Zebrafish as research model to study Gaucher disease: Insights into molecular mechanisms

Lelieveld, L.T.

CHAPTER 4

CRISPR/Cas9 mediated genome-editing in zebrafish

Methodology aspects and protcol

Abstract

The CRISPR/Cas9 system enables a relatively easy manner of specific genome-editing in cells and organisms. This chapter describes the genome-editing approaches underlying the zebrafish loss-of-function mutants characterized in subsequent experimental chapters. First of all, a detailed protocol for the generation of CRISPR/Cas9 modified zebrafish is provided, including notes and considerations. Next, the protocol is applied to generate zebrafish with a loss-of-function mutation in proteins involved in lysosomal storage disorders. For the target proteins Gba1, Gba2, Gpnmb, Asah1a, Asah1b, Npc1 and Cln8, information on chromosome location and obtained mutant gene- and protein sequences of the zebrafish knockouts are provided. RNA expression and lipid analysis of mutant larvae of Gpnmb, Npc1 and Cln8 confirmed that functional gene knockout have been generated, while the impact of defective Gba1, Gba2, Asah1a and/or Asah1b is described in chapters 5, 6 and 7.

In addition, pilot experiments using the Tol2 transposase technique were undertaken to introduce an exogenous target DNA sequence in the zebrafish genome. The coding sequence of human *GBA* was introduced in the zebrafish *gba1*$^{-/-}$ genetic background. In addition, the coding sequence of zebrafish prosaposin was introduced harbouring a point mutation in the saposin C region thought to impair a disulfide bridge.

In conclusion, this chapter describes a relatively easy, straightforward and cheap method to generate genome-edited zebrafish using CRISPR/Cas9 and Tol2 transposase technology. It is advised to closely monitor the experimental steps to achieve high mutagenesis efficiencies with low mortality rates and off-target effects. This could reduce the number of required animals while maintaining a mixed genetic background.

Introduction

In the twentieth century, knowledge on heredity increased with the rediscovery of Gregor Mendel's pea plant research, driving the elucidation of the structure of the double DNA helix and cracking the genetic code[1]. It became apparent that DNA, including smaller or bigger changes, is transmitted from parent to offspring and that these mutations can have detrimental effects on the health state of the offspring. For years, understanding the molecular mechanisms underlying human inherited disorders was limited by the inability to specifically change the endogenous genome of model organisms[2,3]. Early genome-editing endeavours were labour intensive and not specific. Mutagenesis induced by the chemical N-ethyl N-nitrosourea (ENU) required extensive screenings to determine a mutation in a target gene without off-target effects. The spontaneous homologous recombination technique, commonly used in mouse embryo-derived stem cells, was very inefficient and also showed a high frequency of undesired genome-editing events[2,3].

A new era of specific genome-editing commenced when several new programmable nucleases were developed starting with Zinc finger-FokI nucleases (ZFNs), transcription activator-like effector nucleases (TALENS) and, more recently, CRISPR-associated (Cas) enzymes in combination with Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR). The first two techniques utilize protein fusions composed of the DNA cleavage domain of FokI endonuclease with multiple modules of either Zinc fingers[4] or TALE proteins[5]. One particular module could recognize either a 3-bp DNA code (Zinc finger) or one single base (TALE protein), which required combinatorial assembly of multiple modules. Re-design and engineering of new sets of proteins for each target was one of the main limitations of both techniques. Therefore, adaptation of the CRISPR/Cas system has revolutionized the research community, enabling specific and facile genome editing for all researchers.

The CRISPR/Cas system has been adapted from an endogenous prokaryotic immune response. CRISPRs and Cas enzymes are used by prokaryotes as adaptive immune system to cleave invading foreign genetic elements[6]. Prokaryotic Cas enzymes are guided by short CRISPR RNAs (crRNA) which are transcribed from non-repeating spacer DNA sequences originating from viruses and other mobile genetic elements[7]. The spacer sequences acquired by the prokaryote are highly similar in regions called protospacer-adjacent motifs (PAMs) and critical for the CRISPR system[8]. An important step in the discovery and design of the genome-editing technique was the reprogramming of Cas9 from *Streptococcus thermophilus* and *S. pyogenes*[9], which enabled targeting of an intended bacterial DNA sequence[10,11]. These findings were followed by the discovery that 'humanized' versions of *S.pyogenes* Cas9 system also worked in eukaryotic systems[12-14] and the technique expanded quickly to other cell types and whole organisms such as mice, drosophila and zebrafish[15].

The two short RNAs were combined into a single guide RNA (sgRNA)[12]. The sgRNA is composed of a specific sequence of 20 nucleotides at the 5' end of the sgRNA, which binds specifically to the DNA target site by Watson-Crick base pairing, followed by a RNA sequence at the 3' end that forms the complex with the Cas9 endonuclease through RNA hairpins. The target DNA sequence requires the Cas9 specific PAM sequence of 'NGG', immediately adjacent to the 20 bp sequence complementary to the sgRNA, to guide and program the Cas9 protein–sgRNA complex[12].

In general, off-target effects are one of the main problems of genome-editing tools. The 10-12 bp sequence adjacent to the PAM sequence is called the 'seed region' and requires perfect complementarity with the sgRNA. Mismatches in the 8 bp sequence distal to the PAM are tolerated[10] even though it does not necessary lead to DNA cleavage *in vivo*, making off-target effects difficult to predict[16,17]. Therefore both on-target activity (efficiency) and off-target effects (specificity) of the system can be dependent on sgRNA sequence (including GC content, secondary structures and sequence length), cell type, Cas9 enzyme concentration and sgRNA/Cas9 delivery methods[18].

Zebrafish are an attractive vertebrate model to study human diseases, with characteristics such as large progeny, fast embryonic development including functional liver, kidney and tissue barriers and a high occurrence (83%) of functional zebrafish orthologs of human disease-related genes[19]. Generation of a zebrafish knockout using CRISPR/Cas9 technology can be achieved with a relatively straightforward method, a high rate of success and generated in a matter of months instead of years. Therefore many zebrafish disease models have been successfully generated[20].

This chapter provides a detailed workflow, including notes and considerations, for the generation and high-throughput screening of CRISPR/Cas9 mediated gene knockout zebrafish models. The protocol relies on non-homologous end-joining (NHEJ), the most dominant repair mechanism of vertebrates for repairing DNA double-strand breaks. Insertions and deletions (indels) are introduced due to end resection of the complementary strands and misaligned repair. Next, the protocol is used to establish zebrafish models of lysosomal storage disorders. Seven target genes are selected and obtained mutations described. In addition, preliminary biochemical data is described of mutant larvae which are not studied in detail in subsequent chapters. In addition, the Tol2 transposon approach is used to introduce two target sequence, human *GBA* and *prosaposin* with a point mutation, in the zebrafish genome.

Workflow

The procedure is summarized in **Figure 1**. Part 1 requires approximately one week including target design, RNA synthesis and injection of the sgRNA/Cas9 mRNA mixture. Part 2 consists of screening the F_0 and F_1 adult zebrafish for desired modifications using the high resolution melt (HRM) analysis and Sanger sequencing. For the F_2 generation, and subsequent generations, only a HRM analysis is required, accelerating screening of the fish. Several outbred crossings are preferred, before incrossing, to obtain a zebrafish line with a healthy genetic background. The whole procedure of obtaining a zebrafish line with a homozygous mutation comprises approximately 6-16 months: depending on the sexual maturation of the fish and the preference of crossing carrier zebrafish with wildtype (WT) zebrafish (outbred) to the F_3 generation before crossing the carriers with each other (inbred).

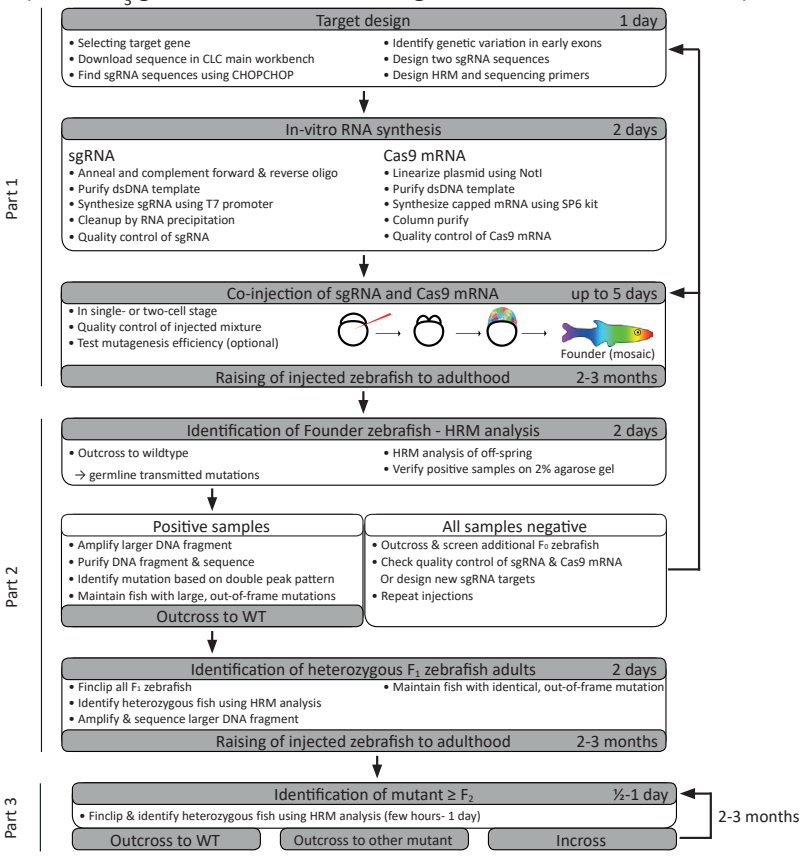


Figure 1 | Protocol for generation of CRISPR/Cas9 mediated knockout zebrafish

Part 1 | Target design, RNA preparation & injection

Genetic variation

As starting point, Ensembl and NCBI databases are used to identify zebrafish orthologues of the target gene. Zebrafish underwent an additional whole-genome duplication and approximately 20% of human genes have co-orthologues in the zebrafish genome[19,21,22]. Therefore, a protein BLAST of the human protein of interest is performed to identify additional zebrafish co-orthologues of the human target. Sequences can be found and downloaded from the NCBI database using the program CLC main workbench, including whole chromosome, mRNA or protein sequences. The chromosome of interest is particularly useful because it includes chromosome positions and annotations such as coding sequences (CDS), exons, introns and non-coding sequences (example in **Figure 2A**). The obtained sequence does not include genetic variations, while genetic variation between different zebrafish lines and within a laboratory stock is large[23]. Therefore, genetic variations of the target region are manually annotated using the Ensembl Variation database before the design of sgRNA sequences and primers (**Figure 2A**). In specific cases, it is recommended to sequence the target region as we have noticed that various indels are not annotated in the variation database. Double peaks in the obtained sequence trace indicate heterozygous SNPs or indels and confirm amplification of both alleles.

SgRNA design

Several CRISPR/Cas9 target design web tools are available and we use the CHOPCHOP web-tool[24,25]. It offers the option to enter via RefSeq, gene ID or fasta sequences as well as various options for target specific regions, PAM sequence, efficiency score, 5' requirements for sgRNA and self-complementarity. We use the gene ID, together with the newest available genome (DanRer11/GRCz11 in 2019) and 'NGG' as Cas9-specific PAM sequence. Our method requires the target sequences to start with 5' GG in order to have efficient *in vitro* RNA transcription due to promoter requirements of the utilized T7 RNA polymerase. Therefore, the option 'GN' or 'NG' is applied as 5' requirements as well as the option to check for self-complementarity when the leading nucleotides are replaced with 'GG'. All other settings of the web-tool are left on default.

Two target sequences are chosen with regard to certain guidelines: high efficiency, low number of off-targets, located in close proximity of each other, in the first few exons common to all known splice variants (exon 1-3) and preferably in the middle of the exon (**Figure 2B**). Two efficient target sequences in close proximity are ideal for screening since both targets can be screened using the same primer pair. Mutants with target sequences in the middle of the exon showed more straightforward prediction of the *in silico* mutation and a lower chance of exon-intron mutations.

