Rescue of Missplicing by CRISPR/*Sp*Cas9-Based Genome Editing Approaches Targeting Deep-Intronic Variants in *ABCA4*

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Pietro De Angeli

aus Urbino, Italien

Tübingen

2022

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

NO ONE SAYS IT WILL BE EASY BUT...

...DO NOT GIVE IN TO MEDIOCRITY...

ABSTRACT

CRISPR/Cas-vermittelte Genom-Editierung stellt eine vielversprechende Technologie dar, mit der sich die meisten bekannten krankheitsverursachenden Mutationen, einschließlich tief-intronischer Varianten, korrigieren lassen. Pathogene tief-intronische Varianten stellen einen erheblichen Teil der krankheitsverursachenden Mutationen im Mutationsspektrum des *ABCA4*-Gens dar. Die molekularen Mechanismen, durch welche sie wirken, sind mit fehlerhaften pre-mRNA-Spleißprozessen verbunden, die schließlich zu fehlgespleißten *ABCA4*-Transkripten führen. Die isolierte *ABCA4* c.5197-557G>T tief-intronische Variante führt zur Retention einer 188-bp intronischen Sequenz im reifen mRNA-Transkript. Durch die Etablierung von drei Standard-CRISPR/*Sp*Cas9-basierten Ansätzen und der neuartigen Enhanced-Deletion-*Sp*Cas9-Strategie konnte der Spleißdefekt, der durch die isolierte c.5197-557G>T verursacht wird, im Minigen-Assay in HEK293T-Zellen sowie in heterozygoten c.5197-557G>T-Zapfen-Photorezeptor-Vorläuferzellen erfolgreich korrigiert werden. Darüber hinaus wurde der neuartige Enhanced-Deletion-*Sp*Cas9-Ansatz in einem Minigen-Assay in HEK293T-Zellen zur Rettung von Fehlspleißen aufgrund der geclusterten tief-intronischen Varianten (c.5196+1013G>A, c.5196+1056G>A, c.5196+1134C>G, c.5196+1137G>A und c.5196+1216C>A) im Intron 30 von *ABCA4* in ersten Experimenten angewendet. Insgesamt konnte in dieser Arbeit zum ersten Mal gezeigt werden, dass Spleißdefekte in *ABCA4*, welche durch tief-intronische Varianten verursacht werden, durch die Anwendung von Genom-Editierungs-Ansätzen erfolgreich behoben werden können. Darüber hinaus wurde ein neuartiges und vielversprechendes Editierungs-Tool (Enhanced-Deletion-*Sp*Cas9) etabliert und für die Rettung von isolierten und geclusterten tief-intronischen-Varianten optimiert.

Genome editing mediated by CRISPR/Cas molecules represent a promising technology applicable to the correction of most of the known disease-causing mutations, including pathogenic deep-intronic variants. Pathogenic deep-intronic variants represent a good portion of the disease causing mutations in the mutational spectrum of the *ABCA4* gene. The molecular mechanisms through which they act is associated with faulty pre-mRNA splicing processes, eventually leading to aberrant *ABCA4* transcripts. The isolated *ABCA4* c.5197-557G>T deep-intronic variants results in the retention of a 188-bp intronic sequence in the mature mRNA transcript. By establishing three standard CRISPR/*Sp*Cas9-based approaches and the novel Enhanced-Deletion *Sp*Cas9 approach, the splicing defect due to the isolated c.5197-557G>T was sucessufully corrected in minigene assay in HEK293T cells, as well as in heterozygous c.5197-557G>T cone photoreceptor precursor cells. Furthermore, the novel Enhanced-Deletion-*Sp*Cas9 approach was preliminarily applied in minigene assay in HEK293T cells for the rescue of missplicing due to the clustered deep-intronic variants (c.5196+1013G>A, c.5196+1056G>A, c.5196+1134C>G, c.5196+1137G>A and c.5196+1216C>A) in intron 30 of *ABCA4*. On the whole, this thesis showed for the first time the successull rescue of *ABCA4*-related splicing defects by implementing genome editing approaches. In addition, a novel and highly promising editing tool (Enhanced-Deletion-*Sp*Cas9) was established and optimized for the rescue of isolated and clustered deep-intronic variants.

CONTENT

Content

1 ABBREVIATIONS

°C	Degree Celsius	**cGMP**	Cyclic guanidine monophosphate
μ	Micro	***Cj*Cas9**	*Campylobacter jejuni* Cas9
μl	Micro liter		
A	Adenosine	**Cl⁻**	Chlorine
aa	Amino acid	**CPC**	Cone photoreceptor precursor cells
AAV	Adeno-associated virus		
ABCA4	ATP-binding cassette transporter gene	**CRD**	Cone-rod dystrophy
		CRISPR	Clustered regularly interspaced short palindromic repeats
ABE	Adenosine base editor		
***As*Cas12a (Cpf1)**	*Acidaminococcus sp* Cas12a	**crRNA**	CRISPR RNA
		CtIP	Carboxy-terminal binding protein interacting protein
ASO or AON	Antisense oligonucleotides		
ATP	Adenosine triphosphate	**dATP**	Deoxyadenosine triphosphate
bp	Basepairs		
C	Cytosine	**dCTP**	Deoxycytidine triphosphate
Ca²⁺	Calcium		
CAG	Cytomegalovirus enhancer fused to the chicken beta-acting promoter	**ddH₂0**	Double-distillated H₂0
		ddNTPs	Didesoxyribonucleosid-triphosphate
Cas	CRISPR-associated protein	**dGTP**	Deoxyguanosine triphosphate
		DIV	Deep-intronic variant
CBE	Cytosine base editor		
cCPCs	Control CPCs	**DMEM**	Dulbecco's modified eagle medium
CD	Cone dystrophy		
cDNA	Coding or complementary DNA	**DMEM-**	DMEM without penicillin streptomycin
		DMEM+	DMEM including penicillin streptomycin
CFTR	Cystic fibrosis transmembrane conductance regulator		
		DMEM+f	DMEM+ containing 20% FBS
		DMSO	Dimethylsulfoxid

DNA	Desoxyribonucleic acid	**IRDs**	Inherited retinal disorders
DSB	Double strand break	**ISE**	Intronic splicing enhancer
DTT	Dithiothreitol	**ISS**	Intronic splicing silencer
dTTP	Deoxythymidine triphosphate	**IVM**	*in vitro* mutagenesis
e.g.	*exempli gratia* "for example"	**K⁺**	Potassium
ECD	Extracellular domain	**kb**	Kilobase (1000 bp)
ED*Sp*Cas9	Enhanced-Deletion *Sp*Cas9	**kDa**	Kilo Dalton
EDTA	Ethylendiamintetraacetate	**LCA**	Leber congenital amaurosis
EGFP	Enhanced green fluorescent protein	**M**	Molar
EPO	Eureopean patent office	**m**	Milli
ESE	Exonic splicing enhancer	**Mg²⁺**	Magnesium
ESS	Exonic splicing silencer	**min**	Minute
FBS	Fetal bovine serum	**ml**	Milliliter
FCS	Fetal calf serum	**mm**	Millimeter
FDA	United States Food and Drug administration	**MMEJ**	Microhomology-mediated end joining
FEN1	Flap endonuclease 1	**MRN complex**	MRE11–RAD50–NBS1
G	Guanine	**mRNA**	Messenger ribonucleic acid
gDNA	Genomic DNA	**ms**	Millisecond
h	Hour	**Na⁺**	Sodium
HDR	Homologous-directed recombination	**NBD**	Nucleotide-binding domain
hFGF	Human fibroblast growth factor	**NEAAs**	Non-essential amino acids
I	Inosine	**ng**	Nanogramm
IGF-1	Insulin-like growth factor 1	**NGS**	Next-generation sequencing
iPSCs	Induced pluripotent stem cells	**NHEJ**	Non-homologous end joining
		NLS	Nuclear localization signal

2

nm	Nanometer		**RPMI**	Roswell Park Memorial Institute medium
NMD	Nonsense-mediated decay		**RPMI-**	RPMI without penicillin streptomycin
***Nme2*Cas9**	*Neisseria meningitides* 2 Cas9		**RPMI+**	RPMI including penicillin streptomycin
N-Ret-PEA	N-retinylidene-PEA		**RT-PCR**	Reverse transcription - PCR
nt	Nucleotide		***Sa*Cas9**	*Streptococcus aureus Cas9*
NUC	Nuclease lobe			
PAM	Protospacer adjacent motif		***Sauri*Cas9**	*Staphylococcus auricularis* Cas9
PBS	Phosphate buffered saline		**SD**	Standard deviation
PCR	Polymerase chain reaction		**sgRNA**	Single chimeric guide RNA
PE	Pseudoexon		***Sp*Cas9**	*Streptococcus pyogenes Cas9*
PEA	Phosphatidylethanolamine			
pegRNA	Prime editing gRNA		**ssODN**	Single-stranded oligodeoxynucleotides
PGK	Phosphoglycerate kinase 1 promoter		**STGD1**	Stargardt disease
PI	PAM-interacting domain		**T**	Thymidine
pmol	Pico mole		**TALEN**	Transcription activator-like effector nuclease
PTC	Premature termination codon		**Taq**	*Thermus aquaticus*
Puro	Puromycin		**TE**	Tris-EDTA
REC	Recognition lobe		**TMD**	Transmembrane domain
RNA	Ribonucleic acid		**tracrRNA**	Transactivating crRNA
RNP	Ribonucleoproteins		**TTR**	Transthyretin protein
RP	Retinitis pigmentosa		**U**	Uracil
RPE	Retinal pigment epithelial cells		**UGI**	Uracil glycosylase inhibitor
RPE65	Retinal pigment epithelium-specific 65 kDa protein		**UV**	Ultra violet
			V	Voltage

v/v	Volume in volume	**ZNF**	Zinc-finger nuclease
WT	Wild-type		

2 INTRODUCTION

2.1 INHERITED RETINAL DISORDERS

Inherited retinal disorders (IRDs) represent a group of genetically and clinically heterogenous diseases characterized by visual dysfunction and impairment. On the basis of the clinical evaluation, non-syndromic IRDs can be divided in various sub-groups: i) congenital blindness (i.e. Leber congenital amaurosis, LCA), ii) IRD with primary peripheral visual loss (i.e. retinitis pigmentosa, RP), iii) IRD with primary central visual loss (i.e. cone dystrophy, CD, cone-rod dystrophy, CRD), iv) macular dystrophy (incl. Stargardt (STGD1) and Best disease), v) IRD with congenital, stationary function loss (i.e. achromatopsia, congenital stationary night blindness), and vi) syndromic forms (i.e. Bardet-Biedl and Usher syndrome) (Berger et al. 2010).

The most common form of non-syndromic IRD is RP, with an estimated prevalence of 1 in 4,000 people, whereas Usher syndrome is the most prevalent syndrome affecting 4 to 17 in 100,000 people, and being characterized by sensorineural hearing loss and RP of different grades of severity (Boughman et al. 1983, Hartong et al. 2006, Kimberling et al. 2010). Decades of molecular and genetic studies have led to the identification >250 genes linked to IRDs (https://sph.uth.edu/retnet/sum-dis.htm#A-genes, last query 17.08.2022), showing different inheritance patterns which include autosomal recessive, autosomal dominant, X-linked, as well as mitochondrial inheritance (Gamwell et al. 2013, Rivolta et al. 2002). Although some genes are predominantly associated to a specific phenotype and inheritance pattern, different mutations in the same gene can result in diverse clinical manifestations. For example, mutations in the ATP-binding cassette transporter gene (*ABCA4*) are mostly associated with STGD1, however, are also commonly found in CRD, fundus flavimaculatus and occationally in RP (Allikmets et al. 1997, Martínez et al. 1998, Maugeri et al. 2000).

Overall, the calculated prevalence of monogenic IRDs is 1 in 2,000 people, signifying that around 5.5 million people worldwide are affected, while 2.7 billion people are estimated to be healthy carriers for IRD-causing mutations (Berger et al. 2010, Hanany et al. 2020).

2.2 STARGARDT DISEASE

STGD1 is the most common form of juvenile macular degeneration, with a reported prevalence of 1 in 10,000 people (Michaelides et al. 2003). The phenotype is characterized by bilateral progressive loss of central vision associated with irregular macular and perimacular yellow-white fundus flecks, and a so-called ''beaten bronze'' atrophic central macular lesion. Clinical manifestation of the early-onset STGD1 is usually observed before 10 years of age and is correlated to the most severe *ABCA4*-mutation associated phenotype. The underlying genetic cause are biallelic pathogenic variants in the *ABCA4* gene and the inheritance pattern is autosomal recessive (Allikmets et al. 1997). Mutations in *ABCA4* can also cause other IRD phenotypes, i.e. CRD, fundus flavimaculatus and RP. These phenotypes may have a later age of onset (>40 years of age) enabling some of the affected individuals to maintain meaningful vision until late in the progression of the disease (Tsang & Sharma, 2018, Westeneng-van et al. 2012).

2.3 *ABCA4* –GENE, PROTEIN, LOCALIZATION AND FUNCTION

ABCA4 is a retina-specific expressed gene that encodes for a transmembrane transporter composed of two transmembrane (TMD), two extracellular (ECD) and two cytoplasmic nucleotide-binding (NBD) domains. The six domains reside within the same ~250 kDA single polypeptide chain arranged to form a complex tridimensional structure organized in two halves (Xie et al. 2021) (**Figure 1**).
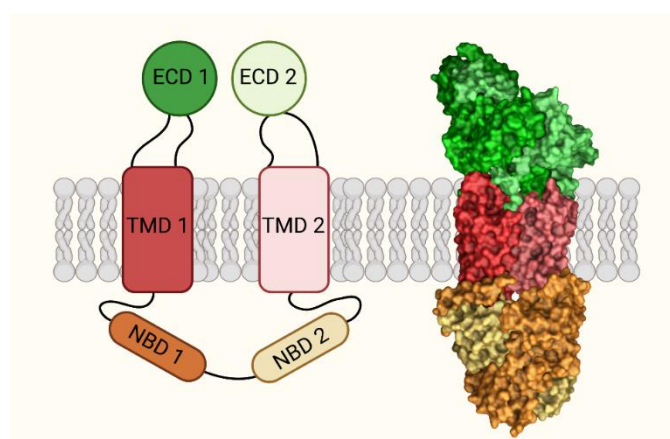
Figure 1: Details of the six domains of the ATP-binding cassette transporter A4 (ABCA4). Left: Schematic drawing of ABCA4 domains: the protein consists of two transmembrane (TMD), two extracellular (ECD) and two cytoplasmic nucleotide-binding (NBD) domains. Right: 3D structure of ABCA4. The two halves are highlighted in dark and light colors. The 3D rendering picture was generated on Maestro, Schrödinger (https://www.schrodinger.com) using the 7E7I PDB file (Capp et al. 2006).

Over the years, the expression of *ABCA4* has been identified in rods and cone photoreceptors, as well as the retinal pigment epithelial (RPE) cells. Limited expression of *ABCA4* has also been found in the mouse brain and human keratinocytes (Bhongsatiern et al. 2005, Tachikawa et al. 2005, Wiley et al. 2016). In photoreceptors, the localization of ABCA4 is confined to the rim region of the outer segment discs, where also the phototransduction cascade initiates (Allikmets et al. 1997, Lenis et al. 2018, Molday, 2007).

The phototrasduction cascade of the visual cycle begins when visual pigments (opsins) adsorb light and induce conformational isomerization of the inactive 11-*cis* retinal chromophore in the active all-*trans* retinal form. Through the subsequent activation of the downstream G-protein mediated reactions of the visual cycle, the intracellular concentration of cGMP is reduced, leading to the closing of ion channels that results in the reduction of intracellular Na^+ and Ca^{2+} levels and hyperpolarization of photoreceptors (Pugh et al. 1990). Most of the all-*trans* retinal is regenerated back to all-*cis* retinal through the visual cycle (Luo et al. 2008). However, a considerable amount of residual all-*trans* retinal and all-*cis* retinal reacts with phosphatidylethanolamine (PEA) to form N-retinylidene-phosphatidylethanolamine (N-Ret-PEA) (Quazi & Molday, 2014). The ABCA4 transporter functions as an importer, utilizing energy from the hydrolysis of ATP to translocate isomers of N-Ret-PEA from the luminal to the cytoplasmic side of the disk membrane, accelerating their clearance (**Figure 2**) (Quazi & Molday, 2014).
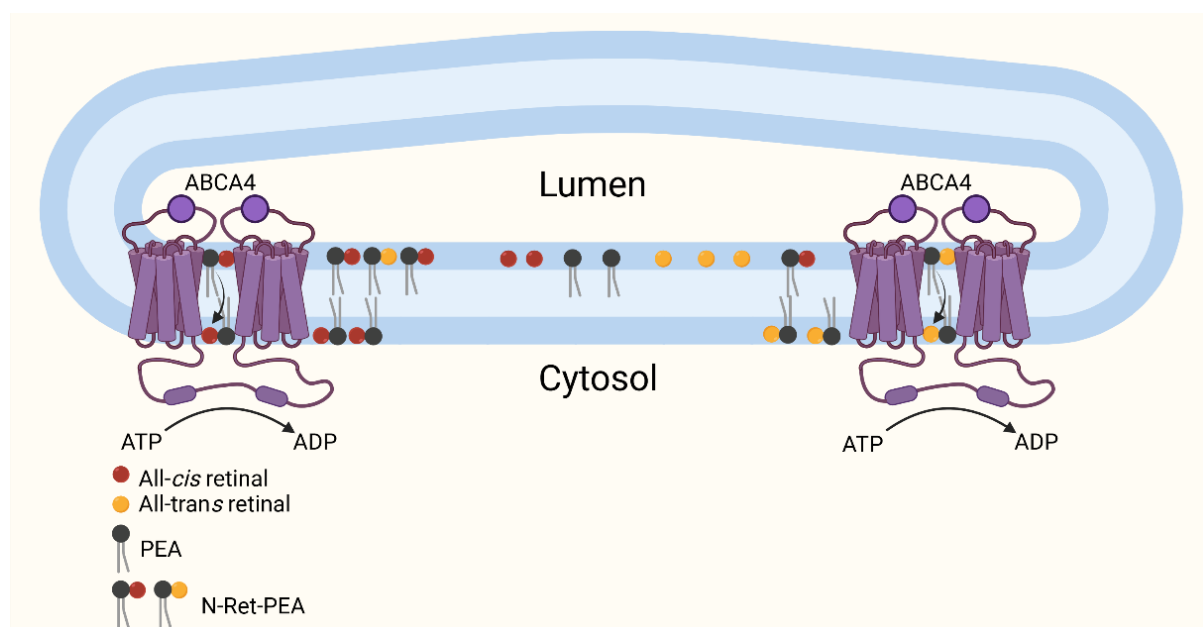
Figure 2: Schematic representation of a photoreceptor disk section highlighting the active transport mechanism mediated by ABCA4. All-*trans* and all-*cis* retinal reacts with phosphatidylethanolamine (PEA) to form N-retinylidene-PEA (N-Ret-PEA). Through the hydrolysis of ATP, N-Ret-PEA is flipped from the luminal to the cytoplasmic side of the disk membrane.

In case of inefficient clearance and accumulation of N-Ret-PEA in the disk membrane, cross reactions lead to toxic build-up of bisretinoids which is thought to alter RPE metabolism and ultimately determines cell death (Molday, 2007, Sparrow et al. 2012). Through the described molecular mechanisms, mutations in ABCA4 that result in the loss of ABCA4 (function) lead to the onset of STGD1 or other ABCA4-related IRDs (Lenis et al. 2018, Molday, 2007).

2.4 MUTATION SPECTRUM OF *ABCA4*

The mutation spectrum of *ABCA4* is highly heterogeneous. More than 1,700 single mutations have been described to date: 62% missense/nonsense, 13% small deletions, 3% small insertions, 1% small insertions and deletions, 1% large deletions, 1% large insertions, 2% complex rearrangements, and 17% splicing mutations (**HGMD® Professional 2021.4**, last query 13.01.2021). Following the discovery of *ABCA4* and association with STGD1 and other *ABCA4*-allied IRDs, the investigation of the mutation spectrum was primarily focused on exonic point mutations, including missense and nonsense mutations, small deletions and/or insertion, and canonical splice site variants (Shroyer et al. 2001, Yatsenko et al. 2001, Zernant et al. 2011). Despite comprehensive genetic testing, numerous clinically-diagnosed STGD1 cases remained unsolved or carried only a single heterozygous *ABCA4* variant (Zernant et al. 2011, Zernant et al. 2017). Recent locus sequencing projects have elucidated the missing heritability in *ABCA4*-related IRD, identifying large deletions, insertions and other structural aberrations, as well as deep intronic mutations (Bauwens et al. 2019, Khan et al. 2020b).

2.4.1 Deep-intronic variants in *ABCA4*

In the attempt of solving cases of STGD1 missing heritability, genetic testing started to focus on the intronic regions of the gene that were farther than 100 nt from intron-exon junctions. These mutations were regarded to as deep-intronic variants (DIVs) (Vaz-Drago et al. 2017). In 2013 Braun *et al* reported the first evidence of STGD1-causing DIVs in *ABCA4* resulting in missplicing (Braun et al. 2013). Thereafter, various important and recurrent *ABCA4* DIVs have been identified and characterized (Bauwens et al. 2019, Khan et al. 2020b, Sangermano et al. 2018, Zernant et al. 2014). Among the identified 287 splicing mutations, 15% are represented by DIVs. Although they are localized throughout the whole gene, two hot spots, in intron 30 and intron 36, are characterized by various DIVs displayed in close vicinity to each other (**Figure 3**). These are denoted as intron 30 or intron 36 clustered DIVs.
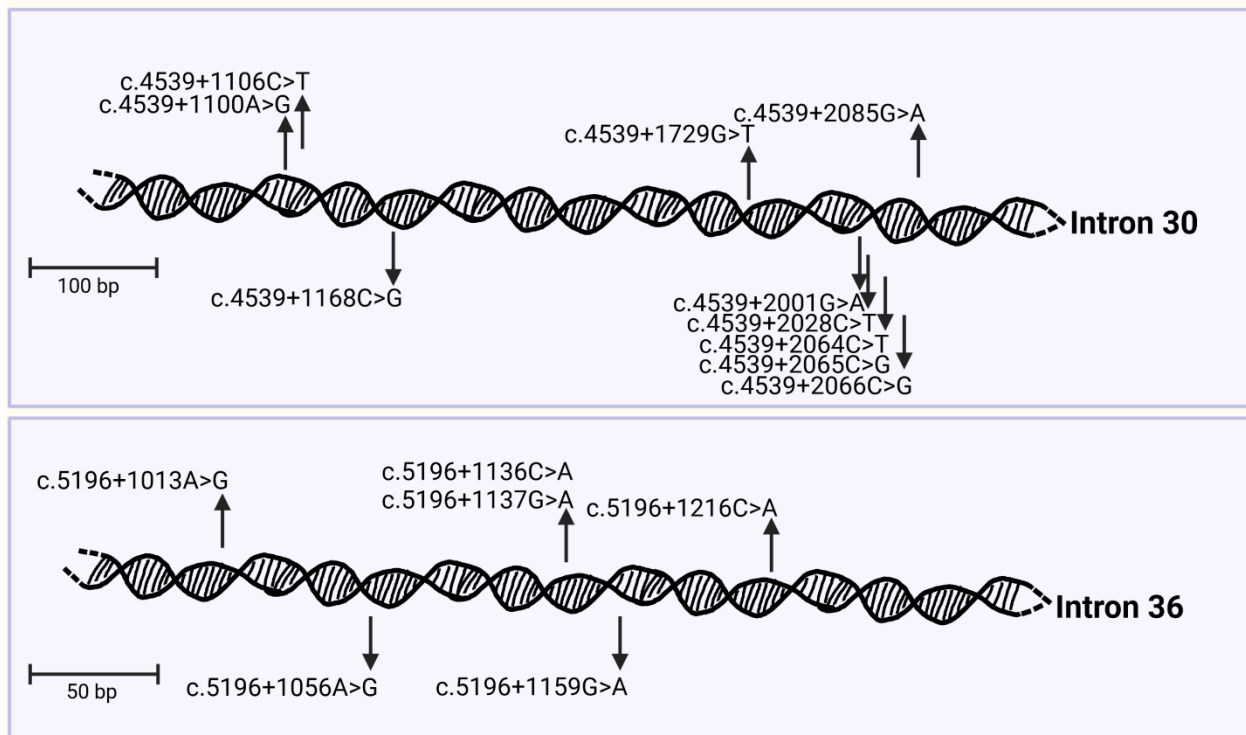
Figure 3: Location of the clustered deep-intronic variants in intron 30 (top part) and 36 (bottom part) of *ABCA4*.

2.4.2   The precursor messenger RNA splicing mechanism

The human genome is estimated to be comprised for 97% of "junk" DNA, which refers to genetic sequences not directly encoding for proteins, part of which is represented by introns (Palazzo et al. 2014). Intronic sequences are intervening parts of a gene, located between exons, that are removed during the high-fidelity splicing process and are not present in the mature mRNA transcript of the gene. In order for the splicing process to correctly take place, the protein complex, known as spliceosome, has to recognize specific splice signals, namely splice acceptor, splice donor, and branch point, that locate at the intron-exon borders and in the intron itself. These 'canonical' splice sites are highly conserved sequences that serve for accurate definition and retention of exons from the pre-mRNA into the mature mRNA transcript. Relative to the intron, the 5' splice donor site is located at the 5' end of an intron and has an invariant -GT- sequence. Whereas, the 3' splice acceptor site is located at the 5' end of an intron and has an invariant -AG- sequence. The branch point is found 15-40 nt upstream of the 3' splice acceptor site and is characterized by a poly-pyrimidine nucleotide stretch (Sheth et al. 2006, Wilkinson et al. 2020). Of note, alternative splice signals have also been characterized, *e.g.* 5'-AT, 3'-AC introns (Wu & Krainer, 1997). Nevertheless, these conserved splicing signal elements alone are insufficient to delineate and discriminate true constitutive exons from the vast majority of

pseudoexons (PE) flanked by pseudo-splice sites; it is the sequence context and presence of additional regulatory sequences that aids the process (Sun & Chasin, 2000). Specifically, by interacting with the spliceosome complex, additional sequences that act as exonic and/or intronic splicing silencers (ESS/ISS) or enhancers (ESE/ISE) allow the incorporation of constitutive exons in the mature mRNA transcript (Blencowe, 2000, Zhu et al. 2001).
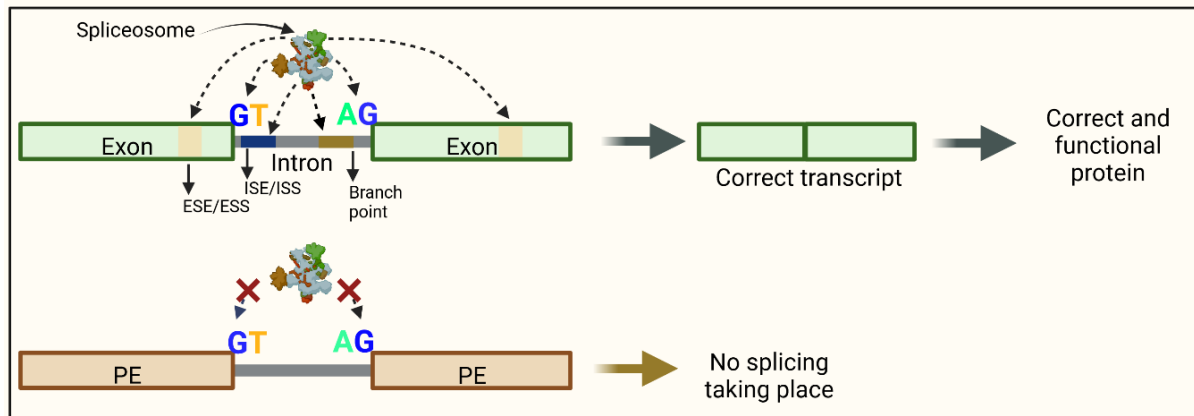
Figure 4: Splicing process highlighting elements involved in the recognition of constitutive exon boundaries. The spliceosome recognizes the splice donor (GT) and the splice acceptor (AG) sites of constitutive exons aided by additional elements in *cis* position, including the branch point, exonic splicing enhancer (ESE) and silencer (ESS), and intronic splicing enhancer (ISE) and silencer (ISS) sequences. Pseudoexons (PE) delimited by cryptic splice sites are not recognized by the spliceosome, and thereby not spliced, due to the lack of additional *cis* elements.

2.4.3 Pathological molecular mechanisms of deep-intronic variants and missplicing

Most commonly, DIVs lead to missplicing by incorporating intronic sequences in the mature mRNA molecule (Vaz-Drago et al. 2017). The incorporated sequence is referred to as a PE. Once the additional intronic sequence is retained, it disrupts the normal reading frame of the protein, often resulting in the formation of a premature stop codon or premature termination codon (PTC). These mutant aberrant mRNA transcripts presenting a PTC are then often degraded via the cellular nonsense-mediated decay pathway (NMD) (Zhang et al. 2009) (**Figure 5A**).

Besides the two highly conserved dinucleotides of the two canonical splice sites, the sequence encompassing the donor and acceptor sites show preferential nucleotides at certain given positions, determining the overall strength of a splice site (**Figure 5B**). As mentioned above, across the genome several sequences resembling canonical splice sites can be identified, namely either pseudo-splice sites or also called cryptic splice sites. Normally, the finely-tuned mechanism of splicing does not primarily recognize cryptic splice sites, thereby largely preventing intronic sequences in form of PEs from being

retained in the mature mRNA. However, the balance can be shifted if a cryptic splice site is strengthened by a variant or mutation, thereby enabling spliceosome recognition of a PE. In such circumstances, an intronic sequence delimited by the same elements as a canonical exon is aberrantly spliced, resulting in a missspliced mRNA transcript introducing a PE into the mature mRNA.
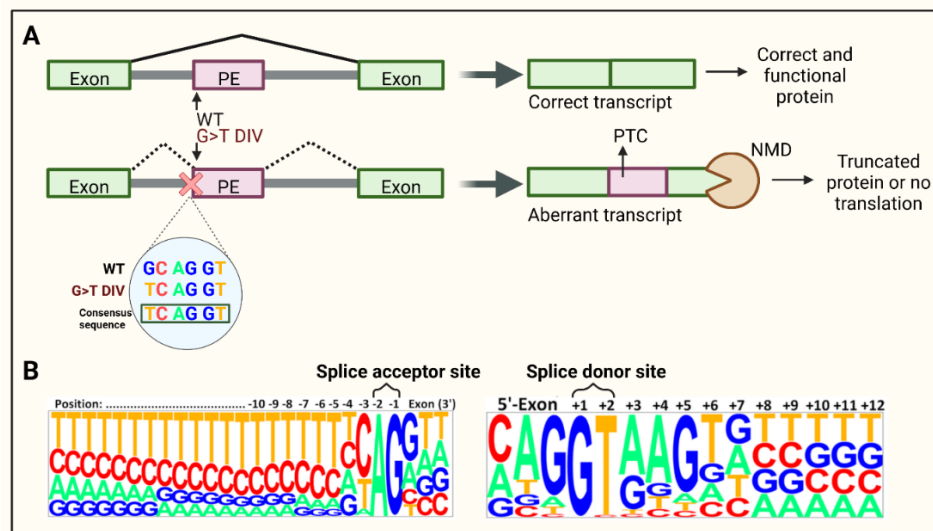
Figure 5: Missplicing process induced by deep-intronic variants (**A**) Representative example of the missplicing mechanism due to a G>T deep-intronic variant (DIV). The wild-type (WT) sequence does not resemble the consensus sequence of the splice acceptor site, while the introduction of a G>T DIV strengthens the WT cryptic splice site resulting in the retention of a pseudoexon sequence (PE). The aberrant transcript contains a premature termination codon (PTC), triggering nonsense mediated mRNA decay (NMD) that degrades the aberrant transcript. (**B**) Preferential composition of the sequences surrounding the splice acceptor and donor sites. The size of the letters indicates the frequency of occurrence of the nucleotides at the given position (Ma et al. 2015).

Consequently, a DIV can induce missplicing either through generation of a cryptic splice site or strengthening of a cryptic splicing enhancer delimitated by opportune cryptic splice sites.

An important example of strengthening of cryptic splicing enhancers are the clustered DIVs in intron 30 of the *ABCA4* gene found in patients with STGD1: the DIVs c.4539+2001G>A, c.4539+2028C>T, c.4539+2064C>T and c.4539+2065C>G all result in the formation of the same 345 bp pseudoexon (Albert et al. 2018, Bauwens et al. 2019) (**Figure 3**). Depending on the sequence features of the single DIVs, the ratio of correct to aberrant mRNA transcripts can largely vary, translating in different phenotype severities (Sangermano et al. 2018).

2.4.4   Clustered deep-intronic variants in intron 36 and the isolated variant c.5197-557G>T

A 203 nt sequence in intron 36, spanning from c.5196+1013 to c.5196+1216 upstream of exon 36, is rich in cryptic splice site sequences. In 2013, Braun *et al* reported and characterized seven intronic variants present in a cohort of 28 STGD1 patients with only one identified *ABCA4* disease-causing allele prior to the analysis of the intronic sequences of *ABCA4*. Three of these variants localized in intron 36: c.5196+1056A>G (V3), c.5196+1137G>A (V1), c.5196+1216C>A (V2) (Braun et al. 2013). Following an augmented molecular genetic testing of *ABCA4* in a cohort of Belgian STGD1 patients, two novel DIVs in intron 36 were identified: c.5196+1136C>A (M8) and c.5196+1159C>A (M10) (Bauwens et al. 2015). Later, Schulz *et al* identified in their cohort another DIV (c.5196+1013A>G, M4), predicted to strengthen a cryptic donor splice site (Schulz et al. 2017). In total, six clustered DIVs have been identified in intron 36. The pathogenicity of the V1, V2, V3 and M4 variants was confirmed by functional splicing experiments, showing clear missplicing patterns (Khan et al. 2020a). Conversely, the presence of M8 and M10 did not alter the splicing pattern, therefore, functional characterization was not confirmative for these (Khan et al. 2020a). In detail, c.5196+1013A>G and c.5196+1056A>G result in the preponderant retention of a 129 bp (PE129) and 177 bp (PE177) pseudoexon, respectively, which is defined by an already existing cryptic acceptor site and the strengthening of two different cryptic donor sites by either DIV. DIVs c.5196+1137A>G and c.5196+1216C>A induce the insertion of the same 77 bp pseudoexon (PE77) by strengthening a cryptic splice acceptor site and a cryptic splice donor site, respectively. For either DIV, the fraction of the aberrant transcript containing the pseudoexon differs considerably in minigene splicing assays performed in HEK293T cells (Khan et al. 2020a).

Besides clustered DIVs, an isolated variant in intron 36 was also identified 557 nt downstream of exon 37 (c.5197-557G>T). This DIV results in the retention of a 188 bp pseudoexon (PE188) by strengthening a cryptic donor site (Bauwens et al. 2019) (**Figure 6**).
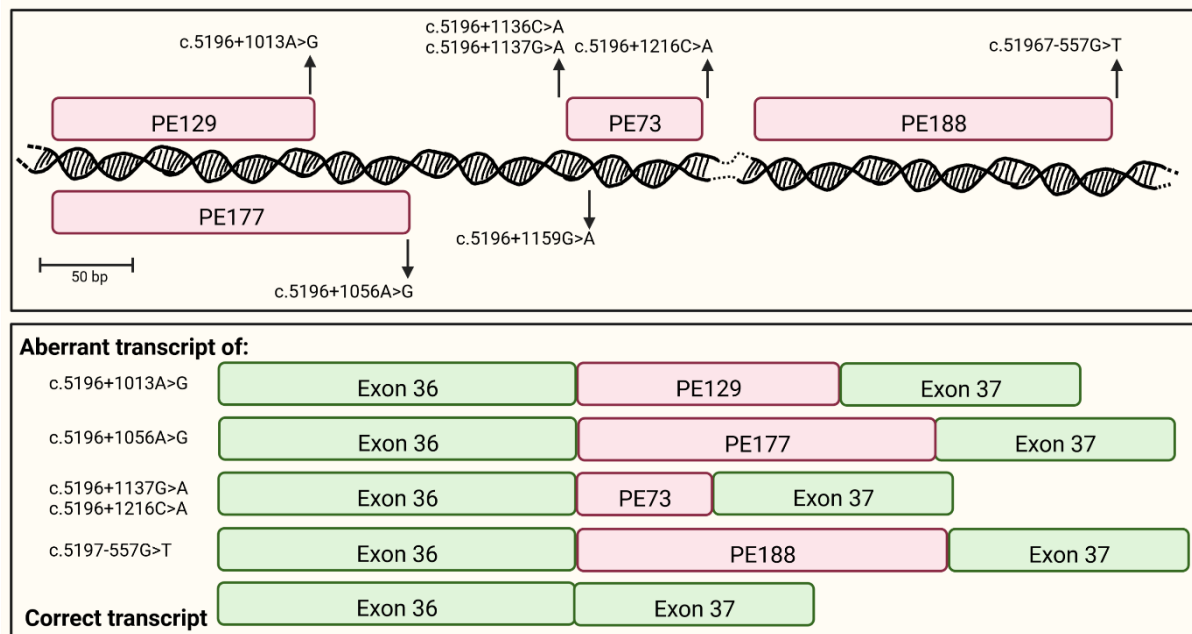
Figure 6: Details on the clustered deep-intronic variants in intron 36. Location and size of the various pseudoexon sequences (PE) in relation to the deep-intronic variants in intron 36 of *ABCA4* (top part). Configuration of the aberrant transcripts determined by the various deep-intronic variants and the expected correct transcript (bottom part).

2.5 GENE THERAPY FOR INHERITED RETINAL DISORDERS

In its broadest definition, gene therapy is defined by the US Food and Drug Administration (FDA) as products "*that mediate their effects by transcription and/or translation of transferred genetic material and/or by integrating into the host genome and that are administered as nucleic acids, viruses, or genetically engineered microorganisms. The products may be used to modify cells in vivo or transferred to cells ex vivo prior to administration to the recipient*" (Wirth et al. 2013).

2.5.1 Gene augmentation therapy in inherited retinal disorders

To date, only one gene therapy has been approved for IRDs, being Luxturna® (voretigene neparvovec). By replacing the function of the endogenous gene with adeno-associated virus (AAV)-delivered genetic material, this treatment aims to treat a form of autosomal recessive IRD caused by mutations in the *RPE65* gene (Pierce & Bennett, 2015). The molecular strategy behind the treatment is gene augmentation therapy, characterized by supplying the cells with a correct copy of the mutated gene that is delivered via viral vectors, most commonly AAVs, in case of Luxturna specifically the AAV2

serotype. The greatest advantage is represented by the mutation independent nature of the therapeutics, therefore overcoming genetic heterogeneity due to different mutations.

However, not all genes relevant for IRDs can be delivered via AAVs as the packaging capacity is limited to ~4.8 kb (Samulski & Muzyczka, 2014). The coding sequence DNA (cDNA) of *ABCA4* (>6.8 kb)*,* for example, exceeds this limit by far, hereby hampering the establishment of gene augmentation therapy for STGD1 and requiring the investigation of mutation-specific approaches (Allikmets et al. 1997). Nonetheless, attempts to deliver *ABCA4* cDNA via dual-AAV strategies, exploiting intein-mediated *trans*-splicing, are underway (Tornabene et al. 2019, Tornabene et al. 2012).

2.5.2   Antisense oligonucleotides for pathogenic deep-intronic variants

Among targets for the establishment of mutation-specific therapeutics, DIVs are attractive candidates. Due to their localization in the non-coding part of a gene, DIVs are particularly amenable to gene therapy as unintended alterations of surrounding sequences are likely to have no detrimental effect.

A particularly appealing therapeutic approach is represented by splicing modulation through antisense oligo-nucleotide molecules (ASO or AON), which can bind the aberrant pre-mRNA generated by DIVs and prompt correct splicing. Cryptic splice sites and cryptic regulatory sequences involved in aberrant splicing (e.g. ESE) are recognized by the spliceosome, however, AONs bind the pre-mRNA, masking essential elements for PE retention and preventing spliceosome from recognizing them (Bacchi et al. 2014, Hammond & Wood, 2011). Splicing modulation via AON molecules allows straightforward preclinical screening and easy scalability in *in vivo* models (Garanto & Collin, 2018). Aberrant splicing induced by clustered and isolated DIVs in *ABCA4* has been proven rescuable by AONs in cellular models of STGD1 (Albert et al. 2018, Bauwens et al. 2019, Khan et al. 2020a). Of particular interest is the ability of some single AONs to rescue different DIVs in the *ABCA4* intron 36 cluster, allowing to target multiple DIVs with the same molecule (Khan et al. 2020a). Splicing modulation by AONs has been also investigated for other IRD genes (e.g. *USH2A*), producing promising results which have enabled to start different clinical trials (NCT03780257, NCT05158296) (Dulla et al. 2021, van Diepen et al. 2019).

Practically, AON molecules are administered intravitreally in the eye, a relative simple procedure that can be performed by general ophthalmologists on an outpatient basis. However, as AONs act at the mRNA level, recurrent administrations are required to maintain therapeutic levels of the molecules, rendering this gene therapy not permanent, while increasing the risks associated with multiple

intravitreal injections (e.g. intraocular inflammation, retinal detachment, ocular hemorrhage, and cataract) (Falavarjani & Nguyen, 2013, Gaudana et al. 2010).

2.5.3 Genome editing

The possibility to permanently correct a genetic defect is a promising field of gene therapy, empowered by molecules able to determine breaks at the genomic levels (Li et al. 2020). Previously, complex molecular scissors, known as zinc-finger nucleases (ZNFs), paved the way towards editing of the genome. ZNFs are composed of a zinc-finger DNA-binding domain, providing target specificity, fused to the non-specific nuclease domain of the *FokI* restriction enzyme (Urnov et al. 2010). Alternatively, transcription activator-like effector nucleases (TALENs) and meganucleases were also used for genome editing. TALENs are composed of the same *FokI* domain as the DNA cleavage part and a series of tandem modules able to recognize single nucleotides (Sun & Zhao, 2013). Meganucleases are sequence-specific endonucleases, which were initially found in *Saccharomyces cerevisiae* and subsequently also characterized in eukaryotes, bacteria and archaea (Pâques & Duchateau, 2007). Despite the potential application of the aforementioned approaches, the need for complex molecule engineering to allow specific recognition of a sequence of interest has made their implementation challenging and laborious (Gaj et al. 2016).

2.5.3.1   *The CRISPR/Cas system*

CRISPR/Cas technologies (CRISPR, clustered regularly interspaced short palindromic repeats; Cas, CRISPR-associated protein) have revolutionized biological and biomedical research by easing targeted genome editing and virtually enabling unlimited genomic manipulation. The system is composed of an RNA molecule, directing the activity of the associated RNA-guided endonucleases (commonly Cas9) towards a specific target sequence, resulting in the generation of a double-strand DNA break (DSB). The sequence of the guide RNA (gRNA) can be easily customized, therefore allowing specific targeting (**Figure 7**). The CRISPR/Cas system was first discovered in microorganisms as an adaptive immune system and only later harnessed for genome editing purposes (Adli, 2018).
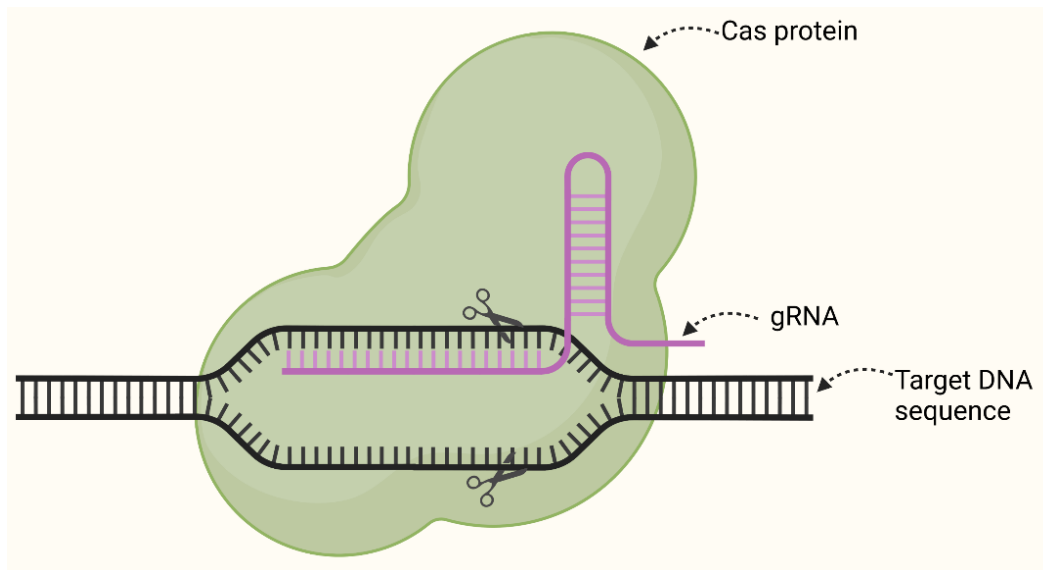
Figure 7: Simplified illustration of the CRISPR/Cas system generating a double-strand break (DSB) at a target DNA sequence. The Cas protein determines the formation of a DSB, whereas the gRNA (guide RNA) specifies the targeted DNA sequence.

2.5.3.1.1   Story of the discovery of CRISPR/Cas systems

The first evidence of the CRISPR array in *Eschericia coli* was reported in 1987, when Ishino and colleagues described an unusual sequence element of a series of repeats of 29 nt (Ishino et al. 1987). In 1995, Francis Mojica described the existence of regularly spaced repeats in archaea (Mojica et al. 1995). Later, in 2002, the term CRISPR was coined for the first time, when also *Cas* genes were identified (Jansen et al. 2002). However, until that point, no scientist could elucidate the function and anticipate the potential associated with the CRISPR/Cas system. It is only in 2007 that Barrangou and colleagues experimentally demonstrated the function of the CRISPR/Cas9 system in protecting bacteria against viruses (Barrangou et al. 2007). The breakthrough in the implementation of CRISPR/Cas9 arrived in 2012, when Doudna and Charpentier harnessed the CRISPR/Cas9 system from *Streptococcus pyogenes* to cut both double-strand DNA and plasmids *in vitro* (Jinek et al. 2012). This represents the first proof that the CRISPR/Cas9 system could be programmed to target specific sequences. In January 2013, the CRISPR/Cas9 system was used for the first time for genome editing in human cells (Jinek et al. 2013). To date, more than 18,000 patent application entries are listed on Google® patent under "CRISPR", highlighting the potential and broad range of applicability of this system (last query 24.08.2022).

2.5.3.1.2   Natural function of CRISPR/Cas systems

The CRISPR/Cas system is an adaptive immune system naturally occurring in some bacteria and archaea, which provides defense against viruses and plasmids. Upon the first infection, the system works by detecting, cleaving and storing foreign DNA into an arrayed DNA locus, namely CRISPR. When a second infection with the same pathogen happens, the invading DNA is recognized and cleaved by the Cas effector protein (Koonin & Makarova, 2009). Depending on unique signature effector genes involved in the defense process, phylogenetic and comparative genomic studies were used to classify the various CRISPR/Cas systems in three main types: Cas3 in type I systems, Cas9 in type II and Cas10 in type III (Brierley et al. 2002). The defense process can be divided into three main stages: 1) acquisition of foreign genetic material and integration in the CRISPR array locus, 2) transcription, processing and maturation of the information stored in the CRISPR array locus (pre-crRNA), resulting in the CRISPR RNA (crRNA), and 3) reconstitution of the crRNA/Cas complex and silencing of the invading DNA or RNA material (**Figure 8**).
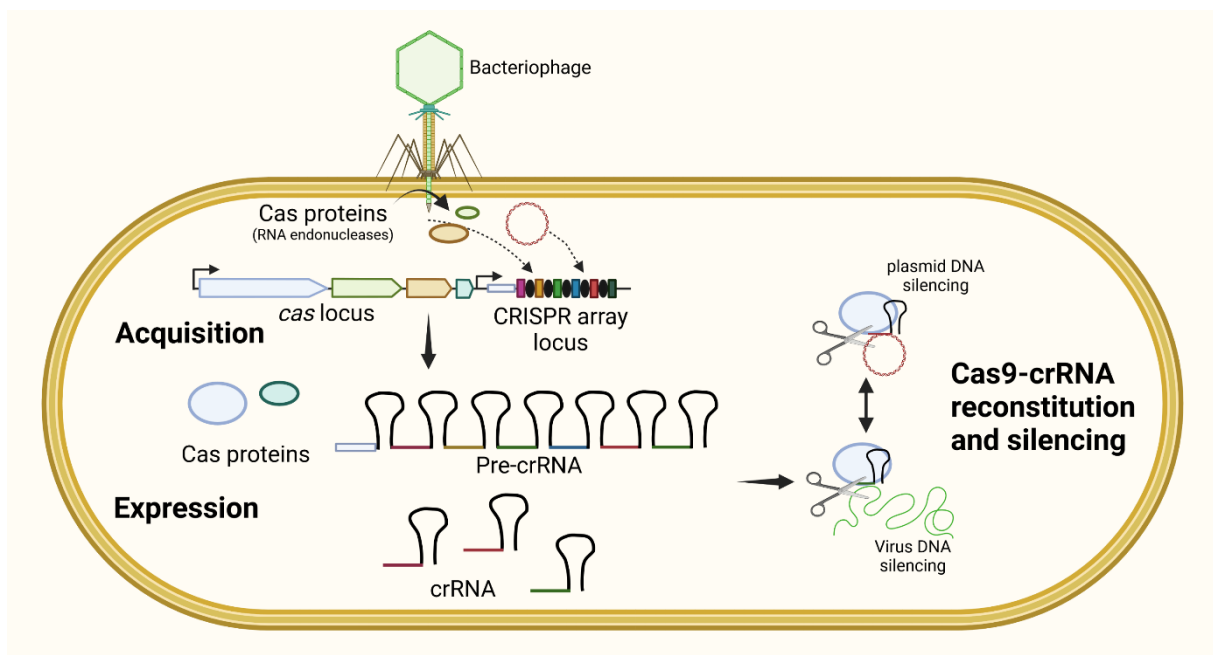
Figure 8: Illustration showing how the CRISPR/Cas system works as an adaptive immune system in bacteria. Upon infection with a bacterial pathogen (e.g. bacteriophage), the acquisition of foreign genetic material (e.g. virus or plasmid DNA) into the CRISPR array locus is mediated by Cas proteins (RNA endonucleases). The expression of Cas proteins and pre-crRNA is carried out by the endogenous cellular machinery. The pre-crRNA molecules are processed to generate crRNA. Upon a second infection, the effector Cas protein complexes with a crRNA and mediated the silencing of the foreign DNA material by specific cleavage.

For each stage, different key players of the CRISPR/Cas interference processes take part. Cas1 and Cas2 are highly conserved across the three different types and they are involved in the acquisition process and integration of foreign genetic material in the CRISPR locus (Makarova et al. 2015, Yosef et al. 2012). In the second step, the maturation of the pre-crRNA into crRNA is carried out by dedicated RNA endonucleases (CasE, Cas6, Csy4) or through mechanisms involving bacterial RNase III (Deltcheva et al. 2011). Depending on the CRISPR/Cas system type, the final step is carried out by different effector proteins: type I and type III systems consist of one large effector unit and several small subunits, whereas, type II only contains one large effector protein. At this stage, the crRNA molecule complexes with a multi-Cas protein complex (for type I and III) or a single effector (for type II), finalizing the interference process (Makarova et al. 2015).

2.5.3.1.3   Features of *Sp*Cas9

Among the most studied and used CRISPR/Cas systems is the Cas9 effector protein from *Streptococcus pyogenes* (*Sp*Cas9, type II system). The system requires three main elements for functioning: the Cas9 protein, a crRNA and a transactivating crRNA (tracrRNA) (Jinek et al 2012). While crRNA contains the sequence complementary to a target protospacer DNA, tracrRNA is a constant ~70 nt sequence forming a tetra loop secondary structure that is needed for *Sp*Cas9 activation and stabilization (Jinek et al. 2012, Nishimasu et al. 2014). The optimal crRNA length is 20 nt, although sequences as short as 17 nt can also be employed (Jinek et al. 2012). crRNA and tracrRNA interact through Watson-Crick base pairing, making the full-length gRNA that drives *Sp*Cas9 nuclease activity (Nishimasu et al. 2014). Of note, by fusing the crRNA and tracrRNA through a 4 bp linker, the RNA features required for the *Sp*Cas9 system to work can be included in a single chimeric gRNA (sgRNA) (Jinek et al. 2012, Nishimasu et al. 2014). DNA cleavage mediated by *Sp*Cas9 is ultimately specified by the occurrence of complementarity between the crRNA and the target protospacer DNA and a short motif (referred to as the protospacer adjacent motif, or PAM) that has to be disposed next to the complementary region in the target DNA (for *Sp*Cas9 is defined by the NGG PAM trinucleotide) (Jinek et al 2012, Shah et al. 2013) (**Figure 9A**).

The *Sp*Cas9 protein is formed of 1,368 amino acids (4,104 bp). Structurally, *Sp*Cas9 shows a bilobed architecture, comprising of a recognition (REC) lobe and a nuclease (NUC) lobe (**Figure 9B**). A PAM-interacting (PI) domain and two separated nuclease domains (RuvC and HNH) can be identified in the NUC lobe. The PI domain is formed of a positively charged surface able to interact with the PAM sequence of the gRNA, therefore providing PAM specificity (**Figure 9C**). The RuvC nuclease domain is located at the amino terminus and is involved in the cleavage of the non-target (or non-

complementary) DNA strand. Oppositely, the HNH domain cleaves the complementary (or target) DNA strand and is located in the middle of the *Sp*Cas9 protein. By targeted mutagenesis of residues of either nuclease domain, the nuclease activity of *Sp*Cas9 can be alternatively inactivated (p.D10A inactivates RuvC, while p.N863A has been found to inactivate HNH), generating a *Sp*Cas9 nickase enzyme only able to nick the target DNA, meaning cutting only one strand (Nishimasu et al. 2014).
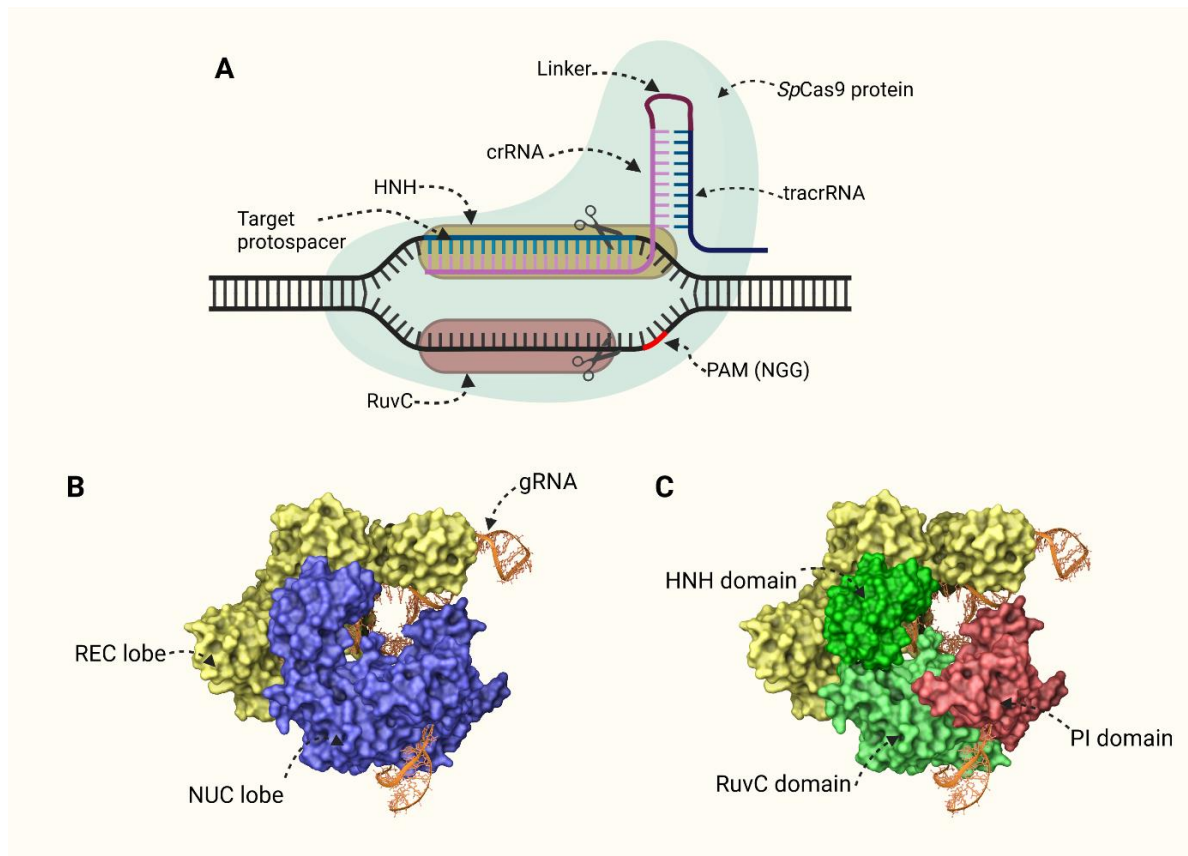
Figure 9: Structural conformation and details of the *Streptococcus pyogenes* Cas9 (*Sp*Cas9) protein (**A**) Detailed illustration of the CRISPR/*Sp*Cas9 system. The CRISPR RNA (crRNA) and the transactivating CRISPR RNA (tracrRNA) complex together forming the full-length guide RNA (gRNA). The two RNA molecules can also be linked through a linker to generate a single chimeric gRNA. The *Sp*Cas9 protein coupled to the gRNA induces the formation of a double-strand break that is three nucleotides upstream of the protospacer adjacent motif (PAM). The DNA sequence complementary to the gRNA is regarded to as protospacer sequence. The HNH and the RucV nuclease domains cleave the target and non-target strands, respectively. (**B**) 3D conformation of *Sp*Cas9. The recognition (REC) and nuclease (NUC) lobes are colored in yellow and blue, respectively. (**C**) The NUC lobe is composed of a PAM-interacting (PI) domain (red), a RuvC domain (light green) and a HNH domain (dark green). (**B**, **C**) The 3D rendering pictures were generated on Maestro, Schrödinger (https://www.schrodinger.com) using the 4OO8 PDB file (Nishimasu et al. 2014).

Besides *Sp*Cas9, other Cas9 effector proteins from others microorganisms have been identified and empowered for gene editing, especially for their smaller size as compared to *Sp*Cas9. Examples are

Staphylococcus auricularis (*Sauri*Cas9 – 1,061 amino acids) and *Streptococcus aureus* Cas9 (*Sa*Cas9 – 1,053 amino acids – 3,159 bp) (Friedland et al. 2015, Hu et al. 2020).

2.5.3.1.4   CRISPR/Cas – double-stranded DNA break repair pathways in mammalian cells

The molecular and cellular mechanisms leveraged by CRISPR/Cas to determine genome editing are multiple (**Figure 10**). Upon DNA cleavage mediated by Cas proteins, the cells react to the damage by activating pathways involved in the repair of the DSB. Three main repair pathways can be described in eukaryotic cells: non-homologous end joining (NHEJ), microhomology-mediated end joining (MMEJ) and homologous-directed recombination (HDR). The repair pathway mainly depends on the DNA sequence end configuration at the breakpoint, stage of the cell cycle and availability of a donor template (for HDR) (Chang et al. 2017, Moynahan & Jasin, 2010). In NHEJ, the two generated ends are recognized by the Ku70-Ku80 heterodimer (Ku), which acts as a holding and recruiting platform for other proteins involved in the repair pathway (Walker et al. 2001). Depending on the DNA end configuration, different key players participate in the DSB repair process: XRCC4 and DNA ligase IV are sufficient to repair DNA blunt end breaks with compatible ends (Grawunder et al. 1997), on the contrary, incompatible DNA ends require more extensive processing by additional proteins (e.g. Artemis endonuclease, and DNA polymerase μ and λ) before DNA ligase IV joins the ends (Capp et al. 2006, Ma et al. 2005, Moon et al. 2014). In MMEJ, the annealing of the two generated ends is mediated by microhomology sequences (2 bp – 20 bp) (Ma et al. 2003, Sfeir & Symington, 2015). While end resection is mostly prevented in NHEJ by Ku bound to the generated ends, in MMEJ, extensive resection is carried out by carboxy-terminal binding protein interacting protein (CtIP) and the MRN (MRE11–RAD50–NBS1 (Nijmegen breakage syndrome protein 1)) complex to display the microhomology sequences (Mimitou & Symington, 2010, Sfeir & Symington, 2015). Ultimately, DNA polymerase θ extends one resected 3' DNA end, stabilizing the annealing, whereas flap endonuclease 1 (FEN1) removes the resulting overhanging flaps, hereby promoting ligation by DNA ligase I or DNA ligase III (Wyatt et al. 2016, Liang et al. 2005). NHEJ and MMEJ result in sequence perturbation at the targeted site, therefore being regarded to as error-prone repair pathways (Allen et al. 2019). Conversely, HDR results in error-free repair of the DNA break. It exploits extensive strand annealing (>100 bp) between the strands at the targeted locus and a donor sequence template homologous to the targeted sequence itself (Lisby & Rothstein, 2015). RAD51 is the key player in HDR that resolves the break by coordinating strand exchange and annealing (Moynahan & Jasin, 2010) (**Figure 10**). The occurrence of the three repair mechanisms is strongly influenced by the stage of the cell cycle.

