker, relying on smell and touch creates a trusting environment for the animal that is a substantial part in animal welfare. Consequently, the homozygous founder sows proved to be excellent mothers, treating their piglets with utmost care despite their impaired senses.

Now that we have a successful breeding herd established, the accidental appearance of the deletion of exon 2 raises an interesting question: is it useful to keep both the R31X mutation as well as the deletion of exon 2 for future experimental animals? Both modifications result in a harmonin null function, therefore the full phenotype is observed. For testing GE therapies or translational read-through drugs, the R31X mutation is of course mandatory, but since the natural occurrence of the *USH1C*

loss of function in human patients is compound heterozygous, keeping both alleles actually reflects better the clinical situation (GROTZ, 2021). Admittedly, it makes genotyping more complicated, especially since we do not know about the exact location and extent of the deletion. Using a copy-number discriminating qPCR for the R31X allele is an adaption of the SCCs screening to reveal the allelic constellation (Figure 22). Taking into account the fast and reliable performance of this qPCR, we have so far not explored the exact dimensions of the deletion, considering that it would be either very time consuming to screen large parts of the *USH1C* locus with end-point PCRs in a trial-and-error approach, or send DNA samples for expensive nanopore sequencing to external service providers.

In conclusion, the USH1C breeding herd is established and a steady supply of experimental animals to scientific partners is initiated. Optimization attempts especially on USH1C piglet survival are ongoing.

Evaluating the tools for Gene Editing

For historical reasons, the USH1C model was established first, leading to testing the more challenging gene reconstitution in the beginning. When using the USH1C model in cell culture for evaluating GE, the future applicability *in vivo* is of course the ultimate goal. Therefore, the question is raised why to start off with the classical CRISPR/Cas plus ssODN approach at all, when it is clear that this system will be very unlikely to be transferred *in vivo* due to low capacity of viral vectors and the single stranded nature of the repair template. The main advantage of this basic approach is to check the accessibility of the targeted locus in a very rapid and easy setting. On the downside, the various dangers of DSBs including additional modifications and off-target effects have to be monitored closely. High rates of HDR with this approach are indicative for good accessibility of the locus itself and represent a scale for other approaches. However, one has to keep in mind that the fibroblasts treated are highly proliferative, therefore presenting a totally different setting from post-mitotic PRs. To create a situation resembling terminally differentiated cells, cultivation in starvation medium or with chemicals that arrest the cell cycle can be an option to have a better prediction of therapeutic GE in post-mitotic cells, as summarized in KUMAR et al. (2007). Such attempts need elaborate establishing work, investigating cell viability and activity of gene repair pathways under these circumstances.

In search for a system working *in vivo* in post mitotic cells, prime editing is postulated to be suited for reconstituting GE in this surrounding, also rendering concerns about DSBs unnecessary. However, considering the substantially lower efficacy that I observed with such an approach in USH1C cells, this shows the importance of having obtained successful results at the same locus with the more robust HDR approach before. Having realized that at least a small effect is achieved when combining prime editing with an additional nick of the unedited strand in close proximity to the intended mutation site, future improvements will include the testing of several different assistant gRNAs and combining the best one with all six variants of pegRNAs designed. If efficacy remains low, other possibilities such as base editing have to be reconsidered. Given the enormous development in the field of therapeutic GE, availability of novel tools in the near future is also very likely. In contrast to the reconstituting approach, efficacy of disruptive GE of up to 80% in the *GUCY2D*^{E837D/R838S} cell lines is already fit for pre-clinical studies. Applicability of disruptive CRISPR/Cas therapy *in vivo* has been demonstrated before (MORETTI et al., 2020), and the ability to conduct pre-clinical eye studies at our institute is being underlined by ongoing investigations (Table 25). Establishing of cohorts of *GUCY2D* pigs and examination of the phenotype is, however, pending.

Any *in vivo* application needs efficient delivery of the repair system into the cell. AAVs, though perfect for the eye, are limited by their packing capacity (VANDENBERGHE & AURICCHIO, 2012). Overcoming this obstacle is possible by splitting the large components and reconstituting the full protein inside the cell with intein-mediated trans-splicing (TRUONG et al., 2015). Promising results were obtained by TORNABENE et al. (2019), who were able to reconstitute large full length proteins *in vitro* in pig retinas and human organoids, as well as *in vivo* in two mouse models of IRDs, where it actually improved the retinal phenotype. Intensifying our collaboration with Dr. Florian Giesert, we aim at packaging the PE along with pegRNA and asgRNA into three AAV genomes and reconstituting the prime editing system upon co-transfection inside the cell. This optimization of virus-compatible self-assembly will be beneficial for the treatment not only of USH, but of various other monogenetic diseases. Other concerns about viral delivery systems, such as immunogenicity and long-term toxicity can be tackled by exploring the prospects of non-viral delivery systems, e.g. lipid nanoparticles (PATEL et al., 2019).

Final conclusion and outlook

In the broad field of generation and treatment of pig models for IRDs, this thesis underlines the potential of GE in large animal models for biomedical research. The work on USH1C pig production and conduction of pre-clinical gene therapy studies to which I contributed in addition to GE experiments, not only illustrates how multidisciplinary work must be organized for effective working on large animal models, but has also led to establishing structures for future preclinical evaluation of therapeutic GE in USH1C and GUCY2D pigs. Close collaboration of different fields and several models and therapeutic strategies complementing each other is of great importance to promote further progress in this area of research.

VI. SUMMARY

Gene Editing in pig models of inherited retinal diseases

Inherited retinal diseases (IRDs) affect approximately one in 2000 people worldwide, with many variants in genetic background and clinical phenotype. For research on the pathophysiological processes as well as on the evaluation of novel therapeutic strategies like Gene Editing (GE), the generation of tailored animal models with established as well as emerging biotechnological tools is an important asset. In my thesis I focused on two complementary IRDs differing in inheritance pattern and disease mechanism, Usher syndrome (USH) and cone-rod dystrophy based on *GUCY2D* (CORD6). All USH subtypes are inherited in a recessive manner, featuring a relatively slow progression starting with rod loss and are accompanied by a congenital vestibular and auditory phenotype. The USH1C pig model is based on the patient specific c.C91T/p.R31X nonsense mutation obliterating harmonin, a scaffold protein involved in mechano-electrical transduction in both photoreceptors and hair cells. In contrast to this, CORD6 displays an autosomal dominant inheritance pattern, characterized by early onset of cone dysfunction followed by rods later on. Disease phenotype is based on the over-activity of mutated retinal-specific guanylate-cyclase 1 (RETGC1), leading to high Ca^{2+} concentration and ultimately resulting in photoreceptor decline.

Generation of a pig model for CORD6 comprising the *GUCY2D*^{E837D/R838S} mutation in the dimerization domain of RETGC1 was achieved using a targeting vector consisting of a porcine BAC that was modified to carry the humanized exon 13 with the causative mutation. A manuscript on protocols for generating large targeting vectors has been published during the writing of this thesis (AUCH et al., 2022). After co-nucleofection of the modified targeting vector with Cas9 and a sgRNA, single cell clones were analyzed in a consecutive manner involving the qPCR-based LOWA-approach, as well as several end-point PCRs and Sanger sequencing, with six of them proving to be correctly modified. Somatic cell nuclear transfer resulted in one pregnancy and the birth of three live founder animals, which also have the entire construct incorporated correctly. Cell lines isolated from these founder animals were used to test disruptive GE as a therapy approach *in vitro*. Seven different sgRNAs were co-nucleofected with Cas9 either alone or in pairs to achieve

inactivation of the dominant negative *GUCY2D* allele, resulting in successful editing of up to 77%.

For evaluating the prospects of reconstituting GE *in vitro*, cell lines from the USH1C pig model were co-nucleofected with Cas9, a sgRNA and different ssODNs as a repair template. Successful editing by homology-directed repair (HDR) reached up to 40%. Aiming at a possible adjustment of reconstituting GE approaches *in vivo*, the same locus was tried with Prime Editing, co-nucleofecting six different pegRNAs with a prime editor consisting of a dCas9 and a reverse transcriptase. Combination with an assistant sgRNA to additionally nick the non-edited strand resulted in successful editing of about 3%.

Apart from testing GE approaches in my thesis, USH1C pigs were produced and provided for various pre-clinical studies and research projects with national and international collaborations, i.e gene therapy studies, a study on photoreceptor transplantation and one focusing on the inner ear phenotype.

In conclusion, tailored pig models of both USH and *CORD6* are important assets to *in vitro* and *in vivo* pre-clinical research. The tools for generation and therapeutical approaches are developing rapidly and will become more and more individually designed in the future. Proof of concept for both disruptive and reconstituting GE holds promise for further application in treatment of IRDs.

VII. ZUSAMMENFASSUNG

Genom-Editierung in Schweinemodellen für erblich bedingte Netzhauterkrankungen

Etwa eine von 2000 Personen weltweit ist von erblich bedingten Netzhauterkrankungen betroffen, wobei der genetische Hintergrund und der klinische Phänotyp sehr vielfältig sind. Für die Erforschung der pathophysiologischen Prozesse sowie für die Evaluierung neuer therapeutischer Strategien wie Genom-Editierung (GE) ist die Erstellung maßgeschneiderter Tiermodelle sowohl mit etablierten als auch mit neu aufkommenden biotechnologischen Werkzeugen ein wichtiger Vorteil. In meiner Dissertation habe ich mich auf zwei komplementäre erblich bedingte Netzhauterkrankungen konzentriert, die sich im Erbgang und im Krankheitsmechanismus unterscheiden: das Usher-Syndrom (USH) und die Zapfen-Stäbchen-Dystrophie auf der Grundlage von *GUCY2D* (CORD6). Alle USH-Subtypen werden rezessiv vererbt und zeichnen sich durch relativ langsames Fortschreiten aus, das mit dem Verlust von Stäbchen beginnt und von einem angeborenen vestibulären und auditiven Phänotyp begleitet wird. Das USH1C-Schweinemodell basiert auf der patientenspezifischen c.C91T/p.R31X-Nonsense-Mutation, die die Produktion von Harmonin verhindert. Harmonin ist ein Gerüstprotein, das an der mechanoelektrischen Transduktion sowohl in Photorezeptoren als auch in Haarzellen beteiligt ist. Im Gegensatz dazu zeigt CORD6 ein autosomal-dominantes Vererbungsmuster und ist durch ein frühes Auftreten von Zapfen- und später von Stäbchenfehlfunktionen gekennzeichnet. Der Phänotyp beruht auf einer Überaktivität der mutierten retina-spezifischen Guanylatcyclase 1 (RETGC1), die zu einer hohen Ca^{2+} -Konzentration und schließlich zum Untergang der Photorezeptoren führt.

Die Erzeugung eines Schweinemodells für CORD6 mit der *GUCY2D*^{E837D/R838S}-Mutation in der Dimerisierungsdomäne von RETGC1 wurde mit Hilfe eines Targeting-Vektors erreicht, bestehend aus einem porcinen BAC, der so modifiziert wurde, dass er das humanisierte Exon 13 mit der ursächlichen Mutation trägt. Ein Manuskript über Protokolle zur Erzeugung großer Targeting-Vektoren wurde während der Erstellung dieser Arbeit veröffentlicht (AUCH et al., 2022). Nach Nukleofektion des modifizierten Targeting-Vektors zusammen mit Cas9 und einer sgRNA wurden Einzelzellklone mit dem qPCR-basierten LOWA-Ansatz und

nachfolgend mehreren Endpunkt-PCRs und Sanger-Sequenzierung analysiert, wobei sich sechs von ihnen als korrekt modifiziert erwiesen. Somatischer Zellkerntransfer führte zu einer Trächtigkeit und der Geburt von drei lebenden Gründertieren, die ebenfalls das gesamte Konstrukt korrekt eingebaut haben. Die aus diesen Gründertieren isolierten Zelllinien wurden verwendet, um destruktive GE als Therapieansatz *in vitro* zu testen. Zusammen mit Cas9 wurden sieben verschiedene sgRNAs entweder allein oder paarweise nukleofiziert, um die Inaktivierung des dominant negativen *GUCY2D*-Allels zu erreichen, was zu einer erfolgreichen Editierung in bis zu 77 % der Fälle führte.

