Chapter 10

Modeling Splicing Variants Amenable to Antisense Therapy by Use of CRISPR-Cas9-Based Gene Editing in HepG2 Cells

Arístides López-Márquez, Ainhoa Martínez-Pizarro, Belén Pérez, Eva Richard, and Lourdes R. Desviat ⓘD

Abstract

The field of splice modulating RNA therapy has gained new momentum with FDA approved antisense-based drugs for several rare diseases. In vitro splicing assays with minigenes or patient-derived cells are commonly employed for initial preclinical testing of antisense oligonucleotides aiming to modulate splicing. However, minigenes do not include the full genomic context of the exons under study and patients' samples are not always available, especially if the gene is expressed solely in certain tissues (e.g. liver or brain). This is the case for specific inherited metabolic diseases such as phenylketonuria (PKU) caused by mutations in the liver-expressed *PAH* gene.

Herein we describe the generation of mutation-specific hepatic cellular models of PKU using CRISPR/Cas9 system, which is a versatile and easy-to-use gene editing tool. We describe in detail the selection of the appropriate cell line, guidelines for design of RNA guides and donor templates, transfection procedures and growth and selection of single-cell colonies with the desired variant, which should result in the accurate recapitulation of the splicing defect.

Key words Splicing, Gene editing, CRISPR/Cas9, HepG2, Inherited metabolic diseases, Phenylketonuria, Cellular models

1 Introduction

Splicing defects account for up to one-third of human disease-causing variants, according to the current estimates [1–3]. Constitutive splicing relies on the recognition of consensus splicing sequences (5′ splice site, 3′ splice site, branch point, and polypyrimidine tract) by spliceosomal components, as well as of other less conserved regulatory elements, referred to as exonic or intronic splicing enhancers or silencers (ESE, ISE, ESS, or ISS), that modulate spliceosome recruitment [4]. These cis-regulatory elements are recognized by trans-acting factors including the serine/arginine-rich domain-containing (SR) protein and heterogeneous nuclear

ribonucleoprotein (hnRNP) families that may act co-ordinately to accurately regulate exon inclusion.

Pathogenic splicing variants disrupt conserved splice sites or regulatory elements or cause aberrant splicing by creating/activating alternative splice sites or by promoting the aberrant inclusion of intronic pseudoexons [4]. Splicing can be modulated therapeutically using antisense approaches, and to date, the clinically approved splice-switching antisense oligonucleotides (SSO) for spinal muscular atrophy, Duchenne muscular atrophy and for an individual patient with a rare, fatal neurodegenerative disease [5–7], represent landmarks in the field, opening new avenues for treatment of patients with defects amenable to splice-mediated correction.

The first requirement for the accurate design and testing of antisense splice correction therapy is the availability of relevant experimental models in which to dissect the underlying molecular mechanisms of pathogenic variants and to test candidate molecules. In this sense, the development of clustered-regulatory interspaced short palindromic repeats (CRISPR)-CRISPR associated nuclease (Cas) genome editing has paved the way to the rapid and easy generation of new and improved cell/animal models of disease. This has facilitated the understanding of the specific pathogenic effect and has allowed efficient testing of targeted therapies, including allele-specific repair for splicing mutations, in tissue types with native expression levels [8–13]. Based on a naturally employed bacterial defense mechanism [14, 15], CRISPR/Cas9 technology was developed as a programmable system of genetic editing that commonly uses the Cas9 nuclease from *Streptococcus pyogenes* and a RNA duplex comprised of a sequence-specific CRISPR RNA (crRNA) and a generic trans-activating CRISPR RNA (tracrRNA) that directs the nuclease to a cut site point, three base pairs upstream of the protospacer adjacent motif or PAM. The PAM is a three-nucleotide motif essential for the nuclease to recognize its DNA target which in the case of Cas9 is NGG. The crRNA and the tracrRNA can be delivered individually or linked in a single RNA molecule. These elements can be delivered to cells as plasmids or as a ribonucleoprotein (RNP) complex [16].

Once Cas9 nuclease cuts the DNA introducing a double stranded break (DSB), the cell can repair this through two different mechanisms: non-homologous end-joining (NHEJ) which usually results in small insertions or deletions, useful for the generation of gene knockouts, or homology driven repair (HDR), used to introduce specific changes via a DNA template with homology arms to our target locus and containing the sequence or point mutation desired [16].

In our laboratory we have used CRISPR/Cas9 technology to introduce splicing mutations causing inherited metabolic diseases (IMD) in cellular and animal models. IMD are monogenic diseases

characterized by dysregulation of the metabolic networks that underlie development and homeostasis [17]. They belong to the category of rare diseases due to their low individual prevalence and are generally enzyme deficiencies of autosomal recessive inheritance, characterized by the toxic accumulation of precursors and of their derivatives or by lack of downstream metabolites. Several of the most frequent and well characterized IMD, e.g. organic acidemias and amino acid disorders, are of major hepatic expression and, as in other genetic diseases, 13–25% of all disease-causing variants interfere with mRNA splicing (HGMD statistics, Professional Release 2019.3). These data warrant further investigation of the therapeutical potential of SSOs in these diseases and the generation of liver specific cellular models for these studies.

The generation of a cell model using CRISPR/Cas9 system can be done in a huge variety of cell lines. In this chapter we describe the protocol for efficient introduction of a specific splicing variant in the *PAH* gene, coding for phenylalanine hydroxylase, and responsible for the well characterized disease phenylketonuria (PKU, MIM#261600), inherited in autosomal recessive fashion. Human *PAH* is exclusively expressed in liver, so in this protocol we use hepatoma cell line HepG2 seeking to attain edition in both alleles (homozygous phenotype). We explain how to select for the appropriate cell line in each particular case, describe the design of RNA guides and donor templates, transfection procedures, growth of single-cell colonies, selection and testing to confirm genomic edition, and accurate recapitulation of the splicing defect (Fig. 1). Appropriate controls to be included in each step are explained, as well as the necessary precautions to be taken especially for intronic splicing variants. We use as example the CRISPR/Cas9-mediated introduction of the recently characterized *PAH* intronic variant, c.1199 + 20G > C, that causes exon skipping due to disruption of a splicing regulatory element [18]. This variant creates a PshAI restriction site, which is used to screen for gene edition in the transfected cells.