Extensive end resection required for MMEJ and HDR is more active during S and G2 phase, whereas NHEJ is active throughout the whole cell cycle (Ira et al. 2004). This translates in very low rate of HDR events occurring in post mitotic cells (Jeggo & Lobrich, 2007).
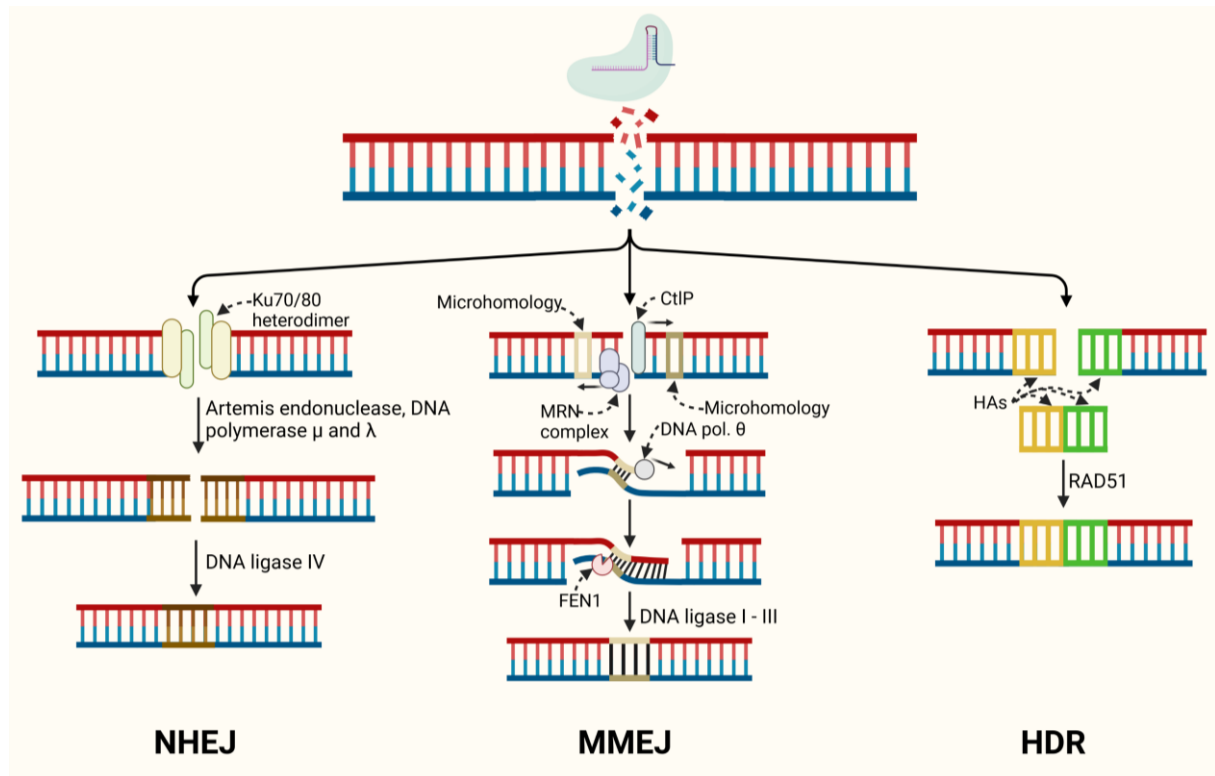
Figure 10: Mechanism of DNA double-strand break (DSB) repair. Non-homologous end joining (NHEJ) involves the initial recognition of the generated ends by the Ku70/80 heterodimer. The DNA ends are further processed by Artemis endonuclease, DNA polymerase μ and λ to generate compatible DNA extremities for ligation by DNA ligase IV. Microhomology-mediated end joining (MMEJ) requires the presence of microhomology sequences in proximity of the DSB. The carboxy-terminal binding protein interacting protein (CtIP) and the MRN (MRE11–RAD50–NBS1 (Nijmegen breakage syndrome protein 1)) complex resect the 5' ends to display the microhomology sequences. DNA polymerase θ extends one resected 3' DNA end, determining a more extensive annealing. The flap endonuclease 1 (FEN1) removes the resulting overhanging flaps and DNA ligase I or DNA ligase III resolve the DSB. Homology-directed repair (HDR) results in the incorporation of a donor template at the DSB through extensive annealing of homology arms (HA). RAD51 is the key player in the recombination process.

2.5.3.1.5   CRISPR/Cas – Targeting of disease-causing mutations in inherited disorders

Due to the ease with which the CRISPR/Cas systems can be programmed to target a genomic sequence to permanently induce change, their implementation to address certain disease-causing mutations would offer the promise to virtually cure any human inherited disorders (Li et al. 2020). By providing the cells with a HDR donor DNA template containing the correct copy of the mutated endogenous sequence, Cas targeted to a genetic locus could be used to determine correction of disease-causing single nucleotide mutations, as well as, small deletions or insertions (Kim H. & Kim J.S., 2014). Dominant and recessive mutations would be amenable to correction through HDR, making the CRISPR/Cas system the holy grail of the fight against any genetic disease (Kocher et al. 2017, Jacków et al. 2019).

Nonetheless, this strategy is not practicable for *in vivo* approaches targeting post mitotic cells, but it is primarily usable for *ex vivo* editing of cells with mitotic potential (e.g. induced pluripotent stem cells or hemopoietic stem cells), reducing by far the scope of CRISPR/Cas application (Orthwein et al. 2015, Yang et al. 2020). On the other hand, NHEJ is largely applicable *in vivo*, therefore representing an attractive therapeutic approach for targeting dominant mutations and DIVs (Cox et al. 2015, Gillmore et al. 2021, Ruan et al. 2017). The implementation of NHEJ triggered by Cas allows the perturbation of the targeted sequences, which, depending on the repair outcomes, can result in gene silencing or alteration of splicing elements (Allen et al. 2019, Shen et al. 2018). To date, two *in vivo* clinical trials using Cas9 and exploiting NHEJ have been initiated NTLA-2001 (sponsored by *Intellia therapeutics* - NCT04601051) and EDIT-101 (sponsored by *Editas medicine* - NCT03872479.) NTLA-2001 aims to treat transthyretin amyloidosis, an autosomal dominant life-threatening disease characterized by progressive accumulation of misfolded transthyretin (TTR) protein in tissues. The *Sp*Cas9 system is programmed to disrupt the activity of the *TTR* gene, and it is delivered in patients via systemic injection of lipoparticles, containing *Sp*Cas9 mRNA and a synthetic sgRNA, that are targeted to liver cells (Gillmore et al. 2021). Of note, NTLA-2001 does not discriminate between mutant and wild-type allele, therefore resulting in complete loss of TTR and TTR activity. TTR is involved in the transport of retinol in the blood stream, however, studies in TTR-deficient mice showed tissue levels of retinol comparable between TTR$^-$ and wild-type mice, therefore suggesting alternative mechanisms able to compensate the lack of the specific transporter for retinol (van Bennekum et al. 2001).

EDIT-101 is a therapeutics for the treatment of a form of LCA10, caused by the common DIV c.2991+1655A>G in *CEP290* located between exon 26 and exon 27 (den Hollander et al. 2006, Maeder et al. 2019). The presence of c.2991+1655A>G results in the retention of a 128 bp PE in the mature

mRNA, which introduces a PTC with the first aberrantly encoded codon (den Hollander et al. 2006). In the case of EDIT-101, *Sa*Cas9 and two sgRNAs are delivered in LCA patients via subretinal injection of AAV serotype 5 particles. The activity of Cas9 is directed to generate an intronic deletion, encompassing c.2991+1655A>G, the pseudoexon sequence and surrounding sequences, therefore rescuing the correct splicing pattern (Maeder et al. 2019, Ruan et al. 2017).

2.5.3.1.6 Advanced CRISPR/Cas9-based molecular tools

Boosted by the unique features of the CRISPR/Cas9 system and fueled by the need to generate approaches able to correct disease-causing mutations in an *in vivo* setting, advanced molecular tools based on CRISPR/Cas9 with expanded features have been established. They are usually composed of Cas9 (most commonly Cas9 nickase), providing the binding specificity, and an additional enzyme that confers additional properties to the editing system.

2.5.3.1.6.1 Base editors

The first greatest breakthrough was represented by base editors formed of *Sp*Cas9 nickase fused to a cytidine or adenosine deaminases (Gaudelli et al. 2017, Koblan et al. 2018, Komor et al. 2016, Komor et al. 2017, Nishida et al. 2016). Cytosine base editors (CBE) can perform gene editing without inducing DSBs, by converting target cytosines to uracil. The uracil intermediate then pairs to an adenosine, resulting in C>G to T>A conversion (Komor et al 2016). In the cell, spontaneous deamination of cytosines in uracils is a frequent undesired event, therefore the cells have evolved a base excision repair process able to remove the incorporated nucleotides, rescuing the correct base pairing (Krokan et al. 2002, Kunz et al. 2009). Despite the importance of the base excision repair pathway, when deamination into uracil is part of the base editing process, the cells will respond alike, therefore reverting the base paring to the original C>G (Komor et al. 2016). To "protect" the U>G intermediate from base excision, an uracil glycosylase inhibitor (UGI) is included in the CBE protein (Komor et al. 2016, Komor et al. 2017). Similarly, to CBEs, adenosine base editors (ABEs) are capable of converting A>T to G>C, through the formation of an inosine intermediate (Gaudelli et al. 2017, Koblan et al. 2018, Richter et al. 2020) (**Figure 11**). Of note, C>G to T>A mutations represent the most common pathogenic single nucleotide variations reported in the ClinVar database (Landrum et al. 2016). The ability of CBEs and ABEs to correct genetic defects in animal models has been proven for several inherited disorders,

including, LCA, phenylketonuria, and Duchenne muscular dystrophy (Chang et al. 2006, Panagiotopoulos et al. 2020, Xu et al. 2020). Therapeutics based on the use of CBEs and ABEs are currently under preclinical investigations by *Beam Therapeutics* (https://beamtx.com).
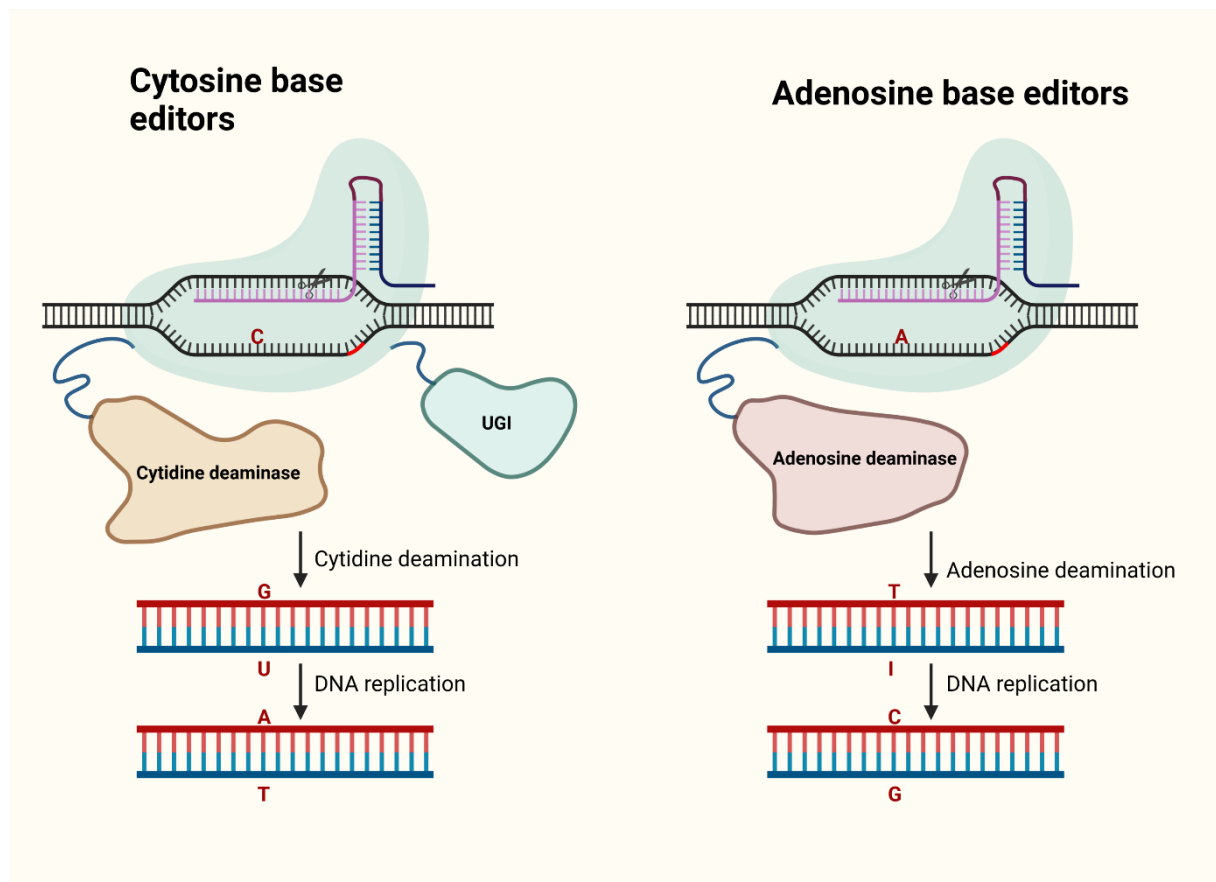
Figure 11: Illustration of the mode of action of base editors. Cytosine base editors (CBE) convert targeted cytosines (C) into uracils (U) mediated by the cytidine deaminase protein fused to Cas9 nickase. During DNA replication U is paired to adenosines (A) inducing C>G to T>A substitution. To prevent the excision of U an uracil glycosylase inhibitor (UGI) is included in the CBE protein. Adenosine base editors (ABE) convert targeted adenosines (A) into inosines (I) mediated by the adenosine deaminase protein fused to Cas9 nickase. During DNA replication, I is paired to C inducing A>T to G>C substitution.

2.5.3.1.6.2   Prime Editing

Besides base editors, another revolutionary DNA editing tool capable of correcting up to 89% of known genetic variants associated with human diseases, including deletion and insertion (in addition to all 12 possible base-to-base conversions), has been recently established: Prime Editing. Prime Editing combines the features of a *Sp*Cas9 nickase with an engineered reverse transcriptase. The system is

complemented with a prime editing gRNA (pegRNA), that contains a sgRNA sequence, a primer binding site that allows hybridization of pegRNA to one of the nicked DNA strands at the targeted locus, and a reverse transcriptase template containing the desired edited sequence. Upon binding of the prime editing system to the target locus, *Sp*Cas9 nickase cleaves one of the DNA strands, the primer binding site anneals to the nicked strand, while the reverse transcriptase copies the RNA template into DNA at the targeted site, resulting in DNA editing (Anzalone et al. 2019) (**Figure 12**). Prime editing is being harnessed by *prime medicine* for the development of therapeutics (https://primemedicine.com/) .
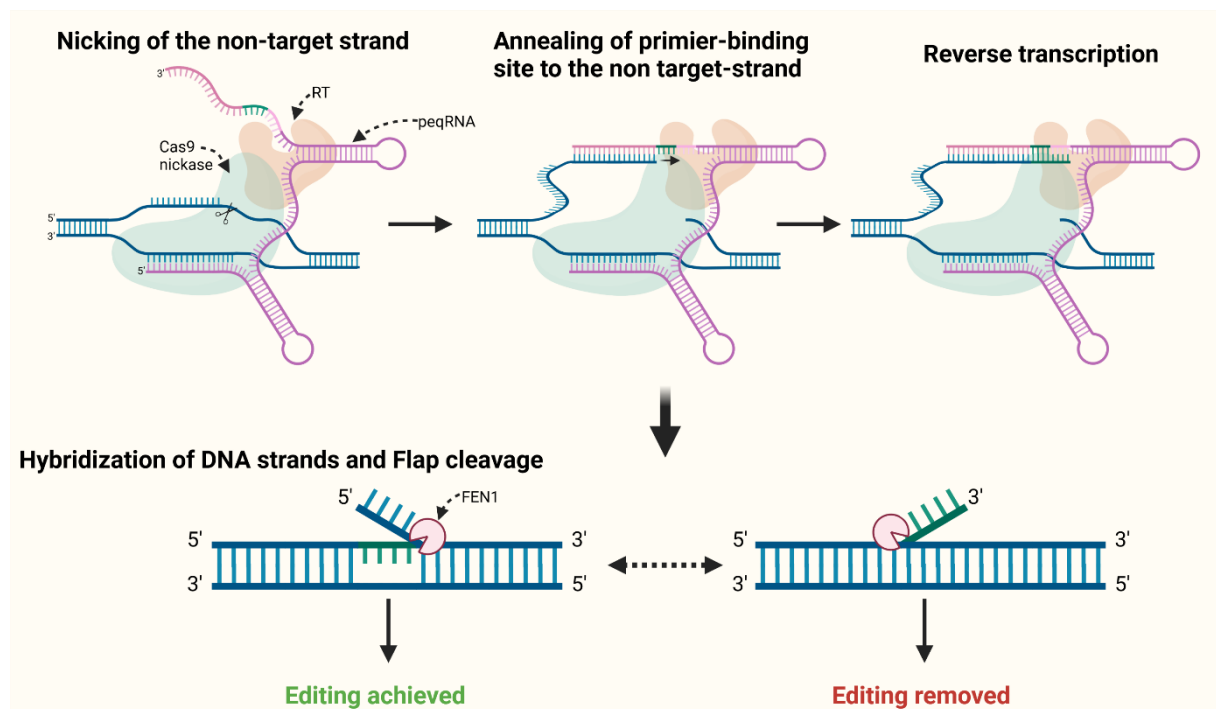
Figure 12: Illustration of the mode of action of prime editing. A Cas9 nickase is fused to a reverse transcriptase (RT) and complexes to a prime editing gRNA (peqRNA) composed of a gRNA, a primer-binding site and a template sequence. The Cas9 nickase mediates the nicking of the non-target strand, allowing its annealing to the primer-binding site. The template sequence is therefore copied on the non-target strand by RT. The balance of hybridization of DNA strands and the resulting flap cleavage by FEN1 determines the overall editing efficiency.

Of note, the described advanced CRISPR/*Sp*Cas9-based tools are only the two most significant examples for targeting genetic disorders, several more exist, including CRISPR activation (for gene transcription activation), CRISPR interference (for gene transcription repression) and dCas9-KRAB (for epigenetic editing) (Gilbert et al. 2013, Thakore et al. 2015). Moreover, other Cas9 variants have also been implemented in the establishment of advanced genome editing tools (e.g. *Nme2*Cas9, *Cj*Cas9 and *Sa*Cas9 etc) (Davis et al. 2022, Zhang et al. 2022).

2.5.3.1.7   CRISPR/Cas genome editing for pathogenic deep-intronic variants

In literature, different examples of use of the CRISPR/Cas systems to permanently address disease-causing DIVs have been reported. Cystic fibrosis is an autosomal recessive progressive disorder characterized by dysfunctional secretory epithelial cells, obstructing the lung airways and pancreatic ducts (Cutting 2015). The causative gene is *CFTR*, and two main DIVs (c.1679+1634A>G and c.3718-2477C>T) and an intronic splicing mutation (c.3140-26A>G) account for ~3% of the patient cases (Stephens & Schneider, 1992). By coupling mutation specific pairs of sgRNAs, Cas9-mediated targeted excision of the three mutations was established and validated in minigene splicing assay, showing substantial splicing rescue. In this case, *Sp*Cas9 delivered with two sgRNAs induced the deletion of genomic fragments (≥50 bp) containing the individual pathogenic variant (Sanz et al. 2017).

Furthermore, a different allele-specific genome editing approach, implementing a Cas protein from *Acidaminococcus sp.* (*As*Cas12a, also known as Cpf1), was designed to target *CFTR* c.3140-26A>G and c.3718-2477C>T DIVs. Single gRNAs overlapping with the mutations were tested for the ability to induce sequence aberrations (insertions or deletions) at the targeted mutant allele, hereby influencing spliceosome recognition of the strengthened cryptic splice sites. The strategy was tested in minigene assay, primary airway cells and intestinal organoids, demonstrating efficient splicing correction in the three experimental models and rescue of the phenotype in organoids (Maule et al. 2019).

With respects to DIVs in IRDs, the DIVs c.254-649T>G in *CLRN1* and c.2991+1655A>G in *CEP290* were addressed by means of CRISPR/*Sp*Cas9 and /*Sa*Cas9, respectively (Maeder et al. 2019, Panagiotopoulos et al. 2020). The DIV c.254-649T>G in *CLRN1* causes Usher syndrome type III, an autosomal recessively inherited disease characterized by progressive hearing impairment, RP, and vestibular dysfunction (Xu et al. 2020). Whereas, c.2991+1655A>G in *CEP290* is responsible for LCA10, an autosomal recessive condition, characterized by congenital or early onset and severe cone-rod dystrophy (Chang et al. 2006, den Hollander et al. 2006). In both cases, two sgRNAs were used to induce genomic deletions encompassing the DIV. The approach for c.2991+1655A>G, based on *Sa*Cas9, is currently investigated within a phase 1/2 single dose escalation clinical trial (BRILLIANCE trial - NCT03872479). The therapeutics consists of a single subretinal administration of AAV5 particles encoding two sgRNAs and *Sa*Cas9. Five patient cohorts are planned, including adult low ($6x10^{11}$ vg/ml), mid ($1.1x10^{12}$ vg/ml), and high ($3x10^{12}$ vg/ml) dose, and pediatric mid and high dose. For the adult low and mid dose cohorts (n=6), early data on safety and efficacy have been made public. Despite the absence of dose-limiting toxicities or serious adverse events, limited signs of potential efficacy were reported in only two of the three subjects in the mid dose cohort (BRILLANCE, Rd meeting, 2021).

3 AIM OF THE THESIS

This thesis purposes to establish and validate *Sp*Cas9-based genome editing for the rescue of DIVs in *ABCA4.* Considering the great numbers of characterized DIVs, representative ones were chosen in intron 36. Specifically, we selected an isolated (c.5197-557G>T) and seven clustered DIVs. Six of the latter (c.5196+1013A>G, c.5196+1056A>G, c.5196+1136C>A, c.5196+1137G>A, c.5196+1159C>A and c.5196+1216C>A) had already been functionally characterized for their effect on splicing, while a newly identified DIV c.5196+1134C>G was characterized by minigene assay in this project.

Three standard *Sp*Cas9-approaches to target the isolated variant were established and validated. In addition, a novel chimeric *Sp*Cas9 protein (ED*Sp*Cas9), able to induce larger deletions at the targeted site by employing sgRNAs, was further improved, tested, and compared to wild-type *Sp*Cas9 for the ability to effectively correct aberrant splicing determined by both the isolated and the clustered DIVs.

Preliminary screening of the designed approaches to target the DIV c.5197-557G>T was carried out through minigene splicing assay in HEK293T cells. The most promising strategies were then validated in a patient-derived cellular model comprised of induce pluripotent stem cells (iPSCs) differentiated in two-dimensional cone photoreceptor precursor cells (CPCs). Relative and absolute splicing rescue, as well as editing efficiency at the genomic levels were used as experimental readout.

Also for the clustered DIVs, minigene splicing experiments were used for preliminary screening of the most effective sgRNAs. Not surprisingly, the minigene splicing assay of c.5196+1136G>A and c.5196+1159C>A DIVs did not show any aberrant splicing patterns in HEK293T, WERI-Rb1 or control CPCs, therefore no additional rescue experiment was pursued for these. For these experiments, iPSCs, harboring individual clustered DIVs, were established either by reprogramming of patient-derived fibroblasts or variant knock-in by CRISPR/Cas in control iPSCs. The established cell lines will be used for future assessment of ED*Sp*Cas9 strategies in CPCs and potentially retinal organoid models.

4 MATERIALS

4.1 BACTERIA STRAINS

Escherichia coli DH10B: NEB® 10-beta Competent *E. coli*   New England Biolabs GmbH,

(High Efficiency), Genotype: Δ(ara-   Frankfurt am Main, Germany

leu)7697araD139 fhuA ΔlacX74 galK16 galE15

e14- φ80dlacZΔM15 recA1 relA1 endA1

nupG rpsL (Str^R) rph spoT1 Δ(mrr-hsdRMS-mcrBC)

4.2 BUFFERS AND SOLUTIONS

Agarose Electrophoresis

Tris-Borate-EDTA-Buffer (TBE)	**Ethidium bromide solution**
89 mM Tris-Base, pH 8.0	100 µg/ml in TE
89 mM Boric acid	
2.5 mM EDTA	

Bacterial culture

LB-Medium	**LB-Agar- and media supplements**
1 % (w/v) Bacto-Trypton	0.15 mg/ml Ampicillin
0.5 % (w/v) Yeast-Extract	0.05 mg/ml Kanamycin
170 mM NaCl	100 µM IPTG
pH 7.5	0.02 mg/ml X-Gal

DNA resuspension buffers

Tris-EDTA buffer (1xTE)	**¼ TE**
10 mM Tris-HCl, pH 7.5	2.5 mM Tris-HCl, pH 8.0
1 mM EDTA	0.25 mM EDTA

PCR: dNTP mix and buffer

dNTP-Mix	**10x buffer 18 (BOM 18)**
5 mM dATP	100 mM Tris-HCl, pH 8.9
5 mM dCTP	500 mM KCl
5 mM dGTP	30 mM $MgCl_2$
5 mM dTTP	
5 mM Tris, pH 7.0	

LB-Agar

1.5 % Agar in LB-Medium

Restriction enzymes Buffers

10x NEBuffer™ 1.1	New England Biolabs GmbH, Frankfurt am Main, Germany
10x NEBuffer™ 3.1	New England Biolabs GmbH, Frankfurt am Main, Germany
10x NEBuffer™ CutSmart® Buffer	New England Biolabs GmbH, Frankfurt am Main, Germany

10x DNA Loading buffer

20 % Ficoll 400

100 mM EDTA

0.2 % Bromphenol Blue

0.2 % Xylene Cyanol

4.3 CHEMICALS

Ampicillin sodium salt	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Bacto-Agar	MP Biomedicals Inc., Ohio, USA
Bromophenol blue	Merck KGaA, Darmstadt, Germany

Ethanol	MERCK, Darmstadt, Germany
Ethidium bromide	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Isopropanol	MERCK, Darmstadt, Germany
LE-Agarose	Scientific GmbH, Oldendorf, Germany
Potassium chloride	Sigma-Aldrich Chemie GmbH, Steinheim, Germany
Potassium hydroxyde	Fluka Chemie GmbH, Buchs, Switzerland
Sodium acetate-trihydrate	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Sodium butyrate	Sigma-Aldrich Chemie GmbH, Steinheim, Germany
Sodium chloride	MERCK, Darmstadt, Germany
Yeast extract	AppliChem, Darmstadt, Germany
Hi-Di™ Formamide	Applied Biosystems, Life Technologies Corporation, Carlsbad, CA, USA

4.4 CELL LINES

4.4.1 Commercial cell lines

HEK293T/17: Human embryonic kidney cells 293T clone 17 is a cell line derivative of the 293 cell line into which the gene for the simian virus 40 (SV40) large T antigen was inserted. The presence of the constitutively expressed SV40 large T antigen guarantee high efficiency of transfection for plasmids containing the SV40 origin of replication. This cell line was used for minigene splicing assay and production of lentiviral particles.

American Type Culture Collection, Manassas, VA, USA

WERI-Rb-1: The WERI-Rb-1 cell line is a human retinoblastoma cell line showing a morphology of grape-like clusters of round cells. Being of tumoral origin, the total

Original supplier unknown

chromosome number is 47 occurring at 38%, and the rate of polyploidy is 9%. It expresses several retina-specific markers such as *ABCA4* and *RHO*. This cell line was used for minigene splicing assay.

4.4.2   Patient-derived cell lines

ZD574-fibroblasts: This cell line was derived from the skin biopsy of a healthy proband heterozygous for the DIV *ABCA4* c.5197-557G>T obtained upon informed written consent complying with the guidelines and approved by the local ethics committee (project number 124/2015BO1).

Established in house. The cell line was tested negative for HIV, HBV and HCV infections.

ZD574-iPSCs: This cell line was obtained by Sendai reprogramming of the ZD574-fibroblasts through a commercial service.

Stem Cell Technology Center, Radboud University Medical Center, The Netherlands

CO-53-iPSCs: This cell line was obtained by episomal reprogramming of a fibroblast cell line obtained from a healthy proband.

Provided by Prof. Ludger Schöls and Dr. Stefan Hauser, DZNE, Tübingen under scientific collaboration terms. The cell line was tested negative for HIV, HBV and HCV infections.

MST433-fibroblasts: This cell line was derived from the skin biopsy of a Stargardt proband (*ABCA4* c.5603A>T; p.Asn1868Ile heterozygous and c.5196+1134C>G;p.? heterozygous) obtained upon informed written consent complying with the guidelines and approved by the local ethics committee (project number 124/2015BO1).

Established in house. The cell line was tested negative for HIV, HBV and HCV infections.

MST433-iPSCs: This cell line was obtained by episomali reprogramming of the MST433-fibroblasts .

Established in house

CO-53-KI-ABCA4-c.5196+1013A>G: This cell line was obtained by homozygous CRISPR/*Sp*Cas9-mediated knock-

Established in house

in of the c.5196+1013A>G deep-intronic variant in the CO-53-iPSCs cell line.

CO-53-KI-ABCA4-c.5196+1056A>G: This cell line was obtained by homozygous CRISPR/*As*Cas12a-mediated knock-in of the c.5196+1056A>G deep-intronic variant in the CO-53-iPSCs cell line.

Established in house

CO-53-KI-ABCA4-c.5196+1216C>A: This cell line was obtained by homozygous CRISPR/*Sp*Cas9-mediated knock-in of the c.5196+1216C>A deep-intronic variant in the CO-53-iPSCs cell line.

Established in house

4.5 Consumables

Commercial consumables (plastic ware, pipette tips, reaction vessels, etc.) were supplied by changing suppliers for laboratory materials.

4.6 Enzymes

Cas enzymes

Alt-R A.s. Cas12a (Cpf1) Nuclease	IDT Technologies, Breda, Netherlands
*Sp*Cas9 2NLS Nuclease	Synthego, Redwood City, CA, USA

DNA-Polymerases

DNA Polymerase I, Large (Klenow) Fragment	New England Biolabs GmbH, Frankfurt am Main, Germany
KAPA HiFi HotStart ReadyMix	Roche Holding, Basel, Switzerland
PfuUltra High-Fidelity DNA Polymerase	Agilent Technologies, Santa Clara, CA, USA

Q5® High-Fidelity DNA Polymerase   New England Biolabs GmbH, Frankfurt am Main, Germany

Taq DNA-Polymerase (5 U/µl)   Genaxxon Bioscience GmbH, Ulm, Germany

Restrictions enzymes

*Afe*I   New England Biolabs GmbH, Frankfurt am Main, Germany

*Afl*III   New England Biolabs GmbH, Frankfurt am Main, Germany

*Age*I   New England Biolabs GmbH, Frankfurt am Main, Germany

*Apa*I   New England Biolabs GmbH, Frankfurt am Main, Germany

*Bam*HI   Thermo Fisher Scientific, Waltham, MA, USA

*Bbs*I   New England Biolabs GmbH, Frankfurt am Main, Germany

*Blp*I   New England Biolabs GmbH, Frankfurt am Main, Germany

*Dpn*I   New England Biolabs GmbH, Frankfurt am Main, Germany

*Eco*RI-HF   New England Biolabs GmbH, Frankfurt am Main, Germany

*Kpn*I   New England Biolabs GmbH, Frankfurt am Main, Germany (discontinued)

*Not*I   New England Biolabs GmbH, Frankfurt am Main, Germany

*Sap*I   New England Biolabs GmbH, Frankfurt am Main, Germany

*Xba*I   New England Biolabs GmbH, Frankfurt am Main, Germany

Ligases

T4 DNA-Ligase (400 U/µl)   New England Biolabs GmbH, Frankfurt am Main, Germany

T7 DNA-Ligase (45 U/µl)   Mc Labor, Albstadt, Germany

Nucleases

Desoxy ribonuclease I   Merck KGaA, Darmstadt, Germany
(DNaseI for cDNA-Synthesis)

DNA Polymerase I   (Fermentas) Thermo Fisher Scientific, Waltham, MA, USA

ExoSAP-IT   Affymetrix Inc., Santa Clara, CA, USA

Plasmid-Safe™ ATP-Dependent DNase   Lucigen Corporation, Middleton, WI, USA

Other enzymes

T4 Polynucleotide Kinase (10 U/µl)   New England Biolabs GmbH, Frankfurt am Main, Germany

CIP (Calf Intestinal Alkaline   New England Biolabs GmbH, Frankfurt am Main, Germany
Phosphatase)

4.7   EQUIPMENT

2100 Bioanalyzer Instrument   Agilent Technologies, Santa Clara, CA, USA

7500 Real time PCR System   Applied Biosystems, Life Technologies Corporation, Carlsbad, CA, USA

8er-Strip-Centrifuge   Labnet International Inc., Woodbridge, NJ, USA

ABI Prism 3130XL Genetic Analyzer   Applied Biosystems, Life Technologies Corporation, Carlsbad, CA, USA

Balance PJ 400   Mettler-Toledo, Giessen, Germany

Bioer Mixing Block MB-102   Bioer Technology, Hangzhou, China

Block thermostat   Gebr. Liebisch GmbH & Co. KG Labortechnik, Bielefeld, Germany

CCD-Camera EASY 429K   Herolab, Wiesloch, Germany

Circulating water thermostat Julabo F30-HC   Labortechnik GmbH, Seelbach, Germany

CO_2-Incubator	Binder CB210 Binder GmbH, Tuttlingen, Germany
CO_2-Incubator	Heracell™ VIOS 160i Thermo Fisher Scientific, Waltham, MA, USA
CoolCell® LX	Corning, New York, NY, USA
DynaMag™-2 Magnet	Thermo Fisher Scientific, Waltham, MA, USA
EVOS™ XL Core Imaging System	Thermo Fisher Scientific, Waltham, MA, USA
Gel chambers	Bio-Rad Laboratories GmbH, Munich, Germany
Heraeus B6060 Incubator	Thermo Fisher Scientific, Waltham, MA, USA
Heraeus Biofuge 15	Thermo Fisher Scientific, Waltham, MA, USA
Heraeus Biofuge fresco	Thermo Fisher Scientific, Waltham, MA, USA
Heraeus Multifuge 1 L-R	Thermo Fisher Scientific, Waltham, MA, USA
Heraeus Freso 21 microcentrifuges	Heraeus®Fresco, Thermo Scientific, Waltham, WA, USA
Heraeus Pico 21 microcentrifuges	Heraeus®Fresco, Thermo Scientific, Waltham, WA, USA
Ice maker MF36 and Ice-storage	Scotsman Ice Systems, Vernon Hills, IL, USA
Magnetic stirrer RCT basic	IKAR-Werke GmbH & Co. KG, Staufen, Germany
Magnetic-Ring Stand (96 well)	Thermo Fisher Scientific, Waltham, MA, USA
Micro balance (Model CPA225D)	Sartorius AG, Gottingen, Germany
Microwave R-93ST-AA	Sharp Electronics GmbH, Hamburg, Germany
MiSeq™-system	Illumina, San Diego, CA, USA
NanoDrop ND-1000 Spektrophotometer	peqLAB Biotechnologie GmbH, Erlangen, Germany
Neon™ Electroporation Transfection System	Thermo Fisher Scientific, Waltham, MA, USA
PCR-Cycler Veriti 96 Well Applied Biosystems	Life Technologies Corporation, Carlsbad, CA, USA
Power supply phero-stab 300	Biotec-Fischer GmbH, Reiskirchen, Germany

Sorvall RC 28S centrifuge	Thermo Fisher Scientific, Waltham, MA, USA
Spark® Microplate reader	Tecan Group, Maennedorf, Switzerland
Sterile workbench BDK-S 1500	Luft- und Reinraumtechnik GmbH,Sonnenbuhl-Genkingen, Germany
Thermomixer comfort	Eppendorf AG, Hamburg, Germany
UV Transilluminator	Bachofer GmbH, Reutlingen, Germany
UV Transilluminator Chromato-VUE	UVP, San Gabriel, CA, USA
UV-Water System	Millipore GmbH, Schwalbach/Ts., Germany
Vortex Genie 2	Bender & Hobein AG, Zurich, Switzerland

4.8 Kits

AccuBlue® NextGen dsDNA Quantitation Kit	Biotium, Fremont, CA, USA
AMPure XP	Beckman Coulter, Pasadena, Ca, USA
BigDye™ Terminator v1.1 Cycle	Applied Biosystems, Life Technologies Corporation
Cell Chip™	Tecan Group, Maennedorf, Switzerland
DNA 1000 Kit	Agilent Technologies, Santa Clara, CA, USA
EASYstrainer™ (100 µm)	Greiner Bio-One GmbH, Frickenhausen, Germany
Maxima H Minus First Strand cDNA Synthesis Kit	Thermo Fisher Scientific, Waltham, MA, USA
MinElute Enzymatic Reaction Cleanup Kit	QIAGEN GmbH, Hilden, Germany
Monarch® Plasmid Miniprep Kit	New England Biolabs GmbH, Frankfurt am Main, Germany
NEBuilder® HiFi DNA Assembly Cloning Kit	New England Biolabs GmbH, Frankfurt am Main, Germany
Neon™ Transfection System 100 µl Kit	Thermo Fisher Scientific, Waltham, MA, USA

Nextera XT DNA Library Preparation Kit	Illumina, San Diego, CA, USA
NucleoSpin Gel and PCR Clean-up	Macherey-Nagel GmbH, Dueren, Germany
peqGold Tissue DNA mini Kit	VWR International, Radnor, PA, USA
peqGOLD, Total RNA Kit	VWR International, Radnor, PA, USA Erlangen, Germany
QuantiTect SYBR® Green PCR Kits	QIAGEN GmbH, Hilden, Germany
TA Cloning Kit	Invitrogen, Carlsbad, CA, USA

4.9 SIZE STANDARDS

DNA standards size markers

100 bp DNA Ladder	New England Biolabs GmbH, Frankfurt am Main, Germany
Gene Ruler 1 kb DNA Ladder	Fermentas GmbH, St. Leon-Rot, Germany
pcDNA3.1 Zeo/Taq I = KEB (in house manufactured)	Established by Katja Köppen, Molecular Genetics Laboratory, Augenklinik Tübingen

4.10 PREPARED REAGENTS AND MEDIA COMPONENTS

Accutase®	Innovative Cell Technologies, San Diego, CA, USA
Adenosine 5'-Triphosphate (ATP – 10 mM)	New England Biolabs GmbH, Frankfurt am Main, Germany
B-27 Supplement (50x)	Thermo Fisher Scientific, Waltham, MA, USA
CloneR™	Stemcell Technologies, Vancouver, Canada
Collagenase, Type IV, powder	Thermo Fisher Scientific, Waltham, MA, USA

Cycloheximide solution (100 mg/ml in DMSO)	Merck KGaA, Darmstadt, Germany
DMEM – Dulbecco's Modified Eagle Medium	Thermo Fisher Scientific, Waltham, MA, USA
DMEM/F12	Thermo Fisher Scientific, Waltham, MA, USA
DMSO - Dimethylsulfoxid	Sigma-Aldrich Chemie GmbH, Steinheim, Germany
dNTP-Mix (5 mM dATP, dTTP dGTP and dCTP each)	peqLAB Biotechnologie GmbH, Erlangen, Germany
DTT (Dithiothreitol), 0.1 M solution	Thermo Fisher Scientific, Waltham, MA, USA
Dulbecco's PBS	PAA Laboratories GmbH, Pasching, Austria
Essential 8™ Medium	Thermo Fisher Scientific, Waltham, MA, USA
Fetal bovine serum (FBS Gold)	PAA Laboratories GmbH, Pasching, Austria
Heparin Sodium	Merck KGaA, Darmstadt, Germany
Human Recombinant bFGF	Stemcell Technologies, Vancouver, Canada
Insulin-like Growth Factor-I human	Merck KGaA, Darmstadt, Germany
KnockOut™ Serum Replacement	Thermo Fisher Scientific, Waltham, MA, USA
LB-Medium	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Lipofectamine™ 3000	Thermo Fisher Scientific, Waltham, MA, USA
Lipofectamine™ Stem Transfection Reagent	Thermo Fisher Scientific, Waltham, MA, USA
MEM NEAA (100x) - minimum essential medium, non-essential amino acid solution	Merck KGaA, Darmstadt, Germany
N-2 Supplement (100x)	Thermo Fisher Scientific, Waltham, MA, USA
Penicillin/Streptomycin (100x)	PAA Laboratories GmbH, Pasching, Austria
Polybrene (10 mg/ml)	Santa Cruz Biotechnology, Heidelberg, Germany
Primocin (500x)	InvivoGen, San Diego, CA, USA
QuickExtract™ DNA Extraction Solution	Lucigen Corporation, Middleton, WI, USA

Recombinant Human COCO Protein	R&D systems, Minneapolis, MN, USA
RevitaCell™ Supplement (100x)	Thermo Fisher Scientific, Waltham, MA, USA
RPMI-1640 Medium	Merck KGaA, Darmstadt, Germany
Trypsin-EDTA (0.05%), phenol red	Thermo Fisher Scientific, Waltham, MA, USA
Versene Solution	Thermo Fisher Scientific, Waltham, MA, USA
Vitronectin (VTN-N) Recombinant Human Protein, Truncated (500 µg/ml)	Thermo Fisher Scientific, Waltham, MA, USA

4.11 PLASMIDS AND VECTORS

4.11.1 Commercially available plasmids and vectors

Vector	Characteristics	Resistance	Manufacturer
pcDNA3.1/ Zeo(+)	Cloning plasmid	Ampicillin, Zeocin	Invitrogen Corporation, Carlsbad, CA, USA
pCR2.1	TA-Cloning plasmid	Ampicillin, Kanamycin	Invitrogen Corporation, Carlsbad, CA, USA
pCXLE-hOCT3/4-shp53-F	Reprogramming vector (*OCT3/4* and P53/shRNA)	Ampicillin	Addgene #27077 (Shinya Yamanaka)
pCXLE-hSK	Reprogramming vector (*SOX2* and *KLF4*)	Ampicillin	Addgene #27078 (Shinya Yamanaka)
pCXLE-hUL	Reprogramming vector (*L-MYC* and *LIN28*)	Ampicillin	Addgene #27080 (Shinya Yamanaka)
pMD2.G	VSV-G envelope expressing vector	Ampicillin	Addgene #12259 (Didier Trono)
psPAX2	Lentiviral packaging vector	Ampicillin	Addgene #12260 (Didier Trono)

PX458	*Sp*Cas9-2A-EGFP	Ampicillin	Addgene #48138 (Feng Zhang)
PX459	*Sp*Cas9-2A-Puromicine	Ampicillin	Addgene #62988 (Feng Zhang)
PX461	*Sp*Cas9n-2A-EGFP	Ampicillin	Addgene #48140 (Feng Zhang)
PX462	*Sp*Cas9n-2A-Puromicine	Ampicillin	Addgene #62987 (Feng Zhang)

4.11.2 Plasmids and vectors cloned in the course of the project

Name	**pcDNA3.1-*ABCA4*-EX36-EX37_c.5197-557T, _c.5196+1013G, _c.5196+1056G, _c.5196+1134G, _c.5196+1137A, c.5196+1216A**, and **Wild-Type**
Characteristics	Minigene plasmids encoding for Exon35-Intron36-Exon37 of *ABCA4*
Comments	9.854 bp

CMV_immearly_promoter
CAG_enhancer

Exon 36

Pe 129
PE 177, c.5196+1013A/G
c.5196+1056A/G
c.5196+1134C/G, PE78, c.5196+1137G/A, P[
c.5196+1216C/A

Ampicillin

Intron 36

pBR322_origin

3 - pcDNA3.1/Zeo(+)_ABCA4Ex36/37-FLAG-T2A-mCherry
9854 bp

bleo
sh ble

Pseudoexon
c.5197-557G/T

EM7_promoter

SV40_promoter

mCherry

Exon 37
FLAG

T2A

f1_origin
bGH_PA_terminator
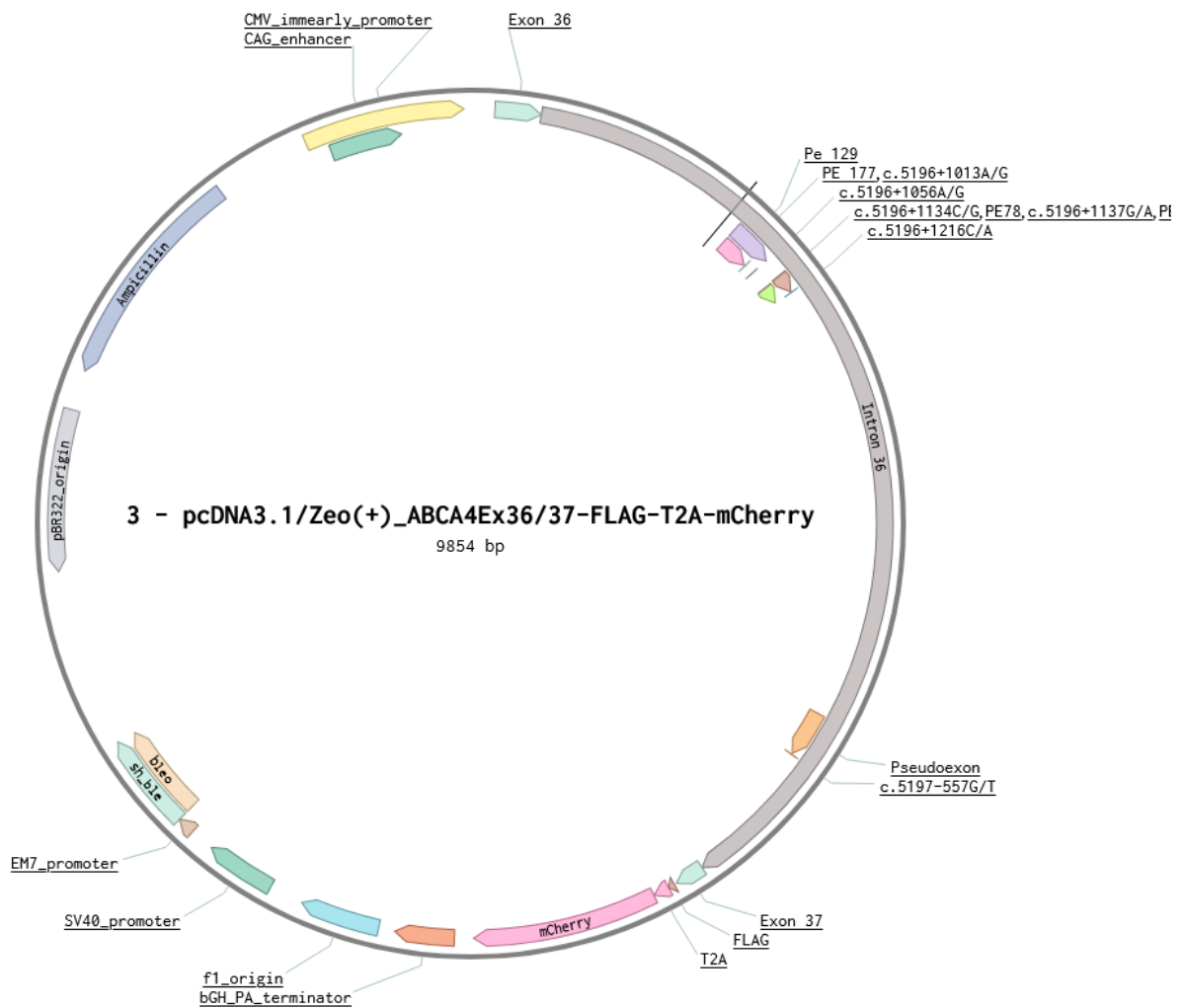
Name	**2sgRNA-PX458 –PX459 –PX461 –PX462**
Characteristics	Plasmids expressing for *Sp*Cas9 or *Sp*Ca9n,encoding for two single gRNAs and a transfection marker (EGFP or Puromycin)
	2sgRNA-PX458: *Sp*Cas9-2A-EGFP (9.703 bp)
	2sgRNA-PX459: *Sp*Cas9-2A-Puromycin (9.509 bp)
Comments	**2sgRNA-PX461:** *Sp*Cas9n-2A-EGFP (9.703 bp)
	2sgRNA-PX462: *Sp*Cas9n-2A-Puromycin (9.589 bp)
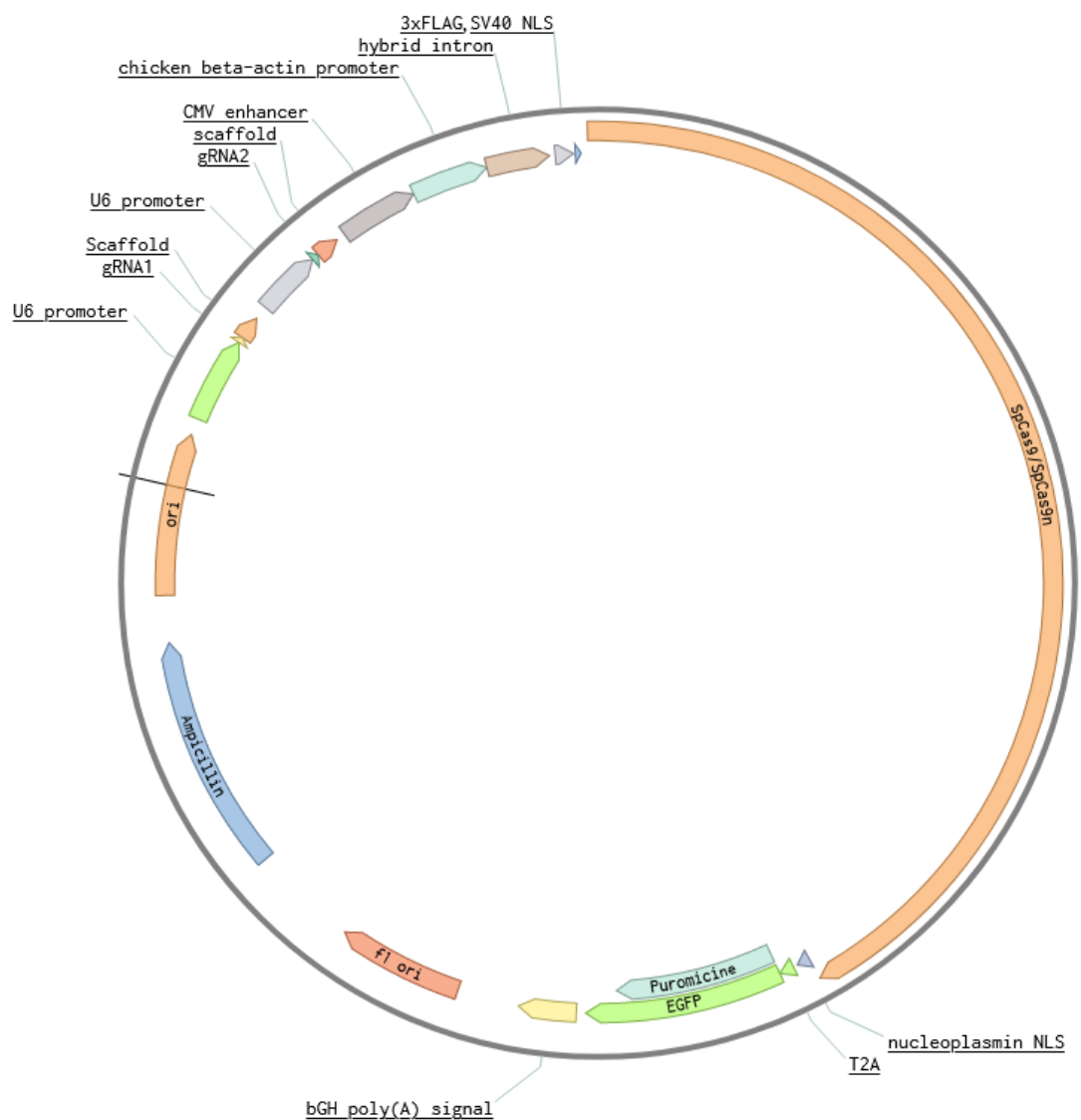
Name	**SpCas9-TREX2-2A-EGFP -PURO**
Characteristics	Plasmid expressing for SpCas9-TREX2, encoding for one single gRNA and a transfection marker (EGFP or Puromycin)
Comments	**SpCas9-TREX2-2A-EGFP:** 9.929 bp
	SpCas9-TREX2-2A-Puro(Puromycin): 9.815 bp
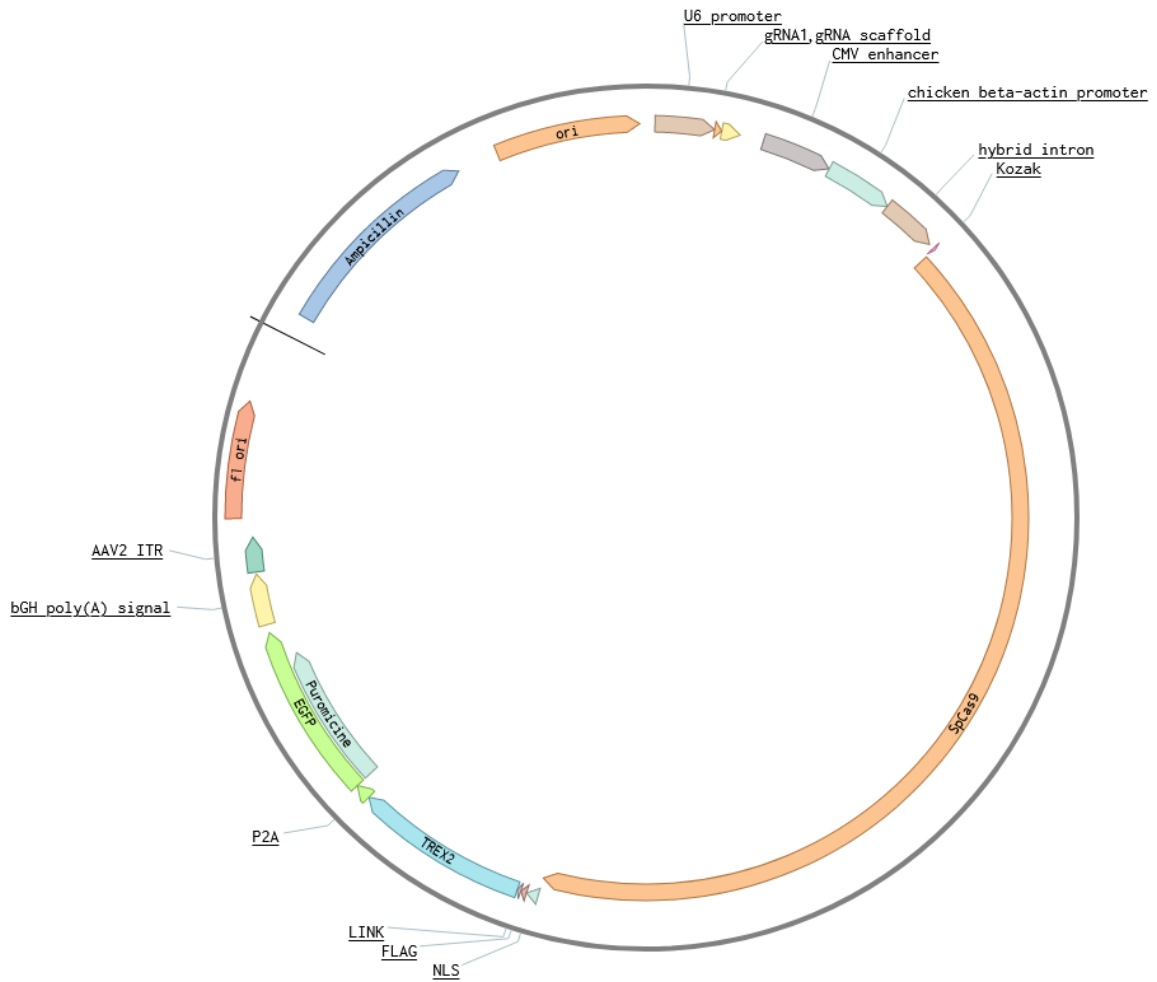
Name	**SpCas9-NLS-2A-TREX2-2A-EGFP**
Characteristics	Plasmid expressing for SpCas9-NLS-2A-TREX2, encoding for one single gRNA and a transfection marker (EGFP)
Comments	9.977 bp
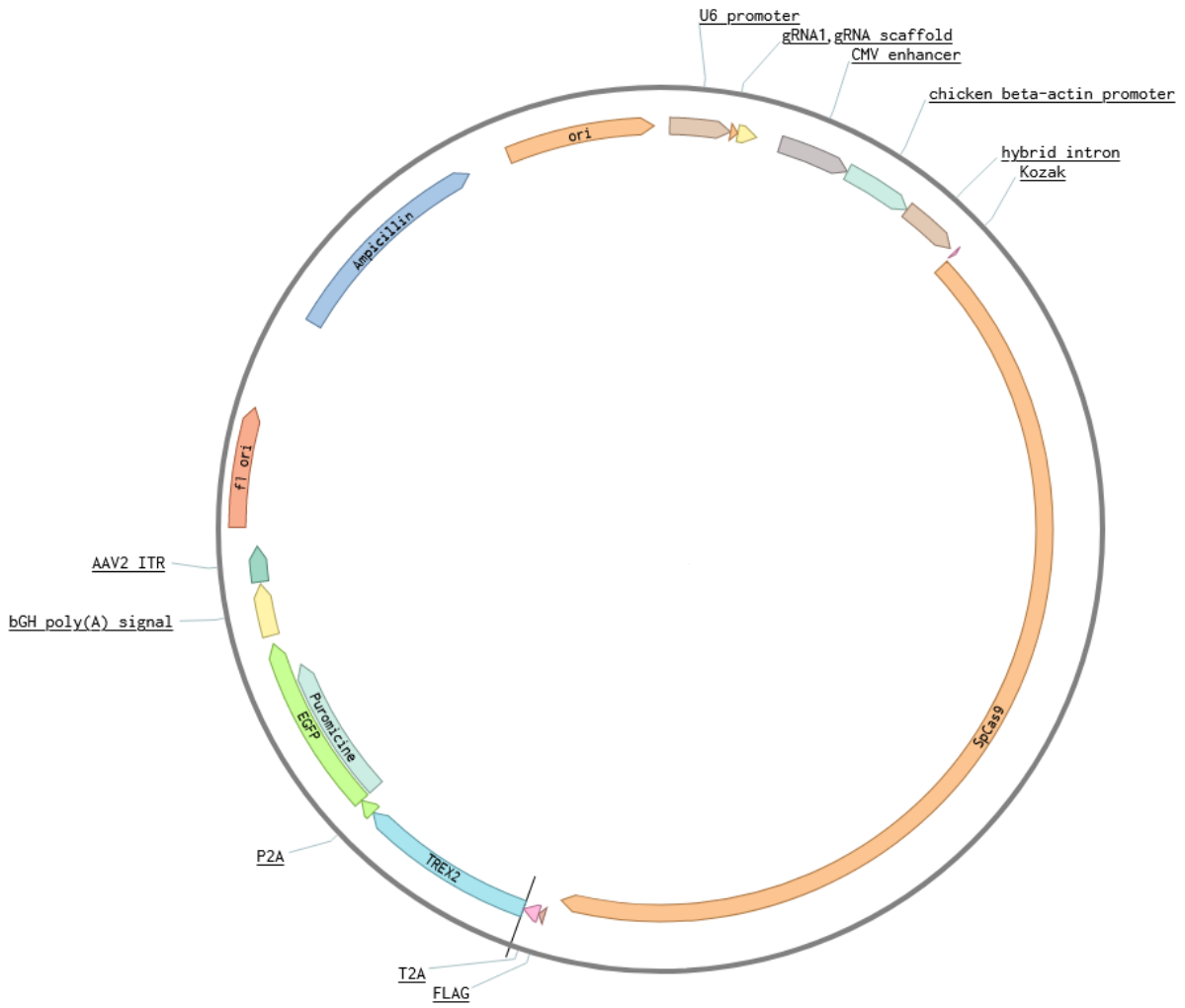
Name	**NLS-*Sp*Cas9-TREX2-2A-EGFP -PURO**
Characteristics	Plasmid expressing for NLS-*Sp*Cas9-2A-TREX2, encoding for one single gRNA and a transfection marker (EGFP or Puromycin)
Comments	**NLS-*Sp*Cas9-TREX2-2A-EGFP:** 10.091 bp
	NLS-*Sp*Cas9-TREX2-2A-Puro (Puromycin): 9.977 bp
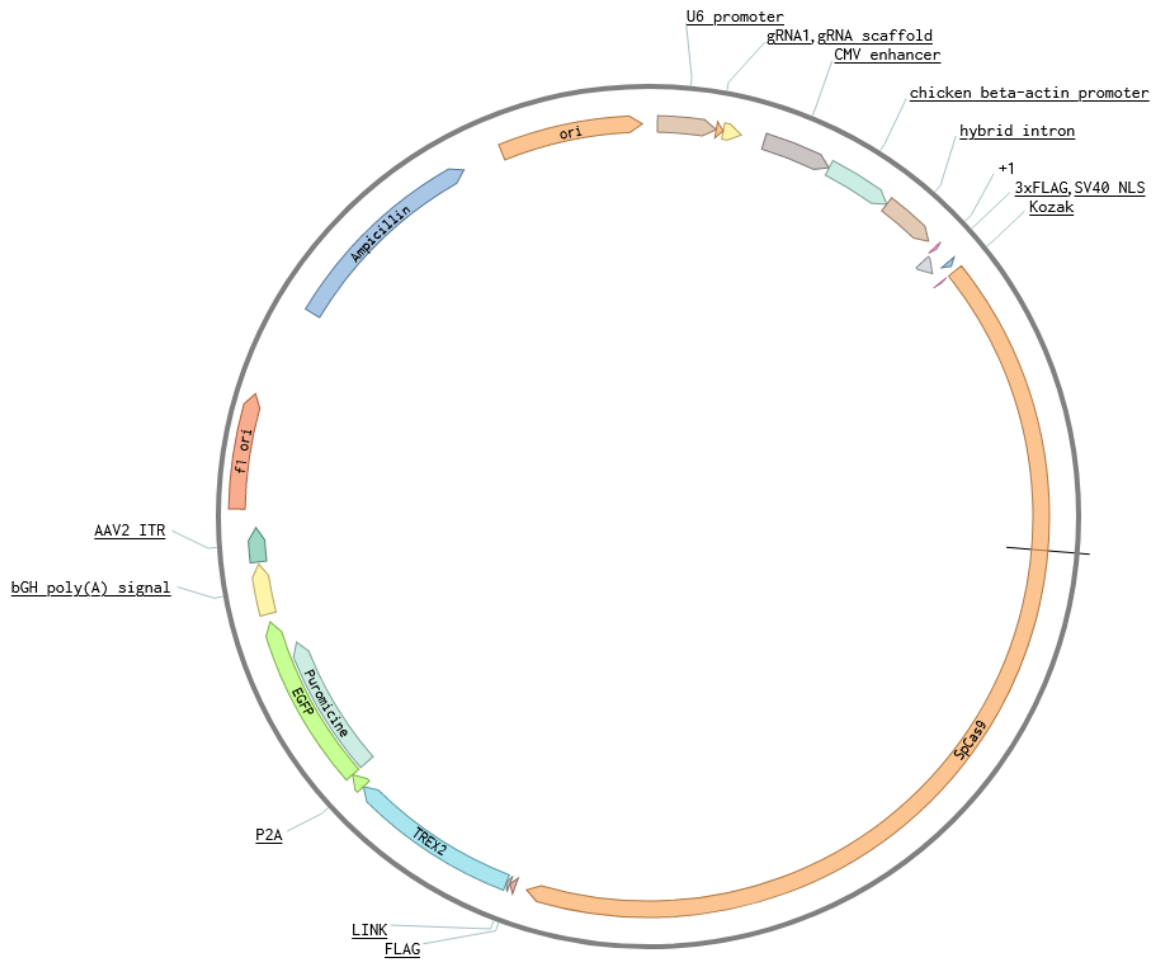
Name   **TREX2-*Sp*Cas9-NLS-2A-EGFP**