Um die Aussichten für wiederherstellende GE *in vitro* zu bewerten, wurden Zelllinien des USH1C-Schweinmodells mit Cas9, einer sgRNA und verschiedenen ssODNs als Reparaturvorlage nukleofiziert. Editierung durch homologe Rekombination (HR) war in bis zu 40 % der Fälle erfolgreich. Im Hinblick auf eine mögliche Anpassung der wiederherstellenden GE-Ansätze *in vivo* wurde Prime Editing auf denselben Genort angewendet, wobei sechs verschiedene pegRNAs mit einem Prime Editor, bestehend aus dCas9 und reverser Transkriptase, nukleofiziert wurden. Die Kombination mit einer assistierenden sgRNA, um den nicht editierten Strang zusätzlich zu schneiden, führte zu einer erfolgreichen Editierung in etwa 3% der Fälle. Zusätzlich zur Erprobung von GE-Ansätzen im Rahmen meiner Dissertation wirkte ich an der Produktion von USH1C-Schweinen für verschiedene präklinische Studien und Forschungsprojekte in nationaler und internationaler Zusammenarbeit mit, z. B. für Gentherapiestudien, eine Studie zur Photorezeptortransplantation und eine Studie mit Schwerpunkt auf dem Innenohrphänotyp.

Zusammenfassend lässt sich sagen, dass maßgeschneiderte Schweinemodelle sowohl für USH als auch für *CORD6* einen wichtigen Beitrag zur präklinischen *in vitro* und *in vivo* Forschung darstellen. Die Werkzeuge für die Generierung und die therapeutischen Ansätze entwickeln sich rasch und werden in Zukunft immer individueller gestaltet werden. Der grundlegende Beweis für die Durchführbarkeit sowohl destruktiver als auch wiederherstellender GE ist vielversprechend für eine zukünftige Anwendung bei der Behandlung von erblich bedingten Netzhauterkrankungen.

VIII. INDEX OF FIGURES

<i>Figure 1: Basic principle of CRISPR/Cas</i>	13
<i>Figure 2: Repair mechanism of the cell</i>	14
<i>Figure 3: Components of Prime Editing</i>	19
<i>Figure 4: Location of the BACs</i>	38
<i>Figure 5: Sequence as ordered for commercial synthesis for the modification plasmid</i>	39
<i>Figure 6: huGUCYfrtNeo plasmid digestion</i>	40
<i>Figure 7: Verification of modified BAC</i>	41
<i>Figure 8: Strategy of generating the GUCY2D pig model</i>	42
<i>Figure 9: Reference loci</i>	43
<i>Figure 10: Analysis of SCCs with LOWA-approach</i>	44
<i>Figure 11: Allelic constellation of poGUCY2D in candidate SCCs</i>	45
<i>Figure 12: Verification of huGUCY2D in SCCs</i>	46
<i>Figure 13: Summary of sequential SCC characterization</i>	47
<i>Figure 14: Validation of the genotype in founder piglets</i>	49
<i>Figure 15: qPCR results of FLPe-treated pig primary cells</i>	50
<i>Figure 16: GUCY2D disruptive gene editing strategy</i>	52
<i>Figure 17: GUCY2D disruptive GE</i>	53
<i>Figure 18: Analysis of Epo 3-5 with Synthego</i>	54
<i>Figure 19: Analysis of Epo 3-5 with FreeHandMXa</i>	55
<i>Figure 20: Analysis of band intensity for Epo 1 and Epo 2</i>	57
<i>Figure 21: Analysis of Epo 1 and Epo 2 with Synthego</i>	58
<i>Figure 22: Genotyping of F2 generation</i>	61
<i>Figure 23: USHIC reconstituting GE strategy</i>	65
<i>Figure 24: Analysis of HDR rates</i>	67
<i>Figure 25: Editing rates in USHIC reconstituting GE experiments</i>	68
<i>Figure 26: Efficacy of USHIC reconstituting GE in SCCs</i>	69
<i>Figure 27: Analysis of individual SCC sequences</i>	69
<i>Figure 28: Analysis of Epo 1-3 with Synthego</i>	71

IX. INDEX OF TABLES

<i>Table 1: Usher subtypes based on clinical features</i>	4
<i>Table 2: Animal models for retinal guanylate cyclase defects</i>	10
<i>Table 3: GUCY2D qPCR</i>	29
<i>Table 4: Neomycin-resistance (neoR) qPCR</i>	30
<i>Table 5: NANOG qPCR</i>	30
<i>Table 6: Cyclor program for GUCY2D, neoR and NANOG qPCR</i>	30
<i>Table 7: USH1C qPCR</i>	30
<i>Table 8: Cyclor program for USH1C qPCR</i>	30
<i>Table 9: Humanized USH1C qPCR</i>	31
<i>Table 10: Cyclor program for humanized USH1C qPCR</i>	31
<i>Table 11: GUCY2D PCR 1</i>	31
<i>Table 12: GUCY2D PCR 2</i>	32
<i>Table 13: GUCY2D PCR 3</i>	32
<i>Table 14: Cyclor program for GUCY2D PCRs 1, 2 and 3</i>	32
<i>Table 15: GUCY2D PCR 4</i>	32
<i>Table 16: Cyclor program for GUCY2D PCR 4</i>	33
<i>Table 17: USHR31X PCR</i>	33
<i>Table 18: Cyclor program for USHR31X PCR</i>	33
<i>Table 19: Sequencing reaction</i>	33
<i>Table 20: Cyclor program for sequencing reaction</i>	34
<i>Table 21: Candidate genes for a complementary model</i>	36
<i>Table 22: Location of the BACs</i>	38
<i>Table 23: Composition of GUCY2D disruptive GE experiments</i>	52
<i>Table 24: Efficiency of GUCY2D knock-out experiments</i>	59
<i>Table 25: Use of animals in research</i>	62
<i>Table 26: Composition of USH1C reconstituting GE experiments</i>	66
<i>Table 27: Composition of prime editing experiments</i>	70

X. REFERENCES

Aigner B, Renner S, Kessler B, Klymiuk N, Kurome M, Wunsch A, Wolf E. Transgenic pigs as models for translational biomedical research. *Journal of molecular medicine* 2010; 88: 653-64.

Anzalone AV, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, Chen PJ, Wilson C, Newby GA, Raguram A, Liu DR. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 2019; 576: 149-57.

Archibald AL, Bolund L, Churcher C, Fredholm M, Groenen MA, Harlizius B, Lee KT, Milan D, Rogers J, Rothschild MF, Uenishi H, Wang J, Schook LB, Swine Genome Sequencing C. Pig genome sequence--analysis and publication strategy. *BMC Genomics* 2010; 11: 438.

Arshavsky VY, Burns ME. Photoreceptor Signaling: Supporting Vision across a Wide Range of Light Intensities *. *Journal of Biological Chemistry* 2012; 287: 1620-6.

Auch H, Klymiuk N, Runa-Vochozkova P. Modifying Bacterial Artificial Chromosomes for Extended Genome Modification. In: *Applications of Genome Modulation and Editing*. Verma PJ, Sumer H, Liu J, eds. New York, NY: Springer US 2022: 67-90.

BACPAC-Resources. <https://bacpacresources.org/library.php?id=124>. 14.06.2022.

Bae S, Park J, Kim J-S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* 2014; 30: 1473-5.

Bainbridge JWB, Smith AJ, Barker SS, Robbie S, Henderson R, Balaggan K, Viswanathan A, Holder GE, Stockman A, Tyler N, Petersen-Jones S, Bhattacharya SS, Thrasher AJ, Fitzke FW, Carter BJ, Rubin GS, Moore AT, Ali RR. Effect of

Gene Therapy on Visual Function in Leber's Congenital Amaurosis. *New England Journal of Medicine* 2008; 358: 2231-9.

Bakondi B, Lv W, Lu B, Jones MK, Tsai Y, Kim KJ, Levy R, Akhtar AA, Breunig JJ, Svendsen CN, Wang S. In Vivo CRISPR/Cas9 Gene Editing Corrects Retinal Dystrophy in the S334ter-3 Rat Model of Autosomal Dominant Retinitis Pigmentosa. *Molecular Therapy* 2016; 24: 556-63.

Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science* 2007; 315: 1709-12.

Bassuk AG, Zheng A, Li Y, Tsang SH, Mahajan VB. Precision Medicine: Genetic Repair of Retinitis Pigmentosa in Patient-Derived Stem Cells. *Scientific Reports* 2016; 6: 19969.

Bennett J. Immune response following intraocular delivery of recombinant viral vectors. *Gene Ther* 2003; 10: 977-82.

Berger W, Kloeckener-Gruissem B, Neidhardt J. The molecular basis of human retinal and vitreoretinal diseases. *Progress in Retinal and Eye Research* 2010; 29: 335-75.

Bibikova M, Golic M, Golic KG, Carroll D. Targeted Chromosomal Cleavage and Mutagenesis in *Drosophila* Using Zinc-Finger Nucleases. *Genetics* 2002; 161: 1169-75.

Birling M-C, Gofflot F, Warot X. Site-Specific Recombinases for Manipulation of the Mouse Genome. In: *Transgenesis Techniques: Principles and Protocols*. Cartwright EJ, ed. Totowa, NJ: Humana Press 2009: 245-63.

Brouns SJJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJH, Snijders APL, Dickman MJ, Makarova KS, Koonin EV, Oost Jvd. Small CRISPR RNAs Guide

Antiviral Defense in Prokaryotes. *Science* 2008; 321: 960-4.

Bujakowska KM, Liu Q, Pierce EA. Photoreceptor Cilia and Retinal Ciliopathies. *Cold Spring Harbor Perspectives in Biology* 2017; 9

Burnight ER, Gupta M, Wiley LA, Anfinson KR, Tran A, Triboulet R, Hoffmann JM, Klaahsen DL, Andorf JL, Jiao C, Sohn EH, Adur MK, Ross JW, Mullins RF, Daley GQ, Schlaeger TM, Stone EM, Tucker BA. Using CRISPR-Cas9 to Generate Gene-Corrected Autologous iPSCs for the Treatment of Inherited Retinal Degeneration. *Molecular Therapy* 2017; 25: 1999-2013.

Collery RF, Cederlund ML, Kennedy BN. Transgenic zebrafish expressing mutant human RETGC-1 exhibit aberrant cone and rod morphology. *Experimental Eye Research* 2013; 108: 120-8.

Colosimo A, Goncz KK, Holmes AR, Kunzelmann K, Novelli G, Malone RW, Bennett MJ, Gruenert DC. Transfer and Expression of Foreign Genes in Mammalian Cells. *BioTechniques* 2000; 29: 314-31.

Conant D, Hsiao T, Rossi N, Oki J, Maures T, Waite K, Yang J, Joshi S, Kelso R, Holden K, Enzmann BL, Stoner R. Inference of CRISPR Edits from Sanger Trace Data. *The CRISPR Journal* 2022; 5: 123-30.

Concordet J-P, Haeussler M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Research* 2018; 46: W242-W5.

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* 2013; 339: 819-23.

Cosgrove D, Zallocchi M. Usher protein functions in hair cells and photoreceptors. *The international journal of biochemistry & cell biology* 2014; 46: 80-9.

Dizhoor AM, Olshevskaya EV, Peshenko IV. The R838S Mutation in Retinal Guanylyl Cyclase 1 (RetGC1) Alters Calcium Sensitivity of cGMP Synthesis in the Retina and Causes Blindness in Transgenic Mice *. *Journal of Biological Chemistry* 2016; 291: 24504-16.

Dizhoor AM, Peshenko IV. Regulation of retinal membrane guanylyl cyclase (RetGC) by negative calcium feedback and RD3 protein. *Pflügers Archiv - European Journal of Physiology* 2021;

Duda T, Koch K-W, Venkataraman V, Lange C, Beyermann M, Sharma RK. Ca²⁺ sensor S100 β -modulated sites of membrane guanylate cyclase in the photoreceptor-bipolar synapse. *The EMBO Journal* 2002; 21: 2547-56.

Edwards A, Fishman GA, Anderson RJ, Grover S, Derlacki DJ. Visual Acuity and Visual Field Impairment in Usher Syndrome. *Archives of Ophthalmology* 1998; 116: 165-8.

Egerer S, Fiebig U, Kessler B, Zakhartchenko V, Kurome M, Reichart B, Kupatt C, Klymiuk N, Wolf E, Denner J, Bähr A. Early weaning completely eliminates porcine cytomegalovirus from a newly established pig donor facility for xenotransplantation. *Xenotransplantation* 2018; 25: e12449.

El-Amraoui A, Petit C. Usher I syndrome: unravelling the mechanisms that underlie the cohesion of the growing hair bundle in inner ear sensory cells. *Journal of Cell Science* 2005; 118: 4593-603.

ENSEMBL. <http://www.ensembl.org/index.html>. 17.6.22.

FDA. approval letter voretigene neparvovec-rzyl. 2017: <https://www.fda.gov/media/109487/download>. 01.04.2022.

Fercher AF, Drexler W, Hitzenberger CK, Lasser T. Optical coherence tomography - principles and applications. *Reports on Progress in Physics* 2003; 66: 239-303.