RNA preparation and microinjection

A cloning- and PCR-free method is used to obtain the double strand template for RNA synthesis[26]. The target-specific oligo consists of a GCG clamp, the T7 promoter, the chosen target sgRNa sequence and a sequence complementary to the constant reverse oligo of 80-nt, that contains the Cas9 binding region (**Figure 2C**). Target- and constant reverse oligo are annealed, extended with DNA polymerase and purified double-strand DNA template

is used for T7 *in vitro* RNA transcription. A smaller batch is transcribed than described in the supplier's protocol and sgRNA is purified by ammonium acetate/ethanol precipitation instead of column purification to increase the yield. The quantity and integrity of the purified sgRNA should always be analysed on agarose gel, whereby one or two bands of around 100 and 200 bp are detected (**Figure 2D,** purified sgRNA). RNA synthesis is repeated when a high RNA concentration is measured, but no bands are detected (low vs high RNA yield in lane 1 vs lane 2 and 3 respectively, **Figure 2E**). A zebrafish codon-optimized Cas9[27] is used containing nuclear localization signals at both ends of the coding sequence of Cas9 with a SP6 promoter for *in vitro* generation of capped and poly-adenylated mRNA. Cas9 mRNA is column-purified and the integrity analysed on agarose gel, detecting multiple bands due to secondary RNA structures (**Figure 2E**).

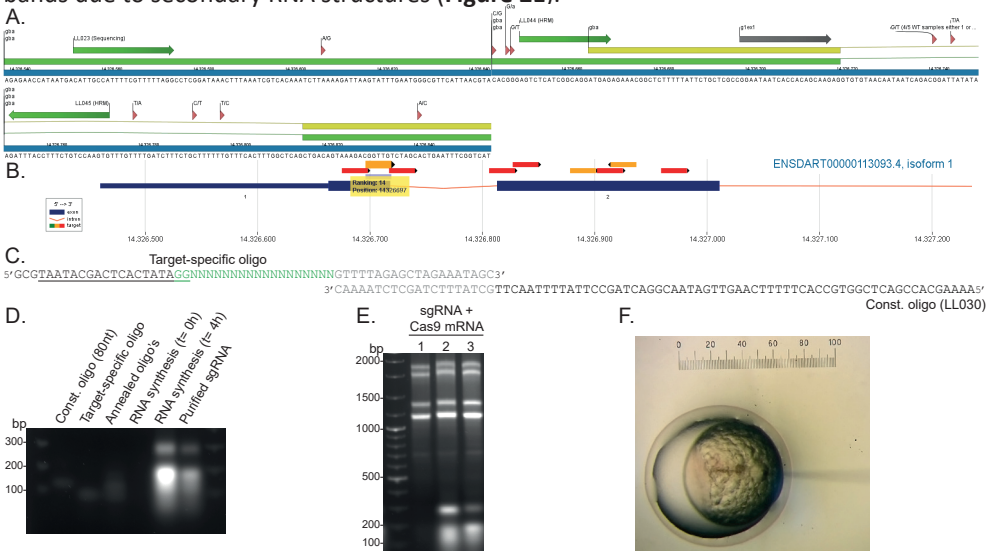


Figure 2 | Target design, RNA preparation and microinjection
(A) Visualization of the *gba1* locus using CLC main workbench, with the genomic sequence (annotated as blue box), mRNA sequence (green box) and coding sequence (CDS, yellow box). Genetic variation is added manually (red triangles) as well as used primer sequences (green) and sgRNA target sequence (grey). **(B)** sgRNA of the zebrafish *gba1* gene determined by the CHOPCHOP web-tool. **(C)** Target-specific oligo (top sequence), with the T7 promoter underlined, target sequence in green and the constant reverse oligo (bottom) with the sequences complementary to each other in grey. **(D)** SgRNA generation on 1% agarose gel with from left to right: only constant oligo, only target-specific oligo, annealed oligos, RNA sample at t= 0 h, RNA sample after 3 h of transcription without degradation of the dsDNA template and purified sgRNA running as two bands between 100 and 200 bp due to secondary structures. **(E)** Different sgRNA/Cas9 injection mixtures with comparable RNA concentrations measured by Nanodrop. **(F)** Microinjection of sgRNA/Cas9 with PhenolRed to visualize proper injection into the single-cell stage of the embryo.

Typically, a mix of 150pg sgRNA/200 pg Cas9 mRNA is injected in the cell cytoplasm or yolk of the single- or two-cell stage embryos (**Figure 2F**). Higher concentrations of sgRNA/Cas9 as well as injection in the cell cytoplasm could give higher mutagenic efficiencies, but might also result in poor survival of founder larvae. After injection, the integrity of the injected sgRNA and Cas9 mRNA should always be analysed on agarose gel. When the intensity of the injected RNA mix is low, synthesis and injections should preferably be repeated before raising the injected larvae to adulthood (**Figure 2E**).

Part 2 & 3 | Germline transmitted mutations & screening

High-resolution Melt (HRM) Analysis

Several methods are available to evaluate sgRNA mutagenic efficiency in injected zebrafish embryos and adult founder zebrafish, including T7 endonuclease I analysis, the Surveyor assay or fluorescence PCR[27,28]. We adapted a high-resolution melt (HRM) analysis method to evaluate mutations, using the specific melting characteristics of a particular DNA fragment[29]. A short fragment around the region of interest (100-150 bp) is amplified in the presence of a dsDNA-intercalating dye. After amplification, the PCR product is heated gradually from 55 °C to 95 °C. With increasing temperature, the strands of the dsDNA sequence are melted and the fluorescence decreases (**Figure 3A**). The specific combination of nucleotides in the DNA sequence results in a characteristic melting curve. NHEJ typically generates small indels which influences the melting curve, although differences between WT and mutant DNA are generally not large enough to analyse the obtained melting temperature. Therefore, fluorescent data of each sample is normalized using the start fluorescence at 55°C (**Figure 3B**) and the difference between the target sample and a reference sample (WT genomic DNA) is calculated and visualized as maximum or minimum (**Figure 3C** and **D** for *gba1*, and **3F** for *gba2*). The deletion or insertion is subsequently verified by loading the PCR amplicon on a 2-2.5% agarose gel (**Figure 3G** for *gba2*).

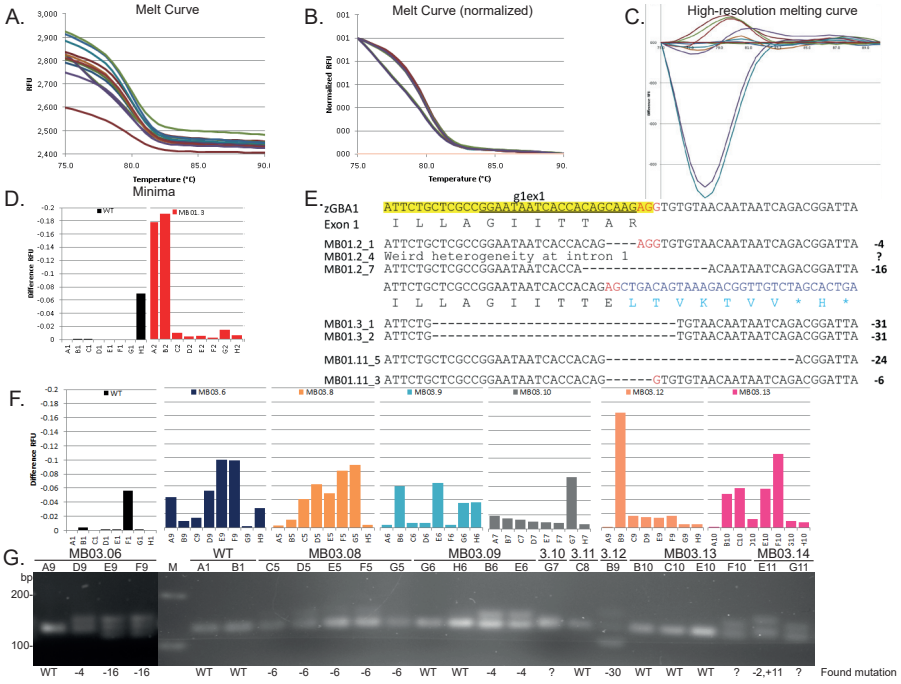


Figure 3 | Screening of founder fish by high-resolution melt (HRM) analysis

(**A-E**) Screening of offspring of founders injected with *gba* sgRNA (LL031). **(A)** The melting curve plotted as fluorescence (in RFU) against temperature. **(B)** Fluorescence is normalized using the fluorescence at 55 °C. **(C)** The difference between the target sample and reference sample is plotted against the temperature and **(D)** the minimum (or maximum) visualized as bar graph. **(E)** The mutation is verified and reconstituted using Sanger sequencing. (**F,G)** Screening of offspring of founders injected with *gba2* sgRNA (LL039) with mutations (minima) visualized in a bar graph **(F)** and verified on 2-2.5 % agarose gel and by Sanger sequencing **(G)**.

Unlike other methods, the HRM assay only requires materials and equipment routinely used in laboratories performing RT-PCR analysis such as multiwell PCR plates, multichannel pipettes, a buffer with dsDNA-binding dye and a real-time PCR instrument. Moreover, a simple excel sheet is used to calculate the differences instead of the available but expensive HRM programs. The simplicity and pace of the method is advantageous because fish can be screened and put back in the aquarium in only a day, thereby reducing the stress of individual housing. A potential disadvantage of this method is that very large deletions might not be detected as these could remove the primer binding site(s) of the small amplicon, generally located 20-25 bp from the target sgRNA cut-site (**Figure 2A**).

Evaluation of injected adult zebrafish (founders) and subsequent generations

Injection of embryos typically generates mosaic mutagenesis in the developing larvae; cells can have unmodified DNA or different mutations. At this stage, the HRM analysis can only be used to evaluate the mutagenic efficiency of the sgRNA. A visible difference and accompanying smeared band on agarose gel, indicates a highly efficient sgRNA as most of the cells of the developing embryo have a mutated target sequence. Subsequently, larvae of the same injected batch are raised to adulthood.

Adult founder fish are crossed with a genetically WT strain, albino for easy crossing. Individual embryos (n= 8) of each outcross are screened by HRM analysis and verified on gel to find founder fish with germ-line transmitted mutations (**Figure 3**). A larger fragment is amplified from positive samples to obtain information of the precise mutation using Sanger sequencing. A fragment of 300-700 bp is amplified with primers, located at a minimum of 50 bp from the target site for optimal sequencing results (**Figure 2A**). If primer design in the flanking introns is difficult, due to genetic variation, an exon in close proximity is used to amplify a larger but more specific fragment. With the determined mutation, the modified *in silico* mRNA sequence is evaluated for the prediction of an early stop codon in the protein sequence (**Figure 4**). We typically keep founder fish with large, out-of-frame mutations leading to an early stop codon and preferably easily detected with the HRM assay.

For the F1 generation, one- to two suitable founder fish are crossed with WT (not of albino background) and raised. Genomic DNA of finclipped adults is used in an HRM assay followed by sequencing of the positive samples. Fish with different mutations are generally found, indicating that germ cells of the founder fish can have different mutations. Adult F_1 zebrafish with the same mutation are maintained, WT fish or fish with suboptimal mutations are discarded.

It is important to be aware of potential off-target effects that might occur. Our WT zebrafish are maintained as genetically heterogeneous outbred stocks and therefore we always cross several generations to WT (>F3 generation, outbred) before generating a homozygous adult zebrafish line (inbred). This common approach results in a diverse genetic background and reduces the selection of fish that show undesired off-target modifications besides the desired genetic modification. Heterozygous adult fish can be crossed with each other and offspring used in pilot biochemical experiments to confirm generation of a functional KO. For experiments, we always use and compare genetically WT zebrafish larvae from crossings of mutant heterozygotes (inbred, ¼ ratio) and from crossings of WT stocks.

Results and preliminary findings

The described CRISPR/Cas9 protocol was used successfully for generation of loss-of-function zebrafish mutants in order to study lysosomal storage disorders. To date, seven targets are selected: lysosomal β-glucosidase (glucocerebrosidase (GCase), gene name *gba1*) impaired in Gaucher disease (GD), the membrane associated β-glucosidase GBA2 (*gba2*), the Gaucher biomarker transmembrane glycoprotein NMB (*gpnmb*), both acid ceramidase co-orthologues (*asah1a* and *asah1b*), NPC intracellular cholesterol transporter 1 (*npc1*) defective in Niemann-Pick disease type C1[30] and CLN8 (*cln8*) which results in neuronal ceroid lipofuscinosis when defective[31] and is reported as genetic modifier in GD[32]. The genetic location of each target in the zebrafish genome is given in **Table 1**, as well as gene and protein annotations.