Characteristics   Plasmid expressing for TREX2-*Sp*Cas9-NLS-2A-TREX2, encoding for one single gRNA and a transfection marker (EGFP)

Comments   9.888 bp
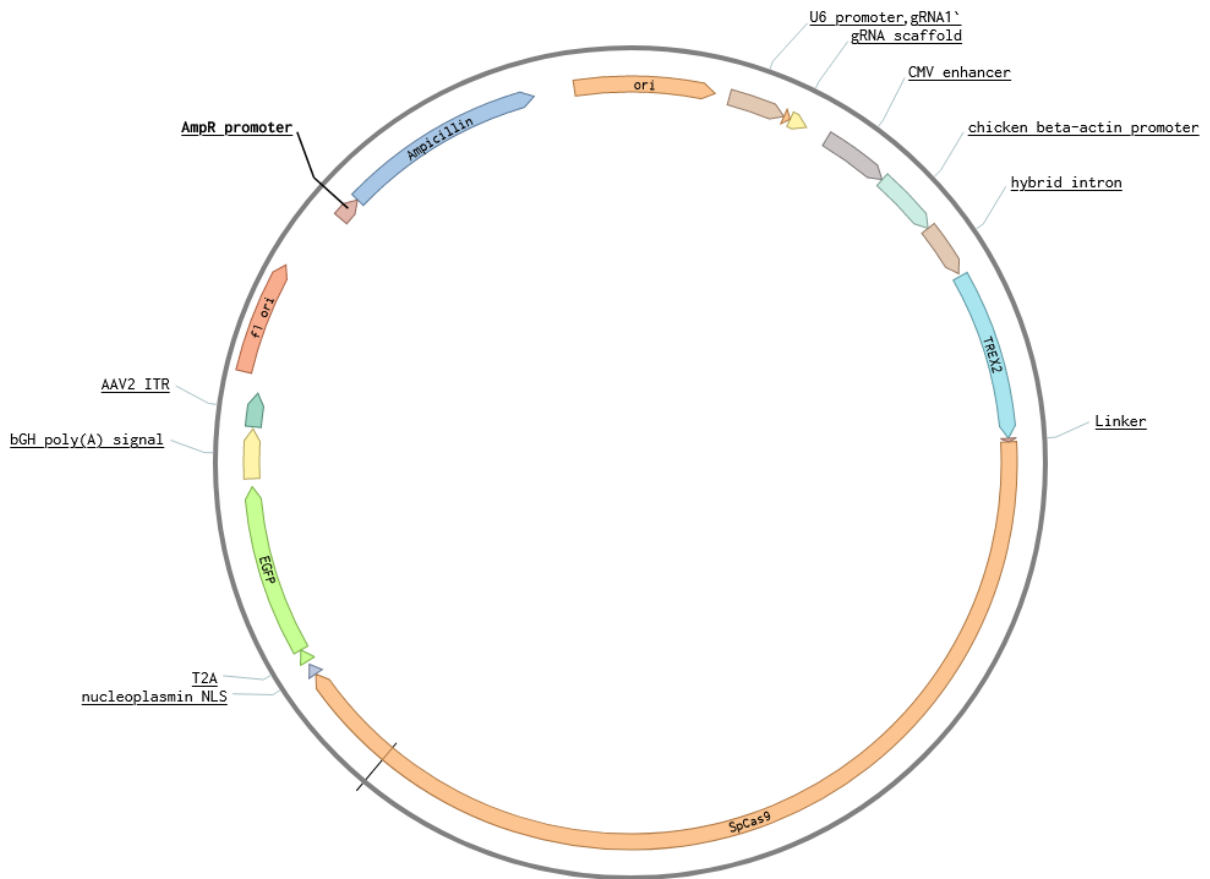
Name	**NLS-TREX2-*Sp*Cas9-NLS-2A-EGFP** and **NLS-EXONUCLEASE*-*Sp*Cas9-NLS-2A-EGFP**
Characteristics	Plasmid expressing for NLS-TREX2-*Sp*Cas9-NLS-2A-TREX2, encoding for one single gRNA and a transfection marker (EGFP)
Comments	10.091 bp (for the other exonucleases the size varies: min 9.454 bp for RAD9 and max 10.150 bp for Exonuclease VI)
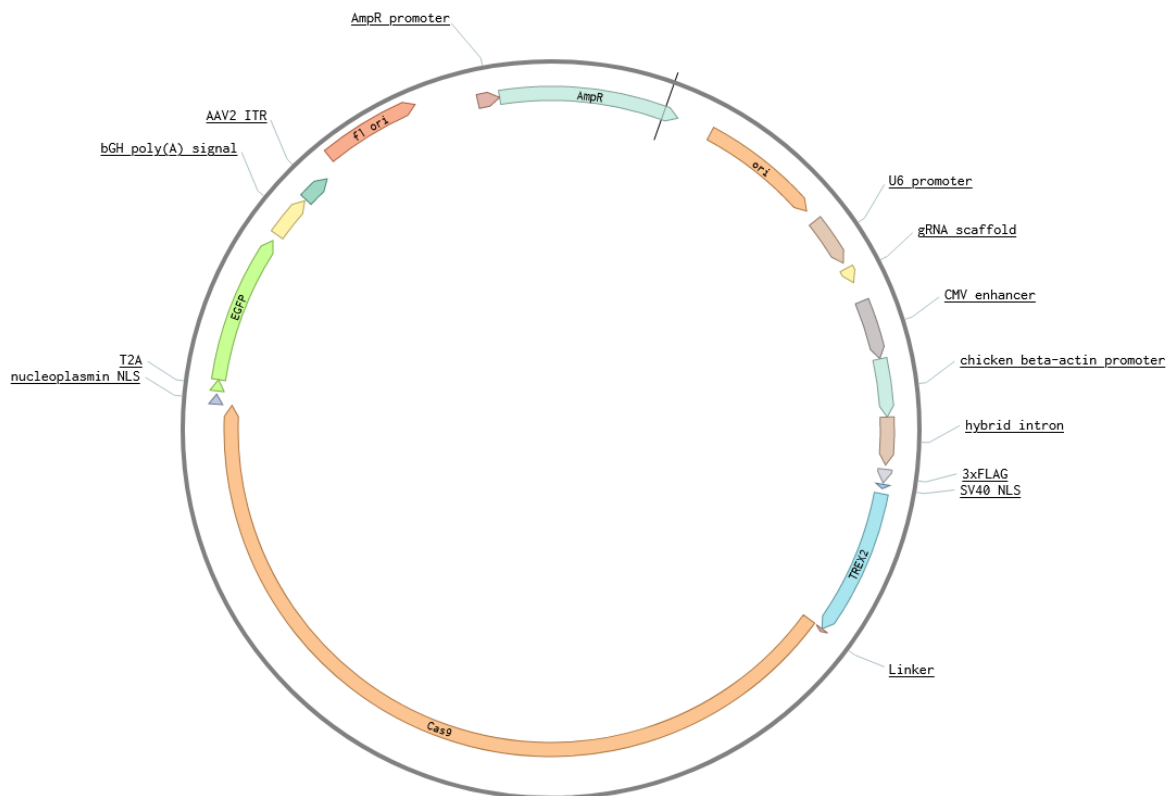
*EXONUCLEASE = either T5 exonuclease (UniProt: P06229), T7 exonuclease (UniProt: P00581), Exonuclease VI (UniProt: P00582 – first 323 aa), phi29 (UniProt: P03680 – first 191 aa), WRN (UniProt: Q14191), TREX1 (UniProt: Q9NSU2), RAD1 (UniProt: O60671), or RAD9 (UniProt: Q99638) were cloned replacing TREX2 by infusion cloning. Detailed information is not provided as expreriments were not conclusive.

Name   **pKLV2.2-CAG-*Sp*Cas9-2A-EGFP**

Characteristics   Lentiviral transfer plasmid expressing for *Sp*Cas9 and a transduction marker (EGFP). Transcription is driven by the CAG (CMV immediate enhancer/β-actin) promoter

Comments   12.211 bp
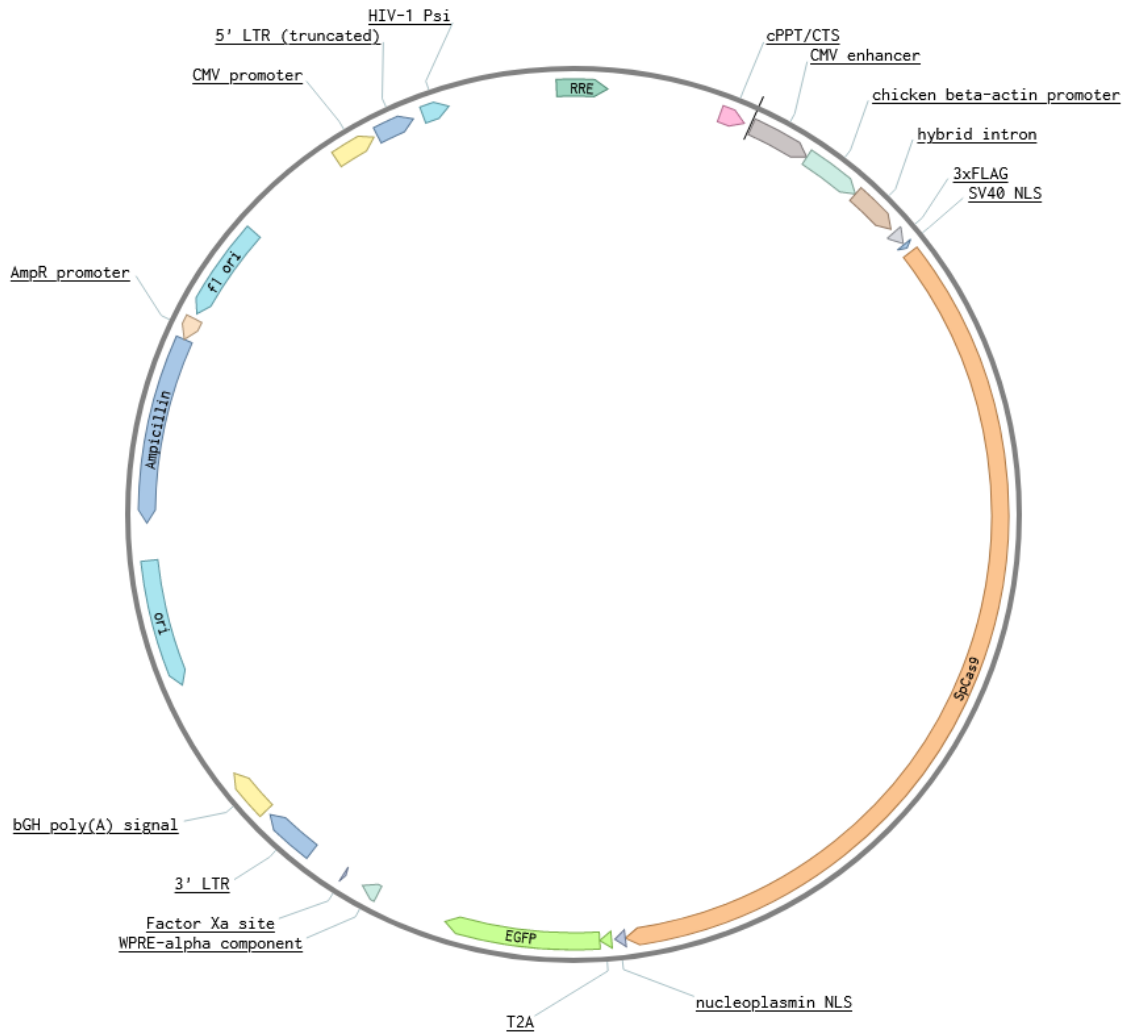
Name	**pKLV2.2-PGK-*Sp*Cas9-2A-EGFP**
Characteristics	Lentiviral transfer plasmid expressing for *Sp*Cas9 and a transduction marker (EGFP). Transcription is driven by the PGK (phosphoglycerate kinase) promoter
Comments	11.931 bp
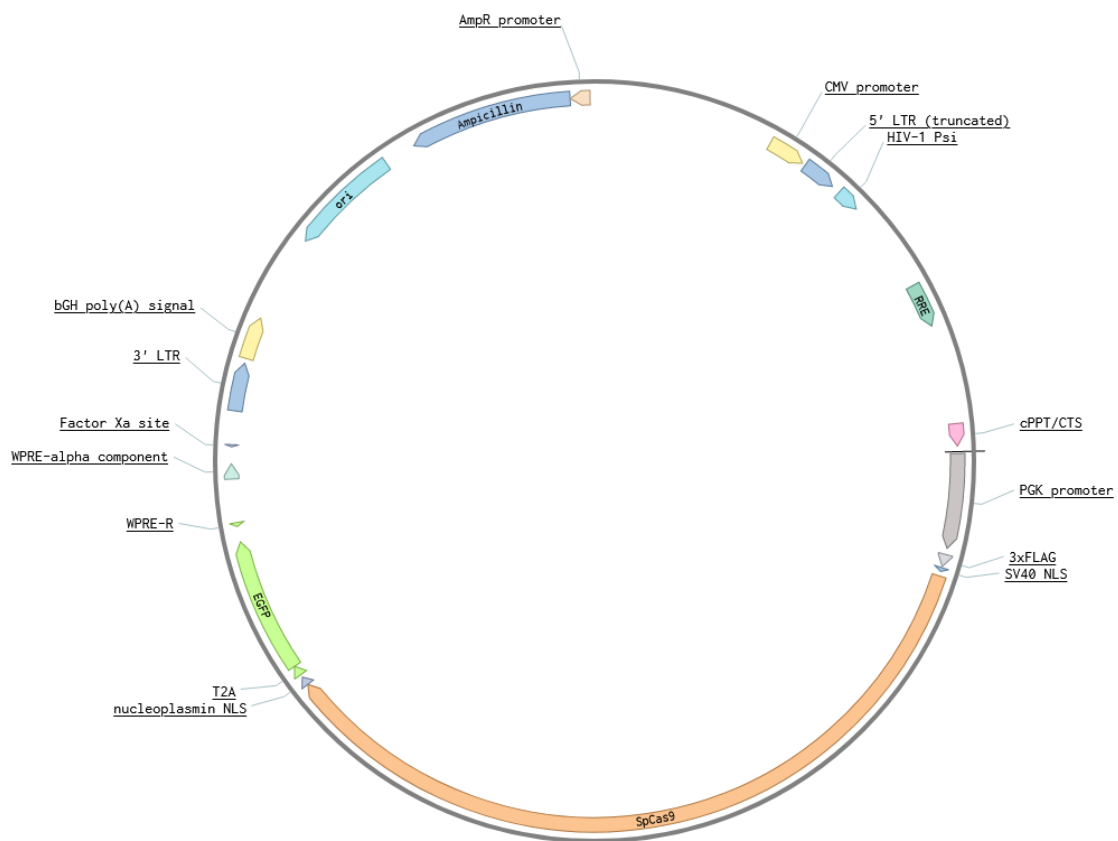
4.12 Primers and synthetic oligonucleotides

All primers were purchased either from IDT Technologies (Breda, Netherlands) or Sigma-Aldrich (Merck KGaA, Darmstadt, Germany) as lyophilized oligonucleotides and resuspended in ¼ TE to a final concentration of 100 µM. The wizard primer tool on benchling.com was used to design primers. The synthetic gRNAs were purchased either from Synthego (Redwood City, CA, USA) or IDT Technologies. ssODNs (donor templates) were purchased from IDT Technologies.

Table 1: List of Primers, synthetic gRNAs and single-stranded oligodeoxynucleotides (ssODN)

Primers for *ABCA4* Ex36-37-FLAG-T2A-mCherry minigene cloning and mutagenesis	Sequence 5' – 3'
ABCA4_infusion_pcDNA3.1-Koz-ATG-Ex36_F	TAGTCCAGTGTGGTGGAATTCGCCACCATGGGGCTGACCACTTCAGTGGATGCTGTG
ABCA4_infusion_Ex37-T2A_R	GCACTGCTCCTGCTGTATGGGAATTCGGATTACAAAGACGA
IVM-ABCA4-c.5196+1013A>G_F	GGTTTTTTAAAATGATGGTGAGTATTGGAATTCCTGAAG
IVM-ABCA4-c.5196+1013A>G_R	CTTCAGGAATTCCAATACTCACCATCATTTTAAAAAACC
IVM-ABCA4-c.5196+1056A>G_F	CTATCTTACTCAGCTTGGTAAGCAGCTATCAC
IVM-ABCA4-c.5196+1056A>G_R	GTGATAGCTGCTTACCAAGCTGAGTAAGATAG
IVM-ABCA4-c.5196+1134C>G_F	CCTCTCTCTTCTGTCTAGACGAGGAAACACTCATAAATG
IVM-ABCA4-c.5196+1134C>G_R	CATTTATGAGTGTTTCCTCGTCTAGACAGAAGAGAGAGG
IVM-ABCA4-c.5196+1137G>A_F	CCTCTCTCTTCTGTCTACACAAGGAAACACTCATAAATG
IVM-ABCA4-c.5196+1137G>A_R	CATTTATGAGTGTTTCCTTGTGTAGACAGAAGAGAGAGG
IVM-ABCA4-c.5196+1216C>A_F	GAGCATCAACTGCAGGTAACACATTGGCCCTGTGATG
IVM-ABCA4-c.5196+1216C>A_R	CATCACAGGGCCAATGTGTTACCTGCAGTTGATGCTC
IVM-ABCA4-c.5197-557G>T_F	CAGCCTCCTGTAAGTATATCTTTCTCATCCCGTTGAGGCC
IVM-ABCA4-c.5197-557G>T_R	GGCCTCAACGGGATGAGAAAGATATACTTACAGGAGGCTG

Primers for 2gRNA-PX458 and 2gRNA-PX459 cloning	Sequence 5' – 3'
Infusion-mU6-hU6_F	GCTGGCCTTTTGCTCACATGGATCCGACGCCGCCATCTCT
Infusion-mU6-hU6_R	TATGTAACGGGTACCTCTAGGGTAGAATTGGATCCAAAAAAAG CACCGACTCGG
Primers for *Sp*Cas9-TREX2-2A-EGFP, *Sp*Cas9-NLS-2A-TREX2-2A-EGFP and *Sp*Cas9-TREX2-2A-PURO cloning	Sequence 5' – 3'
Infusion-vector-SpCas9-TREX2_F	TTTTTTTCAGGTTGGACCGGTTCTAGAGCGCTG
EGFP-SpCas9-TREX2_R	AGCTCCTCGCCCTTGCTCACCGGTCCAGGATTCTCTTCGAC
Infusion-EGFP_F	GTGAGCAAGGGCGAGGAGCT
Infusion_Puro_F	ATGACCGAGTACAAGCCCAC
Infusion_vector-EGFP/PURO_R	GCTGATCAGCGAGCTCTAG
PURO-SpCas9-TREX2_R	GTGGGCTTGTACTCGGTCATCGGTCCAGGATTCTCTTCGAC
Primers for NLS-*Sp*Cas9-TREX2-2A-EGFP, NLS-*Sp*Cas9-TREX2-2A-PURO cloning	Sequence 5' – 3'
Infusion-NLS-SpCas9-TREX2_F	GGTTGGACCGGTTCTAGAGCCGGTGCCACCATGGACTATA
Infusion-NLS-SpCas9-TREX2_R	AGTGTCAGGGTCAGCACGAT
Primers for TREX2-*Sp*Cas9-NLS-2A-EGFP cloning	Sequence 5' – 3'
Infusion_N-TREX2_F	TCACTTTTTTTCAGGTTGGAGCCACCATGTCCGAGGCACCCC
Infusion_N-TREX2_R	GCTGCCGCCTCCTCCGGCCTCCAGGCTGG
Infusion_N-TREX2-SpCas9_F	GGAGGAGGCGGCAGCGACAAGAAGTACAGCATCGGCCTGG
Infusion_N-TREX2-SpCas9_R	TTTCCCCTGGCCAGAGGGCC
Primers for NLS-TREX2-*Sp*Cas9-NLS-2A-EGFP cloning	Sequence 5' – 3'
Infusion_NLS-N_TREX2_F	TCACTTTTTTTCAGGTTGGACCGGTGCC

Infusion_NLS_N_TREX2_R	TCGGCCCGGGGTGCCTCGGAGGCTGCTGGGACTCCGTGGATA
Infusion_NLS-N-TREX2-SpCas9_F	TCCGAGGCACCCCGGGCCGAGA
Infusion_NLS-N-TREX2-SpCas9_R	GCTGCCGCCTCCTCCGGCCTCCAGGCTGG
Primers for pKLV2.2-CAG-*Sp*Cas9-2A-EGFP cloning	**Sequence 5' – 3'**
Infusion_pKLV-CAG-SpCas9_F	TCCTCCGGGCTGTAATTAGC
Infusion_pKLV-CAG-SpCas9_R	CCTTAAACCCTTACCTCTTGCTCAGCTGATCAGCGAGCTCTAGT
Primers for pKLV2.2-PGK-*Sp*Cas9-2A-EGFP cloning	**Sequence 5' – 3'**
Infusion_pKLV-PGK-SpCas9_F	GACCTGCATCCATCTAGATCCGGTGCCACCATGGACTATA
Infusion_pKLV-PGK-SpCas9_R	TTGTAATCCAGAGGTTGAGCGCTGATCAGCGAGCTCTAGT
Primers used to sequence cloned gRNAs	**Sequence 5' – 3'**
mU6 (2gRNA-plasmids)	CAGCACAAAAGGAAACTCACC
hU6 (2gRNA-plasmids)	GAGGGCCTATTTCCCATGATT
CRISPR_insertcheck_R (for single gRNA plasmids)	CGCGCTAAAAACGGACTAGC
Primers for TIDE/ICE analysis	**Sequence 5' – 3'**
ABCA4_CRISPRcheck1_F	CAGGTAGAGGCACTTCAAAG
ABCA4_CRISPRcheck2_F	CCGGACCTATACACCTGAACAG
ABCA4_CRISPRcheck_R	GCCTAGTTCTTGGCTTCACAAA
Primers for splicing assays	**Sequence 5' – 3'**
MinigeneSplicing_ABCA4_Ex36_F	GCTGACCACTTCAGTGGATGC
MinigeneSplicing_ABCA4_Ex37_R	CCATACAGCAGGAGCAGTGC
CPCSplicing_ABCA4-Ex36_F	GGTTTTCTGGAGAAGTGTAGGC
CPCSplicing_ABCA4-Ex37_F	CTGCGTGATTTTCTCCATGTCC
Co-CPCSplicing_ABCA4-Ex33_F	TTACGGGCCAGCCTGCCTAA

Co-CPCSplicing_ABCA4-Ex41_R	CTTCATCAACAATGGGCTCCTT
Primers for gDNA cut efficiency	**Sequence 5' – 3'**
ABCA4_gDNA_DGM_F	TGTATGCAGTCATGGTTCATGTGT
ABCA4_gDNA_DGM_R	GCCTAGTTCTTGGCTTCACAAA
NGS_ABCA4_Nexteradpt_F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTACCGGACCT ATACACCTGAAC
NGS_ABCA4_Nexteradpt_R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGCCTGGTCT CTCACTCTTTCTA
Primers for genomic *ABCA4* sequencing	**Sequence 5' – 3'**
ABCA4_IN36_232-254_F	ACGGGGGCTCTCTATGTCCTGC
ABCA4_IN36_730-752_F	ACCACCTTCCTGACAACCAAGGT
ABCA4_IN36_1208-1230_F	GCAGGTACCACATTGGCCCTGT
ABCA4_IN36_2157-2137_R	AATAACCAGCCTGCAGTCCT
ABCA4_In36_2477-2498_R	GGCCAGCCCCAAGTGTGTAAAT
ABCA4_IN36_2571-2551_R	CACTGACCACAGTGGGGATA
ABCA4_IN36_2742-2763_F	GAAGCCAAGGTCTGAGTGCTAG
ABCA4_IN36_3017-3038_F	CCGGACCTATACACCTGAACAG
ABCA4_IN36_3579-3598_F	TGTTTGCCAACCAAGAACTG
Primers for gRNA cloning (in upper case gRNA sequences, in lower case adapter sequences)	**Sequence 5' – 3'**
ABCA4_gRNA 1_F	caccgACAGGAGGCTGATCTGGTGC
ABCA4_gRNA 1 _R	aaacGCACCAGATCAGCCTCCTGTc
ABCA4_gRNA2_F	caccgATACTTACAGGAGGCTGATC
ABCA4_gRNA2_R	aaacGATCAGCCTCCTGTAAGTATc
ABCA4_gRNA3_F	caccGAGAAAGATATACTTACAGG
ABCA4_gRNA3_R	aaacCCTGTAAGTATATCTTTCTC

ABCA4_gRNA4_F	caccGATGAGAAAGATATACTTAC
ABCA4_gRNA4_R	aaacGTAAGTATATCTTTCTCATC
ABCA4_gRNA5_F	caccgATATCTTTCTCATCCCGTTG
ABCA4_gRNA5_R	aaacCAACGGGATGAGAAAGATATc
ABCA4_gRNA6_F	caccGTTGAGGCCTCCACACCCCA
ABCA4_gRNA6_R	aaacTGGGGTGTGGAGGCCTCAAC
ABCA4_gRNA7_F	caccGCTGTGCCATTGAGCATGG
ABCA4_gRNA7_R	aaacCCATGCTCAATGGCACAGC
ABCA4_gRNA8_F	caccgTCAGGGGGCAACCTGCCGTG
ABCA4_gRNA8_R	aaacCACGGCAGGTTGCCCCCTGAc
ABCA4_gRNA9_F	caccGAGCATGGAGGAGGTCCAG
ABCA4_gRNA9_R	aaacCTGGACCTCCTCCATGCTC
ABCA4_gRNA10_F	caccgCAGGAGGCTGATCTGGTGCA
ABCA4_gRNA10_R	aaacTGCACCAGATCAGCCTCCTGc
ABCA4_gRNA11_F	caccgATCAGAGCCAGAGCAGAACG
ABCA4_gRNA11_R	aaacCGTTCTGCTCTGGCTCTGATc
ABCA4_gRNA12_F	caccgAAAAACTAGAATGAGCCCTG
ABCA4_gRNA12_R	aaacCAGGGCTCATTCTAGTTTTTc
ABCA4_gRNA13_F	caccGGAGAAGGGAGTTACAGATG
ABCA4_gRNA11_R	aaacCATCTGTAACTCCCTTCTCC
ABCA4_gRNAI_F	caccgAATTTATTCTAGGTGTCTTG
ABCA4_gRNAI_R	aaacCAAGACACCTAGAATAAATTc
ABCA4_gRNAII_F	caccgATTTATTCTAGGTGTCTTGA
ABCA4_gRNAII_R	aaacTCAAGACACCTAGAATAAATc
ABCA4_gRNAIII_F	caccgCAATTTATTCTAGGTGTCTT

ABCA4_gRNAIII_R	aaacAAGACACCTAGAATAAATTGc
ABCA4_gRNAVIIa_F	caccgCTCTCTCTTCTGTCTAGACG
ABCA4_gRNAVIIa_R	aaacCGTCTAGACAGAAGAGAGAGc
ABCA4_gRNAVIIb_F	caccgCTCTCTCTTCTGTCTACACA
ABCA4_gRNAVIIb_R	aaacTGTGTAGACAGAAGAGAGAGc
ABCA4_gRNAVIIwt_F	caccgCTCTCTCTTCTGTCTACACG
ABCA4_gRNAVIIwt_R	aaacCGTGTAGACAGAAGAGAGAGc
ABCA4_gRNAVIII_F	caccGGAAACACTCATAAATGCAC
ABCA4_gRNAVIII_R	aaacGTGCATTTATGAGTGTTTCC
ABCA4_gRNAIX_F	caccGGATAAGAGCATCAACTGC
ABCA4_gRNAIX_R	aaacGCAGTTGATGCTCTTATCC
ABCA4_gRNAXVI_F	caccgCATCAACCCCAATTTATTCT
ABCA4_gRNAXVI_R	aaacAGAATAAATTGGGGTTGATGc
Primers for RT-qPCR	**Sequence 5' – 3'**
OCT3/4_F	GTTCTTCATTCACTAAGGAAGG
OCT3/4_R	CAAGAGCATCATTGAACTTCAC
PAX6_F	CCGGCAGAAGATTGTAGAGC
PAX6_R	GCCCGTTCAACATCCTTAGT
ABCA4_F	CATCCTGTTCCACCACCTCA
ABCA4_R	CTGTGTCCTCCAACATGGCT
OPN1SW_F	ACCATTGGTATTGGCGTCTC
OPN1SW_R	GGAGAGAGGCACAATGAAGC
OPN1M/LW_F	GTGGTCACTGCATCCGTCTT
OPN1M/LW_R	ACGGTCTCTGCTAGGTCAGC
RCV1_F	ACACCAAGTTCTCGGAGGAG

RCV1_R	ACTTGGCGTAGATGCTCTGG
GUSB_F	AGAGTGGTGCTGAGGATTGG
GUSB_R	CCCTCATGCTCTAGCGTGTC
POU5F1_F	GGAAGGTATTCAGCCAAACG
POU5F1_R	CTCCAGGTTGCCTCTCACTC
SOX2_F	AGCTCGCAGACCTACATGAA
SOX2_R	CCGGGGAGATACATGCTGAT
NANOG_F	CAAAGGCAAACAACCCACTT
NANOG_R	TGCGTCACACCATTGCTATT
DNMT3B_F	ACGACACAGAGGACACACAT
DNMT3B_F	AAGCCCTTGATCTTTCCCCA
GAPDH_F	TCACCAGGGCTGCTTTTAAC
GAPDH_R	GACAAGCTTCCCGTTCTCAG
P21_F	ATGAAATTCACCCCCTTTCC
P21_R	CCCTAGGCTGTGCTCACTTC
Synthetic gRNAs	**Sequence 5' – 3'**
KnockIn(*Sp*Cas9)-ABCA4_c.5196+1013A>G (from Synthego)	U*U*U*AAAAUGAUGGUGAAUAU- modified scaffold
KnockIn(Cpf1)-ABCA4_c.5196+1056A>G (from IDT)	AltR1/rUrArA rUrUrU rCrUrA rCrUrC rUrUrG rUrArG rArUrU rArUrC rUrUrA rCrUrC rArGrC rUrUrA rGrUrA rArG/AltR2
KnockIn(*Sp*Cas9)-ABCA4_c.5196+1216C>A (from Synthego)	U*U*A*GCAUCACAGGGCCAAUG – modified scaffold
ssODN (Donor Templates)	**Sequence 5' – 3'** In red, the position of the DIV
ssODN- ABCA4_c.5196+1013A>G	GACCAACACAAATGACCTTCTCATCCATGGTTTTTTAAAATGATGGTGAgTATTGGAATTCCTGAAGATATGATTTCTATCTTACTCAGCTTAGTAAGCA

ssODN- ABCA4_c.5196+1056A>G	ATGGTGAATATTGGAATTCCTGAAGATATGATTTCTATCTTACTC AGCTT**g**GTAAGCAGCTATCACTTAACAATACAAAACCAGAGATT ATCAGTAGCAAC
ssODN- ABCA4_c.5196+1216C>A	GGTCAGAACCTGAAAGCCTTTCTTTGGATAAGAGCATCAACTGC AGGTA**a**CACATTGGCCCTGTGATGCTAATATAAAAGGAGCTAG GCCCACCGGTAC

gBLOCK	**Sequence 5' – 3'** In green, the T2A sequence In red, the mCherry sequence the additional sequences are the infusion-compatible overlapping ends
T2A-mCherry	AATTCGGATTACAAAGACGATGACGATAAGGGATCCGGCGAGG GCAGAGGAAGTCTCCTAACATGCGGTGACGTGGAGGAGAATCC TGGCCCAGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATC AAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGA ACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCC CCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGG GTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCA TGTACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGACATCCCC GACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGC GCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCA GGACTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAG CTGCGCGGCACCAACTTCCCCTCCGACGGCCCCGTAATGCAGAA GAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCC GAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAG CTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCT ACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGT CAACATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCA TCGTGGAACAGTACGAACGCGCCGAGGGCCGCCACTCCACCGG CGGCATGGACGAGCTGTACAAGTAAAATTCTGCAGATATCCAG CA

4.13 SOFTWARE AND TOOLS

- **7500 Real-Time PCR Software, Applied Biosystems, Waltham, MA, USA** was used to acquire Real-Time PCR data.

- **Benchling (**http://benchling.com**), San Francisco, CA, USA,** was used to design and select sgRNAs, to design the cloning experiments and primers, as well as to align sequencing results to the reference sequences.

- **BootstRatio, de Heredia Lab, Hospital Duran i Reynals, Spain,** was used for the statistical analysis of RT-PCR results.

- **Cas-OFFinder (**http://www.rgenome.net/cas-offinder/**), Seoul National University College of Medicine, South Korea,** was used for the identification of off-target sites.

- **CRISPResso2 (**https://crispresso.pinellolab.partners.org/**), Pinello Lab, Massachusetts General Hospital and Harvard Medical School, USA** was used to quantify and characterize the mutational profiles generated by the dgRNAs/*Sp*Cas9 and sgRNA/*Sp*Cas9 approaches, out of deep-sequencing data.

- **E.A.S.Y. Win32, Herolab, Wiesloch, Germany,** was used to acquire agarose gel pictures.

- **Excel, Microsoft, Redmond, WA, USA,** was used for analysis of splicing, genomic DNA (gDNA) cleavage and RT-PCR data.

- **FlowJo™, BD Biosciences, Franklin Lakes, NJ, USA,** was used to analyze flow cytometry data and present them in appropriate charts. The software used was under license of the *Flow Cytometry Core Facility – Berg of the Medical Faculty of the University of Tübingen*.

- **Graphpad Prism 7, GraphPad Software, La Jolla, CA, USA,** was used for statistical analysis.

- **ICE analysis (**https://ice.synthego.com/#/**), SYNTHEGO corporation, Redwood City, CA, USA,** was used to preliminary assess the gDNA cleavage efficacy of selected gRNAs by tracking sequence aberrations on elecropherograms of treated samples in comparison to a non-treated control.

- **Maestro, Schrödinger, New York, NY, USA,** was used to generate the rendering 3D images of ABCA4 and *Sp*Cas9.

- **Sequencing Analysis 5.2, Applied Biosystems, Waltham, MA, USA,** was used as a sequencing tool for the processing of sequencing raw data.

- **TIDE (**http://shinyapps.datacurators.nl/tide/**), van Steensel Lab, Netherlands Cancer Institute**, was used to preliminary assess the gDNA cleavage efficacy of selected gRNAs by tracking sequence aberrations on elecropherograms of treated samples in comparison to a non-treated control.

- **Zen 2 Blue Edition, Zeiss Microscopy, Wetzlar, Germany,** was used to take and process microscope pictures.

5 METHODS

5.1 SINGLE GUIDE RNA IDENTIFICATION AND SELECTION

sgRNA sequences were identified and select using a freely-accessible online tool (https://www.benchling.com). The reference genome used for sgRNA design was the human genome assembly GRCh38. Protospacer Adjacent Motif (PAM) requirements for *Sp*Cas9 was set as NGG at the 3' side of the sgRNA sequence. To aid the selection of the best sgRNAs, on-target and off-target scores were also considered. The on-target score scale is based on Doench et al. 2016, whereas, the off-target score scale is based on Hsu et al. 2013. The calculation parameters were left from default.

5.2 CLONING

The detailed description of the plasmids generated and used in this thesis is available in the Material: 4.11.

5.2.1 Infusion cloning

Most of the plasmids described thereafter were cloned through infusion cloning, which is a highly efficient cloning method that does not involve ligation. It is based on the annealing of complementary ends of cloning inserts and linearized backbone plasmid, after that exonuclease directional catalytical removal determines the generation of overlapping overhanging ends (Gibson et al. 2009). The infusion protocol used throughout this thesis combines 100 ng of digested backbone plasmid with 1:2 molar ratio of big fragments (>2000 bp), 1:4 molar ratio of middle fragments (>400 bp and <2000 bp), and 1:6 molar ratio of small fragments (<400 bp). The backbone plasmids and the fragment(s) were combined in a final volume of 10 µl, ddH$_2$0 is used to fill up. 10 µl NEBuilder® HiFi mix was added to the backbone plasmids and the fragment(s) and incubated 30 min at 50°C. 1.5 µl of the mixture was transformed in chemically competent 10-beta *E. coli* bacteria cells by heat shock according to manufacturer's procedure, plated on an agar plates containing ampicillin 100 µg/ml, and incubated at 37°C overnight.

A summary of the plasmid backbones and fragments used for the various infusion cloning is reported in **Table 2**.

5.2.2 Backbone plasmid digestion with restriction enzymes

All the backbone plasmids used for cloning through infusion cloning as well as T4-mediated ligation of fragments with compatible ends were digested with restriction enzymes. This step was needed to linearize the backbone plasmids and generate plasmids ends compatible with the downstream cloning protocol. In general, 1000 ng of backbone plasmid was digested overnight in a 50 µl reaction set as follows:

Reagents	Volume (µl)
Backbone plasmid (1000 ngl)	X
10x reaction buffer	5
Restriction enzyme I (10 U)	X
Restriction enzyme II (10 U)	X
ddH$_2$0	Up to 50

The digested backbone plasmids were purified by NucleoSpin Gel and PCR clean-up kit according to the manufacturer's PCR clean-up protocol and eluted in 15 µl elution buffer. The concentrations of the digested backbone plasmids were quantified by NanoDrop®.

Table 2: Details on the plasmids cloned via infusion cloning. Backbones and restriction enzymes used for digestion are listed, as well as cloned fragments and the original source.

	Backbone	Restriction enzyme	Fragment 1	Amplified from	Fragment 2	Amplified from
pcDNA3.1-ABCA4-EX36-EX37_c.5197-557G (wild-type minigene)	pcDNA3.1/Zeo(+)	EcoRI-HF	ABCA4 (Exon36-Intron36-Exon37)	HEK293T genomic DNA	T2A-mCherry	Chemically synthetized (IDT)
2sgRNA-PX458	PX458					
-PX459,	PX459	AflIII and XbaI	mU6-sgRNA1 _ hU6 sgRNA2	pKLV2.2-mU6sgRNA5(SapI)-hU6sgRNA5(BbsI)-PGKpuroBFP-W		
-PX461	PX461					
-PX462	PX462					
SpCas9-TREX2-2A-EGFP	PX459	AgeI and NotI	SpCas9-TREX2 (EGFP)	pKLV2.2-EF1-SpCas9-TREX2-2A-EGFP	2A-EGFP	PX458
SpCas9-TREX2-2A-PURO	PX459	AgeI and NotI	SpCas9-TREX2 (PURO)	pKLV2.2-EF1-SpCas9-TREX2-2A-PURO	2A-PURO	PX459
SpCas9-NLS-2A-TREX2-2A-EGFP	PX459	AgeI and NotI	SpCas9-2A-TREX2 (EGFP)	pKLV2.2-EF1-SpCas9-2A-TREX2-EGFP	2A-EGFP	PX458
NLS-SpCas9-TREX2-2A-EGFP	SpCas9-TREX2-2A-EGFP	EcoRV and AfeI	NLS	PX458		
NLS-SpCas9-TREX2-2A-PURO	SpCas9-TREX2-2A-PURO	EcoRV and AfeI	NLS	PX458		
TREX2-SpCas9-NLS-2A-EGFP	PX458	AgeI and ApaI	TREX2	pKLV2.2-EF1-SpCas9-TREX2	Part of SpCas9	PX458
NLS-TREX2-SpCas9-NLS-2A-EGFP	PX458	AgeI and ApaI	NLS	PX458	TREX2-SpCas9	TREX2-SpCas9-NLS-2A-EGFP
pKLV2.2-CAG-SpCas9-2A-EGFP	pKLV2.2-CAG	BlbI	SpCas9-2A-EGFP	PX458		
pKLV2.2-2sgRNA-PGK-SpCas9-2A-EGFP cloning	pKLV2.2-mU6sgRNA5(SapI)-hU6sgRNA5(BbsI)-PGKpuroBFP-W	XhoI and NotI	SpCas9-2A-EGFP	PX458		

5.2.3   Cloning of minigene plasmids pcDNA3.1-*ABCA4*-EX36-EX37_c.5197-557T, _c.5196+1013G, _c.5196+1056G, _c.5196+1134G, _c.5196+1137A, c.5196+1216A, and _Wild-Type

A minigene plasmid encoding for a genomic fragment of the *ABCA4* gene spanning from Exon 36 to Exon 37 was amplified from human genomic DNA of a healthy proband, and cloned in pcDNA3.1/Zeo(+) and referred to as pcDNA3.1-*ABCA4*-EX36-EX37_Wild-Type. The plasmid was cloned in the configuration of CMV_*ABCA4*(Exon36-Intron36-Exon37)-T2A-mCherry through infusion cloning of two fragments in the pcDNA3.1/Zeo(+) backbone plasmid. The c.5197-557G>T, c.5196+1013A>G, c.5196+1056A>G, c.5196+1134C>G, c.5196+1137G>A and c.5196+1216C>A deep-intronic variants were introduced in the wild-type minigene via *in vitro* mutagenesis. The resulting plasmid was referred to as pcDNA3.1-*ABCA4*-EX36-EX37_c.5197-557T, _c.5196+1013G, _c.5196+1056G, _c.5196+1134G, _c.5196+1137A, and c.5196+1216A.

5.2.3.1   Cloning of minigene plasmid pcDNA3.1-ABCA4-EX36-EX37_Wild-Type

The pcDNA3.1/Zeo(+) backbone plasmid was digested by *Eco*RI, in NEBuffer™ CutSmart, overnight at 37°C.

The *ABCA4*(Exon36-Intron36-Exon37) fragment was amplified from HEK293T genomic DNA as follows:

Reagents	Volume (µl)	PCR program		
5X *Pfu*Ultra HF reaction buffer	5	98°C	30 sec	
dNTP-Mix (5 mM/ dA,dG,dT,dC)	1	98°C	10 sec	
F primer (10 µM ABCA4_infusion_pcDNA3.1-Koz-ATG-Ex36_F)	1.25	60°C	25 sec	x25
R primer (10 µM ABCA4_infusion_Ex37-T2A_R)	1.25	72°C	2 min 10 sec	
Q5 enzyme	0.25	72°C	2 min	
Healthy proband gDNA (50 ng/µl)	1			
ddH$_2$0	15.25			

The T2A-mCherry gBlock gene fragment was purchased from Integrated DNA Technologies; Germany as lyophilized and resuspended in ¼ TE at a final concentration of 10 ng/µl.

The concentrations of the PCR amplified *ABCA4*(Exon36-Intron36-Exon37) fragment was quantified by NanoDrop®. The backbone plasmid and fragment were assembled by infusion cloning.

5.2.3.2 Cloning of minigene plasmid pcDNA3.1-ABCA4-EX36-EX37_c.5197-557T, _c.5196+1013G, _c.5196+1056G, _c.5196+1134G, _c.5196+1137A, and c.5196+1216A

The mutant minigene plasmid was generated by *in vitro* mutagenesis using the wild-type minigene plasmid as a template. *In vitro* mutagenesis (IVM) enables introduction of single nucleotide changes, as well as, small deletion or insertion changes by PCR amplification. A primer pair containing the desired edited sequence is used to amplify the template plasmid. Following *in vitro* mutagenesis, the PCR reaction is digested by *Dpn*I which selectively degrades methylated adenosines at "GATC" sequences of the plasmid template, therefore leaving the newly generated and mutated PCR products intact (Smith, 1985)

The *in vitro* mutagenesis reaction was set as follows:

Reagents	Volume (µl)
5X *Pfu*Ultra HF reaction buffer	5
dNTP-Mix (5mM/ dA,dG,dT,dC)	1
F primer (10 µM IVM-ABCA4-(*DIV)_F)	1.25
R primer (10 µM IVM-ABCA4-(*DIV)_R)	1.25
Q5 enzyme	0.25
HEK293T gRNA (50 ng/µl)	1
ddH₂0	15.25

PCR program		
98°C	30 sec	
98°C	10 sec	
60°C	25 sec	x25
72°C	2 min 10 sec	
72°C	2 min	

*primers details are avaliable in **Table 1**

To confirm the success of the *in vitro* mutagenesis amplification, 10 µl of the reaction was load on a 1% agarose gel. The plasmid used as template was removed by digestion of the remaining 40 µl by 1 µl *Dpn*I (20 U/µl) incubated for 2 hours at 37°C. 1.5 µl of the mixture was transformed in chemically competent 10-beta *E. coli* bacteria cells by heat shock according to manufacturer's procedure, plated on an agar plates containing ampicillin 100 µg/ml, and incubated at 37°C overnight.

5.2.4   Cloning of *Sp*Cas9 backbone plasmids encoding for two sgRNAs

To transfect cells with plasmids encoding for *Sp*Cas9 or *Sp*Cas9n and two sgRNAs, 2sgRNA-PX458, -PX459, -PX461, and -PX462 plasmids were used and altered by cloning through infusion cloning. Different steps were necessary as described below:

To remove the hU6-sgRNA cassette included in the backbones of PX458, PX459, PX461 and PX462 plasmids, they were digested by *Afl*III and *Xba*I, in NEBuffer™ 3.1, overnight at 37°C.

The 2sgRNA cassette encoding for murine U6 promoter (mU6) – sgRNA1 and human U6 promoter (hU6) – sgRNA2 was amplified from pKLV2.2-mU6sgRNA5(*Sap*I)-hU6sgRNA5(*Bbs*I)-PGKpuroBFP-W as follows:

Reagents	Volume (µl)
5X Q5 reaction buffer	5
ddNTP-Mix (5mM/ dA,dG,dT,dC)	1
F primer (10 µM - Infusion-mU6-hU6_F)	1.25
R primer (10 µM - Infusion-mU6-hU6_R)	1.25
Q5 enzyme	0.25
pKLV2.2-mU6sgRNA5(*Sap*I)-hU6sgRNA5(*Bbs*I)-PGKpuroBFP-W (2 ng/µl)	1
ddH$_2$0	15.25

PCR program		
98 °°C	30 sec	
98 °°C	10 sec	
66 °°C	15 sec	x35
72 °°C	30 sec	
72 °°C	2 min	

The concentrations of the PCR amplified mU6 – sgRNA1 – hu6 – sgRNA2 fragment was quantified by NanoDrop®. The backbone plasmid and fragment were assembled by infusion cloning.

5.2.5   Cloning of *Sp*Cas9-TREX2 backbone plasmids

*Sp*Cas9-TREX2-2A-EGFP, *Sp*Cas9-TREX2-2A-PURO, NLS-*Sp*Cas9-TREX2-2A-EGFP, NLS-*Sp*Cas9-TREX2-2A-PURO, TREX2-*Sp*Cas9-NLS-2A-EGFP, NLS-TREX2-*Sp*Cas9-NLS-2A-EGFP plasmids were cloned through infusion cloning.

5.2.5.1   Sp*Cas9-TREX2-2A-EGFP,* Sp*Cas9-NLS-2A-TREX2-2A-EGFP and* Sp*Cas9-TREX2-2A-PURO*

The PX459 backbone vector was digested by *Age*I and *Not*I, in NEBuffer™ CutSmart, overnight at 37°C. The *Sp*Cas9-TREX2 (EGFP) and *Sp*Cas9-2A-TREX2 (EGFP) fragment was amplified from pKLV2.2-EF1-*Sp*Cas9-TREX2 and pKLV2.2-EF1-*Sp*Cas9-NLS-2A-TREX2, respectively, as follows:

Reagents	Volume (µl)
5X Q5 reaction buffer	5
ddNTP-Mix (5mM/ dA,dG,dT,dC)	1
F primer (10 µM - Infusion-vector-SpCas9-TREX2_F)	1.25
R primer (10 µM – EGFP-SpCas9-TREX2_R)	1.25
Q5 enzyme	0.25
pKLV2.2-EF1-*Sp*Cas9-TREX2-2A-EGFP or pKLV2.2-EF1-*Sp*Cas9-NLS-2A-TREX2-2A-EGFP (2 ng/µl)	1
ddH₂0	15.25

PCR program		
98°C	30 sec	
98°C	15 sec	
60°C	20 sec	x30
72°C	2min 40 sec	
72°C	2 min	

The *Sp*Cas9-TREX2 (PURO) fragment was amplified from pKLV2.2-EF1-*Sp*Cas9-TREX2 as follows:

Reagents	Volume (µl)
5X Q5 reaction buffer	5
dNTP-Mix (5mM/ dA,dG,dT,dC)	1
F primer (10 µM - Infusion-vector-SpCas9-TREX2_F)	1.25
R primer (10 µM – PURO-SpCas9-TREX2_R)	1.25
Q5 enzyme	0.25
pKLV2.2-EF1-*Sp*Cas9-TREX2-2A-EGFP (2 ng/µl)	1
ddH₂0	15.25

PCR program		
98°C	30 sec	
98°C	15 sec	
60°C	20 sec	x30
72°C	2min 40 sec	
72°C	2 min	

The 2A-EGFP fragment was amplified from PX458 as follows:

Reagents	Volume (µl)
5X Q5 reaction buffer	5
ddNTP-Mix (5mM/ dA,dG,dT,dC)	1
F primer (10 µM – Infusion_EGFP_F)	1.25
R primer (10 µM – Infusion_vector-EGFP/PURO_R)	1.25
Q5 enzyme	0.25
PX458 (2 ng/µl)	1
ddH₂0	15.25

PCR program		
98°C	30 sec	
98°C	15 sec	
60°C	20 sec	x30
72°C	40 sec	
72°C	2 min	

The 2A-PURO fragment was amplified from PX459 as follows:

Reagents	Volume (μl)
5X Q5 reaction buffer	5
dNTP-Mix (5mM/ dA,dG,dT,dC)	1
F primer (10 μM – Infusion_Puro_F)	1.25
R primer (10 μM – Infusion_vector-EGFP/PURO_R)	1.25
Q5 enzyme	0.25
PX458 (2 ng/μl)	1
ddH$_2$0	15.25

PCR program		
98°C	30 sec	
98°C	15 sec	
60°C	20 sec	x30
72°C	40 sec	
72°C	2 min	

The PCR amplified fragments were purified by NucleoSpin Gel and PCR clean-up kit according to the manufacturer's PCR clean-up protocol and eluted in 15 μl elution buffer. The concentrations of the PCR amplified fragments were quantified by NanoDrop®. The backbone plasmids and the fragments were assembled by infusion cloning.

5.2.5.2   NLS-SpCas9-TREX2-2A-EGFP, NLS-SpCas9-TREX2-2A-PURO

The *Sp*Cas9-TREX2-2A-EGFP and the *Sp*Cas9-TREX2-2A-PURO backbone vectors were digested by *Nco*I and *Eco*RV, in NEBuffer™ CutSmart, overnight at 37°C.

The NLS fragment was amplified from PX458 as follows:

Reagents	Volume (µl)	PCR program		
5X Q5 reaction buffer	5	98°C	30 sec	
dNTP-Mix (5mM/ dA,dG,dT,dC)	1	98°C	15 sec	
F primer (10 µM - Infusion-NLS-*Sp*Cas9-TREX2_F)	1.25	66°C	15 sec	x30
R primer (10 µM – Infusion-NLS-*Sp*Cas9-TREX2_R)	1.25	72°C	1 min	
Q5 enzyme	0.25	72°C	2 min	
PX458 (2 ng/µl)	1			
ddH$_2$0	15.25			

The PCR-amplified NLS fragment was purified by NucleoSpin Gel and PCR clean-up kit according to the manufacturer's PCR clean-up protocol and eluted in 15 µl elution buffer. The concentration of the PCR amplified fragments was quantified by NanoDrop®. The backbone plasmids and the fragments were assembled by infusion cloning.

5.2.5.3   TREX2-SpCas9-NLS-2A-EGFP

The PX458 backbone vector was digested by *Apa*I and *Age*I, in NEBuffer™ CutSmart, overnight at 37°C.

The TREX2 fragment was amplified from pKLV2.2-EF1-*Sp*Cas9-TREX2 as follows:

Reagents	Volume (µl)
5X Q5 reaction buffer	5
ddNTP-Mix (5mM/ dA,dG,dT,dC)	1
F primer (10 µM – Infusion_N-TREX2_F)	1.25
R primer (10 µM – Infusion_N-TREX2_R	1.25
Q5 enzyme	0.25
pKLV2.2-EF1-SpCas9-TREX2 (2 ng/µl)	1
ddH$_2$0	15.25

PCR program		
98°C	30 sec	
98°C	15 sec	
67°C	20 sec	x30
72°C	40 sec	
72°C	2 min	

Since the digestion of PX458 by *Apa*I and *Age*I leads to the elimination of part of the *Sp*Cas9 sequence, the digested *Sp*Cas9 sequence was re-amplified from PX458 as follows:

Reagents	Volume (µl)
5X Q5 reaction buffer	5
dNTP-Mix (5mM/ dA,dG,dT,dC)	1
F primer (10 µM – Infusion_N-TREX2-SpCas9_F)	1.25
R primer (10 µM – Infusion_N-TREX2-SpCas9_R	1.25
Q5 enzyme	0.25
PX458 (2 ng/µl)	1
ddH$_2$0	15.25

PCR program		
98°C	30 sec	
98°C	15 sec	
72°C	10 sec	x30
72°C	30 sec	
72°C	2 min	

The PCR-amplified fragments were purified by NucleoSpin Gel and PCR clean-up kit according to the manufacturer's PCR clean-up protocol and eluted in 15 µl elution buffer. The concentrations of PCR amplified fragments were quantified by NanoDrop®. The backbone plasmids and the fragments were assembled by infusion cloning.

5.2.5.4 NLS-TREX2-SpCas9-NLS-2A-EGFP

The PX458 backbone vector was digested by *Apa*I and *Age*I, in NEBuffer™ CutSmart, overnight at 37°C.

The NLS fragment was amplified from PX458 as follows:

Reagents	Volume (µl)	PCR program		
5X Q5 reaction buffer	5	98°C	30 sec	
dNTP-Mix (5mM/ dA,dG,dT,dC)	1	98°C	15 sec	
F primer (10 µM – Infusion_NLS-N-TREX2_F)	1.25	72°C	15 sec	x35
R primer (10 µM – Infusion_NLS-N-TREX2_R)	1.25			
Q5 enzyme	0.25	72°C	2 min	
PX458 (2 ng/µl)	1			
ddH$_2$0	15.25			

The TREX2-*Sp*Cas9 fragment was amplified from TREX2-*Sp*Cas9-NLS-2A-EGFP as follows:

Reagents	Volume (µl)		PCR program		
5X Q5 reaction buffer	5		98°C	30 sec	
dNTP-Mix (5mM/ dA,dG,dT,dC)	1		98°C	10 sec	
F primer (10 µM – Infusion_NLS-N-TREX2-SpCas9_F)	1.25		72°C	1 min 10 sec	x35
R primer (10 µM Infusion_N-TREX2-SPCas9_R	1.25				
Q5 enzyme	0.25		72°C	2 min	
TREX2-SpCas9-NLS-2A-EGFP (2 ng/µl)	1				
ddH$_2$0	15.25				

The PCR-amplified fragments were purified by NucleoSpin Gel and PCR clean-up kit according to the manufacturer's PCR clean-up protocol and eluted in 15 µl elution buffer. The concentrations of the PCR amplified fragments were quantified by NanoDrop®. The backbone plasmids and the fragments were assembled by infusion cloning.

5.2.6 Cloning of pKLV2.2-CAG-*Sp*Cas9-2A-EGFP and pKLV2.2-PGK-*Sp*Cas9-2A-EGFP lentiviral plasmids

5.2.6.1 pKLV2.2-CAG-SpCas9-2A-EGFP

The pKLV2.2-CAG-*Sp*Cas9-2A-EGFP was obtained by a 2-step infusion cloning protocol. First, the CAG cassette was cloned into the pKLV2.2 backbone plasmid, resulting in pKLV2.2-CAG. Secondly, to generate pKLV2.2-CAG-*Sp*Cas9-2A-EGFP, the *Sp*Cas9-2A-EGFP cassette was cloned into the pKLV2.2-CAG backbone plasmid.

5.2.6.1.1   pKLV2.2-CAG cloning

The pKLV2.2-mU6sgRNA5(*Sap*I)-hU6sgRNA5(*Bbs*I)-PGKpuroBFP-W backbone vector was digested by *Apa*I and *Not*I, in NEBuffer™ CutSmart, overnight at 37°C.

The CAG fragment was obtained by digesting PX458 with *Kpn*I and *Age*I, in NEBuffer™ 1.1, overnight at 37°C.

The digested pKLV2.2 backbone plasmid and CAG fragment were treated with DNA Polymerase I Large (Klenow) fragment 15 min at 25°C to generate compatible blunt ends:

Reagents	Volume (µl)
digested pKLV2.2 backbone plasmid or CAG fragment reaction	50
dNTP-Mix (5mM/ dA,dG,dT,dC)	1
Klenow (5 U/µl)	1

To prevent self-re-ligation, the digested pKLV2.2 backbone plasmid processed with the Klenow enzyme was further treated with calf intestinal alkaline phosphatase (CIP) by direct adding of 1 µl to the previous reaction and incubation for 30 min at 37°C. The pKLV2.2 backbone plasmid and CAG fragment were ligated by T4 ligation overnight at 16°C as follows:

Reagents	Volume (µl)
pKLV2.2 backbone plasmid (12 ng/µl)	4
CAG fragment (2.5 ng/µl)	10
10x T4 DNA ligation buffer	2
T4 DNA ligase (400 U/µl)	1
ddH$_2$0	3

1.5 µl of the mixture was transformed in chemically competent 10-beta E. coli bacteria cells by heat shock according to manufacturer's procedure, plated on an agar plates containing ampicillin 100 µg/ml, and incubated at 37°C overnight.

5.2.6.1.2   pKLV2.2-CAG-*Sp*Cas9-2A-EGFP

The pKLV2.2-CAG backbone plasmid was digested by *Blp*I, in NEBuffer™ CutSmart, overnight at 37°C.