Fishman GA, Kumar A, Joseph ME, Torok N, Anderson RJ. Usher's Syndrome: Ophthalmic and Neuro-otologic Findings Suggesting Genetic Heterogeneity. *Archives of Ophthalmology* 1983; 101: 1367-74.

Galvin JA, Fishman GA, Stone EM, Koenekoop RK. Clinical phenotypes in carriers of Leber congenital amaurosis mutations. *Ophthalmology* 2005; 112: 349-56.

Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, Liu DR. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* 2017; 551: 464.

Giarmarco MM, Cleghorn WM, Sloat SR, Hurley JB, Brockerhoff SE. Mitochondria Maintain Distinct Ca^{2+} Pools in Cone Photoreceptors. *The Journal of Neuroscience* 2017; 37: 2061-72.

Giraldo P, Montoliu L. Size Matters: Use of YACs, BACs and PACs in Transgenic Animals. *Transgenic Research* 2001; 10: 83-103.

Gregory-Evans K, Kelsell RE, Gregory-Evans CY, Downes SM, Fitzke FW, Holder GE, Simunovic M, Mollon JD, Taylor R, Hunt DM, Bird AC, Moore AT. Autosomal dominant cone-rod retinal dystrophy (CORD6) from heterozygous mutation of GUCY2D, which encodes retinal guanylate cyclase 1 | The authors have no proprietary interests in the materials mentioned in the study. *Ophthalmology* 2000; 107: 55-61.

Grotz S (2021) Characterization of a pig model of Usher syndrome. In: Tierärztliche Fakultät. Ludwig-Maximilian-Universität München, Munich

Grotz S, Schäfer J, Wunderlich KA, Ellederova Z, Auch H, Bähr A, Runa-Vochozkova P, Fadl J, Arnold V, Ardan T, Veith M, Santamaria G, Dhom G, Hitzl W, Kessler B, Eckardt C, Klein J, Brymova A, Linnert J, Kurome M, Zakharchenko V, Fischer A, Blutke A, Döring A, Suchankova S, Popelar J, Rodríguez-Bocanegra E, Długaiczek J, Straka H, May-Simera H, Wang W, Laugwitz K-L, Vandenberghe LH, Wolf E, Nagel-Wolfrum K, Peters T, Motlik J, Fischer MD, Wolfrum U,

Klymiuk N. Early disruption of photoreceptor cell architecture and loss of vision in a humanized pig model of usher syndromes. *EMBO Molecular Medicine* 2022; n/a: e14817.

Gupta PR, Huckfeldt RM. Gene therapy for inherited retinal degenerations: initial successes and future challenges. *Journal of Neural Engineering* 2017; 14: 051002.

Hartong DT, Berson EL, Dryja TP. Retinitis pigmentosa. *The Lancet* 2006; 368: 1795-809.

Hood DC, Bach M, Brigell M, Keating D, Kondo M, Lyons JS, Marmor MF, McCulloch DL, Palmowski-Wolfe AM, For the International Society For Clinical Electrophysiology of V. ISCEV standard for clinical multifocal electroretinography (mfERG) (2011 edition). *Documenta Ophthalmologica* 2012; 124: 1-13.

Hörmanseder E, Simeone A, Allen GE, Bradshaw CR, Figlmüller M, Gurdon J, Jullien J. H3K4 Methylation-Dependent Memory of Somatic Cell Identity Inhibits Reprogramming and Development of Nuclear Transfer Embryos. *Cell Stem Cell* 2017; 21: 135-43.e6.

Hryhorowicz M, Lipiński D, Hryhorowicz S, Nowak-Terpiłowska A, Ryczek N, Zeyland J. Application of Genetically Engineered Pigs in Biomedical Research. *Genes* 2020; 11: 670.

Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O, Cradick TJ, Marraffini LA, Bao G, Zhang F. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology* 2013; 31: 827-32.

Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 2014; 157: 1262-78.

Hu ML, Edwards TL, O'Hare F, Hickey DG, Wang J-H, Liu Z, Ayton LN. *Gene*

therapy for inherited retinal diseases: progress and possibilities. *Clinical and Experimental Optometry* 2021; 104: 444-54.

Jao L, Wenthe SR, Chen W. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proceedings of the National Academy of Sciences* 2013; 110: 13904-9.

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: 816-21.

Kellsell RE, Gregory-Evans K, Payne AM, Perrault I, Kaplan J, Yang R-B, Garbers DL, Bird AC, Moore AT, Hunt DM. Mutations in the Retinal Guanylate Cyclase (RETGC-1) Gene in Dominant Cone-Rod Dystrophy. *Human Molecular Genetics* 1998; 7: 1179-84.

Kitiratschky VBD, Wilke R, Renner AB, Kellner U, Vadalà M, Birch DG, Wissinger B, Zrenner E, Kohl S. Mutation Analysis Identifies GUCY2D as the Major Gene Responsible for Autosomal Dominant Progressive Cone Degeneration. *Investigative Ophthalmology & Visual Science* 2008; 49: 5015-23.

Klymiuk N, Mundhenk L, Kraehe K, Wuensch A, Plog S, Emrich D, Langenmayer M, Stehr M, Holzinger A, Kröner C. Sequential targeting of CFTR by BAC vectors generates a novel pig model of cystic fibrosis. *Journal of molecular medicine* 2012; 90: 597-608.

Klymiuk N, Blutke A, Graf A, Krause S, Burkhardt K, Wuensch A, Krebs S, Kessler B, Zakhartchenko V, Kurome M, Kemter E, Nagashima H, Schoser B, Herbach N, Blum H, Wanke R, Aartsma-Rus A, Thirion C, Lochmüller H, Walter MC, Wolf E. Dystrophin-deficient pigs provide new insights into the hierarchy of physiological derangements of dystrophic muscle. *Human Molecular Genetics* 2013; 22: 4368-82.

Klymiuk N, Seeliger F, Bohlooly-Y M, Blutke A, Rudmann DG, Wolf E. Tailored

Pig Models for Preclinical Efficacy and Safety Testing of Targeted Therapies. *Toxicologic Pathology* 2015; 44: 346-57.

Koblan LW, Erdos MR, Wilson C, Cabral WA, Levy JM, Xiong Z-M, Tavares UL, Davison LM, Gete YG, Mao X, Newby GA, Doherty SP, Narisu N, Sheng Q, Krilow C, Lin CY, Gordon LB, Cao K, Collins FS, Brown JD, Liu DR. In vivo base editing rescues Hutchinson–Gilford progeria syndrome in mice. *Nature* 2021; 589: 608-14.

Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 2016; 533: 420-4.

Komor AC, Badran AH, Liu DR. Editing the Genome Without Double-Stranded DNA Breaks. *ACS Chemical Biology* 2018; 13: 383-8.

Kostic C, Lillico SG, Crippa SV, Grandchamp N, Pilet H, Philippe S, Lu Z, King TJ, Mallet J, Sarkis C, Arsenijevic Y, Whitelaw CBA. Rapid Cohort Generation and Analysis of Disease Spectrum of Large Animal Model of Cone Dystrophy. *PLOS ONE* 2013; 8: e71363.

Kumar N, Gammell P, Clynes M. Proliferation control strategies to improve productivity and survival during CHO based production culture. *Cytotechnology* 2007; 53: 33-46.

Kurome M, Geistlinger L, Kessler B, Zakhartchenko V, Klymiuk N, Wuensch A, Richter A, Baehr A, Kraehe K, Burkhardt K, Flisikowski K, Flisikowska T, Merkl C, Landmann M, Durkovic M, Tschukes A, Kraner S, Schindelbauer D, Petri T, Kind A, Nagashima H, Schnieke A, Zimmer R, Wolf E. Factors influencing the efficiency of generating genetically engineered pigs by nuclear transfer: multi-factorial analysis of a large data set. *BMC Biotechnology* 2013; 13: 43.

Kurome M, Kessler B, Wuensch A, Nagashima H, Wolf E. Nuclear Transfer and Transgenesis in the Pig. In: *Nuclear Reprogramming: Methods and Protocols*.

Beaujean N, Jammes H, Jouneau A, eds. New York, NY: Springer New York 2015: 37-59.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann Y, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blocker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kasprzyk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A,

Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowki J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ, Szustakowki J, International Human Genome Sequencing C. Initial sequencing and analysis of the human genome. *Nature* 2001; 409: 860-921.

Lanza AM, Kim DS, Alper HS. Evaluating the influence of selection markers on obtaining selected pools and stable cell lines in human cells. *Biotechnology Journal* 2013; 8: 811-21.

Laura RP, Dizhoor AM, Hurley JB. The Membrane Guanylyl Cyclase, Retinal Guanylyl Cyclase-1, Is Activated through Its Intracellular Domain (2017);. *Journal of Biological Chemistry* 1996; 271: 11646-51.

Ledford H. Titanic clash over CRISPR patents turns ugly. *Nature* 2016; 537: 460-1.

Ledford H. Major CRISPR patent decision won't end tangled dispute. *Nature* 2022; 603: 373-4.

Li D, Qiu Z, Shao Y, Chen Y, Guan Y, Liu M, Li Y, Gao N, Wang L, Lu X, Zhao Y, Liu M. Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. *Nature Biotechnology* 2013; 31: 681-3.

Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biology Direct* 2006; 1: 7.

Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. RNA-guided human genome engineering via Cas9. *Science (New York, N.Y.)* 2013; 339: 823-6.

Mathur P, Yang J. Usher syndrome: Hearing loss, retinal degeneration and associated abnormalities. *Biochimica et biophysica acta* 2015; 1852: 406-20.

McCullough. Somatic Gene Editing of GUCY2D by AAV-CRISPR/Cas9 Alters Retinal Structure and Function in Mouse and Macaque. *Human Gene Therapy* 2019; 30: 571-89.

Moretti A, Fonteyne L, Giesert F, Hoppmann P, Meier AB, Bozoglu T, Baehr A, Schneider CM, Sinnecker D, Klett K, Fröhlich T, Rahman FA, Haufe T, Sun S, Jurisch V, Kessler B, Hinkel R, Dirschinger R, Martens E, Jilek C, Graf A, Krebs S, Santamaria G, Kurome M, Zakhartchenko V, Campbell B, Voelse K, Wolf A, Ziegler T, Reichert S, Lee S, Flenkenthaler F, Dorn T, Jeremias I, Blum H, Dendorfer A, Schnieke A, Krause S, Walter MC, Klymiuk N, Laugwitz KL, Wolf E, Wurst W, Kupatt C. Somatic gene editing ameliorates skeletal and cardiac muscle failure in pig and human models of Duchenne muscular dystrophy. *Nature Medicine* 2020; 26: 207-14.

NCBI. <https://www.ncbi.nlm.nih.gov/>. 17.6.22.

Newby GA, Yen JS, Woodard KJ, Mayuranathan T, Lazzarotto CR, Li Y, Sheppard-Tillman H, Porter SN, Yao Y, Mayberry K, Everette KA, Jang Y, Podracky CJ, Thaman E, Lechauve C, Sharma A, Henderson JM, Richter MF, Zhao KT, Miller SM, Wang T, Koblan LW, McCaffrey AP, Tisdale JF, Kalfa TA, Pruett-Miller SM, Tsai SQ, Weiss MJ, Liu DR. Base editing of haematopoietic stem cells rescues sickle cell disease in mice. *Nature* 2021; 595: 295-302.

Niu Y, Shen B, Cui Y, Chen Y, Wang J, Wang L, Kang Y, Zhao X, Si W, Li W, Xiang Andy P, Zhou J, Guo X, Bi Y, Si C, Hu B, Dong G, Wang H, Zhou Z, Li T, Tan T, Pu X, Wang F, Ji S, Zhou Q, Huang X, Ji W, Sha J. Generation of Gene-Modified Cynomolgus Monkey via Cas9/RNA-Mediated Gene Targeting in One-Cell Embryos. *Cell* 2014; 156: 836-43.

OMIM. <https://www.ncbi.nlm.nih.gov/omim>. 17.6.22.

Paquet D, Kwart D, Chen A, Sproul A, Jacob S, Teo S, Olsen KM, Gregg A, Noggle S, Tessier-Lavigne M. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* 2016; 533: 125.

Patel S, Ryals RC, Weller KK, Pennesi ME, Sahay G. Lipid nanoparticles for delivery of messenger RNA to the back of the eye. *Journal of Controlled Release* 2019; 303: 91-100.

Perrault I, Rozet JM, Calvas P, Gerber S, Camuzat A, Dollfus H, Châtelin S, Souied E, Ghazi I, Leowski C, Bonnemaïson M, Paslier DL, Frézal J, Dufier J-L, Pittler S, Munnich A, Kaplan J. Retinal-specific guanylate cyclase gene mutations in Leber's congenital amaurosis. *Nature Genetics* 1996; 14: 461-4.