2 Materials

2.1 Cell Culture

1. Laminar flow-hood.
2. Humid CO_2 incubator.
3. Centrifuge.
4. Phase-contrast microscope.
5. Hemocytometer–double chamber with Neubauer rulings.
6. Manual Counter.
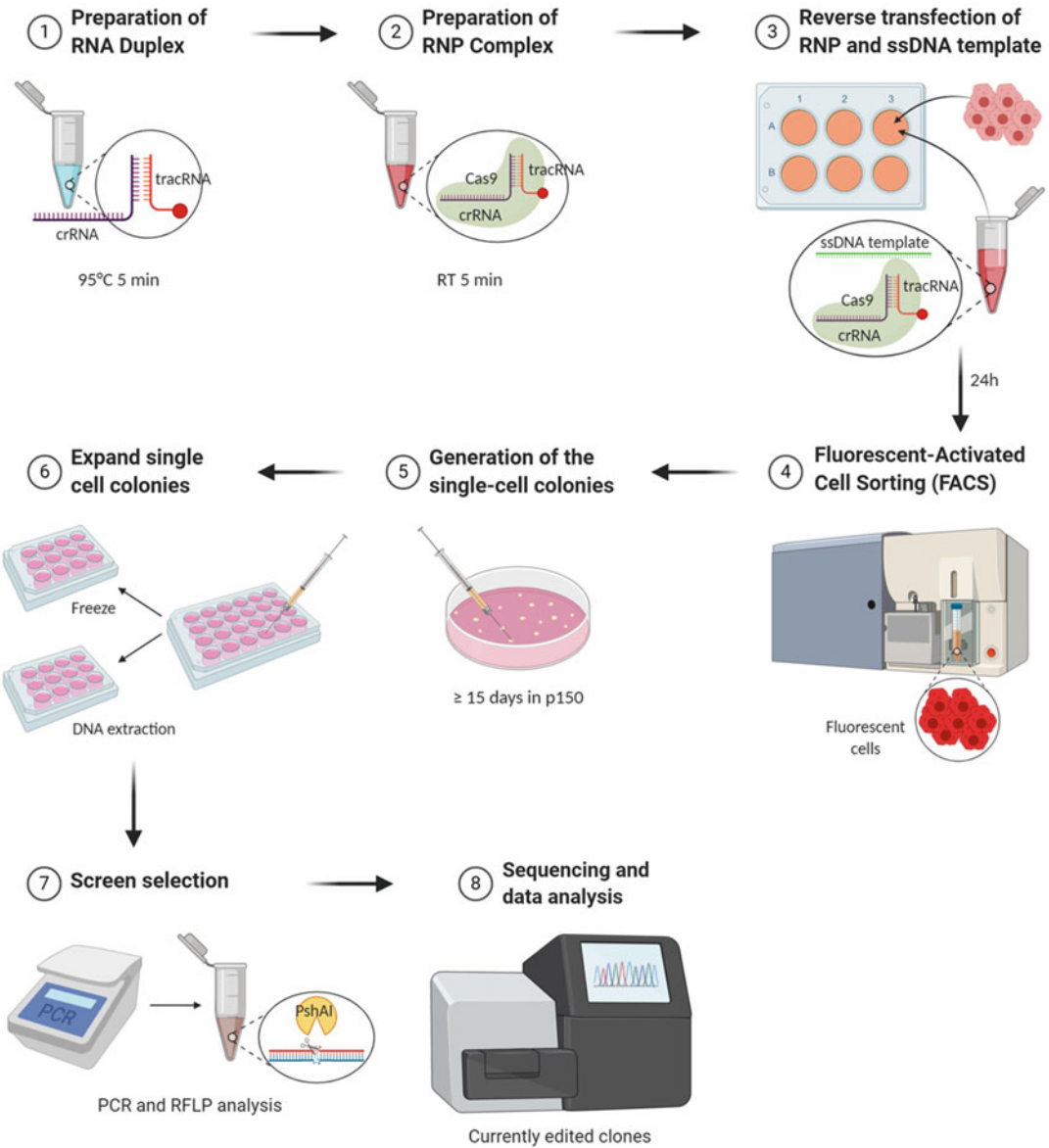7. Consumables: Tissue culture plates, filtered tips, falcon tubes, Eppendorf tubes.

Fig. 1 Outline of the gene editing experimental protocol. (This image was created using BioRender)