The SpCas9-2A-EGFP fragment was amplified from PX458 as follows:

Reagents	Volume (µl)		PCR program		
5X Q5 reaction buffer	5		98°C	30 sec	
dNTP-Mix (5mM/ dA,dG,dT,dC)	1		98°C	10 sec	
F primer (10 µM – Infusion_pKLV-CAG-SpCas9_F)	1.25				x30
R primer (10 µM –Infusion_pKLV-SpCas9_R)	1.25		72°C	2 min	
Q5 enzyme	0.25		72°C	2 min	
PX458 (2 ng/µl)	1				
ddH$_2$0	15.25				

The PCR-amplified *Sp*Cas9-2A-EGFP fragment was purified by NucleoSpin Gel and PCR clean-up kit according to the manufacturer's PCR clean-up protocol and eluted in 15 µl elution buffer. The concentrations of the PCR amplified fragments was quantified by NanoDrop®. The backbone plasmids and the fragment were assembled by infusion cloning.

5.2.6.2   pKLV2.2-PGK-SpCas9-2A-EGFP cloning

Lentiviral transfer plasmids pKLV2.2-PGK-*Sp*Cas9-2A-EGFP was obtained by first cloning the *Sp*Cas9-2A-EGFP fragment, amplified from PX458, into pKLV2.2-mU6sgRNA5(*Sap*I)-hU6sgRNA5(*Bbs*I)-PGKpuroBFP-W. Secondly, the resulting pKLV2.2-2sgRNA-PGK-*Sp*Cas9-2A-EGFP was treated to remove the 2sgRNA cassette, encoding for two separated sgRNAs.

pKLV2.2-mU6sgRNA5(*Sap*I)-hU6sgRNA5(*Bbs*I)-PGKpuroBFP-W was digested by *Xho*I and *Not*I, in NEBuffer™ 3.1, overnight at 37°C.

The *Sp*Cas9-2A-EGFP fragment was amplified from PX458 as follows:

Reagents	Volume (µl)		PCR program		
5X Q5 reaction buffer	5		98°C	30 sec	
dNTP-Mix (5mM/ dA,dG,dT,dC)	1		98°C	10 sec	
F primer (10 µM Infusion_pKLV-PGK-SpCas9_F)	1.25		68°C	15 sec	x35
R primer (10 µM – Infusion_pKLV-PGK-SpCas9_R)	1.25		72°C	3 min	
Q5 enzyme	0.25		72°C	2 min	
PX458 (2 ng/µl)	1				
ddH$_2$0	15.25				

The PCR-amplified *Sp*Cas9-2A-EGFP fragment was purified by NucleoSpin Gel and PCR clean-up kit according to the manufacturer's PCR clean-up protocol and eluted in 15 µl elution buffer. The concentration of the PCR amplified fragment was quantified by NanoDrop®. The backbone plasmids and the fragment were assembled by infusion cloning.

To delete the 2sgRNA cassette, the resulting pKLV2.2-2sgRNA-PGK-*Sp*Cas9-2A-EGFP was then digested by *Apa*I and *Bam*HI, in NEBuffer™ 3.1, overnight at 37°C.

To generate compatible blunt ends for ligation, the digested pKLV2.2-PGK-*Sp*Cas9-2A-EGFP backbone vectors was treated with DNA Polymerase I as follows:

Reagents	Volume (µl)		PCR program	
pKLV2.2-PGK-*Sp*Cas9-2A-EGFP (20 ng/µl)	5		25°C	60 min
5x DNA polymerase I buffer	2		80°C	20 min
DNA polymerase I (20 U/ µl)	1			
ddH$_2$0	2			

The digested backbone vector was then ligated by T4 ligase overnight at 16°C as follows:

Reagents	Volume (µl)
pKLV2.2-PGK-*Sp*Cas9-2A-EGFP (from the previous reaction)	5
10x T4 DNA Ligase Buffer	2
T4 DNA Ligase (20 U/ µl)	1
ddH$_2$0	2

1.5 µl of the mixture was transformed in chemically competent 10-beta E. coli bacteria cells by heat shock according to manufacturer's procedure, plated on an agar plates containing ampicillin 100 µg/ml, and incubated at 37°C overnight.

5.2.7   sgRNA cloning in backbone plasmids

sgRNAs were cloned into the backbone plasmids by Golden Gate cloning of compatible overhanging ends between linearized backbone plasmids and sgRNAs. The Golden Gate cloning protocol enables simultaneous and directional cloning of DNA fragments into a recipient backbone plasmid, by parallel restriction digestion and ligation of the elements that have to be cloned (Engler & Mariollonnet, 2014). sgRNAs were ordered as single-stranded DNA oligo nucleotides from Integrated DNA Technologies

(IDT). When not otherwise specified, sgRNA sequences are 20 bp in length. For each sgRNA, the forward and respective reverse-complement DNA oligo nucleotides were ordered. Since the transcription of sgRNAs in the plasmid is driven by a U6 promoter, which preferentially recognizes a "G" nucleotide as a transcription start (Kunkel et al. 1986), an extra "G" nucleotide was added to the 5' end of the forward sequence for all the sgRNAs not starting with a "G" nucleotide. To the respective reverse-complement sequence, a "C" was added to the 3' end. A flanking Gibson-Assembly cloning adaptor sequence was included to the 5' end of each primer, being "CAGG" and "AAAC" for the forward and reverse-complement sequences of *Bbs*I-compatible sgRNAs, respectively. Whereas, for *Sap*I-compatible sgRNAs, a "TTT" stretch was added to the forward oligo and a "AAC" sequence to the reverse-complement one. Below, the configuration of the sgRNA oligos is exemplified:

	*Bbs*I-compatible oligo sequences 5'-3'	*Sap*I-compatible oligo sequences 5'-3'
Forward	CACCGNNNNNNNNNNNNNNNNNNNN	TTTGNNNNNNNNNNNNNNNNNNNN
reverse-complement	AAACNNNNNNNNNNNNNNNNNNNNC	AACNNNNNNNNNNNNNNNNNNNNC

To clone the sgRNA sequences in the backbone plasmids, forward and reverse-complement DNA oligo nucleotides were first phosphorylated and annealed by setting the following reaction:

Reagents	Volume (µl)		Time and Temperature	Cycle numbers
Forward DNA oligo (100µM)	1		37°C 30 min	1
reverse-complement DNA oligo (100µM)	1		95°C 5 min	1
10X T4 ligation buffer	5		Ramp down to 25°C at 1°C/12 sec	70
ATP 10mM –	5		20°C	∞
T4 Polynucleotide Kinase	1			
ddH$_2$0	37			

Annealed sgRNA DNA oligo nucleotides were then cloned in a backbone vector by simultaneous digestion and ligation (Golden Gate reaction) as follows:

Common step for all backbone vectors	
Reagents	**Volume (µl)**
Backbone Cas9-encoding plasmid (100ng)	X
BbsI-compatible annealed DNA oligo duplex (1:40 diluted)	2
10x NEBuffer™ CutSmart® Buffer	2
ATP 10mM	1
DTT 10mM	1
BbsI-HF (20 U/µl)	0.5
T7 Ligase (45 U/µl)	0.5
ddH$_2$0	Up to 20

To clone the second sgRNA in 2sgRNA-PX458, -459, -461 and -462 backbone vectors, the following reaction was set:

Only for 2sgRNA-PX458, -459, -461 and -462 backbone vectors

Reagents	Volume (µl)	PCR program		
Golden Gate reaction from the previous step	20	37°C	5 min	x6
*Sap*I-compatible annealed DNA oligo duplex (1:40 diluted)	2	21°C	5 min	
*Sap*I - (10 U/µl)	0.5	20°C	∞	

The Golden Gate reaction was treated with an ATP-dependent DNAse exonuclease, an enzyme able to selectively digest linear DNA, therefore preventing carry-over of unwanted recombination products. The reaction was set as follows:

Reagents	Volume (µl)		PCR program	
Golden Gate reaction	11		37°C	30 min
10X PlasmidSafe buffer	1.5		70°C	30 min
ATP 10mM	1.5		20°C	∞
Plasmid-Safe™ ATP-Dependent DNase enzyme	1			

1.5 µl of the mixture was transformed in chemically competent 10-beta E. coli bacteria cells by heat shock according to manufacturer's procedure, plated on an agar plates containing ampicillin 100 µg/ml, and incubated at 37°C overnight. Clones weres were screened by Sanger sequencing (primers used for sequencing in **Table 1**).

5.3 BACTERIAL TRANSFORMATION BY HEAT SHOCK IN CHEMICALLY COMPETENT 10-BETA *E. COLI* BACTERIA CELLS

By introducing a plasmid into a competent host bacterial strain, plasmids can be propagated exploiting the ability of bacteria to exponentially divide, while passaging down the plasmid to daughter cells (Johnston et al. 2014). Bacteria are chemically treated to induce permeability of the cell wall, while heat shock enhanced plasmid uptake by transiently loosening the cell wall (Mandel & Higa, 1970, Chang et al. 2017).

The introduction of the plasmids into recipient 10-beta *E. coli* bacteria cells was carried out as follows: 1.5 µl of plasmid mixture to transform was added to a frozen 25 µl competent bacteria cell aliquot, the mixture was then incubated 30 min on ice. Heat shock was performed on cells by incubation at 42°C for 30 seconds in a water bath. 400 µl pre-warmed SOC medium was added to the heat-shocked cell mixture and this was incubated for 1 hour at 37°C in an air orbital shaker. Thereafter, 100 µl incubated mixture was then plated on an agar plate and incubated overnight at 37°C.

5.4   COLONY PCR

In order to identify positive clones after bacterial transformation colony PCR was carried. Individual colonies were picked with a pipette tip and then dip into PCR cups containing 50 µl PCR mixture. The same pipette tip was used to inoculate clones to a fresh agar plate to create a backup of the screened colonies. Either vector- or insert-specific primers were used to amplify the insert of interest. PCR amplification was performed as follows:

Reagents	Volume (µl)
BOOM 17-Buffer	5
dNTP-Mix (5 mM/ dA,dG,dT,dC)	2
Primer F (10 µM)	2
Primer R (10 µM)	2
Taq DNA Polymerase (5 U/µl)	0.5
ddH$_2$O	38.5

5 µl PCR product was loaded on an agarose gel for visual validation of size of the expected bands, and agarose gel concentration dependent on PCR fragment size.

5.5   NUCLEIC ACIDS ISOLATION AND PURIFICATION

5.5.1   Plasmid DNA isolation and purification

5.5.1.1   MiniPrep plasmid DNA Protocol

Commercial kits for plasmid isolation and purification are based on alkaline lysis of bacteria, followed by plasmid binding to a silica membrane under high salt concentration. Subsequent washing of the silica membrane allows removal of contaminants (e.g. salt, protein, RNA), before plasmid elution. To extract plasmid DNA, the day before extraction single colonies were inoculated in 5 ml LB medium containing 50 µg/ml ampicillin by picking them with a pipette tip from the agar backup plate. The inoculated colonies were incubated overnight at 37°C in an orbital shaker incubator. The following

morning, the grown bacteria were pelleted by centrifugation at 1000xg 10 min in a swinging-bucket benchtop centrifuge. The supernatant was discarded and the pellet was processed according to the manufacturer's protocol. The concentration of purified plasmids was quantified by NanoDrop.

5.5.1.2   Endotoxin-free MaxiPrep plasmid DNA protocol

Transfection of primary cell lines require the use of high quality plasmids free of endotoxin (i.e. most commonly, bacterial lipopolysaccharides). This requirement is mostly needed to achieve high transfection efficiency. Although the connection is not completely elucidated, endotoxins seem to indirectly influence transfection efficiency by reducing cell viability (Weber et al, 1995, Butash et al. 2000).

To prepare endotoxin-free plasmids, the day before extraction single colonies are pre-inoculated in 5 ml LB medium containing 50 µg/ml ampicillin by picking them with a pipette tip from the agar backup plate and incubated 4/6 hours at 37°C in an orbital shaker incubator. The pre-inoculated 5 ml is added to 250 ml LB medium containing 50 µg/ml ampicillin in a 500 ml flask and incubated overnight at 37°C in an orbital shaker incubator. The following morning, the grown bacteria are pelleted by centrifugation at 1000xg 10 min in a swinging-bucket benchtop centrifuge. The supernatant is discarded and the pellet is processed according to the manufacturer's protocol. The concentration of purified plasmids is quantified by NanoDrop.

5.5.2   Genomic DNA isolation and purification

5.5.2.1   Standard protocols

gDNA was isolated and purified from cells cultured in suspension or adhesion either by eluting it from homogenizing columns of the peqGOLD Total RNA Kit during RNA isolation, or using the peqGOLD Tissue DNA Mini Kit according to manufacturer's protocol. DNA samples were quantified by NanoDrop and stored at -20°C. The homogenizing columns of the peqGOLD Total RNA kit bind and remove most of the gDNA from the lysed solution. The bound gDNA can be directly eluted, obtaining good quality gDNA for downstream applications.

5.5.2.2   Quick-and-dirty protocol to isolate genomic DNA

To extract gDNA from iPSC clones during the establishment of knocked-in colonies, a quick-and-dirty protocol was used to facilitate and streamline the characterization of positive clones. In brief, iPSCs were transferred in a 1.5 ml Eppendorf tube and pelleted at 500xg for 3 min. Upon pelleting, supernatant was discarded and 50 µl QuickExtract™ DNA extraction solution was added to the pellet. Cells were incubated at 65°C for 15 min followed by a deactivation step of 2 min at 98°C. 2-5 µl solution was used for PCR amplification. The ready-to-go QuickExtract™ DNA extraction solution is a proprietary formulation. It only requires heat treatment to induce cell lyses, release the DNA, and degrade compounds inhibitory to PCR amplification.

5.5.3   Total RNA isolation and purification

Commercial kits for isolation and purification of RNA are based on the use of a lysis buffer containing RNAse inhibitors – most commonly, guanidinium thiocyanate or guanidinium hydrochloride (Chirgwin et al, 1979). The lysate is passed through a silica membrane that specifically binds RNA. Subsequent washing of the silica membrane allows removal of contaminants (e.g. salt, protein), before RNA elution.

Total RNA was isolated and purified from cells cultured in suspension or adhesion using the peqGOLD Total RNA Kit according to the manufacturer's protocol. RNA samples were quantified by NanoDrop and stored at -80°C.

5.5.4   PCR reaction purification by magnetic beads

For sensitive downstream experiments, such as chip electrophoresis and next-generation sequencing (NGS) library preparation, PCR reactions were purified using magnetic AMPure XP SPRI beads. This technology exploits the ability of carboxyl-coated paramagnetic beads to reversibly bind DNA molecules. The purification was performed according to the manufacturer's protocol on a Magnetic-Ring Stand (96 well) or DynaMag™-2 Magnet. The DNA was eluted in 10 µl ¼ TE buffer.

5.6 QUANTIFICATION OF NUCLEIC ACIDS

5.6.1 Quantification of DNA, RNA and plasmid nucleic acids samples by NanoDrop® measurement

DNA, RNA and plasmids concentration and purity was determined using the NanoDrop® ND-1000 UV/VIS spectrophotometer. The volume loaded on the device for quantification was 1.5 µl/sample. To calculate the concentration and purity of nucleic acids in aqueous solution a wavelength of 260 nm was used. An optical density (OD) of 1 corresponds to 50 ng/µl dsDNA or 40 ng/µl RNA. Proteins show their maximum of absorption at 280 nm of wavelength, whereas organic compounds (e.g. EDTA, carbohydrates and phenol) have their maximal absorption at 230 nm. The ratios OD260/OD280 and OD230/OD280 are used to determine the level of protein and organic compound contaminations in the samples, respectively. High DNA purity is set with an OD260/OD280 ratio of 1.8, whereas high RNA purity is set at 2. Expected OD260/OD230 values are in the range of 2.0-2.2. Lower values may indicate the presence of organic contaminants in the samples.

5.6.2 Quantification of DNA samples by intercalating fluorescent dye

To accurately quantify individual samples before library pooling for NGS, a fluorometric assay was used. The assay is based on a fluorescent dye able to specifically intercalate only in the double-stranded DNA, offering higher sensitivity compared to UV absorbance quantification (NanoDrop). The presence of contaminants (e.g. organic compounds, free nucleotides, and salt) does not alter the quantification of nucleic acids quantified by intercalating fluorescent dye, as opposed to NanoDrop quantifications.

The quantification was performed according to the manufacturer's protocol (AccuBlue® NextGen dsDNA Quantitation Kit) on a Spark® microplate reader.

5.7 SEPARATION AND VISUALIZATION OF DNA PRODUCTS ON AGAROSE GEL BY ELECTROPHORESIS

To test the successful outcome of experiments involving PCR amplification of DNA or cDNA fragments and restriction digestion of plasmids, a small aliquot of these were loaded on an agarose gel and the products were resolved by electrophoresis. The density of the agarose gel matrix varied depending on

size of the products to be separated, typically ranging from 1 to 3%. Gels were prepared by boiling (up to 60°C) the agarose with 1X TBE. Once the agarose was dissolved it was cooled down in order to obtain a matrix of intertwined agarose molecules. Ethidium bromide (500 µg / ml) was added as nucleic acid intercalant to the gel in liquid phase while cooling at the final concentration of 0.5 µg/ml. Upon UV radiation (254 nm), the ethidium bromide intercalated in the DNA products becomes fluorescent, enabling the visualization of the nucleic acid molecules.

For PCR products, typically 5 µl of PCR product or restriction fragment mixtures were mixed with 5 µl of 10x loading buffer and loaded onto the gel pockets. To estimate the size of the nucleic acid molecules 8 µl of standard markers, KEB, the Gene Ruler 1 kb DNA Ladder, or 100 bp DNA Ladder was loaded in parallel to the samples. Visualization and photography of nucleic acid bands was carried out with a UV-Transilluminator Box, CCD camera and the Documentation System E.A.S.Y., Win32.

5.8 POLYMERASE CHAIN REACTION FOR AMPLIFYING DNA OR cDNA FRAGMENTS

PCR is a technique exploiting a DNA dependent polymerase able to amplify a given gDNA, DNA or cDNA molecule contained between a complementary forward and reverse oligonucleotide primers. The protocol for standard PCR is described below, in the case of DNA fragment amplification for subsequent cloning, the protocol is specified for each amplification in the section "Cloning plasmids".

Reagents	Volume (µl)
5X BOOM17 reaction buffer	2.5
dNTP-Mix (5mM/ dA,dG,dT,dC)	1
F primer	1
R primer	1
TAQ polymerase (5 U/ µl)	0.5
Template	2-5
ddH$_2$0	Up to 25

PCR program		
95°C	1 min	
95°C	15 sec	
58-62°C	15 sec	x35
72°C	1 min/kb	
72°C	2 min	

5.9 cDNA SYNTHESIS BY RETROTRANSCRIPTION OF RNA

To amplify RNA for any downstream application, it has to be retrotranscribed in cDNA. Temin and Baltimore discovered an enzyme having RNA dependent DNA polymerase activity (Baltimore, 1970, Temin & Mizutami, 1970). This enzyme, named RNA retrotranscriptase, can convert RNA molecules into their respective cDNAs. The newly synthetized cDNA can be used as template for PCR amplification.

For sensitive downstream applications, such as quantitative reverse transcription-PCR (qRT-PCR), RNA is pre-treated with DNase I to digest residual gDNA contaminations that can interfere with the analysis. Specifically, the protocol used was:

Reagents	Volume (µl)
Total RNA (1 µg)	X
10X DNaseI buffer	1
DNase I (1 U/µl)	1
ddH$_2$0	Up to 12.5
Following the DNase I digestion step	
DNase STOP buffer (50 µM EDTA)	1

PCR program	
25°C	15 min
70°C	10 min

DNaseI digestion step

STOP reaction step

cDNA was synthetized in 2 steps as follows:

Reagents	Volume (µl)		PCR program	
DNase I-treated RNA or RNA	13.5			
10 m M dNTP Mix	1	**1st step**	65°C	5 min
Oligo (dt)$_{18}$ primer	0.25			
Random hexamer primers	0.25			

5X RT Buffer	4		25°C	10 min
Maxima H Minus Enzyme Mix	1	**2st step**	50°C	15 min
			85°C	5 min

cDNA was used for downstream applications or stored at -20°C.

5.10 QUANTITATIVE REAL-TIME PCR

For all the qRT-PCR experiments, the method utilized was intercalant dyes, in particular SYBR® Green. Upon intercalating into double-stranded DNA molecules, the green fluorescence of the SYBR® Green at 530 nm increases up to 1,000 folds, therefore, enabling to track the increasing number of copies of the PCR products at each PCR cycle. Ultimately, a melting curve analysis is performed by progressive DNA denaturation from 50°C to 95°C of the qRT-PCR products. By trailing the fluorescence intensity signal as a function of temperature, the melting temperature of the PCR products can be determined. This analysis allows the profiling of the PCR products. A single peak corresponding to the melting temperature of the PCR products is expected. If multiple peaks are detected, unspecific products are also present, biasing the correct quantification.

The qRT-PCR consists in three different amplification parts. Firstly, the DNA template starts being amplified. In the next phase, the DNA template gets exponentially amplified. During this phase, called exponential phase, the threshold (CT value) is determined. Eventually, during the last phase, the

reagents become a limiting factor for the amplification, reaching the plateau. CT values from the amplification curves, as well as, melting temperatures were calculated using the 7500 System SDS software.

5.10.1   Relative quantification by ΔΔCt method

To assess the expression level of a target gene, the relative quantification method, also known as ΔΔCt, was used. The quantity of mRNA for a given target gene is relatively compared to the quantity of mRNA of a stably expressed gene, known as reference gene, and typically housekeeping genes are used (Livak & Schmittgen, 2001, Pfaffl et al. 2004). The expression of a housekeeping gene is stable throughout experimental conditions, allowing to relatively compare variation in the expression of a gene target when normalized to the housekeeping gene. Data were expressed as fold-change expression of the gene target between condition A and condition B. The following equations was used to calculate the ΔΔCt value and the fold change:

ΔCt = Ct-mean_Reference-gene- Ct-mean_target-gene

ΔΔCt = ΔCt_condition-A – ΔCt_condition-B

fold-change expression = $2^{-\Delta\Delta Ct}$

For the qRT-PCR analysis, three technical replicates per target and housekeeping gene were set. They were used to calculate the Ct mean per each condition. For each condition 50 ng of cDNA was used. As negative controls, duplicates with water instead of cDNA were used.

Reagents	Volume (µl)
SYBR Green Master Mix	10
cDNA (1:5 dilution) or water	5
F primer (5 µM)	2
R primer (5 µM)	2
ddH$_2$0	1

RT-qPCR program			
95°C	15 min		
Amplification	94°C	15 sec	
	58°C	30 sec	x40
	72°C	40 sec	
	72°C	2 min	
	95°C	15 sec	
Melting Curve	60°C	1 min	
	95°C	15 sec	

Interpretation of data on the relative expression of gene targets across different conditions was carried out with Excel.

5.11 SEQUENCING

The sequencing method used for double-stranded DNA is based on the Sanger sequencing chain-terminating technology (Sanger et al. 1977). Sanger-sequencing was performed applying the Big Dye® Terminator v1.1 Cycle Sequencing kit according to manufacturer's protocol and sequencing products were run on an automatic capillary Sequencer type ABI Prism 3130xl (capillary length: 36 cm, polymer: 3130 POP-7™) following manufacturer's manual. Quality check of electropherograms was performed with the software Sequencing Analysis 5.2. The online tool Benchling was used for sequence alignments to reference sequences.

The primers used for sequencing are listed in **Table 1**.

5.11.1   Sequencing of PCR products

To sequence PCR products (amplicons), primers and dNTP residues from the PCR reaction need removal. Therefore, the PCR reaction was cleaned-up with a hydrolyzing reagent (ExoSAP-IT),

determining the removal of nucleotides and degradation of primers to nucleosides and phosphate residues. Reaction mix was prepared as follows:

Reagents	Volume (µl)
PCR reaction	0.4-2
ExoSAP-IT (1:20 diluted with ddH$_2$O)	4
ddH$_2$0	Up to 6

PCR program	
37°C	15 min
80°C	15 min

The sequencing reaction mix is then prepared as follows:

Reagents	Volume (µl)
ExoSAP-IT-digested PCR reaction	6
Big Dye 3.1	1
Sequencing Primer (5 µM)	1
5x Big Dye buffer	1

PCR program		
96°C	4 min	
96°C	15 sec	
55°C	15 sec	x25
60°C	4 min	
8°C	∞	

5.11.2 Sequencing of plasmid DNA

For sequencing of plasmid DNA, the following sequencing protocol was used:

Reagents	Volume (µl)
Plasmid DNA (500 ng)	X
BigDye™ Terminator v1.1 Cycle	2
Sequencing Primer (5 µM)	2
5x BigDye™ Buffer	4
ddH$_2$0	Up to 20

PCR program		
96°C	4 min	
96°C	15 sec	
55°C	15 sec	x25
60°C	4 min	
8°C	∞	

5.11.3 Sequencing reaction precipitation

Subsequently, the sequencing reaction was precipitated by adding 75 µl ethanol 100% and 5 µl 125 mM EDTA to the mixture and centrifuging at 2577xg for 30 min at room temperature. The supernatant was discarded by inverting the PCR cups and centrifugation at 185xg for 10 sec. The pellet was washed by adding 60 µl 75% ethanol and centrifugation for at 2577xg for 10 min. Then supernatant was discarded and the pellet was dried by inverting the PCR cups and centrifugation at 185xg for 1 min. The pellet was dissolved in 50 µl Hi-Di™ Formamide. Seventeen µl was then loaded into the capillary electrophoresis 3130XL Genetic Analyzer for sequencing.

5.12 MINIGENE SPLICING RESCUE ASSAY

To preliminary validate the rescue potential of the various CRISPR/SpCas9 editing approaches, a minigene splicing assay was performed. Minigenes are in vitro models useful to investigate the molecular and genetic mechanisms of pre-mRNA splicing and usually constitute of a minimal gene fragment comprising an intron and flanking exons (Weisschuh et al 2012). In this project, minigene assays were used to determine the effect ofnovel DIVs on mRNA splicing (including the characherization of the novel c.5196+1134C>G DIV), and to evaluate gRNAs for their impact on genome editing and splicing rescue after genome editing. When co-transfecting a minigene plasmid

together with an editing plasmid, the effect of the latter on the splicing process can be studied. Specifically, following editing, the splicing rescue potential of a given CRISPR/SpCas9 strategy can be determined.

HEK293T cells were co-transfected with a total of 500 ng of a minigene, or a minigene and an editing plasmids at a molar ration of 1:10, respectively. Forty-eight hours post transfection, total mRNA was extracted and retrotranscribed using the primer "MinigeneSplicing_ABCA4_Ex37_R". Two µl cDNA was amplified using "MinigeneSplicing_ABCA4_Ex37_F" and "MinigeneSplicing_ABCA4_Ex37_R" and 5 µl were loaded on a 1.5% agarose gel. The remaining PCR reaction was purified with magnetic SPRI beads and analyzed on the Bioanalyzer.

5.13 TECHNIQUES TO ASSESS SPLICING RESCUE AND GENOMIC DNA CLEAVAGE EFFICACY

5.13.1 Tracking indels by decomposition (TIDE) and Inference of CRISPR Edits (ICE)

In order to preliminary validate the genomic cleavage efficacy of the designed gRNAs, the editing plasmids were individually transfected in HEK293T cells or fibroblasts[c.5197-557G/T]. The PX458 backbone plasmid encoding for EGFP was used to transfect HEK293T cells, whereas the PX459 backbone plasmid encoding for the puromycin resistance gene was used to transfect fibroblasts. 72h post transfection the gDNA was extracted and used for assessing the cleavage efficacy of the tested gRNAs. The validation was done by using two Sanger-sequencing-based protocol known as TIDE (Tracking indels by decomposition) (Brinkman et al. 2014) and ICE (Inference of CRISPR Edits) (Conant et al. 2022). In brief, a genomic fragment of ~700 nt encompassing the targeted sequence was PCR amplified and Sanger sequenced. The resulting elecropherograms were loaded on http://shinyapps.datacurators.nl/tide/ or https://ice.synthego.com/ to calculate the genomic cut efficiency. As a control, the electropherogram of the same genomic fragment derived from non-transfected cells was used.

5.13.2 Automated chip electrophoresis (Bioanalyzer)

DNA samples to be loaded on Bioanalyzer 2100 were quantified by NanoDrop. The kit used (Agilent DNA 1000 kit) allows size resolution included within 25 bp to 1000 bp and a quantification range, determined using the respective DNA ladder as sample, from 0.5 up to 50 ng / µl. Samples were diluted when exceeding the upper quantification limit. Chips were prepared and loaded on the reading device according to the manufacturer's protocol. Analysis reports reported fragment size and concentration in ng/µl, as well as, in nmol/l.

5.13.2.1 *Assessment of genomic DNA (gDNA) cut efficiency of the double sgRNA/SpCas9 approach*

The calculation of gDNA cut efficiency of the double sgRNA/*Sp*Cas9 approach was performed through chip electrophoresis. The size difference between unedited and edited fragments allowed consistent detection and quantification across samples. Percentage of cut efficiency was determined as (ED/(WT+ED))*100, where ED was the sum of the molarity of edited peaks and WT was the molarity of unedited or wild-type peaks. The analysis and calculations were performed on Excel.

5.13.2.2 *Assessment of splicing rescue*

The final validation of splicing rescue following the treatment with the various CRISPR/*Sp*Cas9 approaches was performed through chip electrophoresis. Percentage of splicing rescue values was determined as (CM/(CM+AM))*100, where CM was the molarity of correct peaks and AM was the molarity of aberrant peaks. The analysis and calculations were performed on Excel.

5.13.3 Next Generation Sequencing (NGS)

5.13.3.1 *Library preparation for NGS*

5.13.3.1.1 Amplification and indexing of gDNA samples

The calculation of gDNA cut efficiency of the double sgRNA/*Sp*Cas9 nickase and single gRNA/*Sp*Cas9 approaches was performed via high throughput sequencing. The genomic deletions or sequence alterations generated by these two approaches cannot be readily detected by chip electrophoresis, therefore high throughput sequencing was rendered necessary for accurate quantification. To prepare the samples for sequencing, a 2-step amplification protocol was used. In the first step, a genomic sequence of 461 nt, including the region targeted by the approaches, was amplified with degenerated primers containing universal illumina append sequences. The PCR amplification was performed as follows:

Reagents	Volume (µl)
5X Q5 reaction buffer	5
dNTP-Mix (5mM/ dA,dG,dT,dC)	1
F primer (10 µM NGS_illumina-adapter_ABCA4-In36_F)	1.25
R primer (10 µM – NGS_illumina-adapter_ABCA4-In36_R)	1.25
Q5 enzyme	0.25
gDNA template (50 ng/µl)	1
ddH$_2$0	15.25

PCR program		
98°C	30 sec	
98°C	10 sec	
68°C	15 sec	x30
72°C	10 sec	
72°C	2 min	

The second amplification was performed to add i5 and i7 illumina unique barcoding indexes to the individual samples. The reaction was performed as follows:

Reagents	Volume (µl)	PCR program		
2x KAPA HiFi HotStart ReadyMix	5	95°C	3 min	
i7 index primer (N700/N708 - Nextera XT DNA Library Preparation Kit)	1	98°C	20 sec	
I5 index primer (S502/S507 - Nextera XT DNA Library Preparation Kit)	0.3	70°C	15 sec	x35
R primer (10 µM – NGS_illumina-adapter_ABCA4-In36_R)	0.3	72°C	30 sec	
First amplification (1:200 dilution)	1	72°C	5 min	
ddH$_2$0	2.4			

Three µl were loaded on a 1.5% Agarose gel for visual validation of size of the expected band.

5.13.3.1.2 Indexed amplicons purification, quantification and pooling

gDNA amplicon samples were purified following the protocol reported in the "Methods: 5.5.4" section. The purification was performed on a magnetic-ring stand (96-well). Whereas, the quantification of gDNA amplicon samples was performed according to the "Methods: 5.6.2" section. The different samples were then pooled together at the same molarity to obtain a final library concentration of ~20 µM.

5.13.3.2 NGS and data analysis

The library was sequenced on a MiSeq™ System – Illumina using a MiSeq Reagent Kit v2 Nano (2 x 250 bp). Primary data anlysis and de-multiplexing was performed by the c.ATG / NGS Competence Center Tübingen (NCCT). Secondary analysis was performed on https://crispresso.pinellolab.partners.org/.

5.14 CELL CULTURE AND ALLIED PROCECURES

5.14.1 HEK293T cells

5.14.1.1 HEK293T culturing conditions

HEK293T cells were thawed from a frozen vial maintained in liquid nitrogen. The vial was placed on ice for transport and quickly thawed in a water bath set at 37°C. Cells were transferred in a 50 ml Falcon tube filled with 10 ml pre-warmed DMEM+ and centrifuged at 300xg for 3 min. The supernatant was discarded, cells resuspended in 4 ml fresh DMEM+ and plated in a T75 flask in 13 ml DMEM+ final volume. Cells were grown and maintained at 37 °C in a 5% CO_2 humidified atmosphere. Upon 80% confluence, cells were passaged (usually at 1:20/30 dilution ratio) in a T75 flask. To do so, spent medium was aspirated and cells gently washed once in PBS. One ml 0.05% trypsin was added on cells and let incubated 2 min at room temperature. Cells were harvested in 12 ml DMEM+ in a 50 ml Falcon tube. For cell maintenance, 400-600 µl of cell suspension was seeded in a new T75 flask in 13 ml DMEM+ final volume. For cell seeding for downstream transfection, the remaining cell suspension was centrifuged 300xg for 3 min, the supernatant discharged and the cell pellet resuspended in 4 ml DMEM- (DMEM without addition of penicillin and streptomycin). Ten µl cell suspension was loaded on a Cell Chip™ and cells were counted on a Spark® microplate reader. Cells were seeded on DMEM-, the seeding density depended on the protocol used: 250,000 cells per a well of a 24-well plate and 7 million cells per well of a 10 cm dish.

5.14.1.2 HEK293T cell transfection by lipofection

Lipofection transfection enables delivery of nucleic acids into cells by inducing the assembly of liposomes particles where nucleic acids associate to cationic lipids. The liposome particles are internalized in cells by endocytosis, thereby resulting in the transfection of these cells with the nucleic acids (Felgner et al. 1987). Cells seeded for transfection were incubated overnight at 37 °C in a 5% CO_2 humidified atmosphere. Cells were transfected by Lipofectamine™ 3000 Transfection Reagent according to the manufacturer's protocol using plasmids purified with MiniPrep plasmid DNA Protocol.

5.14.1.3 Lentivirus production

Lentiviral production is a biological safety level 2 procedure, and an S2 application was submitted and granted (UNI.TÜK.61.07). To produce lentiviral particles, a 2nd generation protocol was used. A transfer (pKLV2.2), the envelope (pMD2.G), and the packaging (psPAX2) plasmids were co-transfected with lipofection in HEK293T cells at a molar ratio of 2:2:1, respectively. The production was set in 10-cm dishes. In brief, 5 Mio cells were seeded in 12 ml of DMEM- and incubated overnight at 37 °C in a 5% CO_2 humidified atmosphere. The following day the cells were transfected with a total amount of 15 µg of plasmids. Seven hours post transfection the medium was replaced with 12 ml DMEM+ containing 2% FBS. Supernatant was harvested 72 h post transfection, filtered through a 0.45 µm filter, and concentrated via ultracentrifugation in a 20% sucrose cushion. In brief, 12 ml of filtered supernatant was placed in 50 ml ultracentrifuge tube and 10 ml 20% sucrose cushion solution was carefully added at the bottom of the tube. The solution was centrifuged at 82.700xg for 2 h at 4°C. The supernatant was carefully discarded and the pellet resuspended in 150 µl OPTIMEM medium. The lentiviral preparation was stored at -80 °C. Titration of the lentiviral preparations was not performed.

5.14.2 WERI-Rb1

5.14.2.1 WERI-Rb1 culturing conditions

WERI-Rb1 cells were thawed from a frozen vial maintained in liquid nitrogen. The vial was placed on ice for transport and quickly thawed in a water bath set at 37°C. Cells were transferred in a 50 ml Falcon tube filled with 10 ml pre-warmed RPMI + and centrifuged at 300xg for 5 min. The supernatant was discarded, cells were resuspended in 5 ml fresh RPMI+ and plated in a T25 flask. Cells were grown and maintained at 37 °C in a 5% CO_2 humidified atmosphere. Cells were counted every other day until a density of 200,000 cells/ml. At this point, cells were maintained or expanded at 50,000 cells/ml. To do so, cell suspension was transferred in a 50 ml Falcon tube and centrifuged at 300xg for 6 min. Supernatant was discarded and cells were resuspended in 2 ml RPMI+. After cell counting, cells were resuspended at 50,000 cells/ml in fresh RPMI+ in a T25 or T75, depending on the volume of cell suspension, and incubated at 37 °C in a 5% CO_2 humidified atmosphere.

5.14.2.2 WERI-Rb1 transfection by electroporation

WERI-Rb1 are cells considered hard to transfect, especially with commercial kits based on lipofection. However, we successfully established an efficient transfection protocol based on electroporation. Electroporation transiently increases the permeability of cells, allowing internalization of nucleic acids in solution with cells (Chu et al. 1987). In brief, cells were counted when in suspension and the volume of cell suspension corresponding to 500,000 cells per transfection was transferred in a 2 ml Eppendorf tube. Cells were centrifuged 300xg for 5 min and the supernatant discarded. Cell pellet was resuspended in buffer T at a 500,000 cells/100 µl. Eight µg endotoxin-free plasmid to be transfected into WERI-Rb1 cells was mixed with 100 µl cell suspension in buffer S in a 2 ml Eppendorf tube and loaded in a 100 µl electroporation tip using the Neon® electroporation pipette. The Neon® electroporation pipette was placed in the electroporation station loaded with 3 ml electroporation buffer. Electroporation setting used were: 1400 V, 10 ms, 2 pulses. Once electroporated, cells were transferred in 2 ml pre-warmed RPMI- in a well of a 6-well plate and incubated at 37 °C in a 5% CO_2 humidified atmosphere.

5.14.3   Fibroblasts

5.14.3.1   Skin punch biopsy culture for propagation of patient-derived fibroblast cell lines

Patient-derived fibroblast cell lines were established starting from skin punch biopsies collected upon informed written consent complying with the guidelines and approved by the local ethics committee (Project no. 124/2015BO1).

5.14.3.1.1 Skin punch biopsy collection

The collection of skin punch biopsies was performed by trained clinicians. The site of biopsy was disinfected with Kodan® (it contains 2-Propanol, 1-Propanol, and Biphenyl-2-ol). Subsequently, the skin was locally anesthetized by Scandicain® 2%. Skin biopsy was performed with a sterile 4 mm skin punch and cut off by a sterile scissor. The skin punch was then transferred in a 30 ml tube filled with 15 ml DMEM+ and immediately processed.

5.14.3.1.2 Skin punch processing for fibroblast cells isolation and expansion

Skin biopsies were dissected to separate the epidermis and fat layers from the dermis layer. This procedure was aided by using sterile scissors, tweezers, and pipette tips. The dissection was carried out in a 10-cm dish filled with 15 ml PBS. Upon dissection, the dermis layer was transferred in a 2-cm dish filled with 1 ml collagenase and minced to obtain smaller pieces. To prepare the collagenase solution, the proper amount of powdered collagenase was dissolved in DMEM+f (20% fetal bovine serum, 10 U/ml penicillin/streptomycin) at a final concentration of 10 mg/ml. Aliquots of 1 ml were pre-prepared and stored at -20°C for a maximum of 6 months. The skin punch biopsy was incubated overnight at 37 °C in a 5% CO_2 humidified atmosphere. The following day, the collagenase solution was transferred in a 50 ml Falcon tube and centrifuged at 300xg for 5 min. Supernatant was discarded and the skin punch biopsy dissolved cell pellet was resuspended in 2 ml DMEM+f which was then plated in a well of a 6-well plate. The plate was incubated at 37 °C in a 5% CO^2 humidified atmosphere. Medium was refreshed every 4 days until confluent areas of fibroblasts appeared. At this point, spent medium was removed and the well washed once with PBS. 300 µl trypsin 0.05% was added to the fibroblasts and let incubated at 37 °C for 4 min. Fibroblasts were harvested with 2 ml DMEM+f and expanded in a T25 flask with 5 ml DMEM+f final volume. Medium was refreshed every 4 days until confluence and then expanded in a T75 in 13 ml DMEM+f (fibroblasts were detached using 0.5 ml trypsin 0.05%). Upon confluence, fibroblasts were harvested and resuspended in 10 ml DMEM+f in a 50 ml Falcon tube. Half of the cell suspension was expanded in a T175 in 23 ml DMEM+f final volume, whereas, the remaining 5 ml was centrifuged at 300xg for 5 min and resuspended in 2 ml freezing medium (70% DMEM+, 20% FBS, 10% DMSO) which was equally split in two cryogenic vials and placed at -80 °C in a freezing box for 24 hours. The vials were then placed in liquid nitrogen for long-term storage. Upon T175 confluency, four cryogenic vials were made following the same procedure and stored in liquid nitrogen.

5.14.3.1.3 Thawing of fibroblasts

Fibroblasts were thawed from a frozen vial maintained in liquid nitrogen. The vial was placed on ice for transport and quickly thawed in a water bath set at 37°C. One ml pre-warmed DMEM+f was added on cells dropwise. The resulting 2 ml were transferred in a 50 ml Falcon tube using a 5 ml pipette. 8 ml pre-warmed DMEM+f was added on the cell suspension dropwise and centrifuged at 200xg for 2 min. The supernatant was discarded, cells were resuspended in 2 ml fresh DMEM+f containing 2 ng/µl hFGF and plated in a T75 flask pre-filled with 12 ml DMEM+f containing 2 ng/µl

hFGF. Cells were grown and maintained at 37 °C in a 5% CO2 humidified atmosphere. The medium was refreshed every four days with DMEM+f until 80% cell confluency.

5.14.3.1.4 Fibroblast culturing and transfection for TIDE and ICE analysis

Upon establishment of the fibroblast cell line, fibroblasts were cultured in DMEM supplemented with 20% fetal bovine serum, 10 U/ml penicillin/streptomycin at 37 °C in a 5% CO_2 humidified atmosphere. Upon 80% confluency, cells were passaged (usually at 1:4 dilution ratio) in a T175 flask. To do so, spent medium was aspirated and cells gently washed once in PBS. Two ml 0.05% trypsin was added on cells and let incubated for 4 min at room temperature. Cells were harvested in 12 ml DMEM+ in a 50 ml Falcon tube. For cell maintenance, 3 ml of cell suspension was seeded in a new T175 flask in 25 ml DMEM+ final volume. To transfect the cells through electroporation, the remaining cells were centrifuged 300xg for 3 min and resuspended in 3 ml PBS. Ten μl cell suspension was loaded on a Cell Chip™ and cells were counted on a Spark® microplate reader. The cells were centrifuged again. The cell pellet was then resuspended in buffer R at 1 million cells/100μl. Five μg plasmid was mixed with 100 μl cell suspension in buffer R and loaded in a 100 μl electroporation tip using the Neon® electroporation pipette. The Neon® electroporation pipette was placed in the electroporation station loaded with 3 ml electroporation buffer. The electroporation setting used was: pulse voltage 1400 V, 20 ms, 2 pulses. Once electroporated, cells were split in two wells of a 6-well plate filled with 2 ml DMEM medium without antibiotics and incubated at 37 °C in a 5% CO_2 humidified atmosphere. Twenty-four hours post electroporation 2μg/ml puromycine was added to the medium.

5.14.4 Induced pluripotent stem cells establishment, culturing and manipulation

In 2006, Takahashi and Yamanaka discovered that the overexpression of four factors: Oct3/4, Sox2, c-Myc, and Klf4, were able to induce the reprogramming of adult mouse fibroblasts into pluripotent stem cells (Takahashi & Yamanaka, 2006). iPSCs are self-renewal cells able to give rise to differentiated progeny representative of all three germ layers of the body: ectoderm, endoderm, and mesoderm (Shafa et al. 2018). They grow in round compact colonies characterized by distinct boarders and well-defined edges.

iPSCs were obtained from patient-derived fibroblasts reprogrammed by either Sendai viral transduction or episomal protocols. For deep-intronic variants for which fibroblast cell lines were not

available, control iPSCs (co-53 iPS) were used as an isogenic cell line for knock-in of DIVs by means of CRISPR/Cas9. iPSCs were differentiated in 2-dimensional cone photoreceptor precursor cell cultures via a 30-day protocol (Flamier et al. 2016).

5.14.4.1 Coating of plates for 5.13.4 induced pluripotent stem cells culturing and seeding

iPSCs are cultured on a protein-based matrix. For the experiments described in this thesis, vitronectin coated plates were used for iPSCs establishment, manipulation and culturing, whereas, Matrigel – coated plates were used for differentiation in photoreceptor precursor cells (Flamier et al. 2016).

To coat plates with vitronectin, the stock solution (500 µg⁄ml) was diluted 1:100 v/v in PBS in a 50 ml Falcon tube. Depending on the plate or dish format, different volumes of diluted vitronectin were used for coating: 0.5 ml for 24-well plates, 2 ml for 6-well plates and 6 ml for 10-cm dishes. Coated plates or dishes were incubated at room temperature 2 hours before use. Before using coated plates, the coating solution was aspirated and replaced with culturing medium as specified in the relevant sections below.

To coat plates with Matrigel, aliquots of stock solution were thawed on ice and diluted 1:50 v/v in cold DMEM/F12 medium using cold pipette tips in a cold 50 ml Falcon tube. 0.5 ml diluted Matrigel was used to coat 12-well plates pipetted using cold tips. Coated plates were incubated at 37 °C 2 hours. Before using coated plates, the coating solution was aspirated and replaced with culturing medium as specified in the relevant sections below.

5.14.4.2 Reprogramming of patient-derived fibroblast cell lines in 5.13.4 induced pluripotent stem cells

5.14.4.2.1 Sendai reprogramming

The reprogramming of the iPSCs heterozygous for the DIV *ABCA4* c.5197-557G>T (iPSC$^{c.5197-557G/T}$) was outsourced to the Stem Cell Technology Center, Radboud University Medical Center, The Netherlands, as a commercial service. The Het-c.5197-557G>T patient-derived fibroblasts were reprogrammed in iPSCs by CytoTune™-iPS 2.0 Sendai Reprogramming Kit (ThermoFisher Scientific) according to the manufacturer's protocol and were cultured in complete Essential E8 medium supplemented with 10 U/ml penicillin/streptomycin at 37 °C in a 5% CO2 humidified atmosphere. As part of the

reprogramming service, iPSCs were characterized for pluripotency markers by reverse transcription-qPCR (RT-qPCR) and immunohistochemistry. Differentiation potential was evaluated by trilineage differentiation, and the karyotype was confirmed (**Annex 1**).

5.14.4.2.2 Episomal reprogramming

Conversely, the iPS cell line heterozygous for the DIV *ABCA4* c.5196+1134C>G (iPSCs [c.5196+1134C/G]) was generated in-house by episomal reprogramming as described below. Patient-derived fibroblasts were thawed from a frozen vial maintained in liquid nitrogen. The vial was placed on ice for transport and quickly thawed in a water bath set at 37°C. Cells were transferred in a 50 ml Falcon tube filled with 10 ml pre-warmed DMEM + and centrifuged at 300xg for 5 min. The supernatant was discarded, cells resuspended in 4 ml fresh DMEM+ and plated in a T75 flask in 13 ml DMEM+ final volume. Cells were grown and maintained at 37 °C in a 5% CO_2 humidified atmosphere. Upon 70% confluency, cells were expanded in a T175 following the same protocol as described in the "Methods: 5.14.3.1.2" section. Upon 70% confluency, cells were harvested in a 50 ml Falcon tube, resuspended in 4 ml PBS, counted and centrifuged at 300xg for 4 min. Cell pellet was resuspended in buffer R at a 1 Milion cells/100 μl. 4 μg of endotoxin-free plasmid for each of the three reprogramming plasmids (pCXLE-hUL, pCXLE-hSK and pCXLE-hOCT3/4-shp53-F) was mixed with 100 μl cell suspension in buffer R and loaded in a 100 μl electroporation tip using the Neon electroporation pipette. The Neon electroporation pipette was placed in the electroporation station loaded with 3 ml electroporation buffer. The electroporation setting used was: pulse voltage 1650 ms, pulse width 10ms and pulse number 3. Once electroporated, cells were transferred in 15 ml pre-warmed DMEM- in a 10 cm dish and incubated at 37 °C in a 5% CO_2 humidified atmosphere, the day of electroporation was regarded to as day 0. Until day 2, medium was refreshed daily with 15 ml DMEM+. At day 3, cells were harvested, counted and plated in a vitronectin-coated 6-well plate at 80,000 cells/well in 2 ml 75% DMEM+ and 25% E8+. At day 4, medium was refreshed with 50% DMEM+ and 50% E8+. From day 5 on, medium was refreshed daily with 2 ml E8+. Within day 15 to day 28, iPS colonies grew to an appropriated size for picking and transfer. iPS colonies were picked under an EVOS™ XL core microscope. First, the edges of the iPS colonies were cleaned from unreprogrammed fibroblasts with a 10 μl pipette tip, then, the colonies were cut in pieces using a 22-gauge needle and transferred in a new vitronectin-coated well of a 6-well plate filled with 2 ml E8+ using a 20 μl pipette. The colonies were grown and maintained as described in the "Methods: 5.14.4.3" section.

5.14.4.3 Induced pluripotent stem cells culturing conditions

iPSCs were thawed from a frozen vial maintained in liquid nitrogen. The vial was placed on ice for transport and quickly thawed in a water bath set at 37°C. 1 ml pre-warmed E8+ was added on cells dropwise. The resulting 2 ml were transferred in a 50 ml Falcon tube using a 5 ml pipette. 8 ml pre-warmed E8+ was added on the cell suspension dropwise and centrifuged at 200xg for 2 min. The supernatant was discarded, cells were resuspended in 2 ml fresh E8+ containing Revitacell and plated in two vitronectin-coated wells of a 6-well plate pre-filled with 1 ml E8+ containing Rivitacell per well. Cells were grown and maintained at 37 °C in a 5% CO_2 humidified atmosphere. The medium was refreshed daily with E8+ until 80% cell confluency.

For cell maintenance and passaging, spent E8+ medium was discarded and iPSCs were washed once with PBS. 1 ml Versene solution was added on cells and immediately aspirate to only leave a thin layer covering the cells. iPSCs were then incubated at 37 °C for 4 to 10 min until the edges of the iPS colonies appeared slightly rounded up and/or folded away from the plate surface. Two ml E8+ was added on cells and gently pipetted up and down to help detach cells in clumps (use a 5 ml pipette). The cell suspension was transferred in a 50 ml Falcon tube and 4 to 13 ml fresh E8+ medium was added depending on the final dilution at which cells have to be passaged. The splitting ratio is mainly dependent on the cell line handled and can only be assessed by experience. One ml cell suspension was added on a new vitronectin-coated 6-well plate filled with 1 ml E8+. Cells were grown and maintained at 37 °C in a 5% CO_2 humidified atmosphere refreshing the medium daily until 80% cell confluency.

5.14.4.4 Knock-in of deep-intronic variants in a control induced plutipotent stem cell line

The co-53 iPS cell line was thawed and grown in 6-well plate format as described in the "Methods: 5.14.4.3" section. Upon 80% cell confluence, cells were harvested in single cell solution using accutase. In brief, spent medium was discarded and cells were washed in PBS. Five hundred µl accutase was added on cells and incubated at 37 °C for 3 min. Two ml E8+ was added and cells were collected in a 50 ml Falcon tube using a 1000 µl pipette. Cell suspension was centrifuged at 300xg for 5 min and resuspended in PBS. Following cell counting, the volume of cell suspension corresponding to 500,000 cells per transfection was transferred in a 2 ml Eppendorf tube. Cells were centrifuged at 300xg for 5 min and the supernatant discarded. Cell pellet was resuspended in buffer R at 500,000 cells/100 µl. The sgRNA/*Sp*Cas9 or /*As*Cas12a complex was assembled by incubating 120 pmol *Sp*Cas9 or *As*Cas12a

with 250 pmol sgRNA 15 min at room temperature. Prior to the addition of the sgRNA/*Sp*Cas9 or /*As*Cas12a complex to the cell suspension, 400 pmol ssODN was added resulting in the ssODN/sgRNA/*Sp*Cas9 or /*As*Cas12a mixture. The ssODN/sgRNA/*Sp*Cas9 or /*As*Cas12a mixture was mixed with 100 µl cell suspension in buffer R in a 2 ml Eppendorf tube and loaded in a 100 µl electroporation tip using the Neon electroporation pipette. The Neon electroporation pipette was placed in the electroporation station loaded with 3 ml electroporation buffer. The electroporation setting used was: 1400 V, 5 ms, 3 pulses. Once electroporated, cells were transferred on a vitronectin-coated well of a 6-well plate filled with 2 ml E8+ containing 1x CloneR and incubated at 37 °C in a 5% CO_2 humidified atmosphere. Medium was refreshed daily until 80% cell confluence. At this point, cells were detached using accutase, harvested in a 50 ml Falcon tube and resuspended in 4 ml E8+ containing 1x CloneR. After counting, cells were diluted in E8+ containing 1x CloneR at 1000 cells/ml and 2 ml plated on a vitronectin-coated 10-cm dish filled with 13 ml E8+ containing 1x CloneR. Cells were let grow for 72 hours without replacing medium. Thereafter, medium was replaced daily with 15 ml E8+ until iPS colonies were around 2 mm in diameter. Individual colonies were picked using a 20 µl pipette and transferred on a well of a 24-well plate (usually 48 colonies were picked). Medium was refreshed daily starting from 48-hours post picking until the picked colonies were around 2/5 mm in diameter. Grown colonies were detached using versene as described in the "Methods: 5.14.4.3" section using 300 µl Accutase and 500 µl E8+ for harvesting (a 1000 µl pipette was used). 300 µl cell suspension was transferred on a new vitronectin-coated well of a 24-well plate filled with 1 ml E8+, whereas, the remaining 200 µl was transferred in a 1.5 ml Eppendorf tube and centrifuged at 1000xg for 5 min. Supernatant was discarded and the pellet was processed as described in the "Methods: 5.5.2.2" section for genotyping and selection of successfully knocked-in iPS colonies. The extracted gDNA was used to amplify the genomic locus of interest as described in the "Methods: 5.8" section, which was then sequenced following the protocol reported in the "Methods: 5.11.1" section. Upon sequencing, positive clones were expanded and maintained as described in the "Methods: 5.14.4.3" section.

5.14.5   Cone photoreceptor precursor cells

iPSCs can principally differentiate in any cell type (Kolagar et al. 2020, Wang & Farzaneh, 2020, Zhao et al. 2021). Here, a protocol for iPSCs to be differentiated in cells expressing early cone specific markers, commonly named cone photoreceptor precursor cells (CPCs), was used (Flamier et al. 2016). This cell type has been widely used for the validation of aberrant splicing resulting from DIVs and for

the assessment of AON-mediated splicing correction. For the purpose of this thesis, this cell type has been selected as it expresses discreet amounts of *ABCA4* transcripts.

5.14.5.1 Differentiation of photoreceptor precursor cells

iPSCs to be differentiated in CPCs were thawed and grown as described in the *"iPSCs culturing conditions"* section. Upon 80% cell confluence, cells were detached using accutase as described in the "*Knock-in of deep-intronic variants in a control iPS cell line section*". Cells were resuspended in E8+ containing 1x Revitacell, counted and 600,000 cells were plated on a matrigel-coated well of a 12-well plate in 1 ml E8+ containing 1x Revitacell. The following day, cells were expected to be confluent and the medium was switched to differentiation medium which was refreshed daily for 30 days. The differentiation medium was prepared weekly and aliquoted in seven 50 ml Falcon tubes which were stored at -20°C and thawed overnight at +4°C. The differentiation medium was prepared as follows (volumes for 100 ml differentiation medium):

Reagents	Volume
DMEM/F12	95.8 ml
NEAAs (x100)	1 ml
N2 supplement (x100)	1 ml
B27 supplement (x50)	2 ml
Insulin-like growth factor 1 (IGF-1 – 100 ng/µl)	10 µl
Human Fibroblast growth factor basic (hFGF-basic 100 ng/ µl)	25 µl
Heparin (50 µg/ µl)	2 µl
Human recombinant COCO (200 µg/ml)	15 µl
Primocin (x500)	200 µl

5.14.5.2   *Electroporation of cone photoreceptor precursor cells*

At day 30 of differentiation, CPCs were transfected via electroporation with different *Sp*Cas9 plasmids. All the plasmid DNA used were endotoxin-free. Spent medium was discarded and CPCs washed with PBS. One ml accutase was added on cells and let incubate at room temperature for 10 min. Frequently up-and-down pipetting with a 1000 µl pipette helped to disaggregate cell clumps. Cells were harvested in a 50 ml Falcon tube filled with 10 ml differentiation medium and centrifuged at 500xg for 6 min. Supernatant was discarded, cells washed in PBS, counted and centrifuged. The cell pellet was resuspended in buffer R at 2 millions cells/100µl. Twenty µg plasmid was mixed with 100 µl cell suspension in buffer R and loaded in a 100 µl electroporation tip using the Neon electroporation pipette. The Neon electroporation pipette was placed in the electroporation station loaded with 3 ml electroporation buffer. The electroporation setting used was: pulse voltage 1100 V, 30 ms, 2 pulses. Once electroporated, cells were transferred in a Matrigel-coated well of a 6-well plate filled with 2 ml differentiation medium without antibiotics and incubated at 37 °C in a 5% CO_2 humidified atmosphere. Differentiation medium was refreshed daily.

5.14.5.3   *Transduction of cone photoreceptor precursor cells*

An estimated amount of 2 million CPCs were transduced with 50 µl of the concentrated lentiviral vector preparation. Fifty µl of concentrated lentiviral preparation was either directly added to CPCs in 1 ml differentiation medium (for a well of a 12-well plate), or to CPCs which had been detached using accutase (3 min at 37°C), resuspended in 1 ml differentiation medium and plated on a freshly coated well of a 12-well plate. Polybrene was added to enhance transduction at a final concentration of 8 µg/ml. CPCs were analyzed 7-days post transduction.

5.14.6 Freezing

Freezing of cells was performed in the following media:

Cells	Freezing medium
HEK293T	DMEM+ + 10% DMSO
WERI-Rb1	RPMI+ + 10% DMSO
Fibroblasts	DMEM+f + 10% DMSO
iPSCs	KnockOut Serum Replacement + 10% DMSO

Cells to be frozen were harvested in a 50 ml falcon tube according to the specific cell line protocols described above. The cell pellet was resuspened in the corresponsing freezing medium and 1 ml of cell suspension was transferred in a cryovial. Cryovials were placed in a CoolCell LX overnight at -80°C. For long term storage, cryovials were stored in liquid nitrogen. The amount of HEK293T and WERI-Rb1 cells frozen per cryovial varied considerably (~100.000/1 Mio cells), the amount of fibroblasts present in a confluent T175 flask was normally frozen in four cryovials, whereas from a confluent iPSC well of a 6-well plate, ~4/10 cryovials were made. CPCs were never frozen.

5.15 FLUORESCENCE-ACTIVATED CELL ANALYSIS OF TRANSFECTED CELLS

To evaluate the transfection and transduction efficiency of the different delivery protocols achieved on CPCs, fluorescence-activated cell analysis (FACS) was employed. CPCs undergoing the analysis were detached in single cells by 5 min incubation in a thin layer of accutase at room temperature and filtered through a 100 µm cell strainer. CPCs were resuspended in PBS at a final concentration ranging from 250,000 to 2 million cells/ml in a 5 ml FACS tube. Non-transfected CPCs were also included for each experiment as a negative control to set the acquisition parameters. Tubes were placed on ice for transport and cells were analyzed on a BD FACS Canto II at the Flow Cytometry Core Facility – Berg of the Medical Faculty of the University of Tübingen.

To differentiate the cell population of interest while excluding debris, cells were first analyzed by plotting forward (FSC) and side scatters (SSC) signals, allowing the identification and gating of only living cells. To discriminate between single cells (singlets) from cell clumps (duplets) passing through the detector, the area of the FSC signal (FSC-A) was plotted against the height of the same signal (FSC-H). A 488 nm laser was used to detect the intensity of the EGFP signal. EGFP shows maximal excitation

at 488 nm and maximal emission at 507 nm. Non-transfected cells were used to set the fluorescence background at 488 nm. Results were displayed on a histogram showing the intensity of the EGFP signal on the X axis and count of cells normalized to mode on the Y axis.

5.16 STATISTICAL ANALYSIS

Statistical significance of RT-qPCR experiments was performed on BootstRatio: http://rht.iconcologia.net/stats/br/ (Clèries et al. 2012). BootstRatio works through resampling methods without any assumption on the underlying probability distribution for the data analyzed.

Unless specified otherwise, Two-tailed t test was used to assess statistical differences in *ABCA4* % of transcript and gDNA cut efficiencies on GraphPad.