Table 1 | Selected targets related to lysosomal storage disorders

Information of protein and gene names, including representative RefSeq and Uniprot codes with genetic locations, transcript (in bp), number of amino acids (aa), predicted protein size (kDa) and amino acid identity compared to the human orthologue (in %). Some of the generated zebrafish are characterized in other chapters, while results of pilot experiments of other zebrafish lines are presented in this chapter.

Protein name	Gene	Chr	mRNA: refseq	bp	Uniprot	aa	kDa	Identity	
Glucosylceramidase	*gba*	16	XM_682379	4173	E7EZM1	518	58	58%	Ch. 5-7
Non-lysosomal glucosylceramidase	*gba2*	7	XM_005172332	3676	E7F65W0	851	96.5	66%	Ch. 5, 6
Glycoprotein nmb	*gpnmb* X1	19	XM_009294247	5185	E7F7J7	627	70	33%	Pilot exp.
N-acylsphingosine amidohydrolase Acid ceramidase 1a	*asah1a*	14	NM_001006088	1578	Q5XJR7	390	44.6	61%	Ch. 6, 7
Acid ceramidase 1b	*asah1b*	1	NM_200577	1717	Q6PH71	395	44.6	60%	Ch. 6, 7
CLN8, transmembrane ER and ERGIC protein	*cln8*	17	NM_001114588	1473	F1QYC0	289	33.8	61%	Pilot exp.
Niemann-Pick disease, type C1	*npc1* X2	2	XM_005163659	5059	F1QNG7	1276	143	69%	Pilot exp.

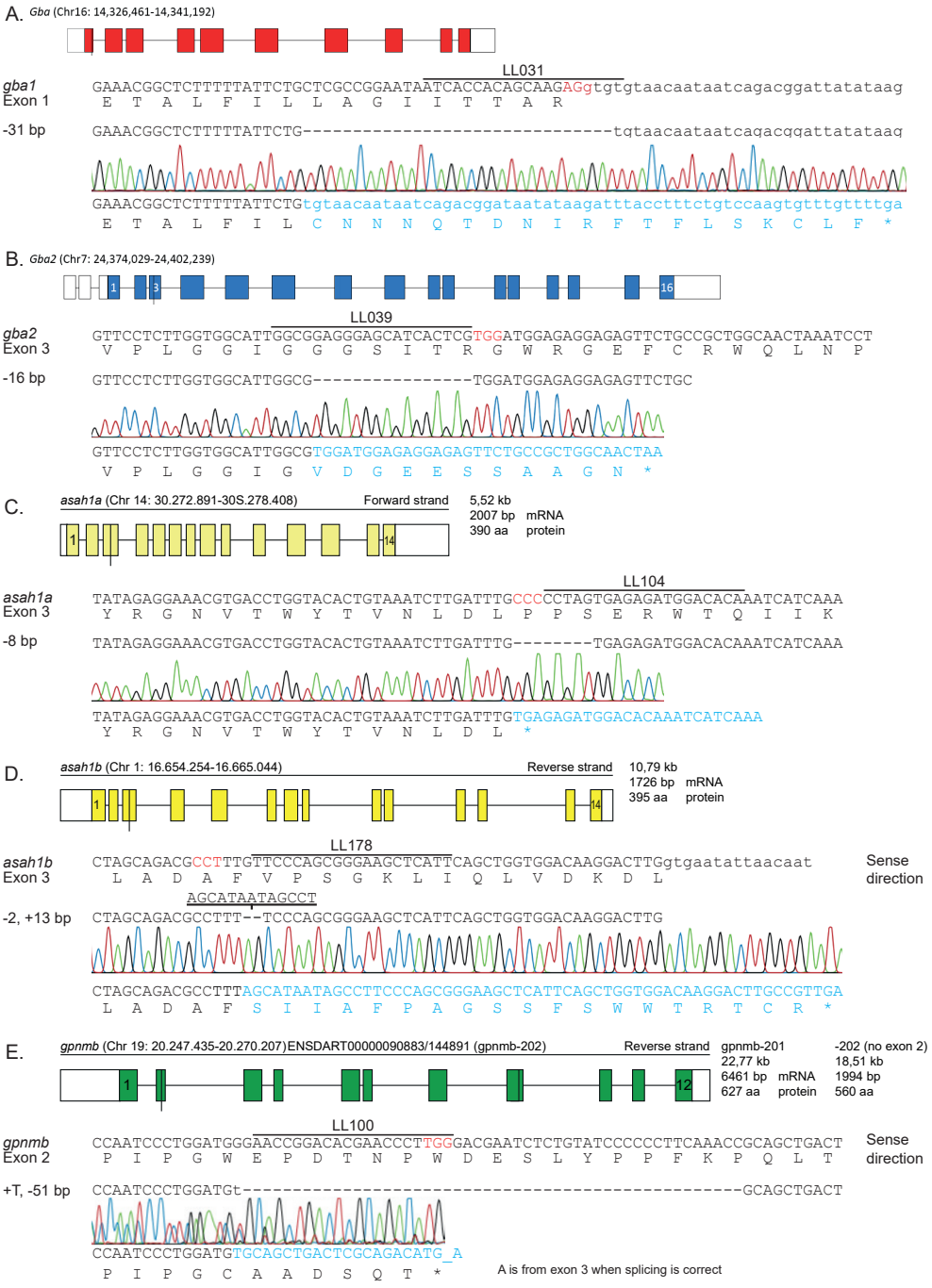
For every target, two sgRNA sequences were considered, given in **Supplementary Table 4**. Injection of sgRNA and Cas9 mRNA was performed as described above and multiple adult founder and F1 zebrafish were screened until a suitable modification was found. To obtain the genetic zebrafish knockouts, more than 500 embryos were injected, approximately 130 founder F_0 fish were crossed and screened, while roughly 750 zebrafish of the F_1 generation were finclipped and screened. From the total screened founder fish, approximately 25% transmitted mutations to the offspring, although not all founder F_0 fish showed suitable modifications. Founder fish transmitting small (-1 bp, TMB010) or in-frame mutations (-6 bp, MB002) were discarded, while founder fish with large and frameshift germline transmitted mutations were selected (**Table 2**). F_1 zebrafish were finclipped, genotyped and selected based on the desired modifications (**Table 2**). Selected F_1 zebrafish were used for subsequent crossings generating both outbred, inbred and double mutant zebrafish lines (**Supplementary Table 1**).

Schematic representations of the target locus, location of the sgRNA, target sequence and obtained mutation with sequence trace of the verified genetic KO zebrafish are given in **Figure 4**.

Table 2 | Overview screening of founder and F1 zebrafish to obtain genetic KO zebrafish

Overview of screened founder and F1 zebrafish, including gene target, location of the sgRNA, sgRNA code as well as the code of the zebrafish line, the number of screened fish, the number of positives and mutations for both the founder fish and F1 zebrafish line. The sgRNA resulting in zebrafish with suitable modifications are given in bold. Mut = mutation, ORF = open reading frame, bp = basepairs

			Founder fish				F1 fish				
Target	**Exon**	**sgRNA**	**Code**	**# screen**	**# of pos.**	**Mutation**	**Code**	**# Screen**	**# of pos.**	**Mutation**	**Notes**
gba1	1	LL031	MB001	21	4	2/4 splice site mut.	TMB003	79	38	17x -31 bp	
gba2	3	LL038	MB002			All -6 bp; no ORF mut.	-				
	3	LL039	MB003	20	8	4x ORF mut.	TMB001	37	13	7x -16 bp	
gpnmb	2	LL098	TMB005	8	3	2x ORF mut.	TMB020	34	0	-	Mutations suboptimal: only kept TMB024
							TMB021	103	19	-6, -7, -8 bp	
	2	LL099	TMB006	2	2	1x mut. in intron	-				
	2	LL100	TMB007	2	1	Not verified	TMB024	94	41	7x +T, -50 bp	
asah1a	3	LL104	TMB008	6	4	3x ORF mut.	TMB022	86	31	20x -8 bp	
	3	LL105	TMB009	8	3	Not verified	-				
asah1b	4	LL108	TMB010	8	1	Only -1 bp	TMB023	84	0		
	4	LL109	TMB011	12	2	1x ORF mt.	TMB025	31	0		
	4	LL178	TMB033	5	1	1x ORF mut.	TMB042	39	2	1x +11 bp	
	4	LL179	TMB034	7	0		-				
npc1	2	LL253	TMB048	4	1	Not verified	TMB064	45	15	4x -3, +16 bp	Kept +16bp F1 fish
	2	LL254	TMB049	1	0		-				
cln8	3	LL206	TMB046	12	2	Not verified	TMB068	88	44	34x -4 bp	TMB068 only -4 bp mutation; too small to screen
							TMB083	24	7	7x -23 bp	
	3	LL207	TMB047	14	0		-				


A.
Gba (Chr16: 14,326,461-14,341,192)
LL031
gba1
Exon 1
GAAACGGCTCTTTTTATTCTGCTCGCCGGAATAATCACCACAGCAAGAGgtgtgtaacaataatcagacggattatataag
E T A L F I L L A G I I T T A R
-31 bp
GAAACGGCTCTTTTTATTCTG----------tgtaacaataatcagacggattatataag
GAAACGGCTCTTTTTATTCTGtgtaacaataatcagacggataatataagatttacctttctgtccaagtgtttgttttga
E T A L F I L C N N N Q T D N I R F T F L S K C L F *
B.
Gba2 (Chr7: 24,374,029-24,402,239)
LL039
gba2
Exon 3
GTTCCTCTTGGTGGCATTGGCGGAGGGAGCATCACTCGTGGATGGAGAGGAGAGTTCTGCCGCTGGCAACTAAATCCT
V P L G G I G G G S I T R G W R G E F C R W Q L N P
-16 bp
GTTCCTCTTGGTGGCATTGGCG----------TGGATGGAGAGGAGAGTTCTGC
GTTCCTCTTGGTGGCATTGGCGTGGATGGAGAGGAGAGTTCTGCCGCTGGCAACTAA
V P L G G I G V D G E E S S A A G N *
C.
asah1a (Chr 14: 30.272.891-30S.278.408)
Forward strand
5,52 kb
2007 bp mRNA
390 aa protein
LL104
asah1a
Exon 3
TATAGAGGAAACGTGACCTGGTACACTGTAAATCTTGATTTGCCCCCTAGTGAGAGATGGACACAAATCATCAAA
Y R G N V T W Y T V N L D L P P S E R W T Q I I K
-8 bp
TATAGAGGAAACGTGACCTGGTACACTGTAAATCTTGATTTG--------TGAGAGATGGACACAAATCATCAAA
TATAGAGGAAACGTGACCTGGTACACTGTAAATCTTGATTTGTGAGAGATGGACACAAATCATCAAA
Y R G N V T W Y T V N L D L *
D.
asah1b (Chr 1: 16.654.254-16.665.044)
Reverse strand
10,79 kb
1726 bp mRNA
395 aa protein
LL178
asah1b
Exon 3
CTAGCAGACGCCTTTGTTCCCAGCGGGAAGCTCATTCAGCTGGTGGACAAGGACTTGgtgaatattaacaat
L A D A F V P S G K L I Q L V D K D L
Sense direction
AGCATAATAGCCT
-2, +13 bp
CTAGCAGACGCCTTT--TCCCAGCGGGAAGCTCATTCAGCTGGTGGACAAGGACTTG
CTAGCAGACGCCTTTAGCATAATAGCCTTCCCAGCGGGAAGCTCATTCAGCTGGTGGACAAGGACTTGCCGTTGA
L A D A F S I I A F P A G S S F S W W T R T C R *
E.
gpnmb (Chr 19: 20.247.435-20.270.207)ENSDART00000090883/144891 (gpnmb-202)
Reverse strand
gpnmb-201
22,77 kb
6461 bp mRNA
627 aa protein
-202 (no exon 2)
18,51 kb
1994 bp
560 aa
LL100
gpnmb
Exon 2
CCAATCCCTGGATGGGAACCGGACACGAACCCTTGGGACGAATCTCTGTATCCCCCCTTCAAACCGCAGCTGACT
P I P G W E P D T N P W D E S L Y P P F K P Q L T
Sense direction
+T, -51 bp
CCAATCCCTGGATGt----------GCAGCTGACT
CCAATCCCTGGATGTGCAGCTGACTCGCAGACATG_A
P I P G C A A D S Q T *
A is from exon 3 when splicing is correct