8. Phosphate Buffered Saline (PBS).

9. Minimum Essential Medium (MEM) supplemented with 10% Fetal Bovine Serum (FBS), 1% L-Glutamine, and antibiotics.

10. Solution of trypsin–EDTA: 0.25% trypsin, 1 mM EDTA.

11. Trypan Blue Solution: 0.4% trypan blue in PBS.

12. Micropipettes.

13. Stripper micropipettes and 150 μm tips (Origio Inc).

2.2 Ribonucle- oprotein (RNP) Transfection	1. Cas9 Nuclease (*see* **Note 1**).
	2. Fluorescently labeled tracrRNA (tracrRNA-ATTO550) (*see* **Note 1**).
	3. Single-stranded (ss) DNA Template (*see* **Note 1**).
	4. crRNA (*see* **Note 1**).
	5. RNA Duplex Buffer supplied by the manufacturer (*see* **Note 1**).
	6. Nuclease-Free Water.
	7. OptiMEM media.
	8. Lipofectamine Transfection Reagent (*see* **Note 2**).
2.3 Fluorescence Activated Cell Sorting	1. Sorting buffer: PBS, 5 mM EDTA, 25 mM Hepes pH 7.0 supplemented with 2% FBS.
	2. 5-mL polystyrene tubes with cell strainer.
	3. Cell Sorter.
2.4 Genomic DNA Isolation	1. QIAamp DNA Mini Kit for DNA purification (Qiagen).
	2. Centrifuge.
	3. NanoDrop One spectrophotometer (Thermo Fisher Scientific).
2.5 Polymerase Chain Reaction (PCR)	1. Thermal cycler.
	2. PCR tubes.
	3. Nuclease-Free Water.
	4. dNTPs.
	5. FastStart Taq DNA Polymerase (Roche) and PCR buffer $10\times$ (25 mM $MgCl_2$).
	6. Target-specific primers.
	7. Agarose gel with ethidium bromide (0.4 μg/mL) and UV transilluminator.
	8. DNA Molecular Weight Marker.
	9. Kit to purify PCR products, e.g., Cycle Pure Kit for PCR product purification (Omega).
2.6 Restriction Fragment Length Polymorphism Assay (RFLP)	1. PshAI restriction enzyme and enzyme reaction buffer.
	2. Agarose gel with ethidium bromide (0.4 μg/mL) and UV transilluminator.
	3. DNA Molecular Weight Marker.
2.7 RNA Isolation and Reverse Transcription	1. Trizol isolation reagent (Ambion).
	2. 2-Propanol and chloroform.
	3. Ethanol 75%.

4. RNase-Free Water and RNase free-consumables.

5. NanoDrop One spectrophotometer (Thermo Fisher Scientific).

6. Thermal cycler.

7. NZY First-Strand cDNA Synthesis Kit.

2.8   Web Resources

1. Sequences and genomes: https://www.ensembl.org/.

2. Sequences alignments: https://blast.ncbi.nlm.nih.gov/Blast.cgi.

3. Design and analysis of crRNAs: https://bioinfogp.cnb.csic.es/tools/breakingcas/.

4. Design, analysis and/or ordering of crRNAs, tracrRNA, ssDNA Templates: https://eu.idtdna.com/site/order/designtool/index/CRISPR_SEQUENCE.

5. PCR primer design: http://bioinfo.ut.ee/primer3-0.4.0/.

6. Primer and PCR product analysis: https://www.ncbi.nlm.nih.gov/tools/primer-blast/.

3   Methods

3.1   HepG2 Cell Culture

1. Culture the selected HepG2 cell line following standard procedures in P-100 culture dishes with MEM supplemented with 10% FBS, antibiotics, and glutamine at 37 °C in an incubator with 95% humidity and 5% CO_2 (*see* **Notes 3–6**).

2. Just before transfection (see Subheading 3.5 below) detach cells by trypsinization. First, aspirate the media and wash the cells with PBS. Once the PBS has been aspirated from the plate, add 0.25% trypsin-EDTA into the plate and incubate at 37 °C for 5 min. Check by microscopy that the cells are rounding up. Once the cells are detached from the plate add 10% FBS in MEM to stop the trypsin reaction. Pipette the cells up and down to dissociate detached cell clumps into single cells. Transfer the cells to a falcon tube and spin them in a centrifuge at $218 \times g$ for 5 min. Discard the supernatant and resuspend the cell pellet in fresh medium.

3. Count the resuspended cells using a hemocytometer. Prepare a dilution 1:8 of the cells in Trypan Blue solution to distinguish dead cells (stained blue). Add the cell suspension to both chamber sides of the hemocytometer and count the cells with the help of a manual counter.

3.2   Design of Guide RNAs and Donor Template

1. Design the specific crRNA guides with the help of bioinformatics software and their potential off-targets (*see* **Notes 7–12**) (Fig. 2).
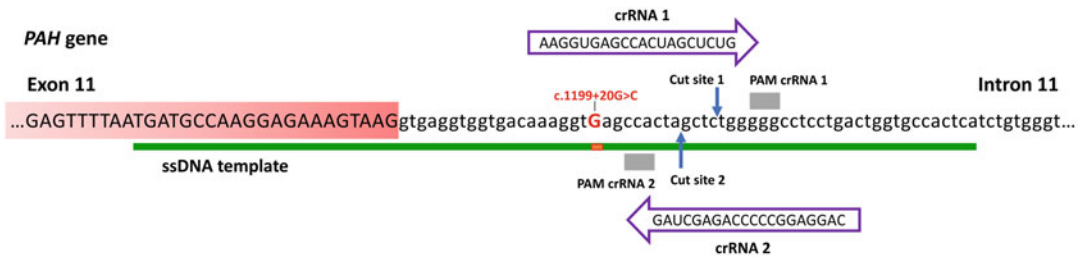
Fig. 2 Schematic representation of the *PAH* gene region surrounding the c.1199 + 20G > C mutation, showing the sequence of the ssDNA template which will include the desired change (green line with red box) and the crRNA guides used (purple arrows), indicating the corresponding PAM sequences (gray rectangles) and the Cas9 nuclease cut sites 3 nucleotides upstream of PAM (blue arrows)

2. Design the ssDNA donor template (*see* **Notes 13–15**) (Fig. 2).

3. Order the crRNA guides, tracrRNA, ssDNA donor template, Cas9 nuclease, transfection reagent, and all the necessary reagents for the CRISPR/Cas9 system (*see* **Note 1**).

3.3 Preparation of RNA Duplex

1. Resuspend the crRNA and tracrRNA-ATTO550 in 20 and 50 μL of Nuclease-Free Duplex Buffer, respectively, resulting in 100 μM stock concentrations (*see* **Notes 16–19**).