6 RESULTS

6.1 MINIGENE ESTABLISHMENT FOR PRELIMINARY VALIDATION OF SPLICING PATTERNS

To study the effect and potency of editing approaches to rescue aberrant splicing induced by either the isolated c.5197-557G>T DIV or the clustered DIVs in intron 36, a minigene construct encompassing *ABCA4* exon 36 and 37, referred to as WT minigene (wild-type), was generated. The isolated or clustered DIVs were individually introduced in the WT minigene by *in vitro* mutagenesis, thereby resulting in eight mutated minigenes (c.5197-557G>T, c.5196+1013G>A, c.5196+1056G>A, c.5196+1134C>G, c.5196+1136C>A, c.5196+1137G>A, c.5196+1159G>A and c.5196+1216C>A). For validation of the minigenes, HEK293T cells were transfected with either the mutant or wild-type minigene. Total RNA was extracted after 24h and the aberrantly and/or correctly-spliced transcripts were assessed by RT-PCR and agarose gel electrophoresis (**Figure 13**). Aberrant RT-PCR fragments were further characterized by sequencing analysis.

Splicing of the minigene plasmids harboring the c.5196+1136C>A or c.5196+1159G>A DIVs did not give rise to aberrant transcripts, as previously also reported in literature (Khan et al. 2020a). Since splicing can be regulated by tissue-specific RNA-binding factors, we additionally attempted the characterization of splicing for the c.5196+1136C>A or c.5196+1159G>A minigene plasmids in a retinoblastoma cell line (WERI-Rb1) and a retina-like cell line (CPCs), produced through differentiation of iPSCs. The rationale behind this extended experiment was that possibly cell-type specific splicing might influence the correct or aberrant splicing (Grosso et al. 2008). In both cases, the DIVs did not induce evident missplicing also in the two additional cell lines tested (**Figure 14**). To finally solve the answer to whether these two variants are true splicing mutations, iPSC lines and thereof derived photoreceptor precursor cells need to be generated and differentiated endogenously expressing (mutant) *ABCA4*.
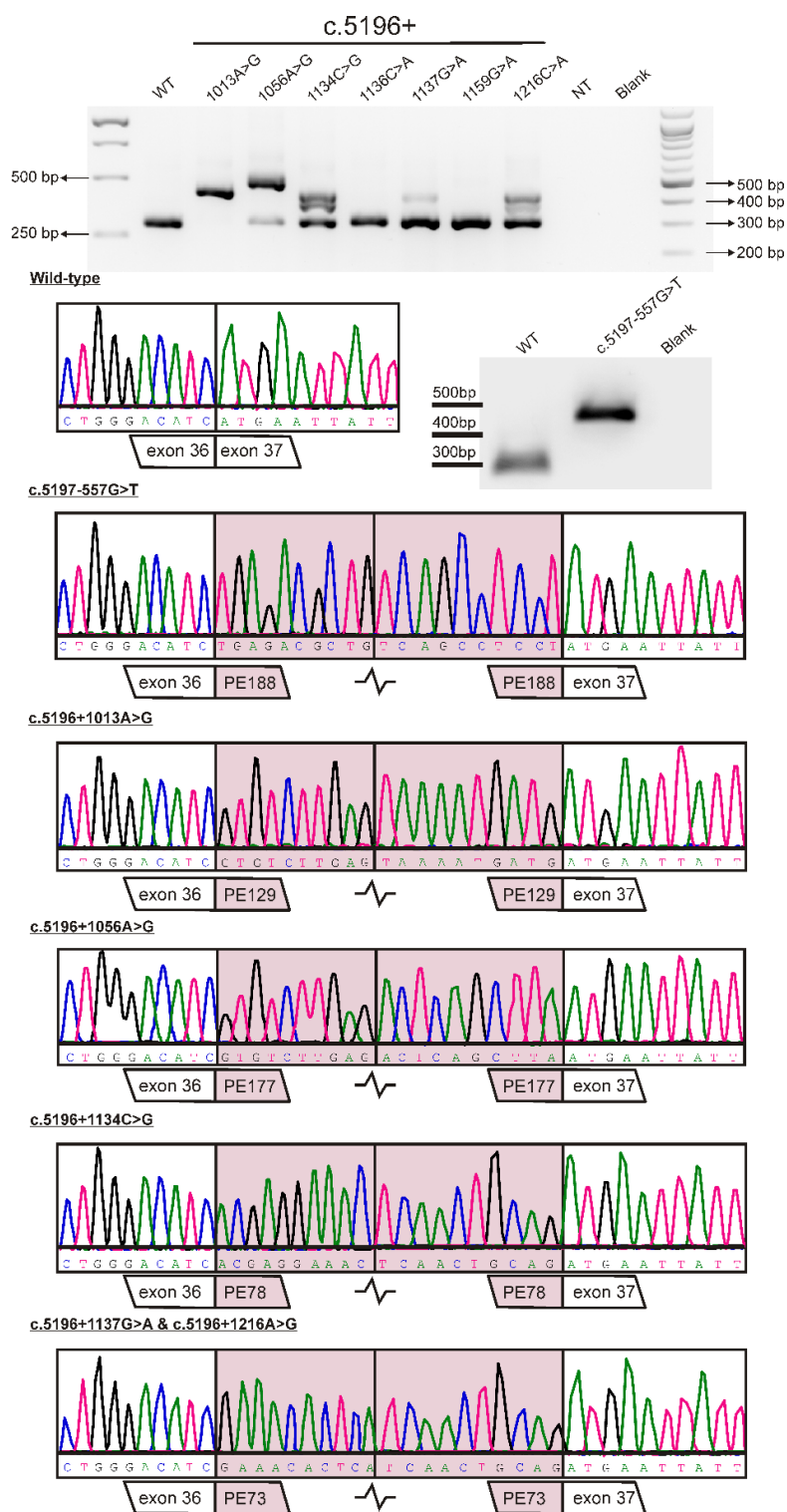
Figure 13: Characterization of the splicing patterns. Aberrant splicing patterns induced by the isolated and clustered deep-intronic variants in intron 36 of *ABCA4* as assessed by cDNA analysis out of minigene assays and visualized by agarose gel electrophoresis (top), and Sanger sequencing electropherograms of misspliced products in comparison to the wild-type cDNA sequence. The splicing patterns and sequences of the retained pseudoexons (PE) are consistent with literature (Bauwens et al. 2019, Braun et al. 2013, Khan et al. 2020a). The minigenes harboring the c.5196+1136C>A or c.5196+1159G>A DIVs do not induce missplicing and retention of a PE sequences as noticed on the agarose gel picture. WT= wild-type. NT=non-transfected. Blank =negative PCR control without template. ***The top agarose gel picture is taken from the Master thesis of Eleonora Roschi's (see contribution section)***
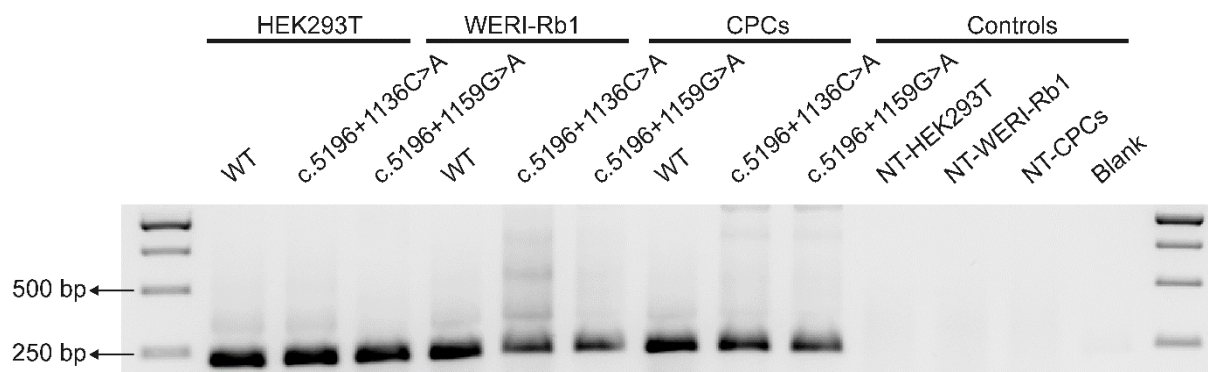
107

Figure 14: Transfection and cDNA analysis of wild-type (WT), c.5196+1136C>A and c.5196+1159G>A minigene plasmids in HEK293T, WERI-Rb1 and CPCs. Splicing analysis was performed 48 h post transfection. Minigene-specific primers were used to selectively amplify the ABCA4 transcript derivative of the plasmid to characterize any missplicing event. Non-transfected (NT) cells were used as controls. Blank =negative PCR control without template.

For the remaining minigene plasmids (c.5197-557G>T, c.5196+1013G>A, c.5196+1056G>A, c.5196+1134C>G, c.5196+1137G>A and c.5196+1216C>A), splicing patterns and identified transcripts were consistent with published data (**Figure 13**). The splicing patterns due to c.5196+1013A>G, c.5196+1056A>G, and c.5197-557G>T DIVs resulted in most of the transcripts being misspliced. Conversely, the c.5196+1134C>G, c.5196+1137G>A, and c.5196+1216C>A-induced splicing patterns was only partially exhibiting aberrant products.

6.2 ESTABLISHING A TRANSFECTION PROTOCOL FOR CONE PHOTORECEPTOR PRECURSOR CELLS.

In order to further validate promising editing approaches in a cellular model that would allow for the expression of *ABCA4* endogenously, culturing of iPSCs and their subsequent differentiation into cone photoreceptor precursors cells was implemented in the lab within the scope of this thesis. Furthermore, although patient-derived CPCs have been widely used to assess the efficacy of AON-mediated splicing modulation, to date, no protocol for successful transfection or transduction of plasmids in CPCs has been described (Albert et al. 2018, Dulla et al. 2021, Garanto et al. 2019, Khan et al.2020a). To deliver the genome editing plasmids in CPCs, we have evaluated a number of commercial transfection reagents (e.g. Lipofectamine™ 3000 and Lipofectamine™ Stem Transfection Reagent) as well as lentiviral particles for their efficiency to deliver the editing plasmids into CPCs. Virtually no transfected cells (as assessed by fluorescence microscopy) were seen using lipofection reagents and only a low number of transduced cells was observed by transduction with a custom-made lentiviral vector batch (**Figure 15**).
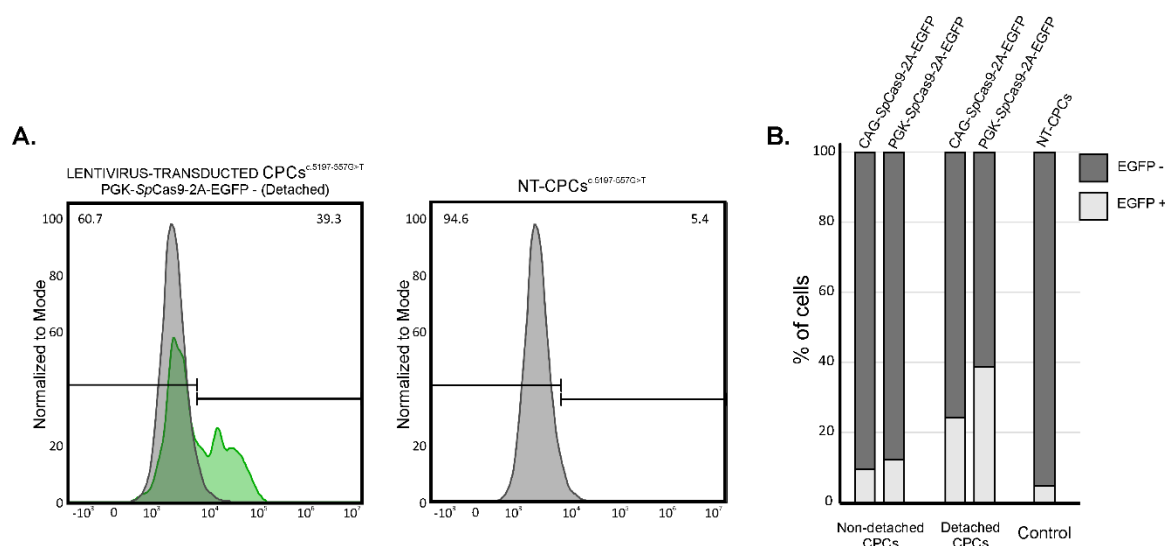
Figure 15: Efficacy of lentiviral delivery in cone-photoreceptor precursor cells (CPCs). (**A**) Flow cytometry analysis of patient-derived, heterozygous c.5197-557G/T cone photoreceptor precursor cells (CPCs[c.5197-557G/T]) transduced with lentiviral particles encoding for PGK-*Sp*Cas9-2A-EGFP. CPCs[c.5197-557G/T] were detached prior to the addition of the lentivirus particles. The analysis was performed five days after transduction. Non-Transduced (NT) CPCs[c.5197-557G/T] were used to set the threshold of background fluorescence. In the top corner of each histogram (normalized to mode), the percentage of transfected cells (left side) and non-transfected cells (right side) in relation to the fluorescence intensity is shown. (**B**) Bar graph showing the percentage of EGFP positive (EGFP +) and negative (EGFP -) of CPCs[c.5197-557G/T] upon transduction with two different lentivirus preparations: PGK-*Sp*Cas9-2A-EGFP and CAG-*Sp*Cas9-2A-EGFP, by detaching (detached CPCs) or non-detaching (Non-detached CPCs) the cells prior to the addition of the lentivirus particles (n=1).

We therefore established a reliable highly efficient electroporation protocol for the transfection of 30-day differentiated CPCs (**Figure 16A**). Reproducible and consistent high transfection efficiencies of CPC[c.5197-557G/T] were achieved (average transfection efficiency: 79±10% transfected cells in 27 independent electroporation rounds, **Figure 16B**), and low levels of cell death were observed 24 h post electroporation by comparing the number of electroporated cells to the number of attached cells 24h post transfection. Specifically, an average of 80±8% re-attached cells were counted 24h post transfection (n=3). This protocol was selected to deliver the genome editing plasmids into any CPCs throughout this thesis.
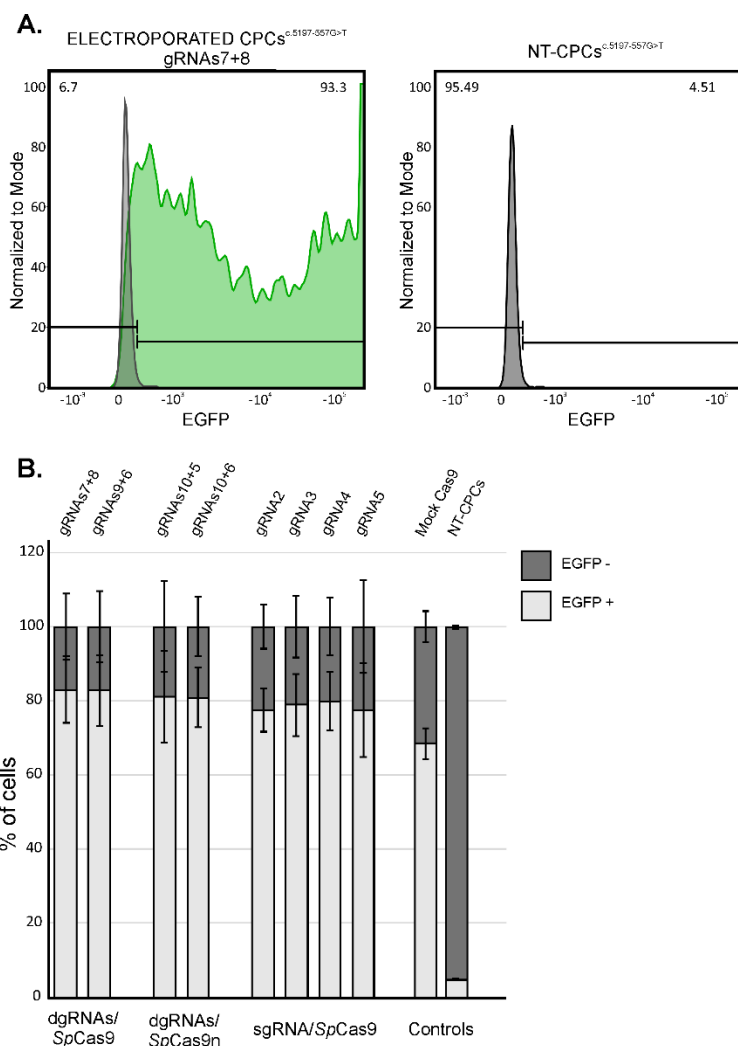
Figure 16: Flow cytometry analysis of electroporated patient-derived, heterozygous c.5197-557G/T cone photoreceptor precursor cells (CPCs$^{c.5197-557G/T}$) (**A**) Flow cytometry analysis of CPCs$^{c.5197-557G/T}$ electroporated with a plasmid encoding for CAG-*Sp*Cas9-2A-EGFP and gRNAs7+8. The analysis was performed two days after transfection. Non-Transduced (NT) CPCs$^{c.5197-557G/T}$ were used to set the threshold of background fluorescence. In the top corner of each histogram (normalized to mode), the percentage of transfected cells (left side) and non-transfected cells (right side) in relation to the fluorescence intensity is shown. (**B**) Bar graph showing the transfection efficiencies achieved on CPCs$^{c.5197-557G/T}$ upon electroporation of the editing plasmids. Results are presented as mean ± standard deviation (n= 3 independent electroporations).

6.3 Establishment of induced pluripotent stem cell lines for the clustered deep-intronic variants

To generate appropriate cellular models for the future validation of the gRNA/ED*Sp*Cas9-mediated splicing rescue results obtained through minigene splicing assay of the clustered DIVs, fibroblast cell lines were established from patients with suitable genotypes who were willing to donate their cells for

subsequent reprogramming into iPSCs and further differentiation into CPCs. In particular, we successfully established a heterozygous c.5196+1056A>G and a heterozygous c.5196+1134C>G patient-derived fibroblast cell lines (both compound-heterozygous for another *ABCA4* pathogenic variant). A heterozygous iPSC line for the c.5196+1137G>A DIV was obtained from collaborators (Khan et al. 2020a). Genotyped patients in our cohort for the c.5196+1013A>G and c.5196+1216C>A DIVs were not available. Therefore, by means of CRISPR/Cas-mediated knock-in, the establishment of homozygous isogenic iPSC cell lines was planned for these two DIVs.

6.3.1 Reprogramming of a c.5196+1056A>G and a c.5196+1134C>G fibroblast cell lines

Transfection of the heterozygous c.5196+1134C>G fibroblast cell line with plasmids encoding for the reprogramming factors *L-MYC*, *LIN28*, *SOX2*, *POU5F1* (*OCT4*), and *KLF4* (Okita et al. 2011) resulted in the generation of a heterozygous c.5194+1134C>G iPSC line (iPSCs $^{c.5196+1134C/G}$). The established iPSCs$^{c.5196+1134C/G}$ showed high expression levels of the selected pluripotency markers, indicating successful reprogramming of the parental fibroblast cell line (**Figure 17A**). Spontaneous integration of the plasmids used for reprogramming, was excluded by RT-PCR profiling for reprogramming factors derived from the plasmids (**Figure 17B**). As for the c.5196+1056A>G fibroblast cell line, two subsequent reprogramming experiments failed, therefore the variant was introduced by knock-in into an iPSC control line (see Results 6.3.2).
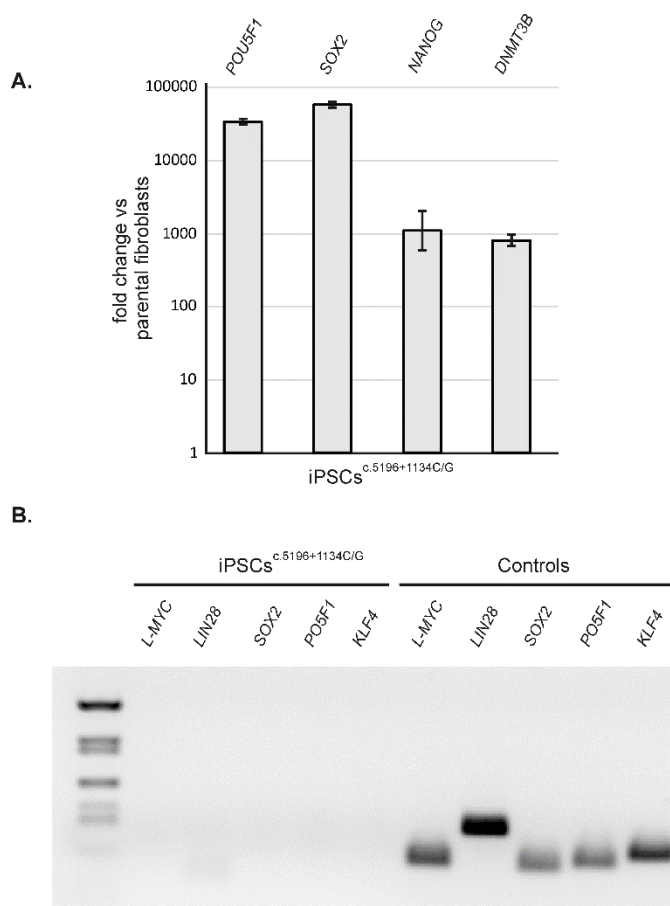
Figure 17: Validation of the established c.5196+1134C/G induced pluripotent stem cell line (iPSCs ^{c.5196+1134C/G}). (**A**) Quantitative PCR analysis of pluripotency markers POU5F1 (*OCT4*), *SOX2*, *NANOG* and *DNMT3B*. *GAPDH* was used as housekeeping gene. Results are presented as mean ± standard deviation of three technical replicates. (**B**) RT-PCR of the reprogramming factors to exclude the spontaneous integration of the reprogramming plasmids in the genome of the established iPSC ^{c.5196+1134C/G}. Plasmid-specific primers were used to amplified the cDNA of the established iPSC ^{c.5196+1134C/G} to detect potential expression of reprogramming factors due to possible integration of the reprogramming plasmids into the genome. No expression could be seen. As controls, the reprogramming plasmids were amplified alongside the cDNA of iPSCs ^{c.5196+1134C/G}.

6.3.2 CRISPR/Cas-mediated knock-in of c.5196+1013A>G, c.5196+1056A>G, and c.5196+1216C>A

An iPS cell line derived from a healthy individual was CRISPR/Cas-edited to generate homozygous c.5196+1013A>G (iPSCs^{c.5196+1013G}), c.5196+1056A>G (iPSCs^{c.5196+1056G}) (subsequently to the reprogramming failures from patient-derived fibroblasts), and c.5196+1216C>A (iPSCs^{c.5196+1216A}) iPS cell lines. Ribonucleoproteins (RNPs), a synthetic gRNA, and a donor, symmetrical single-stranded oligodeoxynucleotides (ssODNs) with 50 bp-symmetrical homology arms were employed to obtain the knock-in of the desired variants. Specifically, to generate iPSCs^{c.5196+1013G} and iPSCs^{c.5196+1216A} *Sp*Cas9 was used, whereas iPSCs^{c.5196+1056G} was established by employing *Acidaminococcus sp.* Cas12a (*As*Cas12a or

*As*Cpf1) due to incompatible PAM profile for *Sp*Cas9. Cas12a recognizes TTTV PAM sequences which matched for variant c.5196+1056G (**Figure 18**).
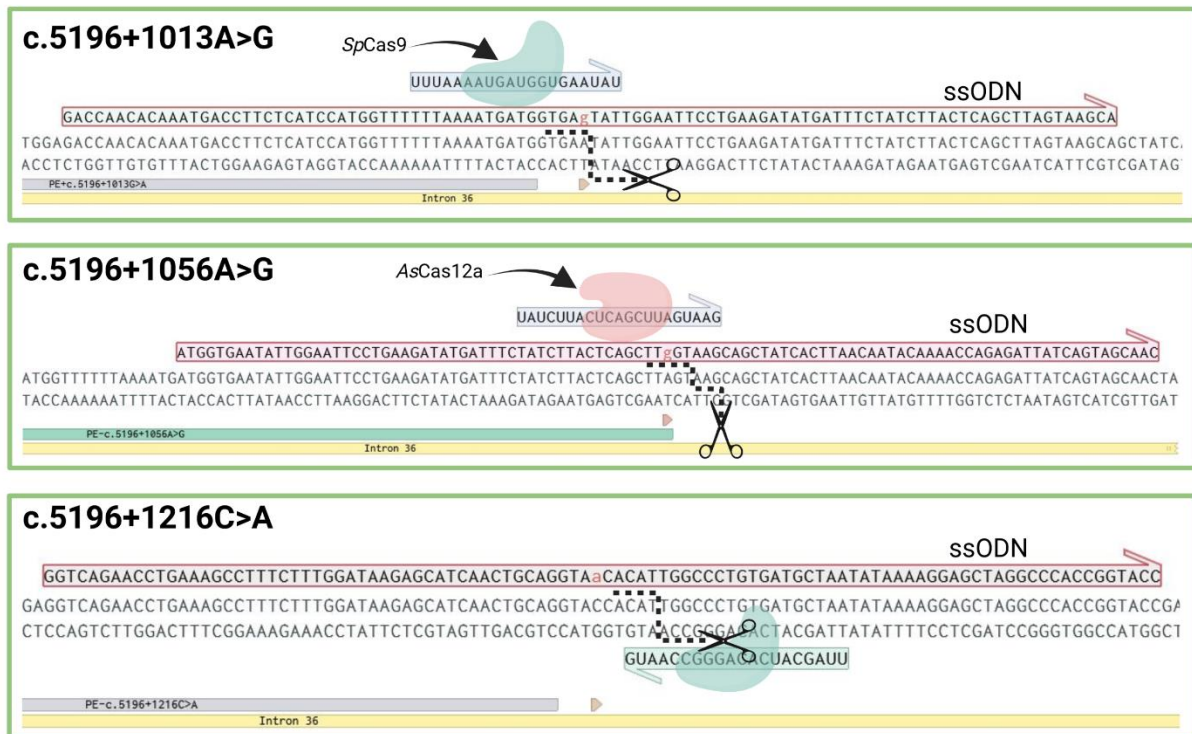

Figure 18: Schematic representation of the CRISPR/Cas-mediated knock-in experiments. The dashed line depicts the genomic cleavage mediated by the implemented genome editing molecules, either *Streptococcus pyogenesis* Cas9 (*Sp*Cas9) or *Acidaminococcus sp.* Cas12a (*As*Cas12a). The triangle annotation indicates the position of the deep-intronic variant to be inserted. ssODN (single-strand oligonucleotides) depicts the location and sequence of the donor template used for knocking-in the specific variant. PE= Pseudoexon.

Forty-eight iPSC clones each for c.5196+1013A>G, c.5196+1056A>G as well as c.5196+1216C>A were sequenced. Successful homozygous knock-in was obtained in one, four, and three clones for c.5196+1013A>G, c.5196+1056A>G, and c.5196+1216C>A, respectively (**Figure 19**).
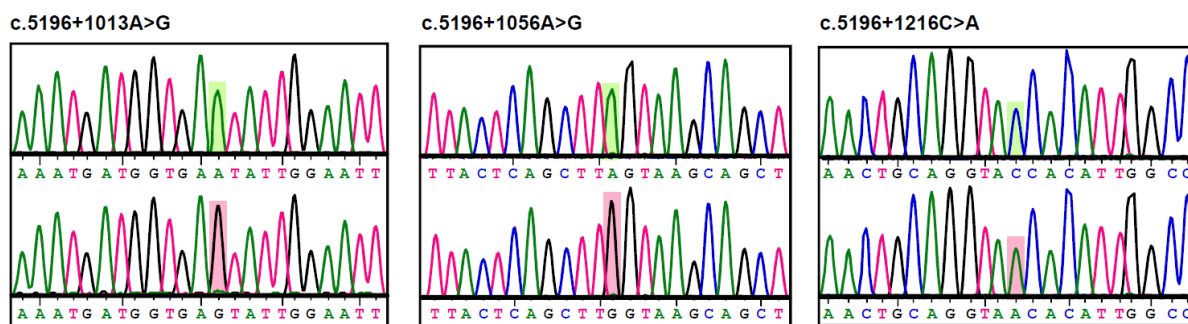
Figure 19: Representative sequencing results of the knocked-in iPS clones for c.5196+1013A>G, c.5196+1056A>G and c.5196+1216C>A. Electropherograms derived from one clone of each iPS line is depicted. The wild-type sequence is provided as a reference. The green and red boxes highlight the location on the wild-type (top part) and edited nucleotide (bottom part), respectively.

6.4 DESIGN OF *Sp*CAS9-BASED STRATEGIES TO RESCUE DEEP-INTRONIC VARIANTS IN INTRON 36 OF *ABCA4*

In order to attempt the rescue of aberrant splicing patterns induced by *ABCA4* DIVs in intron 36, three standard genome editing approaches employing *Streptococcus pyogenes* Cas9 (*Sp*Cas9) were established. Although such approaches have already been described in literature and therefore are regarded as standard (Maule et al. 2019, Sanz et al. 2017, Shen et al. 2014), they have not been applied to the rescue of splicing defects in *ABCA4*. In this thesis, they have extensively and comparatively been assessed for the isolated DIV c.5197-557G>T.

Furthermore, within this thesis, a fourth innovative approach that utilizes a *Sp*Cas9 fusion protein was designed and tested on the isolated c.5197-557G>T DIV in intron 36 of *ABCA4*. In addition, this novel concept was then used to preliminarily assess the rescue of the so-called clustered DIVs (c.5196+1013G>A, c.5196+1056G>A, c.5196+1134C>G, c.5196+1137G>A, and c.5196+1216C>A).

6.4.1 "Standard" dual gRNAs/SpCas9, dual gRNAs/SpCas9 nickase and single sgRNA/SpCas9 approaches to target the isolated DIV c.5197-557G>T

The three standard *Sp*Cas9-based approaches were implemented with the aim to rescue the isolated DIV c.5197-557G>T. *Sp*Cas9 was used to design dual gRNAs/*Sp*Cas9 (**dgRNAs/*Sp*Cas9**), dual

gRNAs/*Sp*Cas9 nickase (**dgRNAs/*Sp*Cas9n**) and single sgRNA/*Sp*Cas9 (**sgRNA/*Sp*Cas9**) approaches (**Figure 20**).
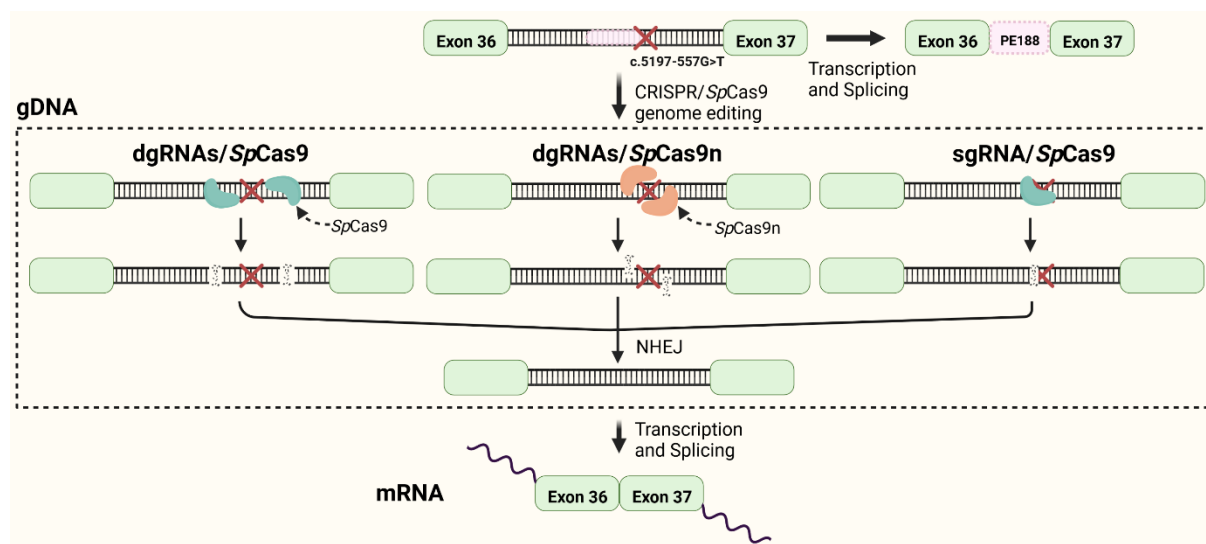

Figure 20: Illustration of the three "standard" CRISPR/SpCas9-based approaches established to target the isolated c.5197-557G>T deep-intronic variant (DIV) localized in intron 36 of *ABCA4*. Upon transcription and splicing, the *ABCA4* c.5197-557G>T DIV induces the retention of a 188 intronic sequence (pseudoexon – (PE)) in the mature mRNA transcript of *ABCA4*. By implementing CRISPR/*Sp*Cas9-mediated genome editing, the splicing is corrected back to normal: dgRNAs/*Sp*Cas9 inducing double strand breaks (DSB) up- and downstream of the PE for the generation of a large deletion of intronic sequences encompassing the entire PE, dgRNAs/*Sp*Cas9n inducing single strand breaks induced by a nickase in close vicinity on opposite strands for the generation of small deletions encompassing parts of the PE, and sgRNA/*Sp*Cas9 inducing a DSB and NHEJ-mediated Indels in the vicinity of the DIV.

dgRNAs/*Sp*Cas9 utilizes two paired sgRNAs targeted to induce two DSBs, resulting in the deletion of a genomic sequence encompassing the isolated DIV. The deleted genomic fragments had an approximate size of 600 nt. The two generated DSBs of the dgRNAs/*Sp*Cas9 approach are repaired by the NHEJ pathway, thereby leading to a large genomic deletion.

The sgRNA/*Sp*Cas9 approach relies on the use of a sgRNA targeted to the genomic sequence in proximity to the isolated DIV. With the aim to induce sequence perturbation of the cryptic donor splice site strengthened by the isolated DIV, sgRNA/*Sp*Cas9 intends to induce a single DSB which is repaired by NHEJ, resulting in small deletions or insertions at the targeted site.

To establish the third approach, the nickase version of *Sp*Cas9, specifically *Sp*Cas9 nickase D10A was applied to the design of the dual gRNAs/*Sp*Cas9 nickase (dgRNAs/*Sp*Cas9n) approach. In this design, the endonuclease activity of two paired sgRNAs coupled to *Sp*Cas9n is directed towards a small (20-80 nt) genomic stretch that also encompasses the isolated DIV. The two sgRNAs target opposite strands,

thereby dgRNAs/*Sp*Cas9n determines the formation of opposite single-strand breaks (nicks), ultimately leading to the generation of a small genomic deletion that contains the isolated DIV and part of the 3' end of the pseudoexon sequence. To generate effective sgRNA combinations for the dgRNAs/*Sp*Cas9n approach, the sgRNAs have to be offset, the PAM sequences have to face outward (PAM-out configuration) and the distance between the nicks has to range from 40 nt to 70 nt apart (Ran et al. 2013).

6.4.2 Development of a novel Enhanced-Deletion *Sp*Cas9 variant and establishment of a novel single gRNA/Enhanced-Deletion *Sp*Cas9 approach to target the isolated DIV c.5197-557G>T and clustered deep-intronic variants in intron 36

To generate a fourth approach that empowered the use of a (single) sgRNA to determine consistent and effective splicing rescue of both isolated and clustered DIVs, a *Sp*Cas9 fusion protein was cloned, improved and validated. The gRNA/Enhanced-Deletion *Sp*Cas9 (**sgRNA/ED*Sp*Ca9**) design includes the human 3' – 5' Three Prime Repair Exonuclease 2 (TREX2) fused to *Sp*Cas9 and interspaced by a "GGGGS" linker. The initial construct was obtained from Dr. Andrew Basset, Sanger Institute, UK (Allen et al. 2019). Upon generation of a single DSB mediated by *Sp*Cas9, the fused TREX2 protein partner enables further processing of the produced break ends, thereby resulting in enhanced deletions at the targeted site (Allen et al. 2019). For this reason, the approach is regarded as "Enhanced-Deletion". sgRNA/ED*Sp*Cas9 can be directed to generate deletions at a cryptic splice site containing a DIV or to any sequence feature that is in common to more clustered DIVs (e.g. donor and acceptor splice sites, splice enhancers and silencers). By doing so, the ED*Sp*Cas9 system determines profound sequence disruption of the element acting on splicing, thereby preventing the spliceosome complex from recognizing it, which ultimately leads to correction of the normal splicing pattern (**Figure 21**).
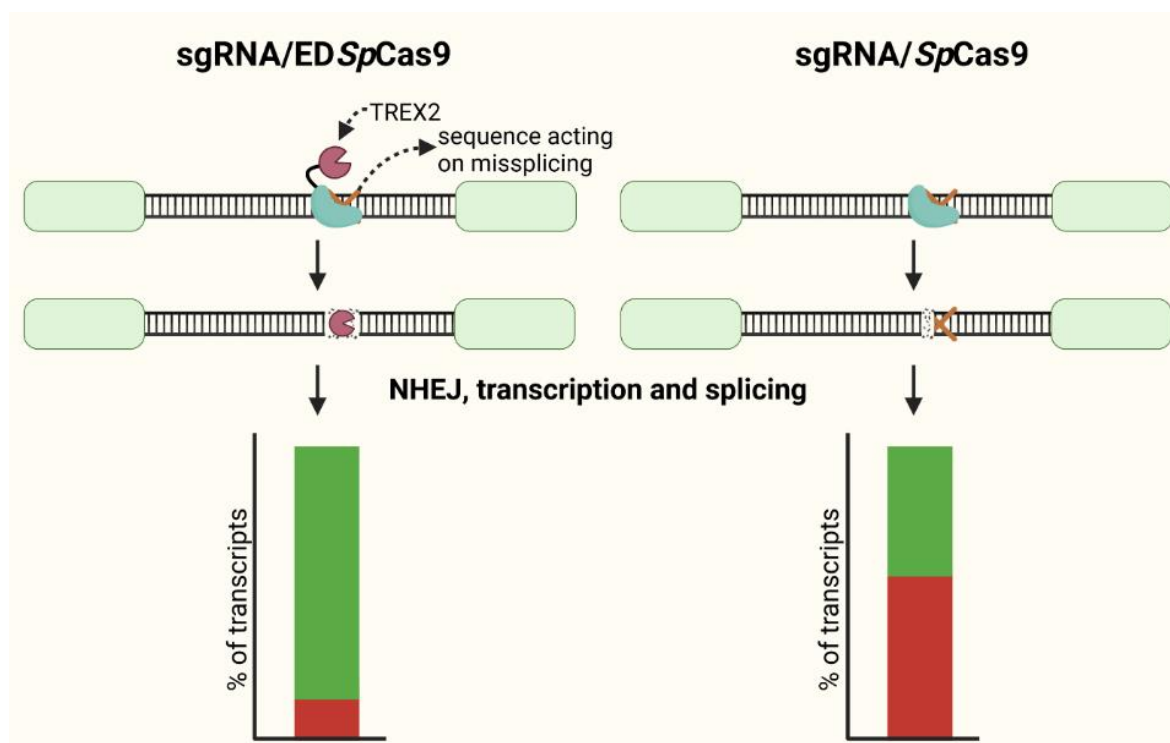
Figure 21: Illustration of the mechanism of action of the novel sgRNA/Enhanced-Deletion *Sp*Cas9 (sgRNA/ED*Sp*Cas9) approach in comparison to the standard sgRNA/*Sp*Cas9 approach. When targeting ED*Sp*Cas9 to a sequence element affecting splicing (e.g. deep-intronic variants or cryptic splice sites) by a single gRNA (sgRNA), the extensive sequence perturbation mediated by the fusion partner TREX2 effectively destroys the identity of the targeted sequence (i.e. with a tendency to generate deletions rather than indels), thereby preventing the spliceosome complex from recognizing and ultimately missplicing the mRNA. Conversely, the sequence perturbation mediated by *Sp*Cas9 only relies on indels (insertion and/or deletions) formation mediated by the cellular repair mechanisms, resulting in minor alteration of the targeted sequences, ultimately resulting in lower splicing rescue.

Several configurations of *Sp*Cas9 and TREX2 were cloned and tested during the optimization steps (**Figure 22**).
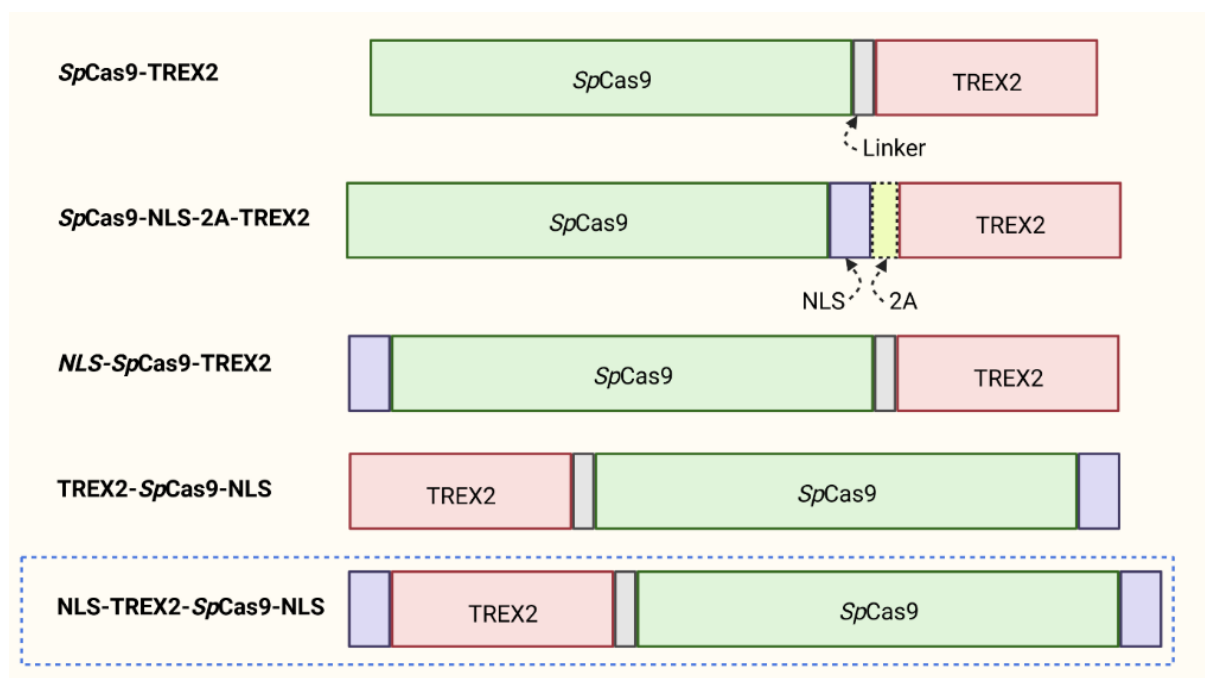
Figure 22: Schematic representation of the different configurations of ED*Sp*Cas9 cloned during the optimization experiments. The linker (GGGGS) allows structural flexibility between the two proteins. The 2A sequence induces ribosomal skipping during translation, thereby resulting in the synthesis of *Sp*Cas9-NLS and TREX2. The Nuclear Localization Signal (NLS) sequence is added to facilitate the translocation of ED*Sp*Cas9 into the nucleus. NLS-TREX2-*Sp*Cas9-NLS represents the finally selected configuration.

The optimization steps that led to the final nuclear localization signal (NLS)-TREX2-*Sp*Cas9-NLS design began with the cloning of *Sp*Cas9-TREX2 and *Sp*Cas9-NLS-2A-TREX2. Although the chimeric *Sp*Cas9-TREX2 protein and the co-expression of *Sp*Cas9-NLS and TREX2 (*Sp*Cas9-NLS-2A-TREX2) were inducing enhanced splicing rescue for different gRNA combinations in the minigene assay experiments targeting the isolated c.5197-557G>T DIV (**Figure 23A**), *Sp*Cas9-TREX2 failed to achieve similar genomic DNA (gDNA) cleavage efficacy in patient-derived fibroblasts compared to NLS-*Sp*Cas9-NLS alone (**Figure 23B**).
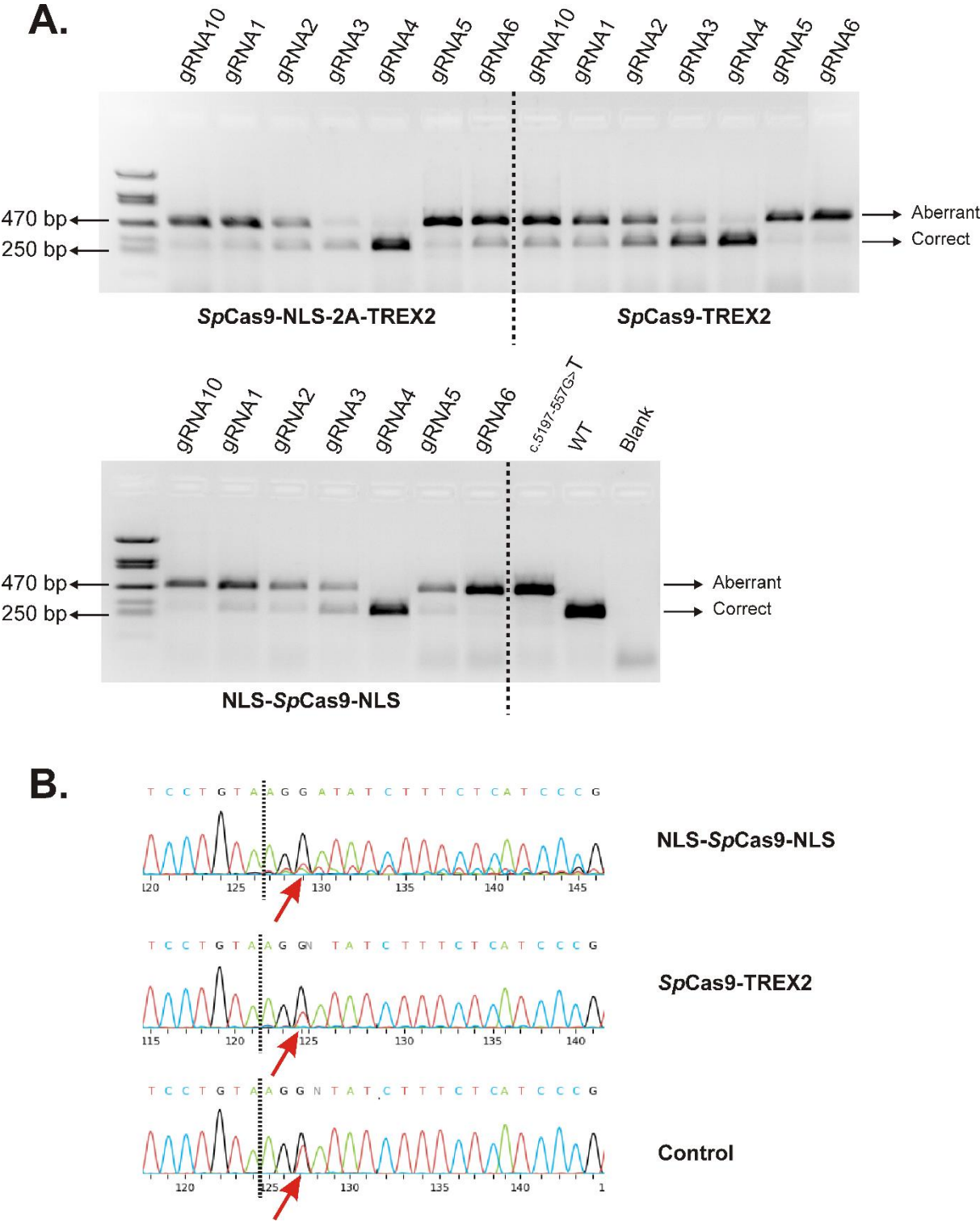
Figure 23: Part of the results produced during the sgRNA/ED*Sp*Cas9 optimization (**A**) Gel picture showing the minigene rescue results of the aberrant splicing induced by the isolated c.51197-557G>T DIV when implementing *Sp*Cas9-NLS-2A-TREX2 and *Sp*Cas9-TREX2 in comparision to NLS-*Sp*Cas9-NLS. WT= wild-type. Blank = negative PCR control without template. (**B**) Representative electropherogram traces upon editing with gRNA4 in heterozygous c.5197-557G/T fibroblasts. The height of the signal of the c.5197-557T nucleotide (red arrow) and the additional traces obtained at the cut site are qualitative indicators of the level of editing that is achieved. The dashed lane indicates the predicted cut site of gRNA4.

As for the co-expression of *Sp*Cas9 and TREX2, the overexpression of an exonuclease (TREX2) as an independent protein might induce DNA damage to the cells, therefore this design was not selected for further experiments. We speculated that the lack of any nuclear localization signal NLS would affect the importing of *Sp*Cas9-TREX2 in the nucleus, thereby the gDNA cut efficiency. To address this point, a NLS signal was added to *Sp*Cas9-TREX2 and TREX2-*Sp*Cas9 designs, resulting in NLS-*Sp*Cas9-TREX2 and TREX2-*Sp*Cas9-NLS, respectively. Also these designs failed to achieve gDNA editing efficacy comparable to NLS-*Sp*Cas9-NLS. Following these results and the fact that the TREX2-*Sp*Cas9-NLS design gave higher rescue in preliminary minigene assay results, a second NLS was added at the N-terminal end of TREX2. The resulting NLS-TREX2-*Sp*Cas9-NLS is composed of two nuclear NLSs that ease the translocation of TREX2-*Sp*Cas9 to the nucleus (Boulikas, 1993), and it was used for the experiments described thereafter. As a reference approach to demonstrate sgRNA/ED*Sp*Cas9 superiority, sgRNA/ED*Sp*Cas9 experiments were performed in parallel to sgRNA/*Sp*Cas9 coupled to the same sgRNAs.

The possible improvement derived from the use of different exonucleases, beside TREX2, was also investigated. Specifically, T5 exonuclease, T7 exonuclease, Exonuclease VI, phi29 (exonuclease domain), WRN, TREX1, RAD1, and RAD9 were individually fused to the amino-terminus of *Sp*Cas9 in the NLS-(EXONUCLEASE)-*Sp*Cas9-NLS configuration and tested for rescueing the missplicing induced by the isolated c.5197-557G>T DIV. Following the quantification of the correctly-spliced band, we concluded that no other exonuclease partner performed as well as TREX2 in the tested experimental conditions, therefore no further optimization step, in this respect, was taken (**Figure 24**).
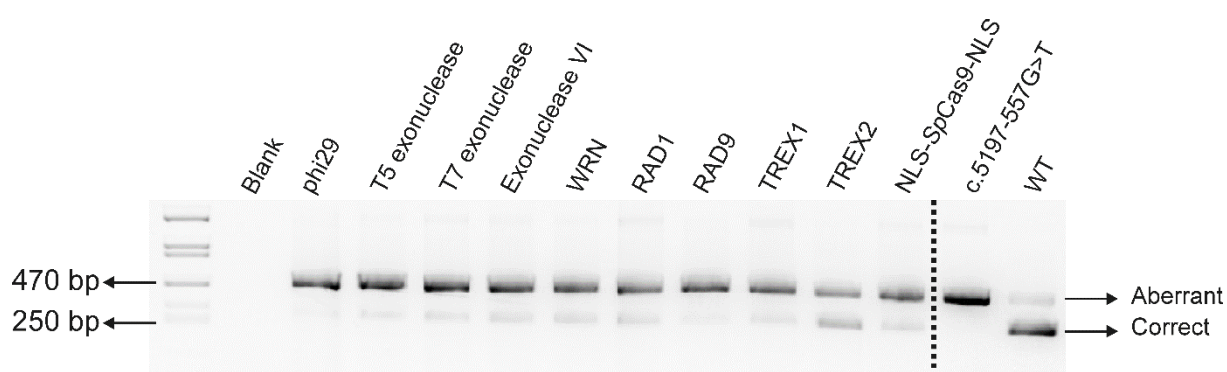
Figure 24: Agarose gel showing the preliminary results of the different exonuclease partners tested for the rescue of the c.5197-557G>T-induced missplicing in a minigene assay*. The configuration of the tested exonuclease was NLS-(EXONUCLEASE)-*Sp*Cas9-NLS. The larger and smaller bands on the agarose gel correspond to the aberrant and correct transcripts, respectively. WT= Wild-type. Blank = Negative PCR control without template. ***The agarose gel is taken from Evangelos Bouris' lab rotation report (see contributions section)**

6.4.3 gRNA design, preliminary validation, and selection for targeting *ABCA4* deep-intronic variant c.5197-557G>T

A total of fourteen sgRNAs directed to rescue c.5197-557G>T-induced missplicing were *in silico* identified using Benchling (www.benchling.com) and selected for preliminary assessment of the gDNA cut efficiency. Eight of them (gRNA1, gRNA7, gRNA9, gRNA10, gRNA11, gRNA12, gRNA13, gRNA14) were located upstream, three (gRNA5, gRNA6, gRNA8) downstream the isolated DIV c.5197-557G>T, and three overlapping with the isolated DIV (gRNA2, gRNA3, gRNA4) (**Figure 25**).
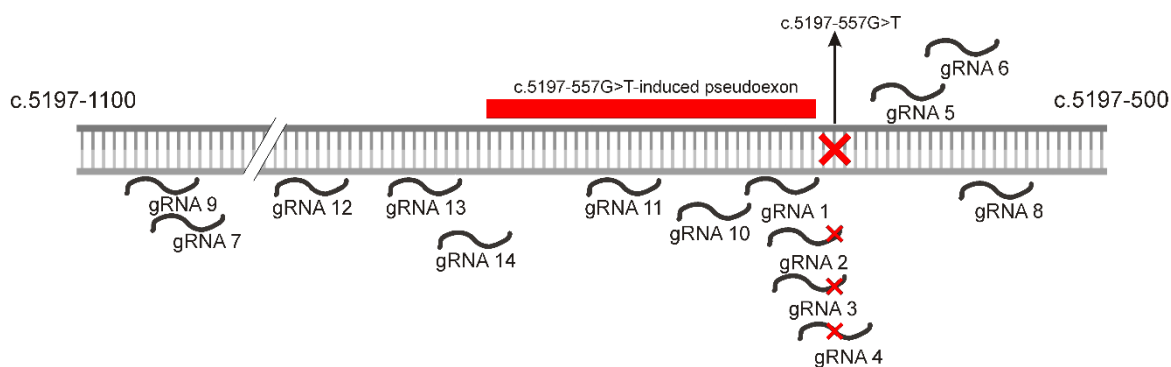
Figure 25: Schematic illustration of the location of the single gRNAs designed to address c.5197-557G>T-induced aberrant splicing.

The gRNAs were individually screened for their ability to generate DSBs when coupled to standard *Sp*Cas9. The gDNA cut efficiency of the sgRNAs located up- and downstream the DIV was assessed in HEK293T cells, whereas for the sgRNAs overlapping with the DIVs, gDNA cut efficiency was tested in patient-derived fibroblasts due to the necessity of a suitable genetic background. The quantification of the gDNA cut efficiency was carried out through sequencing-based validation (TIDE or ICE), a representative alligment of edited samples versus a control sample (non-treated) is presented in **Figure 23B**. A summary of achieved efficiencies is reported in **Table 3**.

Table 3: Preliminary validation of the genomic DNA cut efficiency of the sgRNA used in this study. The gRNA sequence, location compared to c.5197-557G>T (location compared to the DIV), the cell line used for the preliminary validation and the genomic DNA (gDNA) cut efficiency are reported. Data of the gDNA cut efficiency are reported as mean ± standard deviation. The number of replicates performed are indicated. The nucleotide overlapping the DIV in the gRNA sequence is highlighted in bold.

gRNA name	Sequence 5' – 3'	gDNA cut efficiency (%)	Location compared to DIV	Cell line used
gRNA1	ACAGGAGGCTGATCTGGTGC	25.5 ± 10.5 (n=2)	Upstream	HEK293T
gRNA2	AT**A**CTTACAGGAGGCTGATC	14.0 (n=1)	Overlapping	Het. patient-derived fibroblasts
gRNA3	GAGAAAGATAT**A**CTTACAGG	25.3 (n=1)	Overlapping	Het. patient-derived fibroblasts
gRNA4	GATGAGAAAGATAT**A**CTTAC	21.3 (n=1)	Overlapping	Het. patient-derived fibroblasts
gRNA5	ATATCTTTCTCATCCCGTTG	8.5 (n=1)	Downstream	HEK293T
gRNA6	GTTGAGGCCTCCACACCCCA	21.1 ± 3.8 (n=3)	Downstream	HEK293T
gRNA7	GCTGTGCCATTGAGCATGG	36.9 ± 2.9 (n=2)	Upstream	HEK293T
gRNA8	TCAGGGGGCAACCTGCCGTG	19.3 ± 3.7 (n=3)	Downstream	HEK293T
gRNA9	GAGCATGGAGGAGGTCCAG	33.0 ± 3.9 (n=2)	Upstream	HEK293T
gRNA10	CAGGAGGCTGATCTGGTGCA	27.6 ± 8.3 (n=2)	Upstream	HEK293T
gRNA11	ATCAGAGCCAGAGCAGAACG	15.5 ± 1.7 (n=2)	Upstream	HEK293T
gRNA12	AAAAACTAGAATGAGCCCTG	6.0 (n=1)	Upstream	HEK293T
gRNA13	GGAGAAGGGAGTTACAGATG	7.2 (n=1)	Upstream	HEK293T
gRNA14	CAGCGTCTCACTGTTGGAGA	8.5 (n=1)	Upstream	HEK293T

The gDNA cut efficiency achieved by the different sgRNA/*Sp*Cas9 combinations varied as calculated by Tracking Indels by Decomposition (TIDE) or Inference of CRISPR edits (ICE) (Brinkman et al. 2014, Conant et al. 2022) – both are quantification methods that compare electropherogram sequencing traces of sgRNA/*Sp*Cas9-treated samples to untreated samples - thereby estimating gDNA cut efficiencies. gRNA5, gRNA12, gRNA13, and gRNA14 achieved low (< 10%) gDNA cut efficiency, specifically, 8.5%, 6.0%, 7.2% and 8.5%, respectively. Mid gDNA cut efficiency (> 10% and < 25%) was achieved for gRNA2, gRNA6, and gRNA8 which reached 14.0%, 21.1 ± 3.8%, and 19.3 ± 3.7% cleavage efficacy, respectively. The highest gDNA cut efficiency (> 25%) was attained by gRNA1: 25.5 ± 10.5, gRNA3: 25.3%, gRNA7: 36.9 ± 2.9, gRNA9: 33.0 ± 3.9%, and gRNA10: 27.6 ± 8.3% (**Table 3**). Of note,

gRNA2, gRNA3, and gRNA4 overlap with the isolated c.5197-557G>T DIV, thereby they can only effectively target the heterozygous mutant allele in the patient-derived fibroblast cell line. Nonetheless, the obtained value of gDNA cut efficiency derives from the contribution of both alleles, resulting in the underestimation of the actual cleavage efficiency of sgRNAs overlapping with the targeted c.5197-557G>T allele in the heterozygous fibroblast cell line.

Considering the obtained single gDNA cut efficiencies and sgRNA location, eleven sgRNAs and sgRNA combinations were selected for the design of the four different editing approaches as illustrated in **Table 4**.

Table 4: Selected sgRNA and sgRNA combinations for the four different genome editing approaches

dgRNAs/*Sp*Cas9	dgRNAs/*Sp*Cas9n	sgRNA/*Sp*Cas9	sgRNA/ED*Sp*Cas9
gRNAs7+6	gRNAs10+5	gRNA1	gRNA1
gRNAs7+8	gRNAs4+5	gRNA2	gRNA2
gRNAs9+6	gRNAs10+6	gRNA3	gRNA3
gRNAs9+8	gRNAs4+6	gRNA4	gRNA4
	gRNAs11+5	gRNA5	gRNA5
	gRNAs11+6	gRNA6	gRNA6

6.4.4   single gRNA design and selection for the clustered deep-intron variants

With the aim to rescue splicing defects due to the clustered DIVs in intron 36 by the sgRNA/ED*Sp*Cas9 approach, sgRNAs were designed at sequence elements involved in inducing aberrant splicing. In particular, the retention of pseudoexons PE129 and PE177, due to c.5196+1013A>G and c.5196+1056A>G, respectively, was primarily targeted at the shared cryptic splice acceptor site by designing four sgRNAs (gRNAI, gRNAII, gRNAIII and gRNAXIV) (**Figure 26**, **Figure 13, Table 5**). Due to an annotation mistake, gRNAI, gRNAII, and gRNAIII, were eventually tested as reverse complement sgRNA sequences compared to the actual designed sgRNA sequence. Therefore, the PAM sequence of the tested gRNAI became NGT, for gRNAII became NTA, and for gRNAIII became NAG.
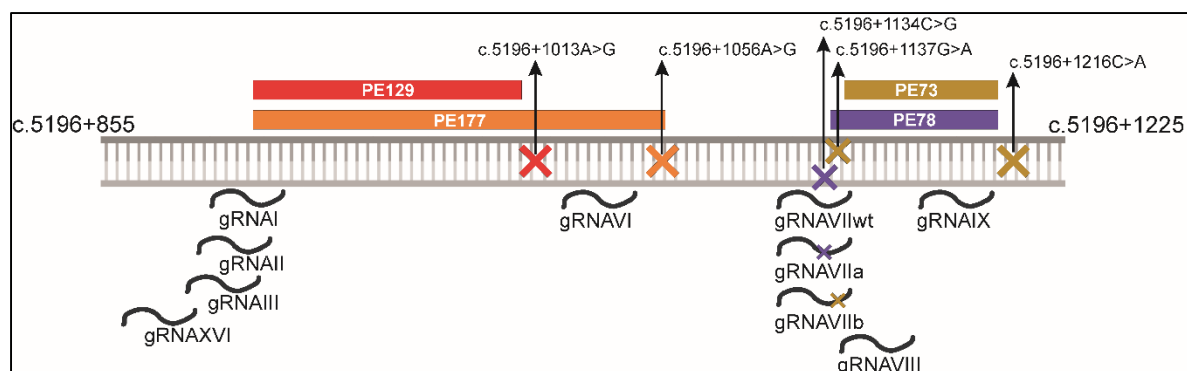
Figure 26: Schematic illustration of the location of the single gRNAs designed to address the aberrant splicing events induced by the clustered deep-intronic variants (DIV) in intron 36 of *ABCA4*. The color code of the pseudoexons (PE) matches the respective DIV(s)-inducing PE.