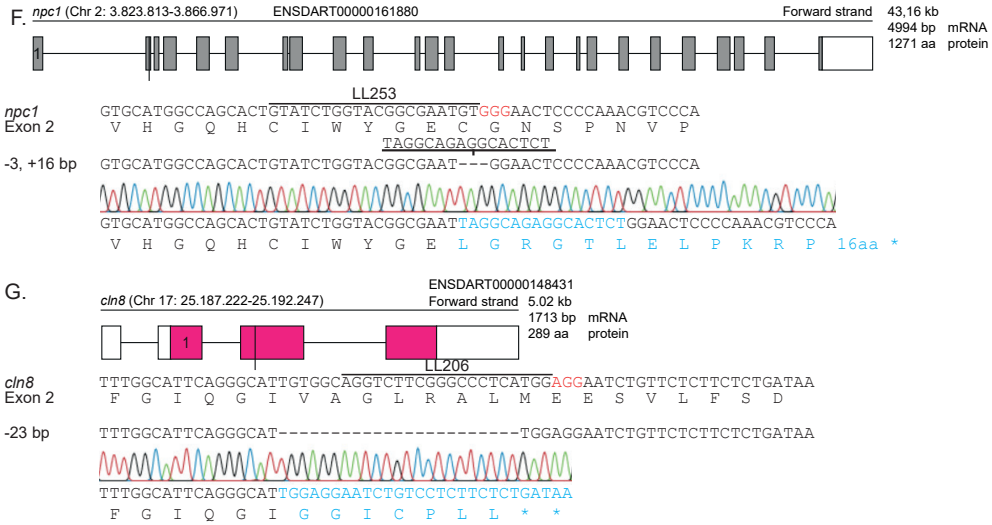


Figure 4 | Schematic representation of the obtained CRISPR/Cas9 mediated genetic modifications.
(A) *gba1* **(B)** *gba2* **(C)** *asah1a* **(D)** *asah1b* **(E)** *gpnmb* **(F)** *npc1* and **(G)** *cln8*. Top panel: schematic representation of the respective gene with the location and sequence of the sgRNA target given in Supplementary Table 1. Middle panel: DNA sequence of the WT target sequence, with the exon sequence in uppercase and intron sequence in lowercase, the sgRNA sequence lined above, PAM site in red and the protein sequence shown below. Lower panel: the mutation as obtained from the sequence trace and the predicted translated protein sequence. Of note, the sequence trace of a heterozygous *gpnmb* sample is displayed.

On average, ten fish are screened to obtain a suitable mutation. The number of screened fish necessary appeared to be dependent on the location of the sgRNA, the quality of the injected sgRNA and the internal sequence of the target. For example, the used sgRNA for *gba1* (LL031, **Figure 2B**) was in close proximity of the exon-intron transition (**Figure 2A**), therefore prediction of the *in silico* mutation using the sequence trace of heterozygous F_1 zebrafish was difficult. The 31 base pair deletion was thought to remove the endogenous splice-site and this hypothesis was confirmed by sequencing cDNA of homozygous samples. Another example (LL038 of *gba2*), resulted in a high ratio of positive samples with the same 6 bp deletion, most likely due to high sequence similarity and microhomology-directed repair using small repeats in the target sequence. Together these findings affirm the importance of part 1 of the described protocol, especially the design and quality of the sgRNA. The possible repair mechanisms, on the other hand, are more difficult to predict, although machine learning is used to predict repair profiles and improve sgRNA design[33].

As described before, mutant zebrafish are crossed with WT fish for several generations to obtain a heterogeneous outbred stock and thereby segregate unlinked mutation, lower off-target effects and improve general health. Typically, multiple individual zebrafish are crossed with different WT zebrafish, preferably from different WT stocks, in order to prevent selection of genetic variations in individuals.

Whole-genome sequencing was not considered to study off-target effects because of the known high genetic variation in laboratory zebrafish. Instead, the top off-targets of the *gba1* and *gba2* sgRNA sequences were sequenced. These off-targets were predicted using the CRISTA algorithm[34]. No mutations were found at the predicted CRISPR/Cas9 cut site in several individual homozygous mutants of different crossings (**Supplementary Figure 1**). Several indels and SNPs were found in WT and mutant fish, confirming the high genetic variation described above.

Glycosphingolipid abnormalities in *npc1* and *cln8* knockouts

Knockouts of *gba1, gba2, asah1a* and *asah1b* are topics of the investigations in chapters 5-7. The *gpnmb*, *npc1* and *cln8* KO larvae have been generated and analysed in this chapter in pilot biochemical experiments. The protein gpNMB is remarkably induced in human Gaucher disease (GD) patients, GD mouse and zebrafish models[35,36]. At present, the physiological function of gpNMB during lysosomal stress is unknown. The ER protein CLN8 has been suggested as candidate modifier of clinical severity of GD. CLN8 is a member of the neuronal ceroid lipofuscinosis (NCL) associated proteins and is thought to be involved in the regulation of ceramide synthesis and glycosphingolid trafficking[37]. Patients with a defect in CLN8 develop progressive epilepsy with mental retardation (EPMR) and brains displayed reduced levels of glycosphingolipids including ceramide, galactosylceramide and lactosylceramide[38]. In Niemann Pick disease (NPC) the efflux of cholesterol from lysosomes is impaired, due to defects in either NPC1 or NPC2. As a result, secondary (partial) deficiencies in sphingomyelinase (ASMase) and glucocerebrosidase cause accumulation of sphingomyelin, glucosylsphingosine (GlcSph), glucosylceramide (GlcCer) and glucosylcholesterol (GlcChol) in these cholesterol-laden lysosomes[39,40].

First, glycosphingolipid levels of offspring of *npc1*$^{+/-}$ and *cln8*$^{+/-}$ inbreds were measured (**Figure 5A**). No significant differences were found for *cln8*$^{-/-}$ larvae compared to heterozygous or WT. Surprisingly, no change in GlcSph, GlcCer or GlcChol levels was found in the *npc1*$^{-/-}$ larvae as in Npc1 mouse models[39,40]. The absence of secondary (glyco) sphingolipid abnormalities might be explained by the presence of maternal Npc1 in the developing offspring, as observed for GCase[41] or the young age of the larvae in general. It suggests that the burden of lysosomal cholesterol is not sufficient enough to develop the proposed secondary deficiencies of ASMase and GCase.

Next, the impact of GCase deficiency in *npc1* or *cln8* genetic backgrounds was determined. Offspring of n*pc1* and *cln8* carriers were incubated with a GCase-specific inhibitor for 5 days. Interestingly, *cln8*$^{-/-}$ zebrafish larvae showed significantly reduced GlcSph accumulation compared to their *cln8*$^{+/+}$ and *cln8*$^{+/-}$ siblings, while GlcCer and GlcChol levels in the Cln8 deficient zebrafish were comparable to their WT and heterozygous counterparts. No apparent differences were found between *npc1*$^{-/-}$ zebrafish larvae and

their *npc1*$^{+/+}$ siblings, both showing increased GlcSph, GlcCer and GlcChol levels upon GCase inactivation. GlcSph levels appeared lower in the npc1$^{-/-}$ larvae, however limited biological *npc1* KO samples were evaluated in these pilot lipid analyses.

RT-qPCR analysis revealed a reduction of *npc1* mRNA expression in *npc1*$^{-/-}$ larvae compared to WT siblings. This finding indicates that the generated mutation has led to a gene knockout, albeit it is not known if a functional protein knockout is generated. Raising the Npc1 deficient zebrafish to adulthood would enable analysis of cholesterol and (glyco) sphingolipid abnormalities in older fish in combination with observation of symptoms such as hepatosplenomegaly and neurological abnormalities as reported for other *npc1* KO zebrafish [42].

No biochemical readout exists to study the endogeneous zebrafish Gpnmb protein and commercially available zebrafish anti-gpnmb antibodies were not reactive (data not shown). Quantification of levels of the *gpnmb* transcript from individual larvae revealed a reduction of mRNA levels in *gpnmb*$^{-/-}$ larvae compared to WT siblings (**Figure 5B**). However homozygous *gpnmb* KO adult fish showed no phenotype (data not shown). Upregulation of gpNMB expression has been reported in Gaucher cells of GD patients as well as mouse macrophages with induced lysosomal stress[36,43]. Crossing these *gpnmb* KO fish with animals having a *gba1* mutation could reveal a particular function of Gpnmb during lysosomal stress of macrophages.

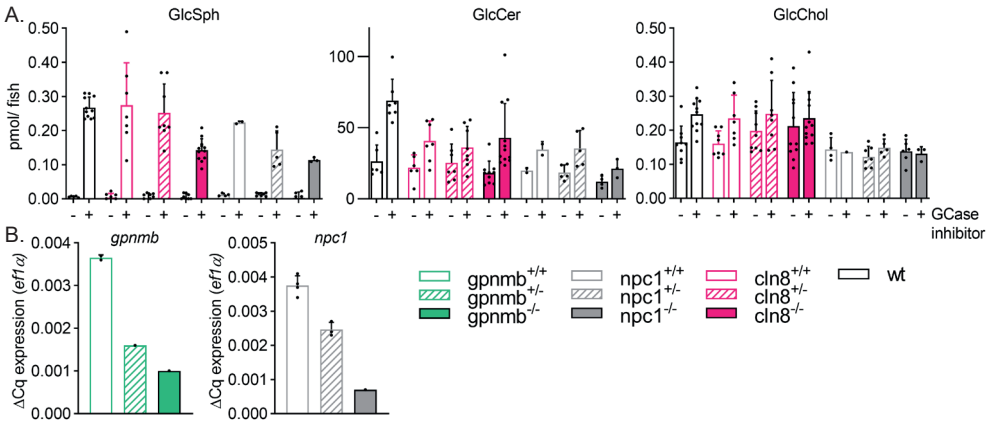


Figure 5 | Biochemical impact of Cln8, Npc1 and Gpnmb deficiency

(A) GlcSph, GlcCer and GlcChol levels were determined of untreated individual zebrafish larvae (5 dpf, - GCase inhibitor) or embryos treated with the GCase specific inhibitor ME656 (10 μM) for 5 days (+ GCase inhibitor) in pmol/fish; WT (n = 9-11); data from [41], *cln8*$^{+/+}$ (n = 7-8), *cln8*$^{+/-}$ (n = 8-9), *cln8*$^{-/-}$ (n = 12), *npc1*$^{+/+}$ (n = 2-4), *npc1*$^{+/-}$ (n = 5-7) and *npc1*$^{-/-}$ (n = 3-6). Data is depicted as mean ± SD. **(B)** mRNA levels of *gpnmb* or *npc1* individual zebrafish larvae (5 dpf) as dermined by RT-qPCR using specific primers. *Gpnmb*$^{+/+}$ (n = 2), *gpnmb*$^{+/-}$ (n = 1), *gpnmb*$^{-/-}$ (n = 1), *npc1*$^{+/+}$ (n = 4), *npc1*$^{+/-}$ (n = 3) and *npc1*$^{-/-}$ (n = 1).

Random integration using the Tol2 transposase technique

As final part of the genome-editing approaches, the Tol2 transposase technique was used to integrate exogenous DNA. Random integration of a donor sequence is mediated by the Tol2 transposon element of the medaka fish (*Oryzias latipes*)[44]. Donor constructs, with the target sequence flanked by the required Tol2 sequences, are integrated by the transposase protein as a single copy with high germline transmission (> 30%)[45,46]. Two targets were chosen for proof-of-concept studies. First, it was appreciated that introduction of human *GBA* in the zebrafish *gba1* KO background could enable evaluating the biological impact of human GCase in the zebrafish GD model. Furthermore, it was rationalized that certain genes would not be suitable for CRISPR/Cas9 mediated genome editing. One such example is the gene encoding the prosaposin precursor protein which is cleaved to four activator lipid-binding proteins: saposin A, -B, -C and –D. Only Saposin C is required for optimal intralysosomal activity of GCase and defects in saposin C cause symptoms similar to Gaucher disease[47-49]. The zebrafish genome appeared to have one prosaposin (*psap*) gene, including annotated sequences for the activator lipid-binding proteins saposin A, -B, -C and –D. It was rationalized that CRISPR/Cas9 genome editing according to the above established protocol would generate a stop codon in the *psap* gene and thereby likely impact the other saposins as well. An alternative approach was considered by introducing a point mutation in the *psap* sequence (T>G, **Supplementary Figure 2**) which would result in a Cys to Gly amino acid substitution and impair one of the predicted disulfide bridges of zebrafish saposin C (predicted Sap C region underlined in **Supplementary Figure 2**)[50].