Perrault I, Rozet J-M, Gerber S, Ghazi I, Ducroq D, Souied E, Leowski C, Bonnemaïson M, Dufier J-L, Munnich A, Kaplan J. Spectrum of retGC1 mutations in Leber's congenital amaurosis. *European Journal of Human Genetics* 2000; 8: 578-82.

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nature protocols* 2013; 8: 2281.

Reiners J, Nagel-Wolfrum K, Jürgens K, Märker T, Wolfrum U. Molecular basis of human Usher syndrome: Deciphering the meshes of the Usher protein network provides insights into the pathomechanisms of the Usher disease. *Experimental Eye Research* 2006; 83: 97-119.

Reiners J, Wolfrum U (2006) Molecular Analysis of the Supramolecular Usher Protein Complex in the Retina. *Retinal Degenerative Diseases, 2006*//. Boston, MA. 349-53

Richardson CD, Ray GJ, DeWitt MA, Curie GL, Corn JE. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nature Biotechnology* 2016; 34: 339.

Salehi Chaleshtori AR, Garshasbi M, Salehi A. A novel deletion mutation in GUCY2D gene may be responsible for Leber congenital amaurosis-1 disease: A case report. *Journal of Current Ophthalmology* 2019; 31: 458-62.

Sanchez I, Martin R, Ussa F, Fernandez-Bueno I. The parameters of the porcine eyeball. *Graefe's Archive for Clinical and Experimental Ophthalmology* 2011; 249: 475-82.

Sato S, Peshenko IV, Olshevskaya EV, Kefalov VJ, Dizhoor AM. GUCY2D Cone-Rod Dystrophy-6 Is a "Phototransduction Disease" Triggered by Abnormal Calcium Feedback on Retinal Membrane Guanylyl Cyclase 1. *The Journal of Neuroscience* 2018; 38: 2990-3000.

Scully R, Panday A, Elango R, Willis NA. DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nature Reviews Molecular Cell Biology* 2019; 20: 698-714.

Sharon D, Wimberg H, Kinarty Y, Koch K-W. Genotype-functional-phenotype correlations in photoreceptor guanylate cyclase (GC-E) encoded by GUCY2D. *Progress in Retinal and Eye Research* 2018; 63: 69-91.

Shizuya H, Birren B, Kim UJ, Mancino V, Slepak T, Tachiiri Y, Simon M. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc Natl Acad Sci U S A* 1992; 89: 8794-7.

Shrivastav M, De Haro LP, Nickoloff JA. Regulation of DNA double-strand break repair pathway choice. *Cell Research* 2008; 18: 134-47.

Slijkerman RWN, Song F, Astuti GDN, Huynen MA, van Wijk E, Stieger K, Collin RWJ. The pros and cons of vertebrate animal models for functional and therapeutic research on inherited retinal dystrophies. *Progress in Retinal and Eye Research* 2015; 48: 137-59.

Song H, Chung S-K, Xu Y. Modeling Disease in Human ESCs Using an Efficient BAC-Based Homologous Recombination System. *Cell Stem Cell* 2010; 6: 80-9.

Tang W, Hu JH, Liu DR. Aptazyme-embedded guide RNAs enable ligand-responsive genome editing and transcriptional activation. *Nature Communications* 2017; 8: 15939.

Toms M, Bitner-Glindzicz M, Webster A, Moosajee M. Usher syndrome: a review of the clinical phenotype, genes and therapeutic strategies. *Expert Review of Ophthalmology* 2015; 10: 241-56.

Toms M, Pagarkar W, Moosajee M. Usher syndrome: clinical features, molecular genetics and advancing therapeutics. *Therapeutic Advances in Ophthalmology* 2020a; 12: 2515841420952194.

Tornabene P, Trapani I, Minopoli R, Centrulo M, Lupo M, Simone Sd, Tiberi P, Dell'Aquila F, Marrocco E, Iodice C, Iuliano A, Gesualdo C, Rossi S, Giaquinto L, Albert S, Hoyng CB, Polishchuk E, Cremers FPM, Surace EM, Simonelli F, Matteis MAD, Polishchuk R, Auricchio A. Intein-mediated protein trans-splicing expands adeno-associated virus transfer capacity in the retina. *Science Translational Medicine* 2019; 11: eaav4523.

Truong D-JJ, Kühner K, Kühn R, Werfel S, Engelhardt S, Wurst W, Ortiz O. Development of an intein-mediated split-Cas9 system for gene therapy. *Nucleic Acids Research* 2015; 43: 6450-8.

Tsai SQ, Joung JK. Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. *Nature Reviews Genetics* 2016; 17: 300-12.

Tucker CL, Woodcock SC, Kelsell RE, Ramamurthy V, Hunt DM, Hurley JB. Biochemical analysis of a dimerization domain mutation in RetGC-1 associated with dominant cone-rod dystrophy. *Proceedings of the National Academy of Sciences* 1999; 96: 9039-44.

Usher CH. On the inheritance of retinitis pigmentosa; with note of cases. *The Royal London Ophthalmic Hospital Reports* 1914; 19: 130-236.

Valera A, Perales JC, Hatzoglou M, Bosch F. Expression of the Neomycin-Resistance (neo) Gene Induces Alterations in Gene Expression and Metabolism. *Human Gene Therapy* 1994; 5: 449-56.

Vandenberghe LH, Auricchio A. Novel adeno-associated viral vectors for retinal gene therapy. *Gene Ther* 2012; 19: 162-8.

Vochozkova P, Simmet K, Jemiller E-M, Wünsch A, Klymiuk N. Gene Editing in Primary Cells of Cattle and Pig. In: *CRISPR Gene Editing: Springer* 2019: 271-89.

Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R. One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. *Cell* 2013; 153: 910-8.

Warming S, Costantino N, Court DL, Jenkins NA, Copeland NG. Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Research* 2005; 33: e36-e.

Wu W-H, Tsai Y-T, Justus S, Lee T-T, Zhang L, Lin C-S, Bassuk AG, Mahajan VB, Tsang SH. CRISPR Repair Reveals Causative Mutation in a Preclinical Model of Retinitis Pigmentosa. *Molecular Therapy* 2016; 24: 1388-94.

Xu J, Morris L, Thapa A, Ma H, Michalakis S, Biel M, Baehr W, Peshenko IV, Dizhoor AM, Ding X-Q. cGMP Accumulation Causes Photoreceptor Degeneration in CNG Channel Deficiency: Evidence of cGMP Cytotoxicity Independently of Enhanced CNG Channel Function. *The Journal of Neuroscience* 2013; 33: 14939-48.

Yeh W-H, Chiang H, Rees HA, Edge ASB, Liu DR. In vivo base editing of post-mitotic sensory cells. *Nature Communications* 2018; 9: 2184.

Zanazzi G, Matthews G. The Molecular Architecture of Ribbon Presynaptic Terminals. *Molecular Neurobiology* 2009; 39: 130-48.

Zhang Y, Buchholz F, Muyrers JPP, Stewart AF. A new logic for DNA engineering using recombination in *Escherichia coli*. *Nature Genetics* 1998; 20: 123-8.

XI. APPENDIX

Publication: Modifying Bacterial Artificial Chromosomes for Extended Genome Modification

Hannah Auch, Nikolai Klymiuk, and Petra Runa-Vochozkova

Methods in Molecular Biology book series (Volume 2495: Applications of Genome Modulation and Editing), 13th of June, 2022.



Chapter 4

Modifying Bacterial Artificial Chromosomes for Extended Genome Modification

Hannah Auch, Nikolai Klymiuk, and Petra Runa-Vochozkova

Abstract

Bacterial artificial chromosomes have been used extensively for the exploration of mammalian genomes. Although novel approaches made their initial function expendable, the available BAC libraries are a precious source for life science. Their comprising of extended genomic regions provides an ideal basis for creating a large targeting vector. Here, we describe the identification of suitable BACs from their libraries and their verification prior to manipulation. Further, protocols for modifying BAC, confirming the desired modification and the preparation of transfection into mammalian cells are given.

Key words Instructions: Bacterial artificial chromosomes (BAC), Bacterial recombination, BAC libraries, BAC fingerprinting, BAC Sanger sequencing, Genome modification

1 Introduction

Bacterial artificial chromosomes (BACs) are huge plasmids comprising pieces of 150–300 kb from vertebrate genomes [1]. BACs have been initially developed for the generation of genomic libraries in the early ages of full genome exploration. Such libraries have been produced by partial digestion of genomic DNA with restriction enzymes and cloning of these fragments into a plasmid backbone of approximately 20 kb. The BACs thus comprised overlapping inserts and systematic restriction enzyme digestions of all BACs from a library allowed their ordered assembling and finally the construction of a chromosome. Reference genomes, such as for human, have then been generated by shot-gun sequencing of BACs with minimal overlap [2–4]. Next-generation sequencing approaches were based on shot-gun sequencing of the whole genome [5, 6] avoiding the intermediate step of BAC libraries, but the assembly of these short sequence reads required the previously defined genomic structure. Present-day sequencing approaches generate extremely large contigs [7, 8] facilitating de novo genome

exploration and making the once precious BAC libraries dispensable. BAC libraries (Table 1) are, however, still accessible and by carrying large genomic regions they provide a source for advanced experimental genetic modification strategies in life science.

In contrast to classical targeting or gene editing by CRISPR/Cas and short single-stranded oligo-deoxy-nucleotides (ssODN), the large genomic segments in BACs allow either the introduction of huge modifications into a genome or provide extended homologous arms (HA) for homologous recombination (HR), which is often helpful if well-characterized target genomes such as in stem cells are missing. The availability of different bacterial recombination systems such as the RecET system [9, 10] or λ Red system [11, 12] promotes the effective manipulation of BACs in *E. coli* and facilitates an unrestricted portfolio of shaping BACs for genetic modification in vertebrate cells.

We have gained extensive experience in tailoring porcine BACs for generating pig models for biomedical research, but the protocol below describes a very general procedure of preparing a BAC as a targeting vector for the site-specific introduction of exogenous DNA into the mammalian genome. For the complexity and multiplicity of a possible modification, however, we cannot refer to the design and construction of the desired modification. Rather, we generally speak about a “modification of interest” (MOI) which shall be representative for any approach on reporter genes, fusion genes, gene mutations, etc. Importantly, for genome manipulation of vertebrate cells, BAC vectors can be combined with gene-editing tools to promote site-directed HR. For gene editing in porcine and bovine somatic cells and their preparation for somatic cell nuclear transfer, we would like to refer to [13].

2 Materials

2.1 Design and Construct a Desired Modification of Interest

1. Vector pL451/pL452 (Addgene).
2. Electrocompetent *E. coli* strain DH10B.
3. Kanamycin (storage concentration 25 mg/mL, working concentration 1000 \times diluted).
4. Restriction enzymes with an appropriate buffer.
5. T4 ligase and ligation buffer.
6. Calf Intestinal Alkaline Phosphatase (CIAP).
7. General equipment for agarose gel electrophoresis: Universal agarose powder (BioSell), Bromophenol blue, 1 kb DNA ladder (GeneRuler 1 kb DNA Ladder), GelRed, 50 \times TAE buffer (2 M Tris base, 50 mM EDTA, 1 M acetic acid) storage solution and 1 \times working solution, agarose gel electrophoresis system, Imager for gel documentation UVP GelStudio Plus (Analytic Jena).

Table 1
BAC libraries

Library	Source	Digest	Average insert size	Coverage	Vector ^a	Constructed by	Distributor
Human	CHORI-17	<i>EcoRI</i>	197 kb	11.3×	pBACGKI.1	BACPAC ^b	BACPAC ^b
	CaITech-D	<i>HindIII</i>	129 kb	17×	pBeloBAC11	CaITech ^c	Invitrogen ^d
	RPCI-11	<i>EcoRI</i> , <i>MhoI</i>	178 kb	32.2×	pBACe3.6, pTARBAC1	[17]	BACPAC ^b
Porcine	CHORI-242	<i>MhoI</i>	173 kb	11.4×	pTARBAC1.3	BACPAC ^b	BACPAC ^b
	SBAB	Fibroblasts	135 kb	5×	pBeloBAC11	[18]	INRA ^e
	RPCI-44	<i>EcoRI</i>	165 kb	10.2×	pTARBAC2	[19]	BACPAC ^b
Bovine	CHORI-240	<i>MhoI</i>	167 kb	10.7×	pTARBAC1.3	BACPAC ^b	BACPAC ^b
	RPCI-42	<i>EcoRI</i>	164 kb	11.9×	pBACe3.6	BACPAC ^b	BACPAC ^b
	BtINRA	Fibroblasts	120 kb	4×	pBeloBAC11	[20]	INRA ^e
Sheep	CHORI-243	<i>EcoRI</i>	184 kb	5.4–14× ^f	pTARBAC2.1	[21]	BACPAC ^b

^a In all backbones, resistance is provided against *cam*, the integration site is flanked by SP6 and T7 primer-binding sites, NotI and *AscI* sites for linearization are provided.