Aberrant splicing due to c.5196+1137G>A and c.5196+1216C>A shares the exact same retained intronic sequence (PE73), whereas c.5196+1134C>G determines the inclusion of an intronic fragment (PE78) that is 5 bp longer at the cryptic splice acceptor site (**Figure 13**). Of note, the DIV c.5196+1134C>G has been recently found in a patient within the Tübingen *ABCA4*-related IRD cohort and the resulting splicing defect was hereby characterized for the first time (**Figure 13**). To target such clustered DIVs, one variant-specific sgRNA per DIV (gRNAVIIa, gRNAVIIb, and gRNAVIIwt), located at the cryptic acceptor site, and two variant-independent sgRNAs (gRNAVIII and gRNAIX), located at the cryptic acceptor and donor splice sites, respectively, were designed (**Figure 26, Table 5**).

Table 5: sgRNA name and sequences of the sgRNA tested to target the clustered DIVs

sgRNA name	Tested Sequence 5' – 3'
gRNAI	AATTTATTCTAGGTGTCTTG
gRNAII	ATTTATTCTAGGTGTCTTGA
gRNAIII	CAATTTATTCTAGGTGTCTT
gRNAVIIa	CTCTCTCTTCTGTCTAGACG
gRNAVIIb	CTCTCTCTTCTGTCTACACA
gRNAVIIwt	CTCTCTCTTCTGTCTACACG
gRNAVIII	GGAAACACTCATAAATGCAC
gRNAIX	GGATAAGAGCATCAACTGC
gRNAXIV	CATCAACCCCAATTTATTCT

6.4.5   Bioinformatic assessment for single guide RNA off-target potential

Despite the predominant ability of CRISPR/*Sp*Cas9 to selectively recognize and cleave only targeted sequences, the editing system can still induce the formation of DSBs at undesired locations, known as off-target edits (Zhang et al. 2015). These locations can be experimentally identified using unbiased protocols such as GUIDE-seq or CIRCLE-seq (Tsai et al. 2015, Tsai et al. 2017). Alternatively, as a preliminary test, off-targets can be bioinformatically predicted through tools that search for sequence homology in a reference genome (Bae et al. 2014). For the sgRNAs selected for further validation, bioinformatic analysis to identify human sequences complementary to a given gRNA displaying up to two mismatched nucleotides was performed. Results are summarized in**Table 6**.

Except for gRNA12 and gRNA13, all the other designed sgRNAs offer high specificity. The few off-target sites (up to two mismatches) individuated for the designed gRNAs are located in introns or intergenic regions (**Annex 2**). Nonetheless, it is challenging to *a priori* evaluate the risk of cleavage at genomic locations having mismatches with the sgRNA, therefore the actual biological significance of off-target sites should always be experimentally validated, in relation also to the target cell line, delivery system and exposure time to the editing system (Corsi et al. 2022, Daer et al. 2017, Huston et al. 2019, Lin et al. 2014, Schep et al. 2021, Zhuo et al. 2021).

Table 6: Bioinformatic off-target assessment allowing for up to 3 nucleotide (nt) mismatches

sgRNA name	Number of mismatched sequences for		
	1 nt	2 nt	3 nt
gRNA1	0	2	16
gRNA2	1*	0	5
gRNA3	1*	1	15
gRNA4	1*	0	19
gRNA5	0	0	5
gRNA6	0	1	17
gRNA7	0	1	48
gRNA8	0	0	1
gRNA9	0	2	110
gRNA10	0	1	33
gRNA11	0	1	12
gRNA12	1	19	74
gRNA13	0	5	72
gRNA14	0	2	22
gRNAI	0	0	14
gRNAII	0	1	11
gRNAIII	0	2	18
gRNAVIIa	1*	0	58
gRNAVIIb	1*	2	112
gRNAVIIwt	0	1	43
gRNAVIII	0	0	11
gRNAIX	0	1	8
gRNAXVI	0	1	11

*DIV-specific gRNA, the 1-nt mismatch corresponds to the wild-type allele

6.5 Minigene assay to preliminary assess the rescue potential of the CRISPR/*Sp*Cas9-based approaches

Minigene plasmids encoding for the deep-intronic variants of interest were co-transfected in HEK293T cells alongside the different editing plasmids for the preliminary selection of efficient editing strategies.

6.5.1 dgRNAs/*Sp*Cas9, dgRNAs/*Sp*Cas9n, sgRNA/*Sp*Cas9 and sgRNA/ED*Sp*Cas9 mediated splicing rescue using an *ABCA4* c.5197-557G>T minigene model

In a first control experiment, HEK293T cells were co-transfected with the mutant minigene and a Mock *Sp*Cas9 targeted to a protospacer not present in GRCh38, and 97.3±0.1% aberrant and 2.7±0.1% correctly spliced *ABCA4* minigene transcripts were detected (**Figure 27A**) Of note, when the same *ABCA4*-specific primer pair was tested for endogenously expressed *ABCA4* transcript in HEK293T cells, no transcript was detected (**Figure 27B**).
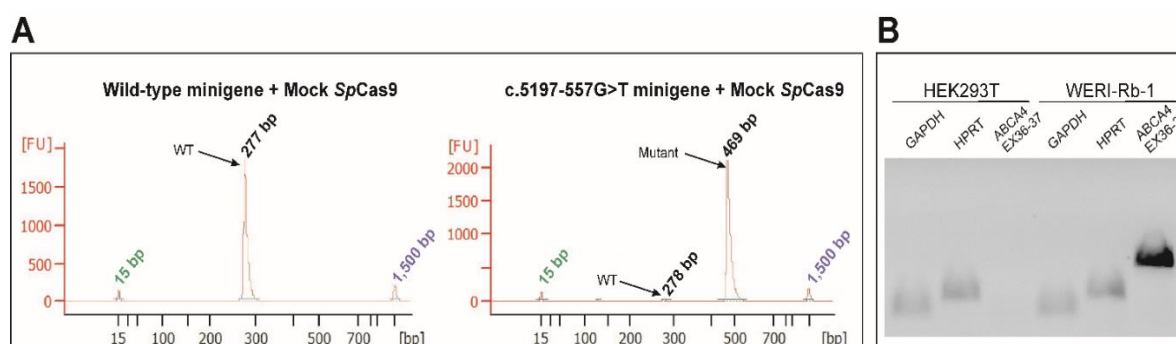

Figure 27: ***ABCA4* c.5197-557G>T-induced splicing pattern** (**A**) Example of the splicing pattern obtained in HEK293T cells transfected with a wild-type minigene + Mock SpCas9 and the c.5197-557G>T minigene + Mock *Sp*Cas9. The results shown are derived from chip-electrophoresis on a bioanalyzer: length of the fragment (bp) is represented on the X axis and fluorescence intensity corresponding to the amount of product plotted on the Y axis (FU). The wild-type (WT) peak and mutant peak are highlighted. The 15 bp and 1,500 bp peaks represent the internal markers. (**B**) Agarose gel showing absence of *ABCA4* expression in HEK293T cells. WERI-Rb1 cells were used as a positive control for *ABCA4* expression. *GAPDH* and *HPRT* were used as positive controls for cDNA synthesis and amplification.

Co-expression of the four different dgRNAs/*Sp*Cas9 combinations with the mutant minigene increased the fraction of correctly spliced transcripts to 98.4%±0.1 − 100.0%±0.0. Similarly, an increase in the

range to 71.9±1.8% and 99.5±0.7% of correctly spliced transcript was observed for the six dgRNAs/*Sp*Cas9n combinations. Specifically, gRNAs4+5 resulted in the lowest increase to 71.9±1.8%, whereas the remaining dgRNAs/*Sp*Cas9 combinations achieved similar rescue values in the range of 90.1±1.4% (gRNAs4+6) and 99.1±0.7% (gRNAs10+6) (**Figure 29**).

The distances to the isolated DIV of predicted sgRNA cut sites employed in the sgRNA/*Sp*Cas9 approach are reported in **Figure 28**. gRNA1 and gRNA6 are the farthest from the isolated DIV, up- and down-stream, respectively.
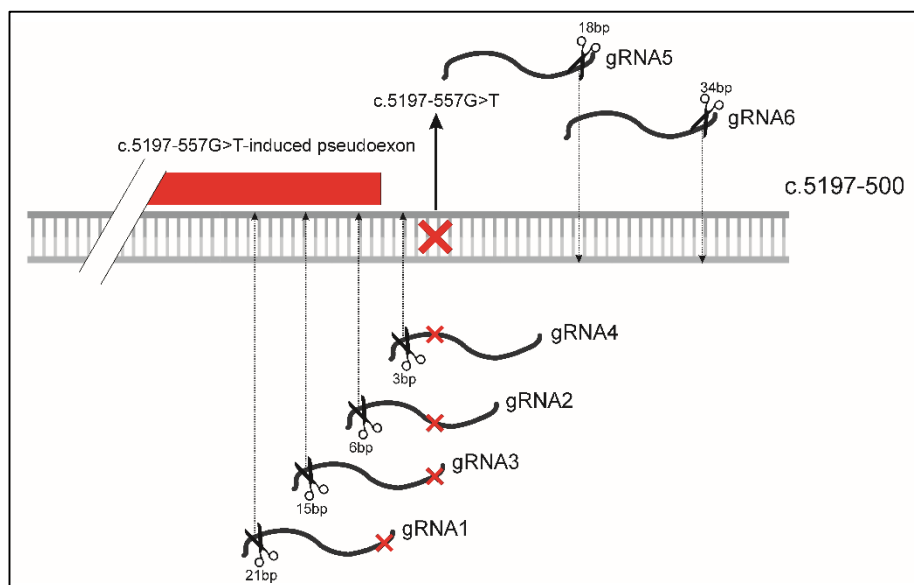
Figure 28: Schematic illustration of the location of the single gRNAs (sgRNA) used in the sgRNA/*Sp*Cas9 and /ED*Sp*Cas9 approaches in comparison to the c.5197-557G>T-induced pseudoexon. For each sgRNA, the distance of the cut site to c.5197-557G>T is depicted. The red cross represents the c.5197-557G>T deep-intronic variant.

When evaluating the six different sgRNA/*Sp*Cas9 combinations, the fraction of correctly spliced transcript strongly varied depending on the sgRNA used. With gRNA4 a complete rescue (100.0±0.0%) was obtained, whereas intermediate rescue levels were observed with gRNA1: 60.9±4.0%, gRNA2: 49.6±4.4% and gRNA3: 51.1±1.1% correctly spliced transcripts, respectively. Only minor rescue levels of 27.1±3.6% and 10.0±2.6% correctly spliced transcripts were obtained with gRNA5 and gRNA6, respectively (**Figure 29A**).

Figure 29: Minigene splicing assay of the CRISPR/*Sp*Cas9 strategies targeting *ABCA4* c.5197-557G>T. (**A**) Minigene splicing assay to evaluate the rescue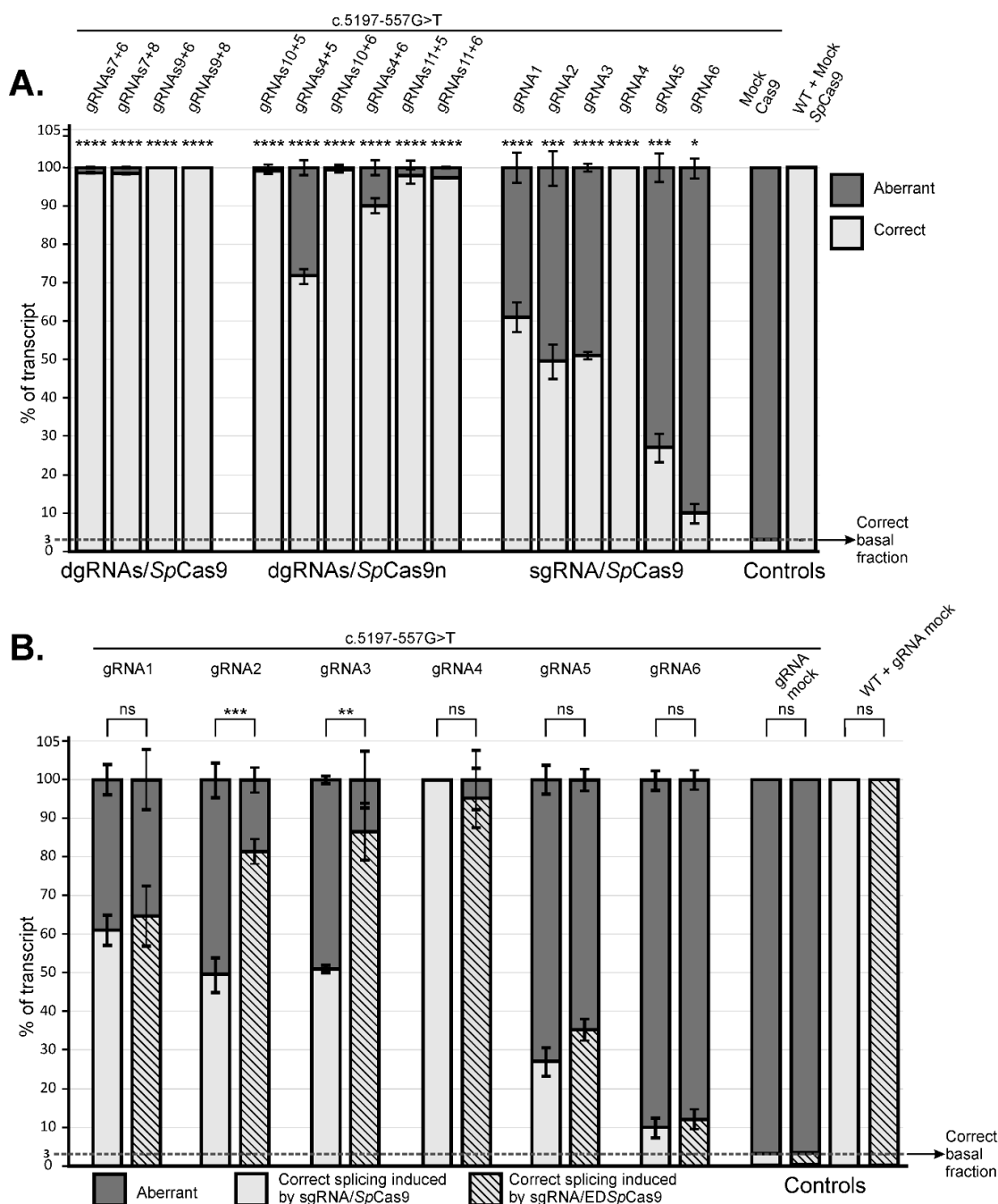 of *ABCA4* c.5197-557G>T-induced splice defects in HEK293T cells of the dual sgRNA/*Sp*Cas9 (dgRNAs/*Sp*Cas9), dual gRNAs/*Sp*Cas9 nickase (dgRNA/SpCa9n), and single sgRNA/*Sp*Cas9 (sgRNA/*Sp*Cas9) approaches. Statistically significant changes in *ABCA4* % of transcript are expressed as * = p≤0.05, *** = p≤0.001, **** = p ≤0.0001, and ns = non-significant compared to Mock Cas9. (**B**) Minigene splicing assay to evaluate the rescue of *ABCA4* c.5197-557G>T-induced splice defects in HEK293T cells of standard sgRNA/*Sp*Cas9 in comparison to the novel sgRNA/Enhanced-Deletion *Sp*Cas9 (sgRNA/ED*Sp*Cas9) approach. Statistically significant changes in *ABCA4* % of transcript are expressed as ** = p≤0.01, *** = p≤0.001, **** = p ≤0.0001, and ns = non-significant. (**A,B**) HEK293T cells were co-transfected with wild-type (WT) or mutant (*ABCA4* c.5197-557G>T) *ABCA4* minigene constructs and plasmids encoding for the four different *Sp*Cas9-mediated editing strategies. Relative proportions (in %) of correctly and aberrantly spliced transcripts as quantified from RT-PCR products. A plasmid expressing *Sp*Cas9 and a scrambled gRNA was used as a Mock *Sp*Cas9 control. (**A,B**) Results are presented as mean ± SD (n= 3 independent transfections).

The use of gRNA2 and gRNA3 in combination with the novel bioengineered ED*Sp*Cas9 substantially enhanced splicing rescue, leading to 81.3±3.3% and 86.5±7.3% of correctly spliced transcripts, respectively. Moderate splicing rescue improvement was obtained by the combination of gRNA5/ED*Sp*Cas9 (35.2±2.8% compared to 27.1±4.5% of gRNA5/*Sp*Cas9), whereas no substantial increment was observed by implementing gRNA1/, gRNA4/ and gRNA6/ED*Sp*Cas9 (**Figure 29B**).

6.5.2 sgRNA/ED*Sp*Cas9 mediated splicing rescue in the *ABCA4* intron 36 clustered deep-intronic variants c.5196+1013A>G, c.5196+1056A>G, c.5196+1134C>G, c.5196+1137G>A and c.5196+1216C>A minigene models

The designed sgRNAs to target the clustered DIVs in intron 36 of *ABCA4* were coupled to ED*Sp*Cas9 and also to *Sp*Cas9 as a reference strategy for comparison. gRNAI, gRNAII, gRNAIII, and gRNAXVI were used to rescue the c.5196+1013A>G- and c.5196+1056A>G-induced missplicing. gRNAVII and gRNAVIII were implemented as DIV-independent sgRNAs to the rescue of the missplicing resulting from the c.5196+1134C>G, c.5196+1137G>A, and c.5196+1216C>A DIVs. Conversely, the location of gRNAXVII overlaps with c.5196+1134C>G and c.5196+1137G>A, therefore three variants had to be designed: gRNAXVIIa, gRNAXVIIb and gRNAXVIIwt, fully matching to c.5196+1134C>G, c.5196+1137G>A, and c.5196+1216C>A, respectively.

All designed sgRNAs were only tested once in a pilot experiment for their genome editing potential in a single minigene splicing assay (**Figure 30**).
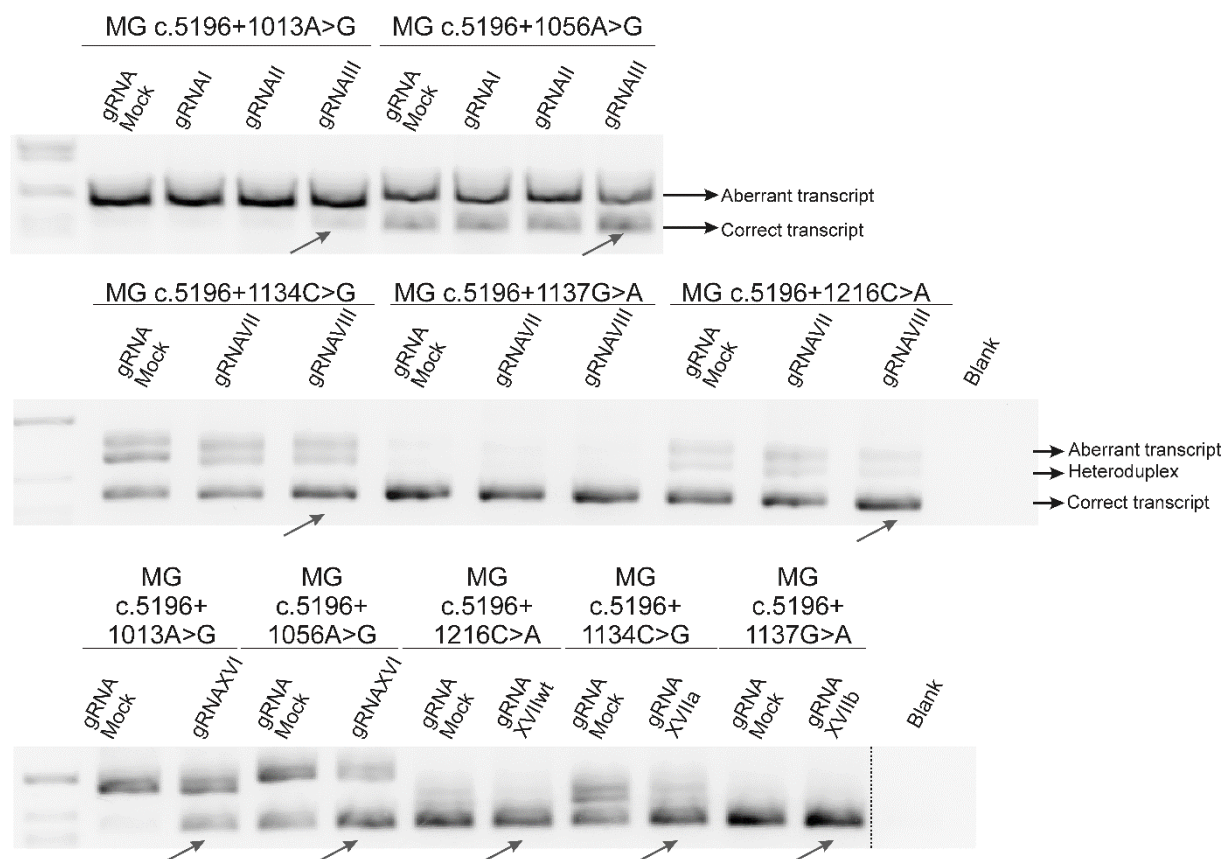
Figure 30: Pilot minigene (MG) splicing assay to determine the most effective CRISPR/Enhanced-Deletion *Sp*Cas9 (sgRNA/ED*Sp*Cas9) combinations in correcting aberrant splicing back to wild-type splicing for the individual clustered DIVs. The Mock lane refers to the untreated situation. The lower product shown on the gel refers to the correct transcript, whereas the higher product represents the aberrant transcript including PE inclusion. Middle products are expected to be artifacts due to heteroduplex formation produced during PCR amplification. The red arrows highlight the increased amounts of correctly spliced transcript upon sgRNA/ED*Sp*Cas9 treatment of the most effective gRNAs. Blank = Negative PCR control without template. The dashed line indicates that the original gel picture was cut.

Visual assessment of the splicing rescue indicated gRNAIII, gRNAXVI, gRNAVII and, gRNAXVIIa/b/wt as being most effective, while gRNAI and gRNAII did not shown any rescue and were not further tested (**Figure 30**). Thereafter, gRNAIII and gRNAXVI were selected for further validation in the rescue of c.5196+1013A>G- and c.5196+1056A>G-induced aberrant splicing, whereas gRNAVIII, gRNAIX and gRNAVIIa/b/wt were chosen for targeting c.5196+1134C>G, c.5196+1137G>A and c.5196+1216C>A, respectively.

In control experiments, HEK293T cells were again co-transfected with a mutant minigene and a Mock ED*Sp*Cas9 or *Sp*Cas9 targeted to a protospacer not present in GRCh38. Most of the basal fractions of correctly spliced transcripts varied considerably between standard *Sp*Cas9 or ED*Sp*Cas9 (see discussion). An example of the different splicing pattern obtained upon co-transfection of the c.5196+1056A>G minigene plasmid and Mock ED*Sp*Cas9 or Mock*Sp*Cas9 is reported in **Figure 31**. The

area of the peaks represents the concentration of the amplicons, which is then corrected for the size to obtain the molarity (number of amplicons) of the different transcripts. The fraction of correctly spliced transcripts is calculated using the molarity (**Table 7**, **Figure 32**).
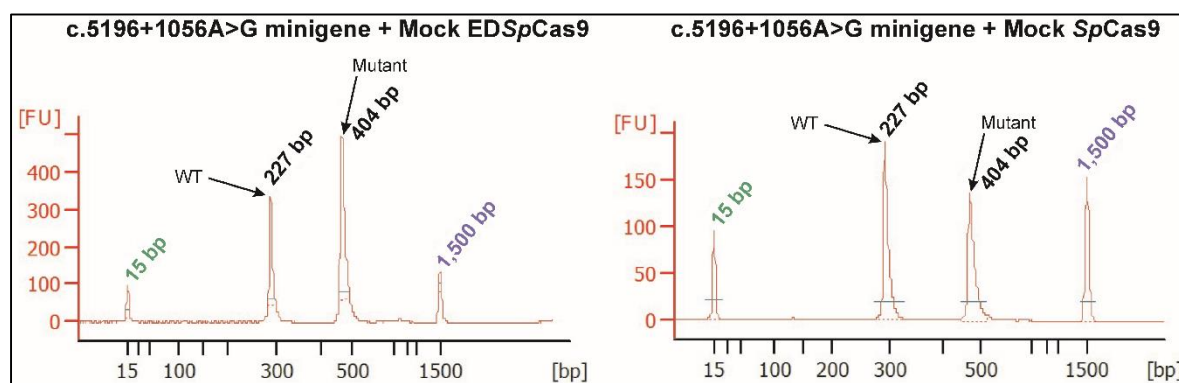

Figure 31: Example of the splicing pattern obtained in HEK293T cells co-transfected with the c.5196+1056A>G minigene + Mock ED*Sp*Cas9 (left) or Mock *Sp*Cas9 (right). The results shown are derived from chip-electrophoresis using the Bioanalyzer device: length of the fragment (bp) is represented on the X axis and fluorescence intensity corresponding to the amount of product plotted on the Y axis (FU). The wild-type (WT) peak and Mutant peak are highlighted. The 15 bp and 1,500 bp peaks represent the internal markers.

Table 7: Basal fraction of correctly spliced transcripts induced by the different clustered intron 36 DIVs as determined by fragment analysis (Bioanalyzer) within the respective minigene assays. Experiments were done in duplicate. Results are expressed as mean ± standard deviation.

Minigene	Fraction of correctly spliced transcripts	Comment
c.5196+1013A>G + Mock ED*Sp*Cas9	9.0 ± 3.0%	
c.5196+1013A>G + Mock *Sp*Cas9	16.6 ± 5.2%	
c.5196+1056A>G + Mock ED*Sp*Cas9	50.3 ± 3.7%	
c.5196+1056A>G + Mock *Sp*Cas9	73.3 ± 6.3%	
c.5196+1134C>G + Mock ED*Sp*Cas9	67.6 ± 4.1%	
c.5196+1034C>G + Mock *Sp*Cas9	80.6 ± 2.6%	
c.5196+1137G>A + Mock ED*Sp*Cas9	99.0 ± 1.4%	Variant does not result in substantial missplicing in minigene assay
c.5196+1137G>A + Mock *Sp*Cas9	100 ± 0%	
c.5196+1216C>A + Mock ED*Sp*Cas9	92.6 ± 3.4%	Variant does not result in substantial missplicing in minigene assay
c.5196+1216C>A + Mock *Sp*Cas9	94.0 ± 2.4%	

Due to intra-experimental variability of the basal fraction of correctly spliced transcripts (**Table 7**) and to obtain and present comparable results, the results out of these experiments and the rescue is presented as relative percentage increase of correctly spliced transcript normalized to the respective *Sp*Cas9 or ED*Sp*Cas9 background of the correctly-spliced transcript (**Figure 32**).
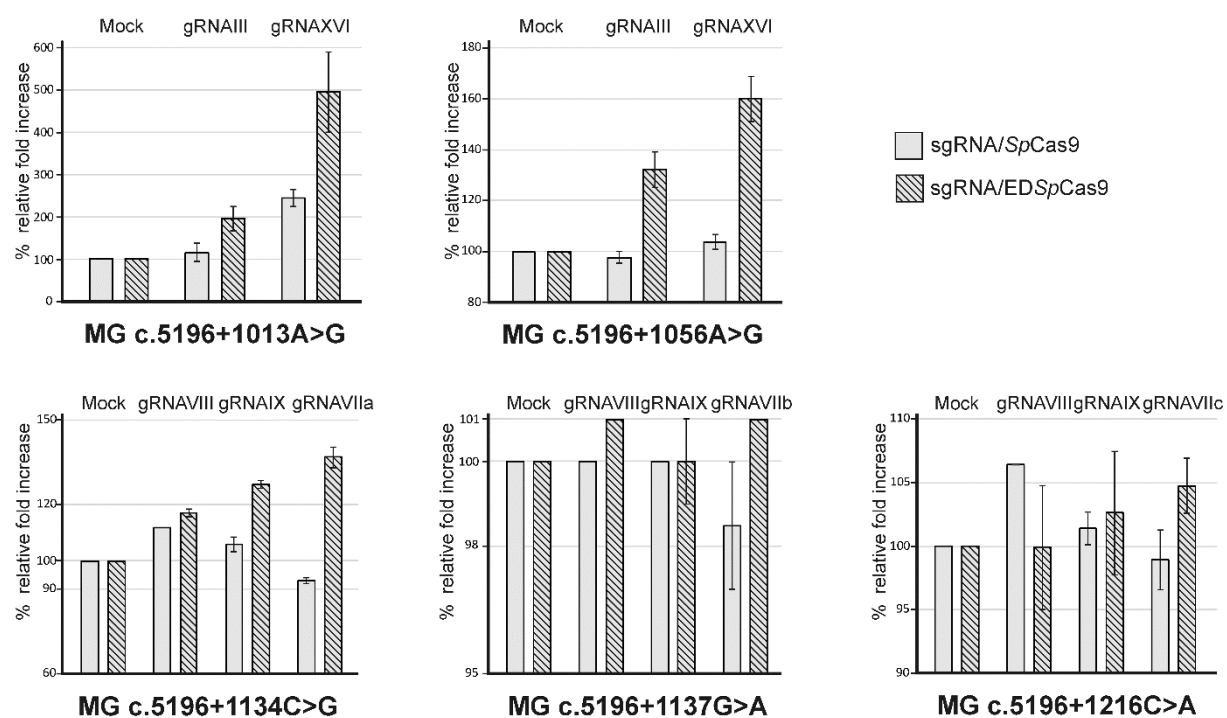
Figure 32: Minigene splicing assay of the CRISPR/Enhanced-Deletion *Sp*Cas9 strategies (sgRNA/ED*Sp*Cas9) targeting clustered *ABCA4* DIVs in intron 36 in comparison to *Sp*Cas9. The minigene (MG) splicing assay was performed in HEK293T cells. Results are presented as mean of the percentage (%) relative fold increase normalized to Mock control ± SD (n= 2 independent transfections).

When complexing gRNAIII to ED*Sp*Cas9, the relative percentage of correctly spliced transcript for c.5196+1013A>G and c.5196+1056A>G minigenes after genome editing increased by 194 ± 28% and 132 ± 7%, respectively, compared to 114 ± 23% and 104 ± 2% obtained with *Sp*Cas9 and the same sgRNA. For gRNAXVI complexed to ED*Sp*Cas9 a considerably higher relative percentage of correctly spliced transcript of 494 ± 94% for c.5196+1013A>G, and 160±9%, for c.5196+1056A>G, compared to 243±20% and 106±3% obtained with *Sp*Cas9 and the same sgRNA was achieved.

Similarly, gRNAVIII, gRNAIX and gRNAVIIa complexed to ED*Sp*Cas9 increased the relative percentage of correct transcript for c.5196+1134C>G of 117±1%, 127±1% and 137±4%, respectively, as opposed to *Sp*Cas9 which attained 112±0%, 106±2% and 93±1%, respectively.

Due to the negligible percentage of c.5196+1137G>A- and c.5196+1216C>A-induced aberrant transcripts available to the rescue, the increase of the correct transcript, upon genome editing, was substantially limited: when targeting c.5196+1137G>A, gRNAVIII, gRNAIX and gRNAVIIb complexed to ED*Sp*Cas9 resulted in 101±0%, 100±0% and 101±0% correctly spliced transcript, respectively. *Sp*Cas9 complexed with the same sgRNAs and targeted to the same DIV induced similar results: gRNAVIII, gRNAIX and gRNAVIIb targeted to c.5196+1137G>A led to 100±0%, 100±0%, and 98.5±1.5% variation of the percentage of the correct transcript.

Similarly, the rescue of the correct transcript obtained upon editing with gRNAVIII, gRNAIX and gRNAVIIc complexed to ED*Sp*Cas9 was minimal: 99±0%, 103±0%, and 105±0% for c.5196+1216C>A, respectively. The same sgRNAs coupled to *Sp*Cas9 showed similar outcomes: 106±0% for gRNAVIII, 101±0% for gRNAIX, and 99±0% for gRNAVIIc correctly spliced transcript (**Figure 32**).

A more accurate evaluation of the splicing rescue potential for c.5196+1137G>A and c.5196+1216C>A will be performed in cone photoreceptor precursor cells in future experiments, which are expected to show higher levels of aberrantly spliced transcript (Khan et al. 2020a, Results: 6.13)

6.6 c.5197-557G>T-INDUCED ABERRANT SPLICING RESCUE IN CONE PHOTORECEPTOR PRECURSOR CELLS

In order to validate the results obtained in the minigene splicing assay for the rescue of aberrant splicing induced by the isolated c.5197-557G>T DIV, patient-derived CPCs were used as a cellular model. Since our ethics approval does not allow the collection of skin punches from underage patients, human material for fibroblast isolation was collected from the father of a young STGD1 patient who is compound heterozygous for *ABCA4* DIV c.5197-557G>T and the nonsense mutation c.5917delG; p.Val1973Ter. An iPS cell line derived from this healthy heterozygous proband for the isolated c.5197-557G>T DIV (iPSC$^{c.5197-557G/T}$) was generated by reprogramming of fibroblasts into iPSCs. iPSC$^{c.5197-557G/T}$ were differentiated into CPCs$^{c.5197-557G/T}$ through a 30-day differentiation protocol.
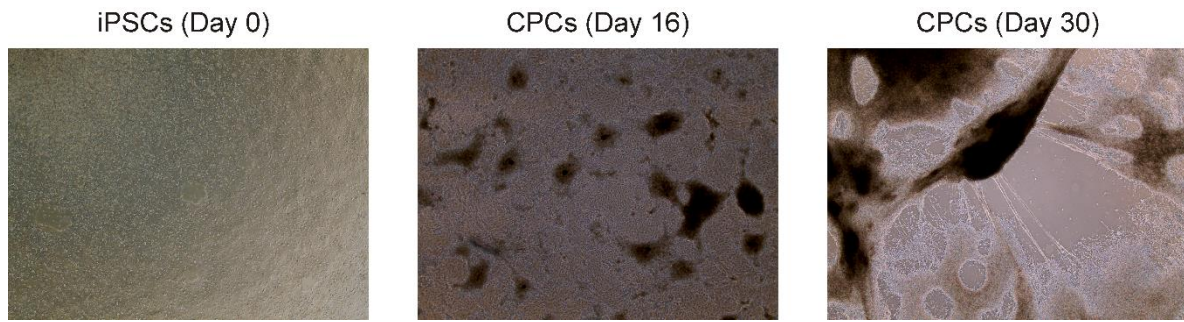
Figure 33: Appearance of cone photoreceptor precursor cells (CPCs) during differentiation. 30-days differentiated cone photoreceptor precursor cells[c.5197-557G/T] (CPCs[c.5197-557G/T]) showing 3D cellular structures and highly packed cellular organization.

After 30 days of differentiation, CPCs[c.5197-557G/T] appeared tightly packed, 3D-like cellular structures formed (**Figure 33**) and different early retina- and photoreceptor-specific markers started to express, including *ABCA4*: 30-days-differentiated CPCs[c.5197-557G/T] showed upregulation of *PAX6*, *RCVRN*, *RPE65*, *ABCA4* and *OPN1SW*, and downregulation of *OPN1LW* and *OCT3/4* (**Figure 34**, **Table 8**).
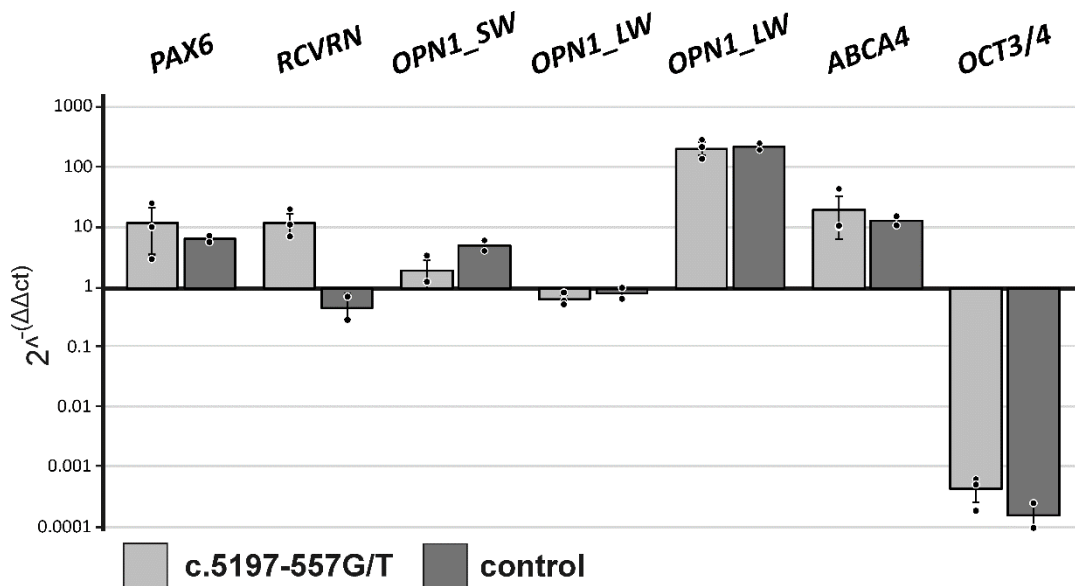
Figure 34: Expression profiling of 30-days cone photoreceptor precursor cells (CPCs) for selected retinal markers. Relative quantification of retinal marker gene expression in CPCs. Data of 30-days CPCs are shown. Data are normalized to *GUSB* and expressed as fold change compared to undifferentiated CPCs (Day 0). Results are presented as mean ± SD including the data points (n=3 for c.5197-557G>T and n=2 for control). Downregulation of the pluripotency marker *OCT3/4* denotes the differentiation status of CPCs.

Table 8: Relative quantification value of retinal marker gene expression in 30-days cone photoreceptor precursor cells (CPCs) compared to day-0 iPSCs. Value are presented as mean of the fold change ($2^{-\Delta\Delta ct}$) ± SD (n=3 for c.5197-557G>T and n=2 for control)

Cell line	Gene						
	PAX6	*RCVRN*	*OPN1_SW*	*OPN1_LW*	*RPE65*	*ABCA4*	*OCT3/4*
c.5197-557G/T	12±8	12±5	2±1	0.65±0.11	213±53	21±13	0.0005±0.0002
Control	7±1	0.5±0.2	5±1	1.0±0.4	210±25	12±3	0.0002±0.0001

Comparable up- and downregulation was observed in 30-days-differentiated control CPCs (**Figure 34, Table 8**). Qualitative splicing pattern on agarose gel showed the presence of correctly spliced, as well as aberrant transcripts, and sequencing confirmed the identity of the transcripts and consistency with the splicing pattern obtained by the minigene assay (**Figure 35**).
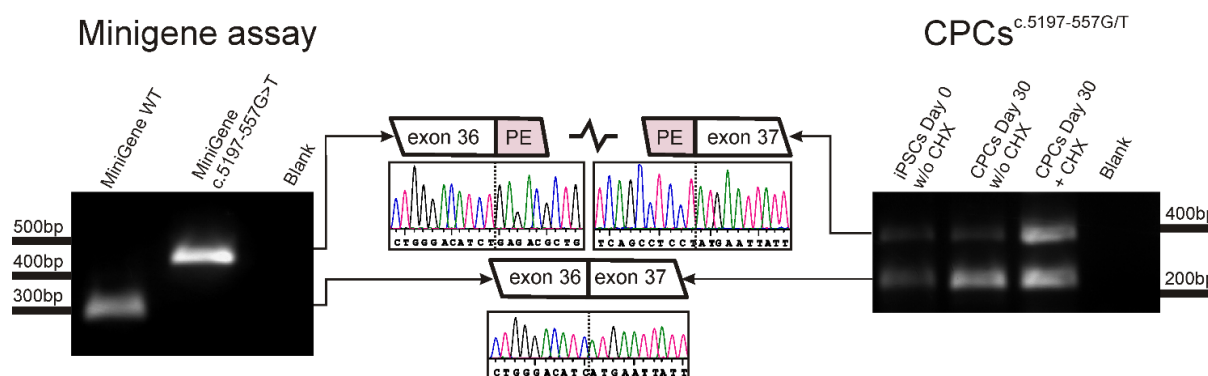
Figure 35: Splicing pattern for the c.5197-557G>T deep-intronic variant in the cellular CPC (cone photoreceptor precursor) models used in this thesis. Left: Agarose gel analysis of RT-PCR products for the wild-type (WT) and mutant (c.5197-557G>T) minigene expressed in HEK293T cells demonstrate correct exon 36 – exon 37 splicing of wild-type and fully penetrant missplicing with the inclusion of the pseudoexon (PE) in mutant transcripts. **Right**: Electrophoretic separation of RT-PCR amplified splicing products in patient-derived heterozygous *ABCA4* iPSCs[c.5197-557G/T] (Day 0) and CPCs [c.5197-557G/T]. Day 0 w/o cycloheximide (CHX) corresponds to the parental iPS [c.5197-557G/T] cell line. CHX was added to suppress NMD. Day 30+CHX and Day 30 w/o CHX correspond to 30-day-old CPCs[c.5197-557G/T] treated overnight with 0.1 mg/ml or untreated; respectively. Middle: Representative electropherograms of RT-PCR products from patient-derived CPCs Day30 + CHX are shown.

6.7 DGRNAs/*Sp*CAS9, DGRNAS/*Sp*CAS9N, SGRNA/*Sp*CAS9 SPLICING RESCUE IN CONE PHOTORECEPTOR PRECURSOR CELLS.

As described above, the eight most-effective *Sp*Cas9(n)-based strategies tested in the minigene assays in HEK293T cells were selected for further validation in CPC$^{c.5197-557G/T}$. Specifically, two sgRNA combinations each for dgRNAs/*Sp*Cas9 (gRNAs7+8 and gRNAs9+6), as well as for dgRNAs/*Sp*Cas9n (gRNAs10+5 and gRNAs10+6) were chosen. Four sgRNAs (gRNA2, gRNA3, gRNA4 and gRNA5) were selected for the sgRNA/*Sp*Cas9 approach.

In non-transfected CPC$^{c.5197-557G/T}$ (NT-CPC$^{c.5197-557G/T}$), only 5.8±0.6% aberrant *ABCA4* transcript could be detected due to degradation of aberrant transcripts by NMD. To overcome this issue, CPC$^{c.5197-557G/T}$ were treated with 0.1 mg/ml cycloheximide (CHX) – a known NMD blocker – 16 h prior to harvesting, hereby increasing the percent of aberrant *ABCA4* transcript to 27.5±1.5%, corresponding to a 4.6±0.4-fold increase. Non-transfected CPCs and CPC$^{c.5197-557G/T}$ transfected with Mock Cas9 (*Sp*Cas9 + scramble sgRNA) were used as controls. Day-30 CPC$^{c.5197-557G/T}$ were transfected and subcultured for 7 days post electroporation before quantification of splicing rescue of *ABCA4* transcripts and editing efficiency on gDNA level. Of note, the 7 days post electroporation time point was selected, as preliminary experiments showed higher rescue compared to 48 hours post electroporation (**Annex 3**).

All tested sgRNA combinations were able to improve correct splicing albeit with variable degree of efficiency (**Figure 36**, **Table 9**).
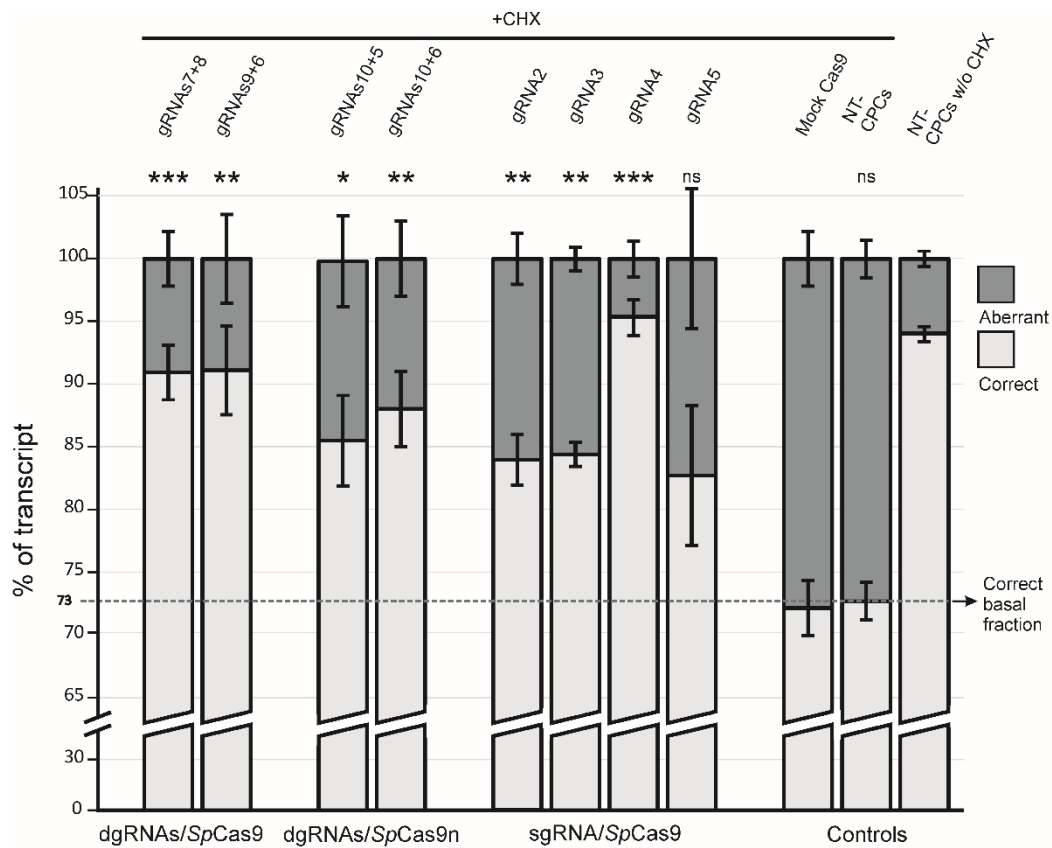
Figure 36: Rescue of the *ABCA4* splice defect in patient-derived heterozygous c.5197-557G/T cone photoreceptor precursor cells (CPC$^{c.5197-557G/T}$). Electroporated CPC$^{c.5197-557G/T}$ were analyzed 7-days post electroporation for assessing relative rescue efficacy of the tested *Sp*Cas9- and *Sp*Cas9n-mediated approaches. 16 h prior to harvesting, 0.1 mg/ml cycloheximide (CHX) was added. Non-transfected CPCs (NT-CPCs) without (w/o) CHX treatment, NT-CPCs treated with CHX, and cells transfected with Mock Cas9 (*Sp*Cas9 + scramble sgRNA) were used as controls. Relative quantification of aberrant and correct transcripts is shown. Results are presented as mean ± standard deviation (n= 3 independent electroporations). Statistically significant changes in *ABCA4* % of transcript are expressed as * = p≤0.05, *** = p≤0.001, **** = p ≤0.0001, and ns = not significant compared to Mock Cas9.

Table 9: Relative splicing analysis of cone photoreceptor precursor cells[c.5197-557G/T](CPC[c.5197-557G/T]) (7-days post electroporation). The rescue of the mutant splicing (%) is calculated normalizing to the mean of aberrant transcript (%) of NT-CPCs + CHX. SD = standard deviation, NTC = non-transfected control, CHX = cycloheximide.

	Mean of correct transcript [%]	SD [%]	Mean of aberrant transcript [%]	SD [%]	Rescue of mutant splicing [%]	SD [%]
dgRNAs/*Sp*Cas9						
gRNAs7+8+CHX	90.9	2.2	9.1	2.2	67.1	7.8
gRNAs9+6+CHX	91.1	3.5	8.9	3.5	67.7	12.8
dgRNAs/*Sp*Cas9n						
gRNAs10+5+CHX	85.4	3.6	14.6	3.6	47.1	13.0
gRNAs10+6+CHX	87.9	3.0	12.1	3.0	56.1	10.8
sgRNA/*Sp*Cas9						
gRNA2+CHX	83.9	2.0	16.1	2.0	41.5	7.3
gRNA3+CHX	84.5	1.0	15.5	1.0	43.9	3.5
gRNA4+CHX	95.2	1.4	4.8	1.4	82.7	5.0
gRNA5+CHX	82.5	5.6	17.5	5.6	36.4	20.3
Controls						
Mock Cas9+CHX	71.9	2.2	28.1	2.2		
NT-CPCs + CHX	72.5	1.5	27.5	1.5		
NT-CPCs w/o CHX	94.2	0.6	5.8	0.6		

The basal fraction of the correct transcript in NT-CPC[c.5197-557G/T] treated with CHX was 72.5±1.5% and is not altered by the electroporation procedure or the expression of *Sp*Cas9 as seen in Mock Cas9 controls (71.9±2.2%) (**Figure 36** and **Table 9**).

For the two dgRNAs/*Sp*Cas9 combinations 90.9±2.2% (gRNAs7+8) and 91.1±3.5% (gRNAs9+6) correctly spliced transcripts were observed indicating a splicing rescue of 67.1±7.8% and 67.7±12.8%, respectively (**Figure 36**, **Table 9**).

The dgRNAs/*Sp*Cas9n combinations also resulted in splicing rescue in CPC[c.5197-557G/T], but yielded lower fractions of correctly spliced transcript (gRNAs10+5: 85.4±3.6 % and gRNAs10+6: 87.9±3.0%) compared to dgRNAs/*Sp*Cas9 (**Figure 36**, **Table 9**). The relative ratio of correctly spliced transcript upon treatment indicates a splicing rescue of 47.1±13.0% and 56.1±10.8%, respectively.

The four sgRNA/*Sp*Cas9 combinations also achieved substantial relative splicing correction, with gRNA4 being outstanding effective as it increased the fraction of correct transcript to 95.2±1.4% (82.7±5.0% splicing correction), showing the highest relative rescue among the eight tested strategies

(**Figure 36**, **Table 9**). The gRNA2/, gRNA3/, and gRNA5/*Sp*Cas9 combinations yielded similar fractions of correct transcript of 83.9±2.0%, 84.5±1.0% and 82.5±5.6%, respectively, corresponding to a splicing rescue of around 40%.

Upon successful application of the genome editing approaches in CPC$^{c.5197-557G/T}$, an increase in the overall *ABCA4* transcript amount, resulting from rescued *ABCA4* transcript, should be observed also without cycloheximide treatment, as the rescued transcripts would not undergo NMD.

To investigate this, one strategy for each approach was selected (gRNAs7+8 for dgRNAs/*Sp*Cas9, gRNAs10+6 for dgRNAs/*Sp*Cas9n and gRNA4 for sgRNA/SpCas9) and the corresponding genome editing plasmids transfected in day-30 CPC$^{c.5197-557G/T}$. The cells were subcultured for 35 additional days to allow full *ABCA4* transcript turnover and to clear transfected CPCs from the editing plasmids, which occurred in the range of 20-25-days post electroporation as determined by visual control of EGFP-positive cells (i.e. loss of EGFP expression conveyed by the editing plasmids). Clearance of the editing plasmids was desired to prevent possible bias of RT-qPCR results due to sustained overexpression of an exogenous protein. Total mRNA was extracted from 65 days old CPC$^{c.5197-557G/T}$, and *ABCA4* and *RCVRN* transcript levels were quantified by RT-qPCR. *RCVRN* was used as an additional experimental control.

For the gRNA4/*Sp*Cas9 combination, as well as the dgRNAs10+6/*Sp*Cas9n combination, a significant 1.8±0.2 and 1.7±0.3 -fold increase in *ABCA4* transcript levels compared to NT-CPCs was detected (**Figure 37**). For the dgRNAs7+8/*Sp*Cas9 combination a tendency towards increased *ABCA4* transcript levels was observed. CPC$^{c.5197-557G/T}$ transfected with Mock Cas9 showed a 0.9±0.3 –fold change in *ABCA4* transcript levels, comparable to NT-CPCs. In contrast, transcript levels of *RCVRN* did not change in the *Sp*Cas9 and *Sp*Cas9n treated CPCs, confirming that increased transcript levels were specific for the target gene *ABCA4* (**Figure 37**).
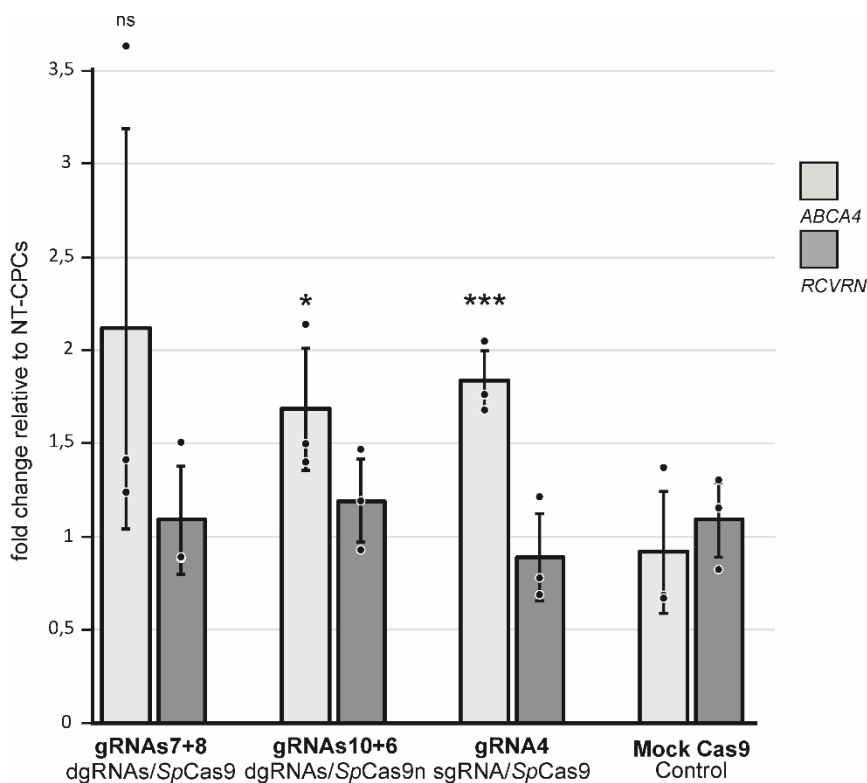
Figure 37: RT-qPCR for *ABCA4* transcript in patient-derived heterozygous c.5197-557G/T cone photoreceptor precursor cells (CPC^c.5197-557G/T) at 35-days post electroporation. Relative quantification of total *ABCA4* and *RCVRN* transcripts is shown for one (d)gRNA(s)/*Sp*Cas9(n) combination per approach in CPC^c.5197-557G/T without CHX treatment. Data are normalized to *GUSB* and expressed as fold change compared to NT-CPSs. Results are presented as mean ± standard deviation including the data points of the three biological replicates (N =3). Statistically significant changes in *ABCA4* and *RCVRN* transcript levels are expressed as * = p≤0.05 and *** = p≤0.001, and ns = not significant compared to Mock Cas9 (*Sp*Cas9 + scramble sgRNA).

6.8 ASSESSMENT OF GENOME EDITING APPROACHES ON WILD-TYPE *ABCA4* SPLICING IN CONE PHOTORECEPTOR PRECURSOR CELLS

To exclude secondary detrimental effects of the applied editing approaches on *ABCA4* canonical splicing, the same eight strategies were tested in CPCs derived from a healthy control (cCPCs). Agarose gel electrophoresis was used to visualize the RT-PCR amplified splicing products spanning exon 35 to exon 41 of *ABCA4,* so as to analyze splicing of adjacent exons as well. With all tested approaches and strategies one single bands corresponding to the expected fragment size of 838 bp was evident (**Figure 38A**). Sanger sequencing of the RT-PCR products confirmed the *ABCA4* transcripts to be correctly spliced (**Figure 38B**).

A.

B.
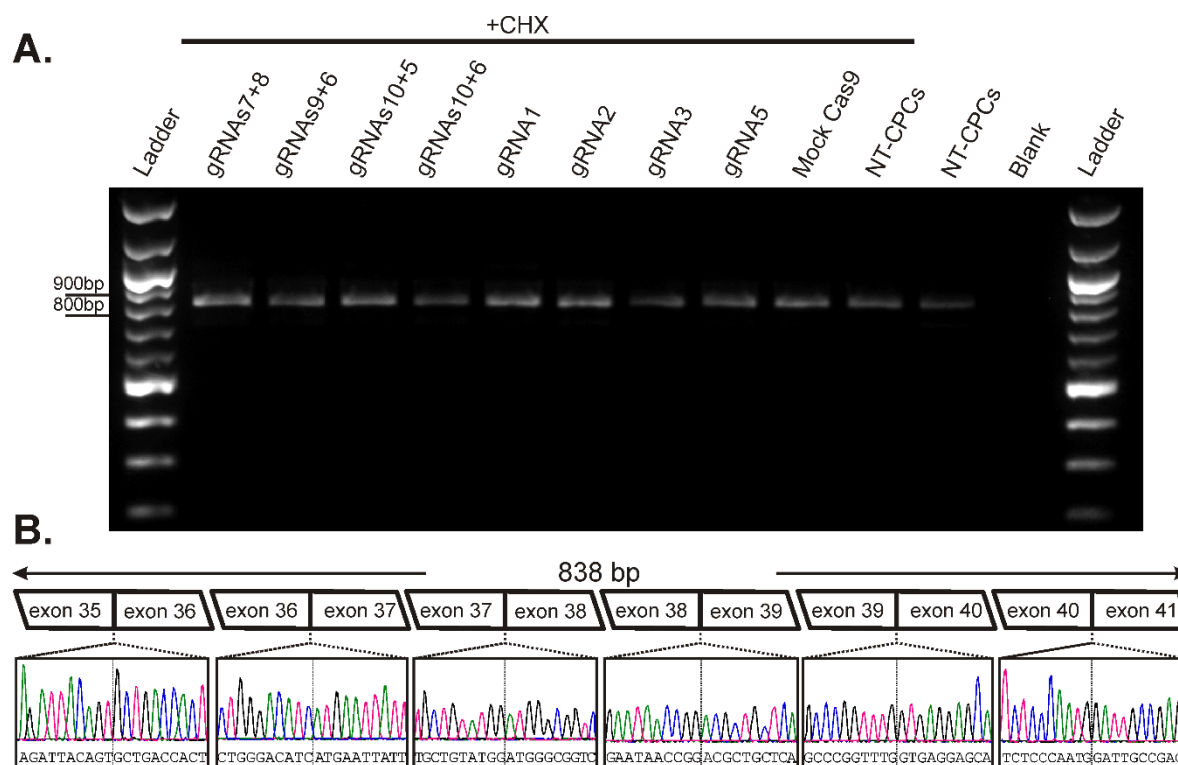
Figure 38: Verification of the correct *ABCA4* splicing pattern in control cone photoreceptor precursor cells (cCPCs) treated with the different *Sp*Cas9- and *Sp*Cas9n-mediated approaches. (**A**) Agarose gel analysis of RT-PCR products spanning exon 35 to exon 41 for cCPCs electroporated with various genome editing constructs. Cycloheximide (CHX) was added to detect possible aberrant transcripts. Only the correct transcript of 838 bp was detected. Blank = Negative PCR control without template (**B**) Sanger sequencing of the RT-PCR products revealed that the *ABCA4* transcripts was correctly spliced in control cCPCs at all exon-exon junctions treated with the different CRISPR/*Sp*Cas9 approaches. As a representative example, electropherograms of the exon-exon boundaries for gRNAs7+8/*Sp*Cas9 are shown.