Tol2 constructs were generated using the reported three-insert Gateway which combines three entry vectors into a destination vector with the required flanking Tol2 sequences[45]. The generated vectors included a ubiquitin promoter, followed by the target sequence and a polyadenylation signal in a destination vector with a cell-specific reporter cassette (γ-cryst:CFP). The reporter cassette allows easy selection of transgenic carriers because positive zebrafish embryos express the cyan fluorescent protein in their lens. A mixture of 20 pg donor vector and 150 pg SP6-generated capped and polyadenylated transposase mRNA was injected in single- or two-cell stage zebrafish embryos. As for the CRISPR/Cas9 method, the balance between high efficiency, by injecting in the cell cytoplasm with a high concentration of donor vector/Tol2 mRNA, and poor survival of injected zebrafish embryos should be experimentally defined. The concentration of the mixture in these pilot experiments was too high since more than 80% of the injected embryos died before they reached the larval age. However the success rate of Tol2 integration final rate was also high, because 2 out of 6 hGBA injected adult founder fish produced offspring with CFP expression in the lens, 1 of the 2 tested $pSAP^{WT}$ zebrafish and 1 of the 3 tested $pSAP^{mut}$ injected founder fish produced offspring showing CFP signal. Glycosphingolipid analysis of the generated h*GBA*:*gba*$^{-/-}$ zebrafish larvae revealed a significant lower accumulation of glucosylsphingosine (GlcSph), indicating that human GCase is expressed and functionally active (Chapter 5). The impact of the $pSAP^{mut}$, on the other hand, has not been studied yet. It is necessary to mutate the endogeneous zebrafish *psap* gene using the CRISPR/Cas9 protocol before an impact of the $pSAP^{mut}$ is expected.

Discussion

A very practical advantage of zebrafish as research model, as compared to mice, is the speed and convenience of generating gene knockouts. This could allow evaluation of zebrafish with multiple genetic traits in a matter of months. This chapter provides a detailed protocol for generation of gene knockouts in zebrafish by means of CRISPR/Cas9 technology and efficient HRM analysis screening. In addition, established mutations are described, including those in the zebrafish *gba1, gba2, asah1a, asah1b, gpnmb, cln8* and *npc1* genes.

The protocol includes several notes and considerations for achieving a high mutagenesis efficiency without off-target effects which allows reduction of the amount of necessary injected founders and F_1 animals. To increase the probability of raising founders with a mutation, it is recommended to inject high quality RNA and evaluate the mutagenesis efficiency of multiple different sgRNAs for the same target, prior to baby raising. In general, it is advised to raise about 20 larvae to adulthood, as not all founder fish show germline transmitted mutations. Moreover, mutations do not always lead to frameshifts and some mutations could be small, which complicates the HRM analysis. When a suitable mutation is obtained in the F_2 generation, a reduction of necessary animals in subsequent generations could be achieved by performing mini finclips on larvae of 3-5 dpf[51]. The HRM analysis is a very suitable method as it allows screening in only a few hours.

The Tol2 transposon technique was successfully used to introduce exogenous DNA in the zebrafish genome. A limitation of this approach is the random integration of the target sequence as well as the regulation of expression by a non-endogenous promoter. These limitations can be circumvented by so-called knockin approaches after the generated CRISPR/Cas9 mediated double strand DNA break. This could allow for subsequent introduction of specific point mutations, generation of endogenous target proteins fused to a fluorescent reporter tag or combining an endogenous promoter with a fluorescent protein to express a reporter under the control of that promoter. Small precise modifications of the zebrafish genome have been achieved by directing the homology-directed repair (HDR) pathway to repair the CRISPR/Cas9 mediated double-stranded DNA break with a single-stranded oligodeoxynucleotide (ssODN) with the intended mutation[52]. For larger and more complex modifications, donor plasmids with long homology arms flanking the insert are frequently used to direct HDR with the donor sequence[53], while another approach uses the dominant NHEJ repair system to insert donor DNA sequences at the target site[54,55]. Both the HDR and NHEJ approach have reported only low efficiency of germline transmission (± 5-7%)[52,53,55]. At present, we have failed to obtain zebrafish with a precise mutation in *gba1* with the reported protocols, although only limited attempts have been performed and ideally many more embryos should be injected and screened.

In all subsequent chapters, zebrafish gene knockouts are used with stable and known out-of-frame mutations. Using the high efficiency of the CRSIPR/Cas9 system, it is also possible to study the biochemical impact of the knockout directly by evaluating the injected F_0 zebrafish embryos. It has been shown that injection of the sgRNA in complex with a Cas9 ribonucleoprotein can achieve sufficiently high levels of bi-allelic gene disruption in the injected F_0 zebrafish embryos, named crispants[56-58]. For example, 5 dpf crispants of the lysosomal enzyme β-hexosaminidase showed a 95% reduction in enzyme activity and similar phenotypic manifestations as larvae with a stable mutation[58]. Therefore, the crispant approach would be very efficient in combination with the sensitive biochemical analysis by means of activity-based probes (Chapter 3) and LC-MS/MS (Chapter 5). This approach could enable studying the impact of gene disruption in a matter of days, thereby accelerating the research on zebrafish models for lysosomal storage disorders.

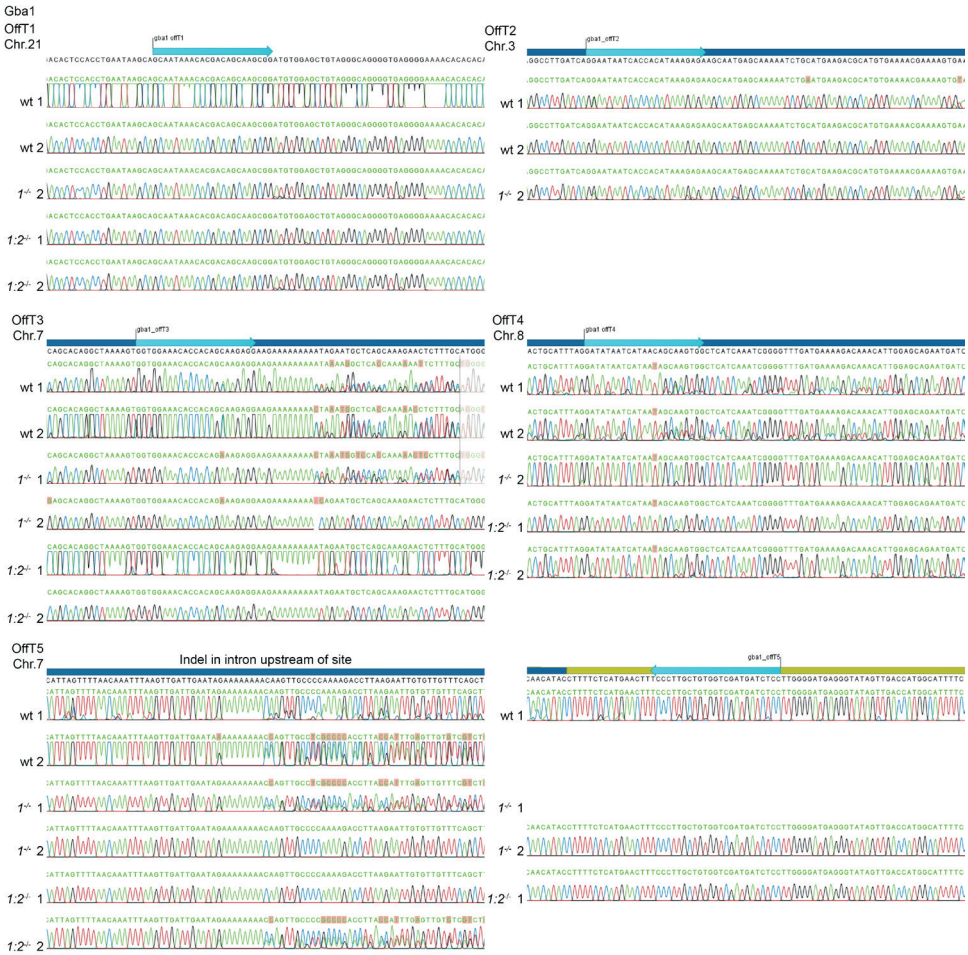
Acknowledgments

Prof. Annemarie Meijer is kindly acknowledged for providing the pCS2-nCas9n plasmid as well as Dr. Vincenzo Torraca for initial CRISPR/Cas9 protocols. Ralf Boland and Dr. Bjørn Koch are acknowledged for their help with initial injections and Dr. Rolf Boot for generation of the tol2 constructs for human *GBA* as well as WT and mutant *psap*.

Supplementary data

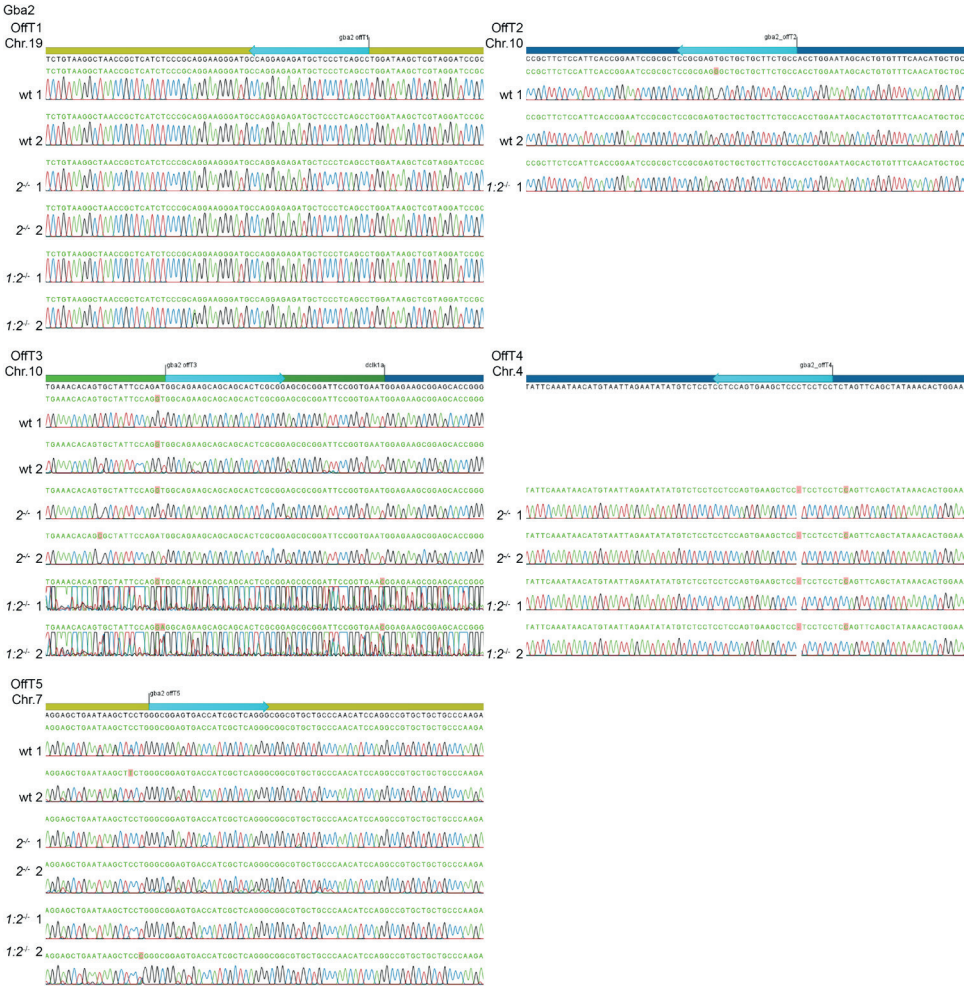
Supplementary Table 1 | Available single-, double- and three-double knockouts

Target	sgRNA	Mutation	Lines
Gba1	LL031	- 31 nt	*gba1*$^{+/-}$, *gba1*$^{+/-}$ *in mpeg:GFP, gba1*$^{+/-}$ *in 2CLIP, gba1*$^{+/-}$*:gba2*$^{-/-}$, *gpnmb*$^{-/-}$*:gba1*$^{+/-}$, *asah1a*$^{-/-}$*:gba1*$^{+/-}$, *asah1b*$^{-/-}$*:gba1*$^{+/-}$, *asah1a*$^{+/-}$*:asah1b*$^{+/-}$*:gba1*$^{-/-}$
Gba2	LL039	- 16 nt	*gba2*$^{+/-}$, *gba2*$^{-/-}$, *gba1*$^{+/-}$*:gba2*$^{-/-}$
gpnmb	LL100	+ T, -50 bp	*gpnmb*$^{+/-}$
Asah1a	LL104	- 8 bp	*asah1a*$^{+/-}$, *asah1a*$^{+/-}$*:asah1b*$^{+/-}$, *asah1a*$^{-/-}$*:gba1*$^{+/-}$, *asah1a*$^{+/-}$*:asah1b*$^{+/-}$*:gba1*$^{+/-}$
Asah1b	LL178	- 2, +13 bp	*asah1b*$^{+/-}$, *asah1a*$^{+/-}$*:asah1b*$^{+/-}$, *gsah1b*$^{-/-}$*:gba1*$^{+/-}$, *asah1a*$^{+/-}$*:asah1b*$^{+/-}$*:gba1*$^{+/-}$ *asah1b*$^{-/-}$*npc*$^{+/-}$
Npc1	LL253	- 3, +16 bp	*asah1b*$^{-/-}$*npc*$^{+/-}$
Cln8	LL206	- 23 bp	*Cln8*$^{+/-}$



Supplementary Figure 1

No mutations were found in *gba1*$^{-/-}$, *gba2*$^{-/-}$ or *gba1*$^{-/-}$*:gba2*$^{-/-}$ samples on the predicted off-targets using the CRISTA algorithm[34] for LL031 (*gba1*) and LL039 (*gba2;* Next page). Predicted off-target sequences are depicted with a light blue box, intron sequence with blue box, mRNA sequence with green box and coding sequence with yellow box.