2. Prepare the RNA Duplex at a final concentration of 1 μM by mixing the tracrRNA and the crRNA in equimolar concentrations in Nuclease-Free Duplex Buffer (add 1 μL of each crRNA and tracrRNA-ATTO550 to 98 μL of buffer) (*see* **Note 20**).

3. Heat at 95 °C for 5 min.

4. Cool to room temperature (25 °C).

3.4 Preparation of the Ribonucleoprotein Complex (RNP)

1. Dilute Cas9 nuclease to a working concentration of 1 μM in OptiMEM (*see* **Note 21**).

2. Prepare the RNP by mixing in independent tubes for each crRNA the following: 24 μL of RNA duplex (1 μM), 24 μL of Cas9 (1 μM), 9.6 μL of Cas9 PLUS reagent from CRISPR MAX kit (*see* **Note 21**), and 342.4 μL of OptiMEM, adding to a total 400 μL.

3. Incubate at room temperature for 5 min.

3.5 Reverse Transfection of RNP and DNA Donor Template

1. Prepare the ssDNA Donor Template at a working concentration of 1 μM in Nuclease-Free Water.

2. Prepare the transfection mixing the following for each well of a 6-well plate: 7.2 μL of 1 μM ssDNA donor template, 400 μL of RNP complex, 19.2 μL of CRISPRMAX Transfection reagent and 373.6 μL of OptiMEM, adding to a total 800 μL (*see* **Notes 22** and **23**).

3. Incubate at room temperature for 20 min.

4. During the incubation of the transfection mix proceed to detach HepG2 cells by trypsinization as explained above (Subheading 3.1) (*see* **Note 24**).

5. Prepare a dilution of 4×10^5 cells/mL with complete MEM without antibiotics.

6. Once the incubation of the transfection mix is complete, add 800 μL to each well of the 6-well plate.

7. Add 1600 μL of the cell suspension to each well containing the transfection mix, for a final volume of 2400 μL. The number of cells should be 6.4×10^5 cells/well; final concentration of RNP is 10 nM and final concentration of the ssDNA donor template is 3 nM (*see* **Note 22**).

8. Incubate the cells in an incubator at 37 °C and 5% CO_2 for 24 h. If you are not using tracrRNA-ATTO550 and performing FACS analysis, incubate for 48 h and skip (Subheading 3.6).

3.6 Fluorescent-Activated Cell Sorting (FACS)

1. Trypsinize the cells as explained above 24 h after transfection.

2. Dilute 1.5×10^6 cells in 300 μL of sorting buffer.

3. For each sample to be collected, a 15 mL Falcon tube with 2 mL of FBS supplemented with 2 μL of antibiotics mix must be prepared.

4. Collect fluorescent cells for each crRNA.

5. Centrifuge at $218 \times g$ for 5 min.

6. Seed 1×10^5 cells per well of 6-well plate. One complete 6-well plate for each crRNA is enough.

7. Incubate the cells in an incubator at 37 °C and 5% CO_2. Change the medium every 2 days. Expand the cell culture and freeze several cryotubes of the total pool of transfected cells (*see* **Note 25**).

3.7 Generation of the Single-Cell Colonies

1. Trypsinize the cells as explained above.

2. Count the cells with the help of a hemocytometer.

3. Seed 150–200 cells in a 150-mm plate.

4. Incubate the cells in an incubator at 37 °C and 5% CO_2 for, at least 15 days (*see* **Note 5**).

5. Once the colonies can be seen with the naked eye, select and pick the colonies with a stripper micropipette and 150 μm tips. This should be done by observing colonies under a microscope inside a laminar flow-hood under sterile conditions (*see* **Notes 26 and 27**).

6. Put each colony in a 1.5 mL Eppendorf tube with 50 μL of trypsin and incubate at 37 °C for 5 min.

7. Individualize the cells by pipetting up and down several times. Seed the cells derived from the colonies into 24-well plates. It is not necessary to centrifuge previously.

8. Expand the culture and change the medium every 48 h. Once the cells are confluent, trypsinize the cells and split them into two wells of a 12-well plate. One of the wells will be used to isolate DNA for analysis, while the other will be used to freeze and/or expand the colony (*see* **Notes 28** and **29**).

3.8 Genomic DNA Extraction and RFLP Analysis

1. Trypsinize the cells as indicated above, centrifuge the cells at $300 \times g$ for 5 min and discard the supernatant.

2. Resuspend the pellet with 200 μL of PBS.

3. Isolate the DNA using QIAamp DNA Mini kit (Qiagen) following the manufacturer's protocol (*see* **Note 30**).

4. Quantify the DNA concentration in the isolate using Nanodrop One spectrophotometer.

5. Design primers using Primer3 software (http://bioinfo.ut.ee/primer3-0.4.0/) for amplification of the region surrounding the desired edited change (500–600 bp) (*see* **Note 31**).

6. Prepare pools of DNA mixing equal amounts (circa 50 ng) of DNA from individual colonies (4 or 5) in a PCR tube to obtain a final amount of 200 ng (*see* **Note 32**).

7. Perform PCR according to standard procedures. Use 200 ng of genomic DNA in a 50 μL PCR reaction with 1 μM of each primer, 200 μM of each dNTP, 2 unit of Taq polymerase and PCR buffer 1×. The PCR amplification program is as follows: 1 cycle with 5 min at 95 °C, 36 subsequent cycles of 25 s at 95 °C, 25 s at 50–60 °C (depending on the primers), and 40 s at 72 °C, with a final 7-min extension at 72 °C.

8. Run 5 μL of each sample in a 2% agarose gel with ethidium bromide (or other safer dye, such as GelRed or SYBR Safe) to confirm amplification.

9. Digest 5 μL of each amplified sample in a final volume of 20 μL with the restriction enzyme using the appropriate buffer and following the manufacturer's indications (*see* **Notes 33** and **34**).