6.9 GENOMIC DNA EDITING PROFILES IN CONE PHOTORECEPTOR PRECURSOR CELLS UPON EDITING MEDIATED BY THE "STANDARD" DGRNAS/*Sp*CAS9 GDNA, DGRNAS/*Sp*CAS9N AND SGRNA/*Sp*CAS9 APPROACHES

To correlate the efficacy of the editing approaches to modify the genomic target site with the splicing rescue potential, the editing efficiency at the gDNA level was analyzed 7-days post electroporation in the $CPC^{c.5197-557G/T}$. For the combinations used to evaluate increased expression of the overall *ABCA4* transcript amount (**Figure 37**), gDNA editing efficiency was also assessed 35-days post electroporation. PCR fragment analysis on a Bioanalyzer 2100 instrument was used to determine editing efficacy (i.e. deletions of ~600 bp) for the dgRNAs/*Sp*Cas9 approaches, whereas for the dgRNAs/*Sp*Cas9n and sgRNA/SpCas9 approaches next generation sequencing (NGS) was performed, since no discreet

fragment size differences were expected. NGS analysis also allowed to determine differences in the editing efficacy of the c.5197-557G>T and the wild-type allele (**Figure 39**, **Table 10**).
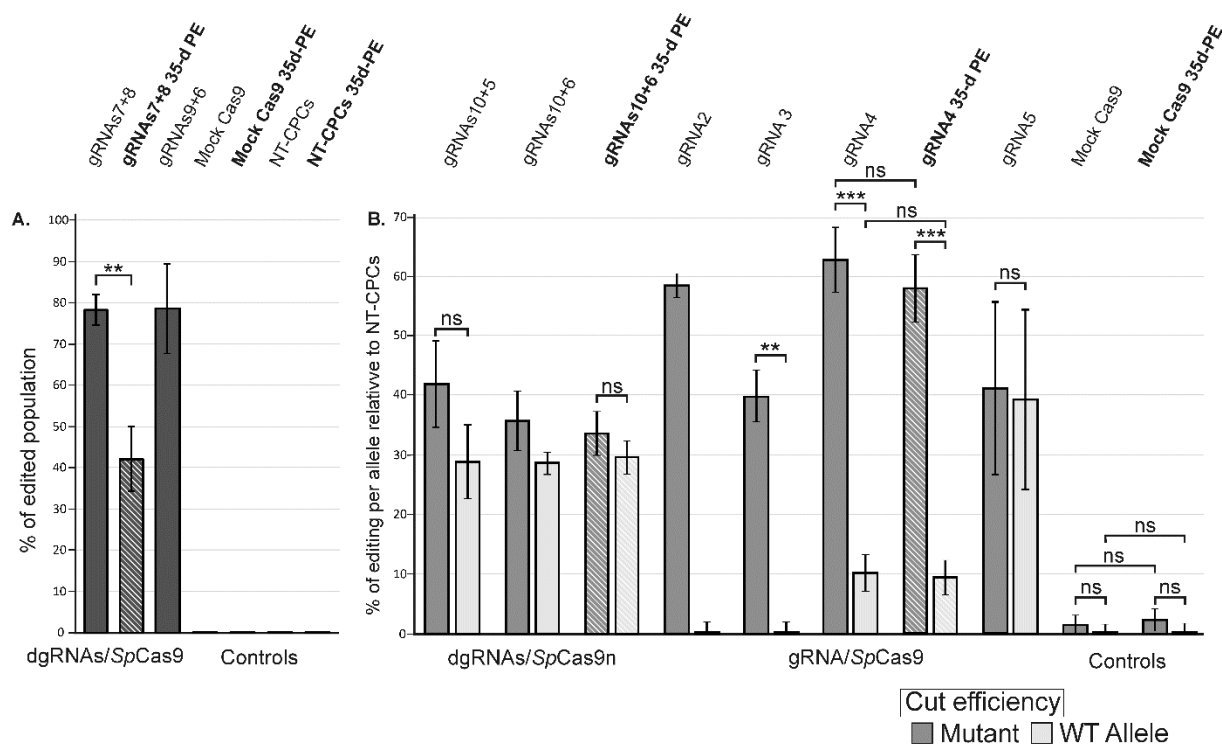

Figure 39: gDNA editing efficiency of the *Sp*Cas9- and *Sp*Cas9n-mediated approaches in patient-derived, heterozygous c.5197-557G/T cone photoreceptor precursor cells (CPC^c.5197-557G/T). (**A**) gDNA editing efficiency both for the gRNAs7+8/ and the gRNAs9+6/dgRNAs/*Sp*Cas9 approach calculated as percentage of edited alleles (not differentiating between wild-type and mutant alleles) at 7-days post electroporation and 35-days post electroporation (highlighted in bold). (**B**) dgRNAs/*Sp*Cas9n and sgRNA/*Sp*Cas9 gDNA editing efficiencies 7-days post electroporation and 35-days post electroporation (highlighted in bold) differentiated for editing of the mutant and wild-type allele, and normalized to non-transfected CPC^c.5197-557G/T (NT-CPCs). (**A,B**) For the samples gRNAs7+8/*Sp*Cas9 gRNAs10+6/*Sp*Cas9n, gRNA4/*Sp*Cas9, Mock Cas9 (*Sp*Cas9 + scramble sgRNA) and Non-Transfected Control (NTC) gDNA editing efficiencies at 35-days post electroporation (35-d PE) are also shown (mesh colors). Results are presented as mean ± standard deviation (n=2 − 3). Statistically significant differences between 7-days post electroporation and 35-days post electroporation as well as between mutant and wild-type (WT) alleles are expressed as ** = p≤0.01, *** = p≤0.001, and ns = non-significant.

Using the dgRNAs/*Sp*Cas9 approach, gDNA editing efficiency of 78.6±8.8% and 78.2±3.0% were achieved in CPC^c.5197-557G/T at 7-days post electroporation for the gRNAs9+6 and gRNAs7+8 combinations, respectively. Notably, the percentage of edited cells dropped to 42.1±6.4% at 35-days post electroporation in CPC^c.5197-557G/T treated with dgRNAs7+8/*Sp*Cas9 (**Figure 39A**) suggesting cellular senescence in CPCs treated with this strategy ([Discussion: 7.2.2](#)).

For the dgRNAs/*Sp*Cas9n approach overall editing efficiencies of 34.4±6.8% or 31.5±3.6% were achieved, respectively, in CPC[c.5197-557G/T] harvested 7-days post electroporation with editing constructs bearing gRNAs10+5 and gRNAs10+6, respectively. For dgRNAs10+6/*Sp*Cas9n editing efficiencies for both alleles were comparable (34.4±2.2% mutant allele and 28.7±5.0 wild-type allele) whereas the mutant allele was more efficiently edited compared to the wild-type allele when treating CPC[c.5197-557G/T] with dgRNA10+5/*Sp*Cas9n (40.2±6.2% mutant allele and 28.7±7.4% wild-type allele).

The outcome of the dgRNAs/*Sp*Cas9n approach applying gRNAs10+6 was also tested at 35-days post transfection. In this case, no difference could be seen comparing efficacies at 7-days and 35-days post transfection (**Figure 39A**).

For the sgRNA/*Sp*Cas9 approach gDNA editing efficiency of 58.5±2.0%, 39.7±5.9% and 62.6±3.1% were observed for the c.5197-557G>T allele with gRNA2, gRNA3 and gRNA4, respectively. As these three sgRNAs overlap with the DIV, low cut efficiency was detected on the wild-type allele (gRNA2: 0.0±2.1%; gRNA3 0.0±4.4%; gRNA4: 10.1±5.5%), implicating decent allelic preference of editing events for the mutant allele. In contrast, gRNA5 does not overlap with the DIV and therefore, as expected, resulted in a similar gDNA editing efficiency at either allele (41.1±15.0% on the mutant allele, and 38.8±14.5% on the wild-type allele).

gDNA editing efficiency of the gRNA4/*Sp*Cas9 combination was also measured at 35-days post electroporation. Similarly, to dgRNA7+8/*Sp*Cas9n, no significant difference was seen when compared to 7-days post transfection (**Figure 39B**).

Table 10: Genomic DNA (gDNA) cut efficiency in patient-derived, cone photoreceptor precursor cells[c.5197-557G/T] using the dual sgRNA/*Sp*Cas9 (dgRNAs/*Sp*Cas9), the dual gRNAs/*Sp*Cas9-nickase (dgRNAs/*Sp*Cas9n) and (single) sgRNA/*Sp*Cas9 approaches.

	Mean edited allele frequency [%]		SD [%]	
dgRNAs/*Sp*Cas9				
gRNA7+8	78.2		3.8	
gRNA9+6	78.6		8.8	
gRNA9+6_35dPE	42.1		6.4	
Controls				
Mock Cas9	0.0		0.0	
Mock Cas9_35dPE	0.0		0.0	
NT-CPCs	0.0		0.0	
NT-CPCs_35dPE	0.0		0.0	
	Allele T	Allele G	Allele T	Allele G
dgRNAs/*Sp*Cas9n				
gRNA10+5	40.2	28.7	6.2	7.4
gRNA10+6	34.4	28.7	2.2	5.2
gRNA10+6_35dPE	33.5	29.8	2.8	3.6
sgRNA/*Sp*Cas9				
gRNA2	58.5	0.0	2.0	2.1
gRNA3	39.7	0.0	5.9	4.4
gRNA4	62.6	10.1	3.1	5.5
gRNA4_35dPE	55.7	9.4	2.9	5.6
gRNA5	41.1	38.8	15.0	14.5
Controls				
Mock Cas9	1.7	0.0	1.8	1.7
Mock Cas9_35dPE	2.4	0.0	2.1	1.9
NT-CPCs	0.4	0.0	5.0	4.6
NT-CPCs_35dPE	0.0	0.5	1.5	1.5

6.10 PRELIMINARY ASSESSMENT OF P53 ACTIVATION UPON GRNA7+8/*SP*CAS9, GRNA10+6/*SP*CAS9N AND GRNA4/*SP*CAS9 TREATMENT IN CONE PHOTORECEPTOR PRECURSOR CELLS

In an attempt to unravel the molecular cause of the observed drop in the percentage of edited cells treated with gRNA7+8/*Sp*Cas9 between 7-days and 35-days post transfection, we considered to profile the available samples for TP53 expression. In literature, it is extensively demonstrated that the formation of CRISPR/Cas-mediated DSBs triggers the activation of the TP53 pathway which is involved in regulating cell growth, apoptosis, as well as senescence (Álvarez et al. 2022, Enache et al. 2020, Haapaniemi et al. 2018, Jiang et al. 2021). The grade of TP53 activation primarily depends on the number of generated DSBs. Therefore, we speculated that the dgRNAs/*Sp*Cas9 approach would result in higher TP53 activation due to two induced DSBs per allele (four in total) compared to single DSBs of the sgRNA/*Sp*Cas9 approach or single-stranded breaks determined by the use of *Sp*Cas9n. Since TP53 is predominantly regulated at the protein level through phosphorylation (Maclaine et al. 2009), we investigated the increase of *P21* transcript level, a direct TP53 downstream target protein involved in cell cycle arrest. The results refer to 48h post transfection samples and are normalized to non-transfected CPCs. *GUSB* was used as a housekeeping gene.
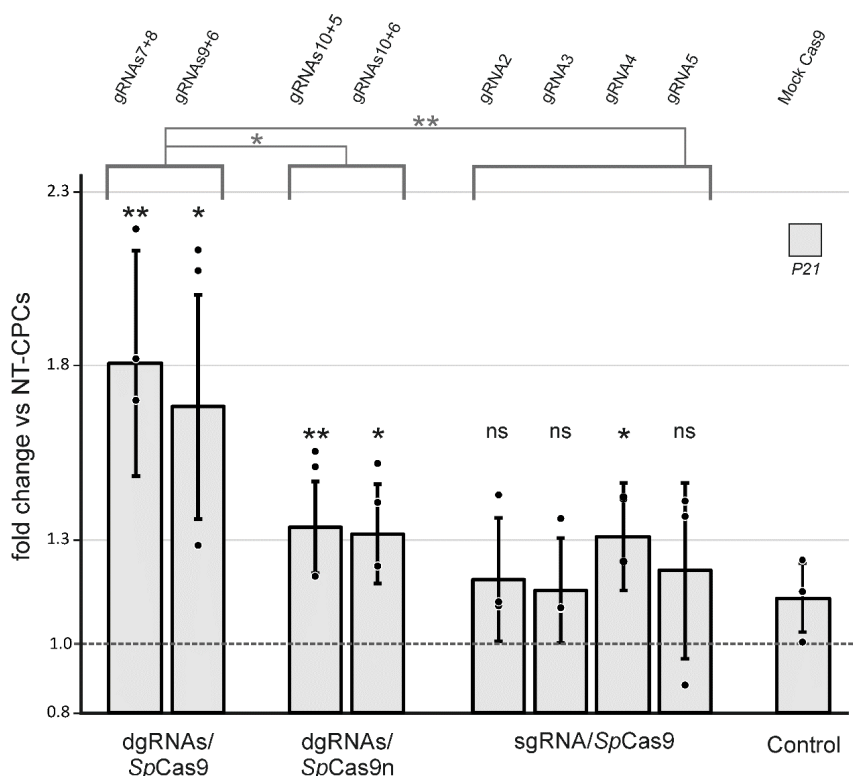
Figure 40: Quantitative PCR analysis for *P21* in patient-derived, heterozygous c.5197-557G>T cone photoreceptor precursor cells (CPCs^c.5197-557G/T) 48h post treatment with the dual sgRNA/*Sp*Cas9 (dgRNAs/*Sp*Cas9), the dual gRNAs/*Sp*Cas9-nickase (dgRNAs/*Sp*Cas9n) and (single) sgRNA/*Sp*Cas9 approaches. Data are normalized to *GUSB* and expressed as -fold change compared to non-transfected NT-CPCs. Statistically significant differences between Mock Cas9 (*Sp*Cas9 + scramble sgRNA) and the (s)gRNA(s)/*Sp*Cas9(n) combinations as well as between the three approaches are expressed as* = p≤0.05, ** = p≤0.01, and ns = non-significant. Results are presented as mean ± standard deviation (n= 3).

Using the dgRNAs/*Sp*Cas9 approach, an increase by 1.81±0.33 and 1.68±0.32 -fold change of *P21* expression for gRNA7+8 and gRNA9+6, respectively, was detected in comparison to non-transfected CPCs (**Figure 40**). The use of the gRNA10+5/*Sp*Cas9n combination induced 1.34±0.13-fold change in *P21* expression, similarly to the gRNA10+6/*Sp*Cas9 combination that led to a 1.32±0.14-fold increase. gRNA2, gRNA3, gRNA4 and gRNA5 of the sgRNA/*Sp*Cas9 approach resulted in 1.31±0.15, 1.19±0.18, 1.15±0.15 and 1.21±0.25-fold change in *P21* level, respectively. Ultimately, the transfection of Mock Cas9 induced a similarily small 1.13±0.10--fold change in *P21* expression. Since the overexpression of the editing system itself can affect the activation of the P53 pathway, the evaluation of statistically significant increases in *P21* expression were calculated in comparison to the Mock Cas9 control. Significant differences were observed for gRNA7+8/ and gRNA9+6/*Sp*Cas9, and gRNA10+5/ and gRNA10+6/*Sp*Cas9n, as well as for gRNA4/*Sp*Cas9.

When comparing the potential of the three different approaches in activating *P21*, dgRNAs/*Sp*Cas9 triggered a statistically significant higher induction compared to dgRNAs/*Sp*Cas9n or sgRNA/*Sp*Cas9.

6.11 GRNA/ED*SP*CAS9 SPLICING RESCUE ENHANCEMENT WITH GRNA2 AND GRNA3 IN CONE PHOTORECEPTOR PRECURSOR CELLS

With the aim to validate whether the higher splicing rescue obtained at the minigene level with ED*Sp*Cas9 coupled to gRNA2 and gRNA3 (**Figure 29**) compared to the standard sgRNA/*Sp*Cas9 approach was retained also when targeting the *ABCA4* endogenous locus, gRNA2/ED*Sp*Cas9 and gRNA3/ED*Sp*Cas9 were tested in CPC[c.5197-557G/T]. gRNA4 was not chosen as the cut site is only located 3 nt upstream of c.5197-557, therefore the full potential of the ED*Sp*Cas9 approach, mediated by the generation of larger deletions aiding the splicing rescue, could not have been fully explored. Indeed, gRNA4 coupled to *Sp*Cas9 already resulted in high rescue when tested in the minigene assay (**Figure 29**), thereby limiting the room for possible improvement mediated by ED*Sp*Cas9. To recapitulate, the results of the minigene splicing assay for gRNA1/, gRNA5/, and gRNA6/ED*Sp*Cas9 in comparison to /*Sp*Cas9 did not show considerable differences in splicing rescue (**Figure 29**), and we therefore decided to not validate these combinations further in the patient-derived system.

The relative ratio of correctly spliced transcript upon ED*Sp*Cas9 treatment was 94.2±1.9% for gRNA2 and 93.3±1.9% for gRNA3, indicating a splicing rescue of aberrant splicing of 79.1±6.8% and 75.5±6.8%, respectively (**Figure 41**). In comparison to *Sp*Cas9, gRNA2 and gRNA3 coupled to ED*Sp*Cas9 achieved nearly double splicing rescue. The cut sites for gRNA2 and gRNA3 are located 6 bp and 15 bp upstream c.5197-557, respectively (**Figure 28**).
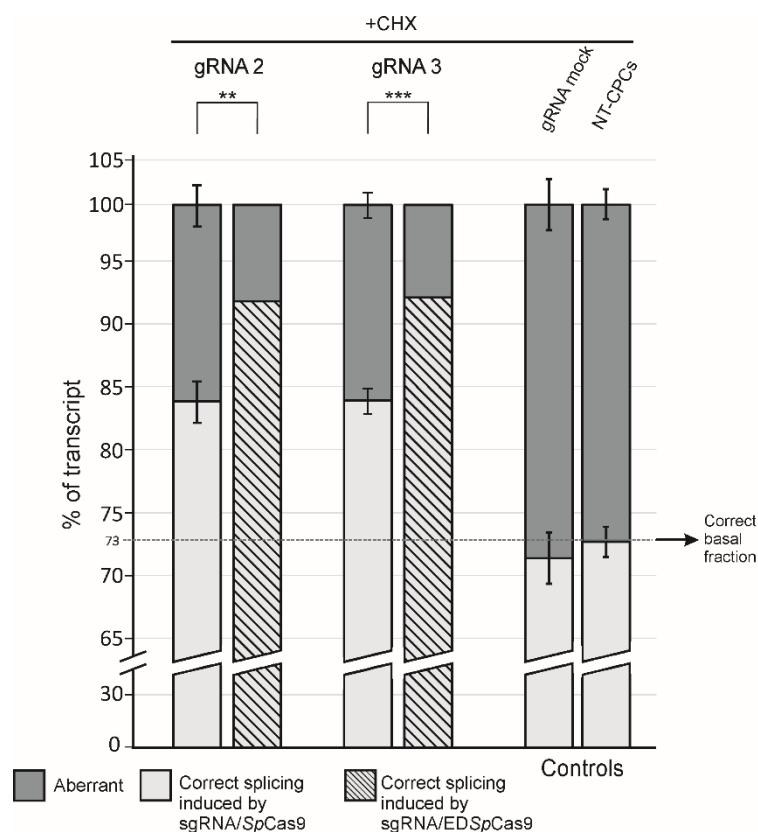
Figure 41: Rescue of the *ABCA4* c.5197-557G>T induced splice defects in patient-derived heterozygous c.5197-557G>T cone photoreceptor precursor cells (CPC^c.5197-557G/T). Electroporated CPC^c.5197-557G/T were analyzed 7-days post electroporation for assessing relative rescue efficacy of the tested sgRNA/*Sp*Cas9 and /ED*Sp*Cas9 approaches. 16 h prior to harvesting, 0.1 mg/ml cycloheximide (CHX) was added. Non-transfected CPCs (NT-CPCs) without (w/o) CHX treatment, NT-CPCs treated with CHX, and cells transfected with a scramble gRNA (gRNA Mock) were used as controls. Relative quantification of aberrant and correct transcripts is shown. Results are presented as mean ± standard deviation (n=3). Statistically significant changes in *ABCA4* % of transcript are expressed as ** = p≤0.01 and *** = p≤0.001.

6.12 PRELIMINARY CHARACTERIZATION OF THE MUTATIONAL PROFILES INDUCED BY GRNA2/ AND GRNA3/ED*Sp*CAS9 IN CONE PHOTORECEPTOR PRECURSOR CELLS

In order to profile the size of the deletion outcomes induced by the ED*Sp*Cas9 approach, high-throughput sequencing was performed on gDNA of CPCs treated with gRNA2/ and gRNA3/ED*Sp*Cas9. Due to the depth of sequencing required for the other samples included in the library being sequenced, only one replicate for gRNA was processed
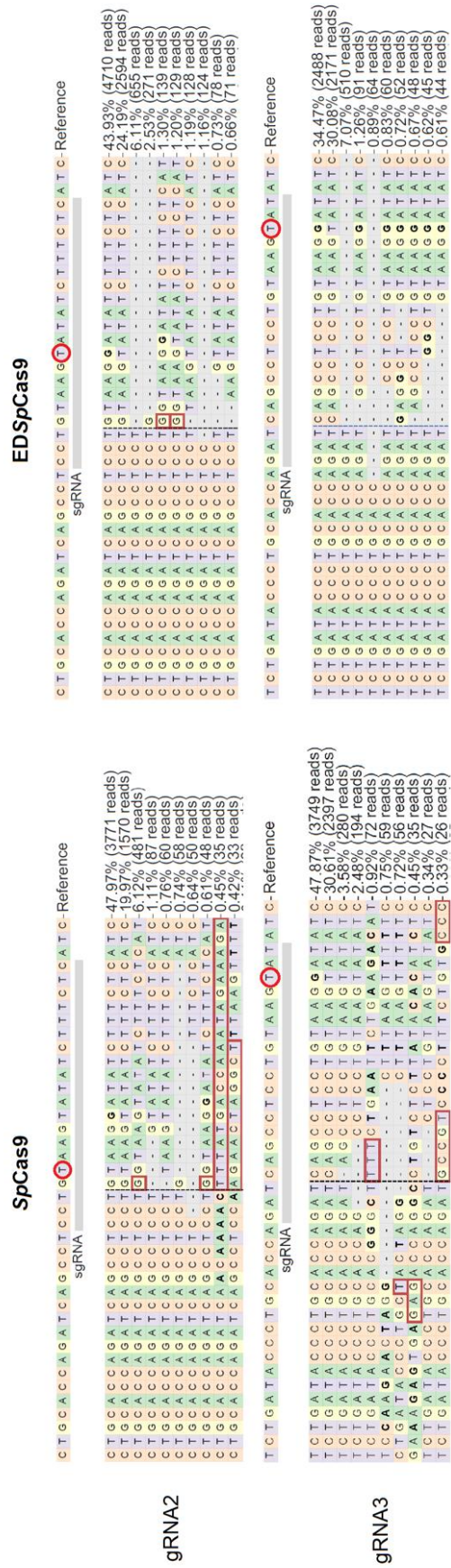
Figure 42: Mutational profile characterization - including frequencies - of the eight most common repair outcomes following gRNA2/SpCas9, gRNA3/SpCas9, gRNA2/EDSpCas9 and gRNA3/EDSpCas9 treatment in patient-derived, heterozygous c.5197-557G>T cone photoreceptor precursor cells (CPCs^{c.5197-557G/T}). CRISPResso2 results are reported for gRNA2/SpCas9 and /EDSpCas9 and gRNA3/SpCas9 and /EDSpCas9. The red circle highlights the variant position. The red square indicates insertions, whereas empty or dashed squares specify deletions. The percentages shown are calculated as fraction of the total number of reads.

While the most prominent repair outcome for gRNA2 and gRNA3 complexed to *Sp*Cas9 was a single nucleotide insertion (6.12% of total reads) on the mutant c.5197-557T allele and a single nucleotide deletion on the mutant c.5197-557T allele (3.58% of total reads), respectively, ED*Sp*Cas9 induced >20 bp deletions in >7% of total reads for the two gRNAs tested. The gDNA cut efficiency on the mutant allele (c.5197-557T) was comparable, indicating that the higher splicing rescue does not depend on superior gDNA cleavage efficacy of the ED*Sp*Cas9 molecule. On the contrary, the editing efficacy achieved on the wild-type allele (c.5197-557G) by ED*Sp*Cas9 – which does not fully match with either of the tested sgRNAs - resulted to be notably higher compared to *Sp*Cas9. Specifically, 43.93% of total reads mapped on the unedited c.5197-557G wild-type allele following gRNA2/ED*Sp*Cas9 treatment, as opposed to 47.97% of total reads for gRNA2/*Sp*Cas9, respectively. Even more marked editing of the non-targeted (wild-type) allele was attained by gRNA3/ED*Sp*Cas9 compared to gRNA3/*Sp*Cas9: 34.47% and 47.87% of total reads mapped on the unedited c.5197-557G, respectively. This finding might highlight enhanced off-target predisposition of the ED*Sp*Cas9 molecule, however, these data are derived from a single biological replicate and will have to be confirmed with further indipendent experiments.
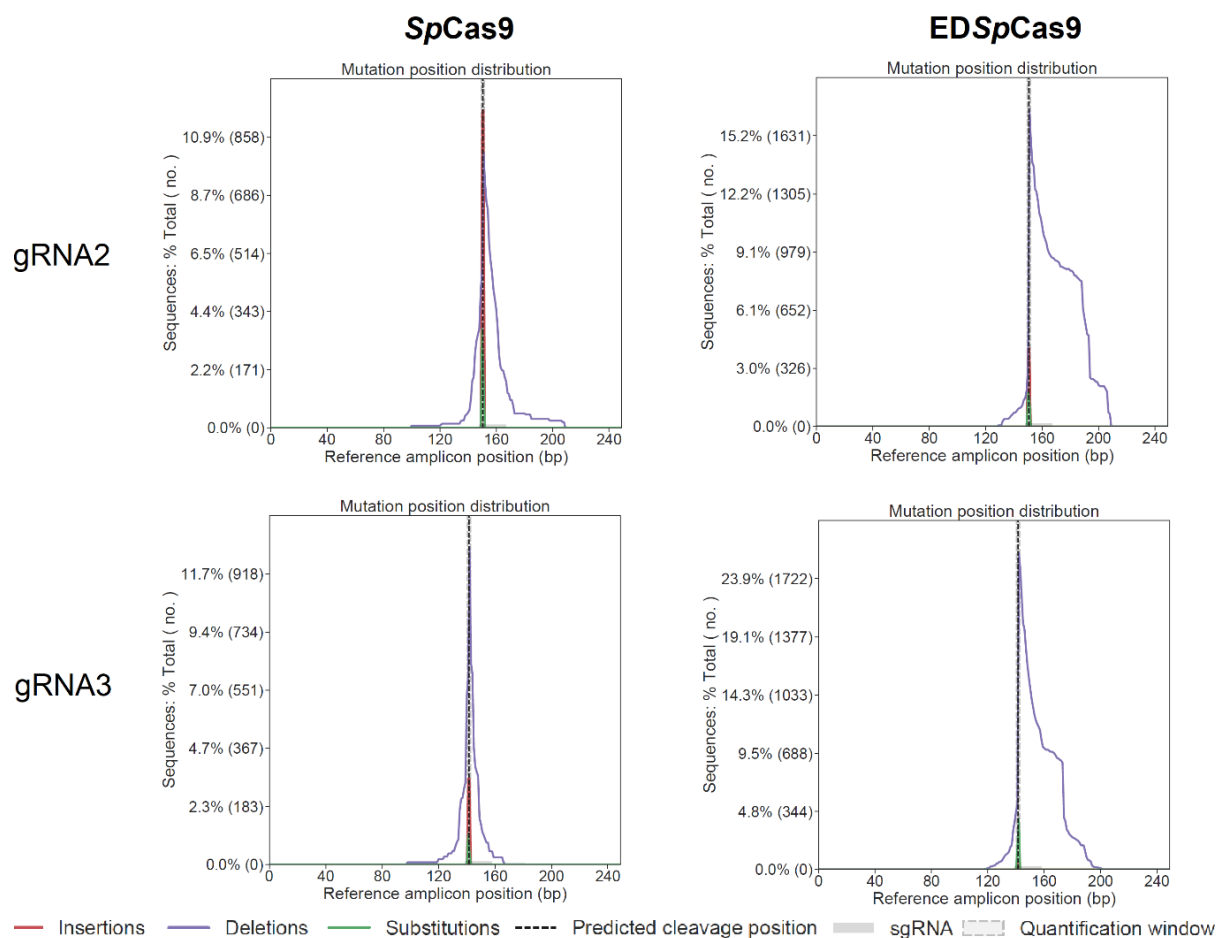
Figure 44: Mutational profile distribution of the editing outcomes following gRNA2/*Sp*Cas9, gRNA3/*Sp*Cas9, gRNA2/ED*Sp*Cas9 and gRNA3/ED*Sp*Cas9 treatment in patient-derived, heterozygous c.5197-557G>T cone photoreceptor precursor cells (CPCs[c.5197-557G/T]). The X axis represents the reference amplicon position in base pairs (bp), whereas the Y axis depicts the relative amount (%) of the specific editing outcomes. The legend box describes how the different repair outcomes are represented on the graph.

In addition to the profiling of the editing outcomes, the mutational profile distribution was plotted on a graph, highlighting the extent and rate of the deletions (violet line) induced by ED*Sp*Cas9 in comparison to *Sp*Cas9. As evident on the graph, ED*Sp*Cas9 induces the formation of enhanced deletions biased towards the 3' end of the target strand at the generated DSB (**Figure 44**). This aspect will be further addressed in the discussion section (Discussion: 7.2.4).

6.13 PRELIMINARY SPLICING RESCUE IN HOMOZYGOUS C.5196+1216C>A CONE PHOTORECEPTOR PRECURSOR CELLS

The established homozygous iPSCs[c.5196+1216A] were differentiated into CPCs[c.5196+1216A] and transfected with gRNAVIII/ and gRNAIX coupled to ED*Sp*Cas9 as well as *Sp*Cas9. The splicing rescue of the splicing defect was analyzed 7-days post transfection. Mock ED*Sp*Cas9 (ED*Sp*Cas9 + scramble sgRNA) and Mock*Sp*Cas9 (*Sp*Cas9 + scramble sgRNA) were used as additional controls to evaluate the impact of the two editing molecules on the analyzed splicing pattern. gRNAVII, being specific for the targeted DIV, was not selected for experiments in CPCs. Non-transfected CPCs[c.5196+1216A] showed 59.8% aberrant splicing. Upon addition of CHX, the percentage of aberrant splicing reached 89.3%, indicating that a proportion of aberrant transcripts is degraded via NMD (**Figure 45**). Upon transfection of the editing plasmids and control, CHX was also added. The gRNAVIII/ED*Sp*Cas9 combination yielded a slightly higher fraction of correct transcript compared to gRNAVIII/*Sp*Cas9. Specifically, 68.7% and 62.4% for ED*Sp*Cas9 and *Sp*Cas9, respectively. Conversely, 30% higher rescue of the correct transcript was relatively obtained when implementing gRNAIX/ED*Sp*Cas9 versus gRNAIX/*Sp*Cas9 (81.6% and 62.6%, respectively). As opposed to what was observed during the minigene splicing assay, the transfection of the Mock controls did not differently effect the c.5196+1216C>A-induced aberrant basal splicing pattern (82.7% Mock ED*Sp*Cas9 and 81.9% Mock *Sp*Cas9 of aberrant transcript).
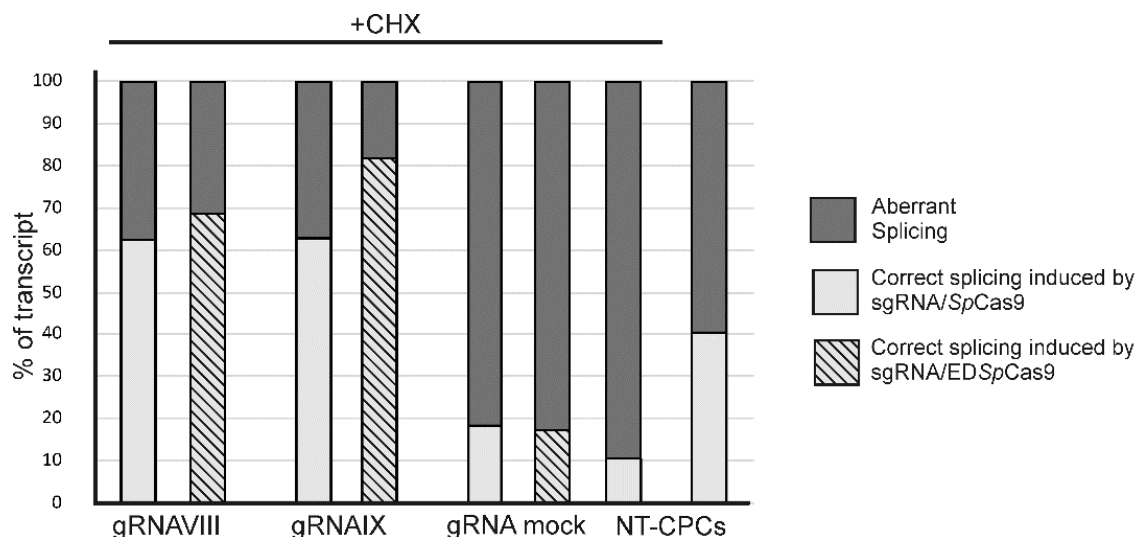
Figure 45: Rescue of *ABCA4* splice defects in homozygous c.5196+1216A cone photoreceptor precursor cells (CPCs[c.5196+1216A]). Electroporated CPCs[c.5196+1216A] were analyzed 7-days post electroporation for assessing relative rescue efficacy of the tested sgRNA/*Sp*Cas9 and /ED*Sp*Cas9 approaches. 16 h prior to harvesting, 0.1 mg/mL cycloheximide (CHX) was added. Non-transfected CPCs (NT-CPCs) without (w/o) CHX treatment, NT-CPCs treated with CHX, and cells transfected with a scramble gRNA (gRNA Mock) were used as controls. Relative quantification of aberrant and correct transcripts is shown. Results are presented as a single experiment results (n=1).

7 DISCUSSION

7.1 STRATEGIES TO ADDRESS DEEP-INTRONIC VARIANTS IN INHERITED RETINAL DISORDERS: WHERE ARE WE?

To date, the most progressed strategies to address DIVs in IRDs are represented by splicing modulation mediated by antisense oligonucleotides (AON). Before their recent decision to suspend and partner the ongoing IRD-focused clinical trials and to shift the company mission on the establishment of RNA-editing platform (https://www.proqr.com/press-releases, August 11, 2022 AT 11:00 AM UTC), the company at the forefront of AON investigation was the Dutch company ProQR (https://www.proqr.com/science-and-pipeline). Part of their decision was due to the unexpected and complete failure of phase 2/3 of their most advanced pipeline addressing the common c.2991+1655A>G DIV in *CEP290* – causative of Leber congenital amaurosis type 10 (LCA10) (https://www.proqr.com/press-releases, February 11, 2022 AT 12:00 PM UTC) –, which also raised concerns in the field and questioned the effectiveness and applicability of the AON approach for IRDs or gene therapy in general. Although it is worth mentioning that the problem lies not only in the question of whether the treatment is effective and applicable, but which time point for treatment can still achieve rescue or halt of the disease, and which clinical endpoints need to be defined and met to assess treatment success (https://www.proqr.com/press-releases, April 12, 2022 AT 11:00 Am UTC).

With the aim to target the very same DIV in *CEP290*, the American company editas MEDICINE established a CRISPR/*Sa*Cas9-based genome editing approach that is currently in phase 1 clinical trial. Pediatric administration has also begun, marking EDIT-101 as the first-ever *in vivo* delivery of an experimental CRISPR gene therapeutics to pediatric patients (https://ir.editasmedicine.com/press-releases, April 11, 2022). Interim results indicated that the therapeutics was well tolerated, however, only minor improvements, that were not in line with the expectations, were seen in two of the three adult patients of the mid dose cohort (XIXth International Symposium on Retinal Degeneration (RD2021), https://ir.editasmedicine.com/press-releases, September 29, 2021).

After these results, speculation can be attempted: LCA10 is a congenital and fast progressing disease. Symptoms appear within the first years of life, resulting in significant vision loss and potential early blindness (Al den Hollander et al. 2006). CEP290 is a centrosomal protein involved in cilia formation, stability and function, thereby assuring photoreceptor outer segment integrity (Sayer et al. 2006). Therefore, therapeutic interventions aiming to target remaining functional cells have to consider the

progression of the disease. Treatment of adult cases may not lead to any beneficial improvement merely because the morphological rescue of the few cells may not be sufficient to achieve a functional response. Furthermore, can the sudden rescue of the CEP290 protein guarantee that the cells are able to let it explicate its function? Cells lacking of this structural protein during genesis and differentiation, therefore never being in contact with such a component, may not be able to integrate it to the place it belongs. In addition, after restoration of all functional components, the visual input may be transmitted to the brain, but the visual cortex may not be capable of processing this input. This may be a result of the too advanced progression of the disease or to defects of crucial structural elements. Ultimately, although functional assessments of LCA10 suggests that degeneration of cone photoreceptors is spared until late in the disease, and thereby the macula stays functional, real-life results from tested therapeutics have hinted at lack of understanding of how therapeutic agents may integrate and solve the underlying pathological mechanisms (Cideciyan et al. 2007, Jacobson et al. 2017).

Conversely, IRDs due to pathological mechanisms involving accumulation of toxic compounds may respond more effectively to gene therapies aiming to resolve the underlying genetic cause. If the onset of the disease is not congenital and the progression of the disease is not overly advanced (i.e. most of the photoreceptor cells are degenerated), the elimination of the pathological insult (e.g. misfolded protein, formation of toxic compounds or byproducts, enhanced activity of an enzyme etc.) would gradually restore a physiological state close to normal, thereby providing halt of degeneration or even restoration of visual input. In this respect, the mutation spectrum of *ABCA4* causative of STGD1 and allied IRDs is ideal for both genome editing and AON splicing modulation approaches, as restoring ABCA4-protein function would stop further accumulation of lipofuscin, and also stop further degeneration. Although no clinical trial has been initiated, the Radboud University Nijmegen (NL) spin-off company ASTHERNA (https://astherna.com/) aims to dose the first STGD1 patients in the next years with their AON therapeutics for *ABCA4* mutations. On the other hand, preclinical investigation of genome editing strategies have not been reported.

7.2   CRISPR/*Sp*CAS9-BASED APPROACHES TO TARGET ISOLATED *ABCA4* DEEP-INTRONIC VARIANTS

Three standard CRISPR/*Sp*Cas9-based approaches (dgRNAs/*Sp*Cas9, dgRNAs/*Sp*Cas9n and sgRNA/*Sp*Cas9) and the novel ED*Sp*Cas9 approach were challenged and compared for the rescue the

isolated c.5197-557G>T DIV. The finding and methods can be applied to the rescue of any DIVs resulting in intron retention (pseudoexon insertion).

7.2.1   Validation tools: Minigene assay and cellular models

Whether composed of minimal (i.e. "minigene", two exons and the intervening intron) or larger (i.e. "midi- / maxigene", several exons and several introns) genomic fragments, minigene splicing assays have been widely used for characterizing the effect of *ABCA4* DIVs, as well as the effect of tailored AON-mediated splicing modulation approaches (Albert et al. 2018, Garanto et al. 2019, Kahn et al. 2020a).

We therefore decided to use minigene splicing assay for preliminary screening of the CRISPR/*Sp*Cas9-based approaches and have successfully used it to validate a number of variants. However, upon transcription and splicing in HEK293T cells, the presence of c.5196+1136C>A and c.5196+1159G>A in the minigene constructs did not result in any aberrant splicing pattern (in line with literature reports (Khan et al. 2020a). Surprisingly, the use of photoreceptor-like (CPCs) or -derived (WERI-Rb1) cell lines for performing the minigene assay did also not induce changes in the splicing pattern, suggesting that c.5196+1136C>A and c.5196+1159G>A may not be pathogenic as predicted or that additional mechanism involving the expression of *ABCA4* may be involved (e.g. distal regulatory elements acting on splicing). What remains to be done to confirm the effect of these two DIVs on splicing is the generation of isogenic iPSCs endogenously harboring the DIVs, their subsequent differentiation in CPCs or retinal organoids and then *ABCA4* splicing patterns evaluation. To support the observation that the genomic context, as well as the cell type play a crucial role in regulating (mis)splicing, we clearly showed the difference in the percentage of aberrantly spliced transcripts between minigene assay and CPCs. Specifically, the minigene assay results for the c.5196+1216C>A DIV showed only minor amounts of aberrantly spliced transcript (**Figure 30**, **Figure 46**). In contrast, cDNA analysis out of RNA extracted from CPCs that had been differentiated from iPSCs harboring the c.5196+1216C>A DIV homozygously at the endogenous *ABCA4* locus provided evidence that most of the *ABCA4* transcript is misspliced in this cellular model, likely reflecting the in vivo situation more closely (**Figure 45**, **Figure 46**).
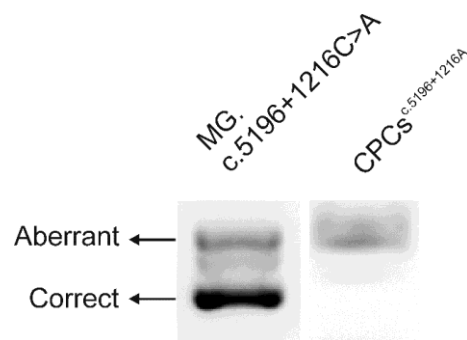
Figure 46: Agarose gel highlighting the different splicing pattern obtained by cDNA analysis of the c.5196+1216C>A minigene assay (MG c.5196+1216C>A) in HEK293T cells and homozygous c.5196+1216C>A cone photoreceptor precursor cells (CPCs[c.5196+1216A]) treated with cycloheximide.

On the whole, the use of a minigene splicing assay allowed us to preliminary screen potential CRISPR/*Sp*Cas9-based splicing rescue, to quickly evaluate strategies and to screen for favorable sgRNAs or sgRNA pairs (Maule et al. 2019, Sanz & Harrison, 2019). Nonetheless, the episomal nature of the plasmid and lack of any chromatin structure limits the minigene splicing assay to testing the effectiveness of introduced indels to correct splicing (Daer et al. 2017). This means that the editing efficiency (intended as the ability to determine genomic breaks) of a given sgRNA/Cas combination may substantially differ from the results obtained in the minigene assay when it is tested at the endogenous genomic locus, thereby affecting the overall efficacy on splicing rescue.

Furthermore, it has been shown that euchromatin is targeted more efficiently by the CRISPR/*Sp*Cas9 system than heterochromatin (Daer et al. 2017, Schep et al. 2021). Therefore, we comparatively evaluated the chosen editing strategies in *ABCA4* expressing CPCs differentiated from a patient-derived iPSC cell line heterozygous for c.5197-557G>T. The implementation of patient-derived cellular models for the assessment of CRISPR/Cas-based therapies represents an important platform for evaluating editing dynamics for mutation-specific approaches targeting rare alleles (Maule et al. 2019, Schwank & Clevers, 2016). While CPCs may serve as surrogate for assessing – for example – the effect of editing approaches in a cellular model endogenously expressing (a) mutant allele(s), they may not completely recapitulate the chromatin structure as *in vivo*, and they do not organize in three-dimensional structures, resembling the retina layer organization, thereby limiting the evaluation of the delivery route of the therapeutic agent. Furthermore, like in the case of *ABCA4*, the level of expression of retina-specific genes may be lower compared to mature photoreceptors which hampers functional read-out on the protein level (e.g. Western blot, immunocytochemistry). The most advanced cellular models available to date, that also overcome the aforementioned limitations encountered with CPCs, are retinal organoids or retina-on-chips linking the retinal organoid to the RPE. Retinal organoids mimic

the three-dimensional structure of the retina and are composed of all cell types of the retina (Meyer et al. 2009, Nakano et al. 2012, Zhong et al. 2014). They allow the validation of therapeutic approaches directly on surrogates of the human retina, harboring the genetic background of interest (e.g. mutations), if patient-derived. Furthermore, the application of such retinal organoids on a microfluidic chip, which also includes a layer of RPE, can simulate the administration of the therapeutic agents through the intravitreal, the subretinal and/or systemic routes (Achberger et al. 2019, Achberger et al. 2022). However, limitations still remain. To better recapitulate the tight interaction that RPE establishes with the retinal layer, the major challenge to overcome is related to the spheroid shape of the retinal organoids that prevents the formation of a continuous interface between the two layers. Major efforts are currently focusing on addressing this issue.

7.2.2   Pros and Cons of the dgRNAs/*Sp*Cas9 approach

Conventionally, the use of two paired sgRNAs targeted to a gDNA area to generate a genomic deletion, including an isolated or clustered DIVs, would represent the most straight forward and probably effective approach (Panagiotopoulos et al. 2020, Ruan et al. 2017). Of note, EDIT-101 applies this design to rescue the splicing defect due to the c.2991+1655A>G DIV in *CEP290* (Maeder et al. 2019). Such design leaves freedom in sgRNA selection, thereby increasing the chance of identifying highly specific and effective paired sgRNAs molecules. The two dgRNAs/*Sp*Cas9 strategies that we designed to target the isolated c.5197-557G>T DIV showed substantial splicing rescue in CPCs (gRNA7+8: 67.1±7.8 % and gRNA9+6: 67.7±12.8 %), demonstrating that the deletion of the pseudoexon sequence and the DIV is able to rescue correct splicing. However, the percentage of gRNAs7+8/*Sp*Cas9-edited CPCs decreased by 50% between 7-days and 35-days post electroporation (78.2±3.0 % vs 42.1±6.4%, respectively), indicating that edited cells died or entered senescence.

Off-target-related genotoxicity of either gRNA7 or gRNA8 may be a tempting argument to explain this finding, however CPCs analyzed 7-days post electroporation did not show signs of enhanced cell death (e.g. transfection efficiency comparable to the other strategies and high percentage of edited cell population, (**Figure 16**,**Figure 36**). Considering that the expression of EGFP, and thereby *Sp*Cas9, was detected as soon as 20 hours post electroporation, and that the cells repair Cas9-mediated DSBs with an average half-life of <10 hours, genotoxicity resulting in cell death would have already been evident at 7-days post electroporation (Brinkman et al. 2018, Joung et al. 2017). Therefore, the drop of the percentage of the edited cell population suggests, to a greater extent, that a fraction of edited cells entered cellular senescence due to response to DNA damage by the DSBs (Hayflick, 1965).

Of note, CPCs are considered immature cells and can still divide, yet slowly. This finding may not be unexpected if considering that the dgRNAs/*Sp*Cas9 approach triggers four DSBs (two per allele) – in comparison to one DSB achieved by the sgRNA approaches using our gRNA2, gRNA3 or gRNA4/*Sp*Cas9 strategies. Thus the dgRNAs/*Sp*Cas9 approach might be more potent in activating the TP53 pathway involved in regulating cell division in response to DNA damage (Enache et al. 2020, Haapaniemi et al. 2018, Hayflick, 1965, Ihry et al. 2018, van den Berg et al. 2018). The induction of TP53 can hamper the growth of Cas9-edited cells and determine positive selective pressure for cells with pre-existing *TP53* mutations, thereby, potentially inducing tumorigenesis (Leibowitz et al. 2021, Sinha et al. 2021).

Although postmitotic cells, like photoreceptors, are terminally differentiated, thus non-dividing cells, the DNA damage and cellular stress induced by several DSBs may still lead to senescence, and this has already been shown for post-mitotic neurons (Jurk et al. 2021, Sedelnikova et al. 2004, von Zglinicki et al. 2020). In an attempt to elucidate the TP53 link to the dgRNA/*Sp*Cas9 approach and the generation of several DSBs, the relative quantification of *P21* expression at 7-days post electroporation revealed a marked increase of *P21* transcripts for the two dgRNAs/*Sp*Cas9 strategies in comparison to the dgRNAs/*Sp*Cas9n and sgRNA/*Sp*Cas9 strategies, therefore supporting the aforementioned line of argument. Nonetheless, further studies are warranted to substantiate these findings, including time course experiments to monitor *P21* expression over time and testing other TP53 downstream targets for independent validation and confirmation to our preliminary findings. In addition, this hypothesis should also be tested in additional dgRNAs/*Sp*Cas9 combinations and targets.

7.2.3 Pro and Cons of the dgRNAs/*Sp*Cas9n approach

To the best of our knowledge, the use of paired Cas nickases to address DIVs has not yet been described in literature. We showed that dgRNAs/*Sp*Cas9n is a practicable approach to achieve substantial splicing rescue of the isolated DIV c.5197-557G>T in CPCs. The advantages of the nickase approach over the standard dgRNAs/*Sp*Cas9 lies on the formation of two opposite single-strand DNA breaks that results in staggered DNA ends and not a DSB. As a consequence, chromosomal instability due to DSBs is not induced, nor abnormal activation of the TP53 pathway (**Figure 40**). Furthermore, the likelihood of the nickase system to target another area of the genome is low as it would need to simultaneously bind two sequence sites defined by the sgRNA pair that are close enough to result in a genomic deletion, thereby reducing the chance of off-target edits (Ran et al. 2013). In contrast, the need of two gRNAs, specifically configured to span from 20 to 80 bp around the desired target and in a specific orientation (i.e. PAM out), as well as the overall lower efficacy of the nickase system, limits its applicability and

scaling up potential to therapeutics. Additionally, it is worth mentioning that the implementation of the dgRNAs/*Sp*Cas9n approach often does not enable allele-specific targeting, therefore extending the editing at both alleles which, depending on the cell line used, may result in recombination between the homologous chromosomes (Brunner et al. 2019).

7.2.4   Promises of the sgRNA/*Sp*Cas9 and further development of the sgRNA/ED*Sp*Cas9 approach

Having described the dgRNAs/*Sp*Cas9 and dgRNAs/*Sp*Cas9n approaches, it is clear that these are not minimalistic approaches and they can easily perturbate genomic sequences beside the intended target sequence. Furthermore, both approaches require the co-delivery of two different sgRNAs, thereby increasing the number of elements required for the system to work and the size of the therapeutic construct to be delivered. The implementation of the sgRNA in the sgRNA/*Sp*Cas9 approach solves most of the described limitations associated with the dgRNAs/*Sp*Cas9 and dgRNAs/SpCas9n approaches. The main advantageous feature of the sgRNA/*Sp*Cas9 approach are:

- There is no generation of multiple DSBs, thereby reducing the chance of large chromosomal rearrangements (Nahmad et al. 2022, Stadtmauer et al. 2020).
- The use of a sgRNA conceivably reduces the chance of off-target events.
- The sgRNA/*Sp*Cas9 approach can be designed to target only the mutant allele, thereby reducing the chance of recombination between the homologous chromosomes (Brunner et al. 2019).
- The overall size of a polynucleotide molecule required for the delivering is smaller, thereby easing scaling up of prospective therapeutics.

The sgRNA/*Sp*Cas9 approach tested for the isolated c.5197-557G>T DIV determined the highest splicing rescue (gRNA4/*Sp*Cas9: 82.7±5%) in CPCs, showing the feasibility of this strategy, alongside high grade of allele-specific genomic cleavage (62.6±3.1% on the c.5197-557T allele compared to 10.1±5.5% on the wild-type allele). However, the results obtained in CPCs may not be translatable to different cell type (i.e. mature photoreceptors) as the repair mutational profiles are variable across different cell lines, thereby arguably impacting on the overall rescuing capacity seen in CPCs (Allen et al. 2019, Shen et al. 2018).

Furthermore, the sgRNA/*Sp*Cas9 approach does not allow for flexibility in sgRNA design, therefore largely restricting the choice of sgRNAs. Indeed, bioinformatic analysis returned 19 potential off-target sequences having three mismatches for gRNA4, while only five were identified for gRNA2 (**Table 6**). Reasonably, gRNA2 would be more specific than gRNA4, yet at the same time, splicing rescue in CPCs

of the gRNA2/*Sp*Cas9 strategy only achieved 41.5±7.3%, translating in 50% lower rescue potential compared to gRNA4/*Sp*Cas9. To tackle these issues, the gRNA/ED*Sp*Cas9 approach was developed.

By fusing the human Three Prime Repair Exonuclease 2 (TREX2) to *Sp*Cas9, we were able to generate an Enhanced Deletion *Sp*Cas9 variant (ED*Sp*Cas9) that increased the occurrence of large deletions (>20 bp) at the targeted site over small deletions, insertions and single nucleotide changes. As shown in **Figure 44**, displaying the mutational profile following gRNA2/ and gRNA3/ED*Sp*Cas9 treatments, deletions are biased towards the 3'-end of the non-target strand. This is due to the mode of action of ED*Sp*Cas9 at the targeted sequence. *Sp*Cas9 tightly engages with the 3' and 5' generated ends on the targeted strand and with the 5' generated end on the non-target strand, whereas the 3' generated end on the non-target strand is promptly releases from the sgRNA/*Sp*Cas9 complex, rendering it accessible to TREX2 processing. (Richardson et al. 2016, Sternberg et al. 2014) (**Figure 47**). This molecular mechanism explains the occurrence of biased deletions.
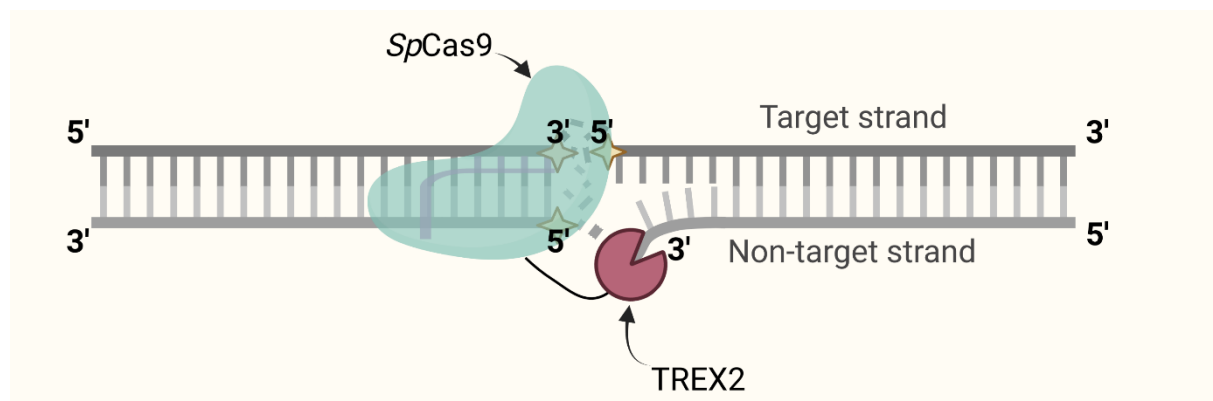
Figure 47: Schematic representation of the possible mode of action of ED*Sp*Cas9 in generating larger deletions. Deletion are biased toward the 3' generated end of the non-target strand as TREX2 binding and processing of the 3' end of the target strand is prevented by *Sp*Cas9.

At the same time, although limited to one single biological replicates, ED*Sp*Cas9 appears to have reduced discrimination between the WT and mutant allele (**Figure 42**). This aspect can be addressed by implementing high-fidelity *Sp*Cas9 variant (e.g. *Sp*Cas9-HF or e*Sp*Cas9) (Kleinstiver et al. 2016, Slaymaker et al. 2016).

The features of the ED*Sp*Cas9 molecule were particularly challenged to restore aberrant splicing due to DIVs. When sequences acting on splicing or the very same DIVs are targeted by sgRNAs coupled to ED*Sp*Cas9 in comparison to standard *Sp*Cas9, the greater grade of sequence perturbation, due to the formation of larger deletions, increases the chance of disrupting the recognition of these *cis* sequence elements by the spliceosome, thereby rescuing the splicing back to normal. In practical terms, coupling

gRNA2 or gRNA3 to ED*Sp*Cas9 results in a splicing rescue of 79.1±6.8% and 75.5±6.8%, respectively, in CPCs, comparable to gRNA4/*Sp*Cas9 (82.7±5.0 %). What remains to be done is to confirm that the overexpression of ED*Sp*Cas9 coupled to a gRNA mock does not affect itself the c.5197-557G>T splicing pattern, as shown for gRNAmock/*Sp*Cas9 (**Figure 41**) and suggested for gRNAmock/ED*Sp*Cas9 when targeting the clustered c.5196+1216C>A DIV (**Figure 45**).

ED*Sp*Cas9 offers increased freedom in sgRNA selection in terms of distance to the DIV and higher splicing rescue potential. Furthermore, Yin and co-workers have recently showed that the implementation of an ED*Sp*Cas9 design reduces the level of Cas9-induced translocations down to the levels of base- or prime-editors (Yin et al. 2022). In the light of these advancements, ED*Sp*Cas9 represents a superior system compared to sgRNA/*Sp*Cas9 to address isolated DIVs, while retaining the advantages of a single-sgRNA approach which includes allele-specific targeting. Furthermore, due to consistent generation of deletions mediated by ED*Sp*Cas9, the repair mutational profiles are less dependent on the cell type, therefore easing the scaling up of the approach to advanced cellular models and *in vivo*.

7.3   NOVEL ED*Sp*CAS9 APPROACH TO TARGET CLUSTERED *ABCA4* DEEP-INTRONIC VARIANTS

Although the generation of two coupled DBSs with the aim to generate a genomic deletion containing all the clustered *ABCA4* DIVs located within one intron would possibly enable splicing rescue, we wanted to explore a novel approach not relying on multiple DSBs and formation of genomic fragments. Therefore, we preliminary applied sgRNA/ED*Sp*Cas9 to the rescue of clustered DIVs in minigene splicing assay. Specifically, ED*Sp*Cas9 was targeted to sequences acting on splicing shared among various clustered DIVs (i.e. cryptic acceptor and donor sites). Interestingly, co-transfection of any of the minigene plasmids for clustered *ABCA4* intron 36 DIVs and mock *Sp*Cas9 or mock ED*Sp*Cas9 resulted in different fractions of correct and aberrant transcript that were not comparable between the ED*Sp*Cas9 and *Sp*Cas9 mock controls (**Table *7***). Nonetheless, such fractions were consistent across replicates and provided meaningful results across the strategies within the same approach (gRNA/*Sp*Cas9 or gRNA/ED*Sp*Cas9) (**Figure 32**).

Results of the minigene splicing assay showed remarkable relative percentage increase of correct *ABCA4* transcript upon sgRNA/ED*Sp*Cas9 treatment compared to sgRNA/*Sp*Cas9, for c.5196+1013A>G (494±94% vs 243±20%, for gRNAXVI) c.5196+1056A>G (160±9% vs 106±3%, for gRNAXVI) and

c.5196+1134C>G (127±1% vs 106±2%, for gRNAIX), outlining the superiority of the sgRNA/ED*Sp*Cas9 approach in the rescue of the clustered *ABCA4* DIVs. Conversely, little or no improvement was seen when targeting c.5196+1137G>A and c.5196+1216C>A with either approach. However, the fraction of aberrant transcript induced by c.5196+1137G>A and c.5196+1216C>A in the minigene splicing assay is minimal, thereby the degree of relative percentage splicing increase is strongly influenced by the background of the correctly-spliced transcript. Indeed, when transfecting homozygous c.5196+1216C>A CPCs with gRNAVIII/ or gRNAIX/ED*Sp*Cas9, better rescue was observed compared to the same sgRNAs complexed to *Sp*Cas9. Additionally, as opposed to the minigene assay results, no substantial difference in the percentage of aberrant transcript was detected between ED*Sp*Cas9 and *Sp*Cas9 when coupled to a gRNA mock. On the other hand, the basal fraction of correct transcript seemed slightly variable between gRNA-mock-transfected CPCs and non-transfected CPCs. This may be due to the overexpression system that affect the transcriptome. Of note, collaborators of ours also noticed the same effect when transducing retinal organoids with mock AAV particles (personal communication, Nuria Suárez Herrera, MSc, Radboudumc). In any case, our preliminary data will have to be confirmed by further biological replicates.

It is worth mentioning that gRNAI and gRNAII, due to an annotation mistake, had PAM sequences (NGT and NTA, respectively) not compatible with *Sp*Cas9 requirements, thereby inhibiting the cleavage at the target sited. On the other hand, gRNAIII possess a NAG PAM sequence, which can be still targeted by *Sp*Cas9 (despite not being the preferential sequence) (Karvelis et al. 2015). For this reason, discreet rescue was observed when coupling gRNAIII to ED*Sp*Cas9 or *Sp*Cas9.

On the whole, the preliminary results obtained through minigene splicing assay for the rescue of clustered *ABCA4* DIVs showed considerable increase in splicing rescue mediated by sgRNA/ED*Sp*Cas9 compared to sgRNA/*Sp*Cas9, providing initial promising scope for further validation in CPCs and possibly retinal organoids.

7.4   FUTURE DIRECTIONS

The sgRNA/ED*Sp*Cas9 results obtained via minigene splicing assay of the intron 36 clustered DIVs will be validated in CPCs as reported for the isolated c.5197-557G>T DIV. gDNA mutational repair outcomes in CPCs both for the isolated c.5197-557G>T and the clustered DIVs will be further analyzed. In this respect, iPSCs harboring the DIVs of interest were already established within the scope of this thesis.

In addition, features of the sgRNA/ED*Sp*Cas9 approach will be further improved. In particular, smaller Cas9 orthologous will be used to replace *Sp*Cas9, resulting in a smaller EDCas9 variant potentially deliverable via AAV particles. In this respect, the recently described small synthetic RNA-guided nuclease 3.1 (sRgn3.1) will represent the first choice (Schmidt et al. 2021). Upon validation in CPCs, the most promising sgRNA/EDsRgn3.1 strategy will be delivered via AAV particles in homozygous c.5196+1013A>G retinal organoids. Since the fraction of aberrant splicing induced by the DIV c.5196+1013A>G is largely prevailing (**Table 7**), we expect to see a clear *ABCA4* protein rescue upon AAV treatment. In parallel, we will focus on the assessment of off-target editing mediated by EDsRgn3.1.