^b Pieter de Jong's laboratory at BACPAC Resources, Children's Hospital Oakland Research Institute (<https://bacpacresources.org/>)

^c California Institute of Technology

^d Invitrogen is shipping the BAC clones as a glycerol stock of bacteria (<https://www.thermofisher.com/de/en/home/life-science/cloning/clone-collections.html>)

^e The BAC-YAC Resource Center of the INRA (<http://dga.jouy.inra.fr/grafra/>)

^f Different information is given in the original publication and an update at <https://www.sheepmap.org/bes.php>

2.2 Verification of BAC Clone Genomic Location

2.2.1 Long-Term Storage of the Received BAC Clones

1. BAC clone in *E. coli* strain DH10B (BACPAC Resources at Children's Hospital Oakland Research Institute).
2. LB-medium: 2% tryptone, 0.5% yeast extract, 10 mM NaCl, ddH₂O to 1000 mL, autoclaved, then supplemented with chloramphenicol (cam, storage concentration 12.5 mg/mL in EtOH, working concentration 1000× diluted).
3. 60% glycerol, sterile filtrated (pores 0.22 μm).
4. Culture tubes 12 mL.
5. Cryo vials 1.5 mL.
6. Shaking incubator GFL 3031 with orbital motion.

2.2.2 Isolation of BAC Clones by Basic Alkaline Lysis Plasmid Miniprep

1. LB-medium: 2% tryptone, 0.5% yeast extract, 10 mM NaCl, ddH₂O to 1000 mL, autoclaved, then supplemented with cam (storage concentration 12.5 mg/mL in EtOH, working concentration 1000× diluted).
2. STE: 10 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA/NaOH pH 8.0.
3. Plasmid A: 50 mM Glucose, 25 mM Tris-HCl pH 8.0, 10 mM EDTA/NaOH pH 8.0.
4. Plasmid B (always prepare freshly): 0.1 M NaOH, 0.5% SDS.
5. Plasmid C: 3 M KOAc pH 4.8 with 9 M HOAc.
6. RNase A (20 mg/mL).
7. PCiA: phenol, chloroform, isoamyl alcohol (25:24:1), store at 4 °C max. 2 months.
8. Isopropanol (iPrOH).
9. Ethanol (EtOH).
10. T-Buffer: 10 mM Tris-HCl, pH 8.0.
11. Culture tubes 12 mL.
12. Shaking incubator GFL 3031 with orbital motion.
13. Centrifuge with swing rotor (Eppendorf, 5910R, 4× Universal).
14. Nanodrop (SimpliNano spectrophotometer, Biochrom).

2.2.3 BAC Clones Fingerprinting

1. General equipment for agarose gel electrophoresis (Midi gel chamber: electrodes in distance of approx. 19 cm and gel box 12 × 14 cm with 12 wells comb).
2. Restriction enzymes with an appropriate buffer.
3. Thermoblock or incubator for 37 °C.

2.2.4 BAC-End Sequencing

1. PEG-MgCl₂: 40% PEG 8000, 30 mM MgCl₂.
2. Commercially synthesized primers.
3. BigDye Terminator v3.1 Sequencing Kit.

4. 125 mM EDTA pH 8.0.
5. EtOH.
6. Thermal cycler.
7. Refrigerated centrifuge (Eppendorf, 5910R, FA-48 × 2).

2.3 Modification of BAC Clone by Bacterial Recombineering

2.3.1 Making Electrocompetent Cells

1. LB-medium: 2% tryptone, 0.5% yeast extract, 10 mM NaCl, ddH₂O to 1000 mL, autoclaved, then supplemented with cam (storage concentration 12.5 mg/mL in EtOH, working concentration 1000× diluted).
2. *E. coli* SW strain (SW102, SW105, SW106 from National Cancer Institute, Frederick, USA).
3. 10% glycerol, sterile filtrated (pores 0.22 μm).
4. ddH₂O.
5. 500 mL Erlenmeyer flask.
6. Glass cuvettes.
7. 50 mL falcon tubes with round bottom.
8. Shaking Incubator GFL 3031 with orbital motion.
9. Photometer GeneQuant Pro (Amersham Biosciences).
10. Centrifuge with fixed angle rotor (Eppendorf, 5910R, rotor FA-6 × 50).

2.3.2 Transfer BAC Clone into Bacteria

1. SOB medium: 2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, ddH₂O to 1000 mL, autoclaved.
2. Petri dishes coated with LB-agar supplemented with appropriate antibiotic.
3. Electroporation cuvettes, gap width 1 mm.
4. Incubator for 32 °C.
5. Electroporator (Eppendorf Eporator).

2.3.3 Preparation of Modification as a Linearized dsDNA Fragment

1. Restriction enzyme with appropriate buffer.
2. Low melting agarose powder (Low Melting Point Agarose, Thermo Scientific).
3. General equipment for agarose gel electrophoresis.
4. Agarase.
5. PCiA: phenol, chloroform, isoamyl alcohol (25:24:1), store at 4 °C max. 2 months.
6. 3 M NaAOc, pH 5.2.
7. EtOH.
8. T-Buffer: 10 mM Tris-HCl, pH 8.0.
9. DNA extraction kit (Double Pure Combi Kit, BioSell).

72 Hannah Auch et al.

10. Nanodrop (SimpliNano spectrophotometer, Biochrom).
11. Refrigerated centrifuge (Eppendorf, 5910R, FA-48 × 2).
12. Thermoblock.

**2.3.4 BAC
Recombineering**

See Subheadings [2.3.1.](#) and [2.3.2.](#)

**2.3.5 Arabinose-Induced
FLPE and Cre
Recombination**

1. LB-medium: 2% tryptone, 0.5% yeast extract, 10 mM NaCl, ddH₂O to 1000 mL, autoclaved, then supplemented with cam (storage concentration 12.5 mg/mL in EtOH, working concentration 1000× diluted).
2. Arabinose.
3. Culture Tubes 12 mL.
4. Glass cuvettes.
5. Shaking Incubator GFL 3031 with orbital motion.
6. Photometer GeneQuant Pro (Amersham Biosciences).

**2.4 Verification of
BAC Recombinant**

**2.4.1 Confirmation of
Recombinants by End-Point
PCR**

1. Commercially synthesized primers.
2. 2 M dNTPs.
3. Herculase II Fusion DNA Polymerase Kit.
4. T-Buffer: 10 mM Tris-HCl, pH 8.0.
5. Thermal cycler.
6. General equipment for agarose gel electrophoresis system.

**2.4.2 Confirmation of
Recombinants by BAC
Fingerprinting**

See Subheading [2.2.3](#)

**2.4.3 Confirmation of
Recombinants by Sanger
Sequencing**

See Subheading [2.2.4](#)

**2.5 Preparing of BAC
Targeting Vector for
Nucleofection into
Cells**

1. LB-medium: 2% tryptone, 0.5% yeast extract, 10 mM NaCl, ddH₂O to 1000 mL, autoclaved, then supplemented with appropriate antibiotic.
2. BAC DNA isolation kit (Large Construct Kit, QIAgen).
3. 3 M NaOAc, pH 5.2.
4. PCiA.
5. iPrOH.
6. EtOH.
7. TE Buffer: 10 mM Tris-HCl pH 8.0, 1 mM EDTA/NaOH pH 8.0.

8. 500 mL Erlenmeyer flask.
9. Refrigerated centrifuge with fixed angle rotors (Eppendorf, 5910R, FA-6 × 50 and FA-48 × 2).
10. Shaking Incubator GFL 3031 with orbital motion.
11. Nanodrop (SimpliNano spectrophotometer, Biochrom).

3 Methods

3.1 Design and Construct a Desired Modification of Interest (MOI)

In principle, any modification of a BAC can be defined as the introduction of an MOI into the genomic segment by bacterial recombination (Fig. 1b). Bacterial recombination without positive selection is theoretically possible, but rather ineffective. As most approaches in mammalian cells require positive selection as well, the usage of a combinatorial positive selection cassette is recommended (*see Note 1*). The opportunity of bacterial recombination and huge genomic regions in BACs provide an unlimited portfolio of potential modifications. For limitations of space and the focus on BAC modification protocols, we can only give very superficial advice on designing and constructing the MOI.

1. Design the modification and assemble it in silico (*see Note 2*). The modification vector in principle comprises the MOI itself, resistance cassette, 5' and 3' HAs complementary to an integration site in BAC clone (*see Note 3*), and additional features such a restriction sites or lox/FRT sites (*see illustration in Fig. 1a, b*).
2. Make available the necessary components (*see Note 4*).
3. Assemble the final modification plasmid from all designed parts, either obtained from gene synthesis or plasmid cloning.
4. Verify the modification plasmid (*see Note 5*).

3.2 Finding BAC Clones Covering the Target Region

Exploring a BAC map is necessary for identifying appropriate BAC clones covering the desired target region. While in the past distinct sources were provided for searching BAC maps, many BAC libraries are available meanwhile in the National Center for Biotechnology Information (NCBI) database (<https://www.ncbi.nlm.nih.gov/>). For this reason, we describe the usage of this flexible resource, while other or similar procedures might work on alternative online or off-line BAC map viewer programs. Depending on the coverage of the BAC library, a number of clones can be selected for a target region of interest. For matters of time and efficacy, we recommend purchasing 3 different BAC clones for a given target region and select one of them upon comprehensive verification for the MOI.

1. Choose the Genome Data Viewer at the NCBI web page (<https://www.ncbi.nlm.nih.gov/genome/gdv/>).

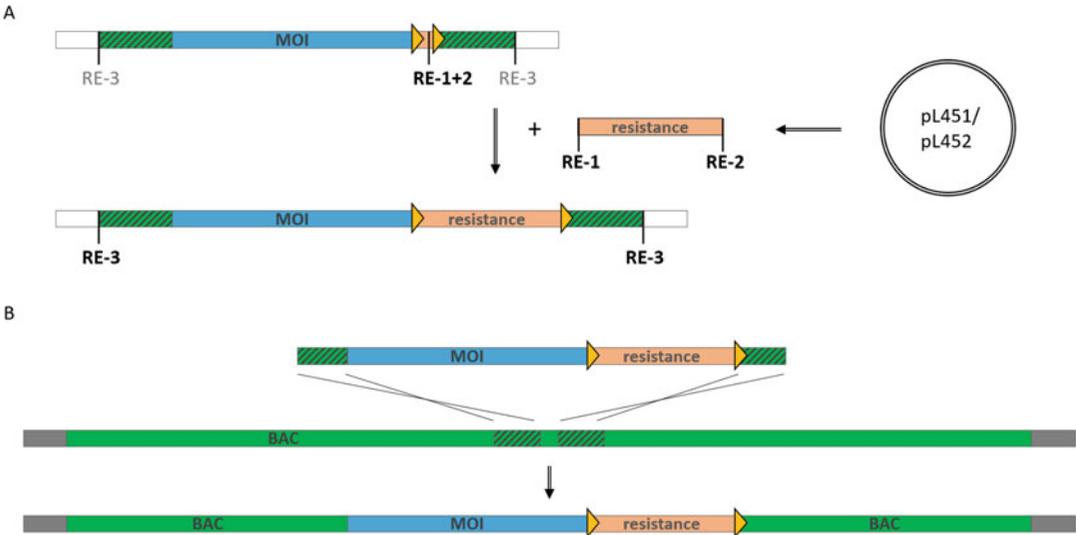


Fig. 1 Schematic illustration of BAC clone modification. **(a)** Assembly of modification vector carrying a desired modification of interest (MOI) in one cloning step. The MOI, homology arms (striped boxes) complementary to target genomic region in BAC clone, lox/FRT sites (arrowheads), and sites for restriction enzymes are synthesized in vitro into standard plasmid (white boxes). The resistance cassette is excised from pL451 or pL452 vector and cloned between lox/FRT sites in pUC vector by restriction enzymes, RE1 and RE2. The sites for restriction enzyme (RE-3) creating the modification fragment used for bacterial recombination into the BAC clone are placed at the ends of modification. **(b)** Introduction of the MOI into the BAC clone. The modification fragment carrying MOI and a resistance cassette with lox/FRT sites (arrowheads) is recombined into the BAC clone via short homology arms (striped boxes)

2. Choose your desired species in the “search organism” window and your target region in the “search in the genome” field (*see Note 6*) and browse the genome.
3. Upon direction to the desired target region, you can zoom out or into the genomic location with the scale bar or move the shown region by the arrows in the control bar.
4. In the same control bar, there is a “Tracks” button with which you can “Configure Tracks” in a newly opened window. By choosing “Genomic Clones” you may select the preferred libraries and confirm by clicking “Configure”.
5. The Genome Data Viewer now contains the BAC clones covering your target region of interest and shows their approximate size (Fig. 2a).