10. Run the restriction reaction volume in a 2% agarose gel with ethidium bromide to visualize the resulting DNA bands.

11. Repeat the PCR and the restriction assay for each individual clone included in the pools for which a positive RFLP analysis is observed (in our example, digestion with PhsAI) (Fig. 3).
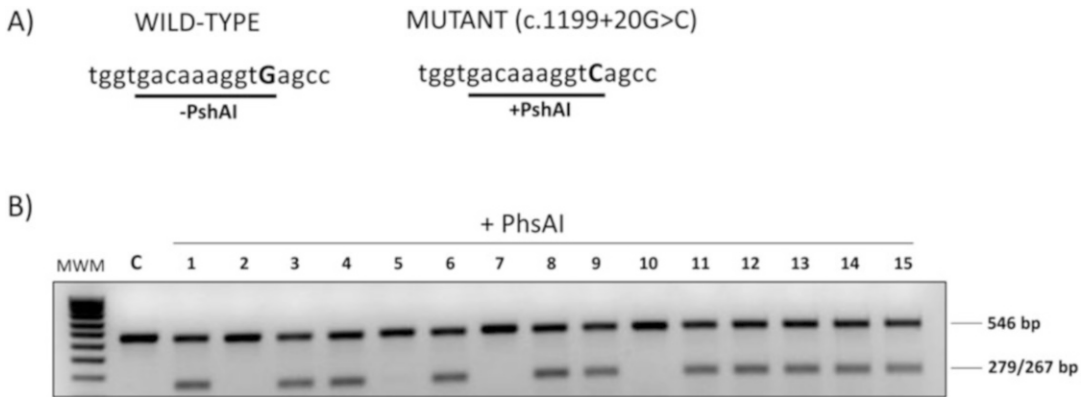
Fig. 3 RFLP analysis to monitor for gene edition. The wild-type and mutant sequences are shown in panel **a** and panel **b** is a representative gel showing RFLP analysis of single-cell colonies. Top bands correspond to the amplified PCR products and lower-sized bands correspond to the products obtained by digestion with PshAI enzyme due to the introduction of a restriction site with the point mutation c.1199 + 20G > C. *C* undigested control. Colonies 1, 3, 4, 6, 8, 9, 11–15 are positive and heterozygous (one allele edited)

3.9 Sequencing Analysis of Candidate Clones and Off-Targets Analysis

1. Using DNA from positive clones, perform a PCR to amplify the edited region and those regions where potential off-targets were identified by the software (*see* **Note 6**), using specific primers designed using Primer3 software (*see* **Note 35**).

2. Purify the PCR products using a PCR purification kit (*see* **Note 36**).

3. Prepare the mix for the sequencing reaction according to the instructions of the genomics facility and/or the sequencer. Carry out the sequencing with the forward and the reverse primers in separate reactions.

4. Analyze the sequences with the help of a chromatogram viewer (*see* **Notes 35** and **37**).

3.10 RNA Isolation

1. Once a correctly edited clone has been identified, expand the culture to obtain enough cells for RNA isolation.

2. Wash with PBS and trypsinize the cells. Centrifuge the cells at $16,000 \times g$ for 5 min and discard the supernatant. Cells can be frozen at -70 °C in this step.

3. Add 1 mL of Trizol per sample. Incubate the homogenate for 5 min at room temperature to achieve complete dissociation of nucleoprotein complexes.

4. Add 200 μL of chloroform. Mix by vortexing for 15 s and incubate for 2 min at room temperature.

5. Centrifuge at $12,000 \times g$ for 15 min at 4 °C to separate the phases.

6. Transfer the aqueous phase (upper and transparent) to a new Eppendorf tube.

7. Add 500 μL of 2-propanol. Mix by vortexing for 15 s.

8. Incubate the samples at room temperature for 10 min followed by an incubation of, at least, 20 min at −20 °C.

9. Centrifuge the samples at 12,000 × g for 30 min at 4 °C. Discard the supernatant.

10. Add 1 mL of 75% ethanol and wash the precipitate by vortexing (it can be stored in 75% ethanol for a week at 4 °C or a year at −20 °C).

11. Centrifuge at 12,000 × g for 5 min at 4 °C and discard the supernatant.

12. Centrifuge at 7500 × g for 1 min at 4 °C. Let the pellet dry at room temperature until they become transparent.

13. Dissolve the dry RNA in 30 μL of "Nuclease-Free Water" by pipetting and incubate it at least 10 min on ice (if it does not dissolve well it can be incubated 10–15 min at 55–60 °C).

14. Measure the concentration of the isolated RNA using Nano-Drop One and keep the samples at −70 °C until use.

3.11   RT-PCR and Sequencing Analysis to Confirm the Splicing Defect

1. For reverse transcription with the NZYRT System, use 1 μg of RNA, following the manufacturer's protocol. Random hexamers, oligo(dT) or vector-specific primer can be used. Mix RNA with NZYRT Master mix and NZYRT Enzyme mix in a final volume of 20 μL in PCR tubes, incubate 10 min at 25 °C, followed by 30 min at 50 °C, 5 min at 85 °C and cool to 4 °C.

2. Add 1 μL of NZY RNase H and incubate at 37 °C for 20 min.

3. Perform a standard PCR reaction using 1 μL cDNA and a final volume of 25 μL.

4. Run 5 μL of each sample in a 2% agarose gel.

5. Purify, quantify, and sequence the PCR product as explained above (*see* **Note 38**).

4   Notes

1. All the specific reagents for gene editing (crRNA, tracrRNA-ATTO550, ssDNA template, Cas9 Nuclease) explained in this protocol were obtained from IDT (Integrated DNA Technology). However, it is important to note that there are several other companies that sell these same products performing equally well. It is the researcher's decision to decide which company he wants to work with.