► **Suplementary Figure 2 |** Coding sequence of prosapin with the (T/G) point mutation in red, leading to the suspected loss-of-function Cys > Gly mutation of saposin C. The saposin C region of zebrafish prosaposin is predicted using a blastp alignment of the reported human saposin C sequence. The predicted saposin C region of the zebrafish is underlined.

```
ATG ATG CTT CTC ACG CTT CTT TTG GTC ACC ACA GCT GTG GCA AGT CCC CTG TTG GGA ACG
M   M   L   L   T   L   L   L   V   T   T   A   V   A   S   P   L   L   G   T
GAG CAG TGT GCC CGT GGT CCC CCC TAC TGG TGC CAG AAT GTC AAG ACC GCT TCC CTT TGT
E   Q   C   A   R   G   P   P   Y   W   C   Q   N   V   K   T   A   S   L   C
GGT GCT GTC CAG CAC TGC CAA CAG AAT GTG TGG AAC AAG CCT CAG ATG AAA ACT GTG CCA
G   A   V   Q   H   C   Q   Q   N   V   W   N   K   P   Q   M   K   T   V   P
TGT GAT CTG TGC AAA GAG GTG TTG GTG GTA GTG GAG CAG CTG CTG AAG GAC AAT GTA ACC
C   D   L   C   K   E   V   L   V   V   V   E   Q   L   L   K   D   N   V   T
GAG AGC GAA CTC CTC GGG TAT TTG GAG AAG GCA TGT CAG CTG ATC CCT GAT GAG GGC CTG
E   S   E   L   L   G   Y   L   E   K   A   C   Q   L   I   P   D   E   G   L
GCT AAC CAG TGC AAG GAG ATT GTG GAC AAC TAC TTC CCA GTT CTC ATG GGC ATC ATC CAA
A   N   Q   C   K   E   I   V   D   N   Y   F   P   V   L   M   G   I   I   Q
GGA GAG CTG GAT GAC CCT GGT GTA GTG TGT GGC GCT TTG GGT CTG TGC GTA TCC CAG CAG
G   E   L   D   D   P   G   V   V   C   G   A   L   G   L   C   V   S   Q   Q
GCA GCT CTG GCT AAA GCT CAG CTC ACT TCC AAC GAG ATC CCT CAA GTG GAC CTG AAT CAA
A   A   L   A   K   A   Q   L   T   S   N   E   I   P   Q   V   D   L   N   Q
CGC GTC AGC CCC TTC CTG CTG AAC ATC CCA CAG CTG CTC TAT CCT GAA GAG AAA AGA GAG
R   V   S   P   F   L   L   N   I   P   Q   L   L   Y   P   E   E   K   R   E
ACC CCT AAA CAG AAA GGT GAT GTG TGC CAG GAC TGC GTT ACG TTC ATC ACT GAC ACT CAG
T   P   K   Q   K   G   D   V   C   Q   D   C   V   T   F   I   T   D   T   Q
GAT GAA GCC AGA ATC AAT TCT TCC TTT ATC AAC ACT CTG ATT GCA CAG GTG GAG AAC CAG
D   E   A   R   I   N   S   S   F   I   N   T   L   I   A   Q   V   E   N   Q
TGT GAG CTT CTG GGA CCT GGA ATG TCT GAT ATG TGC AAG GAG TAC ATC AGC CAG TAC GGG
C   E   L   L   G   P   G   M   S   D   M   C   K   E   Y   I   S   Q   Y   G
CCT CTG GTC TTC CAG CAG CTC ATG TCT ATG CAA CCC AAG GAC ATC TGT GCT CGC GCT GGC
P   L   V   F   Q   Q   L   M   S   M   Q   P   K   D   I   C   A   R   A   G
TTC TGC CCT ACT AAA CAA AAG TCT GTG CCC ATG GAG AAG CTG CTG CCT GCC AAA TCC ATC
F   C   P   T   K   Q   K   S   V   P   M   E   K   L   L   P   A   K   S   I
CCT GCT GTC AAG ATG TTC CCT GCG GTT AAA GTT GAG AAA CCG GTT GCG ACC ATG CCT GCT
P   A   V   K   M   F   P   A   V   K   V   E   K   P   V   A   T   M   P   A
AAG AAC CTG GTG CGT GTG CGT GAC TCT CCT CAG TGT GCC ATC TGT GAA TAT GTG ATG AAG
K   N   L   V   R   V   R   D   S   P   Q   C   A   I   C   E   Y   V   M   K
GAA ATT GAG AAC ATG ATT CAG GAT CAG ACC TCT GAG GCA GAG ATC GTG CAG GCT GTG GAG
E   I   E   N   M   I   Q   D   Q   T   S   E   A   E   I   V   Q   A   V   E
AAG GTC TGC AAC ATT TTG CCC TCC ACA CTC ACT GCT CAG TGC AAG GAC CTG ATT GAA ACC
K   V   C   N   I   L   P   S   T   L   T   A   Q   C   K   D   L   I   E   T
TAC GGC CAG GCC ATC ATC GAT CTC CTT GTG CAG GAG GCC GAT CCC AAA ACG GTC GGC TCT
Y   G   Q   A   I   I   D   L   L   V   Q   E   A   D   P   K   T   V   G(C) S
TTC CTT GCA CTC TGC AGT GGG GTC AGC CAT GTC CCT GTG ATG GAT AAG CAG CAT TTC GCA
F   L   A   L   C   S   G   V   S   H   V   P   V   M   D   K   Q   H   F   A
GCA GGT GGT TTC TGT GAT GTG TGC AAG ATG GCT GTG CGC TAT GTG GAC GGG ATC CTG GAG
A   G   G   F   C   D   V   C   K   M   A   V   R   Y   V   D   G   I   L   E
CAG AAC GCC ACT CAG TCT GAG ATC GAG GAG GCC GTG CTG AAA GTC TGC AGC TTC CTG CCT
Q   N   A   T   Q   S   E   I   E   E   A   V   L   K   V   C   S   F   L   P
GAC GCC GTC AAA GAT GAG TGC AAC CAG CTG ATT GAG CAG TAC GAG CCT CTG CTG GTG CAG
D   A   V   K   D   E   C   N   Q   L   I   E   Q   Y   E   P   L   L   V   Q
CTG CTG CTT CAG ACC CTC GAC CCT GAC TTT GTC TGC ATG AAG CTG GGA GCA TGT CCT GAG
L   L   L   Q   T   L   D   P   D   F   V   C   M   K   L   G   A   C   P   E
GCC GTG CAG AGG CTG CTG GGA TTA AAT CAG TGC AGC TGG GGG CCT GCA TAC TGG TGT AAG
A   V   Q   R   L   L   G   L   N   Q   C   S   W   G   P   A   Y   W   C   K
AAC GTG CAG ACC GCC GCT CGC TGT AAC GCC CTG AAC CAC TGC AGG CGT CAC GTT TGG AGC
N   V   Q   T   A   A   R   C   N   A   L   N   H   C   R   R   H   V   W   S
TAA  < 1563
*
```

Experimental procedures

Zebrafish husbandry and lines - Zebrafish were housed and maintained at Leiden University, the Netherlands, according to standard protocols and described in Chapters 3, 5, 6, 7 and ref 41. Wild-type (WT) zebrafish (ABTL) were a mixed lineage of WT AB and WT TL genetic background. Injections to generate CRISPR/Cas9 mediated knockout (KO) zebrafish were performed in ABTL embryos. Adult zebrafish were outcrossed to ABTL WT zebrafish for several rounds before incrossing. Obtained zebrafish lines were kept as carriers (heterozygous in genotype) and several adult zebrafish of one genetic, mutant line were outcrossed to different WT fish. This type of outcrossing was performed to segregate potential unlinked off-target mutations and to improve a healthy genetic background by preventing sibling mating.

Reagents and equipment

Supplementary Table 2| List of ingredients reagents including supplier and article number

Reagent	Supplier	
Oligonucleotides for sgRNA synthesis	IDT, Newark, USA	Supplementary Table 1
T4 DNA polymerase	Invitrogen, Carlsbad, SA	18005017
dNTPs	Promega, Madison, USA	U1240
Nucleospin PCR and gel clean-up kit	Macherey-Nagel, Duren, Germany	740609250
DEPC treated RNase free H_2O	In-house	
MEGAshortscript™ T7 Transcription kit	Ambion®, ThermoFisher, Waltham, USA	AM1354
Ethanol (HPLC grade)	VWR, Nederland	1009831000
pCS2-nCas9n	Addgene, Watertown, USA	#47929
NotI	New England Biolabs, Ipswich, USA	R0189L
Zymo DNA clean & concentrator™	Zymo, Irvine, USA	D4004
mMessage mMachine SP6 kit	Ambion®, ThermoFisher	AM1340
RNeasy mini kit	Qiagen, Hilden, Germany	74104
Capillaries		
QuickExtract™	EpiCentre®, Madison, USA	QE09050
IQ SYBR green	Bio-Rad, Hercules, USA	172-5006
Clear 96 PCR plate	Sarstedt, Nümbrecht, Germany	721978202
Phusion high-fidelity DNA polymerase	ThermoFisher	F530 L

Supplementary Table 3 | List of equipment including supplier

Equipment	Supplier	Notes	
PCR machine	Bio-Rad		
DeNovix DS-11	DeNovix, Wilmington, USA		
Bio-Rad CFX96 Touch™ Real-Time PCR	Bio-Rad		
P-97 Flaming/Brown micropipette puller	Sutter instruments	Program 9	
M-33 micromanipulator	Sutter instruments		Injection setup
Eppendorf™ Femtojet	ThermoFisher	t= 0.2 s, p= 200-600 Pa To obtain 1nL injections	Injection setup
Stereo microscope	Leica M50, Wetzlar, Germany		Injection setup
Dumont™ forceps #5	Fine Science tools, Foster city, USA	To break the needle	

Reagent preparation

In vitro synthesis of sgsRNA: Adapted from Gagnon and colleagues [26]

Double-stranded DNA (dsDNA) template

- 2 µL Forward oligo (100 µM, **Supplementary Table 4**), 2 µL Reverse oligo (100 µM, LL030) and 16µL MQ
- Anneal: 95°C 5 min, 95°C → 85°C with -2°C/s, 85°C → 25°C with -0,1°C/s, hold at 11°C
- Make premix: 8 µL 5x T4 buffer, 5 µl dNTPs (10 mM), 0.5 µL T4 polymerase and 6 µL Milli-Q for each reaction
- Add 20 µL of T4 DNA mix, incubate for 20 min to 1 h at 12 °
- Add 60 µL Milli-Q and 400 µL MilliQ:NTI (1:1 (v/v))of the Nucleospin PCR and gel clean-up kit
- Purify according to suppliers instruction

- Elute in 20 µL RNase free H_2O
- Determine the concentration of the dsDNA template (± 200-300 ng/µL)

Supplementary Table 4 | Single-guide RNA sequences with the T7 promoter underlined, the target specific sequence (20 nucleotides) in bold and single-nucleotide mutations are given as lower case. RNA synthesis starts from the second-to-last 'G' nucleotide of the T7 promotor.