3.3 Verification of BAC Clone Genomic Location

Our preferred source for BAC clones is BACPAC Genomics (Emeryville, CA, USA), formerly BACPAC Resources at Children’s Hospital Oakland Research Institute, as they provide BAC clones from most of the libraries explorably at NCBI (Table 1). BAC clones are provided in the *E. coli* strain DH10B as an LB-agar stab culture, facilitating shipment at room temperature

(RT) around the globe. Incubation temperature is always at 37 °C, the antibiotic resistance encoded on the BAC backbone is, to the best of our knowledge, always *cam*. For this, we simplify the general term “the appropriate antibiotic” with “*cam*”. Upon arrival, each BAC clone should be prepared for long-term storage before it is verified for its identity.

3.3.1 Long-Term Storage of the Received BAC Clones

1. From each stab culture, create a single cells streak on an LB-*cam* agar plate and incubate plates overnight (o/N).
2. Pick 4 single bacterial clones and inoculate them into a 2.5 mL LB-*cam* medium.
3. Shake at 180 rpm up to 16 h.
4. Transfer 900 µL o/N-inoculum of each clone into a cryo vial, add 300 µL of 60% glycerol, mix by slowly pipetting up and down, and store at –80 °C.

3.3.2 Isolation of BAC Clones by Basic Alkaline Lysis Plasmid Miniprep

1. Use 5 mL o/N-bacterial culture containing BAC clone and growing in LB-*cam* medium.
2. Centrifuge at $1150 \times g$ for 10 min.
3. Resuspend pellet in 750 µL STE and transfer into a 1.5 mL reaction tube.
4. Centrifuge at $4600 \times g$ for 5 min.
5. Resuspend pellet in 200 µL Plasmid A.
6. Add 400 µL Plasmid B, mix by inverting six times, incubate on ice for 5 min.
7. Add 300 µL Plasmid C, mix by inverting six times, incubate on ice for 3 min.
8. Centrifuge at $16,400 \times g$ for 10 min.
9. Transfer the supernatant to a new 1.5 mL reaction tube.
10. Add 4 µL RNase A (20 mg/mL), mix by inverting, and incubate at 37 °C for 45 min.
11. Add 300 µL of PCiA (*see Note 7*), mix by shaking for 2 min, and centrifuge at $16,400 \times g$ for 2.5 min.
12. Transfer the upper, aqueous layer to a new 1.5 mL reaction tube.
13. Add 650 µL iPrOH, mix by shaking for 2 min, and centrifuge at $16,400 \times g$ for 10 min.
14. Remove the supernatant, add 700 µL of 70% EtOH, and incubate at 4 °C o/N.
15. Centrifuge at $16,400 \times g$ for 2.5 min and remove the supernatant.

16. Air-dry the pellet for 3 min and resolve in 100 μL T-Buffer. Resolving can be supported by incubation in a water bath at 42 °C for 1 h with occasional tapping.
17. Measure concentration (*see Note 8*).

3.3.3 BAC Clones Fingerprinting

1. Prepare a total volume of 40 μL per reaction:
 - 1 μL of restriction enzyme (*see Note 9*).
 - 4 μL of the appropriate buffer.
 - 25–40 μg of isolated DNA (*see Note 8*).
2. Digest o/N at 37 °C.
3. Separate the DNA fragments in 0.5–0.6 % agarose gel by gel electrophoresis. Run the gel electrophoresis for 1 h at 10 V, then for another 6 h at 60 V or longer if necessary to separate fragments of similar size (*see Notes 10 and 11*).

3.3.4 BAC-End Sequencing

The positioning of clones in the BAC maps is approximate, based on their assembly by restriction digest fingerprinting. In the case a sequence has been deposited for a given clone to any database, it often comprises an assembly of unordered and incomplete pieces (Fig. 2c, d). For a precise localization of the purchased clones we, therefore, routinely sequence the ends of BACs from both sides with the Sanger method, using primers binding near the cloning sites in the backbones. Then, we BLAST the respective reference genome with the obtained sequence. End sequencing requires high purity of the BAC, but with the following protocol, we commonly achieve >300 bp.

1. Mix 50–150 μg of isolated DNA from miniprep (3.3.2) in a volume of 50 μL (*see Note 8*) with 25 μL of PEG-MgCl₂ by sufficient mixing by pipetting up and down.
2. After 10 min incubation at RT, spin down the DNA at 16400 $\times g$ for 20 min and wash the pellet with 70% EtOH o/N. Resolve DNA in 10–20 μL T-Buffer.
3. Mix sequencing reaction: 1–5 μL DNA (take a minimum of 1 μg BAC DNA per 10 μL reaction, better 2–3 μg DNA per reaction), 5 pmol primer (*see Note 12*), 1 μL BigDye 3.1, 4 μL buffer and fill in dH₂O up to 10 μL .
4. Program: 95 °C for 5 min, 95 °C for 30 sec, 50 °C for 10 sec, 60 °C for 4 min, step 2–4 repeat 50 \times , 4 °C for 10 min.
5. Purify DNA with EtOH precipitation: Add 2.5 μL of 125 mM EDTA to each sequencing reaction, mix with 30 μL 100% EtOH by pipetting up and down several times. Incubate for 15 min on ice, then centrifuge 30 min at 16,650 $\times g$ at 4 °C. Remove supernatant and add 150 μL of 70% EtOH. Vortex, centrifuge 2.5 min at 15,900 $\times g$. Remove carefully absolutely all supernatant and air-dry pellet for 6 min. Resolve in 30 μL dH₂O for separation on capillary electrophoresis.

3.3.5 *Determination of BAC-End Sequences in the Reference Genome*

1. Reference genomes are available at several sites on an open-source basis. We prefer and recommend using Ensembl (<http://www.ensembl.org/index.html>), but other sources should work as well. At [ensembl.org](http://www.ensembl.org), the command line has a “BLAST/BLAT” button tool, directing you to the genome search side. Click the “New job” button, paste your sequence (*see Note 13*), and select the appropriate database at “add/remove” species (*see Note 14*). Use the default settings for the other parameters and “Run”.
2. BLAT will provide an output list for the best matches. Get the list of complementary sequences located on the genome in a new window (Fig. 2b). The data can be re-assorted for a number of parameters. We specifically consider:
 - (a) genomic location—the location of the BAC-ends should at least approximately correlate to the chromosomal position of the BACs in the NCBI Genome Viewer Database.
 - (b) orientation—both end sequences must be in opposite directions.
 - (c) length and %ID—the end sequences are assumed to be (almost) identical to the genome in their entire length. Single nucleotide polymorphism or short insertion/deletion might, however, decrease the % ID below 100%. In the case the identity falls below 90%, you are confronted with an artificial matching.
3. Clicking the link with genomic localization directs you to a new window showing the genomic localization of the BAC-end sequences. The “Region in detail” window illustrates the adjacent region of the BAC locus, confirming the match. By using the genomic localization of both BAC-end sequences, you can specifically zoom into the “localization” covered by the BAC (Fig. 2c).
4. The entire BAC sequence can be extracted from the reference genome by clicking the blue “contig” bar and choosing the “export primary assembly sequence/features” option. When a new window pops up you “Select location” by indicating the chromosome number and the terminal positions of the BAC-ends. Clicking “Next” provides the choice of the “Text” option for providing the entire BAC sequence in a .fasta format.
5. Copy-pasting the sequence into an appropriate viewer facilitates the comprehensive analysis of the region covered by the BAC (*see Note 15*).

3.4 Modification of BAC Clone by Bacterial Recombineering

Our bacterial recombination protocol is based on the “recombineering” system, initially described by [14] and modified by [15]. The BAC clone selected has to be transferred into a SW *E. coli* strain derivative and single cell clones are then prepared for recombination of MOI (*see Note 16*).

3.4.1 Making Electrocompetent Cells

1. Inoculate a 1 mL o/N-culture with the desired SW strain, incubate it at 180 rpm at 32 °C.
2. Transfer 1 mL of o/N-culture in 50 mL LB-medium in a 500 mL Erlenmeyer flask. Continue with incubation at 32 °C and shaking at 180 rpm until $OD_{600} = 0.6\text{--}0.8$.
3. Cool cell suspension with occasional shaking on ice for 10 min.
4. Transfer cell suspension into 50 mL-tube and centrifuge them for 10 min at $2200 \times g$ at 4 °C. Continue with additional centrifugations at 4 °C to wash cells. Besides, always keep cells on ice (*see Note 17*).
5. Resuspend cell pellet in 25 mL ice-cold ddH₂O and centrifuge for 10 min at $2800 \times g$.
6. Resuspend cell pellet in 12.5 mL ice-cold ddH₂O and centrifuge for 10 min at $3200 \times g$.
7. Resuspend cell pellet in 5 mL ice-cold 10% glycerol and centrifuge for 10 min at $4200 \times g$.
8. Resuspend in 0.5 mL ice-cold 10% glycerol.
9. Prepare aliquots of 80 µL. Cells can be used directly or stored at –80 °C.

3.4.2 Transfer BAC Clone into Recombineering-Competent *E. coli* Strain

1. Prepare materials: pre-cool cuvettes and cell aliquots on ice. Pre-warm 1 mL SOB aliquots in 1.5 mL reaction tubes at 32 °C.
2. Prepare reaction tube with BAC DNA amount proportionate to isolation method. BAC integrity should have been checked before electroporation by restriction digest.
3. Mix 80 µL of electrocompetent cells (as prepared in 3.4.1) with 1–6 µL DNA containing 3–6 µg by pipetting 2–3× up and down and transfer it into the cuvette.
4. Dry the cuvette on the outside with a paper wipe (*see Note 18*), put it into an electroporator, and apply a 1.75 kV pulse. Make notes of the actual voltage and time applied (*see Note 19*).
5. Add 1 mL SOB to wash cells from the cuvette and transfer cells back to the reaction tube (*see Note 20*).
6. Incubate the cells at 32 °C for 1 h and clean the cuvette (*see Note 21*).
7. Centrifuge cells for 5 min at $2350 \times g$.

8. Remove most of the supernatant (leave 100–200 μL in the tube).
9. Resuspend cells and plate on appropriate antibiotic LB-agar plates.
10. Incubate plates for 24–48 h (*see* **Note 22**).

*3.4.3 Preparation of
Modification as a
Linearized dsDNA
Fragment*

1. Excise the constructed and verified modification vector (3.1) from the plasmid backbone by appropriate restriction enzyme (s) (*see* **Notes 23** and **24**).
- 2.1 For purification of modification fragments <10 kb:
 - (a) Separate modification fragment from other elements of the plasmid on 0.6–0.7% agarose.
 - (b) Excise the modification fragment from the gel and extract DNA from agarose by conventional column-based DNA kit and elute in TE Buffer (*see* **Note 25**).
- 2.2 For purification of modification fragments >10 kb:
 - (a) Separate restriction enzyme digest on 1% low melting agarose.
 - (b) Excise modification fragment, transfer it into a 1.5 mL reaction tube, weigh and completely resolve agarose, normally 12 min at 70 °C is sufficient.
 - (c) Equilibrate the sample to 42 °C for 5 min, add agarose (1 U/100 mg gel) and digest for 30 min at 42 °C.
 - (d) Centrifuge undigested carbohydrates at 15,000 $\times g$ for 10 min.
 - (e) Add 1/10 volume 3 M NaOAc 5.2 pH and extract with 200 μL PCiA by slowly inverting the reaction tube for 2 min.
 - (f) Centrifuge for 2 min at 15,000 $\times g$ and transfer the upper aqueous phase into the new reaction tube.
 - (g) Repeat the PCiA extraction twice or until the interphase is clear.
 - (h) Add 2.5 \times volumes of 100% EtOH, mix gently, and incubate for 30 min at –80 °C.
 - (i) Centrifuge at 15,000 $\times g$ for 30 min at 4 °C and wash the DNA pellet with 70% EtOH o/N.
 - (j) Resolve the DNA in 10 μL T-Buffer.
3. Determine the concentration of the excised fragment and run a small aliquot on agarose gel for verification of its integrity.
4. Use modification fragment directly for bacterial recombineering into BAC clone (*see* Subheading 3.4.4.) or store DNA fragment at –20 °C.

3.4.4 BAC Recombineering

This step brings together the desired BAC and the modification fragment in the recombineering competent SW cells. The protocol therefore largely refers to previously conducted steps, with some essential variations.