2. The method of delivery and/or the transfection reagent will depend on our cell line or on our preferences. Lipofection with Lipofectamine™ CRISPRMAX™ Cas9 and its Transfection

Reagent (Thermofisher) has been the method and the reagent chosen is this protocol.

3. Before starting the gene editing experiment, it is important to check that the chosen cell line expresses the gene of interest (mRNA and protein) and corresponds to a tissue relevant for your studies, i.e., splicing defect was observed in this type of cells, as splicing outcomes may depend on tissue-specific splice factors. It is also essential to take into account the organism from which the cell line is derived. For example, intronic sequences are not well conserved among species, and this is crucial when, for example, the aim is to study intronic splicing mutations.

4. It is necessary to verify the karyotype of the chosen cell line, to confirm it is normal, at least in relation to the pair of chromosomes where the gene that is going to be edited is located. Most established cell lines show aneuplodies and structural chromosomal alterations that will hinder the desired gene edition if the corresponding chromosome is affected. In our case, we tested a battery of human hepatoma cell lines, Hep3B, HepG2, Huh7 among others and selected an HepG2 cell line with two chromosomes 12 where the PAH gene is located. Karyotype analysis is a routine service offered by many human genetic diagnosis laboratories.

5. The chosen cell line should have the ability to form "single-cell colonies." This is necessary to isolate individual cells after transfection that will be subsequently expanded for genetic characterization to confirm and select correctly gene edited clones. There are different procedures for the generation of "single-cell colonies": (a) cell sorting: 1 cell/96-well-plate well using a cell sorter, (b) serial dilutions, and (c) seeding the cells at a high dilution (approximately 100 cells in one 150 mm plate).

    The election of one method or another will depend on the cell line, so it is advisable to test this before generating the colonies with the edited cells. In our hands, for example, HepG2 cells exhibited high mortality after sorting and plating in 96-well plates, so we selected option c. For some cell lines the use of conditioned medium (filtered culture medium collected from control cells) can aid the growth in the form of a colony derived from a single cell. The time of growth and appearance of single-cell colonies will depend on the type of cells you are working with. With HepG2 cells, colonies emerged and reached the correct size after circa 20 days.

6. It is advisable to have the region sequenced before starting the editing experiment to identify single-nucleotide polymorphisms in the specific cell line used which may affect the design

of RNA guides and DNA templates, as well as result in erroneous interpretation of the sequencing analysis of the edited clones (concluding there has been an extra change introduced during DNA repair after Cas9 reaction when it was already present in the sequence prior to editing).

7. There are multiple softwares for designing RNA guides for CRISPR assays. In our case we have used Breaking Cas software (http://bioinfogp.cnb.csic.es/tools/breakingcas) [19], which we find user-friendly, and the one offered by the company IDT, obtaining nearly identical results. In this sense, it is advisable to use and compare the results from at least two different softwares, to be sure that the selected guides are the most suitable.

8. It is advisable to test at least two RNA guides in a parallel and independent way. Generally, according to IDT, in 2/3 of the cases, sense sequence guides will work better than antisense guides. As we cannot predict which ones will do best for a given locus, we recommend testing both orientations.

9. As an optional step, you can pretest your RNA guides with an in vitro digestion after PCR amplification of the target region to confirm their efficiency (following IDT protocol).

10. SnapGene Viewer (https://www.snapgene.com/) has been the software used for visualization of sequences used in this project, location of crRNA, DNA templates, restriction sites, etc. and for sequence analysis of the individual edited clones. However, other programs and software can be used.

11. Cas9 nuclease cut site should be as close as possible to the sequence (nucleotide) which is to be edited. This requirement limits the region where we will design the RNA guides, especially if we want to introduce a point mutation as is the case here. It should be noted that this does not generally apply for the generation of a *knock-out* model, or in general, if we are not focused on introducing a mutation in a specific DNA position; in those cases the cut site can be in any position, so the design and choice of the RNA guide is much easier.

12. If possible, it is recommended to choose an RNA guide targeting the region that includes the nucleotide we intend to edit. Once the edition of that locus has occurred, the affinity of our RNA guide is reduced (because of a mismatch due to the mutation introduced), thus hindering possible reediting.

13. A ssDNA oligonucleotide containing the desired point mutation to be introduced is used as a template by the cell to repair the double strand break induced by Cas9 through HDR. The mutation of interest included in the ssDNA template should be in the middle of the sequence flanked by the homology arms. The length of the homology arms should be 35–40 nucleotides if it is a single-nucleotide change. Using longer homology arms

does not increase the homologous recombination success rate. However, for longer edits (e.g. insertion/deletion of several nucleotides), the length of the homology arms must also be increased.

14. For small insertions or single-nucleotide changes, ssDNA template is recommended. In other experimental situations (introduction of >100 nucleotide sequences) it may be advisable to use double stranded DNA templates.

15. In most gene editing protocols, introducing translationally silent sequence changes in the DNA template eliminating the PAM sequence is recommended, to avoid reediting of our target which may introduce unwanted changes. However, when dealing with intronic or exonic splice mutations, any extra change may alter the final splicing outcome so this should be avoided.

16. Standard desalting or HLPC are the purification methods recommended when ordering the ssDNA template. Also, especially in rich nucleases environments, phosphorothioate bonds (PS Bonds) at the extremes of the oligonucleotide are advisable, ideally putting at least two for each end of the template.

17. This protocol is written to use separate crRNA and tracrRNA. There is also the possibility of working with single guide RNA, where both are linked together, so this step will be different, refer to manufacturer's recommendations.

18. The resuspension volumes depend on the amount of purchased crRNA and tracrRNA. A table of equivalences for different quantities is available in the IDT protocols. It is important to keep in mind that the resuspended RNA oligonucleotides can be stored at -20 °C. The volumes and quantities referred to in this protocol are calculated for a 6-well plate, which has been the format used by the authors. Refer to the protocols available on the IDT website for other formats (e.g. 96-well plate).