Target	Exon	Code	Sequence 5'->3'
Gba1	1	LL031	GCGTAATACGACTCACTATA**GGAATAATCACCACAGCAAG**GTTTTAGAGCTAGAAATAGC
Gba2	3	LL038	GCGTAATACGACTCACTATA**gGTTCCTCTTGGTGGCATT**GGGTTTTAGAGCTAGAAATAGC
	3	LL039	GCGTAATACGACTCACTATA**GGCGGAGGGAGCATCACT**CGGTTTTAGAGCTAGAAATAGC
gpnmb	2	LL098	GCGTAATACGACTCACTATA**GgCAGCTGCGGTTTGAAGGG**GTTTTAGAGCTAGAAATAGC
	2	LL099	GCGTAATACGACTCACTATA**GGATACAGAGATTCGTCCCA**GTTTTAGAGCTAGAAATAGC
	2	LL100	GCGTAATACGACTCACTATA**GGGAACCGGACACGAACCCT**GTTTTAGAGCTAGAAATAGC
Asah1a	3	LL104	GCGTAATACGACTCACTATA**gGTGTCCATCTCTCACTAGG**GTTTTAGAGCTAGAAATAGC
	3	LL105	GCGTAATACGACTCACTATA**GgTATAGAGGAAACGTGACC**GTTTTAGAGCTAGAAATAGC
Asah1b	4	LL108	GCGTAATACGACTCACTATA**gGCAGACGCCTTTGTTCCCAG**GTTTTAGAGCTAGAAATAGC
	4	LL109	GCGTAATACGACTCACTATA**GGGAAGCTCATTCAGCTGG**GTTTTAGAGCTAGAAATAGC
	4	LL178	GCGTAATACGACTCACTATA**GgGCTTCCCGCTGGGAACAA**GTTTTAGAGCTAGAAATAGC
	4	LL179	GCGTAATACGACTCACTATA**GgATCATGCTCGCCATCTGA**GTTTTAGAGCTAGAAATAGC
Ncp1	2	LL253	GCGTAATACGACTCACTATA**GTATCTGGTACGGCGAATGT**GTTTTAGAGCTAGAAATAGC
	2	LL254	GCGTAATACGACTCACTATA**gGCAGCTCCTGTCCTTCATT**GTTTTAGAGCTAGAAATAGC
Cln8	3	LL206	GCGTAATACGACTCACTATA**gGGTCTTCGGGCCCTCATGG**GTTTTAGAGCTAGAAATAGC
	3	LL207	GCGTAATACGACTCACTATA**GgGAACAGATTCCTCCATGA**GTTTTAGAGCTAGAAATAGC

Transcribe sgRNA using MEGAshortscriptTM T7 Transcription kit

!! Use RNAse free vials, filter tips, gloves, RNase free H_2O.

- 0.5 µL 10x buffer, 0.5 µL each of ATP, GTP, UTP and CTP, 0.5-1.5 µL of dsDNA, 0.5 µL of T7 enzyme and add Nuclease-free water to 5 µL.
- Incubate at 37 °C for 3 h to overnight (run 0.5 µL of RNA sample on 1 % agarose gel to evaluate synthesis)
- Add 14 µL RNase free H_2O + 1 µL Turbo DNase and incubate 15 min at 37 °C
- Add 10 µL 5 M NH_4OAc (supplied in the kit) and 60 µL 100% ethanol (! Use freshly prepared 75% ethanol)
- Freeze for 30 min to 1 h at -80 °C
- Centrifuge 15 min at 4 °C. Remove supernatant & wash with 1 mL 75 % ethanol
- Centrifuge 10 min at 4 °C, remove all supernatant
- Dry 5 min at RT
- Resuspend in 20 µL RNase free H_2O. Determine the concentration & run 0.5 µL on 1% agarose gel
- Store 2 µL aliquots at -80 °C to avoid freeze-thaw cycles

Cas9 mRNA

- Linearize 1 µg pCS2-nCas9n with NotI for 1 h at 37 °C
- Purify using Zymo DNA clean & concentrator™ and determine the concentration
- Transcribe Cas9 mRNA in 20 µL according to manufacturer's instruction (mMessage mMachinery SP6)
- Clean-up with the RNAeasy kit according to the instructions.
- Elute in 20 µL RNase-free water (included in the RNeasy kit), measure the RNA concentration and check the integrity of the Cas9 mRNA on 1% agarose gel.
- Store at -80 °C in 2 µL aliquots to avoid freeze-thaw cycles

Microinjection

- Setup single crossings of WT zebrafish a day prior to injection
- Mix SgRNA, Cas9 mRNA, 0.5 µL phenol red and Milli-Q to a total volume of 5 µL. Final concentration of sgRNA and Cas9 mRNA is dependent on stock concentrations and integrity of the RNA as well as required mutagenic efficiency and mortality rate. Store sgRNA/Cas9 mRNA mix on ice before use.
- Remove the spacer of the single crossings in order to allow mating
- Pull needles with program 9 of the P-97 Flaming/Brown micropipette puller
- Pipet 4 µL in the pulled needle and setup the microinjector setup in order to inject a volume of 1 nL. Change the pressure of the Eppendorf™ Femtojet until a droplet of 1 nL is achieved.
- Place a microscopy slide in a 10 cm plastic petri dish, place the zebrafish embryos along the slide and remove excess water.
- Inject 1 nL in one- or two cell-stage embryos.
- Transfer injected embryos to a petri dish with egg water
- Remove and note the number of unfertilized, dead and deformed embryos
- To test mutagenesis efficiency, mini-finclip embryo at 4-5 dpf or extract genomic DNA from the entire larvae

Genomic DNA extraction

- Use 10µL QuickExtract per embryo and 15-20µL per finclip of an adult fish in PCR tubes or 96-well plate
- Incubate at 65°C for 10 min followed by 5 min at 98 °C
- Vortex samples to mix, dilute to 100 µL with Milli-Q, spin down quickly to pellet all non-processed particles which could interfere in the PCR reaction. Only use supernatant in PCR reactions.

HRM analysis (Work in columns: in order to efficiently use the multichannel & excel sheet)

- Master mix: 5 µL IQ SYBR green, 0.3 µL forward primer, 0.3 µL reverse primer (300 µM final concentration of each primer, **Supplementary Table 5**) and 3.4 µL MQ per sample.
- Use dedicated set of multichannel pipettes, to prevent possible contamination
- Add 9 µL HRM mix to each well of a PCR plate → Add 1 µL of gDNA
- Settings for CFX96 Real-Time PCR machine:
  - 95 °C for 3 min; 40x [95°C for 30 sec, 61°C for 30 sec, image plate]; hold at 55 °C for 30 sec; melt program [55-95°C with 0,5°C per step, image plate every step]
- Export melt curve data and process in excel sheet

Supplementary Table 5| Primers for high-resolution melting (HRM) analysis and fragment size

Target	Forward	Sequence 5'->3'	Reverse	Sequence 5'->3'	Fragment
Gba1	LL044	AGTCTCATCGGCAGGATGAG	LL045	CACTTGGACAGAAAGGTAAATC	123 bp
Gba2	LL042	GTATGTGTTGTTTTTTTCAGGC	LL043	GCAATAACGGTTTTGTAGTGG	141 bp
Gpnmb	LL101	CAATACATTCTTACCATGTCTGC	LL102	CAGCATATGGTGACATGTTCCC	149 bp
Asah1a	LL106	GTCTAGACTCGAATAAGTTCATG	LL107	TGGGAAACAGTTACCTCTGTG	169 bp
Asah1b	LL110	TGCAAAGAGATGTGTTAGATTG	LL111	TCCTTCAGATGGCGAGCATG	128 bp
Npc1	LL255	GGACCGGTGTAATTGCAGTTCAAC	LL256	GAAGTTGAACTGCAATTACACCG	152 bp
Cln8	LL208	TTCTAGGTTTTCTGGGATTTGGC	LL209	AGGACCAGTCTTCCTGTCCGAG	133 bp

PCR reaction for sequencing using phusion polymerase

- Master mix for 20 µL PCR reaction: 4 µL 5x HF buffer, 1 µL forward primer, 1 µL reverse primer (10 µM each, **Supplementary Table 6**), 0.4 µL dNTPs (10 mM), 0.2 µL phusion polymerase, 12.4 µL Milli-Q and 1 µL genomic DNA
- Settings of general PCR machine: Den. at 98 °C for 2 min; 35x [98 °C for 10 s, 65 °C for 15 s, 72 °C for 20 s] and a final extension at 72 °C for 5 min.
- Purify PCR product using Zymo DNA Clean & Concentrator™ and elute in 20 µL MQ
- 125-250 ng purified pcr product (size of 300-700 bp), 1 µL of sequence primer (10 µM, either forward or reverse primer, **Supplementary Table 6**) and Milli-Q to a total volume of 10 µL. Sanger sequencing is performed by Macrogen.

Supplementary Table 6| Primers to obtain DNA fragments for Sanger sequencing and fragment size

Target	Forward	Sequence 5'->3'	Reverse	Sequence 5'->3'	Fragment
Gba1	LL023	CATTGCCATTTTCGTTTTTAGG	LL007	GGAACTGTCCTTGACTCTCCAT	439 bp
Gba2	LL036	AATGGTGGTACCGAAAGACC	LL037	AGTACTACAGACTTCATCTGC	322 bp
Gpnmb	LL145	AGCCTAATCGTTGTAATACTCG	LL102	CAGCATATGGTGACATGTTCCC	295 bp
Asah1a	LL146	TGGGATGTATCCACCTAAAGG	LL147	CAGCAAGCAAAAGATGGACAG	251 bp
Asah1b	LL176	TACGATTTTGGGAGATTTATCTC	LL147	CAGCAAGCAAAAGATGGACAG	451 bp
Npc1	LL321	AGATATTGTCCAACAGAGCATTC	LL258	GTTTCCGATACAAAGCTAACGG	307 bp
Cln8	LL259	TGAATAAGACAGCATTGAAGCAAC	LL260	CAGTTCCACACAACCGCCCCTGC	439 bp

Tol2-mediated transgenesis - The coding sequences of zebrafish prosaposin (NCBI code NM_001309267) was amplified using generated cDNA of a pool of 5 dpf zebrafish larvae (SuperScript II™ reverse transcriptase, Thermo Fisher Scientific) as template and Phusion high-fidelity DNA polymerase using the primers described in **Supplementary Table 7**. The fragment was subsequently cloned into the pDONR entry vector using GATEWAY™ technology (BP reaction) according to the manufacturer's instruction. The point mutation was introduced in the *psap* pDONR vector using the QuickChange Lightning Site-Directed Mutagenesis Kit (Agilent, Santa Clara, USA) according to the suppliers protocol with the primers given in **Supplementary Table 7**. Generation of the human *GBA* pDONR construct described in chapter 2. The pDONR vectors were sequenced before generating the final destination vector. Destination vectors were was obtained by recombining entry vectors of p5E-ubi, pDONR-target and p3E-IRES-GFP-PA with pDEST-Tol2-crystalEye using a LR reaction.
The pCS2FA-transposase plasmid was linearized with *NotI* and purified using Zymo DNA clean & concentrator™. Capped and polyadenylated transposase mRNA was generated using the mMessage mMachine SP6 kit and purified using the RNAeasy mini kit as described for Cas9 mRNA. The concentration was determined, the integrity checked on aragose gel and aliquots stored at -80 °C until use.