1. Prepare o/N-culture with an SW-clone carrying the desired BAC, and inoculate 100 mL culture and grow it to $OD_{600} = 0.4\text{--}0.5$ according to Subheading 3.4.1.
2. Transfer 100 mL of culture to a 42 °C water bath and keep it for 15 min with constant shaking to induce the expression of λ -Red proteins.
3. Cool on ice for 10 min with occasional shaking and prepare electrocompetent cells according to Subheading 3.4.1, **step 3**.
4. For a 5 kb modification fragment, electroporate 100 ng excised vector DNA from Subheading 3.4.3 into 80 μ L recombineering-electrocompetent cells according to Subheading 3.4.2, with the exception that recovery time in SOB at 32 °C after electroporation needs to be 2 h instead of 1 h to guarantee proficient recombination (*see Note 26*).
5. Plate the cells on agar carrying the appropriate antibiotic (*see Note 27*) and grow them for a maximum of 48 h. Prepare back-up streaks on a plate with the same antibiotic (*see Note 28*) to promote analysis of defined colonies.

3.4.5 Arabinose-Induced FLPE and Cre Recombination

In addition to the heat-induced bacterial recombination, the SW105 and SW106 strains provide the opportunity of inducible FLPE and Cre recombinases. This might be used for exchanging or deleting a resistance cassette, in the case it is flanked by appropriate FTR or lox sites. Exchange of resistance cassettes can be carried out according to 3.4.4, using a linearized fragment carrying an alternative resistance cassette, flanked by FTR or lox sites according to the cassette in the BAC. Deletion of a selection cassette can be done according to a simplified protocol.

1. Inoculate 3 mL SOB-cam medium with bacterial cells containing BAC in which the region should be removed.
2. Reach the $OD_{600} = 0.2\text{--}0.3$. Measure the cell culture density by spectrophotometer.
3. Add 10 μ L arabinose (100 mg/1 mL) per 1 mL inoculum.
4. Incubate the culture for 1 h with shaking and transfer bacteria on LB-agar plates by making a streak of single cells or plating a small volume of 5 μ L culture diluted in LB-medium on LB-cam.
5. Pick single colonies (can be four colonies per inoculum as the Cre recombination is very effective) to create a back-up LB-agar plate and 2.5 mL o/N-culture for checking them by relevant restriction digest.

3.5 Verification of BAC Recombinant

Multiple approaches can be followed to verify the correct recombination of the modification fragment into the BAC. We commonly follow a combination of end-point PCR, BAC fingerprinting, and Sanger sequencing. A full BAC exploration by next-generation sequencing is optional but still costly and time-consuming.

3.5.1 End-Point PCR

1. Prepare 100 μ L of T-Buffer in PCR tubes.
2. Transfer a visible amount of bacterial colony/clone into the PCR tube (*see Note 29*). Make sure to wash the bacteria off thoroughly.
3. Disrupt bacteria by 10 min at 95 °C, 5 min at 4 °C
4. Centrifuge at 2000 $\times g$ for 10 min to pellet cell debris.
5. For standard end-point PCR reaction, use 2 μ L of the supernatant as DNA template and primer pairs flanking both HAs. *See* schematic illustration in Fig. 3b.

3.5.2 BAC Clone Fingerprinting

According to Subheading 3.3.3, any enzyme can be chosen for checking the general integrity of the BAC. As the modification normally affects only a rather small region of the BAC, only a few enzymes will indicate whether the modification has been correctly integrated.

1. Perform in silico digestion of the original BAC sequence and the BAC sequence comprising the desired modification for enzymes producing 30–40 bands.
2. Compare digestion patterns and choose an enzyme that produces a significant difference between the two constellations (*see Note 30*).
3. Perform fingerprinting restriction enzyme digest and run it on an agarose gel, according to Subheading 3.3.3 (Fig. 3a).

3.5.3 Sanger Sequencing

The procedure can be carried out according to Subheading 3.3.4, albeit with primers located within the BAC. We recommend verifying the sequences across the HA and across the essential components of the modification (Fig. 3b).

3.6 Preparing of BAC Targeting Vector for Nucleofection into Cells

Once when the BAC clone has been verified to contain all modifications, it can be prepared for transfection into vertebrate cells. Sufficient amounts and purity of BAC vectors can be produced by using a commercially available Large Construct Kit (QIAGEN). For homologous recombination in somatic cells, the BAC needs linearization and co-transfection with CRISPR/Cas components.

1. Use a large construct kit and follow the manufacturers' instructions. Resolve the pellet in an appropriate amount of TE Buffer, for example, 100 μ L in the case of 250 mL o/N-culture. Following the protocol, the concentration of circular BAC is

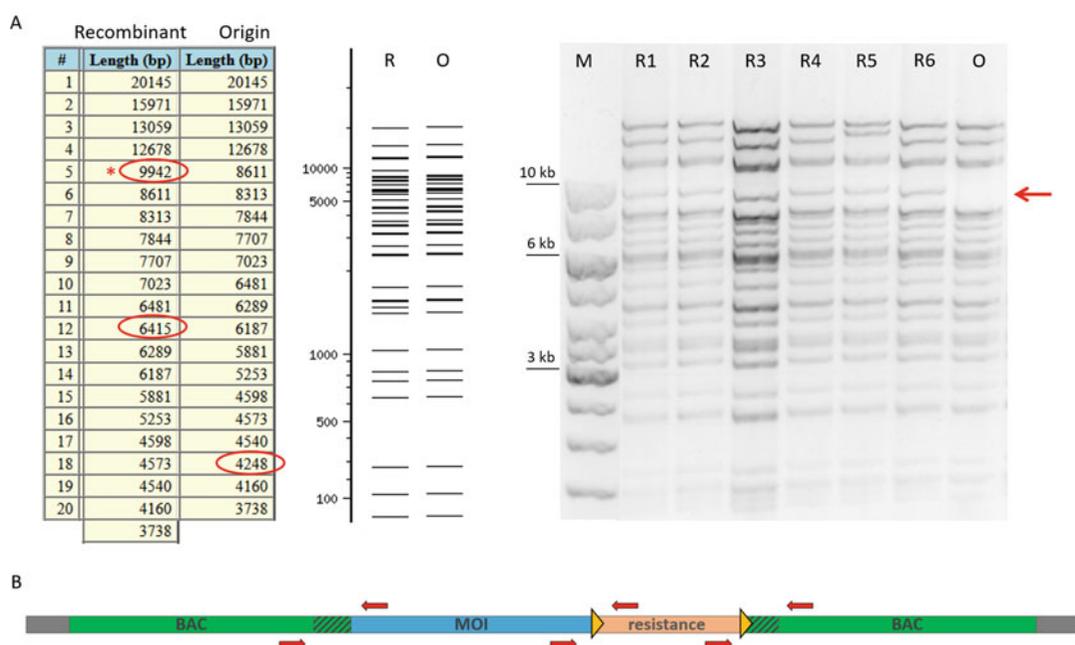


Fig. 3 Verification of modified BAC recombinants. **(a)** BAC clones fingerprinting. The digest patterns of the modified BAC recombinant (R) and origin BAC clone (O) are created in silico and compared to find the distinguishing fragments (right, the fragments are circled). The digestion patterns are also visualized by simulation of separation in agarose gel (the middle) displaying the significance of distinguishing fragments. The fragment of 9942 bp (star) mainly significant for BAC modification is missing in the digest pattern of origin BAC clone (left). This fragment is also detected after DNA fragments separation of digested origin BAC clone (O) and modified BAC recombinants (R1-R6) by agarose gel electrophoresis (right). The other two distinguishing fragments are so close to neighbor fragments that they are not separated on agarose gel. **(b)** Position of primers verifying the correct location of modification of interest (MOI) in modified BAC recombinants by end-point PCR and Sanger sequencing. The primers (arrows) are placed at positions to amplify and sequence the homology arms (striped boxes) and the MOI-resistance cassette border in modified BAC recombinants

mostly 200–600 ng/ μ L. Linearize 20–25 μ g of BAC with an appropriate enzyme o/N at 37 °C in a total volume of 200 μ L (*see Note 31*).

2. Add 20 μ L 3 M NaOAc 5.2 pH and extract with 200 μ L PCiA by slowly inverting the reaction tube.
3. Centrifuge at 16,400 $\times g$ for 2 min and transfer the upper aqueous phase into a new tube (*see Note 32*).
4. Add 400 μ L of 100% EtOH, mix gently, and incubate at -80 °C for 30 min.
5. Centrifuge at 16,400 $\times g$ for 30 min at 4 °C. Carefully remove supernatant and wash DNA pellet in 70% EtOH at 4 °C o/N.
6. Centrifuge at 16,400 $\times g$ for 2 min, carefully remove the supernatant, air-dry for 2–3 min, and resolve in 11 μ L TE buffer (*see Note 33*).
7. Determine concentration and mix with linearized BAC with CRISPR/Cas components (*see Note 34*).

4 Notes

1. Many resistance cassettes use a combination of a murine PGK promoter for expression in mammalian cells and a small EM7 promoter for expression in *E. coli*. Our preferred resistance gene is the acetyltransferase neo, providing kanamycin resistance in bacteria and neomycin/G418-resistance in mammalian cells. Similarly, we had sufficient experience with Sh ble, providing zeocin-resistance in bacteria and mammalia and bsd providing blasticidin resistance. As the performance of both Sh ble and bsd in *E. coli* critically depends on the exact composition of the medium, we make use of commercially available ready-to-use media or agar-media compositions. We also tested puro for resistance against puromycin in a similar way, but this proved not sufficient in bacteria in our hands, even when ready-to-use media were used.
2. We found it helpful to truly assemble the modification at a nucleotide resolution as this provides an excellent basis for analysis by restriction enzyme digest, designing end-point PCR, Sanger sequencing, estimation of splice sites, and open reading frames, etc. BioEdit [16] is our preferred program.
3. Common bacterial recombination protocols claim 50 bp of homologies as sufficient. We presume that this suggestion was based on the ability to create such elements from ssODN in a very cost-efficient manner. We mostly let HAs of 300 bp commercially synthesize, as gene synthesis has become very cost-efficient as well. Gene synthesis also allows us to generate both HAs in the same plasmid, separated by one or several sites for restriction enzymes. This facilitates the integration of MOI and resistance cassette. Further, this strategy provides the opportunity to place further restriction enzyme sites at the terminal ends of the HAs for the final release of the modification vector from the backbone (*see* Subheading 3.4.3).
4. Given the effective gene synthesis services available, the entire modification vector might be commercially synthesized. This would simplify the procedure, but to our assumption, the integration of available components such as resistance cassettes of some 2 kb size or larger from the existing plasmid by cloning is still cheaper than its full synthesis for each new project.
5. Ideally, the entire modification plasmid should be verified by Sanger sequencing. Larger parts that have been acquired from one source, either an established plasmid or a verified gene synthesis, however, we normally do not sequence completely. Rather we make use of several fingerprinting restriction enzyme digests and the verification of the terminal ends of such cassettes by Sanger sequencing.

6. We assume that mostly a gene is the desired target region. As most mammalian genes are properly annotated in the meanwhile, giving the gene name should direct you to the desired locus. In the case the gene name is not known in the species of interest or an intergenic region is the target region, we recommend searching for neighboring genes and to explore the target region by zooming or moving with the control bar. The search for these neighbor sites can be done in the species of interest or, as most loci appear conserved among mammals, in another species; for sure, the information density on genome annotation is highest in human. The current nomenclature of gene names can be explored by using the “gene” database at NCBI (<https://www.ncbi.nlm.nih.gov/gene>), which recognizes also older names or aliases.
7. Phenol is nowadays often seen critical for its harmful properties. Following alternative protocols (<https://bacpacresources.org/protocols.htm>), its usage can be avoided, but the DNA is then more impure, which affects resolving and electroporation into bacteria. The sample quality does, however, not interfere with restriction digest and visualization of band patterns by gel electrophoresis. We use the simplified protocol regularly for fingerprinting of BAC clones, but in this case, we resolve the DNA pellet isolated from only 2.5 mL o/N-culture in 100 μ L T-Buffer.
8. In a common photometer, we normally detect 1000–2500 ng/ μ L from a 5 mL o/N-culture resolved in 100 μ L. This is, however, totally misleading as the method co-precipitates disrupted chromosomal DNA of the *E. coli* genome to a large extent. A more realistic assumption is that you harvest some 1–2 μ g of BAC DNA (<25 ng/ μ L in 100 μ L). Alternative methods might avoid this complication, but they are more complicated to conduct. As co-precipitation of *E. coli* DNA does, in principle, not affect the subsequent steps, we stick to the incorrect but measurable values and talk about “isolated DNA” rather than “BAC DNA.”
9. We found it appropriate to digest BACs in a way that 30–40 bands appear, correlating with a mean band size of approximately 3–6 kb. This fits with many of the commonly used enzymes such as *Hin*DIII, *Spe*I, *Hpa*I, *Eco*RI, and *Xba*I. Our preferred choice for initial fingerprinting is mostly *Eco*RI.
10. Fingerprinting of BACs was initially performed with pulsed-field electrophoresis for creating BAC maps. We found it appropriate to separate band sizes from 3 to 50 kb within this much simpler setting. Care needs to be taken that combs are cleaned with detergent before placing them in the agarose to avoid disruption of the slot surface with debris. Usage of fresh electrophoresis buffer is recommended as well.