19. The use of tracrRNA fused to the ATTO550 fluorophore is not strictly necessary but, in our hands, it was very useful for measuring transfection efficiency and to select transfected cells by fluorescence activated cell sorting (FACS) before clone generation. However, in cell types where transfection efficiency is known to be high/very high this step may be waived. In addition, there are certain cell types that are more prone to damage during the sorting process, so it would not be advisable to use this procedure to avoid increasing cell mortality. The protocol described can also be used for tracrRNA without ATTO550. In addition, it is important not to confuse transfection efficiency rate with editing efficiency, since a cell may have been transfected, but not edited. It is important to keep in mind that the success rate of the gene editing will not

only depend on the quality of the guide, but also on the transfection method, the locus we are editing, the cell type, etc.

20. The RNA Duplex can be prepared at a final concentration >1 μM and stored at −20 °C during, at least, 6 months. Before use, it should be diluted in Nuclease-Free Duplex Buffer to a working concentration of 1 μM.

21. IDT provides Cas9 nuclease at a stock concentration of 62 μM. It can be diluted in different buffers, such as PBS or Cas9 Working Buffer (20 mM HEPES, 150 mM KCl, pH 7.5). This will depend on our cell type. In our case we have used OptiMEM to dilute the Cas9 enzyme. It will be important to take these details into account when purchasing Cas9 nuclease from other companies.

22. The final concentration of the ssDNA template is variable depending on the cell type, delivery method, etc. In this case (transfection of HepG2 cells with Lipofectamine (CRISPR-Max)), a final concentration of 3 nM ssDNA template was used, following the manufacturer's recommendations. Transfecting higher amounts of ssDNA template does not ensure a higher rate of editing success. In addition, large amounts of DNA oligonucleotide can become toxic for the cells and increase cell mortality.

23. In initial experiments, it is advisable to perform the reverse transfection of each crRNA guide in triplicate (three 6-well plate wells/crRNA).

24. As with any transfection assay, it is advisable to split and pass the cells at least once after defrosting before starting the test.

25. Before generating colonies derived from a single cell, it is important to freeze the remaining total pool of cells transfected with each crRNA. In the event of any problem we could defrost those cells to generate the colonies again without the need to repeat the transfection.

26. Once the colonies have grown to a size allowing us to handle them efficiently, they must be expanded for analysis. You can select as many colonies as you can manage. You must consider that expansion, cultivation, and analysis of individual colonies require considerable effort and dedication. Normally we grow around 50–70 colonies for each crRNA used.

27. The system used to select colonies and pick them can be very variable. For example, cloning cylinders can be used or other methods of choice of the researcher.

28. We expanded the single-cell colonies in 24-well plates, but this can be modified according to the researcher's preferences and/or cell line characteristics using plates with different formats. In our case, once the cells are confluent, we divide each

well of the 24-well plate into two wells of a 12-well plate. It is important to keep accurate record of each duplicate, since one of them will be used to extract DNA for analysis, and the other will be used to freeze the colony and, in case it is the one selected, expand it for further characterization and use.

29. The analysis of the colonies derived from a single cell is necessary to identify edited ones. In our case, the point mutation that we are introducing generates a new restriction site for the PshAI enzyme. This is very useful to rapidly and easily screen by RFLP analysis for the presence of the introduced mutation, although the edited region must be verified by sequence analysis. In some applications, translationally silent changes are introduced in the donor template near the mutation to create/destroy a restriction site, thus allowing RFLP screening. However, this is not recommended for splicing mutations as any nearby change may alter the splicing outcome. Alternative approaches to evaluate edition efficiency include next-generation sequencing approaches or digital droplet PCR.

30. Other commercial kits or in-house methods can be used for DNA extraction.

31. Other alternative software and resources can be used with the same objective. Primers are designed to amplify the region with the desired change, which should ideally be in the middle of the amplicon, so after digesting with the corresponding enzyme and running the products in an agarose gel we can easily distinguish digested and undigested DNA bands, which will facilitate the identification of the positive clone. Care should be taken during primer design to ensure that there are no other restriction sites for the corresponding enzyme (in our case PshAI) within the amplicon.

32. Due to the high number of colonies, it is very laborious to analyze all of them individually. Therefore, it is advisable to make pools with DNA extracted from 4 or 5 colonies, mixing them to obtain 200 ng of total DNA. Once edition is observed in the RFLP analysis, colonies will then be analyzed individually.

33. It is not necessary to purify the PCR products before restriction enzyme digestion. Purification of PCR products does not improve digestion efficiency, as the PCR product is diluted enough so that the different components of the PCR reaction do not interfere with the enzymatic activity.

34. The conditions, temperatures, and times of the restriction reaction may depend on the enzyme and/or the trademark.

35. Usually, amplification and subsequent sequencing of the three possible off-targets with the highest scores identified by the software used is enough. Based on our experience we can

conclude that off-targets, although it is important to sequence and validate them, are not the biggest problem. However, we frequently found extra changes in the area near the edited nucleotide (on-target). In this sense, these errors have been the main problem and the cause of having had to discard many clones before finding the final positive one.

36. There are many commercially available kits for purification of PCR products. We routinely use Cycle Pure Kit (Omega).

37. Sequencing is necessary for the validation of the positive clone. And to discard off-target effects in the correctly edited clone. We should confirm that no extra changes have been made in the edited region. IMPORTANT: Do not confuse these random changes that CRISPR introduces when repairing the DSB in the DNA (on-target effects) with potential off-targets, which are locus to which our crRNAs can bind and induce a DSB in the DNA.

38. It is important, once the positive clone is selected and genetically analyzed, to carry out the phenotypic characterization as cellular model of the disease phenotype, to confirm that it accurately recapitulates the splicing defect, resulting (in our case) in the absence of protein and activity. To that aim, RT-PCR and cDNA sequencing, followed by Western blot analysis of PAH protein and PAH activity assay were performed. The specific analyses to be performed will depend on each case according to the aim of the study, but, in the case of splicing mutations they should include at least RT-PCR and subsequent cDNA sequencing analysis.