Supplementary Table 7| Primers to generate pDONR entry vectors (pME)

Target	Sequence 5'->3'	Construct
psap F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAACCACCATGATGCTTCTCACGCTTC	pME-psap WT
psap R	GGGGACCACTTTGTACAAGAAAGCTGGGTcTTAGCTCCAAACGTGACGC	pME-psap WT
psap-mut F	GCCGATCCCAAAACGGTCggcTCTTTCCTTGCACTCTG	pME-psap mut
psap-mut R	CAGAGTGCAAGGAAAGAgccGACCGTTTTGGGATCGGC	pME-psap mut

References

1. Collins F.S., Green E.D., Guttmacher A.E., Guyer M.S.and Institute U.S.N.H.G.R. (2003) A vision for the future of genomics research. *Nature* **422**, 835-847.
2. Komor A.C., Badran A.H.and Liu D.R. (2017) CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. *Cell* **169**, 559.
3. Chen K.Y.and Knoepfler P.S. (2016) To CRISPR and beyond: the evolution of genome editing in stem cells. *Regen Med* **11**, 801-816.
4. Bibikova M., Golic M., Golic K.G.and Carroll D. (2002) Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. *Genetics* **161**, 1169-1175.
5. Christian M., Cermak T., Doyle E.L., Schmidt C., Zhang F., Hummel A.,... and Voytas D.F. (2010) Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* **186**, 757-761.
6. Jansen R., Embden J.D., Gaastra W.and Schouls L.M. (2002) Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* **43**, 1565-1575.
7. Brouns S.J., Jore M.M., Lundgren M., Westra E.R., Slijkhuis R.J., Snijders A.P.,... and van der Oost J. (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960-964.
8. Deveau H., Barrangou R., Garneau J.E., Labonte J., Fremaux C., Boyaval P.,... and Moineau S. (2008) Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. *J Bacteriol* **190**, 1390-1400.
9. Garneau J.E., Dupuis M.E., Villion M., Romero D.A., Barrangou R., Boyaval P.,... and Moineau S. (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**, 67-71.
10. Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A.and Charpentier E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821.
11. Gasiunas G., Barrangou R., Horvath P.and Siksnys V. (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A* **109**, E2579-2586.
12. Jinek M., East A., Cheng A., Lin S., Ma E.and Doudna J. (2013) RNA-programmed genome editing in human cells. *Elife* **2**, e00471.
13. Mali P., Yang L., Esvelt K.M., Aach J., Guell M., DiCarlo J.E.,... and Church G.M. (2013) RNA-guided human genome engineering via Cas9. *Science* **339**, 823-826.
14. Cong L., Ran F.A., Cox D., Lin S., Barretto R., Habib N.,... and Zhang F. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823.
15. Hwang W.Y., Fu Y., Reyon D., Maeder M.L., Tsai S.Q., Sander J.D.,... and Joung J.K. (2013) Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* **31**, 227-229.
16. Kuscu C., Arslan S., Singh R., Thorpe J.and Adli M. (2014) Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Biotechnol* **32**, 677-683.
17. Wu X., Scott D.A., Kriz A.J., Chiu A.C., Hsu P.D., Dadon D.B.,... and Sharp P.A. (2014) Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol* **32**, 670-676.
18. Pattanayak V., Lin S., Guilinger J.P., Ma E., Doudna J.A.and Liu D.R. (2013) High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol* **31**, 839-843.
19. Howe K., Clark M.D., Torroja C.F., Torrance J., Berthelot C., Muffato M.,... and Stemple D.L. (2013) The zebrafish reference genome sequence and its relationship to the human genome. *Nature* **496**, 498-503.
20. Cornet C., Di Donato V.and Terriente J. (2018) Combining Zebrafish and CRISPR/Cas9: Toward a More Efficient Drug Discovery Pipeline. *Frontiers in pharmacology* **9**, 703.
21. Postlethwait J., Ruotti V., Carvan M.J.and Tonellato P.J. (2004) Automated analysis of conserved syntenies for the zebrafish genome. *Methods Cell Biol* **77**, 255-271.
22. Meyer A.and Schartl M. (1999) Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. *Curr Opin Cell Biol* **11**, 699-704.
23. Guryev V., Koudijs M.J., Berezikov E., Johnson S.L., Plasterk R.H., van Eeden F.J.and Cuppen E. (2006) Genetic variation in the zebrafish. *Genome Res* **16**, 491-497.
24. Montague T.G., Cruz J.M., Gagnon J.A., Church G.M.and Valen E. (2014) CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic acids research* **42**, W401-407.
25. Labun K., Montague T.G., Gagnon J.A., Thyme S.B.and Valen E. (2016) CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic acids research* **44**, W272-276.
26. Gagnon J.A., Valen E., Thyme S.B., Huang P., Akhmetova L., Pauli A.,... and Schier A.F. (2014) Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PloS one* **9**, e98186.
27. Jao L.E., Wente S.R.and Chen W. (2013) Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc Natl Acad Sci U S A* **110**, 13904-13909.

28. Varshney G.K., Carrington B., Pei W., Bishop K., Chen Z., Fan C.,... and Burgess S.M. (2016) A high-throughput functional genomics workflow based on CRISPR/Cas9-mediated targeted mutagenesis in zebrafish. *Nat Protoc* **11**, 2357-2375.
29. Talbot J.C.and Amacher S.L. (2014) A streamlined CRISPR pipeline to reliably generate zebrafish frameshifting alleles. *Zebrafish* **11**, 583-585.
30. Pfeffer S.R. (2019) NPC intracellular cholesterol transporter 1 (NPC1)-mediated cholesterol export from lysosomes. *J Biol Chem* **294**, 1706-1709.
31. Tyynela J., Cooper J.D., Khan M.N., Shemilts S.J.and Haltia M. (2004) Hippocampal pathology in the human neuronal ceroid-lipofuscinoses: distinct patterns of storage deposition, neurodegeneration and glial activation. *Brain Pathol* **14**, 349-357.
32. Zhang C.K., Stein P.B., Liu J., Wang Z., Yang R., Cho J.H.,... and Mistry P.K. (2012) Genome-wide association study of N370S homozygous Gaucher disease reveals the candidacy of CLN8 gene as a genetic modifier contributing to extreme phenotypic variation. *Am J Hematol* **87**, 377-383.
33. Shen M.W., Arbab M., Hsu J.Y., Worstell D., Culbertson S.J., Krabbe O.,... and Sherwood R.I. (2018) Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature* **563**, 646-651.
34. Abadi S., Yan W.X., Amar D.and Mayrose I. (2017) A machine learning approach for predicting CRISPR-Cas9 cleavage efficiencies and patterns underlying its mechanism of action. *PLoS Comput Biol* **13**, e1005807.
35. Marques A.R., Gabriel T.L., Aten J., van Roomen C.P., Ottenhoff R., Claessen N.,... and van Eijk M. (2016) Gpnmb Is a Potential Marker for the Visceral Pathology in Niemann-Pick Type C Disease. *PloS one* **11**, e0147208.
36. Kramer G., Wegdam W., Donker-Koopman W., Ottenhoff R., Gaspar P., Verhoek M.,... and van Eijk M. (2016) Elevation of glycoprotein nonmetastatic melanoma protein B in type 1 Gaucher disease patients and mouse models. *FEBS Open Bio* **6**, 902-913.
37. Carcel-Trullols J., Kovacs A.D.and Pearce D.A. (2015) Cell biology of the NCL proteins: What they do and don't do. *Biochimica et biophysica acta* **1852**, 2242-2255.
38. Hermansson M., Kakela R., Berghall M., Lehesjoki A.E., Somerharju P.and Lahtinen U. (2005) Mass spectrometric analysis reveals changes in phospholipid, neutral sphingolipid and sulfatide molecular species in progressive epilepsy with mental retardation, EPMR, brain: a case study. *J Neurochem* **95**, 609-617.
39. Ferraz M.J., Marques A.R., Gaspar P., Mirzaian M., van Roomen C., Ottenhoff R.,... and Aerts J.M. (2016) Lyso-glycosphingolipid abnormalities in different murine models of lysosomal storage disorders. *Mol Genet Metab* **117**, 186-193.
40. Marques A.R., Mirzaian M., Akiyama H., Wisse P., Ferraz M.J., Gaspar P.,... and Aerts J.M. (2016) Glucosylated cholesterol in mammalian cells and tissues: formation and degradation by multiple cellular beta-glucosidases. *J Lipid Res* **57**, 451-463.
41. Lelieveld L.T., Mirzaian M., Kuo C.L., Artola M., Ferraz M.J., Peter R.E.A.,... and Aerts J. (2019) Role of beta-glucosidase 2 in aberrant glycosphingolipid metabolism: model of glucocerebrosidase deficiency in zebrafish. *J Lipid Res* **60**, 1851-1867.
42. Tseng W.C., Loeb H.E., Pei W., Tsai-Morris C.H., Xu L., Cluzeau C.V.,... and Porter F.D. (2018) Modeling Niemann-Pick disease type C1 in zebrafish: a robust platform for in vivo screening of candidate therapeutic compounds. *Dis Model Mech* **11**.
43. van der Lienden M.J.C., Gaspar P., Boot R., Aerts J.and van Eijk M. (2018) Glycoprotein Non-Metastatic Protein B: An Emerging Biomarker for Lysosomal Dysfunction in Macrophages. *Int J Mol Sci* **20**.
44. Kawakami K. (2005) Transposon tools and methods in zebrafish. *Developmental dynamics : an official publication of the American Association of Anatomists* **234**, 244-254.
45. Kwan K.M., Fujimoto E., Grabher C., Mangum B.D., Hardy M.E., Campbell D.S.,... and Chien C.B. (2007) The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Developmental dynamics : an official publication of the American Association of Anatomists* **236**, 3088-3099.
46. Kawakami K., Takeda H., Kawakami N., Kobayashi M., Matsuda N.and Mishina M. (2004) A transposon-mediated gene trap approach identifies developmentally regulated genes in zebrafish. *Dev Cell* **7**, 133-144.
47. Tylki-Szymanska A., Czartoryska B., Vanier M.T., Poorthuis B.J., Groener J.A., Lugowska A.,... and Jurkiewicz E. (2007) Non-neuronopathic Gaucher disease due to saposin C deficiency. *Clin Genet* **72**, 538-542.
48. Tylki-Szymanska A., Groener J.E., Kaminski M.L., Lugowska A., Jurkiewicz E.and Czartoryska B. (2011) Gaucher disease due to saposin C deficiency, previously described as non-neuronopathic form--no positive effects after 2-years of miglustat therapy. *Mol Genet Metab* **104**, 627-630.
49. Vaccaro A.M., Motta M., Tatti M., Scarpa S., Masuelli L., Bhat M.,... and Salvioli R. (2010) Saposin C mutations in Gaucher disease patients resulting in lysosomal lipid accumulation, saposin C deficiency, but normal prosaposin processing and sorting. *Human molecular genetics* **19**, 2987-2997.
50. Vaccaro A.M., Salvioli R., Barca A., Tatti M., Ciaffoni F., Maras B.,... and Pucci P. (1995) Structural analysis of saposin C and B. Complete localization of disulfide bridges. *J Biol Chem* **270**, 9953-9960.
51. Wilkinson R.N., Elworthy S., Ingham P.W.and van Eeden F.J. (2013) A method for high-throughput PCR-based

genotyping of larval zebrafish tail biopsies. *Biotechniques* **55**, 314-316.
52. Armstrong G.A., Liao M., You Z., Lissouba A., Chen B.E.and Drapeau P. (2016) Homology Directed Knockin of Point Mutations in the Zebrafish tardbp and fus Genes in ALS Using the CRISPR/Cas9 System. *PloS one* **11**, e0150188.
53. Eschstruth A., Schneider-Maunoury S.and Giudicelli F. (2020) Creation of zebrafish knock-in reporter lines in the nefma gene by Cas9-mediated homologous recombination. *Genesis* **58**, e23340.
54. Auer T.O., Duroure K., De Cian A., Concordet J.P.and Del Bene F. (2014) Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Res* **24**, 142-153.
55. Kesavan G., Chekuru A., Machate A.and Brand M. (2017) CRISPR/Cas9-Mediated Zebrafish Knock-in as a Novel Strategy to Study Midbrain-Hindbrain Boundary Development. *Front Neuroanat* **11**, 52.
56. Burger A., Lindsay H., Felker A., Hess C., Anders C., Chiavacci E.,... and Mosimann C. (2016) Maximizing mutagenesis with solubilized CRISPR-Cas9 ribonucleoprotein complexes. *Development* **143**, 2025-2037.
57. Trubiroha A., Gillotay P., Giusti N., Gacquer D., Libert F., Lefort A.,... and Costagliola S. (2018) A Rapid CRISPR/Cas-based Mutagenesis Assay in Zebrafish for Identification of Genes Involved in Thyroid Morphogenesis and Function. *Sci Rep* **8**, 5647.
58. Kuil L.E., Lopez Marti A., Carreras Mascaro A., van den Bosch J.C., van den Berg P., van der Linde H.C.,... and van Ham T.J. (2019) Hexb enzyme deficiency leads to lysosomal abnormalities in radial glia and microglia in zebrafish brain development. *Glia* **67**, 1705-1718.

4