11. Visualization of DNA bands with EtBr performed outstandingly well when adding it to agarose that has been cooled to 60–65 °C after boiling. Similarly, the less harmful GelRed can be used, but care must be taken that GelRed has to be used at a concentration of 2.5 µL staining dye to 100 mL agarose gel. GelRed has to be kept in the fridge and cannot be mixed directly into the sample. Otherwise, bands from a BAC digest will run at a different speed and the patterns will be diffuse.
12. On the commonly used BAC backbone, the insertion site of the genomic region is flanked by primer binding sites for T7 (5'-GTAATACGACTCACTATAGGG-3') and SP6 (5'-ATT TAGGTGACACTATAGAA) (Table 1).
13. The BLAST at Ensembl allows the parallel search for several sequences, for example, the 5'-end and the 3'-end of a BAC. This can be done either by choosing “add more sequences” or, very proficient in our hands, by inserting the 5'-end and 3'-end sequence in the same window, separated by “NNNN”. Treating both sides as a single sequence helps if one of the end sequences is completely within a highly repetitive segment. In this case, the connection of the sequence to the complementary end will prioritize the identification of the correct locus in the output list. Theoretically, both end sequences might match highly repetitive elements in the genome, but in practice, this did never happen in our hands.
14. For most of the species, multiple genomes are meanwhile available, but quality often differs substantially. For comparison to other databases such as NCBI, we recommend choosing the respective reference genome. This is, by now, GRCH38 for human, GRCm39 for mouse, ARS-UCD1.2 for cattle, Sscrofa1.1 for pig, Oar_rambouillet_v1.0 for sheep.
15. We normally use BioEdit for identifying the BAC-end sequences and the restriction enzyme sites used for cloning the fragment into the BAC backbone (Fig. 2d). Further, the position of genes and their exons can be defined. For confirming the restriction enzyme digest pattern from the fingerprinting analysis, the BAC sequence should be combined with the sequence of the vector backbone and examined for the used restriction enzyme sites (<http://nc2.neb.com/NEBcutter2/index.php>).
16. The most recent generation of recombinering provides three main *E. coli* strains: SW102, SW105, and SW106. While all of them provide a heat-inducible λ -Red recombination system comprising of the λ -encoded *exo*, *bet*, and *gam* components under a temperature-sensitive repressor. For this cells have to be propagated only at 32 °C, except when the λ -Red system is switched at 42 °C. Strains SW105 and SW106 contain FLPE or

Cre genes in addition. In both cases, the genes can be induced by arabinose.

17. Avoid repetitive centrifugation of cell suspensions. In the case you have more than 50 mL o/N-culture, separate it onto several tubes. Otherwise, the competent cells are of a low quality which might impair electroporation efficacy.
18. Precisely remove any traces of liquid as this might result in a short-cut and therefore loss of the reaction.
19. In our Eppendorf electroporator (Eporator), we usually get a variation of voltage between 1.65 and 1.69 kV, and a pulse time of 4.3–4.9 ms. When checking the quality of new batches of electrocompetent cells by applying a pulse to an empty cell aliquot, a pulse time > 5.0 ms indicates good quality of cells. When using electrocompetent cells of lower quality, clones might still arise, albeit at decreased numbers.
20. The recovery of cells in SOB after electroporation is a critical step. It is said that each additional 30 s of keeping cells in the original suspension might decrease viability by half. Moreover, the usage of pre-warmed medium and placing the cells at 32 °C is seen as an additional heat-shock, supporting the transfer of BAC DNA into the cells. Speed is therefore essential and we normally do not make use of special tips or Pasteur pipets to transfer the cells out of the cuvette. Rather, we spill 1 mL of pre-warmed SOB into the cuvette, pipet up and down 2–3 times, and with the same tip transfer it back to the reaction tube and the reaction tube into the incubator. The advantage of fast processing should overcome the detrimental aspects of leaving small volumes of cell suspension in the cuvette. For convenient and fast work, we placed the electroporator next to the incubator.
21. Cuvettes can be reused several times until visible damage. Washing of the cuvette with dH₂O 10 times is sufficient, but after the end of an electroporation session, we additionally sterilize it for 30 min under UV light.
22. The efficacy of BAC transfer into SW strains strongly depends on the quality of BAC DNA. We only get few colonies when using the simplified protocol of basic alkaline lysis, yield is increased to 20–30 bacterial colonies if the described phenol/chloroform extraction step is included. More than 100 clones can be obtained when BACs are isolated by commercially available time-consuming and costly column-based kits.
23. For normal-sized modifications (approx. 5 kb), we digest 6 µg plasmid from column-based plasmid isolations or 15 µg DNA from plasmid minipreps. Without columns). When modifications of 10 kb are prepared, we use around 40 µg plasmid DNA.

24. In this case, the modification fragment is of similar size as the plasmid backbone, a clear separation of the two bands might be challenging. We found it helpful to use additional restriction enzymes in the digest that cut the backbone into two (or more) fragments. Importantly, separation of bands on agarose is impaired, when DNA concentration is high in the restriction enzyme digestion reaction. It is often helpful to increase the volume of the digest before loading onto the gel, preferentially by filling one or several slots almost completely. Further, the usage of GelRed as DNA stain dye often results in incomplete separation of bands. We normally load GelRed directly into the sample.
25. To get the maximum DNA amount from the purification column, elution should be carried out with pre-warmed (i.e., 70 °C) elution buffer, incubate membrane with elution buffer for 5 min, and elute membrane 2× with the elution buffer. If necessary, concentrate the DNA by EtOH precipitation in a smaller volume. Add 1/10 volume 3 M NaOAc and 2.5 volumes 100% EtOH. Invert a few times and incubate for 30 min at -80 °C. Centrifuge (16,400 × *g*, 30 min, 4 °C) and wash with 70% EtOH o/N.
26. The amount of DNA can be adapted, according to fragment length. We successfully used DNA amounts of 60–1000 ng. It is important, however, to keep the DNA volume lower than 10% of the cell suspension volume of 80 µL.
27. For the growing cells that underwent bacterial recombination, it is essential to select for the resistance encoded by the modification vector, for example, kanamycin, blasticidin, and zeocin.
28. After selection for and verification of the modification, selection can be switched back to cam. In contrast to older protocols that claim this impossible, we do recover cells for 2 h in LB without antibiotic and then plate them on cam. It is particularly recommended to keep bacteria under zeocin selection only temporarily.
29. End-point PCR can be immediately driven on material from the original clone, but it is necessary to prepare a back-up streak on a second plate to regenerate enough material for further analysis.
30. We found it appropriate to choose an enzyme that produces a band of 6–12 kb that is appearing only in the original or the modified BAC. For sure this band should be clearly separated from the other fragments of the digestion pattern. Smaller bands are often difficult to visualize, in the case of larger bands, it is often difficult to estimate the band size. Ideally, two distinct digestion approaches are selected and compared.

31. The appropriate enzyme should produce extended HAs to your MOI. We found it convenient to cut the final BAC targeting vector with *NotI*, whose sites flank the genomic segment in the BAC clone of the most frequent libraries in NCBI and appear very rarely in the genome (Table 1). The second option can be *AscI*, which, unlike *NotI*, only opens the vector. However, you must always check that the enzyme you have selected does not affect your MOI.
32. The transferring of the aqueous phase needs to be done with absolute caution to not have the remainder of phenol in the DNA sample. It could be added next extraction with chloroform to remove phenol residue, but the DNA loss can be too big after two extractions. Unfortunately, we have not seen an evident improvement in cell response to nucleofection.
33. Following the protocol, the concentration of linearized BAC is mostly 1000–1500 ng/ μ L. For further processing, we recommend using fresh BAC DNA (1–2 days old). Longer storage in the fridge or freezing can harm the quality of DNA samples. For longer storage, it is better to keep DNA as a pellet in 70% EtOH.
34. Linearized BAC can be combined with distinct constellations of CRISPR/Cas, either applied as ribonucleoprotein, RNA, or plasmids expressing gRNA and Cas9. For us, it appeared most flexible to use separate plasmids comprising Cas9 and gRNA elements [17], with the latter being commercially synthesized under a human U6 promoter. A common composition of BAC with CRISPR/Cas plasmid for nucleofection into somatic cells is 250 ng/ μ L of Cas9 plasmid, 250 ng/ μ L of gRNA-expressing plasmid, and 500 ng/ μ L of linearized BAC in 5 μ L reaction.

References

1. Shizuya H et al (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc Natl Acad Sci U S A* 89(18):8794–8797
2. Lander ES et al (2001) Initial sequencing and analysis of the human genome. *Nature* 409(6822):860–921
3. Archibald AL et al (2010) Pig genome sequence--analysis and publication strategy. *BMC Genomics* 11:438
4. Bovine Genome S et al (2009) The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science* 324(5926):522–528
5. Wheeler DA et al (2008) The complete genome of an individual by massively parallel DNA sequencing. *Nature* 452(7189):872–876
6. Genomes Project, C et al (2010) A map of human genome variation from population-scale sequencing. *Nature* 467(7319):1061–1073
7. Jain M et al (2018) Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nat Biotechnol* 36(4):338–345
8. Gordon D et al (2016) Long-read sequence assembly of the gorilla genome. *Science* 352(6281):aac0344

90 Hannah Auch et al.

9. Zhang Y et al (1998) A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat Genet* 20(2):123–128
10. Muyrers JP et al (1999) Rapid modification of bacterial artificial chromosomes by ET-recombination. *Nucleic Acids Res* 27(6): 1555–1557
11. Murphy KC (1998) Use of bacteriophage lambda recombination functions to promote gene replacement in *Escherichia coli*. *J Bacteriol* 180(8):2063–2071
12. Yu D et al (2000) An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc Natl Acad Sci U S A* 97(11): 5978–5983
13. Vochozkova P et al (2019) Gene editing in primary cells of cattle and pig. *Methods Mol Biol* 1961:271–289
14. Lee EC et al (2001) A highly efficient *Escherichia coli*-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics* 73(1): 56–65
15. Warming S et al (2005) Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Res* 33(4):e36
16. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, *Nucleic acids symposium series*. Information Retrieval Ltd., London, pp c1979–c2000
17. Osoegawa K et al (2001) A bacterial artificial chromosome library for sequencing the complete human genome. *Genome Res* 11(3): 483–496
18. Rogel-Gaillard C et al (1999) Construction of a swine BAC library: application to the characterization and mapping of porcine type C endoviral elements. *Cytogenet Cell Genet* 85(3–4):205–211
19. Fahrenkrug SC et al (2001) A porcine BAC library with tenfold genome coverage: a resource for physical and genetic map integration. *Mamm Genome* 12(6):472–474
20. Eggen A et al (2001) Construction and characterization of a bovine BAC library with four genome-equivalent coverage. *Genet Sel Evol* 33(5):543–548
21. Dalrymple BP et al (2007) Using comparative genomics to reorder the human genome sequence into a virtual sheep genome. *Genome Biol* 8(7):R152

XII. ACKNOWLEDGEMENTS

I am grateful to Prof. Dr. Eckhard Wolf for giving me the opportunity to contribute to research at the Chair for Molecular Animal Breeding and Biotechnology, LMU Munich and to do my doctoral thesis there.

My greatest appreciation and respect for my mentor Prof. Dr. Nikolai Klymiuk for his excellent support and guidance, for always being there but promoting independency nevertheless.

Special thanks go to Dr. Petra Runa-Vochozková, my direct lead in all experiments, who spent countless hours encouraging, discussing and counselling, introducing me to the world of molecular biology. You are a great teacher! I was your first student, but I am sure there are many more to come.

I also want to thank everybody working at the Moorversuchsgut, the caretakers in the stable for looking after our precious animals as well as the technical assistants in the lab who always give a hand when help or advice is needed. A big thanks goes out to all my fellow doctoral students for being a great team and always helping each other. You make work enjoyable even through stressful times.

I want to thank my family and friends for support and interest in my work, as well as distraction whenever necessary.

Many thanks to all groups collaborating with us to promote research on IRDs that gave me the chance to work in an interdisciplinary and international field of scientists, this is an experience I would not want to miss.

I am grateful to the SPP2127 for providing a network of young scientists, promoting scientific exchange between the groups and new input for everybody. I also appreciate very much the financial support of USHER2020 and Foundation Fighting Blindness, as well as the connection made between science and the patients' world. This made me realize why this research is so important.