References

1. Montes M, Sanford BL, Comiskey DF, Chandler DS (2019) RNA splicing and disease: animal models to therapies. Trends Genet 35(1): 68–87. https://doi.org/10.1016/j.tig.2018.10.002

2. Lim KH, Ferraris L, Filloux ME, Raphael BJ, Fairbrother WG (2011) Using positional distribution to identify splicing elements and predict pre-mRNA processing defects in human genes. Proc Natl Acad Sci U S A 108(27): 11093–11098. https://doi.org/10.1073/pnas.1101135108

3. Sterne-Weiler T, Howard J, Mort M, Cooper DN, Sanford JR (2011) Loss of exon identity is a common mechanism of human inherited disease. Genome Res 21(10):1563–1571. https://doi.org/10.1101/gr.118638.110

4. Scotti MM, Swanson MS (2016) RNA mis-splicing in disease. Nat Rev Genet 17(1): 19–32. https://doi.org/10.1038/nrg.2015.3

5. Aartsma-Rus A (2016) New momentum for the field of oligonucleotide therapeutics. Mol Ther 24(2):193–194. https://doi.org/10.1038/mt.2016.14

6. Kim J, Hu C, Moufawad El Achkar C, Black LE, Douville J, Larson A, Pendergast MK, Goldkind SF, Lee EA, Kuniholm A, Soucy A, Vaze J, Belur NR, Fredriksen K, Stojkovska I, Tsytsykova A, Armant M, DiDonato RL, Choi J, Cornelissen L, Pereira LM, Augustine EF, Genetti CA, Dies K, Barton B, Williams L, Goodlett BD, Riley BL, Pasternak A, Berry ER, Pflock KA, Chu S, Reed C, Tyndall K, Agrawal PB, Beggs AH, Grant PE, Urion DK, Snyder RO, Waisbren SE, Poduri A, Park PJ, Patterson A, Biffi A, Mazzulli JR, Bodamer O, Berde CB, Yu TW (2019) Patient-customized oligonucleotide therapy for a rare genetic disease. N Engl J Med 381(17):1644–1652. https://doi.org/10.1056/NEJMoa1813279

7. Aartsma-Rus A, Corey DR (2020) The 10th oligonucleotide therapy approved: golodirsen for Duchenne muscular dystrophy. Nucleic Acid Ther 30(2):67–70. https://doi.org/10.1089/nat.2020.0845

8. Kemaladewi DU, Maino E, Hyatt E, Hou H, Ding M, Place KM, Zhu X, Bassi P, Baghestani Z, Deshwar AG, Merico D, Xiong HY, Frey BJ, Wilson MD, Ivakine EA, Cohn RD (2017) Correction of a splicing defect in a mouse model of congenital muscular dystrophy type 1A using a homology-directed-repair-independent mechanism. Nat Med 23(8):984–989. https://doi.org/10.1038/nm.4367

9. Schneller JL, Lee CM, Bao G, Venditti CP (2017) Genome editing for inborn errors of metabolism: advancing towards the clinic. BMC Med 15(1):43. https://doi.org/10.1186/s12916-017-0798-4

10. Bollen Y, Post J, Koo BK, Snippert HJG (2018) How to create state-of-the-art genetic model systems: strategies for optimal CRISPR-mediated genome editing. Nucleic Acids Res 46(13):6435–6454. https://doi.org/10.1093/nar/gky571

11. Maule G, Casini A, Montagna C, Ramalho AS, De Boeck K, Debyser Z, Carlon MS, Petris G, Cereseto A (2019) Allele specific repair of splicing mutations in cystic fibrosis through AsCas12a genome editing. Nat Commun 10(1):3556. https://doi.org/10.1038/s41467-019-11454-9

12. Xu S, Luk K, Yao Q, Shen AH, Zeng J, Wu Y, Luo HY, Brendel C, Pinello L, Chui DHK, Wolfe SA, Bauer DE (2019) Editing aberrant splice sites efficiently restores beta-globin expression in beta-thalassemia. Blood 133(21):2255–2262. https://doi.org/10.1182/blood-2019-01-895094

13. Doudna JA (2020) The promise and challenge of therapeutic genome editing. Nature 578(7794):229–236. https://doi.org/10.1038/s41586-020-1978-5

14. Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Soria E (2005) Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J Mol Evol 60(2):174–182. https://doi.org/10.1007/s00239-004-0046-3

15. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337(6096):816–821. https://doi.org/10.1126/science.1225829

16. Knott GJ, Doudna JA (2018) CRISPR-Cas guides the future of genetic engineering. Science 361(6405):866–869. https://doi.org/10.1126/science.aat5011

17. Morava E, Rahman S, Peters V, Baumgartner MR, Patterson M, Zschocke J (2015) Quo vadis: the re-definition of "inborn metabolic diseases". J Inherit Metab Dis 38(6):1003–1006. https://doi.org/10.1007/s10545-015-9893-x

18. Martinez-Pizarro A, Dembic M, Perez B, Andresen BS, Desviat LR (2018) Intronic PAH gene mutations cause a splicing defect by a novel mechanism involving U1snRNP binding downstream of the 5′ splice site. PLoS Genet 14(4):e1007360. https://doi.org/10.1371/journal.pgen.1007360

19. Oliveros JC, Franch M, Tabas-Madrid D, San-Leon D, Montoliu L, Cubas P, Pazos F (2016) Breaking-Cas-interactive design of guide RNAs for CRISPR-Cas experiments for ENSEMBL genomes. Nucleic Acids Res 44 (W1):W267–W271. https://doi.org/10.1093/nar/gkw407