7.5 FORESEEABLE IMPROVEMENTS OF THE ED*Sp*CAS9 EDITING SYSTEM

As mentioned in the previous chapter, the improvement effort will primarily focus on generating an enhanced deletion RNA-guided endonuclease editing system that can fit in AAV particles (< 4.8 kb). Additionally, although current advanced genome editing tools based on the features of Cas9 (e.g. base editors) shows that the activity of the fusion partner (e.g. deaminases) is tightly restricted by the specificity of Cas9 (Gaudelli et al. 2017, Koblan et al. 2018), we want to alter the ability of TREX2 to independently bind the DNA, therefore preventing potential unwanted outcomes. TREX2 residues involved in DNA binding have been characterized (Cheng et al. 2018, Perrino et al. 2005), and it has been demonstrated that their mutagenesis does not affect TREX2 catalysis, thereby supporting the attainability of this improvement without compromising the overall function of the editing system (Yin et al. 2022). Another aspect we will consider is the mutagenesis of TREX2 residues involved in homodimer formation (Cheng et al. 2018, Perrino et al. 2005). Since TREX2 homodimer formation (in the ED*Sp*Cas9 molecule) may lead to aggregation and cellular stress, we will address it by disrupting intermolecular interactions between monomers by targeted residue mutagenesis. Although TREX2 functions as a dimeric unit, when mutating one of the two dimers to retain the binding ability and to disrupt the catalytic activity, the wild-type monomer is still able to process DNA, ableit less efficiently (Perrino et al. 2008). Conversely, when generating a TREX2 mutant not able to bind DNA that forms a dimer with a WT TREX2 monomer, no catalytical activity is observed (Perrino et al. 2008). These findings indicate that the interaction between two TREX2 monomers is essential for DNA binding, but not for DNA processing. Therefore, since the DNA binding ability of ED*Sp*Cas9 is provided by *Sp*Cas9, disrupting the interaction between two TREX2 monomer might still maintain the enhanced-deletion

features of the ED*Sp*Cas9 molecule. Ultimately, we will investigate if the ED system can also work through single-strand breaks (nicks), so as to prevent the formation of DSBs. In particular, we reason that upon the generation of a single nick, TREX2 processes the accessible single DNA strand in direction $3' - 5'$, thereby generating a large single-strand deletion. Although such a deletion should be repaired by DNA polymerase (δ, ε or β), resulting in error-free mutational profiles at the nicked sequence, it has been reported that sgRNA/*Sp*Cas9 nickase treatment may well result in indels formation (Rees et al. 2019). Furthermore, Yin and co-workers also preliminary investigated the feasibility of a nickase coupled to TREX2, showing promising results (Yin et al. 2022). We, therefore, believe that this optimization step may be worth the effort, as it may result in a genome editing tool able to achieve indels formation without the formation of DSBs, and thus no activation of the TP53 pathway (Haapaniemi et al. 2018, Ihry et al. 2018). Eventually, the Enhanced-Deletion RNA-guided endonuclease system could be applied to gene disruption for autosomal dominant inherited diseases, overtaking and replacing the current use of different DSB-inducing Cas9 molecules.

7.6   THE ED*Sp*CAS9 EDITING SYSTEM IN PROSPECTIVE TO OTHER ADVANCED EDITING TOOLS

The rush towards personalized treatments to address genetic conditions finds the greatest hope in genome editing technology. Such technologies enable to permanently change the genetic code of an individual, thereby possibly addressing the genetic cause of a known condition. Due to the ease with which CRISPR/Cas molecules can be programmed to target almost any genomic sequence, an increasing number of laboratories and companies are implementing them as compared to older genome editing technologies such as meganucleases and zinc finger nucleases. Since 2013, when the CRISPR/*Sp*Cas9 system was used for the first time to edit mammalian cells (Jinek et al. 2013), a number of additional editing tools exploiting the unique features of the Cas proteins have been established (Anzalone et al. 2019, Gaudelli et al. 2017, Komor et al. 2016, Komor et al. 2017). These include base- and prime editors. While the wild-type Cas9 molecules induce the generation of DSBs, possibly resulting in chromosomal aberrations and TP53 activation, base- and prime editors work through the installations of single-strand breaks, thereby largely preventing these issues and making the editing tools safer. Furthermore, to place these advanced editing tools in the framework of this thesis, both base- and prime editors may be applied to the rescue of isolated and clustered DIVs. However, beside the discussed superiority of the EDCas design compared to wild-type Cas molecules, argumentations highlighting the advantages of using our EDCas platform compared to base- as well as prime editing can also be put forward.

7.6.1 Comparing EDCas approaches and base editors

Similar to our EDCas approach, the activity of the base editors can be harnessed to perturbate the identity of sequences involved in missplicing, thereby leading to splice correction. In particular, adenosine base editors could be targeted to cryptic acceptor splice sites – composed of the "AG" dinucleotides – resulting in "GG" conversion, or to a cryptic donor site – composed of the "GT" dinucleotide – inducing a "GC" change. Similarly, cytidine base editors can catalyze the conversion of the "AG" and "GT" splice signals in "AA" or "AT", respectively. Although some of the mentioned changes could lead to the generation of rare, non-canonical splice sites, none of the resulting dinucleotide changes would likely be strongly recognized as splicing signals, thereby likely resulting in predominant splicing correction (Thanaraj & Clark, 2001). However, to induce the desired conversion, the targeted nucleotide must be found within the activity window. This varies depending on the base editors, yet it is always found within the sgRNA sequence, specifically from 3 bp to 14 bp (maximum) downstream the 5'-end of a gRNA (Porto et al. 2020). In practical terms, when comparing ED*Sp*Cas9 to *Sp*Cas9-based base editor approaches, a common strategy implementing base editors targeting the clustered c.5196+1134C>G, c.5196+1137G>A and c.5196+1216C>A DIVs cannot be designed, because no suitable sgRNA with an editing window including the common cryptic donor site can be designed. Whereas, to target the clustered c.5196+1013A>G and c.5196+1056A>G DIVs only three sgRNAs could be designed and tested compared to four for the ED*Sp*Cas9 approach. As for the isolated c.5197-557G>T, where mutation-specific sgRNAs were validated for the ED*Sp*Cas9 approach, only one single suitable sgRNA (corresponding to gRNA3) could be tested to convert the "GT" cryptic donor splice site into "GC" or "AT". Additionally, it would not be possible to convert the mutant c.5197-557T allele back to the wild-type c.5197-557G one, using base editors (**Figure 48**).
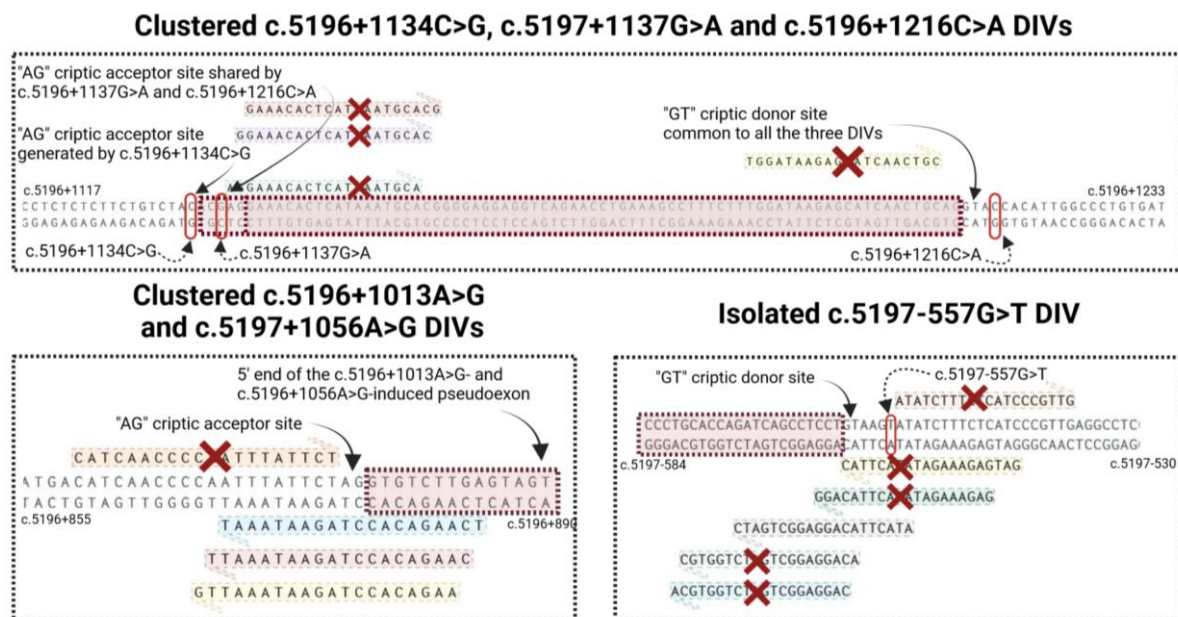
Figure 48 Graphical representation of sgRNAs used in the sgRNA/ED*Sp*Cas9 approach targeting the different clustered deep-intronic variants (DIVs) and the isolated DIV. The red crosses indicate sgRNAs that cannot be implemented by base editors to address the specified DIVs.

7.6.2 Comparing EDCas approaches with prime editing

Beside the four transition mutations (C>T, G>A, A>G, and T>C) also addressable by base editors, prime editors can efficiently perform the eight transversion mutations (C>A, C>G, G>C, G>T, A>C, A>T, T>A, and T>G), as well as targeted deletions and insertions, therefore expanding the scope and capabilities of genome editing by far (Anzalone et al. 2019). Nonetheless, the number of sgRNAs that could potentially be tested for the rescue of the clustered DIVs would only be two (corresponding to gRNAXVI and gRNAIX). Additionally, if implementing gRNAXVI, the primer binding sequence would be matching with a region with low -GC- content, thereby hampering the overall editing efficiency (Anzalone et al. 2019). As for the isolated c.5197-557G>T DIV, no gRNA could be designed to directly result into the c.5197-557T>G transversion, whereas only one sgRNA (corresponding to gRNA4) could be tested to target and disrupt the "GT" sequence of the cryptic donor site (**Figure 49**). Of note, the implementation of prime editors involves additional elements (i.e. primer-binding site and RT template), alongside a gRNA, requiring further optimization. This translates in more laborious and demanding validation experiments.
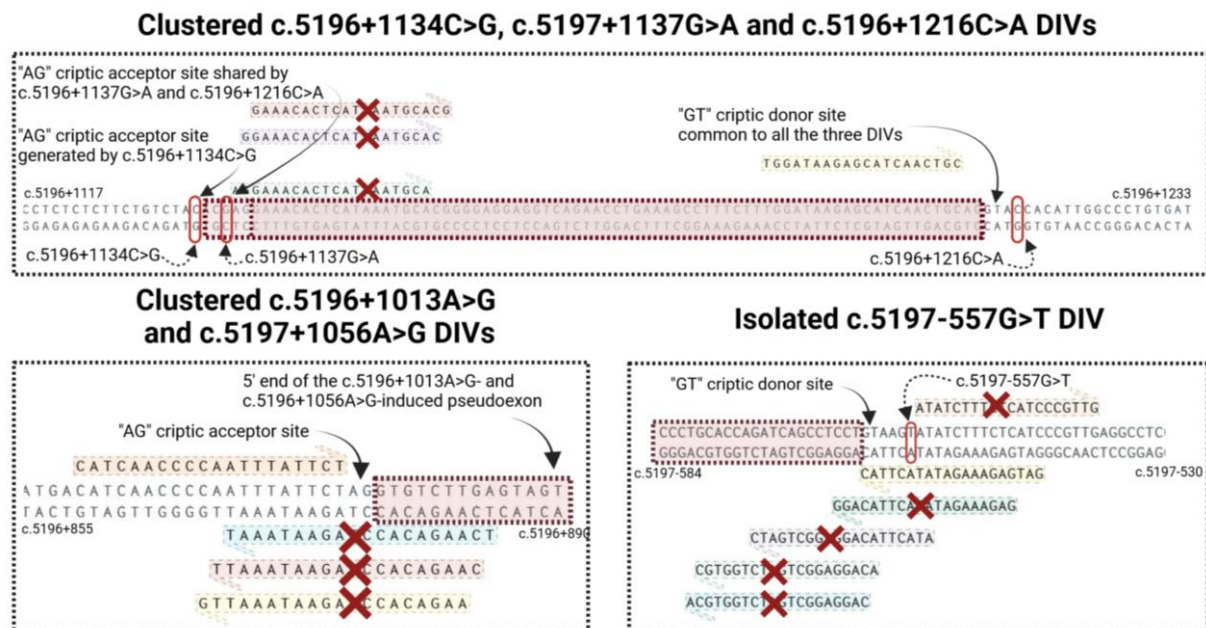
Figure 49 Graphical representation of sgRNAs used in the sgRNA/ED*Sp*Cas9 approach targeting the different clustered deep-intronic variants (DIVs) and the isolated DIV. The red crosses indicate sgRNAs that cannot be implemented by prime editors to address the specified DIVs.

7.7 CONCLUSION AND OUTLOOK

For the first time, this project proved the CRISPR/*Sp*Cas9-mediated rescue of the splicing defect due to the isolated *ABCA4* DIV c.5197-557G>T. Three standard CRISPR/*Sp*Cas9 approaches were compared for their editing efficacy. In addition, a novel *Sp*Cas9 variant, namely ED*Sp*Cas9 was developed, and results compared to the standard approaches. All genome editing approaches were evaluated in minigene assays, HEK293T cells and patient-derived CPCs. Editing events were evaluated by fragment analysis and high-throughput sequencing.

In parallel, lentiviral transduction of CPCs was attempted. Despite the positive outcome, a more efficient transfection protocol for CPCs based on electroporation was established and used for experimental validation in this cell type.

Furthermore, by targeting shared sequences acting on splicing, the potential of the sgRNA/ED*Sp*Cas9 approach was preliminary assessed in minigene assay to rescue aberrant splicing due to clustered DIVs in intron 36 of *ABCA4*, hereby addressing several mutations by one single approach. Such an approach would potentially enable treatment of a larger set of patients with different mutations by the same therapeutic agent. The promising results obtained encouraged us to generate iPSC lines for the

validation in CPCs and potentially retinal organoids. Further optimization experiments of the ED*Sp*Cas9 molecule and the generation of other EDCas molecules are planned. Moreover, the sgRNA/ED*Sp*Cas9 approach is now being tested for other targets, including a common DIV in *USH2A* and for an allele-specific, mutation-independent approach to treat *RHO*-linked IRD. Ultimately, this work paves the way towards permanent correction of DIVs in *ABCA4* by applying genome editing.

8 ANNEX

Annex 1: Characterization reports for iPSC^{c.5197-557G/T}.

Stem Cell Technology Center Genetica

Certificate of Analysis 2020

Invoice number: SCTC2020-00048

Name investigator: Susanne Kohl
Cell line number: IPS20-00052
Project name: ABCA4 c.5197-557G>T

Table 1: Information on the reprogrammed cell line

Information cell line:	
Product description	Fibroblasts transduced with Sendai virus containing the genes OCT3/4, SOX2, KLF4, L-MYC LIN28
Parental cell line	CL20-00007 (ZD574)
Parental cell type	Fibroblasts
Diagnosis	STGD1
Mutation	N/A*
Number of clones	3
Passage (P) of iPSCs reported at submission	P10 (clone 1), P12 (clone 2 and 3)
Culture medium	Essential 8 Flex medium
Culture coating	Matrigel
Feeders during reprogramming	Mouse Embryonic Fibroblasts (MEFs)
Passage method	0.5 mM EDTA
Protocols in Q-portal	046588; 046591

Table 2: Information on the characterization of the reprogrammed cell line

Test description:	Test method:	Test specification:	Result:
Activation of stem cell markers	qPCR	Upregulation of *SOX2, LIN28, NANOG, DNMT3B* in iPSCs compared with fibroblasts	Pass
Expression of stem cell markers	Immunocytochemistry	Expression of OCT4, NANOG, SSEA4, TRA-1-81	Pass
Mycoplasma	PCR	Negative	Pass
Three lineage differentiation	Differentiation assay	Upregulation of germlayer-specific genes	Pass

*N/A: Not applicable

Karyotype: 46,XY
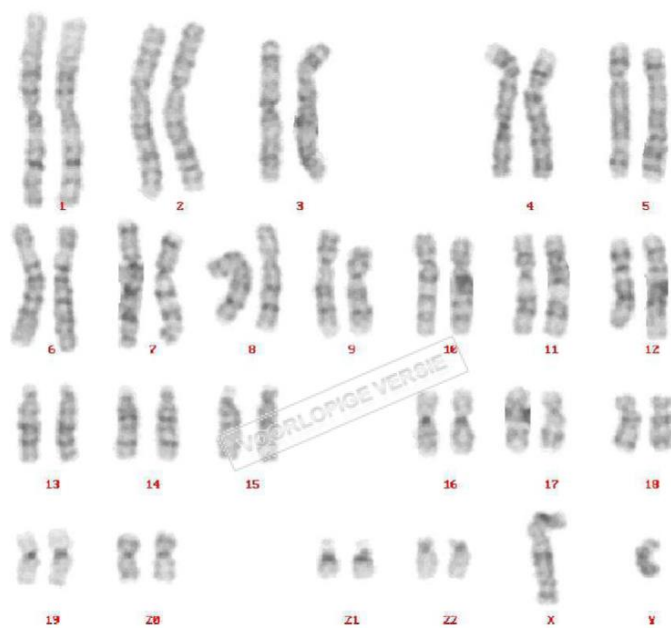
Annex 2: Off-target sites as predicted from Cas-OFFinder. Chromosomal locations are detailed.

gRNA	1 nt Off-target location		2 nt Off-target location(s)	
	Chromosomal location	Annotation	Chromosomal location	Annotation
gRNA1			chr4:89256569	Intronic (*GPRIN3*)
			chr16:81638729	Intronic (*CMIP*)
gRNA2	chr1:94016408	Intronic (*ABCA4*)		
gRNA3	chr1:94016399	Intronic (*ABCA4*)	chr3:11998025	Intergenic
gRNA4	chr1:94016396	Intronic (*ABCA4*)		
gRNA5				
gRNA6			chr16:997843	Intronic (ENSG00000260496.3)
gRNA7			chr20:24603377	Intronic (*SYNDG1*)
gRNA8				
gRNA9			chrX:143448209	intergenic (ENSG00000227303.1)
			chr6:118282735	Intronic (*SLC35F1*)
gRNA10			chr11:48127160	Intergenic
gRNA11			chr4:57649381	intronic (ENSG00000250333.2)
gRNA12	chr1:174933189	Intronic (*RABGAP1L*)	chr8:140200144	intronic (*TRAPPC9*)
			chr15:53736926	Intronic (*WDR72*)
			chr5:50436324	Intronic (*EMB*)
			chr1:121524127	intergenic
			chr22:19191483	Intronic (*CLRCL1*)
			chr2:23776267	Intronic (*ATAD2B*)
			chr2:72164635	intergenic
			chr2:183468391	Intronic (ENSG00000287621.1)
			chr4:61121179	intergenic
			chr4:187074516	intergenic

			chr17:50228625	intergenic
			chr9:99908948	Intronic (*STX17*)
			chr9:109625411	intergenic
			chrX:97677510	intergenic
			chr14:31187688	Intronic (*HECTD1*)
			chr14:45274931	intergenic
			chr6:157797303	intergenic
			chr11:76961229	Intronic (*ACER3*)
			chr18:27917942	intergenic
gRNA13			chr8:91846266	Intronic (ENSG00000253901.1)
			chr21:33131034	intergenic
			chr4:53650480	intergenic
			chr6:4580806	Intronic (ENSG00000282278.1)
			chr3:97808549	Intergenic (ENSG00000282527.2)
gRNA14			chr11:14349564	Intronic (*RRAS2*)
			chr18:10480867	Intronic (*APCDD1*)
gRNAI				
gRNAII			chr2:79076734	intergenic
			chr3:74805122	intergenic
gRNAIII				
gRNAVIIa	chr1:94018441	Intronic (*ABCA4*)		
gRNAVIIb	chr1:94018441	Intronic (*ABCA4*)	chr9:131847491	intergenic
			chr10:116186629	intergenic
gRNAVIIwt			chr9:131847491	intergenic
gRNAVIII				
gRNAIX			chr2:12060775	Long non-coding RNA (MIR3681HG)
gRNAXVI			chr10:110828343	Intronic (RBM20)

Annex 3: Relative splicing analysis of cone photoreceptor precursor cells[c.5197-557G/T] **(48-hours post electroporation).** The rescue of the mutant splicing (%) is calculated normalizing to the mean of aberrant transcript (%) of NT-CPCs + CHX. SD = standard deviation, NTC = non-transfected control, CHX = cycloheximide. Number of biological replicates = 2.

	Mean of correct transcript [%]	SD [%]
dgRNAs/*Sp*Cas9		
gRNAs7+8+CHX	82.5	5.5
gRNAs9+6+CHX	81.0	4.0
dgRNAs/*Sp*Cas9n		
gRNAs10+5+CHX	76.0	0.0
gRNAs10+6+CHX	65.0	7.0
sgRNA/*Sp*Cas9		
gRNA2+CHX	78.5	5.0
gRNA3+CHX	79.0	3.5
gRNA4+CHX	84.0	1.0
gRNA5+CHX	73.5	6.5
Controls		
Mock Cas9+CHX	69.5	4.5
NT-CPCs + CHX	71.0	3.0
NT-CPCs w/o CHX	███	███

9 BIBLIOGRAPHY

- Achberger, Kevin, et al. "Human stem cell-based retina on chip as new translational model for validation of AAV retinal gene therapy vectors." Stem cell reports 16.9 (2021): 2242-2256.

- Achberger, Kevin, et al. "Merging organoid and organ-on-a-chip technology to generate complex multi-layer tissue models in a human retina-on-a-chip platform." Elife 8 (2019): e46188.

- Adli, Mazhar. "The CRISPR tool kit for genome editing and beyond." Nature communications 9.1 (2018): 1-13.

- Albert, Silvia, et al. "Identification and rescue of splice defects caused by two neighboring deep-intronic ABCA4 mutations underlying Stargardt disease." The American Journal of Human Genetics 102.4 (2018): 517-527.

- Allen, Felicity, et al. "Predicting the mutations generated by repair of Cas9-induced double-strand breaks." Nature biotechnology 37.1 (2019): 64-72.

- Allikmets, Rando, et al. "A photoreceptor cell-specific ATP-binding transporter gene (ABCR) is mutated in recessive Starqardt macular dystrophy." Nature genetics 15.3 (1997): 236-246.

- Álvarez, Miguel M., Josep Biayna, and Fran Supek. "TP53-dependent toxicity of CRISPR/Cas9 cuts is differential across genomic loci and can confound genetic screening." Nature communications 13.1 (2022): 1-14.

- Anzalone, Andrew V., et al. "Search-and-replace genome editing without double-strand breaks or donor DNA." Nature 576.7785 (2019): 149-157.

- Bacchi, Niccolo, Simona Casarosa, and Michela A. Denti. "Splicing-correcting therapeutic approaches for retinal dystrophies: where endogenous gene regulation and specificity matter." Investigative ophthalmology & visual science 55.5 (2014): 3285-3294.

- Bae, Sangsu, Jeongbin Park, and Jin-Soo Kim. "Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases." Bioinformatics 30.10 (2014): 1473-1475.

- Baltimore, David. "Viral RNA-dependent DNA polymerase: RNA-dependent DNA polymerase in virions of RNA tumour viruses." Nature 226.5252 (1970): 1209-1211.

- Barrangou, Rodolphe, et al. "CRISPR provides acquired resistance against viruses in prokaryotes." Science 315.5819 (2007): 1709-1712.

- Bauwens, Miriam, et al. "ABCA4-associated disease as a model for missing heritability in autosomal recessive disorders: novel noncoding splice, cis-regulatory, structural, and recurrent hypomorphic variants." Genetics in Medicine 21.8 (2019): 1761-1771.

- Bauwens, Miriam, et al. "An Augmented ABCA 4 Screen Targeting Noncoding Regions Reveals a Deep Intronic Founder Variant in B elgian S targardt Patients." Human mutation 36.1 (2015): 39-42.

- Berger, Wolfgang, Barbara Kloeckener-Gruissem, and John Neidhardt. "The molecular basis of human retinal and vitreoretinal diseases." Progress in retinal and eye research 29.5 (2010): 335-375.

- Bhongsatiern, Jiraganya, et al. "Retinal-specific ATP-binding cassette transporter (ABCR/ABCA4) is expressed at the choroid plexus in rat brain." Journal of neurochemistry 92.5 (2005): 1277-1280.

- Blencowe, Benjamin J. "Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases." Trends in biochemical sciences 25.3 (2000): 106-110.

- Boughman, J. A., M. Vernon, and K. A. Shaver. "Usher syndrome: definition and estimate of prevalence from two high-risk populations." Journal of chronic diseases 36.8 (1983): 595-603.

- Boulikas, Teni. "Nuclear localization signals (NLS)." Critical reviews in eukaryotic gene expression 3.3 (1993): 193-227.

- Braun, Terry A., et al. "Non-exomic and synonymous variants in ABCA4 are an important cause of Stargardt disease." Human molecular genetics 22.25 (2013): 5136-5145.

- Brierley, Andrew S., et al. "Antarctic krill under sea ice: elevated abundance in a narrow band just south of ice edge." Science 295.5561 (2002): 1890-1892.

- BRILLIANCE: A Phase 1/2 Single Ascending Dose Study of EDIT-101, an in vivo CRISPR Gene Editing Therapy in CEP290-Related Retinal Degeneration. September 29, 2021. Rd meeting (https://ir.editasmedicine.com/static-files/22b32d3d-e38f-4e90-899c-a2e701872745)

- Brinkman EK, Chen T, de Haas M, Holland HA, Akhtar W, van Steensel B. (2018). Kinetics and Fidelity of the Repair of Cas9-Induced Double-Strand DNA Breaks. Mol. Cell. 70(5), 801–813.e6.

- Brinkman, Eva K., et al. "Easy quantitative assessment of genome editing by sequence trace decomposition." Nucleic acids research 42.22 (2014): e168-e168.

- Brunner, Erich, et al. "CRISPR-induced double-strand breaks trigger recombination between homologous chromosome arms." Life science alliance 2.3 (2019).

- Butash, Kimberly A., et al. "Reexamination of the effect of endotoxin on cell proliferation and transfection efficiency." Biotechniques 29.3 (2000): 610-619.

- Capp, Jean-Pascal, et al. "The DNA polymerase λ is required for the repair of non-compatible DNA double strand breaks by NHEJ in mammalian cells." Nucleic acids research 34.10 (2006): 2998-3007.

- Chang, Angela Y., et al. "Preparation of calcium competent Escherichia coli and heat-shock transformation." JEMI methods 1 (2017): 22-25.

- Chang, Bo, et al. "In-frame deletion in a novel centrosomal/ciliary protein CEP290/NPHP6 perturbs its interaction with RPGR and results in early-onset retinal degeneration in the rd16 mouse." Human molecular genetics 15.11 (2006): 1847-1857.

- Chang, Howard HY, et al. "Non-homologous DNA end joining and alternative pathways to double-strand break repair." Nature reviews Molecular cell biology 18.8 (2017): 495-506.

- Cheng, Hiu-Lo, et al. "Structural insigNGS into the duplex DNA processing of TREX2." Nucleic acids research 46.22 (2018): 12166-12176.

- Chirgwin, John M., et al. "Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease." Biochemistry 18.24 (1979): 5294-5299.

- Chu, Gilbert, Hiroshi Hayakawa, and Paul Berg. "Electroporation for the efficient transfection of mammalian cells with DNA." Nucleic acids research 15.3 (1987): 1311-1326.

- Cideciyan, Artur V., et al. "Centrosomal-ciliary gene CEP290/NPHP6 mutations result in blindness with unexpected sparing of photoreceptors and visual brain: implications for therapy of Leber congenital amaurosis." Human mutation 28.11 (2007): 1074-1083.

- Clèries, Ramon, et al. "BootstRatio: a web-based statistical analysis of fold-change in qPCR and RT-qPCR data using resampling methods." Computers in biology and medicine 42.4 (2012): 438-445.

- Conant, David, et al. "Inference of CRISPR edits from sanger trace data." The CRISPR Journal 5.1 (2022): 123-130.

- Corsi, Giulia I., et al. "CRISPR/Cas9 gRNA activity depends on free energy changes and on the target PAM context." Nature Communications 13.1 (2022): 1-14.

- Cox, David Benjamin Turitz, Randall Jeffrey Platt, and Feng Zhang. "Therapeutic genome editing: prospects and challenges." Nature medicine 21.2 (2015): 121-131.

- Cutting, Garry R. "Cystic fibrosis genetics: from molecular understanding to clinical application." Nature Reviews Genetics 16.1 (2015): 45-56.

- Daer RM, Cutts JP, Brafman DA, Haynes KA. (2017) The Impact of Chromatin Dynamics on Cas9-Mediated Genome Editing in Human Cells. ACS Synth Biol. 17;6(3):428-438.

- Davis, Jessie R., et al. "Efficient in vivo base editing via single adeno-associated viruses with size-optimized genomes encoding compact adenine base editors." Nature Biomedical Engineering (2022): 1-12.

- Deltcheva, Elitza, et al. "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III." Nature 471.7340 (2011): 602-607.

- den Hollander, Anneke I., et al. "Mutations in the CEP290 (NPHP6) gene are a frequent cause of Leber congenital amaurosis." The American Journal of Human Genetics 79.3 (2006): 556-561.

- den Hollander, Anneke I., et al. "Mutations in the CEP290 (NPHP6) gene are a frequent cause of Leber congenital amaurosis." The American Journal of Human Genetics 79.3 (2006): 556-561.

- Doench, John G., et al. "Optimized ssgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9." Nature biotechnology 34.2 (2016): 184-191.

- Dulla, Kalyan, et al. "Antisense oligonucleotide-based treatment of retinitis pigmentosa caused by USH2A exon 13 mutations." Molecular Therapy (2021).

- Enache, Oana M., et al. "Cas9 activates the p53 pathway and selects for p53-inactivating mutations." Nature genetics 52.7 (2020): 662-668.

- Engler, Carola, and Sylvestre Marillonnet. "Golden gate cloning." DNA cloning and assembly methods. Humana Press, Totowa, NJ, 2014. 119-131.

- Falavarjani KG, Nguyen QD: Adverse events and complications associated with intravitreal injection of anti-VEGF agents: a review of literature. Eye (Lond) 2013;27:787-794.

- Felgner, Philip L., et al. "Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure." Proceedings of the National Academy of Sciences 84.21 (1987): 7413-7417.

- Flamier A, Barabino A, Gilbert B. (2016) Differentiation of human embryonic stem cells into cone photoreceptors. Bio Protoc. 6:e1870.

- Friedland, Ari E., et al. "Characterization of Staphylococcus aureus Cas9: a smaller Cas9 for all-in-one adeno-associated virus delivery and paired nickase applications." Genome biology 16.1 (2015): 1-10.

- Gaj, Thomas, et al. "Genome-editing technologies: principles and applications." Cold Spring Harbor perspectives in biology 8.12 (2016): a023754.

- Gamwell, Lisa F., et al. "Orphanet Journal of Rare Diseases." Orphanet Journal of Rare Diseases 8 (2013): 33.

- Garanto, Alejandro, and Rob WJ Collin. "Design and in vitro use of antisense oligonucleotides to correct pre-mRNA splicing defects in inherited retinal dystrophies." Retinal Gene Therapy. Humana Press, New York, NY, 2018. 61-78.

- Garanto, Alejandro, et al. "Antisense oligonucleotide screening to optimize the rescue of the splicing defect caused by the recurrent deep-intronic ABCA4 variant c. 4539+ 2001G> A in Stargardt disease." Genes 10.6 (2019): 452.

- Gaudana, Ripal, et al. "Ocular drug delivery." The AAPS journal 12.3 (2010): 348-360.

- Gaudelli, Nicole M., et al. "Programmable base editing of A• T to G• C in genomic DNA without DNA cleavage." Nature 551.7681 (2017): 464-471.

- Gibson, Daniel G., et al. "Enzymatic assembly of DNA molecules up to several hundred kilobases." Nature methods 6.5 (2009): 343-345.

- Gilbert, Luke A., et al. "CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes." Cell 154.2 (2013): 442-451.

- Gillmore, Julian D., et al. "CRISPR-Cas9 in vivo gene editing for transthyretin amyloidosis." New England Journal of Medicine 385.6 (2021): 493-502.

- Grawunder, Ulf, et al. "Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells." Nature 388.6641 (1997): 492-495.

- Grosso, Ana Rita, et al. "Tissue-specific splicing factor gene expression signatures." Nucleic acids research 36.15 (2008): 4823-4832.

- Haapaniemi, Emma, et al. "CRISPR–Cas9 genome editing induces a p53-mediated DNA damage response." Nature medicine 24.7 (2018): 927-930.

- Hammond, Suzan M., and Matthew JA Wood. "Genetic therapies for RNA mis-splicing diseases." Trends in genetics 27.5 (2011): 196-205.

- Hanany, Mor, Carlo Rivolta, and Dror Sharon. "Worldwide carrier frequency and genetic prevalence of autosomal recessive inherited retinal diseases." Proceedings of the National Academy of Sciences 117.5 (2020): 2710-2716.

- Hartong, Dyonne T., Eliot L. Berson, and Thaddeus P. Dryja. "Retinitis pigmentosa." The Lancet 368.9549 (2006): 1795-1809.

- Hayflick l. (1965). The limited in vitro lifetime of human diploid cell strains. Exp. Cell Res. 37, 614–636.

- Hsu, Patrick D., et al. "DNA targeting specificity of RNA-guided Cas9 nucleases." Nature biotechnology 31.9 (2013): 827-832.

- Hu, Ziying, et al. "A compact Cas9 ortholog from Staphylococcus Auricularis (SauriCas9) expands the DNA targeting scope." PLoS biology 18.3 (2020): e3000686.

- Huston, Nicholas C., et al. "Identification of guide-intrinsic determinants of Cas9 specificity." The CRISPR journal 2.3 (2019): 172-185.

- Ihry, Robert J., et al. "p53 inhibits CRISPR–Cas9 engineering in human pluripotent stem cells." Nature medicine 24.7 (2018): 939-946.

- Ira, Grzegorz, et al. "DNA end resection, homologous recombination and DNA damage checkpoint activation require CDK1." Nature 431.7011 (2004): 1011-1017.

- Ishino, Yoshizumi, et al. "Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product." Journal of bacteriology 169.12 (1987): 5429-5433.

- Jacków, Joanna, et al. "CRISPR/Cas9-based targeted genome editing for correction of recessive dystrophic epidermolysis bullosa using iPS cells." Proceedings of the National Academy of Sciences 116.52 (2019): 26846-26852.

- Jacobson, Samuel G., et al. "Outcome measures for clinical trials of Leber congenital amaurosis caused by the intronic mutation in the CEP290 gene." Investigative ophthalmology & visual science 58.5 (2017): 2609-2622.

- Jansen, Ruud, et al. "Identification of genes that are associated with DNA repeats in prokaryotes." Molecular microbiology 43.6 (2002): 1565-1575.

- Jeggo, P. A., and M. Lobrich. "DNA double-strand breaks: their cellular and clinical impact?." Oncogene 26.56 (2007): 7717-7720.

- Jiang, Long, et al. "CRISPR/Cas9-induced DNA damage enriches for mutations in a p53-linked interactome: implications for CRISPR-based therapies." Cancer Research (2021).

- Jinek, Martin, et al. "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." science 337.6096 (2012): 816-821.

- Jinek, Martin, et al. "RNA-programmed genome editing in human cells." elife 2 (2013): e00471.

- Johnston, Calum, et al. "Bacterial transformation: distribution, shared mechanisms and divergent control." Nature Riews Microbiology 12.3 (2014): 181-196.

- Joung, Julia, et al. "Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening." Nature protocols 12.4 (2017): 828-863.

- Jurk, Diana, et al. "Postmitotic neurons develop a p21-dependent senescence-like phenotype driven by a DNA damage response." Aging cell 11.6 (2012): 996-1004.

- Karvelis, Tautvydas, et al. "Rapid characterization of CRISPR-Cas9 protospacer adjacent motif sequence elements." Genome biology 16.1 (2015): 1-13.

- Khan, Mubeen, et al. "Detailed phenotyping and therapeutic strategies for intronic ABCA4 variants in Stargardt disease." Molecular Therapy-Nucleic Acids 21 (2020a): 412-427.

- Khan, Mubeen, et al. "Resolving the dark matter of ABCA4 for 1054 Stargardt disease probands through integrated genomics and transcriptomics." Genetics in Medicine 22.7 (2020b): 1235-1246.

- Kim, Hyongbum, and Jin-Soo Kim. "A guide to genome engineering with programmable nucleases." Nature Reviews Genetics 15.5 (2014): 321-334.

- Kimberling, William J., et al. "Frequency of Usher syndrome in two pediatric populations: Implications for genetic screening of deaf and hard of hearing children." Genetics in Medicine 12.8 (2010): 512-516.

- Kleinstiver, Benjamin P., et al. "High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects." Nature 529.7587 (2016): 490-495.

- Koblan, Luke W., et al. "Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction." Nature biotechnology 36.9 (2018): 843-846.

- Kocher, Thomas, et al. "Cut and paste: efficient homology-directed repair of a dominant negative KRT14 mutation via CRISPR/Cas9 nickases." Molecular Therapy 25.11 (2017): 2585-2598.

- Kolagar, Tannaz A., et al. "Human pluripotent stem cells in neurodegenerative diseases: potentials, advances and limitations." Current stem cell research & therapy 15.2 (2020): 102-110.

- Komor, Alexis C., et al. "Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C: G-to-T: A base editors with higher efficiency and product purity." Science advances 3.8 (2017): eaao4774.

- Komor, Alexis C., et al. "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage." Nature 533.7603 (2016): 420-424.

- Koonin, Eugene V., and Kira S. Makarova. "CRISPR-Cas: an adaptive immunity system in prokaryotes." F1000 biology reports 1 (2009).

- Krokan, Hans E., Finn Drabløs, and Geir Slupphaug. "Uracil in DNA–occurrence, consequences and repair." Oncogene 21.58 (2002): 8935-8948.

- Kunkel, Gary R., et al. "U6 small nuclear RNA is transcribed by RNA polymerase III." Proceedings of the National Academy of Sciences 83.22 (1986): 8575-8579.

- Kunz, C., Y. Saito, and P. Schär. "DNA repair in mammalian cells." Cellular and molecular life sciences 66.6 (2009): 1021-1038.

- Landrum, Melissa J., et al. "ClinVar: public archive of interpretations of clinically relevant variants." Nucleic acids research 44.D1 (2016): D862-D868.

- Leibowitz, Mitchell L., et al. "Chromothripsis as an on-target consequence of CRISPR–Cas9 genome editing." Nature genetics 53.6 (2021): 895-905.

- Lenis, Tamara L., et al. "Expression of ABCA4 in the retinal pigment epithelium and its implications for Stargardt macular degeneration." Proceedings of the National Academy of Sciences 115.47 (2018): E11120-E11127.

- Li, Hongyi, et al. "Applications of genome editing technology in the targeted therapy of human diseases: mechanisms, advances and prospects." Signal transduction and targeted therapy 5.1 (2020): 1-23.

- Liang, Li, et al. "Modulation of DNA end joining by nuclear proteins." Journal of Biological Chemistry 280.36 (2005): 31442-31449.

- Lin, Yanni, et al. "CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences." Nucleic acids research 42.11 (2014): 7473-7485.

- Lisby, Michael, and Rodney Rothstein. "Cell biology of mitotic recombination." Cold Spring Harbor perspectives in biology 7.3 (2015): a016535.

- Livak, Kenneth J., and Thomas D. Schmittgen. "Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method." methods 25.4 (2001): 402-408.

- Luo, Dong-Gen, Tian Xue, and King-Wai Yau. "How vision begins: an odyssey." Proceedings of the National Academy of Sciences 105.29 (2008): 9855-9862.

- Ma, Jia-Lin, et al. "Yeast Mre11 and Rad1 proteins define a Ku-independent mechanism to repair double-strand breaks lacking overlapping end sequences." Molecular and cellular biology 23.23 (2003): 8820-8828.

- Ma, Sara L., et al. "Whole exome sequencing reveals novel PHEX splice site mutations in patients with hypophosphatemic rickets." PLoS One 10.6 (2015): e0130729.

- Ma, Yunmei, Klaus Schwarz, and Michael R. Lieber. "The Artemis: DNA-PKcs endonuclease cleaves DNA loops, flaps, and gaps." DNA repair 4.7 (2005): 845-851.

- Maclaine, Nicola J., and Ted R. Hupp. "The regulation of p53 by phosphorylation: a model for how distinct signals integrate into the p53 pathway." Aging 1.5 (2009): 490.

- Maeder, Morgan L., et al. "Development of a gene-editing approach to restore vision loss in Leber congenital amaurosis type 10." Nature medicine 25.2 (2019): 229-233.

- Makarova, Kira S., and Eugene V. Koonin. "Annotation and classification of CRISPR-Cas systems." CRISPR (2015): 47-75.

- Mandel, Morton, and Akiko Higa. "Calcium-dependent bacteriophage DNA infection." Journal of molecular biology 53.1 (1970): 159-162.

- Martínez Mir, Amalia, et al. "Retinitis pigmentosa caused by a homozygous mutation in the Stargardt disease gene ABCR." (1998).

- Maugeri, Alessandra, et al. "Mutations in the ABCA4 (ABCR) gene are the major cause of autosomal recessive cone-rod dystrophy." The American Journal of Human Genetics 67.4 (2000): 960-966.

- Maule, Giulia, et al. "Allele specific repair of splicing mutations in cystic fibrosis through AsCas12a genome editing." Nature communications 10.1 (2019): 1-11.

- Meyer, Jason S., et al. "Modeling early retinal development with human embryonic and induced pluripotent stem cells." Proceedings of the National Academy of Sciences 106.39 (2009): 16698-16703.

- Michaelides, Michel, D. M. Hunt, and A. T. Moore. "The genetics of inherited macular dystrophies." Journal of medical genetics 40.9 (2003): 641-650.

- Mimitou, Eleni P., and Lorraine S. Symington. "Ku prevents Exo1 and Sgs1-dependent resection of DNA ends in the absence of a functional MRX complex or Sae2." The EMBO journal 29.19 (2010): 3358-3369.

- Mojica, F. J. M., et al. "Long stretches of short tandem repeats are present in the largest replicons of the Archaea Haloferax mediterranei and Haloferax volcanii and could be involved in replicon partitioning." Molecular microbiology 17.1 (1995): 85-93.

- Molday, Robert S. "ATP-binding cassette transporter ABCA4: molecular properties and role in vision and macular degeneration." Journal of bioenergetics and biomembranes 39.5 (2007): 507-517.

- Moon, Andrea F., et al. "Sustained active site rigidity during synthesis by human DNA polymerase μ." Nature structural & molecular biology 21.3 (2014): 253-260.

- Moynahan, Mary Ellen, and Maria Jasin. "Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis." Nature reviews Molecular cell biology 11.3 (2010): 196-207.

- Nahmad, Alessio David, et al. "Frequent aneuploidy in primary human T cells after CRISPR–Cas9 cleavage." Nature Biotechnology (2022): 1-7.

- Nakano, Tokushige, et al. "Self-formation of optic cups and storable stratified neural retina from human ESCs." Cell stem cell 10.6 (2012): 771-785.

- Nishida, Keiji, et al. "Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems." Science 353.6305 (2016): aaf8729.

- Nishimasu, Hiroshi, et al. "Crystal structure of Cas9 in complex with guide RNA and target DNA." Cell 156.5 (2014): 935-949.

- Okita, Keisuke, et al. "A more efficient method to generate integration-free human iPS cells." Nature methods 8.5 (2011): 409-412.

- Orthwein, Alexandre, et al. "A mechanism for the suppression of homologous recombination in G1 cells." Nature 528.7582 (2015): 422-426.

- Palazzo, Alexander F., and T. Ryan Gregory. "The case for junk DNA." PLoS genetics 10.5 (2014): e1004351.

- Panagiotopoulos, Anna-Lena, et al. "Antisense oligonucleotide-and CRISPR-Cas9-mediated rescue of mRNA splicing for a deep intronic CLRN1 mutation." Molecular Therapy-Nucleic Acids 21 (2020): 1050-1061.

- Pâques, Frédéric, and Philippe Duchateau. "Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy." Current gene therapy 7.1 (2007): 49-66.

- Perrino, Fred W., et al. "Cooperative DNA binding and communication across the dimer interface in the TREX2 3'→ 5'-exonuclease." Journal of Biological Chemistry 283.31 (2008): 21441-21452.

- Perrino, Fred W., et al. "The human TREX2 3'→ 5'-exonuclease structure suggests a mechanism for efficient nonprocessive DNA catalysis." Journal of Biological Chemistry 280.15 (2005): 15212-15218.

- Pfaffl, Michael W., et al. "Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper–Excel-based tool using pair-wise correlations." Biotechnology letters 26.6 (2004): 509-515.

- Pierce, Eric A., and Jean Bennett. "The status of RPE65 gene therapy trials: safety and efficacy." Cold Spring Harbor perspectives in medicine 5.9 (2015): a017285.

- Porto, Elizabeth M., et al. "Base editing: advances and therapeutic opportunities." Nature Reviews Drug Discovery 19.12 (2020): 839-859.

- Pugh Jr, Edward N., and T. D. Lamb. "Cyclic GMP and calcium: the internal messengers of excitation and adaptation in vertebrate photoreceptors." Vision research 30.12 (1990): 1923-1948.

- Quazi, Faraz, and Robert S. Molday. "ATP-binding cassette transporter ABCA4 and chemical isomerization protect photoreceptor cells from the toxic accumulation of excess 11-cis-retinal." Proceedings of the National Academy of Sciences 111.13 (2014): 5024-5029.

- Quazi, Faraz, Stepan Lenevich, and Robert S. Molday. "ABCA4 is an N-retinylidene-phosphatidylethanolamine and phosphatidylethanolamine importer." Nature communications 3.1 (2012): 1-9.

- Ran, F. Ann, et al. "Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity." Cell 154.6 (2013): 1380-1389.

- Rees, Holly A., Wei-Hsi Yeh, and David R. Liu. "Development of hRad51–Cas9 nickase fusions that mediate HDR without double-stranded breaks." Nature communications 10.1 (2019): 1-12.

- Richardson, Christopher D., et al. "Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA." Nature biotechnology 34.3 (2016): 339-344.

- Richter, Michelle F., et al. "Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity." Nature biotechnology 38.7 (2020): 883-891.

- Rivolta, Carlo, et al. "Retinitis pigmentosa and allied diseases: numerous diseases, genes, and inheritance patterns." Human molecular genetics 11.10 (2002): 1219-1227.

- Ruan, Guo-Xiang, et al. "CRISPR/Cas9-mediated genome editing as a therapeutic approach for Leber congenital amaurosis 10." Molecular Therapy 25.2 (2017): 331-341.

- Samulski, R. Jude, and Nicholas Muzyczka. "AAV-mediated gene therapy for research and therapeutic purposes." Annual review of virology 1 (2014): 427-451.

- Sanger, Frederick, Steven Nicklen, and Alan R. Coulson. "DNA sequencing with chain-terminating inhibitors." Proceedings of the national academy of sciences 74.12 (1977): 5463-5467.

- Sangermano, Riccardo, et al. "ABCA4 midigenes reveal the full splice spectrum of all reported noncanonical splice site variants in Stargardt disease." Genome research 28.1 (2018): 100-110.

- Sanz DJ, Harrison PT. (2019). Minigene assay to Evaluate CRISPR/Cas9-based excision of Intronic mutations that Cause Aberrant Splicing in Human Cells. Bio Protoc. 9(11), e3251.

- Sanz, David J., et al. "Cas9/gRNA targeted excision of cystic fibrosis-causing deep-intronic splicing mutations restores normal splicing of CFTR mRNA." PloS one 12.9 (2017): e0184009.

- Sayer, J. A., Otto, E. A., O'Toole, J. F., Nurnberg, G., Kennedy, M. A., Becker, C., Hennies, H. C., Helou, J., Attanasio, M., Fausett, B. V., Utsch, B., Khanna, H., and 30 others. The centrosomal protein nephrocystin-6 is mutated in Joubert syndrome and activates transcription factor ATF4. Nature Genet. 38: 674-681, 2006.

- Schep, Ruben, et al. "Impact of chromatin context on Cas9-induced DNA double-strand break repair pathway balance." Molecular cell 81.10 (2021): 2216-2230.

- Schmidt, Moritz J., et al. "Improved CRISPR genome editing using small highly active and specific engineered RNA-guided nucleases." Nature communications 12.1 (2021): 1-12.

- Schulz, Heidi L., et al. "Mutation spectrum of the ABCA4 gene in 335 Stargardt disease patients from a multicenter German cohort—impact of selected deep intronic variants and common SNPs." Investigative ophthalmology & visual science 58.1 (2017): 394-403.

- Schwank, Gerald, and Hans Clevers. "CRISPR/Cas9-mediated genome editing of mouse small intestinal organoids." Gastrointestinal Physiology and Diseases. Humana Press, New York, NY, 2016. 3-11.

- Sedelnikova, Olga A., et al. "Senescing human cells and ageing mice accumulate DNA lesions with unrepairable double-strand breaks." Nature cell biology 6.2 (2004): 168-170.

- Sfeir, Agnel, and Lorraine S. Symington. "Microhomology-mediated end joining: a back-up survival mechanism or dedicated pathway?." Trends in biochemical sciences 40.11 (2015): 701-714.

- Shafa, Mehdi, et al. "Human-induced pluripotent stem cells manufactured using a current good manufacturing practice-compliant process differentiate into clinically relevant cells from three germ layers." Frontiers in medicine 5 (2018): 69.

- Shah, Shiraz A., et al. "Protospacer recognition motifs: mixed identities and functional diversity." RNA biology 10.5 (2013): 891-899.

- Shen, Bin, et al. "Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects." Nature methods 11.4 (2014): 399-402.

- Shen, Max W., et al. "Predictable and precise template-free CRISPR editing of pathogenic variants." Nature 563.7733 (2018): 646-651.

- Sheth, Nihar, et al. "Comprehensive splice-site analysis using comparative genomics." Nucleic acids research 34.14 (2006): 3955-3967.

- Shroyer, Noah F., et al. "Cosegregation and functional analysis of mutant ABCR (ABCA4) alleles in families that manifest both Stargardt disease and age-related macular degeneration." Human molecular genetics 10.23 (2001): 2671-2678.

- Sinha, Sanju, et al. "A systematic genome-wide mapping of oncogenic mutation selection during CRISPR-Cas9 genome editing." Nature communications 12.1 (2021): 1-13.

- Slaymaker, Ian M., et al. "Rationally engineered Cas9 nucleases with improved specificity." Science 351.6268 (2016): 84-88.

- Smith, Michael. "In vitro mutagenesis." Annual Riew of genetics 19.1 (1985): 423-462.

- Sparrow, Janet R., et al. "The bisretinoids of retinal pigment epithelium." Progress in retinal and eye research 31.2 (2012): 121-135.

- Stadtmauer, Edward A., et al. "CRISPR-engineered T cells in patients with refractory cancer." Science 367.6481 (2020): eaba7365.

- Stephens, R. Michael, and Thomas Dana Schneider. "Features of spliceosome evolution and function inferred from an analysis of the information at human splice sites." Journal of molecular biology 228.4 (1992): 1124-1136.

- Sternberg, Samuel H., et al. "DNA interrogation by the CRISPR RNA-guided endonuclease Cas9." Nature 507.7490 (2014): 62-67.

- Sun, Hanzhen, and Lawrence A. Chasin. "Multiple splicing defects in an intronic false exon." Molecular and cellular biology 20.17 (2000): 6414-6425.

- Sun, Ning, and Huimin Zhao. "Transcription activator-like effector nucleases (TALENs): a highly efficient and versatile tool for genome editing." Biotechnology and bioengineering 110.7 (2013): 1811-1821.

- Tachikawa, Masanori, et al. "Distinct spatio-temporal expression of ABCA and ABCG transporters in the developing and adult mouse brain." Journal of neurochemistry 95.1 (2005): 294-304.

- Takahashi, Kazutoshi, and Shinya Yamanaka. "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors." cell 126.4 (2006): 663-676.

- Temin, Howard M., and S. Mizutami. "RNA-dependent DNA polymerase in virions of Rous sarcoma virus." Nature 226 (1970): 1211-1213.

- Thakore, Pratiksha I., et al. "Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements." Nature methods 12.12 (2015): 1143-1149.

- Thanaraj, T. A., and Francis Clark. "Human GC-AG alternative intron isoforms with weak donor sites show enhanced consensus at acceptor exon positions." Nucleic acids research 29.12 (2001): 2581-2593.

- Tornabene, Patrizia, et al. "Inclusion of a degron reduces levels of undesired inteins after AAV-mediated protein trans-splicing in the retina." Molecular Therapy-Methods & Clinical Development 23 (2021): 448-459.

- Tornabene, Patrizia, et al. "Intein-mediated protein trans-splicing expands adeno-associated virus transfer capacity in the retina." Science translational medicine 11.492 (2019).

- Tsai, Shengdar Q., et al. "CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR–Cas9 nuclease off-targets." Nature methods 14.6 (2017): 607-614.

- Tsai, Shengdar Q., et al. "GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases." Nature biotechnology 33.2 (2015): 187-197.

- Tsang, Stephen H., and Tarun Sharma. "Stargardt disease." Atlas of Inherited Retinal Diseases (2018): 139-151.

- Urnov, Fyodor D., et al. "Genome editing with engineered zinc finger nucleases." Nature Reviews Genetics 11.9 (2010): 636-646.

- van Bennekum, Ariëtte M., et al. "Biochemical basis for depressed serum retinol levels in transthyretin-deficient mice." Journal of Biological Chemistry 276.2 (2001): 1107-1113.

- van den Berg, Jeroen, et al. "A limited number of double-strand DNA breaks is sufficient to delay cell cycle progression." Nucleic acids research 46.19 (2018): 10132-10144.

- van Diepen, Hester, et al. "QR-421a, an antisense oligonucleotide, for the treatment of retinitis pigmentosa due to USH2A exon 13 mutations." Invest. Ophthalmol. Vis. Sci 60 (2019): 3250.

- Vaz-Drago, Rita, Noélia Custódio, and Maria Carmo-Fonseca. "Deep intronic mutations and human disease." Human genetics 136.9 (2017): 1093-1111.

- von Zglinicki T, Wan T, Miwa S. (2020). Senescence in Post-Mitotic Cells: A Driver of Aging? Antioxid Redox Signal. 34(4), 308–323.

- Walker, John R., Richard A. Corpina, and Jonathan Goldberg. "Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair." Nature 412.6847 (2001): 607-614.

- Wang, Gaifang, and Maryam Farzaneh. "Mini review; differentiation of human pluripotent stem cells into oocytes." Current Stem Cell Research & Therapy 15.4 (2020): 301-307.

- Weber, M., et al. "Short technical reports. Effects of lipopolysaccharide on transfection efficiency in eukaryotic cells." BioTechniques 19.6 (1995): 930-940.

- Weisschuh, Nicole, Bernd Wissinger, and Eugen Gramer. "A splice site mutation in the PAX6 gene which induces exon skipping causes autosomal dominant inherited aniridia." Molecular vision 18 (2012): 751.

- Westeneng-van Haaften, Sarah C., et al. "Clinical and genetic characteristics of late-onset Stargardt's disease." Ophthalmology 119.6 (2012): 1199-1210.

- Wiley, Luke A., et al. "Expression of the retina-specific flippase, ABCA4, in epidermal keratinocytes." F1000Research 5.193 (2016): 193.

- Wilkinson, Max E., Clément Charenton, and Kiyoshi Nagai. "RNA Splicing by the Spliceosome." Annual review of biochemistry 89 (2020): 359-388.

- Wirth, Thomas, Nigel Parker, and Seppo Ylä-Herttuala. "History of gene therapy." Gene 525.2 (2013): 162-169.

- Wu, Q., and ADRIAN R. Krainer. "Splicing of a divergent subclass of AT-AC introns requires the major spliceosomal snRNAs." RNA-CAMBRIDGE- 3 (1997): 586-601.

- Wyatt, David W., et al. "Essential roles for polymerase θ-mediated end joining in the repair of chromosome breaks." Molecular cell 63.4 (2016): 662-673.

- Xie, Tian, et al. "Structural basis of substrate recognition and translocation by human ABCA4." Nature Communications 12.1 (2021): 1-11.

- Xu, Lei, et al. "Clarin-1 expression in adult mouse and human retina highligNGS a role of Müller glia in Usher syndrome." The Journal of pathology 250.2 (2020): 195-204.

- Yang, Han, et al. "Methods favoring homology-directed repair choice in response to CRISPR/Cas9 induced-double strand breaks." International Journal of Molecular Sciences 21.18 (2020): 6461.

- Yatsenko, Alexander, et al. "Late-onset Stargardt disease is associated with missense mutations that map outside known functional regions of ABCR (ABCA4)." Human genetics 108.4 (2001): 346-355.

- Yin, Jianhang, et al. "Cas9 exo-endonuclease eliminates chromosomal translocations during genome editing." Nature communications 13.1 (2022): 1-14.

- Yosef, Ido, Moran G. Goren, and Udi Qimron. "Proteins and DNA elements essential for the CRISPR adaptation process in Escherichia coli." Nucleic acids research 40.12 (2012): 5569-5576.

- Zernant, Jana, et al. "Analysis of the ABCA4 gene by next-generation sequencing." Investigative ophthalmology & visual science 52.11 (2011): 8479-8487.

- Zernant, Jana, et al. "Analysis of the ABCA4 genomic locus in Stargardt disease." Human molecular genetics 23.25 (2014): 6797-6806.

- Zernant, Jana, et al. "Frequent hypomorphic alleles account for a significant fraction of ABCA4 disease and distinguish it from age-related macular degeneration." Journal of medical genetics 54.6 (2017): 404-412.

- Zhang, Han, et al. "Adenine base editing in vivo with a single adeno-associated virus vector." GEN biotechnology 1.3 (2022): 285-299.

- Zhang, Xiao-Hui, et al. "Off-target effects in CRISPR/Cas9-mediated genome engineering." Molecular Therapy-Nucleic Acids 4 (2015): e264.

- Zhang, Zhenguo, et al. "Noisy splicing, more than expression regulation, explains why some exons are subject to nonsense-mediated mRNA decay." BMC biology 7.1 (2009): 1-13.

- Zhao, Na, et al. "Differentiation of human induced pluripotent stem cells into male germ cells." Current Stem Cell Research & Therapy 16.5 (2021): 622-629.

- Zhong, Xiufeng, et al. "Generation of three-dimensional retinal tissue with functional photoreceptors from human iPSCs." Nature communications 5.1 (2014): 1-14.

- Zhu, Jun, Akila Mayeda, and Adrian R. Krainer. "Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins." Molecular cell 8.6 (2001): 1351-1361.

- Zhuo, Chenya, et al. "Spatiotemporal control of CRISPR/Cas9 gene editing." Signal Transduction and Targeted Therapy 6.1 (2021): 1-18.

10 CONTRIBUTIONS

Part of the results included in this thesis have been produced by Master's students during their intership or thesis periods in AG Wissinger. The projects of the students were defined and directly supervised by De Angeli Pietro.

Eleonora Roschi produced part of the data within her 6 months Master project and thesis predominantly related to the the novel c.5196+1134C>G DIV and the clustered intron 36 variants. Specifically:

- Minigene establishment for preliminary validation of splicing patterns
- single gRNA design and selection for the clustered deep-intron variants
- sgRNA/ED*Sp*Cas9 mediated splicing rescue in the *ABCA4* intron 36 clustered deep-intronic variants c.5196+1013A>G, c.5196+1056A>G, c.5196+1134C>G, c.5196+1137G>A and c.5196+1216C>A minigene models

Evangelos Bouris attended an 8 weeks lab rotation within his Master program. He participated in the optimization of the ED*Sp*Cas9 molecule by testing additional exonuclease partners. The information inclued in **Figure 24** is part of his lab rotation report.

Shirley Yu attended an 8 weeks lab rotation within her Master program. She supported the establishment of the c.5196+1134C/G induced pluripotent stem cell line during her lab rotation period.

Dr. Andrew Bassett (Sanger Institute, Cambridge, UK) kindly provided the original *Sp*Cas9-TREX2 original.

Dr. Stefan Hauser and **Prof. Ludger Schöls** (DZNE, Tübingen) have contributed to part of the experiments by providing an established control iPSC line for the differentiation in cCPCs.

Prof. Dr. med. Katarina Stingl and **Dr. med. Laura Kühlewein** (Centre for Ophthalmology, Tübingen) recruited patients and collected skin punches for subsequent fibroblast isolation.

Dr. Peggy Reuter (Centre for Ophthalmology, Tübingen) technically supported part of the experiments.

Parts of the data described in this thesis have been published in Molecular Therapy – Nucleic Acids, in a manuscript titled "Effective splicing restoration of a deep-intronic *ABCA4* variant in cone photoreceptor precursor cells by CRISPR/*Sp*Cas9 approaches".
DOI:https://doi.org/10.1016/j.omtn.2022.07.023.

A patent describing the use of the ED*Sp*Cas9 approach in its broadest application was submitted to the EPO under the provisional **EP22155703.6** application number.